Patterns of population genetic structure among Australian and South Pacific humpback whales (*Megaptera novaeangliae*)

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Biological Sciences,

The Australian National University, Canberra, in collaboration with the

Australian Marine Mammal Centre, Hobart

By

*Natalie Tara Schmitt*

B. Sc. (Hons) in Zoology and Botany

University of Western Australia, Perth

October 2012
There Leviathan,

Hugest of living creatures, on the deep

Stretch’d like a promontory sleeps or swims,

And seems like a moving land; and at his gills

Draws in, and at his breath spouts out a sea.

John Milton, *Paradise Lost*, quoted in title page to the first, English edition of *Moby-Dick*
DECLARATION

The research presented in this thesis is my own original work except where due reference is given in the text. All the chapters were the product of investigations carried out jointly with others but in all cases I am the principal contributor to the work. No part of this thesis has been submitted for any previous degree.

............................................................

Natalie Tara Schmitt

October, 2012
THESIS PLAN

The dissertation is written as a series of four self-contained chapters ready for publication plus a concluding chapter. All chapters are written in the format of the target journal with tables, figures and appendices included at the end of each chapter. As such the pronoun “we” is used to represent co-authors of material intended for publication. Below I outline the contributions of my co-authors.

Chapter 1 - Low levels of genetic differentiation characterize Australian humpback whale (Megaptera novaeangliae) populations.

I was responsible for collecting the eastern Australian humpback skin samples, with Mike Double collecting the samples from western Australia. I was also responsible for conceptual development, laboratory work, statistical analysis and writing, with editorial and analytical assistance from Mike Double, Rod Peakall and Simon Jarman. External co-authors Scott Baker and Curt Jenner provided editorial assistance. I was also given advice in the lab by Simon Jarman and James Marthick.

Chapter 1 has been provisionally accepted for publication subject to revision in an international peer-reviewed journal (Marine Mammal Science). The chapter has also been presented at the International Whaling Commission Stock definition and Southern Hemisphere Scientific Sub-committee meetings held in Panama, 2012 (SC/64/SH15).

Chapter 2 - Mixed-stock analysis of humpback whales (*Megaptera novaeangliae*) on Antarctic feeding grounds.

The South Pacific Whale Research Consortium provided mtDNA and nuclear data from humpback whale breeding populations of the South Pacific. Mike Double, Simon Childerhouse, Rochelle Constantine, Nick Gales, Curt Jenner and Dave Paton assisted in the collection of skin samples from the Southern Ocean. Scott Baker also provided seven samples collected previously in the Southern Ocean. I was responsible for the conceptual development of this chapter as well as the laboratory work, statistical analyses and writing. Editorial advice on the manuscript was also given by Rod Peakall, Mike Double, Scott Baker and Simon Jarman.


Chapter 3 – Re-assessing the genetic evidence for sex-specific migratory route choice in eastern Australian humpback whales (*Megaptera novaeangliae*)

I was responsible for most of the conceptual development and statistical analyses of this chapter with assistance from Rod Peakall and Mike Double. Genetic data was sourced from chapter 1 and chapter 2 with the Evan’s Head samples collected, DNA extracted and sequenced by Andrea Polanowski and Simon Jarman. I received editorial assistance with the chapter from Rod Peakall and Mike Double.

Australian humpback whales (*Megaptera novaeangliae*). This chapter is formatted as a manuscript for submission to *Marine Ecology Progress Series*.

**Chapter 4 - Development and evaluation of single nucleotide polymorphism (SNP) markers for population structure analysis in the humpback whale (*Megaptera novaeangliae*)**.

In this chapter I was responsible for all intronic SNP discovery and Andrea Polanowski ascertained all exonic SNPs, with laboratory advice and assistance from Simon Jarman. I was also responsible for the conceptual design of the study, statistical analyses and writing, with editorial contributions from Simon Jarman, Mike Double and Rod Peakall.

