Evolution of the Mammalian Y Chromosome

Paul Damian Waters

B.Sc. (Hons.)

This thesis is submitted in total fulfilment of the

requirements for the degree of

Doctor of Philosophy.

Comparative Genomics Group,

Research School of Biological Sciences,

The Australian National University,

Acton, A.C.T., 2601

Australia.

December, 2002.

Table of Contents

List of Figures

List of Tables

Abstract

The human X and Y chromosomes evolved from a homologous pair of autosomes. Comparative studies can show how and when this happened. There are regions of the sex chromosomes that are conserved between eutherians and marsupials, which were part of the original therian X and Y chromosomes. Twelve genes conserved between the eutherian and marsupial X chromosomes have also been mapped to the monotreme X chromosome, defining the most ancient region of the human X chromosome that has been conserved since before the divergence of the therian and prototherian lineages. However, genes from human Xp were mapped to autosomes in marsupials and monotremes, mainly to tammar wallaby chromosome 5p and platypus chromosomes 1 and 2, defining a region added after the marsupial-eutherian divergence but before the eutherian radiation.

The human X chromosome represents 5% of the haploid genome. It is highly conserved between species and has undergone little change over its history. On the other hand, the Y chromosome represents only 2% of the haploid genome and is poorly conserved between species. The Y chromosome has been subjected to a high mutation rate and is largely degraded, leaving only a few X/Y-shared genes and small pseudoautosomal regions. Some of the shared genes in the non-recombining region show high conservation, whereas others have diverged in sequence and function.

To study the evolution of the mammalian Y chromosome, I attempted to clone and map genes from the human Y chromosome in *Macropus eugenii* (tammar wallaby), *Ornithorynchus anatinus* (platypus) and *Gallus gallus* (chicken). The mapping of *TB4X/Y* to wallaby 5p, with my previous demonstration that genes on the human Y were autosomal in marsupials, implied that the Y, as well as the X, was composed of conserved and added regions. Indeed most of the genes on the human Y chromosome were part of the added region. This demonstrated that only about 7Mb of the original 115Mb Y chromosome remains on the human Y chromosome.

The autosomal location of *TB4X/Y* in tammar wallaby was surprising, considering that this gene maps to the conserved long arm of the human X chromosome. This showed that the human X chromosome has undergone an inversion or duplication that re-positioned part of the added region on the long arm.

Abstract

Mapping human Y genes in monotremes can test the hypothesis that the genes conserved on the therian sex chromosomes were part of the original proto-X and -Y chromosomes. Location on the sex chromosomes of all mammals would demonstrate that regions of the human Y chromosome were conserved from the original proto-Y. However, mapping of conserved Y genes in *Ornithorynchus anatinus* (platypus) revealed that several genes conserved on the therian sex chromosomes were autosomal (on chromosome 6) in monotremes. The autosomal location of one of these genes, *ATRX*, located on the eutherian and marsupial X chromosomes and with a Y homologue in marsupials, excluded *ATRY* as the ancestral mammalian sex-determining gene. Genes conserved only on the therian sex chromosomes, but on platypus chromosome 6, were redefined as XtCR (X therian conserved region), and represent a block of genes either added to the therian sex chromosomes, or lost from the monotreme sex chromosomes.

Genes conserved on the therian sex chromosomes, but not the monotreme sex chromosomes (XtCR) were therefore mapped in chicken to distinguish between these hypotheses. *SOX3* and *RBMX* were shown to map to the short arm of chicken chromosome 4, co-localising with other genes conserved on the human sex chromosomes. This result, in combination with the map location of human X genes in pufferfish and zebrafish, showed the region previously thought to be the most ancient of the X, and therefore the Y (Stratum 1, 240-320 million years old), was composed of two blocks that had independent origins. One, containing genes on the X in all mammals, was originally on the proto-X and –Y, but has been completely lost from the Y. A second block, containing *ATRY*, *SOX3* and *RBMY* was added to the therian sex chromosomes after the divergence of monotremes, and is therefore only 130-170 millions years old. This addition included *ATRX*, *SOX3* and *RBMX*, which subsequently evolved into *ATRY*, *SRY* and *RBMY* and acquired roles in male sex and reproduction.

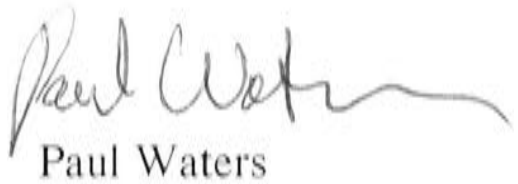
The original mammalian proto-Y chromosome once contained all of the genes that the human X chromosome currently bears. Over time the Y has lost most of these genes. Of those that remain, *SRY* and *RBMY* are now revealed not to be original members of the proto-Y, but to have been part of an addition to the therian Y chromosome.

Statement of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere, or extracted in whole or in part from a thesis by which I have qualified for, or been awarded another degree.

No other person's work has been used without due acknowledgment in the main text of the thesis.

This thesis has not been submitted for award of any other degree or diploma at any other tertiary institution.



Paul Waters

12th December, 2002

Publication list

Waters, P.D., Duffy, B., Frost, C.J., Delbridge, M.L. & Graves, J.A.M. (2001). The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80-130 million years ago. *Cytogenetics and Cell Genetics*, **92**, 74-79.

Waters, P.D., Kirby, P.J. & Graves, J.A.M. (2001). Assignment of the SMARCF1 gene to tammar wallaby chromosome 5p by fluorescence in situ hybridisation. *Cytogenetics and Cell Genetics*, **93**, 315-316.

Kirby, P.J., Waters, P.D., Delbridge, M.L., Svartman, M., Stewart, A. N., Nagai, K & Graves, J.A.M. (in press). Cloning and mapping of platypus *SOX2* and *SOX14*: Insights into SOX group B evolution. *Cytogenetic and Genome Research.*

Graves, J.A.M. & Waters, P.D. (in press). The rise and fall of the mammalian Y chromosome. In *Chromosomes Today*. Kluwer Academic Publishers, Netherlands.

Acknowledgements

I would like to thank first my supervisor, Professor Jenny Graves, for her enthusiasm, patience and the opportunity to do my PhD in her lab. Special thanks goes out to Marg for her words of wisdom over the years. Without Jenny and Marg I would not have made it this far.

I would also like to thank all the people from Jenny's lab, both past and present. Thankyou to my fellow PhD students Patrick, Ian, Rob, Tash, Swathi and Nis, for their advice and friendship over the years. Ian and Rob are from opposite ends of the spectrum, but both have made Canberra life that little bit better.

Patrick Kirby is one I must thank for his sometimes sane, but mostly insane conversation that always kept me guessing. He is a good friend.

Back home there are many I need to thank. Here's cheers to all the boys I grew up with: Bear, Bino, Pork Chops, Faz, Evo, Marno, Peggsy, Combat, and the big man Yinny. There are also many people at Benalla Football Club, too numerous to name, that require a thanks. I'm sure there are many more laughs and drinks to come.

Thanks to my mum and dad for all their love and support. To my brothers (Ben and Barn) and sisters (Bern and Ret), I could not have asked for a better group people to grow up with. I have many fond memories, both recent and distant, which are priceless to me.

Thankyou to Kerrie for her love, support and friendship over what is now most of my adult life. She knew I would make it, even when I didn't.

My love goes out to the Hyland family. The loss of Rhonda has been felt deeply by all. The Hyland's have been a second family to me, something I am sure will not change with time or distance.

Finally, I live for family and friends. To all of you, thanks again.

Chapter 1: Introduction

Sex chromosomes are thought to have evolved from a homologous pair of autosomes. One member became sex specific and was degraded. This hypothesis therefore implies that the mammalian Y chromosome represents a degraded relic of one member of this pair, whereas the X chromosome has changed little from its original state and is remarkably conserved in size and gene content across all mammals. All mammalian groups have an XX female and XY male sex chromosome system, or some variant of it. The Y is usually smaller than the X, and is heterochromatic and gene poor. The aim of this thesis is to determine the origin of the human sex chromosomes, particularly the Y.

Birds and snakes have a sex chromosome system in which males are ZZ and females (the heterogametic sex) are ZW. The W is usually smaller than the Z and is heterochromatic and gene poor. The small heterochromatic bird W is therefore equivalent to the small mammalian Y, both having much less genetic information than their respective partners, the Z and X. The same theory can be applied to account for the evolution of the mammalian X and Y, and the bird Z and W.

This theory, first put forward by Ohno (1967) on the basis of snake sex chromosome variation, has been developed (Charlesworth, 1991) to propose that the first requirement for differentiation of the mammalian X and Y chromosomes was the acquisition of a sex-determining locus by one member of an autosomal pair. Male advantage genes accumulated near this sex-determining region, and recombination between them was suppressed to maintain genes in a male specific region. Further accumulation extended this region of suppressed recombination (Fisher, 1930). When recombination is suppressed, the Y chromosome is open to degradation, as mutations and deletions cannot be repaired or eliminated via recombination. Degradation occurred via genetic drift (Muller's ratchet) and genetic hitchhiking. Thus the Y chromosome became mostly non-recombining and shrank to a mere relic of its former glory, whereas the X chromosome maintained its original state. Small pseudoautosomal regions (PARs) are all that remain of once extensive homology between the two chromosomes.

There is evidence from comparative gene mapping that one or more autosomal regions were added to the X or Y chromosome within the PAR, and then recombined to its partner. This extended PAR was then subjected to suppression of recombination and the Y underwent further degradation (Graves, 1995).

The mammalian Y chromosome determines few phenotypes and is not necessary for survival (as demonstrated by the survival of XO individuals). Uniquely, the Y chromosome displays a "functional coherence" (Lahn and Page, 1997) of male specific characters, such as sex determination, which is controlled by the *SRY* gene (Sinclair *et al.*, 1990; Koopman *et al.*, 1991). Deletion mapping established that there are also at least three (if not four) genes on the human Y chromosome involved in spermatogenesis (Vogt *et al.*, 1996; Kent-First *et al.*, 1999). Other phenotypes associated with the Y chromosome are gonadoblastoma (GBY) and a factor involved in stature (GCY).

In this chapter, I discuss the organization, evolution, function and relationship between sex chromosomes of the three mammalian groups. I also discuss comparisons with homologous regions of the chicken genome. I emphasise the importance of comparing sex chromosomes between different vertebrate groups to determine the evolutionary origin of sex chromosomes, and also to identify regions of functional importance.

1.1 Mammals and Birds

Vertebrates are a diverse group that includes fish, amphibians, reptiles, birds and mammals. Tetrapods (amphibians, reptiles, birds and mammals) diverged from fish about 400 million years before present (MyBP). Amphibians diverged from reptiles about 360MyBP, and mammals diverged from the reptile lineage about 310MyBP. The bird and reptilian lineages diverged from each other approximately 250MyBP (Benton, 1990) (Figure 1.1).

Two primary characteristics are popularly used to define mammals: they are furry animals that suckle their young with milk. Extant mammals are divided into three different groups: eutherians, marsupials and monotremes, arranged into two subclasses: Prototheria and Theria. Subclass Theria (eutherians and marsupials) diverged from Subclass Prototheria (monotremes) about 170MyBP (Hope *et al.*, 1990). Infraclass Eutheria and Metatheria (marsupials) diverged from each other about 130MyBP (Figure 1.1).

Prototheria consists of a single Order Monotremata, which includes only three species, the platypus and two species of echidna, all confined to Australasia. Our model

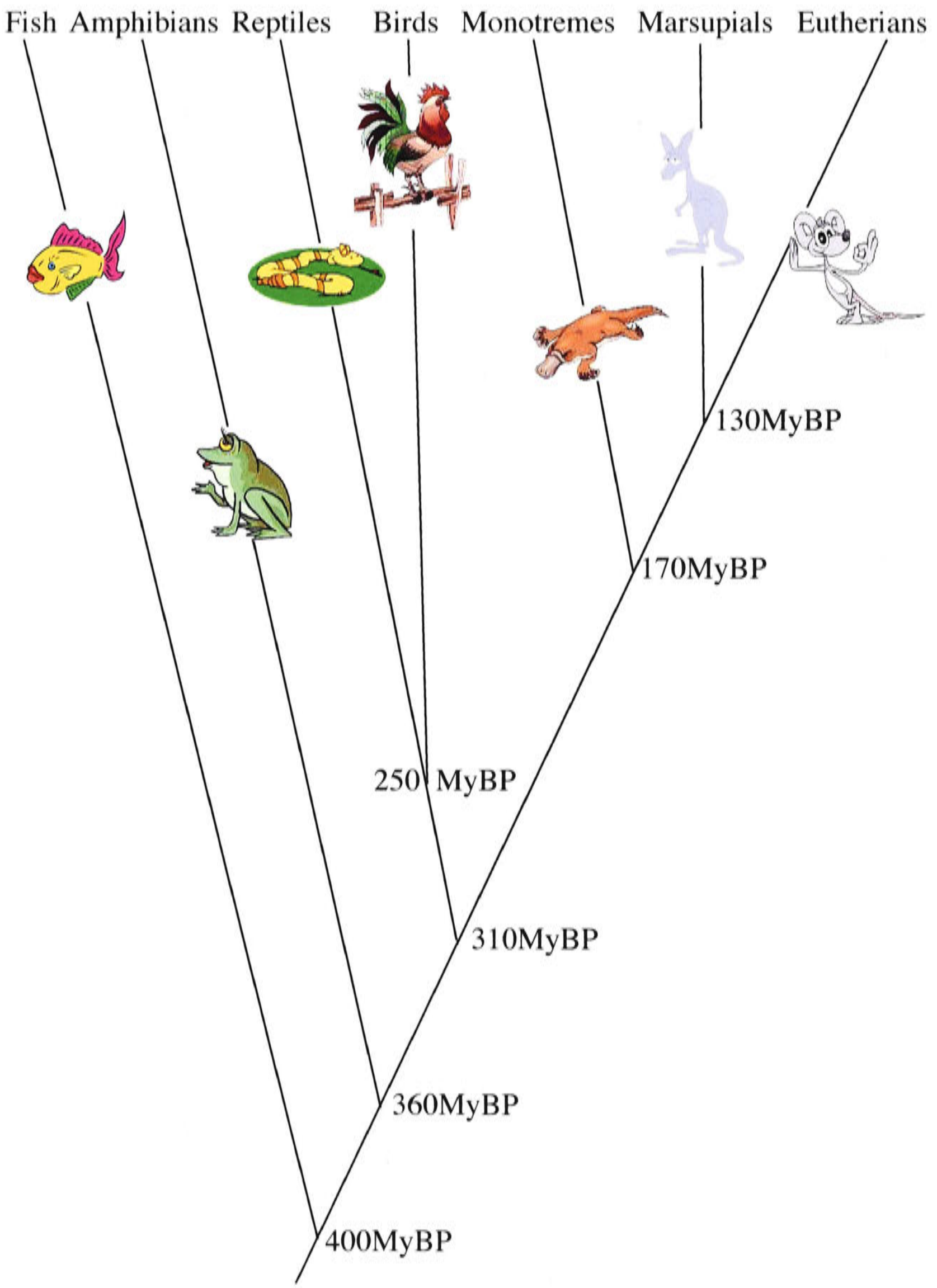


Figure1.1: The evolutionary relationship of vertebrates. The approximate divergence times are indicated in million years before present.

monotreme is the platypus, *Ornithorynchus anatinus*. They retain the ancient reptilian characteristic of laying eggs. Many other anatomical and physiological characteristics also suggest that monotremes form a separate subclass of mammals, and fossil evidence is consistent with this grouping. Thus Monotremes are usually thought to be the mammals most distantly related to humans.

However, mitochondrial (mt) DNA analysis of the echidna and comparison to eutherians and two marsupials (*Macropus robustus* and *Vombatus ursinus)* placed monotremes as a sister group to marsupials, within a mammalian subclass Marsupionta (Janke el al., 1996; Janke et al., 2002). The marsupionta hypothesis is not consistent with the extensive morphological and fossil evidence. Nor is it supported by the limited data on nuclear gene sequences available to date. For instance, study of the large nuclear gene M6P/IGF2R, unambiguously placed eutherians and marsupials as sister groups, supporting the theria hypothesis (Killian. et al., 2001). The different results from nuclear and mitochondrial sequence may be explained by the assumptions of equal substitution rates among sites that were made in this analysis. This assumption can bias test of phylogenetic hypotheses in many situations (reviewed by Killian *et al.*, 2001).

Metatheria (commonly referred to as marsupials) are found in North and South America and Australia. Australia has the most diverse range of marsupials, with 13 families comprising ~200 species, whereas the Americas have only three families consisting of ~70 species. Reproduction in marsupials differs from eutherian reproduction in several major ways. The pregnant marsupial female does develop a placenta, but it is less developed and lacks the endocrine function of the eutherian placenta. The young are underdeveloped when born and finish development attached to a teat, often (but not always) protected in a pouch. Although marsupials are present on two continents, they are considered to be monophyletic on the basis of morphological evidence and the fossil record. The isolation of Australia in the early Cretaceous period is thought to account for the diversification of marsupials in Australia.

The *Macropodidae* (kangaroos and wallabies) and *Dasyuridae* (insectivorous and carnivorous marsupials) lineages diverged from each other about 65MyBP (Kirsch *et al.*, 1997). Of all Australian marsupials *Macropus eugenii*, the tammar wallaby (2n=16), has the best studied genome. *Sminthopsis crassicaudata* (2n=14), a dasyurid, has also been well studied. Both species have large easily distinguishable chromosomes and

would make good marsupial models. For this study our model marsupial is the tammar wallaby.

Eutherians are commonly known as placental mammals because of the well developed placenta that connects the foetus and the female uterus, through which the young gets its nutritional and respiratory needs. Eutherians are diverse and widespread. They radiated from a common stock of insectivorous mammals in the late Cretaceous and early Tertiary periods 70-50MyBP (Kardong, 1995). Of the wide range of eutherian mammals, human and mouse have the best studied sex chromosomes.

Birds are different from other tetrapods because of their feathered bodies and forelimbs that are usually adapted for flying. Their variety outnumbers that of all other vertebrates except fishes (Kardong, 1995). Birds are phylogenetically younger than mammals, having diverged from reptiles 250MyBP. The bird radiation occurred in the middle Cretaceous and early Tertiary periods 80-60MyBP (Kardong, 1995).

Comparisons between the sex chromosomes of these three distantly related mammal groups, and comparisons between mammals and birds, have been used to determine the origin, and regions of conservation, of mammalian sex chromosomes. The same approach has also been important in identifying functionally important regions on the Y chromosome, such as the testis-determining factor.

1.2 Organization of vertebrate sex chromosomes

The mammalian X and Y chromosomes, share regions that are genetically equivalent, and regions that recombine, as does the bird Z and W chromosomes. However, there is a striking difference in gene content of the X and Y. In this section, I discuss the organization of mammalian sex chromosomes, and the differences that distinguish the sex chromosomes of the three different groups of mammals. I also discussed the bird Z and W, and the relationships between mammalian sex chromosomes and the bird genome.

1.2.1 Eutherian sex chromosomes

Eutherian X and Y chromosomes are morphologically very distinct and have quite different gene contents. Usually the Y chromosome is significantly smaller than the X. The human X chromosome is about 165 million (Mega) base pairs (Mb) in length and represents about 5% of the haploid genome. It is euchromatic, though rather gene poor. The human Y chromosome is 60Mb, which represents 2% of the haploid genome (Lander *et al.*, 2001). About half of the human Y is euchromatic and the other half (comprising the distal long arm) is heterochromatic.

1.2.1.1 The Eutherian X chromosome

The human X chromosome contains roughly 1500 genes with a mixture of housekeeping and specialist functions. Most genes on the X chromosome are not specifically concerned with sex. However, there does seem to be an increased concentration of reproduction related genes on the human X (Saifi and Chandra, 1999). This increased concentration of sex related genes on the X was confirmed in mouse (Wang *et al.*, 2001) using subtractive hybridisation to isolate transcripts specific to mouse spermatogonia and absent from somatic tissue. Wang *et al.* found that of the 25 spermatogonia-specific genes (19 novel), 3 were Y-linked and 10 X-linked. If the genes were randomly distributed in the genome, only one or two genes would be expected to be located on the X chromosome. There is also an increased concentration of genes on the human X chromosome involved in cognitive ability (Zechner *et al.*, 2001).

The size and gene content of the X chromosome is remarkably conserved between eutherian species at about 5% of the haploid genome (Ohno, 1967). Even the gene arrangement is largely conserved among distant orders (*e.g.* primates and felids). However, differences in morphology and gene order between the X chromosomes of human and mouse demonstrated that internal rearrangements have occurred in rodent lineages (Blair *et al.*, 2000). Gene mapping in eutherians has demonstrated extremely high conservation of gene content of all eutherian X chromosomes, to the extent that almost every gene that maps to the X of one species will map to the X of all others (Wakefield and Graves, 1996).

Because females have two X chromosomes and males have only one X, equal transcription from both X chromosomes in females, and the single X of males, would

result in doubled activity of X genes in females relative to males. Dosage compensation ensures equal dosage of X-linked genes in males and females. In eutherians, one of the X chromosomes in somatic cells of females is randomly chosen for inactivation. The stable, inactive X is transcriptionally silent with the exception of pseudoautosomal region (PAR) genes and some genes on human Xp that have partners on the Y (Disteche, 1995).

1.2.1.2 Homologies between the eutherian X and Y chromosomes

Most of the eutherian X does not undergo recombination with the Y at male meiosis. However, there is a small pseudoautosomal region (PAR) at the ends of the short arms of the X and the Y (PAR1) that undergo pairing at male meiosis. In humans, PAR1 is 2.6Mb and contains 13 genes (Figure 1.2) and the 5' region of a fourteenth (XGA) (Gianfrancesco *et al.*, 1998; Blaschke and Rappold, 1997). Deletion of PAR1 results in the failure of X-Y pairing at male meiosis and therefore sterility (Mohandas *et al.*, 1992). The mouse X and Y share a 2Mb PAR, deletion of which results in sterility (Burgoyne *et al.*, 1992). The presence of a PAR is crucial for X-Y pairing in all eutherian mammals, although there are some rodents whose sex chromosome lack a PAR and do not pair at male meiosis (Ashley and Fredga, 1994).

The role of gene content in the function of PAR1 can be determined by comparing the PARs of different species. In prosimians (lemurs) the human PAR1 genes *ANT3*, *SHOX* and *IL3RA* are located in the X-Y pairing region along with the human X-specific genes *STS* and *PRKX* (Glaser *et al.*, 1999). The sheep and cow PARs contain *CSF2RA*, *ANT3* and *STS*, and the dog PAR contains *PRK*, *ANT3* and *STS* (Moore *et al.*, 2001; Toder *et al.*, 1997a). The mouse PAR contains only one gene, *Sts* (Salido *et al.*, 1996). A second gene, *Fxy* spans the mouse PAR boundary and is truncated at this point in the non-recombining region of the Y (NRY) (Palmer *et al.*, 1997; Dal Zotto *et al.*, 1998). The human PAR1 genes *PGPL*, *DHRSXY*, *CSF2RA* and *IL3RA* are autosomal in mouse, whereas other PAR genes seem to have no mouse orthologues (Gianfrancesco *et al.*, 2001; Disteche *et al.*, 1992; Milatovich *et al.*, 1993) (Figure 1.2). The variable gene content of PAR1 between mammalian species suggests that the genes borne by the PAR are not crucial to its function.

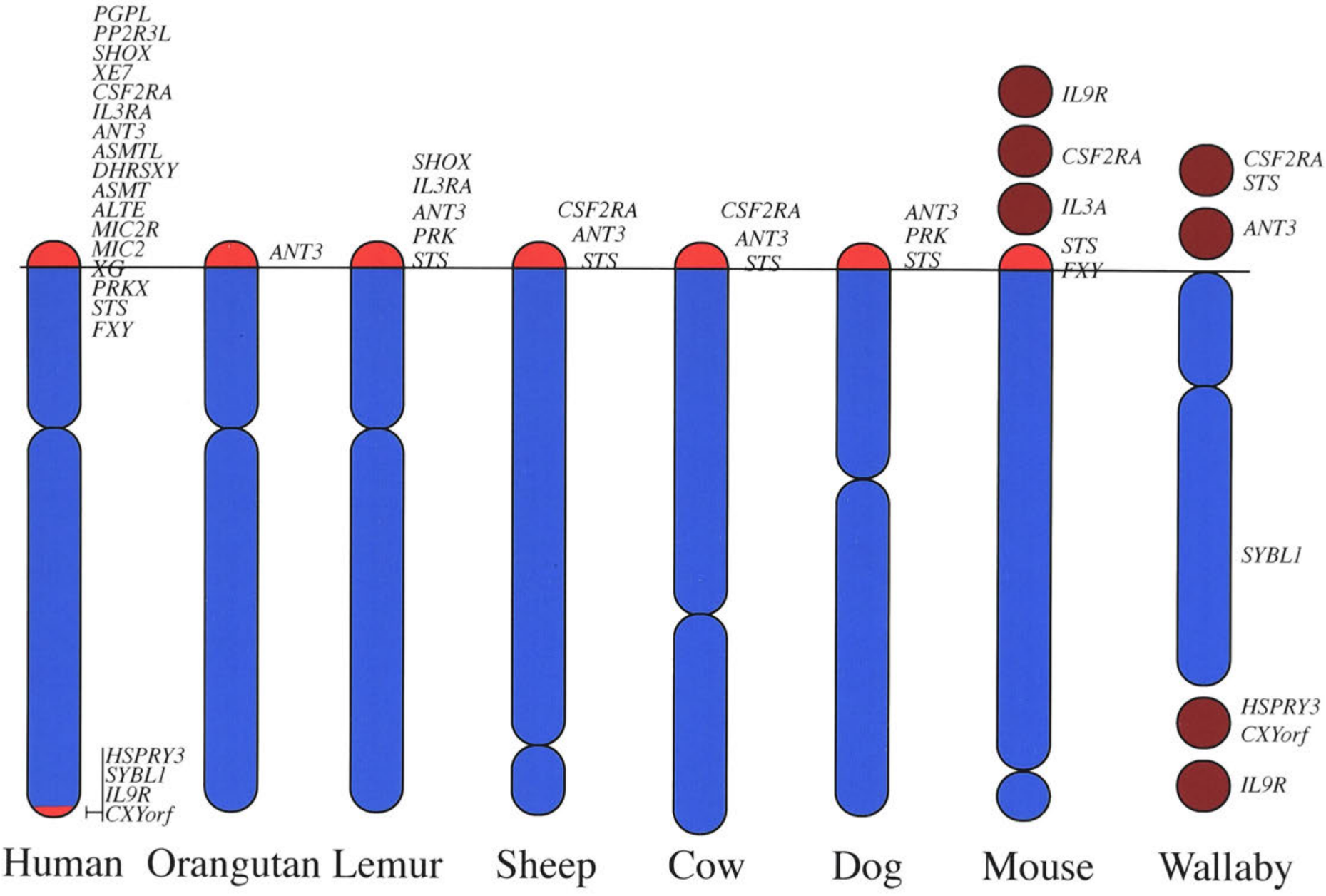


Figure 1.2: The gene content of the PARs of different species. Several genes in human PAR1 are located on autosomes in mouse and tammar wallaby. Genes located in the PAR of some species are located in the X-specific region proximal to the PAR boundary in human.

Human sex chromosomes have a second smaller (0.5Mb) pseudoautosomal region, PAR2, located at the ends of Xq and Yq (Freije *et al.*, 1992) (Figure 1.3). PAR2 contains four known genes *HSPRY3*, *CXYorf1*, *SYBL3* and *ILR9* (D'Esposito *et al.*, 1996; Kermouni *et al.*, 1995; Ciccodicola *et al.*, 2000). PAR2 was the result of a recent illegitimate exchange between the X and Y chromosome. PAR2 also pairs occasionally at male meiosis (much less frequently than PAR1), but the significance of this is unclear (Li and Hamer, 1995). Genes in PAR2 show quite different functional characteristics. *IL9R* and *CXYorf1* show biallelic expression characteristic of PAR1 genes (Vermeesch *et al.*, 1997). On the other hand, *SYBL1* and *HSPRY3* have a unique method of achieving equal male and female dosage by inactivating the Y copy and the inactive X chromosome copy (D'Esposito *et al.*, 1996; Ciccodicola *et al.*, 2000).

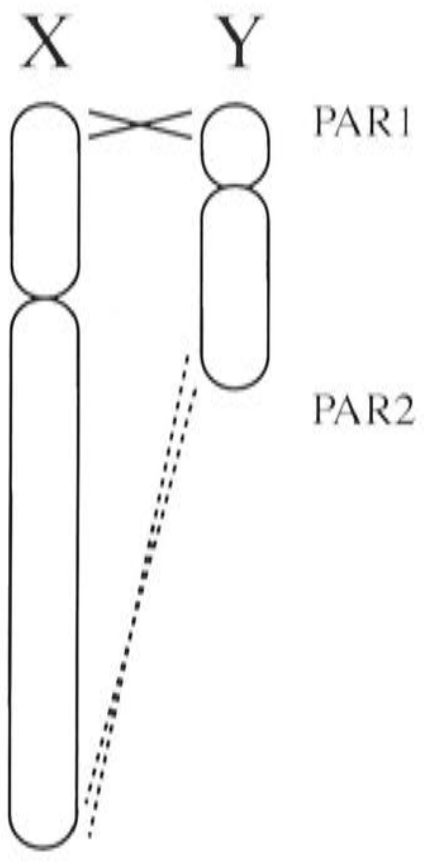
The largest region of homology between the human X and Y is not one of the pseudoautosomal regions, but a region in the non-recombining portion of the sex chromosomes shared between Xq21.3 and Yp11.1, estimated to be 4Mb (Mumm *et al.*, 1997; Sargent *et al.*, 1996; Vacca *et al.*, 1999). This homology was a result of a translocation from the X to the Y (Page *et al.*, 1984) dated to 3-4MyBP. On the Y chromosome its integrity has been disrupted by a LINE-mediated paracentric inversion (Schwartz *et al.*, 1998) that has not been dated, but could be much more recent. The inversion created a small proximal block of homology and a larger distal block.

1.2.1.3 The eutherian Y chromosome

Because the Y chromosome is present in only half the population, it must not contain any genes that are crucial for life, unless these genes are shared with the X chromosome. The phenotypes of XO individuals indicate that shared genes affect at least some physical traits such as stature. X/Y shared genes may also be required for viability, as >95% of XO foetuses rapidly abort. There is some suggestion that surviving XO individuals started off as mosaics.

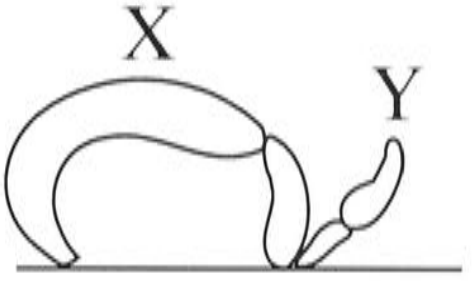
The human Y is roughly half euchromatic and half heterochromatic. The differentiated (non-recombining) region contains a total of 30-35 (Table 1.1) recognisable genes (excluding PAR1 and PAR2) that all reside in the euchromatic region (Lahn and Page, 1997). Because the Y chromosome is confined to males, it is not

Eutherians
(Human)



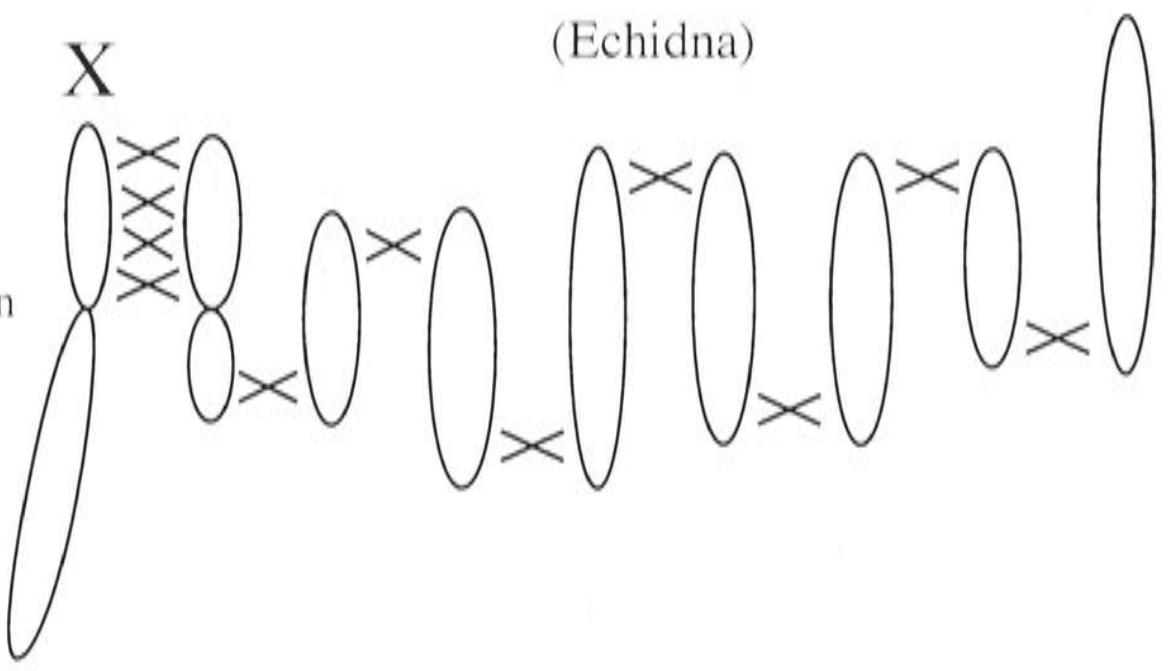
X and Y pair at male meiosis across PAR1 (and occasionally PAR2)

Marsupials
(Tammar wallaby)



X and Y are attached to a basal plate side by side at male meiosis.

Monotremes
(Echidna)



Formation of a translocation chain at male meiosis

Figure 1.3: Pairing of sex chromosomes at male meiosis of human, tammar wallaby and echidna.

surprising to find that it contains genes involved in male specific functions such as spermatogenesis and male determination (Section 1.3).

Table 1.1 Names and functions of genes on the human Y chromosome. Genes are listed in the order they appear an the Y.

Region	Locus	Full Name	Function	Reference
PAR1	*PGPL*	Pseudoautosomal GTP-binding protein-like	GTP-binding	Gianfrancesco *et al.*, 1998
	PPP2R3L	Protein phosphatase 2 regulatory subunit-like	Protein phosphatase	Schiebel *et al.*, 2000
	SHOX	Short staure homeobox	Transcription activator	Ellison *et al.*, 1997
	XE7	Pseudoautosomal gene XE7	?	Ellison *et al.*, 1992
	CSF2RA	Colony-stimulating factor 2 receptor	Receptor signalling protein	Raines *et al.*, 1991
	IL3RA	Interleukin 3 receptor	Receptor signalling protein	Milatovich *et al.*, 1993
	ANT3	Adenine nucleotide translocator 3	Adenine transporter	Slim *et al.*, 1993
	ASMTL	Acetylserotonin methytransferase-like	Methyltransferase	Ried *et al.*, 1998
	DHRSXY	Dehydrogenase/reductase of SDR family on the X and Y	Oxidoreductase	Gianfrancesco *et al.*, 2001
	ASMT	Acetylserotonin methytransferase	Methyltransferase	Rodriguez *et al.*, 1994
	ALTE	Ac-like transposable element	Transposease	Esposito *et al.*, 1994
	MIC2R	MIC2 related	?	Smith and goodfellow, 1993
	MIC2	Cell surface antigen MIC2	Cell adhesion	Banting *et al.*, 1989
	XG	XG blood group system	Pseudogene	Weller *et al.*, 1995
NRY	*SRY*	Sex-determining region Y	Transcription factor	Sinclair *et al.*, 1990
	RPS4Y	Ribosomal protein S4	Protein of small ribosomal subunit	Fisher *et al.*, 1990
	ZFY	Zinc fingure protein Y	Zinc fingure transcription factor	Page *et al.*, 1987
	PCDHY	Protocadherin Y	Cell adhesion	Blanco *et al.*, 2000
	TSPY	Testis-specific protein Y	Cyclin B binding protein?	Arnemann *et al.*, 1987
	GYGP	Glycogenin 2 pseudogene	Pseudogene	Zhai et al., 2000
	CASKP	Calsium/calmodulin-dependent serine protine kinase pseudogene	Pseudogene	Dimitratos el al., 1998
	PRKY	Protein kinase Y	Ser/Thr protein kinase?	Schiebel *et al.*, 1997
	AMELY	Amelogenin Y	Regulation of tooth enamel formation	Lau *et al.*, 1989
	PRY	PTP-BL related Y	Protein tyrosine phosphatase	Lahn and Page, 1997
	USP9Y	Ubiquitin-specific protease 9 Y	Ubiquitin-specific protease	Brown *et al.*, 1998
	DBY	Dead box Y	RNA helicase	Lahn and Page, 1997
	UTY	Ubiquitously expressed tetratricopeptide repeat Y	Protein-protein interaction?	Lahn and Page, 1997
	TB4Y	Thymosin beta 4 Y	Actin sequestration?	Lahn and Page, 1997
	ARSDP	Arylsulfatase D pseudogene	Pseudogene	Meroni et al., 1996
	ARSEP	Arylsulfatase E pseudogene	Pseudogene	Meroni et al., 1996
	KAL1P	Kallmann syndrome pseudogene	Pseudogene	Franco et al., 1991
	STSP	Streriod sulphate pseudogene	Pseudogene	Yen et al., 1988
	VCY1	Variable charge protein Y 1	Variable charge protien 1	Lahn and Page, 1997
	CDY	Chromodomain Y	Chromodomain protein	Lahn and Page, 1997
	XKRY	XK related Y	Membrane transport?	Lahn and Page, 1997
	SMCY	Selected cDNA mouse homologue Y	Transcription factor?	Wang *et al.*, 1995
	EIF1AY	Eukaryotic translation initiation factor 1 A Y	Translation initiation factor	Lahn and Page, 1997
	RBMY	RNA-binding motif protein Y	RNA-binding protein	Ma *et al.*, 1993
	VCY2	Variable charge protein Y 2	Variable charge protien 2	Lahn and Page, 1997
	DAZ	Deleted in azoospermia	RNA-binding protein	Reijo *et al.*, 1995
PAR2	*HSPRY3*	Homologue of *Drosophila* sprouty 3	Cell-cell signal transduction	Coccodicola *et al.*, 2000
	SYBL1	Synaptobrevin-like 1	Synaptic cell membrane transport	D'Esposito *et al.*, 1996
	IL9R	Interleukin 9 receptor	Signal transduction	Kermouni *et al.*, 1995
	CXYorf1	*CXYorf1*	?	Coccodicola *et al.*, 2000

In contrast to the conserved X chromosome, the human Y has suffered multiple internal rearrangements, deduced by comparative mapping of genes in closely related hominid species (Glaser *et al.*, 1997; Glaser *et al.*, 1998a). A simple model involving a single or even a few evolutionary events cannot explain the rearrangements seen between the Y of hominids, let alone mouse. The Y is therefore highly rearranged in all lineages, so it is unlikely that there are many large regions on the Y in which gene content and order are conserved between species. The only report of conserved order between a group of genes on the Y chromosome of mammalian species is a region that contains *USP9Y*, *DBY* and *UTY*, which is conserved on the Y chromosome of cat, mouse and human (Mazeyrat *et al.*, 1998; Murphy *et al.*, 1999).

Table 1.2: Genes in the differentiated region of the Y chromosome of different species. Genes are listed in the order that they appear on the human Y chromosome, from the border of PAR1 to the heterochromatic region. Genes known not to be on the Y chromosome are indicated (—), whereas genes with an unknown location are blank. References: Vogt *et al.*, 1997; Toder *et al.*, 1997a; Lahn and Page, 1997; Mazeyrat *et al.*, 1998; Murphy *et al.*, 1999; Lahn and Page, 1999.

Gene	Human	Orang utan	Mouse	Cat	Sheep
SRY	Y	Y	Y	Y	Y
RPS4Y	Y	Y	—		
PCDHY	Y	Y	—	—	—
ZFY	Y	Y	Y	Y	Y
KALP	Y	Y	—		
TSPY	Y	Y	Y		Y
AMELY	Y	Y	—	Y	
PRY	Y		—		
USP9Y	Y		Y	Y	
DBY	Y		Y	Y	
UTY	Y		Y	Y	
TB4Y	Y		—		
STSP	Y	Y	PAR	PAR	PAR
VCY	Y				
CDY	Y	Y	—	—	—
XKRY	Y		-		
SMCY	Y		Y	Y	Y
EIF1AY	Y		—		
RBMY	Y	Y	Y		
DAZ	Y	Y	—	—	—
BPY2	Y				
UBE1Y	—	—	Y	Y	Y
EIF2S3Y	—	—	Y	Y	Y

Unlike the gene content of the X chromosome, the gene content of the Y chromosome is poorly conserved between species (Table 1.2). There are several genes on the human Y that do not lie on the Y chromosome in other species. For instance, no *PRY*, *XKRY* and *BPY2* homologous sequences were detected on the mouse Y (M. Delbridge, personal communication). Conversely, there are human Y-borne genes (*RPS4Y*, *PCDHY*, *CDY* and *DAZ*) that have copies on the Y chromosome of primates but not more distantly related eutherians. Alternatively, there are two genes, *Ube1y* and *Eif2s3*, lacking from the human Y chromosome that are located on the Y chromosome of other eutherian species.

Genes on the Y chromosome have been claimed to belong to two different classes, Class I and Class II (Lahn and Page, 1997). Class I genes are single copy, ubiquitously expressed, and have homologues on the X that escape X-inactivation. Class II genes are Y specific and multicopy, and have testis-specific expression. However, this model does not account for genes such as *SRY* (Foster *et al.*, 1992), *RBMY* (Delbridge *et al.*, 1999) and *TSPY* (M. Delbridge, personal communication), which have functional characteristics of Class II genes (multiple Y-borne copies and testis-specific expression) but an obvious X-borne partner from which they evolved. The definition breaks down further when genes are considered in other species. For instance, *ZFY* and *USP9Y* are both Class I genes in human but have testis-specific expression in mouse.

Many genes on the human Y chromosome have been shown to have a partner on the X. Of these human Class I genes, many reside on the X and Y chromosomes of other eutherian species, either in the PAR (*STS* and *PRKY*) or the differentiated region (*SRY*, *ZFY* and *RBMY*). However, few human Class II genes (*e.g. PRY*, *XKRY* and *BPY2*) are located on the Y chromosome of other species. Thus the gene content of the Y chromosome in different species is quite variable, and some species have no Y chromosome at all.

1.2.1.4 Lack of a Y chromosome in *Ellobius*

Probably the most important function of the Y chromosome is male determination. In therian mammals, the *SRY* gene is responsible for determination of a testis (Section 1.3.1). The Y chromosome also contains factors involved in sperm production (Section

1.3.2). Despite these important functions, there are some species that have lost their Y chromosomes, and potentially all the genes on it.

There are three species of closely related mole voles, two of which lack a Y chromosome. One species, *Ellobius fuscocapillus*, has the standard mammalian XY male and XX female system (Fredga, 1988) (Table 1.3), and males possess Y-borne *Sry* and *Zfy* genes (Just *et al.*, 1995). *E. lutescens* have 2n=17, and males and females are both XO. The X chromosome remains unpaired at male meiosis. *E. tancrei* has a karyotype of 2n=32-54, in which both males and females are XX. In both sexes, one of the X chromosomes is heterochromatic and presumed to be inactive. The two species that lack Y chromosomes have been demonstrated to lack male-specific *Sry* and *Zfy* genes (Just *et al.*, 1995).

Table 1.3: Species with unusual sex chromosome systems

Species	Female	Male
E. lutescens	XO	XO
E. tancrei	X(X)	X(X)
E. fuscocapillus	XX	XY
Myopus schisticolor	X*X	XY
	X*Y	
	XX	
Akodon azarae	XX	XY
	XY*	

There are other species of rodents that also have strange sex chromosome systems. Wood lemmings have females with three different sex chromosome constitutions: XX, X*X and X*Y. Males retain the standard mammalian XY. The observation that X*Y individuals develop as females suggests that the X* chromosome contains a locus that suppresses the effect of a dominant male determining factor located on the Y chromosome.

South American Akodont rodents have females that are XX and XY* whereas males retain the standard mammalian XY. Y* is evidently a Y chromosome with a defective male determining locus, such that XY* individuals develop as females. Little is know of the gene content of the Y chromosomes in these species.

In contrast to the poorly conserved Y chromosome, the X chromosome is well conserved and shares a large region with the marsupial X chromosome.

1.2.2 Marsupial sex chromosomes

Like eutherians, marsupials have an XX female: XY male sex determining system, or a simple variant of this. The tammar wallaby has a diploid chromosome number of 16 with large easily distinguishable chromosomes. Marsupial sex chromosomes are generally smaller than eutherian sex chromosomes. In many groups, the basic X chromosome represents about 3% of the haploid genome and the Y chromosome is tiny.

1.2.2.1 The marsupial X chromosome

There is practically no size variation of the X amongst the 60 species of the family Dasyuridae, whereas in Macropoddae there is a sixfold difference. Size variation of the X chromosome is generally due to an increase in heterochromatic material, which was derived from addition to the nucleolar organiser region (Hayman and Martin, 1974). The orthologues of many human X-borne genes have been mapped in tammar wallaby. Seventeen, spanning the entire long arm and proximal short arm of the human X, have been mapped to the tammar X (Figure 1.4). The location of human X genes on the marsupial X chromosome demonstrated that a large region of the X from the two groups is homologous. This was also confirmed by chromosome painting. DNA from a flow-sorted tammar wallaby X chromosome was painted onto human chromosomes. The wallaby X DNA hybridised to the long arm and proximal short arm of the human X (Glas, *et al.*, 1999).

However, many human Xp genes map to autosomal locations in marsupials. For instance, *STS*, *AMELX*, *ZFX*, *DMD*, *CYBB*, *USP9X*, *DBX*, *MAOA* and *UTX* are all located on chromosome 5p in the tammar wallaby (Sinclair *et al.*, 1988; Waters *et al.*, 2001; Toder and Graves, 1998; Watson *et al.*, 1992b; Wakefield and Graves, 1996), whereas *SYN1*, *POLA* and *OTC* map to wallaby chromosome 1p (Spencer *et al.*, 1991). The autosomal location of eutherian X-borne genes in marsupials demonstrated that not all genes are shared by both the eutherian and marsupial X chromosomes.

As for eutherians, one of the marsupial X chromosomes in females is subject to X-inactivation. However, unlike eutherians, the inactive X is not randomly chosen. Rather, the inactive X is of paternal origin. Marsupial X-inactivation is also less stable than eutherian X-inactivation.

