



# *PACIFIC LINGUISTICS*

School of culture, History and Language

College of Asia and the Pacific

THE AUSTRALIAN NATIONAL UNIVERSITY

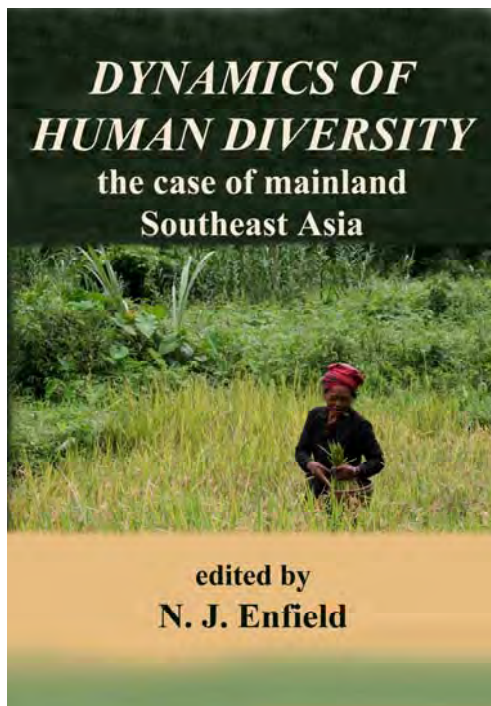
---

## **Out of print books in pdf format**

This PDF document may be copied, printed and stored for use in libraries and for personal use. It may not be reproduced for sale or distribution.

*Pacific Linguistics* Home Page:

<http://www.pacling.com>



## **627 Dynamics of human diversity: the case of mainland Southeast Asia**

edited by N. J. Enfield

Human diversity is the central problem of all the fields of anthropology. Our languages, our genetics, our material cultures, our social organization: these are woven together by the ancient processes of change and diversification that produce the rich diversity we see today. What are these processes and how do they work? Can we know what life was like 10,000 years ago, and how it came to be the way it is today?

*Dynamics of human diversity* looks at these questions with a focus on one of the most fascinating sites of human diversity worldwide: mainland Southeast Asia (MSEA). In this book, experts on MSEA from across the disciplines of anthropology—linguistics, social anthropology, human biology, genetics, archaeology—bring together the latest empirical, methodological, and theoretical advances. Special attention is paid to two case studies of human diversity in MSEA: the Aslian peoples of Peninsular Malaysia, and the origin and diversification of the Austroasiatic languages. These, along with other chapters, show how new techniques for data collection and analysis are radically transforming what we know—and can know—about the past, and about the dynamic processes of human diversification.

The chapters of this book raise challenges for some common assumptions about the dynamics of diversity, especially for the idea that the key event in MSEA was a wave of agricultural colonization by ‘demic diffusion’. New evidence and analysis reviewed here suggests alternatives. By a scenario of population continuity, early resident populations of MSEA played a more agentive role in the social diffusion of ideas, technology, language, genes and cultural practices. The issues are explored here from a range of disciplinary approaches and points of view.

ISBN 9780858836389

396 pp

# *Dynamics of human diversity*

## Pacific Linguistics 627

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

---

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: Malcolm Ross and I Wayan Arka (Managing Editors),  
Mark Donohue, Nicholas Evans, David Nash, Andrew Pawley,  
Paul Sidwell, Jane Simpson, and Darrell Tryon

---

### EDITORIAL ADVISORY BOARD:

Karen Adams, *Arizona State University*  
Alexander Adelaar, *University of Melbourne*  
Peter Austin, *School of Oriental and African Studies*  
Byron Bender, *University of Hawai'i*  
Walter Bisang, *Johannes Gutenberg-Universität Mainz*  
Robert Blust, *University of Hawai'i*  
David Bradley, *La Trobe University*  
Lyle Campbell, *University of Hawai'i*  
James Collins, *Northern Illinois University*  
Bernard Comrie, *Max Planck Institute for Evolutionary Anthropology*  
Matthew Dryer, *State University of New York at Buffalo*  
Jerold A. Edmondson, *University of Texas at Arlington*  
Margaret Florey, *Resource Network for Linguistic Diversity*  
William Foley, *University of Sydney*  
Karl Franklin, *SIL International*  
Charles Grimes, *SIL International*  
Nikolaus Himmelmann, *Universität zu Köln*  
Lillian Huang, *Shih-chien University, Taiwan*  
Marian Klamer, *Universiteit Leiden*  
Harold Koch, *The Australian National University*

Frantisek Lichtenberk, *University of Auckland*  
John Lynch, *University of the South Pacific*  
Patrick McConvell, *Australian Institute of Aboriginal and Torres Strait Islander Studies*  
William McGregor, *Aarhus Universitet*  
Ulrike Mosel, *Christian-Albrechts-Universität zu Kiel*  
Claire Moyse-Faurie, *Centre National de la Recherche Scientifique*  
Bernd Nothofer, *Johann Wolfgang Goethe-Universität Frankfurt am Main*  
Bambang Kaswanti Purwo, *Universitas Atma Jaya*  
Ger Reesink, *Radboud University Nijmegen*  
Lawrence Reid, *University of Hawai'i*  
Jean-Claude Rivierre, *Centre National de la Recherche Scientifique*  
Melenaite Taumoefolau, *University of Auckland*  
Tasaku Tsunoda, *University of Tokyo*  
John Wolff, *Cornell University*  
Elizabeth Zeitoun, *Academica Sinica, Taipei*

# **Dynamics of human diversity**

edited by N. J. Enfield



Pacific Linguistics  
School of Culture, History and Language  
College of Asia and the Pacific  
The Australian National University

Published by Pacific Linguistics  
School of Culture, History and Language  
College of Asia and the Pacific  
The Australian National University  
Canberra ACT 2600  
Australia

Copyright in this edition is vested with Pacific Linguistics

First published 2011

National Library of Australia Cataloguing-in-Publication entry:

Author: Enfield, N. J., 1966-

Title: Dynamics of human diversity : the case of mainland Southeast Asia / N. J. Enfield.

ISBN: 9780858836389 (pbk.)

Notes: Includes bibliographical references.

Subjects: Anthropology--Southeast Asia.  
Ethnology--Southeast Asia.  
Agricultural colonies--Southeast Asia.  
Southeast Asia--Social life and customs.  
Southeast Asia--Languages.  
Southeast Asia--History.

Other Authors/ Australian National University.

Contributors: Research School of Pacific and Asian Studies. Pacific Linguistics.

Dewey Number: 301.0959

Typeset by the editor and assistants  
Cover design by Julie Manley and Addcolour Digital  
Printed and bound by Addcolour Digital, Canberra

# *Table of contents*

---

1	Dynamics of human diversity in Mainland Southeast Asia: introduction N. J. ENFIELD .....	1
---	---	---

## **Part I: Disciplinary Perspectives**

2	Cultural diversity in Mainland Southeast Asia: a view from prehistory JOYCE C. WHITE .....	9
3	Human diversity in Mainland Southeast Asia: the contribution of bioarchaeology SIÂN E. HALCROW, N. TAYLES .....	47
4	Linguistic Diversity in mainland Southeast Asia N. J. ENFIELD .....	63
5	MtDNA variation and southward Holocene human dispersals within Mainland Southeast Asia STEPHEN OPPENHEIMER.....	81
6	Ethnology and the Issue of Human Diversity in Mainland Southeast Asia HIORLEIFUR JONSSON .....	109

## **Part II: Archaeological advances**

7	The role of agriculture in the evolution of Mainland Southeast Asian language phyla ROGER BLENCH .....	125
8	Population history of MSEA: The Two Layer model in the context of Northern Vietnam HIROFUMI MATSUMARA, MARC F. OXENHAM, NGUYEN KIM THUY, NGUYEN LAN CUONG AND NGUYEN KIM DUNG.....	153

- 9 A physical anthropology perspective on a mainland Southeast Asian agrarian population: prehistoric skulls of Ban Chiang  
KORAKOT BOONLOP AND SUREERATANA BUBPHA.....179

### **Part III: Intra-group diversity: the case of Aslian groups**

- 10 Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya  
DAVID BULBECK .....207
- 11 Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula  
NICLAS BURENHULT, NICOLE KRUSPE AND MICHAEL DUNN .....257
- 12 Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model  
ALAN FIX.....277

### **Part IV: Origins and diversification: the case of Austroasiatic groups**

- 13 Austroasiatic word histories: boat, husked rice and taro  
GÉRARD DIFFLOTH .....295
- 14 The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis  
PAUL SIDWELL AND ROGER BLENCH.....315
- 15 The Austroasiatics: East to West or West to East?  
LAURENT SAGART .....345
- 16 Rice and the Austroasiatic and Hmong-Mien homelands  
GEORGE VAN DRIEM.....361

# 1 *Dynamics of human diversity in mainland Southeast Asia: Introduction*

---

N. J. ENFIELD

## 1 Introduction

This book explores human diversity and its dynamic causes and consequences with reference to the mainland Southeast Asia (MSEA) region during the Holocene (the last 10,000 years). Global history combined with a unique human capacity for environmental adaptation, ethnic distinction, and cumulative culture has given rise to the rich diversity of cultures and languages that we now observe worldwide. Mainland Southeast Asia is a good case study, with greater linguistic, genetic and cultural diversity than almost anywhere else outside Africa. Here we ask: What is the nature of this diversity? How has it come about?

MSEA—that is, the present day location of Vietnam, Laos, Cambodia, Thailand, and peninsular Malaysia, along with bordering areas of Burma and China—has long been recognized as having a special degree of cultural and social diversity. It has been both a buffer region and a crossroads between the major modern historical areas of South Asia, China, and the Pacific. For a range of reasons explored in this book, the human population of MSEA shows a remarkable kind of socio-cultural diversity and historical dynamicity in world terms. The book examines the nature of this diversity and its dynamic modes of development, as a case study for all scholars interested in global human diversity today.

Human diversity is a product of dynamic processes of dispersal, interaction, and change in the history of our species. The social, environmental, and cultural causes of diversity are measured by all the tools of anthropology—drawing from archaeology, human biology, linguistics, and ethnography—making the study of human diversity a truly interdisciplinary affair. Recent advances in the different disciplines are helping to define the mechanisms of human history in parts of the world including Europe, the Pacific, and more recently East Asia and areas of Africa and the Americas. While MSEA itself has been less studied than other areas, the last ten years have seen exciting and significant developments, particularly in the rapidly developing fields of archeology (including bioarchaeology), human genetics, and linguistics (see Glover & Bellwood 2004; Sagart et al 2005, Oxenham & Tayles 2006). The key challenge now is to continue bridging the gaps in our understanding of both empirical and theoretical advances **across** anthropological disciplines.

---

N. J. Enfield, editor.

*Dynamics of human diversity, 1-5.* Pacific Linguistics, 2011

© This edition vested with Pacific Linguistics.



MSEA is an excellent focus as a microcosm for universal questions of human diversity and its dynamics because of (a) the nature and degree of its diversity, characterized by the presence of at least five distinct major ethnolinguistic families represented in hundreds of different ethnic groups within a relatively small geographical region, (b) the fact that MSEA is relatively understudied in comparison to neighbouring Island Southeast Asia, and (c) new data and methods that have arisen in each of the anthropological disciplines, enabling more nuanced interpretations.

The chapters of this book are, in part, based on presentations that were given at a closed workshop held in Siem Reap in 2009. This workshop was conceived and organized by the editor together with Joyce C. White (U. Penn Museum of Archaeology and Anthropology), with funding support from a Wenner-Gren Foundation workshop grant awarded to Enfield and White, as well as funding from the Max Planck Institute for Psycholinguistics in Nijmegen. A number of participants contributed with funding from their universities and affiliated institutes. The workshop proposal written by Enfield and White formulated a core set of research questions on the dynamics of human diversity in MSEA. With these questions as a frame for what we envisioned for the workshop, contributions were invited from ideal contributors. After the workshop, several further contributions were invited for the book, to complement those that had been aired at the workshop. The resulting roster of contributors to this book are experts from all fields of anthropology, whose empirical and theoretical work has contributed significantly to key areas of understanding the dynamics of diversity in this well-defined geographical and historical region.

## 2 Questions posed

There are two overarching types of question that have shaped the chapters presented here. First, there are questions about the empirical facts and analyses of MSEA human diversity: What is the nature of human diversity in MSEA? How did it come to be this way? What are the dynamic aspects of this diversity? Second, there are questions concerning the relationships among respective theoretical and empirical approaches to answering these first questions: To what extent do different data and assumptions determine the way we think about the emergence of human diversity in MSEA? How can we best communicate our empirical and theoretical concerns across the sub-disciplines?

Following are the substantive sub-questions that the workshop organizers circulated by way of preparation for the drafting of chapters. While this book does not answer, or even directly address, all of these questions, they are listed here for two reasons. The first reason is to contextualize the chapters and point to connections between them. The second is that we want to register these questions as being among those most important for ongoing interdisciplinary research in this field, where so much work remains to be done.

*State(s) of the art(s)*: For each branch and sub-branch of anthropology concerned with the dynamics of human diversity in MSEA, what is the current state of the art? What is well established in each field, and what remains unknown? What is commonly agreed and what remains controversial? What are the hot topics, and why these? What are the key puzzles? Where are the current gaps in research?

*Peopling* (the process by which people of a social group move into a region they previously did not inhabit): What evidence do we find in the different disciplines for peopling activities? Who moved, to where, when, and from where? Can disciplinary disagreements of fact and interpretation be resolved? A special concern of the Siem Reap

workshop was to critically examine a prevalent macro-scale theory among archaeologists that posits agricultural dispersal as the prime mover for bringing mongoloid populations and Austroasiatic languages to MSEA in the late third millennium B.C. Does the latest evidence in archaeology, genetics, palaeodemography, osteology, linguistics, demography, ethnography, and the palaeoenvironment support this scenario? If not, what alternatives are indicated for the relationships among dispersal of populations, languages, and agricultural technologies, and what evidence is needed in future research to resolve discrepancies among the subdisciplines?

*Genetics:* There are clear genetic parallels for Holocene agriculturally-associated north-south cultural flow within MSEA, providing some possible explanation for the presence of Austroasiatic languages among culturally diverse aboriginal populations far south in the Malay Peninsula. However, these are relatively minor and most of the genetic landscape of MSEA and Malaya (and ISEA) was in place well before the Holocene, thus challenging for a conventional model of agricultural Holocene replacement in MSEA from South China.

*Language:* What is the state of the art of the chronology for processes of development, convergence, and differentiation of language families in MSEA? This volume pays special attention to Austroasiatic language family, since languages of this family had the earliest presence of languages currently spoken in the area. Can current chronologies be correlated with archaeological evidence? MSEA shows the highest degree of structural convergence of languages in the world (and hence lower diversity in one sense)—what is the cause? Current developments in the methodology of analysing historical-comparative language data might affect our interpretation of the linguistic facts of MSEA, as should the flood of new descriptive data. Further questions will require new approaches: Do new theories of language contact force us to change our way of thinking about the MSEA language situation? Are there ways to determine whether past contact situations might have involved stable bilingualism? Are there sensible ways to speak of linguistic processes reaching back more than 5000 years in time? Linguists are increasingly looking to combine methodologies, including well-established approaches to historical-comparative linguistics and new applications of quantitative methods developed in biology.

*Social structure:* What do we know about comparative social structure in MSEA? Is there the same degree of structural convergence in social structure as found in linguistic structure? To what degree do patterns of social organisation such as kinship and marriage in MSEA resist ‘horizontal’ transmission (that is, borrowing through contact between social groups)? Different cultural mechanisms must have influenced the maintenance of bio-cultural diversity versus homogenisation and integration in MSEA. What role has been played by marriage rules, demography, ecological/subsistence adaptation, material culture? And what can be said about the dynamics of inter-ethnic relations in a pre-nationalist MSEA? This last question is particularly important since most of the area’s diversification has taken place prior to the emergence of states, and most of the relevant data from other disciplines (archaeology most obviously) relates to time periods well before nationalism.

*Dynamics of micro-macro relations:* Can an understanding of micro-scale processes (for example, marriage patterns, epidemiology, trade, ritual) be successfully incorporated into larger scale discussions of regional diversity? Dynamics at the scale of small population demography and disease patterns, ritual relations, sub group identity formation, and regional agricultural responses to environmental risk likely underlay larger scale patterns

such as gene flow and population movements. What evidence among the subdisciplines exists for smaller scale processes and their impact on larger scale outcomes? What osteological and archaeological evidence exists for small group identities, settlement and breeding population stability and/or flexibility? How might language formation be related to population-formation processes?

*Interdisciplinarity:* To what extent are the interdisciplinary findings compatible? New empirical findings (for example from new analyses of skeletal remains, genetic data, or new language data) may challenge current wisdom, and they may help us to decide among competing hypotheses. What puzzles arise from incompatibilities? What theoretical syntheses can arise? Empirical and theoretical research on the dynamics of human diversity in MSEA in the coming years will require an ongoing process of scholarly investigation and dialogue.

### **3 Organisation of the book**

Each of the book's four parts contains a set of chapters which approach the above questions in related ways, either in terms of a common 'granularity' of perspective (Part I), a common focus on certain kinds of data and methodology (Parts II and IV), or a common focus on a particular empirical domain as a case study for many of the above questions (Parts II-IV). Part I offers overviews of human diversity in MSEA from complementary points of view across the range of disciplines represented in the book. Part II features recent empirical and analytic advances in archaeology, with some attention to their broader disciplinary consequences, and with a special emphasis on bioarchaeology. Part III focuses on human diversity by tackling a critical case study of local diversity within one broad ethnolinguistic group, namely the Aslian speakers of peninsular Malaysia (see also Chapter 5). The three chapters each delve deeply into new areas of empirical and theoretical work. Part IV deals with the problem of origins and dispersal of human groups, again by taking a common focus on a critical case study, though with notably different views of what happened and how. Each of the four chapters assesses the origins and diversification of the Austroasiatic language family.

### **4 Envoi**

It is our sincere hope that this book will complement other edited volumes of similar orientation that have appeared over the last ten years, books that are oriented more broadly toward East Asia and Island Southeast Asia, mostly with a focus on Austronesian groups (see Jin et al 2001, Glover and Bellwood 2004, Sagart et al 2005, Sanchez-Mazas et al 2008). Additionally, they cover a temporal span going back into the late Pleistocene, while here we are mostly constrained to the Holocene. Another volume focuses on latest developments in one research sub-discipline—bioarchaeology—in the region (Oxenham and Tayles 2006). Here we pay special attention to Austroasiatic groups, with many contributors sharing a special interest in the 'nonstate' peoples of the uplands of MSEA (Scott 2009). While much previous attention has been paid to the origins and diversification of Tai and Austronesian groups, the earlier presence of Austroasiatic groups, descendants of whom are now scattered through the uplands of the MSEA area (excepting those who speak the national languages of Vietnam and Cambodia), has been presupposed but is in need of a good deal more discussion. The better we understand these

earlier groups and their descendants, the better we may understand the dynamics of human diversity in MSEA.

## Acknowledgements

I would like to express my heartfelt thanks to all participants in the Siem Reap workshop, and all contributors to this volume, for their rich ideas and generous engagement throughout the project, as well as for their patience and support during the book's production. I gratefully acknowledge several sources of funding that provided support for various costs along the way: The Wenner-Gren Foundation for Anthropological Research, The Max Planck Institute for Psycholinguistics, Nijmegen, and the European Research Council (Project HSSLU). I thank Christophe Pottier of the Centre of the Ecole française d'Extrême-Orient, for kind assistance in hosting the Siem Reap meeting, and Joyce White for her help in the conception and organization of the workshop, as well as in editorial matters. The anonymous reviewers are gratefully acknowledged for giving their time and expertise to improve the book significantly. For their help during the production of this book, I would like to thank Annelies van Wijngaarden for skilled assistance with preparation of the manuscript in the early stages, Peter Johan Nijland for his expert preparation of many of the illustrations in the book, and Paul Sidwell and Julie Manley at Pacific Linguistics for their patience and helpful attention throughout. And finally, I thank my inimitable assistant Julija Baranova for her brilliant handling of not just a complicated and sometimes confusing project, but also of a complicated and sometimes confusing boss. Julija's unruffled poise improves everything.

Thank you all!

## References

- Glover, Ian, & Peter Bellwood (eds). 2004. *Southeast Asia: from prehistory to history*. London and New York: RoutledgeCurzon.
- Jin, Li, Mark Seielstad, & Chunjie Xiao (eds). 2001. *Genetic, Linguistic and Archaeological Perspectives on Human Diversity in Southeast Asia*. River Edge, NJ: World Scientific Publishing Co.
- Oxenham, Marc F. & Nancy Tayles (eds). 2006. *Bioarchaeology of Southeast Asia*. Cambridge: Cambridge University Press.
- Sagart, Laurent, Roger Blench, & Alicia Sanchez-Mazas (eds). 2005. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London and New York: RoutledgeCurzon.
- Sanchez-Mazas, Alicia, Roger Blench, Malcolm Ross, Ilia Peiros & M. Lin (eds). 2008. *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics, and genetics*. London: Routledge.
- Scott, James C. 2009. *The art of not being governed: An anarchist history of upland Southeast Asia*. New Haven & London: Yale University Press.



# *Part I*

*Disciplinary Perspectives*



# 2 *Emergence of cultural diversity in mainland Southeast Asia: a view from prehistory*

---

JOYCE C. WHITE

## 1 Introduction

The study of cultural diversity is central to the discipline of anthropology (Hannerz 2010), but anthropological archaeologists who specialise in mainland Southeast Asia (MSEA) have not made the study of diversity in the archaeological record a priority. Nevertheless, one can discern in the literature at least two general views on the overall timing and processes of cultural diversification during Southeast Asia's prehistoric period (i.e., before about 500 CE). The most widely held view, here called View One, is that cultural diversification in Southeast Asia is a late Holocene phenomenon, occurring after rice cultivating societies bearing a so-called 'neolithic' cultural package and speaking Austroasiatic languages expanded into the region (Bellwood 2005:132). In this 'Neolithic wave of advance from the north' (Higham et al. 2011:541) model, the ultimate source for the neolithic package is the Yangtze Valley cultures dating 6000–4000 BCE (Rispoli 2007). The chronology for the spread of the proposed neolithic package within MSEA is not yet clear. Dates proposed for the arrival of neolithic societies/package in MSEA include ca. 3000 BCE (Bellwood 2009), late third millennium BCE (Rispoli 2007), and as late as ca. 1650 BCE (Higham and Higham 2009). According to Bellwood (2005:132), regionalisation and diversification of 'cultural style' developed after the initial dispersal of these fairly homogeneous neolithic societies.

Implied in some older literature is another view, one that suggests that processes of cultural diversification in MSEA predate the prominence of rice agriculture in the region. Scholars holding this view, here called View Two, see the region's cultural diversity as embedded in and reflective of the ecological diversity generally inherent in tropical ecosystems (Dunn 1975; Hutterer 1976; Kennedy 1977, 1978). In this second view, rice agriculture and its possible linguistic correlates are de-emphasised and priority is given to investigating the variety and interdependencies among human resource exploitation systems, including various agroecosystems that emerged by the late Holocene (Hutterer 1976, 1983, 1988).

View One is embedded in a high level theory, termed the 'farming/language dispersal hypothesis', that is elegant, elaborate, and global in scope (Bellwood and Renfrew 2002). The hypothesis in its variant forms posits an array of relationships among early cereal



cultivators, the global distribution of language families, and large scale demographic changes that is said to explain the foundation and dispersal from homelands of the agrarian economies, populations, and lifeways that have dominated the earth for the last several millennia (Bellwood 2002, 2005, 2009). With regard to MSEA, this view maintains that pre-rice farming peoples contributed little to the subsequent cultural and biological character of the region (Bellwood 2005:130). Low density foraging populations and pre-farming lifeways were out-competed or subsumed by the rapidly expanding rice-growing societies.

View Two sees MSEA diversity as a fundamentally autochthonous process and a byproduct of tropical subsistence systems, both generally and with some aspects peculiar to the Southeast Asian geographic region. While View One argues that the appearance of agriculture in MSEA represented a fundamental and rapid discontinuity with preceding hunter-gatherer societies (Bellwood 2005:130; Higham 2009:252), View Two expects a greater subsistence and population continuity during a lengthy, though geographically uneven, overall regional transition from predominantly hunting and gathering to predominantly agricultural modes of subsistence (here termed the HG-AG transition; see Kealhofer 2002). Although View One sees rice agriculture as the driver of early 'neolithisation', View Two hypothesises that the prominence of rice agriculture is a relatively recent phenomenon preceded by extensive, multi-faceted poly-cultural agricultural systems including horticulture and shifting dryland cultivation of many crops. [See also Sidwell & Blench (this volume) who posit root crops and access to aquatic transport as key drivers in the transformation of the MSEA landscape.] View Two predicts that cultural differentiation emerged initially in areas with ecological mosaics, as local groups, including hunter-gatherers, exchanged resources that were differentially distributed across the landscape (Kennedy 1978).

Although View One dominates in recent regional archaeological literature, proponents of both views recognise that these hypotheses need more data (Higham et al. 2011; Hutterer 1988). Regional archaeologists acknowledge that data sets in MSEA are biased by small samples and uneven geographic coverage. Most archaeologists in any area are painfully aware of the perpetual gaps in the archaeological record, but in Southeast Asia the gaps pertinent to the topic addressed here are particularly gaping.

Nonetheless, evidence for cultural diversity, such as in the sense of Jonsson's (this volume) 'localised domains of identity', is present in Thailand's archaeological record during the metal age (2000 BCE–500 CE; see White and Eyre 2011). The distinct metal age ceramic subregions discussed in White and Eyre (2011) suggest that sets of settlements shared distinctive styles and technologies in mortuary pottery in contradistinction to other sets of contemporaneous settlements with different shared pottery styles and technologies. The socio-cultural subgroupings implied in the subregional ceramic traditions might at first glance appear to support View One—diversification following neolithisation. But that article does not address the possibility that cultural diversification in MSEA existed before the metal age, nor does that article fully address the genesis of metal age cultural diversity. Can diversity be identified in the earliest appearance of agriculture, or in earlier hunter-gatherer contexts? Are the fundamental processes of the early cultural diversification best characterised as divergence from a homogeneous predecessor? Or did domains of identity constellate from even more diverse predecessor societies? While definitive answers cannot be given with the limited evidence we have now, this chapter will argue that Southeast Asia has an extraordinary diversity of ecological zones and that the region's prehistoric cultural diversity is closely linked to that environmental diversity. Some of the contributing

factors to Southeast Asia's ecological diversity are summarised in the first part of this chapter.

This chapter also argues that variation in technological choice underlies many aspects of cultural diversity, and it can be observed in the record from the initial settlement of anatomically modern humans in Southeast Asia. Technological choice is inferred among the Holocene hunter-gatherers despite a relatively uniform and widespread stone tool tradition, and is seen in the archaeological record from the earliest documented agricultural/horticultural societies. In addition, I propose that early bamboo exploitation was an important medium by which societies made technological choices, developed technological styles, and initially articulated cultural diversity in material culture.

Furthermore, I propose that future studies of the prehistoric period should not assume that rice economies are the *sine qua non* of regional culture history and should rather focus on broader ecological and technological approaches that are geared towards the investigation of diversity. The use of new methods, such as analyses of residues, starches, phytoliths, aDNA, and isotopes of archaeobotanical and archaeozoological remains as well as use-wear of lithic artifacts, all of which are just beginning in Southeast Asia, can and will transform our understanding of human occupation of MSEA and especially of past subsistence practices and lifeways.

### 1.1 Diversity, styles, technologies, and communities

But first I consider briefly what evidence archaeologists might use to ascertain cultural diversity during the prehistoric period. 'Cultural diversity' may be construed in many ways, and several but not all dimensions should be archaeologically visible. The development of 'localised domains of identity' (Jonsson, this volume) is one aspect of diversity of interest to anthropological archaeologists. How to recognise such domains in the pre-literate period from archaeological evidence is neither simple nor straightforward, as different criteria for what constitutes diversity, different scales of analysis, different data, and different questions may result in more than one reconstruction of past social groupings (Hegmon 1998). Nevertheless, archaeologists explicitly or implicitly discuss such domains, when they discuss 'archaeological cultures'. Archaeologists often use geographically patterned variation in decorative styles in material culture such as pottery as a means to differentiate past cultures and help define social boundaries (as reviewed by Stark 1998:2). While not foolproof, variation in morphological and decorative style has been shown to coincide with perceived regional social boundaries in some ethnoarchaeological studies of acephalous agrarian societies (e.g., Graves 1994).

However, it is increasingly recognised that the study of decorative style is insufficient for identification of past social boundaries, and that technological style provides an enhanced avenue for archaeologists to document communities that share ways of life, social interaction systems, and bodies of knowledge (Stark 1998). Whereas decorative styles can span cultural boundaries through trade and imitation, technological styles endure through transmission of technological know-how across space and time via socially constructed learning frameworks. Through the detailed study of past technological systems, identification of 'communities of practice' (Wenger 1998) who shared 'specific ways of doing things' is becoming an important avenue by which archaeologists can distinguish past groupings of peoples (Hegmon 1998; Stark 1998).

If we apply a community of practice perspective to the study of past societies, the spread of agriculture can be seen as the spread of technological systems—systems of know-how and practice whereby biological resources are produced intentionally on land that,

without the intervention of agrarian technologies, would be sustaining natural resources that may be less desirable, useful, numerous, or concentrated for the societies residing on the land. There are of course many variant agricultural systems, each deserving of focused investigation. The systems of know-how are inevitably embedded in group choices and styles of agrarian practice within specific environmental contexts that shape decisions about which lands to cultivate, what kinds of implements to use, which crops can and cannot be grown on particular landscapes, and what crops are fundamental to a lifeway and which are peripheral.

Subsistence technology is a major structural element in organizing any group's way of life. One major problem for understanding past agrarian subsistence systems in MSEA is the paucity of direct archaeobotanical evidence for agriculture (Castillo and Fuller 2010). View One infers the presence and spread of agriculture primarily from indirect evidence such as decorative style on pottery and settlement locations near arable land. It is often assumed that the economy was based on rice, even when many sites and regions have no hard evidence for rice cultivation (e.g., Higham 2002:228). By framing rice agriculture as an economy rather than as a range of technological systems practiced by different communities of learning, View One ignores the different effects that various kinds of cultivation systems for rice or any crop would have had on the processes of spread and adoption of agriculture. Wet rice/dry rice, extensive/intensive, shifting/stationary field, monocrop/polycrop, these technological variants would have had very different implications and outcomes for labor organisation, demography, expandability, environmental impact, settlement system, and other aspects of society and environment that are of central importance to understanding the regional transition to agriculture.

It is particularly important not to conflate cultivation of dry upland rice with cultivation of inundated rice. The two cropping systems require different varieties of rice, and dryland rice generally would have required shifting cultivation in a multi-crop field, whereas wet rice at least in these latitudes was probably raised in stationary plots (White 1995). The investigations of DNA of modern and ancient rice (Fuller et al. 2010) are changing the picture of the palaeogeography of early rice cultivation, with *indica* and *japonica*, and wet and dry variants having a more complex history than the single homeland assumption allows. Although discussions of View One rarely state whether the expansionary rice-based economies are cultivating wet or dry rice, the demographic dynamics inferred, the focus on alluvial plain settlements and other attributes discussed are more consistent with an inference of wet rice.

This raises a key problem with finding evidence to support View Two expectations regarding diversification: societies practicing extensive polycultural agriculture, such as might be postulated for early agricultural groups in MSEA, often do not leave much for the archaeologist to find. Certainly reconstruction of subsistence systems by analyses of plant and animal remains, which in MSEA currently lags other parts of the world, will eventually help resolve the nature and details of the HG-AG transition in MSEA (Castillo and Fuller 2010). For the present, proponents of either view are required to infer subsistence technology largely from models, some faunal and other palaeoenvironmental data (e.g., Kealhofer 2002, White et al. 2004), settlement location, and the handful of archaeobotanical studies available from the region (Castillo and Fuller 2010).

Some of the newer data already available include palaeoenvironmental, technological, archaeobotanical, and settlement evidence. This paper briefly outlines some of the newer ecological and technological evidence pertinent to discussion of prehistoric cultural diversity in MSEA. The linguistic and human biological evidence is not discussed here, as these topics are addressed by appropriate specialists in other chapters in this volume.

Because technological choices are embedded in environmental contexts that provide both constraints and possibilities, the study of past technological systems is inseparable from environmental reconstruction. An appreciation of Southeast Asian environmental variability is needed to provide a contextual foundation for studying human cultural diversity in the region. Recent palaeoenvironmental data are transforming understanding of the profound changes in landscape, vegetation, and other natural resources that occurred in the late Quaternary of MSEA. These changes in biotic resources and environment must have affected human land use and subsistence. While much detailed research relating the new palaeoenvironmental evidence to archaeological data still needs to be undertaken, some preliminary insights are evident. Some of the new data undermine the parsimonious elegance of the View One hypothesis of diversity in MSEA.

## **2 Environmental background to diversity in Southeast Asia**

### **2.1 Inherent regional diversity**

The Southeast Asian environment has long been recognised as complex and diverse simply because of its tropical location. With their intense solar radiation and high rainfall, tropical landscapes are known for their high species diversity relative to more temperate latitudes. High species diversity means that subregions tend to have large numbers of resident species, but proportionally fewer individuals or concentrations of those species, in comparison with geographic areas with lower species diversity such as, for example, temperate latitude grasslands or pinelands.

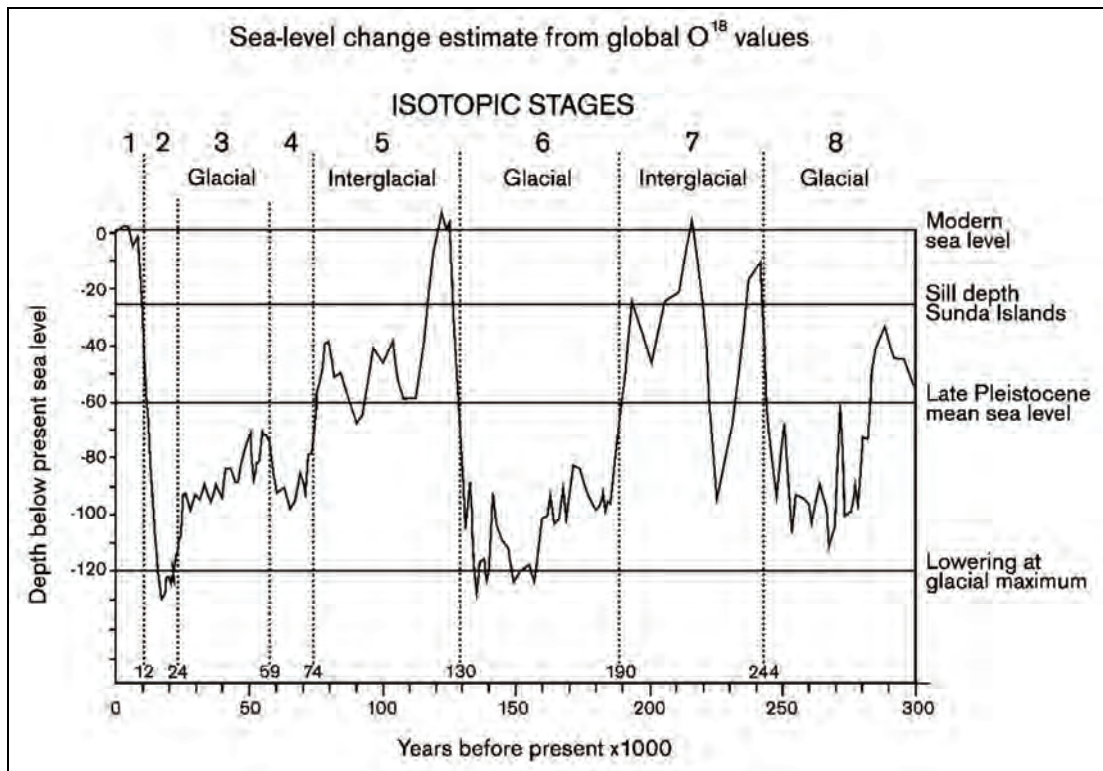
However, there are many other factors contributing to Southeast Asia's biodiversity besides its tropical latitude. The plate tectonics of the region have created a variety of contrasting landforms ranging from mountains as high as 3000 m in northern Myanmar to vast low-lying continental shelves currently underneath the South China Sea. A fan of major rivers drains the eastern Himalaya Mountains that formed from the collision of South Asia with the Eurasian plate. The hub of the fan lies in Yunnan where north to south trending drainages diverge and establish the closely interdigitating highlands and valleys for which Southeast Asian geography is known. The suite of rivers includes the Brahmaputra extending west to the Ganges, the Salween extending south, the Mekong crossing to the southeast across MSEA, and the Yangtze extending to the east across what is today the southern half of China. Other rivers in between these four major 'spokes' include the Irrawaddy, Chao Phraya, and Red Rivers. These riverine systems dissect and expose the underlying geological formations, including remnants of micro-continents dating as far back as the pre-Cambrian (Hutchison 2005). This complex geological and geomorphological history has provided Southeast Asia with a wealth of topographically and geomorphologically differentiated environments (Hutchison 2005) that ultimately provide a more complex natural geographic infrastructure than most other tropical landmasses (Gupta 2005:38).

An additional component of MSEA geographic complexity is the widespread carbonate (limestone) deposits, including 'some of the more extensive karst regions in the world' (Gillieson 2005:157). Differing greatly in age, including parts dating from the Paleozoic, the karst landforms exposed throughout much of MSEA vary greatly among themselves and include limestone plateaus, towers, and swamps. The karstic topography contributes to high regional rates of allopatric speciation (species formed from geographic separation of breeding populations; Gillieson 2005:172). Areas of extensive limestone exposure create isolated niches for biota, as the escarpments and cross-cutting river systems create physical

barriers for breeding populations of many species. The karstic microniches provided refuge areas for biota over periods of climate change from which genetic bottlenecks and founders' events occurred as climate trends fluctuated over time. Hence species in karstic environments often are highly adapted to the local habitat (endemism) and consequently have small ranges. Karsts, which in MSEA often have extensive cave systems, influence agricultural potential in a variety of ways, usually endowing the area with soils vulnerable to erosion and poor moisture retention.

## 2.2 Late Quaternary sea level change

Another outstanding geographic characteristic of Southeast Asia is the extensive and shallow Sunda continental shelf that now lies under the South China Sea and the Gulf of Thailand. Throughout most of the late Pleistocene, most of this continental shelf was exposed land, as sea level vacillated between 40 and 90 meters below its present level (mbpl., Figure 1; Hope 2005:28). The lower sea level meant that throughout the late Pleistocene, continental MSEA included Sumatra, Java, Borneo, and many other land areas that today are considered parts of island Southeast Asia (ISEA). The Southeast Asian post-LGM (last glacial maximum) sea level rise reduced the region's total land area by a third (Voris 2000:1155). More than two million square kilometers of lowlands were inundated, much of it alluvial plain (Sathiamurthy and Voris 2006:3). The rise in sea level reshaped the Southeast Asian landscape, creating several large and many small islands, dramatically expanding the coastline, and reducing the expanse of alluvial lowlands (Figure 2). Sea level change has contributed directly to the intrinsic environmental and biotic diversity of the Southeast Asian region through reconfiguring both terrestrial and aquatic communication corridors and barriers for humans and other species.



**Figure 1:** Diagram showing changes in sea level during late Quaternary global oxygen isotopic stages. Adapted from Hope 2005 and Martinson et al. 1987.

There is not yet full agreement on the exact dating of the LGM and the specific depth of the LGM sea level, but the general picture is well enough known for our purposes (Hanebuth et al. 2011). Figure 2 compiles current data from several sources on the changes in exposed land of the Sunda shelf since the LGM. The most significant changes in land configuration occurred at the Pleistocene/Holocene transition as the sea level passed the 50 mbpl mark, when areas such as the Gulf of Thailand and the Gulf of Tonkin became inundated. Borneo lost its land bridge to Sumatra and the mainland as early as 11,000 BP at roughly 30 mbpl, while Sumatra became disconnected from the Malay Peninsula during the early Holocene after c. 10,000 BP (Bird et al. 2004; Sathiamurthy and Voris 2006). Today's sea level is thought to have been reached only around 6000 BP.

Pollen and other data show that during the LGM, average temperatures may have been slightly cooler and humidity slightly lower than today, and lowland rain forests and lower montane rainforests covered the exposed shelf (Wang et al. 2009). Inundation, although reducing land area and alluvial habitats, dramatically expanded coastline habitats. Such periodic geographic changes in contiguous land mass and habitat configuration due to late Quaternary marine transgressions and regressions are thought to be another important contributing factor to high allopatric speciation in Southeast Asia and a resulting high regional biodiversity (Hanebuth et al. 2011; Sathiamurthy and Voris 2006:2). Human population density would also have been affected, especially in areas such as the Gulf of Tonkin where large areas of continental shelf were inundated.

### **2.3 Late Quaternary climate change**

In addition to the landscape changes brought about by rising sea levels, changes in climate following the LGM and their impact on the Southeast Asian biosphere are increasingly being documented (Cook et al. 2011; Wang et al. 1999). Changes in climate include not only rising temperatures and increased precipitation, but also seasonal concentration of precipitation during the summer from the strengthening of the Indian and East Asian monsoons. These climate changes had profound effects on habitat range and variability in MSEA.

In addition to changes in precipitation and temperature, atmospheric CO<sup>2</sup> concentrations changed (Maslin and Thomas 2003). The higher CO<sup>2</sup> levels during deglaciation (caused by release of methane as tropical wetlands expanded) favoured not only great increases in plant biomass, but the expansion of C3 plants at much higher rates than C4 plants (Maslin and Thomas 2003:1733). C3 and C4 plants differ in photosynthetic pathways, and C4 plants tend to thrive in drier contexts, C3 in moister contexts. The C3/C4 distinction is important to assessing availability of certain foods and in turn subsistence constraints and potentials; for example Job's tears and some millets are C4 plants, and rice and bamboo are C3 plants.

Changes in climate in concert with inundation of the Sunda shelf contributed to highly localised variation in environmental changes for MSEA. The seasonality and total amount of annual rainfall in any particular region has complex relationships with wind patterns, temperature, topography, and distance from oceans. Even today Southeast Asia can have dramatic changes in precipitation over relatively short distances. For example, interior parts of Myanmar can average around 1000 mm annual precipitation, and three to four hundred kilometers to the west at the same latitudes along the Bay of Bengal there are zones with more than 5000 mm. Figure 2 shows that the Bay of Bengal was less affected by post-LGM inundation than the South China Sea, where the inundation of the Gulf of Thailand would have drastically changed the exposure of Cambodia to summer monsoon

moisture systems. This point has been suggested to account for a delay in manifestation of Holocene climate patterns in Cambodia (White et al. 2004). As a result of the variability in the slope and width of the Sunda continental shelf, the effects of inundation on local moisture systems varied from place to place, which in turn further stimulated overall diversification in the regional biosphere. These climate changes must have affected resources available to human societies at very local levels.

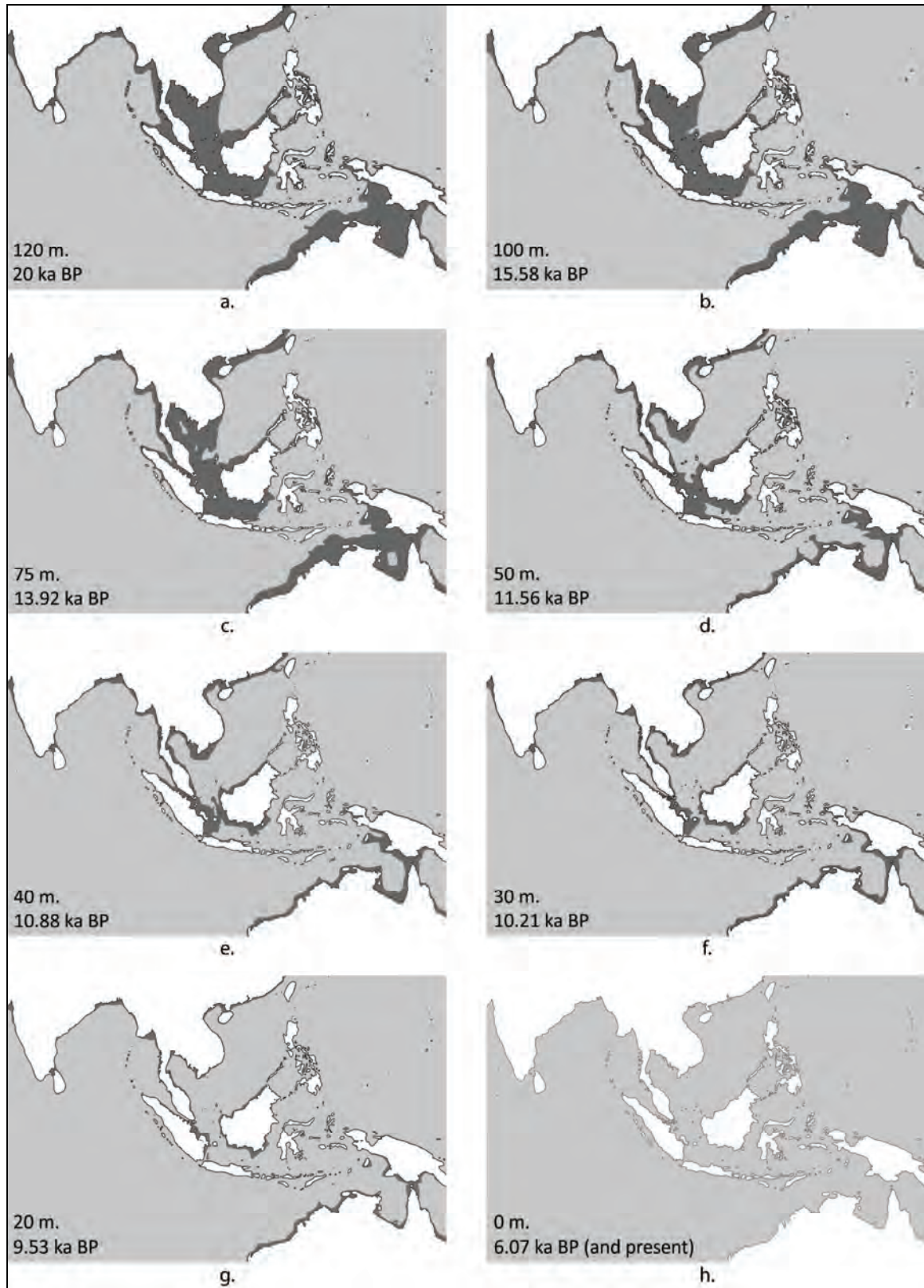
#### **2.4 Late Quaternary vegetation change**

Whereas many larger-scale aspects of the post LGM environmental changes are being clarified as new methodologies such as speleothem records (Wang et al. 2008) are brought into the discussion, the sub-regional details of late Quaternary vegetation changes in MSEA are only beginning to be documented. It is these local stories that are needed to flesh out the biotic resources and subsistence options the human inhabitants of particular regions had as environments of particular places changed over time.

Palaeopalynological and other palaeobiological records indicate that although the late Pleistocene climate in Southeast Asia was generally drier and cooler than today, the type and degree of habitat changes that specific regions experienced during the climate fluctuations were quite different. Northeast Thailand, for example, is one of the few interior areas of MSEA from which there is a palynological record extending back 40,000 years (Penny 2001). The sediment core from Nong Pa Kho indicates that pine/oak vegetation dominated this area during the late Pleistocene. At the beginning of the Holocene about 10,000 years ago, tropical broad leaf forest rapidly replaced the pine/oak forest. In contrast, a core from Phayao at a higher elevation in northern Thailand (Penny and Kealhofer 2005; White et al. 2004) does not reveal dramatic or rapid vegetation change at the Pleistocene/Holocene boundary, instead suggesting greater continuity. Sediment cores from Cambodia suggest a delay of about 1000 years there for Holocene changes in vegetation (Maxwell 1999, 2001).

#### **2.5 Implications of recent palaeoenvironmental research for MSEA archaeology**

There are several points archaeologists need to keep in mind when examining the palaeoenvironmental record in relationship to past human societies. First, the terminal Pleistocene to early Holocene in MSEA was a period of very lively and diverse habitat change, the details for which were highly localised. Second, late Quaternary changes in habitat and resource availability near individual archaeological sites therefore must be studied location by location. Archaeologists cannot necessarily generalise palaeoenvironmental evidence for habitat change in one location in Southeast Asia to another dated to the same time frame. Third, while in some regions habitat change may have been profound and widespread (for example, northeast Thailand) other areas may have had more stable habitats (parts of northern Thailand). Other areas such as northwest Borneo experienced proportional shifts in habitat expansion and contraction (Wurster et al. 2010).



**Figure 2:** Exposure of Sunda Shelf at several depths below modern sea level from the LGM. Inundation with depths and dates adapted from Hanebuth et al. 2011; Sathiamurthy and Voris 2006; Voris 2000.



### 3 Late Pleistocene Occupation by Modern Humans

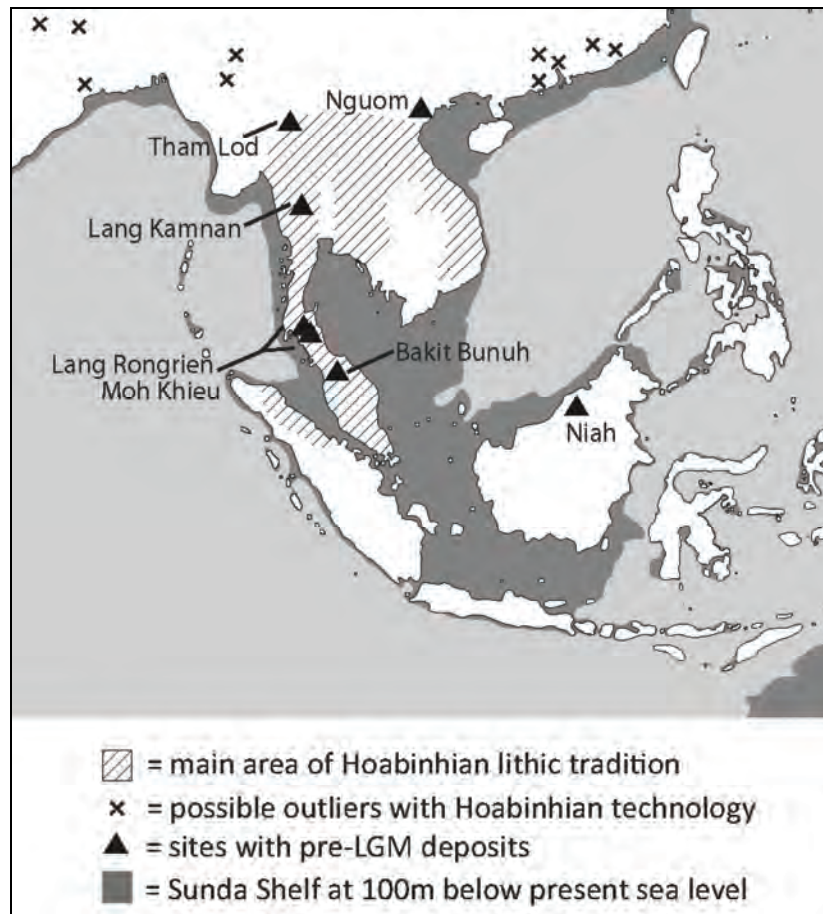
According to many current models for the appearance of *Homo sapiens sapiens* in MSEA (e.g., Bird et al. 2005; Pope and Terrell 2007), modern humans expanded into MSEA during a late Quaternary climate phase referred to as ‘Oxygen Isotope Stage 3’ or OIS3, which dates c. 59,000–24,000 yr BP. Sea level during this period fluctuated between roughly 70 m and 95 m below current levels (Hope 2005:28; see Figure 1). Throughout OIS3, the Sunda continental shelf was exposed and MSEA extended through to Borneo. [Note, although Palawan has evidence of human occupation during OIS3, no land bridge existed with the mainland during the last glacial period (Piper et al. 2011). Thus the late Quaternary archaeology of that island will not be considered here.] The major spread of humans may be further delimited to a relatively stable warm/wet interval between c. 47,000–37,000 yr BP that substantially overlaps with a period of swamp, estuarine, lagoon, and reef development, according to Pope and Terrell (2007:8). They (2007:11) ‘propose that the development of productive coastal ecosystems ... was a major factor in the rapid spread of modern humans from the west along a southern migration route in the interval 47,000–37,000 yr BP’. Others posit human expansion via a savannah corridor that reached from interior Asia into western Borneo (Bird et al. 2005; Wuster et al. 2010).

#### 3.1 OIS3 sites

Only a handful of archaeological sites have dates from OIS3 in MSEA (Fig. 3). The earliest include: (1) Lang Rongrien in southern Thailand (Unit 10 dating to about 43,000 BP and Unit 9 dating to about 37,000 BP re Mudar and Anderson 2007:303); (2) Niah Cave in northern Borneo, with several dates indicating early occupation c. 45,000–38,000 BP (Barker 2005; Barker et al. 2007); (3) Tham Lod in northern Thailand, which has a TL date of about 36,000 years ago from Area 3, Layer 4 (Shoocongdej 2006); and (4) Bukit Bunuh, an open air lithic workshop in Malaya with an OSL date of roughly 40,000 years old (Roberts et al. 2005; Saidin 2006).

Even if most Late Pleistocene MSEA sites now lie under several meters of water, these four excavated sites together demonstrate that the OIS3 humans in MSEA occupied a variety of habitats. Tham Lod in particular indicates that habitats at more than 600 m above modern sea level and hundreds of kilometers from coasts were occupied during the OIS3 in MSEA.

Evidence from these four OIS3 sites suggests that the occupants exhibit a diversity of hunter-gatherer technologies, mobility patterns, and subsistence strategies. Although the quantity and quality of archaeological data available varies among the different excavations, variation between the sites is clear. Regarding lithics from the OIS3 sites, Shoocongdej (2006:35) reports that the context producing the early date at Tham Lod contained only flakes, mostly of sandstone. Saidin (2006) reports a wide range of pebble tools, cores, flake tools, and debitage at the purported lithic workshop at Bukit Bunuh. Lang Rongrien’s layers 10 and 9 show a predominance of small irregular flake tools with some core tools made mostly from local chert (Anderson 1990). At Niah, in addition to stone flakes, a worked bone point was found in the OIS3 layers (Rabett et al. 2006:56). The overall evidence suggests that the inhabitants used expedient flake and core technologies lacking morphological types and probably used lithic resources close to sites (Rabett and Barker 2010:69).



**Figure 3:** Map showing pre-LGM sites mentioned in the text, Hoabinhian main range, and Hoabinhian outliers. Sunda Shelf exposure depicted at 100 mbpl. c.15,500 years ago.

Available faunal evidence from Niah and Lang Rongrien indicates that humans exploited a wide range of habitats and species during OIS3, but that species concentrations varied between sites and sub-periods. Although bones of large animals are present at both sites, medium and small animals were more common. The range of fauna exploited demonstrates that diverse foraging technologies were used in diverse habitats (arboreal, terrestrial, riverine, and more; Barker et al. 2007; Rabett et al. 2006). Rabett et al. (2006) note that the wide range of fauna and flora from Niah demonstrates the existence of a wide range of hunting and gathering skills and behaviours (Barker 2005), from mollusk collection and fishing to hunting of arboreal mammals such as monkeys and small mammals such as rodents and porcupine. However, terrestrial animals predominated, especially pig (Rabett et al. 2006). Mudar and Anderson (2007) found that Lang Rongrien Units 10 and 9 have mostly deer, bovids, turtles, and tortoise. The smaller range of mammals and absence of pig at Lang Rongrien are notable. They suggest that the types of faunal diversity at the two sites reflect differences in local habitat and settlement use. Niah may have been occupied for longer periods of time, more like an intermittent base camp, and the inhabitants must at least periodically have had access to patches of rainforest of sufficient size to sustain pig populations. Lang Rongrien may have been used more sporadically as a short term hunting camp, and remains of Eld's deer (*Cervus eldii*) indicate that dry dipterocarp forest and open plains were within the catchment zone. Evidence of butchery of both complete and partially dismembered carcasses around

hearths was found at both Niah and Lang Rongrien (Barker et al 2007; Mudar and Anderson 2007). No details of faunal remains from the deposits associated with the OIS3 dates at Tham Lod or Bukit Bunuh are reported, and Shoocondej (2006:34) states that faunal remains were rare and fragile in the Tham Lod Area 3 layer 4 deposit which produced the early date.

The most startling indications of OIS3 technological capabilities are the recent findings at Niah of evidence for tubers and nuts requiring extensive processing to remove toxins (Barton 2005; Barker et al. 2007). This evidence demonstrates that early anatomically modern humans in Southeast Asia practiced delayed-return behaviors requiring forward planning and resource processing over the course of days to weeks. Such planning in resource manipulation and consumption is not expected among immediate-return hunter-gatherer societies at the smallest scale, technologically-simplest, and most egalitarian end of the foraging spectrum. The findings of such plant processing suggest a more sophisticated range of social and technological behaviors than heretofore expected for humans in this region at this time depth.

### **3.2 Other pre-LGM sites**

Between 37,000 BP and roughly 20,000 BP (the LGM), a few more dated sites are known from MSEA. During this period sea level fluctuated around 95 mbpl until c. 24,000 BP when sea level lowering accelerated (OIS2) culminating in the LGM. During the LGM, continental shelves reached maximal exposure, alluvial plains were at their maximum extent, and the climate may have been the driest since 130,000 BP. Relict closed canopy rainforest patches likely were at their most reduced, and open habitats were at their most extensive (Wuster et al. 2010). This condition rapidly reversed after roughly 20,000 BP and in little more than 13,000 years, the sea rose about 120 meters to modern levels.

The regions where OIS3 sites have already been noted continue to have evidence of human use into OIS2. Niah (Barker et al. 2007) and Tham Lod (Shoocongdej 2006) have dated deposits showing continued use of areas of northern Borneo and northern Thailand into OIS2. Located near Lang Rongrien, Moh Khieu's Cultural Level 1 dates to about 25,000 BP (Pookajorn 1996), indicating continued human occupation in that area of southern Thailand. Other areas with pre-LGM dated sites include northern Vietnam where a flake-based technology, Nguomien, is associated with a date of older than 32,000 BP (Anisyutkin and Timofeyev 2006). In western Thailand, Lang Kamnan Cave has deposits dating from c. 27,000 BP (Shoocongdej 2010).

### **3.3 Late Pleistocene culture and society**

Although OIS3 evidence is too sparse to argue for the existence of 'localised domains of identity' in the data of late Pleistocene of MSEA, there is evidence of diversity in terms of habitat selection and technological choice. Variability among the OIS3 sites indicates variability in settlement location relative to habitat variation, in fauna exploited, and in lithic technology.

Barker et al. (2007) argue for a tactical interpretation of the OIS3 subsistence evidence whereby resource diversity was strategically targeted. They see OIS3 hunter-gatherers as:

...exploiting a diverse interior landscape using a battery of technologies that may have included mammal and fish trapping, some form of projectile technology, tuber digging, plant detoxification, and forest burning. The levels of resource use, forward planning, and ingenuity that would have been necessary for such strategies would not only parallel many of

the developments seen in Late Pleistocene records of Europe and Africa, but also serve to illustrate human adaptive plasticity with *the emergence in Southeast Asia of strategies directed specifically towards exploiting the structure and diversity of lowland tropical environments* (emphasis added; Barker et al. 2007:259).

Thus even though we do not know about human occupation of the now-submerged Sunda alluvial lowlands, the evidence that we do have of MSEA during the late Pleistocene shows that the flexibility and ‘situational shifting’ observed in more recent foraging groups in Southeast Asia may be a general characteristic of pre-agricultural populations in this area extending back tens of thousands of years (Junker 2002:164). The pre-LGM evidence from MSEA is consistent with local adaptations to diverse mosaic environments by generalist hunter-gatherers who could ‘habitat-tailor’ (Barker et al. 2007).

## **4 LGM to Early Holocene: Emergence of the Hoabinhian**

The drowning of the Sunda shelf between roughly 20,000 and 6,000 BP coincides with the emergence of a distinctive lithic technological tradition recognizable throughout MSEA that archaeologists call ‘Hoabinhian’. As the post-LGM environmental changes are more fully understood, the highly variable and dynamic environmental context in which this technological system developed is becoming clearer.

### **4.1 Summary of Post-LGM environmental changes**

Despite the inevitable climate wavers and events, the overall trend of rising sea levels after the LGM is clear (Figure 1), and ultimately the inundation transformed the extent of the land area, the character and configuration of MSEA, and the routes by land and water connecting one part of Southeast Asia to another. The vegetation on lands remaining above sea level must have been transformed in many areas as the monsoonal system took its present parameters and many stretches of previously dry inlands became coastal lands.

For the purposes of this paper, there are three general points to keep in mind about the post-LGM environment of MSEA. First, the post-LGM period to which Hoabinhian assemblages in Southeast Asia are usually attributed (terminal Pleistocene to early Holocene) was a period of great environmental change for the region. Palaeoenvironmental data currently indicate that when the sea level was low, vegetation formed a mosaic of drier, more open plant communities with refugia of denser wet tropical plant communities (Wurster et al. 2010). As climate changed with warming temperatures and increases in precipitation, vegetation was altered but the specifics of those alterations vary greatly across MSEA (White et al. 2004). As CO<sub>2</sub> concentrations, temperatures, and precipitation increased after the LGM, the wet tropical variants increased in areal extent and drier variants decreased, again with details locally contingent. Closing of tropical forest canopy at the lower latitudes would have favoured a shift at least proportionally from fauna requiring open expanses to fauna needing continuous tree cover or arboreal habitats. Second, total land area above sea level in Sundaland decreased by about half which in itself must have altered total regional population density. However, there would likely have been local variability in the demographic effects of decreasing land area depending in part on the width of the continental shelf. Areas with steeper, narrower continental shelves such as western Sumatra lost little land area while areas with wide gently sloped shelves such as the land now under the Gulf of Thailand lost large expanses. Third, overall, the areal extent of alluvial plains greatly decreased but coastline and hence coastal habitats

greatly increased. All these changes must have meant alterations in habitat ranges, distributions, and densities of biological resources, both plants and animals, available to the human occupants of the region.

#### **4.2 The Hoabinhian lithic tradition**

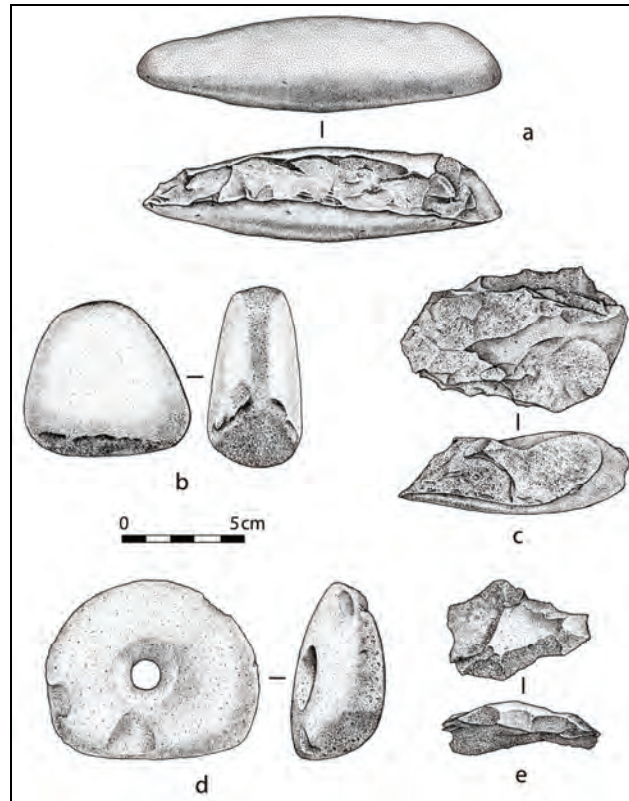
Most MSEA sites whose deposits can be dated to the post-LGM inundation of the Sunda shelf are attributed to the 'Hoabinhian'. Unfortunately, while the term 'Hoabinhian' is widely recognised and used, the phenomenon is ill-defined (vide Shoocongdej 2010). Reports from excavations are usually too vague in their descriptions to know if scholar A's Hoabinhian is truly comparable to scholar B's; descriptions of Hoabinhian lithic assemblages lack the details needed to concretely assess variability within the tradition. Moreover, late Pleistocene/early Holocene deposits may be assigned to the 'Hoabinhian culture complex' primarily on the basis of dates, even if no typical flaked cobbles or debitage are found (Rabett et al. 2011:163).

Nevertheless, Hoabinhian assemblages are widely recognised by the presence of highly characteristic flaked artifacts called 'sumatraliths'. Sumatraliths reveal a characteristic production sequence that begins with apparently careful selection of water worn river cobbles with several desirable qualities of size, shape, and material (Forestier 2000). The cobbles tend to be flattish ovoids of hard, often coarse-grained materials such as quartzite or andesite. One of the flattish surfaces typically is selected as the striking platform, from which the cobble is unifacially flaked around all or most of its perimeter (Figure 4a). The steep working edge produced is thus cortex (the water-worn surface) on one side with the flaked surface on the opposite side. The established perimeter can be reflaked along the same platform surface as needed to resharpen, trim, and rejuvenate the edge (White and Gorman 2004). In this manner, the length of edge obtainable from a river cobble can be maximised. 'Classic' Hoabinhian assemblages are also distinguished by the rarity of retouched flakes and the rarity of bifacial flaking on cobbles. Occasionally other kinds of lithic artifacts have been found in purportedly Hoabinhian assemblages, such as pecked stone rings (Figure 4d).

It appears that the flaked cobbles excavated from Hoabinhian deposits were locally acquired, and access to this raw material may have been one determinant of landscape choice by the users of this tradition. There is no clear evidence that non-local lithic resources were exchanged over significant distances during the post-LGM Hoabinhian.

Cobble beds form at locations along riverine systems that are broadly predictable from hydrodynamic models. Cobble-sized 'sediment' (defined as nodules 64 to 256 millimeters in diameter) is commonly transported by rivers and deposited along drainage basins according to parameters of water velocity, turbulence, river bed slope, and other factors. Sites with Hoabinhian lithics are often in caves and rock shelters close to or within couple of kilometers from rivers and extant cobble beds. The heavy cobbles, weighing as much as a kilogram, were thus carried, often up steep inclines, to these caves or rock shelters. The activities for which they were used there apparently often required considerable on-site resharpening and modification of the cobbles, as high rates of debitage recovery at some Hoabinhian sites have been noted (Corvinus 2007:336–9). The repeated circumferential reduction sequence produces a characteristic Hoabinhian flake debitage with high prevalence of flat flakes with cortical striking platforms (Figure 5). On larger flakes, it is common to observe step flake scars or an 'overhang' of small flake scars along the flake's striking platform (Figure 5e–g). The very density of lithic debris in many Hoabinhian

deposits is one of the characteristics that can differentiate them from pre-LGM deposits (Anderson 2005).



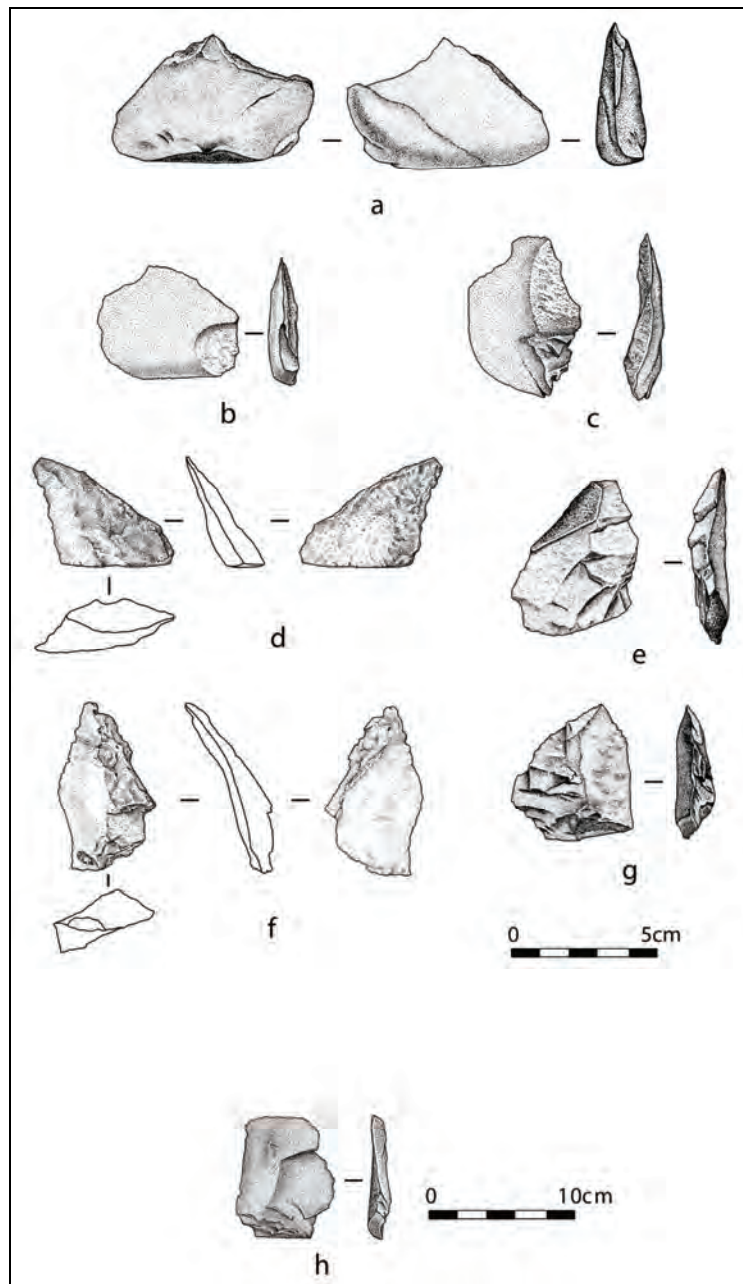
**Figure 4:** Examples of Hoabinhian lithics from Tham Phaa Can, a rock shelter site in northern Thailand.

- a. side view of river cobble depicted before and after initial unifacial flaking;
- b. 'hammerstone';
- c. sumatralith, top and side views;
- d. 'donut' stone;
- e. dorsal flake removal probably from sumatralith; (White and Gorman 2004).

The circumferentially flaked ovoid river cobbles with characteristic reduction debitage seem to define a widespread technological tradition (White and Gorman 2004) that does, however, show some regional variation (Anderson 1990:47). While morphological variation of flaked cobbles is evident, it is widely agreed that a typological approach to classifying morphological variants is inappropriate, since the cobbles are in an ongoing state of reduction and modification (Marwick 2008a; Pawlik 2009; White and Gorman 2004). Yet Southeast Asian archaeologists have not established a systematic methodology to describe full lithic assemblages that have been grouped under the Hoabinhian label. The lack of a systematic recording framework prevents meaningful comparisons and assessments of intra- and inter-assemblage variability.

Nonetheless, review of the various attempts at describing Hoabinhian assemblages reveals that there probably exists significant variation even if its meaning is still far from clear (Nishimura 2005). In southern Thailand and the Malay Peninsula, for example, bifacial flaking is noted as common (Bulbeck this volume; Anderson 1990:47), and slaty rock is frequently used. Splitting a cobble before unifacial flaking has been noted in northern Thailand (Forestier et al. 2008). Variability in intensity of reduction is evident in some debitage studies (Marwick 2008a; Nishimura 2005). Marwick (2008b) proposes that

reduction intensity is related to distance from raw material, with more reduction at sites further from cobble beds in an effort to conserve resources. Variation in type of stone also seems to be related to availability of local raw materials. Further effort to systematically describe the flaked river cobbles may yet reveal important variants within reduction sequences, or perhaps functional variants. As one example from personal experience, I have observed at sites in northern Laos flaked cobble variants similar to Kamminga's (2007) 'flaked unifaces with strongly convex cortex surface' from Sai Yok in western Thailand. In these examples, the knappers did not choose a flat side of the cobble as the striking platform, but rather retained a markedly curved surface on the unflaked side of the implement.



**Figure 5:** Hoabinhian debitage from Tham Phaa Can  
 a.–c. decortication flakes;  
 d.–h. trimming and rejuvenation flakes (White and Gorman 2004).

### 4.3 Geographic distribution of the Hoabinhian

The currently known geographic distribution of Hoabinhian assemblages is depicted in Figure 3. Despite the poor descriptions of individual assemblages available, archaeologists generally agree that assemblages with the characteristics noted above are widely found in cave/rock shelters and some open air sites across modern MSEA. More than 120 sites are known from northern Vietnam, especially in the Ma River basin, and many sites are found in a broad band extending to the west across northern Laos and northern Thailand into the eastern half of Myanmar (Burma). The Hoabinhian industry also extends south along the Annamite Cordillera of Laos and Vietnam. It is found in western Thailand, southern Thailand, the Malay Peninsula, and the northern half of Sumatra. Known sites are usually in hilly, especially karstic landscapes, although this distribution may reflect a sample bias of archaeological research at cave sites. Hoabinhian lithics have not been found in interior parts of large, relatively flat alluvial basins like the Chao Phraya in central Thailand or major tributaries to the Mekong in northeast Thailand and Cambodia. But the industry is found in upland fringes around these basins in eastern Thailand and northwest Cambodia. This distribution within MSEA may reflect the geomorphological distribution and availability of typical source of raw material for Hoabinhian tools—water worn river cobbles—and their rarity and even absence in large relatively flat alluviated basins.

The outer geographical reaches of the Hoabinhian technological tradition are not entirely clear. The outermost assemblages often co-occur at the same or nearby sites with technologically distinctive lithic industries, which implies that Hoabinhian technology came into contact with other lithic traditions. To the north of the main expression of Hoabinhian, possibly related assemblages have been noted in southern China in Yunnan, Guangdong and as far east as Guangxi (Rispoli 2007).

It has been suggested that sites with Hoabinhian lithics can be found as far west as South Asia, often in association with ‘tropical monsoon forests’ along the sub-Himalaya hills. In northeast India, in the Garo hills of Meghalaya, Sharma (1984, 1990) identifies ‘typical Hoabinhian’ with flaked river cobbles of dolerite at sites such as Rongram Alagiri. Mohanty et al. (1997:174) claim that Hoabinian tradition tools have been found in eastern India in the Myurbhanj district of Orissa. Mohanty (1993) notes the unexpected presence of heavy duty tools, principally unifacially flaked dolerite river cobbles, at 39 sites. These implements are in assemblages that are undated but considered ‘Mesolithic’, often dominated by microliths made principally of chert.

In Nepal, a Hoabinhian-like lithic industry dating to earlier than 7500 BP calibrated has been identified at Patu in the Ratu Khola area of eastern Nepal (Corvinus 2007:254, 327). Patu lies in a monsoonal sub-tropical forest belt in the foothills of the Himalayas. Tools are generally unifacially flaked, flattened river cobbles of quartzite, but include some regionally distinctive elements (Corvinus 2007:336–9). The assemblage is more than 95% waste flakes, usually with cortical platforms, and includes typical ‘rejuvenation’ flakes as described by White and Gorman (2004). The heavy duty component of the Patu assemblage is considered quite different from most other contemporaneous South Asian subcontinent and western Nepalese assemblages (Corvinus 2007:285), which tend toward a microlithic tradition. Based on extensive experimentation and observations of gloss, the steep-edged implements are considered suitable for working wood and bamboo (Corvinus 2007). Even further west, Gaillard et al. (2011) suggest that a Hoabinhian-related cobble-based industry can be found in a monsoon forest belt along the southern fringes of the Himalayas in the western Siwaliks.



To the south, shell midden sites with Hoabinhian tools are known along the northwest coast of Sumatra. The cave site Togi Naruwa, on the Sumatran island of Nias (Forestier et al. 2005), is one of the southernmost sites with typical sumatraliths in deposits dating to around 8600 BP (uncalibrated). The Sumatran Hoabinhian is the clearest case for the occurrence of this lithic tradition in island Southeast Asia, and transmission of this technology to the island may have occurred while the Sumatran/Malay Peninsula land bridge still existed, or the Straits of Malacca were still an easily traversed estuarine landscape. Hoabinhian technology has also been claimed for southeast Australia, New Guinea and Japan (Bowdler 2006, 2008; Matthews 1966), but these possibilities would have involved long distance sea-faring and will not be considered here.

#### **4.4 Dating the Hoabinhian**

Based on radiocarbon dates from many sites, the Hoabinhian is commonly dated to the terminal Pleistocene and early Holocene. However, how early it appeared is unclear, in part due to the aforementioned lack of commonly employed standards for the description of MSEA stone age assemblages. There is a tendency to lump together all flake and core assemblages dating from pre-LGM to terminal Pleistocene as Hoabinhian, even if no typical sumatraliths are noted.

Nevertheless, Hoabinhian technology appears to have emerged out of MSEA's pre-LGM more generalised flake and core lithic traditions. In Vietnam, an intermediary tradition called *Son Vi* is dated from the LGM to the end of the Pleistocene, but *Son Vi* as a distinct entity is not recognised outside of northern Vietnam. The interface between a more generalised flake and core tradition and the more focused and systematic lithic tradition emphasizing unifacially and circumferentially flaked ovoid river cobbles is an important issue to be resolved by regional archaeologists. Clarification of this issue is also important for documenting the transmission of the Hoabinhian technology across MSEA as well as to peripheral areas to the north and west.

#### **4.5 Some suggestions on cultural configurations of the Hoabinhian**

The discussion of the Hoabinhian above portrays it as a distinct, enduring, lithic technological tradition with roughly definable spatial and temporal ranges, even if the edges in time and space for this tradition are still fuzzy. We can say from the environmental discussion above that the lithic tradition appears to have been a technology shared across a wide range of environments, tropical to subtropical, uplands to coastal. Although it emerged during a period of great environmental change in many parts of MSEA, the technology is found both in areas showing such change (e.g., Krabi and Trang provinces in southern Thailand) and those that are environmentally more stable (Phayao/Maehongson, Thailand). It is found in sites with differing fauna, at sites of different altitudes, latitudes, and habitats, and in both open air (including shell midden) and cave/rock shelter sites. Some sites may have been seasonally or sporadically occupied, but longer term occupations have been suggested for some locations (Anderson 1997:624). A variety of mobility patterns are in evidence (Rabett and Barker 2010, Shoocongdej 2010). The technology does not seem to have been restricted to particular subsistence or social regimes (in the sense that variation in mobility behavior may correlate with variation in group size, sharing patterns, social investment in delayed-return as opposed to immediate return behaviors, and degree of investment in specific locations in a landscape). In view of the broad expanse over time and space of the Hoabinhian technology, most archaeologists

do not consider the Hoabinhian as a 'culture', but rather consider it a 'techno-complex' or an 'industry' (Gorman 1972; Pookajorn 1990:25; White and Gorman 2004:437). Recently, Rabett et al. (2011) have used the phrase 'Hoabinhian cultural complex'. Here I will consider it a 'lithic technological tradition'.

The expansive reach of Hoabinhian technology over great distances and its endurance over thousands of years does deserve explanation. One view might explain the distribution of the technology in terms of migrations of users, and attribute its endurance to backwardness or 'conservatism' of its users. While movement of populations using Hoabinhian technology cannot be discounted, neither is migration a necessary cause of the widespread dispersal of the technological tradition. Alternatively, the technological tradition may represent an easily learned manufacturing process on a natural resource (river cobbles) widely available in a region rich in rivers draining uplands. A product that is quickly made and easily maintained from abundant local resources and is also highly suited for tasks commonly needed could in theory be transmitted via loose interacting networks of practitioners/learners within the zone of abundance of raw materials. The endurance of the Hoabinhian technology probably represents a highly successful and thus sustainable appropriate technology.

But appropriate for what? What were the uses that spanned thousands of kilometers, thousands of years, and so many environmental contexts? The main other common denominator of the Hoabinhian, in addition to being a definable tradition of selecting and uniaxially reducing river cobbles, seems to be the presumption that the flaked cobbles were extraction tools mainly used to work 'wood', or more specifically bamboo (Bannanurag 1988; Corvinus 2007; Forestier 2003).

Let us postulate that a flaked river cobble/bamboo technological system was the basic Hoabinhian 'techno-complex' (Gorman 1972). The flaked river cobble reduction system appears to have been a shared technology over large regions, suggesting that it was an easily learned and readily transmissible technological system. This technological system was used with diverse subsistence strategies in varied environmental zones in Southeast Asia including upland/karstic and coastal (Gorman 1971). Judging from associated floral and faunal remains, Hoabinhian assemblages show no single subsistence or food resource orientation (Pyramarn 1989; Nguyen Viet 2008; Yen 1977). In other words, there is no reason to infer that users of Hoabinhian tools belonged to a single 'culture', or had a single cultural configuration. The variety of site placements, lengths of occupations, and variety of resources exploited suggest the existence of numerous different cultures, as one would expect from this large and ecologically diverse area. This cultural diversity has been masked by the widespread uniformity of the lithic tradition. I argue that this uniformity in lithics is due largely to these typical Hoabinhian tools being used to exploit an important resource—bamboo—that was not only widespread throughout the Hoabinhian tool tradition area but was also probably relied upon by all the cultural groups, no matter how diverse they were in other respects.

There are many reasons to suggest that Hoabinhian flaked cobbles are ideally suited to bamboo exploitation. The utility of bamboo is of course not limited to particular subsistence regimes. As an industrial material and edible plant, bamboo is as useful for hunter-gatherers as it is for agriculturalists or fishers. But the implications of the probable bamboo focus of users of Hoabinhian lithic technology have not been fully explored.

#### 4.6 Bamboo, the resource

A postulated focus on bamboo exploitation for the Hoabinhian flaked cobble technological system has implications for: 1) the emergence of ‘localised domains of identity’ expressed in part via localised bamboo technological traditions, and 2) the effects of focused and intensified bamboo exploitation on regional ecology and subsistence practices. Of course this proposal that Hoabinhian technology was intimately entwined with bamboo exploitation needs more directed research, including application of suitable use-wear studies, experimental studies, and residue analyses to see if the functions of Hoabinhian lithic technology can be empirically identified (Bar-Yosef et al. 2011). In addition, larger implications for systematic exploitation of bamboo need to be considered, if the study of the Hoabinhian period is to advance. Below I briefly propose some implications of the possibility that the Hoabinhian lithic technology fundamentally represents an intensified use of bamboo.

Bamboo has long been recognized as an extremely important resource for Southeast Asian societies. According to Dransfield and Widjaja (1995:15), there are 200 species from 80 genera of bamboo in Southeast Asia, which is considered a main region of bamboo genetic diversity for Asia (Dransfield and Widjaja 1995:16). While bamboos are found in a wide range of habitats from the alluvial plains to mountains, it is most prevalent in low to medium elevations in the tropics and subtropics (elevations comparable to where cobble beds along rivers commonly can be found).

Botanically bamboos are grasses, not trees, which makes their material properties quite different from wood. The most important attributes for human use are that: a) the stems are hollow but rigid so that they cannot be easily bent unless split; thus bamboo lengths can be as strong as a tree stem of similar dimensions but be much lighter in weight; b) the stems or culms have nodes (horizontal partitions) and hollow internodes, whose spacing can be exploited in various ways; for example a single internode with node can be used as a container; c) soft shoots are edible and in some seasons contribute significantly to local human diets; d) from their subterranean rhizomatous base, bamboo grows and regenerates quickly, so that a year-old culm can be used for handicrafts, and a three-year-old culm may be ready for use as building material.

A summary of key bamboo uses (depending on species) include: a) food (shoots and seeds); b) building construction materials, including thick bamboo culms for house posts, split and flattened culms for walls, floors, and roofing; c) containers using individual internodes for storing, transporting, and cooking; d) containers and traps using split lengths woven into baskets for storage and for trapping fish; e) musical instruments; f) hand tools including weapons (blow pipes, projectiles, and sharp long knives; West and Louys 2007); and g) water transportation by strapping several culms together into rafts as the hollow but sealed stems float.

One bamboo type cannot fulfill all these uses. Different bamboo species have different material qualities that affect their performance characteristics as items of material culture. Variations in culm width, length of internodes, fiber length, node and wall thickness, silica content, and other attributes mean that specific bamboos are exploited for specific properties. For example, the large internodes of *Schizostachyum brachycladum* enable it to be used for cooking, but species with long and flexible fibers such as *Bambusa atra* make this bamboo suitable to weave into baskets (Dransfield and Widjaja 1995:21–22). Bar-Yosef et al. (2011) also found that differing qualities of species and clumps affected ease of exploitation.

To harvest and prepare bamboo for use, two main activities are undertaken: *chopping* across the fibrous stem to detach the culm from the clump or rhizome or to detach a section of a culm, and *splitting* when a longitudinal break parallel to the fibers is required (Bar-Yosef et al. 2011). A single split along the length of a culm section enables the culm to be spread out and flattened and used as a surface such as a floor board. Detachment of slender splits can produce bamboo knives or strands that can be plaited.

The main tool needed to exploit bamboo would be a heavy hard implement with a non-brittle working edge that could cut or chop across the fibrous and silica-rich culm. Hoabinhian flaked cobbles appear eminently suitable for such a task. Bar-Yosef et al. (2011) conducted experiments that demonstrated the utility of unifacially-flaked river cobbles for detaching culms from bamboo stands. Longitudinal splits might require an implement with a steeper edge angle but with a suitable weight (or a flake and cobble can be used together; Bar-Yosef et al. 2011) depending on the size of the culm and the desired thinness of the strand, if detaching pliable strands to make plaited artifacts. Although woods, rattans, and other plant products were likely exploited by users of a Hoabinhian toolkit, the modest labor involved in exploiting hollow bamboo stems in comparison to solid trees, the astonishing versatility presented by bamboo compared to other plants such as rattan, and the rapidity by which bamboo regenerates, all suggest that when available, bamboo was more intensively exploited than solid tree wood or other industrial plant resources.

How available was bamboo to users of Hoabinhian technology? Figure 6 shows that Asiatic woody bamboos have a wide modern day distribution from South Asia to southeast China. However, the centre in Asia of greatest bamboo diversity appears to be mainland Southeast Asia (Dransfield and Widjaja 1995:16; Saha and How 2001:659). The coincidence of the region of high bamboo diversity with the main area of Hoabinhian technology is striking, but also raises the possibility of a ‘chicken and egg scenario’. Users of Hoabinhian technology may have intensively exploited many species of wild bamboo that emerged via general processes of allopatric speciation during the fluctuating environmental conditions in Southeast Asia over millions of years. On the other hand, it is also possible that the genetic diversity of bamboo in Southeast Asia could be at least in part a byproduct of long term human use and manipulation of bamboos in this region, perhaps a co-evolutionary phenomenon. Both human and non-human factors demonstrably contributed to bamboo diversification, but we now expand on the former.

#### **4.7 Bamboo and horticulture**

There is an aspect of bamboo ecology that has not been discussed by archaeologists but was likely very important to bamboo exploitation, namely that disturbance (human or non-human) stimulates its growth (Dransfield and Widjaja 1995:16). ‘Systematic and regular exploitation increases the production of the bamboo stock’ (Dransfield and Widjaja 1995:41). Natural and anthropogenic habitat disturbance not directly related to bamboo exploitation are likely to have had an unintended consequence of stimulating bamboo expansion. Invasive and fast-growing, some bamboos have strong weedy tendencies, are early colonisers of disturbed forests, or are significant in forest succession sequences (Christanty et al. 1997). Younger rhizomes are particularly productive of new shoots. These inherent expansionary characteristics facilitate bamboo propagation.

Planting a cutting of rhizome stock, a technique termed *vegeculture*, is considered the easiest bamboo propagation method, quickly producing new young clumps. One year after transplanting a section of bamboo rhizome, new culms appear, in four years a mature

clump is established, and the patch can continue to produce culms for decades. Bar-Yosef et al. (2011) furthermore demonstrate that it can be more difficult to extract a culm from older, long-established denser clumps in comparison to clumps with fewer culms. Thus establishing new clumps may save total effort at harvest time by creating new stands with easily harvested culms.

These observations lead to a key point, namely that regular exploitation of bamboo likely entailed horticultural and other delayed-return behaviours. Horticulture involves cultivation of plants as individuals, often in mixed plantings of several species (in contrast to aggregate seed cropping methods which are commonly employed in field monocropping of small-seeded cereals). Horticultural subsistence strategies are common in tropical landscapes, partly because the cultivator can take advantage of both natural and man-made micro-niches to position plants for conditions suitable to individual cultivars. Mixed plots can include both vegeticulture and seed propagation. Diversity of environment and diversity of cultivars are both intrinsic to horticultural technology.

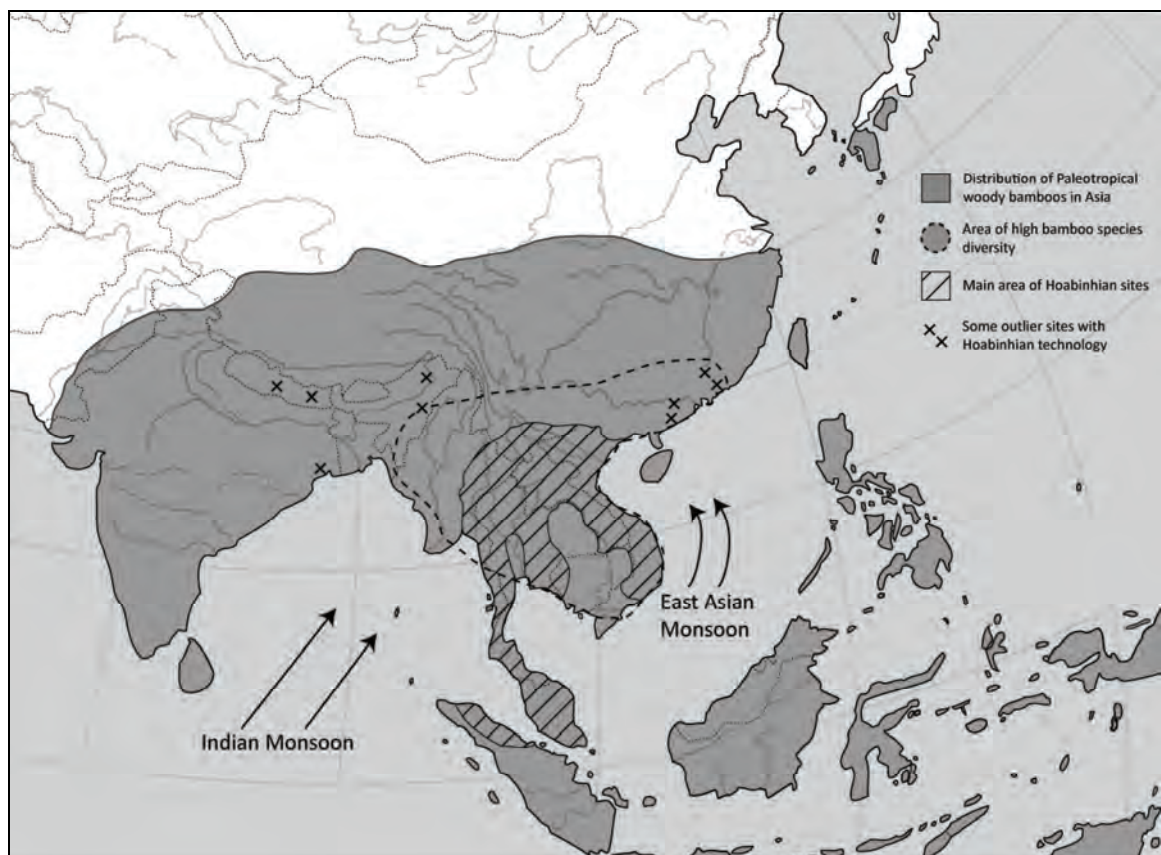
One further implication of the ease of regenerating bamboo stock via vegeticultural techniques is that it is a very small step to manipulating other plant resources such as yams. A society accustomed to the quick replanting of a bamboo rhizome anticipating return of new shoots in several months would have no difficulty in understanding the usefulness of replanting the top of a harvested yam for an anticipated return in the coming year. We have no direct paleobotanical evidence to support this, but then, no one has systematically looked for it, either. However, some sediment cores do show evidence of forest disturbance during the Hoabinhian period, which has been suggested to indicate some degree of human management of the environment during the early Holocene (e.g., Kealhofer 2002, 2003).

Exploitation of bamboo also has potential implications for exploitation of some fauna. While large mammals were exploited by the users of a Hoabinhian tool kit, only rarely are they prominent in Hoabinhian faunal assemblages (Higham 1977). Small to medium terrestrial fauna are more characteristic of Hoabinhian sites. Straightforward use of bamboo in creating projectiles and simple traps can be postulated.

However, mollusc shells are common in many Hoabinhian deposits and studies indicate that molluscs were likely transported in quantity to sites where they were consumed (e.g., Rabett et al. 2011). While transportation of medium to large fauna (large package resources) does not generally require more than one or two pairs of hands, perhaps a carrying pole and some rattan twine, and a strong back, a focus on small scale resources that are collected in aggregates—for example snails, small fish, and potentially seeds and other plant resources—implies the use of containers to collect and transport the resources to settlements such as caves, as only very small quantities can be transported any distance in one's hands. The possible importance of plaited bamboo artifacts as containers and traps for at least some users of Hoabinhian lithic technology has implications for emergence of technological styles and cultural diversity in the sense of 'localised domains of identity'.

Transmission of technological systems over space and time is related to learning frameworks, including social networks and the amount of training needed to acquire the requisite skills. The Hoabinhian flaked river cobble reduction system appears to have been a shared technology over large regions, suggesting that it was an easily learned and readily transmissible technological system. However, once a society relies on plaiting artifacts such as containers, the learning framework and time frame required for practitioners to gain aptitude and fine motor skills needed to consistently construct such artifacts inevitably expands. I suggest that if plaited artifacts that usually have specific functions, design, and performance characteristics—e.g., to trap a particular species of fish in a specific season

with specific water conditions, to winnow husked rice, to carry loads of particular weight several kilometers—are made, then the learning curve to replicate reliable artifacts that may be used over the course of months or years takes more time, and requires more direct instruction than the manufacture of simple bamboo knives and projectiles. Knowledgeable practitioners must train apprentices how to select and prepare appropriate raw materials, and how to manufacture durable goods with specific performance characteristics. Apprentices must practice to gain the motor skills needed to reliably reproduce these artifacts. As a corollary basketry learning frameworks tend to result in differentiation in both technological and aesthetic domains: ‘...it appears to be an established fact that no two populations ever manufacture basketry in precisely the same fashion’ (Silvestre 1994:199). Even in ethnographic contexts in Southeast Asia today, basketry and fish trap technology tend to show ethnically distinguishable forms and design characteristics (Baird and Shoemaker 2008:207).



**Figure 6:** Extent of Hoabinhian technological tradition relative to distribution of Asiatic woody bamboo, area of high bamboo diversity, and summer Asian monsoons. Compiled from map3 in *Bamboo Biodiversity* <http://www.eob.iastate.edu/research/bamboo/index.html> (Saha and Howe 2001).

Societies employing Hoabinhian technology did not necessarily plait bamboo artifacts. Bamboo can be exploited in many ways simply and directly with little stylistic development, technological or decorative. However, once a society commits to a material culture involving plaiting, the level of investment in durable goods suggests a corresponding investment in delayed-return activities. The stage is set for developing an associated learning community with identifying material correlates. Furthermore, the exploitation of specific bamboos for specific plaiting qualities may stimulate exchange

among groups with differential access to the raw materials and/or finished artifacts made from specific bamboo materials. The combination of centripetal activities transmitting technological knowledge within a group, with centrifugal activities such as exchange for necessary or exotic materials between groups creates social contexts from which social boundaries emerge (Welsch and Terrell 1998).

The proposal that Hoabinhian technology represents intensified exploitation of bamboo primarily as an industrial product is, at this stage, little more than a suggestion; much research will be needed to further examine this hypothesis with empirical data. And much research is needed to confirm or refute the proposed implications that investment in bamboo material culture contributed to specific learning networks, technological traditions, horticultural activities, and, as a further consequence, constellation and diversification of cultural identities. If these suggestions lead to more thorough examination of the ‘Bamboo hypothesis’ (Bar-Yosef et al. 2011), the proposals will have accomplished their objective.

## **5 Post-Hoabinhian Diversity**

### **5.1 The Middle Holocene: the missing millennia**

During the middle Holocene (roughly 6000–3500 BP uncalibrated, about 5000–2000 BCE calibrated), the archaeological record in MSEA is particularly sparse. This major gap in archaeological evidence for the region has been called ‘the missing millennia’ (White and Bouasisengpaseuth 2008:39). Judging from what is known archaeologically about the periods from before 5000 BCE and after 2000 BCE, seminal changes emerged during this gap in the archaeological record. Ceramics appeared in many parts of Southeast Asia; domesticated foodstuffs including millet and rice appeared; tool technologies changed with lithic tools transitioning from predominantly flaked to predominantly ground stone tools. At some point in this period, settlement systems changed focus. Locations of known settlements expanded from primarily karstic upland and estuarine landscapes during the early Holocene to include inland alluvial lowland villages by the late Holocene. Caves declined in importance for habitation (Anderson 1997, 2005). The upland/lowland dichotomy said to characterise Southeast Asian agrarian lifeways may have emerged.

Despite the dearth of archaeological evidence in MSEA for the middle Holocene, the *prima facie* case for cultural diversity strengthens, even though the nature of the transition between societies relying on a Hoabinhian lithic tradition and successor societies is not yet clear. However, a maritime/lakeside adaptation involving sedentary communities appears in some coastal areas, including in northern Vietnam. Da But in the Ma valley (Figure 7) is the best documented of the middle Holocene maritime/lakeside-oriented societies (see review in Nguyen 2005), but some less well known sites in southern Thailand (Srisuchat 1989) with similar dates suggest that this subsistence orientation emerged in other coastal areas of MSEA as well, which is not surprising given the great expansion of coastal habitats with the flooding of the Sunda Shelf. At Da But sites, which date generally c. 5500 BCE cal–3500 BCE cal, evidence for ceramic vessels, ground stone tools, a sedentary lifestyle in villages oriented toward exploiting resources of lakes and coastal swamps, and the practice of residential burial of the deceased suggests that some MSEA communities chose a distinct change in lifeways from earlier in the Holocene (Bui Vinh 1991; Nguyen 2005). It is possible that such an adaptation existed earlier in the Holocene, but marine transgression has prevented archaeologists from finding earlier maritime-oriented sites in MSEA. Nguyen (2005:91), however, describes the emergence of Da But as a local development derivative of the local Hoabinhian.



**Figure 7:** Holocene sites mentioned in the text.

Evidence for continuation of the Hoabinhian lithic tradition is found in interior cave sites with dates in the middle Holocene in northern and western Thailand (e.g., Pookajorn 1990; Reynolds 1992; White and Gorman 2004). However, deposits from some inland caves/rock shelters dating to the middle Holocene suggest that ceramics and polished stone technologies were appearing in some otherwise Hoabinhian contexts as well. For example, at Banyan Valley Cave, sherds are in clear association with Hoabinhian lithics in deposits dating approximately 4200–2000 BCE cal, and at the end of the sequence edge-grinding appears (Reynolds 1992). The cultural origins of the earliest ceramics and tool grinding technologies found in interior MSEA sites remain to be investigated. It is unclear if these new technologies are appearing due primarily to culture contact via trade with non-Hoabinhian, perhaps settled societies that may have lived in the north, or if population movements are introducing the new technologies. The answers will require identifying and excavating sites dating to the middle Holocene.

The reasons behind an overall decline in archaeological evidence during the middle Holocene as suggested in Marwick (2008c:13) and noted for most of MSEA (White et al. 2004:127, White and Bouasisengpaseuth 2008:39) are not clear. Possibilities include regional population decline or a shift in settlement systems away from caves and rock



shelters towards more open air contexts that have so far eluded archaeological research. Archaeological evidence from these ‘missing millennia’ to resolve this puzzle is surely a priority for the coming decade of archaeological research in MSEA.

## 5.2 Late Holocene settled societies

By the very end of the middle Holocene in late third millennium BCE, mixed mortuary/occupation sites of settled societies that cultivated plants and raised domestic stock appear in several interior parts of MSEA. Example sites and cultures with late third millennium calibrated dates include Ywa Gon Gyi in the Samon Valley of Myanmar/Burma (Pautreau and Maitay 2009), Non Pa Wai in the Khao Wong Prachan Valley of central Thailand (Natapintu 1991:154; Weber et al. 2010), Ban Kao in west central Thailand (extrapolating from dates in Leong Sau Heng 1991), and Ban Chiang in northeast Thailand (White 2008).

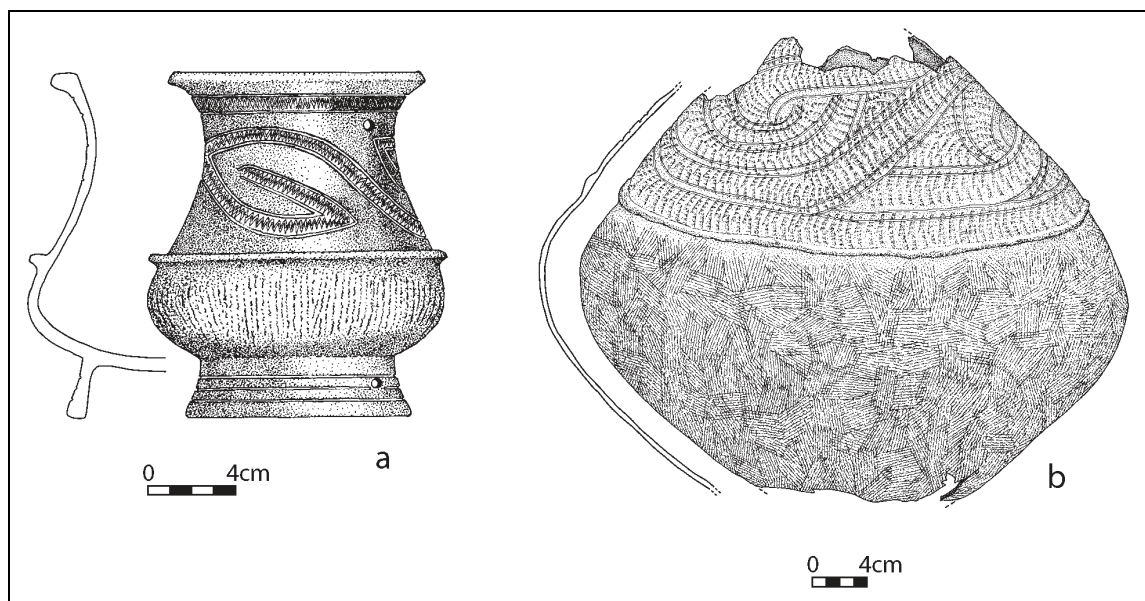
Even though several archaeologists have argued that this late third millennium settlement ‘horizon’ represents an intrusion of rice agricultural societies ultimately deriving from early farmers in the Yangtze Valley (e.g., Bellwood 2005; Higham 2002; Higham and Higham 2009; Rispoli 2007:287), as the data are analysed in more detail, the picture is not as straightforward and homogeneous as initially expected. Diversity in landscape orientation, material culture, and ceramic and subsistence technologies can be seen in these late third, early second millennium BCE societies.

Those who would claim that a homogeneous group of rice-agriculturalists moved down from southern China into Southeast Asia (View One discussed above) need to address the intrinsic environmental diversity of MSEA when they study the region’s prehistoric food producing societies as well as its hunting and gathering societies. The earliest stages of nonindustrial human food production systems tend to be strongly influenced by and to varying degrees mimic and/or manipulate a region’s basic natural ecosystem. Given the remarkable amount of inherent environmental diversity in MSEA, it seems unlikely that a single model of hunter-gatherer subsistence strategy or of agrarian change will fit all or most areas of MSEA. Transformations of natural ecosystems by ecologically dissimilar human economies, such as replacing a diverse tropical ecosystem with mono-crop plantations, are costly in terms of labor and high in risk to establish and maintain. Mono-crop field agriculture in the tropics produces populations of species with increased vulnerability to uncontrollable and hence potentially disastrous disease vectors and weather events. In an environmental context as diverse as MSEA, rapid successful incursion from northern Vietnam to the Malay Peninsula of a homogeneous early food production technological system focusing on a single species of cereal seems unlikely. On ecological grounds alone, a realistic expectation of the earliest prehistoric agriculturalists in Southeast Asia is that they did not overinvest in concentrations of a single species but that they hedged their bets with diverse resources and production strategies responding closely to local conditions and contingencies, just as other species do in natural tropical ecosystems. In other words, risk management in pre-industrial tropical ecosystems, man-made or not, almost always involves diversified investment strategies.

Archaeological evidence of this expected diversity can be seen clearly at the sites of Ban Chiang, Non Pa Wai, and Khok Phanom Di. In northern northeast Thailand, rice remains from Ban Chiang dating to calibrated 2289–1978 BCE (White 2008) show the presence of rice-cultivating societies in that region and time. But in central Thailand, Weber et al. (2010) identified foxtail millet (*Setaria italica*) dating to the late third millennium BCE (calibrated 2,470–2,200 BCE) from Non Pa Wai’s so-called ‘Outlier’.

According to their analysis, rice did not appear in interior central Thailand's Chao Phraya basin until after 1000 BCE. This conclusion meshes well with settlement findings (Eyre 2006, 2011; Mudar 1995; Mudar and Pigott 2003) that pre-metal and bronze age sites of the eastern Chao Phraya basin were preferentially located in upland areas where slopes and soils are not conducive to wet rice cultivation and there is no evidence for ancient terracing. Thus the late third millennium BCE evidence from northeast and central Thailand indicates that at least two distinct and contemporaneous agrarian regimes coexisted in inland Thailand. In coastal Thailand during the late third, early second millennium BCE, another contemporaneous subsistence orientation is evident at the sites of Nong Nor (Higham and Thosarat 1998) and Khok Phanom Di (Higham and Thosarat 2004), which were inhabited by settled societies with a non-agrarian marine adaptation (Higham and Thosarat 2004).

One intriguing aspect of these three areas occupied by settled societies with distinctly different subsistence orientations is that their ceramics appear to share a distinctive ceramic decorative surface treatment, known as the incised and impressed (i&i) style (Rispoli 2007). Pots displaying this decorative style (Figure 8) have a zone, such as a shoulder or pedestal, with incised, usually elaborate designs with impressed infilling (Bellwood 2005:132; Rispoli 2007). The impressions can be applied with a variety of tools and techniques. At Ban Chiang, these impressed designs are made by rocker stamping with a serrated or plain instrument. In other areas, dentate impressions were made by pressing or pricking the surface with a comb, shell edge, or individual stylus (Rispoli 1997). At Ban Non Wat pots with i&i designs are also painted (Higham 2009:205). Despite these differences, there is a common vocabulary and grammar in the range and complexity of the overall composition of the design field. Rispoli (2007:235) expresses the presumption of some that 'these decorative techniques belong to or sprang from a single cultural entity', which until recently had been assumed to be early (neolithic) rice-cultivating societies.



**Figure 8:** Examples of incised and impressed (i&i) pottery from Ban Chiang, Thailand.

- a. Early Period Phase I example dating to about 2100 B.C.E calibrated;
- b. Infant burial jar with dentate i&i design from Early Period Phase IIc dating roughly between 2000–1700 BCE calibrated (White 2008).

Current archaeological evidence demonstrates that i&i style was not *ipso facto* related to the spread of rice agriculture, contrary to the proposals of the View One proponents discussed above (Bellwood 2005:132; Higham 2002). As they are examined in detail, the ceramic assemblages if analysed as a whole exhibit clear diversity in morphology and technology. Specialists are at the early stages of studying and publishing technological details of full ceramic assemblages for many of the pertinent sites, yet it is already evident that the forms on which the i&i decorative system is applied vary from subregion to subregion. At Ban Chiang, i&i decoration is most commonly found on large globular or ovoid infant burial jars (Figure 8b) usually formed by coils added to a base slab (Glanzman and Fleming 1985). Within the Ban Chiang levels that contain i&i pottery, most ceramics lack this decorative treatment and many undecorated forms are present. At Ban Non Wat, i&i decoration is commonly found on flaring pedestals of large unrestricted bowls (Higham 2009: 207). Khok Charoen, Non Pa Wai, and Khok Phanom Di each have their own idiosyncratic set of forms with i&i decoration (Higham and Thosarat 1998; Ho 1984; Rispoli 1997).

The intra-site ceramic variability at these late third, early second millennium settlements is also not consistent with expectations of 'neolithic' homogeneity. The best example of the co-existence of i&i vessels with a completely distinct ceramic technological tradition comes from central Thailand at the same site, Non Pa Wai, with millet seeds dated to the late third millennium BC (Rispoli 1997; Weber et al. 2010). At the site of Non Pa Wai was found a vessel forming tradition of impressing clay into large coarsely woven baskets that produce the so-called 'elephant hide' pottery. This localised ceramic tradition co-occurs with i&i, which is found on the same forms as the elephant hide vessels but with cordmarked exteriors (Rispoli 1997).

How to account for the variability in late third, early 2<sup>nd</sup> millennium BCE ceramic assemblages will be a major challenge for archaeologists and a critical component of any future discussions of MSEA cultural diversity during the prehistoric period. White and Eyre (2011) have proposed that several ceramic subregions in Thailand appear by the bronze age, and it now seems likely that these sub-regions were established in the pre-metal period. What social mechanisms led to the establishment of many distinct ceramic traditions in the third millennium BCE? How can we account for a shared decorative technique and aesthetic that appears across technological traditions? It is not yet clear if the appearance of the i&i design system across a wide geographic area represents the spread of people (and potters), the spread of a full technological system, or the sharing of an aesthetic style by many discrete ceramic technological traditions. Understanding the context for the phenomenon will ultimately require study of whole assemblages.

## **6 Summary and Conclusions**

This chapter has reviewed evidence for prehistoric environmental and cultural diversity in MSEA. The region has inherent geographic diversity stemming from its complex geological history, its tropical latitudes, and its late quaternary climate and sea level changes. This inherent environmental diversity sets the stage for cultural diversity reaching back into the late Pleistocene.

How human societies dealt with the complex environmental context can be discerned in variability in settlement and technological choices discernible even in Pleistocene OIS3 sites. It appears that resources exploited ranging from pig to molluscs would have entailed flexible procurement and planning strategies, tailored to the range of habitats in the vicinity of sites. Delayed-return strategies such as would be needed to detoxify certain nuts

and yams were employed, as well as more immediate-return, opportunistic encounter hunting.

Habitat-tailoring probably continued in the terminal Pleistocene and early Holocene period when the lithic technology took on a distinctive cast known as the Hoabinhian. Although Hoabinhian lithic assemblages have commonalities across a wide area, the sites from which these assemblages came show diversity and flexibility in subsistence evidence and landscape position. I propose that the florescence and durability of Hoabinhian technology reflects intensified exploitation of bamboo, which has great utility across a wide array of landscapes and subsistence orientations. In addition, as explained above, intensified exploitation of bamboo would likely involve horticultural behaviours to exploit and enhance bamboo's natural proclivity to thrive in disturbed contexts.

If the bamboo-focus of this time frame is empirically confirmed, then I would also propose that there was a consequence for societies that invested in plaited bamboo artefacts, like baskets and fish traps. Because the material properties of bamboo species vary, it is plausible that investment in particular basketry technologies and styles would lead to increased investment in particular geographic areas in order to maintain access to preferred species. Transmission of plaiting technology would require more extended learning periods than needed for flaking river cobbles in the usual Hoabinhian manner. Enduring communities of practice reflecting learning networks for plaiting basketry and traps would likely have formed and regional stylistic variation could have emerged. While the 'Bamboo hypothesis' needs much more research, it is proposed that in some areas, 'localised domains of identity' could have constellated during the Hoabinhian period.

The middle Holocene is a period in MSEA with a dearth of archaeological evidence. Nevertheless, diversity in technology and subsistence orientation is undoubted, with the appearance of settled maritime-oriented societies like Da But. Hoabinhian finds decline, but it is possible that the middle Holocene settlement systems have little archaeological visibility with current archaeological methodologies. Ceramics and polished lithics appear in this time frame, but until archaeologists identify and excavate interior middle Holocene settlements, not much will be understood of this period.

By the end of the middle Holocene, at least three subsistence orientations can be identified: millet cultivation in central Thailand, rice cultivation in northern northeast Thailand, and a maritime orientation in coastal Thailand. Some ceramic vessels from these and other areas of MSEA share, however, a decorative treatment known as i&i style. As assemblages where i&i pottery is represented are studied, it is becoming clear that ceramics of this time range (late third to mid second millennium BCE) are highly diverse in form and technology. Therefore the cultural meaning of this widespread decorative treatment remains unclear. It can no longer be simply claimed that it is tied to the spread of rice agriculture, since it is found with at least three different subsistence regimes.

Marked regionality of ceramic technology and style is indisputable in the metal age (c. 2000 BCE–500 CE). Further archaeological research is needed to explore this phenomenon of ceramic subregions. These subregions are evident across the range of subsistence regimes noted above (dryland millet, probably wetland rice, and maritime). It is proposed that these subregions could coincide with enduring 'localised domains of identity' that involved not only development and transmission of specific technological traditions, but aspects of social life like marriage pools, alliance networks, and subregional ritual systems.

In conclusion, evidence for cultural diversity in MSEA extends far into the prehistoric period, probably mirroring the ecological diversity inherent in the region. Evidence for technological choice is even seen in pre-LGM sites. Diversity in subsistence practices throughout the HG-AG transition argues against models prioritizing homogeneous

neolithic waves of advance. As prehistoric technological traditions are more fully defined and studied, the parameters of past cultural diversity will be more convincingly established. The interactions of societies with varying technological styles, such as mobile hunter-gatherers interfacing with settled horticultural communities, is a topic of high priority in the coming decade.

## Acknowledgements

The following are thanked for helpful comments on earlier drafts of this paper: Roger Blench, Marie-Claude Boileau, Elizabeth Hamilton, and Vincent Pigott. Rasmi Shoocongdej and Ben Marwick were consulted on the Hoabinhian extent shown in Figure 3. Ardeth Anderson Abrams prepared Figures 1, 3–5, 7, and 8. Peter Johan Nijland prepared Figures 2, 3 and 6.

## References

- Anderson, Douglas D. 1990. *Lang Rongrien Rockshelter: A Pleistocene-early Holocene archaeological site from Krabi, Southwestern Thailand*. University Museum Monograph No. 71. Philadelphia: The University Museum, University of Pennsylvania.
- 1997. Cave archaeology in Southeast Asia. *Geoarchaeology* 12. 607–38.
- 2005. The use of caves in peninsular Thailand in the late Pleistocene and early and middle Holocene. *Asian Perspectives* 44. 137–53.
- Anisutkin, N. K. & V. I. Timofeyev. 2006. The paleolithic flake industry in Vietnam. *Archaeology, Ethnology and Anthropology of Eurasia* 3(27). 16–24.
- Baird, Ian G. & Bruce Shoemaker. 2008. People, livelihoods and development in the Xekong river basin, Laos. Bangkok: White Lotus.
- Bannanurag, Rachanie. 1988. Evidence for ancient woodworking: A microwear study of Hoabinhian stone tools. In Pisit Charoenwongsa & Bennet Bronson (eds) *Prehistoric Studies: The Stone and Metal Ages in Thailand*, 60–79. Bangkok: Thai Antiquity Working Group.
- Bar-Yosef, Ofer, Metin I. Eren, Jiarong Yuan, David J. Cohen & Yiyuan Li. 2011. Were bamboo tools made in prehistoric Southeast Asia? An experimental view from South China. *Quaternary International*, doi:10.1016/j.quaint.2011.03.026.
- Barker, Graeme. 2005. The archaeology of foraging and farming at Niah Cave, Sarawak. *Asian Perspectives* 44 (1). 90–106.
- Barker, Graeme, Huw Barton, Michael Bird, Patrick Daly, Ipoi Datan, Alan Dykes, Lucy Farr, David Gilbertson, Barbara Harisson, Chris Hunt, Tom Higham, Lisa Kealhofer, John Krigbaum, Helen Lewis, Sue McLaren, Victor Paz & Alistair Pike. 2007. The ‘human revolution’ in lowland tropical Southeast Asia: the antiquity and behavior of anatomically modern humans at Niah Cave (Sarawak, Borneo). *Journal of Human Evolution* 52. 243–61.
- Barton, Huw. 2005. The case for rainforest foragers: The starch record at Niah Cave, Sarawak. *Asian Perspectives* 44 (1). 56–72.
- Bellwood, Peter. 2002. Farmers, foragers, languages, genes: The genesis of agricultural societies. In Peter Bellwood & Colin Renfrew (eds) *Examining the Farming/language Dispersal Hypothesis*, 17–28. McDonald Institute for Archaeological Research, University of Cambridge: Cambridge.

- . 2005. *First farmers: The origins of agricultural societies*. Malden: Blackwell Publishing.
- . 2009. The dispersals of established food-producing populations. *Current Anthropology* 50. 621–6.
- Bellwood, Peter & Colin Renfrew (eds). 2002. *Examining the farming/language dispersal hypothesis*. Cambridge, UK: McDonald Institute for Archaeological Research.
- Bird, Michael I., Geoffrey Hope & David Taylor. 2004. Population PEP II: The dispersal of humans and agriculture through Austral-Asia and Oceania. *Quaternary International* 118–119. 145–163.
- Bird, Michael I., David Taylor & Chris Hunt. 2005. Palaeoenvironments in insular Southeast Asia during the Last Glacial Period: A savanna corridor in Sundaland? *Quaternary Science Reviews* 24. 2228–2242.
- Bowdler, Sandra. 2006. The Hoabinhian: Early evidence for SE Asian trade networks? In Elisabeth A. Bacus, Ian C. Glover, and Vincent C. Pigott (eds) *Uncovering Southeast Asia's past: selected papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists*, 355–9. Singapore: National University of Singapore Press.
- . 2008. Hoabinhian and non-Hoabinhian, In Jean-Pierre Pautreau, Anne-Sophie Coupey, Valéry Zeitoun & Emma Rambault (eds) *From Homo erectus to the living traditions*, 59–66. Chiang Mai: European Association of Southeast Asian Archaeologists.
- Bui Vinh 1991. The Da But culture in the stone age of Viêt Nam. *Bulletin of the Indo-Pacific Prehistory Association* 10. 127–131.
- Bulbeck, David, this volume, Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya.
- Castillo, Cristina & Dorian Q. Fuller. 2010. Still too fragmentary and dependent upon chance? Advances in the study of early Southeast Asian archaeobotany. In Bérénice Bellina, Elisabeth A. Bacus, Thomas Oliver Pryce & Jan Wisseman Christie (eds) *50 Years of Archaeology in Southeast Asia: Essays in Honour of Ian Glover*, 93–111. Bangkok: River Books.
- Christanty, Linda, J. P. Kimmins & D. Maily. 1997. 'Without bamboo, the land dies': A conceptual model of the biogeochemical role of bamboo in an Indonesian agroforestry system. *Forest Ecology and Management* 91. 83–91.
- Cook, Charlotte G., Richard T. Jones, Peter G. Langdon, Melanie J. Leng & Enlou Zhang. 2011. New insights on Late Quaternary Asian palaeomonsoon variability and the timing of the Last Glacial Maximum in southwestern China. *Quaternary Science Reviews* 30. 808–20.
- Corvinus, Gudrun. 2007. *Prehistoric cultures in Nepal: From the early Palaeolithic to the Neolithic and Quaternary geology of the Dang-Deokhuri Dun valleys*. Weisbaden: Harrassowitz Verlag.
- Dransfield, S. & E. A. Widjaja. 1995. *Plant resources of South-East Asia. No. 7: bamboos*. PROSEA Handbook No. 7. Leiden, Netherlands: Backhuys Publishers.
- Dunn, Frederick L. 1975. *Rain-forest collectors and traders: A study of resource utilization in modern and ancient Malaya*. Kuala Lumpur: The Malaysian Branch of the Royal Asiatic Society.
- Eyre, Chureekamol Onsuwan. 2006. *Prehistoric and proto-historic communities in the eastern Upper Chao Phraya River Valley, Thailand: Analysis of site chronology, settlement patterns, and land use*. Unpublished Ph.D. thesis. Philadelphia: University of Pennsylvania, Department of Anthropology.

- . 2011. Social variation and dynamics in Metal Age and Protohistoric central Thailand: A Regional Perspective, *Asian Perspectives* 49(1), 43-84.
- Forestier, Hubert. 2000. De quelques chaînes opératoires lithiques en Asie du Sud-Est au Pléistocène supérieur final et au début de l'Holocène. *L'Anthropologie* 104. 531–548.
- . 2003. Des outils nés de la forêt. In Alain Froment & Jean Guffroy (eds) *Peuplements Anciens et Actuels des Forêts Tropicales*, 315–38. Paris: Institut de Recherche pour le Développement.
- Forestier, Hubert, Truman Simanjuntak, Dominique Guillaud, Dubel Driwantoro, Ketut Wiradnyana, Darwin Siregar, Rokus Due Awe & Budiman. 2005. Le site de Tögi Ndrawa, île de Nias, Sumatra nord: Les premières traces d'une occupation hoabinhienne en grotte en Indonésie. *Comptes Rendus Palevol* 4. 727–733.
- Forestier, Hubert, Valéry Zeitoun, Christophe Maitay & Chinnawut Winayalai. 2008. The Hoabinhian site of Huai Hin (northern Thailand). In Jean-Pierre Pautreau, Anne-Sophie Coupey, Valéry Zeitoun & Emma Rambault (eds) *From Homo Erectus to the Living Traditions*, 67–72. Chang Mai: European Association of Southeast Asian Archaeologists.
- Fuller, Dorian Q., Yo-Ichiro Sato, Cristina Castillo, Ling Qin, Alison R. Weisskopf, Eleanor J. Kingwell-Banham, Jixiang Song, Sung-Mo Ahn & Jacob Van Etten. 2010. Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeological and Anthropological Sciences* 2. 115–31.
- Gaillard, Claire, Mukesh Singh & Anne Dambricourt Malassé. 2011. Late Pleistocene to early Holocene lithic industries in the southern fringes of the Himalaya. *Quaternary International* 229. 112–122.
- Gillieson, David. 2005. Karst in Southeast Asia. In Avijit Gupta (ed.) *The Physical Geography of Southeast Asia*, 157–176. New York: Oxford University Press.
- Glanzman, William D. & Stuart J. Fleming. 1985. Ceramic technology at prehistoric Ban Chiang, Thailand: Fabrication methods. *MASCA Journal* 3. 114–21.
- Gorman, Chester F. 1971. The Hoabinhian and after: Subsistence patterns in Southeast Asia during the Late Pleistocene and Early Recent Periods. *World Archaeology* 2. 300–20.
- . 1972. Excavations at Spirit Cave, north Thailand: Some interim interpretations. *Asian Perspectives (1970)* 13. 79–107.
- Graves, Michael W. 1994. Kalinga social and material culture boundaries: A case of spatial convergence. In William A. Longacre & James M. Skibo (eds) *Kalinga Ethnoarchaeology: Expanding Archaeological Method and Theory*, 13–49. Washington, D. C.: Smithsonian Institution Press.
- Gupta, Avijit. 2005. Landforms of Southeast Asia. In Avijit Gupta (ed.) *The Physical Geography of Southeast Asia*, 38–64. New York: Oxford University Press.
- Hanebuth, Till J. J., Harold K. Voris, Yusuke Yokoyama, Yoshiki Saito & Jun'ichi Okuno. 2011. Formation and fate of sedimentary depocentres on Southeast Asia's Sunda Shelf over the past sea-level cycle and biogeographic implications. *Earth-Science Reviews* 104. 92–110.
- Hannerz, Ulf. 2010. Diversity is our business. *American Anthropologist* 112(4). 539–51.
- Hegmon, Michelle. 1998. Technology, style, and social practices: Archaeological approaches. In Miriam T. Stark (ed.) *The Archaeology of Social Boundaries*, 264–280. Washington, D. C.: Smithsonian Institution Press.
- Higham, Charles. 1977. Economic change in prehistoric Thailand. In Charles A. Reed (ed.) *Origins of Agriculture*, 385–412. The Hague: Mouton.

- . 2002. Languages and farming dispersals: Austroasiatic languages and rice cultivation. In Peter Bellwood & Colin Renfrew (eds) *Examining the Farming/Language Dispersal Hypothesis*, 223–232. Cambridge: McDonald Institute for Archaeological Research, University of Cambridge.
- . 2009. 'Material culture' and 'Summary,' in C. W. F. Higham & A. Kijngam (eds) *The Origins of the Civilization of Angkor, Vol. 3: The Excavation of Ban Non Wat. Part One: Introduction*, 187–257. Bangkok: The Fine Arts Department of Thailand.
- Higham, Charles, Xie Guangmao & Lin Qiang. 2011. The prehistory of a Friction Zone: first farmers and hunters-gatherers in Southeast Asia. *Antiquity* 85. 529–43.
- Higham, Charles, & Thomas Higham. 2009. A new chronological framework for prehistoric Southeast Asia, based on a Bayesian model from Ban Non Wat. *Antiquity* 83(319). 125–144.
- Higham, Charles, & Rachanie Thosarat. 1998. Conclusions and summary of results. In Charles F. W. Higham & Rachanie Thosarat (eds) *The Excavation of Nong Nor, a Prehistoric Site in Central Thailand*, 523–539. Dunedin: Department of Anthropology, University of Otago.
- . 2004. The excavation of Khok Phanom Di, a prehistoric site in Central Thailand. Volume VII: Summary and Conclusions. London: The Society of Antiquaries of London.
- Ho, Chui-Mei. 1984. *The pottery of Kok Charoen and its farther context*. Unpublished Ph.D. thesis. University of London, London.
- Hope, Geoffrey. 2005. The Quaternary in Southeast Asia. In Avijit Gupta (ed.) *The Physical Geography of Southeast Asia*, 24–37. New York: Oxford University Press.
- Hutchison, Charles S. 2005. The geological framework. In Avijit Gupta (ed.) *The Physical Geography of Southeast Asia*, 3–23. New York: Oxford University Press.
- Hutterer, Karl L. 1976. An evolutionary approach to the Southeast Asian cultural sequence. *Current Anthropology* 17(2). 221–242.
- . 1983. The natural and cultural history of Southeast Asian agriculture. *Anthropos* 78. 169–212.
- . 1988. The prehistory of the Asian rain forests. In Julie Sloan Denslow & Christine Padoch (eds) *People of the Tropical Rain Forest*, 63–72. Berkeley: University of California Press.
- Jonsson, Hjorleifur, this volume, Ethnology and the Issue of Human Diversity in Mainland Southeast Asia.
- Junker, Laura L. 2002. Introduction. In Kathleen D. Morrison & Laura L. Junker (eds) *Forager-Traders in South and Southeast Asia: Long Term Histories*, 131–66. Cambridge: Cambridge University Press.
- Kammaing, Johan. 2007. Reinterpretation of the Hoabinhian at Sai Yok, Thailand. In Zuraina Majid and Mokhtar Saidin (eds) *Archaeological heritage of Malaysia*, No 1, Chapter 3. 38–58. Centre for Archaeological Research in Malaysia, Universiti Sains Malaysia, Kuala Lumpur.
- Kealhofer, Lisa. 2002. Changing perceptions of risk: The development of agro-ecosystems in Southeast Asia. *American Anthropologist* 104:178-94.
- Kealhofer, Lisa. 2003. Looking into the gap: Land use and the tropical forests of southern Thailand. *Asian Perspectives* 42(1). 72–95.
- Kennedy, J. I. 1977. A course toward diversity: Economic interaction and cultural differentiation in prehistoric mainland Southeast Asia. Unpublished Ph.D. thesis. Manoa: University of Hawaii, Department of Anthropology.



- 1978. From stage to development in prehistoric Thailand: An exploration of the origins of growth, exchange, and variability in Southeast Asia. In Karl L. Hutterer (ed.) *Economic Exchange and Social Interaction in Southeast Asia: Perspectives from Prehistory, History, and Ethnography*, 23-38. Ann Arbor: University of Michigan, Center for South and Southeast Asian Studies.
- Leong Sau Heng. 1991. Jenderam Hilir and the Mid-Holocene prehistory of the west coast plain of peninsular Malaysia. *Bulletin of the Indo-Pacific Prehistory Association* 10. 150–60.
- Martinson, D.G., N.G. Pisias, J.D. Hays, J. Imbrie, T.C. Moore & N.J. Shackleton. 1987. Age dating and orbital theory of the Ice Ages: Development of a high resolution 0–300 000-year chronostratigraphy. *Quaternary Research* 27. 1-29.
- Marwick, Ben. 2008a. Beyond typologies: The reduction thesis and its implications for lithic assemblages in Southeast Asia. *Bulletin of the Indo-Pacific Prehistory Association* 28. 108–116.
- 2008b. Human behavioural ecology and stone artefacts in northwest Thailand during the terminal Pleistocene and Holocene. In Jean-Pierre Pautreau, Anne-Sophie Coupey, Valéry Zeitoun & Emma Rambault (eds) *From Homo Erectus to the Living Traditions*, 73–80. Chiang Mai: European Association of Southeast Asian Archaeologists.
- 2008c. Stone artefacts and recent research in the archaeology of mainland Southeast Asian hunter-gatherers. *Before Farming* 2008/4 [[wapress.co.uk](http://wapress.co.uk)] [PDF]
- Maslin, Mark A. & Ellen Thomas. 2003. Balancing the deglacial global carbon budget: The hydrate factor. *Quaternary Science Reviews* 22. 1729–1736.
- Matthews, John M. 1966. The Hoabinhian affinities of some Australian assemblages. *Archaeology and Physical Anthropology in Oceania* 1:5-22.
- Maxwell, Andrew Lee 1999. Holocene environmental change in mainland Southeast Asia: Pollen and charcoal records from Cambodia. Unpublished Ph.D. thesis. Baton Rouge: Louisiana State University and Agricultural and Mechanical College, Department of Geography and Anthropology.
- 2001. Holocene monsoon changes inferred from lake sediment pollen and carbonate records, north-eastern Cambodia. *Quaternary Research* 56. 390-400.
- Matthews, John M. 1964. The Hoabinhian in Southeast Asia and elsewhere. PhD thesis. Canberra: Australian National University.
- 1966. A review of the 'Hoabinhian' in Indo-China. *Asian Perspectives* 9: 86-95.
- Mohanty, Pradeep. 1993. Mesolithic hunter-gatherers of Keonjhar District, Orissa, India. *Asian Perspectives* 32(1). 85–103.
- Mohanty, Pradeep, Kishore K. Basa & Balaram Tripathy. 1997. Stone Age cultures of Mayurbhanj District, Orissa. *Man in India* 77(2 and 3). 159–177.
- Moser, Johannes. 2001. Hoabinhian: Geographie und Chronologie eines steinzeitlichen Technocomplexes in Südostasien. Köln, Lindensoft.
- Mudar, Karen M. 1995. Evidence for prehistoric dryland farming in mainland Southeast Asia: Results of regional survey in Lopburi Province, Thailand. *Asian Perspectives* 34(2). 157–191.
- Mudar, Karen M. & Douglas Anderson. 2007. New evidence for Southeast Asian Pleistocene foraging economies: Faunal remains from the early levels of Lang Rongrien. *Asian Perspectives* 46(2). 298–334.

- Mudar, Karen M. & Vincent C. Pigott. 2003. Subsistence changes and community-based craft production in prehistoric central Thailand. In Anna Karlström and Anna Källén (eds) *Fishbones and glittering emblems: Southeast Asian Archaeology 2002*, 149–60. Stockholm: Museum of Far Eastern Antiquities.
- Natapintu, Surapol. 1991. Archaeometallurgical studies in the Khao Wong Prachan Valley, Central Thailand. *Bulletin of the Indo-Pacific Prehistory Association* 11: 153-159. Special issue, P. Bellwood (ed.) *Indo-Pacific Prehistory 1990*, Vol. 2). Papers from the 14<sup>th</sup> IPPA Congress, Yogyakarta.
- Nishimura, Masanari. 2005. Attribute analysis of the Hoabinhian industry: Implications from a comparative study of Bung Cave and Xom Trai Cave, northern Vietnam. *Journal of Southeast Asian Archaeology* 25. 81–104.
- Nguyen Viet. 2005. The Da But culture: Evidence for cultural development in Vietnam during the middle Holocene. *Bulletin of the Indo-Pacific Prehistory Association* 25. 89–93.
- . 2008. Hoabinhian macrobotanical remains from archaeological sites in Vietnam: Indicators of climate changes from the Late Pleistocene to the Early Holocene. *Bulletin of the Indo-Pacific Prehistory Association* 28. 80–3.
- Pautreau, Jean-Pierre & Christophe Maitay. 2009. Birmanie, premières cultures historiques du haut pays. Niveau d'occupation néolithique et datation C14. *Archéologia*, n°471, Novembre 2009. 67–69.
- Pawlik, Alfred. 2009. Is the functional approach helpful to overcome the typology dilemma of lithic archaeology in Southeast Asia? *Bulletin of the Indo-Pacific Prehistory Association* 29. 6–14.
- Penny, Daniel. 2001. A 40,000 year palynological record from north-east Thailand; implications for biogeography and palaeo-environmental reconstruction. *Palaeogeography, Palaeoclimatology, Palaeoecology* 171(3–4). 97–128.
- Penny, Daniel & Lisa Kealhofer. 2005. Microfossil evidence of land-use intensification in north Thailand. *Journal of Archaeological Science* 32(1). 69–82.
- Piper, Philip J., Janine Ochoa, Emil C. Robles, Helen Lewis, & Victor Paz. 2011. Palaeozoology of Palawan Island, Philippines. *Quaternary International* 233.142-158.
- Pookajorn, Surin. 1996. Human activities and environmental changes during the Late Pleistocene to middle Holocene in southern Thailand and Southeast Asia. In Lawrence Guy Straus, Berit Valentin Eriksen, Jon M. Erlandson & David R. Yesner (eds) *Humans at the End of the Ice Age: The Archaeology of the Pleistocene-Holocene Transition*, 201–213. New York: Plenum Press.
- . 1990. Hoabinhian cave excavations in Ban Kao district, west Thailand. In Ian C. Glover & Emily Glover (eds) *Southeast Asian Archaeology 1986*. Proceedings of the First Conference of the Association of Southeast Asian Archaeologists in Western Europe, 11–28. Oxford: British Archaeological Reports International Series 561.
- Pope, Kevin O. & John E. Terrell. 2007. Environmental setting of human migrations in the circum-Pacific region. *Journal of Biogeography* 35(1): 1–21.
- Pyramarn, Kosum. 1989. New evidence on plant exploitation and environment during the Hoabinhian (late Stone Age) from Ban Kao Caves, Thailand. In David R. Harris & Gordon C. Hillman (eds) *Foraging and Farming: The Evolution of Plant Exploitation*, 282–291. London: Unwin Hyman.
- Rabett, Ryan, Joanna Appleby, Alison Blyth, Lucy Farr, Athanasia Gallou, Thomas Griffiths, Jason Hawkes, David Marcus, Lisa Marlow, Mike Morley, Nguyễn Tấn Cao, Nguyễn Sơn

- Van, Kirsty Penkman, Tim Reynolds, Christopher Stimpson & Katherine Szabó. 2011. Inland shell midden site-formation: Investigation into a late Pleistocene to early Holocene midden from Trảng An, Northern Vietnam. *Quaternary International* 239. 153–69.
- Rabett, Ryan J. & Graeme Barker. 2010. Late Pleistocene and early Holocene forager mobility in Southeast Asia. In Bérénice Bellina, Elisabeth A. Bacus, Thomas Oliver Pryce & Jan Wisseman Christie (eds) *50 Years of Archaeology in Southeast Asia: Essays in Honour of Ian Glover*, 67–78. Bangkok: River Books.
- Rabett, Ryan J., Philip J. Piper & Graeme Barker. 2006. Bones from 'Hell': Preliminary results of new work on the Harrisson faunal assemblage from the deepest part of Niah Cave, Sarawak. In Elisabeth A. Bacus, Ian C. Glover & Vincent C. Pigott (eds) *Uncovering Southeast Asia's Past: Selected Papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists*, 46–59. Singapore: National University of Singapore Press.
- Reynolds, Timothy E.G. 1992. Excavations at Banyan Valley cave, northern Thailand: A report on the 1972 season. *Asian Perspectives* 31. 77–97.
- Rispoli, Fiorella 1997. Late third-mid second millennium B.C. pottery traditions in central Thailand: Some preliminary observations in a wider perspective. In Roberto Ciarla & Fiorella Rispoli (eds) *South-east Asian Archaeology 1992. Proceedings of the fourth International Conference of the European Association of South-east Asian Archaeologists. Rome, 28th September - 4th October 1992*, 59–97. Rome: Istituto Italiano per L'Africa e L'Oriente.
- 2007. The incised and impressed pottery style of mainland Southeast Asia: Following the paths of Neolithization. *East and West* 57(1-4). 235–304.
- Roberts, Richard Graeme, Mike J. Morwood & Kira E. Westaway. 2005. Illuminating Southeast Asian prehistory: New archaeological and paleoanthropological frontiers for luminescence dating. *Asian Perspectives* 44(2). 293–319.
- Saha, Sonali & Henry F. Howe. 2001. The bamboo fire cycle hypothesis: A comment. *The American Naturalist* 158:659–63.
- Saidin, Mokhtar. 2006. Bukit Bunuh, Lenggong, Malaysia: New evidence of late Pleistocene culture in Malaysia and Southeast Asia. In Elisabeth A. Bacus, Ian C. Glover & Vincent C. Pigott (eds) *Uncovering Southeast Asia's Past: Selected Papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists*, 60–64. Singapore: National University of Singapore Press.
- Sathiamurthy, Edlic & Harold K. Voris. 2006. Maps of Holocene sea level transgression and submerged lakes on the Sunda Shelf. *Natural History Journal of Chulalongkorn University* 2. 1–43.
- Sharma, Tarun C. 1984. Recent advances in prehistory and archaeology of northeast India. *Journal of the Assam Research Society* 28. 1–28.
- 1990. Discovery of Hoabinian cultural relics in north-east India. In N.C. Ghosh & Subrata Chakrabarti (eds) *Adaptation and Other Essays: Proceedings of the Archaeological Conference, 1988*, 136–9. Santiniketan: Visva-Bharati Research Publications.
- Sidwell, Paul & Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.
- Silvestre, Ramon. 1994. The ethnoarchaeology of Kalinga basketry: A preliminary investigation. In William A. Longacre and James M. Skibo (eds) *Kalinga Ethnoarchaeology: Expanding*

- Archaeological Method and Theory*, 199–207. Washington, D.C.: Smithsonian Institution Press.
- Shocongdej, Rasmi. 2006. Late Pleistocene activities at the Tham Lod rockshelter in highland Pang Mapha, Mae Hong Son Province, northwestern Thailand. In Elisabeth A. Bacus, Ian C. Glover & Vincent C. Pigott (eds) *Uncovering Southeast Asia's Past: Selected Papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists*, 22–37. Singapore: National University of Singapore Press.
- . 2010. Subsistence-settlement organisation during the late Pleistocene-early Holocene: The case of Lang Kamnan Cave, western Thailand. In Bérénice Bellina, Elisabeth A. Bacus, Thomas Oliver Pryce & Jan Wisseman Christie (eds) *50 Years of Archaeology in Southeast Asia: Essays in Honour of Ian Glover*, 51–66. Bangkok: River Books.
- Srisuchat, Amara. 1989. Comments on the results of carbon-14 dating applied to archaeological finds from excavated sites in Thailand. *Silpakorn Journal* 33(5). 44–67.
- Stark, Miriam T. 1998. Technical choices and social boundaries in material culture patterning: An introduction. In Miriam T. Stark (ed.) *The Archaeology of Social Boundaries*, 1–11. Washington, D.C.: Smithsonian Institution Press.
- Voris, Harold K. 2000. Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems, and time durations. *Journal of Biogeography* 27(5). 1153–1167.
- Wang, Yongjin, Hai Cheng, R. Lawrence Edwards, Xinggong Kong, Xiaohua Shao, Shitao Chen, Jiangyin Wu, Xiouyang Jiang, Xianfeng Wang & Zhisheng An. 2008. Millennial- and orbital-scale changes in the East Asian monsoon over the past 224,000 years. *Nature* 451. 1090–3.
- Wang, XiaMei, XiangJun Sun, PinXian Wang & Karl Stattegger. 2009. Vegetation on the Sunda Shelf, South China Sea, during the Last Glacial Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology* 278. 88–97.
- Welsch, Robert L. & John Edward Terrell. 1998. Material culture, social fields, and social boundaries on the Sepik Coast of New Guinea, in Miriam T. Stark, ed. *The Archaeology of Social Boundaries*. 50–77 Smithsonian Series in Archaeological Inquiry Washington, D.C.: Smithsonian Institute Press.
- Wenger, Etienne. 1998. *Communities of Practice*. Cambridge: University Press.
- West, Jolee A. & Julien Louys. 2007. Differentiating bamboo from stone tool cut marks in the zooarchaeological record, with a discussion on the use of bamboo knives. *Journal of Archaeological Science* 34. 512–518.
- White, Joyce C. 1995. Modeling the development of early rice agriculture: Ethnoecological perspectives from northeast Thailand. *Asian Perspectives* 34(1). 37–68.
- . 2008. Dating early bronze at Ban Chiang, Thailand. In Jean-Pierre Pautreau, Anne-Sophie Coupey, Valéry Zeitoun & Emma Rambault (eds) *From Homo Erectus to the Living Traditions*, 91–104. Chiang Mai: European Association of Southeast Asian Archaeologists.
- White, Joyce C. & Bounheuang Bouasisengpaseuth. 2008. Archaeology of the Middle Mekong: Introduction to the Luang Prabang province exploratory survey. In Yves Goudineau & Michel Lorrillard (eds) *Recherches Nouvelles sur le Laos*, 37–52. Paris: École Française d'Extrême-Orient.
- White, Joyce C. & Chureekamol Onsuwan Eyre. 2011. Residential burial and the metal age of Thailand. In Ron L. Adams & Stacie M. King (eds) *Residential Burial: A Multiregional Exploration*, 59–78. *Archaeological Papers of the American Anthropological Association* 20.

- White, Joyce C. & Chester F. Gorman. 2004. Patterns in 'amorphous' industries: The Hoabinhian viewed through a lithic reduction sequence. In Victor Paz (ed.) *Southeast Asian Archaeology: Wilhelm G. Solheim II Festschrift*, 411–441. Quezon City: University of the Philippines Press.
- White, Joyce C., Daniel Penny, Bernard Maloney & Lisa Kealhofer. 2004. Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. *Quaternary International* 113(1). 111–132.
- Wurster, Christopher M., Michael I. Bird, Ian D. Bull, Frances Creed, Charlotte Bryant, Jennifer A. J. Dungait & Victor Paz. 2010. Forest contraction in north equatorial Southeast Asia during the Last Glacial Period. *Proceedings of the National Academy of Sciences of the United States of America* 107. 15508–11.
- Yen, Douglas E. 1977. Hoabinhian horticulture? The evidence and the questions from northwest Thailand. In Jim Allen, Jack Golson & Rhys Jones (eds) *Sunda and Sahul: Prehistoric Studies in Southeast Asia, Melanesia and Australia*, 567–99. London: Academic Press.

# 3 *Human diversity in mainland Southeast Asia: the contribution of bioarchaeology*

---

SIÂN E. HALCROW AND NANCY TAYLES

## 1 Introduction

This chapter focuses on the contribution of bioarchaeology to understanding human diversity in MSEA. Traditionally, the contribution of research on human skeletal remains would have been to describe human *physical* diversity and from this to infer origins of, and relationships among groups, communities, or populations. Developments within bioarchaeology in the past few decades have taken the field beyond this descriptive taxonomy, to the adoption of a biocultural approach to interpreting human *biological* diversity, including adaptation to environment, culture and/or social structure, using demographic, health and disease and migration analysis.

The traditional method of researching human physical diversity relies on variations in phenotypic skeletal size, shape and minor morphological characteristics being expressions of genotype. Genetic relationships between and among populations are inferred from similarities and differences in these characteristics. This approach is known as biological distance, or biodistance analysis (Buikstra et al. 1990; Stojanowski and Schillaci 2006) and it forms the basis for inferences about population history, particularly migration/gene flow. Recently, the identification of individuals within a population who immigrated into a site has become possible through the use of isotopes (particularly strontium and oxygen) in bones and teeth, although this is limited to those who have migrated from a different geological region (Bentley 2006).

The methods used to research human biological diversity include skeletal and dental evidence of health, diet, physical activity and population demography. The state of individual and population health is determined through evidence of stress during the period of growth such as the presence and prevalence of dental enamel defects, sexual dimorphism in stature, and levels of infant and child mortality. Such stress may be a product of malnutrition, response to disease, or less specific, culturally imposed stressors such as imposition of physical activity or lack of access to resources. More specifically, the presence and prevalence of some diseases is identifiable through skeletal and dental pathology. The composition of diet, at least in the form of balance between major components, can be more specifically identified through isotopes in bones or teeth. Physical activity can be examined through the development and patterns of musculo-

skeletal markers, the form of bone at entheses, the attachment sites of muscle on bone. Population demography is determined through the patterns of mortality and fertility, based on age specific mortality rates. The evidence from these sources is then interpreted in the context of both the physical environment, determined through geomorphology, and the cultural environment, defined using archaeological evidence of subsistence mode, occupational activities, level of technology and social structure.