Schmitt, N.T., A.M. Polanowski, M.C. Double, N. Gales, C.S. Baker, D. Steel, R. Peakall, et al. Small numbers of single nucleotide polymorphism (SNP) markers can detect population structure in the humpback whale (*Megaptera novaeangliae*). This chapter is formatted as a manuscript for submission to *Marine Mammal Science*.

**Appendix – Development and application of SNPs for humpback whale population genetics (literature review)**.

This review was written at the onset of my PhD in 2008 as a way of informing myself and my supervisors of the available literature on SNP discovery and SNP genotyping in preparation for my fourth chapter. I was responsible for the collation of the literature and writing of the review. Simon Jarman, Mike Double and Rod Peakall provided some editorial contribution.
This project was a huge but rewarding challenge for me: despite having little experience in the field of molecular ecology I was inspired and motivated to be a part of the ‘golden age’ of genetic research, an age where we can now utilise genetic tools to learn more about elusive species like the humpback whale to inform management decisions. The Australian humpback whale also represents a suitable demographic and genetic model for the management of less tractable species of baleen whales and for the general study of gene flow among long-lived mobile vertebrates, due to their relatively high abundance and the ease with which they can be identified from natural markings. A symbol of global marine conservation, working with humpback whales has taught me what can be achieved in the field of population genetics with elusive species, and have greatly inspired me through the work of some amazing people I’ve had the privilege to meet and collaborate with over the course of the project.

This project would have been impossible to complete without the help of many people. Most of them are co-authors of my chapters or are named in the Acknowledgements section of each chapter. I would therefore like to use this section of the thesis to acknowledge those people that made great contributions to the journey through my PhD study and who are not necessarily colleagues.

I must first start by thanking my three supervisors. Their patience, understanding, encouragement and experience have been so valuable in my growth as a scientist. Thank you to Simon Jarman for your advice in the lab, particularly with the SNP discovery work; having come into this project with very little experience in genetics, your patience with me, expert guidance and ability to help resolve problems very quickly and with little hassle, as well as your friendly ear during times of frustration has made my experience in the lab a pleasurable one. My primary supervisor, Mike Double, thank you.
for firstly believing in my capabilities to undertake this project and thankyou especially for teaching me how to be a good and above all ethical scientist. Although your generous but tough approach was painful for me at times, I have learnt to demand good science from myself as well as from the scientific papers I read, after all, it is critical to get the science right when you’re dealing with species conservation. And finally a special gratitude to my university supervisor, Professor Rod Peakall, your dedication to every one of your students and postdocs is astounding and I was never once left feeling like I was anything less than a top priority. You have always shown a great respect and belief in me and have instilled in me the skills and values that will continue to help me become a more rigorous scientist. You have helped make science exciting for me and taught me the joy of working in a positive, welcoming and professional atmosphere for research.

So keen was I to undertake this PhD, I volunteered in the genetics lab at the University of Tasmania for 6 months under the generous and patient guidance of Kevin Redd, who was doing his PhD at the time. My deep appreciate goes to him for teaching me the basics of genetic lab work and allowing me to make mistakes with his precious samples; I don’t think I would have been permitted to take on the PhD without this experience.

My field work was definitely one of the most enjoyable aspects of the project. We are so privileged to be able to get up close to these magnificent animals. There were some essential people and organisations that were particularly supportive during this time. My appreciation to Jenny Robb, Scotty Sheehan, the Sapphire Coast Marine Discovery Centre and the Eden community for your support and interest in the project; it is so wonderful to have a community involved in conservation research and these guys really do value their whales. Thanks to Dave Paton for passing on his years of wisdom in the field and his ability to get me close enough to the whales that I simply could not miss.
with a biopsy rifle, it was a joy to work with you. And to all the people involved in the Antarctic Whale Expedition in 2010, this was an absolute highlight thanks to an incredible team. I was so privileged to work with such an experienced group of people, where everyone put in an enormous effort to help me obtain enough samples for my study. I learnt a great deal about field work in extreme conditions on this voyage and I still pinch myself that I actually got the opportunity to visit this magical place and see humpback whales on their Southern Ocean feeding grounds, an experience I will never forget.