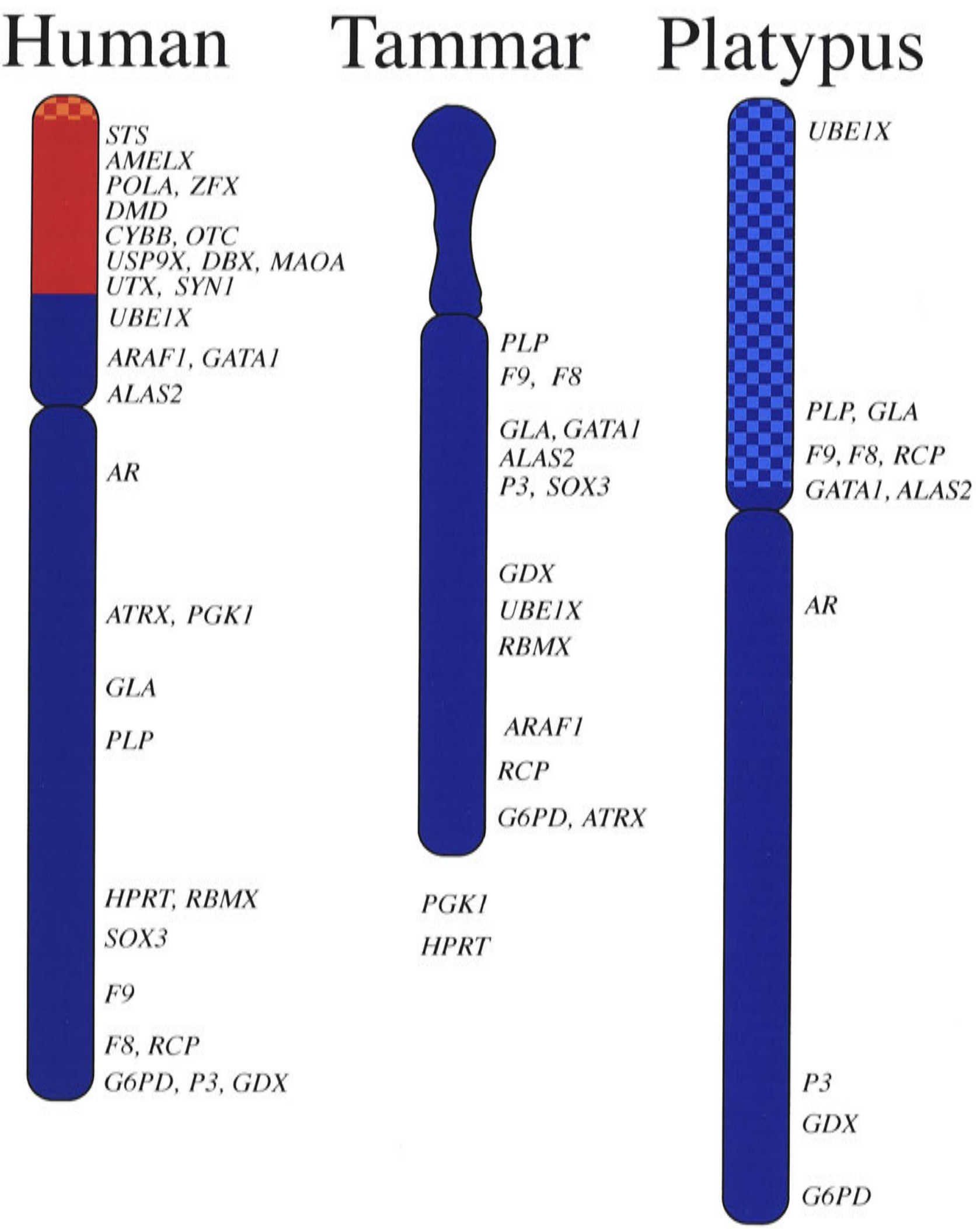


Figure 1.4: Genes shared (blue) by the X chromosome of the three mammalian groups. Checks represent pairing regions. Red on the human the X chromosome represents regions that are autosomal in tammar wallaby and platypus. References: Watson *et al.*, 1990; Mitchell et al., 1992; Foster *et al.*, 1994; Wakefield and Graves, 1996; Wilcox *et al.*, 1996; Toder and Graves, 1998; Pask *et al.*, 2000; Waters *et al.*, 2000.

1.2.2.2 Pairing of marsupial sex chromosomes

The marsupial sex chromosomes have no PAR, so do not undergo recombination at male meiosis. Istead, they show end-to-end alingment and are kept together by attachment to a basal plate structure (Page *et al.*, 2002). Painting *Sminthopsis* chromosomes with a microdissected Y showed no signal on the X, demonstrating that the *Sminthopsis* X and Y chromosomes shared no homology (Toder *et al.*, 2000). However, in tammar wallaby chromosome painting has demonstrated that there are sequences shared between the long arm of the Y chromosome and the heterochromatic, nucleolus-bearing, short arm of the X chromosome (Toder *et al.*, 1997b). No chiasmata were observed between the X and Y of marsupials at meiosis (Sharp, 1982), hence it was concluded that the repetitive sequences shared by the wallaby X and Y probably do not act as a PAR. The association of marsupial sex chromosomes at meiosis suggests that there is a different mechanism governing proper segregation of the X and Y (Page *et al.*, 2002) (Figure 1.3).

1.2.2.3 The marsupial Y chromosome

The *Sminthopsis* Y chromosome is tiny (~10Mb) and share no homology with the *Sminthopsis* X. It has few repetitive sequences and could represent the minimal mammalian Y chromosome (Toder *et al.*, 2000) (Figure 1.5). The best studied marsupial Y chromosome is that of the tammar wallaby. Five genes have been mapped to the tammar wallaby Y chromosome: *SRY*, *RBMY*, *SMCY*, *UBE1Y* and *ATRY* (Foster *et al.*, 1992; Delbridge *et al.*, 1997; Mitchell *et al.*, 1992; Pask *et al.*, 2000). A report of a male specific *RPS4Y* band on a Southern blot of the opossum suggests that there may be a Y-linked copy in marsupials (Jegalian and Page, 1998). All of these genes are X-Y shared in humans with the exception of *ATRY*. *ATRX* is X-borne in human and mouse, but *ATRY* has not been detected in any eutherians It is therefore the only known marsupial-specific Y-borne gene.

1.2.3 Monotreme sex chromosomes

It was originally thought that all monotremes had a different chromosome number in male and female (Bick and Jackson, 1967), suggesting an XX/XO sex determination

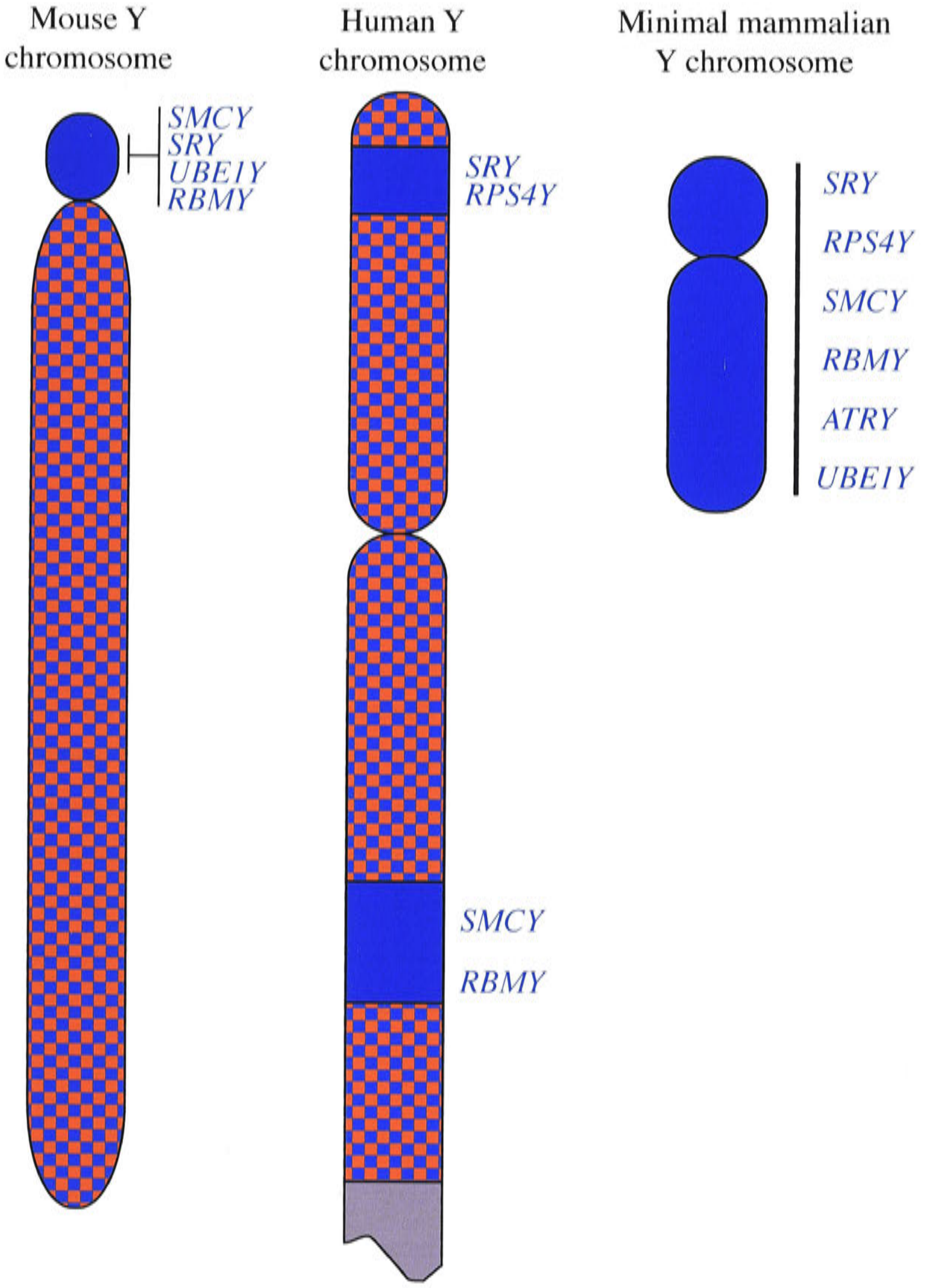


Figure 1.5: Genes shared by the human and/or mouse Y chromosome and the small Y chromosome of *Sminthopsis* (blue), as well as the marsupial-specific *ATRY*. The tiny *Sminthopsis* Y (~10Mb) repersents the minimal mammalian Y chromosome.

system. Both echidna species were confirmed to have 2n=63 in males and 2n=64 in females, but it was discovered that platypus had 2n=52 (Murtagh, 1977). There are six large cytogenetically distinguishable autosomes (1-6) and the X. There are then several groups of cytogenetically indistinguishable chromosomes: chromosomes 7-11, chromosomes 12-17, chromosomes 18-21 and chromosomes 22 and 23. There are four unpaired elements.

1.2.3.1 The monotreme X chromosome

Monotremes have a larger X chromosome than that of therian mammals, which represents 6% of the haploid genome (Watson *et al.*, 1990). Xp pairs completely at meiosis with what has been thought to be the Yq (Watson *et al.*, 1992a). Echidna differs from platypus with females $X_1X_1X_2X_2$ and males X_1X_2Y. X_1 is the original X chromosome, whereas X_2 is derived from an autosome. The Y is a compound chromosome consisting of the putative Y attached to the homologue of X_2 (Wrigley and Graves, 1988a). The platypus X chromosome is cytogenetically and genetically equivalent to the echidna X_1 chromosome (Watson *et al.*, 1992a).

The limited mapping data available for the monotreme X chromosome (achieved by radioactive *in situ* hybridisation) demonstrated that many genes shared by the eutherian and marsupial X chromosomes also mapped to the platypus X chromosome (Watson *et al.*, 1990; Watson *et al.*, 1992a; Wilcox *et al.*, 1996) (Figure 1.4). The echidna X_1 chromosome also bears at least four of the same genes, *F9*, *AR*, *G6PD* and *GDX*, which demonstrated conservation of the monotreme X chromosome (Watson *et al.*, 1992a). This information showed that there was a large region of material that is shared by the X in all three mammalian groups.

As with marsupials, genes located on human Xp were shown to be located on platypus autosomes. *DMD*, *CYBB* and *TIMP1* were located on platypus chromosome 1. *POLA*, *OTC*, *SYN1* and *ARAF1* were located on platypus chromosome 2. *AMELX*, *ZFX* and *MAOA* all had ambiguous locations in platypus. Signal peaks were observed over both chromosomes 1 and 2 for each of these genes (Watson *et al.*, 1991;Wakefield and Graves, 1996).

The monotreme X chromosome therefore shares regions of homology with the eutherian X chromosome. However, as with marsupials, genes on the eutherian X chromosome are located on platypus autosomes.

1.2.3.2 Pairing of monotreme sex chromosomes

Monotremes are unique amongst vertebrates because several unpaired elements in their genome form a multivalent chain at metaphase I of male meiosis (Bick *et al.*, 1973). The meiotic chain consists of eight chromosomes in platypus and nine in echidna (Watson *et al.*, 1992a). The entire short arm of the monotreme X chromosome pairs with the long arm of the first element in the translocation chain at male meiosis (Figure 1.3). This element is male specific (F. Grutzner, personal communication), but since it pairs with another element, it is not clear whether it contains any male-specific sequences. The multivalent chain therefore confuses the identity of the Y chromosome in both platypus and echidna.

Male specific monotreme sequences are yet to be isolated, and no monotreme Y-borne genes have been cloned and mapped. This (along with the multivalent translocation chain at male meiosis) has made the identity of the monotreme Y unclear.

The size and gene content of the X chromosome is remarkably conserved between species. The X chromosomes from all three mammalian groups share a large region of homology. However, part of the eutherian X chromosome is represented by marsupial and monotreme autosomes. In contrast to the X chromosome, the Y displays poor homology between species. Many genes from the human Y chromosome have not been detected on the Y chromosome of other eutherian mammals, and the eutherian and marsupial Y chromosomes share only five known genes.

Because females possess two X chromosomes and males possess only one, equal expression of genes from the two X chromosomes in females, and the one X chromosome in males, would result in gene dosage differences between sexes. A dosage compensation mechanism must be put in place to equalise this dosage difference between males and females.

1.2.3 X inactivation

X inactivation refers to the silencing of one of the X chromosomes in the somatic cells of females so that there is equal dosage of X genes in XX and XY individuals (Lyon, 1961). Inactivation of one X chromosome occurs early in embryogenesis; it is stable and somatically heritable. There are two types of X-inactivation, random and imprinted. The latter is thought to be ancestral because paternal X-inactivation occurs in marsupials as well as in the extra embryonic membranes of rodents (Cooper *et al.*, 1971; Sharman, 1971). Random X-inactivation occurs in the embryo of eutherian mammals. There is thought to be several steps leading to X-inactivation: counting, choice, initiation, establishment and maintenance (Figure 1.6). Each of these steps is genetically separable.

In the mid 1990's transgenic analysis in embryonic stem (ES) cells substantiated the decades old idea of an X-inactivation centre (XIC) that controls initiation of X-inactivation (Figure 1.7) (Lee *et al.*, 1999b). Excluding maintenance, each step of X-inactivation seems to be controlled by the XIC. The size of the XIC is not clear. It may be anywhere from 35Kb (Herzing *et al.*, 1997) to 80Kb (Lee *et al.*, 1999b) or even much bigger. This depends on whether the XIC is considered to be involved only in silencing, or includes all the contiguous X-linked sequences that effect the X-inactivation phenotype.

It is agreed that the XIC contains at least *XIST*, a locus that transcribes a non-coding RNA that is expressed from (Brown *et al.*, 1991) and accumulates around the inactive X (Clemson *et al.*, 1998). *XIST* has an antisense transcript (*TSIX*) (Lee *et al.*, 1999a) that is expressed from the active X. It is thought that *TSIX* represses *XIST* expression and regulates X chromosome choice (Debrand *et al.*, 1999; Lee and Lu, 1999). There is also a region 3' to *XIST* (XCE), alleles of which influence the probability of being chosen as the active X (reviewed by Boumil and Lee, 2001). Maintenance of the imprint is achieved by methylation of CpG islands in DNA (Mohandas *et al.*, 1981), hypo-acetylation of histone H4 and macro-H2A association (Pfeifer *et al.*, 1990; Jeppesen and Turner; 1993; Costanzi and Pehrson, 1998).

Eutherian X-inactivation is a chromosome wide phenomenon, but there are several regions of the human X that are exempt. Escapers of X-inactivation include genes in the PARs such as *CSF2RA*, *IL3RA*, *ANT3* and *ILR9*, which are present in two copies in both

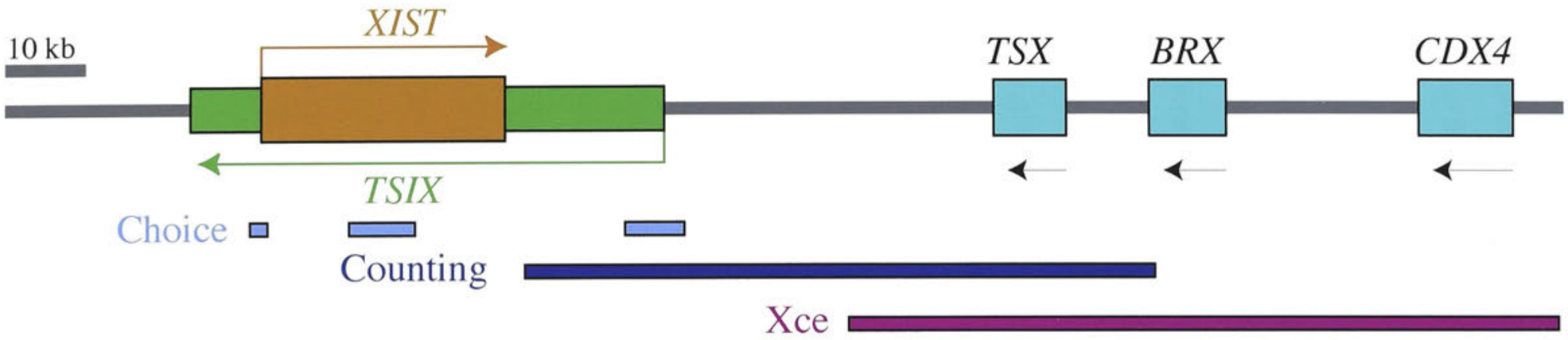


Figure 1.6: The X-inactivation center.

Figure 1.7: The steps involved in X-inactivation

male and female and therefore require no dosage compensation. Genes in the differential region of the X chromosome that have a functional partner on the Y chromosome (*ZFX*, *DBX*, *USP9X* and *RPS4X*) (Page *et al.*, 1990; Lahn and Page, 1997; Fisher *et al.*, 1990) also escape X-inactivation. The presence of two functional copies in males results in equal dosage with females so there is no need for compensation of that particular locus. In mouse, however, many of these loci on the Y chromosome have acquired male-specific functions and longer complement the X-borne partners, so the X allele is inactivated.

There are several factors that appear to be involved in the maintenance of eutherian X-inactivation. The inactive X chromosome displays methylation of CpG islands 5' of the inactive locus (Pfeifer *et al.*, 1990). However, DNA methylation is not critical to X-inactivation. Methylation of CpG islands 5' of the inactive *Hprt* locus, in rodent embryos, is implemented after X-inactivation occurs (Lock *et al.*, 1987). Histone modification may be more critical. Histone-Acetylation enhances the ability of transcription factors to bind to DNA and initiate transcription (reviewed by Wolffe, 1998). It is therefore not surprising to find that the inactive X chromosome is hypoacetylated (Jeppesen and Turner, 1993). In addition, unusual histones such as macro-H2A are associated with the inactive X chromosome (Costanzi and Pehrson, 1998). Macro-H2A localises to the inactive X chromosomes after initiation of X-inactivation, and is therefore not essential for the initiation of X-inactivation (Mermoud *et al.*, 1999). It probably pays a role in condensation of the inactive X. A byproduct of this condensation may be that the inactive X chromosome is late replicating (Willard and Latt, 1976).

The control of imprinted X-inactivation is not as well understood. It is unclear whether marsupials have an XIC, so the existence and roles of *XIST* and *TSIX* in imprinted inactivation are unclear. The nature of the paternal imprint that controls non-random X-inactivation in marsupials is unknown, as for other imprinted regions. This imprint would need to be put into place during male meiosis, and removed from the paternal X chromosome at female meiosis.

The mechanism of imprinted X-inactivation is also unknown. The inactive marsupial X shows hypo-acetylation, but no methylation of CpG islands (Wakefield *et al.*, 1997; Cooper *et al.*, 1993). Thus marsupials appear to share only part of the

eutherian maintenance pathway. The absence of DNA methylation, together with the lower stability of X-inactivation in marsupials (perhaps due to the lack of methylation) is consistent with the idea that marsupial X-inactivation is a less complex system that is ancestral.

Whether X-inactivation occurs in monotremes is yet to be determined. Asynchronous replication (a characteristic of X-inactivation) of the two X chromosomes in females suggested that X-inactivation might occur in monotremes (Wrigley and Graves; 1988b, Murtagh, 1977). However, it was the short arm of the X chromosome that showed asynchronous replication. The short arm of the monotreme X pairs with the second element of the meiotic chain at male meiosis, and therefore should not need dosage compensation. A recent study has indicated that *G6PD* may be subject to X-inactivation. *G6PD*, which has been mapped to the platypus X chromosome by radioactive in situ hybridisation (Watson *et al.*, 1990), showed equal expression in cells from males and females (J. Deakin, personal communication).

Mammals are not the only organisms to have evolved dosage compensation. Dosage compensation systems have been well studied in *Drosophila* and *C. elegans*. The same outcome as mammalian X-inactivation is achieved in two completely different ways. In *Drosophila*, genes on the single X chromosome in males are upregulated. Expression is controlled by five genes: maleless (*mle*), three male-specific lethal (*msl*) genes (*mls-1*, *mls-2* and *mls-3*) and males absent on the first (*mof*). Two non-coding RNA molecules transcribed from *roX-1* and *roX-2* are also involved. The five proteins and two RNA molecules come together to form a complex called a compensasome that mediates upregulation by altering chromatin structure. All but one of the *roX* RNAs are required for the compensasome to operate correctly. Because all gene products are required, the compensasome can be made sex specific by the regulation of just one of its elements. This is exactly what happens; *mls-2* is repressed by the protein of the female-specific sex-determining gene, sex-lethal (*Sxl*) (reviewed by Marin *et al.*, 2000).

In *C. elegans* (XO males and XX hermaphrodites) expression of both the hermaphrodite X chromosomes are down-regulated so that each expresses effectively half. This mechanism for dosage compensation is controlled by another set of unrelated genes (reviewed by Marin *et al.*, 2000). All these organisms have independently

evolved quite different methods for equalising the dosage between individuals with one X, and individuals with two. Equal dosage is clearly so important that it has evolved in several ways.

It is striking to note the parallels of the independently evolved systems of dosage compensation in eutherians and *Drosophila*. Although in *Drosophila* the male X-chromosome is upregulated, and in eutherians one of the female X chromosomes is turned off, they both require genes transcribed into non-coding RNAs as part of the dosage compensation mechanism. In both cases histone acetylation is crucial to maintenance of dosage compensation.

1.2.4 Bird sex chromosomes

The mammalian linage diverged from the reptile/bird lineage about 310MyBP. Thus birds and reptiles are the closest relatives to mammals.

The bird genome consists of large easily distinguishable macrochromosomes, and many small indistinguishable microchromosomes. Of the many bird species, chicken (*Gallus gallus*) has the best-studied genome, and is the most appropriate model species for use as an outgroup to mammals.

Sex determination in birds is accomplished by differentiated sex chromosomes. However, females are the heterogametic sex being ZW, whereas males are ZZ. Chicken has a chromosome number of 2n=78, with cytologically distinct chromosomes, including eight autosome pairs (1-8) and the Z and W. There are also 30 pairs of cytologically indistinguishable microchromosomes (Schmid *et al.*, 2000).

The bird Z chromosome represents 7-10% of the haploid genome (Schmid *et al.*, 2000). The W chromosome is usually small. It only represents about 1.5% of the haploid genome in chicken (Clinton and Haines, 1999). It pairs only over a small PAR at the tip of the short arm containing a single recombination nodule. The W chromosome consists largely (~65%) of repeated sequences, and has a CG-rich long arm. It contains only four known genes, all of which have homologues on the Z chromosome (Schmid *et al.*, 2000; Fridolfsson *et al.*, 1998).

Like the mammal X chromosome, the bird Z chromosome is uniform in size between species. Chromosome painting confirmed complete homology of the Z between chicken and emu, a ratite bird that diverged from chicken 80MyBP (Shetty *et*

al., 1999). However, the W is quite variable in size, ranging from very small in chicken, to sizes similar to the Z in emu and ostrich (Ogawa *et al.*, 1998; Fridolfsson *et al.*, 1998). Cross species chromosome painting in emu shows that the chicken Z is homologous not only to the emu Z, but also to the W, with the exception of a small region around the centromere, which may be responsible for sex determination (Shetty *et al.*, 1999).

Since males have two, and females one Z chromosome, it would be expected that Z gene activity would be equalized in male and females as a result of some form of dosage compensation mechanism. However, early enzyme dosage studies were interpreted as the absence of dosage compensation on the bird Z (Baverstock *et al.*, 1982). In male birds, the two Z chromosomes replicate synchronously. However, real time PCR has recently showed that transcription levels in male and female birds are equivalent for six out of seven Z-borne genes (McQueen *et al.*, 2001). This evidently does not result from mechanisms similar to that of mammalian X-inactivation as RNA fluorescence *in situ* hybridisation (FISH) shows that loci on both Z chromosomes in males were transcribed (Kuroda *et al.*, 2001).

The mechanism of sex determination is not as clear in birds as it is in mammals. It is not even known wether the W is female determining, or whether Z dosage is male determining, or both. Despite extensive searches, no diploids with aberrant sex chromosomes (equivalent to XO and XXY mammals) have been discovered. However, triploid chickens with aberrant Z:W ratios are viable. ZZZ are phenotypically normal males that produce abnormal spermatozoa. ZWW chickens die early during embryonic development. ZZW individuals have a confused sexual identity; they are initially females, but after reaching sexual maturity, begin to crow and adopt a male phenotype (Thorne and Sheldon, 1993). These individuals suggest that the W chromosome has a feminising effect (as the Y chromosome has a male determining effect in mammals), but Z dosage has a masculinizing effect. This suggests that the Z contains one or more dosage-sensitive genes that influence sex. Identification of *DMRT1* on the chicken Z (Nanda *et al.*, 2000), and mapping *DMRT1* to the Z but not the W in emu (Shetty *et al.*, in press), makes it a good candidate for the bird sex-determining gene. This gene is involved in testis differentiation of all vertebrates, and haploinsufficiency in humans could cause male to female sex reversal.

Microchromosomes (chromosomes 9-38 and the W) represent 23% of the haploid chicken genome (Smith and Burt, 1998). The microchromosomes are GC-rich (Auer *et al.*, 1987) and have a higher CpG content than the macrochromosome (McQueen *et al.*, 1996). Because 60-70% of known chicken genes are associated with CpG-islands, it has been suggested that the microchromosomes could be a gene-dense region of the genome. Random sequencing of cosmid clones that mapped to either a microchromosome or macrochromosome by FISH, demonstrated that the microchromosomes contain 1.3 times (Smith *et al.*, 2000) to 2.4 times (Clark *et al.*, 1999) the gene content of the macrochromosomes.

Gene mapping in chicken has revealed many groups of genes conserved with human (Schmid *et al.*, 2000). Of particular interest is the comparative mapping that revealed the relationship of mammal and bird sex chromosomes. Gene mapping failed to find any homology between the bird Z and mammalian X, and limited gene mapping reveals no homology between the bird W and mammalian Y. Genes from the human X chromosome cluster at two locations in chicken, on chromosome 4p and 1p. Human X genes that are located on chicken chromosome 4p include *HPRT1*, *BTK*, *PGX1*, *CUL4B*, *UBE1A* and *FMR1*. Human X genes that are located on chicken 1p include *OTC*, *NROB1*, *ZFX* and *SCML2* (Schmid *et al.*, 2000).

Conversely, most genes from the chicken Z chromosome map to human chromosome 9. This includes the candidate chicken sex determining gene *DMRT1*, which maps to the tip of human 9p, deletion of which causes sex reversal in humans. Several other genes from the chicken Z map to chromosome 5 in humans (Schmid *et al.*, 2000).

Birds have a ZW female: ZZ male sex chromosome system. The Z chromosome is conserved in gene content and size between species. On the other hand, the W chromosome is variable in size between species and can be highly differentiated from the Z (*e.g.* chicken), or it can be almost completely homologous to the Z (emu). The mammalian X and the bird Z share no homology. Rather, genes on the mammalian X chromosome are located on autosomes in bird, and genes on the bird Z are located on autosomes in mammals.

The mammalian Y chromosome is male-specific, whereas the bird W chromosome is female-specific. These two chromosomes therefore bear genes involved in male-

specific and female-specific functions respectively. In mammals the Y chromosome is male determining and contains genes involved in spermatogenesis.

1.3 Function of eutherian Y chromosomes

During the 1950's, most biologists believed the human/mouse Y chromosome to be genetically inactive, as no Y-linked inheritance had been revealed via pedigree studies (Stern, 1957). However, in 1959, sex determination was ascribed to the Y chromosome. More recently several functions have been attributed to the Y. It is now known that there are 30-35 genes on the non-recombining portion of the human Y chromosome (Lahn and Page, 1997), some of which control functions important for reproduction.

1.3.1 Sex determination

Sex determination is a unique and critical function of the mammalian Y chromosome. A dominant male determining gene was deduced to be on the Y chromosome from the phenotypes of human and mouse XO females and XXY males (Jacobs and Strong, 1959; Ford *et al.*, 1959; Welshons and Russell, 1959).

Sexual differentiation is somewhat different in marsupials. The presence of the Y chromosome is essential for development of the testis. However, XO individuals lack a testis, but often have an empty scrotum, no pouch and underdeveloped mammary glands. XXY individuals have internal testes but no scrotum, and may have an underdeveloped pouch. Thus, the Y chromosome in marsupials is necessary for testis determination, but other characteristics, such as a scrotum and pouch differentiation, are independent of embryonic androgen production (Sharman *et al.*, 1990). Instead, development of a pouch or scrotum may be controlled by dosage of a gene on the X chromosome, or the action of a paternally imprinted pouch gene on the X chromosome (Cooper *et al.*, 1993).

The mechanism of monotreme sex determination is unknown. The meiotic chain at male meiosis confuses the identity of the Y chromosome, and no male specific sequences have been isolated.

Classically, there are two steps involved in sex determination. First is the step from chromosomes to the formation of gonads, and the second step is from the gonad to phenotype. The first step is controlled by sex determining genes and the second controlled by production of gonadal hormones (reviewed by Vaiman and Pailhoux, 2000).

In eutherian mammals male and female embryos are indistinguishable until the development of the gonads. A urogenital ridge forms within the mesonephros of the embryo, and can develop into either an ovary or a testis. The mesonephros also develops both Wolffian (male) and Müllerian (female) ducts. The presence of a Y chromosome determines the fate of the genital ridge. Once a testis has been determined, Sertoli and Leydig cells differentiate, Sertoli cells produce anti-Müllerian hormone (AMH), which inhibits the development of the Müllerian ducts. Androgens, produced by Leydig cells, stimulate the differentiation of the male urogenital tract from the Wolffian ducts and also stimulate formation of external genitalia. In females the absence of male hormones results in regression of the Wolffian ducts, proliferation of Müllerian ducts and subsequent feminisation (reviewed by Vaiman and Pailhoux, 2000). The production of AMH in males signals the beginning of the hormonal phase of sexual differentiation. The genes and events leading up to, and involved in the regulation of Sertoli cell differentiation and AMH expression are critical in understanding the genetic component of sex determination.

In XY embryos the testis confers masculinity during critical stages of development. Female development can occur irrespective of the genetic sex of an individual (or the presence or absence of ovaries) if they lack testes (Jost, 1970). The Y chromosome, therefore, contains the critical switch that leads to testis determination, in turn leading to the expression of AMH causing regression of Müllerian ducts.

The subsequent expression of testicular hormones confers a male phenotype, as shown sex reversal phenotypes. For instance, XY embryos that produce androgens but lack the androgen receptor develop as females. Conversely, male testicular hormones can be transferred to a female twin during pregnancy causing masculinization (Jost *et al.*, 1972). Freemartins are XX but were twin pregnancies with an XY embryo. After birth, they are characterised by several variable features: individuals are sterile; ovaries are reduced in size and often contain seminiferous tubules; and the presence of seminal

vesicles are common, but a prostate and external male genitalia are rare. AMH is high at birth in normal males calf and freemartins. After birth the level remains stable in normal males but drops sharply in freemartins. AMH, causing masculinization of the freemartin calf, is therefore essentially of male calf origin (Rota *et al.*, 2002).

In marsupials, sexual development is not entirely dependent on production of gonadal hormones. Scrotal bulges are visible in male but not females, and mammary anlagen were visible in females some days before testis development, and before Wolffian and Müllerian duct differentiation (O *et al.*, 1988). This confirms genetic studies that show that scrotal/ mammary gland development is independent of the Y chromosome, but depends on X gene dosage.

The testis-determining factor (TDF) on the human Y chromosome was hunted for many years by deletion analysis using sex-reversed individuals with fragments of the Y chromosome. XX males with part of the Y chromosome exchanged for the X should have the TDF, whereas XY females should lack the TDF. The TDF was eventually mapped to the distal short arm of the human Y chromosome. The HY antigen (a male-specific minor histoincompatibility antigen), a TDF candidate for many years, was excluded when it was mapped to the long arm of the Y chromosome (Simpson *et al.*, 1987).

The region containing the TDF was narrowed by further deletion analysis to a small region on Yp, from which a zinc finger protein was cloned (*ZFY*) (Page *et al.*, 1987). *ZFY* has an X-linked homologue *ZFX*, which maps to the short arm of the human X chromosome (Schneider-Gadicke *et al.*, 1989). *ZFY* was highly conserved between mammalian species and contained motifs similar to those of known transcription factors. It therefore made a good candidate for the TDF gene that was expected to control the activity of testis-differentiating genes.

ZFY first came into doubt as the TDF when it was shown that it was not on the Y chromosome in marsupials. Instead, it mapped to chromosome 5p in tammar wallaby, and to chromosome 3 in *Sminthopsis* (Sinclair *et al.*, 1988), and was later shown to be autosomal also in monotremes (Watson *et al.*, 1993). The autosomal location of *ZFY* in marsupials and monotremes excluded it as the universal mammalian sex determining gene. The exclusion of *ZFY* as the mammalian testis-determining factor was supported by its inappropriate expression in germcells but not Sertoli cells (Koopman *et al.*, 1989).

Expression in Sertoli cells is a prerequisite of the TDF because Sertoli cells produce AMH. *ZFY* was finally excluded as the TDF when it was found that some human XX males had fragments of the Y chromosome that did not include *ZFY*, re-defining the region of the Y chromosome that contained the TDF as just proximal to the PAR boundary (Palmer *et al.*, 1989).

A new gene, *SRY*, was cloned from this region of the Y chromosome (Sinclair *et al.*, 1990) and an orthologue was shown to map to the marsupial Y chromosome (Foster *et al.*, 1992). Several human XY females were shown to have point mutations in *SRY* (Berta *et al.*, 1990). XX mice transgenic for mouse *Sry* developed as males, demonstrating that *Sry* was indeed the TDF (Koopman *et al.*, 1991). The SRY protein contains a DNA binding domain (HMG box) that binds to DNA at a 6-base consensus target sequence (Harley *et al.*, 1992). Binding with the HMG box of the SRY protein bends DNA at specific angles that might bring sequences, or proteins bound to them, into position for activation (Ferrari *et al.*, 1992). This could act as a master switch, setting off a cascade of events that leads to testis determination and ultimately to male determination in eutherians.

Unexpectedly, *SRY* was found to be poorly conserved between species (Foster *et al.*, 1992). There is moderate conservation of HMG box sequences, but no conservation outside the HMG box. Most sex reversing mutations of human *SRY* occur in the HMG box; the products of mutant alleles either bind poorly or bend DNA incorrectly (Harley *et al.*, 1992). The conserved function of *SRY*, is therefore, thought to be controlled by the HMG box. Hence, the function of *SRY* relies on its ability to bind and bend DNA. This is thought to bring into proximity non-adjacent DNA sequences, or the proteins associated with them to form a productive interaction.

The cloning and characterization of *SRY*, and the confirmation that it was the mammalian TDF, was expected to provide the breakthrough to discover other steps in the sex-determining pathway.

Study of sex-reversed and intersex mammals have identified several genes involved in the pathway leading to testis determination, and subsequent AMH expression. These include *WT1*, *SF-1*, *SOX9*, *DAX-1* and the master switch *SRY*. Four proteins are known to interact with the AMH promoter and regulate its expression (*WT1*, *SF-1*, *SOX9* and *GATA-4*). Tissue and stage of expression for *WT1* and *SF-1* suggests that they are also involved in early development of the undifferentiated gonad from the urogenital ridge.

Loss of one copy of the autosomal *SF-1* gene in humans results in complete XY sex reversal. For instance, a patient with XY sex reversal had a 2bp (frameshifting) deletion in *SF-1* that caused sex reversal in the heterozygote (Achermann *et al.*, 1999). *SF-1* activates AMH by binding to a specific and conserved promoter region about 90bp 5' of the transcription start site. In humans, different mutations of the *WT1* gene result in four different sex-reversal pathologies (Little and Wells, 1997). As a result of alternative splicing, alternative translation initiation and RNA editing, 16 different protein products are produced by *WT1* (Reddy and Licht, 1996), only one of which interacts with *SF-1* to activate AMH (Nachtigal *et al.*, 1998).

SOX9 has a very conserved function in testis determination and AMH regulation. It is upregulated in the testis and downregulated in the ovary of all vertebrates. This gene was suggested to be controlled by *SRY* in a double inhibition pathway (McElreavey *et al.*, 1993), and *SOX3* was suggested to be the gene inhibited by *SRY* and inhibiting *SOX9* (Graves, 1998) (Figure 1.8). A conserved SOX binding site is located 150bp 5' of the AMH transcription site. The importance of this site has been demonstrated in mouse by its disruption (Arango *et al.*, 1999). Defects in *SOX9* indicate that there are controlling elements up to 1Mb upstream (Wunderle *et al.*, 1998). The report of XX sex reversal in human caused by duplication of *SOX9* (Huang *et al.*, 1999), and the report of Odsex (*Ods*) mice also demonstrated the importance of *Sox9* in XX sex reversal. In normal XX individuals *Sox9* expression was downregulated in the foetal gonads. In normal XY individuals, and XX *Ods*/+ individuals, *Sox9* expression was upregulated and maintained. Odsex mice carry a 150kb deletion approximately 1Mb upstream of *Sox9*. It was proposed that this deletion blocked function of a long range, gonad-specific regulatory element that mediated the repression *Sox9* in XX individuals (Bishop *et al.*, 2000). The double inhibition pathway best explains XX males that lack *SRY* (McElreavey *et al.*, 1993); the gene/s (repressed by *SRY*) that act upon *SOX9* to repress its expression may have lost function in these individuals.

DMRT1, located on human chromosome 9, is the most conserved putative vertebrate sex-determining gene isolated. Heterozygous deletions of the chromosomal region that contains *DMRT1* results in male to female XY sex-reversal. This is probably caused by haploinsufficiency of *DMRT1* and/or another gene from this region. Mice that carry homozygous deletions of *Dmrt1* are not XY sex-reversed but are infertile due to severe defects in proliferation of germ cells (Raymond *et al*, 2000). *DMRT1* has also

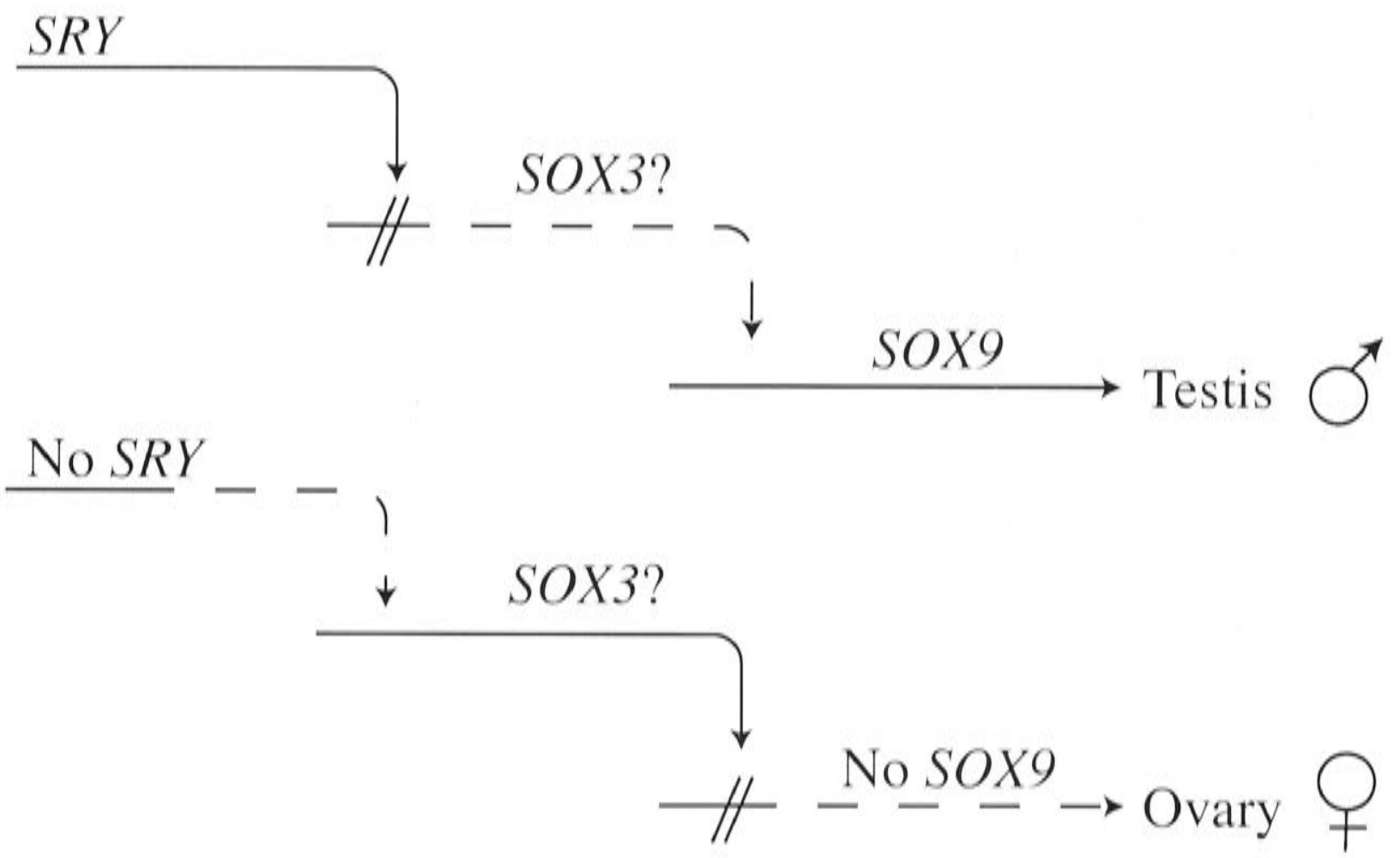


Figure 1.8: The double inhibition pathway. SRY inhibits the expression of SOX3 and SOX3 inhibits the expression of SOX9. In the presence of SRY, SOX9 is expressed and a testis forms. In the absence of SRY, SOX3 is expressed inhibiting SOX9 expression and an ovary forms.

been mapped to the Z chromosome, but is not present on the W in chickens (Nanda *et al.*, 1999). It is therefore present in a double dose in male and a single dose in female birds, making it a strong candidate for a dosage dependent sex-determining locus in birds. However, in the absence of mutation analysis or transgenesis in chicken, it cannot be disregard that differential dosage of another of the thousands of genes on the Z, and absent from the small heterochromatic W (therefore present as two copies in males and one copy in females), is responsible for sex determination.

Important confirmation has therefore come from emu, which has only a small region of differentiation between the virtually homomorphic Z and W chromosomes (Shetty *et al.*, 1999). *DMRT1* in emu is located on the Z and not the W (Shetty *et al.*, in press). It is the only reported Z-specific gene in emu, and is expressed during gonadogenesis (Shetty *et al.*, in press), supporting the hypothesis that *DMRT1* is involved in bird sex determination and therefore plays a conserved role in mammalian sex determination.

Another gene essential for sex determination is *ATRX*. Mutation in *ATRX* causes X-linked male to female sex-reversal in XY embryos, and is also responsible for α-thalassemia and mental retardation (Gibbons *et al.*, 1995). *ATRX* sex-reversed patients have streak gonads but no Müllerian ducts, indicating that AMH is expressed from Sertoli cells in these patients. Gonad formation is therefore inhibited after the determination of testes by *SRY*, *SOX9* and AMH expression (Pask and Renfree, 2001). In marsupials, *ATRX* has a Y-linked homologue (*ATRY*) (Pask *et al.*, 2000) that has been lost from the Y chromosome in eutherian mammals. It was suggested that *ATRY* had a role as an ancestral sex-determining gene in the common ancestor to therian mammals.

Duplications of human Xp21.3 result in XY individuals with ambiguous or female external genitalia (Zanaria *et al.*, 1995). *DAX-1*, in the critical region, has been identified as a candidate gene for this sex-reversal. The expression of *DAX-1* corresponds to the rise and fall of *SRY* expression. *DAX-1* has been suggested to interfere with *SF1* action. *DAX-1* binds to hairpin loops in DNA, which are caused by inverted repeats. Such repeats have been identified close to *SF1* binding sites (reviewed by Swain and Lovell-Badge, 1999). *DAX-1* could therefore prevent the promotion of AMH expression by *SF1*. *SRY* could operate antagonistically to do *DAX-1* by competing for to the same binding sites. A double dose of *DAX-1* could result in sex-reversal

because it out competes *SRY* (Swain *et al.*, 1998). The role of *DAX-1* in sex determination is unclear, but it is accepted that it interferes negatively in testis development (reviewed by Veitia *et al.*, 2001).

Few cases of XX sex reversal have been explained and many cases of XY sex-reversal have not yet been explained, so there are probably still many genes in the mammalian pathway still to be discovered.

1.3.2 Spermatogenesis

Other Y-borne genes are involved in spermatogenesis, the production of sperm. The mammalian testis contains Leydig cells, macrophages, blood vessels, lymphatic vessels and fibroblast-like cells that surround the seminiferous tubules. Sertoli cells are located in the epithelium of the seminiferous tubules, where spermatogenesis occurs (Weiss, 1983).

Spermatogonia are diploid precursors to spermatocytes, which undergo meiosis to produce haploid sperm. Spermatogonia divide mitotically to produce two daughter cells. One remains at the basal lamina of the seminiferous tubule epithelium; the second is committed to differentiate into sperm. The committed cell undergoes two mitotic divisions, producing four diploid daughter cells called primary spermatocytes. Each of the primary spermatocytes undergoes the first meiotic division, producing two diploid secondary spermatocytes. The secondary spermatocytes continue though the second meiotic division producing the haploid spermatids (Weiss, 1983).

The spermatids mature in close association with each other and Sertoli calls via cytoplasmic bridges through which proteins and other molecules diffuse. The sperm head, acrosome, midpiece and tail are assembled, producing mature spermatozoa. Spermatozoa are released into the lumen of the seminiferous tubules, collected in the epididymis and matured further into motile sperm (Figure 1.9) (Weiss, 1983).