This chapter addresses the question of the ability of bioarchaeology to contribute to understanding the basis of human biological diversity in MSEA, which is the focus of our research interests. The focus is on late prehistory, specifically the last 4000 years, because the samples of human skeletal remains prior to this period are minimal, and have been reported only with the context of physical diversity.

## **2 The archaeological context**

Research on later prehistory of Southeast Asia has focused, in large part, on the topics of the introduction of agriculture (Bellwood 2005; Higham 2002; Higham and Lu 1998; Hutterer 1983) and technological developments such as the introduction of bronze and iron, along with the effects of these events on social organisation (Higham 2002; O'Reilly 2003; Stark 2006; Welch 1984), modes of production (White and Pigott 1996), and the environment (Boyd et al. 1999; Boyd and Habberfield-Short 2007; White et al. 2004). Much of this research has the classical archaeological foci on questions of origins and timing of these introductions and subsequent changes, using evidence from linguistics and material culture. This tends to be interpreted using the orthodox, Euro- and technocentric model of chronological development, within a series of simplified, linear evolutionary stepping stones, through the Neolithic, Bronze Age and Iron Age (Higham 2002), despite arguments that these stages are inappropriate for the region (Hutterer 1976; Pietruszewsky and Douglas 2002:223; White 2002:xvi–xvii).

The prevailing archaeological hypothesis is that the development of agriculture in the Yangtze River valley, southern China drove population growth and stimulated migration into Southeast Asia (Bellwood 2005; Higham 2002). Bellwood's (2005) recent book advances this 'early farming dispersal hypothesis' through a thorough synthesis of archaeological, linguistic and biological (genetic) evidence. The main tenet of his hypothesis is that farming spread through migration, with limited indigenous input or negotiation. However, as is apparent from the polarised reviews of the book (Bellwood 2007), Bellwood's (2005) negation of the importance of indigenous agency does not accommodate variation seen in the Southeast Asian archaeological record, such as evidence for site localisation of material culture (White 2006), and evidence of subsistence from palynological records (reviewed in White et al. 2004) around the time of agricultural development.

Some of the prevalent models on Southeast Asian prehistory rely to a large extent on certain linguistic models (Bellwood 2005; Higham 2002). The question is how to more profitably integrate linguistic evidence with bioarchaeological research in this region, which we argue later may be able to be addressed drawing on an areal linguistic approach. While more archaeological and bioarchaeological data are being amassed, there remain inadequacies in explanatory theories and models of social and biological changes over time.

### 3 Bioarchaeological research in mainland Southeast Asia

In the past morphometric data have been viewed as concrete evidence of genetic relationships among populations. Increasingly, publications acknowledge that these are phenetic rather than genetic relationships (for example, Matsumura et al. 2008), but nevertheless go on to interpret the data as if they are purely genetically determined. There is no doubt that head form, to take one of the morphological characteristics used, does have a genetic basis, indeed a strong genetic basis, but the extent to which this is reflected in the phenotype is variable depending on multiple environmental influences from very early in life, including but not limited to diet, nutrition, physical activity and pathology. For example, not only are agriculturalists eating a different diet from their hunter-gatherer-forager predecessors, and so using the masticatory muscles differently, but their nutrition will be different, affecting head size and shape particularly early in life when nutrition has a significant effect during the period of rapid growth. They will also be engaging in different patterns of physical activity, with potential to involve the muscles supporting the head and thereby altering head shape through the action of those muscles. Judicious interpretation of the statistics generated in assessment of biological distance is called for, taking into account the 'lifestyles' of the populations represented, rather than their blanket interpretation as evidence of genetic relationships. Bio-distance studies will not be addressed in this chapter, as an example of this type of research is contained in Matsumura et al.'s chapter.

A focus of bioarchaeological inquiry is on human biological diversity that develops in response to the interaction between biology, behavior and the environment (Larsen 2002). This is reflected in the biological anthropological research in Southeast Asia with the increase over the past 15 years in assessment of health and disease (Domett 2001; Douglas 1996; Halcrow 2006; Halcrow and Tayles 2008; Oxenham 2000; Oxenham and Tayles 2006; Pietrusewsky and Douglas 2002; Pietrusewsky and Douglas 2002a; Tayles 1999; Tayles et al. 2007).

Exemplifying the contribution this research focus can make to the understanding of regional human biological diversity, is the questioning of the universal applicability of the general model of health change with agriculture (Domett 2001; Domett and Tayles 2006, 2007; Halcrow 2006; Halcrow and Tayles 2008; Oxenham et al. 2006; Pietrusewsky and Douglas 2002a; Tayles et al. 2000, 2009). The model is based mainly on skeletal populations from the Western hemisphere, and posits a decline in health with the introduction and intensification of agriculture and associated sedentism (Larsen 1995; Steckel and Rose 2002). The non-adherence of MSEA prehistoric populations to this general model of decline in health has been interpreted by some (Douglas 1996; Douglas and Pietrusewsky 2007; Halcrow 2006; Pietrusewsky and Douglas 2002a) as the result of the retention of broad-based subsistence economy and community-based level of production (White 1995). However, while there is no evidence that the health declines over the period from early agricultural development (Neolithic) through to the Bronze Age (circa 3000BP–2500BP), there is some evidence of an increase in infectious disease in later prehistory (circa 2500BP–1500BP) (Halcrow 2006; Halcrow et al. 2008; Oxenham et al. 2005; Tayles and Buckley 2004). For example, Tayles and Buckley (2004) have found at the site of Noen U-Loke (cemetery dates 2500BC–1500BP) possible evidence of mycobacterial diseases. This is interpreted to be the result of the introduction of the pathogens through either immigration or trading contacts, together with an increase in settlement density during the time of settlement of Noen U-Loke, which provided a large



enough population to maintain these infectious diseases in contrast to the earlier nearby site of Ban Lum Khao (Tayles and Buckley 2004).

Oxenham et al. (2006) reviewed oral pathology from samples spanning from the early to late agricultural development. They found that oral health was homogeneous over time, which is explained, in part, by the low cariogenicity of rice, and also by the archaeological evidence for the increase of agricultural intensification occurring later in prehistory around the time of the introduction of iron (circa 2500BP).

Other specialist areas of bioarchaeology being investigated in MSEA include field anthropology (Willis and Tayles 2009), which is the integration of skeletal position in analysis of mortuary behavior, the integration of grave wealth and interment type into interpretation of health status (Oxenham et al. 2008; Halcrow et al. 2008) and stable isotope analysis for assessing migration (Bentley et al. 2005, 2007, 2009; Cox 2009; Cox et al. 2008). There has been limited work done on diet using stable isotopic analysis in samples from Southeast Asia (King and Norr 2006; King 2008). This is, in part, due to the poor preservation of the organic component of bone in some tropical environments (King et al. in prep).

## **4 Peopling**

The predominant macro-scale theory in MSEA archaeology that agriculture was the prime mover for migration into Southeast Asian prehistory, and therefore agriculture was introduced by migrants rather than through cultural diffusion, and/or local development (Bellwood 2005; Higham 2002) would have identifiable consequences. There would be a large number of migrants into the existing ‘hunter-gatherer’ communities and an increase in population size. Bioarchaeological work investigating kinship and migration through isotopic analysis, biodistance analysis, and paleodemography are main ways to test the theory.

Current work by Bentley and others using stable isotopes is contributing to understanding the nature and scale of migration in mainland Southeast Asian prehistoric societies (Bentley et al. 2005, 2007, 2009; Cox 2009; Cox et al. 2011). The samples so far included in this research are mainly from northeast Thailand, with the exception of Khok Phanom Di, a coastal site in southeast Thailand. It is for a pragmatic reason that these are the samples available to the research group. The sites in the northeast cover a wide temporal range from first settlers to the protohistoric period. Khok Phanom Di is unusual, in that it is coastal and was occupied at the apparent time of introduction of agriculture to the region. Bentley et al. (2007) found that most who were tested at Khok Phanom Di were locals or from a similar geological coastal area, which suggests that there was no large-scale migration. Cox’s (2009) doctoral thesis looked at relationships between three sites (Ban Non Wat, Ban Lum Khao and Noen U-Loke) in the upper Mun river valley of Northeast Thailand. To address this aim of identifying migrants she used dental non-metric traits (discrete morphological variables believed to have a genetic basis) and strontium and oxygen isotope signatures of migration. The non-metric trait analyses did not provide a clear picture of whether there was large-scale migration to the sites, or whether migration occurred from outside the region, and the lack of correlation with the isotope evidence shows the issues of dealing with data that reflects genetic relationships, rather than the direct individual evidence of isotopes (Cox 2009). The isotope evidence suggested that there was a stable population base with few people migrating from outside the region, and a reduction in migration over time. This supports the argument for a diffusion of ideas and a local role in the development of technology and agriculture (Cox 2009).

These isotopic methods can also serve to address hypotheses on genetic homogeneity, which tend to be an underlying assumption in comparative bioarchaeological analysis. An investigation of migration is also important when assessing the occurrence of novel diseases in the community (Tayles and Buckley 2004), and is a useful method to address questions about cultural practices, including postmarital residence.

Recent paleodemographic analysis at the northeast Thai site of Ban Non Wat and other comparative northeast Thai sites is also contributing to understanding population and agricultural change during this time in this region (Tayles and Halcrow in prep). The site of Ban Non Wat provides a unique opportunity to assess social and health changes, with the largest human skeletal sample ever recovered in Southeast Asia ( $n > 670$ ), covering a time span of over 2000 years (3750BP–1500BP) from the development of agriculture through to the intensification of production (Higham 2004; Higham and Higham 2009). Of particular importance is the large well-preserved sample ( $n = 81$ ) of first settlers at this site, dated from 3750–3050 years BP. Until now, only small skeletal samples, some with very poor preservation, have been found in Thailand from this time period. Tayles and Halcrow (in prep) used paleodemography to test the hypothesis that in the highly productive tropical environment, introduction of agriculture had a less profound impact on humans than in some other parts of the world (cf. Bellwood and Oxenham 2008). The comparative analysis of demography at Ban Non Wat and the other Thai sites showed no change in fertility in the earliest period of occupation. However, there is evidence for an increase in fertility in the later samples, including Noen U-Loke and the late Ban Non Wat sample, indicating a demographic change during this time. This argument of demographic change later than the period of agricultural introduction fits with other data from Southeast Asia, including the bioarchaeological evidence for a rise in infectious disease in the later ‘metal ages’ (Oxenham et al. 2005; Tayles and Buckley 2004) and archaeological evidence for increased settlement density on the Khorat Plateau and the development of hierarchical social structures from about 2500BP (O’Reilly 2008).

A factor to acknowledge in assessing paleodemographic change with agricultural intensification is the mode of agricultural development. Some archaeologists pursue grand questions about the ‘origins of agriculture’ (for example, Higham 2002), even though the archaeological record attests to the gradual nature of the change of the subsistence mode (Smith 2001). The origins models are based on an underlying dualistic construction of pre-agricultural (hunter-gatherer) and agricultural societies, with the introduction of agriculture representing a ‘revolution’ on which their research focuses, with the effect that the complexity of development of these practices is ignored (Smith 2001). Recent archaeological work has started to focus on a diverse array of societies, albeit not Southeast Asian societies, with low-level food production, that occupy the middle ground between hunting-gathering and agriculture (for a review see Smith 2001). In fact, it is argued by Chamberlain (2006:351) that only small changes in fertility occur with any transition to low-intensity horticulture or non-intensive agriculture.

Bellwood and Oxenham (2008) have recently looked for evidence of this Neolithic demographic transition, using data from East and Southeast Asia. Assuming that this occurs somewhere between 4500–4000BP, they apply a statistic measuring fertility through the proportion of immature skeletons (5–19 years of age) relative to the overall skeletal sample but excluding children under 5 years of age ( $_{15}P_5$ ) (Bocquet-Appel and Naji 2006), to data from sites in Viet Nam, Northeast Thailand and Southeast Thailand to argue that there is a demographic transition (increase in fertility) as shown by elevated values at two earlier sites (Man Bac in Viet Nam and Khok Phanom Di in Southeast Thailand). Man Bac offers opportunities to assess biological human diversity between 4000–3500BP, but

the sample is biased with a dramatically higher proportion of infants and children than most prehistoric samples (Jackes 1992), making it biologically impossible that it is representative of the cemetery population. We can therefore safely discard this aberrant value. Bellwood and Oxenham (2008) rightfully state that Khok Phanom Di has a  $^{15}P_5$  value that is higher than the majority (7 of the 11) forager populations from North America sampled by Bocquet-Appel and Naji (2006). However, it would be speculative to base an argument for some sort of Neolithic demographic shift at this time on this value alone. Bellwood and Oxenham (2008) also spend some time downplaying the importance of 'low-level food production societies', and instead favor the model of a sharp demographic transition with agriculture, when in fact their data do not indicate this.

We are not denying that the development of agriculture and sedentism in Southeast Asia must have ultimately affected demography. Certainly, there must have been demographic change, as is evidenced by the archaeological evidence of increased population density in the Mun river valley in late prehistory, in the Iron Age (Welch and McNeill 1991). Rather, the mode and duration of change, with an instantaneous leap from a 'hunter-gatherer' lifestyle to full-blown agriculturalism is a myth. As noted, there is much literature (reviewed in Smith 2001) that argues that the shift to agriculture is not a rapid change without intermediate steps between hunter-gatherers and agriculturalists. It may be that there are site-specific environmental factors that also need to be taken into account. An example of the complexities of agricultural adoption is at the site of Khok Phanom Di, which maintained a sedentary settlement with highly specialised pottery production, but was dependent on other agricultural groups for rice (Higham and Thosarat 2004a). This example shows that there is more complexity during agricultural development than the model allows. It also shows that there is no reason to hypothesise population replacement through mass migration of farmers from China (or anywhere else for that matter) because there is no evidence yet for a dramatic population increase with the integration of agriculture into a mixed subsistence mode. This is also supported by the evidence of low levels of migration shown by isotopic analysis (Bentley et al. 2005, 2007, 2009; Cox et al. 2011).

## **5 Social structure/organisation**

Stable isotope analysis of migration with an associated assessment of material culture can also contribute to assessing social structure and diversity in material culture in the region. Bentley et al. (2009) in a study of migration at the late 'Bronze Age' site of Ban Lum Khao found that social identity (as assumed through material artifacts) was associated with place of childhood origin. The isotopic evidence indicates distinct groups of women raised in different communities were associated with specific pottery types, which is not seen in the data from the males. This suggests immigration of women into the site, and there is archaeological evidence of increased craft specialisation associated with this. Bentley et al. (2009) hypothesise based on this social and possibly kin-based differentiation, in what has been interpreted to be an egalitarian community (O'Reilly 2003), that this may represent the start of complexity in the upper Mun River Valley. Bentley et al. (2009:93) state:

An ethos of exogamy and in-migration, as well as increased local specialisation and external exchange, might have been important catalysts in the development of the types of regional power centers and status differentiation that herald the emergence of early chiefdoms and states in Southeast Asian Bronze Age and Iron Age societies.

This work gives insight into how the development of agriculture, migration and increase in complexity of social structure may, as in this case, promote, or alternatively discourage diversity in material culture.

Another bioarchaeological project that has the potential to enhance understanding of the effects of agricultural development on social structure and health change at the site of Ban Non Wat uses a novel approach that incorporates aDNA evidence of kinship, archaeological evidence of burial treatment, and bioarchaeological evidence of sex, and infant and child age and health (Halcrow et al., 2008). Kinship, age, sex and status are fundamental aspects of individual social identity and therefore central in determining social organisation, which in turn affect quality of life. The analysis of aDNA is a powerful tool to identify maternal and paternal kinship within the archaeological record and represents a significant advance for testing hypotheses on social structure. Hitherto, archaeologists have relied on circumstantial evidence of the spatial relationship of burials and biodistance studies to infer kinship among individuals. Assuming aDNA is successfully recovered from the skeletons (a pilot study is currently underway), the proposed study will provide the essential baseline work for assessing the questions of kinship and social organisation.

The study of infant and child bioarchaeology can be very informative of social and environmental changes that were occurring, and therefore useful in assessing the effects that the intensification of agriculture had on past populations. Infants and children are very responsive to environmental stress and are therefore a sensitive index of population health (Lewis 2007) and their burial rites can be particularly informative of social organisation (Tainter 1978). For example, it is argued that the wealthy interment of an infant can indicate either that the community had a hierarchical social structure or that status was ascribed rather than achieved.

With the increase in archaeological excavations in MSEA it is now possible to assess questions of social status and bioarchaeology using the infant and child component of these samples (Halcrow 2006; Halcrow and Tayles 2008; Halcrow et al. 2008; Oxenham et al. 2008). Oxenham et al. (2008) incorporated an analysis of mortuary artefacts to investigate social roles of children in a paper also assessing health and disease from the late Neolithic period of Man Bac, Viet Nam. He argues that by the second year there is some sort of social milestone that children reach in that these children share certain forms of material culture with adult females, especially shell. Halcrow et al. (2008) in a bioarchaeological assessment of infant age-at-death distributions at seven Mainland Southeast Asian sites found an underrepresentation of fetal individuals at one of the sites (Khok Phanom Di), which is interpreted to be the result of the differential burial treatment of infants between this site and the other sites. It seems likely that the age distribution results from different burial rites of pre-term infants as a consequence of social and cultural differences between Khok Phanom Di and the other sites (Halcrow et al. 2008).

## **6 Dynamics of micro-macro relations**

There is an issue in archaeological research of marrying the micro scale data (for example funerary artefacts, prevalence of lesions in a sample and kinship data) with larger regional processes or macro dynamics. The nature of archaeological excavation means that bioarchaeological data is from discrete sites, and ultimately from individual skeletons (or micro-level data). To get a comparative picture of macro-level change it is necessary to build up a regional perspective. Here we present a short case study from the site of Khok Phanom Di to illustrate how multidisciplinary analysis, including different

bioarchaeological analyses, and palynological, pottery and mortuary analyses from a single site can contribute to understanding larger scale patterns in the region.

The bioarchaeological evidence from Khok Phanom Di (including metric and non-metric morphology, stature, health and disease, infant burial treatment and paleodemography) indicates that the community represented a different population from those living at inland Northeast Thai sites (Halcrow 2006; Halcrow et al. 2008; Pietrusewsky and Douglas 2002; Tayles 1996; Tayles 1999). The independence of culture, economy and technology is also well illustrated at this site (Vincent 2004; Halcrow et al. 2008). It is argued that the community traded with inland farmers for rice, with a gradual introduction of rice at this site (Higham and Thosarat 2004a), which would mean that the development of agriculture need not occur in conjunction with the replacement of either people or culture. As noted above, Bentley et al.'s (2007) isotopic investigation of migration at the site identified few migrants, suggesting that most of the sample tested were locals and that migration was small-scale. A reduction in isotope variation in females during mortuary phase 4 was interpreted as a change to matrilocality (Bentley et al. 2007). There is also evidence that these women possibly gained status through their production of pottery, probably for trade with inland farming groups, as indicated from females with rich grave goods in the sample (Vincent 2004). Although there is evidence for development of agriculture in inland sites (Higham 2002), along with the minimal isotopic evidence for immigration (Bentley et al. 2007), there is no ceramic evidence for intrusive populations from further afield than the immediate region (Vincent 2004:731). These different lines of evidence from bioarchaeology and archaeology build up a picture of a distinctive biological population gradually adopting an agricultural subsistence mode without the replacement of people or culture, and therefore contribute to understanding diversity of cultural groups at this time in MSEA. This also illustrates how detailed multidisciplinary analysis at a site, along with comparative data from other sites, can build up a picture of regional diversity. This example is counter to prominent arguments on demic diffusion, but accommodates diffusion of ideas with small-scale migration and indigenous 'agency' by emphasising the diverse social outcomes with adoption of agriculture in the region.

## **7 Linguistics, archaeology and bioarchaeology: levels of analysis**

Despite the recent increase of bioarchaeological research in Southeast Asia, there has been no attempt to synthesise or link this with linguistic research in the region. Bellwood (2007) captures the potential value and approach to drawing on different lines of evidence of specialisation in prehistory:

Biological populations, languages and cultures need not have evolved through time in absolute unison ... But nor are these categories of human variation always completely uncorrelated ... Investigating history requires an understanding of how these three sources of data can be used in a supportive way—not by circular reasoning, but by understanding how one can draw separate lines of historical information and then compare them. (Bellwood 2007:90)

It is interesting that the archaeological theories of demic diffusion models are largely based on linguistic evidence (Bellwood 2005; Higham 2002), into which evidence from material culture is incorporated (for example, Wiriyaromp 2007), or sites are categorised into pre-existing evolutionary frameworks, with no real examination of the diversity found in material culture and mortuary ritual (Higham 2002). The projects often follow an

inductive research paradigm, with the research designed specifically to prove, rather than test, a theory.

Given the palynological evidence of localised site environments (White et al. 2004) and recent archaeological interpretations that emphasise the diversity of material culture and social organisation (White 2006), continued thorough question-driven research and analysis of the data from these recent archaeological investigations is of utmost importance, as White (2006) states:

Just picture some other cultures of roughly the same time period in the Mekong basin—the Plain of Jars, Ban Chiang Late Period, the jar burial sites of the lower Mun and Chi, Phnum [sic] Snay and Noen U-Loke to name a few—and a picture of a riot of regional self-expression in everything from ceramics to burial rites to social organization comes to mind. So much is happening in the Iron Age, yet little scholarship has tried to meaningfully bring out the salient evidence in all its rich diversity. The field cannot ignore this variability much longer. (White 2006:103–104)

Future archaeological work could benefit from moving away from fitting sites into pre-existing neo-evolutionary models of agricultural and social development and assessing this variation to give a better insight into the way these prehistoric people were adapting to their environment. Using these models also ignores the agency of individuals, or this is perhaps better defined as individuals being ‘social actors’ (Cowgill 2000:52). Hodder (2000) has argued that archaeological techniques are designed for dealing with large scale and long-term processes, compared with individual acts, so the ‘agency’ is lost. He also argues that there is not a deterministic relationship between micro- and macro- processes and to do so would treat all variability of the micro-processes as ‘noise’ (Hodder 2000:26), and dismiss this variable data, which is often seen in the archaeological research on MSEA, as noted in White’s (2006:103–104) quote above.

The reliance on linguistic phylogenetic models and trees means that the reconstructions of social change are rather uni-directional and linear. In historical linguistics the analysis of distribution of languages creates an impression of grand migration theories, which work from the very top-down (Richards 2007:100). Meanwhile, prehistorians favoring models in which localised cultural changes are taken into account (White et al. 2004), (while not necessarily discounting migration) largely base their hypotheses on specific site data (for example, White 1986; White et al. 2004; White and Pigott 1996), including comparison of palynological data between sites. These approaches of investigating diversity of archaeological artifacts, environmental and cultural change within and between sites give a chance for small-scale and even individual levels of interaction and cultural processes to be acknowledged and people can be seen as ‘social actors’ in these cultural transmissions, including trade and production.

Areal linguistics (see Enfield, this volume), in contrast to building phylogenetic trees where entire language systems are passed from generation to generation, is a unit-based approach where individuals and groups are the level of analysis. Linguists using this approach are interested in language change at the intersections of people. They are interested in micro-level social dynamics, which underlie macro-level changes. This is where linguistics and bioarchaeology could ‘cross-pollinate’ somewhat. The assessment of micro-level social dynamics in areal linguistics including social contacts and interactions, are related to factors including power structures between and within societies, marriage patterns, and trade. Assessing health and disease in relationship to age and sex, or mortuary goods, can help to identify any disparities of resources within and between populations,

and therefore power structures within and between communities. Also, an analysis of bone and tooth chemistry, to assess marriage patterns (for example, Bentley et al. 2009) can give insights into power structures within and between communities and within families. This can inform gender identity within a society, and the patterns of goods used within particular social groups, and therefore give insight into production and trading of goods between individuals and social groups. These issues of power structures within and between communities, kinship and material production are all informative of social transmission, which has a bearing on language transmission.

## 8 Conclusion

Bioarchaeology, by its definition, has close ties with archaeology, and is contributing significantly to understanding social, cultural and technological change and diversity of these aspects in prehistoric Southeast Asia. This chapter illustrates that bioarchaeology can assess health and demography at the individual and population level and comparative studies of sites can begin to assess larger region-wide analyses of the effects of agricultural development on health and social organisation, migration and kinship patterns and ultimately on human diversity. Future bioarchaeological investigation of dental morphology and genetic and chemical analysis of migration has the potential to contribute further to these questions of migration, kinship and social organisation in the field of Southeast Asian prehistory. There also seems to be a real connection that can be made between bioarchaeological work and the interests of areal linguistics in social and language transmission.

## Acknowledgements

We would like to thank the conference organisers Professor Nick Enfield and Dr Joyce White for inviting us to contribute to this book. Much of our work is based on samples excavated under the direction of Professor Charles Higham, Dr. Rachanie Thosarat, and Dr. Amphan Kijngam with funding from the University of Otago, the NZ Government Marsden Fund, and Earthwatch Corps. We also thank the editors and one anonymous reviewer for their helpful comments on an earlier draft of this chapter.

## References

- Bellwood, Peter. 2005. *First Farmers: the Origins of Agricultural Societies*. Malden: Blackwell.
- . 2007. Review Feature: First Farmers: The origins of agricultural societies. *Cambridge Archaeological Journal* 17. 87–109.
- Bellwood, Peter and Marc F. Oxenham. 2008. The expansions of farming societies and the role of the Neolithic demographic transition. In J. P. Bocquet-Appel and O. Bar-Yosef (eds) *The Neolithic Demographic Transition and its Consequences*, 13–34. Paris: Springer.
- Bentley, R. Alexander. 2006. Strontium isotopes from the earth to the archaeological skeleton: A review. *Journal of Archaeological Method and Theory* 13. 135–187.
- Bentley, R. Alexander, Katharine Cox, Nancy Tayles, Charles F. W. Higham, Colin Macpherson, Geoff Nowell, Matthew Cooper and Tina E. F. Hayes. 2009. Community diversity at Ban Lum Khao, Thailand: Isotopic evidence from the skeletons. *Asian Perspectives* 48. 79–97.

- Bentley, R. Alexander, Michael Pietrusewsky, Michelle T. Douglas and Tim C. Atkinson. 2005. Matrilocality during the prehistoric transition to agriculture in Thailand? *Antiquity* 79. 865–881.
- Bentley, R. Alexander, Nancy Tayles, Charles F. W. Higham, Colin Macpherson and Tim C. Atkinson. 2007. Shifting gender relations at Khok Phanom Di, Thailand. *Current Anthropology* 48. 301–314.
- Bocquet-Appel, Jean-Pierre and Stephan Naji. 2006. Testing the hypothesis of a worldwide Neolithic demographic transition, Corroboration from American Cemeteries. *Current Anthropology* 47. 341–365.
- Boyd, William E., and Jeremy Habberfield-Short. 2007. Geoarchaeological landscape model of the Iron Age settlements of the upper Mun River floodplain. In Charles F.W. Higham, Amphan Kijngam, and Sarah Talbot (eds) *The Origins of the Civilization of Angkor, Volume 2: The Excavation of Noen U-Loke and Non Muang Kao*, 1–27. Bangkok: The Thai Fine Arts Department.
- Boyd, William E., Charles F. W. Higham and Roger J. McGrath. 1999. The geoarchaeology of Iron Age ‘Moated’ sites of the Upper Mae Nam Mun Valley, N.E. Thailand. 1: Palaeodrainage, Site-Landscape relationships and the origins of the ‘moats’. *Geoarchaeology: An International Journal* 14. 675–716.
- Buikstra, Jane E., Susan R. Frankenberg and Lyle W. Konigsberg. 1990. Skeletal biological distance studies in American physical anthropology: recent trends. *American Journal of Physical Anthropology* 82. 1–7.
- Chamberlain, Andrew T. 2006. Response to Bocquet-Appel and Naji: Testing the hypothesis of a worldwide demographic transition. *Current Anthropology* 47. 341–365.
- Cowgill, George L. 2000. Rationality and contexts in agency theory. In Marcia-Anne Dobres and John E. Robb (eds) *Agency in Archaeology*, 51–60. London: Routledge.
- Cox, Katharine J. 2009. Human migration in prehistoric Northeast Thailand. PhD thesis. University of Otago.
- Cox, Katharine J., R. Alexander Bentley, Nancy Tayles, Colin G. Macpherson, Matthew J. Cooper. 2011. Intrinsic or extrinsic population growth in Iron Age northeast Thailand? The evidence from isotopic analysis. *Journal of Archeological Science* 38. 665–671.
- Domett, Kathryn M. 2001. Health in Late Prehistoric Thailand. B.A.R. International Series 0946. Oxford: Archaeopress.
- Domett, Kathryn M. and Nancy Tayles. 2006. Human biology from the bronze age to the iron age in the Mun River Valley of Northeast Thailand. In Marc F. Oxenham and Nancy Tayles (eds) *Bioarchaeology of Southeast Asia, Cambridge Studies in Biological and Evolutionary Anthropology, No. 43*, 220–240. Cambridge: Cambridge University Press.
- . 2007. Population health from the Bronze to the Iron Age in the Mun River Valley, Northeast Thailand. In Mark N. Cohen and Gillian M. M. Crane-Kramer (eds) *Ancient Health: Skeletal Indicators of Agricultural and Economic Intensification*, 286–299. Florida: University Press of Florida.
- Douglas, Michelle T. 1996. Paleopathology in human skeletal remains from the Pre-metal, Bronze and Iron Ages, Northeastern Thailand. PhD thesis. University of Hawaii.
- Douglas, Michelle T. and Michael Pietrusewsky. 2007. Biological consequences of sedentism: Agricultural intensification in Northeastern Thailand. In Mark N. Cohen and Gillian M.M. Crane-Kramer (eds) *Ancient Health: Skeletal indicators of agricultural and economic intensification*, 300–319. Florida: University of Florida Press.



- Enfield, N. J., this volume. Linguistic diversity in mainland Southeast Asia.
- Halcrow, Siân E. 2006. Subadult Health and Disease in Late Prehistoric Mainland Southeast Asia. PhD thesis. University of Otago.
- 2008. An application of ancient DNA methods for understanding health and social changes with agricultural intensification in prehistoric Southeast Asia. *Ancient DNA and Associated Biomolecules Conference Proceedings, Pompeii*, 14 pps. Typescript, University of Otago. Submitted.
- Halcrow, Siân E. and Nancy Tayles. 2008. Stress near the start of life? Localised enamel hypoplasia of the primary canine in late prehistoric mainland Southeast Asia. *Journal of Archaeological Science* 35. 2215–2222.
- Halcrow, Siân E., Nancy Tayles and Vicki Livingstone. 2008. Infant death in prehistoric mainland Southeast Asia. *Asian Perspectives* 47. 371–404.
- Higham, Charles F. W. 2002. *Early Cultures of Mainland Southeast Asia*. Bangkok: River Books.
- 2004. Ban Non Wat: lessons from the Field. In Victor Paz (ed.) *Southeast Asian Archaeology: Wilhelm G. Solheim II Festschrift*, 491–503. Quezon: The University of Philippines.
- Higham, Charles F. W. and Thomas Higham. 2009. Dating Southeast Asian prehistory: New radiometric evidence from Ban Non Wat. *Antiquity* 83. 125–144.
- Higham, Charles F. W., and Tracey L.-D. Lu. 1998. The origins and dispersal of rice cultivation. *Antiquity* 72. 367–377.
- Higham, Charles F. W. and Rachanie Thosarat. Editors. 2004a. The Excavation of Khok Phanom Di: A Prehistoric Site in Central Thailand. Volume VII: Summary and Conclusions. London: The Society of Antiquaries of London.
- Hodder, I. 2000. Agency and individuals in long-term processes. In Marcia-Anne Dobres and John E. Robb (eds) *Agency in Archaeology*, 21–33. London: Routledge.
- Hutterer, Karl L. 1976. An evolutionary approach to the Southeast Asian cultural sequence. *Current Anthropology* 17. 221–242.
- Hutterer, Karl L. 1983. The natural and cultural history of Southeast Asian agriculture: ecological and evolutionary considerations. *Anthropos* 78.169-212.
- Jacks, Mary K. 1992. Paleodemography: problems and techniques. In Shelley R. Saunders (ed.) *Skeletal Biology of Past Peoples: Research Methods*, 189–224. New York: Wiley-Liss.
- King, Charlotte, R., Nancy Tayles, and Keith C. Gordon (in prep). Re-examining the chemical evaluation of diagenesis in human bone apatite. *Journal of Archaeological Science*, doi:10.1016/j.jas.2011.03.023
- King, Christopher A. (2008). *Stable Isotopic Analysis of Carbon and Nitrogen as an Indicator of Paleodietary Change among Pre-State Metal Age Societies in Northeast Thailand*. Oxford: BAR International Series 1788, British Archaeological Reports.
- King, Christopher A. and Lynette Norr (2006). Paleodietary change among pre-state metal-age societies in northeast Thailand: A study using bone stable isotopes. In Marc F. Oxenham and Nancy Tayles *Bioarchaeology of Southeast Asia*. Cambridge: Cambridge University Press. 241–262.
- Larsen, Clark S. 1995. Biological changes in human populations with agriculture. *Annual Review of Anthropology* 24. 185–213.
- 2002. Bioarchaeology: The lives and lifestyles of past people. *Journal of Archaeological Research* 10. 119–166.

- Lewis, Mary E. 2007. *The Bioarchaeology of Children: Perspectives from biological and forensic anthropology*. Cambridge: Cambridge University Press.
- Matsumura, Hirofumi, Marc F. Oxenham, Yukio Dodo, Kate Domett, Nguyen Kim Thuy, Nguyen Lan Cuong, Nguyen Kim Dung, Damien Huffer and Mariko Yamagata. 2008. Morphometric affinity of the late Neolithic human remains from Man Bac, Ninh Binh Province, Vietnam: Key skeletons with which to debate the 'two layer' hypothesis. *Anthropological Science* 116. 95–104.
- Matsumura, Hirofumi, Marc F. Oxenham, Nguyen Kim Thuy, Nguyen Lan Cuong and Nguyen Kim Dung, this volume, Population history of Mainland Southeast Asia: The Two Layer model in the context of Northern Vietnam.
- O'Reilly, Dougal J. W. 2003. Further evidence of heterarchy in Bronze Age Thailand. *Current Anthropology* 44. 300–306.
- 2008. Multivallate sites and socio-economic change: Thailand and Britain in their Iron Ages. *Antiquity* 82. 377–389.
- Oxenham, Marc F. 2000. Health and behaviour during the mid-Holocene and Metal period of Northern Vietnam. PhD thesis. Anthropology Department, Northern Territory University.
- Oxenham, Marc F., Nguyen Kim Thuy and Nguyen Lan Cuong. 2005. Skeletal evidence for the emergence of infectious disease in Bronze and Iron Age Vietnam. *American Journal of Physical Anthropology* 126. 359–376.
- Oxenham, Marc F., Nguyen Lan Cuong and Nguyen Kim Thuy. 2006. The oral health consequences of the adoption and intensification of agriculture in Southeast Asia. In Marc F. Oxenham and Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 263–289. Cambridge: Cambridge University Press.
- Oxenham, Marc F., Hirofumi Matsumura, Kathryn Domett, Nguyen Kim Thuy, Nguyen Kim Dung, Nguyen Lan Cuong, Damien Huffer and Sarah Muller. 2008. Health and the experience of childhood in late Neolithic Viet Nam. *Asian Perspectives* 47. 190–209.
- Oxenham, Marc F. and Nancy Tayles (eds). 2006. *Bioarchaeology of Southeast Asia*. Cambridge Studies on Biological and Evolutionary Anthropology. Cambridge: Cambridge University Press.
- Pietrusewsky, Michael and Michelle T. Douglas. 2002. Ban Chiang, A Prehistoric Village Site In Northeast Thailand I: The Human Skeletal Remains. Thai Archaeology Monograph Series 1, University Museum Monograph 111. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
- 2002a. Intensification of agriculture at Ban Chiang: Is there evidence from the skeletons? *Asian Perspectives* 40. 157–178.
- Richards, Martin. 2007. Is genetics coming between archaeology and linguistics? Review Feature response to *First Farmers: the Origins of Agricultural Societies*, by Peter Bellwood. *Cambridge Archaeological Journal* 17. 98–100.
- Smith, Bruce D. 2001. Low-level food production. *Journal of Archaeological Research* 9. 1–43.
- Stark, Miriam T. 2006. Early Mainland Southeast Asian Landscapes in the First Millennium A.D. *Annual Review of Anthropology* 35:21.1-21.26.
- Steckel, Richard H. and Jerome C. Rose (eds). 2002. *The Backbone of History: Health and nutrition in the Western Hemisphere*. Cambridge: Cambridge University Press.

- Stojanowski, Christopher M. and Schillaci, Michael A. (2006). Phenotypic approaches for understanding patterns of intracemetery biological variation. *Yearbook of Physical Anthropology*, 131(S43), 49–88.
- Tainter, Joseph A. 1978. Mortuary practices and the study of prehistoric social systems. *Advances in Archaeological Method and Theory*. 1. 105–141.
- Tayles, Nancy. 1996. Anemia, Genetic Diseases, and Malaria in Prehistoric Mainland Southeast Asia. *American Journal of Physical Anthropology* 101. 11–27.
- 1999. The Excavation of Khok Phanom Di: A Prehistoric Site in Central Thailand. Vol. V. The People. London: The Society of Antiquaries of London, Research Report L.
- Tayles, Nancy and Hallie R. Buckley. 2004. Leprosy and tuberculosis in Iron Age Southeast Asia? *American Journal of Physical Anthropology* 125. 239–256.
- Tayles, Nancy, Kathryn Domett and Siân Halcrow. 2009. Can dental caries be interpreted as evidence of farming? The Asian experience. In Thomas Koppe, Georg Meyer G, and Kurt W. Alt (eds) *Interdisciplinary Dental Morphology. Frontiers of Oral Biology* 13: 162–166. Basel: Karger.
- Tayles, Nancy, Kathryn Domett and Kirsten Nelsen. 2000. Agriculture and dental caries? The case of rice in prehistoric Southeast Asia. *World Archaeology* 32. 68–83.
- Tayles, Nancy and Siân E. Halcrow. submitted. Demographic change in Southeast Asia: new evidence from the prehistoric settlement of Ban Non Wat, Khorat Plateau, Northeast Thailand (forthcoming).
- Tayles, Nancy, Siân E. Halcrow and Kate Domett. 2007. The people of Noen U-Loke. In Charles F.W. Higham, Ampham Kijngam, and Sarah Talbot (eds) *The Origins of the Civilization of Angkor, Volume 2: The Excavation of Noen U-Loke and Non Muang Kao*, 243–304. Bangkok: The Thai Fine Arts Department.
- Vincent, Brian A. 2004. The Excavation of Khok Phanom Di, a Prehistoric Site in Central Thailand, Vol. IV. The Material Culture (Part II). Volume VI: The Pottery. London: Society of Antiquaries of London, Research Report LXX.
- Welch, David J. 1984. Settlement pattern as an indicator of sociopolitical complexity in the Phimai region, Thailand. In Donn T. Bayard (ed.) *Southeast Asian Archaeology at the XV Pacific Science Congress: The Origins of Agriculture, Metallurgy, and the State in Mainland Southeast Asia*, vol. 16. 129–151. Dunedin: University of Otago.
- Welch, David J. and Judith R. McNeill. 1991. Settlement, agriculture and population changes in the Phimai region, Thailand. *Bulletin of the Indo-Pacific Prehistory Association* 11. 210–228.
- White, Joyce C. 1986. A Revision of the Chronology of Ban Chiang and its Implications for the Prehistory of Northeast Thailand. PhD thesis. University of Pennsylvania.
- 1995. Incorporating heterarchy into theory on sociopolitical development: The case from Southeast Asia. In Robert M. Ehrenreich, Carole L. Crumley and Janet E. Levy *Heterarchy and the Analysis of Complex Societies*, 101–123. Archaeological Papers of the American Anthropological Association, Number 6. Arlington: American Anthropological Association.
- 2002. Preface. In Joyce C. White (ed.) *Ban Chiang, A Prehistoric Village site in Northeast Thailand I: The Human Skeletal Remains*, xvi–xvii. Philadelphia: The University of Pennsylvania.
- 2006. Book review of *And through Flows the River: Archaeology and the Pasts of Lao Pako*. Anna Källén. *Studies in Global Archaeology* 6. Uppsala: Uppsala University, 2004. *Asian Perspectives* 45. 102–104.

- White, Joyce C., Daniel Penny, Lisa Kealhofer and Bernard Maloney. 2004. Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. *Quaternary International* 113. 111–132.
- White, Joyce C. and Vincent C. Pigott. 1996. From community craft to regional specialisation: Intensification of copper production in pre-state Thailand. In Bernard Wailes (ed.) *Craft Specialization and Social Evolution: In Memory of V.G. Childe*, University Museum Symposium Series Volume VI, *University Museum Monographs*, 151–175. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
- Willis, Anna and Nancy Tayles. 2009. Field anthropology: application to burial contexts in prehistoric Southeast Asia. *Journal of Archaeological Science* 36. 547–554.
- Wiriyaromp, Warrachai. 2007. The Neolithic period in Thailand. PhD thesis. University of Otago.



# 4 *Linguistic diversity in mainland Southeast Asia*

---

N. J. ENFIELD

## 1 Introduction

One way to look at human diversity in MSEA is through the prism of language. With an adequate understanding of the nature of linguistic diversity, and of mechanisms of linguistic diversification, we learn what can be presumed and what remains unknown for a proper understanding of the current state of affairs in MSEA. This chapter charts a territory of possibility for the analysis of linguistic diversity grounded in social relations, outlining a particular position and motivating an agenda for research. One aim is to clear the conceptual ground for what needs to be done in research on language in order to understand the special degree of structural convergence that the languages of MSEA appear to have undergone.

Research on human diversity often focuses on historical questions of how a modern state of affairs came to be the way it is. This presumes a satisfactory account of that modern state of affairs. What is the nature of linguistic diversity in mainland Southeast Asia today? While this question is normally answered in terms of patterns of genealogical relatedness among languages of a region, I want to address other measures of diversity as well, including structural diversity of MSEA languages. The situation of linguistic diversity in MSEA appears to be special among regions of the world (Dahl 2008), and for this reason begs explanation. In this chapter I note some of the characteristics of linguistic diversity in MSEA, and consider these in terms of the socially-grounded causal processes that are known to give rise to patterns of linguistic diversity over time. This addresses another goal: to point to the kinds of underlying, social processes of diversification that we want to keep in mind when discussing history and diversity across disciplinary boundaries.

## 2 Linguistics

In the kinds of interdisciplinary collaborations on human history and diversity that have inspired the present volume, ‘linguistics’ means comparative/historical linguistics. A central aim of this branch of linguistics is to establish phylogenies of language diversification, both in the sense of grouping (establishing that languages are related) and subgrouping (establishing how they are related).<sup>1</sup> The establishment of language

---

1 I am grateful to Malcolm Ross for insisting on this distinction and for suggesting the wording that I have used here.

phylogenies is done typically if not exclusively with reference to word forms that are taken to be cognate across modern languages (for example Thai *ram* versus Lao *ham* ‘bran’). If a linguist can establish regularities in sound change across groups of words, this may be used as evidence for a proposed phylogeny, where the evolving entity is the language as a whole. It is a cladistic approach (Moore 1994). The number of words that serve as data points is a subset of the number of words in the language, ranging depending on the researcher and the research method from a small fraction, say one or two hundred words, to a sizable portion of the lexicon, say a few thousand words. While it is always a subset, it is nevertheless taken to be sufficient for inferring the history of the entire language, and, by implication, the history of the community that has spoken the language and its earlier forms.<sup>2</sup> In this way, inferred linguistic phylogenies may serve as hypotheses for phylogenies of human groups, to be tested against other kinds of data (cultural, environmental, biological, genetic, etc.).<sup>3</sup> Such cross-checking requires interdisciplinary collaboration, and such collaboration is one objective of the present volume. It is one step in a long march, and as Blench (2008) advises, we are still at the early stages of achieving true interdisciplinary collaboration.

A phylogenetic approach to language relatedness assumes vertical transmission for languages as whole systems. Through vertical transmission, an entire language structure is passed on from generation to generation, with minor changes in each generational cycle. But like other anthropologists, linguists acknowledge processes of horizontal transmission as well—the borrowing of elements from one language into another **within** generations—and we need to account for these processes too. Accordingly, a special concern of historical/comparative linguistics is to distinguish between the results of vertical versus horizontal transmission (cf. for example Aikhenvald & Dixon 2001). When two languages share a given structure, is it because they each derive it from a common ancestor (common inheritance) or because one or both of the languages has borrowed the structure (diffusion)? One view is that horizontal transmission effects obscure the real signal of interest, the signal of language phylogeny. But the special effects of horizontal transmission are of no less interest in characterising linguistic diversity, particularly where—as is the case in MSEA—a relative **lack** of diversity is part of what needs to be explained. As it happens, historical processes in language can also be readily viewed as rhizotic, as Moore (1994) puts it, that is, involving hybridisation, implying quite different underlying processes from those represented in phylogenetic trees.

As a complement to work in the comparative/historical tradition, this chapter draws on two other areas of linguistics: language typology and sociolinguistics. Language typology asks: Are there universals in language structure? Are there dependencies between types of structure? What generalisations can be made? Typology provides a set of measures of diversity in language (see Comrie 1989; Shopen 1985, 2007; Croft 2003; Haspelmath et al. 2005). Sociolinguistics asks: How is variation in linguistic behaviour related to the position of speakers within a social system? What role do social networks play in determining language variation? What are the causal relations between social life and language

---

2 I make no assumption that the phylogeny of a human community coincides with the phylogeny of a language spoken by members of that community at any point in history.

3 Also of relevance in using linguistics to test hypothesised scenarios of human descent are inferences that can be made from the presence of words with certain meanings, and what their presence implies about the history of the community (e.g. indigenous words for certain species of plant or animal whose geographical distribution may be limited: Evans 2010, Ch. 6; cf. discussion in this volume in chapters by Blench, Sagart, van Driem, Sidwell & Blench, and Diffloth).

structure? (See Weinreich et al. 1968; Labov 1972; Trudgill 1974, 1997; Milroy 1980.) This chapter considers the nature of human diversity in MSEA today through the lenses of linguistic typology and sociolinguistics together.

### 3 Typological convergence of language in mainland Southeast Asia

Convergence in the structure of neighbouring languages due to social contact between communities is a global phenomenon (see Weinreich 1953; Emeneau 1956; Muysken 2008; Thomason 2001; Silva-Corvalán 1994 and many others). It results in what are referred to as linguistic areas. Another common term for linguistic area is *Sprachbund*. In global terms, MSEA is a remarkable example of a linguistic area. Five major language families are found in MSEA: Tai-Kadai, Sino-Tibetan, Hmong-Mien, Austroasiatic, Austronesian. (How these families are related to each other is a matter of discussion by other contributors to this volume, including Blench, Diffloth, van Driem, Sagart, and Sidwell & Blench.) Languages from these different families show massive convergence in structure at every level (Matisoff 1973, 2001, 1991a; Clark 1985, 1989, 1996; Clark & Prasithrathsint 1985; Bisang 1991, 1999; Enfield 2003, 2005; Grant & Sidwell 2005, among many others). The relevant levels of linguistic organisation are as follows:

#### *Levels of linguistic organisation*

- phonetics/phonology: sounds and sound systems
- morphology-syntax: internal structure of words/phrases/sentences
- lexicon: words, their meanings, and relations among them
- pragmatics: patterns in language usage and interpretation

The convergence among languages in MSEA is so thoroughgoing that the typologist Östen Dahl has labelled the area ‘the ultimate Sprachbund’ (Dahl 2008:218). Using data from the *World Atlas of Language Structures* (‘WALS’, Haspelmath et al. 2005), Dahl computed a pairwise measure of typological distance among a set of over 200 languages. The typological distance between two languages is an expression of how similar they are on a range of structural measures. This numerical expression of distance ranges from a low of 10 (Dutch versus German) to a high of 75 (Ju’hoan, a Khoisan language of Botswana and Namibia versus Central Yup’ik, a Yup’ik language of Alaska). The mean figure for typological distance among the world’s languages is 42 (for example English versus Persian). Dahl’s measure spans all domains of linguistic structure for which WALS supplies data. These include features of sound structure as well as presence and nature of grammatical patterns (for example, whether number is obligatorily marked in the grammar—yes for English, no for Khmer—or, for an ‘adjective-plus-noun’ type structure, whether the adjective comes before the noun as in English *new village* or after it as in Khmer *phuum thmei*).

The typological distance measures for MSEA are striking. For example, Dahl compared Hmong, a language of the Hmong-Mien group, spoken in China, Thailand, Laos, and Vietnam, with Khmu, a language of the Northern Mon-Khmer group, spoken in Thailand, Laos, and Vietnam. Even though Hmong and Khmu are entirely unrelated, and have been in direct contact only relatively recently (perhaps a few centuries), on Dahl’s typological



distance measure they score about the same as German and English, two closely related languages:

- (1) Hmong vs. Khmu = 22.5  
German vs. English = 21.1

In a more extreme example, genealogically unrelated languages of MSEA that have had more intensive contact—Thai, Khmer, and Vietnamese—measure as close to each other as Polish and Russian, among the typologically closest pairs on Dahl's global scale:

- (2) Thai vs. Vietnamese = 11.4  
Thai vs. Khmer = 12.3  
Polish vs. Russian = 12.8

These figures indicate a dramatic macro-level convergence in grammatical structure between unrelated or very distantly related languages that has taken place over centuries in MSEA. What kinds of micro-level, real-time causal processes have brought this convergence about? A fair amount is known of the micro mechanisms that must ultimately be involved in the macro aggregation of linguistic change (Weinreich, Labov & Herzog 1968; Keller 1994; Enfield 2008). But it is difficult to get a grip on these enormous processes of language transmission and change in real time, and virtually impossible to do so in historical retrospect (Enfield 2003). We are, however, beginning to get a reasonable idea of the general causal anatomy of language transmission and change (as an instance of cultural transmission and change more generally), both within and across generations, such that we may in due course be able to relate the micro-processes of transmission to their macro-level effects (see Rogers 1995; Ross 1997; Enfield 2008 and references therein). Let us now place the question of MSEA's linguistic diversity in a broader account of the nature of linguistic diversity in general, drawing primarily on Nettle (1999).

## 4 Linguistic Diversity

Linguistic diversity can be measured in three ways (Nettle 1999), always with reference to a geographical area: 1. The **language diversity** of an area is the number of distinct languages spoken in it; 2. The **phylogenetic diversity** of an area is the number of distinct language families found in that area; 3. The **structural diversity** of an area is the degree of typological difference between the languages. These are not three different ways of measuring the same thing, but represent different senses of linguistic diversity that may be independent from one another.

### 4.1 Language diversity

Language diversity in MSEA is high compared to many regions of the world. This is consistent with observed high language diversity in areas with similar environmental conditions to MSEA. There is a statistical association between tropical environments and high language diversity: 'language diversity tends to be greatest near the equator' (Nettle 1999:61). Nettle suggests a causal account: Because a tropical environment affords economic self-sufficiency, human groups can afford to maintain greater socio-economic distance from their neighbours. This independence allows greater differentiation in social identity between groups, of which language distinctness is a primary indicator. Groups will

still come into social contact under these conditions, but they will maintain more distant types of social relation, characterised less by reliance on exchange of fundamental economic resources. Exchange will be more specialised, including ritual exchange. In such conditions, Nettle argues, ethnic distinction is allowed to flourish.

Nettle (1999) insists on a connection between society, economy, and ecology.<sup>4</sup> He generalises two kinds of social bonds: primary and secondary. Primary bonds are ‘relatively enduring, are often formed early in life, and are multivalent’. They are the bonds on which people depend ‘for their basic livelihood’ (Nettle 1999:67). Normally, our first language is learned from people with whom we have primary bonds (Thomason & Kaufman 1988:11), biasing faithful vertical transmission in which genes, culture, and language follow a single historical path, despite their logical distinctness (Boas 1911:12 and *passim*). Secondary bonds ‘are based on specific functions, such as a trade in a specialised good’. These bonds ‘are associated with greater social distance than primary ones and are more typical of the relationships between ethnolinguistic groups than those within them’ (*ibid.*). Such relations are typical of neighbouring ethnic groups (linguistically defined) almost anywhere in MSEA.<sup>5</sup>

Nettle argues that an appropriate theory for explaining the latitudinal trend ‘will be one that links human agents to their ecological setting’ (1999:69). He illustrates with a case study of equatorial horticulturalists in Interior New Guinea, where the ecology affords an ‘enormous potential for self-sufficiency’. This self-sufficiency allows language groups to be very small, and therefore more numerous. Social relations between members of language groups in this sort of setting are of the secondary type. This is contrasted with case studies from sub-Saharan Africa where a very different ecology brings about primary social links between spatially distant households. Nettle writes of Hausa: ‘The wide extent of the language must surely have its origin in the wide extent of these [primary social] links, which are in turn a response to the dangerous highs and lows of the agricultural calendar’ (1999:77).<sup>6</sup> If this account is correct, the hypothesis for MSEA will be that neighbouring groups have tended not to maintain primary social links across ethnolinguistic lines, a possibility arising from the low ecological risk afforded by the area’s tropical environment.<sup>7</sup>

Language diversity in MSEA is observed to different degrees within different language families. In Laos, for example, the Austroasiatic language family shows greater language diversity than other language families—that is, there is a higher number of Austroasiatic languages than other languages, where each of these Austroasiatic languages is spoken by a smaller number of speakers. It should be possible to account for the different degrees of language diversity across language families in terms of socio-economic history and inter-group social relations, with multiple determining factors (including ecology). The problem

4 Nettle restricts his account to ‘the post-Neolithic, but pre-Industrial, world’ (1999:96).

5 There is room for further distinction of types of social relation, beyond Nettle’s primary versus secondary bonds; Evans & McConvell (1997) speak of neighbouring groups in Australia as being either ‘isolated’ or ‘linked’; Thomason & Kaufman (1988) speak of more or less ‘intense’ relations between groups; Ross (1997) speaks of groups being more or less ‘tight-knit’; and so on.

6 Roger Blench (in personal communication) is skeptical, citing languages with 200 speakers spoken in the same area as Hausa.

7 Note that beyond developing certain types of social links with neighbours, there are other strategies for dealing with ecological risk, including livelihood diversification (certainly relevant for MSEA; cf. White 1995, this volume), storage, and mobility. People who specialise in diversity or mobility—hunter-gatherers—may thereby offset any tendencies to rely on primary social bonds outside the group.

is to find evidence independent from language for these historical patterns of social dynamics.

## 4.2 Phylogenetic diversity

The languages of MSEA are from five distinct families, whose distinctness is well established: Tai-Kadai, Sino-Tibetan, Hmong-Mien, Austroasiatic, and Austronesian (cf. Diller et al. 2008; Matisoff 1991b; Thurgood & LaPolla 2003; Ratliff 2010; Blust 2009; Grant & Sidwell 2005). Considerable controversy surrounds the internal structure of these families (how many sub-groups, what are they, etc.), as well as their possible grouping at higher levels (for example, whether Tai and Austronesian share a common ancestor).<sup>8</sup> On this point I defer to the chapters in this volume written by historical/comparative linguists (Blench, Diffloth, Sagart, van Driem, Sidwell & Blench).

Nettle attributes to Nichols (1990) the idea that ‘stock diversity is an increasing function of the time since founding’ of distinct populations (Nettle 1999:119). He takes issue with Nichols’ assumption of constant rate of ramification (citing Nichols 1990:503). Nettle writes, ‘linguists have no rigorous or widely accepted method of dating the split of phylogenetic groupings’, stating that ‘the rate of diversification is actually rather variable’ (Nettle 1999:120). Different types of speech-community event (Ross 1997) clearly have different dynamics, for social reasons.

Nettle’s proposals concerning the causal role of social relations among neighbouring groups are based on an idealised notion of equilibrium, a relatively static social state of affairs with established patterns of economic and ritual contact between interacting groups. But of course social relations are dynamic. How to capture this? Dixon (1997) takes a punctuated equilibrium model—earlier applied to biological speciation (Eldredge & Gould 1972) and then cultural diversification (Bellwood 1996)—and applies it to the diversification of languages. The idea is that during equilibrium periods, ethnic groups live alongside each other with a rather unconstrained process of diffusion of features between languages. Punctuation arises from cataclysmic social events that trigger the split and expansion of groups (Dixon 1997:67; cf. Nichols 1992).<sup>9</sup>

The modern day states of affairs that we are presently trying to disentangle were initially caused by dramatic expansions of Neolithic societies into open or effectively open territory (Nichols 1992; Nettle 1999). Such scenarios provide the bread and butter of the comparative method: sub-families separated by the split and spatial separation of sub-communities. The problem of course is that the split and spatial separation (and subsequent domination by expanding groups over others) is seldom neat and tidy. Moore (1994:15) notes the special nature of the Polynesian situation which so neatly fits a cladistic approach:

---

8 An inherent problem for interdisciplinary research on the dynamics of human diversity arises from limitations in the time depth of historical/comparative linguistics’ reach. Once we are looking further than 8000 years or so into the past, the traditional comparative method ceases to be effective. But archaeologists and biological anthropologists are interested in significantly greater time-depths (see chapters in this volume by White, Halcrow & Tayles, Oppenheimer, Matsumura et al., Bulbeck, and Fix). Many hope that this gap may be bridged by the application in historical linguistics of statistical methods from evolutionary biology (see Burenhult et al. this volume, and references therein).

9 The distinction between equilibrium and punctuation does not imply different underlying processes of social diffusion at the fundamental micro-level (Enfield 2005:194-7).

While an ocean limited contact among neighbouring Polynesian ethnoi [= distinct ethnic groups], thereby preserving the integrity of ethnic boundaries, continental populations are constantly surrounded, buffeted, and confronted by their neighbours. Unless they are contiguous to an ocean or desert, continental populations are normally surrounded by foreign ethnoi, often of several ethnic and linguistic types. (Moore 1994:15)

But despite the continental nature of the MSEA populations with all their interethnic buffeting, there is surprisingly crisp separation of languages in different linguistic families of the region.<sup>10</sup> While higher-level distinctions between families may be controversial, there is virtually no case of an MSEA language whose membership in one or another established family is disputed.

### 4.3 Structural diversity

Every language contains a huge inventory of features, including tens of thousands of words, hundreds of smaller components of words like *-s* for plural in English (as in *dogs* vs. *dog*), scores if not hundreds of grammatical patterns for combining those words (for example active versus passive sentences in English), patterns of word meaning, and habits of language usage (greetings, speech formulas, etc.). The most accessible data for calculating structural diversity are in phonology (sound structure) and morphosyntax (patterns of building words and combining them into phrases and sentences). Here is a list of some technical linguistic features that are common to all or most MSEA languages, demonstrating the area's low degree of structural diversity (see Clark 1985, 1989, 1996; Clark & Prasithrathsint 1985; Bisang 1991, 1999; Enfield 2003, 2005):

*Some phonological features shared across all MSEA language families*

- very high number of vowels (relative to consonants)
- common underlying structure of vowel system (often 9-place, symmetrical)
- systematic distinction in vowel length (long versus short)
- preference for one (major) syllable per word
- laryngeal features lexically contrastive (tone, phonation type)<sup>11</sup>
- many fewer consonants possible in final than initial position
- gap in voiced stop series at velar place of articulation (no voiced 'g')

---

10 I am grateful to Roger Blench for insisting on this point in personal communication.

11 Tone and phonation type are treated here as instances of a single sound system property because they each involve the use of laryngeal features for lexical contrast. While tone and phonation type are often considered to be fundamentally distinct phenomena, in fact most systems that are identified as one versus the other (in phonological terms) actually display properties of both (in phonetic terms; Henderson 1967:171). Pitch contours, distinctions in phonation type, and glottalic effects are all produced in the larynx (specifically, by the vocal folds), and are all articulatorily independent of segmental speech sounds produced with the lips, teeth, and tongue (i.e., typical 'consonants'). Tone and phonation are intimately bound, not essentially distinct, and for this reason I do not regard the sound system of a classical MSEA tone language such as Vietnamese to be of a different species from that of a classical MSEA register language such as Kri (Enfield & Diffloth 2009).

*Some morphosyntactic features shared across all MSEA language families*

- no inflectional morphology (no agreement, no case, no gender/number/definiteness on noun phrases, no tense-marking on verbs)
- zero anaphora (free ellipsis of arguments if contextually retrievable)
- prevalence of verb-object constituent order
- topic-comment structure in clauses
- large set of labile verbs (verbs that can be both transitive or intransitive)
- rich inventories of sentence-final particles
- rich inventories of expressives (or ideophones)
- numeral classifiers and related systems of nominal classification
- verb serialisation, richly multifunctional

Some of the most noteworthy commonalities among MSEA languages are their complete **lack** of values on certain parameters (most notably, ‘no inflectional morphology’, a property which denotes the absence of a great range of structures found in languages of the world, for example agreement, case, gender/number/definiteness on noun phrases, tense-marking on verbs). It might be said that a common **presence** of features would be more convincing evidence of convergence across languages. What is the significance of this shared lack of features? Consider how it is that languages may lose grammatical structure over time. According to one hypothesis, a way to remove structure from a language is to have a community of speakers learn the language as adults. (The inverse is also true: one is likely to **add** structure to a language by learning it as a child; Trudgill 2008.) This effect—where an adult speaker’s native language affects the structure of the new language—is called substratum interference.<sup>12</sup> According to Thomason and Kaufman,

in this kind of interference a group of speakers shifting to a target language fails to learn the target language perfectly. The errors made by members of the shifting group in speaking the target language then spread to the target language as a whole when they are imitated by original speakers of that language. (Thomason & Kaufman 1988:39)

The long-term effect of these changes on the target language comes about because the changes are learned by children in the next generation. In this way, widespread learning of new languages in adulthood (for example in large-scale language shift during migration or under strong domination from another ethnic group) is one driving force of structural convergence.<sup>13</sup>

---

12 My use of terminology is as follows. When there is interference between two languages due to social contact of their speakers, the substrate language or substratum is the language of the politically subordinate group. The superstrate is the language of the politically dominant group.

13 LaPolla (2001) points in this direction in his assessment of the situation in China, with massive convergence between varieties of Chinese arising from massive migration over centuries; but see Ansaldo & Matthews (2001) against the idea that Chinese structure is evidence of earlier creolisation. Part of the problem relates to ideological baggage belonging to the term ‘creolisation’. Ansaldo & Matthews posit ‘hybridisation’ instead.

If we are to understand why MSEA shows an especially low degree of structural diversity we want to know how it **came to be** like this. Processes of language change follow the same causal processes of innovation diffusion that underlie culture evolution more generally (Rogers 1995). An important feature of these processes are the biases that promote or inhibit diffusion (Boyd & Richerson 2005; Enfield 2008).

One of these biases is called a context bias (Enfield 2008:303). A social innovation will diffuse more effectively if the adopter of the innovation has a ready context into which it fits. For example, the necessary context for adopting downloadable ringtones is the presence and use of mobile telephones and the Internet—if members of a social group do not already have mobile telephones and the Internet, they will not adopt downloadable ringtones. The idea of a context bias is applicable to just about any technology. In language, a context bias favours diffusion of less ‘embedded’ linguistic items, since they are less dependent on structures that are specific to a particular language. This means, for example, that interjections like *Wow!* will diffuse more readily than grammatical markers, since the interjections do not have a language-specific grammatical context. For any given linguistic innovation, the grammar of the borrowing language is a potential set of brakes on diffusion and convergence.

Another source of constraints on diffusion is a content bias (Boyd & Richerson 2005). A content bias favours the diffusion of innovations which have some payoff for the adopter that the alternatives lack (for example certain agricultural practices will be of special interest to potential adopters if those practices offer greater yield for less effort; certain handtools will be more readily adopted if they are easier to use). In language, a content bias will favour adoption of ‘unmarked’ structures, that is, structures that are simpler, more transparent, and more frequent.

In general, because they are harder to learn, universally marked features are less likely than unmarked features to be transferred in language contact. ... In shift situations, this works two ways: shifting speakers are likely to fail to acquire marked features of the target language, and marked features carried over by shifting speakers from their original language are relatively unlikely to spread by imitation to the target language as a whole. (Thomason & Kaufman 1988:51)

Of course, simplification is not the only force in language change. Were it so, then all languages would be maximally and equally simple. Languages complexify, too, if given time, and specifically, if given a sufficient number of generations of normal transmission by which children who learn the language as a first language are able to effectively add structure and complexity to it (Thomason & Kaufman 1988; Trudgill 2008).

The MSEA situation of low structural diversity across the languages, and relative absence of many structural features within the languages, is compatible with a long history of widespread language shift with continuous bilingualism. This suggests a conclusion to be drawn from the MSEA linguistic facts. If the effects of large-scale contact are overwhelmingly in the direction of grammatical simplification, then this may have come from a historical context in which adults widely learn and use the languages of neighbouring groups, yet while keeping sufficient distance from those groups such that children are not heavily embedded in multilingual learning settings. This kind of social context may be called **ethnic pluralism**, as it requires the co-presence and interaction of a plurality of ethnic groups, yet where the distinctness of their respective identities is maintained as a matter of common preference.

An important goal is to uncover the causal processes that underlie language change, differentiation, and convergence alike. Any historical account of a region's human diversity must be compatible with a proper understanding of these causal processes. Such processes are critical not only to understanding the MSEA facts, but to understanding cultural processes in general. With this background on the social context of language, we turn to the current MSEA situation.

## **5 Linguistic diffusion in mainland Southeast Asia**

In MSEA today, there are widespread asymmetrical social relations between language communities, with levels of nesting (Blench 2005). One class of language consists of the major national languages: for example, Thai, Lao, Khmer, Vietnamese, Chinese, Malay. These major languages are spoken by huge numbers of people. They are written, formalised, standardised, taught in schools, used in media such as newspapers, television, official correspondence and law. Large numbers of people are monolingual in these languages, but they are also second (non-native) languages for many speakers of ethnic minority languages. The recent existence of national languages is a major factor in obscuring the information we have about the linguistic past of MSEA, both through the acceleration of attrition and loss of ethnic minority languages, and through interference with scientific discourse and analysis thanks to the political ideologies that national languages introduce. In these respects, the major national languages are the least representative of the languages of the area. Vast numbers of people have shifted over recent generations to these languages, and this process continues in full swing.

The remaining languages—that is, most of the languages—of MSEA are greater in number, and are spoken by much smaller communities. Within the 'smaller' languages, there is another level of nesting, where some languages that are relatively widely spoken have de facto or semi-official status as major languages in local terms: for example Karen in Myanmar and Thailand, Khmu in Laos, Cham in Cambodia, Tai Daeng in Vietnam, varieties of Zhuang in Southern China. Each of these second-tier languages has some degree of official and administrative recognition, and each is spoken by large numbers of people (from several hundred thousand in the case of Khmu to millions in the case of Zhuang). Still, the fact that they are not national languages means that their speakers will tend to be bilingual in a national standard language. As noted already, the national language phenomenon has been recently imposed upon the general linguistic scene that was in place before the 20<sup>th</sup> century (cf. Smalley 1994). That said, we have no reason to doubt that nested multilingualism was the norm in pre-nationalist MSEA.

While all minority languages are politically subordinate to the national languages of MSEA, they are not equally subordinate. Ignoring the national languages, we still see political asymmetries between minority languages at local levels. As an example, take Kaleum District in Sekong Province of Laos. Kaleum is a small, isolated upland district. Several languages of the Katuic sub-branch of Austroasiatic are spoken there. These languages are each politically subordinate to Lao, the official national language. So, most natives of Kaleum also speak Lao to some degree, and are in some circumstances obliged to do so (for example using Lao as a lingua franca when travelling outside of their home district or when dealing with visiting officials). But within daily life in their home district, there is a recapitulation of the language dominance relation at a more local level. One of the local languages—Ngkriang (also known as Ngeq)—belongs to a locally dominant ethnic group, and serves as a lingua franca for between-group dealings. So, if you are of the Ngkriang ethnicity, then you will speak Lao with people outside your district and

Ngkriang with people inside your district, including both fellow Ngkriang and people of other local ethnicities. If, however, you are member of one of those other ethnicities, you will speak at least three languages: (1) Lao with people from outside your district, (2) Ngkriang with people of other ethnicities inside your district, and (3) your own language with your own people.<sup>14</sup>

Such nested political asymmetries of neighbouring languages is the general pattern in MSEA. Social relations are typically asymmetric, and the asymmetry tends to be oriented in a downhill direction (Diller 2004; Blench 2005). The higher upstream you live, the less politically dominant you are, and the more likely you will be to accommodate to the language of your downstream neighbours. In the Kri-speaking area at the peak of the Nam Noy valley in eastern Khammouane Province, Laos, Kri speakers inhabit the highest reaches of a major watershed, and they are the most multilingual of all their neighbours (Enfield & Diffloth 2009). Kri men mostly speak proficient Sek, the Northern Tai language of their immediate neighbours downstream. The Sek tend to speak little or no Kri in return. But the Sek, in turn, speak Lao with their downstream compatriots, and the Lao, in turn, who speak no Sek, speak Thai to their neighbours over the Mekong to the West.

One account for the special degree of linguistic convergence observed in MSEA appeals to horizontal diffusion through sustained social contact between language groups, in a more ‘rhizotic’ account of language history (Moore 1994). As suggested by the discussion so far, to understand these patterns of contact, we have to look at the structure and dynamics of social relations, not within language groups, but across language groups. Recent research on language contact and its effects (for example by Thomason 2001; Ross 2003; Aikhenvald 2002; Muysken 2008; *inter alia*) has underlined the primacy of inter-community social factors in determining structural linguistic outcomes. This work offers empirically-based distinctions between types of inter-community relations. Direct social contact between groups—a necessary condition for horizontal transmission—is more likely when there is more loose-knit, open social organisation in a given group. Once two groups come into regular contact, the type and intensity of this contact will be a function of the nature of social relations between the two groups. In MSEA, it is a kind of ethnic pluralism defined above that will most likely account for the observed patterns.

A useful scheme for thinking about differences between types of inter-group contact is a scale from symmetrical to asymmetrical social contact (cf. Thomason 2001; Aikhenvald 2006).<sup>15</sup> Social contact between groups is more symmetrical when each group has a similar degree of control over local power and resources. The languages of the two groups will be mutually influential, each contributing structure to the other, resulting in gradual convergence. By contrast, social contact between groups is more asymmetrical when one social group wields significant power over the other, particularly when one group displaces the other, coming into control of land and other resources through military force, technology, or other means. In these conditions, speakers of a substrate language (the language of the subordinate group) may either find their language being heavily affected by the superstrate language (especially in vocabulary) or they may shift to the superstrate entirely. Language shift of this kind is currently in full swing in MSEA among many minority populations, especially those who speak Austroasiatic languages.

Our knowledge of what determines the likelihood and rate of language shift is limited. One factor concerns the language attitudes of a speech community. Some communities

---

14 People of Kaleum also know Vietnamese because of proximity to the border and contact with traders.

15 I doubt that social contact between different groups is ever truly symmetrical. The scale is perhaps better characterised as a scale of asymmetry, running from maximally to minimally asymmetrical.



view their language as ethnically emblematic, and go to special lengths to make sure that the younger generation learns and uses it (for example Hmong in Laos). Other communities are apparently willing to let their language disappear. Differences in language ideology in communities of the past may account for modern day facts, but unfortunately, while it is possible to determine ethnolinguistic ideology through contemporary ethnography, it is not clear how this can be done for past states of affairs through archaeological or comparative linguistic research.

Large-scale language shift does not mean that a superstrate language simply replaces a substrate. In language shift, large numbers of people learn the superstrate language as adults and as multilinguals. Factors of both adult second-language learning and of multilingual speech contribute to the transformation of a language through contact. As noted above, when adults learn new languages, they do so imperfectly, and often with the result of simplifying the language learnt. If this simpler variety serves as input to the next generation of language learners, a net result can be simplification of the superstrate language as a whole.

Multilingual speech can also have a structural effect on language. In multilingual settings, multilinguals practice code-mixing, that is, they use multiple languages at the same time, often within the same utterance (Muysken 2000). Because such within-utterance language mixing involves the interlocking of components of two or more languages, the languages involved will tend to structurally converge where possible in order to better facilitate this (cf. Weinreich 1953; Silva-Corvalán 1994). The modern MSEA situation is compatible with a past scenario of long-term and widespread practices of code-mixing in multilingual settings. In line with this, Khanittanan (2001) argues that the C14-18 ‘Khmero-Thai’ society of Ayuthya in present day Thailand was fully bilingual, driving convergence of the two languages to an extreme degree (Huffman 1973).

If the above observations are correct, we may hypothesise that the MSEA facts—a very high degree of linguistic convergence combined with relative simplicity of grammatical structure—are to be explained by a history of widespread adult learning of neighbouring languages (indicative, say, of large-scale and repeated migrations) and widespread code-mixing in multilingual environments, as facilitated by a widespread ethnic pluralism.

## 6 Conclusion

The puzzle in MSEA, like in every region, is often seen to be that of distinguishing between two kinds of cause for the existence of common structure in languages: internal vs. external change, vertical vs. horizontal transmission, descent vs. diffusion. But do we not overestimate the extent to which these are competing accounts? Perhaps the reason that it is difficult to detect and maintain these distinctions is because the distinctions are weak or even illusory. There is an essential commonality to the causal processes that underlie them all, namely unit-based diffusion taking place in real time conducted through social interaction (Enfield 2003:368, 2005:190–198, 2008:304). Resolving the issue of ‘vertical’ versus ‘horizontal’ transmission—a distinction that may be fundamentally questioned—is a challenge for future research.

This chapter has addressed the question of human diversity in mainland Southeast Asia from the viewpoint of linguistics. While most linguistic research focuses on phylogenetic diversity (number of language families in an area, and putative relations of common descent between them), I have pointed to two other senses of diversity (Nettle 1999): language diversity (number of languages in an area) and typological diversity (similarity or difference in structure of languages in an area). MSEA appears to show an exceptionally

low degree of typological diversity in world terms (Dahl 2008). A likely explanation combines multiple factors, including (a) a tendency for MSEA peoples to maintain ethnic distinctions through language yet nevertheless cultivate close and sustained social connections across ethnic boundaries, and (b) a tendency for the isolating/analytic type of language found in MSEA to accelerate convergence through heightened facilitation of code-mixing. These hypotheses emerge partly from what is known ethnographically about the linguistic situation in modern MSEA, and partly from what is known about the linguistic effects of different types of social situation on linguistic diversity globally. To be plausible as a natural, causal account, any account must be statable in terms of the basic, micro-level mechanisms that underlie not just linguistic transmission but all cultural transmission. A priority for future research in MSEA is close ethnographic investigation of the micro-level processes of language transmission, and through this, the development of an account that links these micro-level processes to their macro-level products.

### Acknowledgements

I am grateful for comments from Roger Blench, Bernard Comrie, Gérard Diffloth, David Gil, Pieter Muysken, Malcolm Ross, and Chip Zuckerman, as well as all of the participants in the ‘Dynamics of Human Diversity’ workshop in Siem Reap, January 2009. This work was supported by the Max Planck Institute Nijmegen, the Wenner-Gren Foundation for Anthropological Research, and the European Research Council.

### References

- Aikhenvald, Alexandra Y. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.
- 2006. Grammars in contact: a cross-linguistic perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds) *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra & R. M.W. Dixon (eds). 2001. *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford: Oxford University Press.
- Ansaldo, Umberto & Stephen J. Matthews. 2001. Typical creoles and simple languages: the case of Sinitic. *Linguistic typology* 5 (2/3). 311–324.
- Bellwood, Peter. 1996. The origins and spread of agriculture in the Indo-Pacific region: gradualism and diffusion or revolution and colonization? In David R. Harris (ed.) *The origins and spread of agriculture and pastoralism in Eurasia*, 465–498. London: UCL Press.
- Bisang, Walter. 1991. Verb Serialization, grammaticalization and attractor positions in Chinese, Hmong, Vietnamese, Thai and Khmer. In Hansjakob Seiler & Waldfried Premper (eds) *Partizipation: das Sprachliche Erfassen von Sachverhalten*, 509–562. Tübingen: Gunter Narr Verlag.
- 1999. Classifiers in East and Southeast Asian languages: Counting and beyond. In Jadranka Gvozdanovic (ed.) *Numeral types and changes worldwide*, 113–185. Berlin: Mouton de Gruyter.
- Blench, Roger M. 2005. From the mountains to the valleys: Understanding ethnolinguistic geography in SE Asia. In Laurent Sagart, Roger M. Blench & Alicia Sanchez-Mazas (eds) *Perspectives on the Phylogeny of East Asian Languages*, 31–50. London: Curzon Press.
- 2008. Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? In Alicia Sanchez-Mazas, Roger Blench, Malcolm Ross, Ilia

- Peiros & Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics, and genetic*, 105–132. London: Routledge.
- this volume, The role of agriculture in the evolution of mainland Southeast Asian language phyla.
- Blust, Robert. 2009. *The Austronesian Languages*. Canberra: Pacific Linguistics.
- Boas, Franz. 1966 [1911]. Introduction. In Franz Boas (ed.) *Handbook of American Indian Languages*, 1–79. Lincoln, NE: University of Nebraska Press.
- Boyd, Robert & Peter J. Richerson. 2005. *The origin and evolution of cultures*. New York: Oxford University Press.
- Bulbeck, David, this volume, Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya.
- Burenhult, Niclas, Nicole Kruspe & Michael Dunn, this volume, Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula.
- Clark, Marybeth. 1985. Asking questions in Hmong and other Southeast Asian languages. *Linguistics of the Tibeto-Burman Area* 8(2). 60–67.
- 1989. Hmong and areal Southeast Asia. In David Bradley (ed.) *Southeast Asian syntax*, 175–230. Canberra: Pacific Linguistics.
- 1996. Where do you feel? Stative verbs and body-part terms in Mainland Southeast Asia. In Hilary Chappell & William McGregor (eds) *The grammar of inalienability*, 529–564. Berlin: De Gruyter.
- Clark, Marybeth & Amara Prasithrathsint. 1985. Synchronic Lexical Derivation in Southeast Asian Languages. In Suriya Ratanakul, David D. Thomas & Suwilai Premsrirat (eds) *Southeast Asian Linguistic Studies presented to André G. Haudricourt*, 34–81. Bangkok: Mahidol University.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*, 2nd edn. University of Chicago Press.
- Croft, William. 2003. *Typology and universals*, 2nd edn. Cambridge: Cambridge University Press.
- Dahl, Östen. 2008. An exercise in ‘a posteriori’ language sampling. *Sprachtypologie und Universalienforschung* 61(3). 208–220.
- Diffloth, Gérard, this volume, Austroasiatic word histories: boat, husked rice and taro.
- Diller, Anthony V.N. 2004. *Mountains, rivers and seas: Ecology and language history in Southeast Asia*. Paper read at 14th Southeast Asian Linguistics Society Meeting, May 2004, at Bangkok.
- Diller, Anthony V.N., Jerold A. Edmondson & Yongxian Luo (eds). 2008. *The Tai-Kadai Languages*. London: RoutledgeCurzon.
- Dixon, R. M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Eldredge, Niles, & Stephen Jay Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In Thomas J.M. Schopf (ed.) *Models in Paleobiology*, 82–115. San Francisco: Freeman Cooper.
- Emeneau, Murray B. 1956. India as a linguistic area. *Language* 32. 3–16.
- Enfield, N. J. 2003. Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia. London: RoutledgeCurzon.
- 2005. Areal linguistics and Mainland Southeast Asia. *Annual review of anthropology* 34. 181–206.

- . 2008. Transmission biases in linguistic epidemiology. *Journal of Language Contact* THEMA 2. 295–306.
- Enfield, N. J. & Gérard Diffloth. 2009. Phonology and sketch grammar of Kri, a Vietic language of Laos. *Cahiers de Linguistique - Asie Orientale* (CLAO), 38(1). 3–69.
- Evans, Nicholas. 2010. *Dying words: Endangered languages and what they have to tell us*. Chichester: John Wiley and Sons.
- Evans, Nicholas & Patrick McConvell. 1997. The enigma of Pama-Nyungan expansion in Australia. In Roger Blench & Matthew Spriggs (eds) *Archaeology and Language II: Correlating archaeological and linguistic hypotheses*, 174–193. London/New York: Routledge.
- Fix, Alan, this volume, Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model.
- Grant, Anthony & Paul Sidwell (eds). 2005. *Chamic and beyond: Studies in mainland Austronesian languages*. Canberra: Pacific Linguistics.
- Halcrow, Siân & Nancy Tayles, this volume, Human Diversity in Mainland Southeast Asia: the contribution of bioarchaeology.
- Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Henderson, Eugenie J.A. 1967. Grammar and tone in South East Asian languages. *Wissenschaftliche Zeitschrift der Karl-marx-Universität Leipzig* 16(1/2). 171–178.
- Huffman, Franklin E. 1973. Thai and Cambodian - a case of syntactic borrowing? *Journal of the American Oriental Society* 93(4). 488–509.
- Keller, Rudi. 1994. *On language change: The invisible hand in language*. London/New York: Routledge.
- Khanittanan, Wilaiwan. 2001. Khmero-Thai: the great change in the history of the Thai language of the Chao Phraya Basin. In Somsonge Burusphat (ed.) *Proceedings of the 11th Meeting of the Southeast Asian Linguistics Society*, 375–392.
- Labov, William. 1972. *Language in the inner city: Studies in the Black English Vernacular*. University of Pennsylvania Press.
- LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In Alexandra Y. Aikhenvald & R. M.W. Dixon (eds) *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 225–254. Oxford: Oxford University Press.
- Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.) *Southern California occasional papers in Linguistics, No. 1*, 72–95. Los Angeles: University of Southern California.
- . 1991a. Areal and universal dimensions of grammaticalization in Lahu. In Elizabeth C. Traugott & Bernd Heine (eds) *Approaches to grammaticalization*, 383–453. Amsterdam: John Benjamins.
- . 1991b. Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology* 20. 469–504.
- . 2001. Genetic versus contact relationship: prosodic diffusability in Southeast Asian languages. In Alexandra Y. Aikhenvald & Robert M.W. Dixon (eds) *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 291–327. Oxford: Oxford University Press.

- Matsumura, Hirofumi, Marc F. Oxenham, Nguyen Kim Thuy, Nguyen Lan Cuong & Nguyen Kim Dung, this volume, Population history of Mainland Southeast Asia: The Two Layer model in the context of Northern Vietnam.
- Milroy, Leslie. 1980. *Language and social networks*. Oxford: Basil Blackwell.
- Moore, John H. 1994. Ethnogenetic theory. *National Geographic Research and Exploration* 10(1). 10–23.
- Muysken, Pieter C. 2000. *Bilingual speech: a typology of code-mixing*. Cambridge: Cambridge University Press.
- (ed.). 2008. *From linguistic areas to areal linguistics*. Amsterdam: John Benjamins.
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Nichols, Johanna. 1990. Linguistic diversity and the first settlement of the New World. *Language* 66. 475–521.
- 1992. *Linguistic diversity in space and time*. Chicago: Chicago University Press.
- Oppenheimer, Stephen, this volume, MtDNA Variation and Southward Holocene Human Dispersals within Mainland Southeast Asia.
- Ratliff, Martha. 2010. *Hmong-Mien language history*. Canberra: Pacific Linguistics.
- Rogers, Everett M. 1995. *Diffusion of innovations*, fourth edn. New York: The Free Press.
- Ross, Malcolm. 1997. Social networks and kinds of speech-community event. In Roger Blench & Matthew Spriggs (eds) *Archaeology and language I: theoretical and methodological orientations*. London/New York: Routledge.
- 2003. Diagnosing prehistoric language contact. In Raymond Hickey (ed.) *Motives for language change*, 174–198. Cambridge: Cambridge University Press.
- Sagart, Laurent. This volume, The Austroasiatics: East to West or West to East?
- Shopen, Timothy, ed. 1985. *Language typology and syntactic description*, 3 vols. Cambridge University Press.
- (ed.). 2007. *Language Typology and Syntactic Description*, 2nd edn. Cambridge: Cambridge University Press.
- Sidwell, Paul & Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Smalley, William A. 1994. Linguistic diversity and national unity: language ecology in Thailand. Chicago: University of Chicago Press.
- Thomason, Sarah Grey. 2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thurgood, Graham & Randy J. La Polla (eds). 2003. *The Sino-Tibetan languages*. London: Routledge.
- Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- 1997. Typology and sociolinguistics: linguistic structure, social structure and explanatory comparative dialectology. *Folia Linguistica* 23(3/4). 349–360.

- 2008. *Contact, transmission, and sociolinguistic typology*. Paper delivered at conference ‘Transmission and diffusion’, MPI Nijmegen & Radboud University Nijmegen, January 17–19, 2008.
- Weinreich, Uriel. 1953. *Languages in contact*. New York: Linguistic Circle of New York.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In W. Lehmann (ed.) *Proceedings of the Texas conference on historical linguistics*, 97-195. Austin: University of Texas Press.
- White, Joyce. 1995. Incorporating heterarchy into theory on socio-political development: The case from Southeast Asia. In Robert M. Ehrenreich, Carole L. Crumley and Janet E. Levy (eds) *Heterarchy and the analysis of complex societies*, 101-123. Arlington, VA: American Anthropological Association.
- this volume, *Cultural Diversity in Mainland Southeast Asia: a view from prehistory*.



# 5 *MtDNA variation and southward Holocene human dispersals within mainland Southeast Asia*

---

STEPHEN OPPENHEIMER

## 1 Introduction

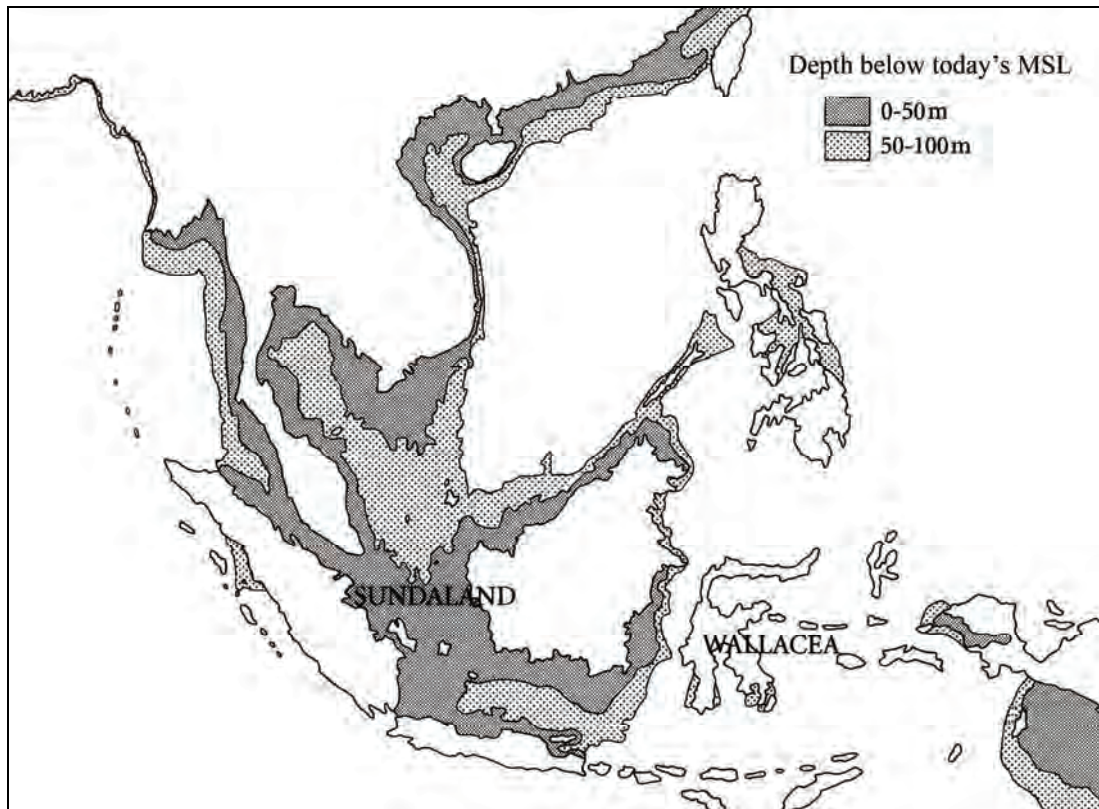
This review concerns genetic evidence for human migrations in Mainland Southeast Asia (MSEA) during the Holocene. The Malay Peninsula is today still connected to the rest of MSEA by a long thin strip of land, at its narrowest point called the Isthmus of Kra. Thus, although the isthmus connects it to the northern part of MSEA, the Malay Peninsula is generally thought of as genetically and culturally more part of Island Southeast Asia (and is referred to as the ‘Malay Peninsula’ here), while the rest of MSEA north of the Malay Peninsula (referred to for brevity and convenience throughout the rest of this chapter as ‘Indo–China’), having long borders with southwest China, is sometimes seen as having its genetic and cultural leanings further north in China, in spite of the fact that the aboriginal languages of both parts of MSEA are predominantly Austro-Asiatic.

When periods prior to the Early Holocene are considered, the geographical (and genetic) relationships between Indo-China, the Malay Peninsula and Island Southeast Asia (ISEA) were much closer, since they were all physically merged as a large Asian landmass. During all of the Late Pleistocene, modern human occupation of Southeast Asia (SEA), MSEA and *most* of Island SEA (the Greater Sundas of Borneo, Sumatra, Java and Bali) were joined together as part of a vast East Asian sub-continent, the so-called Sundaland (Figure 1) (Oppenheimer 1998). The Malay Peninsula was joined to Sumatra and Borneo until 8,400 years ago, and people could also walk straight across from southern Vietnam directly to the east coast of West Malaysia, avoiding the Gulf of Bangkok—then a lake—until just after the beginning of the Holocene. This situation not only provided major land bridges between MSEA and ISEA, but also brought them physically much closer. This proximity has implications for routes of north-south migrations and probably also for the incipient Neolithic of the two regions which, from recent work in Indo-China and Borneo, may have started much earlier than previously recognised.

Recent study of the settlement prehistory of ISEA demonstrates that mitochondrial DNA (mtDNA) diversity in the region is high and includes a large number of unique indigenous Pleistocene clades (genetic haplogroups). Many of these clades date back to the time of first settlement over 50,000 years ago, but the majority appear to mark dispersals in the terminal Pleistocene and Early-Holocene, most likely triggered by postglacial flooding



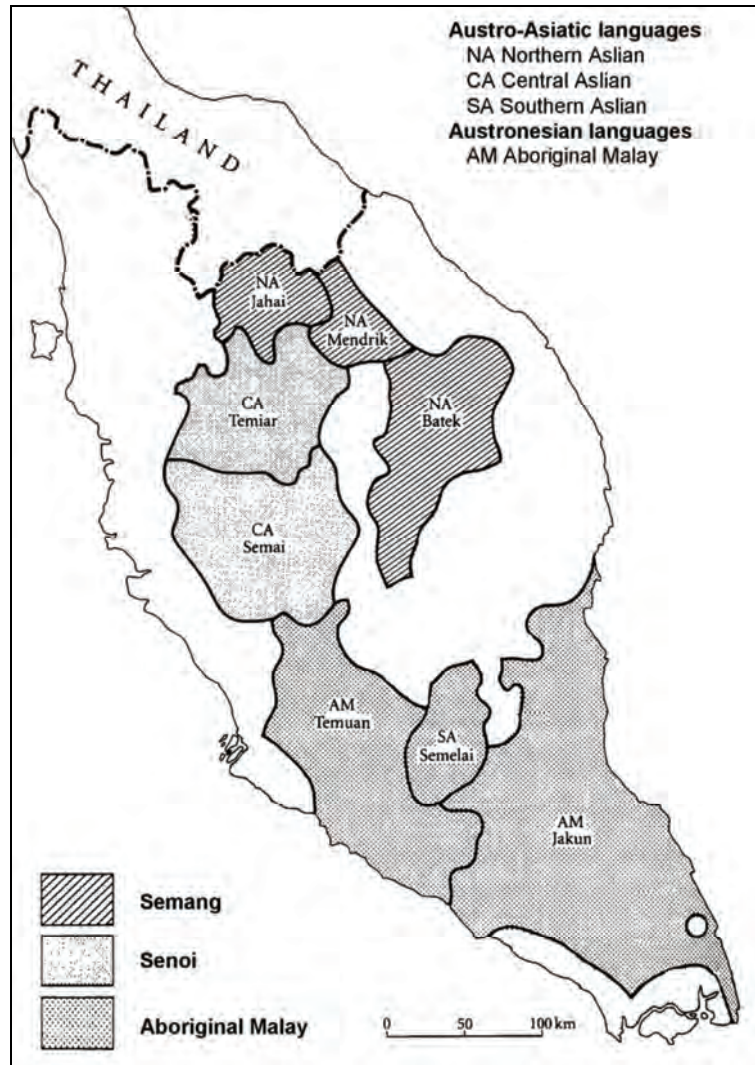
(Hill et al. 2007). A substantial proportion of these dispersals consisted of derivatives from endogenous clades, whose genesis coincided with rapid sea-level rises during both pre-Holocene and Early Holocene periods (Hill et al 2007; Soares et al 2008). There was also substantial genetic interaction between the Philippines and Taiwan but, contrary to popular archaeological models based on the Austronesian linguistic phylogeny, most of this involved gene-flow south-north from ISEA to Taiwan during the Early Holocene (Hill et al 2007; Soares et al 2008). The bulk of the exogenous dispersals into ISEA in fact appear to have arrived from Indo-China in the north, by land bridge down the Sunda shelf (Hill et al. 2007).



**Figure 1:** Map to show geographic extent of the emerged Sunda Shelf (Sundaland) at depths of 50m and 100m respectively below modern Mean Sea Level. Greatest worldwide sea level fall at Last Glacial Maximum, c. 20,000 years ago: 127m. (after Figure 13, Oppenheimer, 1998).

Further north, MSEA comprising Indo-China and the southern Malay Peninsula provides contrasts, both culturally and genetically but also in perceptions of its prehistory. Genetically, culturally and linguistically, Indo-China exhibits extraordinary complexity both prehistoric and historic. Its genetic diversity is so complex that it will be some time before its phylogeographic relationships with South China are resolved in terms of the exact directions and times of multiple migration events of the past 10–20,000 years (see for example Li et al 2007). However, for MSEA in general, as for ISEA, preliminary evidence shows that the bulk of mtDNA lineages had their genesis locally and before the Holocene (Li et al 2007; Mormina 2008). This conservative perspective obviously has implications for any view of substantial migration of rice farmers south from China or India into MSEA during the Neolithic (compare Fix, Diffloth, and Sidwell & Blench, but see Sagart: all this

volume). Picking out lineages entering MSEA during the Holocene from the north or south is likely to be difficult, both because the amount was probably small, and also because of the background noise of to-and-fro movement within MSEA. Thus, looking for movement from Indo-China into the southern Malay Peninsula may offer a filtered perspective on migrations in this complex region during the Holocene.



**Figure 2:** Map to show former distribution of Orang Asli (aboriginals of the Malay Peninsula) and their respective languages and traditional forms of subsistence economy within the past 100 years (after Figure 33, Oppenheimer 1998). *Note: the only Orang Asli groups shown here are those for whom relevant mitochondrial DNA data is available and discussed in the text.*

The Malay Peninsula is more accessible to study genetically than Indo-China. Indigenous populations are older; others are rather less complicated ethnically, and very different. On the one hand, there are the three mainstream groups making up the bulk of the modern population, Chinese and Southern Indians deriving mainly from 20<sup>th</sup> century immigrations, and the largest of the three groups made up of indigenous coastal wet-rice growing Melayu Malays (35%). The latter have been presumed under the current Archaeo-linguistic orthodoxy to be rice farmers from ISEA that arrived in the Late Holocene, although regional genetic variation actually indicates considerable indigenous diversity and

local genetic retention as well as some Holocene migration from Indo-China (see below and Hill et al. 2006, 2007).

On the other hand, there are three broad aboriginal groupings of Orang Asli—literally original people: Semang, Senoi and Aboriginal Malay (see map Figure 2) of acknowledged great local antiquity reaching well over 50,000 years back into the Pleistocene (Macaulay et al. 2005, cf. Part 3 of this book). The Orang Asli are tiny minorities (representing around 0.5% of the Malay Peninsular population—Hill et al. 2006), mainly speaking related Aslian Austro-Asiatic languages and living isolated existences in the central upland jungles. The result of small size and isolation has been tight bottlenecks (a severe reduction in overall size of an isolated population) and genetic drift (a reduction in genetic diversity resulting from a population bottleneck) in all three Aslian populations resulting in preservation of distinctive local polymorphisms at high frequency but relatively low diversity.

Clear differences are seen in the physical/morphological features of these three main recognised Orang Asli groups, which co-categorise largely with differences in their lifestyle and language sub-types, thus suggesting discrete ancestral prehistories, possibly influenced by external intrusion. It is the core aim of this review to explore alternative models of these prehistories with a focus on the Holocene, using a genetic phylogeographic approach and geographically informative mtDNA haplogroups/haplotypes; and from this to reconstruct the Holocene intrusions the Orang Asli may have received from the ancestors of the mainstream populations of MSEA, the Melayu Malays, and the peoples of Thailand, Laos, Cambodia and Vietnam as a photoshot of the latter.

In summary, the aim of this indirect approach is to infer aspects of the prehistory of expansive colonisation by the mainstream regional populations of MSEA during the Holocene, from their distinctive genetic admixtures into the isolated pre-existing Orang Asli gene pools, thus avoiding the afore-mentioned genetic complexity and antiquity, as evident in Indo-China.

While not prejudging or necessarily accepting any of the existing models of Orang Asli ethnogenesis based on more traditional methods, it is important to acknowledge the contribution of physical anthropology, ethnography and linguistics in framing their ethnological categories. Social and physical anthropology are reviewed, with citation of prior studies, in Hill et al (2006), and in Burenhult et al. and Bulbeck (both this volume), but a brief summary is given here.

It should be noted that the morphological terms used below are chosen for their common usage and not as racial stereotypes, nor to under-estimate the range of variation they each represent. Terms such as ‘Mongoloid’ (further division into northern and southern Mongoloids, although valid, is not relevant to this discussion), ‘non-Mongoloid’ and ‘Negrito’ are used here to broadly differentiate the various regional populations of MSEA from each other and those of East Asia and Australasia (Table 1). Full discussion of usage of such terms and their application, and justification based on the morphological literature (for example Bulbeck 2000, 2003; Rayner & Bulbeck 2001) is given in Oppenheimer 2003 (chapter 5 and table 5.1).

The Semang ‘Negrito groups’ are generally acknowledged to be the least changed in all respects, physical and cultural, since their first settlement in the region. Until recently, they practiced a nomadic hunter-gatherer lifestyle and retain more undifferentiated features from common ancestors in Africa (Oppenheimer 2003), notably uniformly very dark skin colour, tightly crinkled hair, ‘non-Mongoloid’ cranial morphology and undifferentiated dental morphology (Bulbeck 2003; Bulbeck et al. 2005; Bulbeck & Lauer 2006; Rayner & Bulbeck 2001; Hill et al. 2006). The Senoi are sedentary and grow hill-rice and, although

‘non-Mongoloid’ and preserving some degree of the aforementioned physical features, show less characteristic archetypes and, without tending towards Mongoloid types, are altogether more pleomorphic (of varying form) with respect to these features (Bulbeck 2003; Bulbeck et al. 2005; Bulbeck & Lauer 2006; Rayner & Bulbeck 2001). The most parsimonious explanation for these differences between Semang and Senoi has been illuminated by the genetics (discussed in detail below) and posits that Semang and Senoi share common local population ancestry but the ancestors of the Senoi have undergone quantitatively more genetic and cultural admixture from outside than the Semang, in particular from Indo-China (see also Fix and Burenhult et al. this volume).

**Table 1:** Morphology key

<b>Terminology used in this article and regional comparison populations of East Asia and Australasia</b>	<b>Terminology focused on regions of geographically related indigenous populations</b>
Negritos	Andaman Islanders, Semang, Philippine ‘Aeta’
Australoids	Aboriginal Australians including Tasmanians
Melanesians	New Guinea, Bismarck Archipelago, Vanuatu, New Caledonia
Mongoloids (Southern)	Southeast Asians (excluding Negritos), Polynesians, Micronesians
Mongoloids (Northern)	Northeast Asians (excluding Ainu), New World populations

Analogies for this apparent difference in intrusive admixture can be seen in the Aslian languages. They are overall Austro-Asiatic and sub-labelled ‘Aslian’ (but presumed Neolithic intrusive from Indo-China—see below), and are divided into related ‘Northern Aslian’ dialects for the Semang and ‘Central Aslian’ for the Senoi. Note however that Burenhult et al. (this volume) warn against too rigid over-interpretation of such received correlations between culture, language and genetics, when all three moieties are at least partly intrusive, and there are thus mismatches at the category edges.

Unlike the Semang or the Senoi, the Aboriginal Malays are horticulturalists (Bulbeck 2003). They typically grow root crops, which probably preceded rice in MSEA, and live in the south of the Peninsula (map Figure 2). They speak several languages according to their location (See Burenhult et al. this volume). In the extreme south of the peninsula they only speak dialects of Malay, although farther north some speak ‘Southern Aslian’, for instance, the Semelai.

## **2 Malayan prehistory**

This review of Holocene movements in MSEA and their impact on aboriginal groups is multidisciplinary in that it addresses evidence from archaeology, linguistics, physical anthropology, ethnography, oceanography and genetics and is intended to be accessible to all those disciplines. However since the focus is on the last discipline, which also seems most arcane to outsiders, a brief simplified overview on methods used here is warranted, and has been requested by the editor.

The greatest advances in genetic tracing and measuring migrations over the past three decades have used samples from living populations to reconstruct past migrations. Such

research goes back to the discovery of blood groups, but male Y-chromosomes (discrete, male-associated packets of DNA found inside the cell nucleus) and female-transmitted mitochondrial DNA (found in small bodies outside the nucleus in the cell plasma) are currently the most fruitful markers to study migrations, since they do not recombine (that is, do not get maternally and paternally contributed DNA sequence data spliced and mixed up at each generation), thus preserving clear uncorrupted uni-parental *lines* of descent. The result of this rare property is that a detailed uni-parental *tree* of descent (for example Figure 3) can be reconstructed for each gender, backwards from many living DNA samples, using accumulated point mutations as branching flags, stretching right back to the earliest days of Anatomically Modern Humans nearly 200,000 years ago.

Geographic study of this detailed uni-parental tree (for example Figure 3) of genetic lineages is the core of genetic phylogeography. Put in simple terms, genetic phylogeography lays the tree (more like an ivy plant) of lineages<sup>1</sup> on the map to see where lineages arose and analyse which ones went where and when. When a population migrates, different individuals carry multiple peripheral ‘twigs’ from the ‘genetic ivy plant’ and establish new colonies; these founding twigs eventually mutate to create unique local sub-clades of new growth. It is the diversity of fresh mutations on an unique founding twig which is used to date the founding time for that sub-clade. Since each migrating lineage carries a suite of unique mutational branch signatures from its immediate geographic source population, the ultimate degree of lineage intrusion or admixture into the recipients, from the recent ancestral source, can be directly counted in the modern population, rather than inferred mathematically. This method, which follows gene lineages rather than populations, is very different and more specific to purpose than traditional population genetics, which was used to compare similarities or differences between whole sample populations by comparing the relative frequencies of numerous unrelated genetic markers. The mitochondrial genetic tree for Southeast Asia is shown schematically in Figure 3.

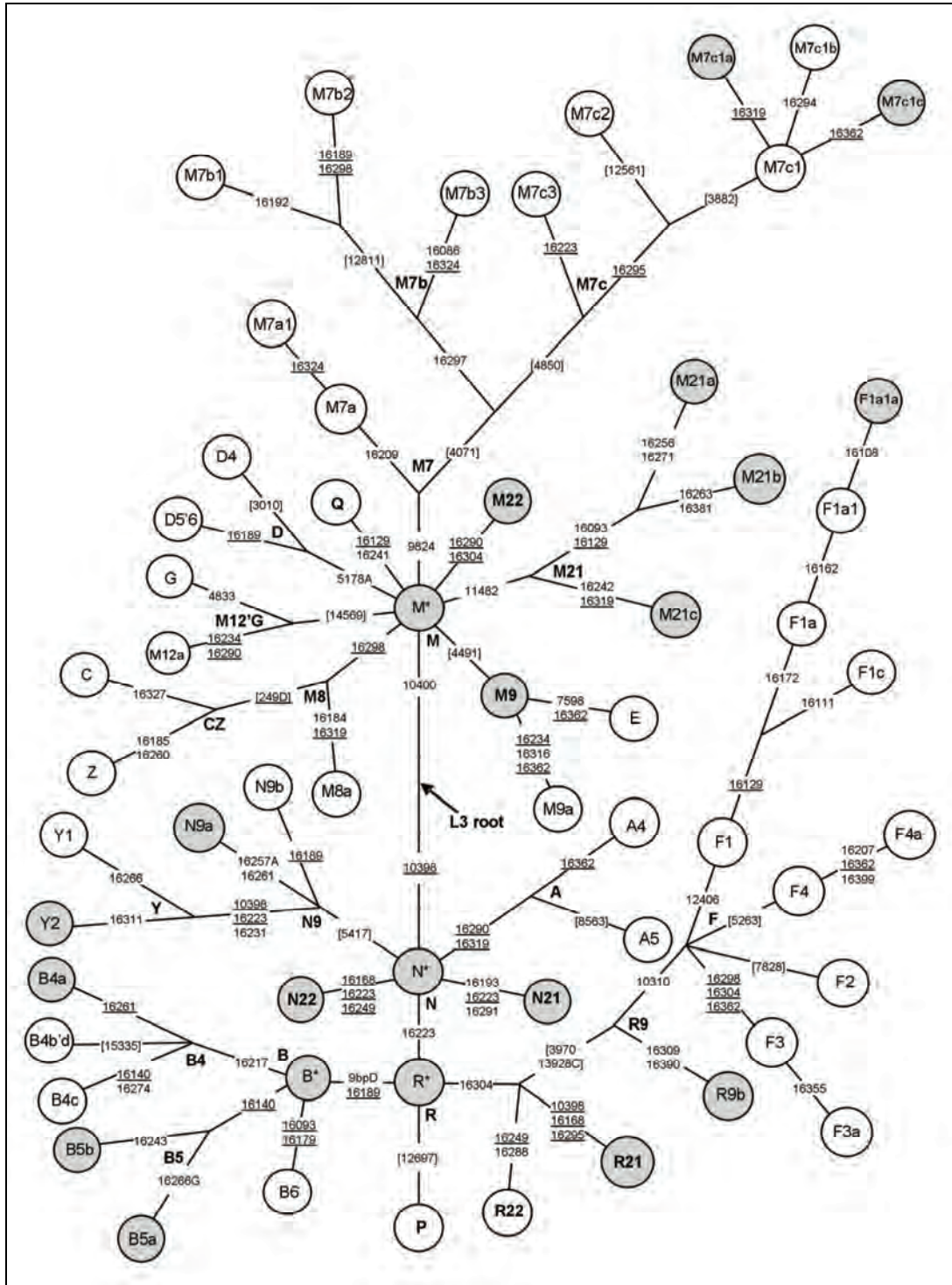
## 2.1 Dating

As implied above, the diversity of accumulated *locally unique* new mutations clustering round a founding mtDNA twig ‘haplotype’ (that is, an mtDNA unique sequence, which can be used to define the root of a founding sub-clade) in a target geographic region can be used to estimate time elapsed since that lineage’s founding event. This approach is based on the observation that new mutations occur randomly, but at a constant rate. Over the past 14 years, two dating methods using different parts of the mitochondrial genome with independently calibrated mutation rates have been used widely, with reasonably consistent and comparable results in the literature.

The older and more widespread of these (Forster et al. 1996; Saillard et al. 2000) addresses and compares the short, non-coding ‘Hypervariable segment 1’ (HVS1) sequence of mtDNA. Its main continuing advantage, apart from cheapness, is that HVS1 mutational information is now available for tens of thousands of samples worldwide (larger numbers of samples inevitably help improve dating precision), including many from Southeast Asian communities, hence nearly all of the dates reviewed here, especially MSEA Holocene ones, use that method.

---

1 Syn: haplogroups, branches or clades and sub-clades defined by uniquely shared mutations, that is, innovations.



**Figure 3:** Schematic tree of major Southeast Asian mtDNA haplogroups, contextualising those found in the Malay Peninsula. Branches are defined by inherited mutations, so diagnostic HVS-I and coding-region markers tested are indicated on each branch. Additional coding-region motif positions are shown in parentheses. Underlined mutations occur more than once in the tree. Discs represent haplogroups, i.e. groups or clades of mtDNA haplotypes sharing a common ancestor with the same unique suite of mutations. Shaded haplogroups are those found in the Orang Asli (after Hill et al. 2006).

Using a non-coding segment means that dating of HVS1 is evidently much less affected by the confounding effects of natural selection than the coding region (see below) and has more mutational information for a given sequence length (although the very 'hypervariable' nature of the segment introduces occasional ambiguities in the branching structure). The most important continuing disadvantage of HVS1, apart from reduced branching information, is that such a very short DNA segment offers much less overall mutational information in a given branch than complete sequencing, thus also partly limiting resolution and precision and accuracy of dating, and offsetting the advantage of sample numbers.

In contrast to the HVS1, the coding region of the mitochondrial genome is 56 times longer thus having much more potential for mutational information; although it is fortunately much less hypervariable at any site than HVS1, thus nearly completely removing branch ambiguity. Taking into account both the longer sequence and a much slower overall mutation rate per coding site, the coding region as a whole still has four times as much mutational information as the HVS1, thus offering more information per sample, and improving precision while at the same time removing ambiguities and increasing branch diversity. For these reasons coding region dating has been regarded as the gold standard for dating for most of the past seven years. Luckily the main calibration used for the coding region (Mishmar et al. 2003) produced parallel dates for the major clades that were largely comparable with HVS1 dating. The main disadvantage until recently was that due to expense and laboriousness of sequencing, only a small number of samples had had their coding region completely sequenced (somewhat over 2000 worldwide in 2009), again limiting geographical application and affecting precision of date estimates.

The initial situation of full confidence in coding region dating has seen two pendulum swings, negative then positive, over the past few years after the theoretical problem, that functional (i.e. coding) sequences are subject to natural selection, became a demonstrated reality in coding mtDNA. Study of existing phylogenetic trees showed that non-lethal mutations, that nevertheless genetically changed protein composition, were commoner in younger peripheral branches than in deeper older branches. The explanation was, as expected, that minimally deleterious so-called non-synonymous mutations were being progressively 'purified out' by natural selection over a few thousand years (Soares et al. 2009). The effect of this process was to introduce a degree of non-linearity in the calibration curve for younger branch ages, which straightened out for older ages.).

## **2.2 Dating correction**

Since the purifying effect, once recognised, was regular, predictable and relatively easy to study on existing phylogenies, it was possible to correct for mathematically without losing effective information. As a result of a large systematic analysis of over 2196 complete mtDNA sequences, recalibration of the entire mitochondrial genome allowed the maximum use of existing data and an improved human mitochondrial molecular clock by correcting for purifying selection (Soares et al. 2009). An Excel-based calculator supplied in the publication allows correction of estimates both from complete sequence data and from mtDNA genome segments such as the coding region and HVS1. This correction may be regarded as analogous to 'corrected Radiocarbon dates', and comparable in size of Holocene corrections, except that the mtDNA correction is based on a mathematically regular process and algorithm, whereas radiocarbon correction factors are anything but regular being based on empirical comparisons of varying prehistoric natural radiocarbon

levels (for example using dendrochronology). Three new date estimates introduced in this review use this correction, since the newly published calibration method is likely to change dating practice in future.

### **2.3. Sampling and surveys**

Field sampling is technically quite easy and non-invasive, since blood sampling ceased to be necessary owing to improvements in sensitivity of laboratory methods. Either non-sterile saliva or inside cheek brushings suffice and DNA extraction can take place later in the laboratory and the DNA code is sequenced. Much more time-consuming than the collection is the process of getting permission and approval for the specified research and sampling at all levels from central government and involved universities to regional and local bodies and of course informed consent from the participant. Within an identified community of interest, samples are taken randomly from unrelated individuals. The reason for this is that multiple samples from individuals closely related on the same locus (site in the genome for example mtDNA) will give a biased estimate of their shared type. So ideally, within a family, only the father and mother are sampled to maximise yield and minimise haplotype duplication, thus giving two different types of mtDNA and one Y chromosome. If the mother is not available, any of her children would suffice for mtDNA. Similarly if the father is not available any male child should carry his Y chromosome, but not his mtDNA. Detailed ethnographic information is taken for each individual including language, and their relationships to other persons sampled (if any).

Systematic sampling and study of mitochondrial DNA in Southeast Asian populations using older indirect techniques such as Restriction Fragment Length Polymorphisms (RFLP; for example Ballinger et al. 1992) goes back nearly two decades, but large-scale sequencing of the mtDNA in MSEA and among Orang Asli is much more recent and has mainly been carried out by the Richards group, whose publications include cited data from other groups, wherever available, to increase sample numbers. These extra data citations are credited in our main cited papers (for example Hill 2005; Hill et al. 2006, 2007), which may thus be regarded as compendia of relevant data. Since dates used here from those publications are often based on data from several studies, for reasons of space the extra data points in those publications are not re-cited here.

## **3 Unique local mtDNA lineages characterising Peninsular Malaysian Aboriginals**

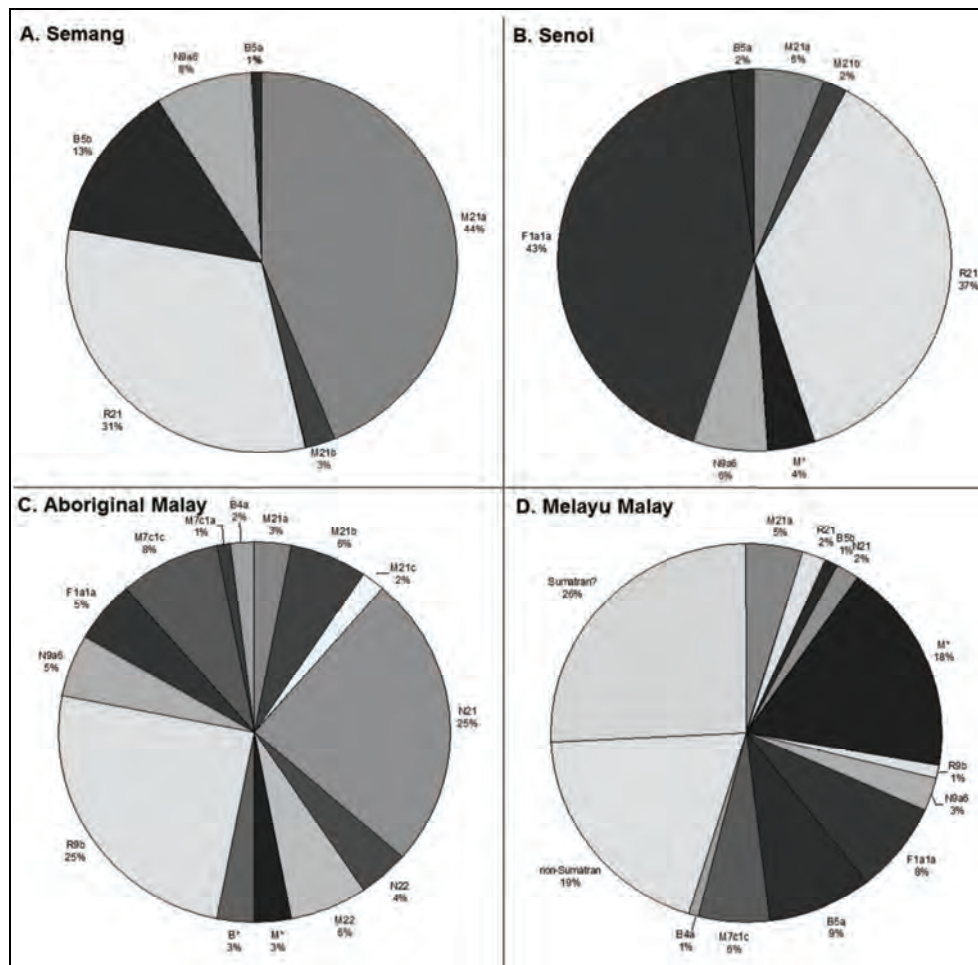
Although less genetically diverse than groups in Indo-China as a result of population bottlenecks and drift among all aboriginal groups, the Orang Asli have older diversity and several unique 'backdrop' genetic features that highlight a clearer window on southward expansions from South China and Indo-China during the Holocene (cf. the chapters in Part 3 of this book). Examination of mtDNA profiles of the three jungle-dwelling peninsular aboriginal groups indicates two unique local Pleistocene founding sources of lineage diversity (Hill et al. 2006).

### **3.1 Identifying indigenous lineages in 'non-Mongoloid' Orang Asli (Semang and Senoi)**

MtDNA profiles reveal distinctive genetic profiles for these groups. One of these profiles of local diversity consists of a combination of isolate haplogroups which are



characteristic of, and represented in, two-thirds (67.1%; 110/164) of ‘non-Mongoloid’ aboriginal groups (Semang and Senoi). This profile consists mainly of derivatives of two isolate indigenous mtDNA lineages, R21 and M21. One of these two lineages, R21, is commonest in the Jahai Semang and the Temiar Senoi. R21 is of great local antiquity in the Malay Peninsula (diverging from the root haplogroup R ancestor 60,000 years ago (Macaulay et al. 2005) and, constituting around a third of lineages in ‘non-Mongoloid’ groups. R21 is apparently unique to the Semang (31.3%; 35/112 samples; Figure 4) and Senoi (36.5%; 19/52 samples; Figure 4). This haplogroup is possibly related to R9, another characteristic Southeast Asian R lineage that shares a basal 16304 mutation (see below).



**Figure 4:** Pie distribution charts to illustrate overall relative percent frequencies of relevant mtDNA haplogroups between the three main divisions of Orang Asli populations sampled (a. Semang, b. Senoi, c. Aboriginal Malay). Shades are consistent between pies and only include relevant haplogroups discussed in text. Also shown, as referred to in the text, a sample of 109 Melayu Malays (4.d.).

The data in Figure 4 can be interpreted as follows. 4.a (Semang) and 4.b (Senoi) share ancestral indigenous Malay Peninsula lineages (M21a&b and R21), but 4.b has a much greater (51%) relative component of intrusive Holocene lineages from the north (N9a6a, B5a and F1a1a). 4.c (Aboriginal Malay) possess circa 50% characteristic indigenous lineages (M21a-c, N21, N22 and M22), 36% of intrusive Holocene lineages from the north (R9b, N9a6a, F1a1a and M7c1a) and 10% of intrusive Holocene lineages from ISEA

(M7c1c and B4a). By inference, the main ancestral Aboriginal Malay population is indigenous to the Peninsula and only 10% derive from ISEA. 4.d (Melayu Malay) shares, at low rates, some indigenous lineages characteristic individually to all three aboriginal groups of the Malay Peninsula (M21a, R21, B5b and N21) and at higher rates, presumed indigenous M\* (18%), 21% Holocene northern intrusive lineages (R9b, N9a6a, F1a1a and B5a), and 33% of intrusive Holocene lineages from ISEA (M7c1c 6%, B4a 1% and 26% 'Sumatran'). Although this ISEA component is higher than in Aboriginal Malays, this figure now places doubt on the assumption that ancestors of Melayu Malays were mainly intrusive during the Late Holocene from ISEA, rather that, while distinct from AM, they may largely be (for example 2/3) indigenous to MSEA and shared in the Holocene dispersals from northern MSEA (data from Hill et al. 2006, which include Melayu Malays from Zainuddin & Goodwin 2003, and Macaulay et al. 2005).<sup>2</sup>

After R21, the other main Southeast Asian isolate lineage characteristic of 'non-Mongoloid' Orang Asli is M21a, a subgroup of M21. Haplogroup M21 is of great local antiquity in the Malay Peninsula being ~57,000 years old (Macaulay et al. 2005), although, given the extreme drift apparent in populations possessing this haplogroup (greatest in the Semang and least in the Aboriginal Malay; Hill et al 2006), the true age may be considerably older. M21 exemplifies half (46.4%; 52/112) of the hunter-gatherer Semang, also being found in 7.7% (4/52) of the swidden hill rice-growing Senoi, and 11.5% (11/96) of the 'Mongoloid' Aboriginal Malay (Hill et al 2006). In the Semang and Senoi, however, M21 sub-groups belong mainly to sub-group M21a, (52/56 M21 haplotypes) while, in the Aboriginal Malays, M21b and M21c haplotypes predominate (8/11 M21 haplotypes). In spite of rare sharing of M21 sub-groups, virtually no M21 haplotypes are shared between Aboriginal Malay and 'non-Mongoloid' groups. M21a is characteristic of 'non-Mongoloid' aboriginal groups, and has spread only minimally into neighbouring Melayu and Aboriginal Malay groups, suggesting its origin amongst 'non-Mongoloid' groups.

As mentioned, M21b and M21c seem more characteristic of Aboriginal Malays. M21b shares a common ancestor with M21a (labelled M21a'b) around 44,000 years ago, and may also be indigenous to the Malay Peninsula, consistent with a deep split between the ancestors of 'Mongoloid' and 'non-Mongoloid' groups in the Peninsula. M21c, a sister clade to M21a'b, is even rarer than M21b, having been sampled in only two Semelai (Aboriginal Malay; Hill et al 2006). However, it is a further indicator of possible long-standing (perhaps even pre-glacial) relationships between the apparently distinct aboriginal groups.

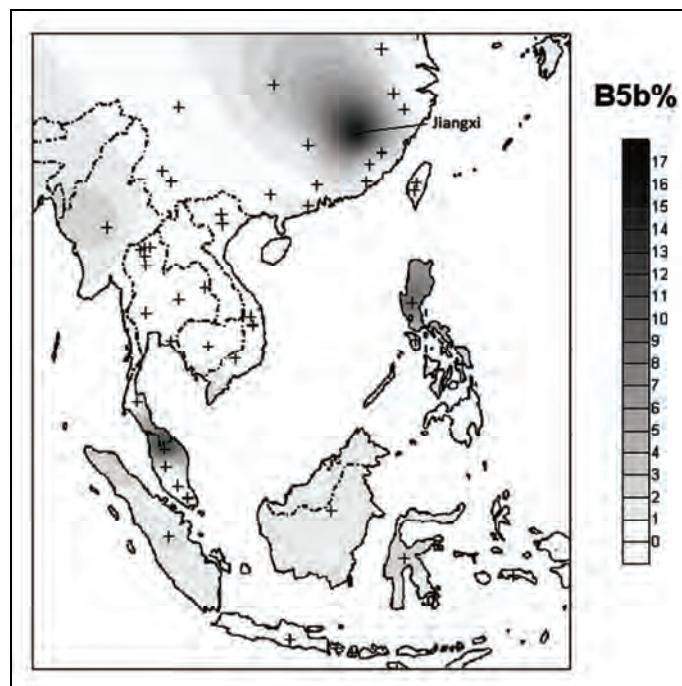
The Batek and Mendriq Semang additionally appear to share a significant frequency of the B5b haplogroup (44.8% and 6.3%, respectively; see also Figures 3 and 4). B5b has great antiquity in East Asia (35,300 years +/- 11,400; Hill et al. 2007) This lineage may also have been present locally among the Semang for some time, as the relevant haplotypes all share a unique HVS1 mutation (at 16294) from the root haplotype. Further, the founding type is outnumbered in both relevant Semang groups, overall 14:1, by its derivative haplotype (mutated at 16354) (Fig 74; Hill 2005). B5b has also previously been identified at high frequency in the Semai Senoi (Melton et al. 1995). Only one Semai was surveyed in the Hill et al. (2006) sample. Only one B (unclassified) was noted in 32 Orang

---

2 In 4.d, other haplogroups found in Melayu Malay, not relevant to Orang Asli, are merged into two arbitrary groups in the d. pie: Sumatran: i.e. haplogroups also found in Sumatra in the Hill et al. 2006 study and thus possibly derived from Sumatra as a Malay traditional homeland; and non-Sumatran: i.e. haplogroups not also found in Sumatra and, in most cases deriving from the Asian mainland.

Asli in the RFLP study of Ballinger et al. (1992), which included five Semai, although which Orang Asli group the haplotype belonged to is not clear.

Hill et al. (2006) argue that B5b is intrusive in these two related groups, most likely deriving from Sumatra in ISEA, where the B5b is present in lower frequency (4/42) including the root type in Medan in the north. B5b is present at variable frequency in both China and ISEA, but excluding Indo-China and Taiwan, including high frequencies in the northern Philippines and in Jianxi province (see Figure 5). There is, however, always the possibility that B5b originated locally in the Malay Peninsula, rather than elsewhere, and lost diversity in the Orang Asli through drift. Singleton putative root B5\* types, which might point towards a homeland are unhelpful having only been found in the Nicobars, Java and the island Malayo-Polynesian-speaking Yami north of the Philippines (Hill 2005). B5b is absent from most of Indo-China, however, ruling out that region as a potential source. Thus, for the purpose of this review of Holocene MSEA, B5b may be regarded as indigenous to Orang Asli.

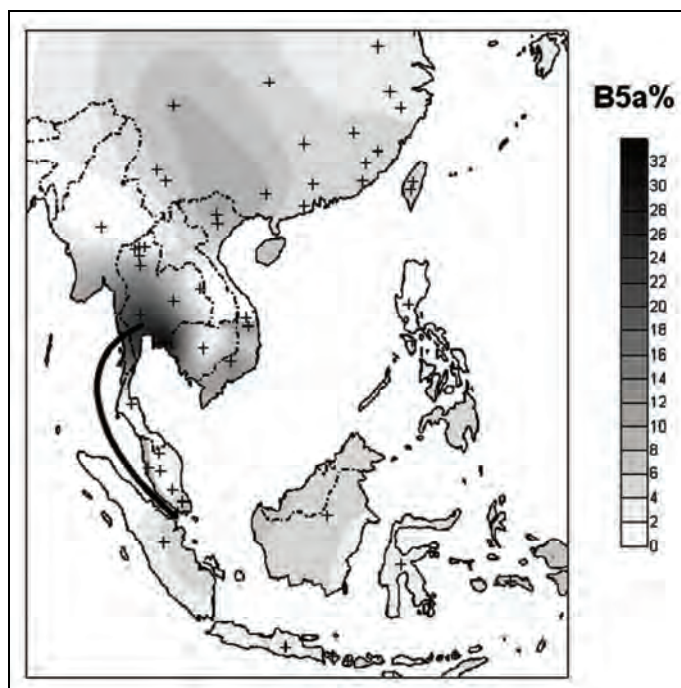


**Figure 5:** Contour map of frequency distribution of mtDNA haplogroup B5b in the Far East. It is difficult to infer any Holocene movement of B5b into or out of MSEA phylogeographically.

The other extant B5 branch, B5a, could also possibly have originated in Pleistocene populations of Mainland Southeast Asia, for that matter, since it is present both among Semang and Senoi. However, in contrast to B5b, it is very common and diverse in China, Taiwan and SE Asia, especially south-western Thailand and southern Vietnam, where it reaches frequencies of up to 16% (Figure 6, and see below) Therefore, in the comparisons of the Orang Asli below, B5a is treated as intrusive from Indo-China to the north, while B5b is treated as indigenous to the Semang/Senoi.

As previously stated, both Semang and Senoi groups share significant proportions of characteristic M21 and R21 lineages, but the degree of preservation of these two signatures differs. The Semang retain these genetic characters in 77.7% (87/112) of those surveyed, while the Senoi retain them in only 44.2% (23/52). The rest of the lineages in both groups

appear to be intrusive from elsewhere in Southeast Asia (but see previous paragraph), mainly from further north in Indo-China (see Figure 4 and below).



**Figure 6:** Contour map of frequency distribution of mtDNA haplogroup B5a in the Far East showing probable coastal spread southwards down western Sundaland from Central Thailand in the Early Holocene.

### 3.2 Identifying indigenous lineages in ‘Mongoloid’ Orang Asli: Aboriginal Malays (AM)

The other local profile of ancient lineages is characteristic of the root-crop-growing Aboriginal Malays, belying their morphological association with other ‘Mongoloid’ so-called ‘Malayic races’ in Southeast Asia. This profile consists of a combination of isolate haplogroups called N21, N22 and M22, all of which are very ancient, specific to the Sunda region and almost unique to Aboriginal Malay (AM) groups. These comprise roughly a third (35.4%; 34/96) of AM mtDNA lineages. This ‘local-indigenous canvas’ percentage rises to over a half (53.1%; 51/96) if the 3 M\*, 3 B\* and the 11 M21 haplotypes mentioned above are included in the tally (Hill et al 2006).

The putative phylogeographic origins of this Aboriginal Malay ‘local profile’ differ slightly in nature from that described above for the Semang and Senoi. While, for the latter, the two unique lineages M21 and R21 have a sterling claim to originate locally in the Malay Peninsula among the ancestors of those groups, the same confidence in origin cannot be applied equally for the rare Southeast Asian lineages N21, N22 and M22 although they each have high representation among the Aboriginal Malays.

Overall, the age of N21 is estimated at 43,000 years, but could possibly be up to 63,000 years old, since it lacks a defining N mutation (Hill et al 2006). Its age, rarity and scattered relict distribution off the southeast Sunda shelf make it a probable vanguard Pleistocene founder lineage in the region. While N21 is present in 31.1% (19/61) and 15.2% (5/33), respectively, of the Semelai AM and Temuan AM, in the Malay Peninsula, it is only represented as two derivative types there. N21 is also found elsewhere in the south of

present-day ISEA. Although appearing at very low rates there, its diversity is greater, including the only examples of the N21 root type (in Bali and Sulawesi: Hill 2005; Hill et al. 2006, 2007). The Aboriginal Malay presence of N21 is thus argued by Hill et al. (2005) to be the result of recent back-migration from ISEA.

N22 is common among Temuan, but found nowhere else in the Asian mainland and is even less common than N21 elsewhere in ISEA. In addition to its presence as two haplotypes in 12.1% (4/33) of Temuan, it is also represented as four individuals and three derived haplotypes in 8.0% of Sumbanese (southeast of the Sunda shelf). Its limited distribution, phylogeography and direct origin from N place its origin near the Sunda shelf, again most likely with the earliest settlers. However, whether this settlement was first in Sumba or the Malay Peninsular region (slightly more likely since the root N22 type is only represented in Aboriginal Malays) is open to question.

As mentioned above, Hill and colleagues (2006; see also Hill 2005) argue that the phylogeography, relict diversity and distribution of N21 and, by implication, N22, could indicate their ISEA origin with migration to the Aboriginal Malay groups in the Malay Peninsula probably in the mid-Late Holocene and just possibly associated with Austronesian speakers. However, this hypothesis does not really fit the evidence of their unique, combined and substantial presence in several Aboriginal Malay groups, especially the Temuan, but not in any other Peninsular groups. It also cannot account for their relict presence in Bali and non-Sunda Sulawesi, Sumba and Alor (i.e. Wallacea: see Figure 1), yet near absence over the entire intervening proximal island region occupying the Austronesian-speaking rump Sunda shelf (that is, the Greater Sundas) (Hill et al. 2007).

In addition, there is the lack of expected parallel similarity in Southeast Asian distribution between N21 and N22 and several other common lineages both characteristically and widely associated with Austronesian-speaking ISEA populations and also known to have expanded in ISEA during the Holocene (Hill et al. 2007; Soares et al 2008). The latter are, namely, (a) haplogroups E1a, E1b and E2, which are absent from all Orang Asli, but not from Melayu Malays (Hill et al 2007; Soares et al 2008), (b) B4a, which is present in Melayu Malays and only two Semelai, but *not* in Temuan or any other Orang Asli (Hill et al 2006 & 2007); (c) M7c1c, which is present only in Semelai AM and Melayu Malays; and the absence of N21 and N22 from the Philippines and Taiwan.

There is a more parsimonious explanation for the unequal and disjunct partition of N21 and N22 in both Wallacea and in Aboriginal Malays (at high rates and low diversity in the latter and at low rates and higher diversity in the former). It is as follows: These are rare ancient Pleistocene founding lineages, as Hill and colleagues surmise, but they were formerly present throughout Sundaland. Furthermore, they have been preserved, by isolation, in both regions and lost by subsequent drift and replacement over the rest of the Sunda shelf. Far away from the Malay Peninsula, Bali, Sulawesi, Sumba and Alor surrounding the Sunda foot, as permanent islands, were relatively protected from replacement by glacial/post-glacial gene flow down the Sundaland coast and, thus, also preserved greater lineage diversity. These two lineages could also have been indigenous and characteristic to the ancestral populations of the Aboriginal Malays isolated in the jungle farther north in Pleistocene Sundaland, but lost their lineage diversity by recent drift. Hill and colleagues (2006) have already provided an explanation for the lack of other indigenous lineage diversity in all Orang Asli groups, resulting not necessarily from immigrant founder effects, but from the severe population bottlenecks and drift they appear to have suffered more recently in their shrinking jungle homeland.

A similar argument can be used for the presence of M22 in 18.2% of Temuan AM, but again at low diversity. M22 diverged directly from the most recent common ancestor

(MRCA) of haplogroup M around 63,000 years ago, with the root type of M22 dating to 31,700 years (SE 20,600 years) (Hill et al 2006; Hill 2005). Apart from the Temuan, M22 is only found in two minority Thais. Hill et al. (2006) argue that M22 in AM probably represents intrusion from the north, that is, from present day Thailand. However, again, a similar argument to that for N21 and N22 can be made, namely, that M22 is a rare Pleistocene founding lineage of the Sunda shelf, whose high rate in Semelai (AM) could more likely represent a persistent, drifted indigenous isolate survival in the jungle.

On these arguments, the Aboriginal Malay populations surveyed thus far may reasonably be regarded as largely aboriginal, and in some way representative of at least one of the Sundaland founding populations, possessing overall 53% Pleistocene lineages indigenous to the Sunda shelf.

In addition and by contrast with Temuan, the Semelai (AM) appear to possess 16.4% intrusive lineages from ISEA, to the south, during the Holocene. As implied above, the significant presence of both M7c1c (13.1%) and B4a (3.3%) as single haplotypes each among the Semelai (AM) and in no other Orang Asli groups (Hill et al. 2006) may clearly be regarded as recent focal Holocene intrusions from ISEA. The same applies to a single instance of M7c1a in the Semelai. Thus, in terms of the main focus of this review, none of these latter lineages are likely to have derived recently from South China or Indo-China.

On the basis of this mtDNA phylogeographic evidence, ancestors to all three broad groups of Orang Asli (Semang, Senoi and Aboriginal Malay) may thus be regarded as indigenous to the Malay Peninsula since well before the Last Glacial Maximum. This indigenous label includes significantly the more southerly-sited, 'Mongoloid', Aboriginal Malay groups, in spite of previous assumptions of their immigrant status as 'Proto-Malays' from ISEA (Harrower 1933; Carey 1976). Furthermore, the ancestors of the 'non-Mongoloid' Semang and Senoi on the one hand, and the 'Mongoloid' AM on the other, may be regarded as originally having two discrete and very different genetic profiles, the former characterised by local lineages M21a, M21b and R21 (and probably B5b), the latter by N21, N22, M22 and M21c.

These two discrete and unique 'indigenous' mtDNA profiles, resulting most likely from geography, isolation and drift, are so characteristic of the relevant aboriginal groups ('Mongoloid' and 'non-Mongoloid'), that they all but constitute that elusive feature, so-called genetic 'ethnic markers'. This means that intrusive lineages from elsewhere (that is, in this review from South China and Indo-China, but also more generally from ISEA or even Taiwan) can be clearly identified against a distinctive indigenous genetic 'backdrop'.

## 4 Identifying intrusion from Indo-China/South China

### 4.1 Semang and Senoi

The Semang clearly have considerably less genetic intrusion than the Senoi, although of a similar nature. Presumably, as a result, they underwent less change of lifestyle and physical morphology. However, two mtDNA lineages that are characteristically intrusive from Indo-China are found in two Semang groups so far surveyed: B5a in Batek Semang (3.4%; 1/29) and N9a6 in Jahai Semang (17.6%; 9/51). By contrast, the Mendriq Semang have none.

As mentioned above, the Senoi have much higher rates of genetic intrusion than the Semang. Indeed, they are the clearest example of admixture, deriving over half of (27/52) their lineages from Indo-China, and sharing both B5a (1.9%; 1/52) and N9a6 (5.8%; 3/52) with Semang, but additionally having a 44.2% intrusion of F1a1a (23/52) (Hill et al. 2006).

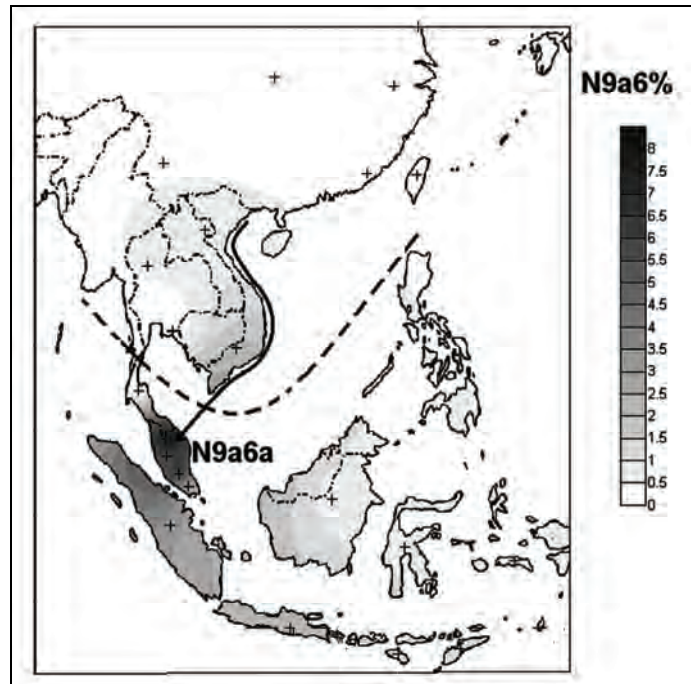
The remaining half (48.1%) of the Senoi population shares characteristic indigenous lineages with the nomadic Semang (4/52 M21; 19/52 R21 and 2/52 M\*, see Figure 4 and above).

The dates of two of these intrusions (N9a6 in Figure 7 and F1a1a in Figure 8) are reasonably straightforward to deduce from the phylogeography. The Southeast Asian lineage N9a6, aged overall 24,300 years (Hill et al 2006), is found at low rates in southeast China and is widespread although uncommon throughout Southeast Asia. In Indo-China, it is commonest in Vietnam. South of Isthmus of Kra (that is, in the Malay Peninsula and Western Indonesia), it is found largely in the form of the derived sub-clade N9a6a, whose root type is only found in Thailand and in Senoi, and has a founding age estimated at ~5,500 (+/- 2600) years (Hill et al. 2006). This sub-clade is commonest in Orang Asli where it is present *only* in the derived form in all three main groups, but is also the main N9a representative in Malays (2.8%) and throughout Western Indonesia (Fig 59, Hill 2005; see Figure 7). This differential distribution makes it possible to apply the same founding date for this lineage to all Orang Asli groups and also, interestingly, for ancestors in Malay and Sumatran populations.

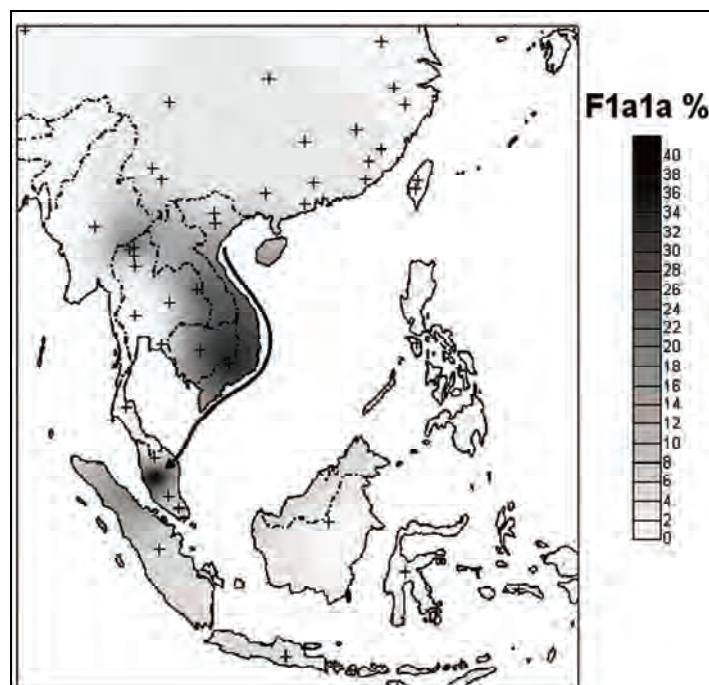
The same principle of dating a shared unique intrusion applies to F1a1a, a sub-clade of F1a, which is common and widespread throughout western Southeast Asia (Hill et al. 2006; see also Figure 8 here). Commoner in the Senoi than in the Aboriginal Malays (below), F1a1a is the only F sub-clade found in Orang Asli (Hill et al. 2006; Fig 83, Hill, 2005). The root type of F1a1a is shared with individuals from Indonesia, Taiwan, and China. However, F1a1a is most frequent in Senoi, Malays, north Thailand and eastern Indo-China, particularly southern Vietnam, Laos and Cambodia (Figure 8).

The MRCA of F1a1a has been estimated to be ~10,700 (+/- 4500) years old from complete sequences (Macaulay et al. 2005), whereas, based on control-region sequence data, it has been dated to ~7700 (+/- 3000) years in Indochina (Hill et al. 2006). This data suggests an arrival of new people in the Malay Peninsula from a northern source (most likely in Indochina) and intermarriage with the ancestors of the Semang, within that time (Hill et al. 2006), that is, in the early Holocene.

Although 17 of the Senoi types belong to the F1a1a root, all of the AM types are derived. Derived F1a1a types are found in five AM (all Semelai), four of these on an F1a1a-16189 sub-clade aged ~ 5500 years (calibrated according to Soares et al. 2009) with some Nicobarese, who also possess that sub-clade at high frequency. Root types for this sub-clade are present in southern Chinese, Aboriginal Malays and the Nicobars, with derived types in the Nicobars and Borneo (Hill et al. 2006). This probably reflects a shared common ancestry in Indo-China, but it may be no coincidence that Senoi and Nicobarese both speak Austro-Asiatic languages (see Diffloth, Sidwell & Blench this volume). A separate F1a1a sub-clade, derived at 16304, with a similar age of ~5500 years (calibrated according to Soares et al 2009), includes seven Senoi.



**Figure 7:** Contour map of frequency distribution of mtDNA haplogroup N9a6 in the Far East. The dashed line indicates the northern limit of distribution of the subgroup N9a6a which exhibits a Mid-Holocene founding event in the Malay Peninsula and nearby Island Southeast Asia. The conspicuous absence of N9a6 from most of Thailand suggests the spread may have been directly across the Gulf of Bangkok.



**Figure 8:** Contour map of frequency distribution of mtDNA haplogroup F1a1a in the Far East and its Mid-Holocene spread down eastern Sundaland. The conspicuous absence of F1a1a from southern Thailand suggests the spread may have been directly across the Gulf of Bangkok.



As mentioned above, the origin and date of arrival of B5a in Orang Asli is more difficult to deduce. It has a widespread and relatively common distribution in China and Southeast Asia, but is most common and diverse in Austro-Asiatic speakers in southern Indo-China (Figure 6; Fig 73, Hill 2005; Fucharoen et al 2001). The overall age of B5a is 17,300 years (+/- 3900 Hill 2005, Table 7), while in ISEA it has been dated at 9200 years (+/- 3000; Hill et al 2007). B5a is present as single root types in the Orang Asli, one each, only in the Batek Semang and Temiar Senoi, making sub-clade dating impossible. There is, however, one sub-clade found mainly in Austro-Asiatic speakers from Indo-China, which is also present in Sumatra, and in two Malays. This branch dates to 8300 years (SE +/- 3400) (calibrated according to Soares et al 2009; data from Hill 2005, Fig 73).

The overall high level of Holocene intrusion into the Senoi (over half (27/52) of their lineages from Indo-China) has been remarked on by Fix (this volume) as being much higher than could be expected from Neolithic 'demic diffusion' on genetic figures derived from elsewhere (for example Richards et al. 2000) and estimates of differential population growth. Fix postulates a possible selective advantage of high rates of Haemoglobin E in immigrants during the first few generations after migration, which could be advantageous in the admixed population following a sedentary lifestyle in a malaria-endemic situation.