Throughout the journey, which has been incredibly tough at times, I was grateful to have the emotional support of some wonderful people. A big thankyou to all the other PhD students at the Australian Antarctic Division, SCUM as we called ourselves; there was both tears and laughter during our time together but your support will always be remembered and valued. My deep appreciation to my good friend Rebecca Leaper, your unwavering support, advice and friendship during this time was invaluable. I will never forget our many conversations on science and conservation, they were always tremendously inspiring and motivating. My family was also instrumental in their support and encouragement of me, particularly during the tough times. In particular, I want to thank my Grandmother for letting me live with her for two years and for being there for me when I needed a shoulder to cry on.

Finally, I want to thank my partner David Donnelly whom I met when he volunteered on my very first field trip off Eden. You have been so patient, supportive, unwavering and understanding and words simply cannot express my gratitude to you. Thankyou for being my rock.

And of course, I cannot forget my study species….it was never difficult to be inspired by you.
ABSTRACT

Humpback whales undertake long-distance seasonal migrations between low latitude winter breeding grounds and high latitude summer feeding grounds. Although arguably one of the best studied of all baleen whales, there remain some critical gaps in our understanding of their population structure, migratory movement and the mixing of putative populations on the feeding grounds. Addressing these uncertainties is important in the development of demographic models that reconstruct the historical trajectory of population decline and recovery following the cessation of commercial whaling.

Utilising both mitochondrial and nuclear genetic markers, this thesis examines the population structure and distribution of humpback whales that migrate to separate winter breeding grounds along the north-western and north-eastern coasts of Australia, and their interaction with the endangered populations of the South Pacific. The project investigated three important gaps in knowledge: population structure among putative breeding populations, the mixing of breeding populations on high latitude Antarctic feeding grounds and evidence for sex-specific migration along the eastern Australian migratory corridor. The thesis also reports the discovery and utility of novel nuclear genetic markers (single nucleotide polymorphisms, SNPs). These markers hold promise for facilitating more effective multi-laboratory collaboration.

Among the Australian putative populations, weak but significant differentiation was detected across ten microsatellite loci and mitochondrial control region sequences. This pattern of low level differentiation is emerging as a characteristic of Southern Hemisphere humpback whale populations indicating extensive movement at least historically, if not presently.
As the first step towards assessing the mixing of Australian and endangered South Pacific humpback whale breeding populations on the Antarctic feeding grounds, a series of simulations were conducted to estimate the statistical power of both mitochondrial and nuclear microsatellite data from these populations for a mixed-stock analysis (MSA). The results of these simulations confirmed that we can draw robust conclusions from our MSA of Antarctic feeding ground samples collected south of eastern Australia and New Zealand in 2010. Using combined mtDNA and microsatellite datasets revealed substantial contributions from both eastern Australia and New Caledonia, but not western Australia; strengthening emerging evidence that these Antarctic waters are utilized by humpback whales from both eastern Australia and the more vulnerable breeding population of New Caledonia, representing Oceania.

There was no compelling evidence for sex-specific migration within the eastern Australian breeding population as indicated by the lack of significant differences detected in the patterns of haplotype sharing, haplotype frequency or haplotype differentiation between males and females. Instead, the significant differentiation revealed between the sexes at the nucleotide level for one sampling location and between sampling locations at the haplotype level suggests that humpback whale migration along eastern Australia may be more complex than previously thought.