The idea that one or more spermatogenesis genes must lie on the human Y chromosome was first put forwards when many sterile, but otherwise normal men were observed to have deletions of the Y chromosome. The loss of all or part of the long arm of human Y chromosome was associated with azoospermia, implying that some kind of azoospermia factor (AZF) was located in this region (Tiepolo and Zuffardi, 1976). Vogt *et al.*(1996) defined three non-overlapping regions (AZFa, AZFb and AZFc) on the long

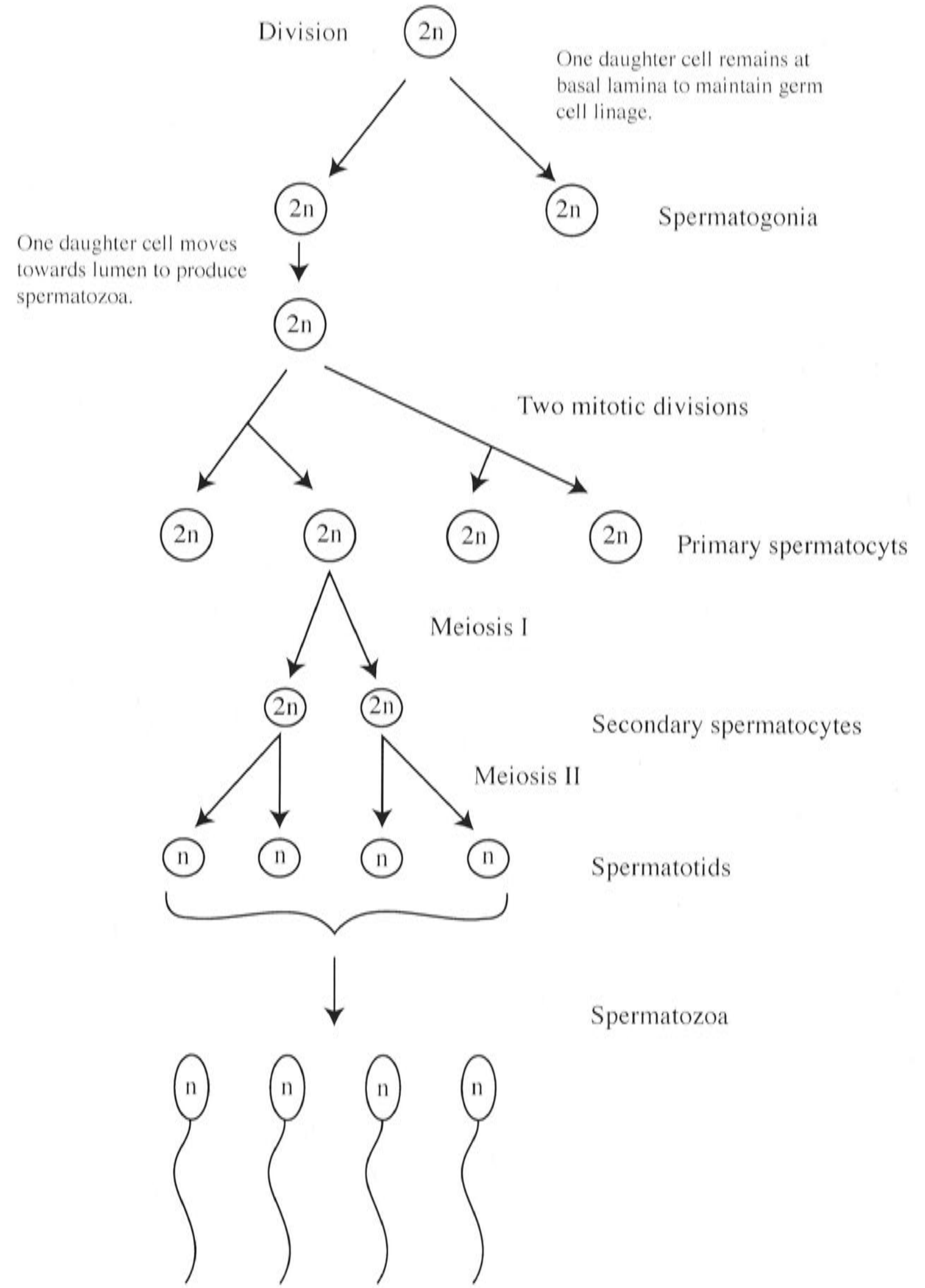


Figure 1.9: The stages of mammalian spermatogenesis. The mitotic and meiotic divisions take place within the wall of the seminiferous tubules. The developing germ cells are in close contact with each other and Sertoli cells (Weiss, 1983).

arm that, when deleted, resulted in spermatogenic failure. A fourth AZF region (AZFd) was defined between AZFb and AZFc (Kent-First *et al.*, 1999). Deletions resulting in spermatogenic failure can be restricted to any one of the four regions or can encompass multiple AZF regions.

Deletions of AZFa result in a severe phenotype, including Sertoli cell only (SCO) type I with a complete absence of germ cells, azoospermia and severe oligozoospermia (absent or few sperm) associated with spermatogenic arrest. Deletions of AZFb, AZFc and AZFd all display a range of phenotypes, from azoospermia associated with complete spermatogenic arrest to mild oligozoospermia (reviewed by Kent-First, 2000).

AZFa is located at proximal Yq in deletion interval 5 (Figure 1.10). It spans approximately 800kb (Sun *et al.*, 1999), and contains three known genes: Ubiquitin-specific protease 9, Y chromosome (*USP9Y*), DEAD/H box polypeptide, Y chromosome (*DBY*) and ubiquitously expressed tetratricopeptide repeat, Y chromosome (*UTY*). All three of these genes are ubiquitously expressed and have X-linked homologues that escape X-inactivation (Jones *et al.*, 1996; Lahn and Page, 1997). However, the mouse orthologue of *Usp9y* is testis specific (Brown *et al.*, 1998). The map location of *USP9X* implicates it in the gonadal phenotype associated with Turner syndrome, in which there is failure of the oocytes to pass through the first meiotic phase (Jones *et al.*, 1996; Cockwell *et al.*, 1991). *USP9X/Y* is the orthologue of the Drosophila developmental gene *fat facets* (*faf*), which, when mutated, results in abnormal oocyte content and inability of the fertilized egg to undergo normal embryogenesis. Therefore, despite its ubiquitous expression, *USP9Y* is an obvious and strong candidate for AZFa that may play a role in the development and maintenance of the male germ cells (Kent-First, 2000).

A role for *USP9Y* in spermatogenesis was confirmed by the discovery of a 4-bp deletion in a splice–donor site. This resulted in skipping of an exon and protein truncation in a patient with premeiotic arrest, meiotic germ cells in most seminiferous tubules and a small number of spermatids (Sun *et al.*, 1999). The *de novo* mutation was present in the patient, but absent in his fertile brother. Another infertile patient with no germ cells, but only Sertoli cells (SCO), was observed with a deletion of most of the AZFa region. The more severe phenotype in the second patient implied that a second spermatogenesis gene could also have been deleted.

DBY produces a long transcript that is ubiquitously expressed, along with a second shorter transcript that is expressed only in the testis (Foresta *et al.*, 2000), suggesting

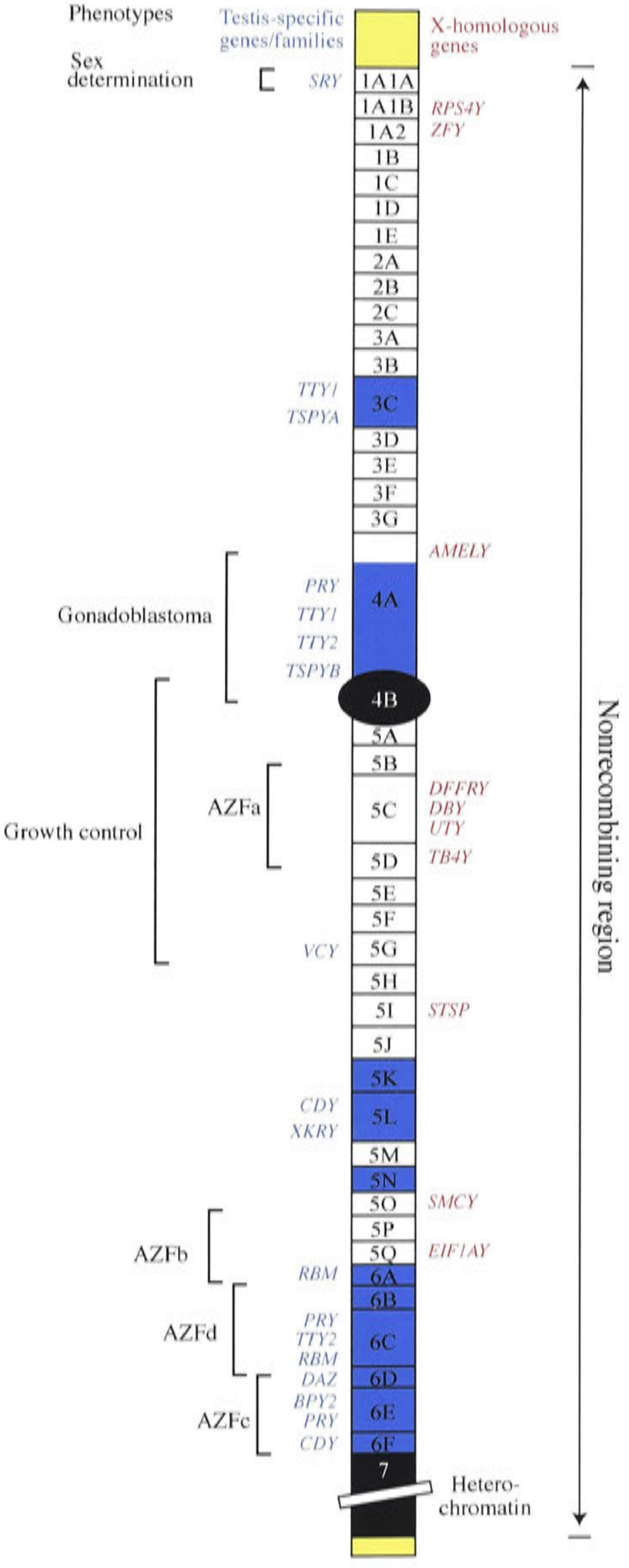


Figure 1.10: Gene map of the non-recombining region of the human Y chromosome according to Lahn & Page (1997). Large NRY is flanked by PARs (yellow). NRY is divided into 43 ordered intervals (A1A1 to 7) by deletion analysis (Vollrath et al., 1992). The intervals known to contain Y specific repeats are shaded blue. Left of the chromosome are genes with functional X homologues, right are testis specific genes or gene families, some with multiple locations (*SRY* and *RBMY* should be rebadged as X/Y homologouse genes, as discussed). Within intervals genes have not been ordered. On the left are regions involved in sex determination, spermatogenic failure, short stature (GCY) and gonadoblastom (GBY).

that the gene has both housekeeping and testis-specific functions. *DBY* is probably responsible for the more severe SCO phenotype in the second patient and *USP9Y* for the milder oligozoospermia (Foresta *et al.*, 2000). Conformation of the role of *DBY* in spermatogenesis awaits the detection of a point mutation in a patient displaying a Sertoli cell only phenotype. In mouse, deletion of the *ΔSxrb* interval, containing *Usp9y*, *Dby* and *Uty*, results in a severe phenotype similar to that of AZFa in human (Sutcliffe and Burgoyne, 1989), strengthening the hypothesis that both these genes are crucial to spermatogenic function (Mazeyrat *et al.*, 1998).

There are several unstable regions on the Y chromosome that contain motifs such as insertions, repeats and palindromes. Microdeletions on the Y chromosome tend to begin at, or extend to these regions (reviewed by Kent-First, 2000). Deletions of the AZFa region are probably mediated by homologous human endogenous retroviral sequences (HERV15) located at either end of the region. Study of men with AZFa deletions revealed that break points of the deletion fell within the HERV15 retroviruses (Sun *et al.*, 2000; Kamp *et al.*, 2000). An intrachromosomal recombination event between these two sequences is probably responsible for most deletions of AZFa and explains why the deletion is consistent between individuals, although all these mutations must be *de novo*.

AZFb extends from the distal part of deletion interval 5 to the proximal part of interval 6 of the human Y chromosome (Figure 1.10). Five genes have been described in this 1-3Mb region (*SMCY*, *EIF1AY*, *RBMY*, *PRY* and *XKRY*). *RBMY* (RNA binding motif on the Y chromosome) (Ma *et al.*, 1993) is considered to be the best candidate for AZFb because it is conserved on the Y chromosome in eutherian and marsupial mammals (Delbridge *et al.*, 1997), indicating that this gene has an important and conserved function. *RBMY* contains an RNA recognition motif that has been shown to mediate an interaction with RNA (Ma *et al.*, 1993). There are about 30 copies of *RBMY* distributed across the proximal short arm and euchromatic portion of the long arm of the human Y chromosome. They have been divided into six subfamilies (*RBMI* to *RBMVI*), of which only *RBMI* is actively transcribed and encodes functional proteins. A total of six *RBMI* copies have been identified (Chai *et al.*, 1997), and expression of these *RBMY* copies relies on a small critical region in the AZFb interval (Elliott *et al.*, 1997), strengthening its credentials as a spermatogenesis factor. *Rbmy* has also been amplified

independently on the mouse Y chromosome, as shown by differences in human and mouse *RBMY* gene structure. Most of its 28 copies can be deleted without causing sterility (Burgoyne *et al.*, 1992). Thus both the human and mouse Y chromosomes have many *RBMY* pseudogenes.

The 1.5Mb AZFc is distal to AZFb, and is flanked by the heterochromatic portion of the Y chromosome. AZFc comprises three subregions; proximal AZFc (AZFd), middle AZFc (*DAZ*) and distal AZFc. AZFc contains the *DAZ* gene cluster, at least two copies each of *PRY*, *BPY2*, *TTY2*, and additional copies of *CDY* and *RBMY* (reviewed by Kent-First, 2000). The function of most of these genes is unknown.

When *DAZ* was originally mapped to the Y chromosome (Reijo *et al.*, 1995) it was proposed to be single copy. However, there are now know to be seven *DAZ* genes or pseudogenes clustered within AZFc (Glaser *et al.*, 1998b). *DAZ* codes for a 366 amino acid germ cell-specific protein that contains an RNA recognition motif. It has 8-24 copies of exon 7, whose copy number varies from one *DAZ* gene to the next within a Y chromosome, and between *DAZ* genes in different individuals (Yen *et al.*, 1997). It has an autosomal homologue, *DAZLA*, on human chromosome 3, which is expressed in the germ line of both males and females (Dorfman *et al.*, 1999; Kent-First, 2000). Its Drosophila orthologue (*boule*) is also required for spermatogenesis but not oogenesis (Eberhart *et al.*, 1996), and its *C. elegans* orthologue (*daz-1*) is required in the hermaphrodite for oogenesis (Karashima *et al.*, 2000). Mice that are homozygous for mutant *Dazla* alleles (*Dazla-/-*) have severe germ cell depletion and meiotic failure. The human *DAZ* gene partially rescues the sterile phenotype; they are still infertile but have a substantial increase in germ cell population that survives to the pachytene stage of meiosis (Slee *et al.*, 1999).

Spermatogenesis is a critical male specific function. Only 3-30% of spermatogenic failure is a result of microdeletions of the human Y chromosome (Kent-First, 2000), suggesting that there could be many autosomal factors involved in spermatogenesis. On the human Y chromosome there are three, if not four, critical regions that contains factors required for proper spermatogenic function. Exactly which genes are involved in spermatogenic function is not yet clear.

1.3.3 H-Y antigen

The first genetic difference attributed to the Y chromosome was the rejection of male tissue by females from the same inbred mouse line. This was accounted for by the existence of a male-specific minor transplantation antigen, Histocompatibility Y chromosome (H-Y) antigen (Eichwald and Silmser, 1955). Originally thought to be involved in sex-determination (Wachtel *et al.*, 1975), the gene(s) that encoded H-Y antigen remained a mystery for some time. It was initially thought that H-Y antigen represented a surface protein that may have had some function in intracellular signalling. However, it now appears that fragments of many intracellular proteins are displayed on the cell surface with major histocompatibility complex (MHC) molecules (Scott *et al.*, 1995). This means that any gene expressed ubiquitously from the Y chromosome that is unique, or distinct, from its X homologue has the potential to act as a H-Y antigen. Female rejection of male tissue is thus the accidental by-product of ubiquitously expressed Y encoded proteins being broken down and presented on the cell surface as male specific self-antigens.

The mouse H-Y antigen localized to the *ΔSxrb* deletion interval on the short arm of the Y chromosome (McLaren *et al.*, 1988), comprises at least four different epitopes. Both the ubiquitously expressed genes, *Smcy* and *Uty*, have been shown to code for different H-Y epitopes in mouse, H-YKK and H-YDD respectively, demonstrating that more than one gene on the Y chromosome is responsible for the H-Y antigen (Scott *et al.*, 1995; Greenfield *et al.*, 1996).

SMCY also codes for two H-Y epitopes in human (Wang *et al.*, 1995; Meadows *et al.*, 1997). More recently, the human *USP9Y* (*DFFRY*) gene was identified as coding for a H-Y epitope. This was deduced by mass spectrometry, and by female rejection of an HLA-phenotypically identical bone marrow graft, received from her father (Pierce *et al.*, 1999; Vogt *et al.*, 2000). Thus the human Y chromosome and the mouse Y chromosome each contain at least two genes that code for different H-Y epitopes. *SMCY*, *USP9Y* and *UTY* are all ubiquitously expressed in human fulfilling the first requirement for genes encoding H-Y epitopes. Even though *UTY* has not been demonstrated to code for a human H-Y epitope, it still has the potential to do so, as does any gene ubiquitously expressed from the Y chromosome. In mouse, *Smcy* and *Uty* are

ubiquitously expressed, whereas *Usp9y* is testis specific and, therefore, cannot code for a mouse H-Y epitope.

Since most Y-borne genes with an X-linked homologue show ubiquitous expression whereas, Y specific genes have testis specific expression, genes that code for H-Y epitopes are likely to be shared with the X. Other human X-Y shared genes that might contribute to human HYA are *ZFY*, *DBY*, *TB4Y* and *EIF1AY*.

H-Y antigen therefore does not appear to have any functional importance. It is only the accidental by-product of proteins, coded for by ubiquitously expressed Y-borne genes, being broken down and presented on the cell surface.

1.3.4 Growth control

Height is a multifactorial trait controlled by many genes and influenced by environmental factors. However, because adult males are generally taller than adult females (Abassi, 1998), it is assumed that the Y chromosome contains genes associated with height. This hypothesis is supported by several observations; XY females are, on average, taller than XX females; XY males are, on average, taller than XX males and XYY males are, on average, taller than XY males (de la Chapelle, 1972). The human Y chromosome therefore bears at least one gene involved in height. This growth control gene(s) on the Y chromosome is termed GCY.

Observations of short males with deletions within the long arm of the Y chromosome confirmed the idea of growth genes on the Y chromosome. The GCY locus was determined, by deletion mapping, to lie in a 4.5Mb pericentric region of the long arm at Yq11 (Figure 1.10) (Salo *et al.*, 1995a; Kirsch *et al.*, 2000). Finer deletion mapping later divided this region into two discrete non-overlapping regions termed GCY I (Ogata *et al.*, 1995) and GCY II (Rousseaux-Prevost *et al.*, 1996; De Rosa *et al.*, 1997). The GCY II region contains *USP9Y*, *DBY*, *UTY* and *TB4Y* (Lahn and Page, 1997). However, FISH mapping of Y cosmids in individuals with short stature, has excluded GCY II (and the genes that reside in it) as the locus responsible for short stature. This leaves a single GCY locus of about 2Mb, to which no candidate genes have yet been isolated (Kirsch *et al.*, 2000). Further study of the GCY interval is required to identify candidate genes. A candidate gene could then be confirmed as the GCY via detection of a point mutation that confers short stature.

1.3.5 Gonadoblastoma

Gonadoblastoma is a rare tumour that comprises aggregates of germ cells intermixed with smaller epithelial cells resembling immature Sertoli and granulosa cells. These aggregates are surrounded by Leydig cells (Scully, 1970). Gonadoblastomas develop in XY females but not XO females. The high frequency of gonadoblastoma in XY females led to the hypothesis that tumourigenesis is promoted by a gonadoblastoma locus on the Y chromosome (GBY) (Page, 1987). The GBY locus was proposed to have a normal function in the testis that acts as an oncogene only in the dysgenic gonad.

Initially, GBY was mapped to deletion interval 3 on the short arm, and intervals 4B-7 on the long arm of the human Y chromosome. Further study sublocalized GBY to 1-2Mb within intervals 3E-3G on the short arm of the Y (Figure 1.10) (Tsuchiya *et al.*, 1995). A second critical region of about 4Mb was mapped between intervals 4B and 5E on the long arm (Salo *et al.*, 1995b), suggesting that more than one gene is involved, or a gene is present in multiple copies. Alternatively, the discrepancy between the GBY regions could be due to inversion polymorphisms on the Y chromosome (Vogt *et al.*, 1997).

Seven genes have been localized to the interval 3 critical region (*AMELY*, *RBMY*, *PRKY*, *PRY*, *TTY1*, *TTY2* and *TSPY*) (Lahn and Page, 1997; Schiebel *et al.*, 1997), all of which are GBY candidates. Of these genes *TSPY* is the most likely candidate. *TTY1* and *TTY2* are unlikely candidates because they consist of repetitive sequence, their transcripts lack protein-coding sequences and probably do not serve any function (Lahn and Page, 1997). *AMELY* has an X-linked homologue (*AMELX*), both of which encode an enamel protein that is expressed only in the developing toothbud (Lau *et al.*, 1989; Salido *et al.*, 1992) so probably serves no function in gonad cell proliferation. Even though there are *RBMY* transcripts detectable in gonadoblastoma tissue (Tsuchiya *et al.*, 1995), *RBMY* does not make a good GBY candidate, because all *RBMY* functional copies are located outside of the critical GBY region (Elliott *et al.*, 1997).

TSPY is a repeated gene whose functional members are mainly located in two clusters, *TSPYA* (interval 3C) and *TSPYB* (proximal to interval 3D, although the order of sequences on Yp is reversed in some individuals due to the inversion polymorphism) (Dechend *et al.*, 2000; Conrad *et al.*, 1996; Vogt *et al.*, 1997), with a single copy on the

proximal long arm (interval 6E) (Ratti *et al.*, 2000). *TSPY* shares strong homology to cyclin B binding proteins (Schnieders *et al.*, 1996). Other cyclin B binding proteins have been shown to be involved in the mitotic process, cell proliferation and carcinogenesis (reviewed by Lau, 1999), and *TSPY* is preferentially expressed in the germ cells of tumour aggregates in gonadoblastoma (Lau *et al.*, 2000). *TSPY* has been suggested to play a normal role in directing spermatogonial cells to enter meiosis (Schnieders *et al.*, 1996). The above circumstantial evidence suggests that inappropriate expression of *TSPY* in dysgenic gonads may play a role in GBY, making it the most likely candidate. However, *PRY* and *PRKY* are yet to be excluded as GBY candidates. Further study is required to determine what gene/s are involved in GBY.

Thus the human Y chromosome has had several functions attributed to it. None of these functions are critical to the survival of an individual. However, two of these functions, sex determination and spermatogenesis, are critical to the survival of mammalian species.

1.4 Evolution of sex chromosomes

The questions surrounding sex chromosome evolution have been studied and debated for decades. Some of the questions that have been asked are: How did sex chromosomes evolve? What did they evolve from? And, perhaps most fundamentally, why did sex and recombination evolve in the first place?

1.4.1 Why sex?

The answer to why sex and recombination evolved is not obvious when the costs involved are considered. It takes time and energy for organisms to find and secure a mate. Plants produce costly flower displays and nectar rewards for pollinators. Sexual reproduction is slower than asexual reproduction. During mating the individual is less likely to gather resources and evade predators. Sexual reproduction also introduces the risk of sexually transmitted diseases and transfer of parasitic genetic elements. In species with different sex or mating types, sexual conflicts may arise. For example, in *Drosophila* seminal fluid contains toxins (which inactivates sperm of rival males) that

reduce female fitness (Chapman *et al.*, 1995). There is also the risk of randomly mixing alleles with another individual to produce offspring. If an asexual population under selection is in equilibrium, recombination is more likely to break up favourable allele associations than create them. Putting alleles together randomly in a selective environment will break up associations proven to be beneficial, and bring together unproven associations that are likely to be less beneficial (Kondrashov, 1993).

Most importantly, there is the well known twofold cost of sex. The reproductive unit in sexual reproduction is the couple; in asexual reproduction it is the individual. If a sexual and an asexual individual have the same number of progeny, the sexually reproducing individual has only half the genetic input to the population as the asexual individual. This is because a sexual parent contributes only half of his genome to each offspring, whereas an asexually reproducing parent contributes all of his.

Considering all of these problems, the question arises as to why the vast majority of animal species undergo sexual reproduction, at least occasionally. Of about 42,300 known vertebrates, only 22 fish, 23 amphibians and 29 reptiles reproduce asexually (reviewed by Otto and Lenormand, 2002).

Sexual reproduction may be asymmetric or symmetric. Asymmetric reproduction occurs when only part of the genome is transferred from a donor to a recipient. Sexual reproduction in bacteria is almost always asymmetric (transduction, transformation and conjugation) (Redfield, 2001). Symmetrical reproduction, an exclusively eukaryotic phenomenon, requires the coming together of two genomes and then their segregation again at meiosis. Recombination produces gametes or haploid offspring that contain different combinations of genes.

Asymmetric sex results in the transfer of genetic elements to other individuals. The elements are copied, transferred and may be spread through a population. They are more likely to spread if they carry characters that increase the fitness of the host, such as a plasmid carrying antibiotic resistance in bacteria. However, they can spread even if they are detrimental, as long as these genetic elements infect new cells faster than they kill or reduce the fitness of the host. Asymmetric sex could be the by-product of mechanisms that are encoded in genetic elements that allow them to spread (Redfield, 2001).

On rare occasions asymmetric (horizontal transfer) processes have been implicated in gene transfer into and between eukaryotic genomes (Kidwell, 1993; Wolf *et al.*,

2000). Cytoplasmic elements often transfer in an asymmetric manner. Transposable elements can transfer from one genome to the other after syngamy. Cytoplasmic or transposable elements that drive sexual reproduction could therefore spread through a eukaryote population if they are represented more often in offspring than in the parents (Hickey, 1982). When resistance to these elements occur, the remaining asexual populations will out-compete the sexual ones. Thus, these elements do not maintain sexual reproduction in a population.

The change from a haploid genome to a diploid genome may have been the result of cannibalism in protists. In some cases of cannibalism the ingestion succeeds and one organism digests the other. Sometimes, the ingested partner is rejected and expelled back into the medium (Margulis and Sagan, 1986). However, in at least one reported case the result of ingestion was the formation of an individual with two nuclei, two sets of chromosomes and two sets of extranuclear organelles (Cleaveland, 1947). Protists have no immune system, but possess a rich assortment of restriction enzymes by which they can protect themselves from auto digestion. If a devoured organism were resistant to the enzymes of the predator, its DNA would be protected from destruction. In some cases, genetic differences between the two individuals could give the new diploid a selective advantage over the haploid, and there will be pressure for reproduction of the double cell (Margulis and Sagan, 1986). Regular fertilization probably evolved from casual cannibalism in unison with the evolution of a reduction division (meiosis) (Margulis and Sagan, 1986).

It is likely that asymmetric sex was the most ancient form of genetic exchange. The evolution of sexual reproduction could be sporadic, with the arrival of new elements driving sex, and then evolution of resistance to them. Rather than being a sporadic phenomenon, sex has evolved and persisted for most of eukaryotic evolutionary history (reviewed by Otto and Lenormand, 2002). Most asexual species of eukaryotes seem to become extinct more frequently than sexual species. This suggests that sex and recombination is beneficial to the long-term existence of a species (Bachtrog and Charlesworth, 2002). This may be because recombination increases the power of natural selection. If there is a beneficial mutation on a genetic background of poor fitness, it will be lost from the population unless the benefit of the new mutation outweighs the poor background (reviewed by Charlesworth, 2002).

Sex can be determined genetically, as with mammals and birds (Section 1.3.1), or by environmental mechanisms. Genetic sex determination (GSD) may involve a single locus, or many genes on differentiated sex chromosomes. Environmental sex determination (ESD) is usually determined by temperature at critical stages of development (temperature dependent sex determination, TDS), as with the American alligator (*Alligator mississippiensis*) (Lang and Andrews, 1994). ESD and GSD may be able to switch from one to the other with relative ease in evolutionary time (Bull, 1983). For instance, sex determination in two GSD species, *C. elegans* and *D. melanogaster*, became responsive to temperature after temperature sensitive mutation at specific loci (Nelson *et al.*, 1978; Belote and Baker, 1982). Also, GSD and TSD were observed in closely related lizard species (Viets *et al.*, 1994), indicating that one had recently evolved a new mechanism of sex determination.

GSD differs significantly between species and has arisen independently in many vertebrates, presumably because it is strongly selected for. GSD probably evolved from a mechanism such as TSD. Several species of reptiles, fish and amphibians have GSD but apparently lack differentiated sex chromosomes. Some species have only slightly heteromorphic sex chromosomes (Bull, 1983). In these species, a pair of autosomes, representing the first step in sex chromosome differentiation, acquired a locus determining sex. Many genes in the vertebrate sex determination pathway are conserved, though the initial switch leading either to male or female development differs. With vertebrates, the basic requirement is the acquisition of a genetic switch that ultimately regulates *SOX9* expression, as does *SRY* in mammals.

Thus sex has evolved in many species. The sex of an individual can be determined by environmental factors or genetic factors. Genetic sex determination requires the acquisition of a sex-specific gene (which determines sex) by one member of a autosomal pair, which can subsequently be differentiated into a sex chromosome pair.

1.4.2 Once an autosomal pair

Sex chromosomes were proposed to have evolved from homologous pairs of autosomes by Ohno (1967), on the basis of the observation of the Z and W chromosomes of various bird and snake species. The Z and W of different species appeared to represent intermediate steps in sex chromosome differentiation. In boid

snakes and ratite birds, the Z and W are identical with the exception of a small differentiated region (Shetty *et al.*, 1999), whereas in vipers and carinate birds the Z and W are highly differentiated. The Z and W of chicken share several homologous genes, supporting the hypothesis that they evolved from a pair of autosomes. The first step required for sex chromosome differentiation was the evolution of a new sex determining gene on one member of the original autosomal pair. Subsequent evolution of snake and bird sex chromosomes was attributed to accumulation of female advantage genes on the female specific W chromosome, and protection from recombination with the Z chromosome. More alleles with a female advantage accumulated nearby and recombination was suppressed further.

The origin of the mammalian X and Y chromosomes from an autosomal pair is, likewise, supported by the considerable homology between them. The human X and Y chromosomes share homology over PAR1 (Section 1.2.1.2), and most genes in the NRY have homologues on the X (Section 1.2.1.3). In the common mammalian ancestor, one member of an autosomal pair acquired a dominant male determining locus. Alleles conferring a male-advantage accumulated nearby, and this group of male-advantage genes on the proto-Y was protected from recombination with the X by suppression of recombination (probably via a chromosomal rearrangement on the Y). From then on, a region of the genome was male specific. Accumulation of male-advantage genes led to the extension of suppression of recombination further along the proto-Y until almost the entire chromosome was male specific. Lack of recombination between the X and Y prevented repair of mutations in this region, leading to the loss of gene function. Mutation, insertion and deletion around the sex determining region resulted in the eventual loss and heterochromatinisation of most of the Y. The Y is therefore cluttered with pseudogenes and repetitive elements.

Did the degradation of genes on the Y chromosome occur progressively or in jumps? Most genes on the human Y chromosome have a partner on the X chromosome. Assessing the dates of divergence of X and Y partners allowed Lahn and Page (1999) to divide the human X chromosome into four evolutionary strata, each of which had been isolated from recombination with the Y chromosome at different times (Figure 1.11).

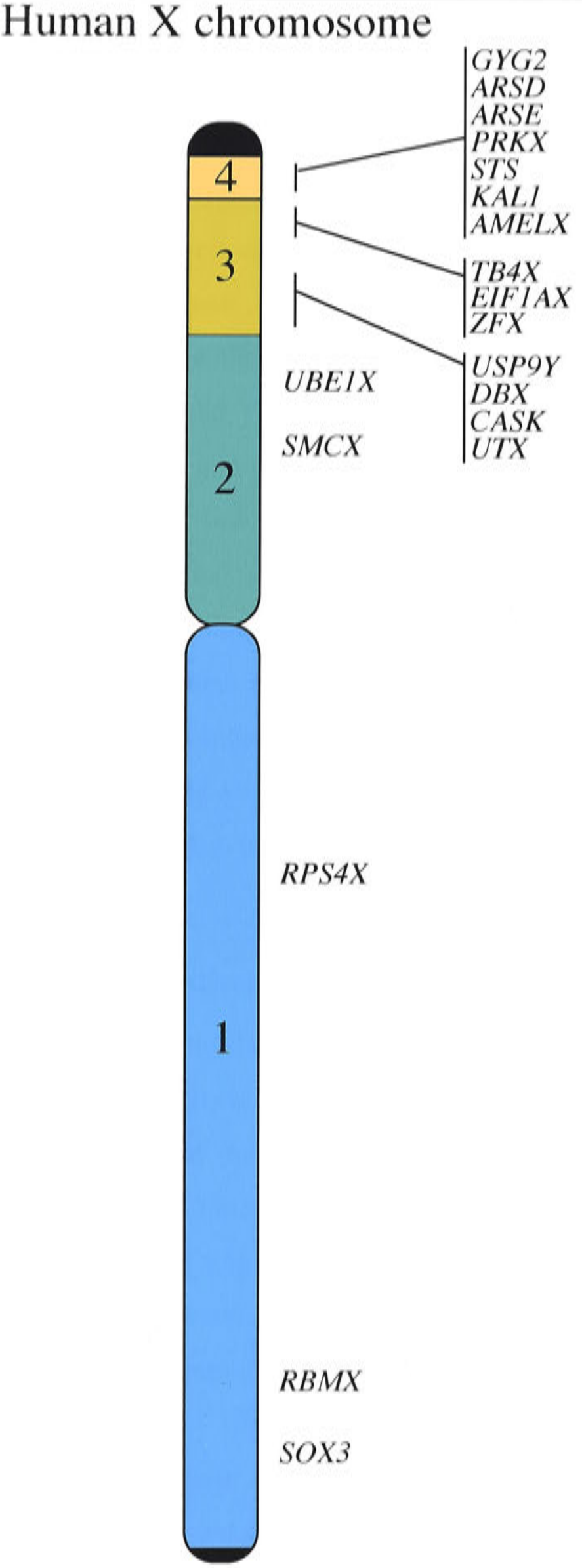


Figure 1.11: The four evolutioary strata of the human X chromosome according to Lahn and Page (1999). The strata were defind by the different times that they were recombinationally isolated from the Y. Estimations of the ages of the strata are: stratum 1, 240-320 million years; stratum 2, 130-170 million years; stratum 3, 80-130 million years; stratum 4, 30-50 million years.

The strata were determined by measuring synonymous nucleotide divergence (K_S) between 19 X-Y shared genes. The longer a gene has been isolated from its partner the greater the synonymous nucleotide divergence should be. The 19 genes studied had K_S values that clustered into four different groups. They concluded that there were at least four major events (probably inversions on the Y chromosome) leading to suppression of recombination between the human X and Y chromosomes, without disrupting gene order on the X. This method of determining the number of events that suppressed recombination between the X and Y chromosomes would miss strata that no longer contain X-Y shared genes (Lahn and Page, 1999).

The mammalian sex chromosomes were once a pair of autosomes. One (the proto-Y) acquired a male determining locus, accumulated male-advantage genes and underwent several inversions, leading to loss of recombination with its partner (the X). The mammalian sex chromosomes should therefore contain genes from the original proto-X and -Y. However, comparative mapping in distantly related mammals (marsupials and monotreme) has shown that the history of the eutherian X and Y is complicated by at least one addition to the proto-sex chromosomes.

1.4.3 Addition to the eutherian X

The addition-attrition hypothesis was first put forward by Graves (1995). On the basis of comparative mapping, Graves (1995) suggested that autosomal blocks of DNA were added to the PAR of one of the sex chromosomes and recombined to its partner, resulting in an extended PAR. Once a region was added to the Y chromosome, it becomes subject to the same degradation pressures. Rearrangements on the Y chromosome suppressed recombination along the extended PAR. The Y chromosome was then subjected to further degradation (attrition); this occurred in cycles (Figure 1.12).

This hypothesis was proposed to account for the results of comparative gene mapping between distantly related mammalian groups. Seventeen genes distributed across the proximal short arm, and entire long arm of the human X chromosome were mapped to the marsupial X chromosome (Section 1.2.2.1). However, nine genes on human Xp, as well as the human PAR1 gene *CSF2RA*, were mapped to tammar wallaby chromosome 5p, and three were mapped to chromosome 1 (Section 1.2.2.1).

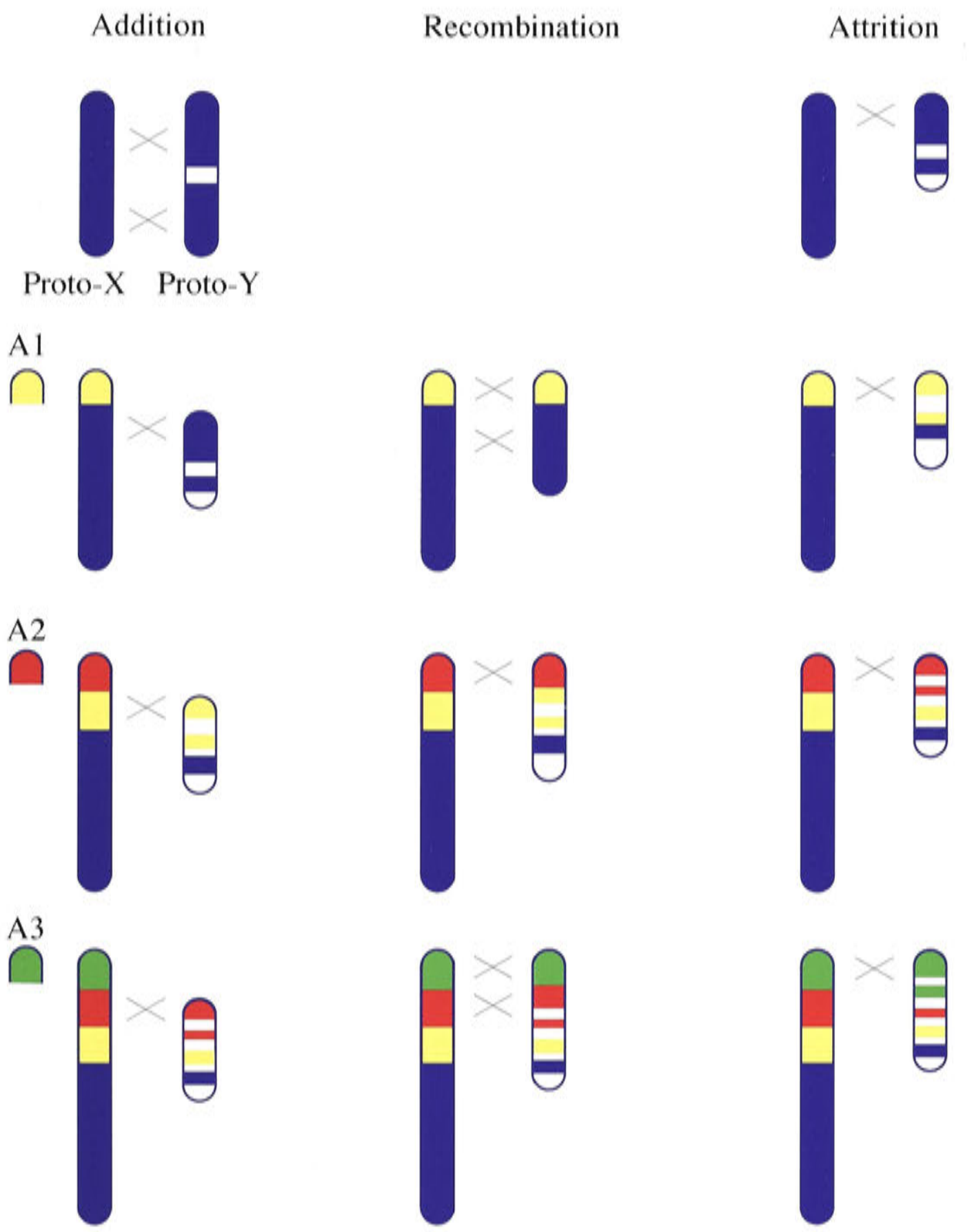


Figure 1.12: Proto-X and proto-Y (purple) were homologous except at a sex determining locus (white). They underwent pairing and crossing over (crosses), except at the sex determining locus. Attrition of the Y reduced the coding region of the Y progressively until most homology was lost, leaving a small PAR at the terminus and a few functional genes (purple stripe). An autosomal segment A1 (yellow) was added to the PAR of the X and recombined to the Y chromosome, extending the size of the PAR. Attrition again reduced the size of the region of crossing over leaving a few functional genes (yellow stripe) and a new PAR (yellow). This cycle was repeated for the autosomal regions A2 (red) and A3 (green), leaving different PRAs each time. The additions do not necessarily have to be to the X, they could have been to the Y and then recombined to the X. From Graves, 1995.

Significantly, genes in the larger of the two clusters (5p in tammar wallaby) cluster together on chromosome 3q in *Sminthopsis crassicaudata*, and map to chromosome 1 in distantly related monotremes (Section 1.2.3.1).

Graves (1995) proposed that the genes mapping to 5p in the wallaby represent a single addition to the eutherian sex chromosomes. The genes from the human Xp that cluster on 1p in the tammar wallaby might represent a separate addition to the eutherian sex chromosomes. These regions of human Xp were added to the eutherian sex chromosomes after the divergence of eutherians and marsupials 130MyBP, and before the eutherian radiation 80MyBP.

An alternative hypothesis is that the eutherian X chromosome represents the ancestral state, and the autosomal clusters in marsupials and monotremes represent regions that were lost from the X chromosomes after these groups diverged from the eutherian lineage. However, because monotremes and marsupials diverged independently from the eutherian lineage, this would require independent loss of the same regions. Therefore, the most parsimonious answer is that the ancestral X chromosome was similar to the marsupial X chromosome, and that autosomal regions were added to the eutherian sex chromosomes after the marsupial divergence (Figure 1.13).

The hypothesis states that the eutherian X chromosome consists of a region (XCR) that has been conserved there since at least before the monotreme-therian divergence 170MyBP. A second region (XAR) was added after the marsupial lineage diverged 130MyBP, but before the eutherian radiation 80MyBP.

Monotremes appear to have had an independent addition to their X chromosome, since it is large, euchromatic and shares a large pairing region with the putative Y (Watson *et al.*, 1992a). However, there is no direct mapping evidence for genes on the monotreme X chromosome that cluster at autosomal locations in marsupials and eutherians. There has also been an independent addition of the nucleolar organizer region to the X and Y chromosomes in the Family Macropodidae (Toder *et al.*, 1997b).

The ultimate origin of mammalian sex chromosomes is determined by mapping genes from them in a vertebrate outgroup. There is no homology between mammal and bird sex chromosomes. Genes on the bird Z map to mammal autosomes, and genes on the mammal X map to bird autosomes. It is therefore clear that the two different sets of

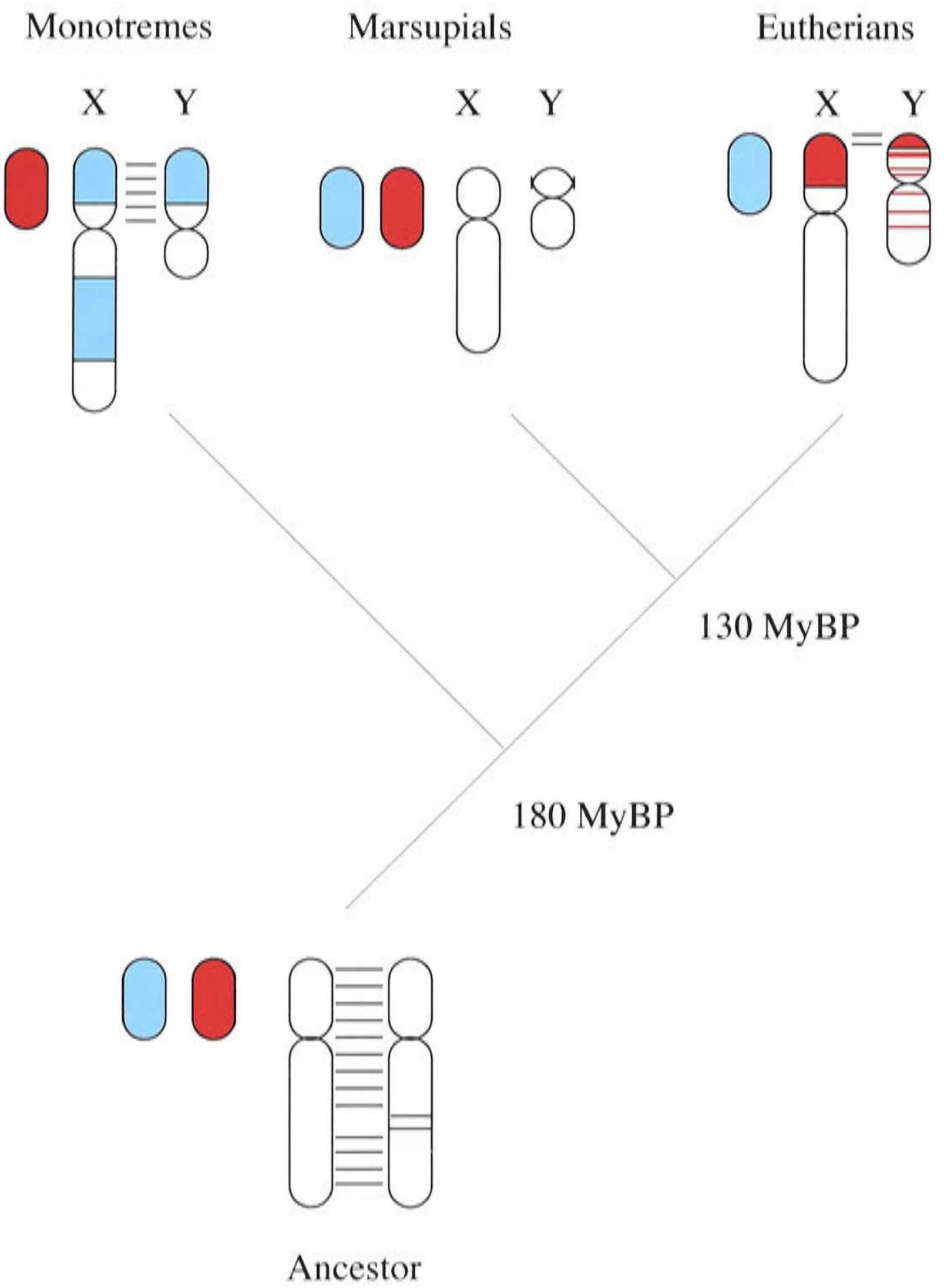


Figure 1.13: Evolution of the mammalian sex chromosomes. In the mammalian ancestor the proto-sex chromosomes were homomorphic. Recombination was suppressed leaving the Y chromosome open to degradation. There were autosomal additions to the eutherian sex chromosomes. The marsupial X chromosome only contains material that is conserved from the common ancestor and represents the XCR of eutherians.

sex chromosomes evolved independently of each other from different ancestral autosomal pairs (Figure 1.14). The chicken mapping data (Section 1.2.4) supports the hypothesis that XCR and XAR had independent origins. The ancient mammalian X chromosome (the marsupial X chromosome) is represented by chicken 4p. The eutherian XAR, autosomal in wallaby, is represented by 1p in chicken.

The eutherian X chromosome has been proposed to consist of a conserved region (XCR) and a recently added region (XAR), which had an independent origin. The theory that the X represents two originally autosomal blocks predicts that their gene content will reflect that of the genome at large. Recent data suggests this is not so.

1.4.4 Gene content of the mammal X

There is convincing evidence that the human X and Y chromosomes harbour a high proportion of genes involved in sex and reproduction (SRR) (discussed in section 1.2.1.1). Was the selection of the pair of mammalian sex chromosomes a random event, or was the proto-X/Y chosen for its gene content? Or did SRR genes accumulate on the X chromosome because it is unpaired in males?

There have been arguments over many decades that the X chromosome should accumulate SRR genes (Hurst and Randerson, 1999; Graves and Delbridge, 2001). The X chromosome is hemizygous in males, so a recessive mutation arising on the X that is beneficial to males would be immediately selected for, even if the allele were detrimental to females (sexually antagonistic allele). Because females have two X chromosomes, a recessive detrimental allele would not be selected against because the wild-type allele would mask its effect. Selection against XX individuals homozygous for the allele would not been seen in the population until the allele had reached a significant frequency. Conversely, a recessive mutation with the opposite effect (beneficial to XX and detrimental to XY individuals) would be strongly, and immediately, selected against. Compared to the fate of an autosomal recessive mutation, which is unlikely to spread because the phenotype would rarely be seen for selection to act upon, the X chromosome acts as a vehicle for recessive, sexually antagonistic alleles to spread though a population (Hurst and Randerson, 1999).