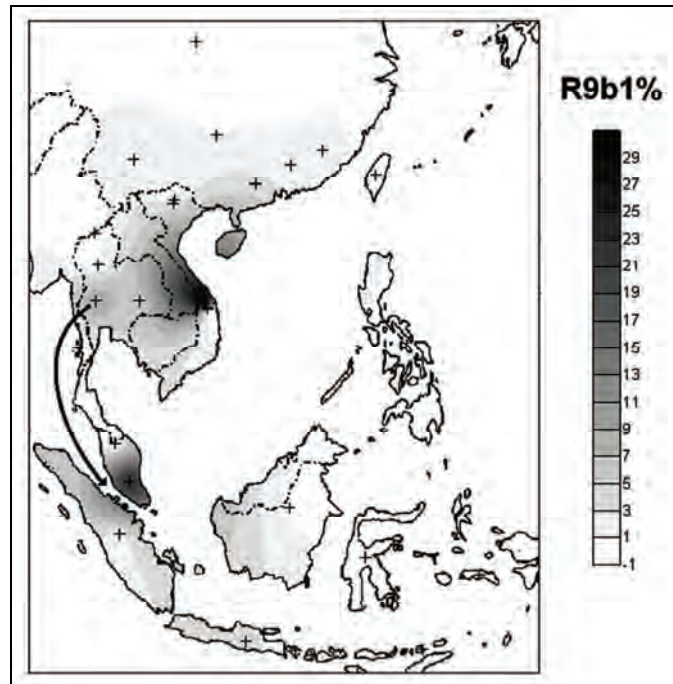
#### **4.2 Identifying intrusion from Indo-China/South China (continued): Aboriginal Malays**

Aboriginal Malays, although differing profoundly in indigenous genetic backdrop from the Semang and Senoi share, as mentioned, two intrusive lineages with the former (namely F1a1a and N9a6) which presumably came from Indo-China at the same time. These northern intrusions are each present at modest rates in at least two discrete AM groupings. N9a6 is present at similar overall frequency (5.2%; 5/96) to the Semang and Senoi; however F1a1a is present at rather lower overall frequency (5.2%; 5/96) than in the Senoi.

Apart from the clear but modest recent intrusions into Semelai AM from ISEA of lineages B4a, M7c1c and M7c1a (discussed above), the main intrusive lineage of note in AM is R9b. It is present in a quarter of Aboriginal Malays (25%; 24/96) and completely absent from Semang and Senoi. R9b is evenly distributed, between the Semelai AM (28%; 17/61) and the Temuan AM (21.2%; 7/33).

R9 is one of the oldest and most important lineages in Southeast Asia. It dates to 53,100 years ago (+/- 5800), while R9b dates overall to 50,700 years ago (+/- 20,100; Macaulay et al 2005; Hill, 2005). As mentioned above, R9 further shares a basal 16304 mutation with R21, which is unique and indigenous to the Semang and Senoi groups. A similar link is found with R22, which is only found in northern Sumatra. The phylogeography of R9 in East Asia is complex, with spread into China and throughout Southeast Asia at a very early stage.

The two main branches of haplogroup R9 are F and R9b. Haplogroup F is very common and widespread in China and Southeast Asia (including F1a1a in the Orang Asli—see above) and the R9 less common and more restricted within Southeast Asia, thus potentially opening a window onto the time of early settlement. Within the Orang Asli, R9b is found only in the Aboriginal Malays (both Semelai and Temuan). It is rare elsewhere but found at low frequencies in Vietnam, Thailand, Indonesia and in the Yunnan and Guangxi provinces of South China (Hill et al 2006).



**Figure 9:** Contour map of frequency distribution of mtDNA haplogroup R9b1 in the Far East. Arrow indicates spread of the sub-clade, R9b1-16288, which appears to have spread south specifically from Central Thailand down the Peninsula to the Aboriginal Malay groups with parallel branches to Western Indonesia in the very Early Holocene.

Detailed complete sequencing of representative R9 haplotypes in these regions (Hill et al. 2006) suggests an ancient divergence of pre-R9b  $\sim 29,000$  ( $\pm 6600$ ) years ago in Indochina and later divergence of R9b  $\sim 19,000$  ( $\pm 5400$ ) years ago in Vietnam/South China. A further derived sub-clade defined by a mutation at 16288 from which the Thai, Aboriginal Malay, and Indonesian R9b lineages all emerge, dates to  $\sim 9000$  ( $\pm 2700$ ) years ago. Although related by common ancestry to the MSEA and AM types, the Indonesian members fall into a parallel derived sub-clade defined by a further mutation at 16192. This overall pattern suggests that R9b diversified in Indo-China particularly in northern Vietnam and Thailand (Figure 9) and then spread southwards into the Malay Peninsula as the R9b1-16288 sub-cluster at least 9000 years ago, with some lineages subsequently dispersing throughout island Southeast Asia, thus arguably returning south to the site of origin of the R21'R22'R9 clade as a pre-Neolithic, possibly Hoabinhian, intrusion to the ancestral Aboriginal Malay population (but see Bulbeck's detailed discussion of the distribution of the Hoabinhian in Southeast Asia and its possible associations with Orang Asli, this volume), with parallel lineages subsequently dispersing throughout island Southeast Asia (Hill et al 2006) (See Figure 9).

However, B5a and B5b, which are found at variable frequencies among Semang and Senoi, are completely absent from all three AM groups surveyed (see above and Figures 4–6). This latter observation is consistent with the suggestion (above) that B5b is indigenous to the Semang/Senoi. As far as B5a is concerned, if it was intrusive from Indo-China during the Holocene, then, it clearly still did not get into the AM. This may be because they were more southerly placed than the Semang or Senoi, or perhaps simply because they were relatively rice-resistant, already having their own sedentary horticultural

subsistence (Bulbeck, 2003), inherited possibly along with R9b, from northeast Indo-China. However, B5a *does* appear in the Peninsular Melayu Malays thus far sampled, at an appreciable rate of 9.2%, which is higher than anywhere in ISEA (Hill et al. 2007). This could suggest an influence on Malays arriving from Thailand to the north, mainly by-passing the Orang Asli (see discussion). More sampling will be required to further explore this hypothesis.

## **5 Summary of Holocene intrusions to Orang Asli from northern Indo-China**

Semang and Senoi may be regarded as having the same indigenous Pleistocene founding mtDNA lineages (M21a+b and R21), while they differ greatly in degree of Holocene cultural and genetic intrusion (9% versus 51%). The ancestors of the Senoi adopted hill rice farming, while those of the Semang were correspondingly less impacted genetically or by agriculture. The Aboriginal Malays have distinct and indigenous Pleistocene founding mtDNA lineages (N21, N22 and M22) different from Semang/Senoi and tend to horticulture rather than rice. Overall the ancestors of the AM received 48% Holocene intrusion, 36% from the north and 10% from ISEA.

Four mtDNA lineages may, on the basis of phylogeographic evidence, be regarded as substantially intrusive from northern Indo-China, in widely varying degrees, to these three Orang Asli groups during or around the beginning of the Holocene. These lineages (N9a6a, F1a1a, B5a, and R9b1-16288), or their immediate ancestors, are common and prominent in Indo-China, with the exception of the derived branch N9a6a, which is absent while its immediate ancestor N9a6 is less common, but widespread in eastern Indo-China.

N9a6 and F1a1a both appear to have taken the same eastern route in the Mid-Holocene from the tip of Vietnam directly south to Malaya, when they were closer to each other, while B5a, and R9b1-16288 by contrast appear to have both moved down the western Sunda coast in the early Holocene.

In all instances, each lineage is intrusive to at least two or more different aboriginal groups. In two lineages (N9a6a and F1a1a), intrusion has occurred in 3–4 groups (including both ‘Mongoloid’ and ‘non-Mongoloid’ groups), while, for 2/4 lineages, intrusion has occurred into two groups belonging to either ‘Mongoloid’ or ‘non-Mongoloid’ types. Given that there is substantial overlap of intrusion across the groups, particularly for the eastern lineages N9a6a and F1a1a, it should be instructive to look at the differences.

F1a1a avoided only the Semang, who were apparently the least culturally impacted, while it is overwhelmingly commonest in the Senoi who are hill rice growers in north-central Malaya. R9b1-16288 is a unique and large intrusion (25%) to the Aboriginal Malays (Semelai and Temuan in the southwest of the Peninsula), being absent from the other (non-Mongoloid) Orang Asli and present only as a trace 1% in Melayu Malays. Given the AM preference for horticulture and the early Holocene date of movement of R9b1-16288, it is not a good candidate for association with rice. R9b1-16288 contrasts with the other Indo-Chinese lineages (N9a6a and F1a1a), which are very much less common (overall 10% in AM). B5a was only found in Batek Semang and Temiar Senoi. This unequal distribution may be less significant in the case of B5a, which is uncommon in Orang Asli anyway, even in Semang and Senoi (1–2%).

Thus, with the exception of the high frequency of R9b1-16288 and the absence of B5a in AM and the absence of R9b1-16288 in the Semang and Senoi, the case can be made

that, with some differences in emphasis, the different Orang Asli population groupings in the jungle regions were to some extent random geographic recipients of Holocene N9a6a and F1a1a gene flow from southeast Indo-China.

The widespread nature of southward migrations from Indo-China is supported by the observation that these four lineages are also collectively found at substantial rates overall (21.2%) among Melayu Malays (B5a: 9.2%, N9a6a: 2.8%, F1a1a: 8.3% and R9b1-16288: 0.9%—Table 2 Hill et al. 2007). Most notable among these is B5a from Thailand, which is much less common among the Orang Asli than the Malays, who also have M7b1 at 3.7% (Table 2 Hill et al. 2007), which was another lineage expanding down the west coast in parallel with B5a from western Indo-China and southwest China at the same time during the Early Holocene, but being absent from Orang Asli. This makes a total of 25% of Malay lineages derived from northern Indo-China.

This large diverse northern element in the Malays is, in itself, a surprising finding. The conventional view is that Malays derive as a whole group from migrations arriving from ISEA, mainly Sumatra, in the Late Holocene, while all these five Indo-Chinese expansions predate that hypothetical event. Other genetic evidence does not support the idea that Malays simply descended from a block recent immigration from ISEA. For instance, one would expect them to be dominated by founder events of characteristic ISEA lineages. Contradicting this view, Melayu Malays have a wide range of non-ISEA lineages, and a lower rate (overall 11%) of lineages known to have expanded in ISEA during the Holocene (E1a, E1b and E2: 5.5%, B4a: 0.9%, M7c1c: 4.6%). They also have a high rate (14.2%) of unclassified local M\* types (Table 7, Hill et al. 2007; Soares et al 2008).

## 6 Prehistoric interpretation and discussion

The two main intrusive Indo-Chinese lineages among the Senoi, F1a1a and N9a6a, both date their expansion to the mid-Holocene (~5500 yr BP). These dates fit an existing hypothesis (Hill et al. 2006) that they are descended from a fusion of local hunter-gatherers with southward-migrating hill-rice farmers from the Ban Kao (and Khok Phanom Di) cultures in Indo-China. The Ban Kao sedentary agricultural assemblage culture first identified by Sørensen in the 1960's (1972), features cordmarked pottery, tripod ware, pedestalled pots, finely polished stone adzes, barkcloth beaters and extended burials and was a landmark discovery at the time. However Bulbeck (this volume) warns that these are not a simple package of elements that necessarily all originated in Ban Kao, nor all at the same time.

The genetic scenario corresponds with evidence-based archaeological hypotheses of southward Ban Kao cultural spread (Sørensen 1972, followed up by Bellwood 1997). The Ban Kao cultures are slightly late, as sources, for the arrival dates estimated for F1a1a and N9a6a in the south. However, this is not a major anachronism since Bulbeck notes (this volume, page 215) several elements predating the Ban Kao culture further south: '... cordmarked pottery was widespread from 5000–6000 years BP onwards, not only in Malaya ... but also across Mainland Southeast Asia'; also tripod ware features in several Peninsular Malaysian sites and in the Isthmus of Kra, 'dated to between 4000 and 5000 years BP', that is, well south of Ban Kao and dating earlier.

There is more of a phylogeographic problem with this unifying hypothesis, namely that F1a1a and N9a6a both derive not from the Ban Kao region, but mainly from southeast and eastern Indo-China, for which no such hypothesis exists (Figures 7 and 8). The absence of a genetic trail of the F1a1a and N9a6a lineages farther down the isthmus, in spite of the

abundant archaeological trail of the Ban Kao cultures, increases the problem and also suggests that F1a1a and N9a6a could have moved directly from southern Vietnam to Malaya, though more sampling is needed in southern Thailand. Geographically, chronologically and archaeologically, an alternative parallel candidate for the cultural spread type for these two lineages would be the Da But cultures emanating from northern Vietnam (Bulbeck, this volume and 2008), and moving south, presumably through southern Vietnam, to the Malay Peninsula and ISEA.

However, geographically speaking, a better candidate lineage for the Ban Kao and Khok Phanom Di locations would be B5a (Figure 6). This candidate is the third northern intrusive lineage among the Semang and Senoi, although it is present as undatable single root types in the Batek Semang and Temiar Senoi and absent in AM. By contrast, the high rate of B5a in coastal rice-growing Melayu Malays (9.5%) and their rich diversity and overall frequency of five different Holocene Indo-Chinese lineages (25%), including M7b1, makes a minor proportion of their ancestors geographically and genetically (although not chronologically) more feasible as main carriers of the Holocene southern spread of rice cultures on the west coast of the peninsula.

The main problem of associating B5a with the spread of Ban Kao cultures is its age. The B5a lineage as a whole dates to the Late Glacial, while the only available peninsular genetic founding date in the Melayu Malays is Early Holocene (8,300 years—calibrated according to Soares et al 2009; see above). Hill et al. (2007) have obtained a similar date of 9200 (+/- 3000) for B5a in ISEA. Of course, there are wide error bars on these estimates, but they do raise the question as to whether the putative rice farmers of western Thailand could have already started moving south before the full flowering of the Ban Kao cultures signalled by ceramic and polished stone tools moving to Malaya. In this context, it has been noted that the Peninsular Malaysian Neolithic burials at Gua Cha provide minimal morphological evidence of any dramatic population incursion at the time of appearance (c. 3000 BP -Cal) of the ceramic/lithic assemblages (Bulbeck 2000, 2005; Bulbeck et al. 2005). Interestingly, one interpretation that Bulbeck makes of this evidence is that if there was a morphological replacement occurring at the Neolithic threshold, it was taking place elsewhere. This idea has some resonance with the concept of B5a intrusions to the ancestors of coastal Melayu Malay populations, rather than to jungle-based Orang Asli.

Speculating further on the B5a date, there is a body of evidence for a single origin of rice domestication in South China around 9000 years ago with rapid spread (Gao & Innan 2008; Liu et 2007; Vaughan et al 2008). In contrast, Londo and colleagues (2006) suggest, on genetic phylogeographic evidence, that *Oryza sativa indica* was domesticated separately. The fascinating aspect of the latter paper is that one of their suggested areas of *Indica* domestication is the western coastal region of Indo-China, and Myanmar, in other words on the eastern coast of the Bay of Bengal. Sagart (this volume) reads Londo et al. as saying ‘south of the Himalayas’ as their suggested area of *Indica* domestication, although what they actually say—‘Thailand, Myanmar, and India’—is more focussed on western MSEA, and their map—Figure 1—includes the whole Malay Peninsula. This observation has resonance both with the distribution of B5a (see Figure 6) and with Gérard Diffloth’s suggestion (personal communication) for that region as an ultimate homeland for Austro-Asiatic languages (but see also alternative interpretations by Sagart, and Sidwell & Blench this volume).

A similar though lesser effect (10–20% Neolithic intrusion) of the same two Mid-Holocene intrusive lineages (N9a6a and F1a1a) to those found in the Senoi is also seen among the Aboriginal Malay populations. However, in this case there were additionally, earlier, rather larger intrusions (~25%) of another lineage R9b1-16288 coming from

northeast Indo-China apparently starting from 9000 years ago. This date, during the Early Holocene period would have been contemporary to the time of the Hoabinhian, Bacsonian horizons but after the Son Vi (for example see Figure 2.3 in Higham 1989). The distribution of this sub-clade (R9b1-16288) (Figure 9) could fit that of Hoabinhian tools and Sumatraliths in Southeast Asia (for example, Figure 33.1 in Bowdler 2006; but see also Bulbeck this volume), including northern Vietnam, Laos, the Malay Peninsula and, to a lesser extent Thailand and Cambodia, but excluding south Vietnam. On the other hand, given the wide error bars on these dates, the movement of R9b1-16288 from northern Vietnam into Aboriginal Malay populations could have been pre-Hoabinhian, or simply have been part of a similar route of east Indo-China coastal gene-flow down the peninsula as F1a1a and N9a6a along with an early Neolithic expansion. Thus, these lineages are geographically complementary, with a slightly different admixture distribution in the target populations.

It is not the purpose of this review to try and answer issues of ultimate origins of morphological traits, for instance why Aboriginal Malay and Melayu Malay can be classified as Southern Mongoloid, while the Semang and the Senoi cannot. Apart from anything else, genetic determinants of these differences are unknown, mtDNA certainly has no known morphological function and the fossil record is still silent as to the geographic and temporal origins of Mongoloid types, particularly whether they originated to the north or south of the Himalayas (Oppenheimer 2003, chapter 5), and when they appeared in MSEA. However, it is worthy of note that there was more than 50% genetic intrusion from eastern Indo-China into the ancestors of the Senoi without the clear appearance of Mongoloid features in their modern populations. This east Sunda lack of morphological impact is consistent with Bulbeck's Neolithic findings in Gua Cha (2000) and contrasts with the association, in west Sunda, of putative intrusions (B5a and R9b1) to the AM and Melayu Malay populations, which are classifiable as Mongoloid today. The implication is that the Neolithic source populations for eastern Sunda intrusions to the ancestors of the Semang and Senoi were not Mongoloid, whereas western Sunda intrusions to ancestors of AM and Melayu Malay populations could have been. These implications have to be taken in the context of Bulbeck's discussion (this volume) of early Neolithic, non-Mongoloid Bac Son and Da But samples from northern Vietnam, and evidence for Neolithic morphological transition to Mongoloid in the late Neolithic Man Bac site, and the early Metal Age Hoa Diem site in northern Vietnam from Matsumura et al. (this volume). A unifying hypothesis might be a linked chain of discrete southward displacements of earlier indigenous non-Mongoloid populations down the east coast of Vietnam towards Malaya during the Neolithic.

It is worth pointing out that the overall Archaeo-genetic reconstruction presented here (Figures 4-9) of Holocene intrusions to the Malay Peninsula is similar in most respects to that of Bulbeck (Figure 5; this volume). Consensus is present in that: 1) M21, R21 and M22 are regarded as Pleistocene indigenous to the Malay Peninsula; 2) B5a is regarded as Holocene intrusive down the western Sunda coast (along with M7b1 not shown in figure); 3) N9a6(a) is seen as Holocene intrusive down the eastern Sunda coast crossing the mouth of the Gulf of Bangkok; 4) F1a1a is seen as Holocene intrusive down eastern Sunda, following the Mekong river according to Bulbeck and taking the same route as N9a6(a) in my reconstruction; 5) M7c1c is seen as ISEA-intrusive during the Holocene. We differ on minor points: a) I see N21 and N22 as originally indigenous to the whole of Pleistocene Sundaland; b) I find no evidence for B5b arriving among the Semang during the Holocene, that is, it could have been there already; c) I argue that R9b1-16288 took the same western

route as B5a from Central Thailand, although ending up among the Aboriginal Malay groups.

Finally, the Malay population of the Peninsula possess the full range of the typical intrusive, Indo-Chinese lineages (described here) at substantial rates (c. ~25%), in particular for B5a (9.5%), including several (for example M7b1) that are not found among the three aboriginal groups. These appear to have expanded in western Indo-China in the early Holocene during the Hoabinhian, Bacsonian and Da But cultures. This pattern could be explained by a coastal population movement by-passing the aboriginal groups of the interior, when the Malay Peninsula was still much broader.

## 7 Conclusions

All three Orang Asli groups appear to descend from geographically discrete original Pleistocene Peninsular populations and to have received substantial multiple (4–5) maternal lineage intrusions from northern Indo-China during the Holocene. Surprisingly, the same northern intrusive scenario can be inferred for the majority of lineage ancestors of the Melayu Malay populations studied here, in spite of their supposed ISEA origin (based on linguistics).

While accepting caveats that genes do not necessarily (or only rarely) move hand-in-hand with culture and language, it is reasonable to suggest that rice agriculture may have been involved in perhaps 2–4 of these north-south lineage expansions (N9a6a, F1a1a (+/- M7b1 & B5a)), and even that the spread of new culture and Austro-Asiatic languages to the Orang Asli and to the Nicobars may be connected to these genetic spreads.

The phylogeographic details of these different spreads raise several speculative questions for archaeologists and geneticists to examine and possibly test. First, are modern-day Aboriginal Malays, who hold the R9b1-16288 lineage, partly descended from Hoabinhian horticulturalists? Second, are modern-day rice-farming Melayu Malays who hold the B5a lineage (and are the main inheritors of that lineage in the Peninsula), partly descended from rice farmers from the region as defined by the Ban Kao cultures or their predecessors? Third, was rice grown in western Indo-China, before the Mid-Holocene as suggested by the early Holocene spread of B5a? Fourth, were there distinctive Eastern Indo-Chinese rice cultures, such as Da But (see Bulbeck this volume), that dispersed directly to Orang Asli using various routes via the east coast of the Malay Peninsula in the Mid-Holocene in parallel with N9a6a and F1a1a?

## Acknowledgements

Many of the results, interpretations and speculations presented in this paper are based on publications of our research group, in particular Hill et al. (2006, 2007). Maru Mormina kindly pointed out more recently published mtDNA data (including from her own thesis) to enable me to make the gene frequency maps more representative in Indo-China. My thanks to David Bulbeck for reading my MS and providing archaeological and morphological advice in relation to interpreting the Orang Asli mtDNA data.

Data for Figures 4–9 were collated from Macaulay et al. 2005, Hill 2005, Hill et al. 2006, 2007, Mormina 2008, Li et al. 2007, Fucharoen et al. 2001, Lertrit et al. 2008, Yao et al. 2002, Trejaut et al. 2005, Black et al. 2006, Maher 2007. On Figures 5–8, crosses (+) indicate sample sites from which contour maps are reconstructed. Arrows are additionally overlaid, not to emulate frequency distribution, but only where phylogeographic analysis

of relevant mtDNA lineage indicates southward Holocene migration mixing into Orang Asli populations.

## References

- Ballinger, Scott W., Theodore G. Schurr, Antonio Torroni, Yik-Yuen Gan, Judith A. Hodge, K Hassan, Kuei Hsiang Chen & Douglas C. Wallace. 1992. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics* 130. 139–152.
- Bellwood, Peter. 1997. *Prehistory of the Indo-Malaysian archipelago*. Hawaii: University of Hawaii Press.
- Black, Michael L., K Duffel, Cheryl A. Wise, Sheena G. Sullivan & Alan H. Bittles. 2006. Genetic ancestries in northwest Cambodia. *Annals of Human Biology* 33. 620–627.
- Bowdler, Sandra. 2006. The Hoabinhian: Early Evidence for SE Asian Trade Networks? In Elisabeth A. Bacus, Ian Glover & Vincent G. Pigott (eds) *Uncovering Southeast Asia's past: Selected papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists*, 355–359. London: The British Museum.
- Bulbeck, David. 2000. Dental morphology at Gua Cha, West Malaysia, and the implications for 'Sundadonty'. *Bulletin of the Indo-Pacific Prehistory Association*. 19. 17–41.
- . 2003. Hunter-gatherer occupation of the Malay Peninsula from the Ice Age to the Iron Age. In Julio Mercader (ed.) *Under the canopy: the archaeology of tropical rain forests*, 119–160. New Brunswick (NJ): Rutgers University Press.
- . 2005. The Gua Cha burials. In Majid Zuraina (ed.) *The Perak Man and other prehistoric skeletons of Malaysia*, 253–309. Penang (Malaysia): Penerbit Universiti Sains Malaysia.
- . 2008. An integrated perspective on the Austronesian Diaspora: The switch from cereal agriculture to maritime foraging in the colonisation of island Southeast Asia. In Jane Balme & Sue O'Connor (eds) *More unconsidered trifles: Papers to celebrate the career of Sandra Bowdler*. *Australian Archaeology* 67. 31–51.
- . this volume, Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya.
- Bulbeck, David, Abdul Kadir Rahimah, Adam Lauer, Zamri Radzi & Daniel Rayner. 2005. Tooth sizes in the Malay Peninsula past and present: insights into the time depth of the indigenous inhabitants' adaptations. *International Journal of Indigenous Research* 1. 41–50.
- Bulbeck, David & Adam Lauer. 2006. Human variation and evolution in Holocene Peninsular Malaysia. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 131–171. Cambridge (MA): Cambridge University Press.
- Burenhult, Niclas, Nicole Kruspe & Michael Dunn, this volume, Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula.
- Carey, Iskandar. 1976. *Orang Asli: the aboriginal tribes of peninsular Malaysia*. Kuala Lumpur (Malaysia): Oxford University Press.
- Diffloth, Gérard, this volume, Austroasiatic word histories: boat, husked rice and taro.
- Fix, Alan, this volume, Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model.



- Forster, Peter, Rosalind Harding, Antonio Torroni & Hans-Jürgen Bandelt. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. *American Journal of Human Genetics* 59. 935–945.
- Fucharoen, Goonnapa, Supan Fucharoen & Satoshi Horai. 2001. Mitochondrial DNA polymorphisms in Thailand. *Journal of Human Genetics* 46. 115–125.
- Harrower, Gordon. 1933. Skeletal remains from the Kuala Selinsing excavations, Perak, Malay Peninsula. *Journal of the Malaysian Branch of the Royal Asiatic Society* 11. 190–210.
- Higham, Charles. 1989. *The Archaeology of Mainland Southeast Asia: From 10,000 BC to the fall of Angkor*. Cambridge: Cambridge University Press.
- Hill, Catherine. 2005. *Mitochondrial DNA Variation in Island Southeast Asia*. PhD Thesis, University of Huddersfield.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, William Meehan, James Blackburn, Douglas Clarke, Joseph Maripa Raja, Patimah Ismail, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2006. Phylogeography and ethnogenesis of Aboriginal Southeast Asians. *Molecular Biological Evolution* 23(12). 2480–2491
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, Douglas Clarke, Petya B. Blumbach, Matthieu Vizuete-Forster, Peter Forster, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2007. A mitochondrial stratigraphy for Island Southeast Asia. *American Journal of Human Genetics*. 80. 29–43.
- Lertrit, Patcharee, Samerchai Poolsuwan, Rachanie Thosarat, Thitima Sanpachudayan, Hathaichanoke Boonyarit, Chatchai Chinpaisal & Bhoom Suktitipat. 2008. Genetic history of Southeast Asian populations as revealed by ancient and modern human mitochondrial DNA analysis. *American Journal of Physical Anthropology* 137. 425–440.
- Li, Hui, Xiaoyun Cai, Elizabeth R. Winograd-Cort, Bo Wen, Xu Cheng, Zhendong Qin, Wenhong Liu, Yangfan Liu, Shangling Pan, Ji Qian, Chia-Chen Tan & Li Jin. 2007. Mitochondrial DNA Diversity and Population Differentiation in Southern East Asia. *American Journal of Physical Anthropology* 134. 481–488.
- Li-zhi Gao & Hideki Innan. 2008. Non-independent Domestication of the Two Rice Subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, Demonstrated by Multilocus Microsatellites. *Genetics* 179. 965–976.
- Jason P. Londo, Yu-Chung Chiang, Kuo-Hsiang Hung, Tzen-Yuh Chiang & Barbara A. Schaal. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Science*. USA 103. 9578.
- Macaulay, Vincent, Catherine Hill, Alessandro Achilli, Chiara Rengo, Douglas Clarke, William Meehan, James Blackburn, Ornella Semino, Rosaria Scozzari, Fulvio Cruciani, Adi Taha, Norazila Kassim Shaari, Joseph Maria Raja, Patimah Ismail, Zafarina Zainuddin, William Goodwin, David Bulbeck, Hans-Jürgen Bandelt, Stephen Oppenheimer, Antonio Torroni & Martin Richards. 2005. Single, Rapid Coastal Settlement of Asia Revealed by Analysis of Complete Mitochondrial Genomes. *Science* 308. 1034–1036.
- Maher, Geoffrey. 2007. *A mitochondrial DNA analysis of ethnic groups in Myanmar*. MRes Dissertation, University of Leeds.
- Matsumura, Hirofumi, Marc F. Oxenham, Nguyen Kim Thuy, Nguyen Lan Cuong & Nguyen Kim Dung, this volume, Population history of Mainland Southeast Asia: The Two Layer model in the context of Northern Vietnam.

- Melton, Terry, Raymond Peterson, Alan J. Redd, Nilmani Saha, Abdul Salam M. Sofro, Jeremy Martinson & Mark Stoneking. 1995. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *American Journal of Human Genetics* 57. 403–414.
- Mishmar, Dan, Eduardo Ruiz-Pesini, Pawel Golik, Vincent Macaulay, Andrew G. Clark, Seyed Hosseini, Martin Brandon, Kirk Easley, Estella Chen, Michael D. Brown, Rem I. Sukernik, Antonel Olckers & Douglas C. Wallace. 2003. Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Science USA* 100. 171–176.
- Mormina, Maru E. 2008. *MtDNA perspectives on human dispersals in Mainland Southeast Asia*. PhD. University of Leeds.
- Oppenheimer, Stephen. 1998. *Eden in the East: the Drowned Continent of Southeast Asia*. London: Weidenfeld and Nicholson.
- . 2003. *Out of Eden: the Peopling of the World*. Constable Robinson (new revised paperback edition 2004).
- Rayner, Daniel & David Bulbeck. 2001. Dental morphology of the ‘Orang Asli’ aborigines of the Malay Peninsula. In Maciej Henneberg (ed.) *Causes and effects of human variation*. Adelaide (Australia): Australasian Society for Human Biology, University of Adelaide.
- Richards, Martin, Vincent Macaulay, Eileen Hickey et al. (37 coauthors). 2000. Tracing European founder lineages in the Near Eastern mitochondrial gene pool. *American Journal of Human Genetics* 67. 1251–1276.
- Sagart, Laurent, this volume, The Austroasiatics: East to West or West to East?
- Saillard, Juliette, Peter Forster, Niels Lynnerup, Hans-Jürgen Bandelt & Søren Nørby. 2000. MtDNA variation among Greenland Eskimos: The edge of the Beringian expansion. *American Journal of Human Genetics* 67. 718–726.
- Sidwell, Paul & Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.
- Soares, Pedro, Jean Alain Trejaut, Jun-Hun Loo, Catherine Hill, Maru Mormina, Chien-Liang Lee, Yao-Ming Chen, Georgi Hudjashov, Peter Forster, Vincent Macaulay, David Bulbeck, Stephen Oppenheimer, Marie Lin, & Martin Richards. 2008. Climate change and post-glacial human dispersals in Southeast Asia. *Molecular Biology and Evolution* 25(6). 1209–1218.
- Soares, Pedro, Luca Ermini, Noel Thomson, Maru Mormina, Teresa Rito, Arne Röhl, Antonio Salas, Stephen Oppenheimer, Vincent Macaulay & Martin Richards. 2009. Correcting for purifying selection: An improved human mitochondrial molecular clock. *American Journal of Human Genetics* 84. 1–20.
- Sørensen, Per. 1972. The Neolithic cultures of Thailand (and north Malaysia) and their Lungshanoid relationships. In N. Barnard (ed.) *Early Chinese Art and its possible influence in the Pacific Basin* 2. 459–506.
- Trejaut, Jean Alain, Toomas Kivisild, Jun Hun Loo, Chien Liang Lee, Chun Lin He, Chia Jung Hsu, Zheng Yuan Li & Marie Lin. 2005. Traces of Archaic Mitochondrial Lineages Persist in Austronesian-Speaking Formosan Populations. *PLoS Biology* 3. 1–11.
- Vaughan, Duncan A., Bao-Rong Lu & Norihiko Tomooka. 2008. The evolving story of rice evolution. *Plant Science* 174(4). 394–408.

Yao, Yong-Gang, Long Nie, Henry Harpending, Yun-Xin Fu, Zhi-Gang Yuan & Ya-Ping Zhang. 2002. Genetic Relationship of Chinese Ethnic Populations Revealed by mtDNA Sequence Diversity. *American Journal of Physical Anthropology* 118. 63–76.

Zainuddin, Zafarina, William Goodwin. 2003. Mitochondrial DNA profiling of modern Malay and Orang Asli populations in peninsular Malaysia. In Christian Doutremepuich & Niels Morling (eds) *Progress in forensic genetics*, 428–430. Amsterdam: Elsevier Science.

# 6 *Ethnology and the issue of human diversity in mainland Southeast Asia*

---

HJORLEIFUR JONSSON

## 1 Introduction

For modern MSEA, much of the public relevance of ethnic identification has revolved around national configurations, but it has roots in colonial-era concerns with the unequal strength of races. Each ‘people’ was presented as a race, whose homeland was either traced or postulated, and often enough such speculations assumed that races were analogous to armies, bouncing up against one another, the strongest race taking the most central location on the terrain (Hallett 1890:21; Cupet 1998 [1893]:148–50). The historical specificity of racialised colonial rule was portrayed as a natural feature of the social landscape of Southeast Asia from the beginning of history. Seidenfaden relates that various people have defined Maeo and Yao linguistically ‘with the Chinese’, as ‘Mon-Khmer’, or as ‘Thai-Kadai-Indonesians [as part of] a common proto-Austroic block’ (1958:129). He remarked that the matter was not solved, the ‘original or “pure” type of any of these groups’ was unknown. He mused that ‘the forefathers of the Maeos and Yaos may have resembled the “Ur”-Thai-Kadai-Indonesians or the “Ur”-Mon-Khmers for all we know!’ He added that ‘the Maeos and Yaos may have formed the vanguard of the Mon-Khmer wave when this, coming from the Brahmaputra valley, broke into Yunnan’ (129).

Looking at such reporting in retrospect, the notion of unequal strength explaining the different position of individual peoples or races appears to express the conceptual framework of the colonial enterprise. Colonials naturalised their presence and their project through references to the long run of regional history in Southeast Asia as being about the uneven power of different races. This act of colonial memory-making was an act of willed forgetting in that it wiped out all traces of the social and historical ruptures involved in the racial structurings of colonial encounters and routines. Any scientific account of human diversity in Southeast Asia that takes for granted contemporary ethnolinguistic classification serves, deliberately or not, to reinforce particular state regimes (colonial, national, etc.) of truth (see Keyes 2002).

What distinctions are drawn, and where, aligns science and hermeneutics in naturalising certain identities and charting particular forms against the flow of time (Gould 1981, 1991). In his examination of European archaeological support for the transition from Neanderthals to Homo Sapiens Sapiens, modern humans, Clark (2002) finds no evidence for replacement and plenty for continuity; ‘it [looks like] the Middle-Upper Paleolithic transition was a monumental “nonevent”, both biologically and culturally’ (2002:64).

---

N. J. Enfield, editor.

*Dynamics of human diversity, 109-122.* Pacific Linguistics, 2011

© This edition vested with Pacific Linguistics.

Instead, it appears that Neanderthal archaeology is very much a European preoccupation that took shape in the context of nationalism and nation-building. Assumptions about the separate subsistence practices, technological and artistic traditions of Chatelperronian, Szeletian, Mousterian, or Aurignacian peoples rest on expectations of discontinuity, ethnicity, and progress. To some American archaeologists, the distinguishing features appear both arbitrary and of exaggerated importance (Clark 2002:55, 63), but Neanderthals are ‘highly mythologised central players in the origin narrative of Europeans’ (Marks 2007:11).

This example suggests the need for caution regarding how we signify the archaeological landscape in terms of peoples. Scientific and vernacular notions of ethnic groups as bounded, biological units reproduce the ‘tautologies of race, ethnicity, and culture [that characterise contemporary] nation-hood’ (O’Connor 2000:441). The critique of such tautologies is common, but it is more important to explore what insights, if any, cultural anthropology can offer toward an alternative. Edmund Leach (1954:36) noted that part of the British control of Further India (Burma) involved separating the Kachin from the Shan, on the grounds that the former were animist highland shifting cultivators and the latter Buddhist members of lowland states. In Indochina, the French spent considerable energy classifying the Empire’s peoples by level of civilisation and in relation to ethnolinguistic categories (Abadie 1924; Bonifacy 1919), to the point of inventing and peopling the category Muong as a primitive form of the Vietnamese (Taylor 2001). In Siam (Thailand), the elite went about classifying the terrain’s peoples as civilised (urbanites), civilisable (rurals), and savage/uncivilisable (forest people) (Winichakul 2000a, b) as it charted an emerging national domain.

In each case, colonial or not, this was about forging national domains in terms of ethnicity, where ethnic identity implied a claim to recognition and rights within the state in relation to heritage, level of civilisation, and the like. This dynamic centred on fashioning subjects in relation to the inequalities of the nation state. As this process played out in Siam, only Thai peoples (including various Tai who were defined as more archaic and less civilised than the Siamese—cultural features were a way to tell time and space) had claims to land and rights, anyone else was by definition an alien and did not qualify as a citizen (Jonsson 2005:45–55).

Much twentieth century ethnography took this ethnicised landscape for granted and elaborated it through studies of individual groups of foragers, highland shifting cultivators (tribes, swiddeners), lowland peasants, and so on (King & Wilder 2003). Leach (1954) challenged the notion that differences in ethnicity, language, and livelihood in northern Burma were the result of separate migrations by distinct peoples. He proposed instead that the separate patterns manifest by Kachin and Shan peoples had developed over time, that peoples had taken shape through deliberate acts of differentiation and that these differences were mapped on the separate landscapes of irrigated rice farming on the plains and shifting cultivation in the highlands. His study showed that there was some movement of people between different categories of identity, and that highland peoples would shift between uniformity/compatibility and diversity/incompatibility in kinship, social relations, and political organisation.

There is nothing inert about whatever patterns of identity and difference we may deduce within the region. Ethnographic and ethnological work offers no support for the linear and progressive narrative that is often imposed on the archaeological record in ethnic and technological (stone, bronze, iron) terms. Any realignment of cultural anthropology with linguistics, archaeology, and biological anthropology on the topic of human diversity in

MSEA rests on finding alternatives to the evolutionary narratives that take ethnic groups and groupings (such as ethnolinguistic families) for granted as the units of analysis.

## 2 From peoples to paradigms

In his study of the ethnic shift in lowland MSEA from Pyu, Mon, Khmer, and Cham to the dominant position of Burmese, Tai, and Vietnamese, between roughly 700 CE and 1700 CE, O'Connor (1995) states that 'there is no direct evidence that an actual influx of immigrants ever displaced earlier peoples ... Awaiting evidence that may never come is unrealistic and leaving racial waves unchallenged is irresponsible' (1995:987). O'Connor argues instead that one agricultural paradigm replaced another, that there was a shift from flood management by garden-farmers to irrigated wet-rice farming (1995:969–74), noting that 'mainlanders link religion to agriculture, agriculture to ritual, and ritual to ethnic identity ... What diffuses is a society-shaping complex, not just agricultural techniques' (1995:986). Language is one aspect of these alternative models, and this suggests a regional and historical approach to the distribution of certain languages as a marker of identity as much as that of difference.

O'Connor's case parallels that made by Leach, in the emphasis on identity and difference as actively created. Both scholars note the region's pervasive multilingualism and show that language is wielded to express social alignments. Between the two cases, one may also come to a historicisation of upland-lowland contrasts. In a separate study, O'Connor (2000) contrasts Khmer and Javanese heartland temple-states to subsequent Tai and Malay hinterland city-states (*muang* and *negeri*). Temple-states assumed a singular centre and stressed the 'allocation and protection of private rights', whereas city-states assumed a fundamental divide between royalty and commoners and rested on the cooperation of a king and local chiefs. 'We need only glance at Khmer and Javanese primary states to see polities that fought over the very distinctions that Tai and Malay polities accepted from the start.' Temple-states elaborated 'localising fertility cults [that] restricted their farmers' mobility [and] their own expansion. Controlling new land required massive investment in temples' (2000:437). In contrast, Tai and Malay city-states were mobile and easily replicated models. 'Distinguishing the state from religion and agriculture, each domain could follow its own principles. A religious crisis or a failed crop need not have threatened the state's integrity' (2000: 438, citing Kirsch 1985).

Archaeologist Bellwood states that the lack of evidence for upland 'agricultural settlement before about 500 BC [could suggest that] the "hill-tribe" phenomenon of mainland Southeast Asia—the expansion into high altitudes from the north of shifting agriculturists—is a relatively recent development mainly limited to the historical period' (1992:120). I assume that his reference to the 'historical period' implies the latter half of the first millennium CE and since, the emergence of inscriptions and other remains involving some extant written remains. Hinterland peoples, identified in relation to forests and mountains, emerge in the historical record as the victims of slave-raids and trade. According to Chinese envoy Chou Ta-Kuan's description of the Angkorean domain at the end of the thirteenth century: 'Wild men from the hills can be bought to serve as slaves. Families of wealth may own more than one hundred; those of lesser means content themselves with ten or twenty; only the very poor have none. These savages are captured in the wild mountainous regions, and are of a wholly separate race called Chuang [brigands]' (Chou 1967:21). He adds that those of the mountain peoples 'who understand the language of the country are sold in the towns as slaves', in contrast with others who

'refuse to submit to civilization and are not familiar with its language. They have no houses but wander about in the mountains' (1967:25).

Both the heartland temple-states and the hinterland city-states centred on control over lowland agriculture and assumed an affinity among their peoples. This is the background to highland adaptations of people not under the state's protection network, and the ethnic and linguistic diversity and complexity of the highlands was in direct contrast to the homogenising tendencies of ritual, language, and identity that states generated and reproduced. Shifting cultivation in the forested highlands became a society-defining framework in relation to and in opposition to state-societies, and states reinforced this by not levying tribute on people making swidden fields (Condominas 1990:60; Durrenberger & Tannenbaum 1990:4–5). As a regional paradigm, shifting cultivation went along with social dynamics that united people through ritual, feasting, and kinship that emphasised people's identity in relation to household, kin-group, and village—units that were in competition with one another (Kirsch 1973).

The adaptation of highland populations was not to the natural environment as such, but to an environment that had been prefigured by the politics of identities and social relations in terms of cultural and political economic dimensions of the state. The state's rhetoric of civility informed ritual practice and social relations, and it was in these cultural terms that ordinary lowland farming populations knew and feared the forest as the abode of evil spirits, dangerous animals, and general lawlessness. This alignment of natural and social landscapes, where forests equalled disorder, was common throughout (and beyond) pre-modern MSEA (Jonsson 2005).

Proposing a regional and historical model for a highland village paradigm anchored to shifting cultivation does not explain any particular upland group either in general or in detail. Rather, it situates the innumerable highland formations of culture, language, and society within the region in a way that emphasises interconnections among divergent but interrelated forms. The rise of city-states actively responded to earlier temple-states and simultaneously created boundaries at the foothills of the forested mountains, beyond which were peoples that were, from the state's perspective, unworthy of even tribute duties. Hinterland populations created and sustained alternative political and cultural formations that were anchored to agriculture as much as feasting. In contrast to the absolutism of this divide, there were various relations of trade, tribute, warfare, slavery. This was particularly pronounced in lowland areas adjacent to the hills, and more 'central' state authorities repeatedly tried to foreclose such intimacies by reasserting boundary mechanisms. In the historical era, many upland settlements were linked to lowland domains through titles and social relations. These were particular relationships, and many highlander identities were place-specific to the histories of certain sets of interrelated kingdoms.

Along with such relationships, upland cultures were partly articulated in terms of warfare, a pattern which disappeared by the early twentieth century. The common image of highland peoples as individual, identifiable groups that could be accounted for in terms of ethnically specific and inward-looking patterns of religion, agriculture, kinship, and the like is in many ways a product of the twentieth century. Karl Gustav Izikowitz (1951), who did field research among Lamet in Laos during 1936–37, suggests that Khmu fit the model of Kha groups which paid tribute to lowland courts, while Lamet had been more independent (1951:28). But the two are not so different. Lamet had earlier paid tax to Yuan rulers in Chiang Khong (1951:346), but as that domain fell or faded away they traded rice to Luang Prabang for salt and metal (1951:311), while Khmu contributed forest products.

In this light, it is unlikely that Lamet as such were any more intrinsically independent than Khmu. Rather, they were able to articulate a cultural autonomy after Chiang Khong fell because they had other trade relations to the lowlands that did not imply or require subservient rank. By the 1930s, French colonial road networks made rice from lowland Vietnam available in northern Laos, but prior to that upland rice had in many places been a valuable commodity in lowland areas, particularly in inland kingdoms with limited valley land that could sustain surplus rice.

With the notable exceptions of ancient Cambodia and Champa, hinterland peoples have been linguistically distinct from their lowland neighbours. Language may accentuate both similarity and difference. Jarai may for instance have become Chamic speakers through their interactions with Cham kingdoms. From O'Connor's work, one expects that the shift from heartland temple-states to hinterland city-states brought an emphasis on increased cultural homogeneity in lowland areas and on differentiation from highland peoples. But this may never have been very simple. The kingdom of Chiang Khaeng, in northern Laos, had ethnically distinct courts; some Tai Khuen (as in Kengtung), some Tai Neua ('Chinese Shans'), and many Tai Lue. In the nineteenth century, each of these courts had relations with distinct groups of highland peoples. Both the uplands and the lowlands were multi-ethnic, while historically there was a simple dichotomy between Tai and Kha peoples (Grabowsky 2003; Grabowsky & Wichasin 2008:6–13). Because this terrain is known as a part of Laos, it is notable that none of the courts were ethnically Lao.

A diversity of peoples coincided with a two-part classification of state-subjects and non-subjects. This historical and ethnographic point is relevant for a long-term view of human diversity on two counts. The first reaffirms the general relevance of White's (2006:103) remark about the complexity of Iron Age Southeast Asia; 'a riot of regional self-expression in everything from ceramics to burial sites to social organization.' The second relates to historicisation and localisation. Particular courts had relations with certain hinterland peoples, and the relationship informed ethnic labels and their scope. Lawa (Lua) formed around and in opposition to Northern Thai and Shan polities, and may have come about originally in relation to the Mon kingdom of Hariphunchai (Lamphun, see Swearer & Premchit 1998; Mangrai 1981; Jonsson 2005). Kha formed around Lao and Lue kingdoms, while the distinction between Khmu and Lamet (Rmeet) may have sharpened because of relations with separate kingdoms.

Island Southeast Asia shows many of the same features of diversity and divisions as does the Mainland. Terms like Batak (Sumatra), Dayak (Borneo), and Toraja (Sulawesi) implied upriver or hinterland populations, but are now ethnic labels. The modern nation state makes them ethnic groups and minorities, but the connections of so-labelled peoples to larger landscapes are a historical problem. In one part of Sulawesi, separate clusters of upland and coastal communities sometimes 'act[ed] in concert [against armed Bugis incursions, but at other times] looked upon one another with suspicion. Above all, these polities were consumed with their own internal rivalries and power struggles [before they collapsed in 1872]' (George 1991:548). In one kingdom on Sumatra; 'two coastal rulers each commanded distinct lines of loyalty with different sections of the Batak population' (Drakard 1990:46). These examples point to intra-elite rivalries in the lowlands and competition among settlements and leader-candidates in the highlands, that came together in various ways in particular settings. This is not to deny various differences between social and cultural formations in relation to shifting cultivation and wet-rice farming, respectively, but to insist that this social landscape does not spell out a firm and lasting binary distinction in relation to the state (contra Scott 2009). Rather, in line with White's



(1995:116) characterisation of the region in terms of heterarchy; ‘hierarchical relationships in politics and society were contextual and flexible.’

Such relations often created inequalities in highland areas, where a leader was able to tax or corvée subordinates, and one settlement might become exempt from duties that were imposed on various others. In this respect, there is considerable similarity to lowland areas, that were differentiated by tribute and service duties. Rulers established themselves by ranking settlements and peoples, uniting and differentiating at the same time, and each kingdom might generate its own set of categories. Any such act of charting the social landscape may erase previous categorisations. That is, the social landscape has continually been remade. The notion of ethnic groups as historical actors risks importing stability and uniformity on a social landscape that was diverse to begin with, and was repeatedly differentiated as it was linked to the prominence of particular kingdoms. Approaching the archaeological or historical landscape in terms of contemporary national categorisations of peoples will miss how profusely identity was tied to smaller kingdoms and determined in tandem with rank and tribute duties.

### **3 Ethnology’s lessons**

One fundamental lesson from ethnology concerns the active production of both sameness and difference, that people make resemblances as well as distinctions. O’Connor highlights how language and livelihood went together with ritual complexes, in ways that account for the apparent shift from Pyu, Mon, Khmer, and Cham to Burmese, Thai, and Viet peoples. Ethnicity, as the clustering of ritual, language, and livelihood, is a manifestation of a society-defining paradigm that is both in history and leaves its mark on the historical landscape.

My research on Thailand’s Mien people, and by extension the Yao of southern China, indicates that rank was a diacritical feature of pre-national identities (Jonsson 2001, 2005). Yao and many others became defined as the state’s outsiders. They were to live in the forested wilderness and owed the state or its officials no respect or taxes. There are signs of similar dynamics in various parts of MSEA over the last millennium. Material on the Yao and Lawa peoples suggests that there was a regional or areal process of people-making that was anchored to specific alignments of political economy, culture, and understandings of the natural environment. The forest became synonymous with the state’s outside, and forest people were those who stood outside the state. In cleared lowland areas associated with state control, there was considerable diversity in how people could negotiate their status, rights, and duties.

Commonly there were considerable lowland populations of slaves, people who had no rights or recognition on their own but belonged to others. Slaves and hinterland peoples stood as the two outsides to the project of defining society in terms of the state. There are many indications of a market in identity; people had to pay through service, tribute, or in another currency to be acknowledged as of a certain rank. One example of such contract is an inscribed silver-plate that dates from 1852. It grants Lawa exemption from corvée duties in return for tribute to the court of Chiangmai. The grant was bestowed upon several leaders, who had the titles Phaya, Khun, and Lam:<sup>1</sup>

---

<sup>1</sup> Phaya and Khun were also official titles in lowland states, while the lower ranked Lam may only have been used for highland leaders.

In consideration of their tribute of 220 pieces of fine silver, the equivalent of 600 shoulder loads (of rice), annually deposited in the royal treasury, we do hereby exempt and release them from all ... duties to our government ... We forbid our [officials], both present and future, to impress their labour ... Fee for enforcing this order 50. Fee for receiving this document 100 (Nimmanhaeminda 1965:235–36).

Taken at face value, this document offered ‘the three million Lawa’ exemption from duties to the court if they handed over a given amount of silver annually. If they did not hand over the silver, they should pay in rice. The responsibility for this contract was placed with the five titled Lawa leaders. It offered protection against duties imposed by lower ranking lowland officials and promised to deliver a curse on the Lawa chief if the Lawa population did not prosper and/or if he oppressed them.

If the contract was respected, it would deliver blessing. But if it was not heeded then the Lawa would be made to perform *corvée* duties like ordinary subjects. Their elephants and women would be appropriated, local lowland authorities would not show any restraint in drawing on Lawa labour, and/or the Lawa rulers would impose too many duties on the commoners, which caused the Lawa to abandon their chief. To establish the protection that the grant promised the Lawa, people had to purchase the silver plate, presumably from the Chiangmai ruler, and then to pay for each occasion that it was activated. The language of this contract suggests that identities and social relations were continually reworked within a larger social and political landscape, in part through service, tribute, and payment. This dynamic contributed to various reconfigurations of the landscapes of identity, livelihood, and social relations, a paradigmatic shift in people-making somewhat analogous to the case O’Connor makes for the lowlands.

Is there a way to reach deeper into the past, to an era prior to states? Villages are fundamental social and ritual units, that draw on local customs for negotiating with the outside world of other peoples, of states, and of spirits. A farming village is the expression of a contract with spirits, for people’s health and well-being as much as for success in farming. There is considerable diversity in how villages define themselves. In Ratanakiri Province in northeast Cambodia in 1992, Tampuan villagers pooled resources for annual offerings to village guardian spirits, while each household was responsible for its own offerings to field spirits. Neighbouring Kreung (Brao) made a very small offering to a village guardian spirit, but pooled resources to purchase an ox for the field spirit. Farming and feasting played out in different ways, with Tampuan emphasising the village in feasting and the household in agriculture while Kreung emphasised the village in agriculture and the household in feasting. Jarai, who lived further to the east, emphasised larger, matrilineal kin-groups in both agriculture and feasting.

One may take the difference in ritual and social life as an expression of ethnicity, and as easily propose that the ethnic reference is a way to highlight local differentiation among peoples who occasionally intermarry and in many cases are multi-lingual. In early 1992, a local man who served as commune headman, under the Cambodian government, made an offering for the prosperity of the commune’s 1700 inhabitants. Among the spirits he invoked was that of King Sihanouk (Jonsson 1998). This ritual may appear new, while the ones about villages and fields may appear traditional. But if the unit of analysis is not ethnicity and community but rather individual actors and acts in historical context, then any one of these ritual frameworks is equally pertinent for how people are fashioned as a unit and in relation to their social universe.

If ritual demarcates community and identity, it is telling that a royal spirit is called upon as a supra-village unit is defined as an entity. Reaching above the village requires connections to supra-local forces, and this may offer some perspective on the attraction and even necessity of domesticating Hindu and Sinitic divinities for the purpose of state-formation in the region (Mus 1975; Wolters 1999). The commune-level ritual simultaneously demarcated a community and established the prominence of an individual. Any definition of community draws on, and expresses, its larger context. A century ago, when this area was involved in raiding and trading slaves, rituals most likely focused more on protection and military prowess than on feasting and agriculture. Warfare brings about particular dynamics of identity work, it entrenches leadership positions, implies supra-local forces (both spirits and states), and generates anxieties about social boundaries that are generally absent in times of peace when households and villages are freer to pursue their own goals of farming, kinship, trade, and the like (Jonsson 2001, 2009).

Evidence of warfare is not prominent in the archaeological record for prehistory (White 1995:111), whereas the historical record indicates extended periods of low-level warfare as well as ongoing raids for slaves within lowland areas as well as into the highlands. The emergence of states appears to generate the motivation for warfare, perhaps mostly to acquire craftspeople and farming populations but also to deny glory to status-rivals. Beemer (2009) argues that slave raids were an important vehicle for cultural transmission in the region, and suggests that historians have overplayed the role of 'foreign groups and ideologies' (2009:483). He gives the examples of dance theatre and lacquer engraving (Chinese and Indian cultural influences), and argues that the region became more culturally uniform as such modes of expression were transmitted, through violent capture, within Southeast Asia. Warfare was 'a force for stimulating a somewhat constant and sometimes large-scale series of forced migrations' (2009:489, see also Grabowsky 2004, 2005).

Beemer's case is important also for its focus on the on-going, historically specific process of creating identity and differentiation, away from the preoccupation with a practice's origins. Burmese and Thai populations had more in common over time because of the repeated practice of waging war and raiding peoples. War evokes oppositional identification and antagonism, and the historical record has innumerable references to wars against ethnic others in this region. In times of peace, social relations are likely to have crossed ethnic lines more easily. It is impossible to write about the region without using ethnic labels, but there is good reason to avoid taking such references for granted as one-dimensional in social, cultural, or biological terms. It is useful to think about ethnicity through kinship. Carstens (1995) shows that people make relations through kinship, there is nothing inert about the process. And she draws attention to two sides of this process; social creativity and coercive incorporation. Both dimensions apply to how ethnicities have been relevant within MSEA.

For the purposes of bringing ethnology to the prehistory of MSEA, I emphasise that ethnicity is a term of engagement, with like units (kin-groups and settlements) as well as with outsiders; be they people, spirits, or states. While Orang Asli has replaced the term Sakai as a reference to the indigenous foragers on the Malay Peninsula, it is worth exploring how the term linked so-called peoples to states and the trade in forest products. In the past, the term Sakai was often taken as synonymous with 'slave'. At different historical moments, this term also referred to 'subjects', 'dependents', and 'allies'. It later became an ethnic label for some hinterland populations; peoples who in recent times are known as Orang Asli, 'original peoples' (Couillard 1984; Porath 2002). These various understandings have always been central to the definition of what is a Malay and a charting

of a Malay(-sian) social universe, issues that in themselves have never been clear or settled (Nagata 1974; Benjamin 2002; Mandal 2003). Sakai does not necessarily mean an ally, a slave, or a specific ethnic group, but it has had any one of those references in particular settings and from particular perspectives—it has never been simply about a people as uniform and self-referential. It is potentially quite specific to time and place who and what a people can be, and how they can relate to others through trade, warfare, kinship and other relations.

Foragers, unlike highland farmers, did not live in permanent settlements, and the absence of village structures and attendant ritual frameworks suggest the lack of formalised relations with outsiders. But from the case of Sakai it is possible to argue that the relationship went along with an identity and a definition of which locals could deal with the outside world. The relationship may even have created the identity, while the local implications of the identity ranged among slaves, subjects, and (more autonomous) allies.

During the colonial era, the trade in forest products dwindled in importance, while bulk goods such as rice and teak became more significant for regional economies. This led to various changes in how lowland and coastal states related to hinterland peoples, both foragers and farmers. It is important to stress the interconnection of identities and relations, be they about raids, servitude, trade, farming, warfare, or kinship. Ethnology's focus on peoples in their context can expand the applicability of historical ontology (Hacking 2002), the focus on how particular kinds of people come about, in ways that bridge disciplinary boundaries. Linguistic research has shown many of the benefits of situating languages in areal context, and moving away from taking a language or a language family as given. A focus on individual linguistic items and on individual speakers brings out unexpected connections among Chamic, Mon-Khmer, and Tai speakers (Enfield 2005), in the same way that historical research brings out cultural commonalities among Burmese and Thai (Beemer 2009).

Identities are created through relations and contrasts. Lowland peoples became an entity in part through a deliberate differentiation from highland farmers. Thai self-fashioning has rested in part on animosity against Burmese people and kingdoms. Identity work is partly ideological, it fashions a people that appears uniform and distinct, in a way that makes unthinkable the manifold entanglements of a people with various others (Herzfeld 1987, 1997; Harrison 2006; Jonsson 2010). Antagonistic and oppositional identifications are ideological, they serve some political purposes and may not come anywhere close to a description of the many entanglements of social life across ethnic and other categories. Because the sense of identity assumes a coherent subject, such as an individual, a village, an ethnic group, or a nation, the process of identification tends to erase traces of whatever relations are formally deemed transgressive.

Among the issues to examine are signs of diversity among prehistoric settlements, whether for instance the archaeological record indicates that social identities became progressively more uniform with increased interaction. In the case of Ancient Greece, there is some indication that contests of various kinds facilitated increasing family resemblance among previously-rival and significantly-varied cities, so that it is more accurate to argue that engagements through athletics created the Greeks than to say that the Greeks had athletics (Renfrew 1988).

Kinship and ethnicity draw on a combination of social creativity and coercive incorporation, but people experience and express them as a natural process of continuity. Taking a regional and historical view brings out the historical contingency of contemporary ethnic classifications, most of which make national majorities seem obvious

and minorities problematic. The process of nation-building brought an identity to Burmans, Vietnamese, Thai, and other such groups that made previous divisions of region, dialect, class, and sub-caste servitude disappear in the name of national unity. National unity and the nation's link to the landscape seem somehow natural, but a regional and comparative perspective can illuminate how the process drew on a combination of social creativity, coercive incorporation, and selective exclusion. Many dimensions of the process concern ethnicity. Modern states recognise ethnicity in ways that reinforce particular national narratives—Thai are historically transcendent through relations to Buddhist monarchs and in a way that makes any non-Thai peoples aliens on the landscape, while the Vietnamese and fifty-three minorities are united through their struggle against foreign aggression.

These are modern narratives that marginalise any sub-national identities and make previous historical realities and attendant identities unthinkable. As O'Connor suggests for earlier periods of the region's history, the modern state rests on a society-defining paradigm through which a national self and various others come into being as particular kinds of subjects. This erases whatever earlier historical realities attached to labels such as Sakai, Lawa, Kha, and Yao, and replaces them with the nation state's terms of engagement. The nation state's anchor in modernity creates expectations about anything that came before as tradition or as *ancien regime*, which risks precluding recognition of continuities in how Chinese as much as Southeast Asian states have repeatedly been deeply involved in projecting particular kinds of identity and rank, partly through images of their ethnic outside on various highland peoples (Jonsson 2005; Fiskesjö 2006). Ethnic notions such as Chinese (Han, Tang, etc.), Thai, and Yao suggest historical continuity, transcendence, and singularity, whereas any such subject is a project that is brought into being and agency through internal and external difference. By not taking seriously the historicity and relational specificity of social categories, we risk undermining our ability to get at the past through a critical juxtaposition with the present.

Enfield (2005) urges linguists to 'work through the implications of the view that "the language" and "the community" are incoherent as units of analysis of causal processes in the historical and areal trajectories of language diffusion and change' (2005:198). Ethnicity is a term of engagement and also a political project. Without the areal, local, and historical context of how identities have acquired resonance through social relations, such as regarding livelihood, trade, kinship, state projects, and ritual, the use of contemporary ethnic labels for past realities is analytically and descriptively suspect. Plotting ethnic groups on the past in terms of linear evolutionary models, with a focus on markers of achievement such as agriculture, chiefdom, state-organisation, or the ladder of Stone, Bronze, and Iron Ages, radically misconstrues the region's prehistory. Such work holds time and space constant and plots people in terms of their relative advancement on a pre-set scale.

If ethnology is to have relevance for interdisciplinary dialogue about human diversity in the past, and to have points of entry, the underlying assumptions must shift from Galilei-Newton and toward Einstein, toward a sense of "space-time as a structural quality of the field" (Einstein 1961:176). Changing our assumptions will facilitate a rethinking of the region in ways that consider people as social actors and points toward the interplay of events and structures that has lent an areal shape to communities, identities, livelihood, and power relations at particular moments (Halcrow & Tayles, this volume; White 1995). Our ability to move between micro-and macro scales may revise the region's periodisation, away from clear and progressive stages within an inert field and toward the ongoing production of identity and difference where forest products, dance drama, warfare, vowels,

inequality, and farming methods have a potentially equal significance in illuminating and exemplifying how the region and its peoples have taken shape through their manifold actions and interactions.

#### **4 Conclusions**

Ethnological work encourages an emphasis on how relatedness and difference are created through a combination of coercive incorporation, social creativity, and selective exclusion. The focus on society-defining paradigms offers a dynamic perspective on how peoples and languages were consolidated around particular adaptations and schemes of livelihood. The process is non-linear, with languages and social forms spreading in expression of alternative paradigms and having different outcomes in upland and lowland areas. Linguistic and social differentiation has an ideological side that conceals the extent to which people are formed through interactions with others. Highland shifting cultivators took shape in contrast to lowland farmers, but the distinction also conceals the frequency and interactions across the upland-lowland divide. Identity is a term of engagement. Villages are frameworks for relations with outsiders, but the case of hinterland peoples on the Malay Peninsula suggests that such identities and relations could take shape in the absence of villages. This examination of the hinterlands shows that warfare, trade, and relations with particular kingdoms were very influential for identity as much as for patterns in social life, livelihood, and ritual. Any people may be quite differentiated internally, and their identity takes shape through relations with others. Thus ethnology resists the notion that ethnicity implies a people as uniform and primarily self-referential. Modern scientific notions of ethnic groups tend to erase the relational and historically contingent dynamics of people-making, and this is where ethnology can make important contributions to interdisciplinary dialogues about human diversity in MSEA over time.

#### **Acknowledgements**

I am indebted to Nick Enfield and Joyce White for the opportunity to take part in the stimulating workshop in Siem Reap, and to them and two anonymous reviewers for very helpful comments on an earlier version of this chapter.

#### **References**

- Abadie, Maurice. 1924. *Les Races du Haut-Tonkin, de Phong-Tho à Lang-Son*. Paris: Société d'Éditions Géographiques, Maritimes et Coloniales.
- Beemer, Bryce. 2009. Southeast Asian slavery and slave-gathering warfare as a vector for cultural transmission: The case of Burma and Thailand. *The Historian* 71(3). 481–506.
- Bellwood, Peter. 1992. Southeast Asia before history. In Nicholas Tarling (ed.) *The Cambridge History of Southeast Asia*, vol. 1, 55–136. Cambridge: Cambridge University Press.
- Benjamin, Geoffrey. 2002. On being tribal in the Malay World. In Geoffrey Benjamin & Cynthia Chou (eds) *Tribal Communities in the Malay World*, 7–76. Singapore: Institute of Southeast Asian Studies.
- Bonifacy, Auguste. 1919. *Cours d'ethnographie Indochinoise*. Hanoi: Imprimerie d'Extreme-Orient.

- Carstens, Janet. 1995. *The Heat of the Hearth: The Process of Kinship in a Malay Fishing Community*. Oxford: Oxford University Press.
- Chou Ta-Kuan. 1967. *The Customs of Cambodia*. Translated by John Gilman d'Arcy Paul, from the French translation of Paul Pelliot. Bangkok: The Siam Society.
- Clark, Geoffrey. 2002. Neanderthal archaeology: Implications for our origins. *American Anthropologist* 104(1). 50–67.
- Condominas, Georges. 1990. *From Lawa to Mon, From Saa to Thai: Historical and Anthropological Aspects of Southeast Asian Social Spaces*. Canberra: Department of Anthropology, Research School of Pacific Studies, Australian National University.
- Couillard, Marie-Andrée. 1984. The Malays and the 'Sakai': Some comments on their social relations in the Malay peninsula. *Kajian Malaysia* 2(1). 81–104.
- Cupet, P. 1998 [1893]. *Among the Tribes of Southern Vietnam and Laos*. Bangkok: White Lotus.
- Drakard, Jane. 1990. *A Malay Frontier: Unity and Duality in a Sumatran Kingdom*. Ithaca: Cornell Southeast Asia Program.
- Durrenberger, Paul & Nicola Tannenbaum. 1990. *Analytical Perspectives on Shan Agriculture and Village Economics*. New Haven: Yale Southeast Asia Studies.
- Einstein, Albert. 1961. *Relativity: The Special and General Theory*. New York: Three Rivers Press.
- Enfield, N. J. 2005. Areal linguistics and Mainland Southeast Asia. *Annual Reviews of Anthropology* 34(1). 181–206.
- Fiskesjö, Magnus. 2006. Rescuing the empire: Chinese nation-building in the Twentieth Century. *European Journal of East Asian Studies* 5(1). 15–44.
- George, Kenneth M. 1991. Headhunting, history, and exchange in upland Sulawesi. *Journal of Asian Studies* 50(3). 536–64.
- Gould, Stephen Jay. 1981. *The Mismeasure of Man*. New York: Norton.
- 1991. Life's little joke. *Bully for Brontosaurus: Reflections in Natural History*, 168–81. New York: Norton.
- Grabowsky, Volker. 2003. Chao Fa Dek Noi and the founding myth of Chiang Khaeng. In Yukio Hayashi & Thongsa Sayavongkhamdy (eds) *Cultural Diversity and Conservation in the Making of Mainland Southeast Asia and Southwest China*, 95–143. Kyoto: Center for Southeast Asian Studies, Kyoto University.
- 2004. *Bevölkerung und Staat in Lan Na*. Wiesbaden: Harrassowitz.
- 2005. Population and state in Lan Na prior to the mid-sixteenth century. *Journal of the Siam Society* 93. 1–68.
- Grabowsky, Volker & Renoo Wichasin. 2008. *Chronicles of Chiang Khaeng*. Honolulu: Center for Southeast Asian Studies, University of Hawaii.
- Hacking, Ian. 2002. *Historical Ontology*. Cambridge, MA: Harvard University Press.
- Halcrow, Siân & Nancy Tayles, this volume, Human Diversity in Mainland Southeast Asia: the contribution of bioarchaeology.
- Hallett, Holt. 1890. *A Thousand Miles on an Elephant in the Shan States*. London: William Blackwood and Sons.
- Harrison, Simon. 2006. *Fracturing Resemblances: Identity and Mimetic Conflict in Melanesia and the West*. New York: Berghahn.

- Herzfeld, Michael. 1987. *Anthropology through the Looking-Glass: Critical Ethnography on the Margins of Europe*. Cambridge: Cambridge University Press.
- . 1997. *Cultural Intimacy: Social Poetics in the Nation State*. New York: Routledge.
- Izikowitz, Karl Gustav. 1951. *The Lamet: Hill Peasants of French Indochina*. Gothenburg: Etnografiska Museet.
- Jonsson, Hjorleifur. 1998. Dead headmen: Histories and communities in the Southeast Asian hinterland. In Ing-Britt Trankell & Laura Summers (eds) *Facets of Power and its Limitations: Political Culture in Southeast Asia*, 191–212. Uppsala: Uppsala Studies in Cultural Anthropology.
- . 2001. Does the house hold? History and the shape of Mien society. *Ethnohistory* 48(4). 613–54.
- . 2005. *Mien relations: Mountain people and state control in Thailand*. Ithaca: Cornell University Press.
- . 2009. War's Ontogeny: Militias and Ethnic Boundaries in Laos. *Southeast Asian Studies* (Kyoto, Japan) 47(2). 125–49.
- . 2010. Mimetic Minorities: National Identity and Desire on Thailand's Fringe. *Identities: Global Studies in Culture and Power* 17(2). 108–30.
- Keyes, Charles F. 2002. The peoples of Asia – Science and politics in the classification of ethnic groups in Thailand, China, and Vietnam. *Journal of Asian Studies* 61(4). 1163–1203.
- King, Victor T. & William D. Wilder. 2003. *The Modern Anthropology of South-East Asia*. New York: Routledge.
- Kirsch, A. Thomas. 1973. *Feasting and Social Oscillation: Religion and Society in Upland Southeast Asia*. Ithaca: Cornell Southeast Asia Program.
- . 1985. *Cosmological Factors in the 'Collapse' of the Khmer and the 'Rise' of the Thai*. Presented to the annual meeting of the Association for Asian Studies, Philadelphia.
- Leach, Edmund R. 1954. *Political Systems of Highland Burma*. Boston: Beacon Press.
- Mandal, Sumit. 2003. Transethnic solidarities in racialized context. *Journal of Contemporary Asia* 33(1). 50–68.
- Mangrai, Sao Saimong. 1981. *The Padaeng Chronicle and the Jengtung State Chronicle Translated*. Ann Arbor: University of Michigan Papers on South and Southeast Asia.
- Marks, Jonathan. 2007. Anthropological taxonomy as subject and object: The consequences of descent from Darwin and Durkheim. *Anthropology Today* 23(4). 7–12.
- Mus, Paul. 1975. *India Seen From the East: Indian and Indigenous Cults in Champa*. Translated by Ian W. Mabbett & David Chandler. Monash: Center for Southeast Asian Studies, Monash University.
- Nagata, Judith. 1974. What is a Malay? Situational selection of ethnic identity in a plural society. *American Ethnologist* 1(2). 331–50.
- Nimmanhaeminda, Kraisi. 1965. An Inscribed Silver-Plate Grant to the Lawa of Boh Luang. *Felicitations Volumes Presented to His Highness Prince Dhanivat*, vol. 2, 233–236. Bangkok: The Siam Society.
- O'Connor, Richard A. 1995. Agricultural change and ethnic succession in Southeast Asian states: A case for regional anthropology. *Journal of Asian Studies* 54(4). 968–996.



- . 2000. A regional explanation of the Tai Muang as a city-state. In Mogens Herman Hansen (ed.) *A Comparative Study of Thirty City-States*, 431–43. Copenhagen: The Royal Danish Academy of Sciences and Letters.
- Porath, Nathan. 2002. Developing indigenous communities into Sakais. In Geoffrey Benjamin & Cynthia Chou (eds) *Tribal Communities in the Malay World*, 97–118. Singapore: Institute of Southeast Asian Studies.
- Renfrew, Colin 1988. The Minoan-Mycenean origins of the Panhellenic games. In Wendy J. Raschke (ed.) *The Archaeology of the Olympics*, 13–23. Madison: University of Wisconsin Press.
- Scott, James C. 2009. *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. New Haven: Yale University Press.
- Seidenfaden, Erik. 1958. *The Thai Peoples*. Bangkok: The Siam Society.
- Swearer, Donald & Sommai Premchit. 1998. *The Legend of Queen Cama*. Albany: SUNY Press.
- Taylor, Keith W. 2001. On Being Muonged. *Asian Ethnicity* 2(1). 25–34.
- White, Joyce. 1995. Incorporating heterarchy into theory on socio-political development: The case from Southeast Asia. In Robert M. Ehrenreich, Carol L. Crumley & Janet E. Levy (eds) *Heterarchy and the Analysis of Complex Societies*, 101–23. Arlington, VA: American Anthropological Association.
- . 2006. Review of *And through Flows the River: Archaeology and the Pasts of Lao Paks*, by Anna Kallen. *Asian Perspectives* 45(1). 102–04.
- Winichakul, Thongchai. 2000a. The others within: Travel and ethno-spatial differentiation of Siamese subjects 1885–1910. In Andrew Turton (ed.) *Civility and Savagery: Social Identity in Tai States*, 38–62. Richmond, Surrey: Curzon.
- . 2000b. The quest for ‘siwilai’: A geographical discourse of civilizational thinking in the late nineteenth and early twentieth-century Siam. *Journal of Asian Studies* 59. 528–549.
- Wolters, Oliver W. 1999. *History, Culture, and Region in Southeast Asian Perspectives*, revised ed. Ithaca: Cornell Southeast Asia Program.

# *Part II*

*Archaeological advances*



# 7 *The role of agriculture in the evolution of mainland Southeast Asian language phyla*

---

ROGER BLENCH

## 1 Introduction

The emergence of mainland Southeast Asian (MSEA) nations from decades of war and the gradual opening up of individual countries has created a new focus on the region. Southeast Asia presents an intriguing mix, combining highly diverse ethnolinguistic groups with generally small populations, and more numerous peoples, such as the Thai, Burmese, Lao, Vietnamese and Khmer, who cover significant swathes of territory and are politically dominant. Historically the vast majority of languages of Southeast Asia were unwritten and remain poorly described, but the existence of scripts has created something of a focus on texts. Recent years have seen the emergence of much new data, often quite difficult to access; nonetheless, the linguistic map is gradually becoming clearer, although many questions remain.

There are five major language phyla in mainland Southeast Asia, Austroasiatic, Austronesian, Daic, Sino-Tibetan and Hmong-Mien and no true isolates, except on offshore islands. Despite a considerable expansion of research in recent years, models for the dates, homelands and engines of expansion of these phyla are markedly absent from the literature, as are convincing correlations with archaeological and genetic research. A claim that has generated considerable discussion in recent years is the importance of agriculture and thus demographic expansion in accounting for ethnolinguistic geography. The chapter considers whether the reconstruction of agricultural terminology in individual language phyla supports this claim, and if so, what can be said about the dating of individual phyla. Since Austroasiatic is discussed at greater length elsewhere in this volume (for example, Diffloth, Sidwell and Blench) its treatment will be abbreviated here.

The five language phyla that dominate MSEA are:

- Sino-Tibetan
- Hmong-Mien     [=Miao-Yao]
- Austroasiatic     [partly = Mon-Khmer]
- Austronesian
- Daic             [=Tai-Kadai, Kra-Dai]

---

N. J. Enfield, editor.

*Dynamics of Human Diversity*, 125-152. Pacific Linguistics, 2011

© This edition vested with Pacific Linguistics.

There are virtually no isolates, except Andamanese, possibly Shompen (Blench 2007, in press) and the extinct Kenaboi (Hajek 1998). The protectiveness of the Indian government has made it difficult to establish clearly the nature of Andamanese and even the relation of the languages to one another. Abbi (2006) has gone some way to remedying this, but still the type of rich lexical and grammatical data which could underlie hypotheses of long-range connections remains elusive. A lack of credible archaeology has meant that there are no dates for first settlement of the archipelago. The Andamanese may have been there for a very long time, or they may have been brought there in the boats of others in the last few thousand years. Other foragers, notably the Orang Asli of the Malay peninsula and the negritos of the Philippines, now speak Austroasiatic and Austronesian languages respectively. Linguistic reconstruction does suggest a substrate vocabulary in their modern-day speech (Reid 1994a, b), but we are far from being able to link this fragmentary lexicon with other language phyla.

Due to long periods of interaction and extensive multilingualism, MSEA language phyla have developed many convergent characteristics, as well as being extensively relexified from dominant or contact languages (Enfield 2003, this volume). As a consequence, numerous macrophyla hypotheses have been advanced linking together almost any combination of phyla, notably Austric (for example Higham 1996; Reid 2005 and references therein), Daic/Austronesian (Schlegel 1901; Ostapirat 2005), Austro-Thai (Benedict 1942, 1975), Sino-Tibetan-Austronesian (STAN) (Sagart 2005a, 2008) and 'Proto-East Asian' (Starosta 2005; van Driem 2008). Despite this, our understanding of the proto-lexicon, dating and patterns of dispersal from the homelands of individual phyla remains both sketchy and controversial. In the case of Sino-Tibetan, a failure to make available comparative materials that purportedly underlie proposed reconstructions has made assessment of the true situation difficult.

In Africa and Oceania there is a relatively long tradition of combining archaeology and linguistics to develop a synthesis of prehistory. This may partly reflect the absence of historical documents and large-scale polities, but it is also a consequence of intellectual traditions which favour interdisciplinarity. With a few exceptions, such an integrated prehistory remains to be created for MSEA. It is, however, clearly needed, if we are to go beyond epigraphy to a more global account of the movements and evolution of present-day populations. The complex patterns of language phyla must correlate with the archaeology of the region in some fashion. Historical linguistics can provide both a general insight into the peopling of the region and also help research both the transition from foraging to farming and the history of individual crops and livestock species. This chapter is intended to present preliminary results of research into the reconstructibility of the agricultural lexicon in the language phyla of Southeast Asia and suggest their interpretation in the light of recent archaeological data. The main hypotheses relating to the peopling of Southeast Asia and the inception of agriculture are outlined, and then each of the major language phyla is reviewed in turn. The conclusions draw together the evidence for the dating and possible expansion of these phyla; it should be emphasised that this is highly preliminary. Many of the speculations presented here will need considerably more work, linguistic and archaeological, to refine their application to archaeological data.

## **2 Archaeology of Mainland Southeast Asia**

The origins of the current populations of Southeast Asia have been the subject of much debate. On the basis that Australoid populations must have passed through the region, a 'two-layers' model has generally been proposed. Broadly speaking, this assumes that there

was an original peopling of Australoids, phenotypically similar to modern Andamanese and other negritos, and that these were replaced by mongoloids, apparently migrating down from present-day China, although not at that point Sinitic-speakers. However, skeletal and mtDNA evidence for this hypothesis has been conspicuous by its absence, although well-preserved finds of the appropriate date are very rare. Matsumura et al. (this volume) report on skeletal material dated to 10,450 +/- 300 years BP at Hang Cho, in Luong Son district, southwest of Hanoi, which they claim shows negrito affinities. Lewis et al. (2008) describe a terminal Pleistocene cremation burial on the island of Palawan dated to 9500–9000 BP, the earliest yet known in the region. The languages of these earlier foragers must remain unknown, apart from Andamanese and possible relic vocabulary in Aslian languages.

Southeast Asian societies today are clearly very agriculturally-based, with rice production a fundamental activity throughout the region. Despite this, there is little or no direct evidence for the transition to agriculture, and even the date at which this took place is disputed (Bellwood 1997). Indeed, Southeast Asian archaeology shows a very distinctive ‘Holocene gap’ with a marked lack of sites between ca. 8–4000 BP, the reasons for which are unknown (Joyce C. White, personal communication). One notable exception is the archaeological site of Da But, in Thanh Hoa province of Vietnam, an early Neolithic cemetery and shell midden, radiocarbon dated to 5085 BC (Vinh 1991). The subsistence strategies of the Da But people were mixed hunting, gathering, and fishing, but evidence for animal husbandry and paddy rice cultivation remains controversial.

The most widespread claim for the dating of the Neolithic transition is that of Higham (2004:47) who notes ‘The pattern of intrusive agriculturalists settling inland valleys in southern China, while the coast continued to be occupied by affluent foraging groups, is repeated in the Red River area and the contiguous coast of Vietnam’. The type-site for this type of agriculture is Phung Nguyen, about 200 km inland from Halong Bay. Dates remain problematic, but the adjacent site of Co Loa has been dated to 2000 BC (Lai Van Toi 1999). In summarising the situation, Higham says:

We find agricultural settlements being founded in the lower Red River valley, along the course of the Mekong and its tributaries, and in the Chao Phraya valley...The dates for initial settlement, as far as they are known, are approximately the same with none earlier than about 2300 BC. Most intriguingly, the pottery vessels in many of the sites over a broad area have a similar mode of decoration. The sites reveal extended inhumation graves and an economy incorporating rice cultivation and the raising of domestic stock. (Higham 2002:352).

In contrast, White (1995) and White et al. (2004:123) say ‘based primarily on dates from basal deposits from Ban Chiang and Ban Tong, two long-term settlements in the Kumphawapi catchment, that societies cultivating plants appeared in the region by the mid-fourth millennium BC cal.’ If so, agriculture would be as much as 1500 years earlier than the Higham model.

Higham and Higham (2009) are now proposing a new chronology for the beginning of the Neolithic in MSEA, based on the new C14 chronology of Ban Non Wat, which derives from a large sample of directly dated shell, analysed with Bayesian statistics. This would imply a revision of all existing dates towards the more recent period, from 1800/1700 to 1100 BC. (Fiorella Rispoli, personal communication) considers the recent C14 dated excavations at An Son and Da Kai in South Vietnam are fully in accord with this. On the basis of comparisons between the sites she has excavated in Central Thailand and most of

the other Neolithic sites in Thailand, Vietnam and Yunnan, the new dates ‘put all the tiles in the right place’, linking MSEA with Yunnan as well as Guangxi/Guangdong.

The claims of White and her collaborators are based on indirect environmental evidence, rather than direct archaeobotanical materials. The conflict between their views and the main body of Southeast Asian archaeologists might be reconciled if farming was preceded by a long period of intensive landscape management, but distinguishing between these two interpretations may be difficult based purely on the archaeological record.

### 3 Sino-Tibetan

The Sino-Tibetan phylum has more speakers than any other language phylum, largely due to the inclusion of the Sinitic branch, composed of the many varieties of Chinese. Despite some two centuries of study and publication, the subclassification of Sino-Tibetan remains highly controversial, as does its external affiliation (Blench 2008). Considering the importance of Sino-Tibetan and its history of scholarship, there is a striking lack of agreement as to its internal classification. Some key questions are:

- whether the primary branching is Sinitic (that is all Chinese languages) versus the remainder (usually called Tibeto-Burman) or whether Sinitic is simply part of one branch, for example Bodic et cetera. Certainly the distinctiveness of Sinitic is far from proven.
- what are the inter-relations of its branches?
- can it be linked with other phyla such as Austronesian or Caucasian (as proposed for example by Sagart 2005a, 2005b; Starostin 1991).

Broadly speaking, the opposing camps are those who consider Sinitic as the primary branching of Sino-Tibetan (Wolfenden 1927; Benedict 1972, 1976; Bodman 1980; Weidert 1987; Matisoff 2003, 2008; Bradley 1997; Thurgood and La Polla 2003) and those who situate it within the remaining languages, consequently applying the name Tibeto-Burman to the whole phylum (Shafer 1966/67; van Driem 1997). Sinitic would thus be incorporated within the group conventionally defined in opposition to it. The two markedly different views are shown in Figures 1 and 2.

The groups represented in Figure 1 are by and large ‘geographic’ categories; Kamarupan and Himalayan have no status as linguistic subgroupings. Even this view has never been justified in print, despite the space afforded by the 800 pages of Matisoff (2003). Moreover, since Matisoff excludes many small branches of Sino-Tibetan and joins many others at a single node, this is far from a fully worked-out theory. An agnostic alternative is represented by van Driem (2005), in his ‘fallen leaves’ schema (Figure 2). van Driem’s model presents no assumptions at all about subgrouping except to map already well-recognised groups. This is a geographical model, which places generally agreed subgroups in proximity, with area of the ellipse corresponding to their size, but advances no hypothesis about their ultimate relationships. Whether this represents progress is debatable, but the ‘fallen leaves’ model has the virtue of treating all branches of Sino-Tibetan as of equal status and requiring that their position be ultimately defined. Van Driem would argue that this is a fair representation of the current state of our knowledge.

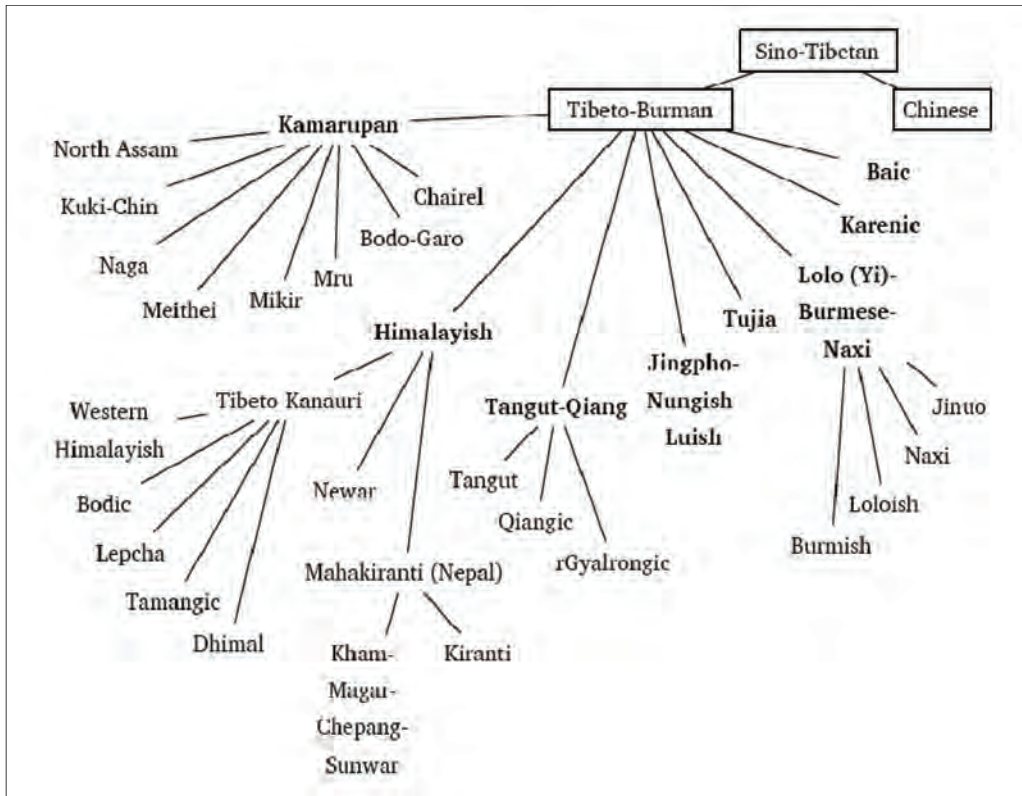


Figure 1: Sino-Tibetan according to Matisoff (2008).

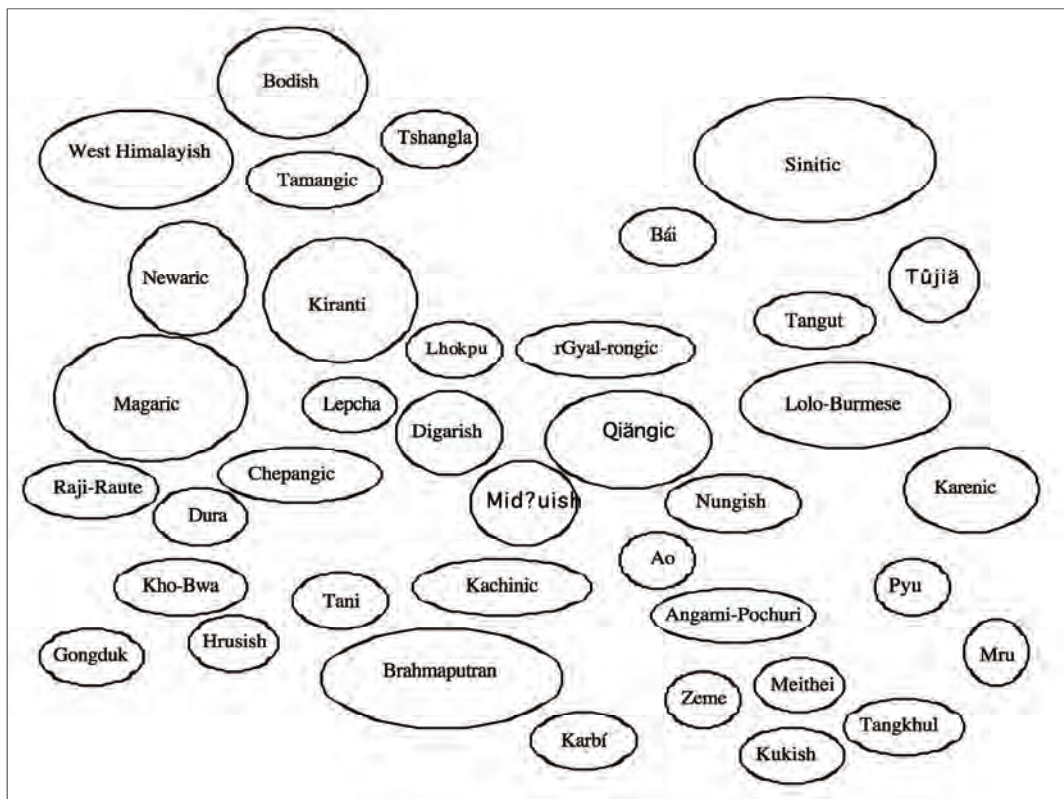


Figure 2: 'Fallen leaves' model of Sino-Tibetan according to van Driem (many places).



Both these classifications essentially show large parallel arrays, with van Driem being the extreme version of the agnostic view. Clearly, the building of a hierarchical model of Sino-Tibetan appears to be a long way off. Neither model seems to give fair weight to the highly diverse languages of Northeast India, for which documentation is gradually becoming available<sup>1</sup>. In terms of internal diversity, the region from the southern flanks of the Himalayas to the Assam region is massively diverse synchronically, with large numbers of small subgroups which appear to be very different from one another. We may have to suppose an original diversity more characteristic of Northeast Asia or parts of the Amazon. As Northeast India and adjacent regions open up, the striking linguistic diversity of this region is becoming more apparent. For example, Sherdukpen, Bugun (Dondrup 1990) Lishpa, and Butpa are listed in the *Ethnologue* as Tibetic languages with no evidence. Examination of the actual data on these languages provides almost no support for such an affiliation; indeed even their membership of Sino-Tibetan is only supported by a small number of lexemes which could well be borrowings. Similarly with Hrusish spoken between Assam and Bhutan in Arunachal Pradesh<sup>2</sup> whose classification remains highly uncertain (Shafer 1947; Simon 1970). It has yet to be proven that some groups are Sino-Tibetan at all rather than isolates with a Sino-Tibetan superstrate.

Related to this diversity is the absence of the classic grain-based agriculture implicit in much of the Sino-Tibetan literature. The supposedly Tani-affiliated Milang emphasise vegetable and hunting strongly in contrast to cereals, and terms for domestic animals, for example, all appear to be recent borrowings. The Sulung, who also speak a language of uncertain affiliation (Tayeng 1990) base their subsistence on sago-exploitation and hunting. Either we assume that the classification of these languages is in error, or that it is not the case that we can confidently reconstruct *any* agricultural terms to Proto Sino-Tibetan, simply because there are no certain attestations in numerous subgroups, especially in the Himalayan and Northeast Indian branches. Logically, therefore, the region of Arunachal Pradesh may well be a remaining zone of high diversity in early Sino-Tibetan from which the later, secondary rice cultivating, pig-producing cultures evolved. Unfortunately, the absence of well-dated, stratified archaeological sites in this region makes it impossible to correlate with archaeology at present.

Northeast India is far from the only problem with Sino-Tibetan. The Sinitic region also includes at least two languages, Tujia and Bai, which are single branches of Sino-Tibetan and which seem to have no particular relationship with Sinitic except for numerous borrowings at many historical levels. Unlike the foragers of Northeast India, these groups are strongly agricultural, but appear to preserve archaic vocabulary pointing to a pre-Sino-Tibetan presence in the region. Tujia has a raft of non-Sinitic agricultural terminology, which has either been innovated or shows links with other Sino-Tibetan subgroups (Table 1).

---

1 Thanks especially to Mark Post, who has kindly collected and made available a wealth of local publications on Northeast India.

2 Shafer (1947:184) says 'A glance the Hruso vocabulary of any author except Campbell reveals almost no similarity to any known Tibeto-Burmic language, and it has been with the greatest difficulty that enough Hruso comparisons have been gathered to show the genetic relationship of that language to Tibeto-Burmic and to establish a few correspondences'.

**Table 1:** Unexpected agricultural vocabulary in Tujia

<b>Tujia</b>	<b>Gloss</b>	<b>Parallels</b>
si <sup>1</sup> li <sup>1</sup>	rice	
ye <sup>3</sup>	cooked rice	
zi <sup>3</sup>	cooked rice	
qie <sup>1</sup> hhe <sup>1</sup>	glutinous rice	cf. Qiang <i>qhə'</i>
oŋ <sup>1</sup> ba <sup>1</sup>	sorghum	
loŋ <sup>1</sup> moŋ <sup>4</sup>	barley	
si <sup>3</sup> tuo <sup>4</sup>	garlic	
kuo <sup>1</sup> su <sup>1</sup>	ginger	
ge <sup>3</sup>	horse	
wu <sup>2</sup>	cow	? reduction of widespread # <i>ŋu</i>
zi <sup>4</sup>	pig	? cf. Chinese <i>shǐ</i> (豕)
ruo <sup>2</sup>	goat	
ha <sup>4</sup> lie <sup>3</sup>	dog	
ra <sup>3</sup>	chicken	cf. Dayang (Qiangic) <i>ro</i>
sa <sup>4</sup>	duck	? Austroasiatic e.g. Bagan <i>mtfa</i> <sup>13</sup>
ŋo <sup>3</sup>	goose	cf. Mantsi Meo Vac (Loloish) <i>ŋɔ</i> <sup>31</sup> <i>ŋɔ</i> <sup>44</sup>

Bai shows many fewer such items, probably because so much of its basic lexicon has been replaced by Chinese loans (Table 2).

**Table 2:** Unexpected agricultural vocabulary in Bai

<b>Bai</b>	<b>Gloss</b>	<b>Parallels</b>
te <sup>42</sup>	pig	cf. Kayah Li <i>thé</i> , Biao Min <i>twə</i> <sup>4</sup>
ky <sup>21</sup>	buckwheat	
me <sup>55</sup> zo <sup>42</sup>	barley	cf. proto-Lolo-Burmese * <i>zəy</i> <sup>2</sup>
tɛũ <sup>33</sup> ŋɔ <sup>21</sup>	sheep	
χua <sup>55</sup> lao <sup>31</sup>	cat	cf. Naxi <i>χua lɛ</i>

These data suggest a pre-Sinitic presence of Sino-Tibetan-speaking agriculturalists throughout much of this region that was largely absorbed following the expansion of the Han Chinese. These *might* be correlated with the earliest Neolithic communities in North China such as the Péilígǎng or Císhān (6500 BP onwards) but Hmong-Mien speakers are equally likely candidates. Fuller et al. (2008) have recently questioned the dating of many of the early rice-producing communities in central China, suggesting that many finds are wild rice and that domestication only really begins by 6500 BP.

Given this situation, we cannot confidently reconstruct *any* agricultural terms to Proto-Sino-Tibetan, simply because there are no certain attestations in numerous subgroups, especially in the Himalayan and Northeast Indian branches. The presence of a term in Lahu unfortunately does not guarantee its reconstructibility to Proto-Sino-Tibetan. Agriculture presumably developed well after the primary dispersal of Sino-Tibetan, which must therefore be considerably earlier than the other language phyla in the region. Given this, there are widespread terms attesting agriculture in well-studied branches such as Sinitic, Karenic and Lolo-Burmese and these must certainly reflect the importance of farming in the secondary expansion of Sino-Tibetan. Table 3 presents my proposals for quasi-

reconstructions (that is, not the result of inspection of systematic sound-correspondences) of crop and livestock terms.

**Table 3:** Widely attested agricultural terms in Sino-Tibetan

	<b>Gloss</b>	<b>Comment</b>
#mei	rice	[also in Daic and Hmong-Mien]
#fan	rice	[also in Daic]
#tʃək	foxtail millet	[also in Mienic and ?Austronesian]
#ŋwV	cow, ox	[also in Daic and Austroasiatic]
#brak	pig	[also in Austronesian]

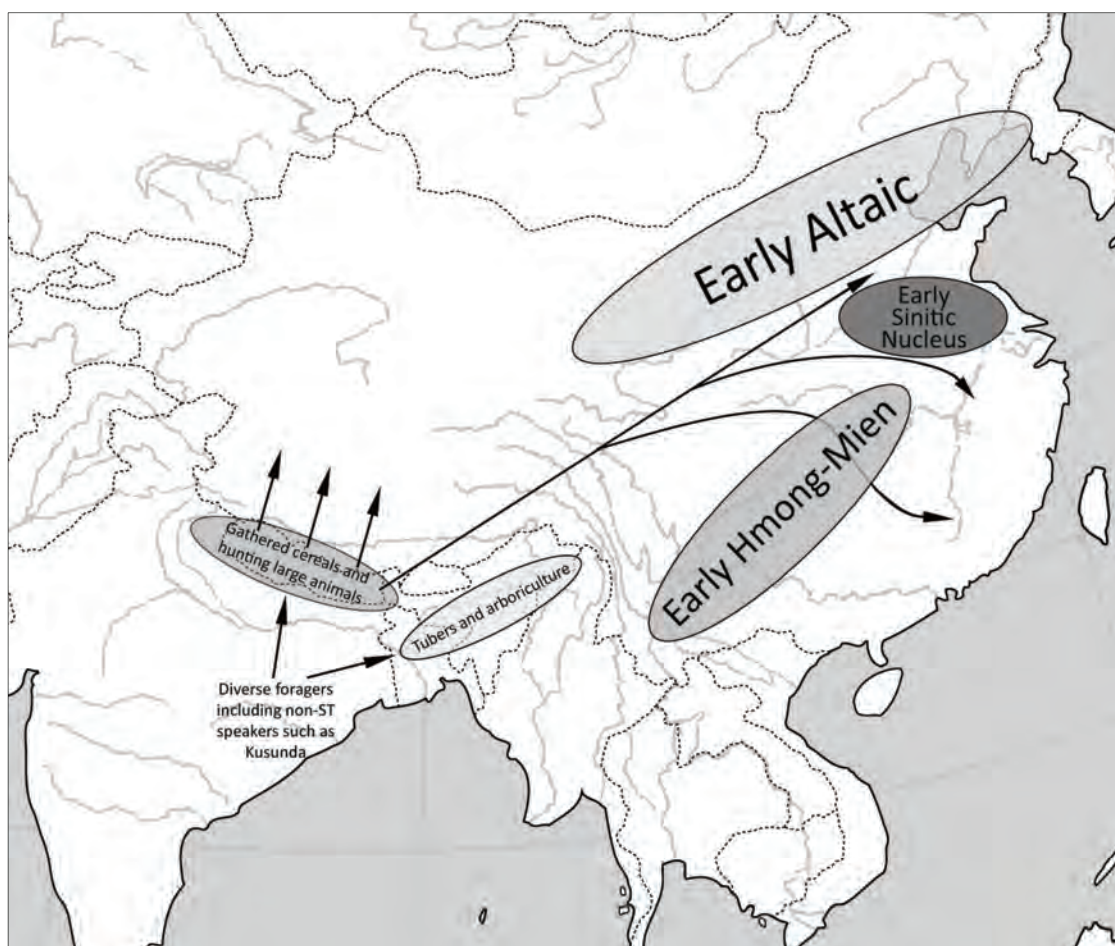
It is notable that all these terms are found outside Sino-Tibetan, especially in Daic, which points strongly to a period of intense interaction in the early phases of the intensification of agriculture.

The evidence for early Sino-Tibetan is marked by gaps in the data; an absence of reflexes for agricultural terms in many of the smaller branches of Sino-Tibetan, a lack of evidence for coherent internal structures and a failure of congruence with archaeology and genetics. Given this, any hypothesis concerning its spread and diversification must be speculative and subject to revision. However, we can do better than any of the claims presently on the table by presenting an account which at least does not contradict the interdisciplinary evidence. With this in mind, the following scenario is put forward as a model of the development of the phylum:

- The earliest speakers of Sino-Tibetan were highly diverse foragers living in an arc between the slopes of the Himalayas and Assam/Arunachal Pradesh up to 10,000 years ago.
- Some spoke early Sino-Tibetan languages, others unknown languages now present only as substrates unless Kusunda is a relic of this period.
- Seasonal foragers exploited the high Tibetan Plateau from 7500 BP.
- By perhaps 6–5000 BP a ‘livestock revolution’ took place. Yak herders moved up and settled the Tibetan Plateau permanently. Pigs were domesticated in China among non-Sino-Tibetan speakers.
- Foragers in Northeast India probably began to practise vegeculture (taro, plantains) and arboriculture (sago) and animal management (mithun) by 6000 BP.
- By 5000 BP diverse early Sino-Tibetan groups spread eastwards to China. Sinitic is not a primary branch but simply one of many migratory groups.
- Proto-Tujia, proto-Bai and probably others met unknown populations (Hmong-Mienic? pre-Austronesians?) with domestic pigs, while also cultivating and beginning to domesticate rice.
- Proto-Sinitic speakers encountered early Altaic speakers with foxtail millet and other crops.
- The Sinitic languages expanded southwards, assimilating or encapsulating many small groups. They encountered Hmong-Mien speakers with rice and switch millet terminology to rice.

- Cold zone cereals (buckwheat, foxtail and Panicum millets) and perhaps also taro were moved from gathering to domestication in the montane areas on the fringes of the Himalayas.
- Rice moved up from India but also westwards from China (hence hybridised types) and overlays older cereals where ecologically possible.
- Ruminants (cows, sheep, goats) spread downwards into China from Central Asia 4400 BP (? Altaic for small ruminants but not cattle).
- Tibetic speakers underwent a major expansion (when?) assimilating linguistic diversity on the Plateau.
- Rice invaded the lowland vegetational zones rather later, pushing taro into residual systems.
- Groups such as early Burmic spread southwards, fragmenting Austroasiatic-speaking peoples.

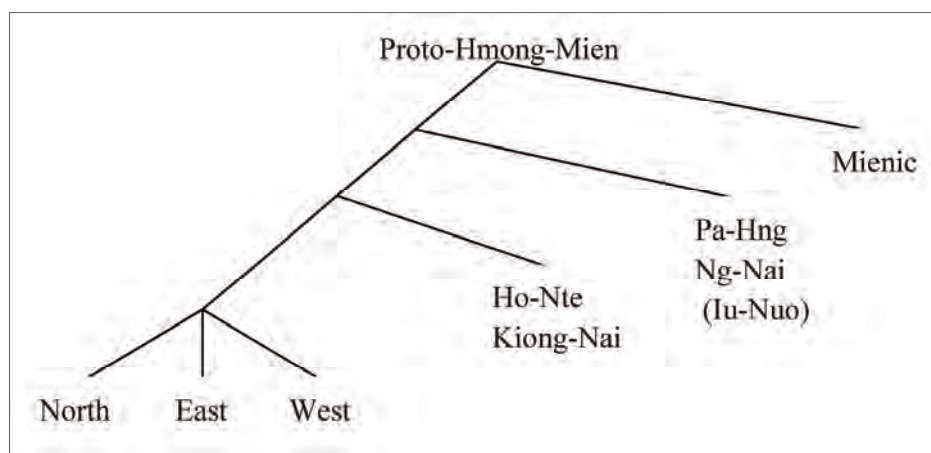
Figure 3 shows a highly simplified map of the early phases of these movements.



**Figure 3:** A possible model of early Sino-Tibetan expansion.

## 4 Hmong-Mien

The Hmong-Mien [=Miao-Yao] languages are spoken mostly in China with some groups also in Laos, Vietnam and Thailand (Niederer 1998, 2004; Ratliff). Their centre of gravity is between the Yangzi and the Mekong rivers and the extensions southwards may be as recent as the last few centuries. Hmong-Mien languages are quite closely related to one another, and although the Ethnologue lists some 32 languages, many of these are mutually intelligible lects. The linguistic geography of Hmong-Mien speakers suggests strongly that they were scattered by the incoming Han and probably forced southwards in the last 3–2000 years. Many agricultural terms can be reconstructed to proto-Hmong-Mien but most of them are either Chinese borrowings or resemble Chinese closely. It seems likely that pre-Hmong-Mien speakers have a long history in China, and that they can be linked with early Neolithic cultures, but Chinese cultural dominance has made this hard to detect. Figure 4 shows a ‘tree’ of Hmong-Mien languages in Niederer (2004).



**Figure 4:** Classification of the Hmong-Mien languages according to Niederer (2004).

The Hmong-Mien proto-language is likely to be older than is apparent from the lexical data. The reconstruction of agricultural vocabulary is clear, but so are borrowings from Old Chinese into the proto-language. Moreover, proto-Hmong-Mien has many fruit-crops and other plants typical of a drier climate which are not generally characteristic of the Southeast Asian region. Table 4 shows proposals for the Hmong-Mien subsistence lexicon adapted from Ratliff.

**Table 4:** Proposals for the Hmong-Mien subsistence lexicon (adapted from Ratliff)

Item	Reconstruction	Source
bean	*dup	< Chinese
buckwheat	*jæu	cf. Chinese
chicken	*Kəi	< Chinese
cucumber	*K <sup>w</sup> a	< Chinese
eggplant	*ja	cf. Chinese
pear	*rəy	< Chinese
plum	*hljəŋX	
rice, cooked	*hnrəŋH	

Item	Reconstruction	Source
rice, husked	*tuX	< Chinese
rice plant	*mbləu	
taro	*wouH	< Chinese
buffalo	*ŋiuŋ	< Chinese
dog	*qluwX	
duck	*ʔap	< Chinese
sheep/goat	*yuŋ	< Chinese

Although Ratliff assumes ‘buckwheat’ is a borrowing from Chinese, this is apparently historically unlikely (Laurent Sagart, personal communication) and it is probable it was adopted into Sinitic. Similarly, ‘eggplant’ resembles terms in many languages of the region and is probably not a borrowing from Sinitic. Whatever the final resolution of the various etymological debates, proto-Hmong-Mien as presently understood is too late to be identified with the earliest agricultural sites in the Yangzi and other regions of Central China. However, it is not unlikely that speakers of pre-*proto-Hmong-Mien* were present in this zone.

## 5 Austroasiatic

Austroasiatic languages remain the most poorly researched of all those in the region. Many are not documented at all and some recently reported in China are still not classified with certainty. Although there have been many proposals, there are no *proto-Austroasiatic* reconstructions with published justifications. The nearest approach to this is the ‘Mon-Khmer’ etymological dictionary of Shorto (2006), which identifies a large number of common roots attested across several branches. However, few of them are pan-Austroasiatic and may reflect regional, local or subgroup innovations. It is therefore as yet unclear whether, for example, we can draw inferences concerning the environment in the homeland of Austroasiatic, as has been claimed (Diffloth 2005, this volume). Sidwell and Blench (this volume) review some of the theories concerning the classification and antiquity of Austroasiatic and argue for a relatively recent dispersal along the Mekong basin, marked by incised and impressed pottery (Rispoli 2008; cf. White, this volume). These arguments will not be repeated here and this section will concentrate on the reconstruction of agricultural vocabulary. Van Driem (this volume) argues almost the exact opposite, for an early domestication of rice in Northeast India and an expansion from west to east. Regrettably, the failure to cite any concrete linguistic data supporting his argument makes it hard to evaluate.

Indirect evidence points to a relatively shallow time-depth for Austroasiatic, since a broad variety of agricultural terminology can be reconstructed to the *proto-language*. The most well-known crop is taro (*Colocasia esculenta*), for which a common root is attested almost everywhere (Table 5). Reflexes of #*trow?* occur throughout Austroasiatic, and Shorto (2006:475) reconstructs a form \**t<sub>2</sub>raw?* for his *proto-Mon-Khmer*. Table 5 shows the distribution of reflexes of #*trow?*.

Although the model presented by Sidwell and Blench for the structure and dispersal of Austroasiatic is quite different, it is consistent with the claim by Diffloth (2005) that Austroasiatic speakers typically spread along river valleys in the early period of their expansion seeking waterlogged soils suitable for taro. Rice terminology is widespread, and includes Muṇḍā, but is not attested in as many branches as taro and therefore should not be

treated as proto-Austroasiatic. Table 6 shows quasi-reconstructions of Austroasiatic names for crops and the numbers of individual branches for which reflexes are attested<sup>3</sup>.

This distribution suggests that almost all other crops were adopted subsequent to the dispersal of Austroasiatic and that after taro, hill-rice and foxtail millet were key crops. Paddy-rice was apparently quite late despite its dominance in agricultural systems today. Ferlus (1996) makes the intriguing suggestion that there has been a *glissement sémantique* [semantic slippage] between taro and rice terms, presumably via the concept of ‘staple crop’ (though see Diffloth, this volume, for a sceptical response). This underlines the relative antiquity of taro in Austroasiatic subsistence systems. If agriculture itself is ca. 4300 BP, the initial dispersal of proto-Austroasiatic would not be earlier than this. If this is the case, then Austroasiatic is unlikely to have an intricate nested structure, because the time would be insufficient for such a structure to develop.

**Table 5:** Reflexes of #*traw*?, ‘taro’ in Austroasiatic

Branch	Language	Attestation	Gloss in source
Palaungic	Riang	sro?	
Palaungic	Palaung	tōh	
Palaungic	Danaw	kāro <sup>1</sup>	
Palaungic	Proto-Wa	kro?	
Palaungic	Lamet	ruə?	
Palaungic	Khang	hō	
Monic	Mon	krao	
Monic	Nyah Kur	traw	
Vietic	Thavung	t <sup>h</sup> oo <sup>3</sup>	
Vietic	Vietnamese	sɔ	
Vietic	Proto-Vietic	*sro?	
Khmeric	Old Khmer	trav	
Khmeric	Khmer	tra:v	
Khmuic	Khmu	sro?	
Pearic	Chong	k <sup>h</sup> re: <sup>A</sup>	
Bahnaric	PSB	*təraw	
Bahnaric	East Bahnar	trəu	amaranth
Katuic	PK	*craw	
Katuic	Bru	ʔarəw	
Katuic	Kuy	ʔaarəaw	
Katuic	Sre	traw	
Katuic	Mlabri	kwaaj	
Katuic	Ong	raw	
Khasian	Khasi	shriew	arum
Muṇḍā	Sora	saro	<i>Caladium esculentum</i>
Muṇḍā	Mundari	saṛu	edible root
Muṇḍā	Santal	saru	

Apart from crops, the speakers of proto-Austroasiatic were also enthusiastic livestock producers. Table 7 shows that almost all the major species found in the region today were already known to speakers at an early period, except horse, donkey and sheep.

3 Full datasets are included in Blench (forthcoming)

The most surprising of these is the goat, which is poorly attested archaeologically but for which the linguistic evidence is very strong. It is also notable that aquatic-adapted poultry, such as geese and ducks, appear to be older than chickens.

This evidence is consistent with a relatively late date for the dispersal of Austroasiatic, which seems to have spread rapidly over a large region, as the ‘flat’ structure of the phylum suggests. Given the importance of taro and other aquatic terminology, it is reasonable to link this with the wide distribution of ‘incised and zone-impressed’ pottery ‘across parts of far southern China, northern Vietnam and Thailand after about 2500 BC’ (Bellwood 2005:132; Rispoli 2008; cf. White, this volume). Rice, millet and chickens would have been adopted midway through the expansion, and paddy rice would have come to replace taro as the principal subsistence crop relatively late.

**Table 6:** Crop reconstructions in Austroasiatic

Gloss	Reconstruction	Comment
rice (general)	#ḡa:ʔ	Found in seven branches
rice-grain	*sḡʔ:ʔ	Reconstructs only to Proto-Mon-Khmer
paddy rice	#srɔ	Found in three branches including Muṅḡā
husked rice	#rəkau	Found in seven branches including Muṅḡā
foxtail millet	#səŋkɔɔy	Found in seven branches
taro	#trawʔ	All branches except Aslian
sesame	#ləŋa	Found in six branches
banana	#tVIVy	Found in six branches
betel pepper	#mpluw	Found in six branches

**Table 7:** Livestock reconstructions in Austroasiatic

Gloss	Reconstruction	Comment
bovid	#ŋwV	Widespread but does not necessarily apply to domestic species
cow	#[rə]mɔɔk	Found in six branches including Muṅḡā
buffalo	#krəpaaw	Found in all branches excluding Muṅḡā
buffalo	#triik	Found in six branches with possible Muṅḡā cognate
pig	#kliik	Found in six branches
pig	#kruul	Found in six branches
goat	#bɛɛŋ	Found in ten branches with doubtful Muṅḡā cognate
dog	#atʃɔ:k	All branches
cat	#miaw	Found in eight branches
chicken	#syiar	Found in six branches
goose/duck	#ŋaŋ	Found in nine branches but referent varies
duck	#tʃaa[k]	Found in nine branches

## 6 Austronesian

One of the most persuasive narratives in recent prehistory has been that of the Austronesian expansion. Deriving from the original hypothesis of the kinship of over a thousand languages in Southeast Asia and the Pacific, it was first given its modern form by Dempwolff (1938). However, Dempwolff omitted to clearly identify and situate the



languages of the indigenous peoples of Taiwan, an omission rectified by the second major figure in Austronesian studies, Isidore Dyen (1963). Blust (1976, 1995, 1999) may have been the first author to clearly establish that the diversity of Formosan languages required them to be ancestral to all others and to constitute an array of primary branches. This hypothesis was adopted by Peter Bellwood (1984/5) to model the archaeological evidence, whence emerged a story about the ancestors of the Austronesians leaving Taiwan by means of highly developed sailing technology and reaching the furthest shores of the Pacific as well as the coast of East Africa. A Neolithic package was deemed to accompany these ocean navigators, consisting of pigs, dogs and chickens, rice, pottery and stone adzes, as well as distinctive types of jewellery, such as the nephrite *linglingo*. Various sub-narratives such as the ‘express-train to Polynesia’ (Diamond 2001) reached high-profile journals and the idea has acquired a certain currency in global prehistory. Blust’s hierarchy of nodes branching from the Austronesian tree up to Oceanic, the branch identified with the Lapita potters and ultimately giving rise to Polynesian, seemed to correlate with this early expansion.

The Austronesian migration has further developed into a more general narrative about demographic expansion in prehistory which has it that the spread of many of the world’s language phyla driven by agriculture (Bellwood 2005; Bellwood and Renfrew 2002). This model has always had its detractors (Meacham 1984/1985; Solheim 1984/1985; Oppenheimer 2004; Szabó and O’Connor 2004; Terrell 2004) but their striking failure to engage with the linguistic evidence has tended to undermine the substance of their arguments. Moreover, in some areas, notably Near Oceania and Polynesia, it would be hard to deny demographic expansion, since this was the colonisation of previously unoccupied territory. Nonetheless, in recent years there has been a rising chorus of discontent; archaeologists are increasingly claiming that the data doesn’t fit a simple demic expansion model. Linguists have been less vocal, but then the number of linguists interested in bigger picture of Austronesian is quite restricted. But with Denham (2004), Donohue and Denham (2010), Blench (2005, 2009, 2010), Lewis et al. (2008) and Bulbeck (2008) the chorus of discontent is too loud to be ignored. The claim, put simply, is that assemblages seem to be rather complex and not to correspond to a simple model of incoming Neolithic farmers replacing foragers. Moreover, some of the key elements in the proposed Austronesian subsistence package are simply not turning up in excavations. The patterns of material culture in prehistory seem to point to earlier and more complex inter-island interactions than the Austronesian expansion model would imply.

One of the key building blocks of the Austronesian expansion hypothesis has been the apparent reconstructibility of key economic terms, both for domestic animals and crops. If the findings (or absences) in the archaeological record are to be taken at face value, then there are problems with these reconstructions. There is a dichotomy between animals and plants, since vegiculture systems could have carried domesticated species across the Austronesian world prior to the expansion from Taiwan, whereas this cannot be the case for animals unless they are attested in the archaeozoological record. There is a specific point concerning pigs, dogs and chickens<sup>4</sup>. It has been shown that the majority of pigs in island Southeast Asia originate not from Taiwan, but from the mainland, probably Vietnam (Hongo et al. 2002). Moreover, they are conspicuously absent from the archaeological record in the main islands until significantly later than the Austronesian expansion (Dobney et al. 2008). There is a small pocket of domestic pig in assemblages in Taiwan,

---

4 Thanks to Phil Piper for discussions on this point.

and the extreme northern Philippines, but this does not spread southwards into the main body of the archipelago<sup>5</sup>. The situation for dogs and chickens is if anything more perplexing; they do not seem to turn up in assemblages at all, until identified much later in Polynesia.

This is in marked contrast to the apparent evidence from linguistics. Blust (2002) conveniently summarises the linguistic evidence for faunal terms in Austronesian. He proposes:

**Table 8:** Domestic animal reconstructions in AN

species	level	proto-form
chicken	PMP	*manuk
cock	PMP	*laluj
dog	PAN	*asu/wasu
puppy	PAN	*titu
domestic pig	PAN	*beRek
? wild pig	PAN	*babuy

In the case of chickens and dogs this sharply contradicts the archaeological evidence; no chickens and dogs have yet been found at this period. Linguistically, they cannot be apical forms which gradually diversify through the Austronesian world, but represent either semantic shifts or widespread loanwords. If this is the case, how do we explain the reconstructions? In the case of chicken, *\*manuk* alternates with reduplicated forms which mean ‘bird’ and indeed in the putative branch of Austronesian represented by Tai-Kadai this is what it does mean (Benedict 1942; Ostapirat 2005). Blust assumes that ‘bird’ is the secondary meaning, but more likely this was the original meaning and it has shifted to chicken with the subsequent spread of the animal. The exact evidence for *\*laluj* is lacking but it occurs in the Northern Philippines and in the Lesser Sundas and is presumably either a widespread loanword or a semantic shift. The case of words for ‘dog’ is more perplexing, since this is well-attested in Taiwanese languages and widely in island Southeast Asia. However, surprisingly it has no reconstruction in proto-Oceanic. In addition, the same root occurs virtually all across Austroasiatic. Yet dogs are again conspicuous by their absence in the archaeology of island Southeast Asia. Dogs were probably domesticated from the Asiatic wolf and appear to be found in early sites in China (Savolainen et al. 2002). Could all those occurrences of the *\*asu* root represent a semantic shift? It seems unlikely. A useful clue is found in the fact that Tai-Kadai languages, despite their evident reflection of PMP in terms of basic numerals, do not have the *\*asu* term for dog. Indeed, it appears that the proto-form in Tai-Kadai (something like *\*hma*) is likely to be a borrowing from Hmong-Mien (see Table 10 below). If so, then this term may have been absent in PMP and all those occurrences of *#asu* are in fact loanwords, reflecting contact with Austroasiatic speakers (and the subsequent spread of the term once borrowed).

The case of the pig is still more perplexing. According to Blust (2002:93), reflexes of *\*beRek* occur in Puyuma, Tsouic, the Northern Philippines and some Borneo languages with the meaning ‘domestic pig’. This is realised in Oceanic as *\*boRok*, a general term for ‘pig’. As with ‘dog’ there is a widespread term in Austroasiatic, *\*C-liik* or *\*C-lek*, which is apparently cognate with Austronesian. The fricatives in Formosan languages may well be

5 There has apparently been an independent domestication of a highly local race on Lanyu (Orchid island) which may account for these finds (Keith Dobney, personal communication).

cognate with forms such as Pear *sru:k*. In this case, the Taiwanese domestic pigs probably came from the mainland of East Asia and the same source also donated the pig to Austroasiatic, hence these terms are related. However, Austronesian has another well-attested form for ‘pig’ *\*babuy*. This term often applies to wild pigs, but Blust (2002:93) finds the meaning of ‘domestic pig’ also widespread. As neither species is subsequently attested in the archaeological record, we have to assume the term devolved to the wild pig and was sporadically shifted back to the domestic pig.

Taro terminology is another complex story. There are two main complexes of terms for taro in Austronesian, *\*talef* and *\*ma*. In Southeast Asia there is a widespread term, *#traw?* which has reflexes throughout Austroasiatic (Table 5) but also appears to be cognate with Austronesian. Although Dempwolff (1938:128–9) reconstructed *\*talə(s)* for proto-Austronesian his evidence did not include either Taiwan or any languages near Formosa. This suggests that Austroasiatic speakers were the original domesticators of taro and that Austronesian speakers borrowed it during an early phase of contact, with the southern Philippines/Borneo the most likely zone for such contact. Since this cannot have been through contact with Negrito hunter-gatherers, Austroasiatic speakers may previously have been resident in insular Southeast Asia. Taro and other vegetable culture had spread east from the mainland, and the expanding Austronesian speakers adopted it from the Austroasiatic speakers whom they subsequently assimilated, but not before borrowing their term for the plant. Ross et al. (2008:266) point out that reflexes of this root are rather scattered in Western Oceanic and that they are possibly borrowings from Eastern Oceanic, where the term is widespread.

The existing paradigm of Austronesian migration is crumbling in the face of a conspicuous absence of archaeological evidence for some of its central claims (for example Donohue and Denham 2010; Blench in press d). Its replacement will be a far more nuanced account of the movement of plants, animals and other types of material culture in the Southeast Asian region. The notion propounded by Bulbeck (2008) of ‘fisher-foragers’ and the emphasis placed by Solheim (1984/5) on trade may well be significant components in any new model. However, these views do not account for the extreme pervasiveness of the Austronesian languages, which must have replaced and assimilated a complex of different language families in numerous different sites in ISEA. This suggests that in addition to sailing technology and trade, the Austronesians must also have had a hugely attractive social, organisational and perhaps religious ideology which persuaded the residents of individual islands to adopt Austronesian culture. This would certainly explain the extraordinary diffusion of certain iconographic elements, such as the *bulul* figure, the *linglingo* and others, noted by art historians but not adopted by archaeologists. The expansion of Chamic on the Vietnamese mainland and the displacement or assimilation of Austroasiatic languages and cultures shows just how effective this type of cultural colonisation can be.

No consensus on a new paradigm for Austronesian is likely to be reached for some time. However, as with Sino-Tibetan, it is possible to outline a speculative model that at least can account for the interdisciplinary data. With all the usual caveats, the following hypotheses are put forward:

- a) The Austronesian phylum must have a structure similar to that outlined by Blust, which has Taiwanese languages as an array of primary branches and Malayo-Polynesian constituting the remainder.

- b) The series of nodes separating Oceanic from Proto-Malayo-Polynesian (PMP) that characterise earlier proposals must be dropped for lack of evidence. Western PMP looks like an array of primary branches, somewhat like Formosan.
- c) The many Austronesian languages in islands with *in situ* residents must have spread largely by adoption rather than demographic spread.
- d) However, this process was by wholesale language switch which would account for the limited evidence for non-Austronesian (NAN) substrates. In Melanesia, these processes broke down, hence both the phenotypic switch of Austronesians to Papuan physiognomy and the occurrence of a range of mixed languages. Traces of 'Papuan' structures can be detected in many Austronesian languages without this being evidence for prior settlement of uninhabited islands.
- e) The primary means of cultural conversion was religious and socio-political, rather than through military conquest or commerce.
- f) The material culture package supposed to be typical of Austronesians was in fact constructed from elements adapted along the way. Only when the Austronesians reach Polynesia and the Solomons do they propagate this package as an apparently coherent ensemble.
- g) As an additional consequence, many of the reconstructions for subsistence-related terms in Austronesian hitherto considered solid must instead be mosaics of ancient loanwords, spreading either east from MSEA or west from Melanesia.

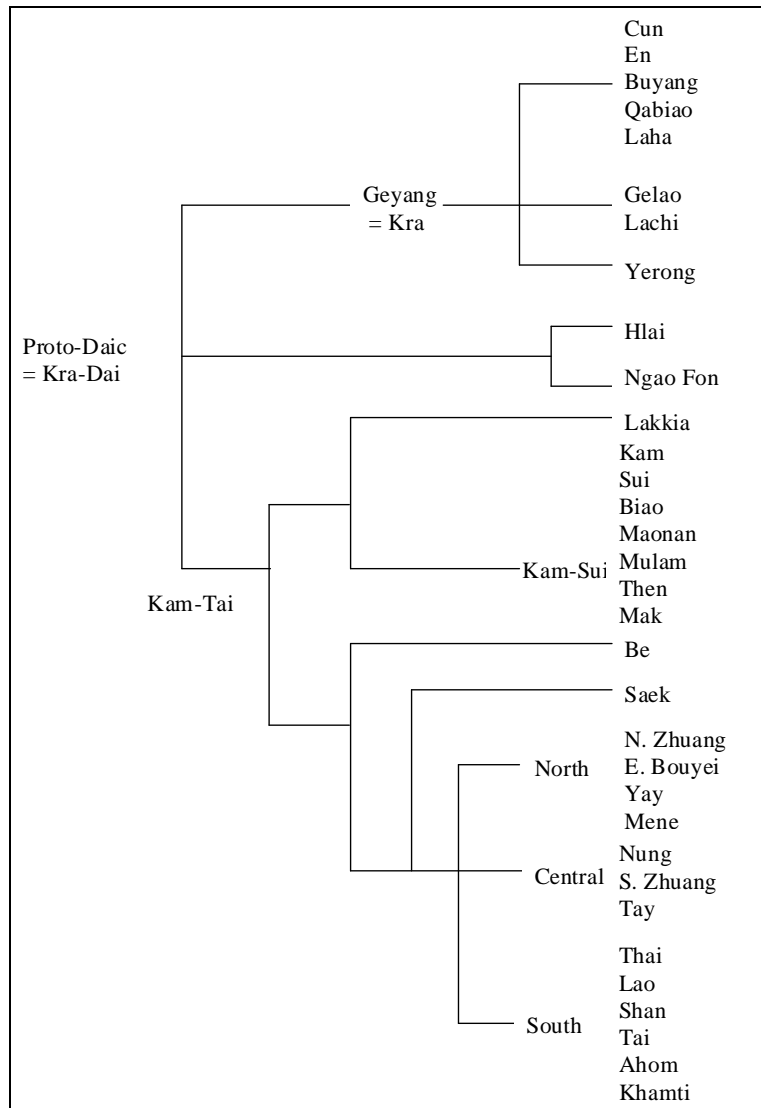
The Bellwood migrationist model has been enshrined in the prefaces of too many dissertations to be easily dislodged. But the mismatch with archaeology has now become too blatant to ignore. Restructuring Austronesian to take account of both its linguistic dominance and problematic presence in the archaeological record will become a major task for the immediate future.

## 7 Daic [=Tai-Kadai]

The Daic or Tai-Kadai languages, of which Thai is the most well-known and widespread representative, are spoken from southern Thailand into Laos, Cambodia, Vietnam and China. Up-to-date maps of their distribution are given in Edmondson and Solnit (1997a) who estimate the number of speakers of these languages as at least 80 million. Overviews of the phylum are given in Edmondson and Solnit (1988, 1997b; and Diller et al, 2008). The Daic languages represent a highly coherent grouping whose structure is well understood; a relatively long list of common glosses make it possible to construct hypotheses concerning the subsistence and migrations of the proto-Daic speakers. All the most diverse Daic languages are in China: despite the marked southward extension of Thai today, the likely origin of Daic is in Guizhou (貴州). Despite this, the Daic expansion has no obvious archaeological correlate, although there are clearly cultural links with Austronesian speakers of Taiwan, for example dental ablation (Blench in press, b).

Surprisingly, however, there is no standard reconstruction of proto-Daic, although branches such as Kra, Tai and Hlai have lists of proto-forms (Li 1977; Hudak 2008; Ostapirat 2000; Norquest 2007). Figure 5 shows the internal relationships of Daic given by

Edmondson and Solnit (1997b) amended with reference to Thongkum (2001) and Ethnologue (2009).



**Figure 5:** Daic subclassification.

The external affiliations of Daic have remained controversial, sharing as it does many features with surrounding language phyla, notably Austroasiatic, Hmong-Mien and Sino-Tibetan. These were used by Benedict (1942, 1975) to erect ‘Austro-Tai’, a macrophylum that would unite Austroasiatic, Hmong-Mien, Daic and Austronesian. Proposals for a linguistic connection between Tai and Austronesian date back at least to Schlegel (1901) and are extended in Wulff (1942). Baoya (1995) reviewed the rather extensive literature in Chinese (some of which argues for a genetic connection with Sinitic) and compiled a rather striking dataset comparing Daic with Malay. A failure to establish regular sound-correspondences meant that his observations have been passed over by later writers. Thurgood (1994) claimed that much of the evidence for hypotheses that link together the major language phyla of MSEA, such as Benedict’s Austro-Tai, derive from ancient loanwords rather than genuine cognacy. However, Ostapirat (2005) set out a series of regular correspondences linking Daic with Austronesian, assuming a simple model of a primary split between Daic and Austronesian; in this model, the two would then be co-

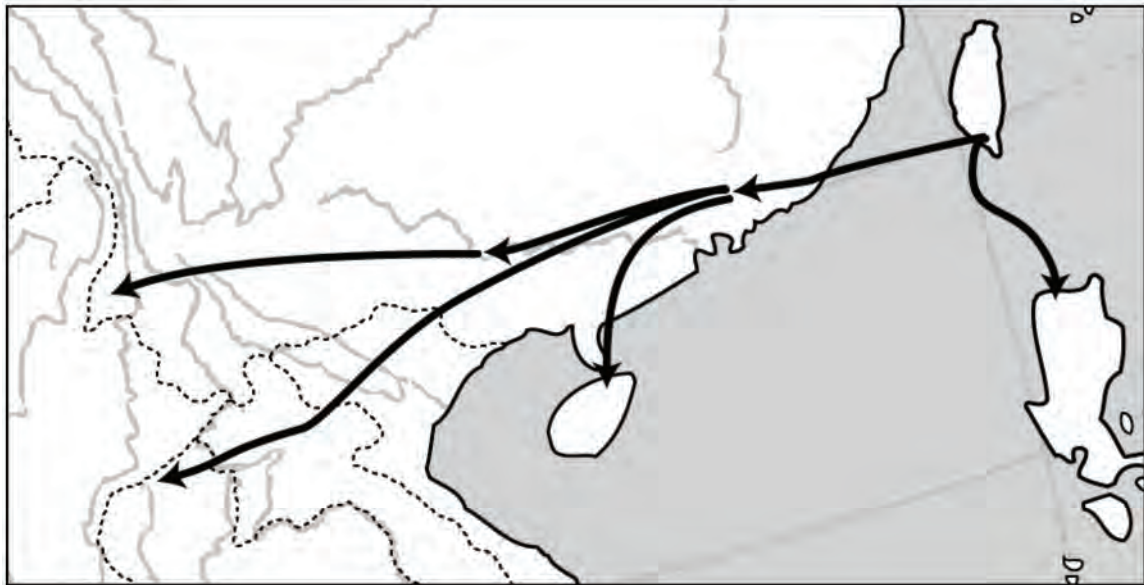
ordinate branches. But this seems unlikely; Daic looks more like a branch of proto-Malayo-Polynesian and does not share in the phonological complexities of Formosan. Sagart (2004, 2005b) proposes that proto-Daic speakers migrated back to the mainland, to Guangdong and the region of Hainan island.

Norquest (2007:413) points out that the Hlai branch shares some striking lexical items with proto-Austronesian which do not occur in the other branches of Daic. A sample of the most convincing comparisons is shown in Table 9.

**Table 9:** Shared lexicon between Austronesian and proto-Hlai<sup>#</sup>

Gloss	Pre-HI	PHI	PAn
slap	*pi:k	*phi:k	*pik
rub rope~weave	*bən	*p <sup>h</sup> ən	*bəl+bəl
pinch	*ti:p	*t <sup>h</sup> i:p	*a-tip (PMP)
seven	*tu:	*t <sup>h</sup> u:	*pitu
three	*tʉ:ʔ	*tʃ <sup>h</sup> u:ʔ	*təru
sharp	*jə:m	*tə <sup>h</sup> ə:m	*tʃəjəm
five	*ma:	*hma:	*rima
six	*nəm	*hnəm	*ʔənəm
maternal grandmother	*na:ʔ	*hna:ʔ	*ina ‘mother’s sister’
that	*C-na:	*C-na:	*i-naʔ
bury	*lɔmfɨ	*hlɔmfɨ	*tələm
fish scale	*C-lə:p	*C-lə:p	*quʂələp
eight	*ru:	*ru:	*waru
sell	*ri:wʔ	*ri:wʔ	*sariw

<sup>#</sup> Source: adapted from Norquest (2007:413)



**Figure 6:** Hypothetical routes of Daic expansion.

This demonstrates neatly that typical Austronesian morphology was retained by Daic *after* the arrival of speakers back on the mainland and that the reduced forms now typical of most Daic languages are a later development. The pattern of morphosyntactic reduction

is identical for the cognates with Kra pointed out by Ostapirat, namely the deletion or assimilation of the first syllable of the Austronesian form in Daic. None of these lexical items are specifically Formosan; they can just as well be PMP, which is certainly the case for Kra-Austronesian cognates identified by Ostapirat. The retention of these forms, in particular the numerals, is a striking testimony to the early diversification of Daic. Hlai must have branched off at the same time as the Kra languages, retaining a specific set of Austronesian lexical items. Daic then became relexified and radically restructured following contact with Hmong-Mien and perhaps other languages of unknown affiliation. Such a migration would be around 4000 BP, in broad conformity with current dates for the first incursions in the Northern Philippines. Figure 6 shows the hypothetical routes of Daic expansion based on this evidence.

Daic languages are not all that diverse and both crops and domestic animals can be reconstructed for proto-Daic. Ostapirat (2000) presents some glosses that appear to be shared across all three branches, including ‘pig’ and ‘dog’ and at least some crops. Table 10 shows items extracted from Ostapirat relating to crops and domestic animals attested across all branches of Daic.

**Table 10:** Daic lexicon illustrative of subsistence\*

Language	chicken	pig	dog	sesame	‘yam’
Gelao	qai	map	mpau	ŋklau	mbø
Lachi	kɛ	mye	m	—	mfiə
Laha	kəi	məu	maa	—	mal
Paha	qai	muu	maa	ŋaa	man
Buyang	ʔai	muu	—	ŋaa	man
Biao	qai	ṃuu	ṃaa	ŋfiua	mfiən
Hlai	khai	pou	pou	keu	man
Sui	qaai	ṃuu	ṃaa	ʔŋaa	man
Tai	kai	muu	maa	ŋaa	man

\* Source: Ostapirat (2000)

**Table 11:** Proposals for the Daic subsistence lexicon

Item	Quasi-reconstruction	Possible source
taro	#pɣaak	< Taiwan names for <i>Alocasia macrorrhizos</i>
cooked rice	#mpVŋ	widespread mV- roots
husked rice	#saan	Sino-Tibetan and Austroasiatic
white rice	#rɔp	Daic innovation
millet	#pfeɛŋ	Daic innovation
ginger	#k <sup>hi</sup> ŋ	< Proto-Hmong-Mien
buffalo	#kwaay	< Austroasiatic
goose	#ɣaan	< Austroasiatic

The exact referent of ‘yam’ is unclear; synchronically, this term can be applied to potato, sweet potato and taro. However, there is no doubt that the true yam, *Dioscorea* (probably *esculenta*), was formerly extensively grown throughout this region and has declined in recent times. Daic #mpaw for ‘dog’ is likely borrowed from Hmong-Mien \*hmaŋ<sup>C</sup> ‘wild dog’ (Ratliff) as it resembles neither Austroasiatic nor Austronesian.

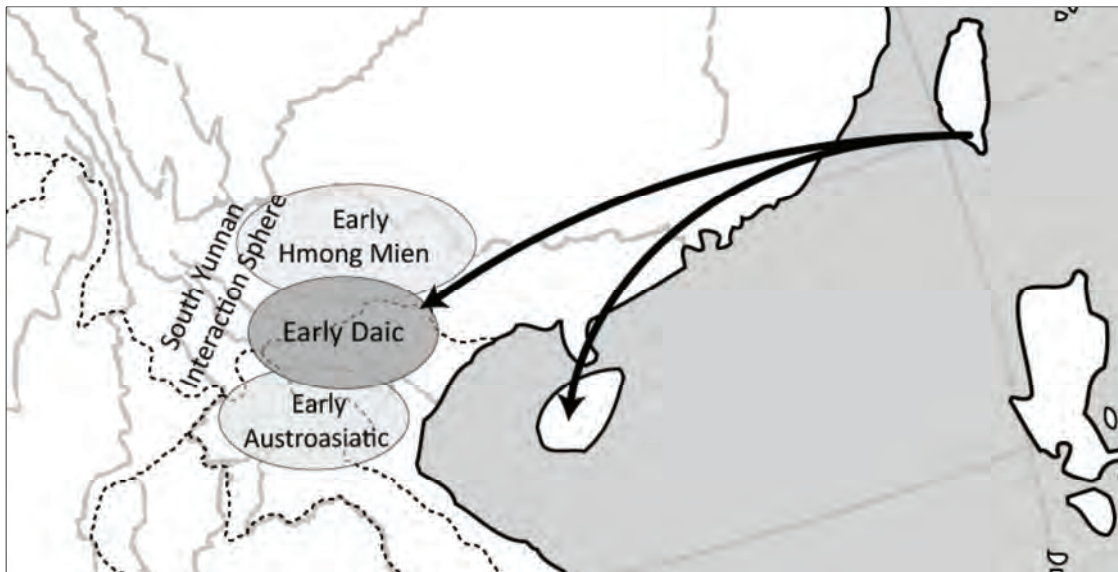
Table 11 presents some additional proposals for the Daic subsistence lexicon based on a more extensive compilation in Blench (in press b).

The argument for deriving the Daic name of taro from a Formosan term for *Alocasia macrorrhizos* is given in more detail in Blench (in press c).

The sheer variety of crops in this inventory strongly suggests that the proto-Daic speakers were established farmers. However, none of these terms except possibly sesame show links with Austronesian farming terminology. This was previously rather perplexing for the argument that Daic was a branch of Austronesian, but if it is the case that the Austronesians were marginal farmers emphasising fisheries and trade, the situation is more explicable. As Table 11 suggests, Daic does borrow terms from other regional phyla. Daic languages apparently underwent a lexical revolution with respect to agriculture on the mainland as a result of interaction with resident language phyla.

## 8 Language phyla and the antiquity of farming

Archaeological coverage of the region is highly skewed, and the quality and density of data from China has a tendency to bias interpretation. Nonetheless, it is apparent that farming began in the colder, drier zones north of MSEA proper. If the argument for the genesis of Sino-Tibetan is accepted, then its earliest phase was foragers in the foothills of the Himalayas, some exploiting sago and lowland fauna, others gathering wild cereals and montane animal species. An agricultural revolution took place ca. 6500 BP, stimulating a wave of expansion eastward into China. Agriculture then spread only slowly further south, presumably because tropical MSEA was such an abundant environment there may have been no need to farm.



**Figure 7:** The South Yunnan Interaction Sphere.

However, by around 4300 BP there was a striking confluence of four quite distinct language phyla in what is now Yunnan and an almost simultaneous adoption of agriculture. This may well have been a result of the expansion of pre-Sinitic Sino-Tibetan speakers such as the Bai. Hmong-Mien, Austroasiatic and Daic speakers had all begun farming at the period when their respective proto-languages began to expand. This early period of intense interaction is provisionally named the South Yunnan Interaction Sphere



(SYIS) and a tentative graphic representation is shown in Figure 7. It cannot be emphasised too strongly that this remains a speculative hypothesis; only more detailed archaeology and linguistics will establish its credibility. It is suggested that the common roots in Sino-Tibetan reflecting agriculture also date from this era, although this must remain controversial due to a lack of properly presented data.

None of this implies that a Bellwood demic expansion model is necessarily appropriate, but it is improbable that these language phyla expanded significantly before the inception of agriculture. In other words, phyla cannot be significantly older than farming unless we reach for a model of ‘extinct branches’, subgroups in the foraging era which have now conveniently disappeared. People can have crops but expand for a variety of reasons; as is suggested in Sidwell and Blench (this volume), improved water transport stimulates dispersal. However, agriculture implies settlement and provide the preconditions for a variety of other important social changes, including technological advance, long-distance trade and more coherent military organisation.

There is a persuasive and pervasive stereotype that the languages of MSEA must be long established and that millennia of intensive interaction are responsible for the widely observed similarities between different language phyla as well as the remarkable interpenetration of fundamental vocabulary. However, a constellation of evidence from different disciplines suggests that this interpretation may be erroneous (Blench forthcoming). In fact it may be that the expansion of the major phyla is associated with the Neolithic and that the observed convergence can occur quite rapidly under specific conditions.

Much remains to be done to add weight to this revised scheme. We urgently await more credible reconstructions for Sino-Tibetan based on attestations from all the diverse branches. An evidence-based approach to the internal classification of Austroasiatic is essential. Direct material remains reflecting the transition to farming, based on actual archaeobotanical and archaeozoological materials would make claims for the dates of the inception of agriculture more plausible. Nonetheless, the potential to correlate different disciplines in creating an integrated prehistory of MSEA is clearly now within reach.

## **Acknowledgements**

The origin of this chapter was first given as a paper at the Workshop ‘Dynamics of Human Diversity in Mainland Southeast Asia’ held at EFEO, Siem Reap, 6-10<sup>th</sup> January 2009 and I am grateful to the organisers, Nick Enfield and Joyce White, for the invitation to attend and present this. I benefited from discussions at the meeting but I would also like to thank Gérard Diffloth, Mark Post, Paul Sidwell and Laurent Sagart and the anonymous referees for additional comments. Thank you to Peter Johan Nijland for preparation of Figures 3, 6, and 7. Doug Cooper in Bangkok is doing remarkable work in making available source materials in electronic form and the data elements have benefited substantially from access to these databases. The present revision incorporates data from papers presented at subsequent conferences in 2009, including Austroasiatic (Bangkok), Sino-Tibetan (Chiang Mai) and the Indo-Pacific Prehistory Association (Hanoi). The argument concerning the origins of Sino-Tibetan was presented at the 16<sup>th</sup> Himalayan Languages Symposium, London September 2010 and the re-interpretation of Austronesian prehistory at the 13<sup>th</sup> EuraSEAA Meeting in Berlin in October 2010.

## References

- Abbi, Anvita. 2006. *Endangered Languages of the Andaman islands*. Munich: Lincom Europa.
- Baoya, Chen. 1995. On the original relationship between Chinese and Kam-Tai. *Linguistics of the Tibeto-Burman Area*, 18(1). 149–171.
- Bellwood, Peter. 1984/85. A hypothesis for Austronesian origins. *Asian Perspectives* 20. 107–117.
- 1997. [2nd ed.] *Prehistory of the Indo-Malaysian archipelago*. Honolulu: University of Hawa'i Press.
- 2005. *First Farmers: The Origins of Agricultural Societies*. Oxford: Blackwell.
- Bellwood, Peter and Colin Renfrew (eds). 2002. *Examining the Farming/ Language Dispersal Hypothesis*. (McDonald Institute Monographs.) Cambridge: McDonald Institute for Archaeological Research.
- Benedict, Paul King. 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *American Anthropologist* 44. 576–601.
- 1972. *Sino-Tibetan: a conspectus* (“STC”). Contributing editor, James A. Matisoff. Princeton-Cambridge Series in Chinese Linguistics #2. New York: Cambridge University Press.
- 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven: Human Relations Areas Files Press.
- 1976. Sino-Tibetan: another look. *Journal of the American Oriental Society* 96(2). 167–197.
- Blench, Roger M. 2005. From the mountains to the valleys: understanding ethnolinguistic geography in SE Asia. In Laurent Sagart, Roger M. Blench and Alicia Sanchez-Mazas (eds) *The peopling of East Asia*, 31–50. London: Routledge.
- 2007. The language of the Shom Pen: a language isolate in the Nicobar Islands. *Mother Tongue XII*, 179–202.
- 2008. Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? In Alicia Sanchez-Mazas, Roger Blench, Malcom D. Ross, Iliia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 105–132. London: Routledge.
- 2009. Remapping the Austronesian expansion. In Bethwyn Evans (ed.) *Discovering history through language. Papers in honour of Malcolm Ross*, 35–59. Canberra: Pacific Linguistics.
- 2010. New evidence for the Austronesian impact on the East African coast. In Atholl Anderson (ed.) *Global origins and the development of seafaring*, 239–248. Cambridge: Macdonald Institute.
- forthcoming. *Reconstructing the Southeast Asian past: archaeology, language, genetics and material culture*. Book ms. under consideration for publication.
- in press a. Is Shom Pen a distinct branch of Austroasiatic? Paper given at the IVth ICAAL, Bangkok, November 2009 and in press for a special edition of Mon-Khmer Studies.
- in press b. *The expansion of the Daic peoples: cultural and iconographic correlations*. Presentation at the XIVth EURASEAA meeting. Leiden, September 2008.
- in press c. Vernacular names for taro in the Indo-Pacific region and their possible implications for centres of diversification. Paper for the special session on taro, IPPA, Hanoi. Special Issue of David Addison and Matthew Spriggs (eds) *Bulletin of the Indo-Pacific Prehistory Association*.