Increasing the statistical power of our genetic datasets through the addition of new informative markers, including the SNPs discovered in this project, and incorporating non-genetic data, will assist in future studies of the population genetic structure and dynamics of Southern Hemisphere humpback whales.
TABLE OF CONTENTS

PREFACE ................................................................................................................................. 1

Chapter 1: Low levels of genetic differentiation characterize Australian humpback whale (*Megaptera novaeangliae*) populations ................................................. 8

ABSTRACT ............................................................................................................................. 9

INTRODUCTION .................................................................................................................... 10

METHODS ............................................................................................................................. 13

Population definition in the Southern Hemisphere ............................................................. 13

Sample collection, DNA extraction and sex identification ................................................. 13

Microsatellite loci ................................................................................................................ 14

Microsatellite validation ................................................................................................. 15

mtDNA ............................................................................................................................... 15

Statistical analysis ........................................................................................................... 16

RESULTS .............................................................................................................................. 19

Sample size and sex ratio ............................................................................................... 19

Genetic diversity .............................................................................................................. 20

Genetic differentiation and population structure analysis .............................................. 21

DISCUSSION ....................................................................................................................... 23
II. Source data ........................................................................................................ 66

Molecular Genetic Analysis .................................................................................. 66

Statistical analysis ................................................................................................. 67

I. Genetic structure ................................................................................................. 67

II. Simulations ......................................................................................................... 68

   Evaluating the resolution of source datasets for individual assignment and MSA
   ......................................................................................................................... 69

   i. Individual assignment .................................................................................. 69

   ii. Mixed-stock analysis (MSA) ...................................................................... 70

Improving the resolution of our source datasets for MSA ................................. 71

III. Mixed-stock estimation of the 62 individual samples collected from

     Antarctic feeding Area V ............................................................................. 73

RESULTS ............................................................................................................ 73

Statistical Analysis ............................................................................................... 74

I. Genetic structure ............................................................................................... 74

II. Simulations ....................................................................................................... 75

   Evaluating the resolution of source datasets for individual assignment and MSA.
   ......................................................................................................................... 75

   i. Individual assignment .................................................................................. 75
ii. Mixed Stock Analysis (MSA) ................................................................. 76

Improving the resolution of our source datasets for MSA ....................... 77

III. Mixed-stock estimation of the 62 individual samples collected from
Antarctic feeding Area V ................................................................. 78

DISCUSSION ...................................................................................... 80

Genetic structure ................................................................. 80

Individual assignment tests .................................................. 81

Feasibility of a mixed stock analysis ........................................ 82

Performance of mixed-stock simulations under the two hypotheses for population
structure .................................................. 84

Improving the accuracy and precision of mixed-stock analyses ............ 84

Simulation limitations .......................................................... 85

Population composition of Area V samples .................................. 86

ACKNOWLEDGEMENTS ................................................................. 88

LITERATURE CITED ................................................................. 90

TABLES AND FIGURES ............................................................. 104
Chapter 3: Re-assessment of the genetic evidence for sex-specific migratory route choice in eastern Australian humpback whales (*Megaptera novaeangliae*)

ABSTRACT

INTRODUCTION

MATERIALS AND METHODS

Samples

mtDNA sequence

Statistical Analysis

I. Analysis of Molecular Variance (AMOVA)

II. Haplotype sharing

III. Haplotype frequency

Influence of small and uneven sizes

RESULTS

I. Between sampling locations

II. Between sampling locations by migration direction

III. Between sexes within sampling locations

IV. Between sexes within sampling locations by migration direction

Influence of small and uneven sizes
Chapter 4: Development and evaluation of single nucleotide polymorphism (SNP) markers for population structure analysis in the humpback whale (*Megaptera novaeangliae*)

ABSTRACT

INTRODUCTION

MATERIALS AND METHODS

Sample collection and DNA extraction

Intronic SNPs

SNP Validation

Data analysis

I. Genetic Diversity

II. Simulation of statistical power

III. Population differentiation

RESULTS

Intron sequencing
Intronic SNP detection

Assay development and SNP validation

Data analyses

I. Genetic Diversity and simulation of statistical power

II. Population differentiation

DISCUSSION

ACKNOWLEDGEMENTS

LITERATURE CITED

TABLES AND FIGURES

CONCLUSIONS

Future recommendations

Measuring gene flow when genetic differentiation is weak

Resolving population structure along the eastern Australian migratory corridor

Enhancing inferences of population structure and improving the resolution of mixed-stock analyses – mitogenomics?