The question of whether the mammalian sex chromosomes were "chosen" because they were suited to the job (*i.e.* already containing genes that have an increased

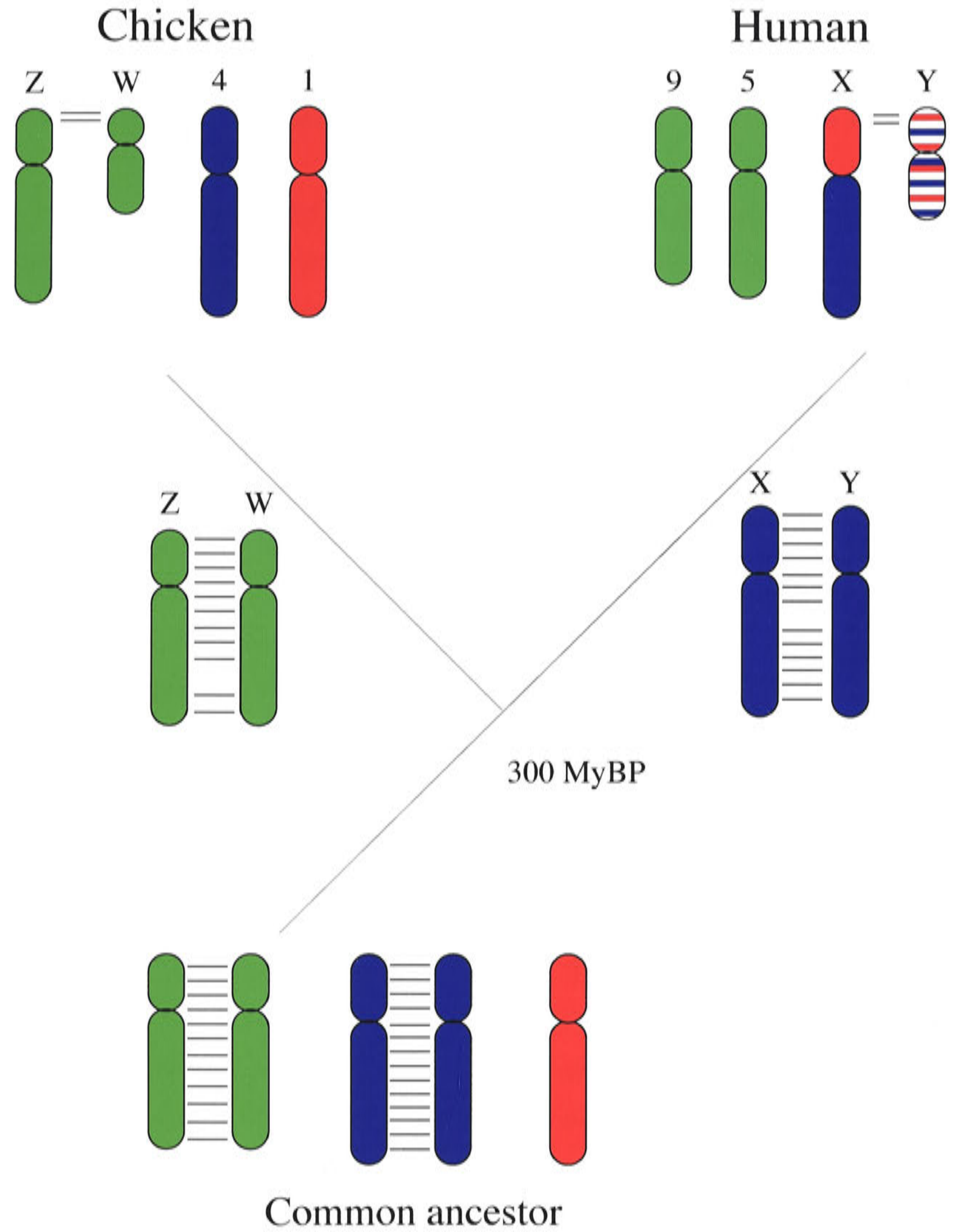


Figure 1.14: Eolution of the bird and eutherian sex chromosomes. The bird and eutherian common anncestor had one pair of autosomes destined to become the mammalian X and Y, and another pair destined to become the bird Z and W. The bird Z and W are represented in human by parts of chromosomes 5 and 9. The Eutherian X and Y are represented in chicken by parts of chromosomes 1 and 4.

proportion of SRR genes), or whether these genes were recruited from autosomes to a selfish X chromosome, or whether they acquired SRR functions after the loss of their Y homologues can be answered using comparative techniques (Graves and Delbridge, 2001).

The hypothesis that the proto-X and -Y chromosomes (and the recent additions) were chosen because they contained SRR genes, can be tested by mapping SRR genes, and studying their expression patterns in distantly related vertebrates. If the original autosome was chosen because it contained SRR genes, SRR genes in human XCR should cluster with other mammalian X-linked genes on chicken chromosomes 4, and still retain their original germ cell-specific expression. If the recent additions were also chosen because they contained SRR genes, then SRR genes in the human XAR should cluster to the same location as other XAR genes in wallaby (chromosome 5p and 1p) and chicken (chromosome 1), and also retain germ-cell-specific expression. On the other hand, if SRR genes were recruited from autosomes to a selfish X, germ cell-specific SRR genes should be scattered throughout the chicken genome, and XAR SRR genes should be scattered throughout the wallaby genome. Finally, if genes on the proto-X and -Y chromosomes evolved new functions as SRR genes after they lost their Y partners, they should all still cluster at the same positions in chicken and wallaby but not show germ-cell-specific expression (Graves and Delbridge, 2001).

The human X chromosome also contains an increased concentration of genes involved in X-linked mental retardation (XLMR). The most recent count showed that there are 202 XLMR syndromes (Chiurazzi *et al.*, 2001) out of a total of 958 entries listed in the Online Mendelian Inheritance in Man database (Zechner *et al.*, 2001). This implies that ~21.1% of mental retardation syndromes are X-linked, although the X chromosome only contains 3.75% of genes in the genome.

It is hard to understand why intelligence should be a sex specific (male-advantageous) character. Zechner (2001) explained this increased frequency of intelligence genes on the X chromosome as being a result of sexual selection. Females selected higher intelligence in males as an indicator of a potential mates genetic fitness. The hemizygous status of the X chromosome in males allows alleles on it to be open to selection (in males) long before they would reach a frequency high enough for selection to act upon an autosomal recessive locus in a homozygote.

Interestingly, about one third of X-linked SRR conditions are associated with mental retardation, and one third of XLMR syndromes are associated with sex determination or gonad development problems (Graves and Delbridge, 2001). Wilda *et al.* (2000) suggested that the same sets of genes could have been recruited for function in the brain, testis and placenta because these are the organs most responsible for human speciation. Alternatively, the same ubiquitously expressed genes may lend themselves to evolution of a specialised function, and have been independently selected for a male specific function, and for increased brain function.

The X, along with chromosomes 17 and 19, also has a disproportionate number of genes that are expressed in skeletal muscle (Pallavicini *et al.*, 1997). Interestingly, there are few tumour suppressor genes located on the X chromosome, probably due to selection against X-linked deleterious recessive mutations in hemizygous males.

Currently, it is unknown whether the proto-sex chromosomes were chosen for their gene content, or whether the gene content was shaped by their special roles. Comparative mapping in marsupials and birds will help answer this question.

1.4.5 Y variation

After the proto-X and -Y were formed, and the Y chromosome became recombinationally isolated from the X, the Y began to vary and waste away. Two processes must be distinguished; the rate of variation and the change in allele frequencies, which lead to Y degradation.

The rate of variation on the Y chromosome is greater than those of the X chromosome and autosomes. Studies of the X-Y shared gene *ZFX/Y* have demonstrated that there is a higher synonymous nucleotide substitution rate at the Y-linked locus than the X-linked locus (Shimmin *et al.*, 1993; Shimmin *et al.*, 1994). Unexpectedly, *SRY*, the sex determining gene, shows a low level of conservation between mammalian species, and poor homology to its X-borne partner, *SOX3*. Conversely, *SOX3* is highly conserved between species. The divergence of *SRY* is therefore due to its location on the Y. Interspecies comparison of the human Y-borne *DAZ* gene, which was transposed from *DAZL* on chromosome 3 after the spit between Old World and New World monkeys, showed that *DAZ* accumulated more mutations than the autosomal *DAZL*.

This was deduced by the wide variation of *DAZ* sequences, but the lack of *DAZL* variation between species (Makova and Li, 2002).

This high interspecific variation, therefore, reflects the evolution of the Y chromosome. Several hypotheses have been suggested to explain this variation. They fall into two categories: Y-driven and male-driven. Under the heading of Y-driven variation is the hypothesis that the Y chromosome is inherently unstable and changes rapidly because it consists of a high proportion of repetitive elements. These repetitive elements permit a high number of intrachromosomal recombination events, leading to inversions and deletions. Because only males posses a Y, the Y chromosome also has a smaller population size than other chromosomes; one quarter the amount for that of an autosome, and one third for that of the X chromosome. With a small population, accelerated drift could therefore account for high interspecific variation (Charlesworth and Charlesworth, 2000). Male-driven variation, on the other hand, relies on observations that the testis and sperm are dangerous environments. The number of germline divisions in spermatogenesis is greater than that of oogenesis; the Y chromosome would therefore be exposed to increased probabilities of replication dependent mutation (Ellegren, 2002). Sperm is an oxidative environment that can cause DNA damage, and with no DNA repair mechanism, *de novo* mutation can occur in the sperm and be passed to the son (Aitken and Graves, 2002).

The increased variation of *DAZ* is hypothesized to be the result of *DAZ* suffering more mutations, because it is confined to the germline in the testis, whereas a *DAZL* locus spends only half of its time in the testis. This indicated a significant male bias in mutation rate. Since a large proportion of mutations are due to replication errors, this bias could be due to the fact that the number of meiotic divisions in spermatogenesis is greater than that of oogenesis (reviewed by Ellegren, 2002). Because spermatogonial cells are continually dividing in men, it would be expected that the male mutation rate to increase with age (Crow, 2000).

The argument that the difference in mutation rate is male driven rather than Y-driven, and due to differences in male and female germlines is supported by studies in birds. In birds the male is ZZ, and the female has a Z and a smaller heterochromatic W. As the W chromosome is comparable to the Y chromosome, it was assumed that it would be highly degraded and poorly conserved between species, as is the mammalian

Y chromosome. However, it was found that in birds the *CAD-Z* gene has evolved faster than its W-linked homologue *CAD-W* (Ellegren and Fridolfsson, 1997). Ellegren accounted for this by observing that because the female is the heterogametic sex, the Z chromosome spends two thirds of its time in the male germline, whereas the W chromosome spends no time in the male germline. Hence, the W is never exposed to the dangers of the testis. The W still acquires mutations that cannot be repaired via recombination and wastes away, although it does not acquire the replication errors as rapidly as the Z chromosome, and mammalian Y chromosome, which spend a disproportionate amount of time in the testis.

Further evidence supporting the dangerous testis hypothesis is that nearly all *de novo* autosomal mutations in humans are of paternal origin. Sporadic mutations with a dominant effect have been investigated to determine the parent of origin. Of 40 sporadic cases of achondroplasia, all were of paternal origin. Similarly, of 57 cases of Apert syndrome, 25 of multiple endocrine neoplasia 2B, 10 of MEN 2A and 22 of Crouzon and Pfeiffer syndromes, all new mutations were of paternal origin. These six conditions are caused by single base substitutions in one of three genes (reviewed by Crow, 2000).

It has recently been suggested that the sperm is susceptible to oxidative stress. Because spermatozoa are deficient in antioxidants and DNA repair mechanisms, the risk of DNA fragmentation is higher. DNA fragmentation would predispose the cell to mutagenic change after fertilisation that would become fixed in the embryo by aberrant recombination. The Y chromosome is highly susceptible to such illegitimate recombination, and therefore fixation of mutation, because of its high frequency of repetitive elements (Aitken and Graves, 2002). Over long periods of time, the mean fitness of the Y chromosome will be reduced relative to the X chromosome.

Thus genes on the Y chromosome diverge rapidly from their X-borne partners or autosomal progenitors, producing a large amount of variation between species. Surprisingly, however, there is little Y chromosome variation within species. For instance, although *ZFY* shows considerable variation between primate species, there are low levels of polymorphism within humans, and also within other primate species. Interspecific comparisons showed variable sites distributed throughout the region of *ZFY* studied, so selection of *ZFY* could therefore not account for the lack of intraspecific variation. The low level of polymorphism is therefore ascribed to selective sweeps, a

recent origin for primate Y chromosomes or recurrent male population bottlenecks (Dorit *et al.*, 1995; Burrows and Ryder, 1997).

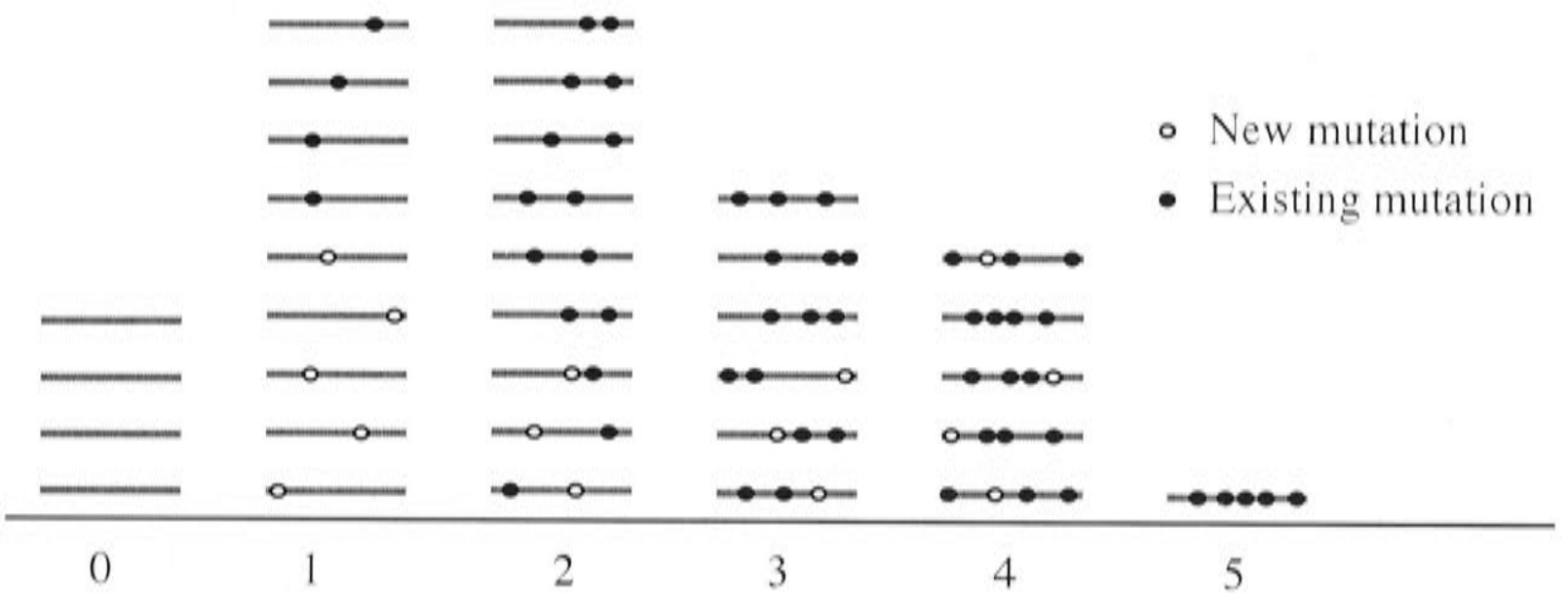
1.4.6 Y degradation

The means by which the Y chromosome degraded, and the forces that drove this degradation have been debated for decades. Several theories have been put forward, all of which depend on the lack of recombination between the X and the Y. Hypotheses include Muller's ratchet, background selection and genetic hitchhiking.

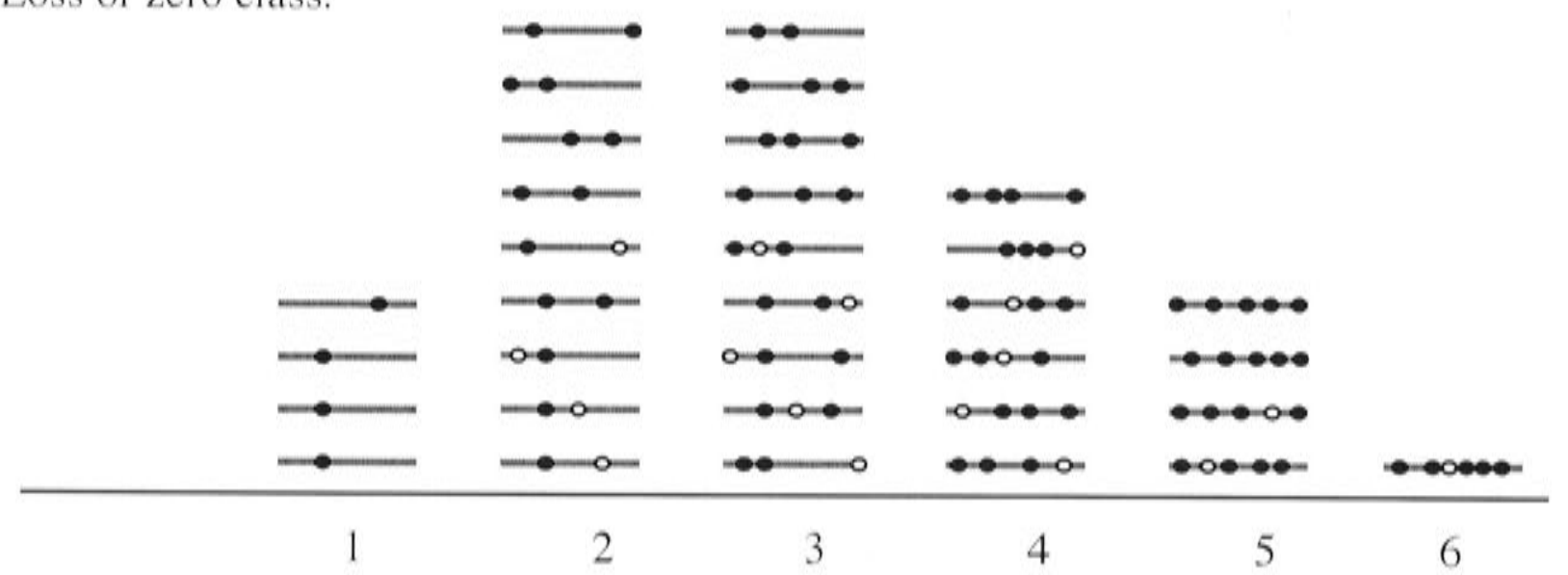
Muller's ratchet describes the random loss of the least mutated class of Y chromosome from a population of finite size (Muller, 1964). This least mutated class cannot be regenerated without rare back mutations, because of the lack of recombination between the X and Y chromosomes. If there is a distribution of Y chromosomes in a population that have different numbers of mutations from none to several, then the least-loaded (zero mutations) class of Y chromosomes can accidentally be lost. The next least loaded class of Y chromosome has one mutation. This class consists of a set of haplotypes, all with one mutation, and if they are selectively equivalent, fixation of one of these haplotypes will occur randomly in a finite population (Figure 1.15) (Charlesworth and Charlesworth, 1997). If this last least-loaded Y chromosome is then lost from the population, Y chromosomes with two mutations become the least-loaded in the population. This ratchet-like process of irreversible losses of least-loaded Y chromosomes continues, leaving Y chromosomes with more and more mutations.

The speed with which the ratchet turns is important for its plausibility in explaining the rapid degeneration of the Y chromosome. All things being equal, the ratchet (as for all genetic drift) should turn faster as population size gets smaller. For mammalian populations, effective population sizes are often only in the tens of thousands, and only half of these (males) have a single Y. The ratchet might therefore be plausible (reviewed by Charlesworth and Charlesworth, 2000). However, for certain insect species (*e.g. Drosophila*), effective population size could be in the millions, bringing into question the effectiveness of the ratchet. However, if a high rate of mutation generates many mutations with small effect on fitness, there will be few mutation free Y chromosomes in the population and the ratchet is plausible (Gordo and Charlesworth, 2000).

a) Some mutation free Y chromosomes present in population.



b) Loss of zero class.



c) Fixation within class with one mutation, and high frequency of this allele in classes with more then one mutation.

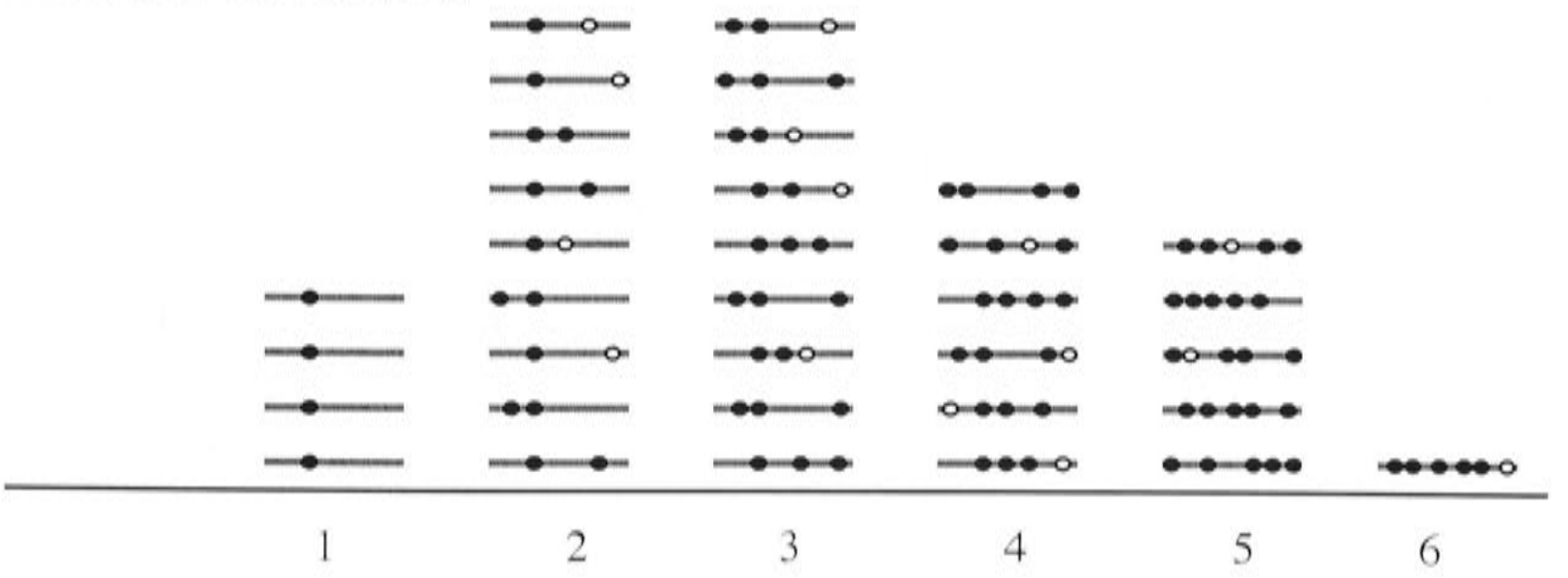


Figure 1.15: Distributions of mutations on Y chromosomes. There are three stages (a-c) in each turn of the Muller's ratchet process. Each line represents a Y chromosome with mutations that are either new (open circle) or inherited from the previous generations (filled circles).

Two hypotheses (genetic hitchhiking and background selection) invoke events occurring within a non-recombining region of the genome (in this case the differentiated region of the Y chromosome). Lack of recombination results in reduced efficiency of natural selection on a single locus. Therefore, selection acts on the overall fitness of a Y chromosome rather than the fitness of an allele at a given locus. Genetic hitchhiking relies on a beneficial mutation that arises on a Y chromosome with many deleterious mutations at other loci. If the benefit of the new mutation outweighs the cost of the deleterious mutations, the haplotype (including deleterious mutations) could be dragged to fixation in the population by strong selection of the beneficial mutation and replace all other haplotypes in a selective sweep. Successive selective sweeps would lead to the fixation of deleterious alleles at many Y-linked loci (Rice, 1987), and reduced intraspecific variation of Y chromosomes, as observed for ZFY (Section 1.4.5).

A converse idea is that of background selection. This hypothesis states that a new mutation has a chance of fixation in a population only when it arises on a Y chromosome that is free of deleterious mutations. Because of the lack of recombination between the X and Y chromosomes, Y chromosomes with deleterious mutations will be rapidly lost from the population, dragging with them new mutations that are beneficial. Background selection can operate with low selection coefficients and large population sizes in which the ratchet will not be effective, resulting in reduced Y chromosome fitness in very large populations (Charlesworth, 1996a). Background selection reduces effective population size of Y chromosomes, and, as a result, slightly deleterious mutations can enjoy a raised chance of fixation in a population due to drift.

Whether background selection, or hitch-hiking, occurs depends on whether the benefit of the new mutation outweighs the effect of the deleterious mutation/s (Charlesworth, 1996a).

With the degradation of the Y chromosome, and the loss of gene function, a dose difference of genes on the X is apparent between males and females. In mammals X-inactivation evolved to equalise this difference.

1.4.7 Evolution of X chromosome inactivation

The loss of function of Y linked alleles leads to a dosage difference of X linked genes between males and females. The evolution of X chromosome inactivation (XCI)

equalized this dosage difference. It is generally thought that XCI evolved as a consequence of selection for dosage equivalence, although it remains possible that variation in the spread of X-inactivation drove compensatory Y degradation (Graves *et al.*, 1998). Dosage compensation is clearly very important, since dosage compensation mechanisms have evolved independently in different species (Section 1.2.3).

Upregulation of the X chromosome in *C. elegans* and mammals may not initially have been sex specific. Increased expression from the X would be beneficial for males, but detrimental to females, who would over-express genes on the X. Therefore, upregulation of the X chromosome in males would, in turn, cause selection for the down regulation of one or both X chromosome in females.

Originally it was proposed that X-inactivation evolved on a regional or chromosomal basis (Lyon, 1974). Lahn and Page (1999) divided the human X chromosome into four evolutionary strata, defined according to the time that they became recombinationally isolated from the Y chromosome (Section 1.4.2). The genes within a stratum would be recruited into the X inactivation system after the loss of complementary functions of the Y partners.

The evolution of XCI can be gauged by comparing the inactivation status of individual X-linked genes. This can be done by comparing methylation patterns across species, since there is a perfect correlation between CpG-island methylation and X-inactivation status in eutherians. Inactivated alleles are methylated, whereas their partners on the active X are undermethylated at this site (Tribioli *et al.*, 1992). For instance, the human X-borne gene *ALD* is known to be subject to inactivation, and showed methylation in females but not males of 18 eutherian species representing nine orders (Jegalian and Page, 1998). Jegalian and Page (1998) also investigated the methylation status (hence X-inactivation status) of X-borne genes with Y homologues (*ZFX*, *RPS4X* and *SMCX*) in the same set of species. *ZFX* was shown to be differentially methylated in myomorph rodents (mouse, rat, hamster and lemming), whereas no methylation was found in the other species studied. *RPS4X* was shown to be differentially methylated in the eight non-primate species tested, but not in primates. *SMCX* was differentially methylated in five out of eleven eutherian orders, not including primates or myomorphs.

These patterns of methylation (hence X-inactivation) exactly mirror the loss of functionally equivalent alleles on the Y. For instance, *ZFX* was not inactivated in the ancestral eutherian but became inactive in myomorph rodents, coincident with the loss of complementary *ZFY* function. *ZFX* has a strongly conserved, and ubiquitously expressed Y homologue in all mammals, except for myomorph rodents, in which *Zfy* sequence has diverged considerably from that of *Zfx*, and become testis specific. Similarly, *RPS4X* is not subject to X-inactivation in primates, where it has retained a Y-linked homologue. *SMCX* is subject to X-inactivation only in guinea pig and cattle, species in which they have no detectable Y-linked homologue. Thus species that have a Y-linked homologue have no need for dosage compensation and therefore do not show X-inactivation.

These results suggest that X-borne genes become subject to inactivation when they lose an active Y-borne partner that complements function. However, in rabbit and goat *SMCX* is subject to X-inactivation even though there is a Y-borne homologue. Although a male-specific rabbit *SMCY* was detected by Southern blotting, it could not be isolated from a cDNA library, suggesting that *SMCY* is not expressed, at least not at high levels.

There are several exceptions of genes that escape X-inactivation, although there is no functional Y partner. For instance, *UBE1X* is not inactive in humans, but there is no Y homologue. Also human *STS* and *KAL1* escape X-inactivation (Disteche, 1995), although the Y-borne partner of each is a non-functional pseudogene. It is likely that, for these X-borne genes, loss of a functional Y-linked locus occurred recently in human evolution, and selective pressure has not yet been strong enough to induce their inclusion into the X-inactivation system. A recent study has shown that 34 genes on the human X chromosome (three in the PAR) escaped X-inactivation. Of these, 31 were located on Xp, suggesting that the two arms of the human X chromosome are epigenetically distinct (Carrel *et al.*, 1999), as would be expected from their separate evolutionary origin (XCR and XAR).

Evidence for the change in X-inactivation status comes from the observation that some genes are polymorphic for inactivation. For instance, *REP1* escaped X-inactivation in cell lines of some human females, and was subject to X-inactivation in cell lines of other females. *REP1* could therefore be in transition from an active gene, to one that is subject to X-inactivation (Carrel and Willard, 1999).

Variation of X-inactivation can also occur within an animal. There is evidence to suggest that, in mouse, some genes are subject to X-inactivation in some tissue types, and exempt from X-inactivation in other tissue types. *Smcx*, *Usp9x*, *Ube1x*, *Eif2s3x*, *Utx* and *Dbx* were all expressed at higher levels in the female brain than the male brain irrespective of their X-inactivation status. X-inactivation of a gene could therefore be specific to tissue type or developmental stages (Xu *et al.*, 2002). These examples of variation of X-inactivation within a population, or even an individual, show that recruitment into the XCI system can be very gradual and require several steps.

Y degradation and X inactivation leads to the halving of gene dosage in males. Dosage compensation must therefore be accompanied by an increase of expression from the single X chromosome so that dosage is maintained in relation to autosomal genes. This is seen most obviously in the upregulation of the X chromosome in male *Drosophila*. There is some evidence for upregulation of an X-borne mammalian gene in both sexes. In *M. spretus*, *Clc4* is X-linked, whereas in *M. musculus* it is autosomal. In backcross individuals from species hybrids that have both loci, the X-linked copy is twice as active as the autosomal copy, indicating that the copy on the X chromosome acquired enhanced gene activity (Adler *et al.*, 1997).

Finally, if dosage compensation evolved on a gene-by-gene or regional basis, how were they recruited into a chromosome wide, X-inactivation system requiring *trans*-acting machinery to alter expression in one sex, and *cis*-acting sequence to attract that machinery to the X chromosome to be inactivated? Gene-by-gene evolution of dosage compensation may be a result of accumulation of local *cis*-acting elements (LINE elements) around genes newly incorporated into the X-inactivation system. This resulted in the chromosome wide mechanisms we see today (reviewed in (Marin *et al.*, 2000).

Thus XCI evolved to equalise dose of X-borne genes in male and females when the Y-borne partner lost its complementary function. There are several regions on the human X chromosome that are exempt from inactivation because the Y allele is still functional. One of these is the pseudoautosomal region.

1.4.8 Evolution of the pseudoautosomal regions

The hypothesis that the Y chromosome is a degraded relic of the original proto-X and -Y, implies that the pseudoautosomal region is a small region that remains from once extensive recombination between the two chromosomes. The PAR genes shared between species, and genes near the PAR, suggested that there was a larger ancestral eutherian PAR that has been reduced in size to varying degrees in different species.

The human PAR1 genes *PGPL*, *DHRSXY*, *CSF2RA* and *IL3RA* are pseudoautosomal in other primates and carnivores, but autosomal or absent in mouse (Gianfrancesco *et al.*, 2001; Disteche *et al.*, 1992; Milatovich *et al.*, 1993). It is more parsimonious to conclude that *PGPL*, *DHRSXY*, *CSF2RA* and *IL3RA* in mouse were independently lost from the sex chromosomes, rather than additions to the PAR in all the others groups (Figure 1.16).

There is evidence that the ancestral PAR was once larger than human PAR1. The PARs of ungulates, carnivores, primates and mice all show homology with the differentiated region of the human X proximal to PAR1 (Figure 1.2). All of these genes together probably represent a larger ancestral PAR in a common ancestor before the eutherian radiation 80MyrBP. This larger ancestral PAR probably included human X-Y shared genes such as *ZFX/Y* and *AMELX/Y* along with genes such as *STS*, *KAL1* and *PRKX* that have Y-borne pseudogenes (Figure 1.16). All of the X-linked copies of these are located close to the human PAR1 and remain in the same order on the mouse X chromosome.

This larger ancestral PAR explains why the pairing region of ungulate and carnivore sex chromosomes is larger than that of humans. Conversely, the PAR of laboratory mice is considerably smaller (~700 kb) (Perry *et al.*, 2001) because it has lost genes to autosomes.

Mapping of three ancestral PAR genes (*CSF2RA*, *STS* and *ANT3*) to autosomes in marsupials showed that this region was not retained from the original proto-X and Y, but was part of the region to added eutherian sex chromosomes. Mapping two of these genes (*CSF2RA* and *STS*) to tammar wallaby chromosome 5p indicated that they were part of the same addition that made up most of the XAR. All but one of the genes in the larger ancestral PAR represents Lahn and Page's (1999) Stratum 4 of the human X

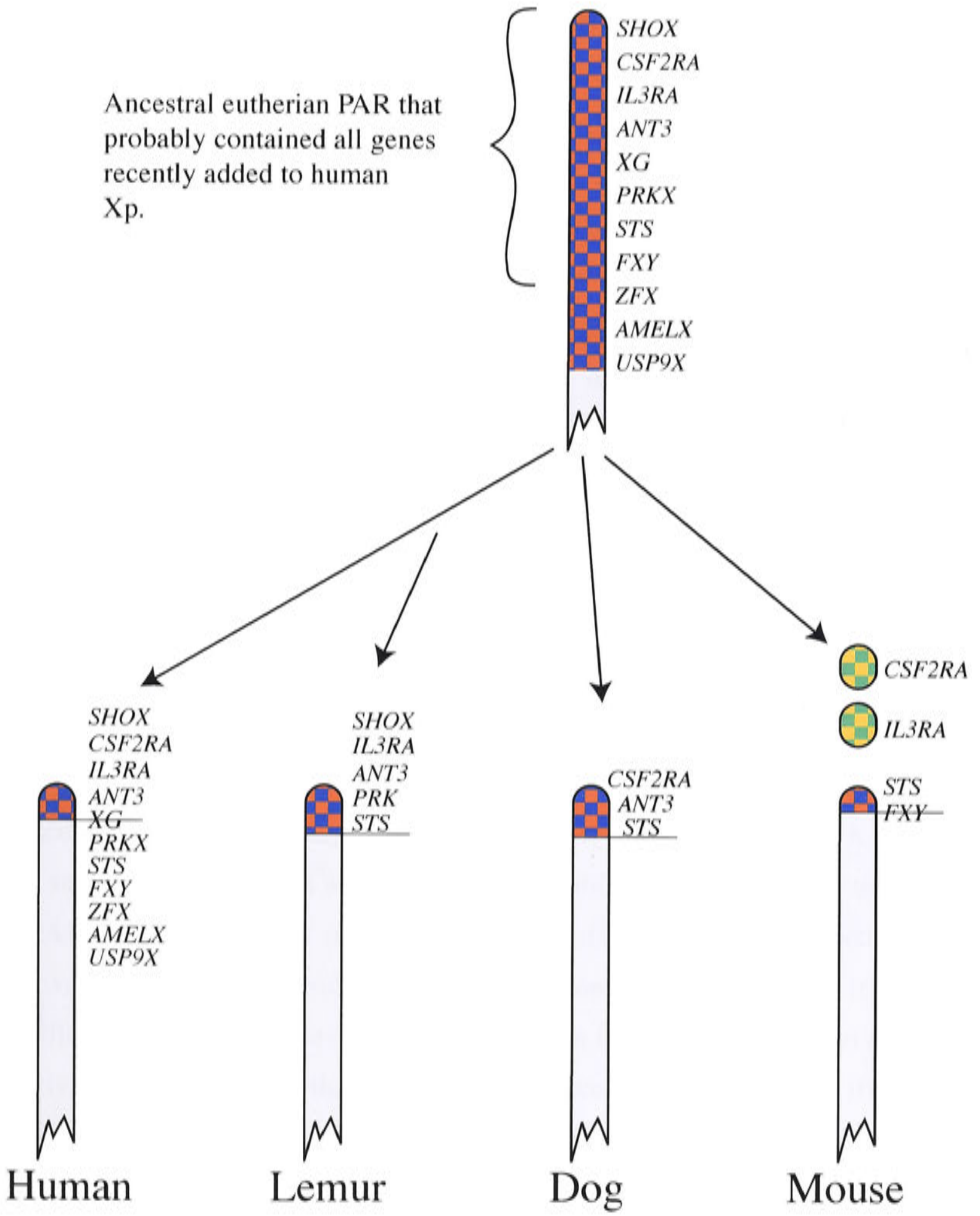


Figure 1.16: Evolution of human PAR1 from a larger ancestral PAR that included recently added genes from human Xp (i.e. *ZFX*, *AMELX* and *USP9X*) that no longer reside in the PAR of any eutherian studied. Indeed all genes on the X chromosome would once have been part an X-Y recombining region.

chromosome. The loss of genes from this ancestral PAR was, therefore, due to a single rearrangement in the human lineage. *ZFX* in Stratum 3 may have been lost previously from a more ancient mammalian PAR that would have contained other Stratum 3 genes. Thus, PAR1 is a relic of recombination left over from previous additions to the sex chromosomes (Graves, 1995), and is maintained because of the necessity for pairing between the X and Y chromosomes at male meiosis.

Most mammal sex chromosomes pair over a single PAR. However, in human there is a second PAR (PAR2) at the tips of the long arms of the X and Y. PAR2 was created more recently in human evolution as a result of a repeat sequence derived X-Y translocation (Kvaloy *et al.*, 1994). The evolution of PAR2 does not appear to be as simple as that of PAR1.

PAR2 is divided into two transcriptionally distinct zones, zone 1 and zone 2. In the proximal 295Kb, the modal CG content is 34.5% and rises sharply in the distal 35Kb to >51%. This difference defines zones 1 and 2. Zone 1 contains 67% LINE sequence in comparison to zone 2, which contains only 29%. Sequence on the PAR 2 boundary suggests that zone 1 may have been added to the X by an illegitimate recombination event between LINE sequences (Kvaloy *et al.*, 1994).

Zone 1 contains the genes *SYBL1* and *HSPRY3*, which are subject to X inactivation. This suggests the two genes were subject to standard X-inactivation prior to the creation of PAR2. Uniquely, both of these genes are also subject to Y-inactivation. The discovery that *SYBL1* was subject to Y-inactivation first suggested that it was a gene-specific phenomenon. However, the demonstration that *HSPRY3* was also subject to Y-inactivation demonstrated that Y-inactivation occurs across at least a 100Kb region, rather than being confined to a specific gene. In the absence of a Y-inactivation centre, no simple sequence can determine inactivation of these genes as they function normally on the X. It has been suggested that the neighbouring heterochromatin on the Y chromosome may inhibit the access of transcription factors to zone 1 (D'Esposito *et al.*, 1996) or could promote inactivation via methylation of *SYBL1* promoter regions (Huber *et al.*, 1999).

Zone 2 shows a further division into two subregions (Ciccodicola *et al.*, 2000), one containing *IL9R* and the other containing *CXYorf1*. It appears that these subregions were added separately during human evolution, as *IL9R* and *CXYorf1* are physically separate

in all other mammals in which they have been mapped (Ciccodicola *et al.*, 2000). The final step required for the creation of PAR 2 is the translocation from the X to the Y by another illegitimate LINE mediated translocation (Kvaloy *et al.*, 1994).

However, studies in marsupials have changed the interpretation of PAR2 evolution. *SYBL1* must have been part of the original conserved eutherian X chromosome because it mapped to the tammar wallaby X chromosome, whereas *HSPRY3* did not. *HSPRY3* and *CXYorf1* co-located on tammar wallaby chromosome 3 (Charchar *et al.*, submitted), suggesting that these gene were added together after marsupial divergence but prior to the eutherian radiation. This addition was followed by an inversion that placed *SYBL1* distal to *HSPRY3* but proximal to *CXYorf1*. *HSPRY3* therefore soon became subject to X-inactivation. The fourth gene, *IL9R*, was located on chromosome 1 in tammar wallaby, so represented an independent addition. A second inversion made *CXYorf1* distal to *IL2R* (Figure 1.17).

Gene content and size of the PAR can vary between species, indicating that these factors are not critical to the crucial X-Y paring function. Human PAR1 is a relic of a larger ancestral PAR, whereas PAR2 was created more recently in human evolution. Genes that reside in the PAR are identical on the X and Y chromosome, and therefore have the same function. However, genes in the human NRY have changed considerably and have different function to their X-borne partners. Some genes on the human Y chromosome have no X-borne homologues. Instead, they show homology to autosomal genes.

1.4.8 Evolution of genes on the Y

The Y chromosome is unique because of its preponderance of genes with a function in male sex determination and reproduction. This 'functional coherence" is likely to have resulted from selection for acquisition of male advantage functions, and from loss of genes with other functions.

Genes on the human Y chromosome have been divided into two classes, Class I and Class II (Section 1.2.1.3). Class I genes have are single copy, ubiquitously expressed and have X-borne homologues from which they evolved. Class II genes are multi-copy, testis specific and have no X homologues; they arrived on the Y chromosome as a result of a transposition event.

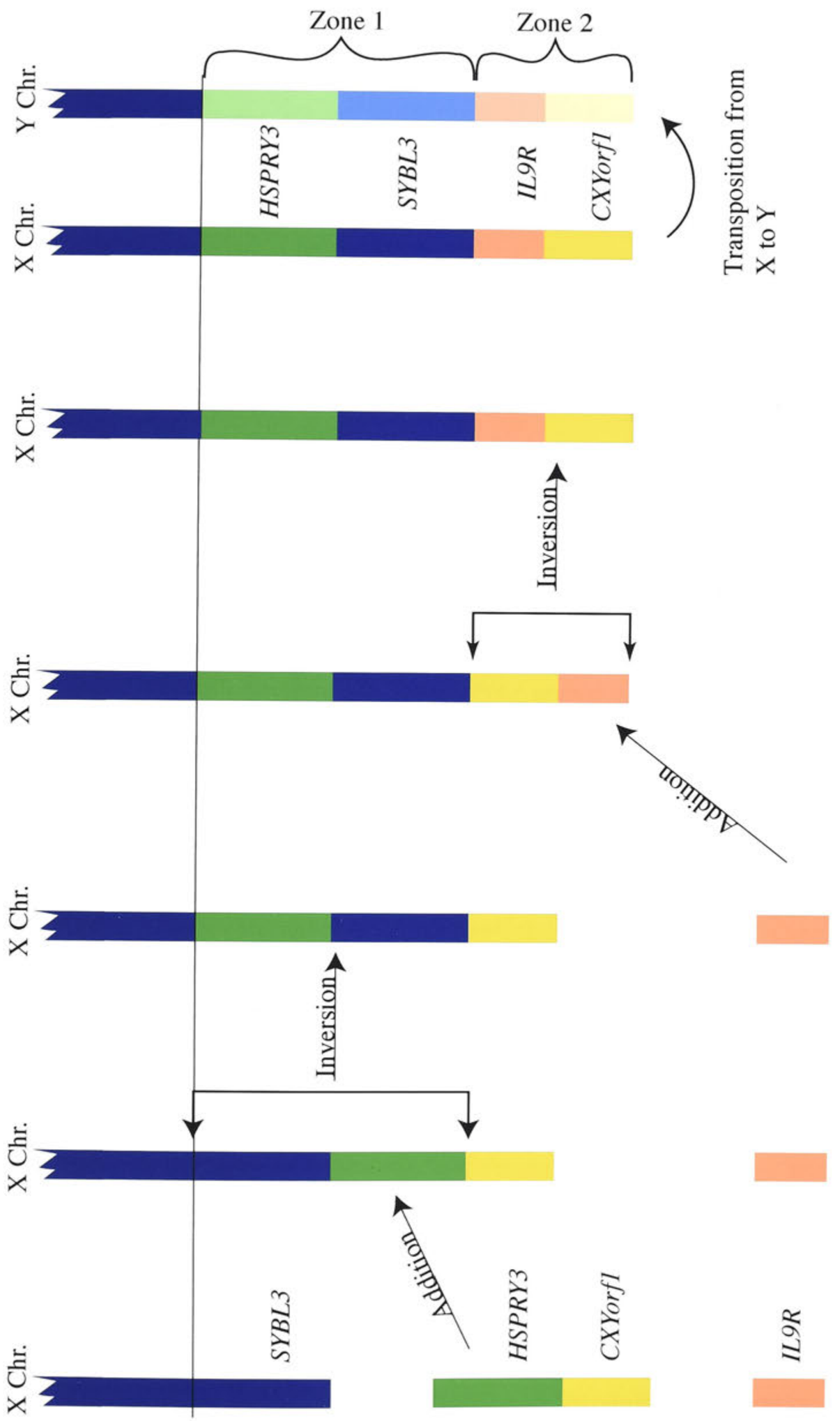


Figure 1.17: Steps in the evolution of PAR 2

Of particular interest are genes that have differentiated from their ubiquitously expressed X-borne partners (Class I) and gained male specific functions. For instance, the testis-specific spermatogenesis gene *RBMY* (Section 1.3.2) evolved from the ubiquitously expressed *RBMX*, which may be involved in brain function. *RBMY* has been amplified on the Y chromosome, and all but two of the 30 copies have become inactive pseudogenes. In the same way, *SRY* evolved from *SOX3* to become the dominant mammalian sex-determining gene (Section 1.3.1). The testis specific *Dby* and *Usp9y* evolved from their ubiquitously expressed X-borne homologues in mouse, but in human, they remain ubiquitously expressed, although both have been shown to play a role in human spermatogenesis (Section 1.3.2).

A different class of genes on the Y chromosome are those that arrived there via transposition or retrotransposition (Class II). The intronless *CDY* was retrotransposed to the Y from *CDYL* on chromosome 13. Alternatively, the intron-containing *DAZ* must have been transposed to the Y chromosome from *DAZLA* (located on human chromosome 3) 30-50MyBP (Saxena *et al.*, 1996). *DAZLA* appears to have a conserved function in germ cell development, which *DAZ* continued to perform after it arrived on the Y chromosome.

Selection for acquisition of male advantage genes on the human Y chromosome, resulted in many genes located on the human Y chromosome that play roles in male sex determination and reproduction. These genes either evolved from X-borne homologues, or jumped to the Y chromosome from an autosome. Genes on the Y chromosome that did not acquire male advantage functions were lost from it.

1.5 The present study

The aim of this study was to determine the origin of the mammalian Y chromosome, and in particular, the human Y chromosome. This was achieved by studying evolutionary intermediates (marsupials, monotremes and birds). Species from these groups are valuable because they are distantly related to humans and mice. Marsupials diverged from the human lineage 130MyBP, whereas monotremes and birds diverged from the human lineage 170MyBP and 310MyBP respectively. The species

studied were *Macropus eugenii* (tammar wallaby), *Ornithorynchus anatinus* (platypus) and *Gallus gallus* (chicken).

I therefore isolated, characterised and mapped the orthologues of genes from the human X and Y chromosomes in distantly related species. Initially, tammar wallaby orthologues of human Y-borne genes (some with X homologues and some Y-specific) were examined by cross-species Southern blot analysis (Chapter 3). Class I (X/Y shared) genes were likely to be much easier than Class II genes (Y-specific) to detect in marsupials because the X homologues are well conserved between species, unlike Y-specific genes. Hybridisation would indicate that a wallaby orthologue existed. Male specific bands would indicate that the wallaby orthologue was located on the Y chromosome, whereas male-female dosage differences would suggest a location on the X chromosome.

Mapping of the genes isolated, demonstrated (along with previous results) that the human Y chromosome retained very little of the ancient Y chromosome, and was largely derived from a single autosomal region added to the sex chromosomes after the divergence of marsupials.