- in press d. Austronesians or Austronesianisation? Paper presented at the 13th International Conference of the European Association of Southeast Asian Archaeologists, Berlin 27th September–1st October 2010 and now submitted for proceedings.
- Blust, Robert A. 1976. Austronesian culture history: some linguistic inferences and their relations to the archaeological record. *World Archaeology* 8(1). 19–43.
- 1995. The Prehistory of the Austronesian-speaking peoples: A view from language. *Journal of World History* 9(4). 453–510.
- 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-Kuei Li (eds) *Selected papers from Eighth International Conference on Austronesian linguistics*, 31–94. Taipei: Academica Sinica.
- 2002. The History of Faunal Terms in Austronesian Languages. *Oceanic linguistics* 41(1). 89–139.
- Bodman, Nicholas Cleaveland. 1980. Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship. In Frans van Coetsem and Linda R. Waugh (eds) *Contributions to Historical Linguistics: Issues and Materials*, 34–199. Leiden: E.J. Brill.
- Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.) *Tibeto-Burman languages of the Himalayas*, 1–72. Pacific Linguistics A-86. Canberra: Australian National University.
- Bulbeck, David. 2008. An integrated perspective on the Austronesian diaspora: the switch from cereal agriculture to maritime foraging in the colonisation of island SE Asia. *Australian Archaeology* 67. 31–52.
- Dempwolff, Otto. 1938. *Vergleichende Lautlehre des austronesischen Wortschatzes, vol. 3: Austronesisches Wörterverzeichnis*. Zeitschrift für Eingeborenen-Sprachen, Supplement 19. Berlin: Reimer.
- Denham, Tim. 2004. The roots of agriculture and arboriculture in New Guinea: looking beyond Austronesian expansion, Neolithic packages and indigenous origins. *World Archaeology* 36. 610–20.
- Diamond, Jared. 2001. Polynesian origins: Slow boat to Melanesia? *Nature* 410. 167.
- Diffloth, Gérard. 2005. The contribution of linguistic palaeontology and Austroasiatic. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 77–80. London: Routledge Curzon.
- this volume, Austroasiatic word histories: boat, husked rice and taro.
- Diller, Anthony V. N., Jerold A. Edmondson and Luo Yongxian (eds) 2008. *The Tai-Kadai Languages*. London: Routledge.
- Dobney, Keith, Thomas Cucchi and Greger Larson. 2008. The pigs of Island Southeast Asia and the Pacific: New evidence for taxonomic status and human-mediated dispersal. *Asian Perspectives* 47. 59–74.
- Donohue, Mark and Tim Denham. 2010. Farming and Language in Island Southeast Asia: Reframing Austronesian History. *Current Anthropology* 51(2). 223–256.
- Dondrup, Rinchin. 1990. *Bugun language guide*. Itanagar: Arunachal Pradesh State Government Directorate of Research.
- van Driem, George L. 1997. Sino-Bodic. *Bulletin of the School of Oriental and Southeast Asian Studies* 60(3). 455–488.
- 2005. Tibeto-Burman vs. Indo-Chinese: implications for population geneticists, archaeologists and prehistorians. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas

- (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 81–106. London: Routledge Curzon.
- 2008. To which language family does Chinese belong, or what's in a name. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*, 219–253. London and New York: Routledge.
- this volume, Rice and the Austroasiatic and Hmong-Mien homelands.
- Dyen, Isidore. 1963. The position of the Malayo-Polynesian languages of Formosa. *Asian Perspectives* 7. 261–71.
- Edmondson, Jerold A. and David B. Solnit. 1997b. Introduction. In *Comparative Kadai: the Tai branch*, 1–32. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- (eds). 1988. *Comparative Kadai: linguistic studies beyond Tai*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- (eds). 1997a. *Comparative Kadai: the Tai branch*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Enfield, N. J. 2003. *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia*. London: RoutledgeCurzon.
- this volume, Linguistic diversity in mainland Southeast Asia.
- Ethnologue. 2009. *A guide to the world's languages*. Dallas: SIL.
- Ferlus, Michel. 1996. Du taro au riz en Asie du Sud-Est, petite histoire d'un glissement sémantique. *Mon-Khmer Studies* 25. 39–49.
- Fuller, Dorian, Lin Qin and Emma Harvey. 2008. Evidence for the late onset of agriculture in the Lower Yangtze region and challenges for an archaeobotany of rice. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 40–83. London: Routledge.
- Hajek, John. 1998. Kenaboi: an extinct unclassified language of the Malay Peninsula. *Mon-Khmer Studies* 28. 137–162.
- Higham, Charles F. W. 1996. Archaeology and linguistics in Southeast Asia: implications of the Austric hypothesis. In: *Bulletin of the Indo-Pacific Prehistory Association* 14. 110–118.
- 2002. *Early cultures of mainland Southeast Asia*. Bangkok: River Books.
- 2004. Mainland Southeast Asia from the Neolithic to the Iron age. In Ian Glover and Peter Bellwood (eds) *Southeast Asia: from prehistory to history*, 41–67. Abingdon: RoutledgeCurzon.
- Higham, Charles. F. W. and Thomas Higham. 2009. Dating Southeast Asian Prehistory: New Radiometric Evidence from Ban Non Wat. *Antiquity* 83. 125–144.
- Hongo, Hitomi, Naotaka Ishiguro, Takuma Watanobe, Nobuo Shigehara, Tomoko Anezaki, Vu T. Long, Dang V. Binh, Nguyen T. Tien and Nguyen H. Nam. 2002. Variation in mitochondrial DNA of Vietnamese pigs: relationships with Asian domestic pigs and Ryukyu wild boars. *Zoological Science* 19. 1329–1335.
- Hudak, Thomas J. 2008. *William J. Gedney's comparative Tai source book*. Honolulu: University of Hawai'i Press.
- Lai Van Toi. 1999. The initial inhabitants of Co Loa [in Vietnamese]. *Khao Co Hoc* 3. 39–54.

- Lewis, Helen, Victor Paz, Myra Lara, Huw Barton, Philip Piper, Janine Ochoa, Timothy Vitales, Jane A. Carlos, Tom Higham, Lee Neri, Vito Hernandez, Janelle Stevenson, Emil C. Robles, Andrea Rraggio, Rojo Padilla, Wilhelm Solheim II, and Wilfredo Ronquillo. 2008. Terminal Pleistocene to mid-Holocene occupation and an early cremation burial at Ille Cave, Palawan, Philippines. *Antiquity* 82. 318–335.
- Li, Fang-Kuei. 1977. *A Handbook of Comparative Tai*. Oceanic Linguistics Special Publication #15. Honolulu: Univ. Press of Hawaii.
- Matisoff, James A. 2003. *Handbook of proto-Tibeto-Burman*. Berkeley: University of California Press.
- 2008. *The Tibeto-Burman Reproductive System: Toward an Etymological Thesaurus*. Berkeley: University of California Press.
- Matsumura, Hirofumi, Marc F. Oxenham, Nguyen Kim Thuy, Nguyen Lan Cuong and Nguyen Kim Dung, this volume, Population history of Mainland Southeast Asia: The Two Layer model in the context of Northern Vietnam.
- Meacham, William. 1984/1985. On the improbability of Austronesian origins in South China. *Asian perspectives* 26(1). 89–106.
- Niederer, Barbara. 1998. *Les langues Hmong–Mjen (Miáo–Yáo)—Phonologie Historique*. Lincom Studies in Asian Linguistics 07. München: Lincom Europa.
- 2004. Pa-hng and the classification of the Hmong-Mien languages. In Nicholas Tapp, Jean Michaud, Christian Culas and Gary Yai Lee (eds) *Hmong/Miao in Asia*, 129–146. Bangkok: Silkworm Books.
- Norquest, Peter K. 2007. *A Phonological Reconstruction of Proto-Hlai*. Ph.D. Department Of Anthropology, University of Arizona.
- Oppenheimer, Stephen. 2004. The ‘Express Train from Taiwan to Polynesia’: on the congruence of proxy lines of evidence. *World Archaeology* 36(4). 591–600.
- Ostapirat, Weera. 2000. Proto-Kra. *Linguistics of the Tibeto-Burman Area* 23(1).
- 2005. Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 107–131. London: Routledge Curzon.
- Ratliff, Martha. 2010. *Hmong-Mien language history*. Pacific Linguistics. Canberra.
- Reid, Lawrence A. 2005. The current status of Austric. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. 132–60. London: Routledge Curzon.
- 1994a. Possible non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics* 33. 37–72.
- 1994b. Unravelling the linguistic histories of Philippine Negritos. In Tom Dutton and Darrell T. Tryon (eds) *Language contact and change in the Austronesian world*, 443–475. Berlin: Mouton de Gruyter.
- Rispoli, Fiorella. 2008. The incised and impressed pottery of Mainland Southeast Asia: following the paths of Neolithization. *East and West* 57. 235–304.
- Ross, Malcolm, Andrew K. Pawley and Meredith Osmond. 2008. *The lexicon of Proto-Oceanic: The culture and environment of ancestral Oceanic society, Part 3: Plants*. Canberra: Pacific Linguistics, Australian National University.

- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444.
- 2005a. Sino-Tibetan-Austronesian: An updated and improved argument. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 161–176. London: Routledge Curzon.
- 2005b. Tai-Kadai as a subgroup of Austronesian. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 177–181. London: Routledge Curzon.
- 2008. The expansion of Setaria farmers in East Asia: a linguistic and archaeological model. In Alicia Sanchez-Mazas, Roger M. Blench, Malcom D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 133–158. London: Routledge.
- Savolainen, Peter, Ya-ping Zhang, Jing Luo, Joakim Lundeberg and Thomas Leitner. 2002. Genetic Evidence for an East Asian Origin of Domestic Dogs. *Science* 298. 1610–1613.
- Schlegel, Gustav. 1901. Review of Frankfurter's Siamese grammar. *T'oung Pao* 2. 76–87.
- Shafer, Robert. 1947. Hruso. *Bulletin of the School of Oriental and African Studies* 12(1). 184–196.
- Shafer, Robert. 1966/1967. *Introduction to Sino-Tibetan*. (Part I: 1966, Part II: 1967) Wiesbaden: Otto Harrassowitz.
- Shorto, Harry L. 2006. *A Mon-Khmer comparative dictionary*. Paul Sidwell, Doug Cooper and Christian Bauer (eds). PL 579. Canberra: ANU.
- Sidwell, Paul and Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.
- Simon, Ivan M. 1970. *Aka language guide*. Itanagar: Government of Arunachal Pradesh.
- Solheim, II Wilhelm G. 1984/1985, The Nusantao hypothesis: the origin and spread of Austronesian speakers, *Asian perspectives* 26(1). 77–88.
- Starosta, Stanley. 2005. Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 182–197. London: Routledge Curzon.
- Starostin, Sergej A. 1991. On the hypothesis of a genetic connection between the Sino-Tibetan languages and the Yenisseian and North Caucasian languages. In Vitaly Sheveroshkin (ed.) *Dene-Sino-Caucasian*, 12–42. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Szabó, Katherine and Sue O'Connor. 2004. Migration and complexity in Holocene Island southeast Asia. *World Archaeology* 36. 621–628.
- Tayeng, Aduk S. 1990. *Sulung language guide*. Guwahati: Government of Arunachal Pradesh.
- Terrell, John Edward. 2004. The “Sleeping Giant” Hypothesis and New Guinea's Place in the Prehistory of Greater Near Oceania. *World Archaeology* 36(4). 601–609.
- Thongkum, Theraphan L. 2001. The Tai-Kadai peoples of Hainan island and their languages. In Kalaya M.R. Tingsabach and Arthur S. Abramso (eds) *Essays in Tai linguistics*, 189–203. Bangkok: Chulalongkorn University Press.
- Thurgood, Graham. 1994. Tai-Kadai and Austronesian: the nature of the relationship. *Oceanic Linguistics* 33(2). 345–368.
- Thurgood, Graham and Randy J. LaPolla (eds). 2003. *The Sino-Tibetan languages*. Routledge Language Family Series. London and New York: Routledge.

- Vinh, Bui. 1991. The Da But culture in the Stone Age of Vietnam. *Bulletin of the Indo-Pacific Prehistory Association* 10. 1271–131.
- Weidert, Alfons. 1987. *Tibeto-Burman Tonology: A comparative analysis*. Amsterdam and Philadelphia: John Benjamins.
- White, Joyce C. 1995. Modeling the development of early rice agriculture: ethnoecological perspectives from northeast Thailand. *Asian perspectives* 34(1). 37–68.
- White, Joyce C., Daniel Penny, Lisa Kealhofer and Bernard Maloney. 2004. Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. *Quaternary International* 113. 111–132.
- Wolfenden, Stuart N. 1927. *Outlines of Tibeto-Burman Linguistic Morphology*. London: The Royal Asiatic Society.
- Wulff, Kurt. 1942. Über das Verhältnis des malayo-polynesischen zum indochinesischen. *Det Küniglige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 27(2).

# 8 *Population history of mainland Southeast Asia: the two layer model in the context of Northern Vietnam*

---

HIROFUMI MATSUMURA, MARC F. OXENHAM,  
NGUYEN KIM THUY, NGUYEN LAN CUONG AND  
NGUYEN KIM DUNG

## 1 Introduction

Since the early twentieth century a considerable number of prehistoric human remains have been uncovered in Southeast Asia, causing researchers to challenge former models of micro-evolutionary change and population movement within and between East and Southeast Asia. The population history of Southeast Asia is somewhat complex due to a number of migratory processes and associated inter-population genetic exchanges throughout prehistory. Southeast Asia is thought to have been initially occupied (the first layer) by anatomically modern humans at least 40,000 years ago (sometimes referred to as Australo-Melanesians) that later exchanged genetic material with later immigrants (the second layer) from North and/or East Asia leading to the formation of present day Southeast Asians (Callenfels 1936; Mijsberg 1940; Von Koenigswald 1952; Coon 1962; Jacob 1967, 1975; Bellwood 1987, 1989, 1996, 1997, 2005; Brace et al. 1991). This population model of the peopling of Southeast Asia is often referred to as the ‘Two Layer’ hypothesis.

Currently the Two Layer model has been validated by both archaeological and historical linguistic research. A great deal of work in these fields has illustrated links between the dispersal of language families, represented by Austroasiatic and Austronesian for instance, and the spread of agricultural technology within Southeast Asia. Both linguistic and archaeological research suggests southern China, with northern Mainland Southeast Asia, and Taiwan as the ultimate sources of these language and population dispersals around 5,000 years ago (Blust 1977; 1984–85, 1996a, b; Bellwood 1987, 1996, 1997, 2005; Renfrew 1987, 1989, 1992; Bellwood et al. 1992; Tryon 1995; Glover & Higham 1996; Higham 1998, 2001; Bellwood & Renfrew 2003; Diamond & Bellwood 2003; Sagart 2008).

The Two Layer model is, however, not without its critics. For instance, Turner’s ‘Sundadont/Sinodont’ dental study (Turner 1989, 1990, 1992) has resulted in a very different interpretation of Asian population history, sometimes referred to as the ‘Local Evolution’ model. Turner’s analysis classified an array of nonmetric dental traits into one

---

N. J. Enfield, editor.

*Dynamics of Human Diversity, 153-178.* Pacific Linguistics, 2011

© This edition vested with Pacific Linguistics.



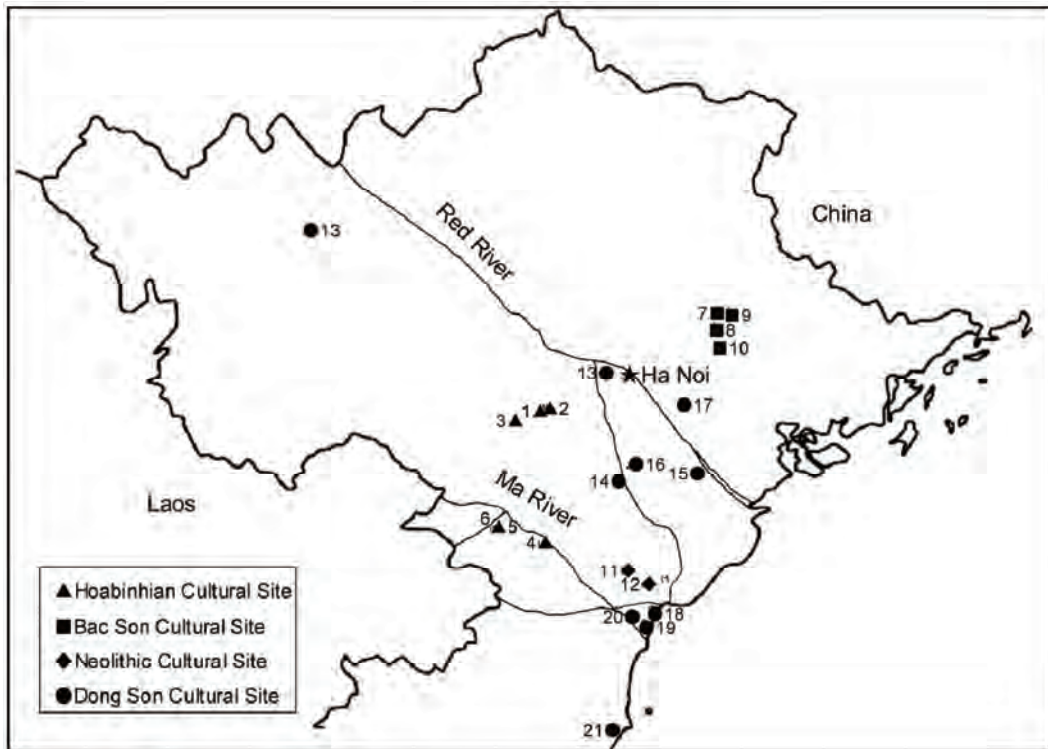
grouping termed Sundadont, characterised by Southeast Asians for the most part, and another group called Sinodont, including Northeast Asians. Turner argues that the Sundadont are the product of long-standing genetic continuity, uninterrupted by any significant admixture with Sinodonts from the north. Multivariate cranial metric analyses by Hanihara (1992, 1993a, b, c, 1994, 2006) and Pietrusewsky (1992, 1994, 1996, 1999a, b, 2005, 2006, 2008, 2010) have also suggested close affinities between early and modern Southeast Asians, coupled with a distinct dissimilarity to Australo-Melanesians, leading them to support a scenario of local evolution in Southeast Asian (including southern China) population history. This hypothesis has recently received some support from genetic research in which Negrito and non-Negrito populations of Southeast and East Asia are historically united via a single primary wave of entry to the region, however, no timeline is provided for this (HUGO Pan-Asian SNP Consortium, 2009). Furthermore, this particular study also suggests that Southeast Asia was a major geographic source of East Asian populations which, to date, is a model that has received little support from prehistorians.

Alternative models explaining the population history of Southeast Asia generally revolve around two main questions. The first is whether the early occupants of Southeast Asia have an Australo-Melanesian affinity. This question is complicated by the nature of the archaeological record. Late Pleistocene and early Holocene human remains are limited in terms of sample size, completeness, and preservation and are often plagued by dating uncertainties, thus preventing a more comprehensive assessment of the biological affinity of these pre-ceramic period peoples. The second question concerns the timing, source and scale of the dispersal of food-producing people often linked with language family expansion from southern China or Taiwan, and whether they mixed with or replaced local, extant populations. This issue is of crucial importance with respect to the validity, or otherwise, of the Two Layer model as a means of understanding the population history of Southeast Asia.

The central aim of this chapter is to address these two major issues by way of examining human skeletal sequences spanning the prehistoric and early historic phases of northern Vietnam. The earliest anatomically modern settlers of this region are often referred to as the Sonvians or Hoabinhians. The Hoabinhian culture-complex is somewhat analogous to what in Europe would be called the Mesolithic and in Southeast Asia is identified as a pebble-tool complex that had a wide distribution over mainland Southeast Asia during the late Pleistocene and early Holocene (Tan 1980). In this region Vietnam is regarded as a centre of Hoabinhian culture, largely represented by numerous cave sites, that on occasion have yielded human remains, excavated and surveyed intensively from the early 20<sup>th</sup> century (for example, see Colani 1927a). Our recent work in northern Vietnam has continued the discovery of well preserved human skeletal remains, providing data crucial for the assessment of the biological affinities of early settlers of mainland Southeast Asia with later populations in the region, that is, Neolithic and early Metal Period communities.

Below, we describe centrally important human remains including those excavated by the authors in northern Vietnam, and discuss how these remains inform debate over the competing Two Layer and Local Evolution models ostensibly characterising the population history of Southeast Asia. Initially, we summarise the archaeological contexts of those sites with relevant skeletal assemblages. Following this, the results of quantitative comparisons of prehistoric and modern samples from East/Southeast Asia and the Southwest Pacific are carried out in order to test the validity of each model as an

explanatory construct. Important site localities and illustrations of representative crania can be seen in Figures 1 and 2.



**Figure 1:** Locality map of the archaeological sites in northern Vietnam, from which well preserved human remains were uncovered.

1 = Lang Gao	8 = Lang Cuom	15 = Doi Son
2 = Hang Cho	9 = Cua Gi	16 = Minh Duc
3 = Hang Muoi	10 = Dong Thuoc	17 = Dong Xa
4 = Lang Bon	11 = Man Bac	18 = Quy Chu
5 = Mai Da Nuoc	12 = Con Co Ngua	19 = Nui Nap
6 = Mai Da Dieu	13 = Vinh Quang	20 = Thieu Duong
7 = Pho Binh Gia	14 = Chau Son	21 = Dong Mom

## 2 Prehistoric Human Remains in Vietnam

### 2.1 Hoabinhian Sequence

The earliest accepted period of anatomically modern human occupation in northern Vietnam, the Son Vi Culture, dates back approximately 30,000 years BP, but these remains are restricted to a few sets of teeth only (Nguyen L.C. 2007). While the Son Vi continues in some areas until some 11,000 years BP, a new culture period, characterised for the most part by lithic artifact morphology, emerged around 18,000 years BP, the Hoabinhian. Of the 100-odd Hoabinhian sites known (Tan 1980), approximately a third have furnished skeletal remains (Nguyen L.C. 1994). Communities attributable to the Hoabinhian, which minimally endured until at least 8000 years BP (although as late as 3000 years BP in Peninsular Malaysia), were involved in hunting and gathering a wide variety of mammals, birds, fish, and shellfish (Higham 1989). The overwhelming majority of material culture consists of unifacial lithic artifacts, the most diagnostic being the oval/almond-shaped

unifacial pebble tools named sumatraliths. The relationship between post-Hoabinhian communities, which are principally coastally oriented, and the Hoabinhian itself is unclear, as the latter is virtually always restricted to inland caves and rock shelters.

A review of early archaeological work on late Pleistocene through early Holocene sites in the region that have yielded human skeletal material is given in Tayles & Oxenham (2006). Here sites immediately relevant to this study are reviewed. Despite the majority of Hoabinhian sites being limestone caves and rock shelters conducive to good preservation of organic materials, only a limited number have produced well preserved skeletal remains. The earliest discoveries include remains excavated by Colani (1927b, 1939) in Lang Gao cave, Hoa Binh Province and Lang Bon, Thanh Hoa Province. The former site produced a substantial number of specimens, approximately 20 individuals, although only two sets were associated with well preserved crania. A further complicating issue is that both sites, while assigned to the Hoabinhian by way of lithic artifact association, lack absolute dates and it is unclear if the skeletal material is earlier or later in the cultural sequence.

One well dated specimen from the earlier phase of the Hoabinhian period was excavated by the Hanoi National University at the cave site of Hang Muoi during work carried out between 1963 and 1965. A fragment of clavicle has a calibrated AMS date of approximately 14,000 BP (Bulbeck et al. 2007), although the associated cranial remains are represented by a fragmentary maxilla and calvaria only.

Two late Hoabinhian rockshelter sites in Thanh Hoa Province, Mai Da Nuoc and Mai Da Dieu, have produced two substantially complete skulls during a 1984 excavation by the Vietnamese Institute of Archaeology, Hanoi (Nguyen L.C. 1986). Lithic artifact analysis suggests Mai Da Nuoc is contemporaneous with Mai Da Dieu. Further, absolute dating using charcoal samples from the burial layer of Mai Da Dieu indicates the human remains date to approximately 8000 years BP.

A further important site, Hang Cho cave, Hoa Binh Province, was originally excavated by Colani in 1926 and then 1932 (no report known). In 1977 the Vietnamese Institute of Archaeology recovered a substantial number of lithics identified as Hoabinhian (Tan 1997; Nguyen K.T. & Doi 1998). More recently, the Hanoi National University and the Seoul National University have published calibrated dates, ranging from  $14,100 \pm 300$  years BP to  $9,710 \pm 50$  years BP, on freshwater shell deposits found in the main chamber (Yi et al. 2004). In 2004 we excavated a well preserved flexed burial, with cranium, AMS dated (using the teeth) to approximately 10,500 years BP (calibrated) (Matsumura et al. 2008a).

## 2.2 Neolithic Sequence

The term 'Neolithic' has been traditionally used to characterise communities with agriculture but without metal (Bellwood, 1992:94). The presence or absence of pottery is not necessarily useful in identifying Neolithic populations; there are good examples of pottery manufacture in hunter-gathering initial and early Jomon, Japan (Habu 2004) and a lack of pottery in the earliest clearly agricultural Neolithic contexts in the Levant (Lev-Yadun et al. 2000). In northern Vietnam at least, pottery appears among Mid-Holocene hunter-gatherer communities (for example Da But) well before any evidence for agriculture. The term 'Neolithic' is used in this chapter to denote pre-metal communities that show clear evidence for agricultural subsistence economies, all of which happen to use pottery.

Overlapping with the terminal Hoabinhian and prior to the appearance of pottery-using hunter-gather communities, is a phase spanning the period c. 9000 to 7000 years BP, termed the Bac Son culture. This culture-period is characterised by polished stone tools,

including edge ground axes, in addition to classic pebble tools characteristic of the Hoabinhian (Nishimura 2006). Bac Son was first described by the French geologist Mansuy when excavating the cave site of Tham Khoach, Lang Son Province. Subsequently, Mansuy and colleagues Colani, Huard and Saurin, undertook a more extensive survey of Bac Son sites in the adjacent area, resulting in the recovery of a considerable number of human remains from Lang Cuom, Pho Binh Gia, Khac Kiem, and Keo Phay (Mansuy & Colani 1925; Huard & Saurin 1938). Among these specimens, several cranial sets remain in very good condition.

The pottery-using hunter-gather period is ushered in with the appearance of the Da But culture. Da But sites are for the most part shell middens that include polished stone artifacts and coarse cord-marked pottery. It has been suggested that the remains of domesticated animals occur in some of these sites, including *Bibos* sp., *Bubalis bubalis* (buffalo) and *Sus scrofa* (pig) (Vinh 1991), although these interpretations have not been corroborated to date. What is clear, however, is a complete lack of floral evidence for any form of agriculture during this period (Higham 1996). Da But communities are foragers, perhaps with a degree of sedentism, with a marked marine/aquatic resource dependency. In terms of human remains, several sets of individuals were uncovered from the Da But type site by Patte (1965), one of which was associated with a nearly complete cranium. The largest Da But skeletal assemblage derives from the cemetery site of Con Co Ngua, Than Hoa province, which is believed to belong to the initial phase of the Da But period c. 6000–5000 BP (Nguyen K.T. 1990; Vinh 1991). Over 100 individuals have been recovered from Con Co Ngua and have been extensively studied both morphologically (Nguyen K.T. 1990; Nguyen L.C. 2003) and bioarchaeologically (Oxenham et al. 2005; Oxenham 2006; Oxenham et al. 2006).

The Neolithic in northern Vietnam is represented by a wealth of archaeological cultures and sites, the most significant in terms of human skeletal remains being the Ha Long and Phung Nguyen. The Ha Long culture, c. 4500 to 3000 years BP, flourished along the coast of Halong Bay in northeastern Vietnam. One of the representative Ha Long sites, Hai Co Tien, has furnished numerous human skeletal remains, although only teeth and jaws are well preserved. Phung Nguyen sites, dated to between 4000–3500 years BP, generally show evidence for food production and are mainly distributed around the fringes of the Red River delta. Phung Nguyen sites are typically characterised by high quality lithic ornaments (including jade) and elaborately decorated pottery. The livelihood of these communities would have been similar to those in surrounding areas and included pig and dog domestication, land clearing, ceramic manufacture, hunting, marine resource gathering and trade (Hiep & Phung 2004; Huffer 2005).

The largest Phung Nguyen period cemetery site excavated to date is Man Bac, Ninh Binh Province, on the southern edge of the Red River delta. Initial excavations by the Institute of Archaeology in Hanoi and Ninh Binh Provincial Museum in 1999 uncovered a substantial amount of lithic and ceramic materials along with several extended burials. Subsequent work in 2001 revealed more of the same in addition to 10 very well preserved sets of skeletal remains (Nguyen L.C. 2001; Phung 2001; Hiep & Phung 2004). The Sapporo Medical University and Australian National University collaborated with Vietnamese researchers for two further excavation seasons at Man Bac in 2004–5 and 2007 revealing a further 82 well preserved burials (see Matsumura et al. 2008b; Oxenham et al. 2011) (Figure 3). Moreover, the rich range of material culture includes a variety of lithics such as axes, adzes, chisels, grinding stones, hammer stones, blades and various types of semi-precious stone ornaments. Ceramics include jars with rounded bases, ring-footed jars, pedestalled dishes, pediform clay support stands, net sinkers, mushroom shaped ceramic

anvils and clay bow pellets. Man Bac is the earliest site in Vietnam to allow an examination of social systems and personhood in any detail. Our work has revealed evidence for an age-based social hierarchy (Oxenham et al. 2008a), sophisticated system of palliative care (Oxenham et al. 2009) and the expression of social identity, including that of children, through complex mortuary practices (Oxenham et al. 2008b). Charcoal above the burial level has been dated to between 1868–1524 BCE ( $2\sigma$  cal., Khao Co Hoc 2007) which is consistent with known dates of the Phung Nguyen, although Man Bac, as a coastal marine oriented site, lacks evidence so far for food production. Domestic ceramic vessel forms and motifs have been interpreted as indicating that the inhabitants of Man Bac maintained connections with surrounding coastal cultures, such as the Ha Long and Hoa Loc groups. Work on the faunal remains indicates a subsistence base rich in both terrestrial and marine resources, as well as evidence for the domestication of pigs (Sawada et al. unpublished).

### **2.3 Early Metal Age sequence**

The close of the Phung Nguyen phase saw the development of later Bronze Age cultures such as the Dong Dau culture, 3500–3000 years BP, which was characterised by a sophisticated bronze-working technology and the appearance of spears and arrowheads, in addition to an increase in the appearance of utilitarian items such as fishing hooks, chisels, and axes. The Vietnamese Bronze Age reaches its peak with the emergence of the Dong Son culture, flourishing in northern Vietnam from 2500 years BP until Han domination and colonising efforts beginning after 111 BCE. Higham (1989:30) described the Vietnamese Dong Son as one of several Southeast Asian examples demonstrating social differentiation and a move ‘from village autonomy towards centralised chiefdoms’. Apart from the diverse range and technological sophistication of material culture, for example, richly decorated massive bronze drums, there is evidence for marked craft specialisation, a complex ritual life, the development of an aristocratic and centralised elite, maritime trade, and sophisticated military skills and equipment. Regarding the Metal period skeletal sample, it is an aggregation of more than 100 individuals from several archaeological sites (including Vinh Quang, Chau Son, Doi Son, Quy Chu, Nui Nap, Minh Duc, Dong Mom, Dong Xa) in northern Vietnam spanning the period from 3000–1700 years BP (Nguyen K.T. 1993; Nguyen L.C. 1996; Oxenham 2000; Oxenham et al. 2005).

## **3 Discontinuity in the Neolithic Population?**

In order to test, and perhaps refine, the Two Layer hypothesis within the regional context of northern Vietnam, it is necessary to examine in some detail the biological relationships, if they in fact exist, between the temporally disparate skeletal sequences discussed above for the pre-Neolithic, Neolithic and Metal Period assemblages available for study. In terms of a qualitative assessment of cranio-morphology (see Figure 2), the pre-Neolithic Hoabinhian and Bac Son specimens tend to exhibit dolichocephalic calvaria, large zygomatic bones, a remarkably prominent glabella and superciliary arches, a concave nasal root and a low and wide face with prominent prognathism. On the other hand, the majority of Metal Period individuals share an array of distinctive cranial characteristics including relatively narrow faces, low glabella, superciliary arches and nasal roots and round orbits. The apparent marked cranio-morphological discontinuity between pre-Neolithic and Metal Period populations suggests that the Neolithic was the turning point in terms of the micro-evolutionary history of northern Vietnam.

As the only large and well-preserved skeletal sample representative of the Neolithic, the Man Bac skeletal sequence is pivotally important in understanding what happened between the pre-Neolithic and later Metal Periods in terms of the profound changes in the biology of the people that has clearly occurred. The Man Bac assemblage is key to testing the Two layer model, for northern Vietnam at least. Are members of the Man Bac community direct biological descendents of earlier Da But and/or Bac Son people, or perhaps even the earlier original Hoabhinian settlers of the region? What is the biological relationship between Man Bac and later Metal Period skeletal series? How does Man Bac sit with respect to other early and modern populations from East and Southeast Asia and the West Pacific? All of these questions are addressed in the context of testing the Two Layer model, which seeks to explain the origins of modern Southeast Asians.



**Figure 2:** Representative human skulls from archeological sites in northern Vietnam Hoabhinian series: Lang Bon, Mai Da Nuoc, Mai Da Dieu, Hang Cho, Bac Son series: Pho Binh Gia, Late Neolithic series: Man Bac, Dong Son series: Minh Duc, all specimens are of males except the female Hang Cho.

#### 4 Materials and Methods

Of the 99 burials excavated from Man Bac, 39 are adult with the remaining being subadult. Of the adult assemblage, 21 male skulls are utilized for metric analysis. Cranial measurements followed Martin's measurement definitions and protocols (see Bräuer 1988) and Yamaguchi's (1973) assessment of facial flatness measurements. Subsequent craniometric comparisons are made using Q-mode correlation coefficients (Sneath & Sokal 1973). The comparative samples are listed in Tables 1 and 2, which also includes a summary of comparative archaeological samples from Laos, Thailand, Malaysia, China and Japan, as well as modern samples from East/Southeast Asia and the Pacific. Inter-sample phenotypic affinities are illustrated using Neighbor-Net Split Tree analysis (Huson

& Bryant 2006) applied to the distance matrix of Q-mode correlation coefficients. This procedure was carried out using the software package Splits Tree Version 4.0 provided by Huson & Bryant (2006).



**Figure 3:** View of the Man Bac site and exposed inhumation burials.

Table 1: Comparative archaeological human remains from East/Southeast Asia

Sample	Locality	Period	Remark	Cranial metrics	N*	Facial flatness	N*	Storage <sup>+</sup>
1	Hoabinhian Vietnam	Late Pleistocene-Early Holocene (Hoabinhian Culture)	Sites of Mai Da Nuoc, Mai Da Dieu, Lang Gao, Lang Bon	H.M.	5	H.M.	5	IAH MHO
2	Bac Son Vietnam	Early Holocene (Bac Son Culture c. 8000–7000 BP)	Sites of Pho Binh Gia, Lang Cuom, Cua Gi, Dong Thuoc	H.M.	7	H.M.	7	MHO
3	Da But Hoa Prov., Nth Vietnam	Middle Neolithic (Da But Culture, sample dated to c.5000 BP)	Sites of Con Co Ngua and Da But	Nguyen L.C. 2003	19	H.M.	7	IAH MHO
4	Man Bac Site in Ninh Binh Prov., Nth Vietnam	Late Neolithic (c.3500– 3000 BP)	-	H.M.	21	H.M.	17	IAH
5	Dong Son Northern Vietnam	Early Metal Age (Dong Son Culture, 3000–1700 BP)	Sites of Vinh Quang, Chau Son, Doi Son, Quy Chu, Thieu Duong, Nui Nap, Dong Mom, Minh Duc, Dong Xa	Nguyen L.C. 1996 and H.M.	32	H.M.	20	IAH CSPH
6	Tam Hang Sites in Laos	Early or Middle Holocene	Mansuy & Colani 1925; Huard & Saurin 1938, including single specimen from Ban Dang site	H.M.	4	H.M.	4	MHO

\* N=sample size

+ Storage=institutions of materials studied by H.M. (H. Matsumura) AST=Academia Sinica of the Republic of China in Taipei; BMNH=Department of Paleontology, Natural History Museum, London; CSPH=Center for South East Asian Prehistory, Hanoi; IAH=Department of Anthropology, Institute of Archaeology, Hanoi; LAM=Long An Museum, Vietnam; MHO=Laboratoire d'Anthropologie Biologique, Musée de l'Homme, Paris; NMP=Department of Archaeology, National Museum of the Philippines, Manila; NTW=National Taiwan University, College of Medicine, SAC=Princess Maha Chakri Sirindhorn Anthropology Centre, Bangkok; UCB=Department of Biological Anthropology, University of Cambridge; UHW=Department of Anthropology, University of Hawaii.



**Table 1 continued:** Comparative archaeological human remains from East/Southeast Asia

Sample	Locality	Period	Remark	Cranial metrics	N*	Facial flatness	N*	Storage <sup>+</sup>
7	Gua Cha Site in Kelantan Prov., Malaysia	Hoabinhian (c.8000–6000 BP)	Sieveking (1954) specimen No. H12	H.M.	1	H.M.	1	UCB
8	Ban Chiang Site in Udon Thani Prov., Thailand	Neolithic-Bronze Age (c. 3500–1800 BP)	Gorman & Charoenwongsa 1976	Pietrusewsky & Douglas 2002 (M51#; Hamihara 1993a)	27	H.M.	15	UHW SAC
9	Weidun Site in Jiangsu Prov., Sth China	Neolithic (Majjiabang Culture, c.7000–5000 BP)	Nakahashi & Li 2002	Nakahashi et al. 2002.	29	Nakahashi et al. 2002.	29	-
10	Anyang Site in Henang Prov., China	Bronze–Iron Age (c. 3300 BP)	IHIA & CASS 1982	Han & Qi 1985	42	H.M.	26	AST
11	Jiangnan Site in Jiangnan Prov., Sth China	Zhou–Western Han (2770–1992 BP)	Nakahashi & Li 2002	Nakahashi et al. 2002	18	Nakahashi et al. 2002	18	-
12	Jomon Japan	Neolithic (Middle-Latest Jomon Culture, c. 5000–2300 BP)	Yamaguchi 1982	Hanihara 1993a	113	Hanihara 2000	36	-

\* N=sample size

# M=Martin's cranial measurement number

+ Storage=institutions of materials studied by H.M. (H. Matsumura) AST=Academia Sinica of the Republic of China in Taipei; BMNH=Department of Paleontology, Natural History Museum, London; CSPH=Center for South East Asian Prehistory, Hanoi; IAH=Department of Anthropology, Institute of Archaeology, Hanoi; LAM=Long An Museum, Vietnam; MHO=Laboratoire d'Anthropologie Biologique, Musée de l'Homme, Paris; NMP=Department of Archaeology, National Museum of the Philippines, Manila; NTW=National Taiwan University, College of Medicine, SAC=Princess Maha Chakri Sirindhorn Anthropology Centre, Bangkok; UCB=Department of Biological Anthropology, University of Cambridge; UHW=Department of Anthropology, University of Hawaii.

Table 2: Comparative modern population samples from East/Southeast Asia and West Pacific

Sample	Cranial metrics	N*	Facial flatness	N*	Remark (M=Martin's number)	Storage <sup>+</sup>
1 Vietnam	H.M.	27	H.M.	25	-	MHO
2 Cambodia	H.M.	12	H.M.	12	-	MHO
3 Laos	Nguyen L.C. 1996	17	H.M.	10	-	MHO
4 Thai	Sangvichien 1971	85	Hanihara 2000	30	-	
5 Myanmar	Pietrusewsky 1981	16 (21)	Hanihara 2000	102	cranial metrics M17,45,48,51 by H.M.	BMNH
6 Java	Pietrusewsky 1981	73 (20)	Hanihara 2000	62	cranial metrics M17,45,48,51 by H.M.	BMNH
8 South Moluccas	Pietrusewsky 1981	78 (8)	Hanihara 2000	28	cranial metrics M17,45,48,51 by H.M.	BMNH
9 Sumatra	Pietrusewsky 1981	43 (8)	Hanihara 2000	13	cranial metrics M17,45,48,51 by H.M.	BMNH
10 Dayak	Yokoh 1940	12	Hanihara 2000	72	Borneo	
11 Philippines	Suzuki et al. 1993	8	H.M.	8	-	NMP
12 Atayal	Pietrusewsky & Chang 2003	36 (41)	H.M.	30	cranial metrics M45,48,51,55 by H.M.	NTW
13 Bunun	Pietrusewsky & Chang 2003	26 (23)	H.M.	16	cranial metrics M45,48,51,55 by H.M.	NTW
14 Hainan	Howells 1989	45 (24)	H.M.	24	cranial metrics M48,51,55 by H.M.	NTW
15 North China #1	Hanihara 1993a	71	Hanihara 2000	56	Manchuria Prov., China	-
16 North China #2	Hanihara 1993a	26	Hanihara 2000	56	Kirin Prov., China	-
17 Japan	Hanihara 1993a	140	Hanihara 2000	138	-	
18 Australia	Hanihara 1993a	53	H.M.	21	-	BMNH
19 Melanesia	Hanihara 1993a	18	Hanihara 2000 (New Guinea)	156	Fiji, Tongans; New Hebrides; New Guinea	-
20 Loyalty	H.M.	17	H.M.	18	-	MHO

\* N=sample size; numbers in parentheses are sample sizes of cranial materials measured by Matsumura (H.M.) in remarks

+ Storage=institutions of materials studied by H.M. (abbreviations: see footnotes in Table 1)

## 5 Results

### 5.1 Phenotypic Variability Within the Man Bac Sample

The majority of Man Bac individuals are characterised by a relatively narrow and flat face with round orbits, whereas some specimens present with quite different features, such as a dolichocephalic cranium with prominent glabella and a low and wide face (see Figure 2). The apparent phenotypic variability in the Man Bac series suggests the possibility of genetic heterogeneity as well. A first step in our analysis included a detailed inter-specimen comparison of Man Bac individuals in order to confirm the apparent within-sample phenotypic variation.

In order to maximise the size of the comparative sample in the face of variable preservation nine cranial measurements (Martin's measurements 1, 8, 9, 45, 48, 51, 52, 54 and 55) were selected for calculating the Q-mode correlation coefficient. Of the 16 Man Bac adult male skulls, 14 met the condition of having a complete data set. The Hoabinhian series was also assessed by individual. The sample, including the nine measurements, for both the Man Bac and Hoabinhian assemblages is presented in Table 3.

The results of Neighbor-Net Split analysis applied to the Q-mode correlation coefficients are illustrated in Figure 4. Concentrated clusters, as expressed through web or net-like connections, indicate morphologically similar groupings. The boxing or web-like connections can be visualized as alternative biological connections, some being genuine linkages with other individuals or groups, while others may be due to other factors such as random similarities arising from parallel micro-evolutionary change.

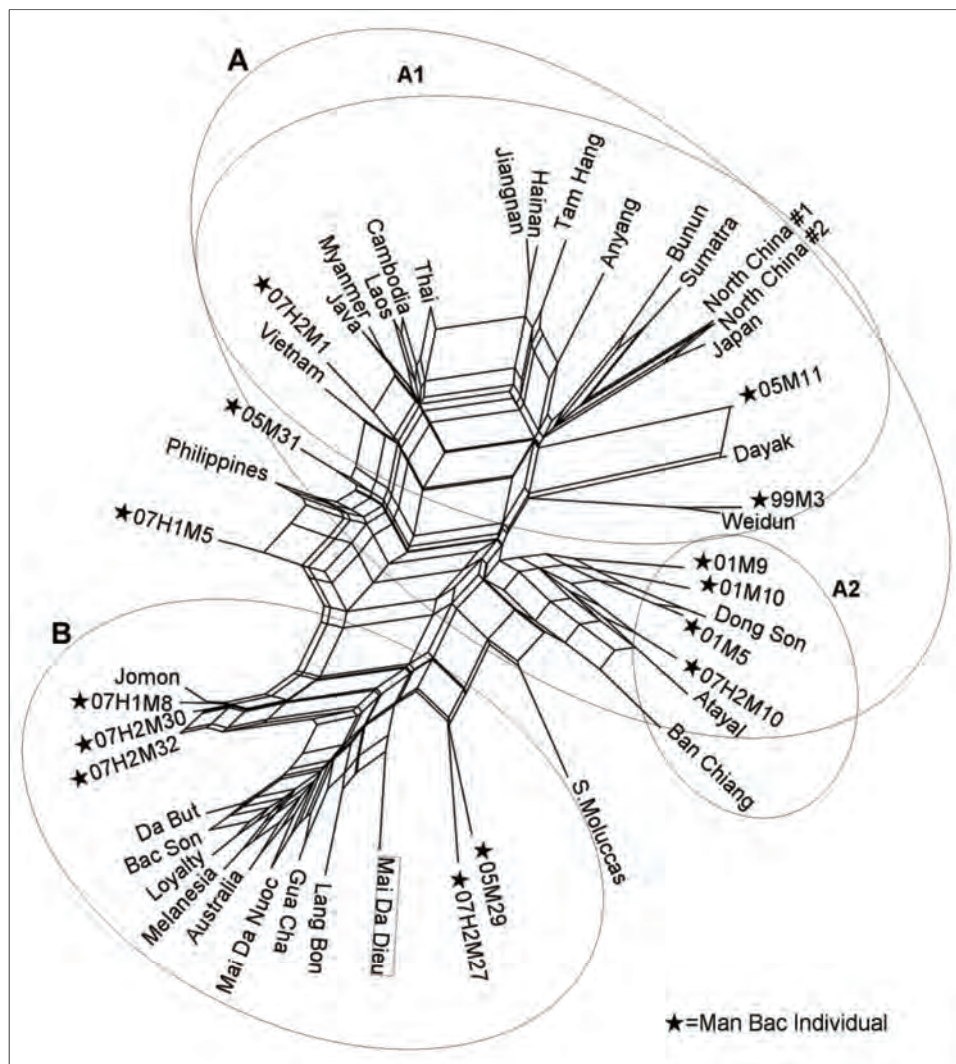
The Neighbor-Net illustrated in Figure 4 indicates a dichotomisation of the samples into two mega-clusters. Mega-cluster A includes the majority of Man Bac individuals (n=8) which are situated among Neolithic, Metal Period and modern samples from East and Southeast Asia. Mega-cluster A consists of two sub-clusters, A1 includes four Man Bac cases more closely associated with modern Vietnamese and a range of modern Southeast Asian samples. The second sub-cluster, A2, which includes the other four Man Bac cases, is situated within a grouping that contains Metal Period Vietnamese and late Neolithic through Metal Period samples from northern Thailand. Mega-cluster B, which includes five Man Bac individuals, is characterised by the presence of Hoabinhian samples and Australian and Melanesian series, in addition to the Jomon. The only Man Bac individual not clearly placed in this neighbor-net is 07H1M5. However, this case, along with the Philippines sample, is in fact more closely associated with A1 than B.

### 5.2 Group Average Comparison

In the next step of the analysis cranial affinities were assessed using a group-average data set for the Man Bac and Hoabinhian Vietnamese samples. The reason for such an approach was to increase the available sample size. In the previous analysis (see Figure 4) nine measurements for 16 male Man Bac crania were used. In the group-average comparison more fragmentary crania could be added to the sample increasing the total Man Bac series to 16 male crania analysed using 16 measurements. As five Man Bac individuals (cluster B in Figure 4) were clearly separated from the others, these specimens were dealt with as a separate group labelled 'Man Bac 1' (n = 5). The remaining eleven individual Man Bac specimens were combined with other incomplete male crania (05M16, 05M20, 07H1M9, 07H2M19) and their group average dataset was designated 'Man Bac 2' (n = 11). The new cranial data set used for the calculation of Q-mode correlation coefficients consists of 16 measurements commonly available among the comparative

samples (Martin's M1, M8, M9, M17, M45, M48, M51, M52, M54, M55; and Yamaguchi's facial flatness measurements: frontal chord, frontal subtense, nasal chord, nasal subtense, zygomaxillary chord, zygomaxillary subtense; see Table 4).

Figure 5 exhibits the Neighbor-Net based on the Q-mode correlation coefficients derived from the new enlarged group-average dataset. Man Bac 2 and the Dong Son sample form a close grouping and are, in turn, connected with modern Vietnamese and prehistoric groups from Tam Hang Laos and Ban Chiang Thailand. These clusters form a mega-cluster that includes other early Metal Period to modern period samples from East/Southeast Asia as well as the Neolithic Weidun (China) sample. Man Bac 1 is tightly associated with the early Vietnamese Hoabinhian, Bac Son and Da But samples, which, in turn, form a mega-cluster with the Australo-Melanesian, Gua Cha Malay (Hoabinhian) and Neolithic Jomon (Japan) samples.



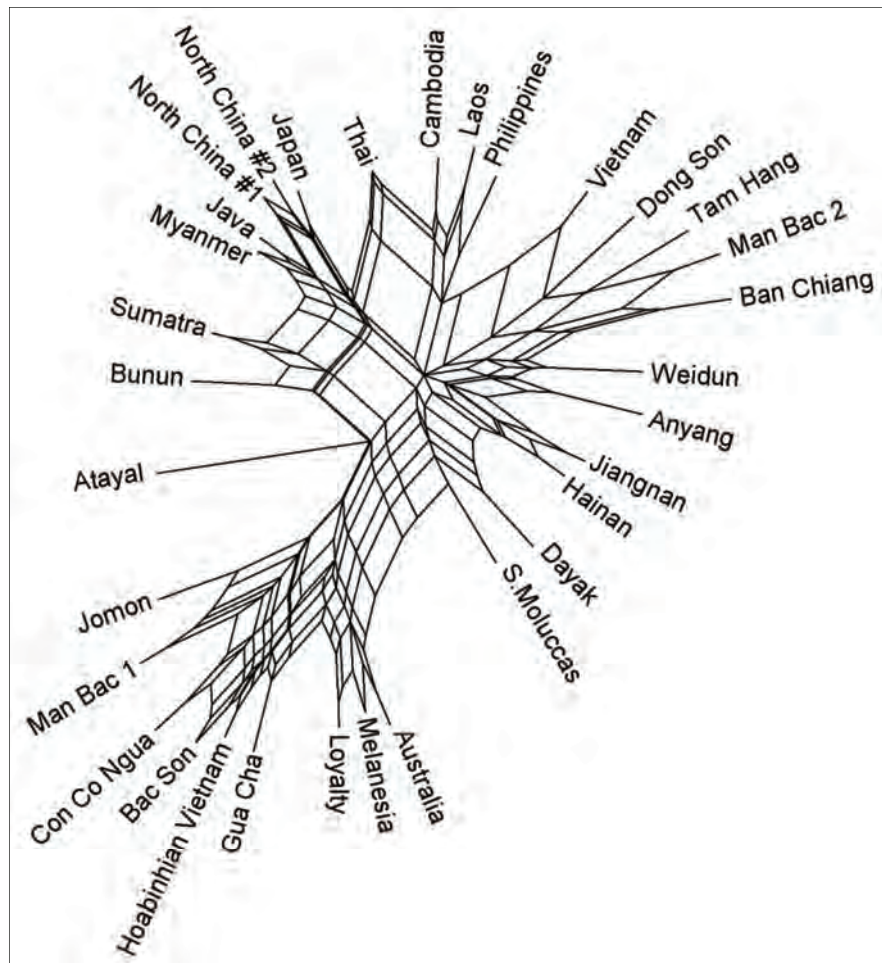
**Figure 4:** A network tree of the Neighbor-Net split analysis applied to the Q-mode correlation coefficients, based on 9 cranial measurements (males).

**Table 3:** Cranial measurements (mm) of Hoabinhian and Late Neolithic Man Bac individuals (males)

Martin's No & measurement	Hoabinhian		Late Neolithic Man Bac														
	Lang <sup>9</sup> Bon	Mai Da Dieu 16	Mai Da Nuoc	99 M3	01 M5	01 M9	01 M10	05 M11	05 M29	05 M31	07H1 M5	07H1 M8	07H2 M1	07H2 M10	07H2 M27	07H2 M30	07H2 M32
1 Maximum cranial length	189	190	185	180	185	174	184	173	186	175	177	185	176	181	192	181	184
8 Maximum cranial breadth	146	144	131	139	143	137	145	136	138	155	140	147	150	141	135	151	147
9 Minimum frontal breadth	90	90	95	96	103	98	102	93	95	101	96	101	94	106	94	108	104
45 Bizygomatic breadth	131	152	134	128	142	147	144	130	138	146	142	146	147	146	133	147	146
48 Upper facial height	60	67	62	70	75	67	68	70	67	74	70	69	71	79	74	66	68
51 Orbital breadth	43	42	42	42	40	40	41	40	42	39	45	44	44	43	43	45	47
52 Orbital height	33	34	32	34	34	35	35	39	30	34	34	35	37	37	32	34	34
54 Nasal breadth	28	29	29	28	29	28	26	28	26	29	29	28	26	27	26	31	33
55 Nasal height	45	49	48	53	56	51	54	50	49	53	54	49	54	63	55	50	52

**Table 4:** Cranial measurements (mm) of archaeological and modern population samples from Northern Vietnam (males)

Population	Hoabinhian		Bac Son		Da But		Man Bac 1		Man Bac 2		Dong Son		Modern Vietnam								
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD						
<b>Martin's No &amp; measurement</b>																					
1 Max. cranial length	5	192.4	7.2	7	190.9	6.1	19	186.7	6.5	5	185.6	4.0	11	179.7	5.7	32	181.6	6.8	27	175.7	4.9
8 Max. cranial breadth	4	141.0	6.8	7	141.6	7.9	19	132.8	6.2	5	143.6	6.8	11	142.4	5.7	31	137.4	6.1	27	140.5	4.9
9 Min. frontal breadth	5	94.0	4.2	7	97.6	3.6	19	99.6	5.1	5	100.4	5.9	11	98.5	4.0	25	96.7	4.4	27	95.1	5.1
17 Basion-bregma height	3	138.3	3.8	4	138.3	5.1	2	130.0	0.7	5	143.8	6.3	6	140.0	8.9	14	138.3	7.5	27	138.0	4.3
45 Bizygomatic breadth	4	137.0	10.1	3	139.7	2.1	5	135.6	6.5	5	142.0	6.2	10	141.6	6.9	17	137.4	6.1	27	136.9	4.8
48 Upper facial height	5	64.6	4.4	5	67.6	4.3	4	63.0	6.8	5	68.8	3.1	10	71.0	4.0	17	69.8	2.9	26	69.2	3.3
51 Orbital breadth	5	42.2	0.8	5	43.4	1.7	4	42.3	3.6	5	44.2	1.9	9	41.6	2.1	20	41.3	2.3	27	41.5	1.4
52 Orbital height	5	33.6	1.5	6	35.3	1.8	4	34.3	3.6	5	33.0	2.0	11	35.5	1.8	21	34.2	1.8	27	34.7	2.0
54 Nasal breadth	5	27.8	1.3	5	28.6	4.1	5	26.2	2.9	5	28.8	3.1	11	28.1	1.4	19	27.3	1.9	27	26.9	1.6
55 Nasal height	5	49.0	3.2	5	50.6	3.9	5	45.7	2.7	5	51.0	2.5	10	54.1	3.5	16	52.4	3.1	27	51.7	2.6
43(1) Frontal chord	5	99.8	3.7	7	102.3	2.6	7	103.6	6.3	5	104.2	5.1	8	103.5	6.6	20	99.4	3.9	25	99.1	3.4
43c Frontal subtense	5	15.6	3.2	7	16.8	4.2	6	14.3	2.3	5	16.3	5.1	8	14.6	3.2	19	16.1	3.6	25	15.6	4.1
57 Simotic chord	5	9.9	2.0	4	9.8	2.4	3	8.0	5.5	5	9.6	1.6	5	10.2	2.6	17	9.6	1.7	25	9.1	2.3
57a Simotic subtense	5	2.3	1.1	4	2.6	0.8	7	1.8	4.0	4	3.2	0.7	5	3.0	1.9	15	2.9	0.7	24	2.6	0.9
46b Zygomaxillary chord	5	103.3	10.1	5	107.8	8.1	6	109.6	0.5	5	107.6	6.1	6	107.0	6.8	17	103.2	4.0	25	103.5	4.6
46c Zygomaxillary subtense	5	24.1	2.7	4	24.8	1.8	2	22.5	2.2	5	22.6	4.1	6	22.5	4.7	15	19.4	3.2	25	20.4	3.0



**Figure 5:** A network tree of the neighbour-net split analysis applied to the Q-mode correlation coefficients, based on 16 cranial measurements (males).

## 6 Discussion

Several sets of human remains are now available to study from late Pleistocene through to early Holocene sites in Southeast Asia. In Malaysia, Niah Cave in Sarawak is the site of the earliest well-dated modern human remains in the region. The so-called ‘Deep Skull’ from Niah Cave has an associated radiocarbon date of ca. 40,000 years BP (Kennedy 1977; Barker et al. 2007). The Tabon Cave on Palawan Island is also a well-known site that has produced the oldest human skeletal remains in the Philippines, consisting of a tibia, frontal bone and two mandibular fragments, the latter of which has been AMS dated to ca. 30,000 years BP (Dizon et al. 2002). However, a very recent discovery from Callao cave in northern Luzon may push this date back to approximately 67,000 years BP (Mijares et al. 2010). The Wadjak skulls, from central Java in Indonesia (Dubois 1922; Weidenreich 1945; Wolpoff et al. 1984), have long been regarded as Late Pleistocene, but AMS dating of the skeletal material indicates that a Middle Holocene date (ca. 6500 years BP) may be more appropriate (Storm 1995). From the Moh Khiew Cave in southern Thailand, an adult female skeleton was discovered by Pookajorn (1991, 1994), with an estimated age of 25,800±600 years BP based on a charcoal sample from the burial layer. Finally, an adult

male skeleton from the cave site Gua Gunung in Lenggong District, Mainland Malaysia, has been dated to 10,000–11,000 years BP by Zuraina (1994, 2005).

The majority of analyses of late Pleistocene and early Holocene human skeletal material, including other early pre-ceramic remains from Malaysia and Indonesia, clearly demonstrate a common Australo-Melanesian cranio-dental morphology, despite issues associated with using subadult or incomplete material (for example, Evans 1918; Duckworth 1934; Mijsberg 1940; Trevor & Brothwell 1962; Jacob 1967; Brothwell 1960; Macintosh 1978; Bulbeck 2000; Matsumura & Zuraina 1999; Matsumura & Pookajorn 2005; Matsumura 2006; Bulbeck et al. 2007). The very well preserved Hoabinhian samples used in the present study were consistently classified as having a close Australo-Melanesian affinity in terms of their cranio-metrically expressed morphology. Because the present study focused on male crania, in order to avoid the confounding affects of sexual dimorphism, the important female skull from the cave site of Hang Cho, Northern Vietnam, was not included in our analysis. However, a previous cranio-metric study of this individual (Matsumura 2008b) found considerable similarities between the Hang Cho material and Australian and/or Melanesian samples, suggesting close biological ties.

It would seem reasonable to suggest that available Hoabinhian material characterises populations that are descended from the first anatomically modern human colonisers of Southeast Asia and the Australian sub-continent. These Hoabinhian individuals may, in turn, share a common ancestry with present-day Australian Aboriginal and Melanesian people. The present study further demonstrates that some cranial traits characterising these early populations in the region were retained through the early Holocene (for example Bac Son) and on into the Middle Holocene (for example, Da But) in northern Vietnam. This suggests that ceramic-using pre-agricultural foraging communities, operating in the region as late as the Mid-Holocene, perhaps even as late as 3000 to 4000 years BP, are likely direct lineal descendants of Hoabinhian settlers. Nonetheless, it is important to note that a broad comparison of dental traits conducted by Matsumura and Hudson (2005) highlighted a close affinity between Hoabinhian and Australo-Melanesian samples on the one hand, but identified a northern source for contemporary Southeast Asians (including modern Vietnamese) on the other hand. This observation supports an immigration, rather than regional continuity, model for the origins of modern Southeast Asians.

In northern Vietnam, examples of sweeping population change, likely associated with large-scale admixture with North/East Asians, are especially evident from the Neolithic (e.g. Phung Nguyen and Ha Long periods) and early Metal Age. The present study has demonstrated a large morphological lacuna between the Metal Period Dong Son on the one hand and early Holocene Bac Son and mid-Holocene Da But series on the other. This biological discontinuity is further corroborated through recent analyses of dental morphology (Matsumura et al. 2001). The close resemblance between Dong Son crania and those of recent East Asians suggests large scale gene flow into northern Vietnam brought about by immigrants from peripheral regions to the north and northeast of Vietnam during the Neolithic and/or early Metal Period. Man Bac, a Neolithic site dated to between 3900–3500 years BP, may be pivotal in resolving the question of the dispersal timing of food-producing populations into northern Vietnam. Cranio-morphologically, the Man Bac series is clearly not a monophyletic group. Some individuals closely resemble the earlier pre-Neolithic inhabitants of the region, while others show a close affinity to the later Dong Son inhabitants. This remarkable intra-group variation in cranial morphology suggests an initial appearance of immigrants at Man Bac with a genetic inheritance located in the northern peripheral region of Vietnam, which includes what is now southern China. Man Bac may be an example of one of those extremely rare archaeological snapshots of a



population in transition; a somewhat cosmopolitan mix of indigenous inhabitants tracing their origins back to the Hoabinhian and new comers with a genetic heritage located outside of the region. The eventual outcome of this integration was a new population that contributed to the modern Southeast Asian morphology.

As noted previously, many prehistorians and linguists have shown that the prehistoric expansion of language families, including Austronesian and Austroasiatic, can often be linked with the Neolithic dispersal of food-producing populations. A number of recent reviews summarise the archaeological evidence for the spread of this Neolithic horizon into Southeast Asia from the third millennium BC (Spriggs 1989; Glover & Higham 1996; Bellwood 2005; Bellwood 1997, 2005; Bellwood et al. 1992) and hypothesise that there was a diffusion of migrants, probably through southern China and down the Mekong and Red river valleys. A recent mitochondrial DNA analysis by Tan (2001) and Lertrit et al. (2008) also corroborates the archaeological and linguistic views for such an early genetic contribution, perhaps by Austroasiatic speaking people. The present analysis supports this latter view and, in fact, the late Neolithic Ban Chiang series is unambiguously associated with modern and Metal Period samples rather than earlier Hoabinhian and/or Australo-Melanesian samples, thus further supporting the Two Layer model. Ban Chiang cannot be used to support a model for regional continuity, as has been done by Pietrusewsky (1981, 1994) and Pietrusewsky and Douglas (2002) in the past. If there is any validity in the regional continuity model, it will need to be sought in the post-Neolithic period, after the major population changes characterised by the Two Layer model have occurred. Finally, it is clear that the Man Bac series is of crucial importance to a more robust understanding of the population history of Southeast Asia.

## **7 Acknowledgements**

We express our sincere gratitude to Prof. Nick Enfield, Max Planck Institute and Radboud University Nijmegen and Dr. Joyce White, University of Pennsylvania for inviting the first author to the workshop entitled 'Dynamics of human diversity in mainland Southeast Asia' and to contribute to this book. We are grateful for the collaborations of Drs. Ha Van Phung, Nguyen Giang Hai, the Vietnamese Institute of Archaeology, and Dr. Peter Bellwood, Australian National University for the Man Bac excavation projects in Vietnam.

Thanks are also due to Dr. Chris Stringer, Department of Palaeontology, the Natural History Museum, London, Mr. Korakot Boonlop, the Princess Maha Chakri Sirindhorn Anthropology Centre, Bangkok, Dr. Michael Pietrusewsky, the University of Hawai'i, Dr. Nguyen Viet, the Center for Southeast Asian Prehistory, Hanoi, Dr. Philippe Menecier, Departement Hommes, Musee de l'Homme, Paris, Dr. Robert Foley, Department of Biological Anthropology, the University of Cambridge, Dr. Tsai Hsi-Kue, National Taiwan University, College of Medicine, Dr. Wang Daw-Hwan, IHP, Academia Sinica, Taipei, and Dr. Wilfred Ronquillio, Archeology Division, National Museum of the Philippines for permission to study the comparative cranial specimens.

This study was supported in part by a Grant-in-Aid in 2003–2011 (No.15405018, No.20370096, No.18520593, No. 20520666, No. 23247040) from the Japan Society for the Promotion of Science, and by the Toyota Foundation (No. D06–R–0035) and an Australian Research Council Grant.

## References

- Barker, Greame, Huw Barton, Michael Bird, Patrick Daly, Ipoi Datan, Alan Dykes, Lucy Farr, David Gilbertson, Barbara Harrisson, Chris Hunt, Tom Higham, Lisa Kealhofer, John Krigbaum, Helen Lewis, Sue McLaren, Victor Paz, Alistair Pike, Phil Piper, Brian Pyatt, Ryan Rabett, Tim Reynolds, Jim Rose, Gary Rushworth, Mark Stephens, Chris Stringer, Jill Thompson & Chris Turney. 2007. The human revolution in lowland tropical Southeast Asia: the antiquity and behaviour of anatomically modern humans at Niah Cave (Sarawak, Borneo). *Journal of Human Evolution* 52. 243–261.
- Bellwood, Peter. 1987. The prehistory of Island Southeast Asia: a multidisciplinary review of recent research. *Journal of World Prehistory* 1. 171–224.
- 1989. The colonization of the Pacific: some current hypotheses. In Adrian V.S. Hill & Susan W. Serjeantson (eds) *The Colonization of the Pacific: a Genetic Trail*, 1–59. Oxford: Clarendon Press.
- 1992. Southeast Asia before history. In Nicholas Tarling (ed.) *The Cambridge History of Southeast Asia: Volume One From Early Times to c. 1800*, 55–136. Cambridge: Cambridge University Press.
- 1996. Early agriculture and the dispersal of the southern Mongoloids. In Takeru Akazawa & Emöke J.E. Szathmàry (eds) *Prehistoric Mongoloid Dispersals*, 287–302. Oxford: Oxford University Press.
- 1997. *Prehistory of the Indo-Malaysian Archipelago, Revised Edition*. Honolulu: University of Hawai'i Press.
- 2005. *First Farmers: the Origins of Agricultural Societies*. Oxford: Blackwell.
- 2005. Examining the farming/language dispersal hypothesis in the East Asian context. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London and New York: Routledge Curzon.
- Bellwood, Peter, Richard Gillespie, Gillian Thompson, John S Vogel, Wayan Ardika & Ipoi Datan. 1992. New dates for prehistoric Asian rice. *Asian Perspectives* 31. 161–170.
- Bellwood, Peter & Colin Renfrew (eds). 2003. *Examining the Farming/Language Dispersal Hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- Blust, R.A. 1996a. Austronesian culture history: the window of language. In Ward H. Goodenough (ed.) *Prehistoric settlement of the Pacific*, 28–35. Philadelphia: American Philosophical Society.
- 1996b. Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. In Ward H. Goodenough (ed.) *Prehistoric settlement of the Pacific*, 117–140. Philadelphia: American Philosophical Society.
- 1977. The proto-Austronesian pronouns and Austronesian subgrouping. *Working Papers in Linguistics* 9. 1–15. Honolulu: University of Hawai'i.
- 1984–85. The Austronesian homeland: a linguistic perspective. *Asian Perspectives* 26. 45–68.
- Brace, Loring C., David P. Tracer & Kevin D. Hunt. 1991. Human craniofacial form and the evidence for the peopling of the Pacific. *Bulletin of the Indo-Pacific Prehistory Association* 12. 247–269.

- Bräuer, Günter. 1988. Osteometrie. In Rudolf Martin, Rainer Knußmann, Ilse Schwidetzky, Hans W. Jürgens & Gerfried Ziegelmayr (eds) *Anthropologie: Handbuch der Vergleichenden Biologie Des Menschen*, 160–232. Stuttgart: Gustav Fisher.
- Brothwell, Don R. 1960. Upper Pleistocene human skull from Niah Caves. *Sarawak Museum Journal* 9. 323–349.
- Bulbeck, David. 2000. Dental morphology at Gua Cha, West Malaysia, and the implications for “Sundadonty”. *Bulletin of the Indo-Pacific Prehistory Association* 19. 17–41.
- Bulbeck, David, Marc Oxenham, Nguyen Lan Cuon & Nguyen Kim Thuy. 2007. Implication of the terminal Pleistocene skull from Hang Muoi, northern Vietnam. *Khao Co Hoc (Vietnamese Archaeology)* 2. 42–52.
- Callenfels, van Stein Pieter V. 1936. The Melanesoid civilizations of Eastern Asia. *Bulletin of the Raffles Museum* 1 (Series B). 41–51.
- Colani, Madeleine. 1927a. L'âge de la pierre dans la province de Hoa Binh. *Mémoires du Service Géologique de l'Indochine* 13.
- 1927b. La grotte sépulcrale de Lan Gao. *Anthropologie* 37.
- 1939. La Civilisation Hoabinhienne Extrême-Orientale. *Bulletin de la Société Préhistorique Française* 36. 170–174.
- Coon, Carleton S. 1962. *The Origin of Races*. New York: Alfred A Knoph.
- Diamond, Jared & Peter Bellwood. 2003. Farmers and their languages: the first expansions. *Science* 300. 597–603.
- Dizon, Eusebio, Florent Détroit, Francois Sémah, Christophe Falguères, Sebastien Hameau, Wilfredo Ronquillo & Emmanuel Cabanis. 2002. Notes on the morphology and age of the Tabon Cave fossil Homo sapiens. *Current Anthropology* 43. 660–666.
- Dubois, Eugène. 1922. The proto-Australian fossil man of Wadjak, Java. *Koninklijke Akademie van Wetenschappen te Amsterdam B23*. 1013–1051.
- Duckworth, Wynfrid Lawrence Henry. 1934. Human remains from rock-shelters and caves in Perak, Pahang and Perlis and from Selinsing. *Journal of Malayan Branch of the Royal Asiatic Society* 12. 149–167.
- Evans, Ivor Hugh Norman. 1918. Preliminary report on cave exploration near Lenggong, upper Perak. *Journal of the Federated Malay States Museum* 7. 227–234.
- Glover, Ian C. & Higham, Charles F.W. 1996. New evidence for early rice cultivation in South, Southeast and East Asia. In David R. Harris (ed.) *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, 413–441. London: UCL Press.
- Gorman, Chester F. & Pisit Charoenwongsa. 1976. Ban Chiang: a mosaic of impressions from the first two years. *Expedition* 18. 14–26.
- Habu, Junko. 2004. *Ancient Jomon of Japan*. Cambridge: Cambridge University Press.
- Han, Kangxin & Pan Qifeng. 1985. The study of the human bones of the middle and small cemeteries of Yin sites, Anyang. In Institute of History and Institute of Archaeology (IHIA) (eds) *Contributions to the Study on Human Skulls from the Shang Sites at Anyang Cultural Relics*, 50–81. Beijing: Publishing House (in Chinese).
- Hanihara, Tsunehiko. 1992. Negritos, Australian Aborigines, and the proto-sundadont dental pattern: the basic populations in East Asia, V. *American Journal of Physical Anthropology* 88. 183–196.

- 1993a. Craniofacial features of Southeast Asians and Jomonese: A reconsideration of their microevolution since the late Pleistocene. *Anthropological Science* 101. 25–46.
- 1993b. Population history of East Asia and the Pacific as viewed from craniofacial morphology: The basic populations in East Asia, IV. *American Journal of Physical Anthropology* 91. 173–187.
- 1993c. Cranial morphological contrasts between Negritos, Australians, and neighbouring populations. *Anthropological Science* 101. 389–404.
- 1994. Craniofacial continuity and discontinuity of Far Easterners in the late Pleistocene and Holocene. *Journal of Human Evolution* 27. 417–441.
- 2000. Frontal and facial flatness of major human populations. *American Journal Physical Anthropology* 111. 105–134.
- 2006. Interpretation of craniofacial variation and diversification of East and Southeast Asia. In Marc F. Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 91–111. Cambridge: Cambridge University Press.
- Hiep, Trinh. Hoang & Phung HaVan. 2004. Man Bac location and its relationship through ceramic data. *Khao Co Hoc (Vietnamese Archaeology)* 6. 13–48 (in Vietnamese with English summary).
- Higham, Charles F. W. 1989. *The Archaeology of Mainland Southeast Asia* (Cambridge World Archaeology Series). Cambridge: Cambridge University Press.
- 1996. *The Bronze Age of Southeast Asia*. Cambridge: Cambridge University Press.
- 1998. Archaeology, linguistics and the expansion of the East and Southeast Asian Neolithic. In Roger Blench & Matthew Spriggs (eds) *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*, 103–114. London: Routledge.
- 2001. Prehistory, language and human biology: is there a consensus in East and Southeast Asia? In Li Jin, Mark Seielstad & Chunjie Xiao (eds) *Genetic, Linguistic and Archaeological Perspectives on Human Diversity in Southeast Asia*, 3–16. Singapore: World Scientific.
- Howells, William White. 1989. Skull Shapes and the Map: Cranio-Metric Analysis in the Dispersion of Modern Homo. *Papers of the Peabody Museum of Archaeology and Ethnology* 79. Cambridge: Harvard University Press.
- Huard, Paul & Edmond Saurin. 1938. État actuel de la craniologie Indochinoise. *Bulletin du Service Géologique de l'Indochine* XXV.
- Huson, Daniel H. & David Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23. 254–267.
- Huffer, Damien G. 2005. *Social Organization at the Neolithic/Bronze Age Boundary in North Vietnam: Man Bac Cemetery as a Case Study*. MA thesis, School of Archaeology and Anthropology. Canberra: Australian National University.
- HUGO Pan-Asian SNP Consortium. 2009. Mapping Human Genetic Diversity in Asia. *Science* 326. 1541–1545.
- Institute of History and Institute of Archaeology (IHIA) and Chinese Academy of Social Science (CASS) (eds). 1982. *Contributions to the Study on Human Skulls from the Shang Sites at Anyang*. Beijing: Cultural Relics Publishing House (in Chinese with English summary).
- Jacob, Teuku. 1967. *Some Problems Pertaining to the Racial History of the Indonesian Region*. Ph.D. Dissertation. Utrecht: University of Utrecht.

- 1975. Morphology and paleontology of early man in Java. In Russel H. Tuttle (ed.) *Paleoanthropology, Morphology, and Paleoecology*, 311–324 Paris: Hague.
- Kennedy, Kenneth A. R. 1977. The deep skull of Niah: An assessment of twenty years of speculation concerning its evolutionary significance. *Asian Perspectives* 20. 32–50.
- Khao, Co Hoc. 2007. Thong bao mhung key qua do tuoi carbon phong xa moi tu Phong thi mghiem NFS (Arizona, USA). *Khao Co Hoc* 1. 101. (in Vietnamese).
- Lertrit, Patcharee, Samerchai Poolsuwan, Rachanie Thosarat, Thitima Sanpachudayan, Hathaichanoke Boonyarit, Chatchai Chinpaisal & Bhoom Suktitipat. 2008. Genetic history of Southeast Asian populations as revealed by ancient and modern human mitochondrial DNA analysis. *American Journal of Physical Anthropology* 137. 425–440.
- Macintosh, Neil William George. 1978. The Tabon Cave mandible. *Archaeological and Physical Anthropology in Oceania* 13. 143–159.
- Mansuy, Henry & Colani Madeleine. 1925. Contribution à l'étude de la pré-histoire de l'Indochine VII. Néolithique inférieur (Bacsonien) et Néolithique supérieur dans le Haut-Tonkin. *Bulletin du Service Géologique de l'Indochine* 12. 1–45.
- Matsumura, Hirofumi. 2006. The population history of Southeast Asia viewed from morphometric analyses of human skeletal and dental remains. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 33–58. Cambridge: Cambridge University Press.
- Matsumura, Hirofumi & Mark J. Hudson 2005. Dental perspectives on the population history of Southeast Asia. *American Journal of Physical Anthropology* 127. 182–209.
- Matsumura, Hirofumi, Lan Cuong Nguyen, Kim Thuy Nguyen & Tomoko Anezaki. 2001. Dental morphology of the early Hoabinhian, the Neolithic Da But and the Metal Age Dong Son Cultural people in Vietnam. *Zeitschrift für Morphologie und Anthropologie* 83. 59–73.
- Matsumura, Hirofumi, Marc F. Oxenham, Yukio Dodo, Kate Domett, Lan Cuong Nguyen, Kim Thuy Nguyen, Kim Dung Nguyen, Damien Huffer & Mariko Yamagata. 2008b. Morphometric affinity of the late Neolithic human remains from Man Bac, Ninh Binh Province, Vietnam: Key skeletons with which to debate the 'Two layer' hypothesis. *Anthropological Science* 116. 135–148.
- Matsumura, Hirofumi & Surin Pookajorn. 2005. Morphometric analysis of the Late Pleistocene human remains from Moh Khiew Cave in Thailand. *Journal of Comparative Human Biology* *Homo* 56. 93–118.
- Matsumura, Hirofumi, Minoru Yoneda, Yukio Dodo, Marc F. Oxenham, Kim Thuy Nguyen, Lan Cuong Nguyen, Lam My Dung, Vu The Long, Mariko Yamagata, Junmei Sawada, Kenichi Shinoda & Wataru Takigawa. 2008a. Terminal Pleistocene human skeleton from Hang Cho cave, northern Vietnam: Implications for the biological affinities of Hoabinhian people. *Anthropological Science* 116. 135–148.
- Matsumura, Hirofumi & Majid Zuraina. 1999. Metric analyses of the early Holocene human skeleton from Gua Gunung Runtuh in Malaysia. *American Journal of Physical Anthropology* 109. 327–340.
- Mijares, Armand S., Florent Détroit, Piper P., Rainer Grün, Peter Bellwood, Maxime Aubert, Guillaume Champion, Nida Cuevas, Alexandra DeLeon & Eusebio Dizon. (2010). New evidence for a 67,000-year-old human presence at Callao Cave, Luzon, Philippines. *Journal of Human Evolution* 59. 123–132.
- Mijsberg, Willem Alphonse. 1940. On a Neolithic Paleo-Melanesian lower jaw found in kitchen midden at Guar Kepah, Province Wellesley, Straits Settlements, 100–118. Singapore: Proceedings of 3rd Congress of Prehistorians of the Far East.

- Nakahashi, Takahiro & Minchang Li. (eds). 2002. *Ancient People in the Jiangnan Region, China*. Fukuoka: Kyushu University Press.
- Nakahashi, Takahiro, Minchang Li & Bin Yamaguchi. 2002. Anthropological study on the cranial measurements of the human remains from Jiangnan region, China. In Takahiro Nakahashi & Minchang Li (eds) *Ancient People in the Jiangnan Region, China*, 17–33. Fukuoka: Kyushu University Press.
- Nguyen, Kim Thuy. 1990. Ancient human skeletons at Con Co Ngua. *Khao Co Hoc (Vietnamese Archeology)* 3. 37–48 (in Vietnamese with English summary).
- 1993. Ancient skulls at Minh Duc. *Khao Co Hoc* 3. 1–8 (in Vietnamese with English summary).
- Nguyen, Kim Thuy & Gia Doi Nguyen. 1998. *Surveying Again Cho Cave (Luong Son, Hoa Binh Province), 1997*. Hanoi: Social Sciences Publishing House.
- Nguyen, Lang Cuon. 1986. Two early Hoabinhian crania from Thanh Hoa province, Vietnam. *Zeitschrift für Morphologie und Anthropologie* 77. 11–17.
- 1994. On the human remains from the Hoabinh culture in Vietnam. *Vietnam Social Sciences* 5. 64–69.
- 1996. *Anthropological Characteristics of Dong Son Population in Vietnam*. Hanoi: Social Sciences Publishing House (in Vietnamese with English title and summary).
- 2001. About human remains at Man Bac site. *Khao Co Hoc (Vietnamese Archeology)* 1. 17–46 (in Vietnamese with English title and summary).
- 2003. Ancient human bones in Da But Culture - Thanh Hoa Province. *Khao Co Hoc (Vietnamese Archeology)* 3. 66–79 (in Vietnamese with English title and summary).
- 2007. Paleoanthropology in Vietnam. *Vietnam Archaeology* 2. 23–41.
- Nishimura, Masanari. 2006. *Archaeological Study in the deltas of Red River Dong Nai River*. PhD Dissertation, University of Tokyo.
- Oxenham, Marc F. 2000. *Health and Behavior During the Mid-Holocene and Metal Period of Northern Viet Nam*. Unpublished PhD Thesis. Darwin, NT, Australia: Northern Territory University,.
- 2006. Biological responses to change in prehistoric Vietnam. *Asian Perspectives* 45. 212–239.
- Oxenham, Marc F., Hirofumi Matsumura, Nguyen, Kim Dung (eds.) 2011. *Man Bac: The Excavation of a Neolithic Site in Northern Vietnam. The Biology*. Canberra: ANU E Press.
- Oxenham, Marc F., Hirofumi Matsumura, Kate Domett, Kim Thuy Nguyen, Kim Dung Nguyen, Lan Cuong Nguyen, Damien Huffer & Sarah Muller. 2008a. Childhood in late Neolithic Vietnam: bio-mortuary insights into an ambiguous life stage. In Krum Bacvarov (ed.) *Babies Reborn: Infant/Child Burials in Pre- and Protohistory. B.A.R. International Series*, 123–136. Oxford: Archaeopress.
- 2008b. Health and the experience of childhood in late Neolithic Vietnam. *Asian Perspectives* 47. 190–209.
- Oxenham, Marc F., Kim Thuy Nguyen & Lan Cuong Nguyen. 2005. Skeletal evidence for the emergence of infectious disease in bronze and iron age northern Vietnam. *American Journal of Physical Anthropology* 126. 359–376.
- Oxenham, Marc F., Lan Cuong Nguyen & Kim Thuy Nguyen. 2006. The oral health consequences of the adoption and intensification of agriculture in Southeast Asia. In Marc Oxenham &

- Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 263–289. Cambridge: Cambridge University Press.
- Oxenham, Marc F., Lorna Tilley, Hirofumi Matsumura, Lan Cuong Nguyen, Kim Thuy Nguyen, Kim Dung Nguyen, Kate Domett & Damien Huffer. 2009. Paralysis and severe disability requiring intensive care in Neolithic Asia. *Anthropological Science* 2. 107–112.
- Patte, Etienne. 1965. Les ossements du kjoekenmodding de Da But. *Bulletin du Service Ethnologie de Indochine* 10. 1–87.
- Phung, Ha Vang. 2001. Man Bac site - Data and perception. *Khao Co Hoc (Vietnamese Archaeology)* 1. 17–46 (in Vietnamese with English summary).
- Pietrusewsky, Michael. 1981. Cranial variation in early metal age Thailand and Southeast Asia studied by multivariate procedures. *Homo* 32. 1–26.
- 1992. Japan, Asia and the Pacific: a multivariate craniometric investigation. In Hanihara Kazuro. (ed.) *Japanese as a Member of the Asian and Pacific Populations*, 9–52. Kyoto: International Research Center for Japanese Studies.
- 1994. Pacific-Asian relationships: a physical anthropological perspective. *Oceanic Linguistics* 33. 407–429.
- 1996. Multivariate craniometric investigations of Japanese, Asians, and Pacific Islanders. In Keiichi Omoto (ed.) *Interdisciplinary Perspectives on the Origins of the Japanese*, 65–104. Kyoto: International Research Centre for Japanese Studies.
- 1999a. The people of Ban Chiang: an early Bronze site in Northeast Thailand. *Bulletin of the Indo-Pacific Prehistory Association* 16. 119–148.
- 1999b. A multivariate craniometric study of the inhabitants of the Ryukyu Islands and comparison with cranial series from Japan, Asia and the Pacific. *Anthropological Science* 107. 255–281.
- 2005. The Physical anthropology of the Pacific, East Asia: A multivariate craniometric analysis. In Laurent Sagart, Roger Blench & Alica Sanchez-Mazas (eds) *The peopling of East Asia Putting together Archaeology, Linguistics and Genetics*, 201–229. Routledge Curzon.
- 2006. A multivariate craniometric study of the prehistoric and modern inhabitants of Southeast Asia, East Asia and surrounding regions: a human kaleidoscope? In Marc F. Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 59–90. Cambridge: Cambridge University Press.
- 2008. Craniometric variation in Southeast Asia and neighbouring regions: a multivariate analysis of cranial measurements. *Human Evolution* 23. 49–86.
- 2010. A multivariate analysis of measurements recorded in early and more modern crania from East Asia and Southeast Asia. *Quaternary International* 211. 42–54.
- Pietrusewsky, Michael & Ching-Fang Chang. 2003. Taiwan aboriginals and peoples of the Pacific-Asia region: multivariate craniometric comparisons. *Anthropological Science* 111. 293–332.
- Pietrusewsky, Michael & Michele T. Douglas. 2002. *Ban Chiang, a Prehistoric Village Site in Northeast Thailand I: The Human Skeletal Remains*. Philadelphia: University of Pennsylvania, Museum of Archaeology and Anthropology.
- Pookajorn, Surin. 1991. *Preliminary Report of Excavations at Moh Khiew Cave, Krabi Province, Sakai Cave, Tran Province and Ethnoarchaeological Research of Hunter-Gatherer Group, so-called Sakai and Semang at Trang Province*. Bangkok: Silpakorn University Press.

- 1994. *Final Report of Excavations at Moh Khiew Cave, Krabi Province; Sakai Cave Trang Province and Ethnoarcheological Research of Hunter-Gatherer Group, so Called Sakai or Semang at Trang Province*. Bangkok: Silpakorn University Press.
- Renfrew, Colin. 1987. *Archaeology and Language: the Puzzle of Indo-European Origins*. London: Jonathan Cape.
- 1989. Models of change in language and archaeology. *Transactions of the Philological Society* 87. 103–155.
- 1992. World languages and human dispersals: a minimalist view. In John A. Hall & Ian C. Jarvie (eds) *Transition to Modernity: Essays on Power, Wealth and Belief*, 11–68. Cambridge: Cambridge University Press.
- Sagart, Laurent. 2008. The expansion of setaria farmers in East Asia: a linguistic and archaeological model. In Alica Sanchez-Mazas, Roger Blench, Malcom Ross, Iliia Peiros & Marie Lin (eds) *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*, 133–157. London: Routledge Studies in the Early History of Asia.
- Sangvichien, Sanjai. 1971. *Physical Anthropology of the skull of Thai* [dissertation]. Bangkok: Faculty of Medicine Siriraj Hospital, Mahidol University, No.2514.
- Sieveling, Gale G. 1954. Excavations at Gua Cha, Kelantan, Part 1. *Federation Museums Journal* 1. 75–143.
- Sneath, Peter H. & Robert R. Sokal. 1973. *Numerical Taxonomy*. San Francisco: WH Freeman and Co.
- Spriggs, Matthew. 1989. The dating of the island Southeast Asian Neolithic: an attempt at chronometric hygiene and linguistic correlation. *Antiquity* 63. 587–613.
- Storm, Paul. 1995. The evolutionary significance of the Wajak Skulls. *Scripta Geologica* 110. National Natuurhistorisch Museum, Netherlands.
- Suzuki, Hisashi, Yuji Mizoguchi & Eduardo Conese. 1993. Craniofacial measurement of artificially deformed skulls from the Philippines. *Anthropological Science* 101. 111–127.
- Tan, Ha Van. 1980. The Hoabinhian in the Context of Vietnam. *Vietnamese Studies* 46. Hanoi.
- 1997. The Hoabinhian and before. *Bulletin of the Indo-Pacific Prehistory Association* 16. 35–41.
- Tan, Soon Guan. 2001. Genetic relationships among sixteen ethnic groups from Malaysia and Southeast Asia. In Li Jin, Mark Seielstad & Chunjie Xiao (eds) *Genetic, Linguistic and Archeological Perspectives on Human Diversity in Southeast Asia*, 83–91. Singapore: World Scientific.
- Tayles, Nancy & Marc F. Oxenham. 2006. Southeast Asian Bioarchaeology: Past and Present. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 1–30. Cambridge: Cambridge University Press.
- Trevor, J. C. & Don R. Brothwell. 1962. The human remains of Mesolithic and Neolithic date from Gua Cha, Kelantan. *Federation Museums Journal* 7. 6–22.
- Tryon, Darell. 1995. Proto-Austronesian and the major Austronesian subgroups. In Peter Bellwood, James J. Fox & Darell Tryon (eds) *The Austronesians: Historical and Comparative Perspectives*, 17–38. Canberra: RSPAS, The Australian National University.
- Turner, Christy G. II. 1989. Teeth and prehistory in Asia. *Scientific American* 260. 70–77.
- 1990. Major features of Sundadonty and Sinodonty, including suggestions about East Asian microevolution, population history and late Pleistocene relationships with Australian Aborigines. *American Journal of Physical Anthropology* 82. 295–317.



- 1992. Microevolution of East Asian and European populations: a dental perspective. In Takeru Akazawa, Knichi Aoki & Tasuku Kimura (eds) *The Evolution and Dispersal of Modern Humans in Asia*, 415-438. Tokyo: Hokusensha.
- Vinh, Bui. 1991. The Da But culture in the Stone Age of Vietnam. *Bulletin of the Indo Pacific Prehistory Association* 10. 127-131.
- Von Koenigswald, Gustav Heinrich Ralph. 1952. Evidence of a prehistoric Australo-Melanesoid population in Malaya and Indonesia. *Southwestern Journal of Anthropology* 8. 92-96.
- Weidenreich, Franz. 1945. Giant Early Man from Java and South China. *Anthropological Paper of the American Museum of Natural History* 40. New York.
- Wolpoff, Milford H., Xinzhi Wu & Alan G. Thorne. 1984. Modern homo sapiens origins: a general theory of hominid evolution involving the fossil evidence from east Asia. In Fred H. Smith & Frank Spencer (eds) *The Origins of Modern Humans*, 411-484. New York: Alan R. Liss.
- Yamaguchi, Bin. 1973. Facial flatness measurements of the Ainu and Japanese crania. *Bulletin of the National Science Museum Tokyo, Series D* 16. 161-171.
- 1982. A review of the osteological characteristics of the Jomon population in prehistoric Japan. *Journal of the Anthropological Society of Nippon* 90 (Supplement). 77-90.
- Yev-Yadun, Simcha, Avi Gopher & Shahal Abbo. 2000. The cradle of agriculture. *Science* 288. 1602-1603.
- Yi, Seonbok, Lee June-Jeong, Kim Seongnam, Yoo Yongwook & Kim Dongwan. 2008. New data on the Hoabinhian: investigations at Hang Cho cave, Northern Vietnam. *Bulletin of the Indo-Pacific Prehistory Association* 28. 73-79.
- Yokoh, Yasuo. 1940. Beiträge zur kranilogie der Dajak. *Japanese Journal of Medical Science, Part I Anatomy* 8. 1-354.
- Zuraina, Majid. 1994. The excavation of Perak Man, an Epi-Palaeolithic burial at Gua Gunung Runtuh. In Majid Zuraina (ed.) *The Excavation of Gua Gunung Runtuh and the Discovery of the Perak Man in Malaysia*, 23-47. Kuala Lumpur: Department of Museums and Antiquities, Malaysia.
- Zuraina, Majid (ed.). 2005. *The Perak Man and Other Prehistoric Skeletons of Malaysia*. Penang: Penerbit Universiti Sains Malaysia.

# 9 *A physical anthropology perspective on a mainland Southeast Asian agrarian population: prehistoric skulls of Ban Chiang*

---

KORAKOT BOONLOP AND SUREERATANA BUBPHA

## 1 Introduction

Ban Chiang is a prehistoric mounded village site located along the Upper Songkhram River in the province of Udon Thani, in the Sakon Nakhon Basin of the Khorat Plateau (Figure 1). The site was first discovered in the early 1960s. In 1967, Vidya Intakosai of the Thai Fine Arts Department (FAD) undertook a small test excavation at Ban Chiang and uncovered a complete inhumed skeleton with bronze and iron artefacts, together with plain and painted pottery vessels. Further excavations by Thai archaeologists in 1972-73 found evidence of bronze and iron metallurgy as well as skeletal remains. The Thai government encouraged further scientific research at the site, and a major research programme was jointly conducted in 1974-75 by the Fine Arts Department and the University of Pennsylvania Museum of Archaeology and Anthropology (UPMAA) under the direction of Pisit Charoenwongsa and Dr. Chester Gorman. This excavation uncovered conclusive evidence for the practice of bronze and iron metallurgy, many artifacts and remains of daily life, and approximately 134 burials. As a result of the 1974-75 excavations, the site achieved international fame and a listing on the UNESCO World Heritage Site List, both because of its beautiful red-on-buff pottery and its evidence for early bronze metallurgy.

The chronology has undergone revision over the years (Gorman & Charoenwongsa 1976; White 1982, 1986), but the most recent chronology (White 2008) dates the Early Period of the site to 2100-900 B.C., the Middle Period to 900 B.C.-300 B.C., and the Late Period to 300 B.C.-A.D. 300. The Protohistoric Period that followed saw the rise of the first states in mainland Southeast Asia. The occupation of Ban Chiang, still inhabited, thus stretches back at least four thousand years into the past.

After the FAD-Penn excavations, Thai and foreign archaeologists conducted surveys in the Sakon Nakhon Basin and adjacent parts of northern Thailand, looking for sites affiliated to the Ban Chiang Cultural Tradition. According to the 1991 survey of Ban Chiang Cultural Tradition (BCCT) sites, there are at least 126 sites with artefacts related to Ban Chiang located in the upper part of northeastern Thailand (Bannanurag & Bamrungwong 1991). Only sixteen of these sites (12.6% of the total discovered) have been

investigated and excavated (Bannanurag et al 1992). Moreover, most of them were small-scale excavations, with only preliminary reports and without intensive study of artifacts recovered. Nonetheless, data from the surveys and the sixteen excavations have led to a more detailed understanding of the Ban Chiang Cultural Tradition, to the extent that the 126 sites can be now divided into seven distinct clusters.



**Figure 1:** Map with location of Ban Chiang (by Ardeth Abrams).

## 2 Osteological studies of prehistoric skeletons in Thailand

The corpus of excavated and analysed human remains from Thailand is relatively small but growing. Published analyses from sites other than Ban Chiang in northern northeast Thailand include Sangvichien for Ban Kao (1966), Sangvichien & Supawan for Ban Thatu and Ban Om Kaew (1977), Pietrusewsky (1974 [Non Nok Tha]), Wiriyaromp for Ban Na Di (1984), Choosiri (1992; 1994 [Ban Tum and Ban Daeng Yai]), Agelarakis for Khao Wong Prachan valley (1997), Pureepatpong for excavation pit # S18 W22 at Ban Mai Chai Mongkhon (1996), Hutangkura for Ban Don Thong Chai (1997), Kae-In (1999 [Prasat Phanom Wan]), Tayles, Domett & Nelsen (2000), Nelsen, Tayles & Domett (2001), Boonlop (2003 [Ban Khok Khon]), Kaewsuwan (2003), Boonlop & Brown (2008 [Ban Khok Khon]). However, published analyses from Ban Chiang Cultural Traditions sites are much fewer; only five sites out of the 127 BCCT sites—Ban Thatu and Ban Na Di in Udon Thani, Ban Don Thong Chai and Ban Khok Khon in Sakon Nakhon, and Ban Chiang itself

have had analyses performed. The skeletal material from Ban Chiang has received the most attention, from Pietrusewsky (1982, 1984, 1997), Douglas (1996, 1997), Douglas & Pietrusewsky (2007) and especially Pietrusewsky & Douglas (2002). In addition, Kaewsuan (2003) analysed the subadult skeletons excavated from Wat Phosi Nai in 2003 to age subadults by teeth. There is an urgent need for more studies of the human remains from prehistoric Thailand, both to understand the morphological diversity of the prehistoric agrarian populations and to understand their genetic differences and similarities to modern-day populations.

### **3 The present study**

For Ban Chiang itself, the metric and non-metric study of human skeletons excavated in 1974–1975 (especially Pietrusewsky & Douglas 2002) has provided us with a great deal of information regarding the interpretation and reconstruction of the ancient population. The analysis of the complete skeletons has allowed us to understand not only details about individuals (age at death, sex, stature, growth and development, physical activities, and biological affiliation) but also details at the population level: diet, nutrition, health and disease, trauma and injury, frequency of warfare, physiological stress, cultural use of teeth and bones, and demography (Larsen 1997, 2000).

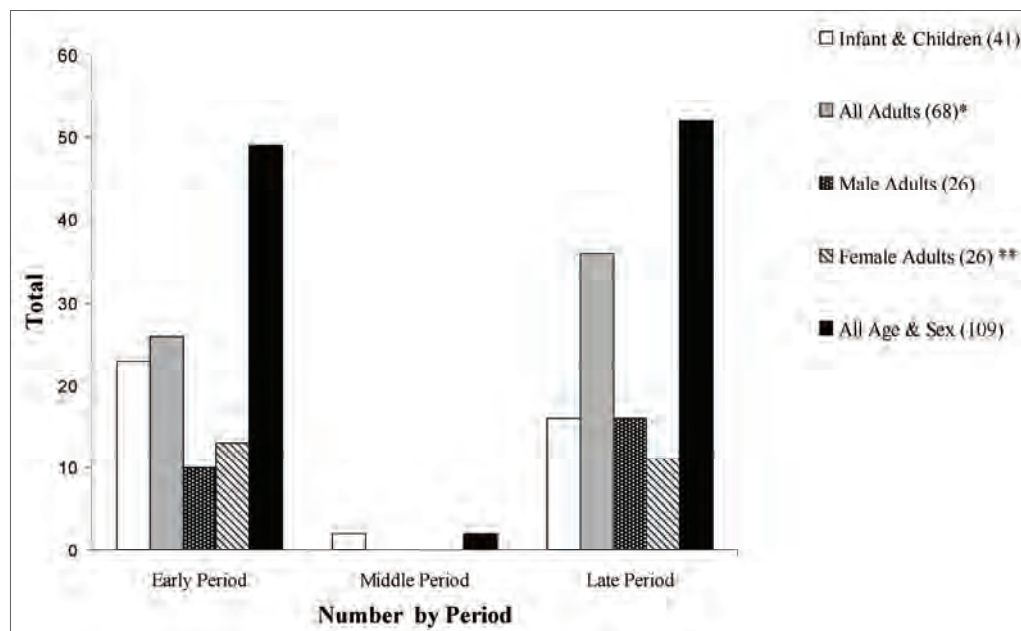
However, in anthropology, there is a long history of studies focusing on the cranial measurements alone. Part of the reason for this focus lies in the relative ease with which the skull may be used to distinguish between individuals and between ethnic populations (Brothwell 1981:77). Metric and non-metric are two major categories of information which have been used traditionally to study and describe skulls. Skull dimensions and traits are influenced by age, diet, muscular development, mechanical pressure, and environment. Moreover, heritability studies have demonstrated an underlying genetic basis to many aspects of skull morphology. Comparisons of skull morphology based on both metric and non-metric data have long been one of the primary avenues for research in physical anthropology in order to characterise and determine relatedness among past and present human groups.

Because of the importance of Ban Chiang, the authors, while employed by the Fine Arts Department of Thailand, returned in 2003–2004 to supervise the excavation of another area of the site, one near Wat Phosi Nai (see Figure 2). This appeared to be a mortuary area, with 109 burials uncovered from 15 x 15 metres excavated area, 3.5 metres deep (see Figures 3 and 4, and Table 1; data tables begin on p. 191). All the previous excavations at Ban Chiang except for the joint 1974–1975 excavations have been limited in scope, so that the Ban Chiang 2003–2004 skeletal sample is the second largest sample to be uncovered at the site.

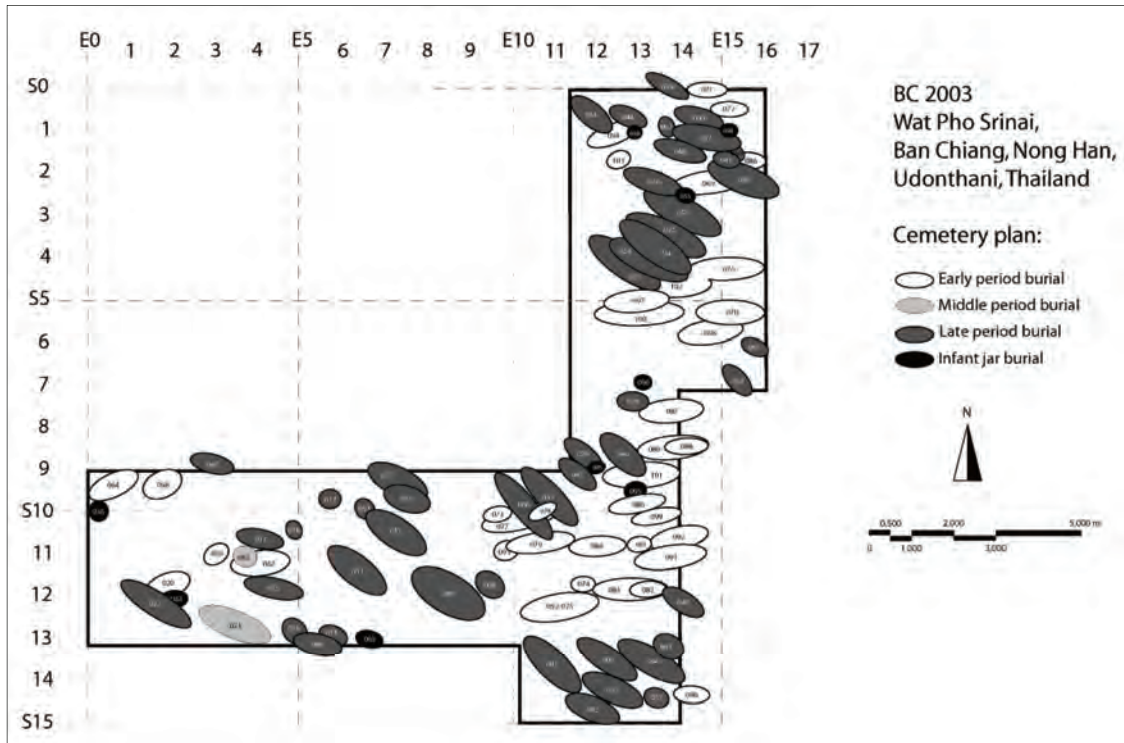
The skulls from these burials form the subject of this study. This chapter reports on preliminary results of a study of skull morphology from this Wat Phosi Nai sample, as well as some observations on mandibles. Our goal is to expand the data available on the diversity of human skull morphology in prehistoric mainland Southeast Asia, as well as to shed a bit more light on the controversy regarding migration of populations into Southeast Asia in the past. Standard methods of physical anthropology, including recording metric and non-metric variations, were applied to the specimens. The chapter also compares the 2003–2004 Ban Chiang human skull series with human skull remains from Ban Chiang 1974–1975 specimens, and the human skull samples from other Ban Chiang Cultural Tradition sites including Ban Na Di and Ban Khok Khon.



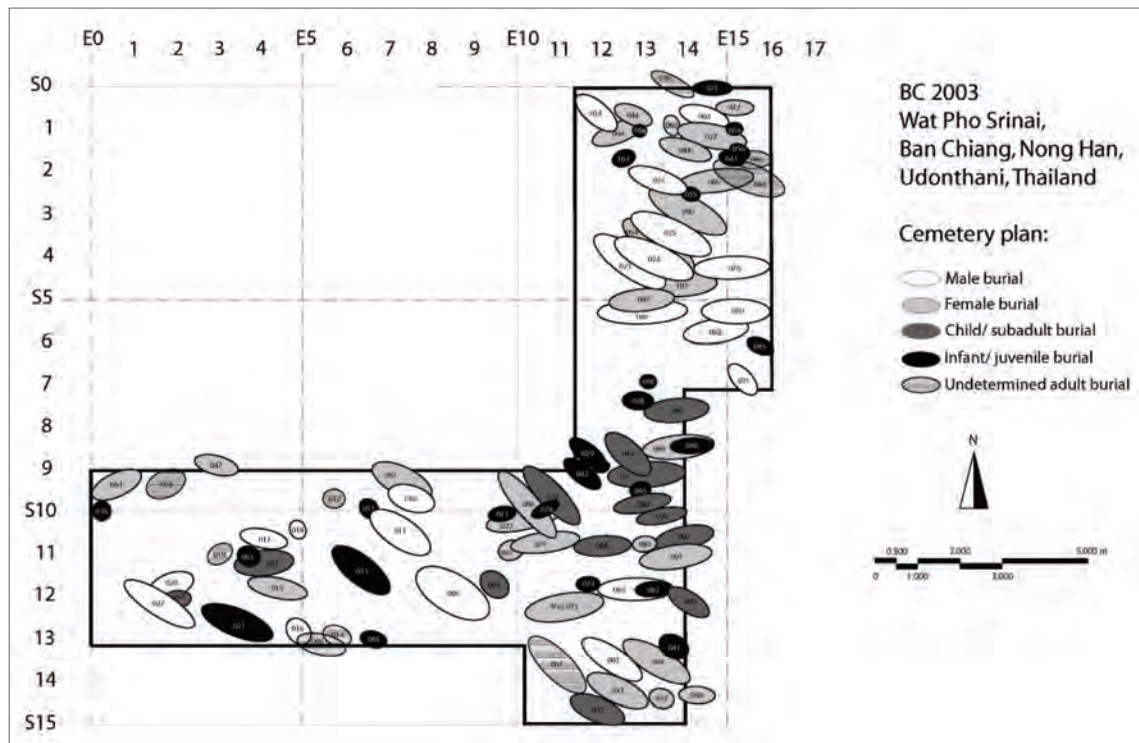
**Figure 2:** The excavation at Wat Phosi Nai, Ban Chiang in 2003 to 2004.



**Figure 3:** The number of individuals broken down by sex, ages at death, and period: \* 6 adults unidentified for period including 4 adults unidentified for sex; \*\* 2 females unidentified for period.



**Figure 4:** Cemetery plan at Wat Phosi Nai showing distribution of burials by cultural periods.



**Figure 5:** Cemetery plan at Wat Phosi Nai showing distribution of burials classified by sex and age.

## 4 Condition of the sample

The complete condition of these skeletons ranges from almost complete skeletons to fragmentary bones. Many burials are disturbed or mixed and the boundaries of the grave pits were difficult to clarify. The mixed and incomplete condition of many skeletons may be caused, in large measure, by recent disturbances. The first excavation of the area was in 1991; workers exposed the highest layer of skeletons and left the bones displayed as part of an open air museum exhibit. It was not until 2003 that the area was fully excavated and the bones and artifacts removed to museum storage, to be replaced in the exhibit by replicas. Overall, the preservation of the 2003–2004 Wat Phosi Nai Ban Chiang human skeletal series is poor to fair; just a few skeletons could be classified as being very well preserved. Excavation timing may have affected skeletal completeness of the lower layers as well; the excavation took place during the rainy season, with some resulting bone damage from the high water table and wet soil.

## 5 Methods

Standard methods of physical anthropology and human osteology were applied to the skeletal remains, including recording of metric and non-metric variations. The methods given in Hillson (1990:180–187, 1992, 1996), Ubelaker (1984:46–54), and White (1991:308–313) were used to determine age at death. The methods provided by Brothwell (1981:62–63), Choosiri (1991:126–132), Ubelaker (1984:42–59), and White (1991:320–327) were used for sex assessments. Methods of craniometry and cranioscopy are given in Boonprakorb (1993), Howells (1973), Krogman (1962), Laughlin & Jorgensen (1956), Martin & Saller (1957), Sangvichien (1970), Sjøvold (1984), Stewart (1940), White (1991), and Wilder (1920).

The excavation at Wat Phosi Nai in 2003–2004 revealed skeletal remains of individuals in both burial and non-burial contexts. A total of 109 skeletons were in sufficiently good condition to be studied. Of these, 41 were subadults with ages ranging from a few months to 18 years of age. Thirty of the subadults were aged between newborn and 4 years, six aged between 5–9 years, four aged between 10–14 years, and six aged between 15–19 years at time of death. Sixty-eight of the 109 remains were determined to be adults. Seventeen died between 20–24 years of age, six were aged between 25–29 years, eight were aged between 30–34 years, fourteen were aged between 35–39 years, one was 40–44 years of age at death, fifteen were aged between 45–49 years, and two were older than 50 years. Among these adults, there are 26 males (38.23 % of all adults), 26 females (38.23 % of all adults), and 16 adults are of undetermined sex (23.52 %, 16/68).

Nevertheless, among the skeletons with fair to poor preservation there are 17 skulls which are complete enough to study. Of these, eight individuals are from the Early Period (2,100 to 900 BC), three males and five females. Nine other individuals are from the Late Period (300 BC to 300 AD), eight males and one female. No well-preserved skulls datable to the Middle Period were unearthed from the 2003–2004 excavation at Wat Phosi Nai cemetery.

### 5.1 Craniometric Data

This section presents the measurements and indices from 17 adult skulls from the 2003–2004 Ban Chiang series, including size and shape characteristics as detailed in Tables 2 and 3. (Data tables are assembled at the end of the chapter.)

Eighteen measurements of the cranium and four measurements of the mandible are shown.

Average sizes and shapes of the 2003-2004 Ban Chiang series skulls have been classified categorically to represent standard ranges in the long tradition of anthropological measurements for human skeletons (Howells 1973; Martin & Saller 1957). The ratios of two or more measurements presented in Tables 4 and 5 are useful expressions of shape that can be used for making comparisons among groups of people (White 1991:292). The average male and female cranium of this series is medium.

## **5.2 Vault Shape**

Cranial vault shape is typically expressed by three indices: the cranial (or length-breadth), height-length, and height-breadth (Pietrusewsky & Douglas 2002). The cranial index (breadth/length X 100) in the average male in this series is mesocranial or medium (range from 69.9 to 85.2), while the females are predominantly mesocranial (range from 74.1 to 83.8). Cranial height expressed as a proportion of cranial length (height-length index) indicates that the average is hypsicrane or high cranium in both males (range from 77.8 to 80.3) and females (range from 69.6 to 80.8). Females, however, have slightly higher crania than males. In the height-breadth cranial index, both males and females have high crania, best classified as acrocrane; the male index ranges from 93.1 to 107.8, and the female index ranges from 87.5 to 107.1. These data suggest that most males and females are similar in the horizontal view, with medium crania in length and breadth. In the vertical view, males and females have mostly high crania. However, female crania overall are smaller than male crania. In addition, similarity in size and shape between male and female crania is affected by diet (Larsen 1997, 2000). This implies that prehistoric males and females at Ban Chiang consumed a similar range of foods.

## **5.3 Face Shape**

Descriptions of the relative shape of the facial skeleton include ratios such as upper facial and total facial (mandible included) indices. Both the upper and total facial indices show medium to high values in males while female facial indices are on the average medium. However, facial proportions are very similar in both sexes. Face morphology is also described by the proportions of the orbit, nasal aperture, and palate. As a rule, orbital shapes (orbital index) are broad or hypericonch in both sexes. The nasal aperture average of males is mesorrhine or medium, which is very similar to the female nasal aperture shape. Some males and females however have broad (chamaerrhine) or very broad (hyperchamaerrhine) nasal apertures. The palatal shape, derived from the external breadth compared to the external length of the palate, is broad in both sexes.

## **5.4 Mandible Shape**

The ramus index and the jugomandibular index are two indices which describe the size and shape of the mandible. In this sample, the jugomandibular index from both sexes suggests a slightly narrower mandible in females than in males. The average ramus index is slightly broader in the males than the females.

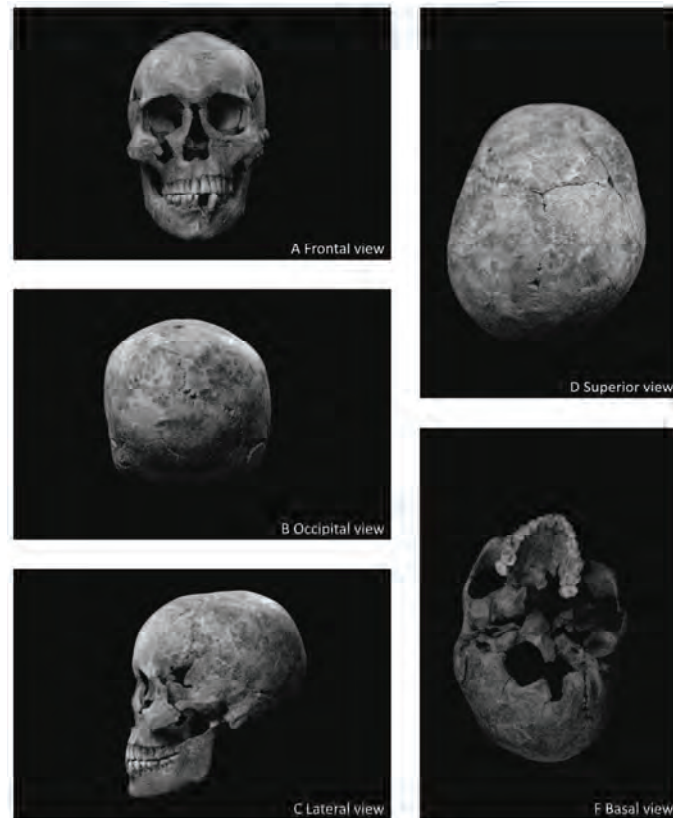
## **5.5 Skull Non-metric Data**

The non-metric characteristics of the BC 2003–2004 Ban Chiang skull series are presented below in Tables 6 and 7 for male and female adults.





**Figure 6:** Male skull of B # 076 from BC 2003 series.



**Figure 7:** Female skull of B # 091 from BC 2003 series.

Two of the most complete skulls, one male and one female, were selected to illustrate skull shapes. A male skull (Figure 6): *Frontal* view shows slightly marked supra-orbital ridges and well-marked robust zygomatics. The upper facial and nasal apertures are both medium in size and shape. The *occipital* view shows the arched form of the vault and the haus-form or the pentagonal shape of the vault. The *left lateral* view shows both slight supra-orbital ridges and prognathic upper face. The cranium is of medium height relative to both breadth and length. The mastoid processes are quite large. *Superior* view shows that the cranial vault is quite symmetrical with a sphenoid shape in average together with ellipsoid and ovoid shapes. The *basal* view indicates a broad palate and moderate to marked dental wear of the maxillary teeth, as well as well-marked shoveling in the upper central incisors.

Male mandible shown by the above pictures: the *frontal* view shows healed premortem tooth loss. The chin shows a bilobate structure and is angled inferiorly. The gonio-condylar index indicates a well-diverged ramus. The *left and right lateral* views show the mandible with a single mental foramen on each side. The ramus index shows a relatively narrow or tall ramus in the low end of the male range. The coronoid process is higher than the mandibular condyle. The *occlusal* or *superior* view of the mandible (not shown here) shows moderate, even tooth wear. The jaw is non-rocker and is quite broad relative to its length.

Female skull (Figure 7): The *anterior* or *frontal* view exhibits a broad nasal aperture. The *occipital* view exhibits a strongly arched form of the vault. The *left lateral* view shows the smooth curvature of the frontal and occipital morphology typical of females in the sample. The upper face is prognathic; the zygomatics are quite broad and robust. The cranium is of medium height relative to both breadth and length. The mastoid process is quite small to medium in size. In the *superior* view, the cranial vault is quite asymmetrical with a sphenoid shape. On the *basal* view of the skull the maxillary molar teeth show moderate dental wear, well-marked shovel shaping of the upper incisors, and a broad palate. The superior and basilar views of the cranium, however, show plagiocephaly and a much skewed cranial base, a result of the reconstruction of the vault.

Female mandible shown by above pictures: the *anterior* view exhibits post mortem tooth break in the left central incisor. The chin has medial and bilateral points and is angled inferiorly. The gonio-condylar index reflects a quite divergent ramus. The left lateral view shows a single mental foramen and the mandible is not rocker-jawed. The ramus index reflects a relatively narrow ramus. In the *occlusal* view (not shown here), the right mandibular condyle is gone. There is moderate tooth wear and premortem tooth loss of the third molar on the right side.

## 6 Comparisons

Table 8 compares metrical measurements from seven different prehistoric skeletal sets from mainland Southeast Asia, including the 2003–2004 sample from Ban Chiang Wat Phosi Nai and the 1973–1974 sample from Ban Chiang (Pietrusewsky & Douglas 2002). Other samples are from the two sites of Ban Khok Khon (BKK, 98–99) (Boonlop 2003), and Ban Na Di (BND) (Wiriyaromp 1984). Both of these sites are in the same area as Ban Chiang, the Sakon Nakhon basin or upper part of the Khorat Plateau.

Cranial measurements suggest that most prehistoric inhabitants in the region, at least from Vietnam, Laos, and Thailand, have no significant difference in their cranial size. Specifically, six cranium measurements (maximum length, maximum breadth, maximum frontal breadth, minimum frontal breadth, biorbital breadth, and interorbital breadth) are not

dramatically different from other prehistoric populations in the Sakon Nakhon Basin of northeast Thailand and mainland Southeast Asia, as the data shows in Table 8.

In addition, the data are compared in Table 9 with modern Thai skull specimens in northeast Thailand from the Department of Anatomy, Faculty of Medicine, Khon Kaen University (KKU) (Boonlop 2003). These skeletal remains came from donated bodies used in medical anatomy courses. To produce a comparative sample, the authors selected 17 skulls similar in sex and age of death to the prehistoric Wat Phosi Nai collection as representative of modern humans in the northeast region.

In Table 9 and Table 10, the prehistoric skulls of BC 2003 (WPSN) are compared with the modern northeast Thai skulls in twenty-two traits for both males and females. The tables indicate that the prehistoric skulls and modern skulls are approximately the same size in both males and females. Some metrical traits in the female skulls, however, suggest a difference between the two female series. These traits, for example maximum cranial length, maximum cranial breadth, maximum frontal breadth, and bimaxillary breadth, show that the female skulls from the prehistoric period are larger than the modern skulls. This data is similar to the results from previous comparative studies between the prehistoric specimens from Ban Khok Khon and modern skulls in northeast Thailand (Boonlop 2003).

## **7 Non-metrical dental traits**

‘Shovel-shaped teeth’ is a term that refers to a condition where the upper or lower incisors have raised ridges along the sides of the inner surface. Because of the general resemblance of this crown form to a shovel, the terminology ‘shovel-shaped teeth’ was proposed to describe this diagnostic tooth shape of East Asian populations (Hrdlicka 1920). Scott & Turner (1997) used the varying expressions of this trait to divide modern East Asian populations into two groups, the Northern and the Southern Mongoloid. The Northern Mongoloid group (Sinodonts) includes the Chinese, Tibetans and other inhabitants of northern and eastern Asia, and is characterised by a very high (60%–90%) incidence of advanced shovel-shaping; the Southern (Sundadont) group, including Southeast Asians, Polynesians, and Micronesians, have a moderate incidence of the trait (20%–50%). Pietrusewsky and Douglas’s analysis of their Ban Chiang sample indicated that the incidence of advanced shovel-shaping for all incisors was 56.8% (Pietrusewsky & Douglas 2002:52).

Other non-metrical traits used to differentiate populations include winging (rotated incisors) and Carabelli’s cusp, an additional cusp on the tongue side of the maxillary molars. Pietrusewsky & Douglas’s sample showed an overall incidence of winging of 6.1% and of Carabelli’s cusp at 7.0%, consistent with other Pacific and East Asian populations.

## **8 Conclusions**

The first aim of this study was to examine a recently excavated skeletal sample from the important site of Ban Chiang in northern northeast Thailand to see how the metric and non-metric traits in this sample compared with analysed samples from nearby sites in Thailand, and then with samples from adjacent areas in Laos and Vietnam.

Our study has shown that the prehistoric inhabitants from Ban Chiang Cultural Tradition villages and modern Thai in northeast Thailand have similarities of skull morphology with other people in adjacent areas, especially in Laos, Vietnam

(Pietrusewsky 2006a:59–90) and Cambodia (Pietrusewsky 2006b:86–95). Similarities include the percentages of the incidence of shovel-shaped incisors, a distinctive non-metric trait of the East Asian population. Moreover, the preliminary result of comparisons between the BC 2003–2004 human skull series and the other series indicates that the prehistoric population of Ban Chiang as revealed in Wat Phosi Nai were similar in cranial size and tooth shape to the early inhabitants of neighboring prehistoric agricultural villages, such as Ban Na Di and Ban Khok Khon, as well as to the inhabitants of Ban Chiang buried at other portions of the site.

The second aim of the study was to compare this Ban Chiang sample with a comparative sample of modern skulls from the same region of Thailand. It was hoped that this comparison would offer some data to help resolve the ongoing controversy as to whether many or most of the population traits in Southeast Asia were brought in by migrants from southern China or belonged to populations long-established in the region. In terms of regional comparisons, the dental morphology of the Northern Mongoloid, or Sinodont, is characterised by very high percentages of shovel-shaped or double shovel-shaped incisors and winged incisors. The Sundadont, or Southern Mongoloid dental complex, show little winging and lower percentages of shovel or double shovel-shaped incisors (Scott & Turner 1997:177–187). The teeth analysed in the Ban Chiang WPN 2003 sample, along with the contemporaneous series from other sites and the modern NE Thailand sample, place the Ban Chiang Wat Phosi Nai sample in the Sundadont category (Boonlop 2003; Matsumura et al 2010; Pietrusewsky & Douglas 2002).

The non-metrical data of crania, teeth, and mandibles indicate that many traits which were present in the prehistoric period continue into the present. Craniometric measures and indices indicate that the skulls in this series are of moderate dimensions comparable to other collections from this area. As a rule, though, the prehistoric skulls from the BC 2003 (WPSN) series are a bit larger than the modern NE Thailand skulls, especially in the breadth and length of the orbital area and cheek bones (Boonlop 2003:319–324). The average male and female skulls show diversity of form, with both ovoid and sphenoid shapes in the superior view, and arched or pentagonal shape of the vault in the occipital view. Prognathic upper facial regions are a long-standing trait. The zygomatic arches are broad and robust, especially in the male, and the nasal apertures are moderate to broad in shape and size. Upper central incisors are marked by shovel shaping, and some characteristics such as extra cusps and Carabelli's cusp are also observed in the upper molar teeth. All of these traits are continuous from past to present.

The study of prehistoric human skulls uncovered from Ban Chiang in 2003–2004 in the Wat Phosi Nai area should help in suggesting alternative explanations for the morphological diversity in variation of the mainland Southeast Asia region, past and present. As Pietrusewsky (2006a) explains, the peoples and cultures of this region are more influenced by its past population than by current geopolitical boundaries. The people of the region, especially in prehistory, are often assumed to be a southern branch of Mongoloid or eastern Asian groups. However, in recent years, new archaeological and linguistic perspectives of the prehistory of mainland Southeast Asia and East Asia have emerged. One idea that has been proposed strongly by some archaeologists and linguists is that the development of agriculture (especially rice domestication) and the dispersal of languages most likely diffused from southern China in an agricultural colonisation model (Higham 1996; Higham & Thosarat 1998; Pietrusewsky 2006a). In contrast to this model of incoming agriculturalists expanding into Southeast Asia, some physical anthropologists have proposed the Population Continuity Model, which argues that the current Southeast Asia inhabitants evolved from earlier groups settled in this region since at least the middle

to late Holocene onward (Turner 1990). The study of prehistoric skulls from the most recent archaeological excavation at Ban Chiang in 2003–2004 has provided support to the latter model as presented in Pietrusewsky & Douglas (2002:234–235, 256–257), as it helps to demonstrate the differences between the characteristic Southeast Asia Sundadont dental complex and the characteristic Chinese Sinodont complex.

This is only a preliminary study. Further work on cranial and dental pathologies and dental metric traits is planned, along with a more advanced comparison with other skeletal samples in Southeast Asia and southern China.

### **Acknowledgements**

A number of institutions and individuals have supported the research on the BC 2003–2004 human skeleton series that made this chapter possible. First of all, we would like to thank the Princess Maha Chakri Sirindhorn Anthropology Centre, Bangkok (SAC), and Dr. Paritta C. Koanantakool, the Director of SAC. We also thank the Cultural Management Programme, College of Innovation, Thammasat University; Ms. Ajara Khaengsarikij, former Head of Ban Chiang National Museum, Thai Fine Arts Department; and Prof. Sanjai Saengvichien (supervisor) and Mr. Naruepol Wangtochchaicharoen, (assistant researcher) of the Thai Physical Anthropology and Bioarchaeology Research Project. We owe thanks to Dr. Elizabeth Hamilton (U Penn Museum) for her advice and English-language editing. We are grateful to Ardeth Abrams (Ban Chiang Project, U Penn Museum) for her help in the map illustration. Prof. Michael Pietrusewsky of the Department of Anthropology, the University of Hawaii at Manoa, USA, and Dr. Siân Halcrow of the Department of Anatomy and Structural Biology, University of Otago, New Zealand, and the anonymous reviewer are also given deep thanks for their comments and suggestions. Finally, we give our thanks to Prof. Nick Enfield (Max Planck Institute) and Dr. Joyce White (U Penn Museum) for their support and encouragement.

**Table 1:** Distribution of the BC 2003–2004 (WPSN) burials by age and sex (listed by burial number)

	Male	Female	Sex Undetermined	Total (N <sup>a</sup> )
0-4.9			13, 21, 28, 29, 36, 41, 42, 45, 53, 55, 56, 59, 61, 63, 65, 67, 71, 73, 74, 78, 80, 82, 85, 88, 95, 96, 98, 103, 107, 113, 114	31
5-9.9			8, 40, 43, 92, 99, 101,	6
10-14.9			32, 49, 108, 116	4
15-19.9	2, 18	18, 77, 97,	84, 109	7
20-24.9	3, 76, 83,	75, 89, 93, 94,	14, 19, 31, 47, 58, 66, 86, 105, 110, 111	17
25-29.9	16	72	46, 51, 62	5
30-34.9	4, 11, 20, 34, 104	57, 115		7
35-39.9	10, 23, 24, 25, 26, 68, 100, 106	37, 44, 64, 91,	1, 38	14
40-44.9		6		1
45-49.9	9, 22, 60, 70, 79	15, 27, 33, 48, 50, 54, 69, 81, 102, 112		15
50-54.9	17			1
55-60+		90		1
Total (N <sup>a</sup> )	25	26	58	109

N<sup>a</sup> = number of individuals**Table 2:** Means and standard deviations for skull measurements in male adults of BC 2003-2004 (WPSN)

Measurement (mm)	Male (N <sup>b</sup> = 11)			
	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range
Maximum cranial length, M-1 / H-GOL	6	177.3	10.7	167–193
Maximum cranial breadth, M-8 / H-XCD	9	137.3	8.5	127–150
Minimum frontal breadth, M-9 / H-FMB	7	95.8	6.6	87–109
Maximum frontal breadth, M-10 / H-XFB	7	114	7.1	107–128
Basion-bregma height, M-17 / H-BBH	4	138.2	8.9	130–151
Biorbital breadth, H-EKB	3	97.3	5.5	92–103
Bizygomatic breadth, M-45 / H-ZYB	4	126.2	6.2	118–132
Bimaxillary breadth, M-46 / H-ZMB	3	104	12	92–116
Facial height (Nasion-gnathion), M-47	4	118	2.8	114–120
Upper facial height (Nasion-alveolare), M-48	4	66.5	11.6	49–73
Posterior interorbital breadth, M-49	2	32	0.0	32
Anterior interorbital breadth, M-50	4	25.7	2.6	22–28
Orbital breadth (left), M-51a	4	39.5	3.1	35–42
Orbital height, M-52	4	36.2	1.7	34–38

<b>Measurement (mm)</b>	<b>Male (N<sup>b</sup> = 11)</b>			
	<b>n<sup>c</sup></b>	<b>Mean</b>	<b>SD<sup>d</sup></b>	<b>Range</b>
Nasal breadth, M-54 / H-NLB	6	27.6	2.5	25–32
Nasal height, M-55 / H-NLH	4	51.2	2.2	48–53
Palatal length, M-62	3	49	4.3	46–54
Palatal breadth, M-61 / H-MAB	4	42.5	3.1	38–45
Bigonial breadth, M-66	6	104.5	8.8	92–114
Mental foramen height, M-69(1)	10	32.6	4.1	26–40
Height of mandibular ramus, M-70	4	59.5	6.8	50–66
Breadth of mandibular ramus, M-71	6	34.5	2.1	32–38

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation.

**Table 3:** Means and standard deviations for skull measurements in female adults of BC 2003–2004 (WPSN)

<b>Measurement (mm)</b>	<b>Female (N<sup>b</sup> = 6)</b>			
	<b>n<sup>c</sup></b>	<b>Mean</b>	<b>SD<sup>d</sup></b>	<b>Range</b>
Maximum cranial length, M-1 / H-GOL	6	173	4.8	167–181
Maximum cranial breadth, M-8 / H-XCD	6	137.5	6.0	126–144
Minimum frontal breadth, M-9 / H-FMB	6	95.1	5.4	87–100
Maximum frontal breadth, M-10 / H-XFB	6	108	8.7	100–122
Basion-bregma height, M-17 / H-BBH	6	133.6	6.3	126–141
Biorbital breadth, H-EKB	4	95.5	1.2	94–97
Bizygomatic breadth, M-45 / H-ZYB	4	125.2	4.2	119–128
Bimaxillary breadth, M-46 / H-ZMB	5	105.6	2.8	102–109
Facial height (Nasion-gnathion), M-47	5	111.2	5.8	105–120
Upper facial height (Nasion-alveolare), M-48	5	70.4	2.8	68–75
Posterior interorbital breadth, M-49	6	28	2.9	25–33
Anterior interorbital breadth, M-50	6	24.3	3.0	21–29
Orbital breadth (left), M-51a	4	40.2	0.9	39–41
Orbital height, M-52	5	37	1.4	36–39
Nasal breadth, M-54 / H-NLB	6	27	2.9	25–33
Nasal height, M-55 / H-NLH	5	49.2	4.4	43–53
Palatal length, M-62	5	50.6	1.5	49–52
Palatal breadth, M-61 / H-MAB	6	42.8	3.4	39–49
Bigonial breadth, M-66	5	96.4	2.4	93–99

Measurement (mm)	Female (N <sup>b</sup> = 6)			
	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range
Mental foramen height, M-69(1)	5	27.6	3.0	24–31
Height of mandibular ramus, M-70	5	53	4.2	50–60
Breadth of mandibular ramus, M-71	5	34.8	1.9	32–37

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation.

**Table 4:** Means and standard deviations for skull indices in male adults of BC 2003–2004 (WPSN)

Indices	Male (N <sup>b</sup> = 11)				Average type	Range
	n <sup>c</sup>	Mean	SD <sup>d</sup>			
Neurocranium						
Cranial or Length-breadth index	6	77.2	5.8	Mesocranial	69.9–85.2	
Length-height index	4	78.8	1.2	Hypsicranial	77.8–80.3	
Breadth-height index	4	99.4	6.7	Acrocranial	93.1–107.8	
Transverse frontal index	7	84.2	6.4		74.2–92.3	
Transverse fronto-parietal index	7	70.2	3.6		65.5–75.5	
Transverse craniofacial index	4	94.0	5.7		89.1–102.3	
Mean height index	4	155.0	7.9	High cranial	147.5–165.6	
Face						
Facial index	4	93.5	3.3	Leptoprosopic	89.3–96.6	
Orbital index	4	91.9	3.9	Hypericonch	87.8–97.1	
Inter orbital index	2	27.7	0.7		27.1–28.2	
Nasal index	4	53.2	6.2	Chamaerrhine	47.1–61.5	
Maxilla						
Palatal index	3	87.6	7.1	Brachystaphyline	82.6–95.7	
Mandible						
Ramus index	4	60.2	11.2		50–76	
Jugomandibular index	3	79.1	6.4		73.4–86.1	

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation



**Table 5:** Means and standard deviations for skull indices in female adults of BC 2003–2004 (WPSN)

Indices	Female (N <sup>b</sup> = 6)			Average type	Range
	n <sup>c</sup>	Mean	SD <sup>d</sup>		
Neurocranium					
Cranial or Length-breadth index	6	79.4	3.1	Mesocranial	74.1–83.8
Length-height index	6	77.3	4.6	Hypsicranial	69.6–80.8
Breadth-height index	6	97.4	7.1	Acrocranial	87.5–107.1
Transverse frontal index	6	88.4	6.8		81.9–98
Transverse fronto-parietal index	6	69.2	2.5		65.2–72.4
Transverse craniofacial index	4	91.5	7.1		82.6–100
Mean height index	6	153.1	5.0	High cranial	144.7–160.4
Face					
Facial index	4	90.0	3.0	Leptoproscopic	86.5–93.7
Orbital index	4	92.6	5.5	Hypericonch	87.8–97.5
Inter orbital index	4	24.0	2.6		21.6–27.3
Nasal index	5	55.8	5.7	Chamaerrhine	49.0–62.2
Maxilla					
Palatal index	5	84.6	8.8	Mesostaphyline	78.8–100
Mandible					
Ramus index	5	65.9	6.2		58.3–74
Jugomandibular index	4	76.4	2.5		73.8–79.8

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation

**Table 6:** Non-metric data for skulls in male and female adults of BC 2003–2004 (WPSN)

Traits	Male		Female	
	Present	N = 11	Present	N = 6
Shape of Skull				
<u>Norma verticalis</u>				
Ellipsoid	√	1 (present in this many skulls)		
Ovoid	√	1		
Sphenoid	√	6	√	6
<u>Norma occipitalis</u>				
Arch	√	6	√	6
Haus	√	3		
Suture				

Traits	Male		Female	
	Present	N = 11	Present	N = 6
Metopic suture	√	2		
Wormain	√	3	√	3
Inca			√	1
Foramen / Notch				
Parietal foramen	√	2		
<u>Supraorbital</u>				
Notch-Single			√	2
Notch - Double	√	1	√	4
Foramen - Single	√	1	√	1
Foramen - Double	√	3	√	1
<u>Infraorbital foramen</u>				
Single	√	4	√	5
Piriform aperture				
Anthropine	√	2	√	3
Paranasal fossa			√	1
Paranasal sulcus	√	5	√	2
Clinocephalia	√	1	√	2
Pterion				
H-Shape	√	6	√	5
Mandible				
Mylohyoid arch/bridge	-	-	-	-
Mandibular torus	-	-	-	-
Mental foramen (single)	√	11	√	6

**Table 7:** Non-metric data for teeth in male and female adults of BC 2003-2004 (WPSN)

Traits	Male		Female		Unidentified	
	Present	N = 16	Present	N = 13	Present	N = 4
Teeth						
Shovel shape ( I )	√	2	√	4	-	-
Groove cingulum ( I )	√	2	√	1	-	-
Carabelli's cusp ( UM )	√	1	-	-	-	-
Protostylid cusp ( LM )	-	-	√	1	-	-

**Table 8:** Means (in mm) and Standard Deviations for 6 cranial measurements recorded in seven prehistoric male sample series in Mainland SEA

Measurement <sup>a</sup>	Con Co Ngua (n = 5) <sup>b</sup>		NE Vietnam (n = 7) <sup>b</sup>		Laos (n = 5) <sup>b</sup>		BC 2003-2004 (WPSN) (n = 11) <sup>b</sup>		Ban Chiang (BC-BCES) (n = 10) <sup>b</sup>		Ban Na Di (n = 3) <sup>b</sup>		Ban Khok Khon (n = 1) <sup>b</sup>	
	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>	Mean	SD <sup>c</sup>
MAXCRANL	187.2	6.0	188.6	6.2	182.4	3.5	177.3	10.7	182.7	8.6	179.7	9.9	188.0	0.0
MAXCRANB	135.8	7.6	133.4	6.0	138.6	4.1	137.3	8.5	139.8	9.0	135.3	3.2	140.0	0.0
MAXFRONB	118.2	9.8	115.7	5.0	118.6	4.0	114.0	7.1	118.8	4.7	117.3	2.9	108.0	0.0
MINFRONB	100.2	7.3	96.0	2.6	95.0	2.1	95.8	6.6	96.8	5.0	99.0	1.0	91.0	0.0
BIORBITB	99.2	6.9	102.6	4.2	98.8	3.3	97.3	5.5	98.9	4.1	97.7	0.6	99.0	0.0
INTERORB	30.2	2.9	29.0	1.4	30.8	1.6	32.0	0.0	26.8	3.0	27.7	2.1	29.0	0.0

**Note:** Table adapted from Pietrusewsky & Douglas (2002:231)

<sup>a</sup> The measurement sources are as follows (a capital letter is followed by the name of the measurement within that source; M = Martin & Saller 1957, H = Howells 1973); MAXCRANL = Maximum cranial length (M-1), MAXCRANB = Maximum cranial Breadth (M-8), MAXFRONB = Maximum frontal breadth (M-10), MINFRONB = Minimum frontal breadth (M-9), BIORBITB = Biorbital breadth (H-EKB), INTERORB = Interorbital breadth (M-49).

<sup>b</sup> n = number of crania.

<sup>c</sup> SD = standard deviation

**Table 9:** Means and standard deviations for skull measurements in male adults of BC 2003-2004 (WPSN) and modern NE Thai at KKU

Measurement (mm) / Indexa	BC 2003 (WPSN) (N <sup>b</sup> = 11)				Modern NE Thai at KKU (N <sup>b</sup> = 9)			
	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range
Maximum cranial length, M-1 ** / H-GOL ***	6	177.3	10.7	167–193	9	166.3	7.1	157–178
Maximum cranial breadth, M-8 / H-XCD	9	137.3	8.5	127–150	9	142.2	7.3	129–154
Minimum frontal breadth, M-9 / H-FMB	7	95.8	6.6	87–109	9	100.7	3.7	95–105
Maximum frontal breadth, M-10 / H-XFB	7	114	7.1	107–128	9	92.2	4.1	84–99
Basion-bregma height, M-17 / H-BBH	4	138.2	8.9	130–151	9	135.6	4.0	130–140
Biorbital breadth, H-EKB	3	97.3	5.5	92–103	9	95.0	1.9	92–98
Bizygomatic breadth, M-45 / H-ZYB	4	126.2	6.2	118–132	9	134.2	2.1	132–138
Bimaxillary breadth, M-46 / H-ZMB	3	104	12	92–116	9	101.2	3.7	93–105
Facial height (Nasion- gnathion), M-47	4	118	2.8	114–120	7	119.0	6.7	110–127
Upper facial height (Nasion- alveolare), M-48	4	66.5	11.6	49–73	8	71.1	4.1	66–77
Posterior interorbital breadth, M-49	2	32	0.0	32	9	27.4	3.0	22–31
Anterior interorbital breadth, M-50	4	25.7	2.6	22–28	9	22.5	1.6	21–25
Orbital breadth (left), M-51a	4	36.2	1.7	34–38	9	33.7	1.8	31–36
Orbital height, M-52	6	27.6	2.5	25–32	8	27.5	2.5	22–30
Nasal breadth, M-54 / H- NLB	4	51.2	2.2	48–53	8	52.3	4.5	46–59
Nasal height, M-55 / H- NLH	3	49	4.3	46–54	8	45.2	2.1	42–49
Palatal length, M-62	4	42.5	3.1	38–45	8	42.0	2.9	36–46
Palatal breadth, M-61 / H-MAB	6	104.5	8.8	92–114	8	100.8	2.94	98–107
Bigonial breadth, M-66	10	32.6	4.1	26–40	8	31.9	2.4	27–35
Mental foramen height, M-69(1)	4	59.5	6.8	50–66	8	62.3	5.9	52–69
Height of mandibular ramus, M-70	6	34.5	2.1	32–38	8	33.8	2.5	31–39
Breadth of mandibular ramus, M-71	4	36.2	1.7	34–38	9	33.7	1.8	31–36

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation

**Table 10:** Means and standard deviations for skull measurements in female adults of BC 2003–2004 (WPSN) and modern NE Thai at KKU

Measurement (mm) / Index <sup>a</sup>	BC 2003 (WPSN) (N <sup>b</sup> = 6)				Modern NE Thai at KKU (N <sup>b</sup> = 6)			
	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range	n <sup>c</sup>	Mean	SD <sup>d</sup>	Range
Maximum cranial length, M-1 ** / H-GOL ***	6	173.0	4.8	167–181	6	160.1	7.3	150–168
Maximum cranial breadth, M-8 / H-XCD	6	137.5	6.0	126–144	6	142.1	4.9	136–149
Minimum frontal breadth, M-9 / H-FMB	6	95.1	5.4	87–100	6	92.3	3.8	86–97
Maximum frontal breadth, M-10 / H-XFB	6	108.0	8.7	100–122	6	96.5	4.0	90–101
Basion-bregma height, M-17 / H-BBH	6	133.6	6.3	126–141	6	131.8	2.4	130–135
Biorbital breadth, H-EKB	4	95.5	1.2	94–97	6	94.5	3.6	90–99
Bizygomatic breadth, M-45 / H-ZYB	4	125.2	4.2	119–128	6	128.1	4.0	122–132
Bimaxillary breadth, M-46 / H-ZMB	5	105.6	2.8	102–109	6	95.1	4.1	92–102
Facial height (Nasion-gnathion), M-47	5	111.2	5.89	105–120	5	112.0	4.6	106–118
Upper facial height (Nasion-alveolare), M-48	5	70.4	2.88	68–75	6	65.5	3.0	61–69
Posterior interorbital breadth, M-49	6	28.0	2.9	25–33	6	27.1	0.9	26–29
Anterior interorbital breadth, M-50	6	24.3	3.0	21–29	6	22.8	0.7	22–24
Orbital breadth (left), M-51a	4	40.2	0.9	39–41	6	39.3	1.7	37–42
Orbital height, M-52	5	37.0	1.4	36–39	6	32.7	1.7	31–36
Nasal breadth, M-54 / H-NLB	6	27.0	2.9	25–33	6	28.8	4.3	24–36
Nasal height, M-55 / H-NLH	5	49.2	4.4	43–53	6	48.4	2.4	44–51
Palatal length, M-62	5	50.6	1.5	49–52	6	44.5	3.2	41–50
Palatal breadth, M-61 / H-MAB	6	42.8	3.4	39–49	6	38.1	3.3	33–42
Bigonial breadth, M-66	5	96.4	2.4	93–99	5	95.2	3.2	90–98
Mental foramen height, M-69(1)	5	27.6	3.0	24–31	5	31.0	2.5	28–35
Height of mandibular ramus, M-70	5	53.0	4.2	50–60	5	55.6	3.3	52–59
Breadth of mandibular ramus, M-71	5	34.8	1.9	32–37	5	30.4	1.9	29–33

<sup>a</sup> The measurement source is given with a capital letter followed by the name or number of the measurement in that source, if available. H = Howells 1973; M = Martin & Saller 1957

<sup>b</sup> = number of adults (> 18 years)

<sup>c</sup> = number of measurements from the BC 2003–2004 (WPSN) series.

<sup>d</sup> = standard deviation

**Table 11:** Selected non-metric observations in BC 2003 (WPSN) permanent dentitions in male and female

Trait		Male, N <sup>a</sup> = 16			Female, N <sup>a</sup> = 13		
Tooth	Variant <sup>b</sup>	A	/O <sup>c</sup>	%	A	/O	%
Shovel-shaped incisors							
Maxillary central							
	absent	7	/8	87.5	4	/6	66.6
	slight	1	/8	12.5	1	/6	16.6
	moderate	0	/8	0.0	1	/6	16.6
Maxillary lateral							
	absent	7	/8	87.5	4	/6	66.6
	slight	1	/8	12.5	1	/6	16.6
	moderate	0	/8	0.0	1	/6	16.6
Mandibular central							
	absent	9	/13	69.2	10	/11	91.0
	slight	4	/13	30.8	1	/11	9.0
Mandibular lateral							
	absent	9	/13	69.2	10	/11	91.0
	slight	4	/13	30.8	1	/11	9.0
Maxillary winging							
	present	0	/9	0.0	0	/10	0.0
Carabelli's cusp							
Maxillary 1st molar							
	present	1	/9	11.1	0	/8	0.0
Maxillary 2nd molar							
	present	0	/9	0.0	0	/8	0.0

<sup>a</sup> Total number<sup>b</sup> Presence versus absence unless other variation noted<sup>c</sup> A = affected teeth, O = observed teeth

## References

- Agelarakis, Anagnostis. 1997. Some Reconstructions of Human Bio-Cultural Conditions during the 3<sup>rd</sup> and 2<sup>nd</sup> millennia B.C. in South-East Asia. In Roberto Ciarla & Fiorella Rispoli (eds) *South-East Asian Archaeology 1992*, 99–118. Proceedings of the Fourth International Conference of the European Association of South-East Asian Archaeologists. Rome: Herder International Book center.
- Bannanurag, Rachanie & Anat Bamrungwong. 1991. A survey of Ban Chiang Cultural Tradition Sites. *Silpakorn* 34(5). 38–59. (In Thai)
- Bannanurag, Rachanie et al. 1992. *A Report of Ban Chiang Cultural Tradition Sites Survey. Vol. I. The Ban Chiang Cultural Tradition Research Project*. Bangkok: Division of Archaeology, Fine Arts Department. (In Thai)
- Boonlop, Korakot. 2003. *An Analysis of Prehistoric Human Skeletal Remains of Ban Chiang Cultural Tradition from Ban Khok Khon site, Sakon Nakhon province*. MA Thesis, Department of Archaeology, Graduate School, Silpakorn University, Bangkok, Thailand.
- Boonlop, Korakot & Kamoltip Brown. 2008. Prehistoric people of Ban Chiang Cultural Tradition from Ban Khok Khon, Sakon Nakhon province. *Proceedings of the 31<sup>st</sup> Annual Conference of the Anatomy Association of Thailand*, 28–34. Bangkok: Kasetsart University.
- Boonprakorb, Yodchai. 1993. The Northeastern Thai Skull: Physical Anthropology Studies. MSc Thesis, Department of Anatomy, Graduate School, Khon Kaen University, Khon Kaen, Thailand.
- Brothwell, Don R. 1981. *Digging up Bones: The excavation, treatment and study of human skeletal remains*. New York: Cornell University Press.
- Choosiri, Prapit. 1991. *Introduction to Human Osteology*. Bangkok: P. Sampanpanit. (In Thai)
- 1992. *A brief history of human skeleton study in Thailand*. Bangkok: Division of Archaeology, Fine Arts Department. (In Thai)
- 1994. *A preliminary report on human skeleton analysis from archaeological sites at Ban Tum and Ban Daeng Yai, Khon Kaen province*. Bangkok: Division of Archaeology, Fine Arts Department. (In Thai)
- Douglas, Michele T. 1996. *Palaeopathology in Human Skeletal Remains from the Pre-Metal, Bronze and Iron Ages, Northeastern Thailand*. PhD Dissertation, Department of Anthropology, Graduate Division, University of Hawaii at Manoa, USA.
- 1997. A Preliminary Discussion of Trauma in the Human Skeletons from Ban Chiang, Northeast Thailand. *Indo-Pacific Prehistory Association Bulletin* 16 (3). 111–117.
- Douglas, Michele T. & Michael Pietrusewsky. 2007. Biological consequences of sedentism and agricultural intensification in northeast Thailand. In Marc N. Cohen & Gillian Crane-Kramer (eds) *Ancient Health: Skeletal Indicators of Agricultural and Economic Intensification (Bioarchaeological Interpretations of the Human Past: Local, Regional, and Global Perspectives)*, 300–319. Gainesville: University of Florida Press.
- Gorman, Chester & Pisit Charoenwongsa. 1976. Ban Chiang: A Mosaic of Impression from the First Two Years. *Expedition* 18(4). 14–26.
- Higham, Charles. 1996. *The Bronze Age of Mainland Southeast Asia*. Cambridge: Cambridge University Press.
- Higham, Charles & Rachanie Thosarat. 1998. *Prehistoric Thailand, from early settlement to Sukhothai*. Bangkok: River Books.