SNPs – the marker of the future for long-term collaborative cetacean research?

Literature cited

Appendix: Development and application of SNP’s for humpback whale population genetics (literature review)
SNP DISCOVERY ................................................................. 204

Expressed Sequence Tags (EST) ........................................... 204

Reduced representation shotgun sequencing (RRS) ..................... 205

AFLPs – a random sequence approach ................................ 206

The targeted gene approach ................................................ 207

SNP Discovery in whales .................................................... 209

SNP GENOTYPING ............................................................. 211

Allele discrimination strategies ........................................... 211

I. Primer Extension (Advantage: Accurate genotyping, Disadvantage: Similar reagents as in PCR) .......................................... 211

II. Hybridisation (Advantage: Widely used, Disadvantage: Limited genotype discrimination) ........................................ 213

III. Ligation (Advantage: Variety of assay formats, Disadvantage: Multiple labelled probes required) .............................. 214

IV. Enzymatic Cleavage (Advantage: PCR amplification avoided, Disadvantage: requires large amount of DNA) .................. 215

Allele detection methods ..................................................... 216

I. Mass Spectrometry .......................................................... 216

II. Fluorescence signal-based detection ................................... 217

III. Chemiluminescence and Pyrosequencing .......................... 219
Two and a half decades have now passed since passage of the International Whaling Commission’s (IWC) moratorium on commercial whaling. Prior to this point, over most of the last century, exploitation of large whales was often immense. In the Antarctic alone, more than 2 million whales were killed, and with the revelations of widespread illegal catches by the Soviet Union (Yablokov 1994, Zemsky et al. 1995, Zemsky et al. 1996), the large Southern Ocean populations were reduced to a fraction of their original size (Clapham et al. 1999).

Our knowledge of the present status of the world’s baleen whales varies considerably between species, with the recovery of some populations proceeding strongly while others remain highly endangered. Of the eleven species of baleen whale (Order Cetacea, suborder Mysticeti) populations of four species are considered ‘critically endangered’ based on abundance estimates, including the blue whale, grey whale, bowhead whale and northern right whale. The northern right whale was amongst the first large whales to be hunted on a systematic, commercial basis with intensive shore whaling during the 17th and 18th century and indiscriminate illegal Soviet whaling in the 19th century making this species the most threatened of all baleen whales throughout all of its range (reviewed in Brownell Jr et al. 1986).

The humpback whale (Megaptera novaeangliae) is arguably the most studied of all the baleen whale species. A coastal species over much of its extensive world-wide range, humpback whales bore the initial brunt of whaling activities in many areas. Often the first species to be taken, it was frequently hunted to commercial extinction, after which other whales were targeted (Tønnessen and Johnsen 1982, Clapham et al. 1997).
Hundreds of thousands of humpback whales were killed during the period of commercial whaling throughout much of their range, particularly in the Southern Ocean, with many populations reduced by more than 90% of their original size. As a result, in 1970 the humpback whale was listed as an endangered species under the Endangered Species Conservation Act of 1969 by the United States government. All populations were listed as one global entity under the act as a precautionary measure, regardless of their individual status.

The impact of whaling and the recovery of whale populations is a key focus of the International Whaling Commission (IWC) scientific committee. Estimating the former abundance and maximum environmental carrying capacity of each of these exploited populations and reconstructing the historical trajectory of their decline are essential to accurately assess the true impact of whaling on the marine ecosystem, and to establish a baseline for a population’s recovery. This measurement of population recovery plays a crucial role in conservation management schemes, which for the most part, can mean the difference between the recovery or decline of a species.