Genes conserved on the sex chromosomes of marsupials and humans were then mapped in monotremes (Chapter 4). The results were not consistent with the hypothesis that the monotreme, as well as marsupial, sex chromosomes represent the ancestral mammalian sex chromosomes. There seemed to be a block of genes missing from the monotreme X, including some thought to be critical to male-specific functions in all mammals.

To determine whether these genes were added to the therian sex chromosomes after the divergence of monotremes, or whether these genes were lost from the sex chromosomes in the monotreme lineage, some were mapped in chicken (Chapter 5). Using chicken as an outgroup, I could deduce that the missing genes were not located on the proto-X and -Y chromosomes in the mammal ancestor, and that they were added later to the sex chromosomes of the therian ancestor.

The mapping of strategically chosen genes in distantly related mammals, and in birds, provided insights into the evolution of the mammalian, and specifically the human, sex chromosome (Chapter 6). These experiments also provided a measure of the time at which genes on the eutherian Y chromosome acquired their male-specific

functions. The organization, function and the similarities and differences of the roles genes play on alternative mammalian Y chromosomes were considered in light of their evolution.

Chapter 2: Materials and Methods

2.1 Probes

Probes used for Southern blot analysis and library screening are listed in table 2.1.

Table 2.1: Probes used in this study

Probe	Size	Source	Type	Reference
PRY	710bp	Human	PCR product	
XKRY	780bp	Human	PCR product	
BPY2	320bp	Human	cDNA	Lahn and Page, 1997
VCX	400bp	Human	cDNA	Lahn and Page, 1997
PCDHX	700bp	Human	PCR product	
RPS4X	440bp	*Monodelphis domestica*	PCR product	
EIF1AX	500bp	Human	cDNA	GenBannk Accession # NM 001412
TB4X	800bp	Human	cDNA	GenBannk Accession # NM 021109
ATRX	400bp	*Monodelphis domestica*	PCR product	Pask *et al.*, 2000
USP9X	5000bp	Mouse	cDNA	GenBannk Accession # NM 009481
DBX	4000bp	Mouse	cDNA	GenBannk Accession # NM 010028
SOX3	900bp	Tammar wallaby	PCR product	
G6PD		Human	cDNA	Persico *et al.*, 1986
RBMX	1000bp	Tammar wallaby	cDNA	Delbridge *et al.*, 1998

2.2 DNA manipulation and Analysis

2.2.1 Restriction Endonuclease Digestion

An aliquot of plasmid, phage or genomic DNA was digested with 5 units of the appropriate restriction enzyme/s (Roche), in 2µl of the appropriate buffer, in a total volume of 20µl for 2 hours at 37 C. This cut the DNA at specific sites either liberating insert DNA from a plasmid or generating fragments of differing size when phage or genomic DNA were digested.

2.2.2 Agarose Gel Electrophoresis

Electrophoresis was used to separated different sized fragments of DNA on a 0.8-1.6% agarose (Gibco BRL) gel, buffered with 0.5xTBE (0.045M Tris-borate (ICN;BDH), 0.001M EDTA (Ajax), pH8) run at a maximum of 6 volts/cm. To visualise bands under UV light, 0.2µg/ml ethidium bromide (Sigma) was added to the agarose. Gels were photographed with 667 Polaroid film. If the gel was post stained, it was submerged in dH_2O with 1µg/ml ethidium bromide and rocked for 15 minutes. Lambda DNA cleaved with *Hind*III (Roche) (λ *Hind*III) and phiX174 DNA cleaved with *Hae*III (Roche) were run as marker DNA. Bromophenol blue (Sigma) was used as a loading dye.

2.2.3 Preparation of Plasmid DNA

A single colony of *E. coli* containing the plasmid with the DNA insert was picked and grown overnight at 37 C in 5 ml of LB (1% NaCl, 0.5% yeast extract, 1% tryptone, pH 7) supplemented with 10µg/ml ampicillin. A 1ml to 4ml aliquot was taken from this overnight culture and mini-preparations were conducted using the Wizard Plus Miniprep DNA purification system (Promega) according to the manufacturer's instructions.

A 5µl aliquot of plasmid DNA containing insert was digested with the appropriate enzyme/s to liberate the insert (Section 2.2.1). Plasmid and insert were then separated on an agarose gel as described previously (Section 2.2.2).

2.2.4 Gel Purification of DNA

Gel purification was used to isolate insert DNA from a plasmid, PCR product or fragment of phage DNA for subcloning.

The insert band to be purified from the gel was excised and the DNA extracted from the agarose with GenElute Minus EtBr Spin Columns (Sigma) according to the manufacturer's instructions.

A 3µl aliquot of DNA was run on an agarose gel to confirm that there was no plasmid contamination and was visually compared to a know amount of marker DNA to approximately determine DNA concentration.

2.2.5 DNA Extraction

Collected tissue samples were snap frozen in liquid nitrogen and stored at –70°C. 100mg-1g of tissue was used for DNA extraction with the phenol/chloroform DNA extraction method (Sambrook *et al.*, 1989). DNA was stored in 1 x TE (10mM Tris, pH8/ 1mM EDTA, pH8) at 4°C.

2.2.6 DNA Precipitation

DNA precipitation was used to increase the concentration of DNA. Ice-cold ethanol (Selby-Biolab) was added to 2.5 times the original volume the DNA was suspended in, and 0.1 times the volume of sodium acetate was added. This was then incubated at -70 C for 2 hours and then spun at 15000 r.p.m. for 15 minutes. The supernatant was discarded, the pellet was washed with ice cold 70% ethanol and spun again for a further 5 minutes. The supernatant was again discarded; the pellet was dried and then resuspended in the appropriate volume of 1xTE pH8 to give the required concentration.

2.2.7 Sequencing

Sequencing reactions were conducted using the BigDye Terminator kit (Applied Biosystems International) according to the manufacturer's instructions in a Peltier Thermal Cycler-200 from MJ Research. Sequencing reactions were run by the Biomolecular Resource Facility, John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia, 2600.

Primers designed and used for sequencing are listed in appendix I.

2.2.8 Polymerase Chain Reaction (PCR)

Amplification from 100ng of genomic, plasmid or phage DNA templates by PCR was conducted in 50µl reactions in a Peltier Thermal Cycler-200 from MJ Research. 2.5 units of *Taq* DNA Polymerase (Roche) was incubated with the template DNA, 100ng of each primer, 0.2mM each of dATP, dCTP, dTTP, and dGTP, in 10mM Tris-HCl, pH

8.3/ 1.5mM $MgCl_2$/ 50mM KCl. Cycling conditions were determined by primer annealing temperature and the length of the amplification product.

Primers designed and used for PCR amplification are listed in appendix I.

2.3 Library and Southern blot construction and screening

2.3.1 Genomic Library Construction

High molecular weight genomic DNA was extracted from female chicken liver as described in section 2.1.6. Serial dilutions of the restriction endonuclease *Sau*3A were set up on ice with the chicken genomic DNA in the following way. 10µg of DNA was added to a 200µl reaction with 0.2 units of *Sau* A. 100µl of this reaction was transferred to another reaction containing 5µg of genomic DNA. This reaction was mixed and 100µl of it was added to yet another reaction containing 5µg of DNA. There were a total of 6 dilutions in this series. Each dilution reduced the amount of enzyme by half.

The reactions were incubated at 37°C for 1 hour, after which 4µl 0.5M EDTA was added to stop the reactions. They were then incubated at 65°C for 15 minutes to heat inactivate the restriction enzyme. 2µl of Alkaline Phosphatase (1unit/µl) was added to the reactions and then incubated for 30 minutes at 37°C, followed by 11µl of TNA (nitriloacetic acid) and incubation for 15 minutes at 65°C.

The DNA was then extracted thrice with phenol/chloroform/isoamyl alcohol (25:24:1) followed by two extractions with chloroform/isoamyl alcohol (24:1). The DNA was precipitated (Section 2.2.6) and resuspended in 10µl of TE. 2µl of each reaction was electrophoresed (Section 2.2.2) on a 0.4% agarose gel. The sample with most fragments in the size range of 10-20kb was chosen.

This sample was ligated into *Bam*HI digested λDASH®II arms (Stratagene) using 2400 NEB units (40 Weiss units) of T4 DNA ligase (New England BioLabs) according to the manufacturer's instructions. The recombinant bacteriophage were packaged using the Gigapack Gold packaging kit (Stratagene) and the library titred on a bacterial lawn of XL1-Blue MRA (P2). For screening, 120 000 bacteriophage were plated per 250x250mm square Petri dish on a lawn of XL1-Blue MRA (P2). This concentration was just high enough to give confluent lysis. The plaques were transferred to Hybond-N^+ membranes according to Sambrook *et al* (1989).

A male tammar wallaby genomic DNA library was constructed by Stephen Wilcox and available for use (Wilcox *et al.*, 1996). A male platypus genomic DNA library was constructed by Margaret Delbridge and available for use.

2.3.2 Genomic Southerns and DNA Transfer

Genomic Southern blots were prepared by digesting completely 10μg of *Macropus eugenii* (tammar wallaby) genomic DNA with *Eco*RI and *Hind*III (both male and female for each enzyme). These digestion reactions were run overnight on an agarose gel at 20 volt/cm. The gel was pretreated by submersing in 0.25M HCl (BDH) for 10 minutes to hydrolyse the fragments. The gel was rinsed in distilled water before transfer. A Southern blot was performed onto a Hybond-N^+ nylon membrane (Amersham) using 0.4M NaOH; after transfer the membrane was rinsed in 2xSSC. A Southern blot relies on capillary action to transfer DNA.

A vacuum blotter was used when pure genomic clones (isolated from genomic libraries) or subclones were digested with a restriction enzyme and run on a gel. The gel was pretreated in 0.25M HCl for 10 minutes. DNA was transferred using 0.4M NaOH from the gel to a Hybond-N^+ nylon membrane using a vacuum blotter. The Tran-vac vacuum blot was set up according to the manufacturer's instructions (Hoefer Scientific Instruments, model number; TE80). After DNA transfer was complete the membrane was washed briefly in 2xSSC and dried.

2.3.3 Radioactive Labelling of DNA

Random primer labelling was used to radioactively label probes. For probing genomic Southerns 100ng of the probe DNA (template) was labelled and 600ng of DNA was labelled for screening genomic libraries. In this system, following denaturation of the DNA, a second strand of DNA is synthesised incorporating labelled nucleotides by the Klenow enzyme from the annealing site of the random primers. The Megaprime labelling kit (Amersham) was used to achieve this. For the labelling reaction of probes, to be used to screen libraries, 50μCi of α-^{32}P-dCTP was added; for probes to be used for genomic Southerns 30μCi of α-^{32}P-dCTP was added.

The labelled probe was spun through a ProbeQuant G-50 Micro Column (Amersham Pharmacia Biotech) at 3000 r.p.m. for 2 minute to remove unincorporated radioactive nucleotides preventing non-specific binding of single nucleotides to Southern blots or library screens.

2.3.4 Hybridisation and Washing

The Southern blots, genomic libraries and plate lifts were prehybridised overnight at 55 C. 0.2 ml of Church buffer (0.5M Na_2HPO_4 (BDH), 1mM EDTA, 7% SDS) was used per square centimetre of membrane as a prehybridisation solution to block non-specific binding of the radioactively labelled probe. The labelled probe was added directly to the prehybridisation solution. A concentration of 10-20ng of radioactively labelled probe per millilitre of hybridisation solution was required. The membrane was then hybridised for two nights at 55 C.

After hybridisation was completed, the hybridisation solution was removed and the membrane was washed twice in 2xSSC (0.3M NaCl, 0.3M sodium citrate (BDH)), 0.1% SDS at the appropriate temperature for 15 minutes (For genomic Southern blots the temperature varied, for library screens all washes were conducted at 55 C). After the washing was completed the membrane was sealed in plastic wrap (Glad Wrap) to keep it moist and protect the film from moisture. Southern blots were exposed to Fugi X-ray film whereas library screens and subsequent plate lifts were exposed to Kodak X-ray film at -70 C for the appropriate time (exposure of Southern blots varied greatly, library screens were exposed overnight) and then developed with Ilford developer according to the manufacturer's instructions.

2.3.5 Preparation of Phage Competent Cells

Bacterial cells were treated chemically causing them to become phage competent allowing the amplification of phage particles. The *E. coli* strains KW251 and XL1 Blue MRA were used to amplify enomic clones in λ Embl 3 and λ Dash respectively.

The *E. coli* strains were first grown overnight in 10-20ml of LB at 37°C. 25ml of LB + 0.2% maltose (Sigma), 10mM $MgSO_4$ was inoculated with 0.25ml of the overnight culture and incubated at 37°C for a further 3-4 hours (= late log phase). The

cells were then spun down at 4000 r.p.m. for 10 minutes and resuspended in 2.5ml 10mM $MgSO_4$ and stored at 4°C for later use.

2.3.6 Library Screens

The genomic libraries contained fragments generated from partial digestion with *Sau*3AI ranging in size from 10-15kb. These fragments were cloned into *Bam*HI sites in the phage vector λEMBL III or λDASH®II. *Sal*I was used to excise the insert from the arms of the phage vector. The phage vectors were plated on phage competent *E. coli* strain KW251. The pouch young and testis cDNA libraries were constructed with the λZAP II phage vector. After excision cDNA inserts were in the plasmid vector pBluescriptSK-, this vector conferred ampicillin resistance. The Hybond-N^+ membranes lifted from the library were prehybridised, hybridised and washed as described above (Section 2.3.4). Autoradiographs and the genomic libraries were lined up and appropriate sections of the library were then cored out and put into 500µl of SM (0.1M NaCl, 10mM $MgSO_4$ (BDH), 50mM Tris.Cl pH8, 0.01% gelatin (Ajax)). As 50-100 plaques are cored out of the library in the primary plugs, serial dilutions (1/10 to 1/1000) were set up from these primary plugs and replated onto a 85mm LB agar plate with phage competent *E.coli* strain KW 251 or XL1 Blue MRA. A concentration of 300-400 plaques was required on these plates in order to purify the positively hybridising plaque.

1µl of each of the dilutions was incubated with 70µl of phage competent KW 251 at 37 C for 10 minutes to allow adsorption of the phage particles into the cells. The cells and phage were then mixed with 3-4ml of top agarose (1% NaCl, 0.5% yeast extract, 1% tryptone, 0.7% agarose) and poured onto the top of an LB agar plate and allowed to set. The plates were incubated overnight at 37 C. Plate lifts were then conducted according to Sambrook *et al* (1989) where DNA from plaques was transferred from the plates to a Hybond-N^+ nylon membrane. The lifts were prehybridised for 2 hours, hybridised overnight, washed and exposed to film under the conditions described above. The plates were then cored in the appropriate positions and the plugs put into 100µl of SM. For secondary plates a concentration of 100-200 plaques was required. For tertiary plates a concentration low enough to pick single

colonies was required. The procedure of diluting the phage, replating and plate lifts was repeated until pure populations of clones were obtained.

2.4 Phage DNA manipulation

2.4.1 Clone Amplification

Pure populations of clones were grown on phage competent KW 251 at a concentration sufficient to achieve confluent lysis on an LB agar plate. 4ml of SM was added to the agar plates and then rocked for 4 hours, after which the SM was collected and 50-100µl of chloroform (Ajax) was added to lyse bacteria cells. Liquid lysates were set up adding 3.5µl of this concentrated solution to 200µl of phage competent cells and 300µl of 10mM $MgCl_2$ (BDH), 10mM $CaCl_2$ and incubated at 37 C for 10 minutes. This was added to LB with 10mM $MgSO_4$ and incubated with shaking at 37 C until lysis occurred. After lysis 6µg/ml of RNase and DNase was added and incubated at 37 C for 30 minutes to destroy bacterial RNA and DNA. The culture was centrifuged at 4000rpm for 10 minutes to pellet cell debris. NaCl and Polyethylene Glycol (PEG) (BDH) was added to the supernatant to final concentrations of 1M and 10% respectively, after which it was incubated on ice for a minimum of 1 hour. 100ml of the supernatant was centrifuged at 4000rpm for 15 minutes and the pellet was resuspended in 400µl of SM. The PEG was extracted by adding 400µl of chloroform (Company).

To extract the DNA from the phage particles EDTA and SDS was added to concentrations of 15mM and 0.1% respectively and then incubated at 68 C for 10 minutes. Protinase K and RNase was added to a concentration of 200µg/ml and incubated at 37 C for 1hr. The solution was then phenol/chloroform extracted twice with 200µl of each phenol and chloroform. After collecting the upper phase the DNA was ethanol precipitated (Section 2.2.5) and resuspended in 50µl of TE. 1µl was run on a gel (Section 2.2.2) to test purity and concentration.

2.4.2 Subcloning

Fragments were subcloned so that they could be sequenced. Fragments of interest were subcloned into pBluescript SK+ (Sratagene). To subclone a fragment, 500ng of the

clone was digested with the appropriate enzyme/s to produce the fragment of interest. The enzyme was then inactivated by incubation at 65 C for 15 minutes, if possible, or an ethanol precipitation was conducted to remove the enzyme.

If the fragment to be subcloned was relatively large compared to other fragments generated by the restriction digest, it was purified by running on a gel and extracting the specific fragment from the agarose as described above and a ligation was performed with this DNA. If the fragment of interest was relatively small, a ligation was performed without purifying the fragment as small fragments are preferentially cloned. The ligation was set up in a total volume of 10µl with 5µl of T4 DNA ligase buffer (Promega), one unit of T4 DNA ligase (Promega) and 200ng of vector DNA, which had previously been digested with the same enzyme/s that produced the fragment to be subcloned. The reaction was incubated at room temperature for 1 hour and then 4 C over night. PCR products were cloned into the pGEM T Easy vector (Promega) according to the manufacturer's instructions. The reaction was transformed into *E. coli* strains XL1 Blue or JM109 as described above (Section 2.4.5) and selected for on an LB agar plate supplemented with 10µg/ml ampicillin, 80µg/ml isopropyl-6-D-thiogalactoside (IPTG) (Roche) and 80µg/ml 5-bromo-3-chloro-2-indolyl-b-D-galactoside (X-gal) (Stratagene). When X-gal is cleaved it produces a blue pigment. *Lac Z*, the gene that cleaves X-gal, is induced by IPTG. When a fragment is cloned into pBluescript it inserts in the *lac Z* gene, interrupting it, so the colony does not produce a blue pigment. White colonies were picked, grown overnight in LB supplemented with 10µg/ml ampicillin at 37 C and mini-preparations were performed. A 4µl aliquot was restricted with an enzyme that cut the plasmid once to linearise it. The sample was then run on a gel to determine if the correct sized of the fragment was.

2.4.3 Shotgun Cloning

Shotgun cloning was undertaken to characterise regions of interest from genomic clones when it was not possible to subclone a relevant fragment generated by restriction endonuclease digestion. The TOPO Shotgun Subcloning Kit (Invitrogen) was used according to manufacturer's instructions. The DNA is sheared to a certain size and the ends are blunted with an enzyme. The DNA is then dephosphorylated and ligated into a plasmid vector.

Transformants were screened to select plasmids with the appropriate inserts for sequencing (Section 2.2.7). Colonies were lifted with a nylon Hybond-N^+ membrane from the selective media and laid colony side up onto a fresh LB plate. A second filter was placed on top thus giving A and B lifts. These plates along with the transformation plates were incubated at 37 C for 4-6hrs to allow regrowth of the colonies. After regrowth the filters were separated and laid colony side up for 3 minutes on 3MM Whatman paper soaked with 10%SDS. The filters were in turn exposed for 5 minutes each to 3MM Whatman paper soaked with denaturating solution, neutralizing solution and 2xSSC. The DNA was then fixed to the membranes by exposing them for 30 minutes to 3MM Whatman paper soaked with 0.4M NaOH after which they were rinsed with 2xSSC. Before hybridisation (Section 2.3.4) filters were prewashed with 5xSSC, 0.5% SDS, 1mM EDTA (pH8.0) for 1 hour at 55 C and then rinsed in 2xSSC.

2.4.4 Preparation of Chemically Competent Cells for Transformation

Plasmid vectors containing an insert were transformed into *Escherichia coli* strain XL1 Blue, which were chemically treated to make the cells competent to take up DNA.

A single colony of *E. coli* was picked and grown overnight at 37 C in 10ml of ψ broth (ψb) (5% yeast extract (Oxoid), 20% tryptone (Oxoid), 5% $MgSO_4$ (BDH), pH 7.6). After incubation, 5ml of the overnight culture was subcultured into 100ml of ψb (prewarmed to 37 C) and grown for a further 2 hours (OD^{550} = 0.48). The culture was then chilled on ice for 5 minutes and spun in 50ml tubes at 6000 r.p.m. for 5 minutes at 4 C to pellet the bacteria. The cells were resuspended in 40% volume of Tfb1 (30mM potassium acetate (BDH), 10mM RbCl (Aldrich), 10mM $CaCl_2$ (BDH), 50mM MnCl (Ajax), 15% glycerol (BDH), pH 5.8). The cells were then incubated on ice for another 5 minutes, spun again at 6000 r.p.m. for 5 minutes at 4 C, resuspended in 4% volume of Tbf2 (10mM PIPES (Sigma), 75mM CaCl, 10mM RbCl, add glycerol to 15% volume, pH 6.5) and left on ice for 15 minutes. Aliquots of 200µl, of cells, were snap frozen in liquid nitrogen and stored at -70 C.

2.4.5 Transformation

Transformations were performed to introduce plasmid vectors, which contained an insert, into bacterial cells. DNA was isolated from bacterial cells as described in section 2.2.3 which yielded much more DNA than was transformed.

An aliquot of the frozen cells (Section 2.4.4) were left at room temperature until just thawed and then placed on ice for 10 minutes. The DNA was then added (no more than 40% the volume of cells and no more than 100ng of DNA per 200µl of cells) and left on ice for 15-45 minutes allowing the cells to take up the DNA. The cells and DNA were then heat shocked at 45°C for 90 seconds and returned to ice for 1-2 minutes. Four volumes of ψb was added (at room temperature) and then incubated at 37°C for 45-60 minutes so that expression of resistance markers could begin. The cells were then spun down and most of the ψb was removed leaving about 40µl so that the cells could be resuspended and spread out onto selective media.

The bacteria containing the DNA (plasmid vector and hence the DNA insert) were then selected for, by growing overnight at 37°C on a Luria broth agar (LB) plate (1% NaCl (BDH), 0.5% yeast extract, 1% tryptone, 1.5% agar (Oxoid)) supplemented with 10µg/ml of ampicillin (Roche). A single colony was then picked from these plates and streaked onto another ampicillin LB agar plate for single colonies and incubated overnight at 37 C.

2.5 Fluorescence *in situ* Hybridisation (FISH)

To localize genes on the mitotic chromosomes, FISH was performed with homologous genomic clones as probes. Genomic clones were only used for FISH on the species they were isolated from.

2.5.1 Preparation of Mitotic Metaphase Spreads

Fibroblast cells from the species of interest were kindly supplied by Mrs Iole Barbieri and grown by Mrs Pat Miethke. Cells were arrested at metaphase by treatment with 50ng/ml–70ng/ml of Colcemid (Roche) for 2-4 hours. Culture media was removed from the flask and the cells were washed with 4ml PBS (calcium-magnesium-free

phosphate buffer) to remove all of the medium. The PBS was removed and 1% trypsin (Gibco) added to collect the cells from the flask. The cells were pelleted by centrifugation at 1200g for 10 minutes. The supernatant was removed and the cells resuspended in 2ml of hypotonic solution (0.1M KCl) pre-warmed to 37 C. Fixative (three parts methanol to one part acetic acid) was then added to a final volume 10ml. The cells were pelleted again by centrifugation at 1200g for 10 minutes, the supernatant discarded and the pellet resuspended again in 10ml fixative. This was repeated three times; after the final spin the pellet was resuspended in 0.5ml-1.5ml of fixative and dropped on to slides cleaned with acetone. Slides were left to air-dry overnight at room temperature and dehydrated by sequentially placing them in 70% ethanol, 90% ethanol and 100% ethanol for 3 minutes each at room temperature. Slides were then air-dried.

2.5.2 Slide pretreatment

Slides were pretreated with RNase to remove RNA from slides and prevent hybridisation of probe to RNA. 200µl of 100µl/ml of RNase was put onto slides, a cover slip was put on top and slides were incubated at 37 C for 1 hour. Slides were then washed three times at room temperature in 2xSSC(pH 7).

The slides were then treated with pepsin to remove excess protein. They were incubated at 37 C for 6 minutes in 0.005% pepsin (pH 2.3). The slides were washed twice for 1 minute in 2xSSC (pH 7) after which they were dehydrated as described above (Section 2.5.1).

Before hybridisation the slides were denatured for between 2 minutes 10 seconds and 3 minutes at 70 C in 70% formamide (Unilab), 2xSSC (pH 7). They were washed in ice-cold 2xSSC for 2 minutes, dehydrated as described above (Section 2.5.1) in ice-cold ethanol and left to air dry.

2.5.3 Probe preparation and hybridisation

Genomic clones were labelled with biotin using the BioNick Labeling System (Gibco BRL) according to the manufacturer's instructions.

200ng of probe was ethanol precipitated (Section 2.2.5) with between 10µg and 100µg of suppressor DNA (Suppressor DNA is prepared from total genomic DNA of the species the FISH is to be performed on. The suppressor DNA was reduced in size to

between 1.5 - 0.5kb by heating it to 100 C for 1 hour. A further 200ng of probe was ethanol precipitated with 60μg of salmon sperm DNA (ssDNA). After precipitation the probe was resuspended in 5μl deionized formamide (Sigma); once resuspended, 5μl of 20% dextran sulfate (Pharmica Biotech), 4xSSC was added and then incubated at 37 C for 30 minutes. The probe was then denatured at 85 C for 6 minutes and pre-annealed at 37 C for 20 minutes. Pre-annealing of genomic DNA to the probe blocks repetitive elements, preventing binding of the probe to repeat sequences on the chromosomes.

The probe precipitated with suppressor DNA, and the probe precipitated with no suppressor DNA, were added to different drops on the same slide. A cover slip was put on top of each drop and sealed with rubber cement (Ta Fong). Hybridisation was conducted in a moisture chamber at 37 C for 36-48 hours.

2.5.4 Washing and detection

Excess probe was washed off the slides by removing the cover slips and washing at 37 C-42 C three times for 5 minutes in 50% formamide/ 50% 2xSSC. Slides were then rinsed for 5 minutes in 2xSSC at 37 C. Depending on the required stringency, slides were washed in 1xSSC to 0.1xSSC at a temperature ranging from 50 C to 60 C for up to 5 minutes. The slides were then blocked in 3% BSA (Sigma), 4xSSC/ 0.1% Tween-20 (Sigma) for 30 minutes at 37 C. After blocking, a 3:500 dilution of 1mg/ml antibiotin (Vector) in antibody solution (1% BSA/ 4xSSC/ 0.1% Tween-20) was incubated on the slide for 45 minutes at 37 C in a moisture chamber. Excess antibiotin was washed off with 4xSSC/ 0.1% Tween-20 at 37 C twice for 5 minutes. A 1:100 dilution of 1mg/ml rabbit anti-goat+FITC (fluorescein isothiocynate) (Vector) (in antibody solution) was incubated on the slide, and washed off, under the same conditions as for antibiotin. FITC fluoresces under light wavelength 488nm. Slides were washed in 2xSSC for 10 minutes at 37 C. Chromosomes were counterstained for 1 minute with 4,6-diamidino-2-phenylindole (DAPI) (Sigma) and washed in dH_2O so they could be visualised under light of wavelength 397nm. Slides were covered with 15μl VectorShield anti-fade (Vector) to prevent signal fading.

Chromosomes were viewed with a Zeiss Axioplan 2 fluorescent microscope. Images were captured with a Spot RT Monochrome camera (Diagnostic Instruments, Inc.). DAPI and FITC images were coloured and overlaid in IP Lab (Scanalytics, Inc.).

Chapter 3: Comparative Mapping of Human Y-linked Genes in Tammar Wallaby

In order to determine wether the human Y chromosome, like the X, is composed of conserved and added regions, comparative mapping was undertaken. The differentiated region of the human Y chromosome contains 30-35 known genes (Lahn and Page, 1997). Genes on the human Y chromosome that also reside on the marsupial Y chromosome must have been located on the Y in a common ancestor 130MyBP, defining an ancient Y (YCR). For instance, *SRY* and *RBMY* are genes on the conserved region of the human Y that have persisted on the Y due to their important, and therefore selectable, functions in male reproduction. Conversely, genes on the human Y chromosome that are located on a marsupial autosome are defined as a part of a recent addition (YAR).

In order to determine which parts of the human Y chromosome were ancient and whether any were added, genes along the length of the human Y chromosome were chosen for mapping in tammar wallaby (Figure 3.1). The evolutionary origin of different regions of the human Y chromosome will give insight into how long they have been part of the Y, and their functional importance.

The Y chromosome has been rapidly degraded, and a gene is unlikely to survive in the long term if it does not acquire a selectable male specific function. For instance, *STSP* is a pseudogene on the added region of the human Y chromosome (YAR) that evidently failed to acquire a selectable male specific function, and as a result was partially deleted therefore inactivated.

Eight genes on the human Y chromosome were investigated by attempting clone and map their wallaby orthologues. The investigation of five X/Y shared Class I genes in wallaby will be discussed in the order of *TB4X/Y*, *EIF1AX/Y*, *RPS4X/Y*, *VCX/Y* and *PCHDX/Y*. The investigation of Y specific Class II genes will then be discussed in the order of *PRY*, *XKRY* and *BPY2*. Where possible, a X-borne homologue was used to screen male tammar wallaby genomic DNA libraries because they are more conserved between species than a Y-borne homologue.

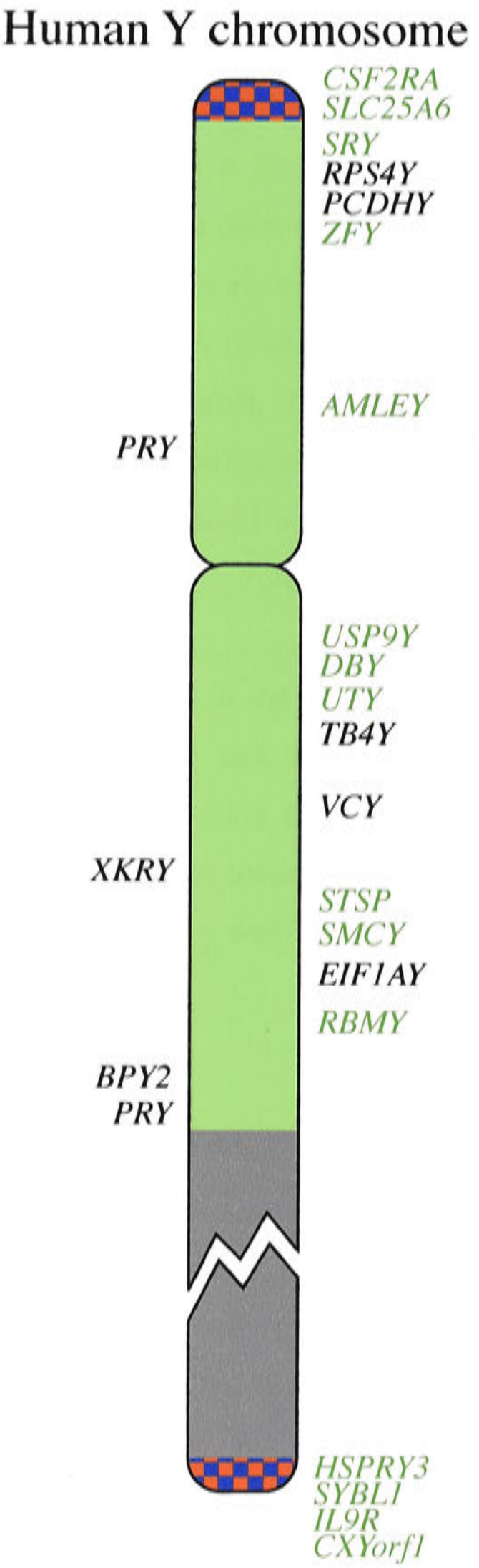


Figure 3.1: Genes on the human Y chromosome investigated in tammar wallaby during this study and previous studies. X/Y shared Class I genes are list on the right. Y-specific Class II genes are listed on the left. Genes investigated in this study are listed in black. Genes investigated in previous and concurrent studies are listed in green. PAR1 and PAR2 are indicated by red and blue checks. The non-recombining region of the Y is pale green and the heterochromatic region is grey.

3.1 Tammar wallaby *TB4X/Y*

Thymosin β4 on the X chromosome (*TB4X*) is an X/Y shared (Class I) gene whose X homologue lies in the XCR. Human *TB4Y* is located in interval 5D of the human Y chromosome and falls just outside the AZFa deletion interval (Lahn and Page, 1997). The mouse *TB4X* orthologue (*Ptmb4*) spans about 2kb of genomic DNA with three exons and two introns. The *Ptmb4* mRNA is about 700bp and codes for an actin monomer binding protein of 33 amino acids. *TB4X* may have a critical role in modulating the dynamics of actin polymerisation and depolymerisation in non muscle cells (Li *et al.*, 1996). *TB4X* maps to human Xq21.3-q22 among other XCR genes in the conserved region of the human X chromosome. *TB4X* was therefore predicted to map to the marsupial X chromosome.

Southern blot analysis was conducted to determine if there were homologous sequences present in the wallaby genome and if there was male: female dosage differences and/or male specific bands. Female dosage and/or male specific bands would indicate that wallaby *TB4X/Y* would be located on the X and/or Y chromosome. A male tammar wallaby genomic DNA library was screened to isolate clones that could be used to map wallaby *TB4X/Y* by FISH.

3.1.1 Southern blot analysis of *TB4X/Y*

A partial human *TB4X* cDNA of approximately 800bp, kindly supplied by Dr. M. Mitchell (INSERM, Marseille, France; GenBank accession number NM021109) was used as a probe. It was liberated from its plasmid vector by digestion with the restriction endonucleases *Sac*I and *Kpn*I for 1 hour at 37°C (Section 2.2.1). The human insert was separated from the plasmid by electrophoresis on a 1% agarose gel (Section 2.2.2). The insert was excised from the gel (Section 2.2.4), purified and labelled with ^{32}P dCTP (Section 2.3.3). The labelled *TB4X* cDNA was hybridised in Church buffer overnight at low stringency (55°C) (Section 2.3.4) to male and female tammar wallaby genomic DNA Southern blots fully digested with *Eco*RI and *Hind*III. The Southern blot membranes were washed at 65°C (Section 2.3.4) and then exposed to X-ray film for 7

days. The *TB4X* Southern blot was prepared and kindly supplied by Dr. Margaret Delbridge.

The Southern blot revealed good homology to *TB4X* in the wallaby genome (Figure 3.2). There were seven bands in the *Eco*RI lanes ranging in size from 2.2kb to approximately 20kb. In the *Hind*III lanes there were five bands ranging in size from 3kb to 7kb. There was no evidence of female dosage or male specific bands, suggesting that tammar wallaby *TB4X* was located on an autosome. The few clear bands indicated one or only a few copies.

3.1.2 Cloning and characterisation of *TB4X/Y*

A tammar wallaby genomic DNA library was screened. The labelled human *TB4X* cDNA was hybridised to tammar wallaby genomic DNA libraries under low stringency (55°C). The library membranes were washed at low stringency (55°C) and exposed to X-ray film for 24 hours. Under these conditions six primary clones were picked from the library, and after a further three rounds of screening, two final clones were isolated (Section 2.3.6). The tammar wallaby *TB4X/Y* clones were isolated from the library and kindly supplied by Natasha Sankovic. Throughout this thesis clones isolated from genomic DNA libraries will be named according to the species from which they were isolated, the gene they represent and the clone number they were assigned. For instance, *MeugTB4X*-2.1 represents *Macropus eugenii*, *TB4X*, clone number 2.1

Clone DNA was amplified (Section 2.4.1) and 1µg of each was digested with a variety of restriction endonucleases for 4 hours at 37°C. The digested samples were subjected to electrophoresis on a 0.8% agarose gel until separation of the fragments occurred. After completion of electrophoresis a Southern blot was performed (Section 2.3.2). The Southern blot was probed with the human *TB4X* cDNA, washed at low stringency (55°C) and exposed to X-ray film for 1 hour. Both clones showed strong hybridisation to the probe (Figure 3.3).

An *Xba*I generated hybridising fragment of about 4kb from *MeugTB4X*-2.1 was subcloned into pBluescriptSK$^+$ (Section 2.4.2), and the insert was sequenced using the vector primers T3 and T7 (Appendix I) (Section 2.2.7). Primers were designed (TB4X286-t3, TB4X354-t3, TB4X392-t3, TB4X81-t7, TB4X397-t7 and TB4X102-t7) (Appendix I) so that the insert could be walked into and a contig constructed. At the

TB4X

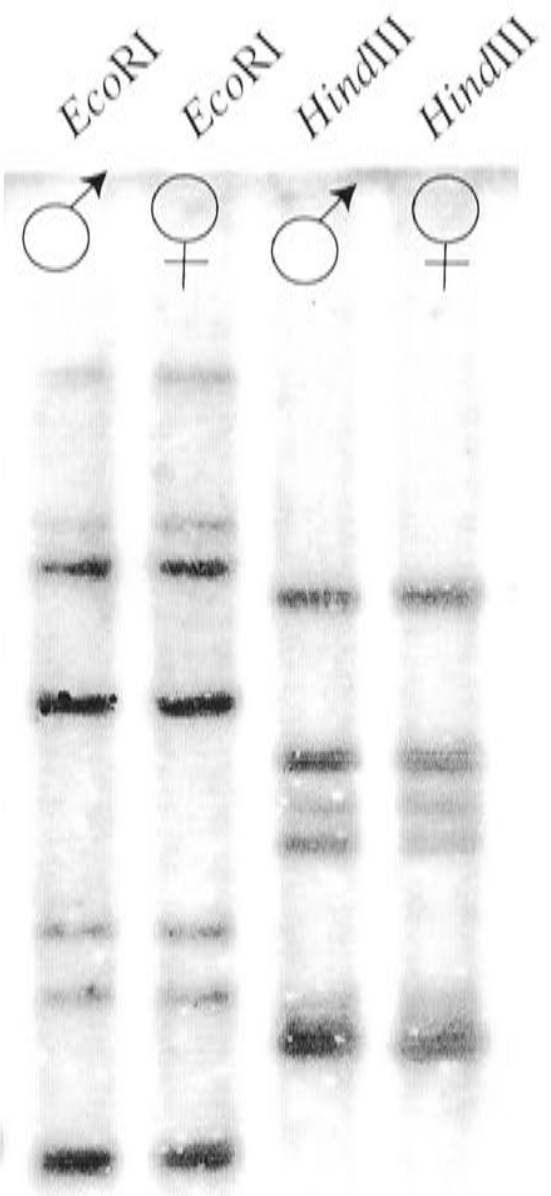
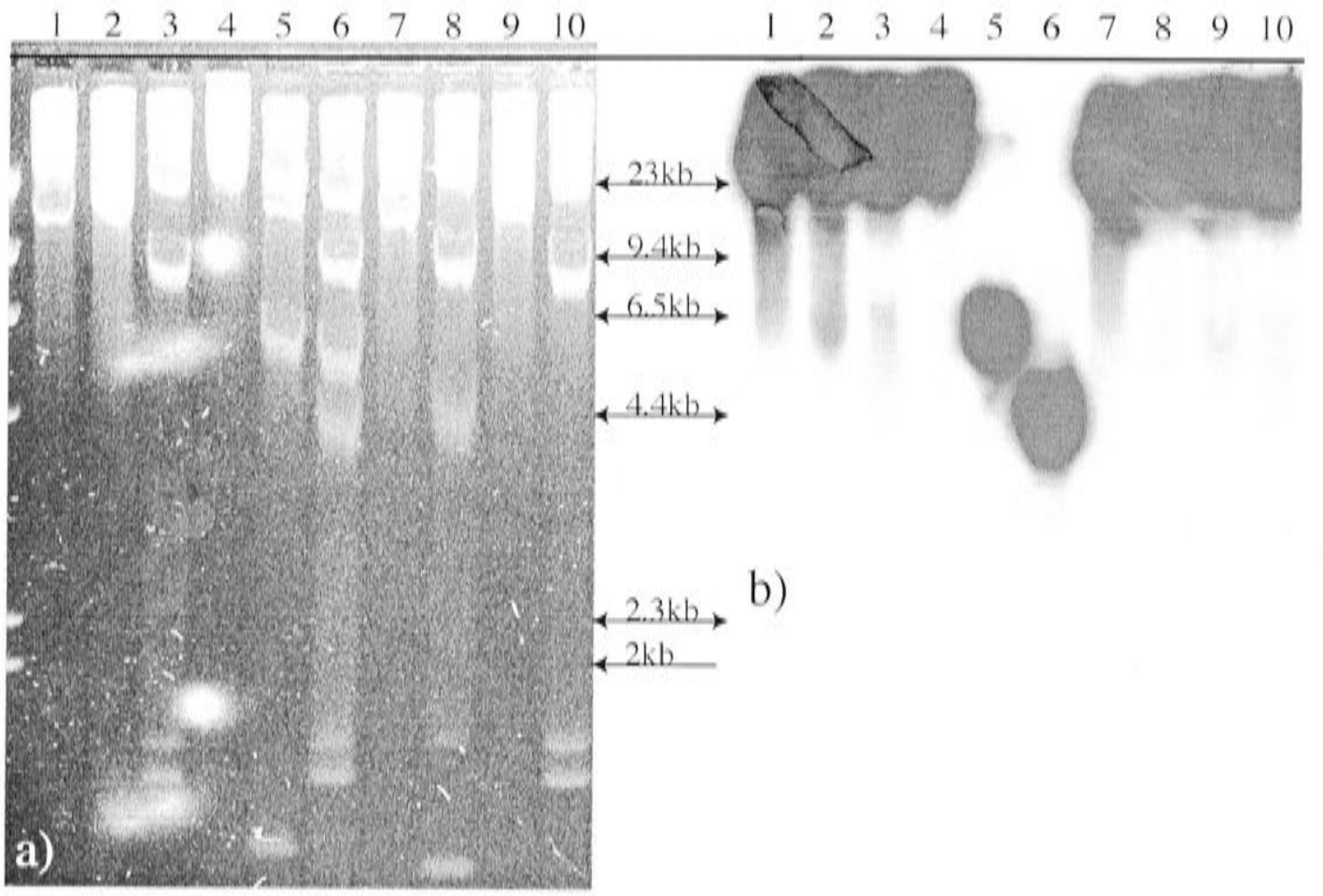
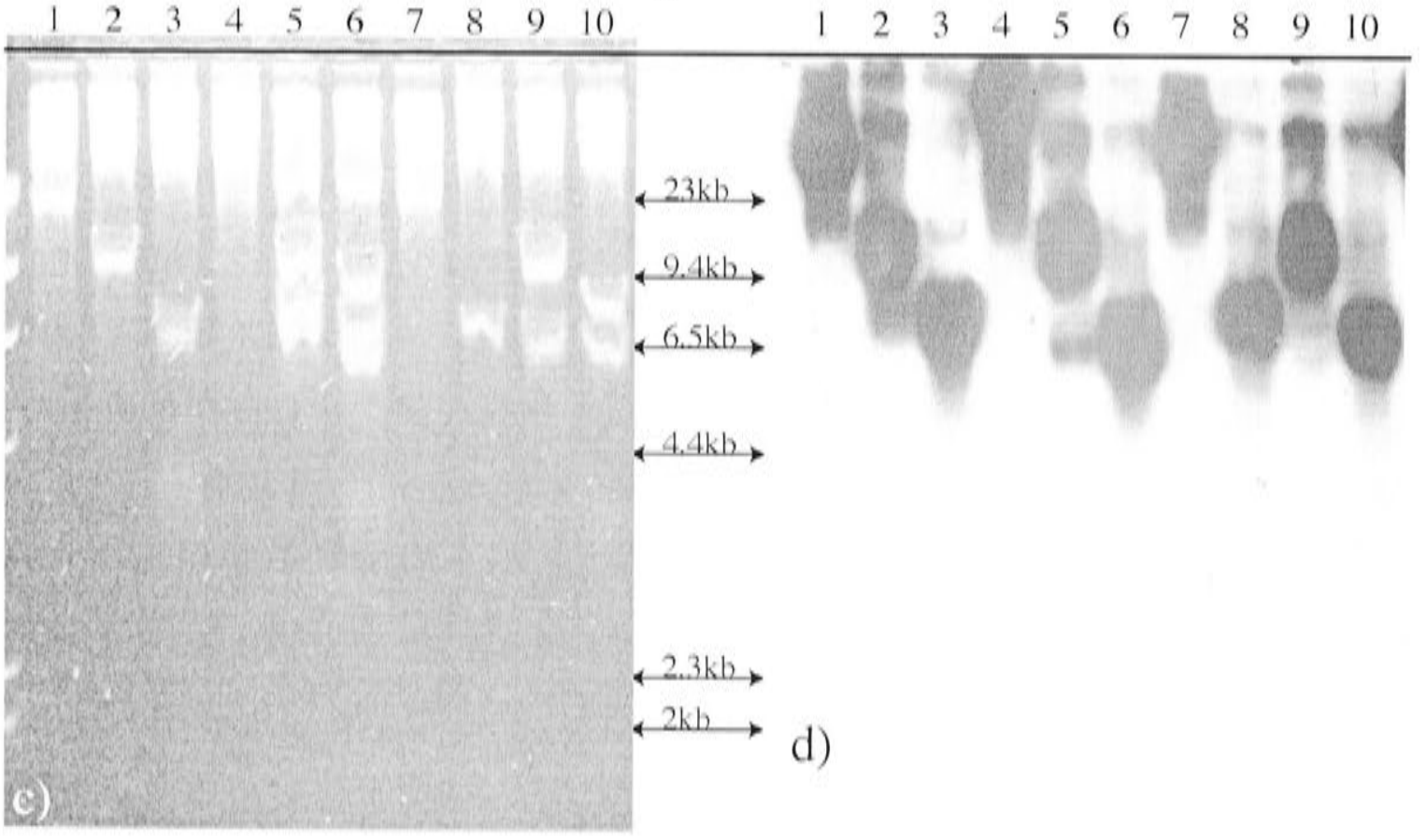


Figure 3.2: Total genomic tammar wallaby male and female DNA fully digested with Eco RI and Hind III. The Southern blot was hybridised with a *TB4X* probe. The autoradiograph was exposed for 7 days.

*MeugTB4X*1.1



MeugTB4X-2.1



Lane	Restriction enzymes		Lane	Restriction enzymes	
1	*Sal*I		6	*Sal*I	*Xba*I
2	*Eco*RI		7	*Sal*I	*Xho*I
3	*Xba*I		8	*Eco*RI	*Xba*I
4	*Xho*I		9	*Eco*RI	*Xho*I
5	*Sal*I	*Eco*RI	10	*Xba*I	*Xho*I

Figure3.3: a) & c) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammar wallaby *TB4X* clone #1.1 and #2.1 respectively. b) & d) Southern blot of restriction endonuclease gel after hybridisation of the *TB4X* probe. The autoradiographs were exposed for 1hr.

nucleotide level *MeugTB4X*-2.1 showed 85-90% homology to human *TB4X* in exon 2 and the first 180bp of exon 3. Homology between human and wallaby *TB4X* in the middle of exon 3, part of the untranslated region, was low. Homology increased to 90-95% in the final 200bp of exon 3. Sequencing demonstrated that *MeugTB4X*-2.1 did contain wallaby *TB4X/Y* (Figure 3.4a). The *MeugTB4X*-2.1 sequence contained introns in the same positions as the human *TB4X*. Human *TB4X* contains a small open reading frame (ORF) that codes for a 33 amino acid protein. The ORF begins in exon two and finishes in exon three. An ORF was observed in the corresponding region of the wallaby *TB4X/Y* (Figure 3.4b).

MeugTB4X-2.1 therefore contained the wallaby *TB4X/Y* orthologue and attempts to map it by fluorescence *in situ* hybridisation were made.