- Hillson, Simon. 1990. Teeth. In D. Brothwell et al. (eds) *Cambridge Manuals in Archaeology*. Cambridge: Cambridge University Press.
- 1992. *Mammal Bones and Teeth. An Introduction Guide to Methods of Identification*. London: University College London.
- 1996. *Dental Anthropology*. Cambridge: Cambridge University Press.
- Howells, William W. 1973. Cranial Variation in Man: A Study by Multivariate Analysis of Patterns of Difference Among Recent Human Populations. *Papers of the Peabody Museum of Archaeology and Ethnology* Vol.67. Cambridge: Harvard University.
- Hrdlicka, Ales. 1920. Shovel-shaped teeth. *American Journal of Physical Anthropology*. Vol. 20. 489–490.
- Hutangkura, Trongjai. 1997. *A preliminary report on human skeleton analysis from archaeological site at Ban Don Thong Chai, Sakon Nakhon province*. Khon Kaen: The 7<sup>th</sup> Regional Office of Archaeology and National Museums. (In Thai)
- Kae-In, Nawarat. 1999. *An Analysis of Human Skeletons and the Mortuary Practice at Prasat Phanom Wan, Nakhon Ratchasima province*. MA Thesis, Department of Archaeology, Graduate School, Silpakorn University, Bangkok, Thailand.
- Kaewsuwan, Siam. 2003. *Estimation of Age at Death for children skeleton from teeth: A case study of the 2003 Wat Po Si Nai Human Skeleton Series, Ban Chiang, Udon Thani*. BA Thesis, Department of Archaeology, Silpakorn University, Bangkok, Thailand.
- Krogman, Wilton M. 1962. *The Human Skeleton in Forensic Medicine*. Springfield: Charles C. Thomas Publisher.
- Laughlin, W. S. & J. B. Jorgensen. 1956. Isolate variation in Greenland Eskimo. *Acta Jenet* 6. 3–12.
- Larsen, Clark S. 1997. *Bioarchaeology: Interpreting behavior from the human skeleton*. Cambridge: Cambridge University Press.
- 2000. *Skeletons in our closet : Revealing our past through bioarchaeology*. Princeton: Princeton University Press.
- Matsumura, Hirofumi, Kathryn M. Domett & Dougald J.W. O'Reilly. 2010. On the origin of pre-Angkorian peoples: perspectives from cranial and dental affinity of the human remains from Iron Age Phum Snay, Cambodia. *Anthropological Science*, Vol 118. (published online)
- Martin, Rudolf & Karl Saller. 1957. *Lehrbuch der Anthropologie*. Band 1, 3rd editon. Stuttgart: Gustav Fischer.
- Nelsen, Kirsten, Nancy Tayles & Kathryn M. Domett. 2001. Missing lateral incisors in Iron Age South-East Asians as possible indicators of dental agenesis. *Archives of Oral Biology* 46. 963–971.
- Pietrusewsky, Micheal. 1974. The Palaeodemography of a prehistoric Thai population: Non Nok Tha. *Asian Perspectives* XVI(2). 126–140.
- 1982. The ancient inhabitants of Ban Chiang: The evidence from the human skeletal and dental remains. *Expedition* 24(4). 42–50.
- 1984. Pioneers on the Khorat Plateau: The prehistoric inhabitants of Ban Chiang. *Journal of the Hong Kong Archaeological Society* X, 1982–1983. 90–106.
- 1997. The people of Ban Chiang: An early Bronze site in northeast Thailand. *Bulletin of the Indo-Pacific Prehistory Association* 16. 119–148.



- 2006a. A multivariate craniometric study of the prehistoric and modern inhabitants of Southeast Asia, East Asia and surrounding regions: a human kaleidoscope? In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 59–90. Cambridge: Cambridge University Press.
- 2006b. The bioarchaeology of the Vat Komnou cemetery, Angkor Borei, Cambodia. *Bulletin of the Indo-Pacific Prehistory Association* 26. 86–95.
- Pietruszewsky, Michael & Michele T. Douglas. 2002. Ban Chiang, A Prehistoric Village Site in Northeast Thailand. I: The Human Skeletal Remains. *University Museum Monograph* 111. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
- Pureepatpong, S. 1996. *An Analysis of Human Skeletons from Ban Mai Chai Mongkhol, Nakhon Sawan Province: A case of S18 W22*. BA Thesis, Department of Archaeology, Silpakorn University, Bangkok, Thailand. (In Thai)
- Sangvichien, Sanjai. 1970. *Thai Skull: Physical Anthropology Studies*. A dissertation for the degree of Doctor of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University.
- 1966. A preliminary report on non-metrical characteristic of Neolithic skeletons found at Ban Kao, Kanchanaburi. *The Journal of the Siam Society* Vol. IV, Part I. 3–18.
- Sangvichien, Sanjai & Wattana Supawan. 1977. Trace of Trephining found on human skull. In Sumit Pitipat & Preecha Kanchanagama (eds) *Ban Thatu and Ban Om Kaew: A Preliminary report of excavations at Ban Chiang and adjacent area in Udon Thani province*. Bangkok: Thammasat University. (In Thai)
- Scott, G. Richard & Christy G. Turner II. 1997. *The anthropology of modern human teeth: dental morphology and its variation on recent human populations*. Cambridge: Cambridge University Press.
- Sjovold, T. 1984. A Report on the Heritability of Some Cranial Measurement and Non-Metric Traits. In Gerrit N. van Vark & William W. Howells (eds) *Multivariate Statistical Methods in Physical Anthropology: A Review of Recent Advances and Current Developments*. D. Reidel, Dordrecht, the Netherlands.
- Stewart, T. D. 1940. Anthropometric nomenclature: II The indices of head height. *American Journal of Physical Anthropology* 29. 23–29.
- Tayles, Nancy, Kathryn M. Domett & Kirsten Nelsen. 2000. Agriculture and dental caries? The case of rice in prehistoric Southeast Asia. *World Archaeology* Vol. 32(1). 68–83.
- Turner, Christy G. II. 1990. Major features of sundadonty and sinodonty including suggestions about East Asian microevolution, population history, and late Pleistocene relationships with Australian Aborigines. *American Journal of Physical Anthropology*, Vol. 82. 295–317.
- Ubelaker, Douglas H. 1984. *Human Skeletal Remains: Excavation, Analysis, Interpretation*. Washington D.C.: Smithsonian Institution.
- White, Timothy D. 1991. *Human Osteology*. San Diego: Academic Press, Inc.
- White, Joyce C. 1982. *Ban Chiang: Discovery of a Lost Bronze Age*. The University Museum, University of Pennsylvania, Philadelphia, and The Smithsonian Institution Traveling Exhibition Service, Smithsonian Institution, Washington D.C.
- 1986. The Ban Chiang chronology revised. In Ian Glover & Emily Glover (eds) *Southeast Asian Archaeology 1986*, Proceedings of the First Conference of the Association of Southeast Asian Archaeologists in Western Europe, 8<sup>th</sup>–10<sup>th</sup> September, 121–130. London: Institute of Archaeology, University College of London.

- . 2008. Dating Early Bronze at Ban Chiang Thailand. In Jean-Pierre Pautreau, Anne-Sophie Coupey, Valery Zeitoun & Emma Rambault (eds) *From Homo erectus to the living traditions*, Proceedings the 11<sup>th</sup> International Conference of the European Association of Southeast Asian Archaeologists, 25<sup>th</sup>–29<sup>th</sup> September 2006, 91–104. Bougon.
- Wilder, Harris H. 1920. *A Laboratory manual of the anthropometry*. Philadelphia: P. Blakiston's son.
- Wiriyaromp, Warrachai. 1984. The Human Skeletal Remains from Ban Na Di, Patterns of Birth, health and death in prehistoric northeast Thailand. MA Thesis, Department of Anthropology, the University of Otago, New Zealand.



# *Part III*

*Intra-group diversity:  
the case of Aslian groups*



# 10 *Biological and cultural evolution in the population and culture history of Homo sapiens in Malaya*

---

DAVID BULBECK

## 1 Introduction

This chapter synthesises current information on the *longue durée* history of the Malay Peninsula south of the Isthmus of Kra, a region referred to here as ‘Malaya’ (Figure 1). Climatically Malaya is humid, and warm to hot except where the mountainous terrain of central Peninsular Malaysia provides for cool temperatures at elevated altitudes. From south to north the northeast monsoon weakens while the southwest monsoon strengthens, resulting in a concomitant shift from equatorial to monsoonal forest and mangroves. In Peninsular Malaysia and southernmost Thailand, Melayu Malays—the Islamic subjects of the Melayu kingdoms (Malay, *kerajaan Melayu*) at the time that Europeans entered the region (Benjamin 2002)—constitute the most numerous ethnic group (Hamilton 2006). However, they share the rural landscape with a medley of other groups collectively known as Orang Asli (Malay for ‘original people’).

The Orang Asli are highly diverse in their biology, linguistics and lifeways. Variation in overt appearance has resulted in scholarly recognition of three broad Orang Asli categories for over 100 years (for example Martin 1905; Carey 1976; Rashid 1995; Hooker 2003), as best described by Cole (1945). The three categories correspond to the Aboriginal Malays who resemble the Melayu in features such as brown skin and straight to slightly wavy hair, the Senoi who are distinguished by wavy hair, and the Semang ‘Negritos’ who have dark skin and woolly hair. This tripartite division of the Orang Asli has been confirmed by studies of their teeth (Rayner & Bulbeck 2001; Bulbeck et al. 2005) and mitochondrial DNA (Hill et al. 2006). Linguistically, the Orang Asli include both Austronesian and Austroasiatic (Mon-Khmer) speakers—respectively, groups speaking dialects of standard Malay, and groups whose languages together constitute the Aslian clade of Austroasiatic (Benjamin 2002). The most recent study of the Aslian languages (Burenhult et al., this volume) confirms the existence of Southern, Central and Northern branches, as found in earlier studies (for example Benjamin 1976), but also finds that one language, Jah Hut, may be sufficiently distinct to constitute a fourth branch. Moreover, the traditional lifeways of the Orang Asli include groups that have met their subsistence needs through swidden agriculture in the lowlands or the highlands, those that have also relied on fishing and foraged forest foods to supplement their farmed produce, those that have specialised in

collecting mangrove and rainforest produce for trade, and those that formerly survived as rainforest foragers (Benjamin 1985, 2002).



**Figure 1:** Malaya, showing distribution of the Orang Asli and location of sites mentioned in the text (Orang Asli distribution taken from Benjamin (1985) and Burenhult et al. (this volume)).

Fittingly diverse for people as variable as the Orang Asli is the range of scholarly opinion on how discrete the three divisions are and whether distinct origins should be sought. Nik Hassan (2005) has proposed that all of the Orang Asli as well as the Melayu are related linguistically and genetically, and that they differentiated during the late Holocene in response to maritime and cross-peninsular trade. Rambo (1988) accepted the

distinctiveness of the Orang Asli from the Melayu, but argued that the three Orang Asli divisions have resulted from biological adaptation to their particular environments. Benjamin (1985, 1987) explicated a model which associated the Aboriginal Malays with Southern Aslian or Malay dialects, ranked tribal social organisation, and an economic focus on collecting forest produce; the Senoi with Central Aslian, egalitarian tribal social organization, and swidden horticulture in the highlands; and the Semang with egalitarian bands of Northern Aslian speakers who foraged in the lowland rainforests. Benjamin emphasised that these three associations lack strict congruity, as further explored by Bulbeck (2003) and Burenhult et al. (this volume), but argued that they provide a window on how the Orang Asli differentiated within a framework of staggered origins in the peninsula. Fix (2002, this volume) has stressed the differentiation aspect of Benjamin's model, whereas Bellwood (1993, 1997) has focused on the implication of successive origin times. In Bellwood's view, as dealt with in due course, the Semang represent the 'Australo-Melanesian' foragers who had occupied Southeast Asia prior to the immigration from south China of 'Mongoloid' farmers who introduced Austroasiatic to the peninsula during the Neolithic, and Austronesian some 2000 years ago.

The question of Orang Asli ethnogenesis has been of great interest to me, and involved me in several recent studies that have presented substantial new information on Orang Asli biological attributes (Rayner & Bulbeck 2001; Bulbeck et al. 2005; Bulbeck & Lauer 2006; Hill et al. 2006). Of particular note is the survey of their mitochondrial DNA (mtDNA), which successfully teased out the complex nature of genetic interactions between populations in the peninsula and those in surrounding regions (Hill et al. 2006). While this study confirmed some of the main points argued by Bellwood (1993, 1997), it also highlighted the simplifications of his three-layer model. Subsequent research on the mtDNA of surrounding populations has additionally allowed fine-tuning of the interpretation of the Orang Asli data, as detailed by Oppenheimer (this volume), who was also one of the protagonists in the original Hill et al. (2006) study. Where my contribution differs from Oppenheimer's is in my slightly different retention of the interpretations advanced in the original study, and my focus on Malaya's archaeology, including its human osteological record. As will be elaborated in due course, this archaeological perspective allows us to appreciate Orang Asli diversity in terms of successful adaptations to the new niches that have been successively created as a result of environmental change, population incursions and long-distance interactions over time (cf. Benjamin 1985, 1987).

Historical linguists agree on tracing the Aslian languages back to an original Austroasiatic language, labelled proto-Aslian, established in Malaya by 4000 years ago (Burenhult et al., this volume). However, the details of Aslian linguistic diversification, and the interaction between Southern Aslian and the later arriving Malayic dialects, are noted as topics for ongoing research (Benjamin 1987, 2002). Fortunately, Malaya has a rich archaeological record, thanks to well over 100 years of research into the region's numerous limestone rockshelters and open-air sites, and these data provide invaluable information about Malaya's past. The sites of relevance to this chapter (Figure 1) date from approximately 75,000 years ago, which is the earliest date for when *Homo sapiens* could have colonised Malaya (Soares et al. 2009; Oppenheimer 2009; see discussion below), to 1300–1400 CE, when the Melayu settled Malaya and in due course established Melaka as their capital (Hooker 2003). Comparisons with the archaeological record in Sumatra are important, both because Sumatra and Malaya were connected by land until c. 8000 years ago (Bulbeck 2003), and because Sumatra was the Melayu homeland (Adelaar 2004).

The widely touted terms 'Palaeolithic' (Old Stone Age) and 'Mesolithic' (Middle Stone Age) have little relevance for Malaya where, as we shall see, the important issues are the



development of the local 'Hoabinhian' cobble-based industry, and the subsequent incorporation of Neolithic technology. This chapter follows Vietnamese terminology (Nguyen et al. 2004) in defining the Neolithic by the presence of polished stone tools and pottery, leaving open the relationship of these archaeologically durable remains to the advent of agriculture, especially 'arboriculture' or the management of economically useful palms and trees (Latinis 2000). As will become clear, the Neolithic transition in Malaya was less an event and more a process of extended interaction with multiple outside regions. Further, the existence of a Bronze Age in Malaya prior to the arrival of iron technology is by no means certain. Accordingly, depending on context, the terms Early Metal Phase, Bronze/Iron Age and Iron Age will all be employed for the proto-historical period corresponding approximately to the first millennium CE. Finally, all cited radiocarbon dates from Malaya (identified as 'years BP') are uncalibrated, and, thus, most are minimum ages, especially with reference to the Pleistocene (whose junction with the Holocene is set at 10,000 years ago).

## **2 Late Quaternary environmental change**

Tropical forest has been the natural vegetation of Malaya throughout most of the Holocene, when the sea has stood at or above its current level. Analysis of pollen and phytoliths from the Nong Thalee Song Hong swamp in southern Thailand documents the local formation of tropical forest between 11,000 and 9000 years BP, and its further expansion during the early Holocene. Beforehand, mosaics of savannah and woodland dominated the vegetation between 21,000 and 11,000 years BP, when sea levels were as much as 120–150m lower than today (Kealhofer 2003). A similar scenario of environmental change, including evidence of cooler Pleistocene temperatures than today's, has been reconstructed by Taylor et al. (2001) from their analysis of plant pollen and spores, dating back to 23,000 years BP, at the Nee Soon swamp in Singapore.

Pookajorn (1996) proposed a transition to moister conditions between the Late Glacial Maximum (LGM) and terminal Pleistocene based on the archaeological evidence from Levels 2 and 3 of the Moh Khiew rockshelter in southern Thailand. Both levels produced pollen taxa reflecting moist conditions and the remains of forest-dwelling fauna. Level 3 is dated to between 9000 and 11,000 years BP, whereas the underlying Level 2 is dated no more precisely than <26,000 years BP, based on a radiocarbon determination near the top of Level 1 (Pookajorn 1996). Given the limited evidence for environmental change between Levels 2 and 3, it would be unlikely that Level 2 is more than a few thousand years older than Level 3. In fact, no site in Malaya has produced positive evidence for habitation between 15,000 and 26,000 years BP, an observation that probably reflects scanty use of the present-day Malaya landmass at the height of the LGM, owing to the relocation of the coastally oriented population to lowlands now inundated by sea (Bulbeck 2003).

Between 26,000 and 43,000 years BP, sea levels were higher than at the LGM, and evidence for habitation comes from the basal archaeological deposits of two rockshelters in southern Thailand, viz. Lang Rongrien as well as Moh Khiew (Figure 1). Mudar & Anderson (2007) discuss the evidence for a climate at the time that was cooler and drier than at present, and an environment dominated by a forest-savannah mosaic, as would be consistent with the pre-LGM faunal remains recovered from both sites. However, the sites' faunal taxa are mutually exclusive, and this observation has been interpreted as indicating sporadic visits to Lang Rongrien by a coastally oriented population, in contrast to more intensive habitation at Moh Khiew (Mudar & Anderson 2007).

Malaya's *H. sapiens* inhabitants have apparently exerted an increasing influence on their environment over time. Starting with the pre-LGM period, we find that Mudar & Anderson (2007) did not see any need to address the topic of controlled burning of the vegetation. However, referring to the period between the LGM and the end of the Pleistocene, both Taylor et al. (2001) and Kealhofer (2003) allude to the possibility of human-mediated fires. Additionally, Taylor et al. (2001) interpret the high levels of mid-Holocene charcoal at Nee Soon swamp as evidence of early farming activities in Singapore. Similarly, Kealhofer (2003) argues for forest management in southern Thailand after 8000 years BP, leading to intentional planting of useful trees by 6500 years BP, more intensive arboriculture by 6000–5000 years BP, and slash-and-burn agriculture by 4000 years BP. This argument is indirectly strengthened by the evidence for woodworking and possible tree felling revealed by the use-wear on Hoabinhian stone tools from northwest Thailand (Bannanurag 1988:77).

A range of economically useful plants was propagated through forest management practices by mid-Holocene times. These include the betel-nut palm, to judge by the observation of betel-stained teeth at the Guar Kepah shell midden (Bulbeck 2005a), and other palms, bananas, and rice, based on Kealhofer's (2003) phytolith evidence. Further evidence of early rice production comes from the reconstruction of the term 'rice' in proto-Aslian by historical linguists (Gianno & Bayr 2009) and the rice grains from Cultural Level 2 at Sakai Cave, in southern Thailand, which Pookajorn (1996) dates to the mid-Holocene. The sites of Nyong and Jenderam Hilir in Peninsular Malaysia are important as indications of Neolithic open-air settlements, complementing the predominance of rockshelters in the inventory of Malaya Neolithic sites (Leong Sau Heng 1991; Bulbeck 2004a). However, neither Nyong nor Jenderam Hilir provides sufficient justification to infer an agriculturally focused economy. Instead, as will be argued below, the range of mid-Holocene adaptations probably ranged from mixed swidden-forager to fully foraging economies.

### 3 Late Pleistocene to mid-Holocene artefact assemblages

The first studies of the Stone Age in tropical Asia were undertaken by European antiquarians, who unsurprisingly promulgated a Eurocentric perspective. Synthesising the available information, Movius (1944) drew a contrast between the Pebble Tool Complex of Southeast Asia and the Acheulian hand-axe assemblages of India and places westward. Movius further distinguished between Southeast Asia's ancient 'chopper/chopping-tool' assemblages and its more recent but 'culturally stagnant' Hoabinhian assemblages. In recent years, archaeologists have been at pains to identify hand-axes in Southeast Asia (for example, Nguyen et al. 2004; Mokhtar 2006; Simanjuntak et al. 2006), by which they mean cobbles, usually bifacially flaked, with flat, relatively wide bases and a pointed shape at the opposing end (labelled as pointed wide-based cobbles in Table 1). Of more direct relevance, the concept of an ancestor-descendant relationship between chopper/chopping-tool and Hoabinhian assemblages has been challenged by the recovery of chronologically intermediate assemblages dominated by cores and/or small flakes (Anderson 1990).

Were these 'core-flake' assemblages related to the arrival of *H. sapiens* in Southeast Asia, and did they later evolve into the Hoabinhian as a cultural adaptation of *H. sapiens* foragers to their forested environment? Questions like these would have been unimaginable during Movius's day or indeed until 1980, when the general assumption was that *H. sapiens* in East Asia had evolved from locally ancestral *H. erectus* populations (for example, Weidenreich 1947; Coon 1962; Wolpoff 1980). However, these questions are

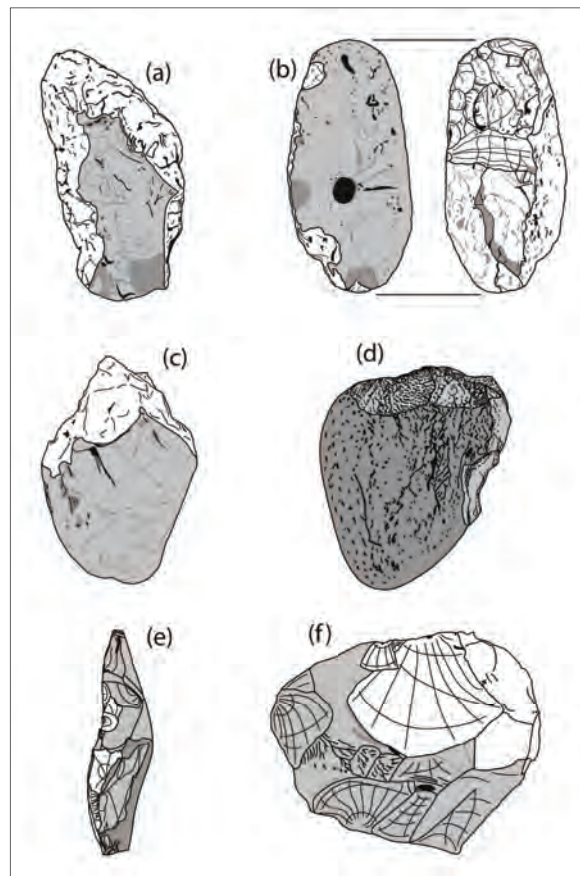
now highly pertinent in view of the new orthodoxy whereby *H. sapiens* had evolved in Africa during the Late Pleistocene prior to colonising other landmasses of the tropical and temperate world (Day & Stringer 1982; Wainscoat et al. 1986; Cann et al. 1987; Stringer 1994; Oppenheimer 2003; Cameron & Groves 2004).

Before addressing the question of associations between species of *Homo* and stone artefact assemblages, certain points of terminology require explanation. First, the nodules extracted from the landscape for manufacturing stone tools of the Hoabinhian and related industries are water-rounded cobbles whose dimensions exceed those of pebbles as defined by geologists. Confusingly, the term 'pebble' is widely used in the archaeological literature when referring to these cobbles, but, in this chapter, it will be substituted by the geologically correct term 'cobble' (White & Gorman 2004). Secondly, the term 'core' will here be restricted to nodules of stone with one or more striking platforms, produced as a knapped surface, for the detachment of flakes at an approximately perpendicular orientation to the striking platform (see Figure 2f). The distinction is made here between cores in this sense and 'flaked cobbles' with the flakes detached at a shallow orientation relative to the flaking surface (see Figure 2d).

Thirdly, this chapter follows Kamminga (2007) in defining 'sumatraliths' as cobbles with a completely flaked perimeter and shallow flaking across most or all of a single face. Kamminga analysed the Hoabinhian assemblage from Sai Yok in south-central Thailand in terms of cobble reduction used to produce sumatraliths, and argued they had been hafted for use as woodworking adzes. One defining characteristic of Kamminga's reduction sequence is the avoidance of bifacial cobble working. Such bifacial flaking would result in cobbles with bifacial circumferential working at earlier reduction stages, and in pieces that resemble hand-axes or picks with comprehensive bifacial flake removal. The second defining characteristic is a completely knapped perimeter (see Figure 2a), in preparation for centripetal flake removal, which distinguishes sumatraliths from pebbles with flakes unifacially detached in a less structured process. The third defining feature is the production of a flattish flaked surface, as opposed to the medially crested surface of 'chopper' forms (Figure 2c). Figure 2b depicts a sumatralith from the Late Pleistocene site of Kota Tampan, which (as we shall see) would appear to be some 65 millennia older than Malaya's other sumatraliths.

In Malaya, bifacial flaking of cobbles is such a common feature of assemblages younger than 15,000 years BP that it could be nominated as a defining characteristic of the Malayan Hoabinhian (Table 1). It has not been documented for assemblages older than 26,000 years BP, except Bukit Bunuh, which is difficult to distinguish from the terminal Pleistocene/Holocene Hoabinhian assemblages except for its pointed wide-based cobbles ('hand-axes'), which rarely occur in the latter assemblages (Table 1). The other three pre-26,000 year BP assemblages (Kota Tampan, and the basal deposits at Lang Rongrien and Moh Khiew) have a further difference from Hoabinhian assemblages, in that cores are at least as strongly represented as flaked cobbles (Table 1). The singularly Hoabinhian aspect of the Bukit Bunuh assemblage certainly cannot be ascribed to site type; whereas the vast majority of the sites listed in Table 1 are rockshelters, Bukit Bunuh and Kota Tampan are both lithic scatters on ancient lakeshores. In addition, Bukit Bunuh is not chronologically distinguishable from the two early rockshelter assemblages; its luminescence dating of around 40,000 years ago would place it in the same age bracket as basal Lang Rongrien (Table 1), while initial occupation at Moh Khiew predated 26,000 years BP by an unknown period of time (Pookajorn 1996). Based on the present evidence, it would be difficult to disagree with Mokhtar's (2006) view that Bukit Bunuh is too Hoabinhian-like

for archaeologists to assume a terminal Pleistocene introduction of the Hoabinhian to Malaya.



**Figure 2:** Stone artefacts from Late Pleistocene lithic scatters in Malaya (not to scale)

- a) Kota Tampan cobble unifacially flaked around most of its perimeter (based on Zuraina 2003:62)
- b) Kota Tampan sumatralith; cobble unifacially flaked across entire face (based on Zuraina 2003:73)
- c) Kota Tampan cobble flaked to produce a chopper (based on Zuraina 2003:63)
- d) Bukit Bunuh quartzite cobble with unifacial edge flaking (based on Zuraina 2003:74)
- e) Bukit Bunuh retouched chert flake with traces of use wear (based on Zuraina 2003:77)
- f) Bukit Bunuh chert core, rotated and multi-platform (based on Zuraina 2003:75).

Kota Tampan also presents interpretative difficulties, but of a different nature to those for Bukit Bunuh. Its age of 74,000 years ago, inferred from the Toba ash fall which seals the site, sits right at the earliest plausible estimates for the arrival of *H. sapiens* mtDNA in Southeast Asia (Soares et al. 2009; cf. Macaulay et al. 2005). It would be convenient to ascribe the Kota Tampan assemblage to an archaic local hominin which had been made extinct by the Toba eruption, and the other Late Pleistocene Malaya assemblages to early *H. sapiens*. However, there is nothing about the Bukit Bunuh and basal Lang Rongrien and Moh Khiew assemblages that would distinguish them as a group from Kota Tampan, with its 20 cores, ten unifacially flaked cobbles, retouch on a small number of flakes, and four flat-based cobbles equipped with a pointed shape at the opposing end (Zuraina 1990). While Zuraina's Figure 8 illustrates two unifacial chopper-like cobbles, unifacial choppers

are also recorded for Bukit Bunuh (Figure 2d), basal Lang Rongrien (Anderson 1990:56) and even some Holocene assemblages (for example Zolkurnian 1998). Especially given Zuraina's (1990) observation of ground edges on one of the Kota Tampan cobbles, and the sumatralith referred to above (Figure 2b), it would be very difficult to dissociate Kota Tampan from an early presence of *H. sapiens*. Accordingly, Kota Tampan couples with Jwalapuram in India (Petraglia et al. 2007) in suggesting that the oldest time estimates for modern human mtDNA in Asia would correctly date to the Toba eruption, and in particular reflect the post-Toba lineage proliferation of the small number of early Asian *H. sapiens* that were able to survive this cataclysmic eruption (Ambrose 1998; Oppenheimer 2009).

Of the four pre-LGM assemblages reviewed here, that from basal Lang Rongrien is the best described (Anderson 1990), and also the least Hoabinhian-like (Table 1). Assessment of the 1181 artifacts from Level 1 at Moh Khiew is particularly difficult in view of these artifacts' peremptory description, and relies on impressionistic appraisals from scattered sources. In addition to the major role of unifacial pebble tools at this level (Chitkament 2006–07), Pookajorn (1996:204) states that the 'typology and raw material of bifacial tools from this cultural level are almost the same as those ... from Lang Rongrien ... older than 37,000 BP', which would suggest cores rather than bifacially flaked pebbles (cf. Anderson 1990). Pookajorn (1994, 1996) also mentions numerous 'flake tools' from Level 1, as well as the other Moh Khiew levels, but this assessment must have been made on the basis of the flakes' shape rather than signs of retouch, because Chitkament (2006-07:141) found very few signs of retouch in his sample of Moh Khiew flakes. Fortunately, illustrations and more complete descriptions of the flaked stone artifacts in the higher levels, and use-wear study on the Level 2 flakes (Pookajorn 1994; Chitkament 2006-07), permit a more certain characterisation (Table 1) of the Hoabinhian assemblages in Levels 2 to 4. (Level 5 at the site is described as having similar flaked lithics to those in Level 4, in addition to pottery and polished stone adzes.)

As shown in Table 1, numerous rockshelters in Peninsular Malaysia have an occupation history restricted to the terminal Pleistocene and Holocene, based on the available radiocarbon dates. A frequent characteristic of these Hoabinhian assemblages in Peninsular Malaysia is the lack of cores, a feature that dates back to the terminal Pleistocene in southeast Malaya (Table 1). Sumatraliths, on the other hand, regularly appear only as of the early Holocene and in western Peninsular Malaysia and Sumatra, with no documented cases in southern Thailand, and very rare instances in central or eastern Peninsular Malaysia (Adi 2000:157; see below). Sumatraliths' restricted chronological and geographical distribution appears to monitor maritime interaction across the Melaka Strait as it underwent flooding during the early Holocene (Bulbeck 2008).

It should be noted here that the radiocarbon chronology of Gua Gunung Runtuh and Gua Singa, both of which have produced sumatraliths, relies on dates from freshwater shell. These dates are likely to be in the order of 2000–3000 years older than the true age of the shell (Adi 2000) and, thus, are shown in Table 1 as overestimates. Togi Ndrawa on Nias Island, immediately west of Sumatra, has been well dated but the lithics are poorly described. From the account provided by Forestier et al. (2005), sumatraliths and other unifacial pebbles do not appear until the third level, which dates to between c. 9500 and 3000 years BP. Further information from Wiradnyana (2008) indicates the presence of cores and retouched flakes, at least in the upper levels. The Togi Ndrawa Hoabinhian assemblage would appear to be broadly similar to that excavated at Gua Pandan, in South Sumatra, which extends the documented distribution of Hoabinhian sites in Sumatra (Forestier et al. 2006; Simanjuntak et al. 2006). Sumatra's Hoabinhian is best represented

in the massive shell middens of coastal northeast Sumatra (Heekeren 1972), associated with early Holocene radiocarbon dates (Edwards McKinnon 1991).

In addition to the Sumatra shell middens, similarly large middens are known across the Melaka Strait. Guar Kepah, previously dated to the fifth millennium BP from consideration of sea-level stands and the artefactual content (Bulbeck 2003), should probably be dated a millennium earlier based on subsequent research into the Peninsula's sea-level changes (cf. Anderson 2005), as would also be consistent with the c. 6000 year BP charcoal dating from the Bukit Perang midden (Bulbeck 2003). In southern Thailand, the Tham Sua midden has a basal radiocarbon determination on marine shell of around 6500 years BP, below the deepest sherd of pottery which otherwise occurs throughout the deposit; the site would have extended over dozens of square metres (Anderson 2005).

Gua Gunung Runtuh also has a very small number of cores and a single 'palaeoadze', described as similar to Kota Tampan examples (Zuraina et al. 1994), while two hand-axes are reported from Gua Bintong, which is referred to as Bukit Chuping in the earlier literature (Matthews 1961:47). Gua Bintong is dated to around 5000 years ago by Rabett (2005), although it also has late Holocene artifacts such as bronze (Leong Sau Heng 1989). The chronology tabulated for Gua Kajang is based on two charcoal determinations of c. 6000 and 9000 years BP (Zuraina 1998) from a virtually aceramic section of the site (Chia 1998:160), although later habitation would also be indicated by the recovery of pottery from above the site's lime-impregnated levels (Matthews 1961).

Polished and ground stone artifacts occur sporadically in contexts older than 5000 years BP in Malaya. Examples include two polished stone fragments from Layer 6 of Lang Rongrien (Anderson 1990:45), the waisted ground adzes from Guar Kepah (Matthews 1961:27–28), and flaked pebbles with polished edges from the deepest levels at Gua Kerbau (Bulbeck 2003) and the c. 6000-year BP levels of Gua Peraling (Adi 2000:114). The virtually aceramic Malaysian site of Gua Madu produced two extensively flaked tools with smoothly ground edges (Matthews 1961:55–56), while Buang Baeb and Khao Khi Chan in southern Thailand have ground stone tools broadly dated between 4000 and 6000 years BP (Srisuchat 1993). Combined with indications that sumatraliths (Kammaing 2007) and other edge-flaked pebbles (Bannanurag 1988) may have had a role in tree felling, the evidence from Malaya's lithics is fully compatible with Kealhofer's (2003) scenario of increasing forest clearance between 8000 and 4000 years BP.

Also consistent with this scenario, based on lithics, is the ceramic evidence. Cordmarked pottery older than 5000 years BP includes a single sherd from Layer 5 of Lang Rongrien (Anderson 1990), the pottery from Tham Sua dated to shortly after 6400 years BP, the cups and bowls from the nearby sites of Buang Baeb and Khao Tau (Anderson 2005:145), the pottery found at all levels in Guar Kepah (Matthews 1961:27), and the Gua Peraling Layer 3 potsherds dated to  $5720 \pm 210$  BP (Adi 2000). Therefore, the date of 4000–5000 years BP for the pottery at Gua Kecil, based on a radiocarbon assay on collagen from pottery-associated bone (Bellwood 1993:35), should not be taken as marking the onset of the Neolithic in Malaya.

Bellwood (1978) nominated Sørensen's 'Ban Kao culture' as the source for Malaya's Neolithic, but subsequent research shows that this 'cultural package' should be decomposed into its multiple constituent elements. One element, cordmarked pottery, was widespread from 5000–6000 years BP onwards, not only in Malaya (as noted above) but also across Mainland Southeast Asia (Higham & Thosarat 1998; Nguyen et al. 2004). A second element, tripod ware, has a focus of distribution and its earliest dates in Malaya (within a Southeast Asian context). Tripod vessels are dated to between 4000 and 5000 years BP (radiocarbon dates on bone from the associated burials) at Buang Baeb and Khao

Khi Chan near the Isthmus of Kra (Srisuchat 1993), 4400 years BP at the nearby site of Khao Kanaab Nam (Anderson 2005:147), and 4000 years BP at Jenderam Hilir in Peninsular Malaysia (Leong Sau Heng 1991). Additional examples from undated contexts occur at Na Ching in southern Thailand (Anderson 1988), Gua Singa on the Thailand-Malaysia border (Azman 1998), and Gua Bintong, Bukit Cangkul, Gua Kodiang, Gua Gergasi, Gua Pasir, Gua Taufan and Gua Baik in western Peninsular Malaysia (Leong Sau Heng 1991). North of Malaya they are known only from Ban Kao (south-central Thailand), where their nominal age of 4000 years BP specifically refers to the age of the deposit into which they had been interred as burial goods (Bellwood 1978:167-68). Thus, while Sørensen (1972) may be correct in deriving the tripods from mid-Holocene China, their point of entry would appear to have been the Isthmus of Kra, with subsequent dispersal both northward and especially southward.

Two other elements of the 'Ban Kao culture', pedestalled pots and finely polished stone adzes, were widespread across Mainland Southeast Asia by 4000–5000 years BP (Higham 2002; Nguyen et al. 2004). This spread probably included southern Thailand (Srisuchat 1993), whereas their appearance in Peninsular Malaysia evidently postdates 4000 years BP (Bulbeck 2004a). The situation with extended burials, reviewed in detail below, is probably similar. In contrast, barkcloth beaters, a sixth element, appear to have arrived earlier in Peninsular Malaysia—one was found in a preceramic context at Gua Madu (Matthews 1961:55–56)—than at Ban Kao, where additionally they were strangely made of pottery rather than stone (Bellwood 1978:168). North Vietnam, where stone barkcloth beaters were present by 4000–5000 years BP (Nguyen et al. 2004:183), is a possible source for the Malaya examples. Thus, while all of the elements listed above came together at Ban Kao, the archaeological evidence refutes any notion that the 'Ban Kao culture' was a discrete cultural package developed in south-central Thailand or transmitted *holus bolus* from there to Malaya.

See Table 1 (Malayan and Sumatran Late Pleistocene and Holocene assemblages in terms of Hoabinhian and other attributes), Table 2 (Prehistoric burial disposal modes in Malaya, along with main grave good associations) and Table 3 (Main Thailand and Peninsular Malaysia sites with extended burials, ordered approximately from north to south).

## 5 Holocene stature reduction – further evidence

Evidence that the prehistoric inhabitants of Malaya were taller than the Orang Asli has been previously reviewed on several occasions (Bulbeck 1996, 2003, 2005b), and these findings can now be supplemented with additional data recently available for Gua Cha (see Acknowledgements). For convenience, the methodology and detailed observations are placed in Appendix A, and only a summary of the evidence is presented in the main text. Collectively, these data show that, for both sexes (Tables 4 and 5), the Orang Asli on average have shorter limb bone lengths and stature than the Peninsula's prehistoric inhabitants, with frequent instances of non-overlapping ranges. Fibula length alone does not follow this pattern, which may be attributed to the very small number of complete prehistoric fibulae available for measurement (Appendix A). When Holocene size reduction is estimated, by expressing the difference between the early Holocene and Orang Asli averages as a percentage of the early Holocene average, the estimate hovers around 10% with a range between 7% and 18%.

**Table 1:** Malayan and Sumatran Late Pleistocene and Holocene assemblages in terms of Hoabinhian and other attributes. Asterisked radiocarbon dates are on freshwater shell.

Site/level	Dating	Shaped cores?	Retouched flakes?	Pointed wide-based cobbles?	Bifacially flaked cobbles?	Unifacially flaked cobbles?	Typological sumatraliths?	Characterisation of assemblage
Kota Tampan (Zuraina 1990, 2003)	c. 74,000 years ago	Yes	Yes	Yes	No	Yes	Yes	Cores and unifacially flaked cobbles, plus dominance of primary flakes
Lang Rongrien units 8–10 (Anderson 1990)	c. 27,000–>43,000 years BP	Yes	Yes	No	No	Yes	No	Core-based assemblage with retouched/utilised flakes
Bukit Bunuh (Mokhtar 2006)	c. 40,000 years ago	Yes	Yes	Yes	Yes	Yes	No	Variety of flaked cobbles plus dominance of primary flakes
Moh Khiew Level 1 (Chitkament 2006–07)	≥ 26,000 years BP	Yes	Not clear	No	No	Yes	No	Flakes from cores and unifacially flaked cobbles
Moh Khiew Level 2 (Chitkament 2006–07)	> 11,000 < 15,000 years BP	Yes	Many utilised	No	Yes	Yes	No	Unifacially and bifacially flaked cobbles with many utilised flakes
Gua Sagu spits 9–13 (Zuraina et al. 1998)	13,000–15,000 years BP	No	No	No	Yes	No	No	Flakes from bifacially flaked cobbles
Moh Khiew Level 3 (Chitkament 2006–07)	8000–11,000 years BP	Yes	Yes	No	Yes	Yes	No	Unifacially and (mainly) bifacially flaked cobbles with retouched flakes
Gua Gunung Runtuh (Zuraina et al. 1994)	<8000*–<13,000* years BP	Yes	Yes	Yes	Yes	Yes	Yes	Predominantly unifacially flaked cobbles and unmodified flakes
Gua Tenggek (Zuraina et al. 1998)	c. 3000–11,000 years BP	No	No	No	Yes	No	No	Flakes from bifacially flaked cobbles



**Table 1:** *continued*

Site/level	Dating	Shaped cores?	Retouched flakes?	Pointed wide-based cobbles?	Bifacially flaked cobbles?	Unifacially flaked cobbles?	Typological sumatraliths?	Characterisation of assemblage
Lang Rongrien units 5–6 (Anderson 1990)	8000–10,000 years BP	Yes	Yes	No	Yes	Yes	No	Diverse cores and flaked cobbles plus many primary flakes
Sakai Cave (Pookajorn 1996)	7500–10,000 years BP	Not clear	No	Not clear	Not clear	Yes	No	Unifacially flaked cobbles
Gua Teluk Kelawar (Zolkurnian 1998)	3000–10,000 years BP	No	Yes	Not clear	Yes	Yes	Yes	Flakes from unifacially and bifacially flaked cobbles
Gua Peraling layers 3–7 (Adi 2000)	5000–<12,000* years BP	No	Not clear	No	Yes	Yes	No	Flakes from unifacially and bifacially flaked cobbles
Gua Chawas spits 26–65 (Adi 2000)	4000–<12,000* years BP	No	Not clear	No	Yes	Yes	No	Flakes from unifacially and bifacially flaked cobbles
Gua Pandan Level 2 (Simanjuntak et al. 2006)	6500–9000 years BP	Yes	Yes	No	No	Yes	Yes	Hoabinhian and ‘Mousteroid’ components
Togi Ndrawa Level 3 (Forestier et al 2005; Wiradnyana 2008)	3000–9000 years BP	Yes	Yes	No	No	Yes	Yes	Poorly described; attributes noted here are provisional
Gua Kajang (Matthews 1961)	6000–9000 years BP	Yes	One utilised	No	Yes	Yes	Yes	Variable lithology and technology
Gua Bukit Taat (Nik Hassan et al. 1990)	3000–9000 years BP	No	Yes	No	Not clear	Yes	Yes	Flakes from unifacially flaked cobbles
Gua Sagu spits 2–8 (Zuraina et al. 1998)	1000–10,000 years BP	No	Yes	No	Yes	Yes	Yes	Flakes from bifacially flaked cobbles
Gua Teluk Kelawar B (Zolkurnian 1998)	3000–<10,000* years BP	No	Yes	No	No	No	No	Small assemblage of flakes and utilised cobbles
Gua Batu Tukang (Zolkurnian 1998)	3500–<9000* years BP	No	Yes	No	Yes	Yes	Yes	Flakes from a variety of cobble-based tools

**Table 1:** *continued*

Site/level	Dating	Shaped cores?	Retouched flakes?	Pointed wide-based cobbles?	Bifacially flaked cobbles?	Unifacially flaked cobbles?	Typological sumatraliths?	Characterisation of assemblage
Gua Singa (Azman 1998; Zuraina 1998)	<9000* years BP	No	Yes	No	Yes	Yes	Yes	Poorly described cobble-based industry
Gua Kerbau (Bulbeck 2003; Matthews 1961)	Early-late Holocene	No	Yes	No	Yes	Yes	Yes	Flakes from unifacially and bifacially flaked cobbles
Moh Khiew Level 4 (Chitkament 2006-07)	5500–7000 years BP	Yes	Yes	No	Yes	Yes	No	Unifacially and (mainly) bifacially flaked cobbles with retouched flakes
Gua Cha Layers 3–4 (Adi 1981)	3000–7000 years BP	No	Yes	No	Yes	Yes	No	Flakes from unifacially and bifacially flaked cobbles
Gua Kecil 14–32" (Dunn 1964)	c. 2000–c. 7000 years BP	No	Some utilised	No	Yes	Yes	Yes	Flakes from unifacially and bifacially flaked cobbles
Guar Kepah (Matthews 1961)	5–6000 years ago	Not reported	Not reported	No	Yes	Yes	Yes	Poorly described cobble-based industry
Gua Ngaum (Zolkurnian 1998)	c. 3000–<6000* years BP	No	Yes	No	Yes	Yes	Yes	Flakes from unifacially and bifacially flaked cobbles
Gua Baik (Callenfels and Noone 1938)	Mid-late Holocene	Not reported	Not reported	No	Yes	Yes	Yes	Ceramic Hoabinhian
Gua Bintong (Matthews 1961)	c. 5000 years ago	No	Yes	Yes	Yes	Yes	No	Flakes from unifacially and bifacially flaked cobbles

**Table 2:** Prehistoric burial disposal modes in Malaya, along with main grave good associations (see text).

<b>Period</b>	<b>Flexed inhumations</b>	<b>Secondary disposals</b>	<b>Extended supine inhumations</b>
Late Holocene Metal Phase	None recorded	Kuala Selinsing (composite canoe burial)	Gua Harimau 3, 5, 7? (bronze and glass grave goods) Kuala Selinsing (majority of burials—pottery and glass grave goods) Gua Peraling 1 (no grave goods) Gua Baik 94–96 (no grave goods)
Late Holocene Neolithic	Lang Rongrien 4 (with pedestalled pot)	Lang Rongrien 2, 3 (cordmarked pots) Gua Harimau 4, 6 (pottery) Gua Harimau 11, 95E (no grave goods) Gua Baik 98, 99 (no grave goods)	Other Lang Rongrien burials (pedestalled and cordmarked pots) Gua Harimau 1, 2, 8, 9, 10 (cordmarked pots) Gua Cha (most have pedestalled and/or cordmarked pots)
Middle Holocene	Moh Khiew 1 Gua Kerbau Gua Cha Gua Peraling 4 Gua Baik 583 (lowest layer) Gunung Cheroh (All lacking grave goods)	Gua Peraling 2, 3 Gua Cha Gua Kepah Gunung Cheroh (All lacking grave goods)	None recorded
Early Holocene	Gua Gunung Runtuh Gua Teluk Kelawar 1 Moh Khiew 2 Gua Kerbau (All lacking grave goods)	None recorded	None recorded

**Table 3:** Main Thailand and Peninsular Malaysia sites with extended burials, ordered approximately from north to south.

Site	Location	Earliest extended burials	Other burial modes	Characterisation of burial goods
Ban Chiang (Higham & Thosarat 1998; Bulbeck 2004c)	Khorat Plateau, northeast Thailand	c. 4300 BP (Neolithic; dates on rice chaff)	Flexed burials; foetus/infant jar burials (till 3500 BP)	Pedestalled and other ceramic vessels, accompanied by bronze ornaments and weapons after 3500 BP, plus iron weapons after 3000 BP
Non Nok Tha (Bayard 1996–97; Higham & Thosarat 1998)	Khorat Plateau, northeast Thailand	c. 4000 BP (Neolithic; dates on rice chaff)	Secondary burials early in the sequence	Pedestalled and other ceramic vessels, occasionally accompanied by bronze bracelets, axes and moulds (3500–3000 BP), as well as polished stone adzes, grindstones, shell jewelry and hematite
Ban Na Di (Higham & Thosarat 1998)	Khorat Plateau, northeast Thailand	c. 3000 BP (Bronze Age; rice present)	Infant jar burials (Iron Age)	Rich assemblage of ceramics (vessels and figurines), shell jewelry, marble and other stone jewelry, and iron items (by 2400 BP)
Ban Non Wat (Higham & Thosarat 2006)	Mun Valley, central Thailand	c. 4000 BP (Neolithic; rice present)	Jar burials including infants and a seated, crouched adult	Decorated ceramic vessels, shell ornaments, pig offerings
Ban Lum Khao (Higham 2002)	Mun Valley, central Thailand	c. 3400 BP (Bronze Age; rice present)	Infant jar burials	Ceramic vessels (some pedestalled), figurines and spindle whorls, accompanied by shell ornaments, polished stone adzes, marble bangles and bronze fragments (ending 2500 BP)
Non Pa Wai (Higham 2002; Pigott et al. 2006)	Chao Phraya Plains, central Thailand	c. 4000 BP (Neolithic; rice, especially after 3500 BP)	Not documented	Ceramic vessels (some pedestalled), polished stone adzes and shell jewelry, with copper axes and moulds, a copper fishhook and ochre after 3500 BP
Non Mak La (Higham & Thosarat 1998; Higham 2002; Pigott et al. 2006)	Chao Phraya Plains, central Thailand	c. 3800 BP (Neolithic; rice, especially after 3500 BP)	Infant jar burials	Ceramic vessels and ornaments of shell, greenstone and marble
Ban Tha Kae (Higham 2002)	Chao Phraya Plains, central Thailand	c. 3800 BP (Neolithic)	Not documented	Ceramic vessels (some pedestalled), polished stone adzes and shell jewelry (Neolithic to Bronze Age), with bronze jewelry, iron tools and weapons, and glass and stone ornaments during the Iron Age
Khok Charoen (Higham 2002; Bulbeck lab. notes)	Chao Phraya Plains, central Thailand	2853 ± 33 years BP (Wk-15038), i.e. c. 3000 BP	Not documented	Ceramic vessels including pedestalled pots, shell and stone jewelry, polished stone adzes, and a bronze ring found with one burial studied by Sood Sangyichien

Table 3: *Continued*

Site	Location	Earliest extended burials	Other burial modes	Characterisation of burial goods
Ban Phu Noi (Higham 2002)	Chao Phraya Plains, central Thailand	Bronze Age	Not documented	Ceramic vessels (some pedestalled); jewelry of shell, ivory, marble and chlorite; and a polished stone adze (no bronze)
Khok Phanom Di (Higham & Thosarat 1998; Higham 2002)	South coastal Thailand	c. 4000 BP (Neolithic; rice throughout sequence)	None	Ceramic vessels (some pedestalled) and anvils and shell jewelry, along with shell knives, bone fishhooks and harpoons, ivory ornaments, and polished stone adzes and hoes (no bronze, ending 3500 BP)
Nong Nor (Boyd et al. 1998)	South coastal Thailand	3100 BP (Bronze Age; rice first appears)	Preceded by Neolithic jar burials including a seated, crouched adult	Pedestalled and other pottery vessels; jewelry of shell, bronze, tin, marble and semi-precious stone; and copper implements (ending 2700 BP)
Ban Kao (Sørensen & Hatting 1967; Bulbeck 2004d)	South-central Thailand	< 4000 BP (Neolithic)	One flexed burial	Ceramic vessels including pedestalled and tripod pots; jewelry of stone, bone and shell; polished stone adzes; polished stone adzes; barbed bone points; ivory disks; iron adzes (postdating 2500 BP)
Lang Rongrien (Anderson 1990, 2005)	Peninsular Thailand	Estimated 3000–5000 BP (Neolithic)	Flexed and secondary burials	Ceramic vessels including pedestalled pots, shell bracelet, polished stone adze; burials sometimes in log-lined coffins
Gua Harimau (Chia & Zolkurnian 2005 – see text)	Peninsular Malaysia	3500 BP (Neolithic)	Secondary burials including tooth burials	Ceramic vessels including pedestalled pots, shell jewelry, barkcloth beaters, and late in the sequence a glass bead and two bronze axes and moulds
Gua Cha (Steveking 1954; Bulbeck 2000)	Peninsular Malaysia	c. 3000 BP (Neolithic)	Preceded by Hoabinhian flexed and secondary burials	Ceramic vessels including pedestalled pots, shell jewelry and spoons, nephrite and marble bracelets, polished stone adzes, and a barkcloth beater
Kuala Selinsing (Bellwood 1997; Bulbeck 1998)	Peninsular Malaysia	2000 BP (Iron Age; rice present)	Canoe burial with bones from multiple skeletons	Ceramic vessels, glass and stone beads; burials sometimes in canoes

**Table 4:** Average male maximum limb bone lengths (in mm) and estimated/recorded stature (in cm), taken from Appendix A

<b>Period</b>	<b>Femur length</b>	<b>Tibia length</b>	<b>Fibula length</b>	<b>Humerus length</b>	<b>Radius length</b>	<b>Ulna length</b>	<b>Stature</b>
Early Holocene	451.5*	395*	–	324*	277*	285*	169*
Middle Holocene	454	355	331.5	321	255*	263*	166
Late Neolithic	434	360	310	320.5*	243*	275*	159
Early Metal Phase	438	362	–	316*	227.5	238	162
Orang Asli	406	342	325	290	226	242	154
Holocene reduction	10 %	13 %	–	10 %	18 %	15 %	9 %

\* Mutually exclusive range with the Orang Asli.

**Table 5:** Average female maximum limb bone lengths (in mm) and estimated/recorded stature (in cm), taken from Appendix A

<b>Period</b>	<b>Femur length</b>	<b>Tibia length</b>	<b>Fibula length</b>	<b>Humerus length</b>	<b>Radius length</b>	<b>Ulna length</b>	<b>Stature</b>
Early Holocene	433*	340*	–	316*	247*	271*	163.5*
Middle Holocene	419*	308	313	298*	222*	262*	158
Late Neolithic	384	327	321	277	–	232	156
Early Metal Phase	397	–	–	278	227*	–	156
Orang Asli	380	317	313	262	206	229	143
Holocene reduction	12 %	7 %	–	17 %	17 %	15 %	13%

\* Mutually exclusive range with the Orang Asli.

Male stature estimates suggest three phases of stature reduction. Gua Gunung Runtuh, suspected to be the oldest male skeleton from Malaya, has limb bone lengths consistent with a taller population than the Gua Cha Hoabinhian burials, which dominate the middle Holocene sample. The middle Holocene burials appear to have been taller on average than the late Neolithic (extended burial) and Early Metal Phase males who, in turn, tended to be taller than Orang Asli males. Humerus and radius lengths are consistent with this trend even if the pattern of steady reduction is less obvious for the other limb bone lengths (Table 4). The scenario of steady stature reduction over three phases is also consistent with the female data on stature and femur, humerus and ulna lengths (Table 5).

Based on the available data, an estimate of around ten percent reduction in stature between the early Holocene and the Orang Asli is strongly supported for both sexes, and appears to have been a process that occurred throughout the Holocene. Several mechanisms can be proposed as pressures selecting for this body size reduction, all of them directly or indirectly related to the warm, humid conditions that prevailed during the Holocene. First, according to Bergmann's Rule, a smaller, leaner physique would have assisted thermal regulation in these conditions (Gilligan & Bulbeck 2007). Secondly, small stature could have assisted with the procurement of tropical rainforest resources, whether these be foraged foods or produce collected for trade, which are distributed sparsely (both horizontally and vertically) in an environment where profuse undergrowth hampers

mobility (Bulbeck 2003). Thirdly, dietary breadth and access to high-quality protein appear to have decreased over time for all Orang Asli groups. In the case of foragers, this dietary shift can be seen in the increase of arboreal species relative to ungulates in the rockshelter faunal refuse (Bulbeck 2003) and, with farming groups, it can be seen in the increasing reliance on low-protein root crops (Fix 2002).

Osteological evidence of the selection pressures favouring reduced stature includes the high proportion of the Gua Cha and Guar Kepah teeth affected by macroscopic dental enamel hypoplasia. This condition registers a temporary cessation to childhood growth, whether due to illness, malnutrition or some other physiological stress, and the child's approximate age at which growth was arrested can be gauged from which teeth and/or crown segments were affected (Hillson 1996). The Gua Cha crown segments involved particularly reflect pronounced stresses on physical growth during mid-childhood, while at Guar Kepah the third molars were most affected, which indicates stresses hindering the growth spurt of adolescence (Bulbeck 2005a, 2005b). These observations may reflect selection pressures operating directly against unfettered childhood growth, or (following the model of Migliano et al. 2007) monitor high childhood mortality rates, and hence, selection pressures for females' younger onset of reproduction and (indirectly) their smaller body size.

## 6 Other relevant studies in physical anthropology

The question of Orang Asli origins is central to the view developed by Coon (1962) and Howells (1973a) in which the foraging populations of Southeast Asia, whether pre-Neolithic or represented by ethnographically recorded Negritos, belong to the Australo-Melanesian 'race'. This term covers the indigenous populations from Tasmania to Melanesia, bordering along eastern Indonesia with 'Mongoloid' populations, which are found across Polynesia, Micronesia, the New World and, apart from isolates such as Negritos and the Ainu, eastern Asia (Bulbeck et al. 2006). Bellwood, in numerous publications (for example, 1978, 1992, 1997), has adumbrated a scenario in which Mongoloid farmers, who originated in South China, expanded across Southeast Asia during the Neolithic, and introduced the proto-versions of Southeast Asia's indigenous languages. Purportedly, the Mongoloids' agricultural economy gave them a crucial demographic advantage over the Australo-Melanesian foragers who were either absorbed or driven into isolated refuges.

The above-mentioned works by Bellwood, and Coon and Howells before him, cited studies on ancient Southeast Asian skeletal remains that inferred an Australo-Melanesian status on the grounds of generally large teeth, robust cranial morphology, squat facial skeleton and elongated braincase. However, as contended by Bulbeck (1981) and Storm (1995), none of these features unambiguously distinguish between Australo-Melanesian and Mongoloid populations. Further, the late Holocene trend amongst Southeast Asians towards smaller teeth, more gracile crania, narrower facial skeleton and broader braincase could be explained in local evolutionary terms, either as adaptations to an agricultural subsistence (Bulbeck 1981) or to the humid heat of the Holocene (Storm 1995).

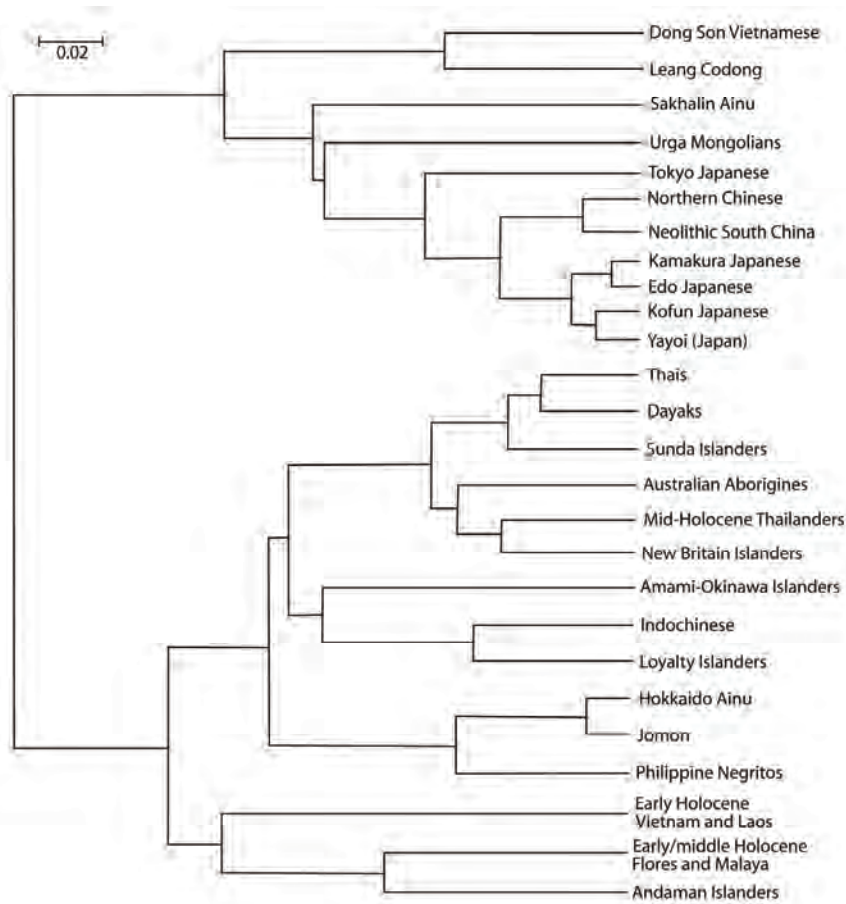
Studies on dental morphology, which covers traits such as shovelled incisors, had suggested that all *H. sapiens* in Southeast Asia, regardless of antiquity, belonged to a single population complex, labelled 'sundadont' by Turner (1990; Turner & Eder 2006). However, Turner based this conclusion on an 'Early Southeast Asian' sample which was dominated by Neolithic and Early Metal Phase teeth, and so would be expected to be similar to recent Southeast Asian teeth even according to Bellwood's scenario of a

Neolithic Mongoloid immigration (Bulbeck 2000). The finding which contradicts Bellwood's scenario is that the Neolithic teeth from Gua Cha appear more Australo-Melanesian than sundadont, which suggests that any China-based migration of Mongoloids across Southeast Asia would appear to have occurred within the last 2000 years (Bulbeck 2000). In this chapter I will test this suggestion against the dental anthropological data provided by Matsumura & Hudson (2004), who interpret their data as supporting Bellwood's Neolithic migration scenario. Their data are crucial in this regard, because they include an analysis of dental metrical shape—that is, the relative size of tooth diameters to each other, with the confounding factor of overall tooth size removed—as well as dental morphology. In addition, I accept the analysis by Matsumura et al. (2008) of the transitional Neolithic/Bronze Age cemetery population of Man Bac, in North Vietnam. That study assigned the majority of the individuals to a Mongoloid population with its origins in China, but also identified a minority component of residual Australo-Melanesians. My point here is that, since North Vietnam abuts South China, any China-derived influences on Southeast Asians should have appeared relatively early there, and so Man Bac represents the earliest possible time for when Southeast Asia more widely could have been affected.

Figure 3 shows the inter-population relationships indicated by the dental morphology distances published by Matsumura & Hudson (2004, Table 7). It differs from the hierarchical dendrogram published by Matsumura & Hudson (2004, Fig. 5) in that the branches have been swivelled so as to seriate the populations, in as orderly a manner as the dendrogram structure allows, along the single major axis of biological variation (see Bulbeck 2000, 2006 on seriated dendrograms). At one end of this axis is a cluster of 'sinodont' groups (cf. Turner 1990) such as Mongolians, Iron Age to recent Japanese, and Neolithic to recent Chinese, and, at the far end, Sakhalin Ainu and the two Bronze/Iron Age Southeast Asian samples (Dong Son in North Vietnam, and Leang Codong in Sulawesi). At the other extreme are the two pre-Neolithic cum Neolithic Southeast Asian samples (from Vietnam/Laos and Flores/Malaya) plus Andaman Islanders, all of which Matsumura & Hudson (2004) consider to be Australo-Melanesian. However, based on where Australian Aborigines and New Britain fall in the seriated order, an Australo-Melanesian status would also apply to 'Mid-Holocene Thailanders' (represented by middle/late Holocene teeth from Ban Kao, Non Nok Tha and Khok Phanom Di), plus the Jomon of pre-Iron Age Japan, as well as recent Hokkaido Ainu, Amami-Okinawa Islanders and Indochinese. That is, the data fail to distinguish Australo-Melanesians from numerous non-Australo-Melanesian populations of eastern Asia, and if any label could be applied to such a diverse grouping it would have to be non-sinodont.

To the degree that we can rely on the dental morphology data in Matsumura & Hudson (2004), evidence for population incursion into Southeast Asia from China would be dated as of 800 BCE, when the Dong Son culture commenced (Nguyen et al. 2004). Further, any such incursion would have affected only some (for example Thais, Dayaks and Sunda Islanders), not all recent Southeast Asians, and certainly not Neolithic/Bronze Age Southeast Asians, given that 'Mid-Holocene Thailanders' cluster with Australian Aborigines and New Britain Melanesians. Additionally, the 'Austroasiatic corridor' described above would have been established at a time when populations along its entire length had a non-sinodont dental morphology. For this reason, there would be little point in looking to dental morphology for evidence of Neolithic/Bronze Age population movement or gene flow along the Mekong valley.



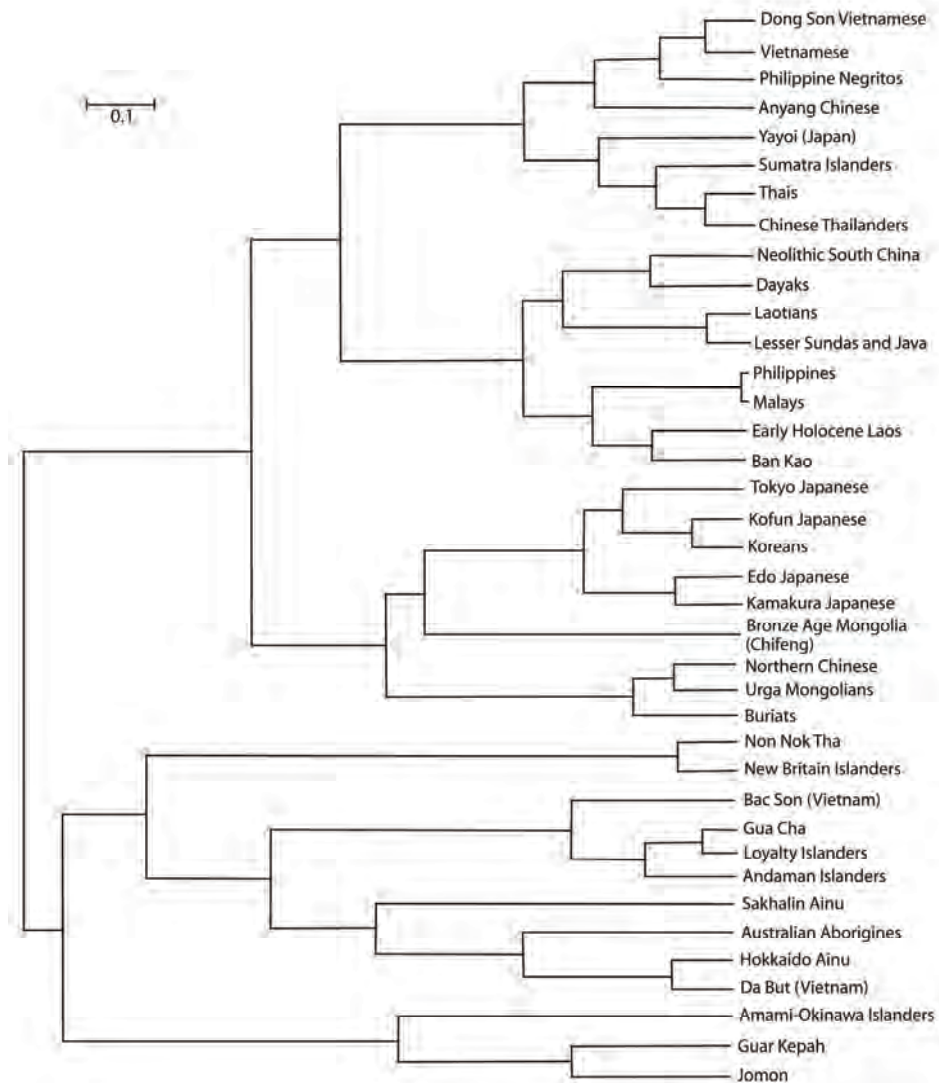


**Figure 3:** Seriated dendrogram (average linkage clustering) of dental morphology distances in Matsumura and Hudson (2004) (Correlation coefficient between the seriated order and a perfect seriation = 57.6%).

Figure 4 shows the seriated dendrogram derived from the dental metric shape distances published by Matsumura & Hudson (2004, Table 5). It suggests two useful improvements on the interpretation that those authors provide of their results. First, the cluster at the bottom of the dendrogram, which they characterise as Australo-Melanesian (including early Southeast Asians), also includes all their samples of Jomon and Ainu, who are indistinguishable from Australo-Melanesians in this analysis. To facilitate discussion, let us label this cluster 'non-Mongoloid', and the other cluster 'Mongoloid'. Secondly, the modern populations which seriate the farthest from the non-Mongoloids are Southeast Asians (here including Philippine Negritos), not Northeast Asians. Thus, whatever explanation might be advanced for the Mongoloid status of recent Southeast Asians in this analysis, genetic input from Northeast Asia might appear to fall short of the mark. Fortunately, as will now be seen, an explanation is possible from considering the point that three Bronze/Iron Age East Asian samples in the study (Dong Son, Anyang and Yayoi) are three of the five most Mongoloid samples (at the very top of Figure 4).

The Dong Son teeth are very distinctive from the Bac Son and Da But teeth (Figure 4) which both represent the Neolithic inhabitants of coastal North Vietnam. This would imply some level of post-Neolithic population incursion into North Vietnam by a people related to Bronze Age Chinese (as represented by Anyang). Similar population expansion from early historical China is evidently represented by Yayoi (with Japanese being the result of Yayoi/Jomon hybridisation), and may also be represented by the Bronze Age Mongolian

sample from Chifeng. Neolithic South China may well be a deep source for the transition to a ‘hyper-Mongoloid’ dental metrical shape across recent Southeast Asia, but its archaeologically attested effects date no earlier than the late Bronze Age (the age of Dong Son), or perhaps the Neolithic/Bronze Age transition at the far north of Southeast Asia (Man Bac, in North Vietnam). What clearly cannot be found in the dental metrical shape analysis of Matsumura & Hudson (2004), or in any currently available osteological data, is evidence for a Neolithic population expansion from China across Southeast Asia.



**Figure 4:** Seriated dendrogram (average linkage clustering) of dental metrical shape distances in Matsumura & Hudson (2004) (Correlation coefficient between the seriated order and a perfect seriation = 65.8%).

Intriguingly, as noted by Matsumura & Hudson (2004), their sample of ‘Early Holocene Laos’ teeth differs from most other prehistoric Southeast Asians in being metrically Mongoloid. They attempt to explain away this result by proposing that their sample possibly includes some Neolithic teeth, thereby obscuring the true non-Mongoloid status of the early inhabitants of Laos. However, there are a number of problems with this rationalisation. First, based on the available accelerator mass spectrometry dates from the remains (Demeter 2006), the time period covered by the Laos burials spans the Late

Pleistocene (Tam Hang, 15,700 years BP) to middle Holocene (Tamp Pong, 5400 years BP)—not early to late Holocene, as Matsumura and Hudson would wish. Secondly, were they justified in allowing for the presence of two distinct populations in the ‘Early Holocene Laos’ collection, this should be evident from other studies of these remains. What we find instead is that all ten early Laos crania studied by Demeter (2006) fell neatly into a single (C2) cluster. Thirdly, even if Neolithic remains have been caught up in Matsumura and Hudson’s ‘Early Holocene Laos’ sample, evidence that Southeast Asia’s Neolithic inhabitants were characterised by Mongoloid dental metrics is yet to be adduced.

Although the Neolithic Ban Kao teeth cluster with ‘Early Holocene Laos’ in Figure 4, along with other ‘Mongoloids’, this reflects a specific similarity between the early Holocene inhabitants of Laos and middle/late Holocene populations along the Mekong valley. Pietrusewsky (2006) found early Laos crania to be virtually indistinguishable from Ban Chiang crania, while Matsumura (2005) also found Ban Kao to be similar to early Laos in its craniometrics (as well as its dental metrics). In addition, Non Nok Tha is closer to early Laos and Ban Kao in its dental metrical shape (Matsumura & Hudson 2004, Fig. 2) than my Figure 4 would suggest. These comparisons all suggest an ancestor-descendant (or perhaps uncle-nephew) relationship between the early Laos and middle/late Holocene Thailand populations.

Osteological analysis accordingly offers strong support for the Mekong Austroasiatic dispersal route proposed by Sidwell and Blench (this volume). The populations involved appear non-Mongoloid in some analyses; for example, in its dental morphology, Ban Kao is most similar to Melanesians from New Britain (Matsumura 2005), and the early Laos teeth also presumably have a non-sinodont dental morphology (Figure 3). However, in other analyses they appear Mongoloid; for example, the Ban Chiang and Khok Phanom Di crania cluster together as marginally East Asian (Pietrusewsky 2006), and the Mongoloid dental metrics of the early inhabitants of Laos has been demonstrated above. These populations’ ambivalently ‘Mongoloid’ status nonetheless distinguishes them from the Guar Kepah and Hoabinhian/Neolithic Gua Cha samples, which all consistently show ‘non-Mongoloid’ tendencies regardless of the analysis (Figures 3 and 4, and see below).

Matsumura (2005) found that the inhabitants of Malaya would appear to have remained Australo-Melanesian till as late as the Early Metal Phase, based on the close similarity between Gua Harimau and Gua Cha in tooth size and shape. This finding is broadly compatible with the results of comparisons between prehistoric Malaya and Orang Asli osteology. Considering dental morphology, Rayner & Bulbeck (2001) show that the Semang join a cluster made up of New Guinea and European/North African samples, whereas the Aboriginal Malays (Temuan and Semelai) and Temiar Senoi are intermediate between this cluster and sundadonts, albeit closer to the latter. In their dental morphology, the Hoabinhian and especially Neolithic/Early Metal Phase samples are closer to the Temiar than any other Orang Asli population, a result interpreted as placing the immediate biological ancestry of the Semang to the north of where most of Peninsular Malaysia’s prehistoric human remains have been found (Bulbeck & Lauer 2006). However, when the Guar Kepah and Gua Cha Neolithic teeth are analysed separately, the former appear equally similar to both Aboriginal Malays and the Temiar (Bulbeck 2005a), even if the distinctive Temiar affinity of the latter is attested (Bulbeck 2005b). Dental metrical analysis, on the other hand, finds the Semang to be very close to all male prehistoric Peninsular Malaysia teeth (except those from Guar Kepah) in terms of shape, but of course not size (the Semang have much smaller teeth, in line with their reduced body size). The mixed-sex Guar Kepah sample, on the other hand, clusters with Khok Phanom Di in terms of dental metrical shape and emerges as more ‘Mongoloid’ than any Orang Asli (Bulbeck

et al. 2005). To combine these results with those of Matsumura & Hudson (2004) (see Figure 4), both prehistoric Malayan and Orang Asli teeth would be characterised as non-Mongoloid in their metrical shape.

In their craniometrics, the three Orang Asli divisions resemble each other in having small crania, a braincase which is modally of middling shape (but varying from narrow to broad), and a short face and broad nasal aperture but tall orbits (Bulbeck & Lauer 2006). Compared to non-Malaya populations, approximately equal numbers resemble Andamanese, Mongoloid Southeast Asians and Africans, with no suggestions of a difference between the divisions except perhaps a specific likeness between Aboriginal Malays and Andamanese (Table 6). Andamanese and Southeast Asian affinities are not found amongst prehistoric Malay Peninsula crania, which instead resemble the crania of Africans, Tasmanians, New Britain Melanesians, and Mongoloid populations in Easter Island and the New World (Bulbeck & Lauer 2006), as well as the Orang Asli. The main shape changes over time—narrowing of the face and nose, and slight broadening of the braincase—may be related to reduction in cranial size, which finally resulted in a proportion of Orang Asli crania metrically similar to recent Andamanese and other Southeast Asians. Certainly, there appears to have been a trend towards less robust crania over time, which can be attributed to cranial- and, ultimately, body-size reduction (Bulbeck & Lauer 2006).

In summary, the osteology of the Orang Asli indicates that their ancestry predominantly lies with Malaya's prehistoric inhabitants, reaching back to the Hoabinhian. However, there are also indications of genetic input from beyond Malaya that has contributed to the Orang Asli, and especially the Senoi and Aboriginal Malay, gene pool. This overall picture is consistent with the genetic distances in Lie-Injo (1976), which, when analysed, show that his Semang, Senoi and Aboriginal Malay samples all seriate together, notwithstanding a Mongoloid-leaning tendency shown by the Senoi (consistent with the findings of Saha et al. 1995) and especially Aboriginal Malays (Bulbeck 1996). The incorporation of extra-Malaya genes in the Orang Asli gene pool would be expected from their linguistic affinities and absorption of mortuary and technological innovations. In addition, as emphasised by Rambo (1988), all Orang Asli populations tend to show a diversity of physical appearances which crosses the three divisions. Whether due to recent 'mixed marriages' (Schebesta & Lebzelter 1928; Noone 1939) or earlier conjugal relations, this observation points to a network of gene flow that has linked all Orang Asli populations (see also Fix this volume).

Mitochondrial DNA (mtDNA) analysis is particularly useful here because it allows for a particulate view on a population's composition, based on its maternal lineages. This approach frees the analysis from the broad-brush classification of groups into macropopulations, which, as we have seen above, leads to a confusing array of terms (Australo-Melanesian, Mongoloid, non-Mongoloid, sundadont, sinodont, non-sinodont) in trying to extract useful information from the complex realities of how groups may resemble or differ from each other depending on the traits being considered (Bulbeck et al. 2006). This approach further releases the analysis from the assumption of episodic prehistoric change (as in Bellwood's Mongoloid migration scenario). The phylogeographical approach detailed by Oppenheimer (2003, this volume) additionally allows donor and recipient regions to be identified, as well as estimation of the time of arrival of a maternal lineage at its recipient destination.

**Table 6:** Classifications of Orang Asli crania relative to the Howells (1973b) populations using Fordisc 2.0 (Ousley & Jantz 1994). The crania's tribal affinities are from Evans (1937:269) for the Siong Semang and otherwise from Bulbeck (1996), while their Fordisc classifications are from Bulbeck (2004b)

Orang Asli group	Andaman Islanders	Sub-Saharan Africans	Southeast Asians	Miscellaneous	Number tested
Kensiw/Kintaq (Maniq)		1 Dogon, 2 Bushman	1 Atayal, 1 Philippines	2 Ainu, 1 Santa Cruz	8
Jahai	1	1 Teita, 1 Zulu	1 Atayal	1 Zalavar (Hungary)	5
Menriq	1				1
Lanoh	1				1
<b>Semang</b>	<b>3</b>	<b>5</b>	<b>3</b>	<b>4</b>	<b>15</b>
Temiar	2		2 Atayal		4
Semai		2 Teita, 1 Dogon, 1 Zulu		1 Easter Island	5
Jah Hut			1 Atayal		1
<b>Senoi</b>	<b>2</b>	<b>4</b>	<b>3</b>	<b>1</b>	<b>10</b>
<b>Aboriginal</b>	<b>2</b>				<b>2</b>
<b>Malays (Jakun)</b>					

## 7 Mitochondrial DNA analysis

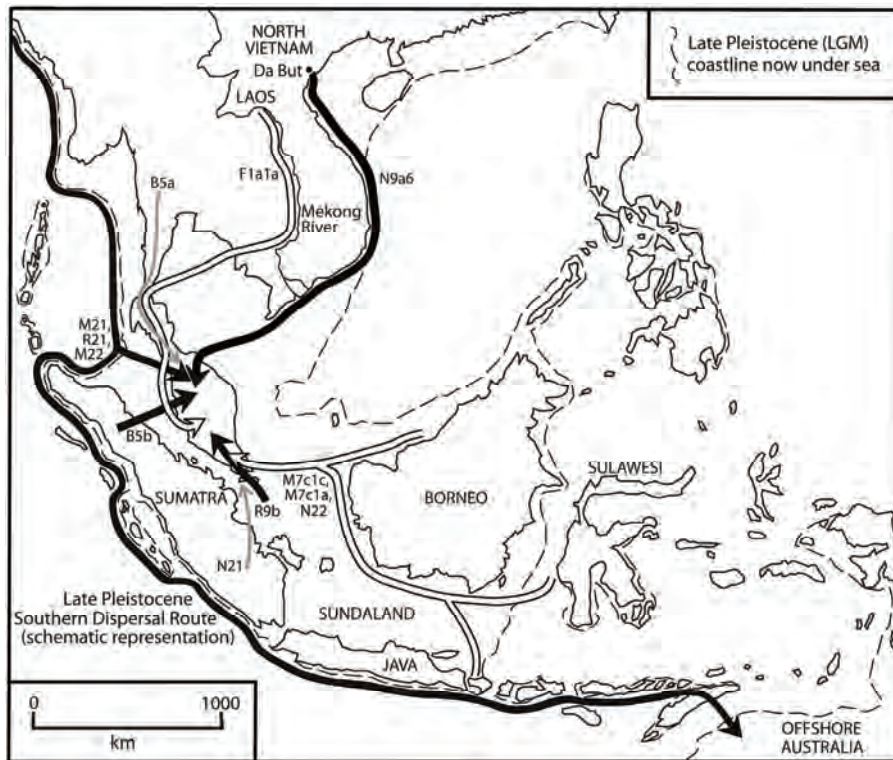
Limited studies on Orang Asli mtDNA were first published by Ballinger et al. (1992), Melton et al. (1995), Oota et al. (2001) and Zainuddin & Goodwin (2003). The study by Hill et al. (2006) has the advantages of extensive sampling of the Orang Asli (three Semang, one Senoi and two Aboriginal Malay groups) as well as Malay populations in Sumatra, which is the likely homeland of the Melayu Malays. As detailed by Hooker (2003), the Melayu sequentially established an early historical capital at Palembang in Sumatra, a medieval capital at Tumasik in Singapore, and their capital at Melaka in the fifteenth century CE.

Hill et al. (2006) identified a small number of haplogroups, particularly well represented among the Semang, which would appear to reach back to the Late Pleistocene dispersal of *Homo sapiens* from Africa along the northern rim of the Indian Ocean to Australia (Oppenheimer 2003; cf. Lahr & Foley 1998). The majority of haplogroups, however, were found to have a terminal Pleistocene to Holocene time depth in Malaya, contrasting with the Late Pleistocene age of their root type elsewhere in Southeast Asia. Hill et al. (2006) related these introduced haplogroups to the Hoabinhian, (late) Neolithic, Malayic and Melayu cultural horizons. Some review of their conclusions may be warranted in light of subsequently published research on Island Southeast Asian mtDNA haplogroups by Hill et al. (2007), which has far better sampling of populations across the region than any preceding study on the topic, and in light of the detailed review of Malaya's archaeology presented here.

Two ancient haplogroups with their root type concentrated amongst the Orang Asli are M21 and R21 (Table 7). Their much lower representation in Sumatra or among the Melayu rules out a Sumatran source. Their cultural counterparts are the core- and cobble-based stone tool assemblages which led to Malaya's 'bifacial Hoabinhian' by 15,000 years ago

(Table 1). The very rare M22 haplogroup may come under the same rubric even though it is a feature of the Temuan, based on present data, whereas the M21 and R21 haplogroups characterise the Semang more than other Orang Asli groups (Table 7).

Both the R9b and N9a6 haplogroups are traced by Hill et al. (2006) to LGM ancestral types in Indochina, and have a stronger representation amongst the Orang Asli (especially Aboriginal Malays) than anywhere in Sumatra, other parts of Island Southeast Asia (Table 7) or Thailand (Hill et al. 2007). However, their hypothesised arrival in Malaya with the introduction of the Hoabinhian (Hill et al. 2006) is contradicted by archaeological evidence for the predominantly local development of the Malayan Hoabinhian. R9b has a disjunct distribution with concentrations in (1) northern Vietnam and central Thailand, and (2) amongst the Temuan and Semelai living to the south of any of the Peninsula's recorded Hoabinhian sites (cf. Oppenheimer this volume with my Figure 1). One likely scenario explaining this distribution would posit a terminal Pleistocene dispersal of R9b from northern Vietnam (then host to Hoabinhian and Bacsonian industries) or central Thailand to central-western Sundaland, and its northward retreat (associated with an undocumented stone tool industry) into southern Malaya as postglacial sea levels rose (Figure 5).



**Figure 5:** Hypothesised routes of entry of Orang Asli mtDNA haplogroups into the Malay Peninsula.

The introduction of the N9a6 haplogroup to Malaya, on the other hand, dates to the same age as the early Neolithic, not only in Malaya but also in North Vietnam. Every characteristic of Malaya's early Neolithic is echoed at North Vietnam's Bacsonian and/or Dabutian sites (Table 8), which suggests a linked introduction of the N9a6 haplogroup and the Neolithic from North Vietnam to Malaya. Archaeologically, the Bacsonian makes the best comparison with Malaya's early Neolithic, but osteologically Guar Kepah and Da But are particularly similar. Guar Kepah compares well with both Bac Son and Da But in tooth size (Matsumura & Hudson 2004, Fig. 3), and particularly well with Da But in terms of

tooth shape (Figure 4), elongated cranial form and Ainu/Jomon-like dental morphology (cf. Bulbeck 2005a and Matsumura 2005). N9a6 is also sporadically present in Island Southeast Asia (Table 7), which correlates with the status of the Da But culture as the likely source for pre-Austronesian Neolithic introductions to Island Southeast Asia, apparently reflecting coastal Vietnam's access to mid-Holocene sailing technology (Bulbeck 2008). Haplogroup N9a6 is therefore inferred to mark a maritime link between coastal Vietnam and Malaya at around 6000 BP (Figure 5).

Sumatraliths have a restricted chronological and geographical presence in Malaya, and probably source to early Holocene Sumatra. Their introduction to Malaya should find its biological echo in an mtDNA lineage more notable in Sumatra (retained there even after the island's colonisation by Austronesian speakers) than elsewhere in Island Southeast Asia (where sumatraliths are lacking). The B5b lineage, which is essentially a Semang-related lineage in Malaya (Table 7), fits the bill. The ancestral type is found in eastern Indonesia and otherwise in Sumatra; more specifically, it is found in Medan (Hill et al. 2006), near Sumatra's concentrated zone of Hoabinhian shell middens (Heekeren 1972). Although Hill et al. (2006) avoid proposing a specific time depth for the arrival of the B5b lineage in Malaya, their data would be consistent with an early Holocene date (see also Oppenheimer this volume).

The F1a1a haplogroup, on the other hand, is well represented at locations along the Mekong valley (see Oppenheimer this volume) as well as among the Senoi. As discussed above, the early inhabitants of Laos appear to have been a suitable biological precursor for Thailand's late Neolithic to Bronze/Iron Age populations. In combination with the Mekong route for early Austroasiatic dispersal (Sidwell & Blench, this volume), these points confirm the inference by Hill et al. (2006) of an association with the (late) Neolithic of Mainland Southeast Asia and the introduction of Austroasiatic to Malaya. However, in that case, the virtual absence of F1a1a from southern Thailand (Oppenheimer this volume) would contrast with its high frequencies amongst the Austroasiatic-speaking Temiar Senoi and the Nicobar Islanders immediately north of Sumatra (Prasad et al. 2001). This line of argument would contradict Bellwood's (1978, 1993) vision of south Thailand's 'Ban Kao culture' as the immediate source of early Austroasiatic speakers in Malaya, and instead imply that they had immigrated overseas across the Gulf of Siam (Oppenheimer, this volume; my Figure 5). Entry into Malaya may have been via the Isthmus of Kra, where tripods arrived at about the same time (see above), and which subsequently acted as a major node in a late prehistoric interaction sphere that extended across the South China Sea (Bellina-Pryce 2009).

Hill et al. (2006) also drew attention to three haplogroups, M7c1c, N22 and N21, with middle to late Holocene antiquity in Malaya, where they particularly characterise Aboriginal Malays (Table 7). The M7c1c and N22 haplogroups possibly arrived from southern Island Southeast Asia east of Sumatra. They are reasonably associated with early Austronesian maritime networks. The linguistic echo is Benjamin's (2002) detection of Malayic (proto-Malay) lexical items among southern Aboriginal Malay groups, while the archaeological marker is southern Malaya's red pottery and early watercraft (see below). The M7c1a haplogroup may also belong here given its trace presence amongst the Semelai (Hill et al. 2006) and geographically closest, external presence in Taiwan and Borneo (Hill et al. 2007).

Haplogroup N21, however, evidently has origins in Sumatra (Hill et al. 2007) although its absence from the Melayu (Table 7) points to a means of introduction separate from the Melayu immigration. The restriction of N21 to Palembang (within Sumatra), the old Srivijaya capital, suggests its introduction via influence that Srivijaya had exerted on

proto-historical Malaya. B5a may also have entered Malaya at this time from south-central Thailand where it is present at concentrated frequencies (Fucharoen et al. 2001). B5a is recorded as root types amongst the Batek and Temiar, and amongst the Melayu and in Sumatra both as root types and a derived lineage otherwise characteristic of Austroasiatic speakers in Indochina (Hill 2005). Northern Malaya was evidently influenced by the early Buddhist Dvaravati civilization of central Thailand, which is traditionally associated by scholars with Mon speakers (Indrawoath 2004). Influences include such aspects as Buddhist votive tablets recorded from the Isthmus of Kra to Gua Chawas, and the monumental architecture, Sanskrit inscriptions and replete Hindu-Buddhist imagery of the Bujang Valley/Sungei Mas state (Bulbeck 2004b). Dvaravati influence on northern Malaya (Benjamin 1987) would explain the introduction of B5a lineages to the Orang Asli, excluding Aboriginal Malays, as well as to the Melayu whose expansion took them well north of the present-day Malaysia-Thailand border.

The signature of haplogroups introduced by the Melayu Malays to the Peninsula would be a significantly more pronounced presence amongst the Melayu, both in the Peninsula and in their Sumatra homeland, than amongst the Orang Asli. The M\* unclassified lineages (Hill et al. 2006) particularly fit the bill, but there are also numerous haplogroups recorded at low frequencies amongst the Peninsula and Sumatra Melayu, which do not appear to have been transmitted to the Orang Asli. These include the B4a\*, B4b, B4c, E1, E2, F1a\*, F1a3, M7b1, M45, M46, N\*, R9\* and Y2 haplogroups, all of which had dispersed across Island Southeast Asia at various times following the LGM (Hill et al. 2007 and references therein).

The present re-analysis of the mtDNA haplogroup data leads to the chronological schema summarised in Table 9. To assist interpretation, we may assume that a group's socio-political marginalisation and/or geographical isolation are marked by its apparent inability to recruit wives from an immigrant group (represented by novel haplogroups) even if the immigrants recruited wives from the marginalised group. As shown in Table 9, over half of the Semang female gene pool dates back to the Late Pleistocene, and virtually all of the remainder dates to around the first half of the Holocene. On that basis, the Semang would appear to have been marginalised from the late Neolithic onwards. The Senoi, to judge by the Temiar, would appear to be closely linked to the late Neolithic developments that led to the marginalisation of the Semang. Approximately half of the Temiar female gene pool can be ascribed to late Neolithic immigrants to Malaya, even if the other half has local roots dated to between the Late Pleistocene and early/mid-Holocene. The Temiar, in turn, appear to have become marginalised in early historical times (Table 9).

The Aboriginal Malays evidently represent a tradition of staying abreast with changing circumstances, as around 20% of their female gene pool dates to the Late Pleistocene, 30% to the early Holocene, 20% to late prehistoric times and 25% to the early historical period. However, they were marginalised with the establishment of the Melayu as Malaya's dominant population. The Melayu clearly mark a major expansion of an immigrant population, with 50% of their female gene pool attributable to their Sumatra-based migration, and 40% reflecting the recruitment of wives from the Orang Asli and Dvaravati-associated groups further north.



**Table 7:** mtDNA haplogroups recorded in Malaya, based on Hill et al. (2006, 2007). ‘Dominant’ means frequencies above 40%; ‘strong’ means frequencies between 11% and 40%; ‘slight’ means presences of 10% or less. Antiquity estimates are approximations; where Hill et al. (2006) provide the standard error (S.E.) for their estimate of a lineage’s arrival in Malaya, antiquity is shown as one S.E. range.

Haplogroup	Semang presence	Temiar Senoi presence	Aboriginal Malay presence	Melayu presence	Representation in Sumatra	Representation in central/eastern Island Southeast Asia (crISEA)	Source	Antiquity in source	Antiquity in Malaya
M21	All 3 (strong to dominant) + Ten'en	Slight	Slight	Slight	Virtually absent	Sporadic slight presence	Ancient Sunda	57 ka	57 ka
R21	All 3 (slight to dominant)	Strong	Absent	Slight	Absent	Absent	Ancient Sunda	60 ka	60 ka
M22	Absent	Absent	Temuan (strong)	Absent	Absent	Absent	Ancient Sunda	63 ka	63 ka (?)
R9b	Absent	Absent	Strong	Slight	Absent to slight	Absent	Indochina	19 ka	6–12 ka
N9a6	Jahai (strong)	Slight	Slight to strong	Slight	Absent to strong	Borneo, Java, Sulawesi (trace)	Indochina	Similar to above	3–8 ka
B5b	Batek (dominant), Menriq (slight)	Absent	Absent	Slight	Medan (strong)	Absent to slight	Sumatra	35 ka	Holocene
F1a1a	Absent	Dominant	Slight	Slight	Absent to strong	Slight (widespread)	Indochina	9 ka	≤7 ka
M7c1c	Absent	Absent	Semelai (strong)	Slight	Absent to strong	Slight to strong	crISEA	8 ka	Mid-Holocene
N22	Absent	Absent	Temuan (strong)	Absent	Absent	Sulawesi (slight)	crISEA	Unclear	Mid/late Holocene
N21	Absent	Absent	Strong	Slight	Slight (Palembang)	Sporadic slight presence	Sumatra	30 ka (ISEA)	Late Holocene
B4*	Absent	Absent	Semelai (slight)	Absent	Absent	Sporadic slight presence	Thailand? (see text)	Unassessed	Late Holocene
M*	Absent	Slight	Slight	Strong	Strong everywhere	Slight to strong	crISEA	Late Pleistocene	Late Holocene
B5a	Slight	Slight	Absent	Strong	Absent to strong	Absent to slight	crISEA	9 ka	Late Holocene
Various (see text)	Absent	Absent	Absent	Slight	Slight to strong	Absent to strong	crISEA	Variable	Late Holocene

**Table 8:** Comparisons between the early Neolithic of Malaya (this chapter) and North Vietnam (abstracted from Nguyen et al. 2004; Nguyen 2005; Bower et al. 2006)

Criterion	Malaya early Neolithic	North Vietnam early Neolithic comparison
Chronology	6000 BP (cf. age of N9a6, 5500 BP)	Dabutian (6500–4500 BP) a better fit than the Bacsonian (11,000–7000 BP)
Pottery	Cordmarked	Cordmarked in the Bacsonian; fibre-impressed and later cordmarked in the Dabutian
Polished/ground stone tools	Incompletely polished (mainly ground edges)	Edge-ground axes in the Bacsonian and early Dabutian; more extensively polished in the late Dabutian
Site type	Rockshelters and large middens	Rockshelters (Bacsonian) and large middens (Dabutian)
Plant cultivation	Arboriculture	Possible simple plant cultivation (Bacsonian); gardens, including rice consumption by 5000 BP (Dabutian)
Burials	Secondary	Bacsonian (secondary, sprinkled with ochre) a better fit than the Dabutian (flexed)
Human osteology	Non-Mongoloid (Guar Kepah)	Dabutian a better match than Bacsonian (see text)

**Table 9:** Combined haplogroup frequencies excluding the unclassified B\*, B4\* (mislabelled B4a), M\* and N\* haplogroups (taken from Hill et al. 2006) amongst the Orang Asli and Melayu Malays

Time of establishment in Malaya	Semang		Senoi	Aboriginal Malays		Melayu Malays	
	Batek	Jahai	Menriq	Temiar	Semelai	Temuan	
Late Pleistocene (M21, R21, M22)	0.52	0.82	0.94	0.45	0.13	0.27	0.07
Early Holocene (R9b, B5b, N9a6a)	0.45	0.18	0.06	0.06	0.30	0.33	0.05
Late prehistoric (F1a1a, M7c1c, N22, M7c1a)	0.00	0.00	0.00	0.43	0.22	0.12	0.14
Early historical (N21, B5a)	0.03	0.00	0.00	0.02	0.31	0.15	0.12
Melaka period (other haplogroups)	0.00	0.00	0.00	0.00	0.00	0.00	0.51

## 8 A synthesis of the evidence

Using Table 9 as a framework, we can now synthesise the biological and cultural evidence on Malaya’s population history. Following the colonisation of Malaya by *Homo sapiens*, which the archaeological evidence would place before 74,000 years ago (see above) but which the mtDNA clock may place as late as 55,000 years ago (Macaulay et al. 2005; Hill et al. 2006; Soares et al. 2009), the region’s inhabitants appear to have remained

in isolation, or at least unaffected by major population incursions, for tens of millennia. The descendants of these original *H. sapiens* colonists are well represented in the female gene pool of all Orang Asli groups, especially the Semang (Table 9). Archaeological remains dating to the LGM are yet to be certified, but this can be attributed to a coastal and lowland orientation of the inhabitants. The introduction of the R9b and B5b haplogroups to the Aboriginal Malays and Semang, respectively, can be related to lowland inhabitants' inland retreat as sea levels rose at the end of the Pleistocene. Early Holocene human remains, which include the flexed Moh Khiew 2, Gua Gunung Runtuh and Gua Teluk Kelawar burials, attest to full-sized people over ten percent taller than the average Orang Asli of ethnographic (British colonial) times.

There is considerable archaeological evidence for an early Neolithic phase in Malaya dated to c. 5000–6000 years ago, linked to the early Neolithic of North Vietnam. In southern Thailand, the early Neolithic is represented by pottery and polished stone tools at several mid-Holocene sites, as well as phytolith evidence for intensified arboriculture. Similarly, the Guar Kepah teeth reveal betel-nut chewing and, based on the high dental caries rate, evidence for high levels of carbohydrate consumption (Bulbeck 2005a). Extension of the early Neolithic to Malaya's remote hinterland is represented by early pottery, edge-ground cobble tools and haematite-coated burials at Gua Peraling. The N9a6 haplogroup, with its origins in Indochina, can be linked to the widespread early Neolithic given that it is recorded amongst all of the Semang, Senoi and Aboriginal Malays (Table 7). The flexed and secondary burials dated to c. 5000 years ago from numerous sites indicate a population of smaller body size than their early Holocene ancestors, probably reflecting biological adaptation to Malaya's equatorial humidity.

The Aslian homeland is parsimoniously located in the vicinity of central-western Pahang and southern Selangor (Figure 1) as this is the area of maximum diversity of the Aslian languages (Bulbeck 2004a:377). This homeland would also be consistent with an Austroasiatic association of the F1ala haplogroup, as it is present amongst the Semelai, Temuan and Jakun to the south as well as the Semai and Temiar to the north (Hill et al. 2006). Neolithic extended burials were interred at Lang Rongrien, Gua Harimau and Gua Cha (and, as argued above, presumably in association with Jenderam Hilir and Nyong), while pollen and palynological evidence is consistent with slash and burn agriculture having been established by 4000 years BP in both southern Thailand and Singapore. The Neolithic Gua Cha burials show clear evidence of arboriculture in the guise of betel-nut stained teeth, plus recourse to cooking in pots by those suffering from poor oral health, and a higher carbohydrate component—possibly honey and forest fruit—in their diet (Bulbeck 2005b). A mixed forager-farmer economy and semi-sedentary lifestyle, similar to that of the Ceq Wong (see Burenhult et al., this volume), would appear to have spread north and south from the Aslian homeland. Wives from resident foraging groups (with or without an arboriculture component) were evidently absorbed by these land-hungry colonists, but without reciprocal spouse exchange to the foragers who were losing access to optimal patches of land.

There is no reason to doubt the cultural and biological discreteness of the main Semang groups, whose traditional ranges extended in a chain from the Ten'en of southern Thailand, across the border to the Kensiw/Kintaq, Jahai, Menriq and Batek of northern and central Peninsular Malaysia (Figure 1). Biologically the Jahai, Menriq and Batek form a discrete group in their mtDNA haplogroup frequencies (Hill et al. 2006) and dental metrics (Bulbeck et al. 2005), and they are also homogeneous in their dental morphology (Rayner 2000) and craniometric affinities (Table 6). With respect to their culture, not only are these Northern Aslian groups linguistically related but also they traditionally foraged the

rainforest in nuclear family groups equipped with blowpipes, iron blades and digging sticks as their main tools. All of this evidence looks very much like a population expansion, the only conceivable objection being that foragers are not supposed to expand. Any other interpretation would posit the unparsimonious scenario of multiple switches to a Northern Aslian language along a chain of closely related people. A scenario of population expansion, on the other hand, would require only one such switch, by a rainforest forager group in contact with early Northern Aslian speakers somewhere between south Thailand and central Peninsular Malaysia, followed by expansion along a lowland rainforest corridor that may have been barely inhabited (almost all of Malaya's Hoabinhian sites lie west of the Semang groups discussed here), aided by acquisition of the blowpipe (Bulbeck 2003). The lack of haplogroups postdating 5000 years ago amongst the studied Semang groups could accordingly be explained by their geographic isolation as well as their socio-political marginalisation.

Having explained the ethnographic distribution of the Northern Aslian Semang, we can turn to the late Neolithic archaeological record to their west to investigate Central Aslian cultural evolution. The Central Aslian speakers due north of the Aslian homeland include the Lanoh and Semnam Negritos as well as Semai and Temiar Senoi. Unfortunately, osteological differences between the c. 4000 BP populations of Thailand and of Malaya, and between the Orang Asli divisions, are too subtle to conclusively assess any osteological sample (for example, Neolithic Gua Cha) as mainly Indochina-derived versus indigenous, or as more Semang- than Senoi-related. Nonetheless, the impression we obtain from the archaeological data is the co-existence of complementary traditions. Gua Harimau, with its minimal habitation evidence, would appear to have been a burial site used by a sedentary population in the vicinity (proto-Semai?). Gua Baik and Gua Kerbau, on the other hand, testify to continuous occupation by foragers (proto-Lanoh/Semnam?) who obtained pottery, iron and other manufactured goods through trade. Finally, Neolithic Gua Cha combines fine grave goods and possible evidence for local manufacture of polished stone adzes (Sieveking 1954) with a predominantly foraging and arboricultural economy (as noted above). This suggests a frontier colony, probably proto-Temiar (Benjamin 1987:118), where old and new traditions were seamlessly combined, especially given the site's status as a focus for the local population both before and during the Neolithic. Despite the material prosperity reflected in the Gua Harimau and Gua Cha grave goods, stature reduction evidently continued between the middle and late Holocene (Tables 4–5). Indeed, Malaya's late Neolithic inhabitants appear to have been shorter than their contemporaries at Ban Kao (Bulbeck 1996:46), Ban Chiang, Non Nok Tha, Ban Na Di, Ban Lum Khao, Khok Phanom Di and Nong Nor in Thailand (Higham 2002:165); the Malaya male average (159 cm) is the shortest recorded, while the female average (156 cm) sits comfortably within the Thailand range of averages (152–157 cm).

Dunn (1964) noted the presence of red-slipped pottery in the upper 14 inches at Gua Kecil, which he interpreted as reflecting late Neolithic or Early Metal Phase influence. Red pottery is also recorded in the early excavation reports from Kota Tongkat and Gua Sagu but not from any sites to their north (Matthews 1961). Red-slipped pottery is one of the markers of Austronesian linguistic expansion (Bellwood 1997; Bulbeck 2008), and so its occurrence in Peninsular Malaysia suggests an Austronesian association. As red-slipped pottery evidently appeared late in the Malaya sequence, and in the vicinity of Gua Kecil southwards (where Aboriginal Malays, both Southern Aslian and Malayic, are located), an association with the introduction of haplogroups M7c1c, N22 and M7c1a appears likely. While the prehistoric archaeology of southern Peninsular Malaysia is scanty, a new type of site—featuring boat remains recovered from waterlogged coastal deposits—dates as of c.

2000 years BP at the sites of Pontian, Kelang and Sungei Lang (Bulbeck 2003). Boat transport is vital in southern Malaya where the rivers are often navigable up to their headwaters and provide natural ports where they debouch (Gianno & Bayr 2009).

Iron Age remains that are probably linked to an Austronesian presence occur to the immediate west of Gua Kecil. The Perak/Selangor foothills include four main locations with stone slab graves which resemble the central Java slab graves with their extended burials and mortuary goods (Bellwood 1997:290). These slab graves cross Temuan and lowland Semai territory (Figure 1). The lowland Semai are the only Orang Asli group numbered amongst Malaya's peasantry (Benjamin 1985). The lowland Semai probably provided a labour pool and agricultural produce to the local elite, who had been attracted by these foothills' tin, gold and high-quality iron ore, and may have been Austronesian speakers (Bulbeck 2004b). Minerals exploited in the Perak foothills may have been exported through the entrepot of Kuala Selinsing to the immediate west (Leong Sau Heng 2000). Kuala Selinsing, which features canoe burials plus pottery decorations particularly similar to those from the 2000-year-old maritime trading site of Sembiran in Bali (Bulbeck 2006), adds to the picture of a maritime proto-historical influence from Island Southeast Asia on southern and coastal Malaya. However, Southern Perak seems to mark the northern limit of this early Austronesian influence, for the associated M7c1c, N22 and M7c1a haplogroups did not travel as far north as Temiar territory (Table 7).

Specific parallels between the proto-historical archaeology of Malaya and Sumatra are sparse. For instance, the megaliths of Iron Age Sumatra are far more diverse than Malaya's, and the slab graves in these two regions have very different architecture (see Hoop 1932). Nonetheless, interaction between these two landmasses which flank the Melaka Strait, so critical to Southeast Asia's maritime trade, is undoubted. Malaya's importance as a source for natural resources, which evidently included tin as well as forest products, is clear from its many early historical toponyms. Inscriptions in Old Malay located in Singapore and Nakhon Si Thammarat, respectively at the immediate south and north of Malaya, suggest that Srivijaya exerted an influence on the southern lowlands and eastern seaboard of Malaya (Bulbeck 2004b). Malaya's fragmentary Dong Son bronze drums, recovered from Kelang on the southwest coast, Terengganu on the east coast, and the upper reaches of the Tembeling River (which debouches on the east coast), may well reflect a shared maritime trade route with South Sumatra where further Dong Son drums, including their depiction in megalithic statues, have been found (see Bellwood 1997). An additional similarity involves the ceremonial bronzes with incised designs found at Kelang and Kampung Pecu in southern Malaya (shaped like bells) and the southern half of Sumatra (shaped like flasks). These observations can explain the presence of the N21 haplogroup, which evidently reflects a connection between Aboriginal Malays and Sumatra prior to the Melayu immigration to Malaya.

The Aboriginal Malays are diverse in terms of (1) their coastal versus hinterland orientation, (2) the manner in which their local economy balances swidden agriculture with the collection and trade of forest produce, (3) their strategies for dealing with local Melayu dominance, and (4) the distinction between Southern Aslian and Malayic speakers (Benjamin 2002; Gianno & Bayr 2009). Their Melayu-like physical appearance (Martin 1905; Cole 1945; Rambo 1988) is a main reason why Aboriginal Malays have been treated as a single ethnological bloc. The presumption of biological homogeneity is supported by the close resemblance between the Malayic-speaking Temuan and Aslian-speaking Semelai in their dentition (Rayner & Bulbeck 2001; Bulbeck et al. 2005) and mtDNA lineages (Hill et al. 2006). The ethnographically recorded short stature of Aboriginal

Malays (Appendix A) is true of the Malayic-speaking Temuan and Jakun as well as the Aslian-speaking Besis (Martin 1905:233–234).

To judge by the historical land-use strategies of the Semelai (Gianno & Bayr 2009), the Aboriginal Malays would have experienced selection pressures favouring small stature owing to a traditional focus on nutrient-poor root crops and a proclivity to seek out primary rainforest when communities resettle or break off. In addition, the Semelai and many other Aboriginal Malays were targeted by the Melayu as sources of slaves as fervidly as were the Semang and Senoi ‘Sakai’ groups (Benjamin 2002; Hooker 2003; Gianno & Bayr 2009). Slave raiding in particular drove the Orang Asli into marginal and/or isolated terrain beyond the reach of the Melayu who increasingly commandeered Malaya’s habitat most conducive to population growth (Bulbeck 2004b). The Melayu were the last pre-colonial episode in a late Holocene population history characterised by new arrivals who absorbed a significant part of their female gene pool from the previously established denizens, and no doubt contributed to the indigenous gene pool through various acts of fathering, but kept their women away from the populations which they had displaced.

## 9 Summary

Over 50% of the Semang gene pool, as represented by the mtDNA of three Northern Aslian groups, has Late Pleistocene roots that reach back to Malaya’s colonisation by *Homo sapiens*. However, it would be a mistake to think of the Semang as the barely changed relics from those early days. The Pleistocene occupants inhabited a cooler drier environment, particularly during the LGM, and were oriented toward the lowlands on or near the coast. Also, all three studied Semang groups have mtDNA haplogroups whose time of introduction to Malaya is dated to around the early Holocene. At that time, based on the available human remains, Malaya’s inhabitants differed from the Semang in being taller and having larger teeth, more robust crania, broader faces and more elongated braincases. As discussed elsewhere (Bulbeck 2003), rockshelter sites bear witness to habitation across Malaya by the early Holocene, in contrast to the evidence for a sporadically distributed Pleistocene population.

To explain how full-sized people could have dispersed across Malaya just as the rainforest expanded, and maintained a viable foraging strategy in an environment where selection pressures for reduced body size would have exerted their influence, a model was developed in which ungulates (well represented in the faunal refuse) were the target of intercept hunting along well-maintained forest trails (Bulbeck 2003). Keeping the forest at bay would have been assisted by the tree-felling functionality of Hoabinhian cobble tools, including sumatraliths which appeared during the early Holocene. However, the most dramatic changes to affect the immediate ancestors of the Semang occurred later, with the arrival of the blowpipe and iron, and the loss of prime terrain to more settled populations. While the Lanoh and Semnam appear to represent an adaptation involving complementary co-habitation with swidden farmers, the Northern Aslian Semang evidently staged a late Holocene population expansion along a sparsely inhabited corridor of lowland rainforest within central Malaya’s interior.

The archaeological evidence suggests two Neolithic phases in Malaya, neither of them sourced to Ban Kao, which instead should be viewed as merely a sister branch of Malaya’s late Neolithic. Malaya’s early Neolithic, documented at Guar Kepah and particularly at various southern Thailand sites, is associated in this study with the Da But culture of North Vietnam. Genetic evidence for a mid-Holocene population incursion comes from the N9a6a haplogroup, present amongst the Jahai and Temiar as well as Aboriginal Malays.

The advent of arboriculture by the mid-Holocene, indicated by phytolith and osteological evidence, as well as cordmarked pottery and lightly polished cobble tools are all aspects of Malaya's early Neolithic.

A late Neolithic immigration from the Mekong delta region, associated with the introduction of extended burials, pedestalled pots, and by association the F1a1a haplogroup and proto-Aslian, then set the scene for trenchant differentiation between Orang Asli populations in competition for the same terrain. Evidence for slash-and-burn agriculture by 4000 years ago, which included rice within the repertoire of crops, comes from northern Malaya (phytolith evidence, and the Sakai Cave Neolithic rice grains), central Malaya (open-air settlements, and a proto-Aslian term for rice), and Singapore (increased charcoal deposition). The F1a1a haplogroup, which strongly characterises the Temiar Senoi, appears not to have been picked up by any Northern Aslian groups although it occurs among Aboriginal Malays. It is likely that early Aslian-speaking communities with a mixed economy dispersed widely across western and southern Malaya, more or less as proposed by Bellwood (1997), but their impact in the north and south was masked by proto-historical developments (Benjamin 1987). Accordingly the Temiar, whose ethnographic location sits virtually in the centre of Malaya, provide the clearest genetic signal we have for a late Neolithic Austroasiatic incursion. Nonetheless they can hardly be treated as a 'Neolithic fossil' as they subsequently evolved culturally, for instance, replacing pottery with bamboo containers (Benjamin 1987:118) and practising tree burials as well as extended inhumations (this chapter).

Following the Neolithic, Mainland Southeast Asia apparently had minimal genetic influence on the Orang Asli, apart from the Dvaravati-associated introduction of the B5a haplogroup. However, Dunn's (1964) documentation of red-slipped pottery late in the Gua Kecil sequence correlates with genetic evidence for the Iron Age incursion of haplogroups from Island Southeast Asia (east of Sumatra) affecting the Aboriginal Malays. Boat remains dating as of 2000 years BP in southern Malaya further point to a maritime Austronesian influence whose linguistic traces include the local Malayic dialects. Subsequent maritime connections between southern Malaya and southern Sumatra are indicated by loose archaeological parallels between these two regions, as well as by the Sumatra-derived N21 haplogroup found amongst Aboriginal Malays.

Overall, the female gene pool of Aboriginal Malays appears to reflect these populations' active participation in the growth of long-distance maritime trade, to which Malaya contributed metals (especially tin) and forest produce, during the first millennium CE. Maritime trade polities were also established in central and northern Malaya, but the female gene pools of the Semang and Senoi were barely affected by these developments. While these latter groups participated in the technological advances of the Common Era, as for instance in the wide-scale replacement of stone tools with iron implements, they were kept at a distance through socio-political marginalisation (combined in many cases with geographic isolation)—a fate that befell all Orang Asli with the expansion of the Melayu across Malaya after c. 1300 CE.

## **Acknowledgements**

My thanks to Ryan Rabett and Mokhtar Saidin who helped me restore and record the postcranial remains from Gua Cha B2, H1, H4, H5, H6, H7, H8, H9, H10, H11, H12, S6 and S7 (Appendix A) in 2006 when they were temporarily stored at the McDonald Institute for Archaeological Research at Cambridge University (prior to their repatriation to Malaysia), and my particular thanks to Ryan Rabett who organised the occasion. My

thanks also to Stephen Oppenheimer for his advice on interpreting the Orang Asli mtDNA data, and to Nick Enfield, Joyce White and an anonymous referee for their comments on the originally submitted manuscript.

## **Appendix A Detailed tables of limb bone length calculations and compilations, and stature estimates.**

---

This study utilises maximum limb-bone lengths estimated from their segments, which is necessary to make use of the often-fragmentary nature of the limb bones from sites in Malaya (Table A1). Such length estimation can be performed for the femur, tibia and humerus following the methodology of Steele & McKern (1969), who calculated the regressions between the lengths of these bones and their segments amongst Mesoamericans. Although Steele (1970) performed a similar regression exercise for American Whites and Blacks, the Mesoamerican-based study is relied on here, for two reasons. Firstly, Steele (1970) combined two of the humerus segments recognised by Steele and McKern (1969), but these segments need to be recognised separately to apply the technique to certain fragmentary prehistoric humeri from Malaya. Secondly, as with all correlation coefficient regression exercises the method has the weakness of ‘regression to the mean’, which means that the resulting estimates will tend to approach the mean values of the population being regressed to. Since the limb bones used in these studies are on average longest for American Blacks, intermediate for American Whites and shortest for Mesoamericans (Steele & McKern 1969; Steele 1970), the resulting estimates also follow this order, sometimes to only a minor degree but sometimes with major discrepancies. We do not know the average lengths of the limb bones of prehistoric Malaya populations, but in using the segments to help estimate these lengths we should adopt the shortest resulting estimates. This is to ensure that we are not creating a spurious difference between the Orang Asli (short people, represented by complete limb bones) and their prehistoric forerunners (represented mainly by limb bone fragments) through our choice of methodology, but that as far as possible we use a methodology consistent with the ‘null hypothesis’ of no reduction in limb-bone length over time (see Tables A2 and A3).

Living stature of prehistoric Malaysians can be estimated from limb-bone lengths and/or extended skeleton length (Tables A4 and A5). This study uses the Javanese regression formulae relating limb bone lengths (maximum lengths in the case of the humerus, ulna and fibula) to corpse length as Javanese are the population geographically closest to Malaya for whom such formulae are available (Bergman & The 1955). After people die, their corpse extends slightly, and so 2 cm is subtracted from corpse length to estimate living stature (Snell 1949). Skeleton length is slightly shorter than corpse length, owing to the loss of soft tissue, and so can be used as a direct estimate of living stature. Anthropometric data on Orang Asli living stature are taken from sources compiled during the British colonial era, rather than post-colonial surveys (for example Wagenseil 1967; Fix 2002) which would be affected by secular effects on growth associated with the various socio-political programs of the Malaysian state, such as the forced settlement of the more mobile of the Orang Asli.



**Table A1:** Maximum limb bone lengths estimated from segment lengths (mm), author's measurements

<b>Specimen</b>	<b>Sex</b>	<b>Limb bone</b>	<b>Segment(s)</b>	<b>Length</b>	<b>Estimated limb-bone length</b>
Gua Teluk Kelawar <sup>(1)</sup>	♀	Femur	2 & 3 (L)	242, 86	433.2 ± 5.1
Gua Teluk Kelawar <sup>(1)</sup>	♀	Humerus	1 & 2 (L)	31, 64	315.5 ± 8.3
Gua Peraling 4 <sup>(2)</sup>	♀	Humerus	2 & 3 (L)	59.6, 169	290.0 ± 4.7
Gua Kerbau Box No. 4 <sup>(3)</sup>	♂	Humerus	3, 4 & 5 (L)	192, 21, 18	339.3 ± 8.5
			2 & 3 (L)	240, 74	428.3 ± 7.5
Gua Cha 1 <sup>(3)</sup>	♂	Femur	2 (R)	222	417.4 ± 13.1
			2, 3, 4 & 5 (R),	64, 156, 30, 19	306.7 ± 2.6
Gua Cha 1 <sup>(3)</sup>	♂	Humerus	3 & 4 (L)	164, 28	312.9 ± 9.2
Gua Cha H10 <sup>(3)</sup>	♂	Femur	3 (R)	89	448.5 ± 16.7
Gua Cha S6 <sup>(3)</sup>	♂	Humerus	3 & 4 (L)	198, 19	327.5 ± 9.2
Gua Cha S7 <sup>(3)</sup>	♂	Femur	4 (R)	39	456.3 ± 15.6
Gua Cha H5 <sup>(3)</sup>	♀	Femur	2 & 3 (L)	217, 94	416.2 ± 5.1
				70, 168, 19,	
Gua Cha H12 <sup>(3)</sup>	♂	Humerus	2, 3, 4 & 5 (R)	14	303.9 ± 2.6
Gua Cha H8 <sup>(3)</sup>	♀	Tibia	2 & 3 (L)	53, 131	308.0 ± 6.4
			3 (L), 4 & 5 (R)		
Gua Cha H8 <sup>(3)</sup>	♀	Humerus		176, 18, 15	296.0 ± 6.9
Gua Cha H9 <sup>(3)</sup>	♂	Humerus	4 (R)	15	301.0 ± 12.3
Guar Kepah B289 <sup>(4)</sup>	♀	Femur	1 (R), 2 (L)	74, 221	422.4 ± 5.1
Guar Kepah B296 <sup>(4)</sup>	♂	Humerus	4 (L)	28	324.4 ± 12.3
Guar Kepah B321 <sup>(4)</sup>	♂	Humerus	4 (R)	29	326.2 ± 12.3
Guar Kepah C82 <sup>(4)</sup>	♂	Humerus	4 & 5 (R)	27, 18	330.1 ± 11.6
Gua Cha H11 <sup>(3)</sup>	♂	Tibia	1, 2 & 3 (L)	32, 50, 136	327.5 ± 8.3
Gua Cha A3 <sup>(3)</sup>	♀	Humerus	3 & 4 (L)	146, 25	288.0 ± 9.2
Gua Harimau 1 <sup>(3)</sup>	♂	Femur	2 & 3 (L)	236, 103	455.4 ± 7.5
Gua Harimau 1 <sup>(3)</sup>	♂	Humerus	3 (R)	166	306.0 ± 9.8
Kuala Selinsing 7 <sup>(3)</sup>	♂	Tibia	3 (R)	165	366.3 ± 13.8
				254, 90,	
Kuala Selinsing 8 <sup>(3)</sup>	♂	Femur	2, 3 & 4 (R)	40.5	451.9 ± 5.8

(1) Bulbeck & Zuraina 2007; (2) Bulbeck & Adi 2005; (3) Author's measurements (see also Acknowledgments). (4) Bulbeck 2005a, modified here by combining Guar Kepah B289 left and right measurements to estimate femur length.

**Table A2:** Male limb bone lengths and estimates (mm), maximum length except where otherwise specified

Period	Specimen	Femur	Tibia	Fibula	Humerus	Radius	Ulna
Early Holocene	Gua Gunung Runtuh <sup>(1)</sup>	449, 454	395		323, 325	277	285
Middle Holocene	Gua Kerbau Box No. 4 <sup>(2)</sup>	474, 480			339, 350		≥260
Middle Holocene	Gua Bintong 1 <sup>(3)</sup>	470 est.					
Middle Holocene	Gua Cha 1 <sup>(4)</sup>	417, 428	355 est.	345 est.	307, 313	248	266.5
Middle Holocene	Gua Cha H10 <sup>(2)</sup>	448.5					
Middle Holocene	Gua Cha S6 <sup>(2)</sup>	448 est.			327.5		
Middle Holocene	Gua Cha S7 <sup>(2)</sup>	456				254	
Middle Holocene	Gua Cha H9 <sup>(2)</sup>			318	301	262 est.	
Middle Holocene	Gua Cha H12 <sup>(2)</sup>				304		
Middle Holocene	Guar Kepah B296 <sup>(5)</sup>				324		
Middle Holocene	Guar Kepah B321 <sup>(5)</sup>				326		
Middle Holocene	Guar Kepah C82 <sup>(5)</sup>				330		
Late Neolithic	Gua Cha H11 <sup>(2)</sup>	413.5	327.5, 342	310 est.			
Late Neolithic	Gua Cha H4 <sup>(2)</sup>		385 est.				275
Late Neolithic	Gua Cha B2 <sup>(2)</sup>					243 est.	
Late Neolithic	Gua Cha H1 <sup>(2)</sup>				335		
Late Neolithic	Gua Harimau 1 <sup>(2)</sup>	455			306		
Iron Age	Gua Baik 94 <sup>(6)</sup>					217	238
Iron Age	Kuala Selinsing 2 <sup>(7)</sup>	423, 426	357, 358			238	
Iron Age	Kuala Selinsing 7 <sup>(2)</sup>		366		317, 317.5		
Iron Age	Kuala Selinsing 8 <sup>(2)</sup>	452			315		
Orang Asli	Semang (n = 1–3) <sup>(8)</sup>	424.0 (411.5–432)	362.8 (345–374.5)	363.5	303.3 (302–305.5)	236.0 (231–240)	250.8 (249.5–252)
Orang Asli	Senoi (n = 1–2) <sup>(9)</sup>	378.5 (368, 389)	311 (299, 323)	305.5 (299, 312)	263.5 (249.5, 277.5)	209.8 (202, 217.5)	228 (222, 234)

(1) Matsumura & Zuraina 1999, excludes lengths of atrophied left forearm bones. (2) Author's measurements (see also Table A1). Note that the total right tibia length of Gua Cha H11 is 334 mm. (3) Duckworth 1934. (4) Bulbeck 2005b, amended where appropriate to apply Mesoamerican regression formulae (see Table A1). (5) Bulbeck 2005a (see also Table A1). (6) Snell 1949. (7) Harrower 1933; tibia lengths are total lengths. (8) Author's measurements for the Pangan Semang (Duckworth Laboratory) combined with the measurements from Schebesta & Lebzelter (1926); averages of left and right limb bones shown. (9) Martin (1905); averages of left and right limb bones shown.

**Table A3:** Female limb bone lengths and estimates (mm), maximum length except where otherwise specified

Period	Specimen	Femur	Tibia	Fibula	Humerus	Radius	Ulna
Early Holocene	Moh Khiew 2 <sup>(1)</sup>					247	271
Early Holocene	Gua Teluk Kelawar <sup>(2)</sup>	433	340 est.		316		
Middle Holocene	Moh Khiew 1 <sup>(1)</sup>				308		
Middle Holocene	Gua Peraling 4 <sup>(3)</sup>				290		
Middle Holocene	Gua Cha H5 <sup>(4)</sup>	416					
Middle Holocene	Gua Cha H8 <sup>(4)</sup>		308		296	222	
Middle Holocene	Gua Cha H6 <sup>(4)</sup>			313			262
Middle Holocene	Gua Cha H7 <sup>(4)</sup>						262 est.
Middle Holocene	Guar Kepah B289 <sup>(5)</sup>	422					
Late Neolithic	Gua Cha B8 <sup>(6)</sup>	384	325, 329	320, 322	265, 268		232
Late Neolithic	Gua Cha A3 <sup>(4)</sup>				288		
Iron Age	Kuala Selinsing 1 <sup>(7)</sup>	394, 400			278	227	
Orang Asli	Semang (n = 1–4) <sup>(8)</sup>	386.6 (373.5–395)	322.3 (313–334)	318.3 (309–330)	267.4 (265–269)	214.2 (211–219.5)	243.5
Orang Asli	Senoi (n = 1–2) <sup>(8)</sup>	388.3 (383, 393.5)	309.8 (300, 319.5)	306.3 (302, 310.5)	269.8 (269, 270.5)	205.8 (197.5, 214)	221.8 (211.5, 232)
Orang Asli	Aboriginal Malay <sup>(8)</sup>	338			228	185	

(1) Bulbeck 2003. (2) Bulbeck & Zuraina 2007 (see also Table A1). (3) Bulbeck & Adi 2005 (see also Table A1). (4) Author's measurements (see also Table A1). (5) Bulbeck 2005a (see also Table A1). (6) Trevor & Brothwell 1962; tibia lengths are total lengths; the published left ulna length of 265 mm is excluded as it replicates the left humerus length and is clearly in error. (7) Harrower 1933, re-sexed as female based on the author's study of the skull; humerus length is total length. (8) Martin (1905), averages of left and right limb bones; the Jakun Aboriginal Malay humerus length is total length.

**Table A4:** Estimates of male living stature in cm

<b>Period</b>	<b>Specimen</b>	<b>Stature</b>	<b>Method</b>
Early Holocene	Gua Gunung Runtuh <sup>(1)</sup>	169	Regression formulae (skeleton length affected by lordosis)
Mid-Holocene	Gua Kerbau Box No. 4 <sup>(2)</sup>	172	Regression formulae (average of 171 cm femur and 173 cm humerus estimates)
Mid-Holocene	Gua Bintong 1 <sup>(3)</sup>	169	Regression formula (using Duckworth's 465 mm minimum femur length estimate)
Mid-Holocene	Gua Cha S6 <sup>(4)</sup>	168	Regression formula (humerus length)
Mid-Holocene	Gua Cha 1 <sup>(4)</sup>	166	Regression formulae (average of 164 cm humerus and 168 cm femur estimates)
Mid-Holocene	Gua Cha S7 <sup>(4)</sup>	166	Regression formula (allowing 448 mm for femur physiological length)
Mid-Holocene	Gua Cha H10 <sup>(4)</sup>	164	Regression formula (allowing 440 mm for femur physiological length)
Mid-Holocene	Gua Cha H12 <sup>(4)</sup>	162	Regression formula (humerus length)
Mid-Holocene	Gua Cha H9 <sup>(4)</sup>	158	Regression formulae (average of 154.4 cm femur and 161 cm humerus estimates)
Mid-Holocene	Guar Kepah C82 <sup>(4)</sup>	169	Regression formula (humerus length)
Mid-Holocene	Guar Kepah B321 <sup>(4)</sup>	167	Regression formula (humerus length)
Mid-Holocene	Guar Kepah B296 <sup>(4)</sup>	167	Regression formula (humerus length)
Late Neolithic	Gua Cha H1 <sup>(4)</sup>	171	Regression formula (humerus length)
Late Neolithic	Gua Cha H4 <sup>(5)</sup>	165	Skeleton length (169 mm would be estimated based on ulna length)
Late Neolithic	Gua Cha B9 <sup>(5)</sup>	164	Skeleton length
Late Neolithic	Gua Cha H13 <sup>(5)</sup>	158	Skeleton length
Late Neolithic	Gua Cha B2 <sup>(5)</sup>	157	Skeleton length
Late Neolithic	Gua Cha B7 <sup>(5)</sup>	155	Skeleton length
Late Neolithic	Gua Cha H11 <sup>(4)</sup>	155	Regression formulae (average of 157 cm tibia and 153 cm fibula estimates)
Late Neolithic	Gua Cha B1 <sup>(5)</sup>	152	Skeleton length
Late Neolithic	Gua Harimau 1 <sup>(6)</sup>	160	Skeleton length (162 cm would be the estimate based on humerus length)
Late Neolithic	Gua Harimau 9 <sup>(7)</sup>	155	Skeleton length
Iron Age	Kuala Selinsing 7 <sup>(4)</sup>	165	Regression formulae (humerus lengths)
Iron Age	Kuala Selinsing 8 <sup>(4)</sup>	165	Regression formula (humerus length)
Iron Age	Kuala Selinsing 2 <sup>(4)</sup>	162	Regression formulae (tibia lengths)
Iron Age	Gua Baik 94 <sup>(4)</sup>	155	Regression formula (ulna length)
Orang Asli	Semang <sup>(8)</sup>	154	Anthropometry (range 142–164 cm)
Orang Asli	Senoi <sup>(9)</sup>	153	Anthropometry (range 138–166 cm)
Orang Asli	Aboriginal Malays <sup>(9)</sup>	154	Anthropometry (range 139–164 cm)

(1) Matsumura & Zuraina 1999, who provide femur and radius physiological lengths and tibia total length, as used by Bergman & The (1955), as well as the maximum lengths shown in Table A2. The estimate employed is the median estimate from the different limb bones (which range from 166 cm using the femur to 175 cm using the radius). (2) Based on author's measurements, including right femoral physiological length. (3) See Duckworth 1934:165. (4) Based on measurements in Table A2. (5) Sieveking 1954. (6) Chia & Zolkurnian (2005). (7) Measured from Chia & Zolkurnian 2005:Fig. 18.5. (8) Bulbeck 1996. (9) Martin 1905:232–234.

**Table A5:** Estimates of female living stature in cm

Period	Specimen	Stature	Method
Early Holocene	Moh Khiew 2 <sup>(1)</sup>	167	Regression formula (ulna measurement)
Early Holocene	Gua Teluk Kelawar <sup>(2)</sup>	160	Regression formulae
Mid-Holocene	Moh Khiew 1 <sup>(1)</sup>	165	Regression formula (humerus measurement)
Mid-Holocene	Gua Peraling 4 <sup>(1)</sup>	157	Regression formula
Mid-Holocene	Gua Cha H7 <sup>(1)</sup>	163	Regression formula
Mid-Holocene	Gua Cha H8 <sup>(1)</sup>	157	Regression formula (humerus measurement)
Mid-Holocene	Gua Cha H5 <sup>(1)</sup>	156	Regression formula (allowing 408 mm for femur physiological length)
Mid-Holocene	Gua Cha H6 <sup>(1)</sup>	156	Regression formulae (average of 163 cm ulna and 149 cm fibula estimates)
Mid-Holocene	Gua Kajang <sup>(3)</sup>	155	General comparison
Mid-Holocene	Guar Kepah B289 <sup>(4)</sup>	153	Regression formulae
Neolithic	Gua Cha A3 <sup>(1)</sup>	156.5	Regression formula
Neolithic	Gua Cha B8 <sup>(5)</sup>	155	Skeleton length (150 cm would be estimate based on limb bone lengths)
Iron Age	Kuala Selinsing 1 <sup>(6)</sup>	156	Regression formula
Orang Asli	Semang <sup>(7)</sup>	143	Anthropometry (range 135–152 cm)
Orang Asli	Senoi <sup>(8)</sup>	143	Anthropometry (range 132–156 cm)
Orang Asli	Aboriginal Malays <sup>(8)</sup>	143	Anthropometry (range 131–155 cm)

(1) Based on measurement in Table A3. (2) Bulbeck & Zuraina 2007 (see also Table A3). (3) Duckworth 1934:154. (4) Bulbeck 2005a. (5) Sieveking 1954. (6) Harrower 1933:203–204. (7) Bulbeck 1996. (8) Martin 1905:232–234.

## References

- Adelaar, Alexander. 2004. Where does Malay come from? Twenty years of discussions about homeland, migrations and classifications. *Bijdragen tot de Taal-, Land- en Volkenkunde* 160. 1–30.
- Adi, Haji Taha. 2000. *Archaeological Excavations in Ulu Kelantan, Peninsular Malaysia*. Ph.D. dissertation. Canberra: Australian National University.
- Ambrose, Stanley H. 1998. Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *Journal of Human Evolution* 34. 623–51.
- Anderson, Douglas D. 1988. Excavations of a Pleistocene rockshelter in Krabi and the prehistory of southern Thailand. In Pisit Charoenwangsa & Bennett Bronson (eds) *Prehistoric Studies: The Stone and Metal Ages in Thailand*, 44–59. Bangkok: Papers in Thai Antiquity Volume 1.
- . 1990. *Lang Rongrien Rockshelter: A Pleistocene-early Holocene Archaeological Site from Krabi, southwestern Thailand*. Philadelphia: University of Pennsylvania Museum.
- . 2005. The use of caves in Peninsular Thailand in the Late Pleistocene and early and middle Holocene. *Asian Perspectives* 44. 137–153.

- Azman, Mohd Noh. 1998. Ekskavasi Gua Singa dan Gua Putih di Gerik, Perak. *Malaysia Museums Journal* 34. 221–239.
- Ballinger, Scott W., Theodore G. Schurr, Antonio Torroni, Yik-Yuen Gan, Judith A. Hodge, K. Hassan, Kuei Hsiang Chen & Douglas C. Wallace. 1992. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics* 130. 139–152.
- Bannanurag, Rachanie. 1988. Evidence for ancient woodworking: A microwear study of Hoabinhian stone tools. In Pisit Charoenwanga & Bennett Bronson (eds) *Prehistoric Studies: The Stone and Metal Ages in Thailand*, 62–79. Bangkok: Papers in Thai Antiquity Volume 1.
- Bayard, Donn. 1996-7. Bones of contention: the Non Nok Tha burials and the chronology and context of early Southeast Asian bronze. In David Bulbeck & Noel Barnard (eds) *Ancient Chinese and Southeast Asian Bronze Age Cultures, volume II*, 889–940. Taipei: Southern Materials Center Inc.
- Bellina-Pryce, Bérénice. 2009. *Late prehistoric South China Sea coastal people: Peninsular Thailand and Sa Huynh communities*. Paper presented at the Sa Huynh Conference, Quang Ngai, July 2009.
- Bellwood, Peter. 1978. *Man's Conquest of the Pacific*. Sydney: Collins.
- 1992. Southeast Asia before history. In Nicholas Tarling (ed.) *The Cambridge History of Southeast Asia Volume I*, 51–136. Cambridge: Cambridge University Press.
- 1993. Cultural and biological differentiation in Peninsular Malaysia: the last 10,000 years. *Asian Perspectives* 32. 37–60.
- 1997. *Prehistory of the Indo-Malaysian Archipelago*. Revised edition. Honolulu: University of Hawai'i Press.
- Benjamin, Geoffrey. 1976. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds) *Austroasiatic Studies*, Part I, 37–128. Honolulu: The University Press of Hawaii.
- 1985. In the long term: three themes in Malayan cultural ecology. In Karl L. Hutterer, Terry Rambo & George Lovelace (eds) *Cultural values and human ecology in South East Asia*, 219–278. University of Michigan: Center for South and Southeast Asian Studies.
- 1987. Ethnohistorical perspectives on Kelantan's prehistory. In Nik Hassan Shuhaimi bin Nik. Abd. Rahman (ed.) *Kelantan Zaman Awal: Kajian Arkeologi dan Sejarah di Malaysia*, 108–53. Kota Bahru: Perbadanan Muzium Negeri Kelantan Istana Johor.
- 2002. On being tribal in the Malay world. In Geoffrey Benjamin & Cynthia Chou (eds) *Tribal Communities in the Malay World: Historical, Cultural and Social Perspectives*, 7–76. Leiden: International Institute for Asian Studies/Singapore: Institute for Southeast Asian Studies.
- Bergman, R.A.M. & The Tjong Hoo. 1955. The length of the body and long bones of the Javanese. *Documenta Medicina Geographica et Tropica* 7. 197–214.
- Bower, Nathan W., Yuichiro Yasumoto, Marc F. Oxenham, Nguyen Lan Cuong & Nguyen Kim Thuy. 2006. Preliminary reconstruction of diet at a Neolithic site in Vietnam using stable isotope and Ba/Sr analyses. *Bulletin of the Indo-Pacific Prehistory Association* 26. 79–85.
- Boyd, William E., Nigel J. Chang, Joss Debreceny, Kate Domett, R. Ewan Fordyce, Charles Higham, Tom F. G. Higham, Alan Hogg, V. Hunt, Bob F. J. Manly, Greame M. Mason,

- Dougald J. W. O'Reilly, C. Pailles, Anthony Reay, Nancy Tayles & Rachanie Thosarat. 1998. *The Excavation of Nong Nor*. Otago: University of Otago Studies in Prehistoric Anthropology No. 18.
- Bulbeck, David. 1981. *Continuities in Southeast Asian Evolution since the Late Pleistocene*. Unpublished MA thesis. Canberra: The Australian National University.
- . 1996. Holocene biological evolution of the Malay Peninsula Aborigines (*Orang Asli*). *Perspectives in Human Biology* 2. 37–61.
- . 1998. *Description and preliminary analysis of the human remains from Pulau Kelumpang (Kuala Selinsing), Perak, Malaysia*. Report to the Department of Museums and Antiquity, Kuala Lumpur, Malaysia.
- . 2000. Dental morphology at Gua Cha, West Malaysia, and the implications for “Sundadonty”. *Bulletin of the Indo-Pacific Prehistory Association* 19. 17–40.
- . 2003. Hunter-gatherer occupation of the Malay Peninsula from the Ice Age to the Iron Age. In Julio Mercader (ed.) *The Archaeology of Tropical Rain Forests*, 119–160. New Brunswick: Rutgers University Press.
- . 2004a. An integrated perspective on Orang Asli ethnogenesis. In Victor Paz (ed.) *Southeast Asian Archaeology: Wilhelm G. Solheim II Festschrift*, 366–399. Quezon City: University of the Philippines Press.
- . 2004b. Indigenous traditions and exogenous influences in the early history of Peninsular Malaysia. In Ian Glover & Peter Bellwood (eds) *Southeast Asia: Origins to Civilization*, 314–336. London: RoutledgeCurzon.
- . 2004c. Ban Chiang. In Ooi Keat Gin (ed.) *Southeast Asia: A Historical Encyclopedia from Angkor Wat to East Timor*, 205–07. Santa Barbara, CA: ABC Clio.
- . 2004d. Ban Kao culture. In Ooi Keat Gin (ed.), *Southeast Asia: A Historical Encyclopedia from Angkor Wat to East Timor*, 208–11. Santa Barbara, CA: ABC Clio.
- . 2005a. The Guar Kepah human remains. In Zuraina Majid (ed.), *The Perak Man and Other Prehistoric Skeletons of Malaysia*, 383–423. Penang: Penerbit Universiti Sains Malaysia.
- . 2005b. The Gua Cha burials. In Zuraina Majid (ed.), *The Perak Man and Other Prehistoric Skeletons of Malaysia*, 253–309. Penang: Penerbit Universiti Sains Malaysia.
- . 2006. A test of Ambika Flavel’s results with an enlarged Sa Huynh-Kalanay data set. Contribution to Wilhelm G. Solheim II, *Archaeology and Culture in Southeast Asia: Unraveling the Nusantara*, 238–270. Diliman, Quezon City: The University of the Philippines Press.
- . 2008. An integrated perspective on the Austronesian diaspora: the switch from cereal agriculture to maritime foraging in the colonisation of Island Southeast Asia. *Australian Archaeology* 67. 31–51.
- Bulbeck, David & Adi Taha. 2005. A description and analysis of the Gua Peraling human remains. In Zuraina Majid (ed.) *The Perak Man and Other Prehistoric Skeletons of Malaysia*, 311–343. Penang: Penerbit Universiti Sains Malaysia.
- Bulbeck, David & Adam Lauer. 2006. Human variation and evolution in Holocene Peninsular Malaysia. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 133–171. Cambridge: Cambridge University Press.
- Bulbeck, David, Pathmanathan Raghavan & Daniel Rayner. 2006. Races of *Homo sapiens*: if not in the southwest Pacific, then nowhere. *World Archaeology* 38. 109–32.

- Bulbeck, David, Abdul Kadir Rahimah, Adam Lauer, Zamri Radzi & Daniel Rayner. 2005. Tooth sizes in the Malay Peninsula past and present: insights into the time depth of the indigenous inhabitants' adaptations. *International Journal of Indigenous Research* 1. 41–50.
- Bulbeck, David & Majid Zuraina. 2007. Gua Teluk Kelawar 1 and Holocene human evolution in Peninsular Malaysia. In Saidin Mokhtar & Stephen Chia (eds) *Archaeological Heritage of Malaysia*, 17–39. Penang: Centre for Archaeological Research Malaysia, Universiti Sains Malaysia.
- Burenhult, David, Nicole Kruspe, & Michael Dunn, this volume, Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula.
- Callenfels, Pieter V. van Stein & H. D. (Pat) Noone. 1938. Report of an excavation in the rock-shelter Gol Ba'it, near Sungei Siput. In Frederick N. Chasen & Michael W.F. Tweedie (eds) *Proceedings of the Third Congress of Prehistorians of the Far East*, 199–130. Singapore: Government of the Straits Settlements.
- Cameron, David W. & Groves, Colin P. 2004. *Bones, Stones and Molecules: "Out of Africa" and Human Origins*. Amsterdam: Elsevier Academic Press.
- Cann, Rebecca L., Stoneking Mark & Wilson, Allan C. 1987. Mitochondrial DNA and human evolution. *Nature* 325. 31–36.
- Carey, Iskandar. 1976. *Orang Asli: The Aboriginal Tribes of Peninsula Malaysia*. Kuala Lumpur: Oxford University Press.
- Chia, Stephen. 1998. Indigenous prehistoric pottery and technology in Peninsular Malaysia. *Malaysia Museums Journal* 34. 153–192.
- Chia, Stephen & Zolkurnian Hasan. 2005. Gua Harimau: a prehistoric cemetery in Lenggong, Perak. In Zuraina Majid (ed.) *The Perak Man and Other Prehistoric Skeletons of Malaysia*, 363–382. Penang: Penerbit Universiti Sains Malaysia.
- Chitkament, Thanon. 2006–07. *Lithic Analysis of Moh Khiew Rockshelter (Locality I) in Krabi River Valley, Krabi Province, Southwestern Thailand*. *Erasmus Mundus Master in Quaternary and Prehistory*. Tarragona: Universitat Roviri I Virgili.
- Cole, Fay-Cooper. 1945. *The Peoples of Malaysia*. Princeton, N.J.: D. Van Nostrand Co. Inc.
- Coon, Carleton S. 1962. *The Origin of Races*. New York: Knopf.
- Day, Michael H. & Stringer, Christopher B. 1982. A reconsideration of the Omo Kibish remains and the *erectus-sapiens* transition. In Henry de Lumley (ed.) *L'Homo erectus et la Place de l'Homme de Tautavel parmi les Hominides Fossiles, Volume 2*, 814–16. Nice: Jean-Louis.
- Demeter, Fabrice. 2006. New perspectives on the peopling of Southeast and East Asia during the late upper Pleistocene. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 112–132. Cambridge: Cambridge University Press.
- Dentan, Robert K. 1968. *The Semai: a non-violent People of Malaya*. New York: Holt, Rinehart and Winston.
- Duckworth, Wynfrid L. H. 1934. Human remains from rock-shelters and caves in Perak, Pahang and Perlis and from Selinsing. *Journal of the Malayan Branch of the Royal Asiatic Society* 12. 149–167.
- Dunn, Frederick L. 1964. Excavations at Gua Kechil, Pahang. *Journal of the Malaysian Branch of the Royal Asiatic Society* 37. 87–124.
- Edwards McKinnon, Edward. 1991. The Hoabinhian in the Wampu/Lau Biang Valley of northeastern Sumatra: an update. *Bulletin of the Indo-Pacific Prehistory Association* 10. 132–142.