Standard demographic models rely heavily on historical catch records from commercial whaling as well as estimates of biological parameters such as the rate of reproduction, age at sexual maturity and natural mortality (Best 2001, Clapham 2001, Jackson et al. 2008). However, there are a number of uncertainties associated with these models that can substantially over or under-estimate pre-exploitation abundance including discrepancies in taxonomic boundaries, population structure and contemporary abundance.

Modern molecular markers have been useful in helping us address these problems. Specifically, the use of highly variable mitochondrial and nuclear genetic markers have
provided high-resolution genetic information, allowing us to distinguish between individuals and populations within species with little phenotypic variation. They also offer clues on aspects of whale migration and movement from high latitude feeding grounds to low latitude tropical breeding grounds.

Utilising both mitochondrial and nuclear genetic markers, this thesis investigated the patterns of population structure among humpback whales that migrate along the east and west coast of Australia, and the neighbouring endangered populations of the South Pacific. Of important relevance to population assessments of Southern Hemisphere humpback whales, the thesis examines population structure among putative breeding populations, the mixing of breeding populations on high latitude Antarctic feeding grounds and evidence for sex-specific migration along the eastern Australian migratory corridor. The thesis also looks at the discovery and utility of novel nuclear genetic markers (single nucleotide polymorphisms, SNPs) that are easier to ascertain, have a well characterised mutation rate and are more universally comparable than the commonly used microsatellite markers.

Field work, which involved the collection of humpback whale skin biopsy samples, was conducted along the migratory corridors of eastern and western Australia and in the Southern Ocean south of eastern Australia and New Zealand as part of a six week Australian-New Zealand Antarctic Whale Expedition (AWE). I was also privileged to have access to an extensive mtDNA and microsatellite database from the breeding grounds of the South Pacific provided by the South Pacific Whale Research Consortium.

This dissertation represents the partial fulfilment of the requirements for the Degree of Doctor of Philosophy (PhD) at the Research School of Biology at the Australian
National University (ANU) and the Australian Marine Mammal Centre (AMMC) at the Australian Antarctic Division.

The outline of the thesis is as follows:

**Chapter 1** - Low levels of genetic differentiation characterize Australian humpback whale (*Megaptera novaeangliae*) populations.

This chapter evaluated the population genetic structure among Australian humpback whales at both maternally inherited mtDNA and biparentally inherited microsatellite markers, as well as extend previous analyses of mtDNA variation in a comparison of Australian humpback whales and the endangered populations of Oceania.

**Chapter 2** - Mixed-stock analysis of humpback whales (*Megaptera novaeangliae*) on Antarctic feeding grounds.

Using a series of simulations, this chapter evaluated the statistical power of microsatellite and mtDNA datasets from the putative humpback whale populations of Australia and Oceania for individual assignment and mixed-stock analysis given available samples size and the patterns of genetic divergence, and 2) Determined ways in which we can improve the accuracy and precision of mixed-stock analyses for these priority populations for future studies. In light of the simulation outcomes, the study then estimated the population composition of Antarctic feeding ground samples collected south of eastern Australia and New Zealand.

**Chapter 3** - Re-assessment of the genetic evidence for sex-specific migratory route choice in eastern Australian humpback whales (*Megaptera novaeangliae*).
This chapter investigated evidence for sex-specific migratory behaviour in humpback whales that migrate along the east coast of Australia by assessing the patterns of mitochondrial haplotype sharing, haplotype frequency and haplotype differentiation among males and females.

**Chapter 4** - Development and evaluation of single nucleotide polymorphism (SNP) markers for population structure analysis in the humpback whale (*Megaptera novaeangliae*).

In this chapter I developed a suite of informative SNP markers from intron sequences and estimate by computer simulation the statistical power of a combined panel of intronic and exonic SNPs to detect population genetic structure in humpback whales compared with a suite of microsatellite markers.

**Appendix** – Development and application of SNPs for humpback whale population genetics (literature review).

A review on SNP marker discovery and genotyping techniques as background research to my final chapter on the development and evaluation of SNPs for humpback whales.
LITERATURE CITED