3.1.3 Mapping of wallaby tammar *TB4X/Y*

In order to test whether TB4X/Y was part of the conserved region of the human sex chromosomes, *MeugTB4X*-2.1 was biotin labelled in a nick translation reaction at 16°C for 3 hours (Section 2.5.3). The labelled probe was precipitated with 11μg of suppressor DNA and 60μg of salmon sperm DNA (ssDNA). The biotin labelled probe was hybridised (Section 2.5.3) to tammar wallaby metaphase chromosomes prepared from fibroblast cells (Section 2.5.1). The slides were pretreated with RNase and pepsin to remove excess RNA and protein (Section 2.5.2). After hybridisation, excess probe was washed off at 37°C in 50% formamide/50% 2 X SSC (Section 2.5.4). Biotin labelled probe that hybridised to target sequences on the metaphase chromosomes was detected with goat anti-biotin and then a rabbit anti-goat+FITC conjugate (Section 2.5.4). Chromosomes were counterstained with DAPI.

Signals were consistently observed just proximal to the centromere on the short arm of chromosome 5 in the tammar wallaby. No signals were observed on the X or the Y chromosomes (Figure 3.5).

3.1.4 Summary

Southern blot analysis of wallaby *TB4X/Y* revealed bands shared between male and female lanes. This demonstrated that *TB4X/Y* homologous sequences were present in the

(a)

(b)

```
HsapTB4X      MSDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES
HsapTB4Y      MSDKPGMAEIEKFDKSKLKKTETQEKNPLSSKETIEQERQAGES
MeugTB4X/Y    MSDKPDMGEIQKFNKSKLKKTETQEKNPLPSKETIEQEKQAGES
```

Figure 3.4: a) Sequence alignment of tammar wallaby and human *TB4X* exon sequence . Intron positions (indicated by arrows) are conserved between species and no homology is evident in them. Intron 1 is 950bp in mouse and unknown in wallaby. Intron 2 is 398bp in mouse and 415bp in wallaby. b) Alignment of human *TB4X* and *TB4Y* protein with the wallaby *TB4X/Y* protein.

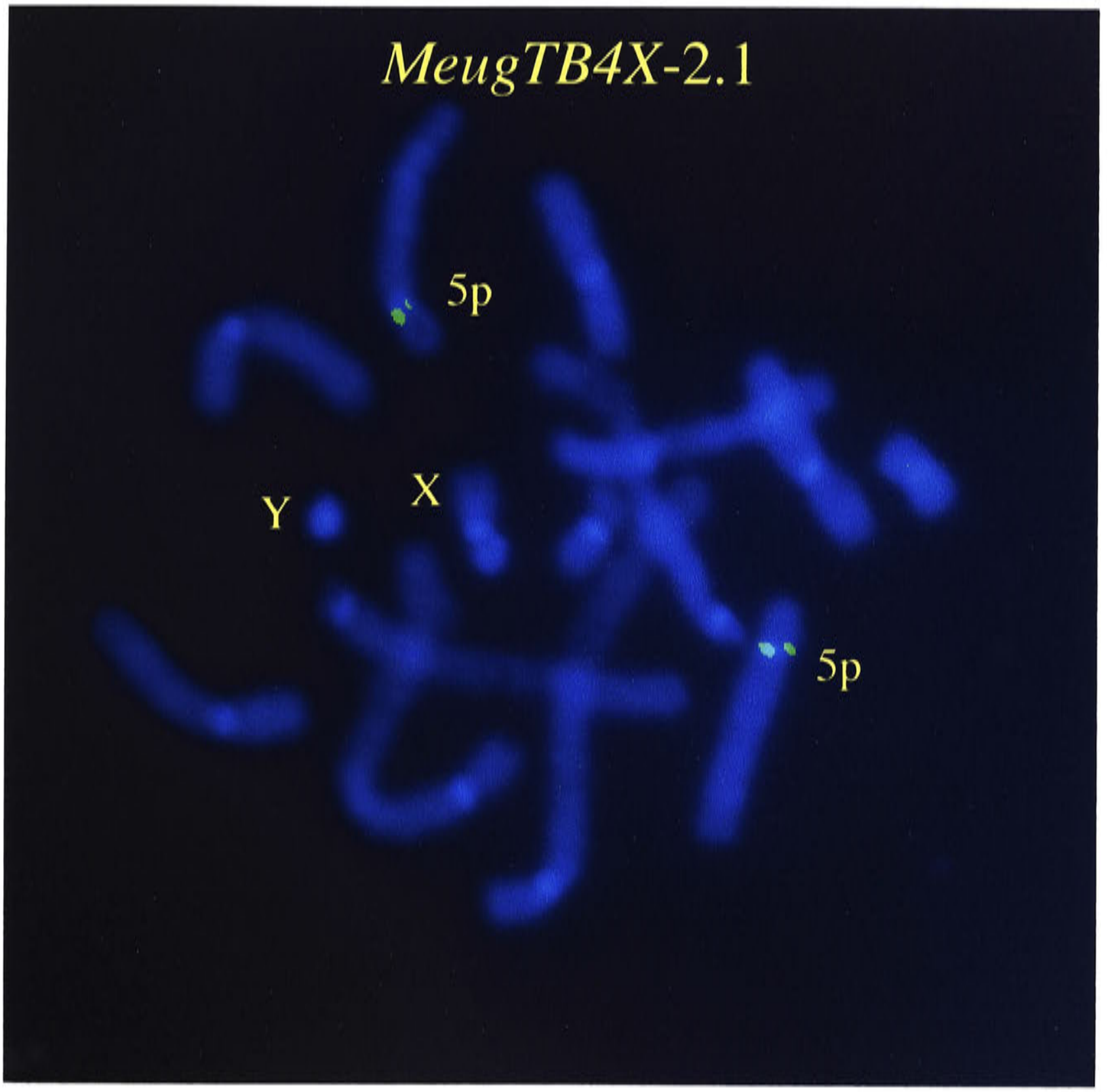


Figure 3.5: FISH mapping of *MeugTB4X*-2.1. Consistent signals were observed on the short arm of chromosome 5, near the centromere. Chromosomes were counterstained blue with DAPI.

wallaby genome and indicated that they were probably located on an autosome. Two clones were isolated from a male genomic DNA library and *MeugTB4X*-2.1 was confirmed as containing the wallaby *TB4X/Y* orthologue. *MeugTB4X*-2.1 was mapped to wallaby chromosome 5p. *TB4X* must therefore be a recent addition to the human X chromosome and *TB4Y* is a recent addition to the human Y chromosome.

3.2 Tammar wallaby *EIF1AX/Y*

Eukaryotic translation initiation factor 1A on the X chromosome (*EIF1AX*) is an X/Y shared gene in humans, but X-specific in mice. It spans approximately 17.3kb of genomic DNA. The *EIF1AX* mRNA is approximately 4.5kb and codes for an essential initiation factor of 144 amino acids. There are seven exons and six introns (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). Its location on the human X chromosome is ambiguous, so there is no indication of its potential location in tammar wallaby. *EIF1AY* is located in Y deletion interval 5Q, part of the deletion interval responsible for AZFb. It is a Class I gene that is single copy and ubiquitously expressed. The X-linked homologue escapes X-inactivation (Lahn and Page, 1997).

Southern blot analysis was conducted to determine if there were homologous sequences present in the wallaby genome, and if there were male: female dosage differences and/or male specific bands. This would indicate a location on the X and/or Y chromosome for wallaby *EIF1AX/Y*. A male tammar wallaby genomic DNA library was screened to isolate *EIF1AX/Y* clones that could be mapped for FISH.

3.2.1 Southern blot analysis of wallaby *EIF1AX/Y*

A partial human *EIF1AX* cDNA of approximately 500bp, kindly supplied by Dr. M. Mitchell (INSERM Unite 406, Marseille, France; GenBank accession number NM001412), was used as a probe. It was liberated from its plasmid vector by digestion with the restriction endonuclease *Eco*RI for one hour at 37°C. This partial *EIF1AX* cDNA was separated from the plasmid by electrophoresis on a 1% agarose gel. The insert was excised from the gel, purified and labelled with ^{32}P dCTP. The labelled probe was hybridised to tammar wallaby Southern blots containing male and female genomic

DNA fully digested with *Eco*RI and *Hind*III. The hybridisation and washing were conducted at low stringency (55°C). The Southern blot was exposed to X-ray film for seven days. Under these conditions some hybridisation was detected. In the *Eco*RI digests there were three faint bands of different size. In the *Hind*III digests there did not appear to be any clear bands in either the male or female lanes. This result suggested that there was one or a few copies of related sequences in the tammar wallaby, but their homology was low. There was no consistent evidence of male: female dosage differences or male specific bands (Figure 3.6).

3.2.2 Cloning and characterisation of wallaby *EIF1AX/Y*

A tammar wallaby genomic DNA library was screened in order to isolate clones containing *EIF1AX/Y* that could be mapped by FISH. A freshly labelled *EIF1AX* probe was hybridised to tammar wallaby genomic DNA library membranes at low stringency (55°C). The membranes were washed at 55°C and exposed to X-ray film for 24 hours. Under these conditions six primary clones were picked from the library. After a further three rounds of screening four final clones were purified. These clones were isolated from the library and kindly supplied by Natasha Sankovic.

The clone DNA was amplified and 1μg of each was digested with a variety of restriction endonucleases at 37°C for 4 hours. The fragments generated by the digestion were separated by electrophoresis on 0.8% agarose gel. A Southern blot was prepared, to which the *EIF1AX* cDNA probe was hybridised before washing at low stringency. The Southern blot was exposed to X-ray film for 24 hours. Strong hybridisation was observed only to *MeugEIF1AX*-1.1; there was no evidence of hybridisation to any of the other clones (Figure 3.7).

Subcloning into pBluescriptSK$^+$ of a 2kb hybridising fragment from *MeugEIF1AX*-1.1 generated by *Sal*I and *Xba*I was attempted. This was unsuccessful, so *MeugEIF1AX*-1.1 was characterised by shotgun-cloning (Section 2.4.3). The clone DNA was sheared in a nebulizer, and the ends of the sheared DNA pieces were repaired. The sheared DNA was subcloned into a TOPO blunt-end plasmid vector. The recombinant plasmids were transformed into a bacterial host (Section 2.4.5) and the colonies were grown on a LB agar plate supplemented with 10μg/ml of ampicillin and 40μg/ml of X-gal. The colonies were transferred to Hybond-N$^+$ membrane and screened for the relevant inserts,

EIF1AX

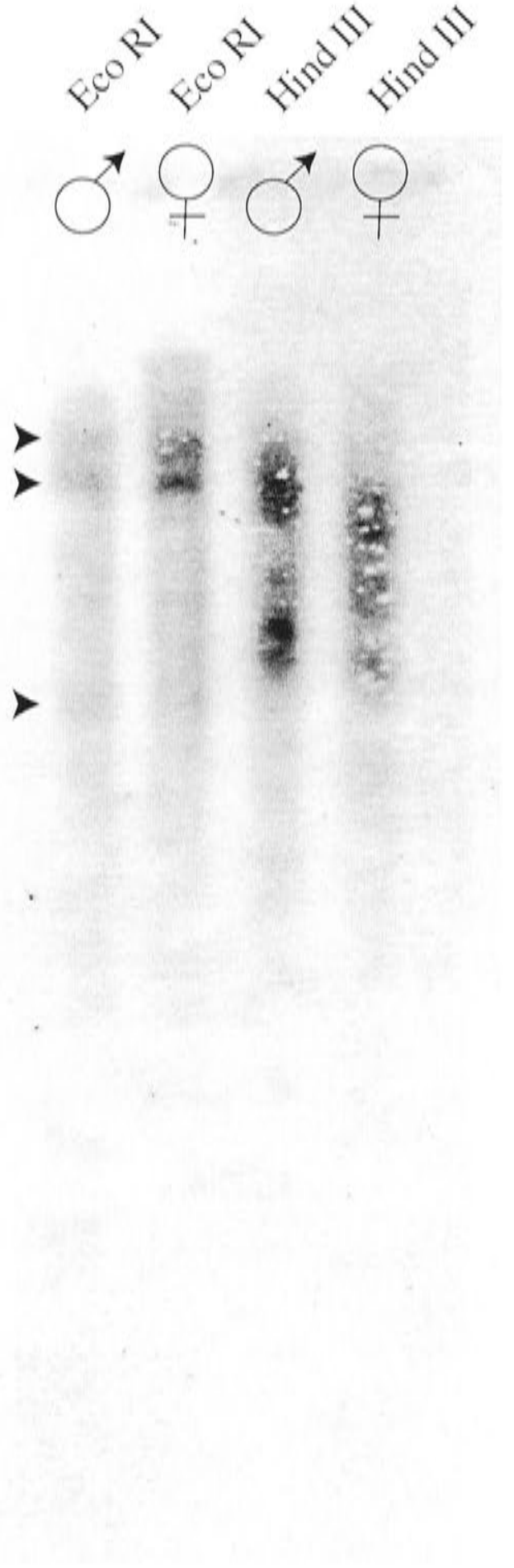
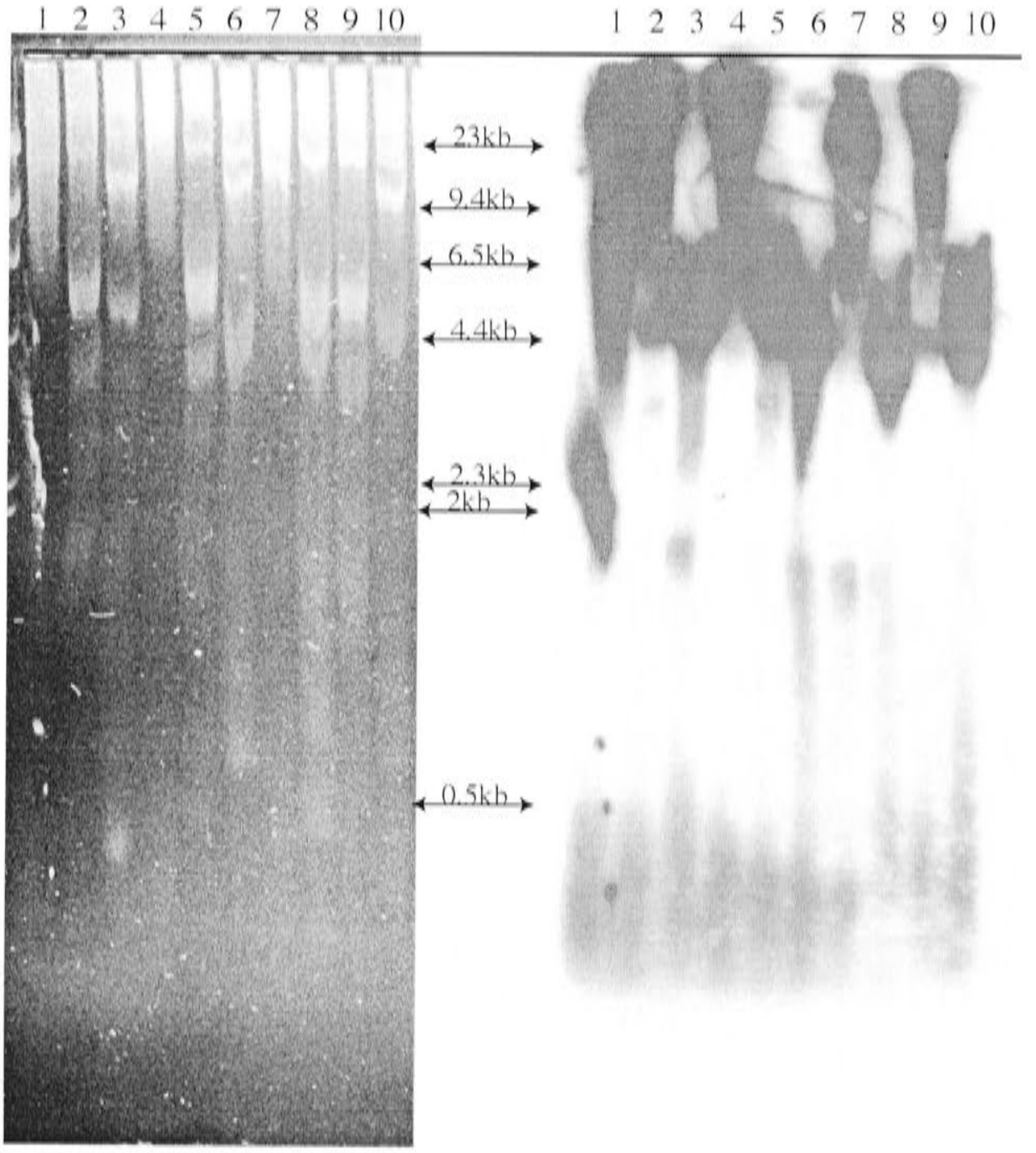


Figure 3.6: Total genomic tammar wallaby male and female DNA fully digested with *Eco*RI and *Hind*III. The Southern blot was hybridised with a *EIF1AX* probe. The autoradiograph was exposed for 7 days. Arrows indicate bands shared between male and female. The *Eco*RI bands were not considered dosed because more DNA was loaded in the female lanes.

MeugEIF1AX-1.1



Lane	Restriction enzymes		Lane	Restriction enzymes	
1	*Sal*I		6	*Sal*I	*Xba*I
2	*Eco*RI		7	*Sal*I	*Xho*I
3	*Xba*I		8	*Eco*RI	*Xba*I
4	*Xho*I		9	*Eco*RI	*Xho*I
5	*Sal*I	*Eco*RI	10	*Xba*I	*Xho*I

Figure3.7: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammsr wallaby *EIF1AX* clone #1.1. b) Southern blot of restriction endonuclease gel after hybridisation of the *EIF1AX* probe.The autoradiograph was exposed overnight.

which were sequenced with the vector primers T3 and T7. Sequencing revealed that this clone had 80-85% homology to human *EIF1AX* (Figure 3.8). The sequence obtained extended from exon 1 through to exon 7. Homology was strong in the coding region and dropped away significantly in the non-coding regions. No introns were present in the wallaby sequence and two frame shift mutations were observed. This indicated that *MeugEIF1AX*-1.1 contained a processed *EIF1AX/Y* pseudogene, here named *MeugEIF1AP* (*M. eugenii, EIF1A pseudogene*).

3.2.3 Mapping of tammar wallaby *EIF1AP*

MeugEIF1AP was biotin labelled in a nick translation reaction at 16°C for 3 hours. The labelled probe was precipitated with 25μg of suppressor DNA and 60μg of ssDNA. This was hybridised to tammar wallaby metaphase chromosomes prepared from fibroblast cells. The slides were pretreated with RNase and pepsin to remove excess RNA and protein. The hybridisation was conducted at 37°C for 48 hours after which the slide was washed at 37°C in 50% formamide/50% 2 X SSC. Biotin labelled probe that hybridised to the target sequence on the metaphase chromosomes was detected with goat anti-biotin and then a rabbit anti-goat+FITC conjugate. Chromosomes were counterstained with DAPI.

The FISH preparations always had high background. No consistent signals were observed on any chromosome, although there did appear to be signals on the middle of the long arm of the X chromosome of some spreads.

3.2.4 Summary

Southern blot analysis of wallaby *EIF1AX/Y* indicated that there was one, or a few copies of related sequences in the wallaby genome. There was no indication of male-specific bands or female dosage. Screening of a male genomic DNA library yielded one clone that was shown to contain a processed *EIF1AX/Y* pseudogene. This pseudogene was tentatively mapped to the X chromosome. However, no conclusion could be made about the evolution of *EIF1AX/Y* because the intron containing copy was not isolated and mapped.

```
HsapEIF1AX        CCCTCGTGTCGCCCCTCGGAGCAGCAGCCGCCGCGGTCGCCGCTACCCGGAAAGAAGTCA
MeugEIF1AX-1.1    ------------------------------------------------------------

                                                START          Exon 1          ▼ Exon 2
HsapEIF1AX        GAGACGCCGCGAGTCGCCGCCACCGCCATGCCCAAGAATAAAGGTAAAGGAGGTAAAAAC
MeugEIF1AX-1.1    -------------------------CCATGCCCAAGAATAAAGGTAAGGGATGTAAAAAT

HsapEIF1AX        AGACGCAGGGGTAAGAATGAGAATGAATCTGAAAAAAGAGAACTGGTATTCAAAGAGGAT
MeugEIF1AX-1.1    AGGCGTGGAGGAAAGAATGAAAATGAATCTGAAAAAAGAGAACTGGTCTTCAAGGAAGAT

                  ▼ Exon 3
HsapEIF1AX        GGGC-AGGAGTATGCTCAGGTAATCAAAATGTTGGGAAATGGACGGCTAGAAGCAATGTG
MeugEIF1AX-1.1    GGACCAGGAATATGCCCAGGTGATCAAAATGTTGGGAAATGGACAATTAGAAACAATGTG

                                                                  ▼ Exon 4
HsapEIF1AX        TTTCGATGGTGTAAAGAGGTTATGTCACATCAGAGGAAAATTGAGAAAAAAGGTTTGGAT
MeugEIF1AX-1.1    TTGTGATGTTGTGAAGAGGT--TGTCACATCAGAGGGAAATTAAGAAAAAAAGGT-----

                                                       ▼ Exon 5
HsapEIF1AX        AAATACCTCGGACATTATTTTGGTTGGTCTCCGAGACTACCAGGATAACAAAGCTGATGT
MeugEIF1AX-1.1    -AATACATCGAATATTATATTGATTGGTCTCCAAGACTACCAGGATAACAAAGCCGATGT

HsapEIF1AX        AATTTTAAAATACAATGCAGACGAAGCTAGAAGTCTGAAGGCATACGGCGAGCTTCCAGA
MeugEIF1AX-1.1    TACACTAAAATACAATGCAGATGAACCAAGAAGTTTAAAGGCATATGAAGAACTTCCAGA

                  ▼ Exon 6
HsapEIF1AX        GCATGCTAAAATCAATGAAACTGATACATTTGGTCCTGGAGATGATGATGAAATTCAGTT
MeugEIF1AX-1.1    GCACACCAAAATCAATGAAACAGACACATTTGTTCCTGGAGGTGATGATAAAATCCAGTT

                                                ▼ STOP      Exon 7
HsapEIF1AX        TGATGACATTGGAGATGATGATGAAGATATTGATGACATCTAAATTGAACTCAACATTTT
MeugEIF1AX-1.1    TGATGATGTCGAGGATGATAATGAAGACTTTGATGATATCTGAATTGAACT---------

HsapEIF1AX        ACATTCCATCTTTTCTGAAGATTGTCCTACAATTTGGATTTTGATCATGACAAAGAAGAT
MeugEIF1AX-1.1    ------------------------------------------------------------
```

Figure 3.8: Sequence alignment of human *EIF1AX* and tammar clone *MeugEIF1AX*-1.1. The locations of start/stop codons and introns in the human gene are indicated. No introns are evident in *MeugEIF1AX*-1.1 and homology outside of the coding region

3.3 Tammar wallaby *RPS4X/Y*

Ribosomal protein small subunit 4 (*RPS4*) codes for a small ribosomal protein that is located on the small ribosomal subunit at the 40S/60S subunit interface and is involved in mRNA binding (Nygard and Nika, 1982). *RPS4X* covers about 4.5kb of genomic DNA with seven exons and six introns (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). The *RPS4X* mRNA is about 900bp and codes for a 263 amino acid protein.

Human *RPS4Y* is a Class I gene that has an X-linked homologue located at Xq13.1 (Fisher *et al.*, 1990), which is part of the XCR. Two isoforms are encoded by X and Y-borne genes. Human *RPS4X* and *RPS4Y* encode proteins that share 92.8% amino acid identity (Fisher *et al.*, 1990). Both are expressed and incorporated into functional ribosomes (Zinn *et al.*, 1994). Human males therefore have two copies of *RPS4*, although *RPS4Y* is expressed at only 10% of *RPS4X*. The significance of the lower expressed Y-linked homologue is unknown (Zinn *et al.*, 1994). There are also many *RPS4* pseudogenes in the human genome.

Previous Southern blot analysis of *RPS4X* in tammar wallaby (Spencer, 1991) produced strong hybridisation. Many bands were evident on Southern blots containing *M. eugenii* and *M. rufus* genomic DNA, which indicated many autosomal pseudogenes. Also at least one band showed male: female dosage differences, which indicated an X-borne homologue. Mapping *RPS4X* by radioactive *in situ* hybridisation in tammar wallaby revealed significant peaks on chromosomes 4, 7 and the X (Spencer, 1991). The Southern blot and mapping results suggested that tammar wallaby and kangaroo had multiple copies of *RPS4X/Y*. Southern blot analysis in opossum, on the other hand (Jegalian and Page, 1998), revealed few copies, and demonstrated that both X- and Y-linked copies were present. *RPS4X* was therefore predicted to map to the wallaby X chromosome, and have a Y-borne partner.

3.3.1 Cloning and characterisation of wallaby *RPS4X/Y*

A probe for screening the tammar wallaby genomic DNA library was generated by PCR (Section 2.2.8) from male genomic *Monodelphis domestica* (opossum) DNA using the primers RPS4X-12f and RPS4X-700r (Appendix I). The *RPS4X* primers were

designed from sequence conserved between human and *M. domestica RPS4X/Y*. The cycling parameters to amplify the expected tammar wallaby *RPS4X* product of approximately 700bp were 35 cycles of (94°C, 1' / 55°C, 1' / 72°C, 1') following a 2' denaturation at 94°C. The product was subjected to electrophoresis on a 1% agarose gel, extracted from this gel and subcloned into pGEM T Easy. The insert was confirmed as *RPS4X* by sequencing from the vector primers. The insert was liberated from its plasmid vector by digestion with *Eco*RI for 1hr at 37°C.

A tammar wallaby genomic DNA library was screened and washed at low stringency (55°C) with the tammar wallaby *RPS4* probe, which was radioactively labelled with ^{32}P dCTP. Six primary clones were isolated from the library under these conditions; five were purified to single clones by a further three rounds of screening.

Southern blots were prepared from the DNA of each of these clones. Each was digested with the restriction endonucleases *Eco*RI and *Sal*I. The fragments generated were separated by electrophoresis on a 0.8% agarose gel. The Southern blots were hybridised to the same *RPS4X* probe used to screen the libraries under the same low stringency conditions. Of the five final clones isolated, only two showed good hybridisation to the probe (Figure 3.9). The strong hybridisation of the probe to two of the clones (*MeugRPS4X*-5 and *MeugRPS4X*-6) suggested that they were more likely to contain wallaby *RPS4X/Y* orthologues.

Because there were no small hybridising fragments in the strongly hybridising clones, primers were designed that were conserved between human *RPS4X/Y* and opossum *RPS4X/Y*. These primers were used for PCR on all five of the wallaby genomic clones to generate products to sequence and confirm their identity. The primers RPS4X-142f and RPS4X-585r were used in an attempt to amplify an *RPS4X/Y* product. The cycling parameters to amplify the expected wallaby *RPS4X/Y* product of approximately 440bp were 35 cycles of (94°C, 45'' / 57°C, 45'' / 72°C, 45'') following a 2' denaturation at 94°C. Of the five final clones tested by PCR, only *MeugRPS4X*-5 gave a single band of the expected size (Figure 3.10). This band was excised from the gel and subcloned into pGEM T Easy. The insert was sequenced using the vector primers SP6 and T7 (Appendix I).

Of the remaining uncharacterised clones, faintly hybridising bands from *MeugRPS4X*-2 (~2.5kb, *Eco*RI generated), *MeugRPS4X*-3 (~2kb, *Eco*RI/*Sal*I generated)

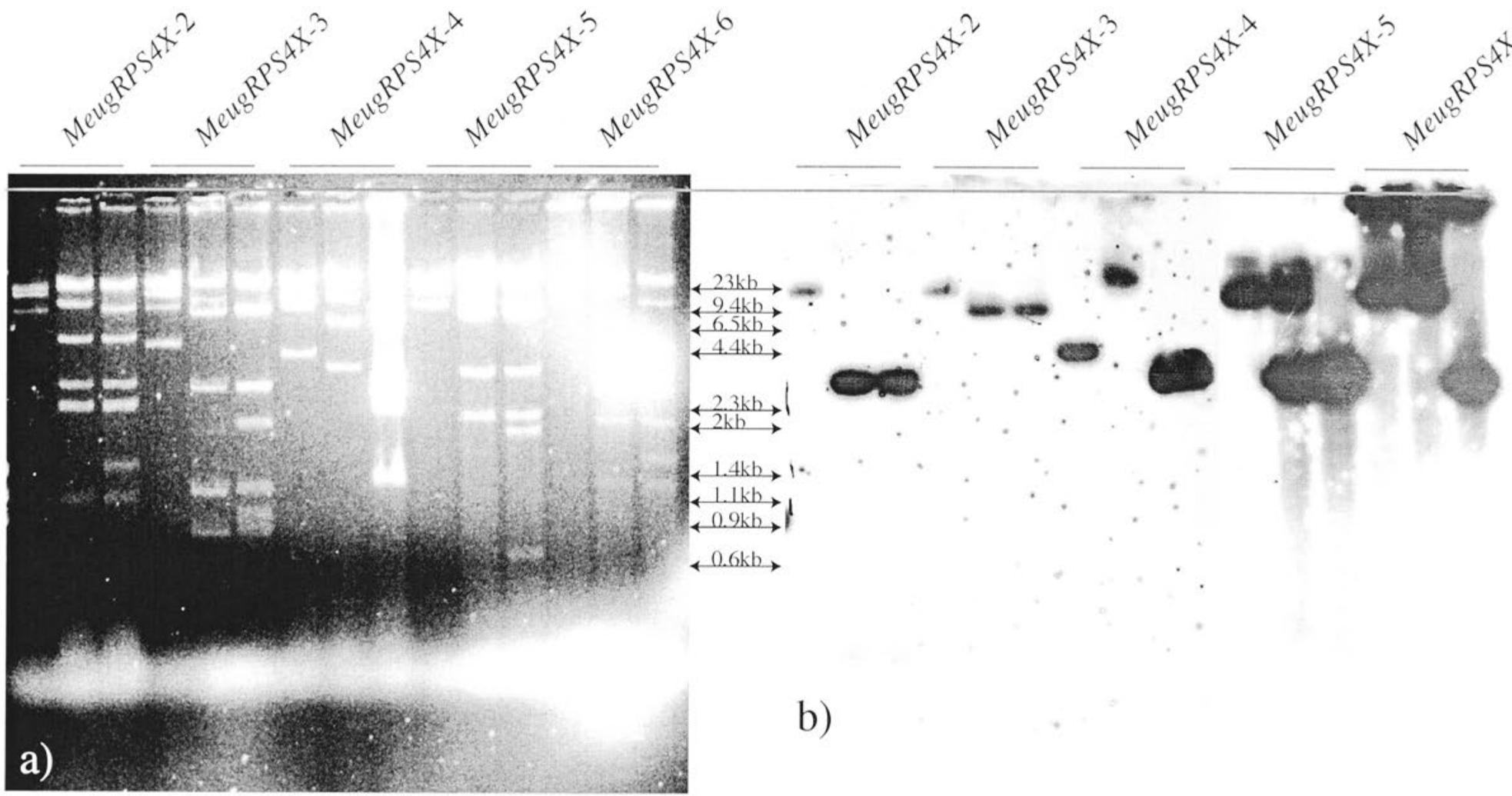


Figure3.9: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammar wallaby *RPS4X* clones #2, 3, 4, 5 and 6. b) Southern blot of restriction endonuclease gel after hybridisation of the *RPS4X* probe. The autoradiograph was exposed overnight for clones #2, 3 and 4. The autoradiograph was exposed for 5hrs for the strongly hybridising conles #5 and 6.

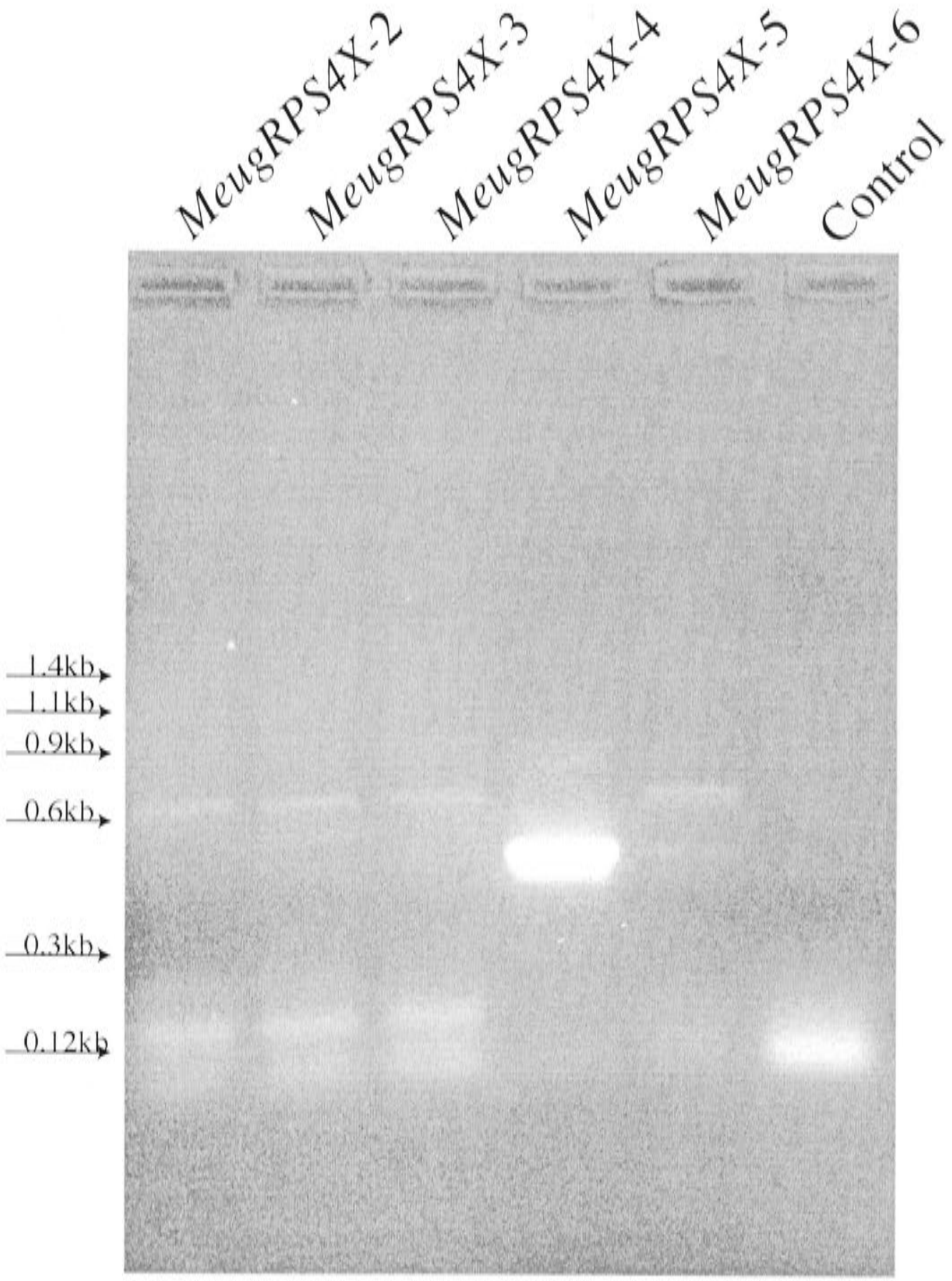


Figure 3.10: Amplification *RPS4X* from tammar wallaby genomic clones. The 440bp product was strong in *MeugRPS4X*-5. A faint band was evident in *MeugRPS4X*-6.

and *MeugRPS4X*-4 (~3kb, *Eco*I generated) were subcloned into pBluescriptSK$^+$ (Figure 3.9). These inserts were sequenced using the vector primers T3 and T7 and showed no homology to *RPS4X/Y*. *MeugRPS4X*-6 showed strong hybridisation to the probe and was therefore more likely to contain *RPS4X/Y*. This clone was characterised by shotgun-cloning.

The PCR product from *MeugRPS4X*-5 and the sequence obtained from *MeugRPS4X*-6 by shotgun-cloning revealed that both had homology to *RPS4X/Y*. A BLAST-N search showed that, at the nucleotide level, *MeugRPS4X*-5 and *MeugRPS4X*-6 had 80-85% sequence identity to human *RPS4X* (Figure 3.11). However, neither had an open reading frame. *MeugRPS4X*-5 and *MeugRPS4X*-6 shared only 85-90% sequence identity and were therefore different pseudogenes. Both wallaby *RPS4X/Y* clones lacked introns, suggesting that they were retrotransposed from an intron-containing copy located at a different region of the genome.

The pseudogenes were not mapped. If used for FISH mapping, and if the processed wallaby *RPS4X/Y* pseudogenes are about the same size as the human *RPS4X* cDNA (approximately 1kb), the wallaby pseudogenes are unlikely to hybridise to the intron containing copy from which they were derived. This is because there is a further 15-20kb of non-coding sequence in the genomic clones homologous to the region of the genome that the pseudogenes are located. Because of the high number of wallaby *RPS4X/Y* pseudogenes, no further attempts were made to clone the intron containing copy of *RPS4X/Y*.

3.3.2 Summary

Previous Southern blot analysis of *RPS4X/Y* in wallaby indicated that there were many related sequences located in its genome, at least one of which lay on the X chromosome (Spencer, 1991). The isolation of two pseudogenes from a male genomic DNA library were consistent with the results of Southern blot analysis, which indicated that wallaby has multiple copies of *RPS4X/Y* like genes.

In the presence of many processed pseudogenes, it is often difficult to isolate a single intron-containing gene. Copies of *RPS4X/Y* that did not contain introns probably presented as the best positives on the library screens because sequences homologous to

```
                                          RPS4X-12f
                                          ---------->
HsapRPS4X       TGGCTCGTGGTCCCAAGAAGCATCTGAAGCGGGTGGCAGCTCCAAAGCATTGGATGCTGG
MeugRPS4X-5     TGGCTCGTGGTCCCAAGAATCATCTGAAGTGTGTAGCAGCTCCAAAGCATTGGATGCTGG
MeugRPS4X-6     TGGCTTGTGGTCCCCAAAAGCATCTGAAGCGTGTAGCAGTTCCAAAGCACTGGATGCTGG

HsapRPS4X       ATAAATTGACCGGTGTGTTTGCTCCTCGTCCATCCACCGGTCCCC-ACAAGTTGAGAGAG
MeugRPS4X-5     ATAAATTAACAGGAGTTTTTGCGCCAAGACCATCCACAGGTCCCCCACAAGCTGAGGGAA
MeugRPS4X-6     ATAAATTAATGGAAGT-TTTGCTCCAAGACCATCCACAGGTCCCT-ACAAGCTGAGTGAA
                                          RPS4X-142f
                                          ---------->
HsapRPS4X       TGTCTCCCCCTCATCATTTTCCTGAGGAACAGACTTAAGTATGCCCTGACAGGAGATGAA
MeugRPS4X-5     TGTCTGCTGCTCATCATCTTCCTGAGAAACACACTCAAGTATGCCCTGACTGGAGATGAA
MeugRPS4X-6     TGTCTGCCACTCATTATCAACCTGAAAAGCAGACTCAAGTATGCCCTGATGGGACATGAA

HsapRPS4X       GTAAAGAAGATTTGCATGCAGCGGTTCATTAAAATCGATGGCAAGGTCCGAACTGATATA
MeugRPS4X-5     GTCAAGAAGACCTGCCTGCAATGATTCATCAAGATTGATGGCAAGGTCCGCACTGATGTT
MeugRPS4X-6     GTAAAAAAGATT----------------------GATGGCAAGGTCCGCACTGACATT

HsapRPS4X       ACCTACCCTGCTGGATTCATGGATGTCATCAGCATTGACAAGACGGGAGAGAATTTCCGT
MeugRPS4X-5     ACATACCCTGCTGGCTTTATGGATGTC------ATTGAGAAGACAG-CGAACATTTCCAT
MeugRPS4X-6     ACATACCCTGCTGCCTTTATTGATGTCATTAGCATTGAGAAGAAGGGTGAACATTTCCAT

HsapRPS4X       CTGATCTATGACACCAAGGGTCGCTTTGCTGTACATCGTATTA---CACCTGAGGAGGCC
MeugRPS4X-5     CTGATTTATGACACAAAAGGGTGTTTTGCTGTTCATTGTATTA---CAGCTGAGGAGGCT
MeugRPS4X-6     TCCGTTTAATAGACAAAGTGGTGCTTTGCTGTTCGTGGTATTATTACAGCTGAGGAAGCT

HsapRPS4X       AAGTACAAGTTGTGCAAAGTGAGAAAGATCTTTGTGGGCACAAAAGGAATCCCTCATCTG
MeugRPS4X-5     AAATATAAATTGTGCA--GTGAGAAAAATCTTTGTGGGTACAAAAGGTATTCCCCCTCTT
MeugRPS4X-6     AAATATAAATTGTGCAAAGTCAGAAAAATCTTTGTGGGTACAAAAGGTATTCCCCATTTG

HsapRPS4X       GTGACTCATGATGCCCGCACCATCCGCTACCCCGATCCCCTCATCAAGGTGAATGATACC
MeugRPS4X-5     GTCACCCATGGTGCCCATACTATCTGATACCCAGATCCTCTCATCAAAGTGAAGGGTACA
MeugRPS4X-6     GTGACCCATGATGCCTG----------------ATTCTCTCATCAAAGTGAATGATATA

HsapRPS4X       ATTCAGATTGATTTGGAGACTGGCAAGATTACTGATTTCATCAAGTTCGACACTGGTAAC
MeugRPS4X-5     GTTCAGATTGATTTAGAAGCTGGCAAGAT-------TTTATCAAGTTTGATACCGGTGAC
MeugRPS4X-6     GTTCAGATCGATTTAAATGTTAGCAAGATCACTGATTTTATAAAGTTTGATGCTGATAAT
                                                                RPS4X-585r
                                                                <----------
HsapRPS4X       CTGTGTATGGTGACTGGAGGTGCTAACCTAGGAAGAATTGGTGTGATCACCAACAGAGAG
MeugRPS4X-5     CTATGTATGGTGACCAGTGGGGCTAATTTGGGTCAAATTGGTGTGATCACAAACAGGGAG
MeugRPS4X-6     CTGTGTATGGTGACCAGTGGGGCTAATTTCAGTCGAATTGGGAGAATT-------GGGAG

HsapRPS4X       AGGCACCCTGGATCTTTTGACGTGGTTCACGTGAAAGATGCCAATGGCAACAGCTTTGCC
MeugRPS4X-5     AAACATCCTGGCTTTTTTGATGTTGTCTGTGTAAAAGATGCCAATGGCAATAGTTTTGCC
MeugRPS4X-6     AAACATCCTGACTCTTTTATTGTTGTCCATGTACAAGATGTCAATGTCAATAGTTTTGCC
                                                                RPS4X-700r
                                                                <----------
HsapRPS4X       ACTCGACTTTCCAACATTTTTGTTATTGGCAAGGGCAACAAACCATGGATTTCTCTTCCC
MeugRPS4X-5     ACTAGGCTCTCCAACATTTTTGTTATTGGTAAAGGTAATAAGGCTTG-ATCTCTCTTCCC
MeugRPS4X-6     ATTAGGCTCTCCAATATTTTTGTTATTGGTAAAGGTAATAAGCCTTGGATCTCTCTCTTC

HsapRPS4X       C--GAGGAAAGGGTATCCGCCTCACCATTGCTGAAGAGAGACAAAAGACTGGCGGCCA
MeugRPS4X-5     T--AAGGAAAGGGTATCTGCCTTACAACTGCTGAAGAAAGAAACAAGAGATTGGCAGCTA
MeugRPS4X-6     CAGGAAGAAAGGGTATCCGCCTTACAGTTGGTAAAGAAAGAGACAAGAGACTGACAGCTA

HsapRPS4X       AACAGAGCAGTGGGT
MeugRPS4X-5     AGCAGAGCAGTGGGT
MeugRPS4X-6     AACAGAGCAGTGGCT
```

Figure 3.11: Sequence alignment of human *RPS4X*, tammar wallaby *MeugRPS4X*-5 (obtained via PRC) and tammar wallaby *MeugRPS4X*-6 (obtained via shotgun-cloning). *RPS4X/Y* primers used in this study are indicated by arrows.

the *RPS4X/Y* probe were uninterrupted by introns, therefore giving stronger hybridisation. The *RPS4X/Y* probe was generated, by PCR, from tammar wallaby genomic DNA. These PCR products did not contain the expected introns. The products were smaller than expected for the intron containing *RPS4*, and sequencing revealed that only coding sequence was present. *RPS4X/Y* intron and exon structure is conserved in human and mouse, so is probably conserved in wallaby. The probe, generated from the genomic DNA, was therefore a PCR product amplified from an *RPS4X/Y* pseudogene, so would have had 100% homology to at least one pseudogene.

No conclusion could be made about the evolution of *RPS4X/Y* because the intron containing copy was not isolated and mapped.

3.4 Tammar wallaby *VCX/Y*

The family of *VCX/Y* (Variably charged protein on the X/Y chromosome) genes on the human Y code for small positively charged proteins. *VCY* is a Class I gene with an X-linked homologue (*VCX*) but is present in multiple copies on the Y. It is located among XAR genes at human Xp22.3. *VCX* and *VCY* contain only one intron of 192bp. The *VCX* mRNA is about 800bp and codes for a protein of 206 amino acids (Lahn and Page, 2000). There are at least two copies of *VCY* on the human Y chromosome, no more than 140kb apart, and there are approximately 12 copies of *VCX* in close proximity to each other on the human X chromosome (Lahn and Page, 2000). Both *VCY* copies along with at least three *VCX* copies are expressed exclusively in the testis, and expression is most probably restricted to the male germ cells, indicating a potential role in spermatogenesis. *VCX* and *VCY* share a high degree of sequence identity, with the exception of a 30bp unit that is tandemly repeated in X derived cDNA and present only once in Y derived cDNA (Lahn and Page, 2000).

Southern blot analysis was conducted to determine if there were homologous sequences present in the wallaby genome. Female dosage and/or male specific bands would indicate that wallaby *VCX/Y* was located on the X and/or Y chromosome. A male tammar wallaby genomic DNA library was screened to isolate *VCX/Y* clones that could be mapped by FISH.

A 400bp partial human cDNA of *VCY*, kindly supplied by Dr. David Page (Whitehead Institute, Boston, USA), was used as a probe to detect wallaby *VCY*. The *VCY* probe was liberated from its plasmid vector by digestion for 1 hour at 37°C with the restriction endonuclease *Not*I.

3.4.1 Southern blot analysis of *VCX/Y*

Hybridisation of the human *VCY* probe to tammar wallaby genomic Southern blots under low stringency (55°C) was conducted to detect homologous sequences. The Southern blot analysis would detect sex dosage difference and male-specific bands that would indicate a locus on the wallaby X and/or Y chromosomes. The Southern blot contained male and female tammar wallaby genomic DNA totally digested with *Eco*RI and *Hind*III. The probe was liberated from the plasmid vector using the restriction endonuclease *Not*I. The insert was purified from the agarose gel after electrophoresis and then labelled with ^{32}P dCTP. After hybridisation, the Southern blot was washed at 65°C and exposed to X-ray film for 48 hours.

The human *VCY* probe detected four bands shared between male and female in the *Eco*RI digests and two bands in the *Hind*III digests. There were no male-specific bands and no evidence of male-female dosage difference. The few discrete bands detected indicated wallaby orthologues of *VCX/Y* existed in one or a few copies (Figure 3.12).

3.4.2 Attempts to clone and sequence wallaby *VCX/Y*

A genomic DNA library was screened to isolate clones to use for FISH mapping. The same human *VCY* cDNA, used to probe Southern blots, was also used to screen a male tammar wallaby genomic DNA library. The library was hybridised under low stringency conditions and washed at 55°C. The library membranes were exposed to X-ray film overnight. Five primary clones were identified with the low stringency washes from the library of approximately 3 x 10^5 clones; four were purified to final clones by three further rounds of screening.