- Endicott, Kirk. 1979. *Batek Negrito Religion*. Oxford: Clarendon Press.
- Evans, Ivor H. N. 1927. *Papers on the Ethnology and Archaeology of the Malay Peninsula*. Cambridge: Cambridge University Press.
- 1937. *The Negritos of Malaya*. Cambridge: Cambridge University Press.
- Fix, Alan. 2002. Foragers, farmers, and traders in the Malayan Peninsula: origins of cultural and biological diversity. In Kathleen D. Morrison & Laura L. Junker (eds) *Forager-Traders in South and Southeast Asia, 185–202*. Cambridge: Cambridge University Press.
- this volume, Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model.
- Forestier, Hubert, Truman Simanjuntak, Dominique Guillaud, Dubel Driwantoro, Ketut Wiradnyana, Darwin Siregar, Rokus Due Awe & Budiman. 2005. Le site de Tögi Ndrawa, île de Nias, Sumatra nord: les premières traces d'une occupation hoabinhienne en grotte en Indonésie. *Comptes Rendus Palevol* 4. 727–733.
- Forestier, Hubert, Dubel Driwantoro, Dominique Guillard, Budiman & Darwin Siregar. 2006. New data for the prehistoric chronology of South Sumatra. In Truman Simanjuntak, M. Hisyam, Bagyo Prasetyo & Titi Surti Nastiti (eds) *Archaeology: Indonesian Perspective. R.P. Soejono's Festschrift*, 177–192. Jakarta: Indonesian Institute of Sciences.
- Fucharoen, Goonnapa, Supan Fucharoen & Satoshi Horai 2001. Mitochondrial DNA polymorphisms in Thailand. *Journal of Human Genetics* 46. 115–125.
- Gianno, Rosemary. 1990. *Semelai Culture and Resin Technology*. New Haven: Connecticut Academy of Arts and Sciences.
- Gianno, Rosemary & Klaus J. Bayr. 2009. Semelai agricultural patterns: toward an understanding of variation among indigenous cultures in southern peninsular Malaysia. *Journal of Southeast Asian Studies* 40. 153–185.
- Gilligan, Ian & David Bulbeck. 2007. Environment and morphology in Australian Aborigines: a re-analysis of the Birdsell database. *American Journal of Physical Anthropology* 134. 75–91.
- Hamilton, Anette. 2006. Reflections on the 'disappearing Sakai': a tribal minority in southern Thailand. *Journal of Southeast Asian Studies* 37. 293–314.
- Harrower, Gordon. 1933. Skeletal remains from the Kuala Selinsing excavations, Perak, Malay Peninsula. *Journal of the Malayan Branch of the Royal Asiatic Society* 11. 190–210.
- Heekeren, H. Robert Van. 1972. *The Stone Age of Indonesia*. Second edition. The Hague: Martinus Nijhoff.
- Higham, Charles. 2002. *Early Cultures of Mainland Southeast Asia*. Bangkok: River Books.
- Higham, Charles & Rachanie Thosarat. 1998. *Prehistoric Thailand from Early Settlement to Sukothai*. Bangkok: River Books.
- Higham, Charles & Rachanie Thosarat. 2006. Ban Non Wat: the first three seasons. In Elisabeth A. Bacus, Ian C. Glover & Vincent C. Pigott (eds), *Uncovering Southeast Asia's Past: Selected Papers from the 10<sup>th</sup> International Conference of the European Association of Southeast Asia Archaeologists*, 98–104. Singapore: National University of Singapore Press.
- Hill, Catherine E. 2005. Mitochondrial DNA Variation in Island Southeast Asia. Unpublished PhD thesis. Huddersfield: University of Huddersfield.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, William Meehan, James Blackburn, Douglas Clarke, Joseph Maripa Raja, Patimah Ismail, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2006. Phylogeography and ethnogenesis of Aboriginal Southeast Asians. *Molecular Biological Evolution* 23(12). 2480–2491.

- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, Douglas Clarke, Petya B. Blumbach, Matthieu Vizuete-Forster, Peter Forster, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2007. A mitochondrial stratigraphy for Island Southeast Asia. *American Journal of Human Genetics* 80. 29–43.
- Hillson, Simon. 1996. *Dental Anthropology*. Cambridge: Cambridge University Press.
- Hooker, Virginia M. 2003. *A Short History of Malaysia*. Sydney: Allen & Unwin.
- Hoop, Abraham N.J. van der. 1932. *Megalithic Remains in South-Sumatra*. Trans W. Shirlaw. Zutphen: W.J. Thieme & Cie.
- Howells, William. 1973a. *The Pacific Islanders*. London: Weidenfeld and Nicolson.
- 1973b. *Cranial Variation in Man*. Cambridge, Mass.: Peabody Museum.
- Indrawoath, Phasook. 2004. The archaeology of the early Buddhist kingdoms of Thailand. In Ian Glover & Peter Bellwood (eds) *Southeast Asia: Origins to Civilization*, 314–336. London: RoutledgeCurzon.
- Jennings, Sue. 1995. *Theatre, Ritual and Transformation: the Senoi Temiars*. New York: Routledge.
- Kamminga, Johan. 2007. Reinterpretation of the Hoabinhian at Sai Yok, Thailand. In Mokhtar Saidin & Stephen Chia (eds) *The Archaeological Heritage of Malaysia*, 40–60. Penang: Centre for Archaeological Research Malaysia.
- Kealhofer, Lisa. 2003. Looking into the gap: land use and the tropical forests of southern Thailand. *Asian Perspectives* 42. 72–95.
- Lahr, Marta M. & Robert Foley. 1998. Toward a theory of modern human origins: Geography, demography and diversity in recent human populations. *Yearbook of Physical Anthropology* 41. 137–176.
- Latinis, Kyle D. 2000. The development of subsistence systems models for Island Southeast Asia and Near Oceania: the nature and role of arboriculture and arboreal-based economies. *World Archaeology* 32. 41–67.
- Lee, Kok Joo. 1976. *Report No. 5. Kampung Lubok Bandung: a Temuan Community of Malacca State*. Penang: Universiti Sains Malaysia.
- Leong Sau Heng. 1989. Satu perbincangan mengenai peninggalan objek-objek gangsa kecil dari zaman pra-sejarah Malaysia. *Jurnal Arkeologi Malaysia* 2. 1–8.
- Leong Sau Heng. 1991. Jenderam Hilir and the mid-Holocene prehistory of the west and coast plain of Peninsular Malaysia. *Bulletin of the Indo-Pacific Prehistory Association* 10. 150–160.
- Leong, Sau Heng 2000. The chronology of the Bernam cist graves in Peninsular Malaysia. *Bulletin of the Indo-Pacific Prehistory Association* 19. 65–72.
- Lie-Injo, Luan Eng. 1976. Genetic relationships of several aboriginal groups in South East Asia. In Robert L. Kirk & Alan G. Thorne (eds) *The Origin of the Australians*, 277–306. Canberra: Australian Institute of Aboriginal [and Torres Strait Islander] Studies.
- Macaulay, Vincent, Catherine Hill, Alessandro Achilli, Chiara Rengo, Douglas Clarke, William Meehan, James Blackburn, Ornella Semino, Rosaria Scozzari, Fulvio Cruciani, Adi Taha, Norazila Kassim Shaari, Joseph Maria Raja, Patimah Ismail, Zafarina Zainuddin, William Goodwin, David Bulbeck, Hans-Jürgen Bandelt, Stephen Oppenheimer, Antonio Torroni & Martin Richards. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308. 1034–1036.

- Martin, Rudolf. 1905. *Die Inlandstämme der Malayischen Halbinsel: Wissenschaftliche Ergebnisse einer Reise durch die Vereinigten Malayischen Staaten*. Stuttgart: Gustav Fischer Verlag.
- Matsumura, Hirofumi. 2005. Bioanthropological significance of prehistoric skeletal remains in Malaysia. In Zuraina Majid (ed.) *The Perak Man and Other Prehistoric Skeletons of Malaysia*, 425–456. Penang: Penerbit Universiti Sains Malaysia.
- Matsumura, Hirofumi & Majid Zuraina. 1999. Metric analysis of an early Holocene human skeleton from Gua Gunung Runtuh, Malaysia. *American Journal of Physical Anthropology* 109. 327–340.
- Matsumura, Hirofumi & Mark J. Hudson. 2004. Dental perspectives on the population history of Southeast Asia. *American Journal of Physical Anthropology* 127. 182–209.
- Matsumura, Hirofumi, Marc Oxenham, Yukio Dodo, Kate Domett, Nguyen Kim Thuy, Nguyen Lan Cuong, Nguyen Kim Dung, Damien Huffer & Mariko Yamagata. 2008. Morphometric affinity of the late Neolithic human remains from Man Bac, Ninh Binh Province, Vietnam: Key skeletons with which to debate the ‘two layer’ hypothesis. *Anthropological Science* 116. 135–148.
- Matthews, John. 1961. *A Check-list of “Hoabinhian” Sites Excavated in Malaya, 1860-1939*. Singapore: Eastern Universities Press Ltd.
- Melton, Terry, Raymond Peterson, Alan J. Redd, Nilmani Saha, Abdul Salam M. Sofro, Jeremy Martinson & Mark Stoneking. 1995. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *American Journal of Human Genetics* 57. 403–414.
- Migliano, Andrea B., Lucio Vinicius & Marta M. Lahr. 2007. Life history trade-offs explain the evolution of human pygmies. *Proceedings of the National Academy of Sciences* 104. 20216–20219.
- Mokhtar Saidin. 2006. Bukit Bunuh, Lenggong, Malaysia: new evidence of Late Pleistocene culture in Malaysia and Southeast Asia. In Elisabeth A. Bacus, Ian C. Glover & Vincent Pigott (eds) *Uncovering Southeast Asia’s Past*, 60–64. Singapore: National University of Singapore Press.
- Movius, Hallam L. Jr. 1944. Early man and Pleistocene stratigraphy in southern and eastern Asia. *Papers of the Peabody Museum of Archaeology and Ethnology* 19(3).
- Mudar, Karen & Douglas Anderson. 2007. New evidence for Southeast Asian Pleistocene foraging economies: faunal remains from the early levels of Lang Rongrien rockshelter, Krabi, Thailand. *Asian Perspectives* 46. 298–334.
- Nik Hassan Shuhaimi bin Nik Abd. Rahman, Mohd. Kamaruzaman A. Rahman & Mohd. Yusof Abdullah. 1988. Tapak prasejarah Gua Bukit Taat Hulu Terengganu (8920 ± B.P. – 2630 ± 80 B.P.). *Jurnal Arkeologi Malaysia* 3. 1–9.
- Nik Hassan Shuhaimi bin Nik Abd. Rahman. 2005. Views on the origin of the Orang Asli from archaeological and ehnohistorical perspectives. *International Journal of Indigenous Research* 1. 25–30.
- Nguyen Khac Su, Pham Minh Huyen & Tong Trung Tin. 2004. Northern Vietnam from the Neolithic to the Han period. In Ian Glover & Peter Bellwood (eds) *Southeast Asia from Prehistory to History*, 177–208. London and New York: RoutledgeCurzon.
- Nguyen Viet. 2005. The Da But culture: evidence for cultural development in Vietnam during the Middle Holocene. *Bulletin of the Indo-Pacific Prehistory Association* 25. 89–94.

- Noone, H. D. (Pat). 1939. Some vital statistics of the lowland Senoi of Perak. *Journal of the Federated Malay States Museum* 15. 195–215.
- Oota, Hiroki, Kunihiro Kurosaki, Surin Pookajorn, Takafumi Ishida & Shintaroh Ueda. 2001. Genetic study of the Palaeolithic and Neolithic Southeast Asians. *Human Biology* 73. 225–231.
- Oppenheimer, Stephen. 2003. *Out of Eden: The Peopling of the World*. London: Constable & Robinson.
- 2009. The great arc of dispersal of modern humans: Africa to Australia. *Quaternary International* 202. 2–13.
- this volume, MtDNA Variation and Southward Holocene Human Dispersals within Mainland Southeast Asia.
- Ousley, Stephen D. & Richard L. Jantz. 1994. *Fordisc 2.0 Personal Computer Forensic Discriminant Functions*. Knoxville: University of Tennessee, Department of Anthropology.
- Petraglia, Michael, Ravi Korisettar, Nicole Bovin, Cristopher Clarkson, Peter Ditchfield, Sacha Jones, Jinu Koshy, Marta M. Lahr, Clive Oppenheimer, David Pyle, Richard Roberts, Jean-Luc Schwenninger, Lee Arnold & Kevin White. 2007. Middle Paleolithic assemblages from the Indian subcontinent before and after the Toba super-eruption. *Science* 317. 114–116.
- Pietrusewsky, Michael. 2006. A multivariate craniometric study of the prehistoric and modern inhabitants of Southeast Asia, East Asia and surrounding regions: a human kaleidoscope? In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 59–90. Cambridge: Cambridge University Press.
- Pigott, Vincent C., Karen M. Mudar, Anagnostis Agelarakis, Lisa Kealhofer, Steven A. Weber & Judy C. Voelker. 2006. A program of analysis of organic remains from prehistoric copper-producing settlements in the Khao Wong Prachan valley, central Thailand: a progress report. In Elisabeth A. Bacus, Ian C. Glover & Vincent C. Pigott (eds) *Uncovering Southeast Asia's Past: Selected Papers from the 10<sup>th</sup> International Conference of the European Association of Southeast Asia Archaeologists*, 154–167. Singapore: National University of Singapore Press.
- Pookajorn, Surin. 1994. *Final Report of Excavations at Moh Khiew Cave, Krabi Province; Sakai Cave, Trang Province and Ethnoarchaeological Research of Hunter-Gatherer Group, socall Mani or Sakai or Orang Asli at Trang Province*. The Hoabinhian Research Project in Thailand, vol. 2. Bangkok: Silkaporn University Press.
- 1996. Human activities and environmental changes during the Late Pleistocene to middle Holocene in southern Thailand and Southeast Asia. In Lawrence G. Straus, Berit V. Eriksen, Jon M. Erlandson & David R. Yesner (eds) *Humans at the End of the Ice Age: The Archaeology of the Pleistocene-Holocene Transition*, 201–213. New York: Plenum Press.
- Prasad, Ravi B.V., Chris E. Ricker, Scott W. Watkins, Mary E. Dixon, Baskara B. Rao, Mastan J. Naidu, Lynn B. Jorda & Michael Banshad. 2001. Mitochondrial DNA variation in Nicobarese Islanders. *Human Biology* 73. 715–725.
- Rabett, Ryan J. 2005. The early exploitation of Southeast Asian mangroves: bone technology from caves and open sites. *Asian Perspectives* 44. 154–179.
- Rambo, Terry A. 1988. Why are the Semang? In Terry A. Rambo, Kathleen Gillogly and Karl L. Hutterer (eds) *Ethnic Diversity and the Control of Natural Resources in Southeast Asia*, 19–35. Ann Arbor: University of Michigan Papers on South and Southeast Asia.
- Rashid, Razha. 1995. Introduction. In Razha Rashid (ed.) *Indigenous Minorities of Peninsular Malaysia: Selected Issues and Ethnographies*, 1–17. Kuala Lumpur: Intersocietal and Scientific Sdn. Bhd.

- Rayner, Daniel. 2000. *The Dental Morphology of the Indigenous Peoples of Peninsular Malaysia (Orang Asli)*. Bachelor of Arts Honors dissertation. Canberra: Australian National University.
- Rayner, Daniel & David Bulbeck. 2001. Dental morphology of the “Orang Asli” Aborigines of the Malay Peninsula. In Maciej Henneberg (ed.) *Causes and Effects of Human Variation*, 19–41. Adelaide: Australasian Society for Human Biology/The University of Adelaide.
- Saha, Nilmani, Joon W. Mak, JS Tay, Y Liu, JA Tan, PS Low & M Singh 1995. Population genetic studies among the Orang Asli (Semai Senoi) of Malaysia: Malayan Aborigines. *Human Biology* 67. 37–57.
- Schebesta, Paul & V. Lebzelter. 1926. Schädel und Skelettreste von drei Semang-Individuen. *Anthropos* 21. 959–990.
- Schebesta, Paul & V. Lebzelter. 1928. Anthropological measurements in Semangs and Sakais in Malaya (Malacca). *Anthropologie* 6. 183–251.
- Sieveking, Gale de G. 1954. Excavations at Gua Cha, Kelantan, 1954. *Malayan Historical Journal* 1-2. 75–134.
- Sidwell, Paul & Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.
- Simanjuntak, Truman, Hubert Forestier, Dubel Driwantoro, Jatmiko & Darwin Siregar. 2006. Industri-industri yang paling kuno di Sumatera: bukti-bukti zaman Acheulian. In Dominique Guillaud (ed.) *Menyelusuri Sungai, Menurut Waktu: Penelitian Arkeologi di Sumatera Selatan*, 23–24. Jakarta: PT Enrique Indonesia.
- Snell, C.A.R.D. 1949. Human skeletal remains from Gol Ba’it, Sungai Siput, Perak, Malay Peninsula. *Acta Neerlandica Morphologica Normalis et Pathologicae* 6. 353–377.
- Soares, Pedro, Luca Ermini, Noel Thomson, Maru Mormina, Teresa Rito, Arne Röhl, Antonio Salas, Stephen Oppenheimer, Vincent Macaulay & Martin Richards. 2009. Correcting for purifying selection: An improved human mitochondrial molecular clock. *American Journal of Human Genetics* 84. 1–20.
- Sørensen, Per. & T. Hatting. 1967. *Archaeological Excavations in Thailand, Volume 2*, Ban Kao Neolithic Settlements with Cemeteries in the Kanchanaburi Province, part 1: The Archaeological Material from the Burials. Copenhagen: Munksgaard.
- Sørensen, Per. 1972. The Neolithic cultures of Thailand (and north Malaysia) and their Lungshanoid relationships. In Noel Barnard (ed.) *Early Chinese Art and its Possible Influence in the Pacific Basin, Volume 3*, 459–501. New York: Intercultural Arts Press.
- Srisuchat, Amara 1993. Comments on the results of Carbon-14 dating applied to archaeological finds from excavated sites in Thailand. In Tharapong Srisuchat & Amara Srisuchat (eds) *An Application of Technology and Science in Archaeological Work in Thailand*, 41–50. Bangkok: Department of Fine Arts, Archaeology Division.
- Steele, Gentry D. 1970. Estimation of stature from fragments of long limb bones. In Thomas Dale Stewart (ed.) *Personal Identification in Mass Disasters*, 85–96. Washington D.C. Smithsonian Institute.
- Steele, Gentry D. & Thomas W. McKern. 1969. A method for assessment of maximum long bone lengths and living stature from fragmentary limb bones. *American Journal of Physical Anthropology* 31. 215–228.
- Storm, Paul. 1995. The Evolutionary Significance of the Wajak Skulls. *Scripta Geologica* 110. Leiden: National Natuurhistorisch Museum.

- Stringer, Chris B. 1994. Out of Africa: a personal history. In Matthew H. Nitecki & Doris V. Nitecki (eds) *Origins of Anatomically Modern Humans*, 149–174. New York: Plenum Press.
- Taylor, David, Oh Hwee Yen, Peta G. Sanderson & John Dodson. 2001. Late Quaternary peat formation and vegetation dynamics in a lowland tropical swamp: Nee Soon, Singapore. *Palaeogeography, Palaeoclimatology, Palaeoecology* 171. 269–287.
- Trevor, J. C. & Don R. Brothwell. 1962. The human remains from Mesolithic and Neolithic date from Gua Cha, Kelantan. *Federations Museum Journal* 7. 6–22.
- Turner, Christy G. II 1990. Major features of sundadonty and sinodonty, including suggestions about East Asian microevolution, population history, and late Pleistocene relationships with Australian aboriginals. *American Journal of Physical Anthropology* 82. 295–317.
- Turner, Christie G. II & James F. Eder. 2006. Dentition of the Batak people of Palawan Island, the Philippines: Southeast Asian Negrito origins. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 172–187. Cambridge: Cambridge University Press.
- Wagenseil, Ferdinand. 1967. Anthropologischer Untersuchung ostmalayischer Negrito. *Zeitschrift für Morphologie und Anthropologie* 59. 1–25.
- Wainscoat, James S., Adrian V. S. Hill, S. L. Thein & J. B. Clegg. 1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. *Nature* 319. 491–493.
- Weidenreich, Franz. 1947. Facts and speculations concerning the origin of *Homo sapiens*. *American Anthropologist* 49. 187–203.
- White, Joyce C. & Chester F. Gorman. 2004. Patterns in “amorphous” industries: the Hoabinhian viewed through a lithic reduction sequence. In Victor Paz (ed.) *Southeast Asian Archaeology: Wilhelm G. Solheim II Festschrift*, 411–441. Quezon City: University of the Philippines Press.
- Wiradnyana, Ketut. 2008. Gua Togi Bogi, hunian berciri Mesolitik di Nias. Viewed April 16 2009 at <http://balarmedan.wordpress.com/category/drsketut-wiradnyana/page/12/>.
- Wolpoff, Milford H. 1980. *Paleoanthropology*. New York: Knopf.
- Zainuddin, Zafarina. & Will Goodwin. 2003. Mitochondrial DNA profiling of modern Malay & Orang Asli populations in peninsular Malaysia. In Christian Doutremepuich & Niels Morling (eds) *Progress in Forensic Genetics*, 428–430. Amsterdam: Elsevier Science.
- Zolkurnian, Hassan. 1998. *Urutan Kebudayaan Prasejarah Lembah Lenggong, Hulu Perak, Perak pada zaman Holosen*. Sarjana (M.A.) dissertation. Centre for Archaeological Research Malaysia, Universiti Sains Malaysia.
- Zuraina Majid. 1990. The Tampanian problem resolved: archaeological evidence of a late Pleistocene lithic workshop. *Modern Quaternary Research in Southeast Asia* 11. 71–96.
- 1998. Radiocarbon dates and culture sequence in the Lenggong Valley and beyond. *Malaysia Museums Journal* 34. 241–249.
- 2003. *Archaeology in Malaysia Arkeologi di Malaysia*. Penang: Centre for Archaeological Research Malaysia.
- Zuraina Majid, Mohd. Mokhtar Saidin, Stephen Chia Ming Soon & Zolkurnian Hassan. 1994. Artifacts from Gua Gunung Runtuh excavation. In Zuraina Majid (ed.) *The Excavation of Gua Gunung Runtuh and the Discovery of Perak Man in Malaysia*, 149–168. Kuala Lumpur: Department of Museums and Antiquity Malaysia.
- Zuraina Majid, Ang Bee Huat & Jeffri Ignatius. 1998. Late Pleistocene-Holocene sites in Pahang: excavations of Gua Sagu and Gua Tenggek. *Malaysia Museums Journal* 34. 65–115.



# 11 *Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula*

---

NICLAS BURENHULT, NICOLE KRUSPE AND MICHAEL DUNN

## 1 Introduction

The Malay Peninsula is a crossroads for people, languages and cultural influences, apparent in today's vibrant mix of Malay, Chinese, Indian, Thai and European. Yet this modern state of affairs all but conceals signals of much older situations of diversity. Thus, some 140,000 people grouped together under the label Orang Asli (Malay for 'aboriginal people') represent a range of cultural and biological adaptations and linguistic diversifications with roots far back in prehistory. These 20-plus ethnolinguistic groups represent a unique and vanishing window on the history of human diversity in the region, and they offer intriguing examples relevant to more general issues of the dynamics of human societies.

By synthesising the current ethnographic, linguistic and genetic body of knowledge about these groups with our own quantitative analyses of new lexical data from 27 language varieties, we explore the local historical relationships and interaction between languages and cultures. Specifically, we look at the relationship between a particular subsistence mode, namely nomadic foraging, and the Aslian branch of the Austroasiatic language stock. While foraging has been considered in many previous accounts to have a historically close connection to one particular sub branch of Aslian (Northern Aslian), we highlight several mismatches in this correlation and take a step toward disentangling a complex picture of linguistic history and contact.

## 2 The forager problem

The Northern Aslian (Aslian, Austroasiatic) languages of the Malay Peninsula have long been considered to be closely associated with a particular societal and economic tradition of nomadic foraging, upheld by the so-called Semang (Benjamin 1985; Rambo 1988; Fix 1995, 2002; Bulbeck 2004). The Semang comprise ten or so ethnolinguistic groups scattered over parts of the Malay Peninsula (Peninsular Malaysia and Isthmian Thailand), and together they number about 3000–4000 individuals. Along with proposed similar correlations of language and societal-economic system in other Aslian-speaking settings in the peninsula (Central Aslian with Senoi swidden cultures and Southern Aslian



with Aboriginal Malay collector-traders), the notion of a forager-Northern Aslian connection has shaped much of the discussion of peninsular prehistory.

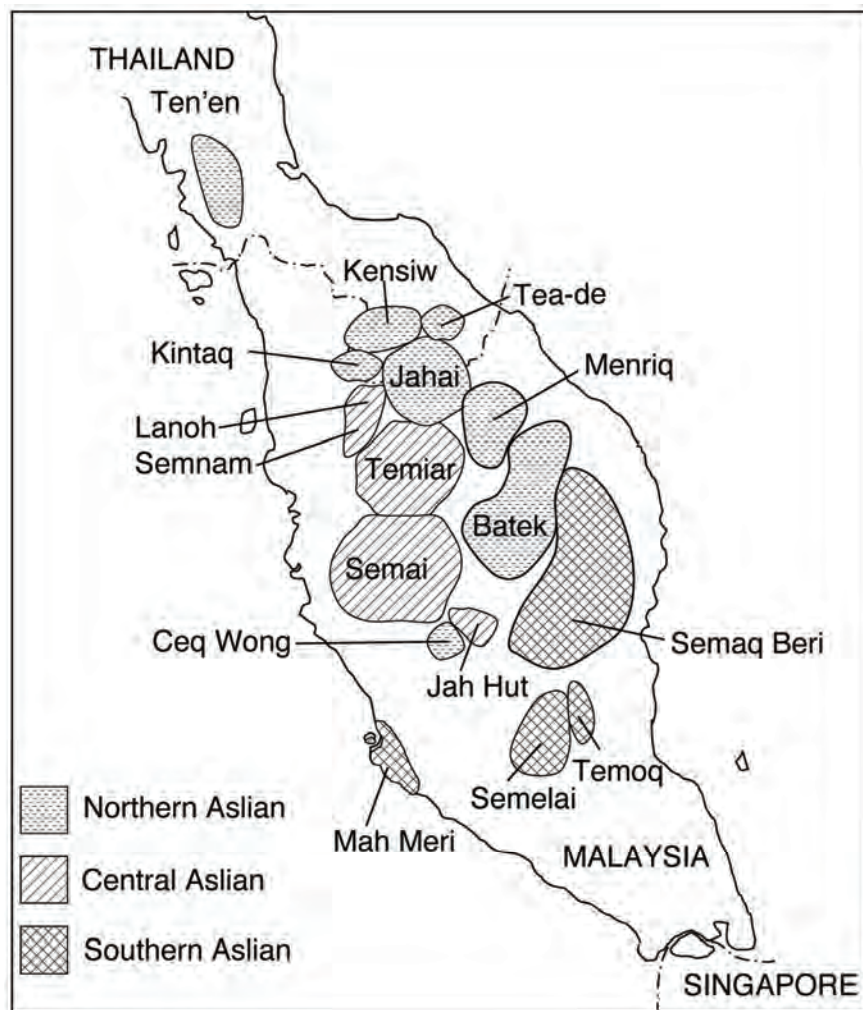
But the forager-Northern Aslian connection is only a near-match. A small group of foragers (c. 300) speak dialects of Lanoh, which belong to the Central Aslian branch of Aslian. Conversely, Ceq Wong, a geographical outlier of Northern Aslian, is spoken by a group of about 300 individuals who are not classified as Semang and whose subsistence is not focused on foraging. Furthermore, Semaq Beri is a Southern Aslian language spoken by a group of c. 2300 people who are not Semang but whose economy contains a significant component of nomadic foraging. Some earlier models tend to treat these ‘mismatches’ as exceptions, mixed societies which are presumably the results of recent departures from the main societal-linguistic categories.

The forager-Northern Aslian near-match is overlapped by a biologically defined category: for a long time, observers have maintained that there is a close connection between the Semang societal sphere and the physical features of its bearers. The reportedly short stature, dark skin and curly hair of the Semang led early anthropologists to classify them physically as ‘Negritos’ (see for example Schebesta 1952), a term which is still used to some extent. They were considered to represent the oldest human stratum in the peninsula. Superficial physical characteristics have been played down or questioned in subsequent accounts, and it has been shown that Semang history goes no further back than that of other indigenous groups in the peninsula (Bulbeck 2003). Still, recent genetic studies confirm that there is a close association between the Semang and an ancient local genetic lineage which goes far back into the Pleistocene and is represented only partially in other peninsular populations (Hill et al. 2006; Oppenheimer, this volume). Burenhult (forthcoming) proposes that continuous mobile foraging has been the best subsistence niche for preserving this ancient genetic lineage, and that today’s forager groups represent the current cultural exploiters of that niche. Yet the Semaq Beri—speakers of a Southern Aslian language who are not classified as Semang and whose relationship with the ancient genetic lineage has not been examined—also lead a predominantly mobile, foraging way of life.

Considering these categorical discrepancies, which genealogical histories and interactional dynamics can account for the relationship between language, culture and genes currently observed in this setting? With a focus on linguistic issues, the following sections revisit Aslian history and research (§3), introduce the ethnolinguistic categories concerned (§4), explore the nature and degree of linguistic contact between them (§5), and discuss the findings in light of current genetic and ethnographic knowledge (§6).

### **3 Aslian history and research**

Aslian is a genealogically and geographically well-defined branch of the Austroasiatic language stock (see Figure 1). It is made up of some 20 languages spoken by minority groups in the Malay Peninsula, mainly in the rainforested areas of the interior. Lexicostatistics (Benjamin 1976) and comparative historical phonology (Diffloth 1975) have produced similar family trees for Aslian, with three main subgroups: Northern, Central and Southern Aslian. The three are considered to have branched off early from each other, soon after their common ancestral language arrived in the peninsula an estimated 4000–5000 years ago.



**Figure 1:** Map of the Malay Peninsula showing the approximate distribution of Aslian languages and subgroups (adapted from Benjamin 1976 and Burenhult, forthcoming).

The three Aslian clades coincide broadly with three ethnographically defined subgroupings of indigenous societal and economic features (Benjamin 1980, 1985). Thus, according to this widely accepted classification, Northern Aslian is by and large associated with the nomadic foragers known ethnographically as the Semang; Central Aslian is associated with semi-sedentary swidden horticulturalists referred to as the Senoi; Southern Aslian is linked to groups of collector-traders called Aboriginal Malay (and, in Benjamin's work, to a larger Malayic societal pattern). Benjamin's tripartite societal division is based primarily on distinct categories of specific institutionalised kinship regimes, namely marital patterns and kin-avoidance rules. Other societal and economic features map more or less well onto these categories, forming the more general and less robust Semang-Senoi-Malay patterns (Benjamin 1985, forthcoming). Importantly, some Aslian ethnolinguistic groups are difficult to classify according to this more general ethnographic division, either because they display a mix of societal-economic features, or because their linguistic identity does not match the expected societal-economic one (see further below).

Despite its complexities, the tripartite model of peninsular ethnography has provided the analytical categories and the basis for sampling in a number of subsequent studies of

Malayan indigenous cultures and history (see especially Rambo 1988; Fix 1995, 2002, this volume; Bulbeck 2004, this volume; Hill et al. 2006; Oppenheimer, this volume). Benjamin has always cautiously pointed to the complexities of his model. However, some authors' subsequent analytical employment of it has frequently focused on the main patterns and has taken the broad association between societal type and language group for granted. As a result, those ethnolinguistic groups which do not conform to the main patterns of the model have typically attracted less attention.<sup>1</sup>

Accordingly, most interpretations of the prehistory of the Malay Peninsula rely on Benjamin's 1985 classification and the associated Aslian family tree. Models by Rambo (1988) and Fix (1995) imply that Benjamin's categories can be projected into the past and suggest a common genetic, linguistic and cultural origin of all of the peninsula's indigenous groups (cf. Bulbeck 2004). This 'indigenist' perspective has largely developed in response to 'migrationist' models which emphasise demic diffusion and linguistic colonisation from the Southeast Asian mainland to the north during the Neolithic (see especially Bellwood 1985, 1993).<sup>2</sup>

Recent genetic studies pose a problem for the indigenist paradigm, essentially because they show that genetic lineages of the indigenous groups display various geographical origins and varying antiquity within the peninsula (Hill et al. 2006; Oppenheimer, this volume). Locally ancient (Pleistocene) haplotypes are present in all groups but show a particular association with the Semang. Furthermore, there is evidence of a considerable influx of lineages from Mainland Southeast Asia around the time of the arrival of agriculture, c. 5000 years ago. These lineages are particularly apparent in the Senoi group.

These advances have prompted significant reinterpretations of genetic and linguistic prehistory in the peninsula. For example, Fix's modelling (this volume) demonstrates how limited genetic influx during the Neolithic has affected the subsequent biological history of the indigenous groups. Also, Burenhult (forthcoming) elaborates a scenario combining elements of local genetic continuity with demic diffusion and language shift to account for the current distribution of Aslian languages, as well as genetic diversity of their speakers. According to this hypothesis, when the ancestor of the Aslian languages entered the peninsula in connection with the introduction of agriculture, some local foragers adopted the new economy and Aslian language, and they intermixed to a greater or lesser degree with the Neolithic immigrants (cf. Bellwood 1985). The hypothesis further proposes that it was in this diverse setting of intermixing that the Aslian sub branches split from the introduced proto-language, the clades coming into being in distinct situations of cultural and linguistic contact. Some foragers retained their economy and nomadic lifestyle (the predecessors of today's Semang), but at some point the pre-existing ties between them and the more settled, intermixed groups led to eventual language shift such that also the still foraging people spoke Northern Aslian languages.

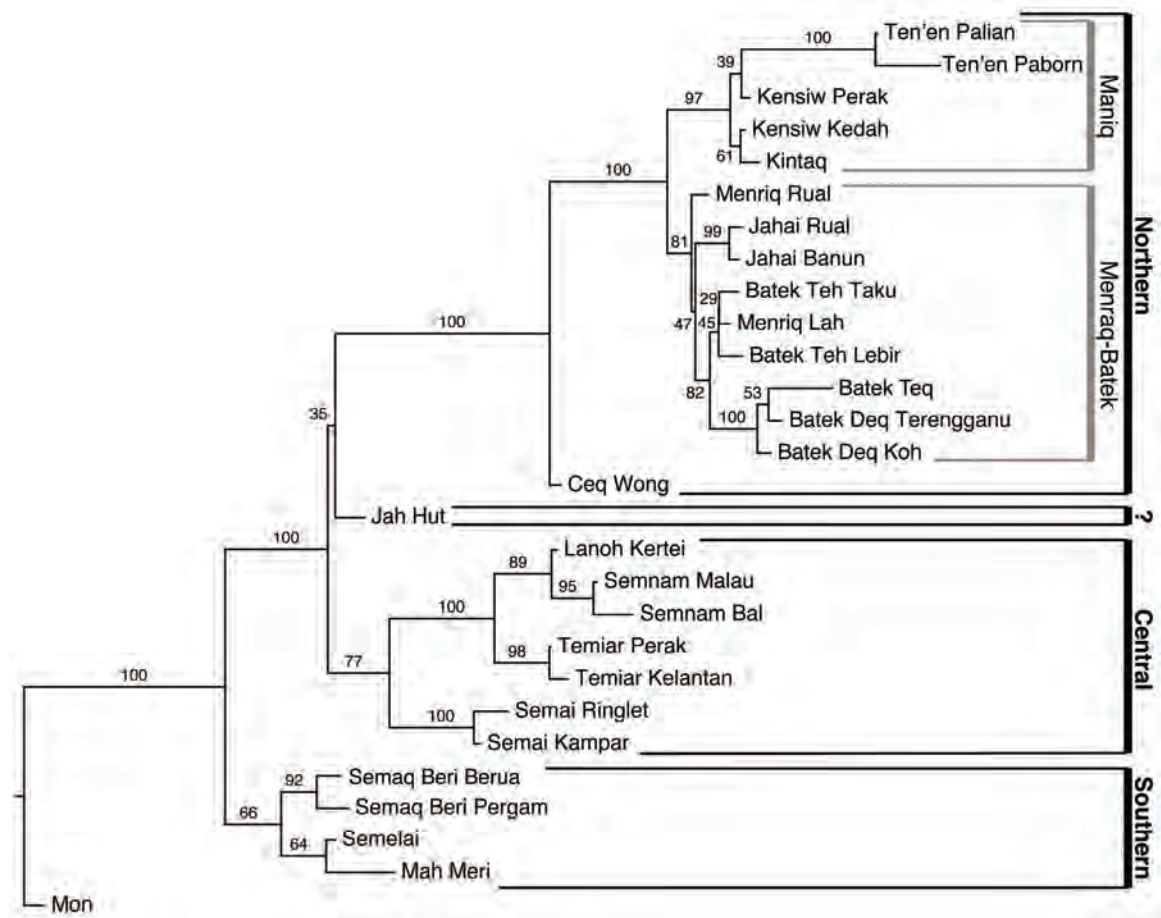
Burenhult argues that the initial splitting up of three Aslian sub branches should not be automatically connected to the formation of the three societal-economic subgroupings and he places emphasis on the ethnolinguistic groups which do not match the general societal-

---

1 Bulbeck (2004) is a notable exception, where the significance of 'mismatches' like the Ceq Wong and Jah Hut is given prominence in interpreting peninsular prehistory. Benjamin himself has explicitly stated the need for ethnographic and linguistic research on some of these groups (Benjamin 1976, 1989).

2 Again, Benjamin has remained cautious about projecting the tripartite model back in time and has also called for a clear analytical distinction between linguistic, cultural and biological history (Benjamin 1989), a suggestion reiterated by Burenhult (forthcoming).

linguistic correlations. Thus, the idea that Northern Aslian crystallised in a non-forager setting is supported by the existence of one distinct and conservative Northern Aslian language—Ceq Wong—spoken by a semi-sedentary group with a mixed economy which does not belong to the Semang forager sphere. Another mismatch is a small group of Central Aslian languages—Semnam and Lanoh—spoken by foragers with both Semang and Senoi-like cultural features: most likely the result of a later language shift from Northern Aslian. The rest of the Central Aslian languages, as well as most Southern Aslian languages, are spoken by non-Semang. However, one Southern Aslian language (Semaq Beri, not discussed in Burenhult, forthcoming) is spoken by people with a clear focus on foraging, although they are not included in the Semang forager grouping (see §4.3).



**Figure 2:** Aslian family tree, rooted on Mon (from Dunn et al.). This is a Maximum Clade Consistency tree, summarising the 750 post-burn in trees of the Bayesian phylogenetic tree sample with branch length equal to the median length of all congruent branches found in the sample. Numbers on the branches indicate percentage of the tree sample supporting each bifurcation (for details, see Dunn et al.).

In a recent paper, Dunn et al. revisit Aslian history by analysing newly collected lexical data from 27 Aslian varieties with quantitative methods to produce measures of linguistic divergence as well as phylogenetic hypotheses.<sup>3</sup> While broadly reproducing the earlier

3 The same dataset is used in the present chapter to analyse post-split contacts, see §5.2.

proposed clades of Aslian genealogy, the phylogenetic aspect of this study also reveals that the three major clades show very unequal rates of lexical divergence: Southern Aslian is the most conservative branch; Central Aslian shows a bit more divergence; and most of Northern Aslian is contained within a clade which is highly divergent externally, but which has low internal diversity (suggesting a recent diversification; see Figure 2). It is the Northern Aslian languages spoken by Semang foragers which show great external divergence whereas the geographic and cultural outlier Ceq Wong comes out as an early and conservative split. So while the degree of lexical divergence coincides with societal distinctions, neither degree of divergence nor societal features coincide with the Northern Aslian clade. Specifically, the conservative nature of Ceq Wong suggests that the initial branching of Northern Aslian did not occur in a Semang-type cultural setting, nor did it coincide with its development. Instead, the Semang-Northern Aslian link and its accelerated lexical divergence is likely to be a later phenomenon. This lends support to the idea that the prehistoric environment in which Aslian spread and branching occurred was culturally, linguistically and biologically diverse.

#### **4 The ethnolinguistic categories**

In this chapter we examine the relationship between language and culture with a particular focus on the role of linguistic contact in analysing the history and development of peninsular societal-economic patterns. Acknowledging that today's broad Northern Aslian connection with foraging does not necessarily have deep historical pertinence, we sidestep analytically the traditional categories 'Northern-Central-Southern Aslian' and 'Semang-Senoi-Aboriginal Malay'. Instead, we examine a broader range of ethnolinguistic groups which cross-cuts these categories. Our dataset represents the Aslian branch as a whole, but our main area of analysis and discussion is groups whose subsistence mode has been described ethnographically or historically as displaying some degree of mobile foraging (see below). These include ethnolinguistic groups speaking languages of the Northern, Central and Southern Aslian branches. They also cross-cut the three conventional societal-economic groupings: Semang, Senoi and Aboriginal Malay. One advantage of this ethnolinguistically based sample is that groups which are difficult to classify according to the conventional tripartite paradigm (because they are in some sense 'mixed') are not anomalies which are best avoided, but can be treated on a par with those whose societal-economic characteristics are more faithful to the overall classification.

The following sections briefly outline the main linguistic and ethnographic features of four ethnolinguistically defined groups which are particularly pertinent to the ensuing analysis: Maniq/Menraq-Batek (§4.1), the Lanoh complex (§4.2), Semaq Beri (§4.3), and Ceq Wong (§4.4). The authors have conducted first-hand linguistic and ethnographic field work among these groups, of which most have not previously received linguistic attention. The descriptions are based on a variety of sources, including the authors' recent and unpublished findings.

##### **4.1 Maniq/Menraq-Batek (MMB)**

On linguistic grounds, this group ('Northern Aslian-speaking Semang') can be divided into two subgroups (Benjamin 1976; Dunn et al.). One subgroup comprises languages and dialects in Isthmian Thailand, such as Ten'en, Kensiw and Tea-de, as well as Kensiw and Kentaq in northernmost Peninsular Malaysia. Following Burenhult (forthcoming), these

varieties will here be referred to generically as Maniq (the preferred endonym for many of these groups, from their word for human being, *mani?* or *مني?*). The other subgroup is found predominantly in Malaysia (Perak, Kelantan and Pahang states) and comprises three languages: Jahai, Menriq and Batek. While Jahai and Menriq share a term for human being (*mnra?*), the Batek term is the same as the ethnonym (*batek*). This group is referred to here as Menraq-Batek. The Maniq/Menraq-Batek (MMB) are the only groups which combine the Semang societal tradition with Northern Aslian language, and they have a particularly close connection with locally ancient genetic lineages (as shown by samples analysed by Fucharoen et al. 2001 and Hill et al. 2006). The MMB languages show the greatest lexical divergence of all Aslian languages (Dunn et al., see Figure 2).

Traditionally, the MMB live in groups of about 15–50 people. Temporary camps of lean-tos or huts are inhabited for a few days to several weeks or months (sometimes even years), depending on the sustenance circumstances. The economy is based on foraging in a broad sense. Hunting, fishing and gathering form the backbone of their subsistence. But the MMB also make occasional swiddens (especially the Menraq-Batek groups), collect rainforest products for trade with outsiders, and seize any opportunity to engage in wage-labour, if such activities are considered to be economically advantageous at the time. Nowadays many MMB are permanently settled in resettlement villages established by the Malaysian and Thai governments, but some groups in both countries still pursue a mobile existence.

Semang society promotes this mobile lifestyle, its social structures encouraging dispersal and flux in space, time and human relations (Benjamin 1985). The conjugal family is the only persistent social unit, and bands and camps consist of several such families which co-exist on a voluntary basis. Strict cross-sex in-law avoidance rules apply, the filiative bias is patrifocal and residence is virilocal. Marriage is strictly exogamous. Society is egalitarian, and there is a moral obligation to share food with other members of a camp (van der Sluys 2000). For detailed accounts, see for example Schebesta (1952), Endicott (1979) and Lye (1997). Benjamin's 1985 classification suggests the Batek tend toward a Malayic societal pattern, have no dominant subsistence mode, and are indeed 'mixed' rather than Semang. Depending on which features are considered critical to classification, it could also be argued that other groups represent a mixed pattern at the present time, for example the Jahai and the Menriq.

## 4.2 The Lanoh complex

Lanoh is a generic label for a cluster of dialects spoken historically along portions of the middle and upper Perak river, Peninsular Malaysia. Most of these dialects are extinct (for example Sabüm) or moribund (for example Yir); Semnam, spoken in one village, still has some 300 speakers (Burenhult & Wegener 2009). They belong to the Central Aslian sub branch of Aslian and are closely related to Temiar, one of the main languages of the Senoi societal sphere. Until recently the subsistence system of Lanoh speakers was based on nomadic foraging very similar to that pursued by the MMB (Dallos 2011; Burenhult, field notes 2006). In some early ethnographic accounts they are described as a particularly reclusive and mobile Semang group (see for example Schebesta 1928). Like the MMB, the Lanoh have a band-based society and the conjugal family is the basic social unit of

production. Their cross-sex in-law avoidance rules are similar to those of the MMB.<sup>4</sup> However, their filiative bias, which is cognatic rather than patrifocal, is shared with the Senoi rather than with the Semang. Benjamin therefore classifies them as belonging to the Senoi tradition instead of the Semang (Benjamin 1985:251, forthcoming).

While no genetic analyses have been carried out on Lanoh speakers, unsystematic phenotypic observations in the literature suggest a ‘Negritoid’ appearance similar to MMB speakers and thus possibly a close connection to the locally ancient genetic lineages. Their territory abuts on that of some MMB groups, and historically they were in close contact with the Kensiw, Kentaq and Jahai. They are currently undergoing assimilation by the Temiar. The fact that Lanoh belongs to the Central Aslian sub branch can possibly be attributed to a language shift among some previously Northern Aslian-speaking Semang (Burenhult, forthcoming).

### 4.3 **Semaq Beri**

The Semaq Beri speak a language belonging to the Southern Aslian sub branch and traditionally inhabited an area north of the Pahang River in the upper reaches of the Kuantan, Kemaman, Dungun, Tembeling, Terengganu and Lebir Rivers. They have been characterised as ‘mixed’, an ‘ill-defined and heterogeneous group’ (Benjamin 1985). Southern Semaq Beri reportedly display the attributes of Benjamin’s Malayic grouping (see §3), while northern Semaq Beri are nomadic foragers whose significantly variable phenotypic features have led observers to suggest a mixed genetic lineage (Endicott 1975:4–5). The Semaq Beri-speaking foragers are traditionally not included in the Semang ethnographic grouping. However, it is clear that the northern Semaq Beri live like their MMB-speaking neighbours (the Batek) in small camps of lean-tos combining hunting, fishing and foraging with the occasional collection of forest produce for trade, or waged labour (Kuchikura 1987; Morris 1996). Societal features also have much in common with those of the Semang. Thus, the Semaq Beri society is egalitarian, band-based, and there is a strong moral obligation to share food with other members of one’s group. The conjugal family is the primary social unit. There is no fixed pattern of post-marriage residence, although there is some preference for virilocal residence in established marriages (Kuchikura 1987:23). Strict cross-sex in-law avoidance is observed. In addition, there is strict cross-sex avoidance between parents and children and between cross-sex siblings after the onset of puberty (Kruspe, field notes 2009).

---

4 In Benjamin’s typology of cross-sex relations (1985:252) the Lanoh pattern with Senoi groups, ‘sister-in-law’ relations being characterised by restraint with one’s spouse’s older siblings and by joking with one’s spouse’s younger siblings. However, recent work shows that the Semnam subgroup of Lanoh has ‘sister-in-law’ relations which are identical to those given by Benjamin for the MMB groups Kensiw and Kentaq. Here, avoidance is observed in relations with one’s spouse’s older siblings while relations with one’s spouse’s younger siblings are neutral (Burenhult, field notes 2008). Benjamin’s typology also characterises Lanoh brother/sister relations as neutral, in line with a common Senoi pattern and in contrast with a unified Semang pattern of avoidance. This is the case also in Burenhult’s recent Semnam data. However, with regard to brother/sister relations among the Semang Burenhult’s data diverge from Benjamin’s: the Jahai observe restraint rather than avoidance, which suggests that brother/sister avoidance is a less common Semang pattern than Benjamin’s classification indicates, at least at present.

Exogamous marriage forms an integral component of the foraging mode, and first cousin marriage is typically avoided.<sup>5</sup> All the Semaq Beri maintain close social relations with other bands across their extensive range, as well as close relations with their immediate Orang Asli neighbours. Presently, the northern Semaq Beri have close relations with the Batek Deq (Kuchikura 1987:9, 16–17; Kruspe, in prep), resulting in significant gene flow and cultural and linguistic influences, for example, inter-Aslian bilingualism which is absent in the southern group. The northern variety of Semaq Beri exhibits lexical and grammatical influence from Northern Aslian, which is less prevalent in southern varieties of Semaq Beri (§5 below, Kruspe, in prep). In the south the Semaq Beri have established social relations with their Aboriginal Malay neighbours the Jakun (Ramle 1993:43), giving rise to different genetic inflow, and cultural-linguistic influences. There are no genetic studies of the Semaq Beri.

There has been speculation as to the origins of foraging by the northern Semaq Beri. Benjamin (1976, 1985) treats their nomadism as a deviation from the defined cultural-linguistic division (collector-trader Southern Aslian) proposing that the Semaq Beri are ‘secondarily nomadic’ having shifted to foraging after splitting off from Semelai (another Southern Aslian language; Kruspe 2004). Endicott (1975) similarly provides a recent-convert scenario. However, ethnohistorical accounts from southern Semaq Beri confirm long-term engagement with nomadic foraging activities, only becoming sedentary in response to external pressure (Kruspe, in prep.; Evans 1915), suggesting that nomadic foraging may have been the dominant mode for all Semaq Beri in the past.

#### 4.4 Ceq Wong

The Ceq Wong are unique within Northern Aslian, geographically isolated from the remainder of the group and not part of the Semang societal sphere. They may represent a relic Northern Aslian population: phylogenetic analysis of the Ceq Wong lexicon clearly identifies the language as a conservative/relic variety (Dunn et al.; Burenhult, forthcoming; see also §5.2. below). All the other attested Northern Aslian languages are considerably divergent from it. The Ceq Wong practise a ‘mixed’ economic adaptation combining both foraging and swiddening, however their residence pattern is semi-sedentary like that of swidden horticulturalists. Furthermore, unlike the MMB, some of whom also engage in occasional swiddening but whose main activity is foraging, the Ceq Wong place emphasis on both activities. Despite this, they perceive themselves as subsistence foragers or ‘digging people’ (Howell 1989:13), in reference to their dependence on wild yams and ability to survive in the forest, unlike their Central Aslian neighbours the Jah Hut. Their simple swiddens in which manioc is grown appear to provide a supplement to foraging, and are often abandoned in favour of other economic activities. Ethnohistorical accounts from the Ceq Wong suggest that until recently they were more mobile (Howell 1989:21–22).

Ceq Wong society is egalitarian and there is a strong moral obligation to share food with other members of the group. The primary social unit is the conjugal family, which functions as an autonomous entity. Residence was traditionally in a camp or swidden in the forest with one or more families. These days some people choose to live in a government

---

5 Kruspe’s ethnographic data differ from those of Benjamin (1985), according to whom cousin marriage is permitted among the Semaq Beri and cross-sex relations are characterised by restraint rather than avoidance.



settlement, but residence remains fluid. There are frequent movements between settlements, and between settlements and the forest, which may last from a few nights to several months or longer. Temporary lean-tos provide shelter in the forest, while more permanent houses are constructed in the swiddens.

In practice, most marriages are endogamous and first cousin marriage is currently the most common union, but curiously the Ceq Wong maintain the cultural ideal that marriage should be exogamous (with someone unrelated) otherwise it is incestuous (Howell 1989:28; Kruspe, field notes 2002). Both polygyny and polyandry are practised. There is uxorilocal residence immediately following marriage, after which residence becomes bilocal. The Ceq Wong observe restraint in cross-sex parent-in-law relations.

There are no genetic studies of the Ceq Wong. Unsystematic observation has seen the Ceq Wong classified phenotypically as Senoi, like the Central Aslian Jah Hut. There is evidence of intrusive genetic flow from sporadic intermarriage with the Jah Hut. A colonial report notes some Ceq Wong showing ‘much Negrito blood’ which was attributed to intermarriage with Batek Nong, a Semang group to the north (Ogilvie 1948:15, 29). The present day Ceq Wong are not aware of the existence of the Batek Nong.

## **5 Contact: lexical and other evidence**

### **5.1 Sociolinguistics**

Evidence suggests that some foraging societies of the Malay Peninsula share linguistic characteristics which cross-cut Aslian genealogical boundaries. For example, Benjamin (1976:74–76, 1980:4, 1985:234–235, 2001:111) has long argued that the Semang display distinct sociolinguistic features (see also Endicott 1997). The mobile lifestyle of the Semang, manifested in their system of intermarriage between individuals of widely dispersed bands, as well as in their pattern of group disintegration and regrouping into new constellations in response to changing subsistence conditions, is linked to particular patterns of individual language use. A speaker may move through several linguistic environments throughout his or her lifetime, which leads to a high rate of idiolectal change. At the same time, the diverse linguistic origins of members of a band also lead to marked variation in the language use of different speakers. Benjamin (2001:111) discusses this in terms of a mesh-like relation between language varieties which is idiolectal as much as dialectal. All of the MMB languages form a continuum of such linguistic interaction and, in the sociolinguistic sense, represent a unitary linguistic constellation. The participation of Lanoh, Semaq Beri and Ceq Wong in this constellation has so far been less clear.

### **5.2 Lexicon**

In a previous study of linguistic divergence in Aslian, Neighbor-Net clustering of the lexicostatistical distance data, which has the advantage of being able to show ‘conflicting signal’, revealed a notable split between MMB-Lanoh on the one hand, and the rest of Aslian on the other (Dunn et al.).<sup>6</sup> This split partially cross-cuts genealogical boundaries,

---

6 The lexical data used in Dunn et al., also used in the new analysis of post-split contacts presented below, comprise lists of basic vocabulary from twenty-seven Aslian varieties. The bulk of these lists were collected recently (1990–2008) in the field by Burenhult and Kruspe, with additional wordlists provided by Neele Becker and Sylvia Tufvesson. For details, see Dunn et al.

showing evidence clustering the Lanoh complex with MMB (that is, Northern Aslian languages apart from Ceq Wong). This, it seems, is a lexically traceable variant of Benjamin's sociolinguistic category, and one that includes Lanoh. Although displaying some obvious MMB loans, Southern Aslian Semaq Beri does not participate in this clustering (see Figure 3).

The Neighbor-Net network is being used to represent the lexicostatistical distance, that is, proportion of shared and unshared cognates between each pair of languages in the list (where a distance of 0.0 indicates identity and 1.0 indicates that the lists are completely different; note that in lexicostatistics it is common to report proximity, the inverse of distance). In this chapter we take the analysis a step further by comparing this measure to a more realistic measure<sup>7</sup> of evolutionary distance, calculated from a phylogenetic tree. The Aslian family tree proposed in Dunn et al. is our best estimate of the phylogenetic relationships within the family given the data we have available. This language sample is large—containing data on more Aslian languages than have ever been presented together before<sup>8</sup>—but varies considerably in the amount of data which is feasibly obtained for each language. The common baseline is a form of the 200-word Swadesh list (used previously by Benjamin 1976), which was collected for all languages of the sample, as well as some outgroup languages (other Austroasiatic languages, not part of the Aslian family). These lists were then coded for probable cognacy on the basis of explicit criteria for identifying similarities in form-meaning mapping.<sup>9</sup>

The phylogenetic relations between the Aslian languages were estimated using computational methods from a family of techniques known as Bayesian phylogenetic inference. These methods allow rich inferences from lexical cognate data, by modelling the evolution of a language family as the gain and loss of reflexes of cognate sets, and are applicable without requiring for example a detailed understanding of regular sound change within the family. The result of this phylogenetic inference includes not only a tree topology (as would be produced by the linguistic comparative method, whenever an exhaustive reconstruction of Aslian should become available), but also includes a measure of statistical confidence for each branch in the tree, and (crucially for the argument below) a measure of the amount of evolutionary change on each branch. The details of our proposed family tree for the Aslian languages are presented in Dunn et al., and a summary tree of our best estimate of the phylogeny (rooted on Mon, a member of a sister clade to the Aslian family) is presented in Figure 2. From the tree sample underlying Figure 2, we calculated the median evolutionary distance between each pair of languages by summing the branch lengths between them in each tree in the sample.

We are contrasting the model of evolution used in this attempt at phylogenetic inference—the Bayesian approach using explicit evolutionary models—with the lexicostatistical model. The Bayesian tree is a better model of the history of the languages under consideration. It considers the evolution of each individual cognate set, allows for different rates of changes, and implicitly can even handle low levels of family-internal borrowing (through allowing low rates of 'spontaneous recreation' of reflexes of cognate

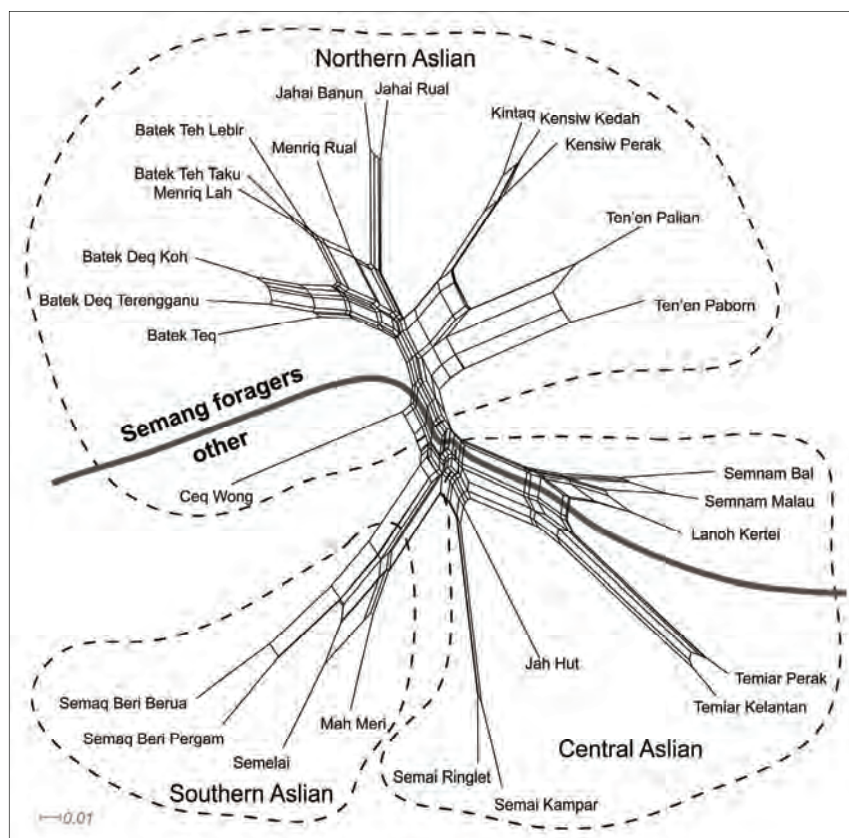
---

7 It is a character-based method, which models the history of the individual reflexes of cognate sets, rather than a distance measure, which just models aggregate (dis)similarity.

8 The pioneering lexicostatistical study of Aslian by Benjamin (1976) involved a similar sample. Both samples are broadly representative of Aslian but differ in detail. For a full account of the differences, see Dunn et al.

9 We refer to Dunn et al. for the criteria used in identifying probable cognates.

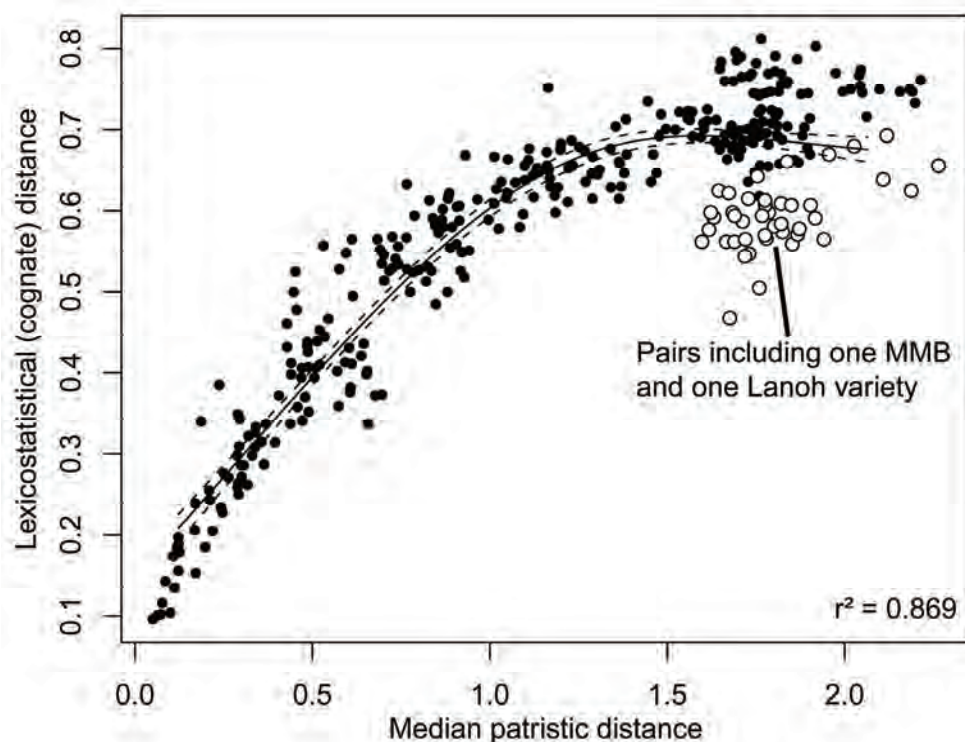
sets in branches where the set is extinct). Lexicostatistics measures pairwise similarity of languages in terms of proportion of shared cognates (or possible cognates) in a word list. As a model of historical relatedness, it assumes that similarity between these word lists is directly proportional to how closely the two languages are related. This is a very simple model of language change, and one that we know to be misleading in many circumstances. Lexicostatistical distances are a poor model of the historical processes involved in language change, and a set of lexicostatistical distances are only historically meaningful if one is willing to allow that lexical similarity has decayed at a constant rate, such that the proportion of non-cognate (or not recognisably cognate) words in a pair of languages should be proportional to the amount of time those languages have been historically distinct. The decay in lexical similarity can be due to lexical replacement, and (at deeper levels) also to processes of regular sound change, where the sum of sound changes has obscured the cognacy relationships. In particular, there is no space for family-internal borrowing in this model: borrowings of cognate terms cannot be distinguished from true, inherited cognate vocabulary.



**Figure 3:** Neighbor-Net clustering of lexicostatistical distances calculated from Aslian vocabulary lists (from Dunn et al.). The Neighbor-Net graph represents a matrix of distances without forcing the resolution of the major conflicts in the data.

In Figure 4 we compare lexicostatistical distance measures between each pair of languages to the sum of the branch lengths between these languages in a phylogenetic tree. If the lexicostatistical model of constantly decreasing similarity between languages accurately reflected the history of these languages, we would expect lexicostatistical and evolutionary patristic distance to be closely correlated.

Unsurprisingly, the correlation is strong ( $r^2=0.87$ , meaning that 87% of the variance in one distance measure can be predicted from knowing the value of the other), since these are two different models of the same data. Thus, in Figure 4, the data points (representing language pairs) mostly fall along a smooth curve representing the phylogenetically expected similarities. But there are also clusters of points which do not fit the curve very well. The white dots mark pairs of languages which include one MMB language and one language from the Lanoh complex. These languages tend to have more terms identified as cognates than their phylogenetic relationships would predict, a clear signal of intra-Aslian borrowing and a pattern restricted to a MMB-Lanoh sphere. Importantly, these pairwise similarities are consistent throughout MMB and the Lanoh complex, and are not restricted to, say, members of each group which are currently contiguous or cohabitant (although such pairs indeed represent the most extreme examples of phylogenetically unexpected similarity). This suggests either (1) a distinct pattern of long-term lexical diffusion/exchange within MMB-Lanoh or (2) an ancient contact situation between proto-varieties of Lanoh and MMB leading to lexical congruence subsequently inherited throughout both groups (including a scenario of language shift from MMB to Proto Lanoh). The former explanation is probably the more likely one, since it largely links up to Benjamin's characterisation of sociolinguistic interaction. Also, the ancient contact scenario faces a chronological obstacle, since Proto-MMB probably is considerably older than Proto Lanoh. Any traces of contact-induced transfer of lexicon in Lanoh would therefore presumably not be shared with MMB as a whole, but with some subsection of it.

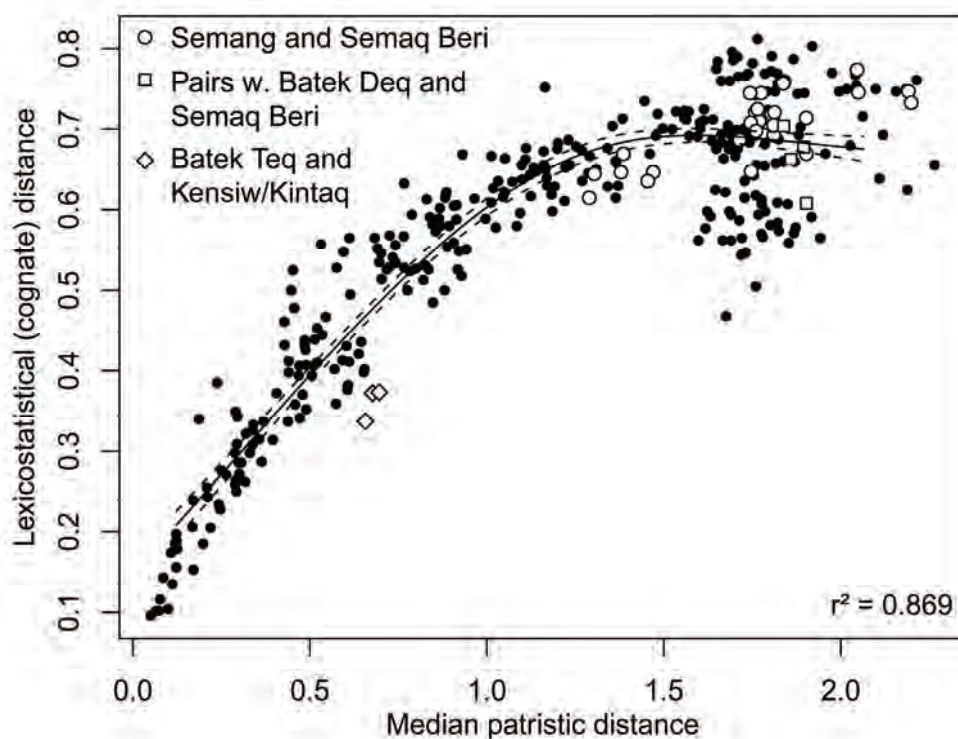


**Figure 4:** Phenetic x patristic distance (with pairs including one MMB variety and one Lanoh variety highlighted).

In Figure 5 we show the same data, but have marked pairs of languages including one MMB or Lanoh variety and one Semaq Beri (forager Southern Aslian) variety. For the

most part, these pairs occur along the expected curve, showing no indication of borrowing. However, a pair including Semaq Beri Berua and Batek Deq Terengganu shows a high degree of contact, whereas the pair including Semaq Beri Berua and another variety of Batek Deq does not. Pairs of one Batek Deq variety and one Semaq Beri variety are shown with squares instead of circles. This suggests very shallow contact between MMB and Semaq Beri, traceable only in one pair of varieties which are currently spoken in the same village.

Figure 5 also highlights a cluster of points which represents another set of pairs whose members are more similar than expected phylogenetically. These all include Batek Teq and some variety of Kensiw/Kintaq. This suggests the existence of a presumably old contact situation between these currently widely separated varieties, subsequently broken up geographically by the Menraq branch of MMB. Benjamin (1976:77) notes a similar geographic intrusion, but in the form of Jahai breaking up ancient contact between Kensiw and Menriq.



**Figure 5:** Phenetic x patristic distance (with pairs including one Semang variety and one Semaq Beri variety, as well as pairs with one Kensiw/Kentag variety and Batek Teq, highlighted).

Another thing to note is that Ceq Wong shows no major identifiable pattern of post-split contact with any other language, instead behaving as expected from the phylogenetic analysis. This lends further support to the idea that Ceq Wong represents a conservative Northern Aslian relic with few traces of secondary contact. There is no evidence, for example, that the language has made an exit from the Semang sphere and subsequently undergone distinct patterns of lexical change due to a new contact situation. Claims of borrowing between Ceq Wong and Kensiw (Benjamin 1976:78) and between Ceq Wong

and Semaq Beri (Endicott 1975:7, citing Diffloth, p.c.) are not supported by the current dataset.

### 5.3 Grammar

There is less evidence for contact in grammar, perhaps stemming partly from the fact that detailed grammatical data are still lacking for many of the languages. However, at least one morphemic category present in both MMB and Lanoh is yet to be discovered elsewhere in Aslian. This is a causative infix surfacing in some MMB languages as <ri> and in the Semnam variety of Lanoh as <yi> (the y in the latter is an expected reflex, given the lack of a phoneme /ɾ/ in Semnam). Another candidate is plural inflection in human nouns by means of an infix <ra> in some MMB languages and a corresponding <ya> in Semnam. A similar infix is present in Semaq Beri, which may have borrowed it from a MMB language, presumably Batek (Kruspe, in prep). It is also present in some Austronesian languages, though not currently productively in Malay. There is emerging evidence that some cognates of forms which belong to the class of expressives in other Aslian languages behave formally as stative verbs in MMB and Lanoh. However, this is a feature which may be shared by Ceq Wong (Kruspe, in progress). In fact, preliminary comparison based on Jahai and Ceq Wong suggests that the genealogical subgrouping Northern Aslian is a better predictor of grammatical similarities than sociolinguistic and societal-economic categories (Burenhult, forthcoming).

## 6 Conclusions

The indigenous communities of the Malay Peninsula represent a microcosm of human dynamics and complexity, and they provide a fascinating analytical setting for disentangling historical relationships between language, culture and genes. This paper has been especially concerned with the historical relationship between linguistic phenomena and a particular type of subsistence mode, namely nomadic foraging. The following conclusions can be drawn:

Firstly, there is evidence against a correlation between linguistic phylogeny (Northern Aslian) and societal type (Semang), so the current dominant societal mode of Northern Aslian speakers should not be taken as a reliable indicator of what the Proto Northern Aslian-speaking society was like. Our analysis of post-split contacts confirms that the Northern Aslian, non-Semang outlier Ceq Wong is a conservative relic which has not experienced significant secondary lexical exchange, and there is no indication that Ceq Wong made an exodus from the Semang cultural sphere. This suggests that the Northern Aslian clade crystallised in a non-Semang and possibly Ceq Wong-like setting, only secondarily spreading to the ancestral Semang (in connection with the branching off of MMB; see Burenhult, forthcoming). Yet, the forager-like aspects of Ceq Wong society point to complexities in the cultural settings in which Aslian was established and spread, and they evince the need to view some of the ‘mixed’ cultures as perhaps more archetypal and conservative than has typically been the case (cf. Bulbeck 2004).

Secondly, the lexical patterns revealed by the present study highlight a category which has typically gone unrecognised and unlabeled in recent literature. The MMB languages and those of the Lanoh complex show evidence of considerable secondary exchange of vocabulary, more so than any other Aslian setting. This pattern of linguistic exchange cross-cuts genealogical boundaries and Benjamin’s Semang-Senoi societal distinction, but

it is consistent with some other characterisations of the Lanoh as Semang-like. Incidentally, it also coincides with observed phenotypical characteristics. In all likelihood, the pattern reflects a distinct configuration of interaction, in line with Benjamin's notion of a Semang sociolinguistic entity but with the notable addition of Lanoh. Along with our new ethnographic data on kinship, these results call for a re-evaluation of the position of the Lanoh in relation to their neighbours.

Southern Aslian Semaq Beri, although partly spoken by foragers with very Semang-like societal features and currently co-existing with MMB neighbours, shows no deep traces of similar lexical contact. This, incidentally, suggests that Semaq Beri contact with the MMB is comparatively recent and that their foraging mode of subsistence possibly developed in response to this contact, potentially providing support to Endicott's (1975) and Benjamin's (1985) scenarios of secondary adoption of nomadic foraging. However, the presence of the foraging subsistence mode in all three branches of Aslian again brings into question the customary alignment of language and subsistence type. Semaq Beri foraging may be more conservative than previously assumed and recent linguistic contact with MMB may not necessarily be congruent with a switch in subsistence mode. The Semaq Beri may have traditionally foraged further to the south, possibly in the area where some southern Semaq Beri still reside. Recall the southern Semaq Beri ethnohistorical account of forager subsistence discussed in §4.3. The suggested recent Semaq Beri and Batek contact is possibly the result of the displacement of indigenous peoples in the historical period precipitated by the influx of Malay immigrants up the Pahang River and its tributaries and along other east coast rivers. Note again in this context the difficulty in defining the Lanoh in relation to the Semang and Senoi categories in purely societal terms—the Lanoh share societal features with both.

In addition to the lexically motivated category proposed here, genetics may offer similarly clear reflections of such a category, as indicated by Hill et al. (2006). However, in the absence of genetic data from the Ceq Wong, Lanoh and Semaq Beri, this cannot be substantiated. Genetic studies in these settings will be crucial to furthering our understanding of Aslian prehistory.

Thirdly, our analysis has revealed traces of ancient contact between geographically distant languages. The more-than-expected lexical similarity between Kensiw/Kentaq and Batek Teq suggests ancient geographical contiguity between Maniq and Batek varieties, possibly pointing to a past distribution of Aslian languages in the northeast of the peninsula (north Kelantan, southern Thailand) which was much wider than can be seen in the historical record. This contact seems to have come to an end with the expansion of Menraq languages northwards, and of Malay southward along the Kelantan river.

Finally, we hope to have shown that our conceptual approach—which sets out from the actual well-defined ethnolinguistic groups rather than generalised categories—provides a robust framework for analysis, and does justice to groups which have traditionally played a minor role in the exploration of Malayan indigenous history.

## **Acknowledgements**

We are grateful to David Bulbeck, Nick Enfield, Stephen Oppenheimer, Joyce White, and two anonymous reviewers for their helpful comments on earlier versions of the paper, as well as participants in the 19<sup>th</sup> International Conference on Historical Linguistics (Nijmegen, 2009) and the Dynamics of Human Diversity in Mainland Southeast Asia workshop (Siem Reap, 2009) for valuable comments on part of this work. The long-term

field research carried out by Burenhult was funded by the Max Planck Society, a European Community Marie Curie Individual Fellowship, and the Volkswagen Foundation's DOBES programme; and that of Kruspe by the Max Planck Society, the Research Centre for Linguistic Typology, the Hans Rausing Endangered Languages Programme, and the Volkswagen Foundation's DOBES programme. Dunn's research was sponsored by the Max Planck Society. We are also grateful to the Economic Planning Unit, the Jabatan Hal Ehwal Orang Asli, and Universiti Kebangsaan Malaysia for their helpful assistance, and we are especially indebted to the numerous Aslian-speaking communities who have generously shared their languages with us. Thanks also to Stephen Levinson, director of the Language and Cognition group at the Max Planck Institute for Psycholinguistics, Nijmegen, for providing the stimulating interdisciplinary research environment that enabled the authors to carry out this collaboration.

## References

- Bellwood, Peter. 1985. *Prehistory of the Indo-Malaysian archipelago*. Sydney: Academic Press.
- . 1993. Cultural and biological differentiation in Peninsular Malaysia: the last 10,000 years. *Asian Perspectives* 32. 37–60.
- Benjamin, Geoffrey. 1976. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds) *Austroasiatic studies*, Part I, 37–128. Honolulu: The University Press of Hawaii.
- . 1980. *Semang, Senoi, Malay: Culture-history, kinship, and consciousness in the Malay Peninsula*. Unpublished manuscript.
- . 1985. In the long term: three themes in Malayan cultural ecology. In Karl L. Hutterer, Terry Rambo & George Lovelace (eds) *Cultural values and human ecology in South East Asia*, 219–278. University of Michigan: Center for South and Southeast Asian Studies.
- . 1989. Achievements and gaps in Orang Asli research. *Akademika* 35. 7–46.
- . 2001. Orang Asli languages: From heritage to death? In Razha Rashid & Wazir Jahan Karim (eds) *Minority cultures of peninsular Malaysia: Survivals of indigenous heritage*, 101–122. Penang: Academy of Social Sciences.
- . forthcoming. The current situation of the Aslian languages. In Hein Steinhauer & James T. Collins (eds) *Endangered languages of Southeast Asia*. London: School of Oriental and African Studies.
- Bulbeck, David F. 2003. Hunter-gatherer occupation of the Malay Peninsula from the Ice Age to the Iron Age. In Julio Mercader (ed.) *The archaeology of tropical rain forests*, 119–160. New Brunswick: Rutgers University Press.
- . 2004. An integrated perspective on Orang Asli ethnogenesis. In Victor Paz (ed.) *Southeast Asian archaeology: Wilhelm G. Solheim II Festschrift*, 366–399. Quezon City: The University of the Philippines Press.
- . this volume, Biological and cultural evolution in the population and culture history of Homo sapiens in Malaya.
- Burenhult, Niclas. (forthcoming). Foraging and the history of languages in the Malay Peninsula. In Tom Güldemann, Patrick McConvell & Richard Rhodes (eds) *Historical linguistics and hunter-gatherer populations in global perspective*.



- Burenhult, Niclas & Claudia Wegener. 2009. Preliminary notes on the phonology, orthography and vocabulary of Semnam (Austroasiatic, Malay Peninsula). *Journal of the Southeast Asian Linguistics Society* 1. 283–312.
- Dallos, Csilla. 2011. *From equality to inequality: social change among newly sedentary Lanoh hunter-gatherer traders of Peninsular Malaysia*. Toronto: University of Toronto Press.
- Diffloth, Gérard. 1975. Les langues mon-khmer de Malaisie: classification historique et innovations. *Asie du Sud-Est et Monde Insulinde* 6(4). 1–19.
- Dunn, Michael, Niclas Burenhult, Nicole Kruspe, Neele Becker & Sylvia Tufvesson. Aslian linguistic prehistory: a case study in computational phylogenetics. *Diachronica*.
- Endicott, Kirk M. 1975. A brief report on the Semaq Beri of Pahang. *Federated Museums Journal (new series)* 20. 1–23.
- . 1979. *Batek Negrito religion: the world-view and rituals of a hunting and gathering people of Peninsular Malaysia*. Oxford: Clarendon.
- . 1997. Batek history, inter-ethnic relations and subgroup dynamics. In R. Winzeler (ed.) *Indigenous peoples and the state: politics, land, and ethnicity in the Malayan Peninsula and Borneo*, 46. 30–50. New Haven: Yale Southeast Asia Studies Monograph.
- Evans, Ivor H. N. 1915. Notes on Some Aboriginal Tribes of Pahang. *Journal of the Federated Malay States Museum* 5(4). 192–219.
- Fix, Alan G. 1995. Malayan paleosociology: implications for patterns of genetic variation among the Orang Asli. *American Anthropologist* 97(2). 313–323.
- . 2002. Foragers, farmers, and traders in the Malayan Peninsula: origins of cultural and biological diversity. In Kathleen D. Morrison & Laura L. Junker (eds) *Forager-traders in South and Southeast Asia*, 185–202. Cambridge: Cambridge University Press.
- . this volume, Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model.
- Fucharoen, Goonapa, Supan Fucharoen & Satoshi Horai. 2001. Mitochondrial DNA polymorphisms in Thailand. *Journal of Human Genetics* 46(3). 115–125.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, William Meehan, James Blackburn, Douglas Clarke, Joseph Maripa Raja, Patimah Ismail, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2006. Phylogeography and ethnogenesis of Aboriginal Southeast Asians. *Molecular Biological Evolution* 23(12). 2480–2491.
- Howell, Signe. 1989. *Society and Cosmos: Chewong of Peninsular Malaysia*. Chicago and London: University of Chicago Press. Original publ. 1984, Oxford: Oxford University Press.
- Kruspe, Nicole. 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press
- . (in prep.). *A preliminary report on Semaq Beri*.
- . (in progress). *A grammar of Ceq Wong*.
- Kuchikura, Yukio. 1987. Subsistence Ecology among Semaq Beri Hunter-gatherers of Peninsula Malaysia. Hokkaido. *Behavioral Science Report Series E*, No. 1. Sapporo: Hokkaido University.
- Lye, Tuck-Po. 1997. *Knowledge, forest, and hunter-gatherer movement: the Batek of Pahang, Malaysia*. Unpublished doctoral dissertation. University of Hawai'i: Department of Anthropology.

- Morris, Kimball. 1996. Forest Products for Market: Rattan and the Semaq Beri of Ulu Tembeling. *Akademika* 48. 93-112.
- Ogilvie, Charles S. 1948. More of the Che Wông. *Malayan Nature Journal* 3(1). 15–29.
- Oppenheimer, Stephen, this volume, MtDNA Variation and Southward Holocene Human Dispersals within Mainland Southeast Asia.
- Rambo, A. Terry. 1988. Why are the Semang? Ecology and ethnogenesis of aboriginal groups in Peninsular Malaysia. In A. Terry Rambo, Kathleen Gillogly & Karl L. Hutterer (eds) *Ethnic diversity and the control of natural resources in Southeast Asia*, 19–35. Ann Arbor: University of Michigan Press.
- Ramle, Abdullah. 1993. *Semaq Beri, Komuniti Orang Asli di Terengganu*. Kuala Terengganu: Hakcipta Kolej Agama Sultan Zainal Abidin.
- Schebesta, Paul. 1928. *Among the forest dwarfs of Malaya*. Kuala Lumpur: Oxford University Press.
- . 1952. *Die Negrito Asiens: Geschichte, Geographie, Umwelt, Demographie und Antropologie der Negrito*. Vienna: St. Gabriel Verlag.
- Sluys, Cornelia M.I. van der. 2000. Gifts from the immortal ancestors: cosmology and ideology of Jahai sharing. In Peter P. Schweitzer, Megan Biesele & Robert K. Hitchcock (eds) *Hunters and gatherers in the modern world: conflict, resistance and self-determination*, 427–454. New York: Berghahn.



# 12 *Origin of genetic diversity among Malaysian Orang Asli: An alternative to the demic diffusion model*

---

ALAN FIX

## 1 Introduction

Demic diffusion by agriculturalists has been postulated to be the primary mechanism for the expansion and dispersal of peoples and languages (Cavalli-Sforza et al. 1993). Originally proposed as the explanation for the spread of agriculture from its origin in the Near East through Europe (Ammerman & Cavalli-Sforza 1971), demic diffusion depends on a very high rate of population growth and much higher carrying capacity of agriculturalists as compared to foraging people. Rapidly growing farmer populations need new land and are driven to expand into the ranges of the thinly distributed foragers. In the original formulation (for example Menozzi et al. 1978), the farming expansion led to admixture with resident foragers and a clinal gradation in gene frequencies along the direction of spread. Such clines, then, serve to mark the demic wave of expansion.

Recently the notion of ‘farming dispersals’ (Bellwood & Renfrew 2002) has been expanded to account for language dispersals and the biological effects have been less emphasised. Indeed, the hypothesis has been so generalised as to be almost necessarily true. Thus Renfrew (Bellwood & Renfrew 2002:3) states: ‘some language families’ were ‘at least in part’ dispersed through ‘demographic and cultural processes ... which accompanied the dispersal ... of the practice of food production’. In this guise, the concept has been extended to many regions of the world including Africa and Asia (Bellwood & Renfrew 2002). On this hypothesis for Asia, the origin of rice agriculture along the Yangtze River in China caused farmer populations to increase, creating an expanding wave of people that replaced indigenous foraging populations across the continent.

At the extreme tip of Southeast Asia, the Malaysian Peninsula was occupied by indigenous (Orang Asli) farming populations (Senoi, Melayu Asli) but unlike most of the rest of Asia, a number of foraging groups as well (Semang). On the farming/demic diffusion hypothesis, the Semang foragers are seen as survivors of the original hunter-gatherers of Southeast Asia; Senoi and Melayu Asli populations represent the later intrusion of rice farmers (Bellwood 1993; Higham 2002). Here at least in the Peninsula the expanding wave of farmers did not entirely replace the foragers.

A recent study of mitochondrial DNA in Malaysian Orang Asli seemingly supports the farmer dispersal model. Macaulay et al. (2005) argue that the Orang Asli represent an

isolated 'relict' population of the original dispersal of modern *Homo sapiens* from Africa some 65,000 years ago. However, half their Senoi sample comprises haplogroups common in mainland Southeast Asian groups such as the Vietnamese suggesting that an important component of Orang Asli populations derived from migrants bringing agriculture to the Peninsula. Also similar to mainland Asian farming populations, the frequency of hemoglobin E, a malarial protective allele, is high in the Senoi (Fix & Lie-Injo 1975).

However, as will be demonstrated in this chapter, demic diffusion is not the only process that can account for the current biological diversity of humans in Asia. The alternative model proposed here, 'trickle effect' colonisation, can explain the data equally well while having the advantage of being more compatible with demographic and cultural patterns in early farming and foraging groups.

The farming dispersal model depends on a very high farmer population growth rate (3% per annum in the original formulation) and a very great difference in carrying capacity between farmers and foragers (Rendine et al. 1986). This extreme differential seems highly unlikely to have existed in the context of the initial development and spread of farming. Surely intensive farming systems such as Asian padi rice agriculture support very large populations, but extensive land use swidden farming more likely to be representative of first farmers show much lower growth rates and densities (see Fix 1999 for examples). More commonly among present day small-scale populations, growth rates are much lower and gene flow is a 'trickle' of mate exchanges (Fix 1999).

The continuing presence of both a foraging economy among the Semang and the high frequency of 'ancient' mtDNA haplogroups among all groups of Orang Asli (including farming Senoi and Melayu Asli) also support a network model of gene flow among farmers and foragers rather than the demic wave of advance.

This chapter presents a computer simulation model incorporating migration into the Peninsula of a small group of initial farmers who maintained gene flow with their parental population and with the already resident foraging groups. The farmers introduced new haplogroups as well as hemoglobin E to the Peninsula. Based on reasonable parameter values, the simulation shows that within some 20 generations the current genetic distribution could have evolved through intermarriage of farmers with foragers as well as natural selection for hemoglobin E as farming increased the prevalence of malaria.

## 2 Malayan Prehistory

The indigenous peoples of the Malaysian Peninsula (Orang Asli) are traditionally divided into three main groups: Semang (or 'Negritos'), Senoi, and Melayu Asli ('proto' or 'aboriginal' Malay). Figure 1 and Table 1 provide the location and a capsule summary of the major differences (excluding the biological) characterising each group.

The traditional explanation for the presence of three distinct cultural patterns coexisting on the Peninsula was migrational. That is, each group originated elsewhere and migrated at different times to Malaya, each subsequent wave displacing some of the previous inhabitants and each being deposited like the layers of a cake (see Carey 1976 for a summary of this literature). The foraging Semang (together with other 'Oceanic Negritos' such as the Andaman Islanders, and various Philippine groups like the Aeta and Agta), represent descendants of a previously widely distributed population now mostly replaced by later migrants. The second wave included the swidden-farming Senoi peoples followed later by the Melayu Asli.

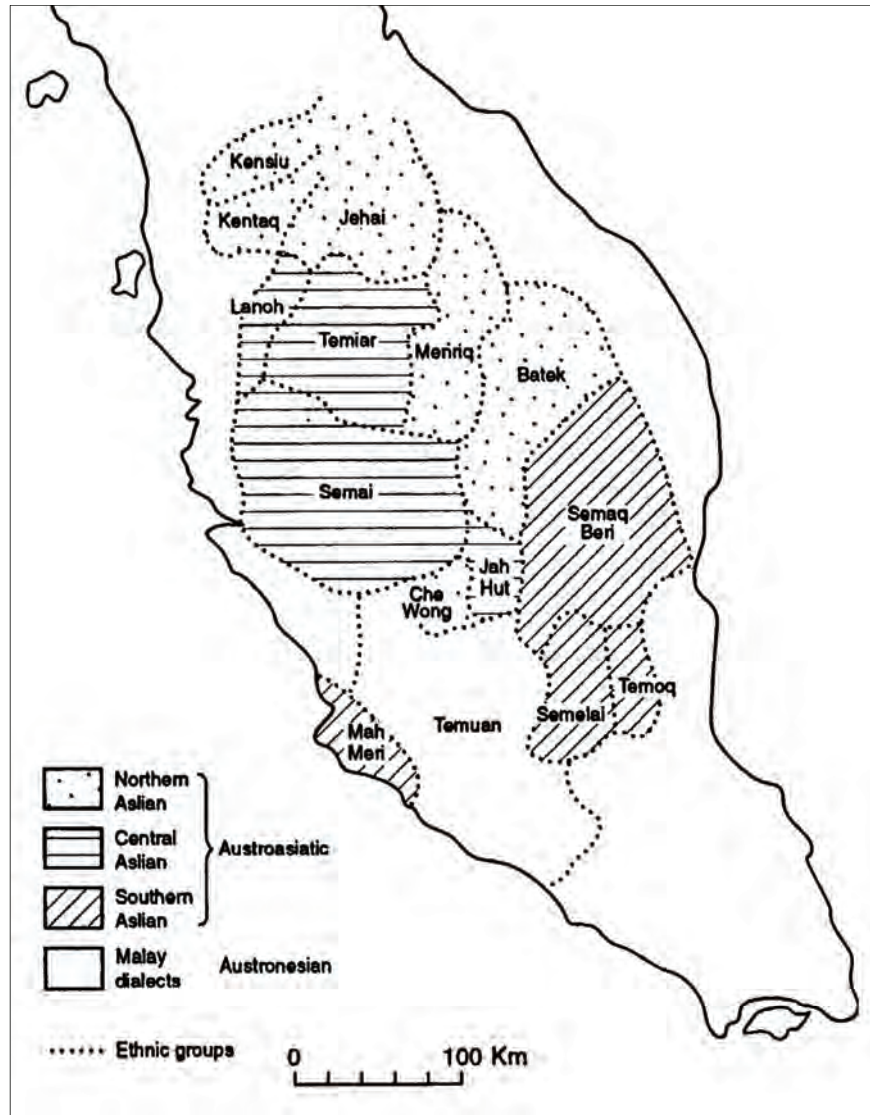


Figure 1: Distribution of Orang Asli groups and languages in the Malayan Peninsula.

Table 1: The Three Orang Asli Traditions

Tradition	Language	Technology/Economy
Semang	Northern Aslian	nomadic foragers
Senoi	Central Aslian	sedentary swiddeners
Melayu Asli	Southern Aslian	sedentary farmers, traders

As an alternative to the wave theory of successive colonisation of the Peninsula, Geoffrey Benjamin (1976, 1980, 1985, 1986) has proposed an *in situ* model for the origin and diversification of Orang Asli cultures. The three traditions differentiated from a common linguistic and cultural matrix within the last few thousand years as the northern Aslian groups continued (and became more committed to) foraging, central groups adopted swidden farming and became increasingly sedentary, and southern groups became

involved in trade relations with outsider groups (probably Austronesian-speaking peoples of the islands).

Clearly, these polar hypotheses do not exhaust the possibilities for explaining present-day cultural diversity. Peter Bellwood (1993), for instance, has proposed one such intermediate position maintaining that both regional continuity and successive flows of people into the Peninsula have occurred. He sees the population increases that followed from adopting agriculture as the cause for the spread of farmers to new regions (Bellwood 1996). This argument extends the demic diffusion model of Cavalli-Sforza and colleagues (1993) originally applied to the spread of the European Neolithic through population expansion from the Near East. This idea (also called the 'Farming/language dispersal hypothesis' [Bellwood & Renfrew 2002]) has come to be widely invoked to explain population dispersals including the major language families of Asia and specifically Austroasiatic, of which Aslian, the languages of the Orang Asli, is a member (Higham 2002).<sup>1</sup>

### 3 Evaluating Scenarios

How might these different scenarios for Orang Asli prehistory be tested? Each is consistent with the broad outlines of the situation; each depends on different mechanisms to produce current cultural, linguistic, and genetic similarities and differences among the three groups. The most direct test (that is, one based on evidence from the past rather than inference from current distributions) would be archaeological—are there clear indications of population replacements and/or new influxes of peoples arriving in the Peninsula bearing diagnostic cultural markers of the traditions? However the evidence for such a migration is not unequivocal. Similarly, skeletal remains from archaeological sites with diagnostic morphological and/or metrical traits might provide direct evidence of migration. Again, such data as exist seem unable to resolve the issue. David Bulbeck (this volume) at least is prepared to say that there is no evidence from dental metrics or osteology to substantiate a Neolithic population expansion from China across Southeast Asia (see also Bulbeck & Lauer 2006; Pietrusewsky 2006).

Turning to inference from present-day distributions, linguistics has been used to substantiate both the *in situ* and the demic diffusion models. Benjamin's (1976) reconstruction of Malayan culture history was based on a comparative analysis of Aslian languages. The pattern of diversification of northern Aslian (mostly Semang) suggested a split between their ancestors and those of central Aslian speakers (Senoi) some 5000 years ago. This analysis links Semang and Senoi ancestry. Bellwood (1993), on the other hand, interprets the linguistic evidence to support population movement. He sees ancestral Senoi-speaking Austroasiatic languages arriving from the north into the Peninsula beginning around 3000 BP as a result of expanding populations of rice farmers. Already resident ancestral Semang, speaking some now unknown languages, through a process of interaction with the migrant farmers, adopted their language. Meanwhile, similar pressures were driving expanding populations of Austronesian speakers, who somewhat later entered the Peninsula from the south as the ancestors of the Melayu Asli. Since several of the southern Melayu Asli now speak dialects of Malay, the Benjamin model requires a language switch on the part of these peoples as they assimilated to their maritime trading partners from the islands. Much more could be said about both these models' congruence

---

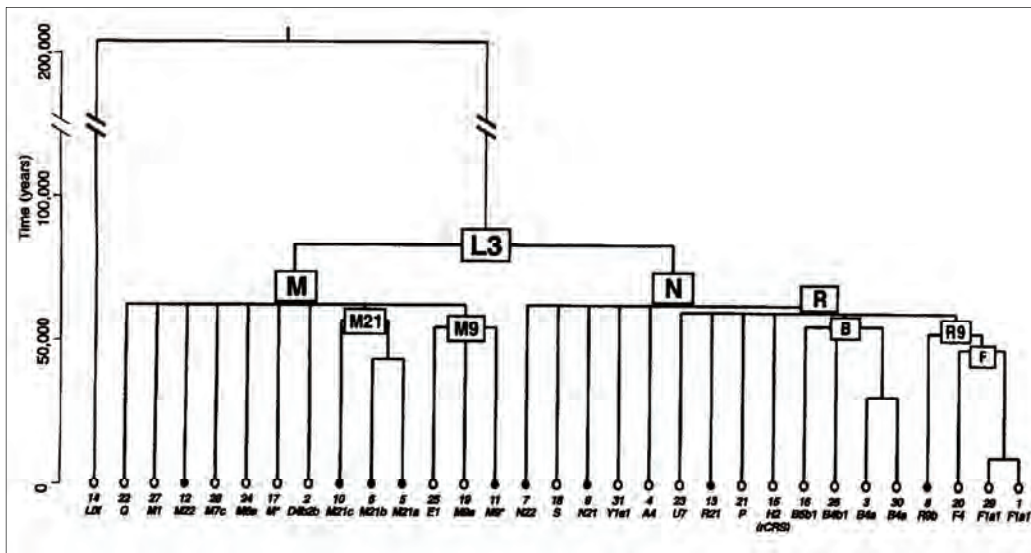
1 The implications of these models for biological diversification have been explored in Fix (1995, 2002, see also 1999).

with linguistic patterns but it should be obvious that a definitive test has not been achieved. Burenhult et al. (this volume) discuss some of the complexities of Orang Asli linguistics particularly addressing the Northern Aslian/Semang relationship (see also Burenhult, forthcoming).

Similarly ‘classic genetic marker’ data do not provide a clear picture of the history of the Orang Asli populations. Cavalli-Sforza et al. (1994) produced a dendrogram based on 31 such ‘markers’ that showed the Semai (Senoi) clustering with the Zhuang, a Tai speaking South Chinese population and quite distant from Khmer who are linguistically and geographically closer to the Semai.

An alternative tree (Saha et al. 1995) based on a different set of markers showed a very different result. Now Semai do group with the Khmer and are distant from other populations. It should be noted that Saha and colleagues included hemoglobin E (Hb\*E), one of the many structural variants of hemoglobin, similar to Hb\*S (sickle cell), as one ‘marker’. Both the Semai and Khmer possess high frequencies of Hb\*E almost surely due to malarial selection in both populations, about which more below.

On the other hand, the scenario presented by a recent mtDNA study includes both an indigenous origin of the Orang Asli in the Peninsula but augmented by significant influxes from outside Malaysia (Macaulay et al. 2005; Hill et al. 2007; Oppenheimer, this volume).



**Figure 2:** Dendrogram of Asian mtDNA haplotypes. Haplogroup L3 is the African clade from which the two non-African clades, M and N, arose. Filled circles show Malaysian haplogroups (from Maccaulay et al. 2005).

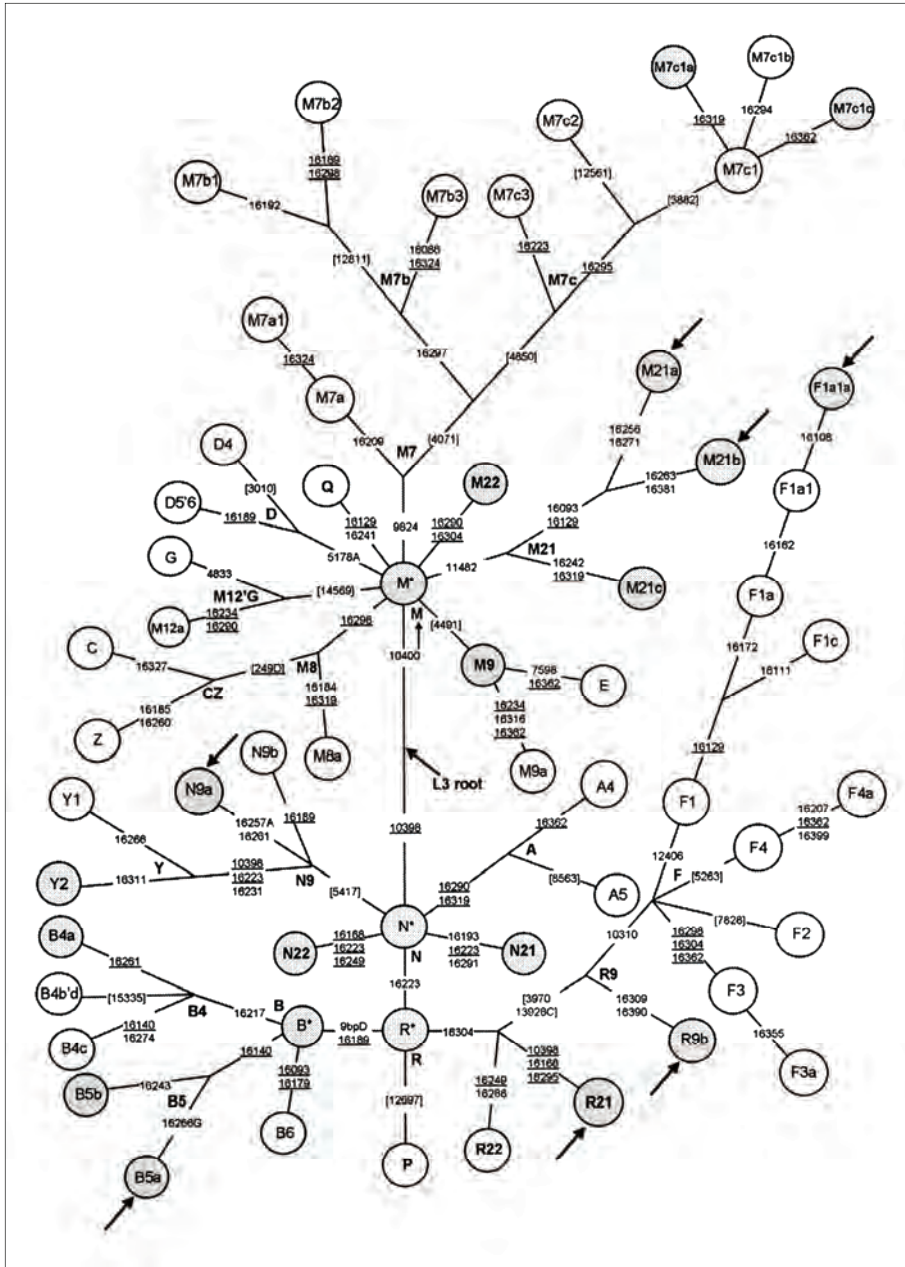
Figure 2 shows a tree of mtDNA divergence from the ancestral L3 ‘Out-of-Africa’ clade. Some of the earliest branches from the root within both M and N superhaplogroups are found in Orang Asli (the filled circles in Fig. 2). Figure 3 shows the mtDNA haplogroup network (with Orang Asli haplogroups marked with an arrow).

The basal lineages are dated to c. 60 kya; these Orang Asli haplogroups seem to branch directly from the founding Asian population. Both haplogroups R21 and M21 are found in the Semang and Senoi (with other ‘ancient’ roots in Melayu Asli). However, half of their sample of Senoi (50 Temiar and one Semai) are F1a1a, a widespread Southeast Asian clade absent from Semang and most Melayu Asli but common in Thailand and Vietnam (where it has a frequency of 20%; Hill et al. 2006). This haplogroup is dated at 9000 years



ago suggesting a dispersal from mainland sources to the Peninsula sometime after this date. This scenario would be consistent with an agricultural spread (a la Bellwood 1993) but would not exclude an earlier Holocene movement of an equally widespread non-agricultural ‘culture’, the Hoabinhian.

While Macaulay and colleagues (2005) emphasise the isolated, ‘relictual’ status of the Orang Asli, the conclusions of Hill et al. (2007) suggest a more complicated history. They state that the traditional ‘layer cake’ theory now seems ‘completely unfounded’. All three Orang Asli groups have local roots that reach back to ~50,000 years ago, and all have been affected to a greater or lesser extent by subsequent migrations to the Peninsula (p. 19).



**Figure 3:** Network of Asian mtDNA haplotypes. The L3 African root is in the center. M and N haplogroups have diverged by mutation into numerous additional haplogroups (R, F et al.). Haplogroups present in Malayan Orang Asli are indicated by arrows (from Hill et al. 2006).

## 4 Simulation Model

In this chapter I present a model that replicates these data in terms of the evolutionary processes of colonisation, migration, and also natural selection to show that the pattern of genetic diversity in present-day Orang Asli reflects neither long term isolation nor massive waves of farming migrants. Instead a small colonising agricultural group maintaining gene flow with their parental population and with the already resident foraging groups would over time produce the mix of mtDNA types and Hb\*E frequencies characterising the current populations.

The model simulates the population history of Peninsular Malaya beginning with residents descended from the initial colonisation of coastal Southeast Asia from Africa (following Macaulay et al. 2005; Hill et al. 2006) into whose range a small group of farmers enters bringing both a new technology but also transforming the environment, thereby encouraging malarial transmission. The farmer/forager demographic differential usually assumed in demic diffusion models (for example Cavalli-Sforza et al. 1993) is not the driving force in this model and therefore it represents an alternative to the demic diffusion/agricultural dispersal hypothesis. The subsequent history of this immigrant population involves both gene flow with resident foragers and with the parental farming groups from which they derived. This group represents the ancestral Senoi farmers of the Peninsula and its entrance initiated the differentiation of Semang/Senoi traditions. These immigrants (derived from northern mainland Southeast Asian agricultural groups) would presumably also have been descendants of the original coastal migrants now after several thousand years having differentiated from the southern Malayan populations. Farming introduced selection pressure for Hb\*E and the immigrants would be expected to carry this allele as well as mtDNA haplogroups such as F1a1a characteristic of mainland populations (Hill et al. 2007). The simulation proceeds with foragers entering the farmer population at a low rate each generation as well as a small amount of continuing gene flow from the parental farming populations. Selection for Hb\*E continued in the farming population throughout the duration of the runs.

The simulation initial conditions, then, included a resident hunter-gatherer population lacking the Hb\*E allele and characterised by deep-rooted mitochondrial DNA clades (haplogroup M). A small group of farmers ( $N_e = 50$ ) migrating from established agricultural regions to the north of the Peninsula entered the territory of the hunter-gatherers. The date of entry of agriculture into Malaya is thought to be circa 3000 years ago (Bellwood 1993). The mutation to Hb\*E has been dated to around 4000 years ago (Ohashi et al. 2004); however, it is assumed that the migrant farmers already would have been subject to malarial selection for many generations and the frequency of Hb\*E in their population would be relatively high ( $q_e = 0.3$ ).

Although in different runs a variety of parameter values were employed, the following table shows a representative set.

**Table 2:** Simulated population parameters

	<b>Population Effective Size (<math>N_e</math>)</b>	<b>Hb-E frequency</b>	<b>mtDNA haplogroup</b>
Migrant Farmers	50*	0.3	F
Resident Foragers	1000*	0.0	M

\* The effective size of the mtDNA gene pool comprises females only and is therefore  $\frac{1}{2}$  the total size; i.e., 25 farmer women and 500 forager women. Census size is usually three or more times effective size.

The sequence of operations each generation in the simulation included:

- 1) Selection at the hemoglobin locus according to the fitnesses  $W_{AA} = 1.10$ ,  $W_{AE} = 1.38$ , and  $W_{EE} = 0.73$ . These fitness values were adjusted to allow a low rate of population growth,  $r$ , of the farmer population (average  $r$  over 20 generations was 0.01; for the full 30 generations of the runs, the average  $r$  was 0.007) and were damped as the overall population grew according to  $W_i = W_i - (0.2(N/N_{\max}))$ , where  $W_i =$  fitnesses  $i=1$  (AA); 2 (AE); 3 (EE);  $N =$  current population size, and  $N_{\max} =$  the maximum size of the farmer population (approximately 'carrying capacity').
- 2) Migration. A proportion of the mates of the farmer population was derived both from neighbouring forager populations and from continuing connections with the parental farming population. Although a variety of rates were used in different runs of the simulation program, the results presented below were based on a migration rate of ten per cent per generation from resident foragers and ten per cent migration from the parental farming population. As for selection, the migration rate was also damped as the population grew through time beginning with ten per cent from each community (forager/parental farmer) and damping to one per cent as the Malayan farmer population reached maximum size ( $N_e$  female =4000).

It should be noted that these experiments do not include simulated genetic drift, which in these small populations would increase the variation among different runs. This omission makes the model less realistic, however, since the goal is to test whether a migration/selection process can replicate the current genetic profile of the Orang Asli, the lack of stochastic variability among runs is not relevant. In the case of mtDNA, which is presumed to be neutral with respect to survival (no selection), the mix of haplogroups will be a function of the percentages of forager and farmer individuals in the growing farmer population. Since Hb\*E offers protection against malaria, natural selection should determine the frequency of this allele in the evolving population. Although genetic drift (and mutation in the case of mtDNA) could affect the frequency of haplotypes and alleles, migration and selection are the determinants in this model.

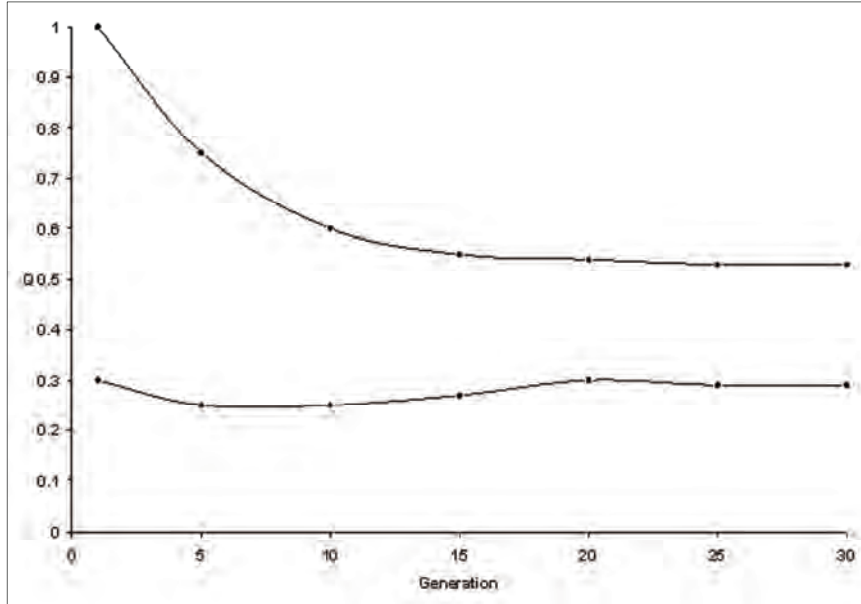
For this reason, results for a wide range of parameter values will not be presented but rather those that are based on the best expectations for the demography/population structure of the simulated populations. The question is, whether it is possible on the model to reproduce Orang Asli genetic patterns using reasonable parameter values. Although opinions may vary on what constitutes 'reasonable', the values used in this simulation model are based on the anthropological genetics of well-studied small-scale human populations (Fix 1999). For instance, a migration rate of 20 per cent per generation fits with an extensive study of 21 ethnographically known societies (Adams & Kasakoff 1976) which discovered that these groups all showed an approximately 80 per cent endogamy rate even though the size of the endogamous unit varied across technologies.

Other parameters will be further discussed following the presentation of the results.

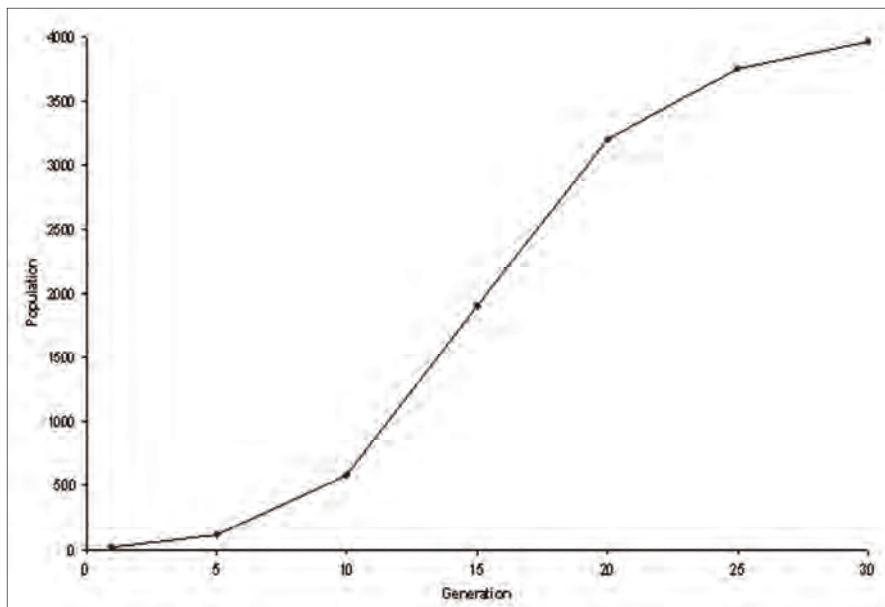
## 5 Results

Figure 4 shows the results of 30 generations of the simulation (although note that 20 generations were sufficient to establish the frequencies that continue to characterise the modern population).  $Q$  refers to the frequencies of the F haplogroup (initially 1.0 in the pioneer migrant group) and the gene frequency for Hb\*E (initial  $q_E = 0.3$ ) in the simulated farmer ('Senoi') population. Haplogroup F shows a steady decline as intermarriage with

resident foragers (in whose population  $F = 0.0$ ;  $M = 1.0$ ) introduces  $M$  haplotypes into the population. Since the farmers introduced not only agriculture but also the conditions fostering malarial endemicity (sedentary village populations opening up the forest for swiddens; Livingstone 1958), selection for  $Hb^*E$  maintains this allele at near equilibrium (0.3) in spite of a steady influx of forager mates.

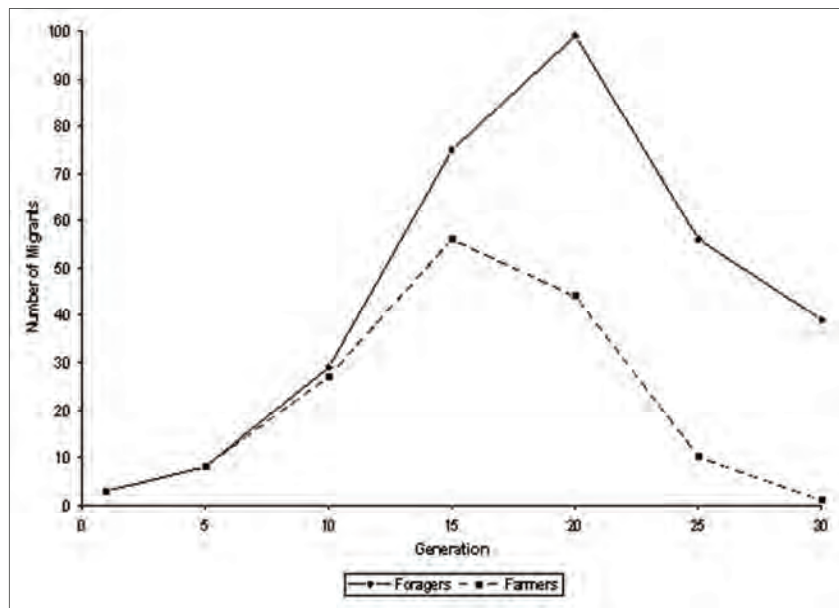


**Figure 4:** Results of simulation. Upper line represents the decline in frequency ( $Q$ ) of the  $F$  haplogroup in the original colonising population to a frequency of 0.5 after 30 generations. Lower line is the frequency of  $Hb^*E$  over the same period of 30 simulated generations.



**Figure 5:** Increase of farmer population over 30 generations.

Figure 5 shows the increase in the female farmer population over the 30 generations of the simulation. Beginning with 25 in the initial colonising group, the population grew fairly rapidly ( $r = 0.01$ ) for 20 generations, then began to level off to achieve equilibrium size ( $N_e$  female = 4000 implying a total census population size of around 12,000).



**Figure 6:** Numbers of migrants (forager and farmer) into farmer population.

Figure 6 presents the contribution of migrants both from the resident forager population (shown as a solid line) as well as continued contact and migration from the parental farming population (dashed line). Again, the number of migrants is initially low but as the population grows, the number increases to a maximum of 143 in-migrants at generation 20. As for overall population growth, in-migration diminishes as the total population approaches equilibrium size. Although the number of migrants increased through time, the rate of migration did not exceed 20 per cent per generation.

## 6 Discussion and Implications

The demonstration that a small colonising farmer group connected by a network of gene flow among small-scale farmer and forager populations can account for the current diversity of mtDNA haplogroups among Malayan Orang Asli has wider implications for understanding human population history. That the model fits reality, however, is not a verification of its historical accuracy; that test depends on evaluating the model on other, hopefully independent, criteria. For the same reason, arguments invoking large-scale population movements or demic diffusion cannot be validated by present-day genetic distributions alone. The demic diffusion hypothesis provides a plausible mechanism for the spread of genes and languages but with equally plausible alternative explanations available, it cannot be accepted without additional evidence. Thus the fact that Senoi possess both a 'farmer' mtDNA haplogroup (F1a1a) and hemoglobin allele (Hb\*E) implies some form of population relationship but not the exact form of that relationship. Indeed, Hb\*E may have involved the spread of the gene and of the technology of farming, however the frequency in modern Senoi depended more on natural selection for malarial protection. Thus the evolutionary forces determining the distribution of Hb\*E in Southeast

Asia include both gene flow and selection, the conditions for selection being produced by agriculture.

Ideally archaeology and linguistics should be able to provide the independent historical evidence distinguishing between large-scale (demic diffusion) and small-scale 'trickle' gene flow. Unfortunately, as already pointed out, data from these domains are not unequivocal (see previously cited references).

In the absence of a direct test, a closer look at the assumptions of the models might help resolve the issue. The demic diffusion model as developed by Ammerman and Cavalli-Sforza (1971) was based on the assumption that agriculture allowed for a much greater population density than foraging. This idea goes back to V.G. Childe (1958) who saw agriculture as a revolutionary innovation in human history, one that allowed increased population growth rates and ultimately the increased population density associated with civilisations. Rapid population growth led to pressure on land resources and the need to colonise new lands. Ammerman and Cavalli-Sforza (1971) postulated that such colonisation (or invasion) would penetrate into territory occupied by foragers, whose less intensive land use supported lower population density and thus 'empty' land suitable for exploitation by farmers. Over time, sparse foraging populations would be assimilated by burgeoning farmer groups and agriculture would predominate (as for Europe).

The key dynamic for demic diffusion is the presumed great differential in population growth between farmers and foragers. This presumption, however, depends on an implicit ethnographic analogy; modern farmers and surviving foragers. For Cavalli-Sforza, a very salient model of foraging societies was the Aka Pygmies of central Africa, a group with whom he and his colleagues did research (Cavalli-Sforza et al. 1986). It is surely true that more intensive land use economies have greater population density potential than less intensive economies (Fix 1999). Hunter-gatherers that survived into the ethnographic present mostly are thinly spread on the land (with some very notable exceptions such as the American Northwest Coast). The question of interest, however, is how do **first farmers** compare with contemporary foragers in their demographic profiles since that is the time period relevant to the hypotheses. To compare the intensive farmers of highland New Guinea or modern Asian rice farmers with the San or Aka accentuates the contrast. A more likely demography for early farmers might be that of less intensive hill farmers such as the Semai Senoi populations in Peninsular Malaysia where the average annual rate of population growth ( $r$ ) is approximately 0.007 (Fix 1977). This low rate of increase takes on significance when compared to the only concrete parameterisation of the demic diffusion model, that of Rendine et al. (1986) in which a simulation model for the spread of agriculture in Europe was presented. This formulation depended on a nearly 30-fold increase in population size between foragers and farmers (foragers,  $N=300$ ; farmers,  $N=8000$ ). Population growth rates for farmers were very high, an annual rate of increase of 2.7 per cent (0.027). While it is true that some modern human populations have grown at a three per cent per annum rate, it seems more likely that the initial stages of farming might have been more similar to the Semai rate. If so, then rather than an overwhelming wave of advance, earliest farmers more likely followed the 'trickle' scenario, small colonising groups intermarrying with not-so-very-much smaller foraging groups into whose ranges they entered. This, of course, is the model presented in this chapter.

This is not to say that **migration** was not an important factor in accounting for Orang Asli origins; only that demic diffusion, a particular migratory mechanism, need not have been involved. According to the 'trickle' model presented here, a small group of migrants introduced agriculture and their genes to the Peninsula. As Figure 5 shows, this population grew considerably over 30 generations (approximately 600 to 750 years depending on the

generation length of the population). That growth comprised both natural increase (overall  $r = 0.007$ ) and a small influx of both forager and other farmer migrants (see Figure 6). As a result, a substantial portion of the Senoi mtDNA haplogroups were derived from in-marrying foragers. Note also that the differential population density of foragers and farmers in the model exists but is 8- rather than 30-fold.

The trickle model is also consistent with the reconstruction of Orang Asli cultural traditions by Benjamin (1986; see also Fix 1995). Benjamin's model posits the forager/farmer (Semang/Senoi) traditions arising by a sort of cultural 'character displacement' (see Brown & Wilson 1956 for the biological model). The basic idea is that species (or by extension, cultures) become more different from one another when they are potentially competing for the same niche. To avoid such competition, each 'displaces' against the other, partitioning the niche into separate components and specialising in that pattern. Thus, in the Orang Asli case, each tradition adopted social and cultural patterns to emphasise differences with the other (for example, Semang foragers become hyper-nomadic stressing dispersal versus the more sedentary Senoi). This view suggests greater population equality between initial migrant farmers and resident foragers with mutual social and cultural accommodations and adjustments to allow coexistence of both groups. Rather than being overwhelmed by waves of invading farmers, foragers would have solidified and maintained their own tradition as well as contributing migrants to the alternative life-way of the farmer population.

While the shift to an Austroasiatic language in the Peninsula might seem to support a demic expansion model of numerically dominant intrusive farmers overwhelming indigenous foragers, the intermarriage of foragers and farmers on the trickle effect model could also account for the language change. Burenhult (forthcoming) provides a cogent Malaysian example: currently speakers of several dialects of Lanoh, one of the Semang foraging groups, after extensive intermarriage with Temiar (Senoi farmers) speakers, have adopted the Temiar language. Burenhult (forthcoming) explicitly states: 'It [the Lanoh case] is also a possible scenario for the early introduction of Austroasiatic by immigrants who intermarried with locals.' (p 15)

Although it does not bear on the issue of demic diffusion versus trickle effect, it might be noted that the model comprises basically one-way gene flow from the resident foragers to the farmers. A common forager/farmer marital pattern involves forager women marrying into farmer populations but not vice-versa (see Wood et al. 2005 for a case from Africa).

This would account for the mixture of mtDNA haplogroups among the Senoi and the predominance of 'ancient' M haplogroups among the Semang foragers. It would be extremely interesting to study the Y chromosome frequencies among these populations since mtDNA traces only maternal ancestry. Since the model depends on female foragers marrying with male farmers, males of both groups should remain in their natal groups. As a consequence, male-specific Y chromosome frequencies should be more homogeneous within each population and would be expected to retain ancestral frequencies.

It should also be noted that the model proposed here refers only to the first entrance of farmers into the Peninsula. Surely continuing networks of gene flow with mainland Southeast Asian populations were maintained and provided avenues for further genetic additions to the Peninsular gene pool (see Oppenheimer, this volume). Later contacts with sea-faring Austronesian-speakers would also have contributed genes to the Orang Asli. A likely example is Southeast Asian Ovalocytosis, a malarial-protective allele widely distributed from coastal New Guinea through parts of Indonesia and in some Orang Asli populations (Fix 1995).

## 7 Conclusions

The model presented here involving a small colonising farmer group augmented by continuing gene flow from the parental population and indigenous foragers—the ‘trickle effect’,—shows that waves of migration, whether the traditional ‘layer-cake’ sequence or a demic expansion of agriculturalists, are not required to explain the origins of genetic diversity among the Malaysian Orang Asli. Although this demonstration does not disprove the demic diffusion hypothesis, demographic and cultural considerations of early founding farmers and foragers support the network of gene flow of the trickle effect model. That is, the extreme differentials in growth rates and carrying capacities required by the demic diffusion model are unlikely to have characterised the **earliest** farmers and foragers contemporary with them. The trickle effect model avoids this problem, depending on demographic rates similar to less intensive modern farming populations and recent foragers inhabiting favourable habitats. At least for the Malayan Peninsula, the trickle effect model also fits a plausible reconstruction of the dynamics of cultural differentiation among the Orang Asli.

As Sewall Wright (1931) pointed out long ago, evolution is a shifting balance among the several evolutionary forces. Migration and gene flow, genetic drift, and natural selection (as in the case of hemoglobin E considered here) may all play a role in shaping diversity. Macro-models that focus on one process such as demic diffusion may obscure the actual dynamics of human biological diversity in Southeast Asia.

## Acknowledgements

Thanks to Nick Enfield and Joyce White for organising the January 2009 workshop on Dynamics of Human Diversity in Siem Reap and to the other participants for much stimulating discussion. The Research Committee of the Academic Senate, University of California, Riverside provided some travel funds to attend the conference.

## References

- Adams, John.W. & Alice B. Kasakoff. 1976. Factors underlying endogamous group size. In Carol A. Smith (ed.) *Regional Analysis, Vol. 2, Social Systems*, 149–173. New York: Academic Press.
- Ammerman, Albert J. & Luigi L. Cavalli-Sforza. 1971. Measuring the rate of spread of early farming in Europe. *Man* 6. 674–688.
- Bellwood, Peter. 1993. Cultural and biological differentiation in Peninsular Malaysia: The last 10,000 years. *Asian Perspectives* 32. 37–60.
- . 1996. The origins and spread of agriculture in the Indo-Pacific region: Gradualism and diffusion or revolution and colonization? In David R. Harris (ed.) *The origins and spread of agriculture and pastoralism in Eurasia*, 465–498. Washington: Smithsonian Institution Press.
- Bellwood, Peter & Colin Renfrew. 2002. Examining the farming/language dispersal hypothesis. *McDonald Institute Monographs*. Oxford: Oxbow Books.
- Benjamin, Geoffrey. 1976. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds) *Austroasiatic Studies*, 37–128. Honolulu: University of Hawaii Press.



- 1980. *Semang, Senoi, Malay: Culture-history, kinship and consciousness in the Malay Peninsula*. Unpublished manuscript .
- 1985. In the long term: Three themes in Malayan cultural ecology. In Karl L. Hutterer & Terry A. Rambo (eds) *Cultural Values and Tropical Ecology in Southeast Asia*, 219–278. Ann Arbor, MI: Center for South and Southeast Asian Studies.
- 1986. *Between isthmus and islands: Reflections on Malayan palaeo-sociology*. Working Paper No. 71, Sociology Department, University of Singapore.
- Brown, William L. & Edward O. Wilson. 1956. Character displacement. *Systematic Zoology* 5. 49–64.
- Bulbeck, David, this volume, Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya.
- Bulbeck, David. & Adam Lauer. 2006. Human variation and evolution in Holocene Peninsular Malaysia. In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 33–171. Cambridge: Cambridge University Press.
- Burenhult, Niclas. forthcoming. *Foraging and the history of languages in the Malay Peninsula*.
- Burenhult, Niclas, Nicole Kruspe & Michael Dunn, this volume, Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula.
- Carey, Iskandar. 1976. *Orang Asli: The aboriginal tribes of Peninsular Malaysia*. Kuala Lumpur: Oxford University Press.
- Cavalli-Sforza, Luigi L. 1986. *African Pygmies*. Orlando: Academic Press.
- Cavalli-Sforza, Luigi L., Paolo Menozzi & Alberto Piazza 1993. Demic expansions and human evolution. *Science* 259. 639–646.
- 1994. *The history and geography of human genes*. Princeton: Princeton University Press.
- Childe, Gordon V. 1958. *The Prehistory of European Society*. Harmondsworth: Penguin.
- Fix, Alan G. 1977. The demography of the Semai Senoi. *Anthropological Papers* 62. Ann Arbor: University of Michigan Museum of Anthropology.
- 1995. Malayan paleosociology: Implications for patterns of genetic variation among the Orang Asli. *American Anthropologist* 97. 313–323.
- 1996. Gene frequency clines in Europe: Demic diffusion or natural selection? *Journal of the Royal Anthropological Institute (N.S.)* 2. 625–643.
- 1999. *Migration and colonization in human microevolution*. Cambridge: Cambridge University Press.
- 2002. Foragers, farmers, and traders in the Malayan Peninsula: Origins of cultural and biological diversity. In Kathleen D. Morrison & Laura L. Junker (eds) *Forager-Traders in South and Southeast Asia: Long Term Histories*, 185–202. Cambridge: Cambridge University Press.
- Fix, Alan G. & Luan Eng Lie-Injo. 1975. Genetic microdifferentiation in the Semai Senoi of Malaysia. *American Journal of Physical Anthropology* 43. 47–55.
- Higham, Charles. 2002. Languages and farming dispersals: Austroasiatic languages and rice cultivation. In Peter Bellwood & Colin Renfrew (eds) *Examining the Farming/language Dispersal Hypothesis*, 223–231.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, William Meehan, James Blackburn, Douglas Clarke, Joseph Maria Raja, Patimah Ismail, David Bulbeck, Stephen

- Oppenheimer & Martin Richards. 2006. Phylogeography and ethnogenesis of Aboriginal Southeast Asians. *Molecular Biology and Evolution* 23. 2480–2491.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, Douglas Clarke, Petya B. Blumbach, Matthieu Vizuete-Forster, Peter Forster, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2007. A mitochondrial stratigraphy for island Southeast Asia. *American Journal of Human Genetics* 80. 29–43.
- Livingstone, Frank B. 1958. Anthropological implications of sickle cell gene distribution in West Africa. *American Anthropology* 60. 533–562.
- Macaulay, Vincent, Catherine Hill, Alessandro Achilli, Chiara Rengo, Douglas Clarke, William Meehan, James Blackburn, Ornella Semino, Rosaria Scozzari, Fulvio Cruciani, Adi Taha, Norazila Kassim Shaari, Joseph Maripa Raja, Patimah Ismail, Zafarina Zainuddin, William Goodwin, David Bulbeck, Hans-Jürgen Bandelt, Stephen Oppenheimer, Antonio Torroni, Martin Richards. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308. 1034–1036.
- Menozzi, Paolo, Alberto Piazza & Luigi L. Cavalli-Sforza. 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201. 786–792.
- Ohashi, Jun, Izumi Naka, Jintana Patarapotikul, Hathairad Hananantachai, Gary Brittenham, Sornchai Looareesuwan, Andrew G. Clark & Katsushi Tokunaga. 2004. Extended linkage disequilibrium surrounding the hemoglobin E variant due to malarial selection. *American Journal of Human Genetics* 74. 1198–1208.
- Oppenheimer, Stephen, this volume, MtDNA Variation and Southward Holocene Human Dispersals within Mainland Southeast Asia.
- Pietrusewsky, Michael. 2006. A multivariate craniometric study of the prehistoric and modern inhabitants of Southeast Asia, East Asia and surrounding regions: A human kaleidoscope? In Marc Oxenham & Nancy Tayles (eds) *Bioarchaeology of Southeast Asia*, 59–90. Cambridge: Cambridge U.P.
- Rendine, Sabina, Alberto Piazza & Luigi L. Cavalli-Sforza. 1986. Simulation and separation by principal components of multiple demic expansions in Europe. *The American Naturalist* 128. 681–706.
- Saha, Nilmani, Joon W. Mak, JS Tay, Y Liu, JA Tan, PS Low & M Singh. 1995. Population genetic studies among the Orang Asli (Semai Senoi) of Malaysia: Malayan Aborigines. *Human Biology* 67. 37–57.
- Wood, Elizabeth T., Daryn A. Stover, Christopher Ehret, Giovanni Destro-Bisol, Gabriella Spedini, Howard McLeod, Leslie Louie, Mike Bamshad, Beverly I. Strassmann, Himla Soodyall & Michael F. Hammer. 2005. Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes. *European Journal of Human Genetics* 13. 867–876.
- Wright, Sewall. 1931. Evolution in mendelian populations. *Genetics* 16. 978–159.



# *Part IV*

*Origins and diversification:  
the case of Austroasiatic groups*



# 13 *Austroasiatic word histories: boat, husked rice and taro*

---

GÉRARD DIFFLOTH

## 1 Introduction

Sad to say, the full extent of linguistic diversity in Southeast Asia will soon be a thing of the past. And in several important ways, it already is. The fairly recent linguistic expansion of Southwestern Tai languages (for example Siamese, Lao, Shan, Phu Thai), and also of Burmese, Vietnamese and Malay, to mention only the main actors, has blanketed with relative uniformity what must have been a far richer linguistic landscape in previous times. The more ancient expansion of Khmer probably also erased an even older diversity, before being itself retired to its current, more confined location. We may never know how many, and more importantly what kinds of languages have been wiped out in Southeast Asia in historical times, to say nothing of prehistory.

Even today, the unrelenting march of linguistic uniformity shows no sign of abating. Quite the contrary: bolstered by new technologies and one-dimensional ambitions, it is gaining momentum. Southeast Asian linguistic diversity still somehow exists, but the full picture is now found, huddled and insecure, only in a few areas difficult of access. The rare field-linguists who are not only qualified but also able, or even allowed to witness and document this on-going tragedy can readily attest to this.

Examples are everywhere, and only one will suffice here: the Pearic branch of Austroasiatic ('AA' for short) that is, or should be, famous for having vowel systems with four phonation-types in phonological contrast. It is the sole witness in Southeast Asia of an extremely rare type of sound-system that is perhaps unique in the world's currently living languages. Seven of its eight languages will probably be forgotten in the next decade or so, and the eighth, Chong, is being rapidly supplanted by Thai in spite of current efforts to encourage its use.

This paper addresses several other issues concerning Austroasiatic linguistics. The reason for discussing them here is their importance for wider claims that are presently being made in the domains of economic history and population movements in India and Mainland Southeast Asia. The various linguistic issues are all ultimately connected to each other in the larger scheme of Austroasiatic history. However, I do not present them here in the form of a single sequential narrative; rather, I let them stand alone, each on their own. I do not try to unite them into one conclusion either; this, I feel, would be premature in the current state of Austroasiatic linguistic history.

In the first section of this paper, I note that Pearic languages contain a feature of glottalised phonation that is fundamental to their phonology. As the position of the Pearic branch within Austroasiatic remains unknown, accounting for this glottalisation will have important consequences in determining etymologies, establishing the AA family-tree, and proposing migration theories for the whole family. At present, Pearic glottalisation remains unaccounted for.

In the main part of the paper, I concentrate on the history of certain Austroasiatic words that, because of their meanings, play an important role in the current debates in other historical disciplines represented in this volume, beyond linguistics. But there are two methodological issues that have to be addressed first: one is the necessity to cover all languages and account for every piece of information they may provide; the other is the imperative need to involve the comparative method and historical phonology<sup>1</sup> in such research.

Moving then on to specific words, I choose three cases: ‘boat’, ‘husked rice’ and ‘taro versus rice plant’ that illustrate, each in their own way, what historical linguistics can offer other disciplines, and the pitfalls to be avoided.

Firstly, the etymon for ‘boat’, important to migration theorists, does not go back to proto-Austroasiatic; its geographic coverage is limited, and it shows signs of having been borrowed in some of the branches where it is represented; at the most, it reconstructs back no further than the Khmero-Vietic division of AA, perhaps even later. It is therefore of limited use in arguments about migrations.

Secondly, the etymon for ‘husked rice’ provides a different picture: the word has now been found in nine AA branches, covering the entire spectrum of Austroasiatic on both sides of the Brahmaputra; it is absent in only four branches that do not form a historical unit, and where its absence can be understood. At the proto-branch levels, the nine reconstructions differ somewhat, suggesting that the word does go back to Proto-Austroasiatic times. This must have consequences for the ancient history of rice cultivation in the region.

Lastly, for the etymon ‘taro’, there is a claim that it is assimilated with, and has been replaced by, the etymon ‘rice’. I show, on the contrary, that we have here two independent names, and two plants which have been consistently distinguished throughout Austroasiatic history: taro on the one hand, and a certain graminea on the other, which may not originally have been rice since AA has another etymon for it. This contains an appeal for detailed ethno-botanical research in AA societies, and also a word of caution.

## 2 Precious phonetic details

As mentioned above, the languages of the Pearic branch of Austroasiatic have four phonation-types in phonological contrast throughout their vowel systems.

There is a precedent for this in Austroasiatic. It has been pointed out some time ago (Diffloth 1989), that a two-way distinction in rime-glottalisation is found, in addition to Pearic, in three other branches of Austroasiatic: Bahnaric (Sedang), Katuic (Ta-oih, and also Ngkriang) and Vietic (several languages). More recently, two more branches, South

---

1 This is a discipline which has been able to build daring hypotheses, such as reconstructing Proto-Indo-European laryngeals, and then to test them, in this case with Hittite inscriptions deciphered several decades later. Historical phonology has reached, in this case, a level of scientific validity that few other social and historical sciences can boast of. It should also be said that it has been able to sail through many years of turmoil in linguistic formalism, practically unscathed.

Munda (Sora)<sup>2</sup>, and Khmuic (Iduh)<sup>3</sup>, have been found to have glottalised rimes as well, all of them unaccounted for historically. The implication is that Proto-Austroasiatic itself had a phonological contrast of this sort, although its history appears to be long and complicated. Accounting for glottalisation may well result in a revamping of the AA family-tree.

The Pearic branch consists of eight languages; they are currently spoken in Cambodia and in a small adjacent portion of Thailand. They are: Chong, Chu-ng (= Sa'och), Samre (= Baradat's West Pear<sup>4</sup>), Samrai, East Pear, Kasong (= Chong of Trat, Baradat's Chong), So-ong (Baradat's Souei) and Peuar (= Baradat's Pear of Kg. Thom).

Of the eight, all but Peuar<sup>5</sup> are languages with a four-register system, where the vowels are pronounced with four possible phonation-types called 'registers', all in phonological contrast. The four registers, labelled: Clear, Breathy, Tight and Creaky<sup>6</sup>, can be analysed as forming two sets of two registers each: the Clear and the Breathy forming a non-glottalised set, the Tight and the Creaky a glottalised set.

For the feature of glottalisation of the second set there is no satisfactory historical explanation at the moment. One hypothesis (Ferlus 2004) proposes that the monosyllabic vs. sesquisyllabic structure of the whole word is the conditioning factor for glottalisation of the vowel nucleus.

As support for this proposal, Ferlus cites three Chong verbs that are monosyllabic and non-glottalised, while the three nouns derived from these verbs are disyllabic and glottalised. These three pairs had been earlier cited in Diffloth (1989) to show that certain cases of glottalisation might be explained as innovations within the Pearic branch, while others had to be more ancient. For example, Diffloth also cites nine examples of Chong words having glottalisation and monosyllabic AA etymologies. These are not mentioned in Ferlus (2004), nor are the many other examples to be found in Huffman (1985), and in Thongkum (1991). More recent collections of Pearic vocabulary (Choosri 2002, 2009), some covering the entire branch (Diffloth field-notes 1982–2009), confirm the argument advanced here.

A full discussion would be beyond the scope of this chapter, but a rough evaluation can be proposed by counting<sup>7</sup> how many words have one register or another. In a So-ong vocabulary<sup>8</sup> where obvious Khmer borrowings have been set aside, in the remaining total of 1550 entries, we find the following lexical-type percentages:

Among words with the non-glottalised registers, 42% are strictly monosyllabic and the rest are sesquisyllabic.

---

2 A. Zide's dissertation, Rau ICAAL 2009.

3 A Khmuic language spoken on both sides of the Vietnam-Laos border, called Óđu in Vietnamese sources, and Thai Haat in older Thai sources (GD fieldnotes 1997).

4 A description of the various Pearic groups is found in Baradat (1941).

5 The remaining Kasong speakers, fluent in the Chanthaburi dialect of Thai, now use the rising-falling tone in place of gottalisation, but they maintain the breathy vs. clear phonation contrast; older notations taken 30 years ago (GD notes) indicate both glottalisation and breathiness.

6 Thongkum (1991) uses slightly different labels for Chong registers, but with the same pattern of bi-partition and glottalisation.

7 Here, we are counting the number of distinct lexical items; the frequency of each lexical item cannot be counted here, as it would require large quantities of texts that simply do not exist for these languages.

8 I have collected So-ong materials over several years, starting in 1992 when much of the community had been relocated near the city of Kampong Speu (Cambodia), and later in several visits to their original village of Tei Do:n Pa:, south of Phnom Aural, in Kg. Speu province; my latest visit was in March 2009.



With the glottalised registers, the result is 53% of strictly monosyllabic words.

We therefore see no correlation in So-ong today between syllabicity and glottalisation in the original vocabulary. If syllabic patterning had been the historical conditioning for Pearic glottalisation, this would have left behind a much more uneven distribution. Similar counts in other Pearic languages can extend this conclusion to the whole Pearic branch.

Pearic glottalisation may appear at first as a puzzling phonetic detail, but it cannot be dismissed in a historical account of the Austroasiatic family.

### 3 Word-histories

Two methodological remarks to begin with: first, we need to cover and account for every language involved, and second we cannot dispense with historical phonology.

The Austroasiatic family still comprises about 164 different languages today, but this large number is misleading; at least half of the total are in situations not much brighter than those of the Pearic branch just mentioned. This includes small languages that provide vital and irreplaceable evidence for the history of the more famous and secure languages: for example Nyah Kur for the ancient history of Mon (Diffloth 1984), Kri for that of Vietnamese (Enfield and Diffloth 2009), Iduh for Khmuic, Mnar for Khasi, Che'Wong for Aslian, etc. Historical linguistics relies heavily on such information. At the same time, the work of historical linguists offers a boost of confidence to small communities that are totally absent from the histories that are taught in schools, or at best relegated to a quasi-mythical past.

As I intend to show below, every language counts, every dialect has something to contribute, and the goal is not reached, at least ideally, until every detail is accounted for.

Something also needs to be said about the choice of techniques. Due in part to rapid developments in other sciences, historical linguistics is presently recovering some of the attention it had gradually been losing to linguistic formalism and typology. The prime movers in this recovery have been archaeology and genetics. Southeast Asian archaeology has now expanded far beyond its previous preoccupation with kingdoms and monuments, and into broader concerns such as foods and habitat. In genetics, the discovery of DNA has opened entirely new vistas regarding the domestication of certain cultivated plants, and more generally the history of cultivation; in all these areas, etymology (historical lexicology) certainly has quite a few things to say. Human geneticists are also raising questions about demic expansions and residual populations, questions where historical linguists feel perhaps less at ease but nonetheless concerned. There is also a regain of interest in such matters as finding linguistic homelands, a topic that had been sidelined for some time.

These are stimulating developments, but their rapidity is disconcerting. Historical linguistic research at the level of large language families consumes enormous amounts of research-hours and takes years, even generations to mature, especially when the devotees are few and dispersed across the world. But now, there is pressure to rapidly produce some answers to the very large and looming questions. In such a context, the temptation is great to have recourse to simple techniques that produce precise numerical results.

Lexicostatistics seems to be an obvious choice in this situation. In a nutshell, it consists in counting lexical replacements in basic vocabulary, in order to estimate the relative times of separation<sup>9</sup> of related languages. Unfortunately, it has faced strong objections from its

---

<sup>9</sup> 'Relative times', as distinct from estimating the absolute times of separation; the latter was the goal of glottochronology, a technique that was abandoned very early on.

very inception, for example Sauvageot (1951), Hoijer (1956), and its use has been regularly denounced since that time by many historical linguists.<sup>10</sup>

As a quick reminder, allow me to summarise here some of its flaws. The idea that some lexicalized notions are basic and others are not remains vague. The glosses chosen for these notions are highly polysemic, and their translation into the glosses, themselves polysemic, of other languages entails arbitrary decisions. The claim that these lexicalised notions would be subject to a constant rate of replacement over time and across languages has been abandoned, but variations in that rate remains unknown and may well be erratic and beyond the scope of statistics. Counting replacements also requires being able to identify cognates, and this creates further problems:<sup>11</sup> cognacy-decisions demand a solid knowledge of linguistic history, including a clear picture of language separations within the family, which is precisely what the technique was supposed to discover in the first place. Overall, lexicostatistics provides numerically precise results; but such figures are deceptive, being based on vagueness, arbitrary decisions, unknown variability, and circular accounting; it should not be accepted as a reliable technique.

Fortunately, historical linguistics and especially historical phonology has better tidings to offer. In the current state of Austroasiatic language history, it is possible to move beyond simple-minded lists of potential cognates. In many cases, we can now give specific reasons for eliminating certain look-alikes that are not cognates, for identifying borrowings and tracing their source and relative timing. Usually, we can justify our reconstructions of the phonetic and of the semantic history of cognates with some precision, or at least by approximation, from the deepest level of Proto-AA down to the present. And in the best cases, we can move up and down the historical tree throughout the family, noticing changes, identifying replacements, and describing the precise nature of outside contacts, as if travelling in a sort of Wellsian time machine.

Let us now move on to the word histories.

### 3.1 Boat

If a word for ‘boat’ could be reconstructed to Proto-AA times, this could be used in an argument concerning ancient navigation and migration; unfortunately, the only AA word for ‘boat’ that has some antiquity in the family, does not fit.

Shorto (2006: No.336) reconstructs a form \**duuk* and assigns it to Proto-MK.

Reflexes are found with that meaning in practically all the languages of five branches of Mon-Khmer as the term is traditionally understood: Vietic, Katuic,<sup>12</sup> Bahnaric, Khmeric and Pearic, but in no other branch so far in Mon-Khmer or in the rest of Austroasiatic.

Except for Pearic, whose historical position remains undecided, the four other branches constitute the Khmero-Vietic (East Mon-Khmer) division of Austroasiatic (Diffloth 2005).

---

10 Campbell (1998): ‘...unreliable and discounted by most historical linguists’, p.185, ‘... should be rejected’ p.186, Lemaréchal (2010): ‘...ought to be declared null and void’. Such vigorous condemnations reflect impatience with the users of this crude and flawed technique.

11 Bayesian phylogeny, imported from genetics to the domain of linguistics, defines ‘cognate-sets’ by using typological, not historical criteria. Bayesian phylotypy would be a better term for it.

12 The only Katuic language that lacks a reflex of this etymon is Katu, where Costello (1971) gives seven words for ‘boat’, none of them related to \**du:k*.

The reconstruction therefore probably does not go back to a historical horizon that would be as early as Proto-MK,<sup>13</sup> even less Proto-AA.

And even within Khmero-Vietic, its antiquity is in doubt.

In the Vietic branch, all the languages support a reconstruction with an initial implosive \*ɗ-; for example:<sup>14</sup> Ahlau: /dɔ:k/, Kri: /dɔ:k/, Vietnamese (Vinh dialect):<sup>15</sup> <noóc>.

But in the other two branches, Katuic and Bahnaric evidence definitely supports a Proto \*d-, not an implosive \*ɗ-. For example in Katuic: Kuay-Ndua: /tɯ:k/, Kuay-Ntaw: /tɯaʔ/, Bru<sup>16</sup> and So: /tɯək/, Pacoh:<sup>17</sup> /tɯək/, and in Bahnaric: Bahnar:<sup>18</sup> /du:k/. If the proto-forms had an implosive \*ɗ-, Kuay would need a /d-/ with either a clear or a breathy phonation, depending on the dialect, Bru and Pacoh would also need a /d-/, and Bahnar would have transmitted an implosive /ɗ-/ until today.