DNA was amplified and a Southern blot was prepared for each clone. The clones were digested with a variety of restriction endonucleases for 4 hours. The fragments generated were separated by electrophoresis on a 0.8% agarose gel. The Southern blot

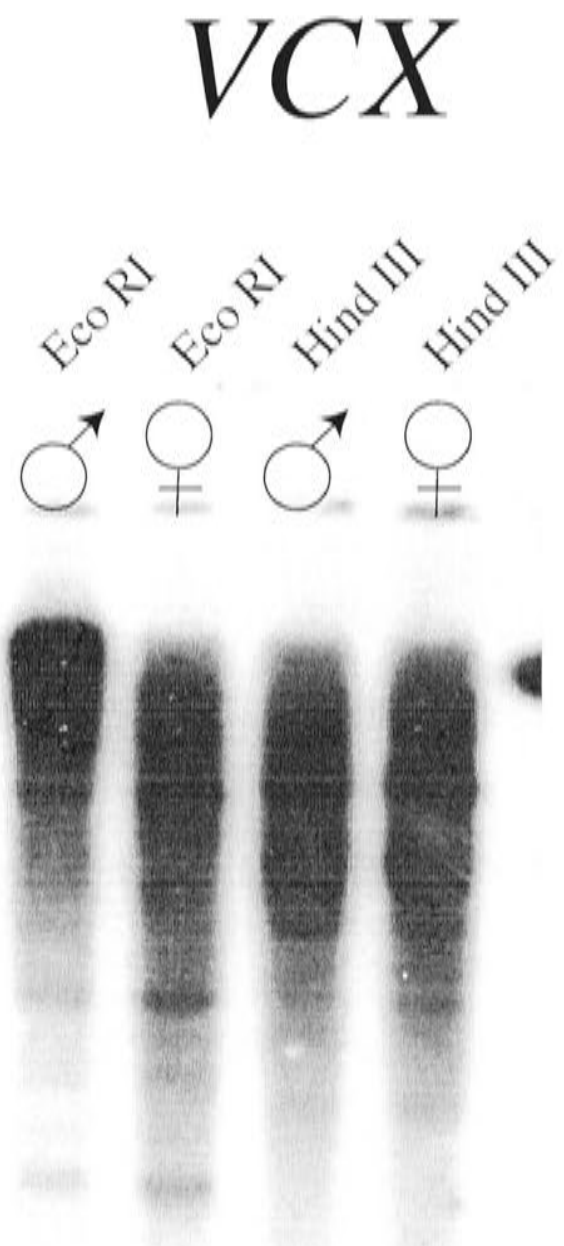


Figure 3.12: Total genomic tammar wallaby male and female DNA fully digested with *Eco*RI and *Hind*III. The Southern blot was hybridised with a *VCX* probe. The autoradiograph was exposed for 2 days. The male *Eco*RI lane is incompletely digested, therefore, female bands appear darker than male bands.

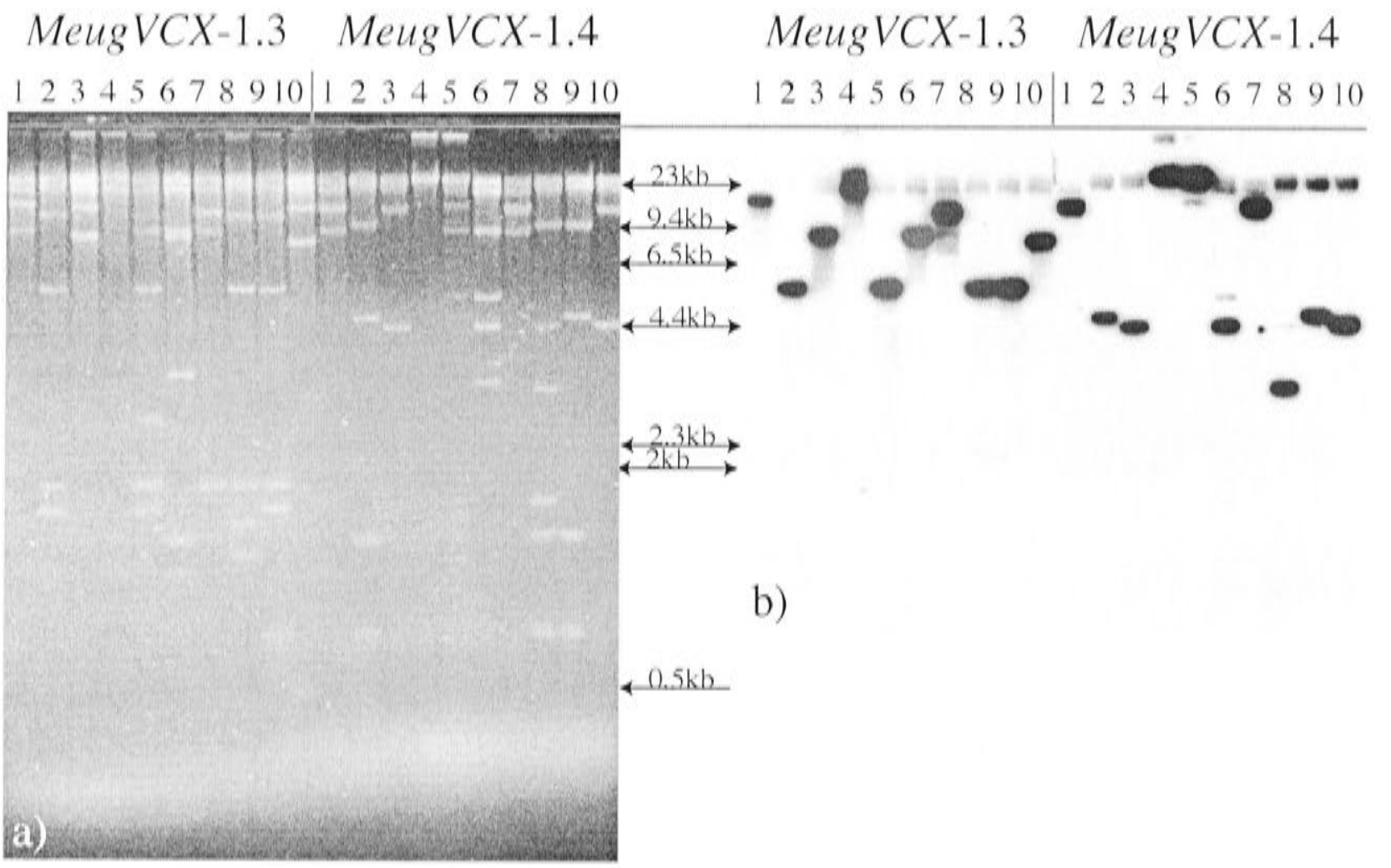
was probed with the human *VCY* cDNA labelled with ^{32}P dCTP. Hybridisation was strongest to clones *MeugVCX*-1.3 and *MeugVCX*-1.4 (Figure 3.13). A 3kb hybridising fragment from *MeugVCX*-1.4, generated by *Eco*RI and *Xba*I, was sub-cloned into pBluescriptSK$^+$.

The insert was sequenced from either end using the plasmid primers. Primers BPY1#4t3(345) and BPY1#4t7(470) were designed from the sequence generated by the plasmid primers to walk into the sub-clone and construct a complete contig. A nucleotide blast search of the insert sequence revealed that clone *MeugVCX*-1.3 had no homology to human *VCX*. Instead it shared 85-90% homology with human *SMARCF1* (Figure 3.14). An open reading frame was observed across the region sequenced (Figure 3.15).

A fresh plating of the same genomic DNA library was screened to account for the possibility that wallaby *VCX/Y* was not represented on the first plates screened. The same *VCY* cDNA probe was used to isolate six new primary clones under low stringency hybridisation and washing conditions. All six of the primary clones were purified to final clones by a further three rounds of screening. A Southern blot was prepared from the DNA of each clone. The clones were digested with the restriction endonucleases *Eco*RI and *Sal*I. The *VCY* probe was hybridised to the Southern blot on which it hybridised to five of the six clones (Figure 3.16). Hybridising fragments from clones *MeugVCX*-2.2, *MeugVCX*-2.4 and *MeugVCX*-2.6 were sub-cloned and sequenced to confirm their identities. However, none of the hybridising fragments that were sequenced showed homology to *VCX/Y* or *SMARCF1*.

PCR was conducted on all of the clones isolated from the library to try and confirm one of the clones as *VCX*. Primers conserved between human *VCX* and *VCY* were designed. Cycling parameters with 35 cycles of (94°C, 1' / 55°C ± 5°C, 1' / 72°C, 1'30'') following a 2' denaturation at 94°C were used to try and amplify fragments of varying expected size. There were no products of expected size amplified. At lower annealing temperatures there was some non-specific amplification.

VCX/Y must therefore have diverged between human and tammar wallaby enough so that primers conserved between human *VCX* and *VCY* were ineffective. However, this did not explain why no *VCX/Y* clones were isolated from the tammar wallaby genomic DNA library.



Lane	Restriction enzymes		Lane	Restriction enzymes	
1	*Sal*I		6	*Sal*I	*Xba*I
2	*Eco*RI		7	*Sal*I	*Xho*I
3	*Xba*I		8	*Eco*RI	*Xba*I
4	*Xho*I		9	*Eco*RI	*Xho*I
5	*Sal*I	*Eco*RI	10	*Xba*I	*Xho*I

Figure3.13: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammsr wallaby *VCX* clones #1.3 and 1.4. b) Southern blot of restriction endonuclease gel after hybridisation of the human *VCX* probe. The autoradiograph was exposed overnight.

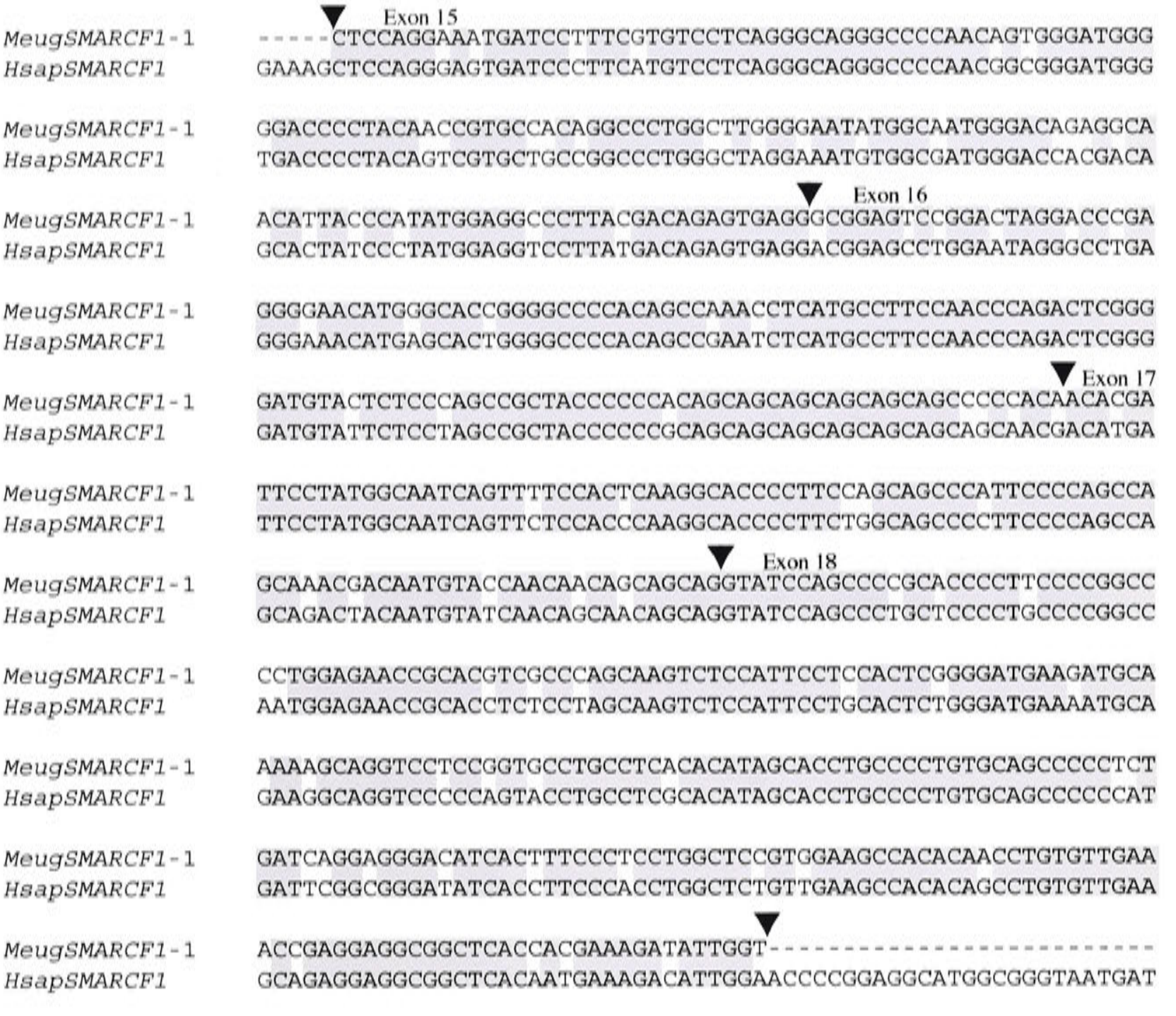


Figure 3.14: Sequence alignment of human *SMARCF1* and tammar wallaby *SMARCF1* exons 15-18. Intron positions (indicated with arrows) were conserved between species. No homology was evident in introns. High homology was evident in exons, extending from the begining of exon 15 to the end of exon 18. This region was translatable, and an open reading frame was observed in the wallaby sequence.

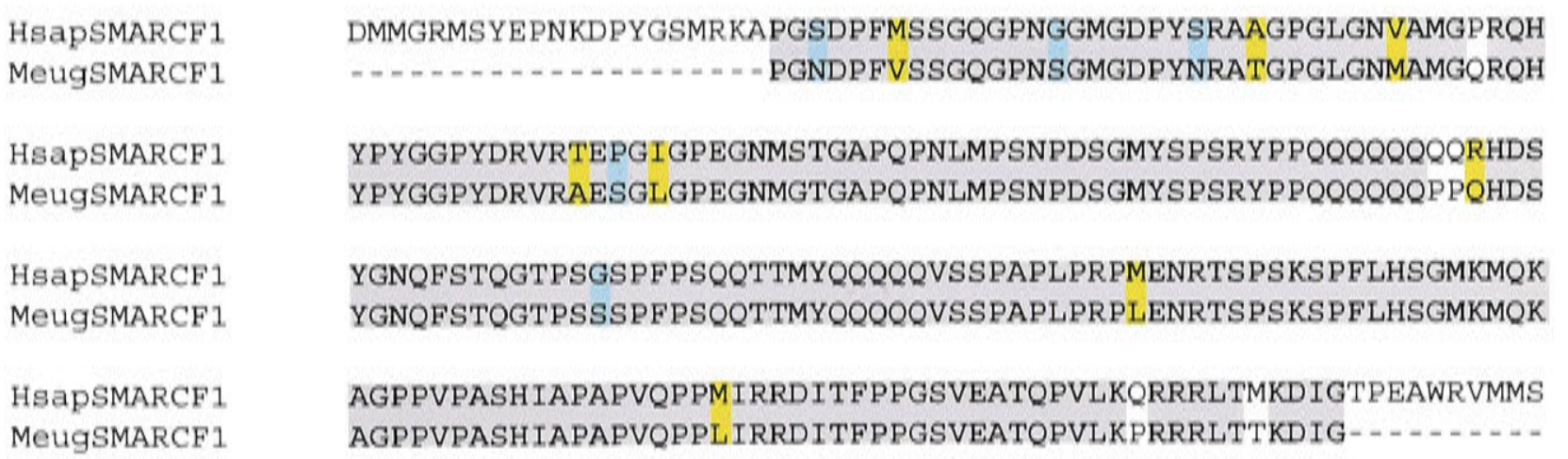


Figure 3.15: Amino acid alignment of human and wallaby *SMARCF1*. Conservative changes are highlighted in yellow. Semi-conservative changes are highlighted in blue.

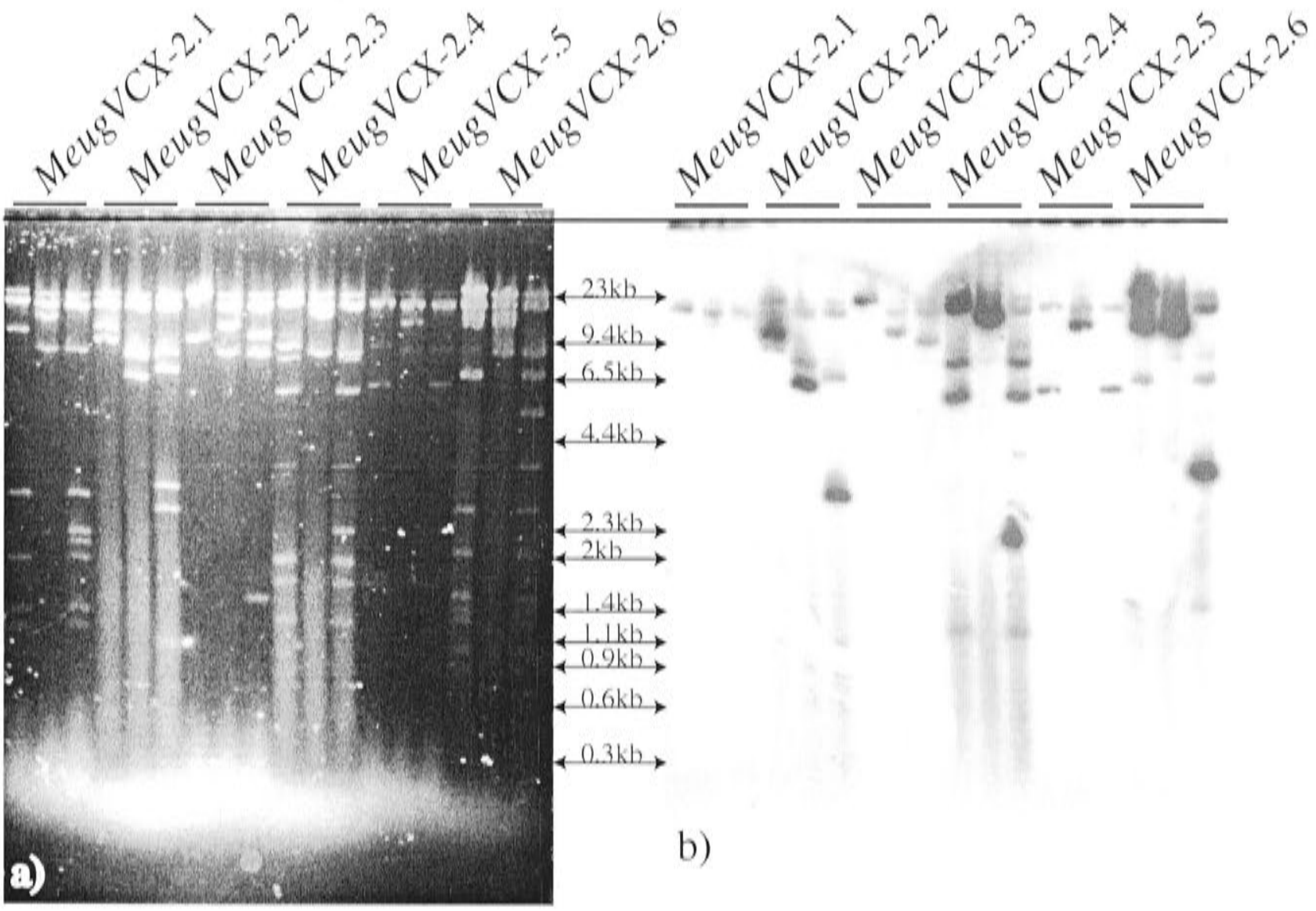


Figure3.16: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammar wallaby *VCX* clones. Each clone was digested with *Eco*RI, *Sal*I and *Eco*RI/*Sal*I respectively. b) Southern blot of restriction endonuclease gel after hybridisation of the human *VCX* probe. The autoradiograph was exposed overnight.

3.4.3 Localisation of wallaby *SMARCF1* by fluorescent *in situ* hybridisation

SMARCF1 (SWI/SNF-related, matrix associated, actin-dependent regulator of chromatin, subfamily f, member 1) is a member of proteins distinguished by a DNA binding motif call ARID (AT-rich interactive domain) (Dallas *et al.*, 2000). SWI/SNF complexes have been suggested to play fundamental roles in the regulation of gene expression during cell growth and development via altering chromatin structure (Kingston and Narlikar, 1999; Kadonaga, 1998). *SMARCF1* maps to human chromosome 1p36.1-p35 (Takeuchi *et al.*, 1998), where it has been suggested that at least two tumour suppressor genes are located (Caron *et al.*, 1995) for which *SMARCF1* is a candidate. *SMARCF1* spans approximately 86kb of genomic DNA with 20 exons and 19 introns. The *SMARCF1* mRNA is approximately 7.9kb and codes for a protein of 2068 amino acids (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly).

The sequence obtained from wallaby *MeugVCX*-1.4 (renamed *MeugSMARCF1*-1), which was homologous to human *SMARCF1*, spaned exons 15 through to 18. Intron positions were conserved and an open reading frame was observed.

MeugSMARCF1-1was labelled for five hours in a nick translation reaction with biotin. 200ng of the labelled probe was precipitated with 25µg of suppressor DNA and a further 200ng of the labelled probe was precipitated with 60µg of ssDNA. The two precipitation reactions containing the biotin labelled probe were hybridised to tammar wallaby metaphase chromosomes prepared from fibroblast cells. The slides were pretreated with RNase and pepsin. The hybridisation was conducted at 37°C for 48 hours. After hybridisation the slide was washed to remove excess probe. Biotin labelled probe that hybridised to the target sequences was detected with two antibody layers: a goat anti-biotin antibody followed by a rabbit anti-goat+FITC conjugate. The slides were visualised as described in section 2.5.4.

Under these conditions, consistent hybridisation signals were observed on the distal region of tammar wallaby chromosome 5q (Figure 3.17). *SMARCF1* is the first gene to be mapped to this region of the wallaby genome, and represents a new region of homology between the short arm of human chromosome 1 and the long arm of wallaby chromosome 5.

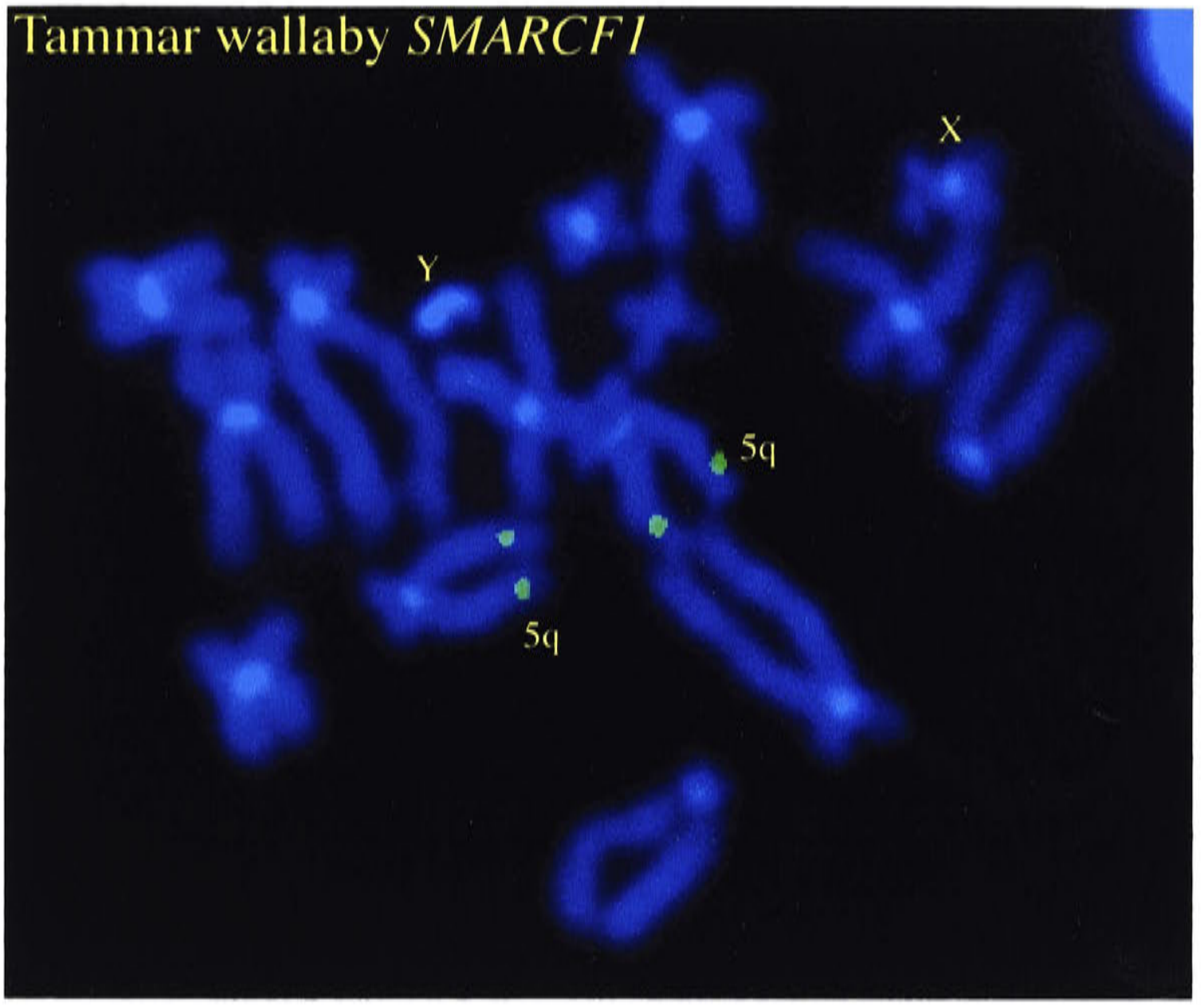


Figure 3.17: FISH mapping of Meug*SMARCF1*-1. Hybridisation signals were detected in the distal region of the long arm of chromosome 5.

3.4.4 Summary

Southern blot analysis of *VCX* revealed that there were homologous sequences in the tammar wallaby genome. There was no indication of female dosage and no male specific bands, indicating that wallaby sequences homologous to *VCX* were autosomal. However, screening of a male genomic DNA library yielded no clones containing a *VCX/Y* orthologue. A clone containing the unrelated gene *SMARCF1* was isolated and mapped to the long arm of chromosome 5.

A partial human *VCY* cDNA was used as a probe to screen the genomic DNA library. A Y chromosome derived cDNA is not the ideal probe to use, as they are likely to be less conserved between species than an X chromosome derived cDNA. This could explain why no wallaby clones containing *VCX/Y* were isolated. However, human *VCX* and *VCY* share high sequence homology (95%).

There is no sequence homology shared between *VCX* and *SMARCF1*, so the isolation of *SMARCF1* from the library was probably due to non-specific hybridisation.

3.5 Tammar wallaby *PCDHX/Y*

Protocadherin genes are members of the cadherin superfamily, and are thought to be involved in cell-cell recognition in the central nervous system. Cadherin genes can be divided into two groups, classical and non-classical. Protocadherin genes are the major non-classical cadherins (Suzuki, 1996). Human protocadherin on the X chromosome (*PCDHX*) is located in the X-Y homology region of Xq21.3 (Yoshida and Sugano, 1999). The location of *PCDHX* in this region of the conserved human X chromosome predicts that it should map to the marsupial X chromosome. However, Blanco *et al.* (2000) suggested that *PCDHX* was part of a recent addition to the eutherian X. Since *PCDHY* was the result of a very recent (3-4MyBP) transposition to the human Y, a conserved Y-borne partner would be unexpected.

Human *PCDHY* is located at Yp11.2 and shares 98.1% nucleotide and 98.3% amino acid homology with *PCHDX*. They have an identical gene structure of six exons and five introns. *PCDHX* spans about 100kb of genomic DNA. The *PCDHX* mRNA is about 4.7kb and codes for a protein of 1025 amino acids (Blanco *et al.*, 2000). Both

PCDHX and *PCDHY* are predominantly expressed in the brain with differential regional expression, consistent with the idea that protocadherins play a role in the central nervous system. The X and Y-linked homologues are expressed at differing levels in different tissues. Both X and Y transcripts are evident in the heart and the subregions of the brain, except cerebellum (predominantly *PCDHX*). Kidney, liver, muscle and testis all have predominantly *PCDHY* transcripts, indicating that *PCDHX* and *PCDHY* are differentially regulated (Blanco *et al.*, 2000).

Southern blot analysis was conducted to determine if there were homologous sequences present in the wallaby genome. Female dosage and/or male specific bands would indicate that wallaby *PCDHX/Y* would be located on the X and/or Y chromosome. A male tammar wallaby genomic DNA library was screened to isolate *PCDHX/Y* clones that could be mapped by FISH.

3.5.1 Southern blot analysis of wallaby *PCDHX/Y*

Primers were designed form sequence conserved between human *PCDHX* and *PCDHY* in order to amplify a human probe to perform a Southern blot analysis and screen the tammar wallaby genomic DNA library. The primers PCDHX-2020F and PCDHX-2735R (Appendix 1) were used to achieve this. The cycling parameters to amplify the expected human *PCDHX* product of approximately 700bp were 30 cycles of (94°C, 1' / 55°C, 1' / 72°C, 1'30'') following a 2' denaturation at 94°C. The PCR product was subcloned into pGEM T Easy and its identity confirmed by sequencing from the vector primers.

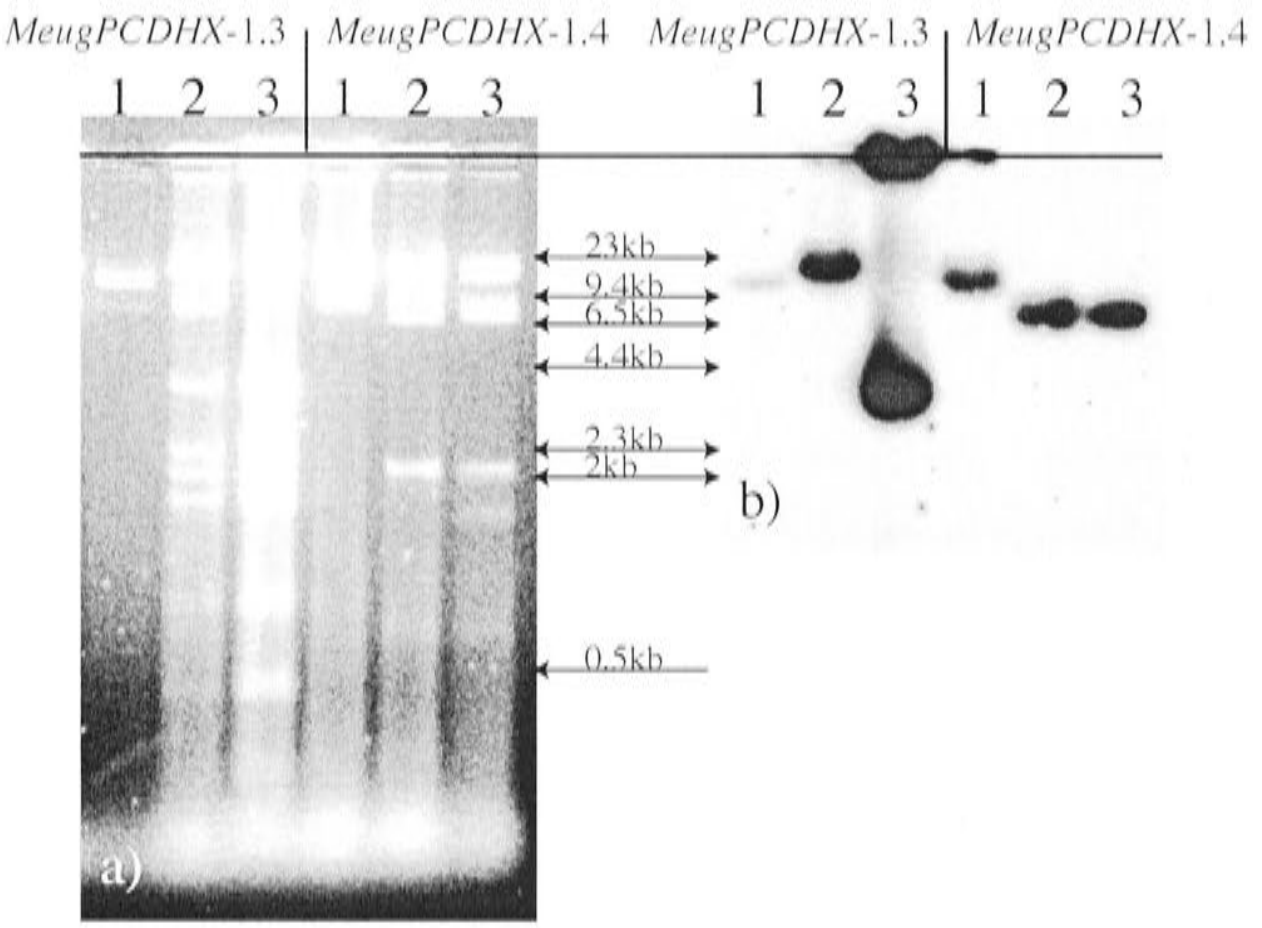
The *PCDHX* probe was radioactively labelled with ^{32}P dCTP and hybridised to tammar wallaby genomic DNA Southern blots. The overnight hybridisation and washing was conducted at low stringency (55°C). The Southern blot membrane was exposed to X-ray film for seven days. Under these conditions very poor hybridisation was detected. There was no indication of bands either shared between male and female, dosed or male specific. Despite the poor hybridisation of the *PCDHX/Y* probe to the Southern blot, isolation of genomic clones from a tammar wallaby genomic DNA library was attempted.

3.5.2 Attempts to clone and characterise wallaby *PCDHX/Y*

The human *PCDHX* probe was used to screen the tammar wallaby genomic DNA library at low stringency (hybridisation and washes conducted at 55°C). The X-linked homologue (rather than the Y-linked homologue) was used as a probe because it was more likely to be conserved between species. Under these conditions six primary clones were picked from the library. After a further three rounds of screening, two final clones were isolated. DNA was prepared from the two final clones and digested for 4 hours with the restriction endonucleases *Eco*RI and *Sal*I. The fragments generated were separated by electrophoresis on a 0.8% agarose gel. A Southern blot was performed using the same probe used to isolate the clones from the library, also under low stringency conditions. *MeugPCDHX*-1.3 had a hybridising band generated by *Eco*RI and *Sal*I that was approximately 3kb in size. *MeugPCDHX*-1.4 did not have any hybridising bands of manageable size to subclone (Figure 3.18).

The 3kb hybridising band from *MeugPCDHX*-1.3 was subcloned into pBluescriptSK$^+$ and sequenced using the vector primers. Sequencing revealed that *MeugPCDHX*-1.3 contained part of exon 2 of the gene *PCDH13*, an autosomal relative of *PCDHX* (Figure 3.19a & b). An open reading frame was observed in the *PCDH13* sequence and 90-95% homology was observed at the amino acid level between human and wallaby (Figure 3.20). *PCDH13* and PCDHX display poor homology to each other with only about 50% sequence identity.

The second clone contained different restriction fragments, so an effort to characterise *MeugPCDHX*-1.4 was made in the anticipation that it could contain wallaby *PCDHX*. *MeugPCDHX*-1.4 was digested with a variety of restriction endonucleases to identify a hybridising fragment that could be subcloned. A hybridising fragment generated by *Xba*I and *Bgl*II, which was approximately 1.7kb, was identified and subcloned into pBluescriptSK$^+$. Sequencing with the vector primers revealed that this clone was also *PCDH13*. The primers PCDHX4(508) and PCDHX4(566) (Appendix I) were designed from the sequence generated by the vector primers to walk into the subclone and construct a contig (Figure 3.19a & b).



Lane	Restriction enzymes
1	*Sal*I
2	*Eco*RI
3	*Sal*I *Eco*RI

Figure3.18: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammsr wallaby *PCDHX* clones #1.3 and #1.4. b) Southern blot of restriction endonuclease gel after hybridisation of the *PCDHX* probe. The autoradiograph was exposed overnight.

```
MeugPcdh13      ----------------------------AACGGGATCACCTACTCCTACAGCCAGAAGGTAC
HsapPCDH13      CTGTGGATAAAGACTTGGGGACCAATGCTCAAATTACTTATTCTTACAGTCAGAAAGTTC

MeugPcdh13      CACAGGTATCAAAAGACTTGTTCTACCTGGATGAAATCACTGGAGTCATTAAACTCTTCA
HsapPCDH13      CACAAGCATCTAAGGATTTATTTCACCTGGATGAAAACACTGGAGTCATTAAACTTTTCA

MeugPcdh13      AAAAGATTGAGGGCAATGTTCTTCGACTACATAAGCTCACTATACTTGCCAATGGACCTG
HsapPCDH13      GTAAGATTGGAGGAAGTGTTCTGGAGTCCCACAAGCTCACCATCCTTGCTAATGGACCAG

MeugPcdh13      GCTGCATTCCAGCTGTAATCACTGTCCTGGTAACCGTCATCAAAGTAGTTTTTCGTCCAC
HsapPCDH13      GCTGCATCCCTGCTGTAATCACTGCTCTTGTGTCCATTATTAAAGTTATTTTCAGACCCC

MeugPcdh13      CTGAGATAGTTCCTCGTTATATTGCTAATGAAGTTGAAGGAGTTGTTTACCTAAAGGAGT
HsapPCDH13      CTGAAATTGTCCCTCGTTACATAGCAAACGAGATAGATGGTGTTGTTTATCTGAAAGAAC

MeugPcdh13      TGGAACCCATCAATACCCCAGTTGCATTTTTTACTATCAGAGATCCAGAAGATAAATATA
HsapPCDH13      TGGAACCCGTTAACACTCCCATTGCGTTTTTCACCATAAGAGATCCAGAAGGTAAATACA

MeugPcdh13      AAGTGAATTGCCACCTGGATGGAGATGGACCATTTCGATTATCTCCCTATCAACCATACA
HsapPCDH13      AGGTTAACTGCTACCTGGATGGTGAAGGGCCGTTTAGGTTATCACCTTACAAACCATACA

MeugPcdh13      ATAATGAATACTTATTAGAAACTGCAAAATCTTTGGATTATGAGACACAGCAGCTCTATG
HsapPCDH13      ATAATGAATATTTACTAGAGACCACAAAACCTATGGACTATGAGCTACAGCAGTTCTATG

MeugPcdh13      AGATAACTGTGGTGGCCTGGAACTCAGAGGGATTTCATGTTAAGAAGGTGATTAAAATAC
HsapPCDH13      AAGTAGCTGTGGTGGCTTGGAACTCTGAGGGATTTCATGTCAAAAGGGTCATTAAAGTGC

MeugPcdh13      AGATTCTAGATGACAATGACAATGCACCTGTTTTCACTCAACCACTGATAGAAGTATCTA
HsapPCDH13      AACTTTTAGATGACAATGATAATGCTCCAATTTTCCTTCAACCCTTAATAGAACTAACCA

MeugPcdh13      TTGAAGAAAATAATGCACCAAATGCCTTTCTGACAAAGTTGCATGCCATAGATGCTGACA
HsapPCDH13      TCGAAGAGAACAACTCACCCAATGCCTTTTTGACTAAGCTGTATGCTACAGATGCCGACA

MeugPcdh13      GTGGAGAAAGAGGTCAAGTTTCTTATTTCCTTGGAGCTGATGCACCATCATATTTTTCAT
HsapPCDH13      GCGAGGAGAGAGGCCAAGTTTCATATTTTCTGGGACCTGATGCTCCATCATATTTTTCCT

MeugPcdh13      TAGACAAGATTACAGGGATTCTTACTGTGTCCACTCAGTTGGACAGAGAAGAAAAAGAGA
HsapPCDH13      TAGACAGTGTCACAGGAATTCTGACAGTTTCTACTCAGCTGGACCGAGAAGAGAAAGAAA

MeugPcdh13      GGTACAGGTACACAGTAAAAGCAGTTGACTCTGGGATACCACCTCAAGAGTCAATAGCTA
HsapPCDH13      AGTACAGATACACTGTCAGAGCTGTTGACTGTGGGAAGCCACCCAGAGAATCAGTAGCCA

MeugPcdh13      CAGTTGCCATCACGGTGTTGGACAAAAATGACAATAGCCCTAGGTTTATCAACAAGGATT
HsapPCDH13      CTGTGGCCCTCACAGTGTTGGATAAAAATGACAACAGTCCTCGGTTTATCAACAAGGACT

MeugPcdh13      TCAGCTTTTTTGTGCCAGAAAATTTCCCAGGGTTTGGTGAAATTGGAGTAATCAGTGTCA
HsapPCDH13      TCAGCTTTTTTGTGCCTGAAAACTTTCCAGGCTATGGTGAGATTGGAGTAATTAGTGTAA

MeugPcdh13      CAGATGCAGATGCAGGACGGAATGGATGGGTTGCCCTTTCAGTGATGAATCAGAGTGACA
HsapPCDH13      CAGATGCTGACGCTGGACGAAATGGATGGGTCGCCCTCTCTGTGGTGAACCAGAGTGATA

MeugPcdh13      TTTTTGTCATAGACACTGGAAAAGGCATGTTGAGAGCAAAAGTCTCTCTGGATAGGGAGC
HsapPCDH13      TTTTTGTCATAGATACAGGAAAGGGTATGCTGAGGGCTAAAGTCTCTTTGGACAGAGAGC

MeugPcdh13      AACAAAGTTCCTATATTTTGTGGGTTGAAGCTGTTGATGGAGGTGAACCTGCCCTCTCCT
HsapPCDH13      AGCAAAGCTCCTATACTTTGTGGGTTGAAGCTGTTGATGGGGGTGAGCCTGCCCTCTCCT

MeugPcdh13      CTACAGCAAAAATAACAATTCTTCTTCTTGATATCAATGACAACCCTCCTCTTGTCCTGT
HsapPCDH13      CTACAGCAAAAATCACAATTCTCCTTCTAGATATCAATGACAACCCTCCTCTTGTTTTGT

MeugPcdh13      TTCCTCAGTCAAATATGTCTTATCTATTGGTCCTACCTTCTACTCTACCTGGCTCACCAG
HsapPCDH13      TTCCTCAGTCTAATATGTCTTATCTGTTAGTACTGCCTTCTACTCTGCCAGGCTCCCCGG

MeugPcdh13      TTACAGAAGTCTATGCTGTTGACAAAGACACTGGTATGAATGCAGTCATAGCTTATAGCA
HsapPCDH13      TTACAGAAGTCTATGCTGTCGACAAAGACACAGGCATGAATGCTGTCATAGCTTACAGCA
```

```
MeugPcdh13      TCATAGGAAGAAGAGGTCCTCGGCCTGAATCCTTTAAGATAGACCCCAAAACTGGTAACA
HsapPCDH13      TCATAGGGAGAAGAGGTCCTAGGCCTGAGTCCTTCAGGATTGACCCTAAAACTGGCAACA

MeugPcdh13      TTACTTTGGAAGAGACATTGATGCGGAATGATTATGGGCTTTATCGCTTGCTGGTTAAAG
HsapPCDH13      TTACTTTGGAAGAGGCATTGCTGCAGACAGATTATGGGCTCCATCGCTTACTGGTGAAAG

MeugPcdh13      TGAGTGACCACGGTTATCCTGAGCCCCTCTACTCTACTGTCATGGTAAATCTATTTGTCA
HsapPCDH13      TGAGTGATCATGGTTATCCCGAGCCTCTCCACTCCACAGTCATGGTGAACCTATTTGTCA

MeugPcdh13      ATGACACTGTTAGCAATGAGAGCTACATTGAGAGTCTATTGCGAAAAGAACCTGAGATCA
HsapPCDH13      ATGACACTGTCAGTAATGAGAGTTACATTGAGAGTCTTTTAAGAAAAGAACCAGAGATTA

MeugPcdh13      GGATTGAGGAAAAAGAACCACAAATTTCAATGGAACCTACTCACCGAAAAGTGGAATCAG
HsapPCDH13      ATATAGAGGAGAAAGAACCACAAATCTCAATAGAACCGACTCATAGGAAGGTAGAATCTG

MeugPcdh13      CCTCTTGTGTGCCCACTTTAGTCGCCCTGTCTGTGATAAGTTTGGGGTCCATCACATTAG
HsapPCDH13      TGTCTTGTATGCCCACCTTAGTAGCTCTGTCTGTAATAAGCTTGGGTTCCATCACACTGG

MeugPcdh13      TCACTGGGATGGGAATATACATATGTCTAAGGAAGGGGAAAAAGCATCACCAGGAAAATG
HsapPCDH13      TCACAGGGATGGGCATATACATCTGTTTAAGGAAAGGGGAAAAGCATCCCAGGGAAGATG

MeugPcdh13      ACAATTTGGAAGTACAAATTCCACTGAAAGGAAAGCTTGACTTACACATGATAGAGAGGA
HsapPCDH13      AAAATTTGGAAGTACAGATTCCACTGAAAGGAAAAATTGACTTGCATATGCGAGAGAGAA
                                                STOP
MeugPcdh13      AACCAATGGAGATTTCTAATATTTGATGTTTCCAAGGGAAAAAATACTAAGACAGATATT
HsapPCDH13      AGCCAATGGATATTTCTAATATTTGATATTTCATGGTGGAATAACACAGAGAAATGTTTT

MeugPcdh13      ---TGGCCTTGGAGAGGCCACAAGCATAATCATAAGTGTGCTGATTGGCAGTTCCAATGA
HsapPCDH13      AACTGACTTTGGATCTTCATCA--CCTAAAAAAGAGTGTGTTGAT-GGCAGTTCCAATGA

MeugPcdh13      AGGACAAATAATTTATAAATTATATTATATTGTAAATAACTGTTTACAAGTTTTATAAAA
HsapPCDH13      AGGACAACTAATTTATAACTTGTTCTATATTGTAAATAGCTGTTTACAGGTTTT--TAAA

MeugPcdh13      TTCAAATTCAGGTCTTATAAAATGTGTACAGCATTTTTTAAAATGAAAAATTAGTATTAA
HsapPCDH13      TTTAAATTCAGAGGTTATAAAATGTGTACAGCATTTTT--AAGTGAAAA-TTAGTACTAA

MeugPcdh13      CAGTTAGTGTACTGCTGAT---------------AAACGGCTTTC-CATCATTTGCAGCT
HsapPCDH13      CAGCTATAGGACTTGTATTTAAAAAAAAAAAAAAAAAAGCTTGGACATGGTTTGCAGCT

MeugPcdh13      TTCATACATCAACCAGGTCATTGAACA-------GAGTAAGGTAAGACATT----ACAG-
HsapPCDH13      TTCATACACCAAGCAGTTGATTGATAAAACCTGGGAGTAAGGTAAGAAAAATGGAACAAA

MeugPcdh13      --TTTGTCTAAGAAGTTATTTTAGCCACCAGAGGGCATCAGTTGCTCCCATGTAGGAAGA
HsapPCDH13      TTTTTATCTAA-AAATTCCTGTCACCACAAGAGGGCATCAGCTGCTCCTTTGCAGGAAAC

MeugPcdh13      TG---TTTGAGATCCCCCCGGGCTGCAGGAATTCGAT----------------------
HsapPCDH13      TGGGGTATTGTACTTGGCAGTTGTACATGAAATTAATGAAAGAGTATATTTTAAATATAT
```

Figure 3.19a: Sequence alignment of human and tammar wallaby *PCDH13* exon 2. No introns are present in this region of human (or tammar) *PCDH13*. This sequence was obtained from *MeugPCDH13*-1.3, *MeugPCDH13*-1.4 and *MeugPCDH13*-2.5. An open reading frame was observed up until the stop codon, after which homology between human and wallaby *PCDH13* decreased.

```
MeugPcdh13      ----------------------------------------------------------GCTG
HsapPCDH13      TACTGGTTGTTTTAGTTGGGCCTTTTATATGAAAGTTATACTGTACTTATCTTTTTGTTG

MeugPcdh13      TTGTTTAGGTTAACTGGCTACTTTTTCTGGCTTTTGTTCTTTGGCTTTGGTTTAGGGCCA
HsapPCDH13      TTGTTTAGGTTTGTTGACTACCATTTCTGGCTTCCTTTCTTTGGTTTTGGATTAAATCCA

MeugPcdh13      AGAGCTGCCTTTTAAGCCTGGTGGCAGATTTCTGGGATAATTTATTCTTTACTACATCCT
HsapPCDH13      ACAGTTTATTTGTAAGCCCTGCAGCAGATTTCTTTGATAATTTAGCTTTTACTGAATCTC

MeugPcdh13      TTGCAATGCAAACAGCAGGTTCATCATCGATGAATCCTTTTCACCTCATTGTGCAAGTGG
HsapPCDH13      TTGCAATGAAGAAAGCTATTTCATCAGTGATTTATCACTTTCAATTCATTGTGTGAGCTG

MeugPcdh13      AGAAAAAAATTCTTTCTTATGAGGGTTATGACAAATATATCTATGTAAACAAGGAAAATG
HsapPCDH13      ---GAAATATTATTTTATATGAGAGCTATAGCAAAATAATCTGTATAAACAAGGA--ATG

MeugPcdh13      TGTCAGCTTAAAAATGGGTCATTTTACTTTCCAGCATCTTATGTCTG-ATTGTACCAAAA
HsapPCDH13      TGTTAGCTTAAAC-TGGATCATTTTACTTTTTGGCATCATGCATCTGTACTGTACCAAAA

MeugPcdh13      GTGTTTATATNTCTGCAAATTAAAGCTATATATTTTCATAACTTCTTTCTGTTGGTTACC
HsapPCDH13      GTGTTTATATGTCTGCAAATTAAAGGTATATATTTTCATAATTTCTTTCTCACTTTTAGC

MeugPcdh13      CTTCTGACTATAAGATTCACTGAACTGCTTTCTGAAGGTATCTGTTAGCTCTGTCTTATT
HsapPCDH13      ATTTTATATTTGAACATGGGCTTGTTATTTCACTGAAGTGCCTGTCA--TTTGTTTGTTT

MeugPcdh13      TAAATCTTAAGAAAGTGATATCCTAAAAGCAAATATGTTGAAGAATAGAGTTCACTTTAA
HsapPCDH13      TTAGGGTTAAATGTG-GGTAACCTGTAG--TATTCTGCT--ATACTAAAGTTTATATTAA

MeugPcdh13      AATAGTCAATCCTTACATGTGTGAAGATGTGTGCACATATCTGAAAGA
HsapPCDH13      ATGAATTCATTACTCAAAAA--GAAAATGGGTAAAT-TATTAAAAAAA
```

Figure 3.19b: Sequence alignment of human and tammar wllaby *PCDH13* 3' untranslated region. The sequence was obtained from *MeugPCDH13*-1.3 and *MeugPCDH13*-2.5. This homology is approximately 600bp 3' of the sequence from figure 3.19a.

```
HsapPCDH13      TVYGNATVGTPIAAVQAVDKDLGTNAQITYSYSQKVPQASKDLFHLDENTGVIKLFSKIG
MeugPCDH13      ----------------------------NGITYSYSQKVPQVSKDLFYLDEITGVIKLFKKIE

HsapPCDH13      GSVLESHKLTILANGPGCIPAVITALVSIIKVIFRPPEIVPRYIANEIDGVVYLKELEPV
MeugPCDH13      GNVLRLHKLTILANGPGCIPAVITVLVTVIKVVFRPPEIVPRYIANEVEGVVYLKELEPI

HsapPCDH13      NTPIAFFTIRDPEGKYKVNCYLDGEGPFRLSPYKPYNNEYLLETTKPMDYELQQFYEVAV
MeugPCDH13      NTPVAFFTIRDPEDKYKVNCHLDGDGPFRLSPYQPYNNEYLLETAKSLDYETQQLYEITV

HsapPCDH13      VAWNSEGFHVKRVIKVQLLDDNDNAPIFLQPLIELTIEENNSPNAFLTKLYATDADSEER
MeugPCDH13      VAWNSEGFHVKKVIKIQILDDNDNAPVFTQPLIEVSIEENNAPNAFLTKLHAIDADSGER

HsapPCDH13      GQVSYFLGPDAPSYFSLDSVTGILTVSTQLDREEKEKYRYTVRAVDCGKPPRESVATVAL
MeugPCDH13      GQVSYFLGADAPSYFSLDKITGILTVSTQLDREEKERYRYTVKAVDSGIPPQESIATVAI

HsapPCDH13      TVLDKNDNSPRFINKDFSFFVPENFPGYGEIGVISVTDADAGRNGWVALSVVNQSDIFVI
MeugPCDH13      TVLDKNDNSPRFINKDFSFFVPENFPGFGEIGVISVTDADAGRNGWVALSVMNQSDIFVI

HsapPCDH13      DTGKGMLRAKVSLDREQQSSYTLWVEAVDGGEPALSSTAKITILLLDINDNPPLVLFPQS
MeugPCDH13      DTGKGMLRAKVSLDREQQSSYILWVEAVDGGEPALSSTAKITILLLDINDNPPLVLFPQS

HsapPCDH13      NMSYLLVLPSTLPGSPVTEVYAVDKDTGMNAVIAYSIIGRRGPRPESFRIDPKTGNITLE
MeugPCDH13      NMSYLLVLPSTLPGSPVTEVYAVDKDTGMNAVIAYSIIGRRGPRPESFKIDPKTGNITLE

HsapPCDH13      EALLQTDYGLHRLLVKVSDHGYPEPLHSTVMVNLFVNDTVSNESYIESLLRKEPEINIEE
MeugPCDH13      ETLMRNDYGLYRLLVKVSDHGYPEPLYSTVMVNLFVNDTVSNESYIESLLRKEPEIRIEE

HsapPCDH13      KEPQISIEPTHRKVESVSCMPTLVALSVISLGSITLVTGMGIYICLRKGEKHPREDENLE
MeugPCDH13      KEPQISMEPTHRKVESASCVPTLVALSVISLGSITLVTGMGIYICLRKGKKHHQENDNLE

HsapPCDH13      VQIPLKGKIDLHMRERKPMDISNI
MeugPCDH13      VQIPLKGKLDLHMIERKPMEISNI
```

Figure 3.20: Amino acid alignment of human and wallaby *PCDH13*, exon 2. Conservative changes are highlighted in yellow. Semi-conservative changes are highlighted in blue.