Katuic and Bahnaric disagree in this respect with Vietic, and this sort of disagreement is not one that would indicate a more ancient history, as there are no correspondences that would support this. On the contrary, it is far more likely that borrowing must have taken place in one direction or another among these three branches, and at a period later than Khmero-Vietic.

The Khmeric evidence: Standard and Surin Khmer: /tu:ʔ/, Chanthaburi Khmer: /tɯ:k/, provides no clue in this regard since \*ɗ- and \*d- have merged to \*d- very early in the history of that branch, and this merged \*d- then conditions a Middle-Khmer breathy voice that is preserved in Chanthaburi, but disappears in modern times elsewhere in Khmeric.

As for Pearic, the register of So-ong: /tɯ:k/ shows that the word was borrowed from Middle Khmer, and so were the Chong, Chu-ng and Kasong forms: /tɔʔ/, perhaps borrowed at an even later date, as the final /-ʔ/ suggests. In this way, the only branch that might have supported some antiquity for the etymon ‘boat’, the Pearic branch, fails to do so.<sup>19</sup>

In the end, close attention to laryngeal features and glottalisation, and in view of its geographic spread, this word for ‘boat’ cannot be very ancient, and so is not useful in reconstructions of remote cultural prehistory.

### 3.2 Husked rice

Rice cultivation and the possible role of rice in ancient diets have taken enormous importance in recent discussions. It is therefore necessary to look into the detailed history of the vocabulary relating to rice. The subject would require a whole volume, and we will examine here only one item, ‘husked rice’ because its meaning is central to these issues.

Proto-AA #rəŋko:ʔ has been proposed for ‘husked rice’ (Diffloth, 2005).

In this notation, the prefixed #, instead of the customary \*, indicates a provisional and impoverished reconstruction for this very ancient period; the proto-consonants are fairly

13 It should be said also that Shorto (2006) operates without branch reconstructions (p.xvii); hundreds of his proposed Mon-Khmer etyma do not actually reconstruct back to Proto-MK times, and in many cases they pertain only to one branch of MK, or even to one sub-branch.

14 Unless otherwise indicated, examples cited in this paper come from my own fieldwork

15 Ferlus (1991).

16 Thongkum and Peungpa (1980).

17 Watson and Cubuat (1979).

18 Banker (1979).

19 The Nyah Kur form /tɯ:k/, without any cognate in the remaining Mon branch of Monic is probably another borrowing from Middle-Khmer, of which Nyah Kur has a large number.

certain, the proto-vowel quality is only an approximation, and suprasegmentals such as phonation-types remain undecided.

As we follow the history of this word from Proto-AA times down to the later divisions and branches of the family, and eventually to the languages and their dialects, reconstructions, with the prefixed \*, become more precise and better justified.

### 3.2.1 Pearic

*Proto-Pearic*<sup>20</sup> \*rək<sup>h</sup>ɔ

The aspiration of medial /-k<sup>h</sup>-/ is a diagnostic Pearic innovation from an earlier \*-k-, as shown by cognates outside Pearic with a \*-k- (see below), and by the Pearic forms: Chong: /kək<sup>h</sup>ɔ/ and /kək<sup>h</sup>ɔ:/, Samrai (West Pear): /rək<sup>h</sup>ɔ/, Kasong (Trat): /lək<sup>h</sup>ɔ:/ and So-ong (Souei): /rək<sup>h</sup>ɔ:/.

Only Peuar (Pear of Kg. Thom): /rəkɔ/, has an unaspirated /-k-/ which is probably not an archaism<sup>21</sup> but rather a secondary development from P-Pearic \*-k<sup>h</sup>-.

This early aspiration is an important historical signal, suggesting that the Pearic words in question were not borrowed from Khmer, but directly transmitted from earlier Proto-Pearic times. The lack of a final -ʔ is also ancient, but not as old as Proto-AA, being the result of a loss that also took place in the entire Khmero-Vietic division,<sup>22</sup> and in some other languages outside it. The Proto-Pearic clear-voice register is consistent with a Proto-AA voiceless \*-k- in the onset; however, the lack of glottalisation in this item contradicts Ferlus' hypothesis according to which sesquisyllabic words would acquire a glottalised register in Pearic.

### 3.2.2 Khmero-Vietic (Eastern Mon-Khmer) division

In the Khmero-Vietic division, only the Bahnaric branch totally lacks reflexes of our etymon; in that branch, we find a form like \*phɛ:, a Proto-Bahnaric semantic innovation. Otherwise, in the other three branches of Khmero-Vietic: Khmeric, Katuic and Vietic, most of the languages of have reflexes of PAA #rəŋko:ʔ 'husked rice'.

#### 3.2.2.1 Khmeric

*Proto-Khmeric* \*rəŋkɔ

Modern Standard Khmer has /ŋkɔ:/; this word is spelled in Khmer with a final <-r> in the accepted orthography of this word. The Khmer dialect of Surin also has /ŋkɔ:/, without a final /-r/; if the Khmer spelling convention with <-r> had been historically faithful we would have expected Surin to have a final /-r/. The very marginal Khmer dialect of Chanthaburi, where registers contrasts have survived, has /ŋkɔ:/ with a Clear register, and

20 In a register language, if breathy phonation is marked on the vowel (thus: [V̤]), clear phonation will also have to be marked. I will choose a subscript line [v̤] for this purpose. Similarly, if vowel glottalisation is marked (thus: [v̤']), then non-glottalisation will also have to be marked (thus: [v̤%]), since this absence of glottalisation is a positive piece of information when a contrast is possible.

21 Whether Peuar has lost aspiration, or never had it, remains a difficult question; most of the few remaining speakers of Peuar are trilingual with Kuay and Khmer, and the inherited Pearic vocabulary of Peuar is steadily fading into oblivion. The Peuar word for 'husked rice' could well be a borrowing from Kuay Ndreu, see below.

22 There is a small number of words with final -ʔ in Bahnaric and Katuic that require a separate explanation.

no final /-r/ either. The Middle Khmer reconstruction is therefore \*ŋkɔ̄, with clear-voice register and no final consonant.

In the Old Khmer inscriptions of the Angkor period (10<sup>th</sup> cy) we find <raŋko> in several lists of gifts to temples, where a meaning ‘husked rice’ seems appropriate; all of this taken together supports a Proto-Khmeric reconstruction \*rəŋkɔ̄, with a voiceless \*k-, without register distinction at that early period, and no final consonant.

### 3.2.2.2 *Katuic*

#### *Proto-Katuic*<sup>23</sup> \*rəŋkau

Most of the reflexes are found in the West Katuic sub-branch which includes the Kuay complex and the Bru complex. In East Katuic, only one language has a reflex.

#### *West-Katuic* \*rəŋkau

Kuay complex: Kuay Â' (Cambodia): /rəŋkɔ̄u/; Kuay Yeu (Sisaket): /rəkɔ̄u/; Kuay of Sukhuma (Laos): /lakɔ̄u/, Kuay Ndua and Kuay Ndreu: (Cambodia) /rəŋgɔ̄u/, Kuay Mlâ (Cambodia and Surin): /rəŋkɔ̄u/, Kuay of Lao Ngaam (Laos), of Kratie (Cambodia), and Kuoy (Sisaket): /ŋkɔ̄u/, Kuuy (Surin): /ŋkɔ̄:w/.

Only in the last item, Kuuy of Surin, do we find a long /a:/; this dialect is spoken in a small area of Surin province, near Baan Tɛel, which is the source of an often-quoted dictionary (Sriwises 1978); lengthening of short \*a in certain conditions is a very recent development specific to this sub-dialect.<sup>24</sup>

#### *Souei (Laos, Ferlus 1974) /haŋkaw/*

Bru (Miller 1976): /rakɔ̄u/; other dialects of the Bru complex lack this word and have replaced it with an innovation: Bru (Thongkum and Peuangpa 1980), Sô: /ʔasɔ̄ʔ/, Truk (Phu Phaen mountains): /ʔasəʔ/; this word is also found in Pacoh and Ta-oih (East-Katuic).

In the East-Katuic sub-branch, our etymon has been found so far only in one language, Ngkriang:<sup>25</sup> /harkɔ̄:/.

#### *Other East-Katuic languages have one of two etyma*

##### 1) \*cəněh (Thongkum, 2001:263):

Chatong: /caněh/, Kantu: /ciněh/, Tariw and DakKang:<sup>26</sup> /canih/, Katu:<sup>27</sup> /cəněh/, Phueang: /ciněh/. This is a Noun derived by -n- infixation from a verb \*cěh found in Katu and in Phueang: /cěh/ ‘to sow by tossing’, and also in West-Bahnaric (Jacq and Sidwell 2000: No.371; Thongkum 2001:533) where it is not a verb but a noun: ‘rice grain (unhusked)’.

23 The no-glottalisation sign (%) indicates that the rime in this Proto-Katuic word belongs to a set of correspondences where a non-glottalised rime would be expected in Ta-oih, if Ta-oih had a reflex.

24 Without the other varieties of Kuay at his disposal, Shorto (2006: No.1820) was led to reconstruct a ProtoMK ‘variant’ with a long vowel: \*rkaaw? (followed by a question mark), entirely on the basis of this Kuuy form; both Peiros (1996) and Sidwell (2005) omit this item from their reconstructions of Proto-Katuic.

25 This language is also called Kriang (Thongum 2001), and in Lao: Ngeq.

26 Thongkum (2001:263)

27 Costello (1971)

- 2) \*ʔasəʔ, of unknown provenance, found in Pacoh and Ta-oih (and in some dialects of the Bru complex, as mentioned above).

As there are no reflexes, other than in Ngkriang, in the East-Katuic languages, the exact nature of the proto-Katuic vowel is not certain for this word, but Proto-Katuic \*rəŋkau is one possibility. There is also a possibility that the Ngkriang word may have been borrowed from an early form of Old-Khmer, with interesting implications for the early ethno-history of the region.

### 3.2.2.3 Vietic

#### *Proto-Vietic \*rəŋkoʔ*

Reflexes are found in all five sub-branches of Vietic.

West-Vietic: Ahlau (Thavung): /hakɔʔ/, Aheu (Pon Sung): /ŋkɔʔ/. The initial /ha-/ of Ahlau has P Vietic \*rə- as one possible source, so the West-Vietic sub-branch contains by itself most of the information needed to reconstruct the Proto-Vietic form.

SWest-Vietic: Maleng (Paak Atan): /ʔakɔʔ/. The other South-West Vietic language, Kri, does not have a reflex of the Vietic etymon; instead, it has a word /ca:wʔ/, from another P-Vietic etymon, whose meaning covers the rice seed in all its forms.

SEast-Vietic: Rục: /rəko: 35/. The [35] rising tone ends in a glottal constriction that is not counted as a consonant (Lợi 1993) but considered a redundant part of the tone. Rục is the only Vietic language to have kept intact the initial \*rə- of Proto-Vietic in this word.

NWest-Vietic: Tum: /ka:oʔ 35/. Tum and the other languages of the NWest-Vietic sub-branch (Phong, Poong, Cuôi), have become monosyllabic and so retain no trace of the minor syllabic \*rəŋ- of Proto-Vietic. The final glottal stop is found following other Tum tones and cannot be considered a predictable feature of the rising tone; it has a segmental, consonantal value in this language.

Viet-Muong: Muong Khoi:<sup>28</sup> /ka:w 33/, Muong Khen:<sup>29</sup> /ka:wʔ 24/, Vietnamese <gạo>. The various dialects of Muong have lost the minor syllable \*rəŋ-, but their tones include glottalisation as a predictable feature, and also indicate a voiceless \*k- initial. Vietnamese <gạo> has a surprising initial fricative [ɣ], and a nặng tone that implies a \*voiced initial; both apparent anomalies have been explained as regular Vietnamese developments (Ferlus 1976) from an earlier sesquisyllabic form having a voiced minor-syllable onset and a \*k- initial in the main syllable.<sup>30</sup>

The nặng tone of VN <gạo> would also require the reconstruction of a final stop. This anomaly had been noticed early on by Haudricourt who sought an explanation in the final glottal stops of Khmu' (Haudricourt 1953). However, there are many examples of Khmu words with a final -ʔ having Vietnamese cognates with the ngang and huyền tones instead, implying no final glottal stop there (Diffloth 1989), so the explanation does not stand. The solution I propose is that words like <gạo> did not have a final segmental consonant, but the rime had a supra-segmental glottalised phonation-type, perhaps tight voice or even creaky voice, noted here as a final [-ʔ] for convenience. It is this phonation-type, not the much older Proto-AA consonantal \*-ʔ, which conditions not only the Vietnamese nặng

28 Sokolovskaja and Tài (1987).

29 Barker and Barker (1967).

30 This explanation can help to settle speculation that VN gạo would have a Tai origin; in fact it is Thai /khâaw/, and Lao /khawʔ/ whose history is unclear and causes difficulties in comparative Tai.

tone,<sup>31</sup> but also the glottal features of the Mưong and Rưc tones, as well as the segmental final /-ʔ/ of Tum, Maleng, Ahlaw and Aheu shown above.

This proposal also implies that Proto-Vietic had actually lost the earlier Proto-AA final segmental \*-ʔ, found in words like our ‘husked-rice’ etymon, as did the other three branches of Khmero-Vietic: Katuic, Bahnaric,<sup>32</sup> and Khmeric. This loss of final P-AA \*-ʔ would constitute one of the Proto-Khmero-Vietic (Eastern Mon-Khmer) diagnostic innovations.

### 3.2.3 Nico-Monic (Southern Mon-Khmer) division

The P-AA etymon for ‘husked rice’ would appear to be absent from the entire Nico-Monic division, were it not for the Aslian evidence; this Nico-Monic division includes, in my view, the Monic, the Aslian and the Nicobarese branches.

Nicobarese lacks the etymon entirely: Nancowry uses a word /ʔaruəs/<sup>33</sup> of Portuguese origin, and Car has /sa:ŋ/ (Whitehead 1925), of unknown ancestry.

Monic also lacks this etymon.

Mon uses a word /həʔ/, spelled <snu>, and found as <sno’> in medieval Old Mon inscriptions (Shorto 1971:390), for which we could reconstruct an early \*sno:ʔ.

But as this word lacks a Nyah-Kur cognate, there is no guarantee that it descends from proto-Monic. Cognates outside Monic are found in nearly every Palaungic and Khmuic language, where it means either ‘husked rice’ or the ‘rice-plant’ or both: for example Khmu’ /həʔ/, Khabit: /səŋ:/ for Khmuic, and Ta-ang /həʔ:/, Imok /səŋəʔ 44/ for Palaungic. It is not found anywhere else in AA. These two closely related branches Palaungic and Khmuic form a node inside the Khasi-Khmuic (Northern MK) division. The possibility then arises that this would be not be a P-AA etymon, but a local Northern word which the Mon would have borrowed as they entered Lower Burma in medieval times, coming from their earlier Dvaravati location<sup>34</sup> in what is now Central Thailand.

Nyah Kur has /ŋkə:/, where the absence of a final glottal stop betrays its Middle-Khmer borrowed origin.<sup>35</sup>

#### 3.2.3.1 Aslian

##### *Proto-Aslian \*raʔkuʔ*

Only Aslian has a reflex of the P-AA etymon, and only one of the 16 Aslian languages, Jah Hut, maintains the original meaning: Jah Hut /rəʔkwəʔ/ ‘husked rice’.

31 And also the sác tones in other etyma with VN open syllables.

32 Some Bahnaric languages have infrequent cases of final glottal stops, usually preceded by a short vowel; these would require a full reconstruction of the proto-Bahnaric vowel system that would fully take into account the crucial and diverse languages of the North-Bahnaric branch; an unpublished reconstruction of Bahnaric (Sidwell 1999) does not do this.

33 In Roepstorff (1884): <arosh>.

34 Diffloth (1984)

35 Nyah Kur borrows Old Khmer words over a period of time, starting in the 9th century AD when the Khmer language became dominant in the Northern Isarn region where Nyah Kur is still spoken, and ending in the 15th cy with the sacking of Angkor. The Nyah Kur word /ŋkə:/ for ‘husked rice’, lacking the initial \*rə- found in 10th cy Angkorian Khmer inscriptions, must have been borrowed by Nyah Kur at a somewhat later date. The Nyah Kur vowel quality /ə:/ agrees with the Old-Khmer and Middle-Khmer values.

There is also a possible reflex in Semai: /rako:ʔ/, but the word designates a weed, commonly found on roadsides and as regrowth in rice-fields after harvest.<sup>36</sup> The Aslian onsets, Semai /ra-/ and JH /rəʔ-/ , raise interesting questions concerning the historical morphology of the Aslian branch; we have no solution for this at the moment. The Semai /o:/ vowel, corresponding to Jah Hut /wə/, indicates a Proto-Aslian \*uə diphthong, an acceptable reflex of P-AA \*o:. The final /-ʔ/ is what is expected historically of Aslian, and of the Proto-Nico-Monic (Southern M-K) division as a whole. Historical phonology shows that there is no reason to suggest that Jah Hut should have borrowed this word with its proto-meaning simply because it is unique in the entire Nico-Monic division. The nearest AA branch on land, Monic, is far away and has lost this etymon in proto-Monic times; besides, the phonology of this word in Jah Hut does not closely resemble anything found in the AA family as a borrowing would require, and its geographic location, far south into the centre of the Malay peninsula, is not a likely place for contact-induced borrowing.

Aslian languages have a variety of other etyma for ‘husked rice’: Semai has \*cənɾə:y, a word with cognates in Katuic meaning ‘glutinous rice’, and in Khmuic meaning ‘broken rice’. Temiar and Semnam have a word /ba:ʔ/ which also designates the rice-plant, this being its original P-AA meaning. North-Aslian languages usually borrow the Malay word <beras>. The South-Aslian languages Mah-Meri, Semelai and Smaq-Bri do not have a specific word for ‘husked rice’,<sup>37</sup> and the P-AA etymon is also lost in that branch.

The Aslian historical picture for ‘husked rice’, which looks confusing at first, is a fair reflection of the various life-styles of societies where rice is not an indispensable element of the diet, and among which only the Jah Hut are lowland farmers. The Semai situation is puzzling and suggests a not-so-ancient shift from some kind of rice cultivation to a more jungle-oriented subsistence where the original referent of the etymon would have become unimportant, and then forgotten; but the word itself was evidently remembered and assigned to another useful plant for some reason unknown to us.

In any event, the survival of the word in Jah Hut, suggests that it has been continuously transmitted through the centuries with its original meaning, via Proto-Nico-Monic, and Proto-Aslian as well, but replaced in other Aslian languages.

In the rest of Nico-Monic, its loss among the Nicobarese fishermen-islanders is much easier to explain; but its total loss in the whole Monic branch is surprising since the people involved, Mon and Nyah Kur, have probably been growing rice for centuries. More on this below.

### 3.2.4 Khasi-Khmuic (Northern Mon-Khmer) division

#### 3.2.4.1 Khmuic

*Khmuic*: \*rəŋkǝʔ

The Khmuic branch is historically far more diverse than is suggested by the only Khmuic language, Khmu’, that has been extensively studied.

Mla-Bri (Phi Tong Lueang) distinguishes itself, perhaps typically for a hunting-gathering group, in lacking a reflex of the Proto-AA term for ‘husked rice’. Instead, it has a

---

36 This word was spontaneously quoted for me during a visit to Angkor by A. Williams-Hunt, an endless source of knowledge on his native language, Semai. I have no identification for this plant at the moment. The sample I was shown was a very common small plant (50 cm) with cordate serrated leaves (5 cm long), glabrous on both sides; the leaves are medicinal and used to stop bleeding; perhaps a species of Eupatorium or Sida. The Surin-Khmer name of the plant is /kənrəaŋ kɛ:t/.

37 Nicole Kruspe, personal communication.



word /pi:ʔ/ (Rischel 1995) which is probably cognate with the Bahnaric \*phɛ: mentioned above, where the loss of P-AA medial \*-h- is diagnostic for the Khmuic branch. Rischel also reports an expression /yu:k thurba:ʔ/, where the second word recalls P-AA \*ɬa:ʔ ‘the rice-plant’.

Khabit (Bit, Phsing) also lacks a reflex, and uses a word /piəŋ/ for ‘husked rice’; this etymon has a wide distribution in AA including: Kharia /ɔmpɛŋ/, Mon /həpɛəŋ/, Khmer /səbiəŋ/.<sup>38</sup> It is also found in several Austronesian languages Aceh: /əmpriəŋ/, Cham: /ʔapiəŋ/, Malay: <emping>, Tagalog: <piping>. It generally refers to certain food items, especially those made with glutinous rice; these are available in local markets, which probably explains the far-flung distribution of the word.

Otherwise, the P-AA etymon for ‘husked rice’ has reflexes in all Khmuic languages. But there is a distinction between the languages that have lost the older final \*-ʔ and those that maintain it.

Final \*-ʔ is lost in the eastern areas of Khmuic: Khang /ŋkɔ:/, Bumang /kɔ: 51/, Ksing Mul (Puok): /harkɔ:/, Pong Laan: /ŋkɔɔ/, Pong Piat /rəkɔ:/, and in Iduh (Ođu, Thai Haat): /ŋkɔ:o/.

Final \*-ʔ is maintained in the more western Khmuic languages: Khmu’: /rəŋkɔʔ/, Thai-Then: /harkɔʔ/, Prai: /khuʔ/ and Mal: /ŋkhuʔ/. A reconstruction \*rəŋkɔʔ is probable for Khmuic.

### 3.2.4.2 *Palaungic*

#### *Proto-Palaungic \*rəŋko:ʔ*

Except for Lamet,<sup>39</sup> all Palaungic languages maintain the P-AA etymon for ‘husked rice’, in a great variety of forms.

In West-Palaungic, where I would reconstruct \*rəkɔ:ʔ, the final \*-ʔ and medial voiceless \*-k- are maintained in some languages: Riang /koʔ 55/ and Kano’ (Danau):<sup>40</sup> /koʔ/. But in the very diverse Ta-angic (Palaung-Pale) sub-branch, the final \*-ʔ is lost in every language after \*long vowels; in addition, some languages maintain the original medial voiceless \*-k-: Ta-ang (Palaung proper): /rəkɔ:/, Ka-ang: /həkɑ:w/, Rumai: /ləkɑ:o/, but others have voiced the older \*-k- in a systematic reversal of voicing values (Diffloth 1991): Pale (Silver Palaung): /diga:w/, Na-ang: /nəgɛ:o/.

In the diverse East-Palaungic branch, where we can reconstruct \*rəŋko:ʔ, final \*-ʔ is always maintained:

In the Waic sub-branch, only Lawa maintains the original initial \*rə-, either exactly: La-up: /rəkɔʔ/, Umphai /rakɔʔ/, or as a trace: Pa-Phae: /ləkɔʔ/, Bo-Luang: /ʔakɔʔ/. The remaining Waic languages have a monosyllable in which some retain the original \*-ŋ- and others lose it: Paraok: /ŋgəʔ/, La: /kauʔ/, Phalok: /gɔʔ/, Bulang: /ŋkuʔ 55/, Phang: /kuʔ/.

In the Angkuic sub-branch, the languages are tonal and have all changed the originally voiceless \*-k- into an aspirated \*-k<sup>h</sup>-, a diagnostic innovation of this sub-branch. In addition one language, Imək, has maintained a trace of the initial \*r-: Imək: /ləŋk<sup>h</sup>oʔ 55/, while the others have lost it: U: /ŋghuʔ 13/ (Svantesson: /ŋkhù/), ManMet /k<sup>h</sup>ɔʔ 55/. The

38 Borrowed by Thai: /sabiəŋ/ ‘food and other items useful for survival’.

39 In spite of its location next to the Khmu’-speaking area, and much social interaction with Khmu’ speakers, Lamet belongs historically to the East-Palaungic branch, and not to Khmuic.

40 GD fieldnotes, 1980; Luce noted: kɔ<sup>l</sup>, where the tone mark is probably meant to include a final glottal stop.

tones of the Angkuic languages are conditioned, in part, by the older contrast in proto-vowel length (Svantesson 1988).

### 3.2.4.3 *Pakanic*

The three languages of this branch: Mang, Palyu (Bo Liu, Lai) and Pakan (Bu Geng) have all lost the P-AA etymon for husked rice: Mang has /ba:k 23/, Palyu has /ʔja:ŋ55/, all without known AA etymologies. This is not surprising as these languages have large numbers of borrowings from their various unrelated neighbours.

### 3.2.4.4 *Khasian*

*Proto-Khasian: #rək<sup>h</sup>aw*

The Khasian branch of AA is far more diverse than was thought until recently.

Standard Khasi has /k<sup>h</sup>a:w/ for ‘husked rice’, while Pnar (Jaintia) has /k<sup>h</sup>o:/ and Lyngngam has /k<sup>h</sup>aw/; these forms suggest that the well-known Tai word for rice may have been borrowed early on by Khasi.

In that case, Amwi /rəhia/ and Lakadong /rəhaw/ ‘husked rice’ would be unrelated to the first set. Systematic comparison tells, however, a different story and both sets actually derive from the same P-AA source; the Tai-family look-alikes will require another explanation.

The medial /-h-/ of Lakadong and Amwi corresponds regularly to the Standard, Pnar and Lyngngam /k<sup>h</sup>-, and both descend from P-AA \*k-, as we can see for example in Khasi /k<sup>h</sup>o:n/ vs. Amwi /hun/ ‘child’. The initial /rə-/ of Lakadong and Amwi was lost in the rest of Khasian. Even the /ia/ of Amwi corresponds, in other etyma, with Khasi /-a:w/, for example: Khasi /ma:w/ ‘stone’ versus Amwi /ʃmia/ ‘stone’. All the Khasian forms descend therefore from a single Proto-Khasian form where the proto-rime is still uncertain, but could be \*-aw.

### 3.2.5 *Munda*

*Proto-Munda: #ruŋkuʔ*

Reflexes are found in Kharia-Juang and in the Southern Munda languages.

Kharia: /rəmkuʔb/, Juang: /ruŋkub/. The labial final, unusual in this etymon, appears to be due historically to the combination of a final glottal with an /u/ vowel.

The other Munda languages do not have this labial final:

Sora: /rɔŋkɔ-n/, Gorum: /rũŋk<sup>h</sup>/, Gutob: /rukug/, Remo: /ruŋku/, Getaʔ: /rkoʔ/.

These cognates are important in establishing the antiquity of the Austroasiatic etymon for ‘husked rice’ as there is nothing in their phonology or geographic distribution that would suggest borrowing.

There has been a suggestion (N. Zide, personal communication) that the first syllable of these words could be an independent etymon, found in a verb meaning ‘to pound (rice)’: Kharia: /duruŋ/, Santali: /huŋuŋ/, Mundari: /ruŋuŋ/, Ho: /ruŋ/. In that case, the word for ‘husked riced’ would be a compound #ruŋ + kuʔ, and the proto-form for ‘rice’ would be a monosyllabic #kuʔ. In the wider context of Austroasiatic, this is problematic as there is no trace of such a monosyllabic form. As shown above, only a few Vietic and Palaungic languages have a monosyllabic reflex, and in every case it can be shown that these monosyllables are the result of sound changes that operate throughout the languages in question and have reduced original sesquyllables to monosyllables.

The /u/ vowel of the first syllables in Munda languages is probably in harmony with the vowel of the second, unstressed syllable.

We now have the following nine branch-level probable reconstructions for Proto-AA #rəŋko:ʔ:

Proto-Pearic *rək <sup>h</sup> o:	Proto-Khmeric *rəŋkə:	Proto-Katuic: *rəŋkau
Proto-Vietic: *rəŋko:ʔ	Proto-Aslian: *raʔkuəʔ	Proto-Khmuic: *rəŋkǝʔ
Proto-Palaungic *rəŋko:ʔ	Proto-Khasian: #rək <sup>h</sup> aw	Proto-Munda: #ruŋkuʔ

The consonantal segments, \*r and \*k do not present serious problems for reconstruction, and the loss of \*-əŋ- in unstressed position is unsurprising and probably took place several times independently. This is not where the problems are; it is the rime that contains unexplained differences among most branches. Not only are the vowels different, but more importantly, the laryngeal features disagree. Notice for example that the glottalised phonation of P-Vietic \*rəŋko:ʔ differs from the non-glottalised register of Pearic \*rəŋk<sup>h</sup>o:, and of Katuic \*rəŋkau. One explanation could be borrowing from one branch to another. But we need to keep in mind the enormous geographic area of distribution of this word and these languages. Also, it has often been remarked that Austroasiatic languages have very large vowel systems with a complicated history. This is not due simply to the inherent flexibility of vowels; it is due, in my opinion, to the fact that vowels and laryngeal features have been intricately involved in various configurations ever since P-AA times. This is where the meat of Austroasiatic history is to be found, and the etymon for ‘husked rice’ still has a long history to tell us.

In conclusion, we can say that the Proto-Austroasiatic language had a word whose phonological shape was approximately \*rəŋko:ʔ, and that it referred to the processed seeds of a graminea, possibly rice. We could also add that the processing in question probably included pounding, mortars, pestles and winnowing, objects and actions for which the proto-AA language also had specific lexical items. But this would require further justification and detailed history.

### 3.3 Taro and rice plants

The name of the plants used in ancient diets is also a large domain for future research; here I will bring up only one hypothesis which has attracted some attention.

Ferlus (1996) proposed that a Palaungic-Khmuic word referring to the taro plant, generally sounding like /sroʔ/, was borrowed by Katuic, Khmeric and Monic in a form sounding like /srɔ:/ and referring to the rice plant. No details are given as to precisely where or when this would have happened, what language group(s) would have shifted the meaning, separately or jointly, or what the phonetic forms might have been at the time of borrowing on either side of the transaction. There is however acknowledgement that the vowel qualities do not fully agree.

Given this vagueness, the proposal cannot be assessed in a direct manner. But one can see two sets of difficulties with it: one has to do with the lexical distribution of these etyma among the Austroasiatic languages, the other concerns their phonological history. A more likely scenario can also be proposed for these important plant-names.

### 3.3.1 Lexical history of rice versus taro

Lexically, the neat distribution between Palaungic-Khmuic on one side, where the word means ‘taro’, and Katuic, Khmeric and Monic on the other where it means ‘rice’ does not hold.

There is one Khmuic language, Thai-Then, spoken in the Luang Prabang area, where we find a word /sroʔ/, in the expression /bat sroʔ/ where /bat/ means ‘grass’, referring not to the rice plant but to another graminea ‘growing near mountain streams, having a panicle similar to that of rice, but eaten only by animals’, perhaps *Aria arundinacea* or some other grass-like plant. In addition, the Thai-Then term for ‘taro’ is /cro:/, without contest the cognate to the word for ‘taro’ in the rest of Khmuic. This is puzzling because if we follow Ferlus’ hypothesis, Thai-Then /sroʔ/ for the grass would then need to have been borrowed back from either Monic, or Khmeric, or Katuic, all located far away from the heartland of Khmuic where Thai-Then is located, and at some time after the alleged shift of meaning from ‘taro’ to ‘rice’ would have taken place.

Not impossible, but complicated. This word immediately evokes another possibility: Thai-Then /sroʔ/ ‘a graminea sp.’ could represent not a reappropriated doublet, but a distinct etymon, etymologically independent from Thai-Then /cro:/ ‘taro’, and a true cognate with the Monic, Khmeric and Katuic words for the rice-plant, not returned by them. In that scenario, no borrowing and no meaning shift from ‘taro’ to ‘rice’ would have taken place.

Lexical support for this second hypothesis is easily found in Khasian and Munda, and possibly even in Bahnaric.

Khasian has an etymon: Standard Khasi /re:w/, Langrin /raə/, which refers to another graminea, *Coix lacryma-jobi*; the Bhoi (Tyrso) cognate /saru:/ refers to maize, and the Pnar: /saru/ refers to yet another graminea, used in making brooms; all these forms reconstruct to Proto-Khasian \*saru:, and this reconstruction is confirmed by Amwi: /hərəu/ and Lakadong: /hərəu/ both of which also refer to *Coix lacryma-jobi*.

This Khasian \*saru: is most probably cognate to the Thai-Then /sroʔ/, and to the Monic, Khmeric and Katuic forms, if we accept a scenario with two distinct etyma without borrowing.

By contrast with \*saru: ‘Coix’, the Khasian forms for ‘taro’ are systematically different and equally well attested throughout Khasian: Standard /ʃre:w/, Langrin: /c<sup>h</sup>aə/, Bhoi (Tyrso) /c<sup>h</sup>ru:/, Amwi /c<sup>h</sup>rəu/. It is not possible to confuse the two plants or to derive one name from the other in Khasian.

One could argue that words for the Khasian gramineae were also borrowed in the same way and from the same source as the Monic, Khmeric, Katuic forms were; but this time the meaning shift would be from ‘taro’ to several other gramineae, complicating the semantic part of the borrowing scenario.

Evidence from the Munda sub-family argues again in the same direction as the Khasian evidence does: Mundari has /huɽu/ for the ‘rice-plant’ and /saɽu/ for ‘taro’, and Santali has /huɽu/ or /hoɽo/ for ‘rice’ and /saru/ for ‘taro’. Here, geography makes it difficult to follow Ferlus’ argument: the evidence of two distinct and ancient etyma with systematically different meanings becomes obvious. Unfortunately, the South Munda evidence is non-committal on this point: Sora apparently has a single phonological form /saro:/, but with both meanings: either taro or rice; the two are not confused but kept apart by suffixes: /-n/ for ‘rice’: /-gai-ən/ for ‘taro’, a probable case of two etyma that have merged phonologically but have been kept distinct by other means.

The Bahnaric evidence goes again in the same direction: Alak has /harɔ:/ for rice, and /trau/ for taro; and Bahnar even introduces a morphological element where /hədrɔʔ/ is a species of slightly early rice; the Bahnar /hədr-/ initial goes back historically to \*hənr-, with an -n- infix which was evidently not present in the Palaungic-Khmuic form if borrowing is the explanation; the geographic route of the alleged borrowing now becomes more difficult to follow, though, here again, not impossibly so. Alak and Bahnar both have a distinct word /trau/ for the taro plant.

### 3.3.2 Phonological history of rice vs. taro

For Palaungic-Khmuic the proto form for ‘taro’ would have a form \*cro:ʔ, with a \*cr-onset required in Palaungic by Wa /kraoʔ/ and U: /cʰrúʔ/, and in Khmuic by Yuan Khmu’ /cróʔ/ and Khabit /cəɔ:/, and with a proto \*-o:ʔ rime, just as in the etymon for ‘husked rice’ discussed earlier. The three branches Monic, Khmeric and Katuic that are alleged to borrow this word with a meaning ‘rice’ would all need as an input a word with a \*sr- onset, implying that by the time the borrowing took place, a shift from \*cr- to \*sr- in the donor language had already occurred. Such a change has occurred several times independently in languages of the Palaungic-Khmuic super-branch, so there is no serious problem here, just a question of timing the borrowing at a date later than Proto-Palaungic-Khmuic, later in fact than the proto of either branch, not a very ancient period.

There are also problems in the rime: the three borrowing branches would need to borrow the word with quite different rimes, each in their own way: Monic needs an \*-o:ʔ rime, which the alleged Northern donor can provide, but Khmeric needs an \*-u: open rime, and Katuic would need a glottalised rime with a more open vowel:<sup>41</sup> \*-ɔ:’. It then becomes difficult to imagine a relatively late Palaungic or Khmuic donor language with an \*-o:ʔ rime providing such diverse rimes to the borrowing branches where this diversity is, in each receiving branch, quite old. The rime for ‘rice’ in Monic, Khmeric, and Katuic is not just old in each of the three branches, it differs in ways which are historically normal.<sup>42</sup> The form can therefore be reconstructed as far back as the proto-node of these three branches: proto Nico-Vietic (proto South-MK plus East-MK); the alleged borrowing would then have to take place at one point of this very early period, not only for this reason but also because the unusual semantic shift from ‘taro’ to ‘rice’ can be expected to occur once, but hardly several times independently. So, the borrowing hypothesis is caught here in a kind of time warp, a late Northern form being lent to a very ancient proto-language.

On the other hand, the two-etyma hypothesis has no such problem: the proto-Nico-Vietic etymon ‘rice’ is ancient, and it would actually have even more ancient cognates in Khasi-Khmuic and in Munda, as shown above, making it a P-AA etymon. The proto-meaning would not be the rice-plant itself, for which there are better candidates, for example \*(-)ɬa:ʔ, but probably a similarly useful grass, cultivated or not.

It would also avoid supposing the existence of a recently appeared and now defunct Northern-AA language with ancient and widespread influence as a cultural lender, that

41 An open rime with a glottalised register would be required to account for Chatong /harɔ:/ and Ta-oih (Talan) /srɔʔ/ ‘the rice plant’.

42 See for example the etymon ‘grass, *Imperata cylindrica*’ where we have: Monic \*cwo:ʔ, Khmeric \*spu:, and Katuic \*srəmpɔ:’ with the same rime-correspondances as in the word for the rice plant.

would have left no other trace than this odd meaning shift from ‘taro’ to ‘rice’.<sup>43</sup> The old Occam, and his legendary razor, has to be invoked here.

There will be room for many other hypotheses and refutations of this kind for the names and uses of plants in this language family. But at the present stage of AA history we have to admit that our knowledge of lexico-botany remains very superficial, or even absent for most of these languages. The phonological side of the picture has been improving, but progress at the deeper historical levels will be slow, as it is in other language families whose antiquity is similar to that of Austroasiatic.

## References

- Baradat, R. 1941. Les Samrê ou Péar, population primitive de l’ouest du Cambodge. *Bulletin de l’Ecole Française d’Extrême-Orient* 4,1–150.
- Banker, John, Elizabeth Banker and Siu Mo’. 1979. *Bahnar Dictionary, Plei Bong-Mang Yang Dialect*. Huntington Beach, Summer Institute of Linguistics.
- Barker, Milton E. and Muriel A. Barker 1976. *Muong-Vietnamese-English Dictionary*. Dallas, Summer Institute of Linguistics (microfiche).
- Chosri, Isara. 2002. Dialects of Chong. *The Mon-Khmer Studies Journal* 32. 55–70.
- 2009. Chung (Saoch) of Thailand and Cambodia: phonological and lexical comparisons. *MKS* 38, 69–85.
- Campbell, L., 1998. *Historical linguistics: an introduction*. Edinburgh University Press, Edinburgh.
- Costello, Nancy A. 1971. *Ngũ-vang Katu* 12. Department of Education.
- Diffloth, Gérard. 1984. *The Dvaravati-Old-Mon language and Nyah Kur*. Bangkok: Chulalongkorn University.
- 1989. Proto-Austroasiatic creaky voice. *Mon-Khmer Studies* 11. 47–82.
- 1991. Palaungic vowels in Mon-Khmer perspective. In Jeremy H.C.S. Davidson (ed.) *Austroasiatic Languages, Essays in honour of H. L. Shorto*, 13–28. School of Oriental and African Studies, University of London.
- 2005. The contribution of linguistic paleontology to the homeland of Austroasiatic. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The peopling of East Asia: putting together archeology, linguistics and genetics*, 79–82. London: Routledge Curzon.
- Enfield, N. J. and Gérard Diffloth. 2009. Phonology and sketch grammar of Kri, a Vietic language of Laos. *Cahiers de Linguistique–Asie Orientale* 38(1). 3–69.
- Ferlus, Michel. 1974. Lexique souei-français. *Asie du Sud-Est et Monde Insulindien* 5.1, 141–159.
- 1976. *Les Traits Sémantiques Dans Le Vocabulaire Phou Noy*.
- 1991. *Vocalisme du Proto-Viet-Muong*. Paper circulated at the Twenty-fourth ICS-TLandL. Chiang Mai University, Oct. 10–11, 1991.
- 1996. Du taro au riz en Asie du Sud-Est, petite histoire d’un glissement sémantique. *Mon-Khmer Studies* 25. 39–49.

---

43 The odd semantic shift would also have to occur in a language where all its Northern relatives already have an ancient word for the rice plant: Proto Palaungic-Khmuic \*sɲɔːʔ, which was pervasive in this super-branch and even borrowed by Mon in medieval times with the meaning ‘unhusked rice’, a word which does not reconstruct to Proto-Monic.

- . 2004. The origin of tones in Viet-Muong. *Papers from the Annual Meeting of the South East Asian Linguistic Society*, 297–313. Arizona State University.
- Hojjer, Harry. 1956. The chronology of the Athapaskan languages. *International Journal of American Linguistics* 22(4). 219–232.
- Haudricourt, André-Georges. 1953. La place du viêtnamien dans les langues austroasiatiques. *Bulletin de la Société de Linguistique de Paris* 49.1, 122–128.
- Huffman, Franklin E. 1985. The phonology of Chong. In Suriya Ratanakul, David Thomas and Suwilai Premsirat (eds) *Southeast Asian Linguistic Studies presented to André-G. Haudricourt*, 355–88. Bangkok: Mahidol University.
- Jacq, Pascale and Paul J. Sidwell. 2000. *A Comparative West Bahnaric Dictionary*. LINCOM Europa.
- Jenner, Philip N., 1976. Les noms de nombre en Khmer. *Linguistics* 174. 39–60.
- Lemaréchal, A. 2010. *Comparative grammar and typology: Essays on the Historical Grammar of the Austronesian Languages*. Orbis/Supplementa 35. Leuven: Peeters.
- Lợi, Nguyễn Văn, 1993. *Tiếng Rục (The Ruc language)*. (In Vietnamese.) Hanoi: Social Science Editions.
- Miller, John D. 1976. *Bru Vocabulary*. Summer Institute of Linguistics, Huntington Beach.
- Peiros, Ilia. 1996. *Katuic Comparative Dictionary*. Canberra: Pacific Linguistics.
- Rischel, Jørgen. 1995. *Minor Mlabri: A Hunter-Gatherer Language of Northern Indochina*. Museum Tusulanum Press, Copenhagen.
- Roepstorff, F. A. de 1884. *Dictionary of the Nancowry Dialect of the Nicobarese language*. Calcutta.
- Sauvageot, Aurélien. 1951. Review of Morris Swadesh. 1950. ‘Salish internal relationships’ IJAL. *Bulletin de la Société de Linguistique de Paris* 47/2. 25–8.
- Shorto, Harry L. 1971. *A dictionary of the Mon inscriptions, from the sixth to the sixteenth centuries, incorporating materials collected by the late C.O. Blagden (London Oriental Series 24)*. London, Oxford University Press.
- . 2006 (posthum.). *A Mon-Khmer Comparative Dictionary*. In Paul Sidwell (main ed.), Doug Cooper and Christian Bauer (asst. eds) Canberra: Pacific Linguistics .
- Sidwell, Paul. 1999. *A Reconstruction of Proto-Bahnaric*. PhD thesis, University of Melbourne.
- . 2005. *The Katuic Languages: classification, reconstruction and comparative lexicon*. Munich, Lincom Europa.
- . 2009. The Austroasiatic Central Riverine Hypothesis. Keynote address read to the 19th annual meeting of the Southeast Asian Linguistic Society. University of Social Sciences, Ho Chi Minh City, May 2009.
- Sokolovskaja, N.K. and Nguyễn Văn Tai, 1987. *The Muong Language: Materials from the Soviet-Vietnamese Expedition of 1979*. (In Russian) Moskva, Nauka.
- Sriwises, Prasert 1978. *Kui (Suai)-Thai-English dictionary*. Bangkok, Chulalongkorn University.
- Svantesson, Jan-Olof 1988. U. *Linguistics of the Tibeto-Burman Area* 11: 64-133.
- Thongkum, Theraphan. L. 1991. An instrumental study of Chong registers. In Jeremy H.C.S. Davidson (ed.) *Mon-Khmer studies in honour of Harry Shorto*, 141–160. London.
- . 2001. *Languages of the Tribes in Xekong Province Southern Laos*. The Thailand Research Fund. Chulalongkorn University, Indigenous Languages of Thailand Research Project.

- Thongkum, Theraphan. L., and Peungpa, 1980. *A Bruu - Thai - English Dictionary*, Chulalongkorn University, Bangkok.
- Thomas, David D. and Robert K. Jr. Headly. 1970. More On Mon-Khmer Subgroupings. *Lingua* 25. 398–418.
- Watson, Richard, Sandra K. Watson and Cubuat. 1979. *Pacoh Dictionary: Pacoh-Vietnamese-English*. Trilingual Language Lessons, No. 25, part 1. Manila, Summer Institute of Linguistics.
- Whitehead, George. 1925. *Dictionary of the Car-Nicobarese language*. Rangoon, American Baptist Mission Press.





# 14 *The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis*

---

PAUL SIDWELL AND ROGER BLENCH

## 1 Introduction<sup>1</sup>

The Austroasiatic language phylum is situated in the heartland of MSEA and yet today is remarkably fragmented, its individual branches scattered from Northeast India to the Malay Peninsula. Interwoven territorially with much more geographically coherent phyla such as Daic and Hmong-Mien, the narrative of its dispersal is central to our general understanding of the ethno-cultural history of Southeast Asia. Although comparative Austroasiatic linguistics is now more than a century old, limited progress has made towards a consensus on the homeland or *Urheimat* of Austroasiatic languages. The Austroasiatic phylum is generally considered to be the oldest identifiable language grouping of that region (excluding perhaps Andamanese). A model for its origins and migration paths that could account for the present distribution of the languages is crucial for the linguistic history and ethnography of Southeast Asia. However, the linguistic literature relating to this too often presents confident claims that invoke unpublished materials, paying little heed to evaluating alternative hypotheses. Published studies are not always transparent, especially problematic when they lack adequate data that readers might assess and analyse for themselves.

In this chapter we focus on linguistic arguments for a likely Austroasiatic homeland, and possible correlations with the—still emerging—archaeological record. The orientation of the chapter is linguistic; we assume no necessary equation between linguistic entities and archaeological assemblages or genetic profiles. It is evident that the ethno-history of Southeast Asia has often involved multilingualism and various radical language shifts among communities large and small, and this must also have occurred among prehistoric communities, especially in the context of the early expansion of agriculture. Consequently, when we talk about a linguistic homeland or *Urheimat*, we do not wish to imply that the cultural complex which radiated from that centre necessarily largely *originated* in that location, only that it began to diversify and spread from there. In other words, it is the last location in which the speaker community presented a linguistic unity.

Among the various suggestions for the Austroasiatic centre of dispersal offered over the years, there are three broad trends:

---

<sup>1</sup> Acronyms and conventions used in this chapter: C (any consonant), V (any vowel), # (quasi-reconstruction, i.e. form based on rapid inspection of cognates), MSEA (Mainland Southeast Asia).

N. J. Enfield, editor.

*Dynamics of Human Diversity*, 315-343. Pacific Linguistics, 2011

© This edition vested with Pacific Linguistics.

1. A western origin, in northeastern India or in the vicinity of the Bay of Bengal (van Driem 2001)
2. A northern origin, in central or southern China (Norman and Mei 1976)
3. A central origin, within Southeast Asia (Von Heine-Geldern 1923)

None of these proposals has been supported by sufficient evidence, and there has been a broad failure to correlate the antiquity and location of the homeland of Proto-Austroasiatic with known archaeological data. It is taken as a given that the reconstructibility of lexical items in Austroasiatic related to subsistence must be congruent with current understanding of the archaeology, ecology and palaeoclimatology of the region. This chapter will argue that:

- a) None of the internal sub-families proposed for Austroasiatic have been demonstrated unambiguously and that the provisional model for its internal structure must be a largely flat array.
- b) The widely-accepted division between Muṅḍā and the remainder of Mon-Khmer is spurious. Although Muṅḍā shows significant typological differences from other Austroasiatic languages, this cannot be taken as evidence for the antiquity or primacy of the split.
- c) A flat array in turn points to the diversification of a dialect chain and therefore implies a relatively younger age (ca. 4000 BP) for Austroasiatic than is usually advanced.
- d) The proto-Austroasiatic lexicon has terms for crops requiring a humid climate (taro and rice) as well as other items (boats, river fauna) which suggest an aquatic/riverine environment.
- e) Congruence between these elements can be achieved with the assumption that early Austroasiatic initially diverged somewhere in the Middle Mekong, and its initial dispersal was along river valleys, exploiting both aquatic resources and humid soils.

Ultimately we suggest that various facts concerning Austroasiatic languages can well be explained by convergence and contact that continued after a relatively late break-up into distinct branches. We propose a model to correlate the linguistic results with evidence from archaeology and anthropology, which we call the Southeastern Riverine Hypothesis. This model has precursors, first advanced nearly a century ago. Von Heine-Geldern (1923), anthropologist and student of Wilhelm Schmidt (founder of comparative Austroasiatic studies) advanced a *Kulturkreise* theory in the 1920s which modelled the dispersal of Austroasiatic out of Southeast Asia. In the 1970s, on the basis of lexicostatistical studies (for example Thomas 1973, Huffman 1978), it was suggested that the phylum dispersed from Indo-China/Northeast Thailand as recently perhaps as 3800 BP.

## 2 General considerations

The three basic proposals for the Austroasiatic homeland have been advanced on the bases of the following types of argument:

- a) lexical isoglosses and/or typological affinities interpreted as indicating contact, and thus proximity, of otherwise unrelated language families

- b) correlation of lexical reconstructions with archaeological facts/hypotheses or features of the natural world ('paleo-linguistics')
- c) 'centre-of-gravity' arguments that identify supposed zones of higher linguistic diversity with greater time depth

Lexical and/or structural similarities between Austroasiatic and other languages of insular and mainland Asia are readily found (Enfield, this volume). This has been the source of a series of genetic hypotheses relating it to other phyla. For example, Shorto (1976a) argued that Austroasiatic-Austronesian comparisons demonstrate the validity of Austric, a hypothesis which goes back at least to Schmidt (1905). Blench (this volume) discusses the history of various macrophyla hypotheses in more detail, but we believe that none of them has been reliably demonstrated and they will not be further considered here.

Another source of quasi-genealogical hypotheses has been extensive lexical borrowing between phyla. It has long been argued that Austroasiatic was once much more widespread in China and was driven south by the expansion of the Han (Norman and Mei 1976). It is claimed that some names of zodiacal animals, and the Old Chinese words for 'river' and 'tiger' are borrowings (Norman 1988:18). Schuessler (2007) expanded this in his *Etymological dictionary of Old Chinese* which asserts that Austroasiatic forms underlie many Sinitic etyma. The subtext is a general identification of northern regions as the homeland of Austroasiatic, but these authors fail to note that many roots are typically widespread in the region occurring in multiple phyla. For example, tigers were historically common across the region, although today they are confined to a few small reserve areas. Tigers have a crucial role in spiritual beliefs of many peoples, which may account for the distribution across language phyla of a key lexeme, #kVla. Table 1 shows the regional reflexes of #kVla.

It is clear that the name for 'tiger' has been freely borrowed between phyla and is apparently ancient in Sino-Tibetan, Daic and Austroasiatic. Such words cannot be used in genealogical classifications and certainly not in arguments about the location of homelands.

Along similar lines, it has been widely claimed that, especially since Norman and Mei (1976), the name of the Yangtze itself is of Austroasiatic origin. This is based on the casual resemblance between Old Chinese \*k<sup>ʰ</sup>roŋ<sup>2</sup> and Austroasiatic forms such as Old Mon *kruñ* /*kruŋ*/ 'river'. While suggestive, no compelling reasons have been put forward to show that this is anything other than a lexical coincidence, such as one may find comparing any pair or group of languages. In a similar vein van Driem (2001:290) reminds us that 'Toponyms and especially river names [...] have suggested to researchers such as Hermann Berger and Manfred Mayrhofer that Austroasiatic is an old ethnic substrate in the north of the Indian subcontinent.' Despite these claims, no specific comparisons have yielded a decisive body of isoglosses. In the Indian context, it is often no more than the claim that various Indic words appear to show prefixes, which is also an Austroasiatic characteristic, although without even suggesting that the forms have specific Austroasiatic cognates (see for example Witzel 1999).

---

2 As presently reconstructed by Baxter and Sagart, see:

[http://sitemaker.umich.edu/wbaxter/old\\_chinese\\_reconstructionsandrecord](http://sitemaker.umich.edu/wbaxter/old_chinese_reconstructionsandrecord)

**Table 1:** The #kVla root for ‘tiger’ in Southeast Asian languages

Phylum	Branch	Language	Attestation
Sino-Tibetan	Sinitic	OCM	*hlâ?
Sino-Tibetan	Burmic	Old Burmese	klya
Sino-Tibetan	Loloish	Mantsi Meo Vac	qua <sup>53</sup>
Sino-Tibetan	Luish	Cak	ka θa
Sino-Tibetan	Luish	Lushai	sa kei
Sino-Tibetan	Bodish	Monpa	khai-la
Sino-Tibetan	Naga	Mao Naga	okhe
Austroasiatic	Palaungic	Shinman	kaʔ <sup>4</sup> vai <sup>3</sup>
Austroasiatic	Khmeric	Angkorian Khmer	khlaa
Austroasiatic	Bahnaric	Sedang	klá
Austroasiatic	Katuic	Pacoh	kulaa
Austroasiatic	Katuic	Ir	kala?
Austroasiatic	Katuic	So	kula
Austroasiatic	Monic	Proto-Monic	*klaa?
Austroasiatic	Aslian	Sakai	kla
Austroasiatic	Khasian	War Jaintia	k <sup>h</sup> la
Austroasiatic	Muṅḍā	Muṅḍā	kula
Daic	Tai	Thai	k <sup>h</sup> aan (Year of ~)

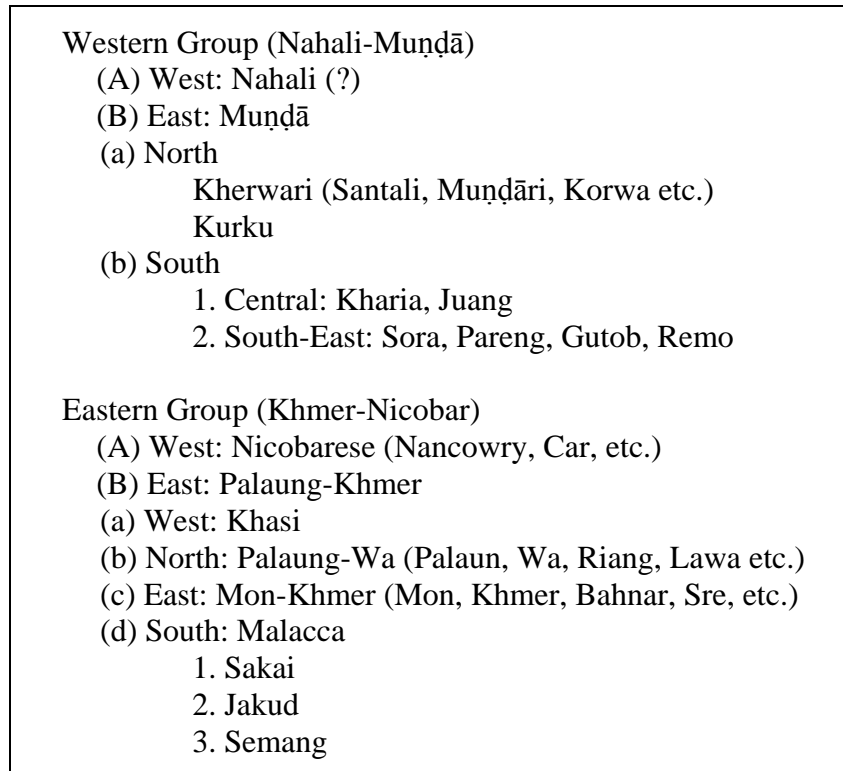
Scholars have also cited characteristics of the palaeo-environment when seeking the likely homeland. Diffloth (2005) observed that since names for various tropical species can be reconstructed for the proto-language, a high humidity ecology is indicated. Hence the suggestion of the shores of the Bay of Bengal urged by van Driem (2001:290). Yet Sagart (2008, citing Chang Kwang-chih 1986) notes that the mid-Holocene climate of central China was 2° to 5° warmer than today. Consequently, the potential zone of tropical flora and fauna likely encompassed any and all proposed Austroasiatic homelands. By contrast, Peiros (1998) and Peiros and Shnirelman (1998) assert that the Austroasiatic lexicon indicates a non-tropical, non-coastal location, but present no evidence to support such a claim.

There is also a related tendency to assert that Austroasiatic is of great antiquity. Peiros (2004) offers a glottochronological calculation of 8300 BP for the initial branching. Diffloth (2005) proposes approximately 7000 BP, and likewise Blust (1996:132) estimates that ‘By 7000–7500 BP, PAA had separated into western (Muṅḍā) and eastern (Mon-Khmer) dialect areas’. Now it seems the antiquity of Austroasiatic falls into the category of facts so well known it can be repeated without qualification or explanation (for example van Driem 2007:10). Since for each of these authors the homeland was located somewhere significantly removed from where the majority of Austroasiatic languages are spoken today, by necessity a long period of time must have elapsed for the languages to disperse over a wide area, and for others (Sinitic, Indo-Aryan, etc.) to occupy the original site.

A common feature of the diverse claims for, variously, China or South Asian origins, is that they are vague; failing to make predictions that may be readily tested. Similar types of evidence, such as resemblances in place names, are invoked to support mutually incompatible models. In response we claim that it is better to begin with reviewing what we know about the Austroasiatic languages, and seek the simplest explanation (the ‘fewest moves’) which may account for those facts. Such a model should be the starting point for discussion of Austroasiatic linguistic origins.

### 3 The Received Classification of Austroasiatic Languages

Current thinking about the classification of the branches of Austroasiatic can be traced to the comparative and typological studies of Pinnow (1959, 1960, 1963). While earlier studies (for example Grierson 1919; Przyluski 1924) treated Muṅḍā ('Kolarian') as a distinct division, scholars in the latter part of the 20th century relied directly upon Pinnow for making explicit the notion of a binary split between Muṅḍā and Mon-Khmer. Pinnow's (1963:278) scheme of a Western group (Nahali-Muṅḍā) and Eastern Group (Khmer-Nicobar) is reproduced in Figure 1.



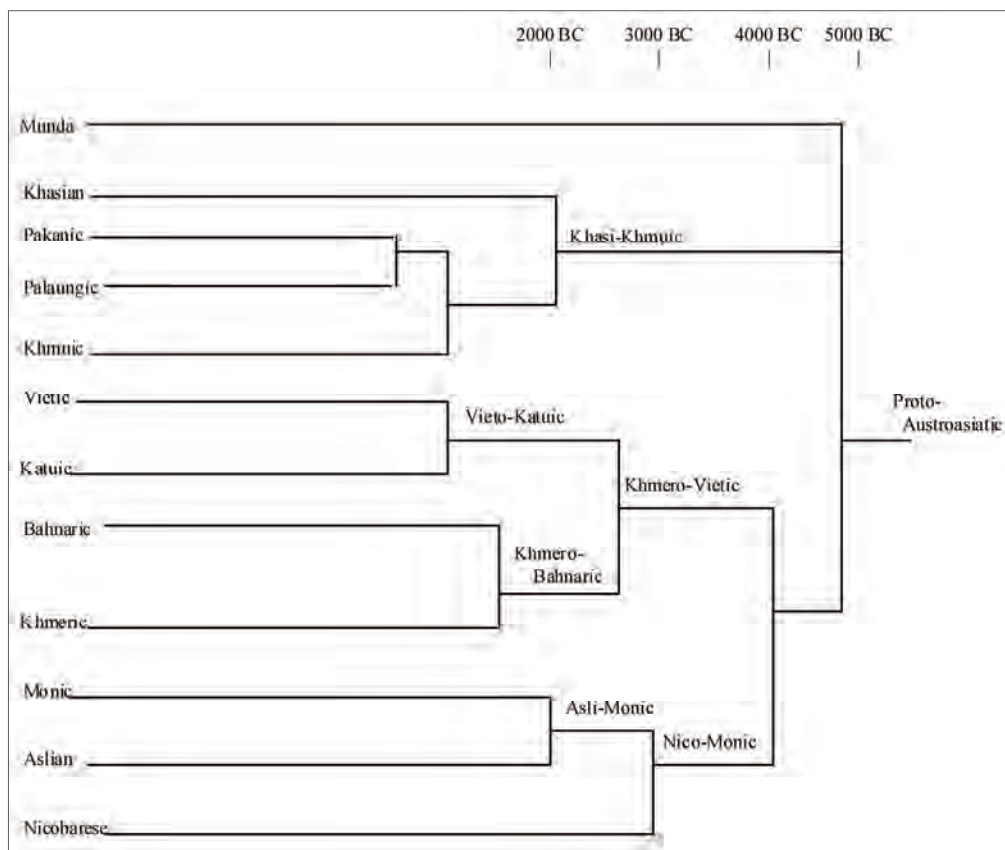
**Figure 1:** Pinnow's Austroasiatic classification.

Nahali is now recognised by almost all scholars to be a language isolate, albeit one that has come under Muṅḍā (Korku) influence (Blench 2008).

Pinnow (1963:150) recognised the limitations of the typological approach as a means of discerning historical relations. He was not especially confident that the 'Khmer-Nicobar' languages formed a unity in the same way as Muṅḍā, and stated that they may even be historically 'independent of one another and traceable solely to Proto-Austroasiatic'. In this case Khmer-Nicobar 'would have only structural and geographical justification'. But this caveat was almost universally ignored in favour of a genealogical reading. Before Pinnow, there had been no clear guide to the likely structure of the family; suddenly we could speak with confidence and authority about Austroasiatic, and a new orthodoxy was adopted. Since then a primary Muṅḍā Mon-Khmer split has largely been assumed. For example, recently we find: 'The primary split in the family is between the Muṅḍā languages in central and eastern India and the rest of the family.' (Anderson 2008:598) And with a twist: 'The Austroasiatic language family is conventionally divided into three branches or sub-families, viz. the Muṅḍā, the Nicobarese and the Mon-Khmer languages.' (van Driem 2001:262).

Programmatically the desirable approach would have been to build on Pinnow (1959) and other comparative studies (Schmidt 1904, 1905, Skeat and Blagden 1906, Schafer 1952, 1965, Haudricourt 1965, etc.) to produce a working model of proto-Austroasiatic in the 1960s–70s, so that a classification based upon shared innovations could have been discussed. Shorto attempted this, but only published a fragment at the time (Shorto 1976b).<sup>3</sup> Diffloth began the task but has not released his results. Consequently the field turned to other strategies, specifically lexicostatistics. Lexicostatistics is a widely used heuristic for revealing likely language relationships, and is especially favoured when one is more or less restricted to using lexical data (cf. Burenhult, Dunn & Kruspe this volume). The method counts shared forms on a limited basic wordlist to produce a crude index of similarity. The method has been strongly criticised, especially in respect of studies that have placed a high value on lexicostatistical results for creating family trees, and the present writers share those concerns.

Notwithstanding the poor image and well understood limitations for the method, it was lexicostatistical studies that distinguished the dozen or so Austroasiatic branches recognised today, and confirmed by subsequent comparative analyses. The most important early lexicostatistical study was by Thomas and Headley (1970). They concluded (p. 405): ‘The Austroasiatic phylum would appear to be composed of at least four families: Muṅḍā, Mon-Khmer, Malacca, Nicobarese.’

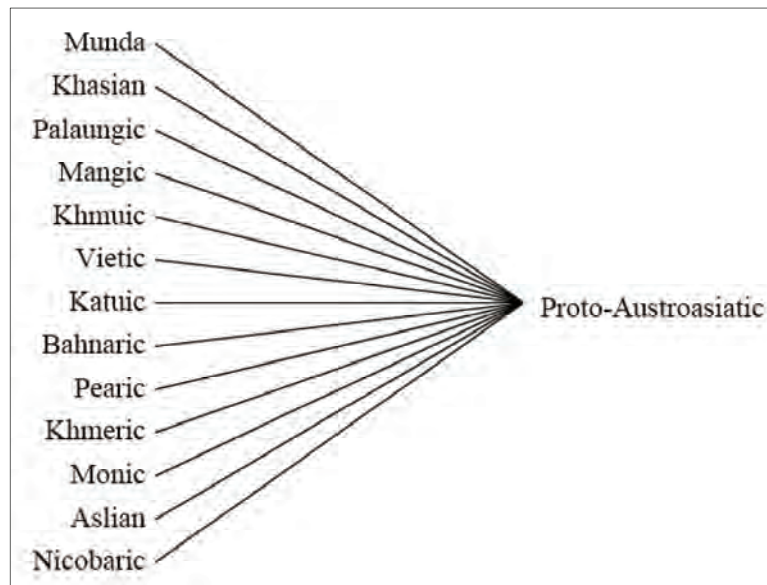


**Figure 2:** Diffloth's model of Austroasiatic (adapted from Diffloth 2005).

3 Shorto pursued a comprehensive comparative reconstruction, which was published posthumously in 2006. In the course of that effort he was unable to justify any nested sub-grouping of AA branches by historical phonology, and instead conducted several lexicostatistical investigations (at least four attempts) which are discussed in Sidwell (2009).

Concretely, the separation of Malacca (Aslian) and Nicobarese from Mon-Khmer was shown by cognate scores of 11~16% between Temiar and Mon-Khmer, and 6~12% between Nicobarese and Mon-Khmer. Above these scores varying between 18% and 35% were taken as indicating nine distinct Mon-Khmer branches: Pearic, Khmer, Bahnaric, Katuic, Khmuic, Monic, Palaungic, Khasi, and Viet-Muong. Another widely cited source (Parkin 1991:6) presents Thomas and Headley's scheme, quoting their lexicostatistical figures, but credits it to Diffloth, saying, 'This breakdown is based mainly on Diffloth, though he included Aslian in with Mon-Khmer.' However, a general view later emerged that the very low percentages found for Aslian and Nicobarese are caused by special factors, and that these are Mon-Khmer groups isolated from the rest of the family. The notion that a similar explanation might also account for Muṅḍā has also surfaced recently.

The present situation is that there are two main competing classifications of Austroasiatic, that given by Diffloth (2005) which can be contrasted with Sidwell (2008) shown in Figure 3. Figure 2 shows a version of Diffloth (2005) with sub-branches omitted so that it can be directly compared with Figure 3. Diffloth proposed dates attached to individual nodes, but these are only applicable to the earlier period once individual sub-branches are merged.



**Figure 3:** Flat-array structure for Austroasiatic (Sidwell 2008).

Diffloth has not offered a general rationale for his scheme, only fragments. For example, he (personal communication)<sup>4</sup> has suggested that the metathesis of proto-Austroasiatic *\*pti(i)s* 'mushroom' to Khmer *psət*, Bahnaric, for example Sre *bsit*, and Pearic, for example Chong *psiit*, must be a unique historical accident indicative of common ancestry. But subgrouping arguments that hinge on a single form rather than a collection of common but unrelated events must admit various possible explanations. After all, if Khmer, Bahnaric and Pearic did share a unique period of common development, we would reasonably expect to find indications in the core vocabulary, much as can be found, for example, in respect of Khasi-Palaungic (discussed below). But one or even several isoglosses found amongst thousands of comparanda cannot be a convincing sub-grouping argument.

4 Diffloth also briefly discussed this at the 2008 SEALS meeting in Malaysia.



	Mundā		Khasian		Palaungic		Khmuic		Vietic		Katuic		Baharic		Monic		Pearic		Khmeric		Aslian		Nicobarese				
	Mun	Sor	Khs	War	Dea	Wa	Kmu	Mal	Muo	Ruc	Bru	Nge	Kat	Bah	Jeh	Jru	Mon	Nya	Cho	Pea	Sur	Cam	Sem	Sml	Jah	Car	Nan
Mundāri	37		18	13	15	17	18	22	21	23	25	24	22	17	20	23	24	23	18	20	16	16	16	16	11	10	15
Sora		37	14	10	13	14	20	20	21	24	23	19	19	17	18	18	21	20	11	15	12	13	17	13	12	8	12
Khasi	18	14		60	23	20	22	26	22	21	24	24	23	21	23	22	21	20	19	20	20	21	20	15	16	13	16
War-J	13	10	60		17	16	16	24	14	14	19	21	17	13	16	16	16	15	16	18	17	17	17	14	10	8	10
De'ang	15	13	23	17		53	26	20	15	18	19	22	22	19	24	17	23	24	16	19	17	18	20	14	14	12	14
Wa	17	14	20	16	53		26	22	16	19	26	26	26	20	26	21	23	26	22	27	17	18	21	18	16	13	18
Khmu'	18	20	22	16	26	26		35	17	22	27	26	24	20	21	21	21	23	21	25	18	19	21	16	15	12	14
Mal	22	20	26	24	20	22	35		18	23	31	31	28	22	27	22	21	24	19	25	21	21	22	20	17	12	13
Muong	21	21	22	14	15	16	17	18		50	27	23	25	24	22	25	22	22	15	18	14	14	20	16	14	11	14
Ruc	23	24	21	14	18	19	22	23	50		30	29	27	26	27	30	23	23	17	20	19	20	20	18	14	14	17
Bru	25	23	24	19	19	26	27	31	27	30		57	57	41	40	43	28	31	30	34	23	24	29	28	20	17	21
Nge'	24	19	24	21	22	26	26	31	23	29	57		61	38	36	38	30	33	30	34	27	26	25	31	23	18	20
Katu	22	19	23	17	22	26	24	28	25	27	57	61		43	39	40	29	30	30	34	26	26	28	28	20	16	19
Bahnar	17	17	21	13	19	20	20	22	24	26	41	38	43		52	51	23	26	30	34	24	22	29	30	18	13	19
Jeh	20	18	23	16	24	26	21	27	22	27	40	36	39	52		55	23	24	27	27	22	23	31	31	23	14	18
Jru'	23	18	22	16	17	21	21	22	25	30	43	38	40	51	55		27	29	29	33	27	26	27	30	21	15	19
Mon	24	21	21	16	23	23	21	21	22	23	28	30	29	23	23	27	84	23	28	24	23	24	27	23	18	15	18
Nyahkur	23	20	20	15	24	26	23	24	22	23	31	33	30	26	24	29	84	23	28	24	23	24	27	24	19	15	17
Chong-W	18	11	19	16	16	22	21	19	15	17	30	30	30	30	27	29	23	23	61	22	20	22	22	22	11	14	18
Pear-KPT	20	15	20	18	19	27	25	25	18	20	34	34	34	34	27	33	28	28	61	30	28	27	29	18	16	22	
Surin	16	12	20	17	17	17	18	21	14	19	23	27	26	24	22	27	24	24	22	30	89	22	21	15	12	14	
Cambodian	16	13	21	17	18	18	19	21	12	20	24	26	26	22	23	26	23	23	20	28	89	20	19	17	13	15	
Semai	16	17	20	17	20	21	21	22	20	20	29	25	28	29	31	27	27	27	22	27	22	20	38	36	14	19	
Semelai	16	13	15	14	14	18	16	20	16	18	28	31	28	30	31	30	23	24	22	29	21	19	38	28	11	16	
Jahai	11	12	16	10	14	16	15	17	14	14	20	23	20	18	23	21	18	19	11	18	15	17	36	28	12	14	
Car	10	8	13	8	12	13	12	12	11	14	17	18	16	13	14	15	15	15	14	16	12	13	14	11	12	49	
Nancowry	16	12	16	10	14	18	14	13	14	17	21	20	19	19	18	19	18	17	18	22	14	15	19	16	14	49	

Figure 4: Austroasiatic lexicostatistical matrix (Sidwell 2009)

A challenge to Diffloth's model has been put forward by Sidwell (2008), who argues instead for a flat array, rejecting not only the Muṅḍā/Mon-Khmer split, but other proposed internal nodes. This is based on a review of the phonological and lexical correspondences at branch level, which failed to find innovations that would justify the kind of nested sub-branching shown in the model in Figure 2. Sidwell's revised model is shown in Figure 3 with the addition of Mangic (Pakanic in Diffloth's terminology), now also considered a distinct branch.

Blench (in press a) argues that the language of the Shom Pen, foragers in the Nicobar Islands, may also constitute a separate branch of Austroasiatic, but this has yet to be assessed by the scholarly community and will not be included in this analysis.

The flat array suggests that Austroasiatic first fanned out into a dialect chain in which neighbours would have remained intercomprehensible while cumulative differences would have been greatest at the geographical extremes. Such a model has implications for both the dating of Austroasiatic and its likely directions of spread.

#### **4 Sidwell's Lexicostatistics**

Having failed to identify phonological innovations that would support Diffloth's nested branching model, Sidwell revisited the lexicostatistical analysis of Austroasiatic. There were 28 languages compared using the standard Swadesh 100 list and a matrix (Figure 4) generated using the programme *glottpc.exe*.<sup>5</sup> The results show remarkable similarities to those of Huffman (1978) in terms of the relative distributions of percentages. The biggest difference is in the direct Katuic-Bahnaric comparison, with an average cognacy of only 40%, rather than the 47% found by Huffman. This is still high, and all other things being equal, it would be consistent with a Katuic-Bahnaric sub-grouping. Yet comparative reconstruction makes it clear that Katuic and Bahnaric do not sub-group (see discussion in section 8, below).

Both Sidwell's and Huffman's figures show a remarkable pattern, highlighted in the shaded parts of Figure 4. Rather than inter-branch percentages indicating neat patterns of nested sub-branching, distinct branches show elevated scores with Katuic-Bahnaric, but a much flatter spread of scores if Katuic-Bahnaric is removed from consideration. The common factor is geography, with scores declining as one gets further afield from the middle Mekong. There is also a weaker but discernibly similar effect between Mon and Nyah Kur and the rest of Austroasiatic. Among the Aslian languages it is the southern sub-branch that show more cognates with Katuic-Bahnaric (and to a lesser extent Monic). For example, Semai and Semelai show 31% agreement with Jeh (Bahnaric), yet within Aslian a lower score (28%) is counted between Jahai (North Aslian) and Semelai (South Aslian).

How are these figures to be interpreted? Since lexicostatistical methods were first pioneered by Swadesh in the early 1950s (for example 1950, 1952 and *passim*.) they have been consistently savaged by critics who have focussed on unreliability in sub-grouping and dating results (for example Hoijer 1956, Bergsland and Vogt 1962, Holm 2003 and others). Those criticisms focus mainly on the fallacy of a constant rate of change, and the problem of distinguishing inherited from borrowed vocabulary, which we acknowledge here. We offer these lexicostatistical results specifically as a heuristic for diagnosing and investigating borrowing between languages, an otherwise crucial but under-investigated aspect of the Austroasiatic homeland question. We do not attempt to construct a traditional

---

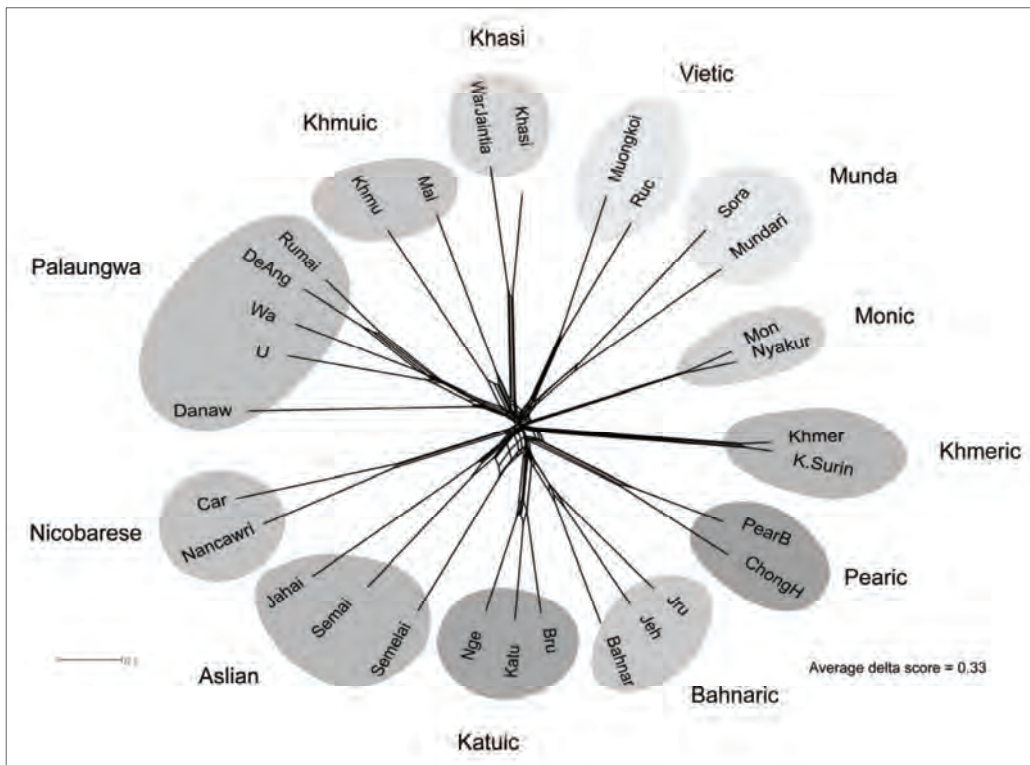
5 The data, cognate assignments, and computational analyses of this and other lexicostatistical studies by Sidwell can be retrieved online at: <http://people.anu.edu.au/~u9907217/lexico>.

lexicostatistical tree, but instead compare the differences in cognate percentages with the indications of historical phonology. Our underlying assumption is that the latter is necessarily the gold standard for determining branching relations, especially in the present case where it is not possible to compare, for example, morphological paradigms of the complexity found in, say, Indo-European. It is clear that our present results, similar to Huffman's results of a generation before, show a clear geographical correlation: cognate percentages are generally higher when compared to languages located in the middle Mekong zone, and higher the closer one gets to that zone. Our interpretation is that the figures are so contaminated by borrowings it is unsafe to posit anything other than roughly equidistant branches on lexical grounds.

At Sidwell's request the lexicostatistical matrix for 28 AA languages was subjected to computational phylogenetic analyses by Russell Gray and Simon Greenhill at the University of Auckland. Subsequently they calculated the Neighbor-Net tree reproduced here in Figure 5. The method is different in kind to a traditional lexicostatistical analysis, and is widely used today in biology and genetics in particular, and increasingly in linguistics (cf. Burenhult, Dunn and Kruspe this volume).

The analysis is strongly consistent with our hypothesis of roughly equidistant branching. The Neighbor-Net shows a strongly tree-like signal with 12 branches unambiguously distinguished. The lexical proximity of Katuic-Bahnaric is found, although one may note the cross-linking lines in the net that indicate signal interference (probably borrowing) between Aslian, Katuic, Bahnaric and Pearic. There is a weak indication of sub-branching between Palaungic, Khmuic and Khasi, and also between Vietic and Muṅḍā. However, these are merely 'best-fits' and have extremely low statistical weight. To put it another way, while a Northern sub-group of Palaungic, Khmuic and Khasi would be an unremarkable—even pleasing—result, it is on the basis of this data no more likely than a Muṅḍā-Vietic sub-branch, a geographically absurd prospect.

Both our interpretations of the matrix (Figure 4) and the phylogenetic analysis would be unremarkable if the languages formed a geographic contiguity, such that borrowings could readily propagate through the speaker community. Yet Austroasiatic branches are today distributed in discontinuous pockets, sometimes hundreds of kilometres apart. The especially high agreement between Katuic and Bahnaric can be explained by ongoing borrowing, since they were never really separated, but the same is not evident for the elevated cognate scores they share with Monic, Pearic, Palaungic, Khmuic, Khmer, and even South and Central Aslian. We suggest that in the initial stages of its dispersal, Austroasiatic had the character of a contiguous dialect chain, along which borrowing could readily spread. As various branches became more isolated, the mechanism of lexical convergence ceased to function, and lexical change would then have been driven by internal and novel external factors. The centre of that chain was located on the middle Mekong, with the most northerly and southerly extremities ultimately becoming the Muṅḍā and Nicobaric branches respectively (see Figures 6 and 7 for a broad representation). On balance, the above hypothetical scenario readily explains these lexical data.



**Figure 5:** Neighbor-Net based on data in Figure 4 (by Russell Gray and Simon Greenhill).

## 5 Is Muṅḍā a primary division?

In relation to Muṅḍā, the most important question must be whether its typological character (suffixing, highly synthetic) is innovative or conservative. If the latter, the Mon-Khmer languages (non-suffixing, analytical) must constitute the innovative group, since they are unlikely to have all independently undergone the same typological restructuring. Pinnow wrote:

... the Muṅḍā languages are undoubtedly more similar to Proto-Austroasiatic than the other members of the family. From a morphological viewpoint they are far more conservative than Nicobarese and Khasi, and from the standpoint of vocabulary they surpass the Mon-Khmer languages in their preservation of ancient word stems and word forms. (Pinnow 1963:150)

Subsequent writers have appropriated these views, for example van Driem (2001:299). But it is much more likely that Muṅḍā is the innovator, and that the other languages retain, more or less unchanged, the typological character of Austroasiatic. The most persistent advocates of this latter view are Donegan and Stampe (for example 1983, 2004; Donegan 1993). They argue that the characteristics that make Muṅḍā distinctive are innovative, and that the restructuring was from isolating to synthetic typology, a reversal of Pinnow's formulation. Donegan and Stampe posit a shift from rising to falling accent in pre-Muṅḍā, which would explain the restructuring of Austroasiatic sesquisyllables into disyllabic roots and the rise of suffixation in Muṅḍā languages. In the course of restructuring, new vowels were inserted to break up initial clusters, forming new initial syllables in Muṅḍā. The original contrast of long versus short vowel was lost as long vowels were effectively split in two to create the new syllables. Where the proto-vowel was short it did not change, and

the quality of the new vowel was determined by other factors. Consequently Muṇḍā initial syllable vowels are predictable when the proto-main vowel was long; this is much easier to explain if the MK pattern is original. The alternative hypothesis, to derive MK word structure from Muṇḍā-type disyllabic roots, defies the usual rules of sound change. Examples showing this root structure correspondence are compiled from Shorto (2006) and are presented in Table 2.

**Table 2:** Root structure correspondence: Mon-Khmer and Muṇḍā

PMK	Mon-Khmer	Muṇḍā
*jliin ‘long’	Old Mon: <i>jliñ</i>	Muṇḍāri: jiliŋ
*kluu? ‘tortoise’	Mon: klao	Kharia: kulu
*brii? ‘forest’	Bahnar: bri:	Muṇḍāri: bir
*kjaal ‘air, wind’	Old Mon: <i>kyāl</i>	Muṇḍāri: hojo
*rk[aw]? ‘husked rice’	Khmer : ʔəŋkə:	Sora: ‘ruŋku:
*kmuu? ‘dirty’	Khmer: khmau	Muṇḍāri: humu
*smuul ‘shadow, soul’	Khmer: srəmaol	Muṇḍāri: umbul
*kra? ‘road, way’	Praok: kra	Muṇḍāri: hora
*kla? ‘tiger’	Old Mon: <i>kla(’)</i>	Muṇḍāri: kul ~ kula
*[hj]mu? ‘name’	Old Mon: <i>jamo’, himo’</i>	Kurku: jumu ~ jimu
*ʔə[ə]ŋ ‘foot/leg’	Old Mon: <i>juñ</i>	Muṇḍāri: jaŋga
*ba? ‘paddy’	Bahnar: ba:	Muṇḍāri: baba

Interestingly, Donegan and Stampe (2004) favour a South Asian origin for Austroasiatic. Their supposition is that such profound structural change within Muṇḍā must have taken a long time, perhaps even more time than Austroasiatic languages appear to have been in Southeast Asia. We speculate that a rapid restructuring could well have occurred if Muṇḍā had gone through a bottleneck event, perhaps as a small population of emigrants arriving in South Asia.

Both Sagart (this volume) and Sidwell (2009) have suggested that the Donegan-Stampe model, characterising Muṇḍā as a restructured Mon-Khmer type language, bears precisely upon the sub-classification of Austroasiatic, since it removes the rationale for Pinnow’s West-East division. It would thus appear there is no strong basis to the widely received notion of Muṇḍā versus Mon-Khmer coordinate branches. Put more strongly, it appears to represent a fundamental methodological error to assume that a branch with the most complex typology must somehow be its most ancient representative. This points strongly to a cultural classification based on fragmentary early documentation being erected into a genealogical theory without any strong evidential base. The parallel with the spurious division between Sinitic and Tibeto-Burman that is supposed to characterise Sino-Tibetan comes to mind (see, for example, discussion in van Driem 2001:316).

## 6 Proposals for Northern Mon-Khmer

Beyond the position of Muṇḍā, the most important characteristic of the various published classifications has been the identification of a Northern Mon-Khmer division, consisting principally of Palaungic, Khasi and Khmuic. The idea was hinted at by Thomas

and Headley (1970:404) who found that a case ‘might possibly be able to be made for a northern vs. southern grouping on the basis of the Khmuic figures, but this would be hard to sustain in the face of the rest of the figures.’ And later lexicostatistical studies, especially Headley (1976) and Peiros (1998, 2004:23), have weakly suggested a grouping of Palaungic and Khmuic counting some 26% cognates, while the same studies show lower percentages with Khasi (21~23%). However, the most important evidence for a Northern clade appears to be phonological, with two specific sound changes discussed.

Diffloth (1977) argued for a loss of medial \*-h- in Palaungic, Khmuic and Khasi, and this appears to have informed his initial formulation of Northern Mon-Khmer. That sound change is now known to be restricted specifically to Khmuic (for example Khmu *maam* ‘blood’ cf. Semelai *maham* ‘id.’, Khmu *bi?* ‘sated’ cf. Danaw *əi'*<sup>3</sup>, Palaung *hu?*, Semai *bahe* ‘id.’) and has been abandoned by Diffloth. Then, for a seminar delivered in Moscow in 1989,<sup>6</sup> Diffloth discussed a correspondence of preconsonantal \*s- in those languages to a \*t- in the rest of Mon-Khmer. Shorto had also discussed the same correspondence in a note drafted in the 1970s, quoted in the introduction to his posthumous 2006 handbook:

The whole of this group is characterized by a shift of \*t in initial position in structures \*CCVC (in some cases) to a sibilant, *prima facie* via an affricate stage. This minor shift is interesting because it apparently extends to Muṅḍā. Its incidence may be conditioned by the lost (in Mon-Khmer) V1 of Proto-Austroasiatic \*CVCVC, or it may entail reconstructing an additional proto-phoneme (\*t<sub>1</sub>) \*t<sub>2</sub>. Thus we find ‘taro’, Khm. tra:v, Ste traw; RL ʔsəroʔ, Khs. shriew, Sora ‘saro:-gai-ən, Muṅḍāri sārū, Santali saru; ‘sun, day’, Old Mon tñey, Khm. thñay; KY səŋiʔ, RL ʔsəŋiʔ, Khs. sngi, Muṅḍāri singi. Contrast (a bad example since it uses infix forms, but with a Muṅḍā cognate) ‘new’, Khm. thmxy ~ Middle Mon t/a/mi, RL ʔt/ən/meʔ, Khs. th/ym/mai, Kharia ‘t/ən/me. (Shorto 2006:x–xi)

This appears to be a real correspondence, although of the 14 examples in Shorto (2006) perhaps only five are viable. According to Sidwell’s provisional analysis, Shorto’s \*t<sub>1</sub>-/\*t<sub>2</sub>- is unnecessary, and the correspondence in question is the regular outcome of \*t- before a non-labial continuant. The suggestion by Shorto that vowel assimilation may be involved runs directly counter to the Donegan-Stampe model of word-structure, and would also contradict the Pinnow inspired treatment of Muṅḍā as a separate division. Presently, it would seem to be a case of the lenition of a stop in a particularly weak position, which may or may not have occurred independently. As it is, the number of tokens is so small, it is difficult to assess its significance.

The general issue of a Northern or Khasi-Palaungic sub-family, and whether it includes Khmuic, and/or the Mangic/Pakanic languages (Mang of Vietnam, Bolyu and Bagan of China) is crucial to the question of Austroasiatic diversity. Returning to our principle view that real sub-grouping ought to be evident in the basic vocabulary, one can readily offer significant observations. Of most immediate importance, it is apparent that between Khasi and Palaungic there are some eight isoglosses on the Swadesh 100 list that can be treated as innovations (lexical replacements, semantic shifts, loans). This is approximately 1/3 of the basic vocabulary they have in common, which we take as strongly indicative of sub-grouping, especially given the great geographical isolation between these two groups. These isoglosses are given in Table 3. Also shown are Mangic and Khmuic data, demonstrating that they do not share these innovations. Unfortunately our lexical sources

6 ‘Sub- and supra-classification of Mon-Khmer’ at the Institute of Far-Eastern Studies in Moscow. A copy of the handout is kept in the Cornell Library manuscript collection.

for Mangic (Ferlus ms.), are not as extensive as we would like, yet even with the sparse data available a clear pattern of independent lexical evolution is evident.

**Table 3:** Innovations suggesting Khasi-Palaungic

Gloss	Khasi	Palaungic <sup>#</sup>	Mang	Bolyu	Bugan	Khmu Chuang	PMK (Shorto)	
blood	snam	*snaam	ham	sa:m			-	
					kaŋ <sup>31</sup>		-	
claw/nail	tirsim	*rnsiim	dɔj			ma:m	*jhaam	
					ma:i <sup>13</sup> ti <sup>55</sup>		-	
					biaŋ <sup>35</sup>		-	
hair	ʃno:ʔ	ɲok (Danaw)	hók	suk <sup>53</sup>	sak <sup>55</sup>	tm <sup>h</sup> mɔ:ŋ	*t <sub>1</sub> m[uə]ŋʔ	
		*suk				gləʔ	*snuuk *suk	
man/male	trme (Amwi)	*-meʔ (Riang k <sup>ə</sup> rmeʔ <sup>2</sup> )	cuj				-	
					qɔ <sup>31</sup> pɔ <sup>33</sup>	piau <sup>35</sup>	-	
rain	slɛ (Amwi)	*sələʔ	maʔ			cmbroʔ/ gleʔ	-	
					qɔ <sup>55</sup>	ʔa <sup>33</sup>	kmaʔ	*gmaʔ
swim	dʒŋi:	*ŋɔj					-	
two	ʔa:r	*lʔaar				kljɔ:ŋ	*[l]ŋuj	
			zuəi					-
water	ʔum	*ʔoom		mbi <sup>55</sup>	bi <sup>31</sup>	ba:r	*ʔaar (?)	
			ʒum					*ʔ[ɔ]m (?)
				nde <sup>53</sup>	da <sup>35</sup>			*ɗaak

<sup>#</sup> Proto-Palaungic reconstruction by Sidwell, published online at sealang.net/monkhmer.

The Khasi-Palaungic innovations are neither shared with Khmuic nor Mangic. Furthermore, where Mangic and Khmuic agree, it is in retention of AA vocabulary. Two etyma stand out as requiring special comment: ‘blood’ and ‘water’.

- Ferlus (2009) reconstructs a proto-Vietic root \**saam* ‘to bleed’ (for example Viet. *tuom* ‘ooze, exude’) which has a direct cognate in Mangic, and is uniquely infixed in Khasi and Palaungic. It is not clear how this connects to other Austroasiatic forms which indicate a prevocalic /h/ (note regular loss of /h/ in Khmuic).

- the root *\*ʔ[o]m* ‘water’ replaced proto-Austroasiatic *\*daak*, perhaps by development from a root meaning ‘to bathe’ (cf. proto-South Bahnaric *\*ʔum* ‘bathe’). Yet reflexes of *\*ʔoom* are not general in Khmuic, but restricted to Khmu’, Khang, and Bit (Bit may be Khmuic or Palaungic, sources conflict). Other Khmuic languages have diverse forms for ‘water’, for example Iduh *paj*, Ksingmul *hɔ:t*, Mlabri *wək*, Pray *ʔɔ:k*.

There are no unambiguous indications in the basic lexicon that might link either Mangic or Khmuic to Khasi-Palaungic. Comparative phonology suggests a sound change in which proto-Austroasiatic *\*tC-* shifted to *\*sC-* in Muṅḍā, Khasi, Palaungic and Khmuic (Mangic language regularly lose this segment so we cannot take them into account here), but the implications of treating it as a single change are too great; it would reduce Muṅḍā to a branch of one of several Mon-Khmer sub-families, and challenge too many other facts. On balance it appears that we are obliged to abandon the notion of a Northern-Mon-Khmer clade beyond a likely Khasi-Palaungic sub-grouping. And it seems appropriate at this stage to treat Mangic as an independent branch.

## 7 Proposals for Nuclear Mon-Khmer

The third coordinate division of Diffloth’s (2005) *stammbaum* includes those groups previously characterised as Southern and Eastern Mon-Khmer or Nuclear Mon-Khmer, Khmero-Vietic and Nico-Monic on Figure 3. This resembles the ‘Central Branch’ of Peiros (2004) although he also includes Nicobarese. Again there is little in print to support such a grouping, although there has been some discussion of the putative lower level groups. One of these is Vieto-Katuic, first proposed in Diffloth (1991). The evidence is supported by an *\*-h-* / *\*-s-* correspondence based on six apparent cases. Alves (2005) compiled some 40 Vieto-Katuic isoglosses, some of which may be loans, or retentions from a higher node of Austroasiatic. Nonetheless, it is clear that Vieto-Katuic presents a promising line of inquiry.

In side remarks concerning Nicobarese and Aslian in a paper on Palaungic vowels, Diffloth (1991:14) explicitly mentions that: ‘...Nancowry Nicobar (Radakrishnan 1981:25) is described even today as also having five diphthongs: /iá/, /ía/, uá/, /úa/ and /úua/, which seem to correspond with what we can reconstruct for Proto-Aslian.’

Diffloth (personal communication 2008) gave several etyma that appear to show these correspondences. Unfortunately it is not possible to assess the significance of this correspondence in isolation from a complete reconstruction of the respective vocalic systems. Further to this, Diffloth (personal communication 2009) indicated an isogloss for ‘wife’ showing a supposed lexical innovation uniting Monic and Aslian, for example: Old Mon *kəndɔr*, Semelai *kərdɔ:r*. So far as we can tell, there is only one isogloss in the basic lexicon that unites all three branches of the supposed southern clade: *\*btəm* ‘night’ (Shorto 2006), for example Semelai *pətəm*, Written Mon *btam*, Car Nicobar *hatəm*. Such isoglosses might equally well be explained by contact or exclusive retention, keeping in mind that the lexicostatistics indicates that we should find some examples of inter-branch borrowing. We would expect that if a southern clade is real, it should have a strength of lexical support similar to what we find for Khasi-Palaungic.

The Bahnaric and Katuic branches, located more or less on the Khorat plateau region, have linguistic histories that are now relatively well researched, with numerous comparative studies since the 1960s (for example Blood 1966; Diffloth 1982; Efimov 1990; Efimov 1983; Gainey 1985; Peiros 1996; Shorto 2006; Sidwell and Jacq 2003;



Sidwell 1998, 2000, 2005b; Smith 1972; Theraphan L-Thongkum 2001; Thomas and Smith 1967; Thomas 1967). So far comparative research reveals only one apparent lexical innovation linking these two branches: the etymon for which Shorto (2006) reconstructs PMK *\*knʔiəs* ‘nail, claw’, and suggests that it may be an infixated form of *\*kiəs* ‘to scrape, scratch’. Sidwell (2005b) reconstructs proto-Katuic *\*krnias*, and Theraphan (2001) reconstructs proto-Bahnaric *\*k(ə)rnias*.

There are many examples of transparent loans back and forth between the two groups, indicated by the asymmetries in their geographical distributions. For example; Bahnaric *\*liam* ‘good’ and *\*ŋkee* ‘horn’ are restricted in Katuic to Katu *liem* and *tage*, while Katuic *\*səəŋ* ‘five’ and *\*kmɔɔ* ‘year’ have replaced proto-Bahnaric *\*pdam* and *\*cnam* in West Bahnaric. Such examples are numerous, the real problem being to identify early-stage loans. Various prominent etyma illustrate the dramatically different lexical, phonological and morphological histories of Bahnaric and Katuic. For example, Katuic *\*ŋhaaŋ* ‘bone’ and Bahnaric *\*ktsiiŋ* ‘bone’ (for example. Bahnar *ktiij*, Laven *ktiəŋ*, Jeh *ksiaŋ*) each show idiosyncratic developments from proto-Austroasiatic *\*jʔaaŋ*; neither Katuic *\*ktiak* ‘earth’ nor Bahnaric *\*teh/tneh* ‘earth’ can be reconciled with proto-Austroasiatic *\*tiʔ* by regular phonological correspondences; Katuic innovates a medial /s/ in *\*ksaj* ‘moon’ without obvious motivation (cf. Bahnaric *\*khaj*, Nancowry Nicobar *kahé* etc.); and Bahnaric *\*ʔuŋ* ‘fire’ has a short vowel and a final nasal while Katuic and the rest of Austroasiatic indicate a long vowel and a final /s/ (for example proto-Katuic *\*ʔuus*, Semai *ʔɔɔs*, Car Nicobar *ʒh* etc.).

In addition to the above, there is admittedly speculative evidence for additional branches of Austroasiatic which no longer exist. Their existence can be inferred indirectly from vocabulary in modern languages which seems to be etymologically distinct from the historical lexicon of those languages as currently classified. Such vocabulary would be borrowings or assimilations from languages no longer spoken. Blench (2009) has presented the idea that there were once three or perhaps four such subgroups. These are:

- a) The language of the Shom Pen on the Nicobar islands. Some documentation for this language has only recently become available, and a preliminary publication argues that it may be a language isolate (Blench 2008). However, additional evidence (Diffloth personal communication) points to some cognates with mainland Austroasiatic not shared with other Nicobarese languages (Blench in press a). Shom Pen might therefore represent an earlier and distinct migration to the islands.
- b) Acehese. This language is usually classified as Chamic (Thurgood 1999; Sidwell 2005a). However, it does have a great deal of ‘residual’ vocabulary whose origins are unclear. This might either result from substrate languages of unknown affiliation or possibly from the Chamicisation of a prior Austroasiatic language (Diffloth personal communication).
- c) Pre-Chamic. In the same way, Chamic itself has distinctive unetymologised vocabulary, which could result from the absorption of resident populations on the Vietnamese mainland.
- d) Borneo Austroasiatic. There is considerable archaeological and cultural evidence for intensive maritime contact between western Borneo and the SE Asian mainland. Adelaar (1995) says ‘The Land Dayak languages have a few striking lexical and phonological similarities in common with Aslian languages. This suggests that Land Dayak originated as the result of a language shift from Aslian to Austronesian, or that both Land Dayak and Aslian have in common a substratum from an unknown

third language'. These connections were observed as early as Skeat and Blagden (1906) and the question remains unresolved. Blench (in press c) has reviewed the lexical evidence for this hypothesis and finds some striking similarities between Borneo languages and mainland Austroasiatic.

It should be underlined that these hypotheses have yet to be fully substantiated, and it may well be that, for example, not all these substrates are Austroasiatic, or that c) and d) turn out to be the same 'lost' branch. However, in the broader picture, they all point in the same direction, to linguistic diversity in the Southeast Asian heartland, which is the underlying thrust of the current argument.

## **8 Centre of diversity**

There is no evidence of shared innovations that would justify grouping Mon-Khmer branches into one or two families co-ordinate with Muṅḍā. The most parsimonious explanation for the typological divergence of Muṅḍā is that it was restructured from isolating to synthetic type due to South Asian contact influences. This conclusion alone would place the centre of Austroasiatic diversity outside South Asia.

This chapter has argued that the default classification of Austroasiatic is an essentially flat array. With no basis for treating the non-Muṅḍā Austroasiatic languages as a single clade, we must abandon the idea of two main branches with a notional geographic centre towards South Asia. Even if we accept some or all the specific sub-grouping proposals outlined above, the centre of diversity would still be radially aligned within Indo-China and the lower Mekong. Groups with few or no indications of such borrowings moved away first, and came under other unrelated influences. Those in an intermediate position, such as Monic or Khmuic, either left later, or migrated early but remained in intermittent contact with the heartland. Perhaps Huffman (1978) was on the right track when he suggested this as the homeland location. Bahnaric and Katuic, resident in that region for the longest period, show the highest level of mutual borrowings. In this context it strikes us as relevant that the languages of this region exhibit strong marks of contact-driven lexical convergence.

## **9 The lexicon of subsistence**

Diffloth (2005) has claimed that reconstructions of words for fauna support a tropical western origin for Austroasiatic. Notwithstanding the difficulty in assessing those reconstructions, such an argument must include considerations of the range and distribution of such environments at the time depths being considered for proto-Austroasiatic (that is, 4–8000 BP). Crucial also to the argument concerning the origin and dispersal of Austroasiatic is the reconstruction of the subsistence lexicon. The other uniting feature is the reckoning of the role of rice in proto-Austroasiatic society. Sagart (this volume) represents the common position when he suggests that the Austroasiatic homeland question is intimately linked with the domestication of rice. Zide and Zide (1976) pointed out that Muṅḍā rice vocabulary shows cognates with other branches of Austroasiatic, and Diffloth (2005:78) has added additional lexical items to show rice cultivation was present in the earliest period (see also Blench 2005). Below in Table 4 we list Diffloth's reconstructions, plus the equivalents from Shorto (2006). (Cf. Diffloth, this volume.)

**Table 4:** Proposed PMK reconstructions for rice terminology

<b>Diffloth (2005)</b>	<b>Shorto (2006)</b>
#(kə)ba:ʔ ‘rice plant’	*baʔ ‘paddy’
#rəŋko:ʔ ‘rice grain’	*rk[aw]ʔ (and *rkaawʔ?) ‘husked rice’
#cəŋka:m ‘rice outer husk’	*skaamʔ ‘chaff, husks of paddy’
#kəndək ‘rice inner husk’	*lʔək ‘rice-bran’
#phe:ʔ ‘rice bran’	*[p]heʔ ‘husked rice’

Diffloth (2005:80) asserts that the reconstruction of such a lexicon indicates ‘that PAA is a very old language whose speakers were well acquainted with rice cultivation (at least dry hillside cultivation)’. But this does not necessarily follow: we would expect that much or all of this terminology was simply transferred, by semantic extension, from previous usage that need not have involved rice previously. For example, Ferlus (1996) discusses at length the adaptation of the Austroasiatic word for ‘taro’ (reconstructed *\*t<sub>2</sub>rawʔ/\*t<sub>2</sub>raaw[ ]* by Shorto) in Mon, Khmer and Katuic, to designate the rice associated with intensive irrigated paddy, as it became increasingly important in Northeast Thailand from around 2000 BP. It is also surely significant that, despite numerous studies (Zide and Zide 1976; Revel 1988; Ferlus 1996; Bradley 1997; Peiros and Shnirelman 1998; Vovin 1998; Sagart 2003; Blench 2005), no convincing correspondence have been found between rice-related vocabulary in Austroasiatic and the terminologies of rice in other language phyla.

Moreover, a major controversy exists over the antiquity of rice cultivation in East Asia. Normile (1997) reports rice remains before 11,500 BP in Hunan and Hubei in central China and dates of similar antiquity regularly occur in the literature. Jiang and Liu (2006) review a series of dates for this region between 8000 and 13,500 BP. However, Fuller et al. (2008) argue that many of these finds are either poorly dated or refer to wild rice and arise from a misunderstanding of the phenotypic characters of the grain, and that true domestic rice only occurs from about 6000 BP onwards. However, as Zong et al. (2007) point out, by 7700 BP there is good evidence for landscape management, through fire and flood control, consistent with paddy cultivation. Whatever the situation in China, it is important to note that no remains of unambiguously domesticated rice have been found in MSEA; these have only been dated to around 4000 BP.

One claim made by Diffloth (2005) appears to us to be uncontroversial; that Austroasiatic speakers typically spread along river valleys, seeking swampy ground to cultivate taro (although they obviously also became seagoing at least once, viz. settlement on the Nicobars). This is consistent with the suggestion made in 1943 by Haudricourt and Hédin, when they proposed that rice in Southeast Asia began as a weed in those boggy taro fields. A close association between taro and rice persists among Austroasiatic farmers, some of whom still ritually plant taro in their rice fields even though they now depend on the rice crop (such as Condominas 1957) described for the Mnong Gar). Generally the indications are strong that taro was the original crop and that rice was superimposed upon it. The extension of rice agriculture into new niches over time, such as the steep hillsides, would have greatly extended the potential range of those early communities.

Another point arises from consideration of Diffloth (2005:79). Discussing faunal terms, he remarks: ‘These words are morphologically opaque, suggesting long-term familiarity with the items in question.’ The same logic must necessarily apply generally, and significantly we find that terms associated with the farming and processing of rice for consumption are not morphologically opaque. The following analyses can be offered in respect of the following quasi-reconstructions:

**#təmpal** ‘mortar’

This word is clearly an instrumental nominalisation with the \*-p- infix. The root is \*tal with a meaning associated with pounding/striking, reflected in modern Khmer /dal/ ‘to punch, pound’.

**#jənrəʔ** ‘pestle’

Ferlus (2008) reconstructs this word as an infixed derivative of proto-Vietic \*tʃeʔ ‘to dig, excavate’ (via \*tʃ-r-eʔ), which diffused through Austroasiatic in association with the pestle itself.

**#jəmpiər** ‘winnowing tray’

The medial -mp- suggests a nominalisation of a hypothetical \*jiər; possible reflexes include Khasi *jiar* ‘to drain off, to filter’ and Prao *ci:a* ‘to fall, shed’. On the other hand, Shorto reconstructs the primary meaning as verbal, on the basis of forms such as Central Nicobarese *ifuə* ‘to blow’.

**#jərmuəl** ‘dibbling-stick’

Diffloth and Shorto disagree on whether the primary meaning is nominal or verbal. The medials -rm- strongly suggest a nominalisation of a hypothetical \*juul, very likely reflected in proto-Pearic \*co:l ‘to plant’ (Headley 1985).

The above are consistent with the idea that methods of farming and preparing harvested rice for consumption were relatively new to proto-Austroasiatic speakers. These words could even have been coined, and diffused through the speaker community, after the linguistic break-up had begun, but while speakers were still in contact (the dialect chain stage).

## 10 Aquatic subsistence

If early Austroasiatic speakers were spreading up river valleys, then their way of life must have been strongly associated with boats and the exploitation of aquatic resources. Potential reconstructions suggesting a riverine environment for Austroasiatic are numerous. Table 5 compiles PMK reconstructions proposed by Shorto, consistent with an original aquatic subsistence strategy.

**Table 5:** PMK reconstructions from Shorto suggesting an aquatic subsistence strategy

Gloss	PMK	Selected Reflexes
boat/canoe	*d̥uuk	Khmer <i>tù:k</i> , Bru <i>tùəʔ</i> , Viet. <i>nó̌c</i>
boat	*d̥ <sub>2</sub> luŋ	Old Mon <i>dluñ</i> , Lawa <i>ʔloŋ</i> , Bahnar <i>pluŋ</i> , Khmu <i>cə̀l̥:ŋ</i>
crab	*kt <sub>1</sub> aam	Khmer <i>kda:m</i> , Khasi <i>tham</i> , Bahnar <i>kə̀taam</i> , Khmu <i>kt̥a:m</i>
prawn	*[k]nt <sub>1</sub> a[i]s	Khmu <i>cntah</i> , Riang-Lang <i>kə̀ntas</i>
prawn, shrimp	*suum	Katu <i>suam</i> , Viet. <i>tôm</i>
shrimp	*knbis	Khmu <i>kə̀mpùh</i> , Chrau <i>kə̀mvih</i>
catfish	*[t]kəʔ	Mon <i>həkəʔ</i> , Chrau <i>kə:</i>
eel	*nduŋ	Khmu <i>ʔə̀ntùəŋ</i> , Chrau <i>nduŋ</i> , Mon <i>daluñ</i> , Chong <i>kə̀ml̥:ŋ</i>
serpent headed fish <sup>#</sup>	*knl(ua)n	Bru <i>kluàn</i> , Bahnar <i>rə̀l̥:n</i> , Mon <i>kanan</i> , Nancowry <i>lúan</i>
otter	*bheʔ	Khmer <i>phè:</i> , Semelai <i>bə̀heʔ</i> , Khasi <i>kə̀siʔ</i>

<sup>#</sup>From Diffloth (1979).

One term for boat is attested in both Austroasiatic and Austronesian, reflexes are laid out in Table 6.

**Table 6:** A SEA regional term for ‘boat’

Phylum	Subgroup	Language	Attestation	Gloss
Austronesian	PAN		* <i>qabaŋ</i>	boat, canoe
	Taiwan	Siraya	<i>avaŋ</i>	canoe
	Taiwan	Favorlang	<i>abaŋu</i>	boat
	Philippines	Magindanao	<i>kaban</i>	boat
	Philippines	Sulu	<i>guban</i>	boat
	Malayic	Malay	<i>kěbang</i>	vessel
	Barrier	Nias	<i>owo</i>	boat
	Barrier	Sichule	<i>ofo</i>	boat
	Austroasiatic	Bima-Sumba	Sawu	<i>kowa</i>
PMK			* <i>kbaŋ</i>	ship
Monic		Old Mon	<i>kḥaṅ</i>	ship
Bahnaric		Biat	<i>baŋ</i>	coffin
Aslian		Jaehai	<i>kupon</i>	boat
Nicobarese			<i>kopòk</i>	boat

The lack of Muṅḍā and Khasi cognates makes it difficult to assign this term to proto-Austroasiatic; nonetheless the Nicobarese and Aslian forms are clearly not just Malay borrowings, and it must be assigned to an early period in Austroasiatic expansion. Mahdi (1999) has identified the links, both cultural and lexical, between coffins and boats, such as is attested in Bahnaric. The cognacy with Austronesian is puzzling but must be evidence for significant early contact between these two phyla (Blench in press c). An aspect of a fisher-forager strategy is that boat-using populations can move far and fast. This is accepted for Austronesian, where the rapid diversification of proto-Malayo-Polynesian is strongly associated with a major upstep in sailing technology, something reflected in the terminology for sails and other parts of the boat (Pawley and Pawley 1994). A similar innovation in boat-forms on the Mekong would account both for the rapid dispersal of Austroasiatic and the absence of nested subgroups.

These examples from the lexical data should be regarded as provisional. Shorto’s evidence is often rather more scattered than a proposed reconstruction to a proto-language might warrant. Nonetheless, it points strongly to the importance of aquatic subsistence in conjunction with wet-zone agriculture. As a consequence, speakers must have been agriculturalists and thus any date or place proposed for the expansion of Austroasiatic must be congruent with archaeological evidence for the MSEA Neolithic. The extent to which this was demic diffusion as opposed to language expansion is largely irrelevant; whether small numbers of speakers spread and influenced in situ foragers to adopt their subsistence strategies or whether all Austroasiatic groups are the result of relatively large population movements is yet to be determined.

The archaeological evidence points to a rapid expansion of the Neolithic in the Yunnan/Northern Vietnam borderland, some 4000 years ago (Higham 2002:85 ff.). Higham (2004:47) notes: ‘The pattern of intrusive agriculturalists settling inland valleys in southern China, while the coast continued to be occupied by affluent foraging groups, is repeated in the Red River area and the contiguous coast of Vietnam.’

The most well-known site of this type is Phung Nguyen, about 200 km inland from Halong Bay. Dates remain problematic, but the adjacent site of Co Loa has been dated to 2000 BC (Lai Van Toi 1999). In summarising the situation, Higham says:

We find agricultural settlements being founded in the lower Red River valley, along the course of the Mekong and its tributaries, and in the Chao Phraya valley...The dates for initial settlement, as far as they are known, are approximately the same with none earlier than about 2300 BC. Most intriguingly, the pottery vessels in many of the sites over a broad area have a similar mode of decoration. The sites reveal extended inhumation graves and an economy incorporating rice cultivation and the raising of domestic stock. (Higham 2002:352)

And Bellwood (2005:132) remarks on the wide distribution of 'incised and zone-impressed' pottery 'across parts of far southern China, northern Vietnam and Thailand after about 2500 BC'. In relation to the spread of this tradition: 'Peninsular Neolithic pottery has cord-marked decoration with rare incision and red-slipping, often with tripod feet or pedestals...Gua Cha in Kelantan also has fine incised pottery with zoned punctuation dating to about 1000 BCE.' (Cf. White, this volume.)

Finally, Rispoli in the most recent, wide-ranging review of 'incised and impressed' pottery says:

The main peculiarity of the incised and impressed pottery style is its sudden appearance around the second half of the 3rd millennium BCE in Neolithic sites distributed in the major river plains of mainland Southeast Asia .... Incised and impressed pottery style, moreover, does not appear in isolation, but it is associated recurrently with: small polished stone tools; stone or shell bracelets and necklace beads. (Rispoli 2008:238)

We suggest that the sudden expansion of this distinctive pottery style and associated toolkit and decorative elements is a marker of the Austroasiatic expansion.

A new chronology is now being developed for the beginning of the Neolithic in MSEA, based on the new C14 chronology of Ban Non Wat (Higham and Higham 2009). This proposes that the older radiocarbon dates need to be revised and the period 1800/1700 to 1100 BC is more credible. Rispoli (personal communication) observes that the recent C14 dated excavations at An Son and Da Kai in South Việt Nam are fully consonant with this. On the basis of comparisons between our sites in Central Thailand and most of the other Neolithic sites in Thailand, Vietnam and Yunnan the new dates 'put all the tiles in the right place', linking MSEA with Yunnan as well as Guangxi/Guangdong.

By what mechanisms could Austroasiatic have spread so far and so fast that all the apparent branches were effectively dispersed and in place by 2–3000 years ago? Probably the answer lies in the kind of transition the Austroasiatic speakers went through as they expanded fisher/forager/vegeticulturalist practices to incorporate irrigated and dry paddy. But they are not congruent with a date of 7000 BP. If rice agriculture in Indo-China is ca. 4200 BP, the initial dispersal of proto-Austroasiatic should not be earlier than this. If this is the case, then the phylum is unlikely to have an intricate nested structure, because this would not allow sufficient time for such a structure to develop. The 'flat array' model of Austroasiatic is thus more plausible and congruent with the lexical data.

The Southeastern Riverine Hypothesis propounded here attributes the original homeland of Austroasiatic, or at least a principal phase of its expansion, to a region along the Mekong River. This is partly a 'centre of gravity' view, permitting the different groups of Austroasiatic to have become established in their various secondary homelands by a

series of least radical moves, that is moves which are characterised by short distances and plausible directions. But it also suggests an important transformation in our views of the subsistence strategies of early speakers of Austroasiatic languages. It seems reasonable to hypothesise that they were based along the river and were primarily fisher-forager populations<sup>7</sup>.

The cultural innovation of adopting rice into their repertoire, including the facility to farm dry rice in areas upland from main waterways, could well have facilitated the outward East-West spread overland, as opposed to the mainly North-South orientation of the Mekong (and Chao Phraya, Irrawaddy etcetera) of peoples who had previously established themselves along the riverine environment as fisher-forager-tuberculturalists. The stimulus for this innovation could have been contact with another population practising agriculture, and the obvious candidate is the Daic speakers, who would have been moving into the relevant area, from the Pearl River Delta to the Red River, Northern Laos and the Upper and Mid-Mekong at around four thousand years ago (Blench in press b). This is not to say that Austroasiatic speakers adopted an agricultural package wholesale, but rather that by stimulus diffusion, the concept of a more intensive use of riverbank and immediate upland environments would have been transmitted.

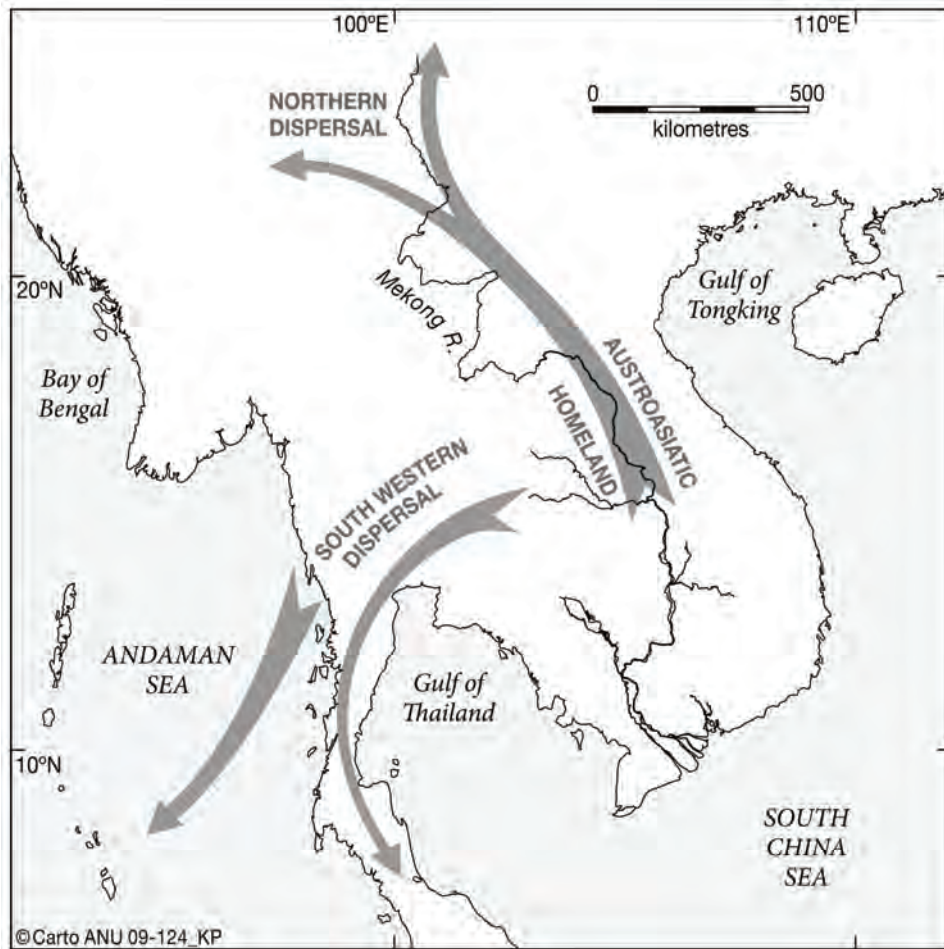
## 11 A model for the early history of Austroasiatic

Taken together, these elements suggest that we can reconstruct the early history of Austroasiatic as follows;

- a) ca. 3800 years ago, a new pottery style begins to spread rapidly throughout the region. This is associated with beginnings of the Neolithic in the region.
- b) early Austroasiatic speakers, already practising taro cultivation, are situated on the middle Mekong and adopt rice and also get access to improved types of boat.
- c) this subsistence revolution stimulates them to move both up and down the Mekong but also to spread westward to parallel river systems, seek new areas for their taro fields.
- d) a significant movement westward (perhaps to the Tonle Sap system and/or Chao Phraya Basin) allows the development of a south-western nucleus, the origin of Monic, Nicobarese and Aslian.
- e) the rapidity of this movement accounts for the difficulty in finding well-supported nested structures in the phylogenetic tree.
- f) subsequent expansions, particularly of the Daic, Sino-Tibetan and Austronesian language phyla fragmented the chain of Austroasiatic languages leading to their comparative geographic isolation in many outlying areas.
- g) Munḍā languages underwent a typological shift in contact with South Asian languages, but this was limited to a single branch rather than indicative of an early two-way division in the phylum.

---

<sup>7</sup> It is intriguing that this is also an evolving characterisation of the Austronesians (Bulbeck 2008) in the light of the conspicuous absence of archaeological evidence for the ‘Neolithic’ agricultural package previously held to drive their expansion.



**Figure 6:** The Southeastern Riverine hypothesis for the Austroasiatic dispersal.

Figure 6 presents a map of the possible pattern of the dispersal of AA according to the Southeastern Riverine Hypothesis. The model posits an early broad North-South differentiation along the Mekong valley, followed by a pattern of outward migrations, represented here by broad-brush arrows.

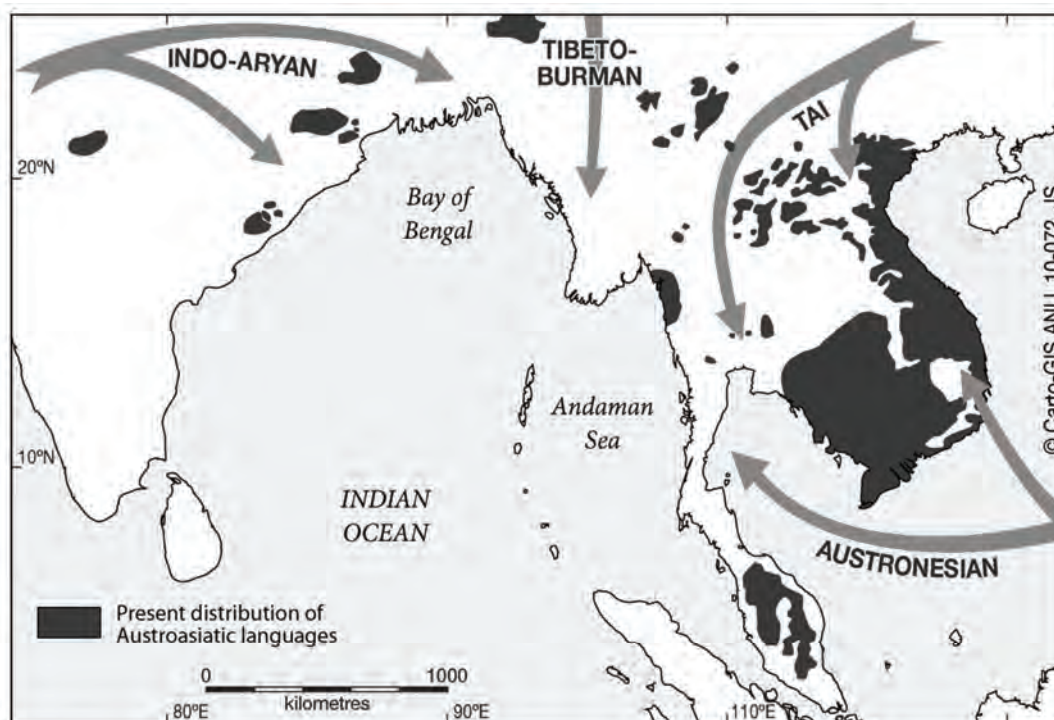
Figure 7 presents the possible subsequent movements of other language phyla which would have acted to fragment Austroasiatic branches, according to this scenario.

See also the paper by Blench (this volume) for the broader picture of language phylum interaction in MSEA.

It is important to underline the provisional nature of this hypothesis. The direct archaeological evidence for agriculture in mainland Southeast Asia is sparse and almost entirely confined to rice, the domestic status of which is a subject of debate. Linguistic pointers to other cultigens and broader indications of subsistence remain without archaeobotanical correlates.

The other aspect of this hypothesis, which cannot be emphasised too strongly, is to develop strong arguments either for or against the unification of individual branches of Austroasiatic. This chapter has reviewed a number of proposals and found them wanting. Nonetheless, this does not exclude the possibility that Monic-Nicobarese-Aslian, Vietic-Katuic or Palaungic-Khasian subgroups will eventually be demonstrated. But both lexicostatistics and the absence of shared phonological innovations suggest that Austroasiatic will never prove to have the complex nested structures of previous proposals, nor will it have the antiquity these are taken to imply.





**Figure 7:** The intrusion of outside language groups leading to isolation of Austroasiatic communities.

## Acknowledgements

Sidwell's contribution to this chapter was facilitated by assistance from National Endowment for the Humanities (Washington). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily represent those of the National Endowment for the Humanities.

## References

- Adelaar, Alexander K. 1995. Borneo as a Cross-Roads for Comparative Austronesian Linguistics. In Peter Bellwood, James J. Fox and Darell Tryon (eds) *The Austronesians: Historical and Comparative Perspectives*, 81–102. Canberra: ANU.
- Anderson, Gregory (ed.). 2008. *The Munda Languages*. London/New York: Routledge.
- Alves, Mark. 2005. The Vieto-Katuic Hypothesis: Lexical Evidence. In Paul Sidwell (ed.) *SEALS XV Papers from the 15th Annual Meeting of the Southeast Asian Linguistics Society 2003*, 169–176. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Bellwood, Peter. 2005. *First Farmers: The Origins of Agricultural Societies*. Oxford: Blackwell.
- Bergsland, Knut and Hans Vogt. 1962. On the validity of glottochronology. *Current Anthropology* 3. 111–153.
- Blench, Roger M. 2005. From the mountains to the valleys: understanding ethnolinguistic geography in SE Asia. In Laurent Sagart, Roger M. Blench and Alia Sanchez-Mazas (eds) *The peopling of East Asia*, 31–50. London: Routledge.

- 2008. Re-evaluating the linguistic prehistory of South Asia. In Toshiki Osada and Akinori Uesugi (eds) *Occasional Paper 3: Linguistics, Archaeology and the Human Past*, 159–178. Kyoto: Indus Project, Research Institute for Humanity and Nature.
  - 2009. *Are there four additional unrecognised branches of Austroasiatic?* Presentation at ICAAL-4. Bangkok, October 29–30, 2009.
  - in press a. *Is Shom Pen a distinct branch of Austroasiatic?* Paper given at the IVth ICAAL, Bangkok, November 2009 and in press for a special edition of *Mon-Khmer Studies*.
  - in press b. *The expansion of the Daic peoples: cultural and iconographic correlations*. Presentation at the XIVth EURASEAA meeting, Leiden, September 2008.
  - in press c. *Was there an Austroasiatic presence in island SE Asia prior to the Austronesian expansion?* Paper given at the 19<sup>th</sup> IPPA, Hanoi December 2009, and subsequently submitted for publication.
  - in press d. *Vernacular names for taro in the Indo-Pacific region and their possible implications for centres of diversification*. Paper submitted for a special IPPA volume on wet taro systems. David Addison and Matthew Spriggs (eds).
  - this volume, The role of agriculture in the evolution of mainland Southeast Asian language phyla.
- Blood, Henry F. 1966. *A Reconstruction of Proto-Muong (Including Tentative Reconstruction of Proto-South-Bahnaric)*. M.A. Thesis, Department of Linguistics Indiana University. Published by SIL in 1976.
- Blust, Robert. 1996. Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology. In Ward H. Goodenough (ed.) *Prehistoric Settlement of the Pacific*, 117–137. Transactions of the American Philosophical Society 86.5.
- Bradley, David. 1997. What did they eat? Grain crops of the Burmic peoples. *Mon-Khmer Studies* 27. 161–170.
- Bulbeck, David. 2008. An integrated perspective on the Austronesian diaspora: the switch from cereal agriculture to maritime foraging in the colonisation of island SE Asia. *Australian Archaeology* 67. 31–52.
- Burenhult, Niclas, Michael Dunn and Nicole Kruspe, this volume, Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula.
- Condominas, Georges. 1957. *Nous avons mangé la forêt [...]*. Paris: Mercure de France.
- Diffloth, Gérard. 1977. Mon-Khmer Initial Palatals and “substratumized” Austro-Thai. *The Mon-Khmer Studies Journal* 6. 39–57.
- 1979. Aslian languages and Southeast Asian prehistory. *Federation Museums Journal* 24. 3–16.
  - 1982. Registres, dévoisement, timbres vocaliques: leur histoire en Katouïque. *Mon-Khmer Studies* 11. 47–82.
  - 1991. Palaungic Vowels in Mon-Khmer Perspective. In Jeremy H.C.S. Davidson (ed.) *Austroasiatic Languages, Essays in honour of H. L. Shorto*, 13–28. School of Oriental and African Studies, University of London.
  - 2005. The contribution of linguistic palaeontology to the homeland of Austro-asiatic. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 79–82. Routledge/Curzon.
- Donegan, Patricia J. 1993. Rhythm and Vocalic Drift in Muṅḍā and Mon-Khmer. *Linguistics of the Tibeto-Burman Area* 16/1.

- Donegan, Patricia J. and David Stampe. 1983. Rhythm and holistic organization of language structure. In John F. Richardson, Mitchell Marks and Amy Chukerman (eds), 337–353.
- . 2004. Rhythm and the synthetic drift of Muṅdā, in Rajendra Singh (ed.) *The Yearbook of South Asian Languages and Linguistics 2004*, 3–36.
- Driem, George van. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region: Containing an Introduction to the Symbiotic Theory of Language*. Brill.
- . 2007. Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies. *Mon-Khmer studies: a journal of Southeast Asian languages and cultures* 37. 1–14.
- Efimov, Aleksandr. 1983. *Problemy fonologicheskoy rekonstrukcii proto-katuicheskogo jazyka*. Kandidat Dissertation, Institute of Far Eastern Studies Moscow.
- . 1990. *Istoricheskaja Fonologija Juzhnobaxnaricheskix Jazykov*. Moskva: Nauka.
- Ferlus, Michel. 1996. Du taro au riz en Asie du Sud-Est, petite histoire d'un glissement sémantique. *Mon-Khmer Studies* 25. 39–49.
- . 2008. Proto-Vietic reconstruction available online at <http://sealang.net/monkhmer/>.
- . 2009. Le mot “sang” en austroasie. *Mon-Khmer Studies* 38. 25–39.
- Fuller, Dorian, Lin Qin and Emma Harvey. 2008. Evidence for the late onset of agriculture in the Lower Yangtze region and challenges for an archaeobotany of rice. In Alicia Sanchez-Mazas, Roger M. Blench, Malcom D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 40–83. London: Routledge.
- Gainey, Jerry. 1985. *A comparative study of Kui, Bruu and So phonology from a genetic point of view*. Bangkok, Thailand: Chulalongkorn University thesis.
- Grierson, George A. 1919. *Linguistic survey of India*. vol. VIII, Part 1. Calcutta.
- Haudricourt, André-Georges. 1965. Mutation consonantique en Mon-Khmer. *Bulletin de la Société Linguistique de Paris* 60. 160–72.
- Haudricourt, André-Georges and Louis Hédin. 1943. L'homme et les plantes cultivés. In *Geographie humaine* 19. Paris: Gallimard.
- Headley, Robert K. 1976. Some Considerations on the Classification of Khmer. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austroasiatic Studies*, 431–452. Honolulu: The University Press of Hawaii.
- . 1985. Proto-Pearic and the Classification of Pearic. In Thomas Ratanakul and Premsirat (eds) *Southeast Asian Linguistic Studies presented to André-G Haudricourt*, 428–478. Bangkok: Mahidol University.
- Heine-Geldern, Robert von. 1923. Südostasien. In Georg Buschan (ed.) *Illustrierte Völkerkunde*. Stuttgart: Strecker und Schröder.
- Higham, Charles. 2002. *Early cultures of mainland Southeast Asia*. Bangkok: River Books.
- . 2004. Mainland Southeast Asia from the Neolithic to the Iron age. In Ian Glover and Peter Bellwood (eds) *Southeast Asia: from prehistory to history*, 41–67. Abingdon: RoutledgeCurzon.
- Higham, Charles and T. Higham. 2009. Dating Southeast Asian Prehistory: New Radiometric Evidence from Ban Non Wat. *Antiquity* 83. 125–144.
- Hojjer, Harry. 1956. Lexicostatistics: a critique. *Language* 32. 49–60.

- Holm, Hans. 2003. The proportionality trap or what is wrong with lexicostatistical subgrouping. *Indogermanische Forschungen* 108. 38–46.
- Huffman, Franklin E. 1978. *On the centrality of Katic-Bahnaric to Austroasiatic*. Communication à la 2ème Conférence Internationale sur les langues Austroasiatiques, Mysore (Inde) In *Proceedings of the International Congress of Anthropological and Ethnological Sciences*, 10(2). 78–79.
- Jiang, Leping and Li Liu. 2006. New evidence for origins of sedentism and rice domestication in the Lower Yangzi River. China. *Antiquity* 80. 355–361.
- Lai Van Toi. 1999. The initial inhabitants of Co Loa [in Vietnamese]. *Khao Co Hoc* 3. 39–54.
- Mahdi, Waruno. 1999. The dispersal of Austronesian boat forms in the Indian Ocean. In Roger Blench and Matthew Spriggs (eds) *Archaeology and Language III*. 144–79. London: Routledge.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Norman, Jerry and Mei Tsu-lin. 1976. The Austroasiatics in Ancient South China: some lexical evidence. *Monumenta Serica* 22. 274–301.
- Normile, Dennis. 1997. Archaeology: Yangtze seen as earliest rice site. *Science* 275. 309–310.
- Pawley, Andrew and Medina Pawley. 1994. Early Austronesian terms for canoe parts and seafaring. In Andre K. Pawley and Malcolm D. Ross (eds) *Austronesian terminologies: continuity and change*, 329–61. Canberra: PL Series C-127.
- Peiros, Ilia. 1996. *Katic comparative dictionary*. Pacific Linguistics C-132.
- 1998. *Comparative Linguistics in Southeast Asia*. Pacific Linguistics Series C-142. Canberra: Australian National University.
- 2004. *Genetičeskaja klassifikacija avstroaziatskix jazykov*. Moskva: Rossijskij gosudarstvennyj gumanitarnyj universitet (doktorskaja dissertacija).
- Peiros, Ilia and Shnirelman, Victor. 1998. Rice in Southeast Asia: a regional interdisciplinary approach. In Roger Blench and Matthew Spriggs (eds) *Archaeology and Language, archaeological data and Linguistic hypotheses* Vol. II. 379–389.
- Pinnow, Heinz-Jürgen. 1959. *Versuch Einer Historischen Lautlehre Der Kharia-Sprache*. Wiesbaden: Otto Harrassowitz.
- 1960. Über den Ursprung der voneinander abweichenden Strukturen der Muṇḍā- und Khmer-Nicobar Sprachen. *Indo-Iranian Journal* 4(1). 81–103.
- 1963. The position of the Muṇḍā languages within the Austroasiatic language family. In Harry L. Shorto (ed.) *Linguistic Comparison in Southeast Asia and the Pacific*, 140–152. London: SOAS.
- Przyluski, Jean. 1924. Les langues austroasiatiques. In Meillet, Antoine and Marcel Cohen (eds) *Les Langues du Monde (Collection linguistique publiée par la société de linguistique de Paris)*, 16. 335–403. Paris: Librairie Ancienne Edouard Champion.
- Radhakrishnan, R. 1981. *The Nancowry word: phonology, affixal morphology and roots of a Nicobarese language*. Current inquiry into language and linguistics 37, Edmonton, Alberta, Linguistic Research Inc.
- Revel, Nicole (ed.). 1988. *Le riz en Asie du sud-est*. [3 vols.] Paris: EHESS.
- Rispoli, Fiorella. 2008. The incised and impressed pottery of Mainland Southeast Asia: following the paths of Neolithization. *East and West* 57. 235–304.

- Sagart, Laurent. 2003. The vocabulary of cereal cultivation and the phylogeny of East Asian languages. *Bulletin of the Indo-Pacific Prehistory Association* 23 (Taipei Papers, Volume 1). 127–136.
- . 2008. The expansion of *Setaria* farmers in East Asia: a linguistic and archaeological model. In Alicia Sanchez-Mazas, Roger M. Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*. 133–158. London: Routledge.
- . this volume, The Austroasiatics: East to West or West to East?
- Schmidt, Wilhelm. 1904. Grundzüge einer Lautlehre der Khasi-Sprache in ihren Beziehungen zu derjenigen der Mon-Khmer-Sprachen. Mit einem Anhang: die Palaung-Wa-, und Riang-Sprachen des mittleren Salwin. *Abh. Bayerischen Akademie der Wissenschaft* 1.22.3. 677–810.
- . 1905. Grundzüge einer Lautlehre der Mon-Khmer-Sprachen. *Denkschrift der Akademie der Wissenschaften, Wien, Philologisch-Historische Klasse* 51. 1–233.
- Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. University of Hawaii Press.
- Shorto, Harry L. 1976a. In Defense of Austric. *Computational analyses of Asian and African languages* 6. 95–104.
- . 1976b. The Vocalism of Proto-Mon-Khmer. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austroasiatic Studies*. Honolulu: University of Hawaii (Oceanic Linguistics, Special Publication, No. 13). Part II. 1041–1067.
- . 2006. *A Mon-Khmer Comparative Dictionary*. Canberra, Pacific Linguistics 579.
- Sidwell, Paul. 1998. *A Reconstruction of Proto-Bahnaric*. PhD Thesis, University of Melbourne.
- . 2000. *Proto South Bahnaric: a reconstruction of a Mon-Khmer language of Indo-China*. Canberra: Pacific Linguistics.
- . 2005a. Acehnese and the Aceh-Chamic Language Family. In Anthony Grant and Paul Sidwell (eds) *Chamic and beyond: Studies in mainland Austronesian languages*, 211–246. Canberra: Pacific Linguistics PL 569.
- . 2005b. *The Katuic Languages: classification, reconstruction and comparative lexicon*. Munich: Lincom Europa.
- . 2008. Is Mon-Khmer dead? Long live Austroasiatic! Paper read to *18th Annual Meeting of the Southeast Asian Linguistics Society*, Bangi: Universiti Kebangsaan Malaysia.
- . 2009. *The Austroasiatic Central Riverine Hypothesis*. Keynote address read to the *19th Annual Meeting of the Southeast Asian Linguistics Society*, University of Social Sciences and Humanities, Vietnam National University-Ho Chi Minh City, 28 May 2009.
- Sidwell, Paul and Pascale Jacq. 2003. *A Handbook of Comparative Bahnaric: Volume 1, West Bahnaric*. Canberra: Pacific Linguistics 551.
- Skeat, Walter Willian and Charles Otto Blagden. 1906. *Pagan Races of the Malay Peninsula*. 2 Vols. London: Macmillan.
- Smith, Kenneth, D. 1972. *A phonological reconstruction of Proto-North-Bahnaric*. Dallas: Language Data Series, Summer Institute of Linguistics.
- Swadesh, Morris. 1950. Salish internal relationships. *International Journal of American Linguistics* 16. 157–167.

- . 1952. Lexico-statistical dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96. 452–463.
- Theraphan, Luang-Thongkum. 2001. *Languages of the Tribes in Xekong Province Southern Laos*. The Thailand Research Fund.
- Thomas, David. 1973. A Note on the Branches of Mon-Khmer. *Mon-Khmer Studies* 4. 139–40.
- Thomas, David and Marilyn Smith. 1967. Proto-Jeh-Halang. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 20. 157–175.
- Thomas, David and Robert Headley. 1970. More on Mon-Khmer subgroupings. *Lingua* 25. 398–418.
- Thomas, Dorothy M. 1967. *A Phonological Reconstruction of Proto-East-Katuic*. MA thesis, University of North Dakota.
- Thurgood, Graham. 1999. *From ancient Chamic to modern dialects: Two thousand years of language contact and change*. Oceanic Linguistics Special Publications No. 28. Honolulu: University of Hawai'i Press.
- Vovin, Alexander. 1998. Japanese rice agriculture terminology and linguistic affiliation of Yayoi culture. In Roger M. Blench and Matthew Spriggs (eds) *Archaeology and Language II*, 366–378. London: Routledge.
- Witzel, Michael. 1999. Early sources for South Asian substrate languages. *Mother Tongue Special Issue*, 1–76.
- Zide, Arlene R.K. and Norman H. Zide. 1976. Proto-Muṅḍā cultural vocabulary: evidence for early agriculture. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austro-Asiatic Studies, Part II*, 1295–1334. Honolulu: University of Hawai'i.
- Zong, Y, Z Chen, JB. Innes, C Chen, Z Wang and H Wang. 2007. Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature* 449. 459–462.



# 15 *The Austroasiatics: East to West or West to East?*

---

LAURENT SAGART

## 1 Introduction

The Austroasiatic language family extends from Northeast India to Southeast Asia. In which of these two regions did it originate? Both an eastern and a western homeland have been proposed. In this chapter, after reviewing the linguistic, archaeological and genetic evidence, I argue that an array of grammatical and phonological characteristics shared by Austroasiatic, Austronesian and Sino-Tibetan at the earliest level argues for an eastern origin of Austroasiatic. I further suggest that while Diffloth's reconstruction of tropical animal names in proto-Austroasiatic does point to a tropical homeland, a language *ancestral* to proto-Austroasiatic may still have participated in the domestication of rice in the mid Yangzi region. This hypothesis not only accounts for the existence of an original proto-Austroasiatic vocabulary of rice agriculture, it also helps make sense of the phonological and grammatical characteristics Austroasiatic shares with Sino-Tibetan and Austronesian—two groups which I have argued are descended from a common ancestor spoken in north China (Sagart 2005, 2008a, b): whether the shared characteristics are due to contact or to a distant genetic relationship, a period of geographical proximity between languages ancestral to Austroasiatic and Sino-Tibetan-Austronesian must be supposed.

## 2 East Asian linguistic diversity, language families and farming

By common consent, leaving aside the languages of northeast Asia—Korean, Japanese, Ainu, the 'Altaic' (Tungusic, Mongolic, Turkic) and the 'Paleo-Siberian' languages (Chukchi, Yukaghir, Nivkh, Ket etc.)—there are no more than five distinct language families in East Asia: Austronesian (AN), Sino-Tibetan (ST), Tai-Kadai (TK), Austroasiatic (AA) and Hmong-Mien (HM).<sup>1</sup> There are in addition on the southern edges of the region three isolates: Andamanese, Kusunda and Burushaski. Disagreement among linguists as to what languages belong to what group is marginal, although the relationships between groups are disputed; see Sagart, Blench & Sanchez-Mazas (2005). Considering the size of the area and the time depth of modern human presence in the region—perhaps

---

1 I have argued that these are in fact reducible to three, as Tai-Kadai turns out to be a subgroup of Austronesian (Sagart 2004) and since I regard Sino-Tibetan and Austronesian as two branches of a larger family, as mentioned in the introductory paragraph.



50,000 years—this diversity is not high. Smaller areas on the periphery of East Asia, such as New Guinea and eastern Siberia show more diversity for similar time depths. It is likely that East Asia as defined here has experienced a reduction of its linguistic diversity on a large scale. When considered in the broad regional context, it is undoubtedly significant that each of the language families in the region has a vocabulary of rice that reconstructs at proto-language level, while much, if not all, of the rice vocabulary of the three isolates consists of loanwords from one East Asian family or other. This illustrates the reproductive advantage that being spoken by a community of farmers has conferred to east Asian languages since Neolithic times. Expansion of farmer languages in neolithic times is probably the main factor behind the reduction of language diversity in East Asia.

A model linking the formation of language families to the spread of farming was put forward by Renfrew (1987) and Bellwood (1984–85). According to that model, hunter-gatherer populations having shifted to farming will experience an increase in their population density: this will cause these populations to expand in space, together with their languages, at the expense of their hunter-gatherer neighbours and their languages. Although this model in its stronger versions is controversial, especially as it applies to the Indo-European expansion, it may help explain the formation of East Asian language families with a reconstructible farming vocabulary, the reduction of East Asian diversity to no more than five modern language families and three isolates, as well as the reproductive success of farmer languages and the correlative demise of hunter-gatherer languages. Such a process, it should be noted, does not imply the physical extinction of palaeolithic populations and their total replacement by expanding farmers: as the neolithic way of life spreads, members of hunter-gatherer groups (some of which may have already become sedentary) may integrate farming communities as spouses; entire groups may shift to agriculture. Integration into expanding farming societies will in turn provide an incentive for such groups to become bilingual; ultimately bilingual speakers will find an advantage in shifting to the dominant language.

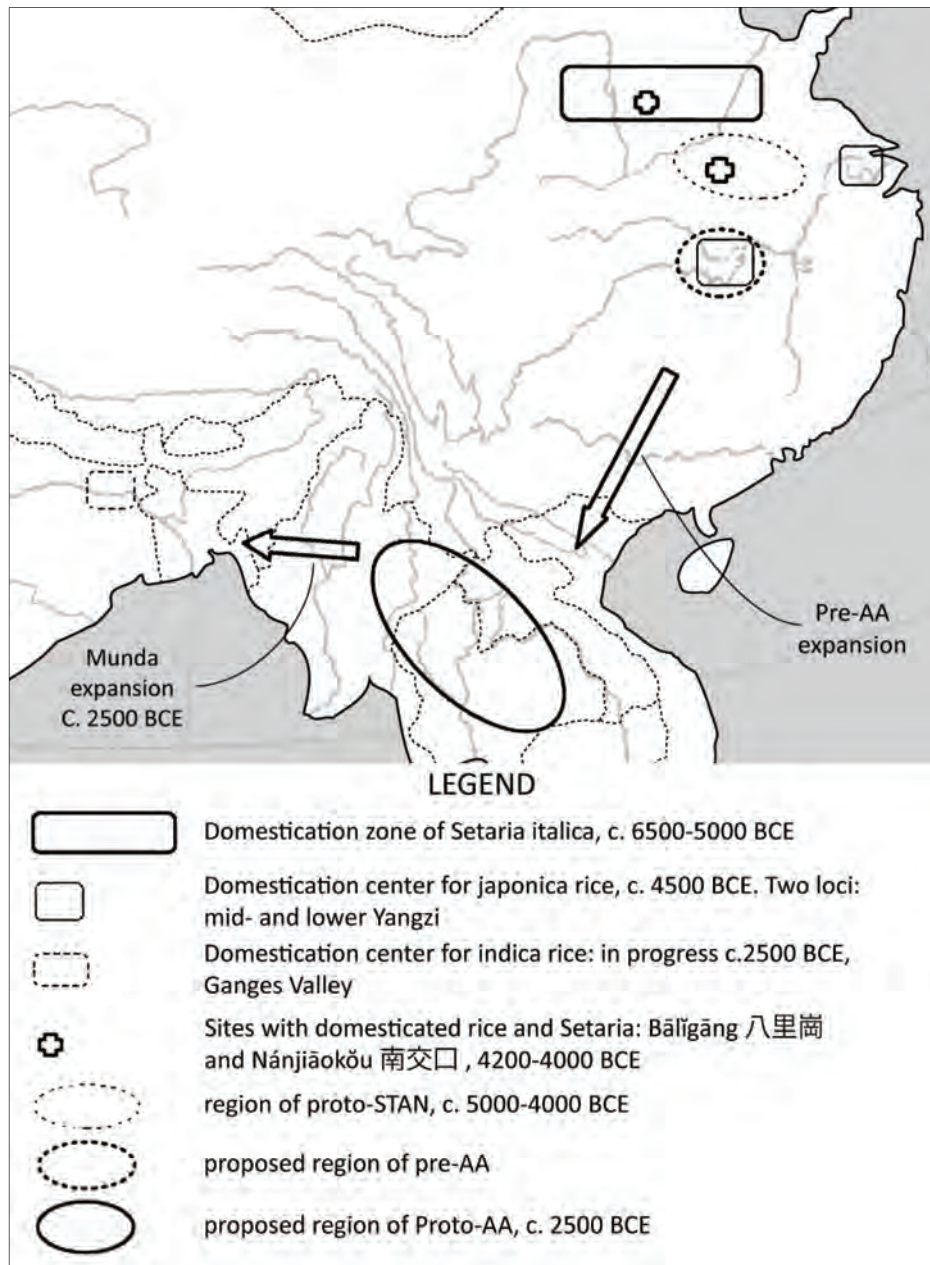
Bellwood (2005) pays particular attention to East Asia, arguing that the AN expansion was driven by rice cultivation. Glover & Higham (1996) and Blust (1998) propose that only one start of cultivation underlies the spread of both AN and AA<sup>2</sup>; Blust hypothesises that the transition occurred in northwestern Yunnan near the borders of Burma, Tibet and Sichuan around 9000 BP. From there, the AAs would have expanded into southeast Asia by following the Mekong and into south Asia along the Brahmaputra, eventually introducing rice cultivation to India. Meanwhile another group, ancestral to proto-Austronesian, would have followed the Yangzi to the east China sea, expanding south along the China coast into Fujian and ultimately Taiwan, where the most recent common ancestor of all living AN languages is thought to have been spoken c. 3500–3000 BCE. The upper Yangzi region, however, has not, so far, yielded any early dates for agriculture: Rispoli (2007) argues from pottery styles that neolithic expansion from the mid Yangzi into southeast Asia was southward rather than upstream along the upper Yangzi, implying that Yunnan was a cul-de-sac.

In Sagart (2008a) I proposed a different Asian version of the farming/language model. I claimed that the expansion of a macrophylum consisting of Sino-Tibetan and Austronesian ('Sino-Tibetan-Austronesian' or STAN) was driven by the possession of *two* domesticated cereals: *Oryza sativa japonica* and *Setaria italica*. I placed the epicentre of the expansion in the Císhān-Péilǐgāng culture area of North and Northeast China around 8500 BP. From

---

2 The idea that the AA and AN families together constitute the 'Austic' macrophylum (Schmidt 1906; Shorto 1976; Hayes 1992; Reid 1994; Blust 1998) remains controversial.

there a western branch led to Sino-Tibetan and a south-eastern branch to PAN and Taiwan. However, following the demonstration by Fuller, Harvey & Ling (2007) that Chinese rice was not fully domesticated before 6500 BP, and that currently the earliest sites with both *domesticated* rice and *Setaria* appear no earlier than 6200 BP in the mid-Yangshao culture of Henan, the dates of proto-STAN must now be regarded as considerably less ancient—perhaps as late as c. 6200 BP in or near Henan (Sagart 2008b)—yet still in time for an entry in Taiwan in the late 4th mill. BCE.



**Figure 1:** Proposed pre-AA and proto-AA homelands in relation to Asian centres of cereal domestication.

A disputed issue is whether rice was domesticated once, or more than once. On archaeological evidence the Yangzi Valley has the oldest sites with domesticated rice, with two subloci, in the lower and the mid Yangzi (Figure 1). Until recently the orthodoxy was

that rice was domesticated in the Yangzi valley as early as 9000 or 10,000 BCE but Fuller, Harvey & Ling (2007) and Fuller & Ling (2008) have convincingly argued that the earliest Chinese sites with rice remains had sedentism without agriculture: rice was collected wild, together with other foods like acorns and foxnuts. They argued that rice was put to cultivation later, and domesticated only considerably later—domestication requires a history of selection of desirable characters and their fixation in the plant, so that the domesticated plant is different from its wild progenitor. In particular, non-shattering, a key domestic trait, only became fixed in lower Yangzi valley rice in the late 5th mill. BCE; the mid Yangzi following an approximately parallel temporal course. It may have taken 2000 or 3000 years between the moment when rice was first put into cultivation until non-shattering was fixed. If so, cultivation may have started around 7000–6000 BCE in the Yangzi valley. Before non-shattering was fixed, rice had to be harvested unripe, as in Hemudu c. 5000 BCE (Fuller, Harvey & Ling 2007). The Yangzi may not be the only region where rice was domesticated: Fuller (2006) argues that Gangetic India may represent another region, with rice cultivation prevalent in a large area around the Ganges Valley by 3000 BCE.

Languages with a reconstructible vocabulary of cereal cultivation, especially rice, have names for such notions as the rice plant, different plant parts, especially its grains (with and without the husk, polished or not, cooked or not), the field in which it is cultivated (irrigated/naturally flooded/swidden), the tools used in rice cultivation (dibble stick/reaping knife/sickle) and verbs for actions such as sowing/planting, weeding, reaping, threshing, pounding and winnowing. Sagart (2003) compared the reconstructible vocabularies of cereal cultivation (rice and *Setaria italica*) across East Asian families. He showed that this vocabulary is shared to a significant degree by AN and ST, not as a result of borrowing but as inherited cognates—in line with the evidence for a genetic relationship between them (Sagart 2005 for a recent statement of that evidence), while this vocabulary is completely different in the AA family, it now appears that the Dravidian vocabulary of rice (as described in Krishnamurti 2003 and Southworth, forthcoming) is also completely distinct from the others. This probably indicates at least three independent starts of rice cultivation, one to account for the AA vocabulary, another for the ST-AN vocabulary and a third for the Dravidian vocabulary. TK appears to have lost much (but not all) of its original agricultural vocabulary after coming into contact with AA and borrowing AA rice vocabulary. This picture does not support an Austric (AA+AN) rice-driven expansion. If AN and AA had gone through rice domestication as one language, one would expect to find at least traces of a common vocabulary of rice agriculture in them. The picture is consistent with the McCouch-Kovach model which sees separate domestications of *Oryza sativa indica* and *Oryza sativa japonica* coupled with limited introgression of key domestication features from japonica into indica (Kovach, Sweeney & McCouch 2007). The complete separation of the Dravidian and Asian rice vocabularies is an unexplained element in the snowball model of Vaughn, Lu & Tomooka (2008), which supposes only one domestication, in China, to account for the genetic variety of Asian domesticated rice.

With this background in mind, we now turn to the circumstances surrounding the spread of Austroasiatic.

The Austroasiatic language family is split into two geographical clusters: the Munda group in Northeast India and the Mon-Khmer group in southeast Asia. Khasi is a Mon-Khmer outlier in Eastern India. Munda and Mon-Khmer are generally portrayed as the two

primary branches of the family.<sup>3</sup> Was the AA homeland in India or in southeast Asia? Both positions have been defended. Other proposals include a central homeland, in a region encompassing parts of Burma and Bangladesh and a northern homeland, or at least early presence, in the Yangzi valley and southeast China coast.

### 3 Proposed AA homelands

#### 3.1 Southeast Asia

Glover & Higham (1996:419),<sup>4</sup> Higham (2002, 2009) and Bellwood (2005) have argued for a southeast Asian origin of the AAs: a migration in the 3rd mill. BCE would have brought the Mundas to Eastern India and, with them, rice cultivation. However, if agriculture had been introduced into the subcontinent by the Mundas, one would expect the techniques to have been transmitted together with the attendant vocabulary: the complete absence of Austroasiatic words in the rice vocabulary of south Asian languages, whether Indo-Aryan or Dravidian, is a serious problem for the view that rice cultivation was introduced into India by the AAs. The idea of an AA homeland in southeast Asia does however find solid support from population genetics (Kivisild et al. 2003; Sahoo et al. 2006; Sengupta et al. 2006; Chaubey et al. 2008): while the mtDNA of the Mundas is indistinguishable from that of neighbouring populations speaking Indo-Aryan or Dravidian languages, the Mundas have a high incidence of the Y-chromosome haplogroup O-M95, which is otherwise largely southeast Asian. Kayser et al. (2003) place the expansion of the O-M95 haplogroup in southeast Asia at 2400 BCE, a reasonably good fit with the dates proposed by of Glover & Higham, and Bellwood, for the AA movement into India.

#### 3.2 South Asia

Certain linguists have argued in favour of a very early AA presence in NW India. Kuiper (1948, 1950) found evidence for a non-Indo-Aryan, non-Dravidian substratum in the earliest Rgveda, arguing these words came from a prefixing language spoken in India before the Rgveda: this fits the Munda languages, which are prefixing, while excluding the suffixing Indo-Aryan and Dravidian languages. Witzel (2000), following Kuiper, calls his hypothetical prefixing language ‘para-Munda’, meaning a third, far-western branch of AA. Kuiper and Witzel identified few actual Munda or Austroasiatic words in the Rgveda, however, and the mysterious prefixing language could belong to other prefixing languages in the region: Burushaski, Tibeto-Burman, or an extinct group. An Indian homeland for AA receives little support from population genetics, despite Kumar et al. (2007) who defend it based on the absence among the Munda of Y-chromosome haplogroup O-M122, a southeast Asian marker. An alternative explanation for the absence of haplogroup O-M122 could be a population bottleneck followed by drift eliminating this haplogroup. The dates Kumar et al. envision for the split between the Munda and the Mon-Khmer are so early—65,000 BP—as to make little sense for a reconstructable language family like AA anyway.

---

3 Diffloth (2005) identifies a third, northern branch, which dominates Khasi and Khmuic. The evidence for his 3-branch analysis has not been published.

4 ‘In summary we can say that, towards the end of the 3rd millennium BC, rice, including domesticated varieties, appeared among the small-scale neolithic farming communities of the central and eastern parts of the Ganga valley, perhaps brought by communities of farmers speaking Proto-Munda languages expanding down the Brahmaputra valley from a homeland in the region of Yunnan and upper Burma’.

### 3.3 Burma/Bangladesh

Diffloth (2005) reconstructed an extensive PAA vocabulary of rice. He found no vocabulary relating to irrigation to be common to the Mundas and MKs. He also reconstructed the names of several tropical animals (like the tree monitor and pangolin), which he says argue for a tropical homeland in Burma or Bangladesh. He feels that this invalidates the idea of a Yangzi origin. Diffloth may well be right that the PAA homeland was located in the tropics,<sup>5</sup> yet strictly speaking this only applies to the most recent common ancestor of attested AA languages: PAA itself could be the daughter of a language ('pre-AA') spoken earlier by a Yangzi valley population having participated in the transition to cultivation of rice. Thus Diffloth's PAA homeland and the view that the population who spoke PAA had migrated south from the mid-Yangzi region are not incompatible.

### 3.4 South China

The idea of an early presence of AA languages in south China and more specifically on the Southeast China coast was first defended by Norman & Mei (1967). These authors claimed to have identified AA loanwords (or possibly substratum words) in Southeast Chinese dialects like Min, and as well as observations of a lexical nature in early Chinese texts dealing with the Southeast China coast. Pulleyblank (1983) pushed their argument to include the entire East China coast, including the Northeast (Shandong), on rather flimsy grounds (a couple of phonetically similar place names). Norman (1985) developed an argument that some of the Chinese twelve year names and associated animal referents have their sources in AA animal names. Schuessler's recent etymological dictionary of Chinese lists AA forms which are treated as AA loans into Chinese (Schuessler 2007, in particular page 23). The treatment of foreign loanwords into Chinese by Wiebusch & Tadmor (2009) reflects these views. The idea of an AA presence in early south China has nothing impossible to it, yet current evidence for a layer of AA loanwords into Chinese does not stand up to scrutiny. Sagart (2008a) found problems with the most salient elements of the lexical evidence marshalled by Norman & Mei: thus the Min word for 'shaman' is really 童, a Chinese word for 'child', with semantic extension to 'boy, man-servant' (of a god); the 'southern Yue' word for 'dog' looks Austronesian more than Austroasiatic; and the Min word for 'to know' is really 別 'to separate' with semantic extension to 'distinguish, discriminate, know'. Norman's AA comparisons for the Chinese year names also look quite dubious especially when formulated in more recent OC reconstruction systems than Li (1971, 1976): for instance *yōu* 酉 'year of the cock', Old Chinese \*rəgwɣ in Li's reconstruction, is compared to an AA form *raka* 'fowl' by Norman. More recent systems of reconstruction give l-, j- or ɣ- for the initial and -u? for the final. The AA-Chinese comparisons in Schuessler's dictionary are unconstrained by sound correspondences: it is not clear what makes them better regarded as loanwords than as look-alikes. Some of them are demonstrably wrong: for instance 昵, Middle Chinese *nrit* 'close-standing, familiar, intimate' is compared with Khmer /cumnit/ (an infix form: j<mn>it) 'nearness, closeness, proximity'. But the Chinese word belongs to the word-family of 暱 'near, familiar', which ends with -k: Middle Chinese *nrik*. Alternation between MC *-it* and *-ik* unambiguously points to Old Chinese \*-ik (Baxter 1992): here we must reconstruct OC \*n<r>ik, which invalidates the comparison with AA. The observation in Norman & Mei

---

5 The mean annual temperatures in China in the period 8000–3000 BP were 2 to 5 degrees C° higher than today (Chang Kwang-Chih 1986).

(1967) that the Chinese name of the Yangzi river: 江, Old Chinese<sup>6</sup> \*k<sup>h</sup>roŋ (Mandarin *jiāng*) sounds almost identical to a widespread AA word for ‘river’ remains intriguing, however, and may support an early AA (or pre-AA) presence in south-central China, though not on the Southeast coast.

#### 4 Rice

It is generally accepted that domesticated rice, *Oryza sativa*, consists of two main subspecies, *Oryza sativa japonica* and *Oryza sativa indica*; the Aus rices of Bangladesh are genetically close to *indica* while the west Asian aromatic rices (‘basmati’) are genetically within the *japonica* group (Garris et al. 2005). *Japonica* rices are cultivated in temperate regions (northern China, Korea, Japan) and in elevated areas in tropical regions (‘tropical japonicas’) in mainland and island southeast Asia. *Indicas* are cultivated in south Asia and in the valley bottoms in Southeast Asia. The lowland versus upland pattern of *indica* and *japonica* rices in southeast Asia, with upland areas isolated by interconnected lowlands, suggests that *indicas* spread in the valley bottoms at an unknown date, in effect confining the *japonica* rices to the upland areas.<sup>7</sup> The Mon-Khmers conform to the Southeast Asia pattern by cultivating *japonica* in upland fields and *indica* in lowland fields, while the Mundas conform to the south Asian pattern by cultivating only *indica* rices.

The dendrogram of domesticated rices in Garris et al. (2005) shows two main clusters, corresponding to the traditional *japonica* versus *indica* distinction. The divergence, it is now recognised, occurred between 100,000 and 400,000 years ago, far earlier than the period during which rice was domesticated. This argues in favour of separate domestications from already distinct wild varieties of wild rice. Based on a comparison of wild and domesticated forms, Londo et al. (2006) placed the domestication of *indica* south of the Himalayas and that of *japonica*, in south China. Kovach et al. (2007) showed that the domestication processes were not completely independent: three crucial domestication genes are common to the two subspecies, clearly the result of introgression.<sup>8</sup> In each case so far, introgression was from *japonica* into *indica*. Moreover, other domestication genes are confined within each subspecies, as would be expected if there had been two or more domestications. The McCouch-Kovach model (Kovach et al. 2007) supports multiple domestications coupled with limited introgression that transferred key domestication alleles primarily, and perhaps exclusively, from *japonica* into *indica*. According to the model proposed here, Tibeto-Burman speakers (who were *indica* rice farmers) are present at the northern edge of the Indian subcontinent at least since the 1st millennium BCE. Contact between them and Ganges Valley *indica* rice farmers provides a possible channel for the transmission of those genes from the *japonicas* into the still outcrossing *indicas*.

A more linear scenario is presented by Vaughn et al. (2008). They argue that the selection of the *sh4* non-shattering allele is the key event of rice domestication and that it occurred only once, in the Yangzi valley. This scenario is consistent with one-domestication scenarios like Bellwood’s (Bellwood 2005), although it must be kept in mind that the language/farming hypothesis is in principle compatible with scenarios involving any number of independent domestications. As mentioned earlier, from a

---

6 OC reconstructions are in the Baxter-Sagart 0.99 system.

7 D. Fuller informs me that there are lowland tropical *japonicas* in Southeast Asia and Island Southeast Asia.

8 Infiltration of the genes of one species into the gene pool of another through repeated backcrossing of an interspecific hybrid with one of its parents.

linguistic point of view, the existence of at least three independent pools of rice vocabulary (Sino-Tibetan-Austronesian; Austroasiatic; Dravidian) is more in line with multiple domestications.

## 5 Linguistics: a typological argument for an eastern AA homeland

There have been many attempts at demonstrating possible genetic links between AA and Austronesian ('Austric' theory). Unfortunately, the lack of a PAA reconstruction has made the issue difficult to tackle. The recent publication of Shorto's (2006) reconstruction of Mon-Khmer has improved the situation somewhat, providing MK material to extend existing PAn-OC sound equations (Sagart 1993), see Table 1:

**Table 1:** Sound equations between AN, OC and MK

	<b>crossbow/shoot</b>	<b>spit/vomit<sup>#</sup></b>
P-Austronesian (consensus forms) *-aq	*panaq 'shoot' (< prefix pa- + naq 'bow' ?)	*utaq 'vomit'
OC (Baxter-Sagart 2008) *-ʰaʔ	*nʰaʔ 'crossbow'	*tʰaʔ 'eject from mouth'
PMK (Shorto 2006) *-aʔ	*s-naʔ 'crossbow'	*s-taʔ 'vomit'

<sup>#</sup> Also reflected in TB: Lushai *chhak* < \*tha:k 'spit'.

Other observations, bearing on lexical items and morphological processes shared by AA and AN, or by AA and ST, can be gleaned from the literature (for instance Reid 2005:146–147 for Austric morphological parallels; Sagart 1995:214 for AA-AN-ST morphological parallels). In the absence of a reconstruction of PAA (not just Mon-Khmer), the hypothesis of a genetic relationship of AA with Sino-Tibetan and Austronesian (Starosta 2005) remains impossible to establish or to falsify, however. A more immediately rewarding approach of the AA homeland question is typological.

### 5.1. East Asian character of PAA structural characteristics

Structural features of languages are known to behave in an 'areal' way. Whether related or not, languages spoken in the same region for a long period tend to acquire similar structural characteristics: word structure, prosody, morphology, constituent order in the sentence, etc. If two groups of languages in different regions of the world exhibit highly similar structural types, one reason can be that they once occupied the same region.

On the surface, the MK and Munda groups have very different structural characteristics: the MK languages are of a markedly southeast Asian type: words are either monosyllabic or iambisyllabic (that is, consisting of an unstressed and reduced syllable followed by a full, stressable syllable), like Sino-Tibetan; nontonal (tones are recent in East Asia, probably no more than 2000 years old); morphology is derivational, making use of prefixes, infixes (inserted between the first consonant and the first vowel: this is like Sino-Tibetan and Austronesian). Word order is Verb-Object, with prepositions, like Austronesian, Chinese and a handful of TB languages such as Maru. Munda is the reverse of MK on many points. Its structural features are as unremarkable in the south Asian context as are the indica rices the Mundas cultivate: words have initial stress, word order is Object-Verb with postpositions, and while morphology has prefixes and infixes like MK, it also has suffixes, like the neighbouring Indo-Aryan and Dravidian languages. Suffixes,

however, appear to have been innovated separately within the various branches of Munda: they are not reconstructible to proto-Munda, let alone PAA (Stampe & Donegan 2004). This is what one would expect if the Munda languages had become suffixing after the breakup of Proto-Munda. From the fact that the order of morphological formatives in Munda verb stems is parallel to the order of constituents in the verb phrase in Mon-Khmer, Pinnow (1960) built a convincing case that the order of constituents in PAA sentences was like MK, not like Munda. His conclusions are accepted by Lehmann (1973) and Stampe & Donegan (2004). It sounds very likely, then, that Munda shifted towards south Asian structure as a result of becoming immersed among south Asian populations speaking languages with south Asian characteristics, a point made by Pinnow and Lehmann.<sup>9</sup> Anderson (2004:174) similarly argues that proto-south-Munda had subject prefixes on the verb, and that these were replaced by suffixes in some languages as a result of influence from Dravidian or Indo-Aryan. This, then, indicates that proto-Munda had a more marked east Asian typology than modern Munda, which makes East Asia a more likely homeland for PAA than south Asia.

## 5.2 Another shared structural characteristic of AA and AN: final syllables as roots

We will now present evidence for yet another typological similarity between Austroasiatic and an East Asian language. Anderson (2004) observed that in the AA languages, whether in Munda or in MK, the last syllable of words often behaves as if it was a stable, meaning-associated element, while the first syllable varies from language to language, cannot be assigned any meaning and is not reconstructible. He argues that the AA languages have a constraint requiring words to have two morae at the phonetic level; when words underlyingly have only one mora-carrying vowel, the AA languages make use of various strategies to add material, often at the beginning of the word, and make the word bimoraic. These strategies, he claims, result in families of words having the same final syllable, related meanings, and different initial syllables which cannot be identified to known morphemes.

Essentially the same situation exists in the AN languages, where stable final syllables of disyllables, known to Austronesianists as ‘roots’, occur preceded by meaningless, non-morphological and highly variable first syllables. For instance according to Blust (1988:83), root \***baw** ‘high; top’ occurs as the second syllable of three reconstructible words: \***babaw** ‘upper surface’, \***Sa(m)baw** ‘high, up, on top’ and \***ti(m)baw** ‘height, be high’, as well as in at least eight isolated words in eight different languages, with similar meanings: Bikol *labaw* ‘protruding above’, Cebuano *tugbaw* ‘lofty, high up’, *usbaw* ‘rise in degree or quantity’, etcetera. Blust (1988) reconstructs over 200 of these roots. He regards them as submorphemic, sound-symbolic units, of the same kind as *gl-* in Eng. *glow*, *gleam*, *gloom* etcetera, but a preferable explanation is that roots are the missing monosyllabic words of Austronesian: they occur almost exclusively at the end of disyllables because Austronesian has a similar constraint on words as Austroasiatic. This constraint is known as the ‘drive towards disyllabism’ in Austronesian studies. Wolff

---

9 Stampe & Donegan dispute this. They observe that the traits relating to word order, use of pre- or postpositions, prefixing or suffixing morphology and iambic or trochaic word stress are linked; that view is not controversial. They argue that Munda acquired south-Asian-like structural features as a result of a change in stress, which triggered all the other changes. They think the AA homeland was in south Asia (2004 and p.c., 2008). However, the original change in stress pattern may have been induced by a move into a different typological environment.



(2007) describes the strategies used by AN languages to make disyllables from monosyllabic roots.

One particular strategy used by Austroasiatic languages is to break the main vowel of a monosyllable by a glottal stop (hypothetical example: mak > maʔak). This strategy is used by Munda languages like Sora and by Mon-Khmer languages like Kontoi (Anderson 2004), which makes it a likely candidate for PAA antiquity. The same is true of Austronesian languages like Bunun and Cebuano: in Bunun, a language of Taiwan (Wolff, forthcoming), monosyllables (except those formed recently through loss of a medial consonant) are expanded by a glottal stop inserted in the vowel, for example *laʔas* ‘nutritious substance’, opposite *tilas* ‘rice or millet’. The Bunun root *-las* reflects a Proto-AN monosyllabic root \*Ras which occurs with different elements on the left in PAN \*beRas ‘dehusked rice’, and in other words relating to rice in the northern Philippines. Cebuano, a Philippine language, systematically expands monosyllables through glottal stop insertion (Wolff 2007; Wolff uses the term ‘stretching’). Aside from Bunun and Cebuano, monosyllables expanded with -ʔ- occur here and there in Formosa, the Philippines and elsewhere in insular southeast Asia, showing the process has much antiquity in the AN family.

Like the typological features described earlier, the bimoraic/disyllabic word constraint and attendant strategies are shared structural characteristics of AA (including Munda) and AN. See Table 2 below (the Tai-Kadai and Hmong-Mien families, not shown, behave like Chinese):

**Table 2:** Features shared by Austroasiatic and the STAN languages  
(Austronesian, Chinese, Tibeto-Burman)

	prefix	infx	suffix	derivational morphology	inflectional morphology	order of constituents	bimoraic/ -ʔ-insertion into disyllabic constraint	-ʔ-insertion into monosylls.
AA	++	+	-	+	-	VO	+	+
AN	++	+	-	+	-	VO	+	+
Chinese	++	+	+	+	-	VO	-	-
TB	++	?	+	+	-	OV	-	-

Table 2 shows that Mon-Khmer and Munda structural characteristics are close to those of AN and ST. True, expanding monosyllables into disyllables or bimoraic words through insertion of a glottal stop is not limited to East Asian languages: Danish original monosyllabic morphemes have an inserted glottal stop (known as ‘stød’: Basbøll 2005) which does not appear in compound words; a similar process exists in Latvian.<sup>10</sup> Yet, this characteristic is not common world-wide: its sharing by AA and AN seems significant.

Moreover, while some of the features in Table 2—VO versus OV word order, use of prepositions or postpositions, prefixation versus suffixation—are interdependent features, inserting a glottal stop in the middle of monosyllables to satisfy a bimoraic constraint appears to be unrelated to, and independent from, VO constituent order and related characteristics.

<sup>10</sup> I am indebted to Bill Baxter for pointing out the relevance of the Danish stød to me.

## 6 Conclusion

The shared linguistic typology just described can be due to a very old genetic relationship between AA and STAN—impossible to substantiate or to reject in the current state of our knowledge—or to diffusion. In either case, a period of geographical closeness between languages ancestral to AA and STAN must be assumed. The area where this typology originated was most likely East Asia, since, as I have argued (Sagart 2008a), the homeland of the STAN family was in northern China.

Although AA and STAN must at some point in their history have occupied neighbouring regions, the fact that AA and STAN have entirely distinct rice vocabularies (Sagart 2003) probably means that the two families went through rice domestication independently. As mentioned earlier, in addition to rice, the proto-STAN people had a domesticated millet: *Setaria italica*. That places the STAN homeland between the Yangzi and Huang He valleys, where both cereals are found together in domesticated form from the late 5th millennium BCE in sites like Bālǐgāng 八里崗 in south Henan (Dorian Fuller, personal communication, May 9, 2008) and Nánjiāokǒu 南交口 in northwest Henan (Wei et al. 2000). Baligang is located in the northern part of the Yangzi catchment. It is possible that a population speaking a language ancestral to proto-Austroasiatic had been participating in the domestication of rice in the lakes area of the mid Yangzi valley, not far to the south-west of Henan, some time before rice and *Setaria* began to overlap further north, in Henan. Rice could have reached Henan from other centres of domestication than the mid Yangzi: this, at any rate, is what the lack of contacts in the rice vocabularies of AA and the STAN languages suggests.

From the mid-Yangzi lakes region, these early rice cultivators would expand in a southerly or southwesterly direction, presumably down the Xiāng 湘 or Yuán 沅 river valleys, in line with the observations of Rispoli (2007). They would eventually reach tropical southeast Asia. There they would evolve into the most recent common ancestor of modern-day Austroasiatic languages: PAA. At some point, perhaps in the third millennium BCE, a group of Austroasiatic speakers would migrate from tropical southeast Asia to northeastern India, where the indigenous population was already engaged in domesticating a local variety of rice (see Figure 1).

This admittedly speculative scenario is presented because it allows one to make sense of linguistic facts relating to reconstructible lexicon (rice, tropical animals) and to phonological and grammatical typology, in a manner consistent with currently available evidence from archaeology and population genetics. As always, falsification of its predictions is invited.

## Acknowledgements

Thanks go to Susan McCouch for numerous conversations on rice domestication and to Dorian Fuller for useful observations on a preliminary version of this paper. They should not be held responsible for the views developed here. Thanks also to workshop participants for comments and criticisms. Figure 1 was prepared by Peter Johan Nijland.

## References

- Anderson, Gregory D.S. 2004. Advances in Proto-Munda reconstruction. *Mon-Khmer Studies* 34. 159–184.
- Basbøll, Hans. 2005. *The Phonology of Danish*. Oxford: Oxford University Press.

- Baxter, William H. 1992. *A Handbook of Old Chinese phonology*. Trends in Linguistics Studies and Monographs 64. Berlin: Mouton de Gruyter.
- Baxter, William H. & Laurent Sagart. 2008. *Old Chinese: the Baxter-Sagart reconstruction (Version 0.98)*. 41st International Conference on Sino-Tibetan languages and Linguistics. London: September 20, 2008.
- Bellwood, Peter. 1984–1985. The Austronesian dispersal and the origin of languages. *Asian Perspectives* 26. 107–117.
- . 2005. Examining the Farming/language hypothesis in the East Asian context. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia*, 17–30. London: RoutledgeCurzon.
- Blust, Robert. 1988. *Austronesian root theory*. Amsterdam: John Benjamins.
- . 1998. Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. In Ward H. Goodenough (ed.) *Prehistoric settlement of the Pacific*. Philadelphia: American Philosophical Society.
- Chaubey, Gyaneshwer, Monika Karmin, Ene Metspalu, Mait Metspalu, Deepa Selvi-Rani, Vijay Kumar Singh, Jüri Parik, Anu Solnik, B. Prathap Naidu, Ajay Kumar, Niharika Adarsh, Chandana Basu Mallick, Bhargav Trivedi, Swami Prakash, Ramesh Reddy, Parul Shukla, Sanjana Bhagat, Swati Verma, Samiksha Vasnik, Imran Khan, Anshu Barwa, Dipti Sahoo, Archana Sharma, Mamoon Rashid, Vishal Chandra, Alla G. Reddy, Antonio Torroni, Robert A. Foley, Kumarasamy Thangaraj, Lalji Singh, Toomas Kivisild & Richard Villems. 2008. Phylogeography of mtDNA haplogroup R7 in the Indian peninsula. *BMC Evolutionary Biology* 8. 227. Published online August 4, 2008.
- Chang, Kwang-Chih. 1986. *The archeology of ancient China*. Newhaven and London: Yale University Press. 4th edition.
- Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austroasiatic. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*, 77–89. London: RoutledgeCurzon.
- Fuller, Dorian Q. 2006. Agricultural Origins and Frontiers in South Asia: A Working Synthesis. *Journal of World Prehistory* 20. 1–86.
- Fuller, Dorian Q., Emma Harvey & Ling Qin. 2007. Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81. 316–331.
- Fuller, Dorian Q. & Ling Qin. 2008. Evidence for a late onset of agriculture in the Lower Yangtze region and challenges for an archaeobotany of rice. In Alicia Sanchez-Mazas, Roger Blench, Malcom Ross, Iliia Peyros & Marie Lin (eds) *Past Human Migrations in East Asia*. Routledge studies in the Early History of East Asia, 40–83. London: Routledge.
- Garris, Amanda, Thomas Tai, Jason Coburn, Steve Kresovich & Susan R. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169(3). 1631–8.
- Glover, Ian & Charles Higham. 1996. New evidence for early rice cultivation in South, Southeast and East Asia. In David Harris (ed.) *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, 413–441. Smithsonian Institution Press: Washington, DC.
- Hayes, La Vaughn H. 1992. On the Track of Austric, Part I: Introduction. *Mon-Khmer Studies* XXI. 143–77.

- Higham, Charles. 2002. Languages and farming dispersals: Austroasiatic languages and rice cultivation. In Peter Bellwood & Colin Renfrew (eds) *Examining the Farming/language dispersal hypothesis*, 223–232. Cambridge: McDonald Institute.
- Higham, Charles. 2009. East Asian Agriculture and its impact. In Chris Scarrre (ed.) *The Human past, Second edition*, 234–263. London: Thames and Hudson.
- Kayser, Manfred, Silke Brauer, Gunter Weiss, Wulf Schiefenhovel, Peter Underhill, Peidong Shen, Peter Oefner, Mila Tommaseo-Ponzetta & Mark Stoneking. 2003. Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *The American Journal of Human Genetics* 72. 281–302.
- Kivisild, Toomad, Siiri Rootsi, Mait Metspalu, Sarabjit S. Mastana, Katrin Kaldma, Jüri Parik, Ene Metspalu, Maarja Adojaan, Helle-Viivi Tolk, V Stepanov, Mukaddes Golge, Esien Usanga, Surinder S. Papiha, Cengiz Cinnioglu, Roy King, Luigi Cavalli-Sforza, Peter A. Underhill & Richard Villems. 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *The American Journal of Human Genetics* 72. 313–332.
- Kovach, Michael J., Megan T. Sweeney & Susan R. McCouch. 2007. New insights into the history of rice domestication. *Trends in Genetics* 23(11). 578–587.
- Krishnamurti, Bhadriraju. 2003. *The Dravidian languages*. Cambridge: Cambridge University Press.
- Kuiper, Fransiscus B. J. 1948. *Proto-Munda words in Sanskrit*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- 1950. *An Austro-Asiatic myth in the RV*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- Kumar, Vikrant, Arimanda N.S. Reddy, Jagedeesh P. Babu, Tipirisetti N. Rao, Banrida T. Langstieh, Kumarasamy Thangaraj, Alla G. Reddy, Lalji Singh & Battini M. Reddy. 2007. Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BMC Evolutionary Biology* 7. 47.
- Lehmann, Winfred P. 1973. A structural principle of language and its implications. *Language* 49. 47–66.
- Li Fang-kuei. 1971. Shang-ku In Yen-chiu. *Tsinghua Journal of Chinese Studies, New Series IX, numbers 1 and 2*. 1–61.
- Li Fang-kuei. 1976. Chi-ke Shang-ku Sheng-mu Wen-ti. In Szu Ch'i-liang et al. (eds), *Tsung-tung Chiang Kung Shih-shih Chou-nien Chi-nien Lun-wen-chi*, 1143–50. Taipei: Academia Sinica.
- Londo, Jason P., Yu-Chung Chiang, Kuo-Hsiang Hung, Tzen-Yuh Chiang & Barbara A. Schaal. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences* 103(25). 9578–583.
- Norman, Jerry. 1985. A note on the origin of the Chinese duodenary cycle. In Graham Thurgood, James A. Matisoff & David Bradley (eds) *Linguistics of the Sino-Tibetan area: the state of the art*. 85–89. Pacific Linguistics: Canberra.
- Norman, Jerry. & Tsu-lin Mei. 1976. The Austroasiatics in ancient south China: some lexical evidence, *Monumenta Serica* 32. 274–301.
- Pinnow, Hans-Jürgen. 1960. Über den Ursprung der voneinander abweichenden Strukturen der Munda und Khmer-Nikobar Sprachen. *Indo-Iranian Journal* 4. 81–103.

- Pulleyblank, Edwin G. 1983. The Chinese and their Neighbors in Prehistoric and Early Historic Times. In David N. Keightley (ed.) *The origins of Chinese Civilization*, 411–466. Berkeley, Los Angeles, London: University of California Press.
- Reid, Lawrence. 2005. The current status of Austric. A review and evaluation of the lexical and morphosyntactic evidence. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*, 134–160. London: RoutledgeCurzon.
- 1994. Morphological evidence for Austric. *Oceanic Linguistics* 33(2). 323–344.
- Renfrew, Colin. 1987. *Archaeology and Language*. London: Jonathan Cape.
- Rispoli, Fiorella. 2007. The incised and impressed pottery style of mainland southeast Asia: following the paths of neolithization. *East and West* 57. 235–304.
- Sagart, Laurent. 1993. Chinese and Austronesian: evidence for a genetic relationship. *Journal of Chinese Linguistics* 21(1). 1–62.
- 1995. Some remarks on the Ancestry of Chinese. In William S.-Y. Wang (ed.) *The Ancestry of the Chinese Language*. Journal of Chinese Linguistics monograph series no. 8. 195–223.
- 2003. The vocabulary of cereal cultivation and the phylogeny of East Asian languages. *Bulletin of the Indo-Pacific Prehistory Association* 23(1). 127–136.
- 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444.
- 2005. Sino-Tibetan-Austronesian: an updated and improved argument. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*, 161–176. London: RoutledgeCurzon.
- 2008a. The expansion of setaria farmers in east asia: a linguistic and archaeological model. In Alicia Sanchez-Mazas, Roger Blench, Malcom Ross, Ilia Peiros & Maria Lin (eds) *Past human migrations in East Asia: matching archaeology, linguistics and genetics*. Routledge Studies in the Early History of Asia, London: Routledge.
- 2008b. *Reconstructable vocabulary and the time, homeland and way of life of proto-Sino-Tibetans*. Workshop on New Directions in Historical Linguistics, Lyon, May 12–14.
- Sagart, Laurent, Roger Blench & Alicia Sanchez-Mazas. 2005. Introduction. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The peopling of East Asia: putting together archaeology, linguistics and genetics*, 1–14. London: RoutledgeCurzon.
- Sahoo, Sanghamitra, Anamika Singh, G Himabindu, Jheelam Banerjee, T Sitalaximi, Sonali Gaikwad, R Trivedi, Phillip Endicott, Toomas Kivisild, Mait Metspalu, Richard Villems & VK Kashyap. 2006. A prehistory of Indian Y chromosomes: evaluating demic diffusion scenarios. *Proceedings of the National Academy of Sciences of the United States of America* 103. 843–848.
- Schmidt, Wilhelm. 1906. *Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentralasiens und Australasiens*. Braunschweig: Friedrich Vieweg und Sohn.
- Schuessler, Axel. 2007. *ABC etymological dictionary of Chinese*. Honolulu: University of Hawaii Press.
- Sengupta, Sanghamitra, Lev A. Zhivotovsky, Roy King, Seyed Quasem Mehdi, Christopher A. Edmonds, Cheryl-Emiliane Chow, Alice A. Lin, Mitashree Mitra, Samir K. Sil, A Ramesh, MV Usha Rani, Chitra M Thakur, Luigi L. Cavalli-Sforza, Partha P. Majumder & Peter A. Underhill. 2006. Polarity and temporality of high-resolution Y-chromosome distributions in

- India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *The American Journal of Human Genetic* 78. 202–221.
- Shorto, Harry L. 1976. In Defense of Austric. *Computational analyses of Asian and African languages* 6. 95–104.
- 2006. *A Mon-Khmer comparative dictionary*, Paul Sidwell, Doug Cooper & Christian Bauer (eds). Canberra: Australian National University. Pacific Linguistics 579.
- Southworth, Franklin C. (forthcoming) *Proto-Dravidian Agriculture*. Online at <http://ccat.sas.upenn.edu/~fsouth/Proto-DravidianAgriculture.pdf>.
- Stampe, David, & Patricia Donegan. 2004. Rhythm and the synthetic drift of Munda. In Rajendra Singh (ed.) *The Yearbook of South Asian Languages and Linguistics*, 3–36. Berlin and New York: De Gruyter.
- Starosta, Stanley. 2005. Proto-East Asian and the origin and dispersal of the languages of East and southeast Asia and the Pacific. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The peopling of East Asia: putting together archaeology, linguistics and genetics*, 182–197. Oxon: RoutledgeCurzon.
- Vaughn, Duncan A., Baorong Lu & Norihiko Tomookoa. 2008. Was Asian rice (*Oryza sativa*) domesticated more than once? *Rice* 1. 16–24.
- Wei, Xingtao, Kong Zhaochen & Liu Changjiang. 2000. Discovery and significance of Yangshao cultivated rice at the Sanmenxia Nanjiaokou site. *Nongye Kaogu* 3. 77–9.
- Wiebusch, Thekla. & Uri Tadmor. 2009. Loanwords in Mandarin Chinese. In Martin Haspelmath & Uri Tadmor (eds) *Loanwords in the World's languages*, 575–598. Berlin: De Gruyter Mouton.
- Witzel, Michael. 2000. *The languages of Harappa*. Online at <http://www.people.fas.harvard.edu/~Witzel/IndusLang.pdf>.
- Wolff, John U. 2007. Petrified Prefixes in the Austronesian Languages. *Studies in Philippine Languages and Cultures* 15. 25–53.
- Wolff, John U. (2010) Proto-Austronesian phonology with glossary Vol 2. In Southeast Asia Program Publications, Cornell University.



# 16 *Rice and the Austroasiatic and Hmong-Mien homelands*

---

GEORGE VAN DRIEM

## 1 Introduction

The locations of the Austroasiatic and Hmong-Mien homelands have been tied to the origins of Asian rice cultivation. The lexical evidence supports the view that ancient Austroasiatics and ancient Hmong-Mien were the first rice cultivators. Historical linguistic insights shed light on the possible geographical location of the two respective homelands. Controversy regarding the origins of rice cultivation that has been waged since 1883 has, by and large, been resolved, though the story is complex. The palaeobotanical evidence is evaluated in light of molecular genetic findings on Asian rice. The overall picture is enhanced by insights from the genetics of Asian human populations, including modern Austroasiatic and Hmong-Mien language communities.

## 2 Phylogeny and linguistic palaeontology

On the basis of linguistic and ethnographic arguments, the presence of Austroasiatic populations in South and Southeast Asia has long been held to antedate the advent of Indo-European and Kradai, alias Daic, and perhaps even Tibeto-Burman and Dravidian in this region. Most conceivable theories about the ancestral homeland of Austroasiatic have already been put forward. Scholars have sought to situate the Austroasiatic *Urheimat* as far west as the Indus valley and as far east as the Yangtze delta or insular Southeast Asia. Today the principal contenders for the Austroasiatic homeland are the Indian subcontinent, mainland Southeast Asia and the middle Yangtze.

From a purely linguistic point of view, the location of the Austroasiatic ancestral homeland can be argued principally on several grounds. These include the geographical centre of gravity of the family based on the distribution of modern Austroasiatic language communities and the deepest phylogenetic divisions in the family tree, as well as historical grammar and linguistic palaeontology. The phylogeny of Austroasiatic presented by Gérard Diffloth is shown in Figure 1. Paul Sidwell has recently presented a rival model (see Sidwell and Blench, this volume), but a comparative appraisal of the two competing phylogenetic models would fall well beyond the scope of this article.

In view of the deepest historical division in the family's linguistic phylogeny, between Munda in the west and Khasi-Aslian in the east, the geographical distribution of the modern language communities would put the geographical centre of the family on either



side of the Ganges and Brahmaputra delta. Even the deepest division within the eastern or Khasi-Aslian trunk, that is the split into Khasi-Pakanic and Mon-Khmer, would suggest a point of dispersal for Khasi-Aslian between South Asia proper and Southeast Asia proper, somewhere in the northern littoral of the Bay of Bengal.

Historical grammarians have long observed typological differences between the Munda and the Khasi-Aslian branches of Austroasiatic. These structural differences raise the historical linguistic question of which grammatical features represent the original state and which structural traits are innovations. Donegan & Stampe (1983, 2004) propose that Austroasiatic spread from the Indian subcontinent to Southeast Asia, but argue, paradoxically, that the synthetic head-final typology of Munda languages resulted from an innovative process of drift which unfolded within South Asia after the linguistic ancestors of modern Khasi-Aslian language communities migrated towards Southeast Asia. In their view, the typological changes in Munda were triggered by a prosodic shift to a falling rhythm, whereas the analytic head-initial typology observed in Khasi-Aslian languages reflects the more original Austroasiatic state of affairs. By contrast, Zide & Anderson (1999, 2003) have argued that Munda verbal morphology is a conservative retention, and that older Austroasiatic grammatical systems were secondarily lost in the Khasi-Aslian languages of Southeast Asia. It is interesting to note that these opposing views of Austroasiatic historical grammar both presume a similar view with regard to the whereabouts of an original Austroasiatic homeland.

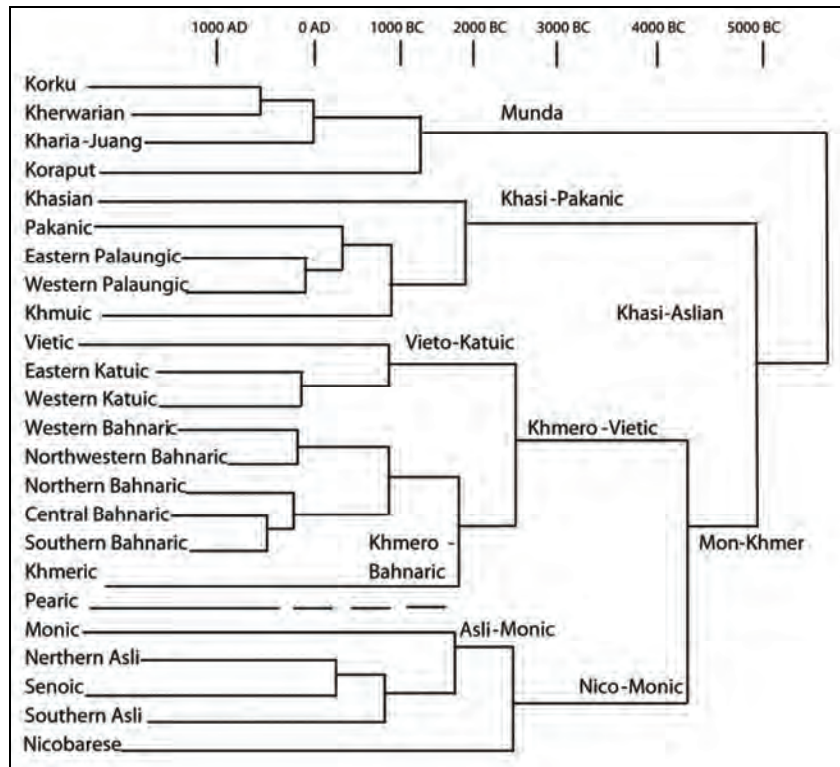
Alongside linguistic phylogeny and historical grammar, linguistic palaeontology offers a window onto the Austroasiatic past. Linguistic palaeontology, a term introduced by Adolphe Pictet in 1859, is an attempt to understand the ancient material culture of a language family on the basis of the lexical items which can be reliably reconstructed for the common ancestral language. The nature of the reconstructible Austroasiatic lexicon can be assessed in relation to the findings of palaeoecology, palaeobotany and archaeology.

Diffloth has shown that the reconstructible Austroasiatic lexicon paints the picture of a fauna, flora and ecology of a tropical humid homeland environment. The reconstructible lexicon possesses three salient isoglosses diagnostic for the faunal ecology of the Proto-Austroasiatic homeland reconstructible all the way to the Austroasiatic level and reflected in all branches of the family. The etyma *\*mra:k* ‘peacock *Pavo muticus*’, *\*tərkuət* ‘tree monitor lizard *Varanus nebulosus* or *bengalensis*’ and *\*tənyu:ʔ* ‘binturong’ or the ‘bear cat *Arctitis binturong*’, a black tropical mammal that is the largest of the civet cats (Diffloth 2005:78). All of these species are not native to areas that currently lie within China, and, to our present knowledge, these species were never native to the area that is today China. Such linguistic palaeontological evidence therefore appears to render the middle Yangtze homeland hypothesis less likely.

Diffloth also adduced additional Proto-Austroasiatic roots indicative of a tropical or subtropical climate, viz. *\*(bən)jo:l* ~ *\*j(ərm)o:l* ‘pangolin, *Manis javanica*’, *\*dəkan* ‘bamboo rat, *Rhizomys sumatrensis*’ (an Austroasiatic root which has found its way into Malay as a loan), *\*kaciaŋ* ‘the Asian elephant, *Elephas maximus*’, *\*kiaə* ‘mountain goat, *Capricornis sumatrensis*’, *\*rəma:s* ‘rhinoceros, *Dicerorhinus sumatrensis*’ and *\*tənriak* ‘buffalo, *Bubalus bubalus*’.

Osada (1995) and others have argued that the reconstructible lexicon strongly qualifies the ancient Austroasiatics as the most likely candidates for the first cultivators of rice. Diffloth (2005:78) adduces a rich repertoire of reconstructible roots representing rice, robustly reflected in all branches of Austroasiatic, viz. *\*(kə)ʔa:ʔ* ‘rice plant’, *\*rəŋko:ʔ* ‘rice grain’, *\*cəŋka:m* ‘rice outer husk’, *\*kəndək* ‘rice inner husk’, *\*phe:ʔ* ‘rice bran’, *\*təmpal* ‘mortar’, *\*jənreʔ* ‘pestle’, *\*jəmpiər* ‘winnowing tray’, *\*gu:m* ‘to winnow’, *\*jərmuəl*

‘dibbling stick’ and \*kəntu:ʔ ‘rice complement’, that is accompanying cooked food other than rice.



**Figure 1:** Gérard Diffloth’s (2009) Austroasiatic phylogeny with his tentative calibration of time depths (Diffloth 2001, 2005) for the various branches of the family. The two main trunks of the Austroasiatic tree are Munda in eastern, northeastern and central India and Khasi-Aslian, which stretches from the Meghālaya in the northeast of the subcontinent to the Nicobars, Malay peninsula and Mekong delta in Southeast Asia. The precise phylogenetic position of Pearic remains uncertain.

Nicole Revel (1988) contributed one of the most elaborate ethnobotanical studies on rice, rice cultivation practices and rice terminology in various Asian language communities. Her copious data already suggested the ancestral Hmong-Mien as another possible candidate for the early cultivators of rice. Progress has been made in our understanding of Hmong-Mien historical phonology, viz. Haudricourt (1954), Purnell (1970), Wáng & Máo (1995), Niederer (1998) and Ratliff (2010). The reconstructible lexicon specific to rice cultivation in Hmong-Mien is just as impressive as the Austroasiatic repertoire.

Martha Ratliff points out that, in Hmong-Mien, ‘[u]nlike the terms for the major non-rice crops..., terms for rice are less likely to show similarities with forms with similar meaning in neighboring families. Quite remarkable is the poor correlation with terms from Kra-dai and Austronesian, the languages of major rice growing peoples’ (2004:158–159). She identifies three rice cultivation etyma as being unambiguously original to the Hmong-Mien phylum: Proto-Hmongic \*ntsuw<sup>C</sup> ‘husked rice’, Proto-Hmong-Mien \*hnrəŋH ‘cooked rice’ and \*hnɔn ‘rice head, head of grain’. For the latter two items, Ratliff (2010) provides Baxter and Sagart’s Old Chinese reconstructions 饌 \*ŋaŋ(?)<sup>s</sup> and 囊 \*nʃaŋ.<sup>1</sup>

1 Hmong-Mien forms are given here in Ratliff’s new phonological reconstruction.

Ratliff (2004) suggests that the reconstructible Proto-Hmong-Mien rice cultivation etyma \*mblut ‘glutinous’, \*ljiŋ ‘paddy field’, \*ljim ‘sickle’, \*ŋkjuəX ‘rice cake’, \*tuX ‘husk/pound rice’ and \*tsjeŋH ‘rice steamer’ might have a Sinitic origin. In her 2010 handbook, she relates these roots to the Old Chinese forms 秣 \*m.lut, 田 \*lʃiŋ, 鎌 \*[r]em and 粃 \*[g](r)aʔ, 擣 \*tʃuʔ and 甑 \*s-təŋ-s respectively. Yet none of these Sinitic terms are well reflected across Tibeto-Burman, and none are currently reconstructible to the ancestral Tibeto-Burman proto-language. Rather, these etyma may all very well represent loans from Hmong-Mien into Sinitic, the one branch of Tibeto-Burman with the longest history of contact with Hmong-Mien.

Additional rice cultivation etyma reconstructible to Proto-Hmong-Mien include \*mbləu ‘rice plant, paddy’ and \*mphiek ‘chaff’. Reconstructible to the Proto-Hmongic level are the etyma \*S-phjæ<sup>C</sup> ‘chaff’, \*mbljæ<sup>C</sup> ‘have food with rice’, \*ʔrin<sup>A</sup> ‘dry (rice) in sun’ and \*tshɛŋ<sup>B</sup> ‘husked rice or millet’, and the rice measure etymon \*hrau<sup>A</sup> is reconstructible to the Proto-Mienic level.

The direction of some borrowings may have been from Sinitic into Hmong-Mien. Ratliff relates Proto-Hmongic \*ʔjeŋ<sup>A</sup> and Proto-Mienic \*ʔjaŋ<sup>A</sup> ‘seedling’ to Middle Chinese 秧 ʔjang and Old Chinese \*ʔaŋ, Proto-Hmongic \*ljeŋ<sup>A</sup> ‘rice measure’ to Old Chinese 量 \*[r]aŋ, and Proto-Mienic \*hmei<sup>B</sup> ‘husked rice’ to Old Chinese 米 \*[m]ʃijʔ. Yet of these, only the latter etymon appears to be reflected in Tibeto-Burman outside of Sinitic, viz. in Bodo-Koch languages.

Rice cultivation terminology could have been borrowed into Sinitic from ancient Hmong-Mien rice cultivators at a time when Proto-Sinitic millet growers came into increasing contact with their southern neighbours. The main split in the Hmong-Mien family is between Hmong and Mien, and the scattered distributions of modern communities of either branch have roughly the same geographical range, which is roughly bisected by the Pearl River. On the basis of the historical sources, it has long been mooted that the geographical centre of gravity of the family would originally have lain further north along the middle Yangtze (Cushman 1970). There is currently no palaeobotanical evidence for the co-cultivation of rice and foxtail millet along the middle Yangtze until around 3800 BC (Nasu et al. 2006).

### 3 Rice domestications and the archaeological record

The rice story is complex, and the plot of the story has changed more than once in the course of time. Alphonse-Louis-Pierre Pyrame de Candolle, director of the botanical garden in Geneva, argued that the origin of cultivated rice lay in China and that rice was introduced to India from China (1883:285, 309–311). Nikolai Ivanovič Vavilov (1926) paid tribute to Pyrame de Candolle, but argued against his claim of a Chinese origin for rice. In his elaborate model of the centres of plant domestication,<sup>2</sup> Vavilov contended instead that the origin of Asian rice lay in India, whence the crop had spread to China and Japan. The controversy has continued until recently, when molecular genetics was applied to the study of rice.

At times, South Asia would have the upper hand, as when Haudricourt & Hédin (1987:159–161, 176) proclaimed that the origin of rice cultivation lay ‘incontestablement’ in the Indian subcontinent. Afterwards, the consensus of scholarship moved the homeland

2 Themes repeat themselves in history, and Purugganan and Fuller’s map ‘Centres of plant domestication’ (2009:844) harks back to Vavilov’s 1926 map.

of rice agriculture from the Ganges to the Yangtze, and for a number of years conventional wisdom dictated that rice was domesticated in the Middle Yangtze, perhaps as early as the sixth millennium BC. Yet this idea was challenged by champions of an Indian homeland for rice. Palaeobotanical remains recovered at archaeological sites were adduced as evidence. I have discussed these shifting stances in the older rice literature elsewhere (van Driem 2001:324–327).

Archaeologists favouring a homeland for rice cultivation either in India or in China have been prone to exaggerate the antiquity and nature of the evidence. Morphometric criteria for distinguishing between actual rice domestication as opposed to mere rice cultivation were also not strictly observed. Both tendencies prompted Fuller, Harvey and Qin to exhort archaeologists ‘to cease and desist in presuming that all finds are domesticated and equate to agriculture’ (2007:328). Actual rice domestication is evinced by phenotypically observable genetic modification. Features such as long vs. short awns, shattering vs. non-shattering panicles, open vs. densely packed panicles and high dormancy vs. uniform germination are not always easy to establish by studying the microanatomy of archaeologically attested rice remains.

The chilly Younger Dryas, which lasted from ca. 10,700 to 9500 BC, is sometimes thought to have compelled the people behind the early Natufian culture in the Levant to turn to agriculture and cereal cultivation. Further east at this time, on the semi-arid Gangetic plain at the end of the mid-Holocene wet period, habitats for wild rices increasingly shifted to oxbow ponds as palaeochannels dried up and turned into oxbows. This shift favoured monsoonal rather than marshland rice species, including *Oryza nivara*, the wild progenitor of *Oryza indica* (Fuller 2006a).

The proponents of an original homeland for rice cultivation on the Indian subcontinent pointed at the presence of domesticated rice and ceramic culture from the Gangetic basin and at Doab sites such as Koldihawa and Mahagarha, dating from the seventh millennium BC (Sharma et al. 1980; Pal 1990; Agrawal 2002). Later, newer sites with rice remains and more reliable dates were reported at Lahuradewa (Lahurādevā), Tokwa (Ṭokuvā) and Sarāī Nahar Rāī.

At Lahuradewa, the early farming phase of period 1A in the site’s stratigraphy has radiocarbon dates ranging from ca. 5300 to 4300 BC. Carbonised material from this period was collected by the flotation method, yielding *Setaria pumila* and *Oryza rufipogon* as well as a morphologically distinct, fully domesticated form of rice ‘comparable to cultivated *Oryza sativa*’ (Tewari et al. 2003). More recently, accelerator mass spectroscopy dates were obtained on the rice grains themselves, corroborating the antiquity of rice agriculture at the site (Tewari et al. 2006, 2009). However, the preserved awn bases and attached rachillae in the Lahuradewa material suggest wild rice, probably *Oryza nivara*. Moreover, husk patterns are not a good diagnostic for distinguishing between *rufipogon*, *nivara* and *sativa* rice (Fuller 2006a, 2008; Fuller & Qin 2009).

Rice was certainly cultivated in the Gangetic basin by 7000 BC, but the current evidence for the actual domestication of *nivara* rice in the middle Ganges dates from no earlier than the third millennium BC. Rice agriculture is reported by Pokharia (2008) from the third millennium BC at the Ganges site of Tokwa near Allahabad. Fuller stresses that meticulous study will have to verify whether Tokwa is of greater antiquity than the second millennium BC, and problems arise from the allegedly disturbed stratigraphy of a few middle Ganges sites. At the same time, a reassessment of radiocarbon datings has tended to move crop domestication in the Indian subcontinent and Southeast Asia from a hoary past into the fifth millennium BC and may have chronological implications for the dawning of

the Neolithic horizon in some areas, for example Higham (2002, 2004), Fuller, Boivin & Korisettar (2007), Fuller, Harvey & Qin (2007).

Moreover, a fine distinction is being made between the cultivation and the actual domestication of crops, based on precise morphometric studies of phenotypical evidence in the form of grain size and anatomical changes accompanying the reduction of seed shattering. Yet the transition manifests itself as a longer and more gradual process in the archaeological record for Asian rice than seen for comparable morphological changes in barley and einkorn wheat in the Levant (Purugganan & Fuller 2009). Indeed, the genetic pathways which led to reduced seed shattering in different domesticated grasses such as wheat, barley, rice, etc. are many and species-specific (Li & Gill 2006). On the Gangetic plain, therefore, groups were using rice by 7000 BC, but between this time and the appearance of better evidence for agriculture, dating from 2500 to 2000 BC, the current archaeological record in the Indian subcontinent offers little evidence from which to infer cultivation practices or the evolution of rice domestication traits.

Similarly in East Asia, claims have been published dating the earliest rice cultivation in the Yangtze basin to as long ago as 10,000 BC, cf. Chén (1999). Certainly, sedentary settlements of the 彭頭山 Péngtóushān culture flourish along the middle Yangtze between 8000 and 6000 BC. However, the currently available evidence indicates that immature morphologically wild rice was cultivated by foragers before the actual domestication of the crop. This state of affairs is in evidence along the middle Yangtze at the Péngtóushān culture site 八十擋 Bāshídàng (7000–6000 BC) as well as in the Hángzhōu Bay littoral around the Yangtze delta at sites such as 跨湖橋 Kuàhúqiáo, 馬家浜 Mǎjiābāng (5000–3000 BC) and 河姆渡 Hémǔdù (5000–4500 BC).

The appearance of a minority of domesticated-type spikelet bases alongside the palynological and micro-charcoal data from Kuàhúqiáo suggest clearance by burning and water management already between 6000 and 5700 BC (Fuller, Harvey & Qin 2007; Fuller & Qin 2009). Yet the actual domestication of *japonica* type rice through genetic modification by selective breeding was probably effectuated along the lower Yangtze in the fifth millennium BC by people, who previously relied far more heavily on the collecting of acorns, water chestnuts and foxnuts before becoming reliant on rice cultivation. Morphological studies of rachides (spikelet bases) found at 田螺山 Tiánluóshān, however, show that cultivated rice first underwent domestication in the Lower Yangtze between 4900 and 4600 BC. This transition is evident in the increased proportion of rice versus other plant remains and the increased proportion of rice with non-shattering panicles. Dates were obtained by direct accelerator mass spectrometry radiocarbon dating on nuts and rice grains (Fuller et al. 2009; Nakamura 2010; Zhao 2010).

The picture which emerges is that the beginnings of rice agriculture were gradual, and the domestication process was initially slow and finished significantly later than often assumed. Over the course of the Chinese Neolithic, nut-gatherers switched to being farmers. The wild progenitor *Oryza rufipogon* was not fully domesticated in the lower Yangtze to yield early *Oryza japonica* until ca. 4000 BC (Fuller 2005a, 2005b, 2005c, 2006a, 2006b, 2006c, 2007a, 2007b; Yasuda 2002; Zong et al. 2007). Unambiguous morphological evidence of domesticated rice appears in the middle Yangtze ca. 4500 BC and in the lower Yangtze ca. 4000 BC (Fuller, Harvey & Qin 2007).

The cultivation of *japonica* rice spread from the Yangtze basin core area only after 4600 BC (Fuller & Qin 2009; Ruddiman et al. 2008). Rice reached the Yellow River basin during the third millennium BC (Crawford & Shen 1998). If the conjectural dates from the Chulmun site are tentatively accepted, then rice cultivation could also have entered the

Korean peninsula in the course of the third millennium BC (Ahn 2010), whereas paddy agriculture appears in the Japanese archaeological record in the course of the first millennium BC (Tanaka et al. 2010). Rice cultivation reaches Formosa and Vietnam between 2500 and 2000 BC (Higham & Lu 1998), but only spreads throughout the Indochinese peninsula between 1500 and 500 BC (Weber et al. 2010). Yet, whereas palaeobotanical investigations yield highly valuable and insightful evidence, molecular genetic findings have provided even more crucial insights to the issue of rice domestication.

#### 4 From morphometrics to rice genetics

By the 1990s, it had been understood that there were two principal domesticated varieties of rice, *Oryza sativa*, var. *indica*, and *Oryza sativa*, var. *japonica*. Some held that the *javanica* cultivar represented yet a third type of cultivated Asian rice.<sup>3</sup> The two cultivars *indica* and *japonica* were held to be phylogenetically distinct and to have distinct wild progenitors, which were identified as being *Oryza nivara* and *rufipogon* respectively. *Oryza rufipogon* is a perennial swamp species ranging throughout South and Southeast Asia and southern China. *Oryza nivara* is an annual which grows in moist soil or shallow water and ranges throughout South Asia and mainland Southeast Asia. Both species grow sympatrically and naturally hybridise with each other as well as with cultivated rice. Early genetic studies corroborated this view (Vaughan 1994).

The old controversy between an Indian versus a Chinese homeland for the origin of cultivated rice was rendered obsolete, although the new view did not dampen the fervour of those who strove to turn up the earliest evidence for rice domestication in India and in China. In fact, this archaeological contest was given a new impetus by the realisation amongst rice specialists that *Oryza rufipogon* actually constitutes a single diverse species encompassing both the annual self-pollinating *Oryza nivara*, adapted to disturbed shallow-water environments, and the perennial *Oryza rufipogon* in the strict sense, which can out cross and is adapted to stable deep-water environments. Some researchers still prefer to treat the two as distinct species, whereas others consider them to be distinct ecotypes under a single *rufipogon* species complex.

Genetically, there is no significant barrier to reproduction, and the two varieties essentially share a common genome. The different ecological niches aid and abet isolation and diversification within the species, but hybridisation is common. On this issue, Chen et al. (2008) have identified the S5 gene, which encodes for an aspartic protease which conditions fertile embryo sacs, as a major locus enforcing hybrid sterility between *indica* and *japonica* rice, and thereby contributing to the observed high rate of post-zygotic reproductive isolation of the two cultivars. Yet gene flow is facilitated by a third neutral allele which, when it occurs, yields fertile hybrids with either species.

In the 1970s, rice dwarf virus devastated many high-yield paddy fields in India and Indonesia. When resistance genes were found in *Oryza nivara*, the wild *nivara* was crossed with the most widely used of the afflicted *sativa* strains, yielding a hybrid which now

---

3 In contrast to Asian rice, African cultivated rice (*Oryza glaberrima*) is genetically related most intimately to the wild species *Oryza barthii* and, to a lesser extent, to *Oryza longistaminata*. African rice represents an independent domestication of a cultivated rice species geographically and genetically distinct from Asian rice, as shown by analyses of mitochondrial, chloroplast and nuclear DNA markers (Duan et al. 2007). Linguistic and genetic evidence both indicate that African rice was domesticated independently of Asian rice, probably in the upper Niger river delta (Blench 2006; Sweeney & McCouch 2007).

covers over 100,000 km<sup>2</sup> of Asian paddy fields. By the same token, so-called ‘weedy rice’ arises from the natural hybridisation of cultivated rice with wild *Oryza rufipogon* or *nivara*. Such de-domesticated rice was first detected in Sri Lanka in 1992, and today much yield loss in Ampara and Puttalam districts is attributed by farmers to the prevalence of weedy rice. Thus, it is defensible to view *Oryza rufipogon*, including *nivara*, as a single species.

Notwithstanding arguments for the single species status of *rufipogon* rice *sensu lato*, studies of the rice genome indicate independent domestications for the *indica* and *japonica* varieties of rice. This conclusion is based both on the time depth of the calculated divergence of their respective wild progenitors, viz. somewhere between 200,000 and 400,000 years ago (Ma & Bennetzen 2004; Vitte et al. 2004; Zhu & Ge 2005), and the independent adaptive selection acting on the *Hd6* photosensitivity gene, which enabled rice to spread beyond the confines of the tropics and subtropics to be cultivated globally from 53°N and 40°S (Yamane et al. 2009).

Londo et al. (2006) conducted a phylogeographic analysis of 203 cultivars of domesticated rice and over 129 populations of wild rice, that is *Oryza rufipogon* and *nivara*, spanning the entire natural geographical range of wild rice. The three distinct gene regions investigated (that is the maternally inherited chloroplast marker *atpB-rbcL*, the neutral nuclear pseudogene *p-VATPase*, presumed to be under little selective constraint, and the functional nuclear gene *S*-adenosyl methionine synthetase) yielded a consistent signal of domestication from distinct wild ancestral gene pools in two distinct geographical regions. Whilst the origin of the *Oryza sativa indica* lay in the region encircling the Bay of Bengal, which stretches from eastern and northeastern India to modern Burma and Thailand, the origin of the *japonica* cultivar lay in the Middle Yangtze basin in what today is southern China. Moreover, Londo et al. (2006) opine that *ghaiyā* or upland rice could represent a third distinct domestication event in or near the Indian subcontinent.

A major quantitative trait locus on the rice genome, *sh4*, is a single nucleotide substitution (SNP) encoding for a transcription factor. The consequent substitution of the amino acid lysine by asparagine in the Myb3 DNA-binding domain led to only partial development of the abscission zone where the mature grain detaches from the pedicle. Diminished gene function, expressed in the reduced brittleness of the rachides, is responsible for the reduction of grain shattering. The wild progenitors of domesticated rice exhibiting the latter trait were selected by early farmers to yield the non-shattering rice cultivars which could be efficiently harvested without the loss of grain. Yet it appears that the gene *sh4* had a single origin (Li et al. 2006a, 2006b; Lin et al. 2007; Onishi et al. 2007), whereas the phylogenetic evidence indicates separate domestications for *japonica* and *indica* rice. Therefore, the trait *sh4* is believed to have spread by introgression from its initial source, probably an early variety of domesticated rice, to other rice cultivars and also to the wild progenitors (Sweeney & McCouch 2007; Sang 2009).

Current genetic data therefore indicate the occurrence of two independent domestications, which were not entirely independent of each other. The available genetic data indicate overall gene flow from *japonica* to *indica*, but the set of domestication alleles studied to date is still limited, and this impression could change in the future (Kovach et al. 2007). Assuming that future research does not undermine this impression,<sup>4</sup> the most parsimonious model would assume that the *sh4* single nucleotide mutation first arose in

---

4 Since the synthesis of the genetics literature presented here was first circulated in 2009, portions of this reconstruction have been echoed in two subsequent discussions of the rice genetics literature (Purugganan 2010; Fuller et al. 2010).

*japonica* and then later spread back into the gene pool by introgression. It turns out that this rather complex genetic picture makes fairly good sense of the currently available palaeobotanical evidence.

The available morphometric evidence for Neolithic Gangetic rice, as collected and assessed by Harvey et al. (2006), indicates that grains are on the small side, congruent with morphologically wild rice, and perhaps containing immature grains. The cumulative evidence could be seen as indicating a long history of cultivation in the Ganges. This process initially involved the gathering of morphologically wild rice, with the evolution of non-shattering forms occurring quite late after the introduction of domesticated *japonica* rice carrying the key mutations, especially the recessive *sh4* allele, and perhaps the white pericarp *Rc* mutation (Kovach et al. 2007; Sweeney & McCouch 2007).

## 5 Rice domestications and geography

The reconstructible lexical evidence warrants us entertaining the hypothesis that the Proto-Hmong-Mien were the initial cultivators of *japonica* rice, and that ancient Austroasiatics were the initial cultivators of *indica* rice. The Austroasiatic and Hmong-Mien homelands need not have been coeval. A homeland is just a geographical region where a proto-language is thought to have been spoken at one point in time. The homelands of two language phyla may not necessarily have co-existed in time, as the homelands of different language families could be attributed to distinct time depths in the past. On a grander scale, no putative linguistic homeland was the original cradle of mankind.

However, for the first time since 1883, a coherent story about the domestication of Asian rice has emerged that could tell us where these homelands may have lain. Yet the plot of this tale may change again, and we can anticipate whence possible changes in the story line might come. Today both the palaeobotanical picture and the available genetic evidence remain incomplete.

Palaeobotanical evidence for an early domestication of Asian rice might as yet turn up in India. More crucially, despite the archaeological work conducted in the Ganges and Yangtze basins, much of the archaeology of ancient rice agriculture remains simply unknown because no substantive archaeological work has been done on the Neolithic in the most relevant areas, for example northeastern India, Bangladesh, the Indo-Chinese borderlands and Burma. The absence of evidence does not constitute evidence of absence, and the sheer dearth of archaeological research in these areas leaves entirely open the possibility that rice cultivation may have originated in this region.

Epistemologically, therefore, the situations in the Indochinese peninsula versus northeastern India and Burma are essentially contrasting. Whereas rice agriculture currently appears in the archaeological record of the Indochinese peninsula between 1500 and 500 BC (Weber et al. 2010), the absence of Neolithic sites is not the crucial empirical issue in Burma and northeastern India. Rather, the pivotal issue is the absence of archaeological research to the present day in this vast region. Whilst a Lower Mekong homeland hypothesis for Austroasiatic is, in view of the linguistic palaeontological evidence, severely challenged by the lack of early sites for rice agriculture in mainland Southeast Asia, the sheer archaeological neglect of the Salween, Irrawaddy and Lower Brahmaputran basins simply does not permit any inferences one way or the other.

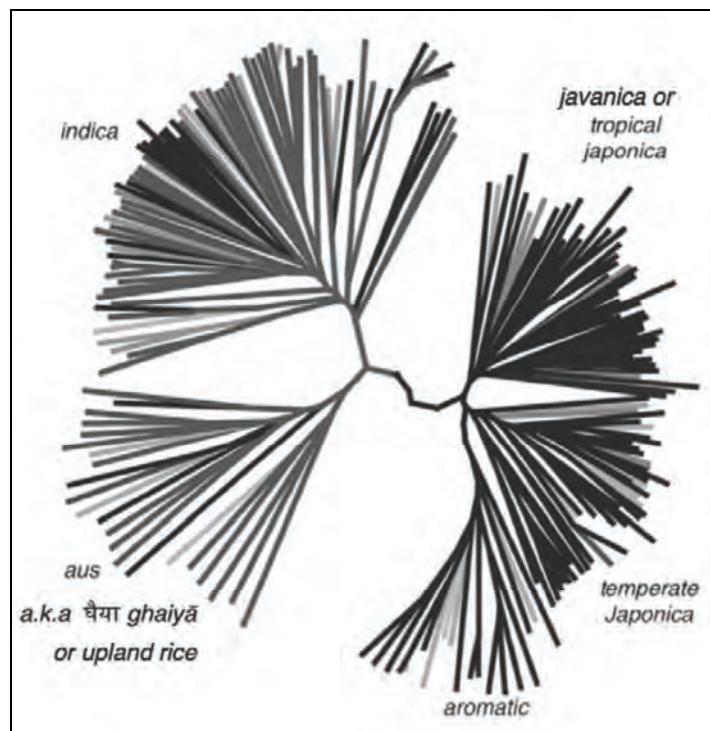
A related but distinct empirical issue is the archaeological recoverability of rice agriculture sites. The traces of ancient farming communities would tend to have been



better preserved in the hill tracts surrounding the Brahmaputran flood plains than on the fertile fields themselves.<sup>5</sup> Yet the earliest rice-based cultures may first have developed on those very flood plains. Perhaps the remains of the first rice cultivating cultural assemblages lie buried forever in the silty sediments of the sinuous lower Brahmaputran basin. Maybe the palaeobotanical evidence for the earliest domestication of rice was washed out by the Brahmaputra long ago and now lies submerged in the depths of the Bay of Bengal.

In future, genetic studies may likewise recast the rice story in a new light as more high-resolution genotyping and resequencing enables the identification of relevant markers of selection by domestication. Already the available genetic data on Asian rice yields information directly relevant to the hypothesis of ancient Austroasiatics and Hmong-Mien as the earliest rice cultivators of Asian rice. These considerations emerge from the diversity of the rice genome itself.

Genetic studies have provided a new view of Asian rice cultivars. Both *javanica*, now often called ‘tropical *japonica*’, as well as ‘temperate *japonica*’ issued from the same domestication process (Figure 2).



**Figure 2:** An unrooted phylogenetic tree of the diversity of *Oryza sativa* based on 169 nuclear simple sequence repeats (SSRs) and two chloroplast markers in 234 cultivars of *Oryza sativa*, adapted from Garris et al. (2005) and Kovach et al. (2007).

Despite the length of the grains and the Indian associations evoked by some aromatic cultivars in the minds of rice lovers today, the aromatics such as *bāsmati* rice likewise ultimately originate from the *japonica* domestication in the Yangtze basin (Parsons et al. 1999; Garris et al. 2005). By contrast, *ghaiyā* or upland rice is most closely affiliated with *indica* rice. However, some geneticists have mooted the possibility that upland rice or

5 Even in East Asia, many salvageable rice agriculture sites are in the foothills or at the base of the foothills (Nakamura 2010).

‘haplotype D’ rice represents a third distinct domestication event which unfolded in or near the Indian subcontinent (Londo et al. 2006).

In the rice genetics literature, upland or dry-cultivated rice sometimes currently goes by the Bengali name আউশ *āuš*, anglicised as ‘aus’, an orthography which belies the pronunciation.<sup>6</sup> The Dzongkha name for upland rice in Bhutan is simply *kambjā* ‘dry paddy’, whereas in Nepal upland rice goes by the proper name of घैया *ghaiyā* (Figure 3). Whether or not *ghaiyā* represents a third domestication event, upland rice is genetically more closely affiliated to *indica* rice. Upland rice has a short growing season, maturing early in the Nepali month of Āṣādh (Asār), that is June-July. Upland rice is tolerant to drought and requires little or no irrigation. These adaptive traits have rendered *ghaiyā* ideally suited to hillside niche environments in Nepal, Bhutan, the Himalayan foothills, Bangladesh and northeastern India, where it is still commonly grown.

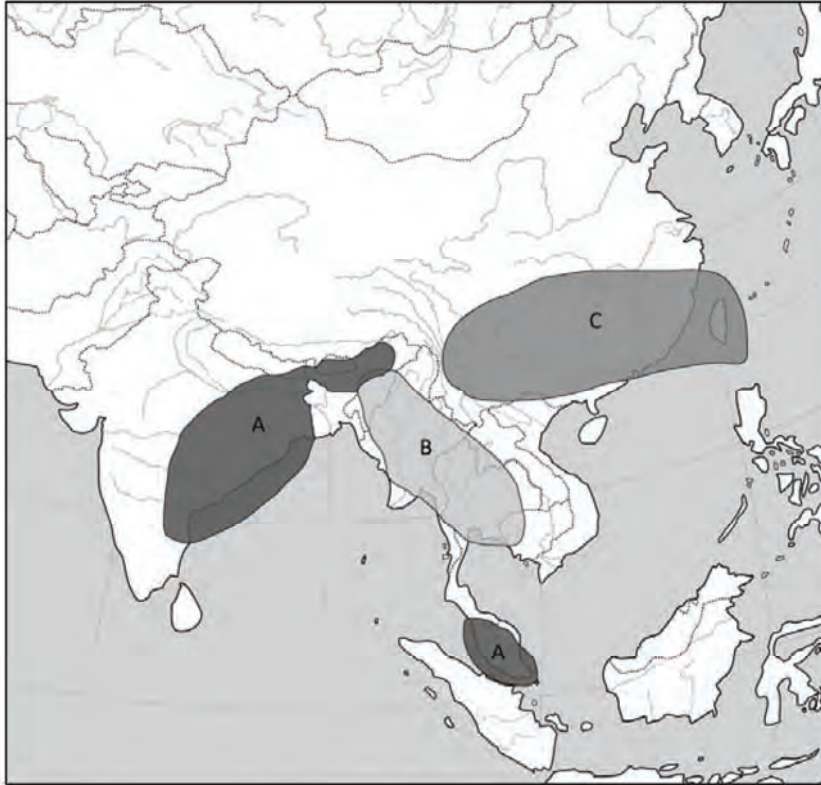
Based on the genetic markers in the wild precursor *Oryza rufipogon*, geneticists have identified the likely geographical ranges for the domestication of (A) *ghaiyā* or upland rice, (B) wet *indica* rice and (C) the *japonica* cultivar, shown in Figure 4. Whereas *japonica* rice is likely to have been domesticated somewhere in the Yangtze basin and its periphery, the most likely region for the domestication of wet *indica* rice stretches from the Brahmaputra to the Mekong. The likely geographical range for an independent domestication of upland rice, however, stretches from the Brahmaputra basin in the northeast to Pulicat Lake on the Coromandel Coast in the southwest and includes a discontinuous patch at the end of the Malay peninsula.



**Figure 3:** घैया *Ghaiyā* or upland rice (courtesy of Hem Bahādur Thāpā, alias Himāl bhānjā).

If Asian rice was indeed domesticated thrice instead of twice, then this finding would raise a whole set of new questions. However, if wet *indica* and upland rice diverged after a single initial domestication process, then the primordial *indica* cultivar may have been domesticated in a complex habitat which included both wet lowlands and dryer hills. Northeastern India and the Indo-Burmese borderlands straddle the area where the geographical range for the first domestication of wet *indica* rice and dry upland rice overlap, that is areas A and B in Figure 4.

6 An English speaker could come close to the Bengali pronunciation by pronouncing the name as ‘ouch’, but substituting an *sh* sound for the *ch* at the end of the word. Alternatively, the Nepali name *ghaiyā* might be a better alternative.



**Figure 4:** The geographical ranges for the domestication of (A) *ghaiyā* or upland rice, (B) wet *indica* rice and (C) the *japonica* cultivar, based on the genetic markers in the wild precursor *Oryza rufipogon* (adapted from Londo et al. 2006).

What else can be said about this area of overlap? First, northeastern India falls within the area whence linguistic phylogeny suggests that the ancient Austroasiatics could have originated. Second, this area forms part of a vast archaeologically largely unexplored region. Third, northeastern India and the Indo-Burmese borderlands have maintained highly diverse rice cultures to the present day.

At least five species of wild rice are native to northeastern India, viz. *Oryza nivara*, *Oryza officianalis* (*O. latifolia*), *Oryza perennis* (*O. longistaminata*), *Oryza meyeriana* (*O. granulata*) and *Oryza rufipogon*, and reportedly over a thousand varieties of domesticated rice are currently in use in the region (Hazarika 2005, 2006a). The different varieties of rice in northeastern India are cultivated in three periods by distinct cultivation processes. In the process of *āhu kheti*, the rice is sown in the months of *Phāgun* and *Sot*, that is mid February to early April. The seedlings are not transplanted but ripen in just four months in fields which must be constantly weeded. In *bāu kheti*, the rice seedlings are sown from mid March to mid April in ploughed wet fields and likewise do not need to be transplanted. In *śāli kheti*, the rice is sown from mid May to mid June, and the seedlings are transplanted. *Śāli kheti* rice varieties are suspected to derive from the wild *officianalis* rice still widely found in swampy village areas. The wild *rufipogon* rice cannot be used for human consumption because the plants shed their seeds before they ripen. Instead, wild *rufipogon* rice is used in Assam and other parts of northeastern India as cattle feed (Hazarika 2006b).

Beyond linguistic phylogeny, historical grammar, linguistic palaeontology, palaeobotanical findings and rice genetics, the homeland question can be argued on the basis of which genetic gradient or haplogroups in modern human populations might be

credibly related to either an ancient linguistic intrusion or to a demic spread of the language family in prehistory.

## **6 Father tongues and paternal homelands**

Often languages appear to show a correlation with the geographical distribution of genetic markers in the populations speaking these languages. Yet often enough there is a disconnect between the genetic affiliation of a population for a certain marker and the linguistic affinity of their language. The linguistic ancestors of a language community were not necessarily the same people as the biological ancestors of that community. Such discrepancies can often be as informative for our understanding of the sociolinguistic past as are the grand correlations. Moreover, the time depth accessible to geneticists stretches far beyond the ken of the linguistically reconstructible past based on systematic correspondences between related languages studied by means of the comparative method.

Human population genetics has made vast strides forward in recent decades. Classical population genetics developed mathematical models such as the Hardy-Weinberg equilibrium, quantified genetic diversity and distance, measured admixture and the effects of population size on selection and drift. In an attempt to reconstruct prehistoric migrations, Cavalli-Sforza, Menozzi & Piazza (1994) famously developed principal component analysis. However, since then, such summary statistical methods have been superseded by the more refined instrumentarium of modern human genetic phylogeography, which is not constrained to dealing with hypothetical prehistoric populations as discrete entities.

Human genetic phylogeography is largely based on the study of STRs and SNPs. Short tandem repeats (STRs) or ‘microsatellites’ are repeats of short segments of DNA generally less than five nucleotides in length. Microsatellites have a high rate of mutation and a large number of alleles. Single nucleotide polymorphisms (SNPs) are generally biallelic and tend to have a lower rate of mutation. Paternal ancestry of a population can be traced through polymorphisms on the paternally inherited Y chromosome, whereas the maternal ancestry of a population can be reconstructed on the polymorphisms in the maternally inherited mitochondrial DNA.<sup>7</sup> Ever more salient markers are being identified on the more numerous autosomes that will be as informative for reconstructing population prehistory as the polymorphisms on the non-recombinant portion of the genome. Human genetic phylogeography employs laborious formal methods, such as founder analysis, in order to date and measure prehistoric human migrations.

The linguistic affinity of a language often correlates better with the paternally inherited polymorphisms of the populations which speak that language than do other markers. I have called this tendency, recognised by Poloni et al. (1997, 2000), the father tongue hypothesis. Our genetic ancestry shows what I have described with an anatomical metaphor as sexual dimorphism, and so an uneven correlation with our linguistic prehistory. Some languages appear to be mother tongues, whereas others show up as father tongues. At many times and in many places in prehistory, the father tongue appears to have been the guiding mechanism in language shift (van Driem 2005, 2006, 2007). The dynamics of a process whereby mothers passed on the language of their spouses to their offspring has major

---

7 Only rarely does paternal mitochondrial DNA from the base of the sperm flagellum survive the process of selective destruction, dilution and inactivation during early embryogenesis by the vast surplus of oocyte mitochondria (Schwartz & Vissing 2002; Williams 2002). The reported case remains an anomaly in humans.

implications for our understanding of language change and so for historical linguistics and historical sociolinguistics in general.

Genetic studies tend often to corroborate models of ethnolinguistic prehistory previously mooted by linguists and ethnographers. For example, the spread of Indo-Aryan languages unambiguously attests to an ancient linguistic intrusion into the Subcontinent from the northwest. The Indo-Iranian intrusion onto the Iranian plateau and into the Subcontinent from the northwest appears also to have left a corresponding genetic legacy in the form of Y haplogroups M11-L, M17-R1a and M124-R2 spreading from the northwest across northern India and to Ceylon, whereas mitochondrial lineages prevalent in India are overwhelmingly indigenous to the Subcontinent (Kivisild et al. 1999a, 1999b; Wells et al. 2001; Cordaux et al. 2003; Kivisild et al. 2003; Baig et al. 2004; Cordaux et al. 2004; Metspalu et al. 2004; Quintana-Murci et al. 2004; Thangaraj et al. 2005; Sahoo et al. 2006).<sup>8</sup>

Similarly, a population genetic study of 23 Hân populations (Wen et al. 2004) has corroborated the picture which linguists and historians have long entertained, of a martial and therefore male-biased Hân expansion southward during the sinification of what today is southern China. The southern Hân paternal lineages, identified by Wen et al. (2004) with haplogroups M122-O3 and M134-O3e (O3a3c), show preponderant northern Hân penetration alongside a faint pre-Sinitic signature. Males from the north were the primary contributor to the paternal gene pool of southern Hân populations, whereas the mitochondrial DNA of southern Hân populations contains roughly equal contributions from pre-Sinitic and Hân maternal ancestors.

Yet certain languages are clearly mother tongues in that the mitochondrial markers of their speakers correlate with those of populations speaking related languages, whereas their Y chromosomal markers show no such correlation. The conservative Tibetan languages or mother tongues of Baltistan are a case in point (Poloni et al. 1997, 2000; Zerjal et al. 1997; Quintana-Murci et al. 2001; Qamar et al. 2002). Sometimes the genetic markers of a population studied to date show no clear correlation with the language spoken by a population. The evident lack amongst Hungarians of the TatC deletion defining the Y chromosomal haplogroup Tat-N3 (N1c),<sup>9</sup> despite its prevalence amongst other Uralic language communities, sheds some additional light on the complex sociolinguistic history of the migrations which have passed through Pannonia and the Magyar linguistic assimilation. Modern language communities have diverse ancestries, resulting from different initial waves of peopling as well as subsequent migrations (Li et al. 1999).

Any first successful human colonisation of a new land mass necessarily involves women. If an island were first settled exclusively by men, then these colonists would leave no progeny. Once women are in place, however, migrations made up mainly or even exclusively of men can follow and successfully introduce new genetic lineages into resident populations if there are any present. Linguistic intrusions need not involve women. The introduction of Indo-Iranian languages into the Indian subcontinent has already been discussed in this regard. In a similar vein, studies of mitochondrial DNA

---

8 Claims that Y chromosomal lineages, particularly R haplogroups, of Indo-Aryan high caste populations are instead autochthonous to India (e.g. Sengupta et al. 2006; Sharma et al. 2009; Reich et al. 2009; Chakravarti 2009) appear to be based on interpretations of modern frequency gradients which fail to take into account the genetic aftermath of known migrations which altered the genetic landscape of Central Asia subsequent to the Aryan ethnolinguistic intrusion. The topic is still under debate.

9 Alongside the familiar Y chromosome haplogroup labels, I provide between parentheses the newer haplogroup labels introduced in 2008 by the Y Chromosome Consortium (Karafet et al. 2008).

lineages show that any involvement by women in the linguistic and cultural colonisation of insular Southeast Asia by Austronesians from Formosa must likewise have been demographically minor (Ballinger et al. 1992; Macauley et al. 2005; Hill et al. 2007).

Fine-grain genetic studies of human populations often show that the maternal ancestry of groups is more complex. The maternal lineages of *orang asli* groups, based on studies of control-region and coding-region markers in the mitochondrial DNA, show that *orang asli* groups represent a microcosm of demographic processes that no doubt took place throughout Asia and must often have had a comparable degree of complexity. (cf. Oppenheimer, Fix, Bulbeck, this volume.) The maternal ancestry of the Semang dates back to the initial settlement of the Malay Peninsula from Africa 50,000 years ago. The Senoi on the other hand trace half of their maternal ancestry back to the same initial settlement, whereas the other half derives from women who came to the southern part of the peninsula as part of an early Austroasiatic speaking agriculturalist settlement 4000 years ago. The maternal ancestry of the Malays is more diverse and shows affinities with both insular and mainland Southeast Asia at time depths suggesting Late Glacial, early Holocene and Neolithic dispersals (Hill et al. 2006).

The mitochondrial ancestry of language communities in what today is China is likewise complex. However, on the whole the maternal ancestry appears older and more sedentary than the more dynamic episodes in the spread of Y chromosomal lineages against this mitochondrial background (for example Yao et al. 2002; Kong et al. 2003). The mitochondrial DNA pool in East Asia tends to be locally region-specific and largely derives from two super-haplogroups M and N. Mitochondrial lineages reflect early settlement patterns in East Asia and support an early phylogenetic partitioning between northern and southern populations (Kivisild et al. 2002). The majority of Hmong-Mien populations belong to haplogroups that tend to be prevalent in southern China, for example B, R9, N9a, and M7. Historically, the Hmong have had more contact with Chinese than have the Mien, and this contact has left genetic traces in the Hmong heritage (Wen et al. 2005). Mitochondrial haplogroup M7 characteristically has a southern distribution in East Asia, especially in the Yellow Sea littoral.<sup>10</sup>

Similarly, the maternal lineages in the Indian subcontinent generally show that there was a relatively rapid dispersal out of eastern Africa, at which time the mitochondrial heritage of the Subcontinent was first firmly established. The Subcontinent shows the effects of having been at a crossroads, but mitochondrial lineages often tend to be old and to antedate the last glacial maximum (Endicott et al. 2007). On the Indian subcontinent, a study of mitochondrial lineages indicates that Austroasiatic tribal groups show a far higher diversity of maternal lineages than Dravidian tribal populations, and that their maternal lineages more closely represent those of earlier settlement. The calculated expansion times

---

10 The M7 mitochondrial daughter groups M7a and M7b2, specific to Japanese and Korean populations, attest to a pre-Jōmon contribution to the modern Japanese mitochondrial DNA pool. The estimated coalescence times for the subclades M7a, M7b, and M7c range between 6000 and 18,000 years ago. This date suggests either that these star-like clades reflect a resettlement process around the Sea of Japan from the south after the Last Glacial Maximum, contemporary with the spread of Suyanggae microblades before the onset of the Jōmon culture, or that M7a and M7b entered Japan during initial settlement over 30,000 years ago and underwent a genetic bottleneck during the Last Glacial Maximum. By contrast, the mitochondrial haplogroups A5, B5, C, F1a, N9a, and Z, which are shared between Koreans and Japanese and are virtually absent in Ryūkyūans and in the Ainu, testify to later migrations through the Korean peninsula to Japan, probably during the Yayoi agricultural intrusion 2300 years ago. The presence of the mitochondrial lineage Y1 amongst the Ainu testifies to the migration of Siberian populations to the Japanese archipelago from the north (Kivisild et al. 2002; cf. Igawa et al. 2009).

for Austroasiatic mitochondrial lineages is 17,000 years older than those of Dravidian tribal groups in the Subcontinent and 22,000 older than Indian Tibeto-Burman maternal lineages (Roychoudhury et al. 2001). Whether and how long this picture will be upheld by emergent genetic data remains to be seen.

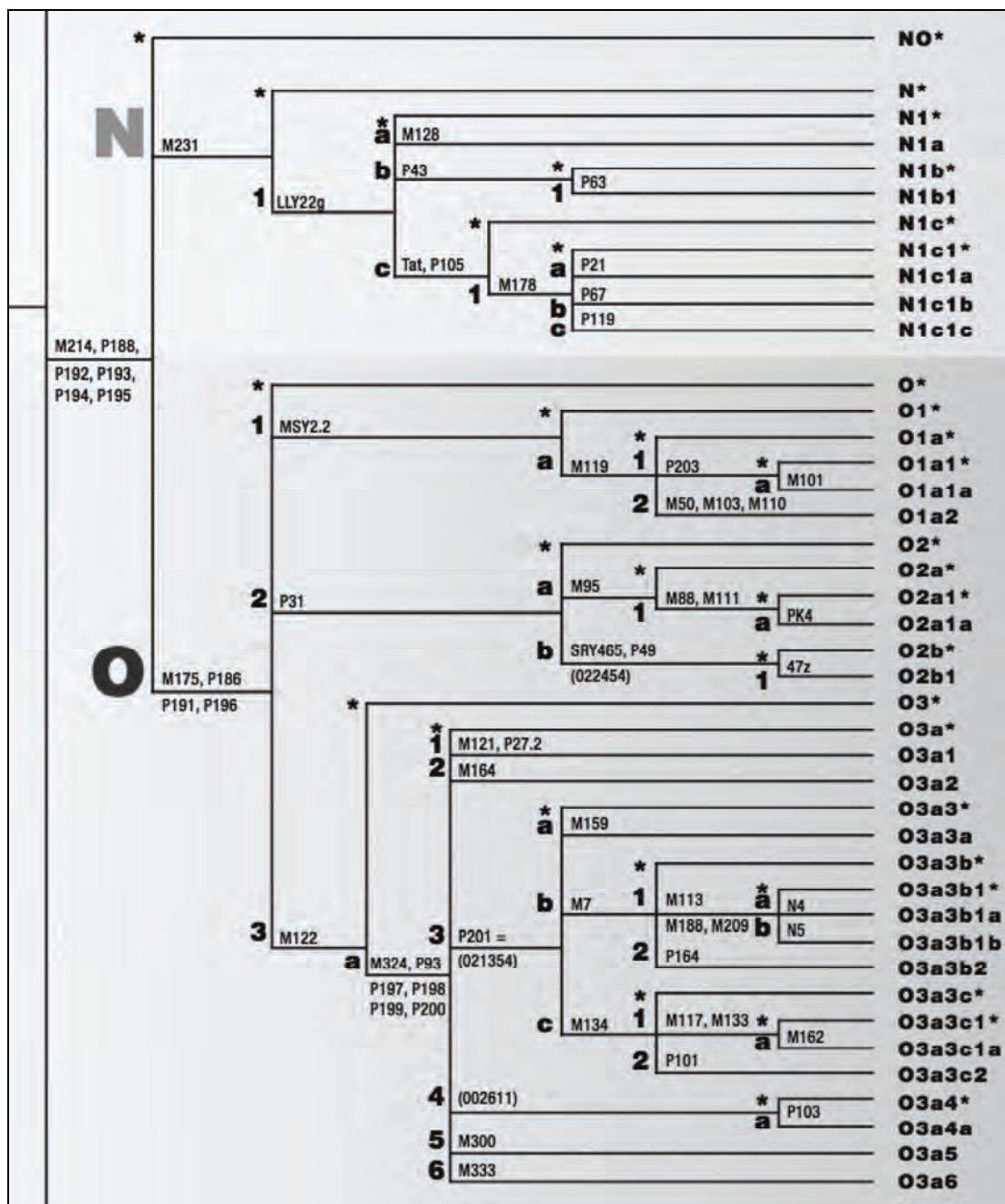
In contrast to this prehistory of maternal lineages, the paternal ancestry of modern language communities generally shows a far more dynamic history subsequent to the first settlement of Asia by anatomically modern humans out of Africa. Single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs) on the Y chromosome support two distinct northern and southern migration routes into East Asia. The northern populations expanded before the Last Glacial Maximum, presumably exploiting Siberian megafauna between 34,000 and 22,000 years ago, whereas southern populations expanded after the Last Glacial Maximum, between 18,000 and 12,000 years ago, but did so more vigorously than the northern populations. The southern expansion was more reliant on the exploitation of bountiful plant resources in the warmer more stable climates (for example Zerjal et al. 1997; Chũ et al. 1998; Sù et al. 1999, 2000; Ding et al. 2000; Shi et al. 2005; Xue et al. 2006; Karafet et al. 2008). In East Asia, the Y haplogroup M231-N has been interpreted by some geneticists as representing the legacy of ancient settlement by way of a boreal route, whereas the haplogroup M175-O and its many branches represent the more robust southern settlement via a littoral expansion. The arguments with respect to the spread of the Y haplogroup RPS4Y-C are undecided.

The correlation of the Y chromosomal haplogroup M134-O3e (O3a3c) with the spread of Tibeto-Burman language family has already been hypothesised (van Driem 2006). On the assumption of the veracity of the father tongue hypothesis for Austroasiatic, I identified the Y chromosomal haplogroup M95-O2a as the marker for the spread of Austroasiatic on the basis of the then available genetic data (van Driem 2007). This view has been corroborated by subsequent genetic studies, for example Kumar et al. (2007). The available genetic data also enable us now to propose a correlation of the Y chromosomal haplogroup M7-O3d (O3a3b) with the spread of Hmong-Mien. The relevant portion of the Y chromosomal tree is shown in Figure 5.

The undifferentiated Y chromosomal haplogroup M122-O3 is ancestral to both haplogroups M134-O3e (O3a3c) and M7-O3d (O3a3b), whereas the undifferentiated Y chromosomal haplogroup M175-O is ancestral to both haplogroups M122-O3 and M95-O2a. The undifferentiated haplogroups M175-O and M122-O3 currently appear at the highest frequency in Southeast Asian populations. At the same time, the high frequency of undifferentiated M122-O3 in assayed Southeast Asian Austroasiatic language communities appears to corroborate the view that the linguistic differentiation into the known linguistic phyla of Austroasiatic, Hmong-Mien and Tibeto-Burman was posterior in time to the differentiation of the haplogroups M95-O2a, M7-O3d (O3a3b) and M134-O3e (O3a3c) respectively. The precise geographical areas where these mutations took place cannot be pinpointed, but these areas no doubt lay between the Indus and the Yangtze delta, and more probably between the Brahmaputra and the Mekong delta, where the incidence of the undifferentiated haplogroups M175-O and M122-O3 is still relatively high.

Whereas the mitochondrial lineages in Austroasiatic groups are locally region-specific, and some of these lineages date back roughly to the advent of the first anatomically modern humans (Reddy et al. 2007), all Nicobarese assayed to date show the Y chromosomal haplogroup M95-O2a. Sahoo et al. (2006) find haplogroup M95-O2a to occur at a frequency of 77% in Austroasiatic groups in India, and Reddy et al. (2007) show a decreasing frequency gradient of this signature haplogroup from Mundari (53%) to Khasi (30%) to Southeast Asian Austroasiatics (ca. 23%). Assuming the veracity of the father

tongue hypothesis for the spread of Austroasiatic, the available data could be interpreted as pointing towards the Brahmaputra basin as the point of origin for this language family.



**Figure 5:** The portion of the Y chromosome phylogenetic tree relevant to the Father Tongue hypothesis with regard to Austroasiatic, Hmong-Mien, Tibeto-Burman and Austronesian and the theory regarding the peopling of eastern Asia via a boreal and a meridional route, reproduced from Karafet et al. (2008) with the kind permission of the Cold Spring Harbor Laboratory Press.

Reddy et al. (2007) measure an overall frequency of 7% for the Y haplogroup M95-O2a in Tibeto-Burman populations on the Indian subcontinent. The highest frequency of the M95-O2a haplogroup is found in tribal populations in Orissa, Chattisgarh and Jharkhand (Sengupta et al. 2006). However, the 47% frequency of the Y haplogroup M95-O2a found



by Sahoo et al. (2006) in Tibeto-Burman groups of northeastern India suggests that a subset of the paternal ancestors of particular Tibeto-Burman populations in northeastern India, for example certain Bodo-Koch communities, may originally have been Austroasiatic speakers who either married into Tibeto-Burman communities or who were linguistically assimilated by ancient Tibeto-Burmans. Indeed, median-joining network analysis of haplogroup M95-O2a microsatellite haplotypes suggested a division in the Indian subcontinent between Tibeto-Burmans vs. Austroasiatic and Dravidian language communities. The Austroasiatics and Dravidians also show greater Y-chromosomal microsatellite (that is short tandem repeat) diversification than is found in the Tibeto-Burman language communities.

Continuing to operate on the same set of hypotheses, the high incidence of the Y chromosomal haplogroup M7-O3d (O3a3b) in Austroasiatic language communities of Southeast Asia could indicate that some of the people who were assimilated by the spread of Austroasiatic into Southeast Asia may have been ancient Hmong-Mien or some related group long forgotten. The low incidence of haplogroup M95-O2a amongst Hmong-Mien speakers could indicate an Austroasiatic contribution to ancient Hmong-Mien populations, and would certainly have implied contact in the past. The incidence of the Y haplogroup M7-O3d (O3a3b) in Austroasiatic communities of the Indian subcontinent, by contrast, is undetectably low.

Perhaps the M175-O haplogroup first split up in northeastern India, with P31-O2 initially staying behind in the eastern part of the Indian subcontinent, and M122-O3 moving into the Indo-Burmese borderlands. Thence the derivative clade M134-O3e (O3a3c) headed north, and M7-O3d (O3a3b) headed east. Subsequently, Austroasiatic bearers of the M95-O2a haplogroup also continued to the east, where they came into contact with the Hmong-Mien whose ancestors had preceded them, and to the southeast, where their numbers expanded and they went on to colonise much of insular Southeast Asia. The significant incidence of the M95-O2a paternal haplogroup in Hmong-Mien populations and the M7-O3d (O3a3b) haplogroup in the paternal ancestry of eastern Austroasiatic populations of Southeast Asia suggests an intimate contact relationship between ancient Austroasiatics and ancient Hmong-Mien.

In his posthumously published theory of an East Asian superphylum comprising distant linguistic relationships, Starosta proposed a Proto-Yangtzean supergroup consisting of Austroasiatic and Hmong-Mien (van Driem 2008). Could Proto-Yangtzean correspond to a real linguistic taxon? If we assume the father tongue hypothesis for the spread of language, then the phylogeny of East Asian Y chromosomal haplogroups gives no reason to assume a closer relationship between these two language families. Instead a distant relationship between Tibeto-Burman and Hmong-Mien might be mooted to have existed at a hoary time depth beyond the linguistically reconstructible past.

Between the ancient Austroasiatics and the ancient Hmong-Mien, on the other hand, both human genes and knowledge about rice domestication were shared in the eastern half of a long and complex interaction sphere which extended from the Ganges in the west to the Yangtze and Mekong in the east. In fact, this more complex view is at least partially corroborated by our own findings in the realm of human genetic phylogeography, relating to how sex-specific admixture in Austroasiatic populations of the Indian subcontinent may have been correlated with a spread from east to west (Chaubey et al. 2010).

When we turn to the Kra-dai groups that now also inhabit this region, linguistically the Kra-dai languages of mainland Southeast Asia and southern China appear to be an old offshoot of Austronesian, dating from just after the Formosan exodus (Ostapirat 2005; Sagart 2004). In view of the continuing refinements to the 2002 and 2005 models of the Y

chromosomal haplogroup tree (Karafet et al. 2008), the currently emerging Y chromosomal picture based on single nucleotide polymorphisms suggests that Kradai peoples could descend mainly from ancient Hmong-Mien and Austroasiatic language communities which were linguistically assimilated by ancient Austronesian remigrants to the East Asian mainland. The Y chromosomal haplogroup M119-O1 (O1a) occurs at a high frequency amongst the Austronesian aboriginal peoples of Formosa and also, albeit in a much lower frequency, in the Philippines and southeastern China, especially in Kradai language communities (Abdulla et al. 2009). The paternal genetic imprint of the ancient Austronesians is but faint in comparison to their linguistic and cultural impact, which has extended from Formosa across half the planet.<sup>11</sup>

The slight disconnect between the immediate linguistic affinity of Kradai peoples and their predominant Y chromosomal signature indicates that the Kradai languages probably spread across populations which by and large retained their original genetic make-up. The ancient mtDNA recovered in northeastern Thailand from the Bronze Age site Noen U-loke, dating from 1500 BC to 500 AD, and from the Iron Age site Ban Lum-Khao, dating from 1200 to 400 BC, most closely resembles the mitochondrial lineages of a local Austroasiatic group, i.e. the Chaobon who speak the Monic language Nyahkur (Lertrit et al. 2008).<sup>12</sup> In a similar vein, the numerous insular Southeast Asian Austronesian language communities, where the M95-O2a haplogroup is also prevalent, may represent ancient Austroasiatics who were linguistically assimilated after the Formosan exodus.

Both linguists and ethnographers have for various reasons presumed that Austroasiatic is a particularly old language family. The prehistoric range of the language phylum could have been vast. If we assume the father tongue hypothesis and a correlation of the Y chromosomal haplogroup M95-O2a with this phylum, then the ancient range of Austroasiatic could have stretched from the Godāvārī to the Halmahera Sea. The theory of an Austroasiatic substrate in insular Southeast Asia dates back at least to Skeat and Blagden (1906). The view was explicitly or implicitly entertained in much of the anthropological literature of the period. However, actual linguistic evidence has seldom been adduced, with the notable exception of a study by Adelaar (1995). The culture of Enggano off the west coast of Sumatra has recently been studied in the light of an Austroasiatic substrate theory (Blench 2009). The strongest evidence for this old theory, however, appears to be prevalence of the Y chromosomal signature haplogroup that could have been associated with the ancient Austroasiatics.

Assuming the father tongue hypothesis, it appears that we can at least paint the linguistic map of much of eastern Asia with a broad brush, based on our understanding of the phylogeny of Y chromosomal haplogroups. Detailed, careful correlations of linguistic and population genetic findings based on a far more fine-meshed and a far more ethnolinguistically well-informed population genetic sampling throughout Asia, in light of the higher-resolution Y haplogroup phylogeny currently available, will hopefully enable us

---

11 A number of the roughly one hundred alleged correspondences adduced as evidence for a genetic relationship between Austronesian and the southernmost Andamanese languages Ōnge and Jarawa (Blevins 2007) are not compelling. Yet it cannot be excluded that some of the forms might reflect the residue of contact influence.

12 Apparently the Thai research group observed numerous caveats which have appeared in the literature regarding the processing and treatment of ancient mtDNA. For example, Yao et al. (2003) stress that unambiguous conclusions regarding mitochondrial prehistory must be based not just on hyper-variable segment 1 (HVS 1) sequences which are held to be indicative for geographical or ethnolinguistic affinity, but also on coding region polymorphisms diagnostic for mtDNA haplogroup membership.

to reconstruct early language contact situations and ancient cases of language shift and linguistic intrusions with far greater detail. In a careful interdisciplinary approach lies the promise of sociolinguistic reconstructions of episodes of our past which could account for the phenotypical differences readily observable and long noted by ethnographers between Munda speakers and Khasi-Aslian language communities or between Aslian negrito populations, Aslian non-negrito populations and the Nicobarese. An understanding of prehistory from the different vantage points of linguistics, archaeology and palaeobotany, and cultural and physical anthropology will provide a richer picture of our shared past.

## Acknowledgements

This chapter was first presented at the 4th International Conference on Austroasiatic Linguistics at Mahidol University on 29 October 2009. I should like to thank Martha Ratliff for sharing an electronic copy of the index of her then forthcoming handbook on Hmong-Mien historical phonology (2010). Figure 4 was prepared by Peter Johan Nijland.

## References

- Abdulla, Mahmood Ameen, Ikhlaq Ahmed, Anunchai Assawamakin, Jong Bhak, Samir K. Brahmachari, Gayvelline C. Calacal, Amit Chaurasia, Chien-Hsiun Chen, Jieming Chen, Yuan-Tsong Chen, Jiayou Chu, Eva Maria C. Cutiongco-de la Paz, Maria Corazon A. de Ungria, Frederick C. Delfin, Juli Edo, Suthat Fuchareon, Ho Ghang, Takashi Gojobori, Junsong Han, Sheng-Feng Ho, Boon Peng Hoh, Wei Huang, Hidetoshi Inoko, Pankaj Jha, Timothy A. Jinam, Li Jin, Jongsun Jung, Daoroon Kangwanpong, Jatupol Kampuansai, Giulia C. Kennedy, Preeti Khurana, Hyung-Lae Kim, Kwangjoong Kim, Sangsoo Kim, Woo-Yeon Kim, Kuchan Kimm, Ryosuke Kimura, Tomohiro Koike, Supasak Kulawonganuchai, Vikrant Kumar, Poh San Lai, Jong-Young Lee, Sunghoon Lee, Edison T. Liu, Partha P. Majumder, Kiran Kumar Mandapati, Sangkot Marzuki, Wayne Mitchell, Mitali Mukerji, Kenji Naritomi, Chumpol Ngamphiw, Norio Niikawa, Nao Nishida, Bermseok Oh, Sangho Oh, Jun Ohashi, Akira Oka, Rick Ong, Carmencita D. Padilla, Prasit Palittapongarnpim, Henry B. Perdigon, Maude Elvira Phipps, Eileen Png et al. (30 co-authors). 2009. Mapping human genetic diversity in Asia. *Science* 326. 1541–1545.
- Adelaar, Karl Alexander. 1995. Borneo as a crossroads for comparative Austronesian linguistics. In Peter Bellwood, James J. Fox & Darrell Tryon (eds) *The Austronesians: Historical and Comparative Perspectives*, 75–95. Canberra: Australian National University.
- Agrawal, D. P. 2002. The earliest pottery and agriculture in South Asia. In Yoshinori Yasuda (ed.) *The Origins of Pottery and Agriculture*, 81–88. New Delhi: Lusre Press and Roli Books, for the International Center for Japanese Studies.
- Ahn, Sung-Mo. 2010. The emergence of rice agriculture in Korea: archaeobotanical perspectives. *Archaeological and Anthropological Sciences* 2(2). 89–98.
- Baig, M. M., A. A. Khan & K. M. Kulkarni. 2004. Mitochondrial DNA diversity in tribal and caste groups of Maharashtra (India) and its implication on their genetic origins. *Annals of Human Genetics* 68. 453–460.
- Ballinger, S. W., T. G. Schurr, Antonio Torroni, Y. Y. Gan, J.A. Hodge, K. Hassan, K. H. Chen & D. G. Wallace. 1992. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. *Genetics* 130. 139–152.
- Blench, Roger. 2006. *Archaeology, Language and the African Past*. Lanham: Alta Mira Press.

- . 2009. *The Enggano: archaic foragers and their interactions with the Austronesian world*. Manuscript 9 September 2009.
- Blevins, Juliette. 2007. A long lost sister of Proto-Austronesian? Proto-Ongan, mother of Jarawa and Onge of the Andaman Islands. *Oceanic Linguistics* 46(1). 154–198.
- de Candolle, Alphonse-Louis-Pierre Pyrame. 1883. *Origine des plantes cultivées*. Paris: Librairie Germer Ballière et Compagnie.
- Cavalli-Sforza, Luigi Luca, Paolo Menozzi & Alberto Piazza. 1994. *The History and Geography of Human Genes*. Princeton, New Jersey: Princeton University Press.
- Chakravarti, Aravinda. 2009. Tracing India's invisible threads. *Nature* 461. 487–488.
- Chaubey, Gyaneshwer, Mait Metspalu, Ying Choi, Reedik Mägi, Irene Gallego Romero, Siiri Rootsi, Pedro Soares, Mannis van Oven, Doron M. Behar, Siiri Rootsi, Georgi Hudjashov, Chandana Basu Mallick, Monika Karmin, Mari Nelis, Jüri Parik, Alla Goverdhana Reddy, Ene Metspalu, George van Driem, Yali Xue, Chris Tyler-Smith, Kumarasamy Thangaraj, Lalji Singh, Mait Metspalu, Martin B. Richards, Marta Mirazon Lahr, Manfred Kayser, Richard Villems and Toomas Kivisild. 2010. Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution*, doi: 10.1093/molbev/msq28.
- Chen, Jiongjiong, Jihua Ding, Yidan Ouyang, Hongyi Du, Jiangyi Yang, Ke Cheng, Jie Zhao, Shuqing Qiu, Xuelian Zhang, Jialing Yao, Kede Liu, Lei Wang, Caiguo Xu, Xianghua Li, Yongbiao Xue, Mian Xia, Qing Ji, Jufei Lu, Mingliang Xu & Qifa Zhang. 2008. A triallelic system of *S5* is a major regulator of the reproductive barrier and compatibility of *indica-japonica* hybrids in rice. *Proceedings of the National Academy of Sciences of the United States of America* 105(32). 11436–11441.
- Chén Xīngcàn. 1999. On the earliest evidence of rice cultivation in China. In Peter Bellwood & Ian Lilley (eds) *Bulletin of the Indo-Pacific Prehistory Association Bulletin, 18 (Indo-Pacific Prehistory: The Melaka Papers, Vol. 2)*, 81–94. Canberra: Australian National University.
- Chǔ Jiāyòu, W. Huang, S. Q. Kuang, J. M. Wang, J. J. Xu, Z. T. Chu, Z. Q. Yang, K. Q. Lin, P. Li, M. Wu, Z. C. Geng, C. C. Tan, R. F. Du & Jīn Lì. 1998. Genetic relationship of populations in China. *Proceedings of the National Academy of Sciences of the United States of America* 95. 11763–11768.
- Cordaux, Richard, Nilmani Saha, Gillian R. Bentley, Robert Aunger, S.M. Sirajuddin & Mark Stoneking. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. *European Journal of Human Genetics* 11. 253–264.
- Cordaux, Richard, Gunter Weiss, Nilmani Saha & Mark Stoneking. 2004. The northeast Indian passageway: A barrier or corridor for human migrations? *Molecular Biology and Evolution* 21. 1525–1533.
- Crawford, Gary W. & Chen Shen 1998. The origins of rice agriculture: Recent progress in East Asia. *Antiquity* 72. 858–866.
- Cushman, Richard D. 1970. *Rebel Haunts and Lotus Huts: Problems in the Ethnohistory of the Yao*. Ithaca, New York: Cornell University doctoral dissertation.
- Diffloth, Gérard. 2001. *Tentative calibration of time depths in Austroasiatic branches*. Paper presented at the Colloque 'Perspectives sur la Phylogénie des Langues d'Asie Orientales' at Périgueux, 30 August 2001.
- . 2005. The contribution of linguistic palaeontology to the homeland of Austroasiatic. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia:*

- Putting Together the Archaeology, Linguistics and Genetics*, 77–80. London: Routledge Curzon.
- 2009. *More on Dvaravati Old Mon*. Paper presented at the Fourth International Conference on Austroasiatic Linguistics, Mahidol University at Salaya, 29 October 2009.
- Ding Yuan-Chun, Stephen Wooding, Henry C. Harpending, Chi Han-Chang, Li Hai-Peng, Fu Yun-Xin, Pang Jun-Feng, Yao Yong-Gang, Xiang Yu Jing-Gong, Robert Moyzis & Zhang Ya-ping. 2000. Population structure and history in East Asia. *Proceedings of the National Academy of Sciences of the United States of America* 97. 14003–14006.
- Donegan, Patricia J. & David Stampe. 1983. Rhythm and holistic organization of language structure. In John F. Richardson, Mitchel Marks & Amy Chukerman (eds) *Papers from the Parasession of the Interplay of Phonology, Morphology and Syntax*, 337–363/ Chicago: Chicago Linguistics Society.
- 2004. Rhythm and the synthetic drift of Munda. In Rajendra Singh (ed.) *The Yearbook of South Asian Languages and Linguistics 2004*, 3–36. Berlin: Mouton de Gruyter.
- Driem, George van. 2001. *Languages of The Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region containing an Introduction to the Symbiotic Theory of Language*, vol. 1, Brill, Leiden.
- 2005. Tibeto-Burman vs. Indo-Chinese: Implications for population geneticists, archaeologists and prehistorians. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*, 81–106. London: Routledge Curzon.
- 2006. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1(2). 211–270.
- 2007. Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies. *Mon-Khmer Studies* 37. 1–14.
- 2008. To which language family does Chinese belong, or what's in a name. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros & Marie Lin (eds) *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*, 219–253. London and New York: Routledge.
- Duan Shihua, Baorong Lu, Zhong Li, Jiping Tong, Jin Kong, Wen Yao, Shaoqing Li & Yingguo Zhu. 2007. Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial and nuclear DNA sequences. *Biochemical Genetics* 45(1/2). 113–129.
- Endicott, Phillip, Mait Metspalu & Toomas Kivisild. 2007. Genetic evidence on modern human dispersals in South Asia: Y chromosome and mitochondrial DNA perspectives: The world through the eyes of two haploid genomes. In Michael D. Petraglia & Bridget Allchin (eds) *The Evolution and History of Human Populations in South Asia: Interdisciplinary Studies in Archaeology, Biological Anthropology, Linguistics and Genetics*, 341–244. Dordrecht: Springer.
- Fuller, Dorian Q. 2005a. Ceramics, seeds and culinary change in prehistoric India. *Antiquity* 79. 761–777.
- 2005b. Formation processes and palaeoethnobotanical interpretation in South Asia. *Journal of Interdisciplinary Studies* 2(1). 93–115.
- 2005c. The Ganges on the world Neolithic map: The significance of recent research on agricultural origins in northern India. *Prāgdhārā* 16. 187–206.

- 2006a. Agricultural origins and frontiers in South Asia: A working synthesis. *Journal of World Prehistory* 20. 1–86.
- 2006b. Dung mounds and domesticators: Early cultivation and pastoralism in Karnataka. In Catherine Jarrige & Vincent Lefèvre (eds) *South Asian Archaeology 2001: Proceedings of the Sixteenth International Conference on South Asian Archaeology, European Association of South Asian Archaeologists, Paris, 2-6 July 2001*, 117–127. Paris: Éditions Recherche sur les Civilisations.
- 2006c. Silence before sedentism and the advent of cash crops: A status report on early agriculture in South Asia from plant domestication to the development of political economies (with an excursus on the problem of semantic shift among millets and rice). In Osada Toshiki (ed.) *Proceedings of the pre-Symposium of the Research Institute for Humanity and Nature and 7th Ethnogenesis of South and Central Asia Round Table Harvard-Kyoto Roundtable*, 175–213. Kyoto: Research Institute for Humanity and Nature.
- 2007a. Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the Old World. *Annals of Botany*, 1–22.
- 2007b. Non-human genetics, agricultural origins and historical linguistics in South Asia. In Michael D. Petraglia & Bridget Allchin (eds) *The Evolution and History of Human Populations in South Asia: Interdisciplinary Studies in Archaeology, Biological Anthropology, Linguistics and Genetics*, 393–443. Dordrecht: Springer.
- Fuller, Dorian Q, Nicole Boivin & Ravi Korisettar. 2007. Dating the Neolithic of South India: New radiometric evidence for key economic, social and ritual transformations. *Antiquity* 81. 755–778.
- Fuller, Dorian Q., Emma Harvey & Ling Qin. 2007. Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81. 316–331.
- Fuller, Dorian Q. 2008. Neolithic Cultures. In Deborah M. Pearsall (ed.) *Encyclopaedia of Archaeology, Vol. 1*, 756–768. New York: Academic Press.
- Fuller, Dorian Q. & Ling Qin. 2009. Water management and labour in the origins and dispersal of Asian rice. *World Archaeology* 41(1). 88–111.
- Fuller, Dorian Q, Ling Qin, Yunfei Zheng, Zhijun Zhao, Xugao Chen, Leo Aoi Hosoya & Guo-Ping Sun. 2009. The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323. 1607–1610.
- Fuller, Dorian Q, Yo-ichiro Sato, Cristina Castillo, Ling Qin, Alison R. Weisskopf, Eleanor J. Kingwell-Banham, Jixiang Song, Sung-Mo Ahn & Jacob van Etten. 2010. Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeological and Anthropological Sciences* 2(2). 115–131.
- Garris, Amanda J., Thomas H. Tai, Jason Coburn, Steve Kresovich & Susan McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169. 1631–1638.
- Harvey, Emma L., Dorian Q. Fuller, R.K. Mohanty & Basanta Mohanta. 2006. Early agriculture in Orissa: Some archaeobotanical results and field observations on the Neolithic. *Man and Environment*, XXXI(2). 21–32.
- Haudricourt, Andre-Georges. 1954. Introduction à la phonologie historique des langues miao-yao. *Bulletin de l'École Française d'Extrême-Orient* 44. 555–574.
- Haudricourt, André-Georges & Louis Hédin. 1987. *L'homme et les plantes cultivées*. Paris: Éditions A.-M. Métailié.

- Hazarika, Manjil. 2005. *Neolithic culture of Northeast India with Special Reference to the Origins of Agriculture and Pottery*. Pune (Poona): Unpublished Master's thesis, Deccan College.
- . 2006a. Neolithic culture of northeast India: A recent perspective on the origins of pottery and agriculture. *Ancient Asia* 1. 25–43.
- . 2006b. Understanding the process of plant and animal domestication in northeast India: A hypothetical approach. *Asian Agri-History* 10(3). 203–212.
- Higham, Charles 2002. *Early Cultures of mainland Southeast Asia*. Bangkok: River Books.
- . 2004. Mainland Southeast Asia from the Neolithic to the Iron age. In Ian Glover & Peter Bellwood (eds) *Southeast Asia: From Prehistory to History*, 41–67. Abingdon: Routledge Curzon.
- Higham, Charles & Tracey Lu. 1998. The origins and dispersal of rice cultivation. *Antiquity* 72. 867–877.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macaulay, William Meehan, James Blackburn, Douglas Clarke, Joseph Maripa Raja, Patimah Ismail, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2006. Phylogeography and Ethnogenesis of Aboriginal Southeast Asians. *Molecular Biology and Evolution*, 23(12). 2480–2491.
- Hill, Catherine, Pedro Soares, Maru Mormina, Vincent Macauley, Dougie Clarke, Petya B. Blumbach, Matthieu Vizuete-Forster, Peter Forster, David Bulbeck, Stephen Oppenheimer & Martin Richards. 2007. A mitochondrial stratigraphy for Island Southeast Asia. *American Journal of Human Genetics* 80. 29–43.
- Igawa, Kazunari, Yoshitaka Manabe, Joichi Oyamada, Yoshikazu Kitagawa, Katsutomo Kato, Kazuya Ikematsu, Ichiro Nakasono, Takayuki Matsushita & Atsushi Rokutanda. 2009. Mitochondrial DNA analysis of Yayoi period human skeletal remains from the Doigahama site. *Journal of Human Genetics* 54. 581–588.
- Karafet, Tatiana M., Fernando L. Mendez, Monica B. Meilerman, Peter A. Underhill, Stephen L. Zegura & Michael F. Hammer. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research* 18. 830–838.
- Kivisild, Toomas, Katrin Kaldma, Mait Metspalu, Jüri Parik, Surinder S. Papiha & Richard Villems. 1999a. The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World. In Surinder S. Papiha, R. Deka & R. Chakraborty (eds) *Genomic Diversity: Applications in Human Population Genetics*, 135–152. New York: Kluwer Academic, Plenum Publishers.
- Kivisild, Toomas, Michael J. Bamshad, Katrin Kaldma, Mait Metspalu, E. Metspalu, M. Reidla, S. Laos, Jüri Parik, W. Scott Watkins, Mary E. Dixon, Surinder S. Papiha, S. S. Mastana, M.R. Mir, V. Ferak & Richard Villems. 1999b. Deep common ancestry of Indian and Western-Eurasian mitochondrial DNA lineages. *Current Biology* 9(22). 1331–1334.
- Kivisild, Toomas, Helle-Viivi Tolk, Jüri Parik, Yiming Wang, Surinder S. Papiha, Hans-Jürgen Bandelt & Richard Villems. 2002. The emerging limbs and twigs of the East Asian mtDNA tree. *Molecular Biology and Evolution* 19(10). 1737–1751.
- Kivisild, Toomas, S. Rootsi, Mait Metspalu, S. Mastana, Katrin Kaldma, Jüri Parik, E. Metspalu, M. Adojaan, H.-V. Tolk, V. Stepanov, M. Gölge, E. Usanga, Surinder S. Papiha, C. Cinnioglu, R. King, Luca Luigi Cavalli-Sforza, Peter A. Underhill & Richard Villems. 2003. The genetic heritage of the earliest settlers persist both in Indian tribal and caste populations. *American Journal of Human Genetics* 72. 313–332.

- Kong Qing-Peng, Yong-Gang Yao, Chang Sun, Hans-Jürgen Bandelt, Chun-Ling Zhu & Ya-Ping Zhang. 2003. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *American Journal of Human Genetics* 73. 671–676.
- Kovach, Michael J., Megan T. Sweeney & Susan McCouch. 2007. New insights into the history of rice domestication. *Trends in Genetics* 23(11). 578–587.
- Kumar, Vikrant, Arimanda N. S. Reddy, Jagedeesh P. Babu, Tiprisetti N. Rao, Banrida T. Langstieh, Kumarasamy Thangaraj, Alla G. Reddy, Lalji Singh & Battini M. Reddy. 2007. Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BioMed Central Evolutionary Biology* 7: 47. 14 pp. <http://www.biomedcentral.com/1471-2148/7/47>.
- Lertrit, Patcharee, Samerchai Poolsuwan, Rachanie Thosarat, Thitima Sanpachudayan, Hathaichanoke Boonyarit, Chatchai Chinpaisal & Bhoom Suktitipat. 2008. Genetic history of Southeast Asian populations as revealed by ancient and modern human mitochondrial DNA analysis. *American Journal of Physical Anthropology* 137(4). 425–440.
- Li Changbao, Ailing Zhou & Tao Sang. 2006a. Genetic analysis of rice domestication syndrome with wild annual species *Oryza nivara*. *New Phytologist* 170. 185–194.
- 2006b. Rice domestication by reducing shattering. *Science* 311. 1936–1939.
- Li Wanglong & Bikram S. Gill. 2006. Multiple genetic pathways for seed shattering in the grasses. *Functional Integrated Genomics* 6. 300–309.
- Lì Jīn, Peter A. Underhill, Vishal Doctor, Ronald W. Davis, Peidong Shen, L. Luca Cavalli-Sforza & Peter J. Oefner. 1999. Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations. *Proceedings of the National Academy of Sciences of the United States of America* 96. 3796–3800.
- Lin Zhongwei, Megan E. Griffith, Xianran Li, Zuofeng Zhu, Lubing Tan, Yongcai Fu, Wenxu Zhang, Xiangkun Wang, Dauxin Xie & Chuanqing Sun. 2007. Origin of seed shattering in rice (*Oryza sativa* L.). *Planta* 266. 11–20.
- Londo, Jason P., Yu-chung Chiang, Kuo-Hsiang Hung, Tzen-Yuh Chiang & Barbara A. Schaal. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* 103(25). 9578–9583.
- Ma Jianxin & Jeffrey L. Bennetzen. 2004. Rapid recent growth and divergence of rice nuclear genomes. *Proceedings of the National Academy of Sciences of the United States of America* 101(34). 12404–12410.
- Macauley, Vincent, Catherine Hill, Alessandro Achilli, Chiara Rengo, Douglas Clarke, William Meehan, James Blackburn, Ornella Semino, Rosario Scozzari, Fulvio Cruciani, Adi Taham, Norazila Kassim Shaari, Joseph Maripa Raja, Patimah Ismail, Zafarina Zainuddin, William Goodwin, David Bulbeck, Hans-Jürgen Bandelt, Stephen Oppenheimer, Antonio Torroni & Martin Richards. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308. 1034–1036.
- Metspalu, Mait, Toomas Kivisild, Ene Metspalu, Jüri Parik, Georgi Hudjashov, Katrin Kadma, Piia Serk, Monika Karmin, Doron M. Behar, M. Thomas P. Gilbert, Phillip Endicott, Sarabjit Mastana, Surinder S. Papiha, Karl Skorecki, Antonio Torroni & Richard Villems. 2004. Most of the extant mtDNA boundaries in South and Southeast Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BioMed Central, Genetics* 5: 26. 25 pp. doi:10.1186/1471-2156-5-26.



- Nakamura Shin-ichi. 2010. The origin of rice cultivation in the Lower Yangtze region. *China. Archaeological and Anthropological Sciences* 2(2). 107–113.
- Nasu Hirō, Arata Momohara, Yoshinori Yasuda & Jiejun He. 2006. The occurrence and identification of *Setaria italica* (L.) P. Beauv. (foxtail millet) grains from the Chengtoushan site (ca. 5800 cal B.P.) in central China, with reference to the domestication centre in Asia. *Vegetation History and Archaeobotany* 16(6). 481–494.
- Niederer, Barbara. 1998. *Les langues hmong-mjen (miao-yao): Phonologie historique*. München: Lincom Europa.
- Onishi Kazumitsu, Kyoko Takagi, Manabu Kontani, Takumi Tanaka & Yoshio Sano. 2007. Different patterns of genealogical relationship found in the two major QTLs [i.e. quantitative trait loci] causing reduction of seed shattering during rice domestication. *Genome* 50. 757–766.
- Osada Toshiki. 1995. *Mundajin no Nōkō Bunka to Shokuji Bunka: Minzoku Gengogakuteki Kōsatsu*. [‘The Rice and Food Culture of Munda in Eastern India: An Ethnolinguistic Study’]. Kyōto: Kokusai Nihon Bunka Kenkyū Sentā.
- Ostapirat, Weera 2005. Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 107–131. London: Routledge Curzon.
- Pal, J.N. 1990. The early farming culture of northern India. *Bulletin of Deccan College Post-Graduate and Research Institute* 49. 297–304.
- Parsons, B. J., H. J. Newbury & M. T. Jackson. 1999. The genetic structure and conservation of aus, aman and boro rices from Bangladesh. *Genetic Resources and Crop Evolution* 46(6). 587–598.
- Pictet, Adolphe. 1859, 1863. *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique (première partie, seconde partie)*. Paris: Joël Cherbuliez.
- Pokharia, Anil. 2008. Palaeoethnobotanical record of cultivated crops and associated weeds and wild taxa from Neolithic site, Tokwa, Uttar Pradesh, India. *Current Science* 94(2). 248–255.
- Poloni, Estella Simone, Ornella Semino, Giuseppe Passarino, A.S. Santachiara-Benerecetti, I. Dupanloup, André Langaney & Laurent Excoffier. 1997. Human genetic affinities for Y chromosome P49a,f/*TaqI* haplotypes show strong correspondence with linguistics. *American Journal of Human Genetics* 61. 1015–1035 (cf. the erratum published in 1998 in the *American Journal of Human Genetics* 62. 1267).
- Poloni, Estella Simone, Nicolas Ray, Stefan Schneider & André Langaney. 2000. Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes. In Jean-Louis Dessalles & Laleh Ghadakpour (eds) *Proceedings: Evolution of Language, 3rd International Conference 3-6 April 2000*, 185–186. Paris: École Nationale Supérieure des Télécommunications.
- Purnell, Herbert C., Jr. 1970. *Toward a Reconstruction of Proto Miao-Yao*. Ithaca: Cornell University Ph.D. dissertation.
- Purugganan, Michael D. & Dorian Q Fuller. 2009. The nature of selection during plant domestication. *Nature* 457. 843–848.
- Purugganan, Michael D. 2010. The evolution of rice: Molecular vignettes of its origin and spread. *Archaeological and Anthropological Sciences* 2(2). 61–68.

- Qamar, Raheel, Qasim Ayub, Aisha Mohyuddin, Agnar Helgason, Kehkashan Mazhar, Atika Manoor, Tatiana Zerjal, Chris Tyler-Smith & S. Qasim Mehdi. 2002. Y-chromosomal variation in Pakistan. *American Journal of Human Genetics* 70. 1107–1124.
- Quintana-Murci, Lluís, C. Krausz, Tatiana Zerjal, S. Hamid Sayar, Michael F. Hammer, S. Qasim Mehdi, Qasim Ayub, Raheel Qamar, Aisha Mohyuddin, U. Radhakrishna, Mark A. Jobling Chris Tyler-Smith & Ken McElreavey. 2001. Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *American Journal of Human Genetics* 68. 537–542.
- Quintana-Murci, Lluís, Raphaëlle Chaix, R. Spencer Wells, Doron M. Behar, Hamid Sayar, Rosaria Scozzari, Chiara Rengo, Nadia al-Zahery, Ornella Semino, A. Silvana Santachiara-Benerecetti, Alfredo Coppa, Qasim Ayub, Aisha Mohyuddin, Chris Tyler-Smith, S. Qasim Mehdi, Antonio Torroni & Ken McElreavey. 2004. Where West meets East: The complex mtDNA landscape of the Southwest and Central Asian corridor. *American Journal of Human Genetics* 74. 827–845.
- Ratliff, Martha. 2004. Vocabulary of environment and subsistence and in the Hmong-Mien protolanguage. In Nicholas Tapp, Jean Michaud, Christian Culas & Gary Yia Lee (eds) *Hmong/Miao in Asia*, 147–165. Chiang Mai: Silkworm Books.
- . 2010. *Hmong-Mien Language History*. Canberra: Pacific Linguistics.
- Reddy, B. Mohan, B.T. Langstieh, Vikrant Kumar, T. Nagaraja, A. N. S. Reddy, M. Aruna, K. Thangaraj, A.G. Reddy & Lalji Singh. 2007. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. *Public Library of Science One* 2(11). e1141. doi:10.1371/journal.pone.0001141.
- Reich, David, Kumarasamy Thangaraj, Nick Patterson, Alkes L. Price & Lalji Singh. 2009. Reconstructing Indian population prehistory. *Nature* 461. 489–495.
- Revel, Nicole. 1988. *Le riz en Asie du sud-est* (3 vols.). Paris: Éditions de l'École des Hautes Études en Sciences Sociales.
- Roychoudhury, Susanta, Sangita Roy, Analabha Basu, Rajat Banerjee, H. Vishwanathan, M.V. Usha Rani, Samir K. Sil, Mitashree Mitra & Partha P. Majumder. 2001. Genomic structures and population histories of linguistically distinct tribal groups of India. *Human Genetics* 109. 339–350.
- Ruddiman, W. F., Z. Guo, X. Zhou, H. Wu & Y. Yu. 2008. Early rice farming and anomalous methane trends. *Quaternary Science Reviews* 27. 1291–1295.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444.
- Sahoo, Sanghamitra, Anamika Singh, G. Himabindu, Jheeman Banerjee, T. Sitalaximi, Sonali Gaikwad, R. Trivedi, Phillip Endicott, Toomas Kivisild, Mait Metspalu, Richard Villems & V. K. Kashyap. 2006. A prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios. *Proceedings of the National Academy of Sciences of the United States of America* 103(4). 843–848.
- Sang, Tao. 2009. Genes and mutations underlying domestication transitions in grasses. *Plant Physiology* 149. 63–70.
- Schwartz, Marianne & John Vissing. 2002. Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine* 347(8). 576–580.
- Sengupta, Sanghamitra, Lev A. Zhivotovsky, Roy King, S.Q. Mehdi, Christopher A. Edmonds, Cheryl-Emiliane T. Chow, Alice A. Lin, Mitashree Mitra, Samir K. Sil, A. Ramesh, M. V. Usha Rani, Chitra M. Thakur, L. Luca Cavalli-Sforza, Partha P. Majumder & Peter A.

- Underhill. 2006. Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *The American Journal of Human Genetics* 78. 202–221.
- Sharma, G.R., V. D. Misra, D. Mandal, B. B. Misra & J. N. Pal. 1980. *Beginnings of Agriculture: From Hunting and Food Gathering to Domestication of Plants and Animals*. Allahabad: Abhinav Prakashan.
- Sharma, Swarkar, Ekta Rai, Prithviraj Sharma, Mamata Jena, Shweta Singh, Katayoon Darvishi, Audesh K. Bhat, A.J.S. Bhanwer, Pramod Kumar Towari & Rameshwar N.K. Bamezai. 2009. The Indian origin of paternal haplogroup R1a1\* substantiates the autochthonous origin of Brahmins and the caste system. *Journal of Human Genetics* 54. 47–55.
- Shi Hong, Dong Yong-li, Wen Bo, Xiao Chun-Jie, Peter A. Underhill, Shen Peidong, Ranajit Chakraborty, Jīn Lì & Sù Bīng. 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3–M122. *American Journal of Human Genetics* 77. 408–419.
- Skeat, Walter William & Charles Otto Blagden. 1906. *Pagan Races of the Malay Peninsula* (2 vols.). London: Macmillan and Company.
- Sù Bīng, Xiao Junhua, Peter Underhill, Ranjan Deka, Zhang Weiling, Joshua Akey, Huang Wei, Shen Di, Daru Lu, Luo Jingchun, Chu Jiayou, Tan Jiazhen, Shen Peidong, Ron Davis, Luca Cavalli-Sforza, Ranajit Chakraborty, Xiong Momiao, Du Ruofu, Peter Oefner, Chen Zhu & Jīn Lì. 1999. Y-chromosome evidence for a northward migration of modern humans into eastern Asia during the last Ice Age. *American Journal of Human Genetics* 65. 1718–1724.
- Sù Bīng, Chunjie Xiao, Ranjan Deka, Mark T. Seielstad, Daoroong Kangwanpong, Junhua Xiao, Daru Lu, Peter Underhill, Luca Cavalli-Sforza, Ranajit Chakraborty & Jīn Lì. 2000. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. *Human Genetics* 107(6). 582–590.
- Sweeney, Megan T. & Susan R. McCouch. 2007. The complex history of the domestication of rice. *Annals of Botany* 100. 951–957.
- Tanaka Katsunori, Takeshi Honda & Ryuji Ishikawa. 2010. Rice archaeological remains and the possibility of DNA archaeology: Examples from Yayoi and Heian periods of northern Japan. *Archaeological and Anthropological Sciences* 2(2). 69–78.
- Tewari, Rakesh, R. K. Srivastava, K. K. Singh, K. S. Saraswat & I. B. Singh. 2003. Preliminary report of the excavation at Lahuradewa, District Sant Kabir Nagar, U.P. 2001–2002: Wider archaeological implications. *Prāgdhāra, Journal of the Uttar Pradesh Archaeological Department* 13. 37–76.
- Tewari, Rakesh, R. K. Srivastava, K. K. Singh, K. S. Saraswat, I. B. Singh, M. S. Chauhan, A. K. Pokharia, A. Saxena, V. Prasad & M. Sharma. 2006. Second Preliminary Report of the excavations at Lahuradewa. District Sant Kabir Nagar, U. P. 2002–2003–2004 & 2005–2006. *Prāgdhāra, Journal of the Uttar Pradesh Archaeological Department* 16. 35–68.
- Tewari, Rakesh, R. K. Srivastava, K. K. Singh, K. S. Saraswat, I. B. Singh & K.K. Singh. 2009. Early farming at Lahuradewa, Ganga Basin. *Prāgdhāra, Journal of the Uttar Pradesh Archaeological Department* 18. 347–374.
- Thangaraj, Kumarasamy, Vempati Sridhar, Toomas Kivisild, Alla G. Reddy, Gyaneshwer Chaubey, Vijay Kumar Singh, Suminder Kaur, Pooja Agarawal, Amit Rai, Jalaj Gupta, Chandana Basu Mallick, Niraj Kumar, Thrimulaisamy P. Velavan, Rajanbabu Suganthan, Divi Udaykumar, Rashmi Kumar, Rachana Mishra, Arif Khan, Chitikineni Annapurna & Lalji Singh. 2005. Different population histories of the Mundari- and Mon-Khmer-speaking

- Austro-Asiatic tribes inferred from the mtDNA 9-bp deletion/insertion polymorphism in Indian populations. *Human Genetics* 116. 507–517.
- Vaughan, Duncan A. 1994. *The Wild Relatives of Rice: A Genetic Resources Handbook*. Manila: International Rice Research Institute.
- Vavilov, Nikolai Ivanovič. 1926. Studies on the origin of cultivated plants. *Trudy po Prikladnoi Botanike, Genetike i Selekcii (Bulletin of Applied Botany, Genetics and Plant Breeding)* 16. 1–248.
- Vitte, C., T. Ishii, F. Lamy, D. Brar & O. Panaud. 2004. Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). *Molecular Genetic Genomics* 272. 504–511.
- Wáng Fūshì & Máo Zōngwǔ. 1995. *Miáo-Yáo Yǔ Gǔyīn de Gòuni* [‘Phonological Reconstruction of Hmong-Mien’]. Peking: Zhōngguó Shèhuì Kēxué Chūbǎnshè.
- Weber, Steve, Heather Lehman, Timothy Barela, Sean Hawks & David Harriman. 2010. Rice or millets: Early farming strategies in prehistoric central Thailand. *Archaeological and Anthropological Sciences* 2(2). 79–88.
- Wells, R. Spencer, Nadira Yuldasheva, Ruslan Ruzibakiev, Peter A. Underhill, Irina Evseeva, Jason Blue-Smith, Jīn Lì, Sù Bīng, Ramasamy Pitchappan, Sadagopal Shanmugalakshmi Karupiah Balakrishnan, Mark Read, Nathaniel M. Pearson, Tatiana Zerjal, Matthew T. Webster, Irakli Zholoshvili, Elena Jamarjashvili, Spartak Gambarov, Behrouz Nikbin, Ashur Distiev, Oganazar Aknazarov, Pierre Zalloua, Igor Tsoy, Mikhail Kitaev, Mirsaid Mirrahimov, Ashir Chariev & Walter F. Bodmer. 2001. The Eurasian heartland: A continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences* 98. 10244–10249.
- Wen Bo, Xie Xuanhua, Gao Song, Li Hui, Shi Hong, Song Xiufeng, Qian Tingzhi, Xiao Chunjie, Jin Jianzhong, Sù Bīng, Lu Daru, Ranajit Chakraborty & Jīn Lì. 2004. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans. *American Journal of Human Genetics* 74(5). 856–865.
- Wen Bo, Hui Li, Song Gao, Xianyun Mao, Yang Gao, Feng Li, Feng Zhang, Yungang He, Yongli Dong, Youjun Zhang, Wenju Huang, Jianzhong Jin, Chunjie Xiao, Daru Lu, Ranajit Chakraborty, Sù Bīng, Ranjan Deka & Jīn Lì. 2005. Genetic structure of Hmong-Mien speaking populations in East Asia as revealed by mtDNA lineages. *Molecular Biology and Evolution* 22(3). 725–734.
- Williams, R. Sander. 2002. Another surprise from the mitochondrial genome. *New England Journal of Medicine* 347(8). 609–612.
- Xue Yali, Tatiana Zerjal, Weidong Bao, Suling Zhu, Qunfang Shu, Jiujiu Xu, Ruofu Du, Songbin Fu, Pu Li, Matthew E. Hurles, Huanming Yang & Chris Tyler-Smith. 2006. Male demography in East Asia: A north-south contrast in human population expansion times. *Genetics* 172. 2431–2439.
- Yamane, Hiroko, Tomoko Ito, Harumi Ishikubo, Masaki Fujisawa, Harumi Yamagata, Kozue Kamiya, Yukiya Ito, Masao Hamada, Hiroyuki Kanamori, Hiroshi Ikawa, Yuichi Katayose, Jianzhong Wu, Takuji Sasaki & Takashi Matusmoto. 2009. Molecular and evolutionary analysis of the *Hd6* photoperiod sensitivity gene within genus *Oryza*. *Rice* 2. 56–66.
- Yao Yong-Gang, Long Nie, Henry Harpending, Yun-Xin Fu, Zhi-Gang Yuan & Ya-Ping Zhang. 2002. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *American Journal of Physical Anthropology* 118. 63–76.

- Yao, Yong-Gang, Qing-Peng Kong, Xia-Yong Man, Hans-Jürgen Bandelt & Ya-Ping Zhang. 2003. Reconstructing the evolutionary history of China: A caveat about inferences drawn from ancient DNA. *Molecular Biology and Evolution* 20(2). 214–219.
- Yasuda Yoshinori. 2002. Origins of pottery and agriculture in East Asia. In Yoshinori Yasuda (ed.) *The Origins of Pottery and Agriculture*, 119–142. New Delhi: Lusre Press and Roli Books, for the International Center for Japanese Studies.
- Zerjal, Tatjana, B. Dashnyam, A. Pandya, Manfred Kayser, Lutz Roewer, F. R. Santos, W. Schiefenhovel, N. Fretwell, M. A. Jobling, S. Harihara, K. Shimizu, D. Semjidmaa, Antti Sajantila, P. Salo, M. H. Crawford, E. K. Ginter, Oleg V. Efgrafov & Chris Tyler-Smith. 1997. Genetic relationships of Asian and Northern Europeans, revealed by Y chromosomal DNA analysis. *American Journal of Human Genetics* 62. 1174–1183.
- Zhao, Zhijun. 2010. New data and new issues for the study of the origin of rice agriculture in China. *Archaeological and Anthropological Sciences* 2(2). 99–105.
- Zhu Qihai & Song Ge. 2005. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* 167. 249–265.
- Zide, Norman H. & Gregory D.S. Anderson. 1999. The Proto-Munda verb system and some connections with Mon-Khmer. In *Working Papers of the International Symposium on South Asian Languages: Contact, Convergence and Typology, Tokyo, December 6-9, 1999*, 400–419. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- 2003. Towards an analysis of the South Munda verbal system. In Karen L. Adams, Thomas John Hudak & F.K. Lehman (eds) *Papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society*, 243–259. Tempe, Arizona: Arizona State University Program for Southeast Asian Studies.
- Zong, Y., Z. Chen, J. B. Innes, C. Chen, Z. Wang & H. Wang. 2007. Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature* 449. 459–462.