A different plating of the same tammar wallaby genomic DNA library was screened with a low stringency hybridisation (55°C) and washed at a higher stringency (65°C) than the previous screening attempt. This would reduce the probability of detecting related genes. Seven primary clones were isolated from the library, all were purified to final clones after a further three rounds of screening. DNA was prepared from the final clones, which were digested with the restriction endonucleases *Eco*RI and *Sal*I. A Southern blot was performed on the 0.8% agarose gel used to separate the fragments generated. Hybridisation of the *PCDHX* probe to the Southern blot identified hybridising bands for all clones (Figure 3.21).

Clone *MeugPCDHX*-2.2 and *MeugPCDHX*-2.4 had the same restriction pattern and were therefore identical clones. The clones *MeugPCDHX*-2.3 and *MeugPCDHX*-2.7 did not have hybridising fragments generated by *Eco*RI and/or *Sal*I that were small enough to subclone and were not further characterised. The smallest hybridising fragments from each remaining clone were pursued. *Eco*RI generated fragments from *MeugPCDHX*-2.1 and *MeugPCDHX*-2.4, of about 3.5kb and 2.5kb respectively, were subcloned into pBluescriptSK$^+$. *Eco*RI and *Sal*I generated hybridising fragments from *MeugPCDHX*-2.5 and *MeugPCDHX*-2.6 of about 3.5kb in size, these were also subcloned into pBluescriptSK$^+$. The inserts were all sequenced using the vector primers. Nucleotide blast searches revealed that *MeugPCDHX*-2.1 and *MeugPCDHX*-2.4 had no significant homology to anything in the database. *MeugPCDHX*-2.5 was shown to contain *PCDH13* (Figure 3.19a & b).

However, a nucleotide blast search of the *MeugPCDHX*-2.6 sequence revealed that there was a small region of homology (22bp) to *PCDHX* in exon five. Exon five is the largest exon at 2493bp, so if *MeugPCDHX*-2.6 contained the true wallaby orthologue much more sequence, homologous to exon five, should be present.

To completely sequence this region of homology to human *PCDHX*, *MeugPCDHX*-2.6 was shotgun-cloned. Sequencing of the relevant recombinant plasmids after shotgun-cloning allowed a contig to be constructed. The contig contained about 150bp of sequence both upstream and downstream of the small *PCDHX* hit, and contained no exonic sequence other than the 22bp hit to human *PCDHX*. This indicated that *MeugPCDHX*-2.6 did not contain the true *PCDHX*.

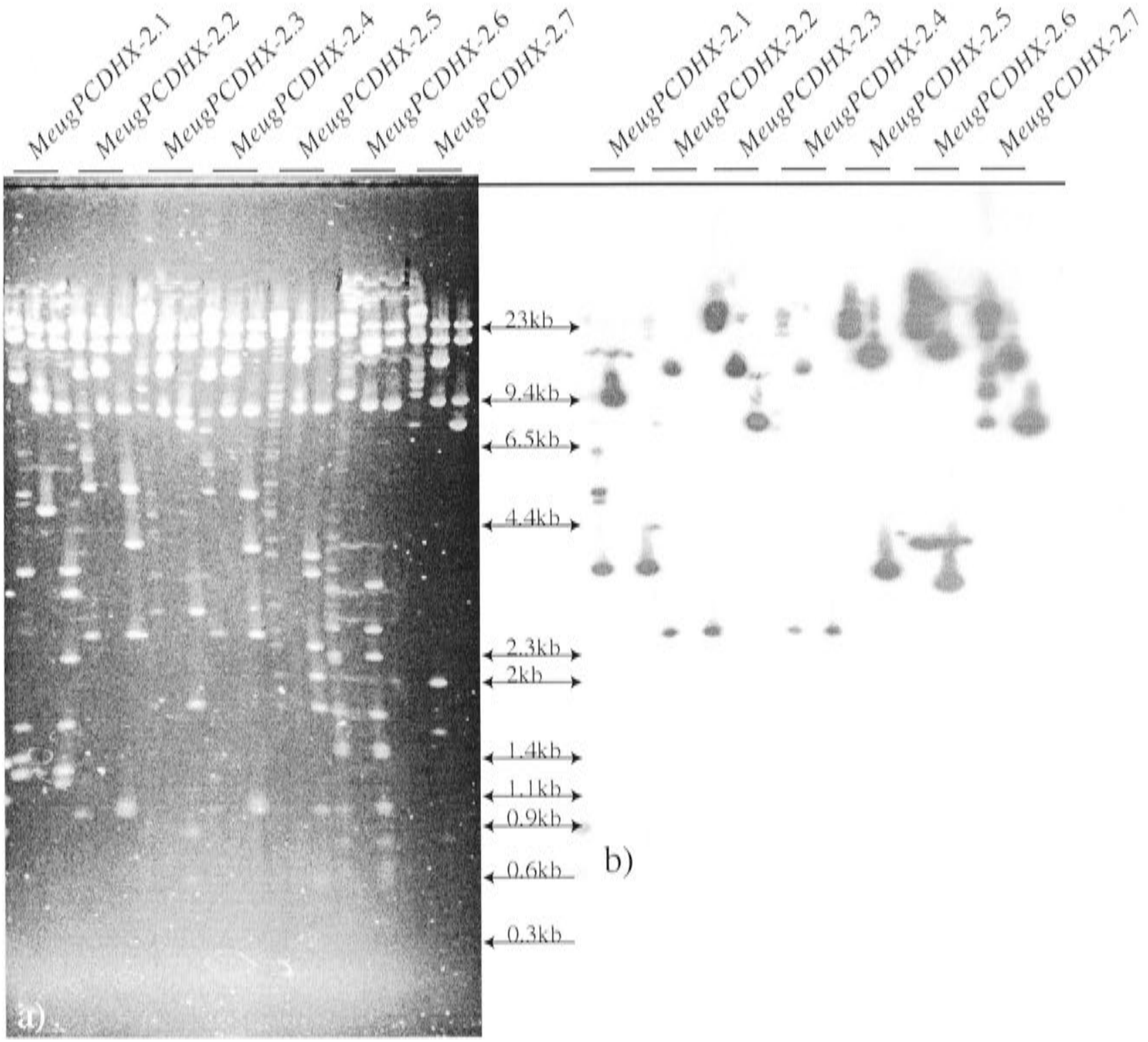


Figure3.21: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of tammar wallaby *PCDHX* clones. Each clone was digested with *Eco*RI, *Sal*I and *Eco*RI/*Sal*I respectively. b) Southern blot of restriction endonuclease gel after hybridisation of the *PCDHX* probe. The autoradiograph was exposed for 2hrs.

3.5.3 Mapping of wallaby *PCDH13* by FISH

Along with *PCDHX*, *PCDH13* is a member of the cadherin superfamily that are thought to be involved in cell-cell recognition in the central nervous system. Mouse *Pcdh13* is expressed in the developing brain along with other protocadherin genes (Hirano *et al.*, 1999). In humans, *PCDH13* maps to chromosome 13q21 (Yagi and Takeichi, 2000). It contains only one intron of about 1kb. The mRNA is approximately 4.5kb and codes for a protein of 925 amino acids (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). *PCDH13* is also known to map to mouse chromosome 18, rat chromosome 18p12-p11 and Chinese hamster chromosome 2q17 (Hirano *et al.*, 1999; Ono *et al.*, 2000).

The clones *MeugPCDHX*-1.3 and *MeugPCDHX*-1.4 were labelled with biotin in a nick translation reaction at 16°C for 4 hours. The labelled probe was precipitated with 40μg of suppressor DNA and 60μg of ssDNA. The biotin labelled probe was hybridised to tammar wallaby metaphase chromosomes prepared from fibroblast cells. The slides were pretreated with RNase and pepsin to remove excess RNA and protein. The hybridisation was conducted at 37°C for 48 hours. After hybridisation, excess probe was washed off with 50% formamide/50% 2 X SSC at 37°C. Biotin labelled probe that had hybridised to the target sequence on the metaphase chromosomes was detected with two antibody layers: a goat anti-biotin followed by a rabbit anti-goat+FITC conjugate.

Signals were consistently detected on the long arm of chromosome 6 for both *MeugPCDHX*-1.3 and *MeugPCDHX*-1.4 (Figure 3.22). *PCDH13* is the first gene, from human chromosome 13 to be mapped in tammar wallaby, and represents a new region of homology between human chromosome 13 and tammar wallaby chromosome 6.

Secondary signals were also observed near the centromere on the long arm of the wallaby X chromosome, implying that a related sequence must also exist at this location, probably another protocadherin gene, and perhaps *PCDHX*.

3.5.4 Summary

Southern blot analysis of wallaby *PCDHX/Y* indicated that the wallaby genome contained only sequences that were poorly homologous. Screening of a male genomic DNA library yielded no clones that contained *PCDHX/Y*. However, a related gene,

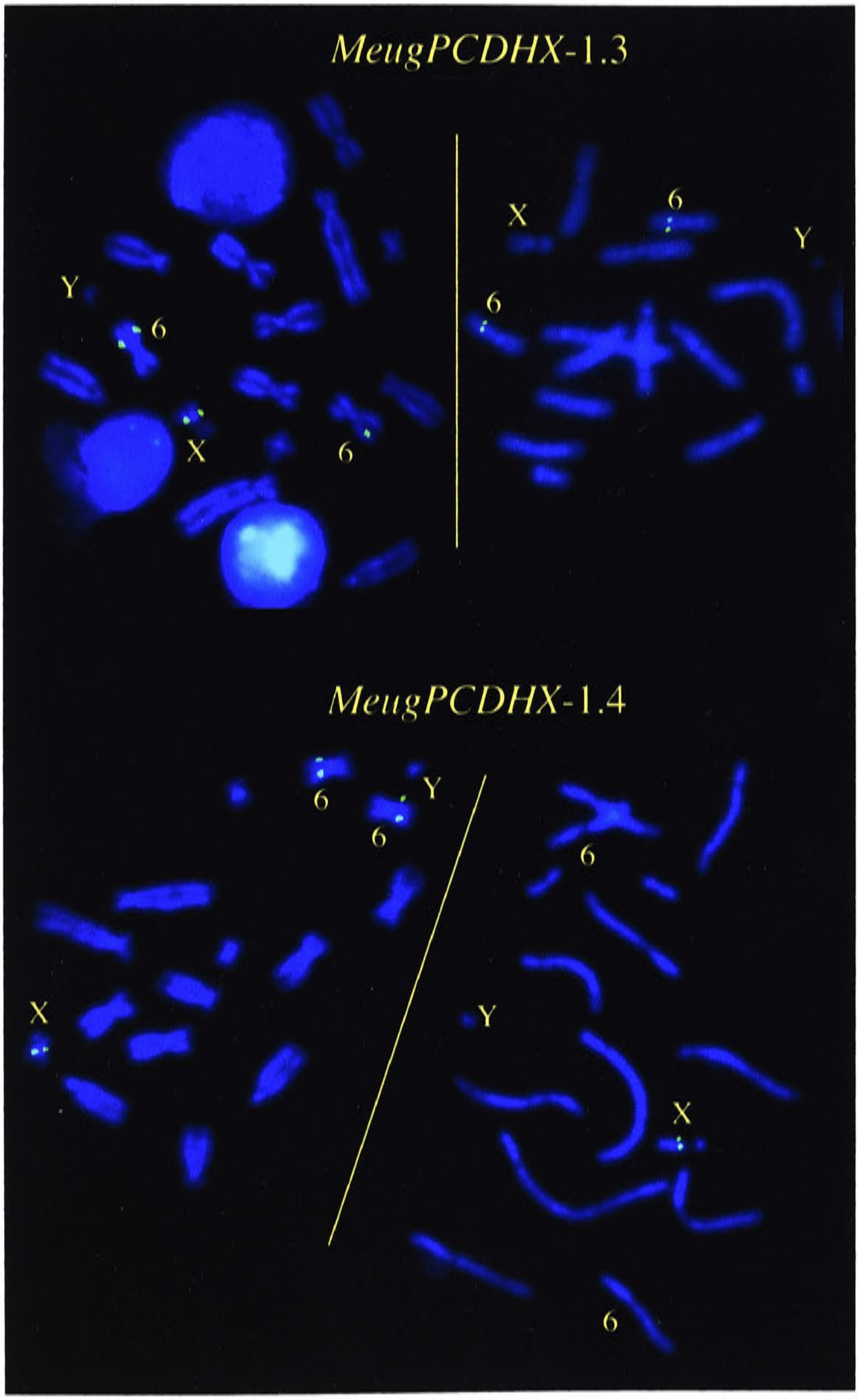


Figure 3.22: FISH mapping of a) *MeugPCDHX*-1.3 and b) *MeugPCDHX*-1.4.Both clones had hybridisation signals on the long arm of chromosome 6. Secondary signals

PCDH13, was repeatedly isolated and mapped to chromosome 6q. The faint hybridisation, detected on the Southern blot, may therefore have represented cross hybridisation to tammar wallaby *PCDH13*.

Human *PCDHX* maps to a conserved region of the X chromosome (Xq21.3), but was suggested to be a recent addition to the human X (Blanco *et al.*, 2000). *PCDHY* is present on the Y in gorilla, orangutan and humans, but in no other primates. Blanco (2000) concluded that there was an X-Y translocation including *PCDHY* in the earlier Great Apes. *PCDHY* was lost from the Y chromosome in the ancestor of the human and chimpanzee lineages, after divergence from gorilla. *PCDHY* was then added again to the human Y chromosome in another X-Y translocation that included all of the Xq21.3 / Y chromosome homologous region (Blanco *et al.*, 2000). The current copy of human *PCHDY* is the third to have existed in the human lineage. The first *PCDHY* was as a result of an autosomal addition to both the X and Y, which was subsequently lost from the Y chromosome (in accordance with the addition-attrition hypothesis); the second *PCDHY* resulted from an X-Y translocation in early Great Apes, which was lost in the human and chimpanzee ancestor; and the third *PCDHY* was the result of another X-Y translocation in the human lineage. The similar degree of nucleotide homology between the introns and exons of the X- and Y-linked copies shows that they have been separated for only a short period of time. The loss of *PCDHY* a third time is likely as it has been lost twice before, never obtaining a selectable male specific function.

Alternatively, it is possible that *PCDHY* confers a slight advantage to the bearer of that Y chromosome. The expression of *PCDHX/Y* in the brain suggested that it could have a function in this tissue. Although *PCDHX* and *PCDHY* were similar in sequence, the potential difference in control of expression, and the small differences in amino acid sequence, indicated that they could function differently. Different functions for *PCDHX* and *PCDHY* could cause a sexually dimorphic trait that influences brain phenotype. This brain phenotype in males could increase their ability to attract a mate (Blanco *et al.*, 2000).

Cross-species hybridisation to a zoo blot failed to detect *PCDHX* in rodents, armadillo or kangaroo (Blanco *et al.*, 2000), reflecting loss or extensive divergence of the gene between species. However, male: female dosage differences were observed for rabbit, cow, goat, sheep, pig, horse, lemur and both new and old world monkeys

implying that *PCDHX* is on the X chromosome in these species (Blanco *et al.*, 2000). More recently, it was shown that there is a *PCDHX* orthologue in the mouse genome located on the X chromosome. At the sequence level mouse and human *PCDHX* share 85% homology. The inability to detect or clone wallaby *PCDHX/Y* is not surprising when you take into account that previous attempts to detect *PCDHX* in armadillo and kangaroo failed.

The secondary signals on the wallaby X chromosome indicated that there were *PCDH* related sequences in this region. The relationship of these sequences to the *MeugPCDHX*-1.3 and *MeugPCDHX*-1.4 genomic clones was unclear. They could have represented the wallaby orthologue of *PCDHX*, therefore demonstrating that *PCDHX* was a conserved gene on the human X chromosome, contradicting the suggestion that *PCDHX* was recently added to the X.

Other than mouse *PCDHX*, the non-primate gene most closely related to human *PCDHX* is a chicken cadherin mRNA, showing 69% sequence homology and 82% amino acid homology. Over the 144 amino acid coding region, the chicken cadherin gene also shows 68% sequence homology and 79% amino acid homology to human *PCDH9*. Human *PCDHX* and *PCDH9* show 56% sequence homology and 73% amino acid homology to each other. The chicken cadherin gene could represent the precursor to both of these genes in human.

3.6 Tammar wallaby Class II genes *PRY*, *XKRY* and *BPY2*

PTP-BL Related Y (*PRY*), *XK Related Y* (*XKRY*) and *Basic Protein Y2* (*BPY2*) all fall into the Class II genes defined by Lahn and Page (1997). They have testis specific expression, no homologue on the X chromosome and are multicopy on the Y (Figure 1.10).

PRY has similarity to *PTP-BL* (located on an autosome), which codes for a protein tyrosine phosphatase. It has five exons and four introns, and its 1238bp transcript codes for a 148 amino acid protein (Stouffs *et al.*, 2001). Different copies of *PRY* fall into the deletion intervals responsible for AZFb, AZFc and GBY (Chapter 1). Stronger candidates have been identified for all of these phenotypes but *PRY* is yet to be excluded from playing a role in fertility or gonadoblastoma. *XKRY* shows similarity to

XK (located on an autosome). It contains no introns and has a 1579bp transcript that codes for a putative membrane transport protein of 117 amino acids (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). It has a copy that falls into the deletion interval responsible for AZFb. *BPY2* spans 21.2kb of genomic DNA with nine exons and eight introns. It has a 1201bp transcript that codes for a 101-residue protein (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). A copy falls into the AZFc deletion. The roles of these three genes on the human Y chromosome are unclear.

3.6.1 Southern blot analysis of *PRY*, *XKRY* and *BPY2*

Low stringency hybridisation of heterologous human probes for *PRY*, *XKRY* and *BPY2* to tammar wallaby genomic Southern blots was used to detect related sequences, and to determine whether these sequences showed male: female dosage or were male-specific. Male-specific sequences would identify genes that are located on the Y chromosome in marsupials, as they are in human. Evidence of male-female dosage difference would indicate an X-linked locus. Equivalent signals in male and female lanes would indicate an autosomal location.

Probes for *PRY* and *XKRY* were generated by PCR with primers designed from their respective human sequences (PRY-418f, PRY-1138-r, XKRY-695f and XKRY-1481r) (Appendix I). Human male genomic DNA was used as template to generate the expected products of 710bp and 780bp respectively. The cycling parameters used were 35 cycles of (94°C, 1' / 51°C, 1' / 72°C, 1') following a 2 minute denaturation at 94°C. The products were cloned into the plasmid vector pGEM T Easy and the inserts were sequenced using the vector primers to confirm their identity. The PCR generated probes were liberated from their plasmid vectors by digestion for 1 hour at 37°C with the restriction endonuclease *Eco*RI.

A 320bp partial human cDNA probe for *BPY2* was kindly supplied by Dr. David Page (Whitehead Institute, Boston, USA). It was liberated from its plasmid vector with the restriction endonuclease *Not*I. The inserts were purified from the gel after electrophoresis on a 1% agarose gel and were radioactively labelled with ^{32}P dCTP.

The radioactively labelled inserts were hybridised to tammar wallaby genomic Southern blots containing male and female wallaby DNA, fully digested with *Eco*RI

and *Hind*III. The hybridisation was conducted at low stringency (55 C) overnight in Church's buffer The Southern blot membranes were then washed at low stringency.

No clear evidence of specific hybridisation was detected for any of *PRY*, *XKRY* or *BPY1*. Apparent hybridisation was non-specific, permitted by the low stringency hybridisation and washing conditions. Thus, Southern analysis did not reveal any sequences with strong homology to these genes in the wallaby genome. There was no indication of male-specific bands or male: female dosage difference to suggest genes that lie on the tammar wallaby sex chromosomes.

3.6.2 Attempts to clone marsupial *PRY*, *XKRY* and *BPY2*

Despite these negative Southern blot results, attempts were made to clone *PRY*, *XRY* and *BPY2* by homologous screening of a wallaby genomic DNA library. The probes used for Southern blot analysis of *PRY*, *XRY* and *BPY2* were used to screen a male tammar wallaby genomic DNA library. The libraries were screened and washed at the same low stringency conditions as the Southern blots. Screening at lower stringency would result in non-specific binding of the probe to the library membranes. No clear positives were detected, and no clones were isolated from the genomic library for any of the three genes.

A tammar wallaby pouch young cDNA library was also screened to try and identify any transcripts of the three genes expressed early in development. These homologous cDNAs could then be used as probes for the genomic library, which would be more successful than their human counterparts as they would be 100% homologous to the target sequence. No clones were isolated for these genes from the pouch young cDNA libraries. Since *PRY*, *XKRY* and *BPY2* all show testis specific expression in adult human, a tammar wallaby testis cDNA library was therefore screened under the low stringency conditions outlined above. However, no clones were isolated from this library either.

3.6.3 Summary

Southern blot analysis of *PRY*, *XKRY* and *BPY2* indicated that there was poor homology to these genes in the wallaby genome. Screening of a male genomic DNA library did not yield any clones.

The three human Y-borne genes *PRY*, *XKRY* and *BPY2* appeared to have no X-borne copies, so they fit into Lahn and Page's (1997) Class II genes. Therefore, all three genes are likely to have arisen recently on the human Y chromosome by (retro) transposition from an autosome. (Retro)transposed copies often change rapidly, and therefore appear to be different from their autosomal progenitors, as shown by the divergence of *DAZ* from *DAZL*. This occurs because once on the Y chromosome these genes are subject to a higher mutation rate (Section 1.4.5 and 1.4.6). Their recent origin would explain why *PRY*, *XKRY* and *BPY2* are specific to the human or primate Y, and why autosomal copies would be undetectable via Southern blot analysis and homologous screening of DNA libraries. Failure to isolate *PRY*, *XKRY* and *BPY2* in tammar wallaby is therefore likely to reflect the rapid divergence of Y-specific genes.

3.8 Discussion of human-wallaby comparisons

Out of the eight human Y-borne genes I attempted to isolate in tammar wallaby, only one (*TB4X/Y*) was successfully cloned and mapped. Pseudogenes were isolated for *EIF1AX/Y* and *RPS4X/Y*. No clones for *PCDHX/Y* were isolated, but a related protocadherin gene (*PCDH13*) was isolated and mapped, and some indication was obtained of a related gene on the X. No clones for *VCX/Y* were isolated, but an unrelated gene (*SMARCF1*) was isolated and mapped.

Southern blot analysis of tammar wallaby *TB4X/Y* revealed a few discrete bands. There was no indication of gene dosage or male specific bands, suggesting an autosomal location. Cloning and mapping of wallaby *TB4X/Y* demonstrated that it was located on chromosome 5p (Figure 3.23). *TB4X* and *TB4Y* are therefore recent additions to the eutherian sex chromosomes. This was unexpected since *TB4X* maps to Xq21.3 in human, among genes that are conserved on the marsupial X chromosome and therefore defined as a part of the human XCR. As a recent addition, *TB4X* disrupts this conserved material.

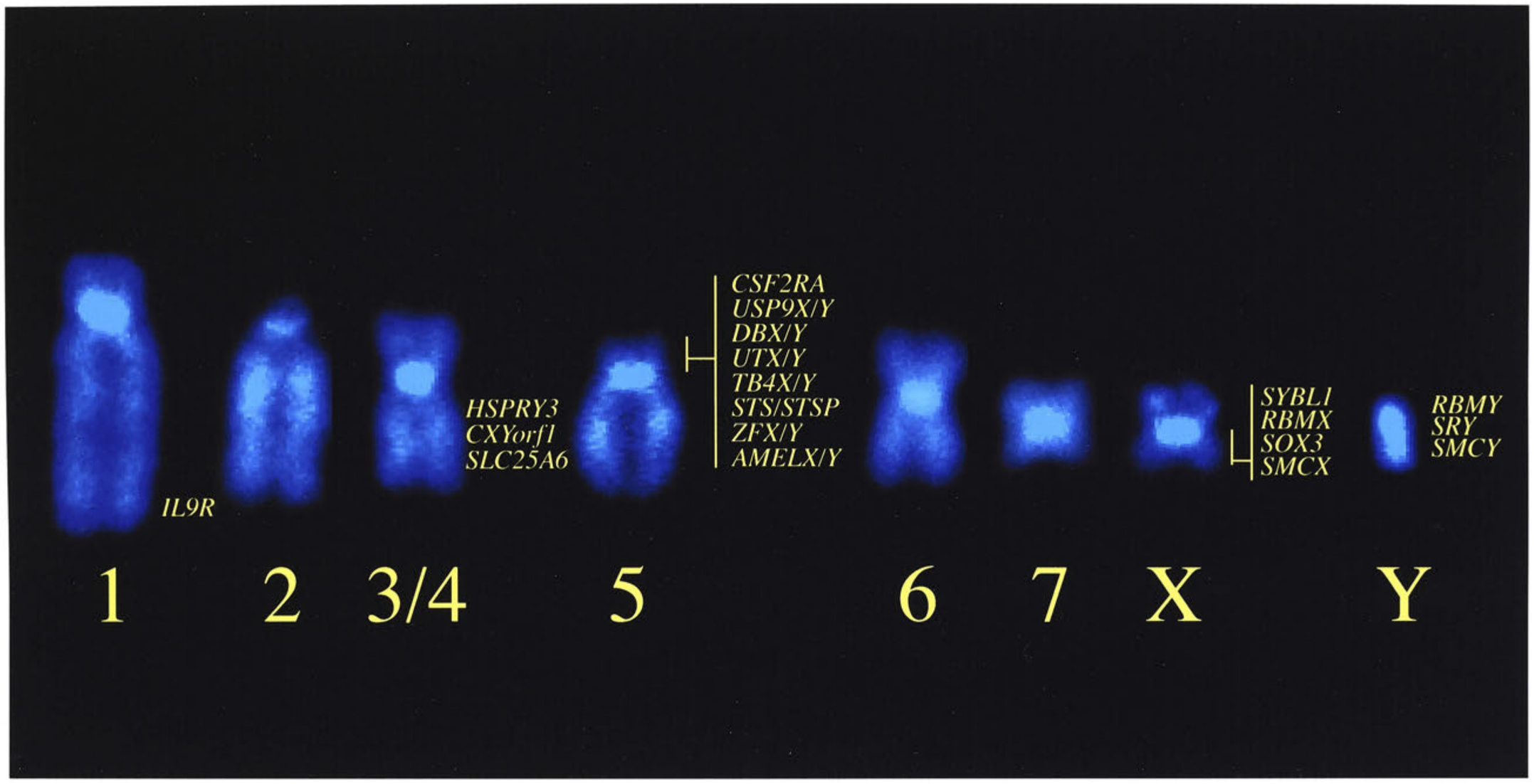


Figure 3.23: Locations of genes from the human Y chromosome in tammar wallaby. There are a total of 16 genes from the human Y that have been mapped in wallaby (two from PAR1, four from PAR2 and ten from the NRY). Chromosomes 3 and 4 are cytogenetically indistinguishable and therefore grouped together.

Southern blot analysis of tammar wallaby *RPS4X/Y* revealed that there were many copies of the gene in the wallaby genome. Without clones for the intron containing copies of *RPS4X/Y* in tammar wallaby, only tentative conclusions can be made that it has a copy on the X but not the Y, on the basis of previous Southern blot data. If this is so, *RPS4* was a part of the original proto-X and -Y, but was lost from the tammar wallaby Y. This is consistent with evidence that in other marsupial species, intron-containing copies of *RPS4X/Y* were located on both sex chromosomes (Jegalian and Page, 1998); it is therefore a conserved gene on the therian X and Y chromosomes. This suggested that *RPS4Y* must have a conserved selectable function on the human Y chromosome.

Fifteen genes shared by the human X and Y chromosomes have been mapped previously in tammar wallaby by myself and others in the Graves' lab (Figure 3.23). Two of these genes, *CSF2RA* and *SLC25A6* (previously known as *ANT3*), were located in PAR1 and four in PAR2. The remaining nine were located in the non-recombining region of the Y chromosome.

Ubiquitin-specific protease 9, Y chromosome (*USP9Y*), previously know as *DFFRY*, was shown to play a role in the AZFa phenotype (Sun *et al.*, 1999) and is therefore likely to be critical for spermatogenesis. It functions as an ubiquitin C-terminal hydrolase and is ubiquitously expressed, an unexpected pattern of spermatogenesis factors, which are usually gonad-specific. However, in mouse it showed testis specific expression (Brown *et al.*, 1998). *USP9Y* has an X-borne homologue in human that was mapped to Xp11.4 (Jones *et al.*, 1996), *USP9Y* was classified as a Class I gene (Lahn and Page, 1997).

DEAD/H box polypeptide, Y chromosome (*DBY*) has also been shown to play a critical role in the AZFa phenotype. *DBY* produces a long transcript that is ubiquitously expressed, as well as a shorter transcript that is expressed specifically in the testis (Foresta *et al.*, 2000). It has an X-borne homologue that was mapped to Xp.11.3. *DBY* was classified as a Class I gene (Lahn and Page, 1997).

USP9X/Y and *DBX/Y* tammar orthologues were cloned previously as part of work completed in fulfilment of an Honours degree and subsequently published (Waters *et al.*, 2001) (Appendix II). Both mapped to the short arm of chromosome 5 in tammar wallaby. Five other human Y-borne genes have also been mapped to tammar wallaby

chromosome 5p: the PAR1 gene *CSF2RA* and the Class I genes, from the differentiated region of the human Y chromosome, *ZFY*, *AMELY*, *UTY* and *STSP* (Waters *et al.*, 2001; Toder and Graves, 1998; Watson *et al.*, 1992b). *SLC25A6* was mapped to the long arm of chromosome 3/4 (Toder and Graves, 1998) (Figure 3.23).

The PAR2 genes *HSPRY3* and *CXYorf1* have been mapped to wallaby chromosome 3. *IL9R* was mapped to wallaby chromosome 1 and *SYBL1* was mapped to the wallaby X chromosome (Charchar et al, submitted) (Figure 3.23).

In addition to the 12 genes mentioned above, three genes (*SOX3/SRY*, *SMCX/Y* and *RBMX/Y*) have been mapped to the X and Y chromosomes of tammar wallaby (Foster *et al.*, 1992; Foster and Graves, 1994; Waters *et al.*, 2001; Delbridge *et al.*, 1999) (Figure 3.23). There has also been a report of a male specific *RPS4Y* band on an opossum Southern blot (Jegalian and Page, 1998), indicating that *RPS4X/Y* is located on the sex chromosomes of marsupials.

SMCY and *RPS4Y* both have X-borne homologues and were classified as Class I genes (Lahn and Page, 1997). *SRY* and *RBMY* were both initially classified as Y-specific Class II genes (Lahn and Page, 1997) but have both been found to have X-borne homologues. Thus *SRY* and *RBMY* are in fact Class I genes with testis specific expression and roles in sex-determination and spermatogenesis respectively.

Two genes, *ATRX* and *UBE1X*, absent from the human Y chromosome have also been mapped to tammar wallaby X and Y chromosomes. *ATRY* is located only on the marsupial Y chromosome, and has been proposed to represent the ancestral sex-determining gene (Pask et al., 2000). *Ube1y* is located on the mouse Y chromosome, but absent from the Y chromosomes of old world monkeys and hominids. *Ube1y* is also located in the X-Y pairing region of the platypus sex chromosomes.

Mapping genes conserved on the therian sex chromosomes in monotremes could test the hypothesis that the marsupial sex chromosomes represent the proto-sex chromosomes. A finding that some genes conserved on the therian sex chromosomes were located on the monotreme X, but that others were located on autosomes, could mean one of two things. First, the therian X and Y were equivalent to the original proto-sex chromosomes, and a region was lost from the monotreme X. Conversely, the monotreme sex chromosomes could have been equivalent to the original proto-sex chromosomes, and a region was added to the therian sex chromosome.

Chapter 4: Comparative Mapping of Human X- and Y-borne Genes in Platypus

The human Y chromosome consists of regions that are conserved with the marsupial Y chromosome, and regions that are autosomal in marsupials. Genes conserved on the therian sex chromosomes were mapped in a monotreme (platypus) to test that hypothesis that these genes were also present on the original proto-Y chromosome. If they were on the ancient mammalian Y chromosome 170MyBP they should be on the sex chromosomes in monotremes, whereas if they were acquired by the sex chromosomes of therian mammals they should be absent from the sex chromosomes of distantly related mammals.

The mammals most distantly related to therian mammals are the monotremes. The platypus, *Ornithorynchus anatinus*, was used for these studies because they diverged from Theria ~170MyBP, and therefore represent a mammalian outgroup to therians. To determine whether genes shared by the marsupial and eutherian Y chromosome were part of the proto-Y, or added to the therian Y, human X-Y shared genes were mapped in the platypus.

ATRX/Y, present on the human and marsupial X, and with a Y homologue in marsupials is therefore a gene conserved on the therian X(Y). *USP9X/Y*, which is X/Y shared in human and mouse, but autosomal in marsupials, and therefore recently added to the eutherian sex chromosomes. These two genes were mapped in platypus to further test the hypothesis that part of the human X and Y represent ancient sex chromosomes, and part a recent addition to the eutherian X and Y chromosomes.

4.1 Platypus *ATRX/Y*

ATRX (α-thalassemia and mental retardation associated with the X chromosome) is a member of a helicase superfamily subgroup that contains genes involved in DNA recombination, repair and regulation of transcription (Picketts *et al.*, 1998). It spans 280kb of genomic DNA with 36 exons and 35 introns. The ATRX mRNA is approximately 10.5kb and codes for a protein of 2492 amino acid (NCBI, www.ncbi.nih.gov/IEB/ Research/Acembly). In humans, mutations in *ATRX* cause X-linked α-thalassemia and mental retardation. XY individuals with mutations or deletions in *ATRX* also display varying degrees of sex reversal, implicating this gene in human testis development.

There is no evidence of a Y-linked homologue in human or mouse, but there is a homologue on the tammar wallaby Y chromosome (Pask *et al.*, 2000). The marsupial *ATRY* displays testis specific expression, whereas *ATRX* is ubiquitously expressed. Along with the observation that mutations in human *ATRX* cause sex-reversal, the gonad-specific expression of marsupial *ATRY* suggests it is involved in testis development. Since no male-specific *SRY* orthologue can be detected in platypus, it has been suggested to represent an ancestral testis-determining mechanism that predates *SRY* (Pask *et al.*, 2000). The rise of *SRY* as the primary mammalian male determining gene may have supplanted the role of *ATRY*, therefore permitting its demise in eutherian mammals.

4.1.1 Cloning and characterisation of platypus *ATRX/Y*

The platypus genomic DNA library was constructed by Dr. Margaret Delbridge and made available for use. An *ARTX* probe was amplified from genomic DNA of male *Monodelphis domestica* with the primers ATRX(f) and ATRX(r) (Appendix I). To amplify the expected product, the cycling parameters of a 2' denaturation at 94°C followed by 35 cycles of (94°C, 1' / 58°C, 1' / 72°C, 1') and then held at 4°C (Pask *et al.*, 2000). A product of approximately 400bp was generated that represented a region of exon 10, the largest of the *ATRX* exons.

The *M. domestica ATRX* probe was radioactively labelled with ^{32}P dCTP and hybridised to a platypus genomic DNA library under low stringency (55°C). The library membranes were washed at low stringency (55°C) and exposed to X-ray film overnight. Under these conditions five primary clones were isolated from the library. After a further three rounds of screening two final clones were obtained.

DNA from these final clones was amplified and digested with a variety of restriction endonucleases for 4 hours at 37°C. Upon digestion, the fragments generated were separated by electrophoresis on a 0.8% agarose gel. A Southern blot was performed on this gel, to which the radioactively labelled probe was hybridised under low stringency. The Southern blot was also washed at low stringency and exposed to X-ray film overnight (Figure 4.1). Good hybridisation was observed to the final clones *OanaATRX*-1 and *OanaATRX*-3. They shared both hybridising and non-hybridising bands. The restriction endonuclease *Eco*RI generated a hybridising band of about 1.8kb.

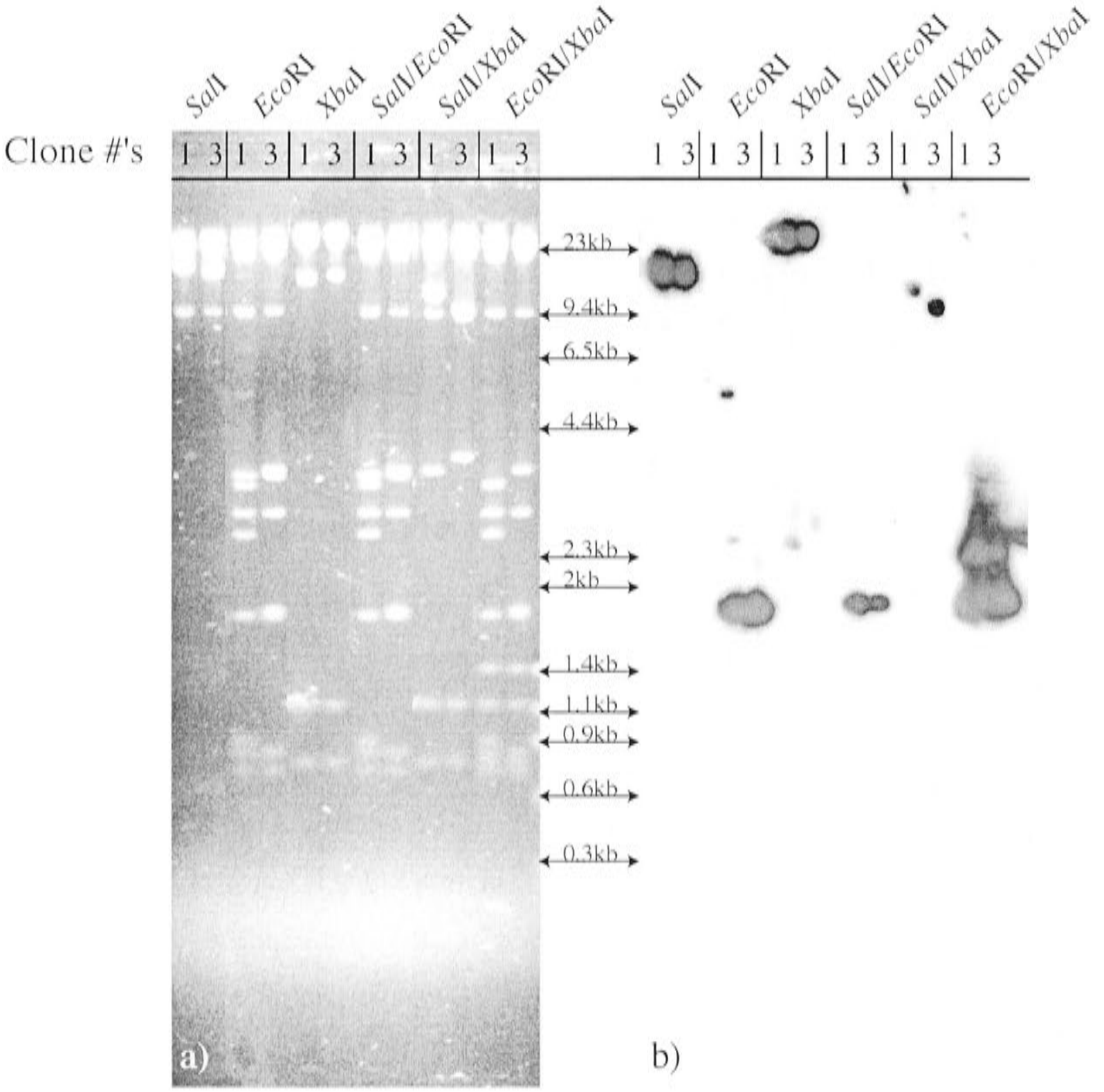


Figure4.1: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of platypus *ATRX* clones #1 and #3. b) Southern blot of restriction endonuclease gel with the *ATRX* probe hybridised. The autoradiograph was exposed overnight.

An attempt to subclone this hybridising fragment into pBluescriptSK$^+$ was unsuccessful. Shotgun-cloning was therefore attempted on *OanaATRX*-1 and *OanaATRX*-3. Sequencing of the relevant shotgun-clones confirmed that both *OanaATRX*-1 and *OanaATRX*-3 were platypus *ATRX* orthologues, representing exon 9, part of exon 10 and the intron between them.

Sequence from the two clones was identical and displayed 70-80% homology to human *ATRX* (Figure 4.2). Primers were designed to walk into plasmid clones. Sequence from the clones stretched from exon 9 through a small intron of approximately 280bp and into exon 10. An open reading frame was observed across the entire region sequenced in platypus (Figure 4.2). At the amino acid level 80-85% homology was observed (Figure 4.3).

4.1.2 FISH mapping of Platypus *ATRX*

The clones *OanaATRX*-1 and *OanaATRX*-3 were labelled with biotin in a nick translation reaction at 16°C for 4 hours. 200ng of the labelled probe was precipitated with 30µg of suppressor DNA, and a further 200ng was precipitated with 60µg of ssDNA. The biotin labelled probe was hybridised to platypus metaphase chromosome spreads prepared from fibroblast cells. The slides were pretreated with RNase and pepsin to remove excess RNA and protein. The hybridisation was conducted at 37°C for 48 hours. After hybridisation excess probe was washed off. Biotin labelled probe that had hybridised to the target sequence was detected with two antibody layers: a goat anti-biotin followed by a rabbit anti-goat+FITC conjugate.

Consistent signals were observed near the centromere on the long arm of chromosome 6. Platypus chromosome 6 contains the nuclear organising region (NOR), which was frequently differentially contracted causing the two homologues to look cytologically dissimilar in different spreads. In most metaphase spreads, signals were observed on only one member of the pair, leading to the initial suspicion that there were different alleles of *ATRX* on the two chromosome 6 homologues. This could indicate that there was a polymorphism of *ATRX*, perhaps equivalent to an *ATRX* and *ATRY*. However, in a few spreads three or four signals were observed so this was probably the result of low hybridisation efficiency (Figure 4.4). The difference in condensation of the NOR region could have caused differential hybridisation to the pericentric region.

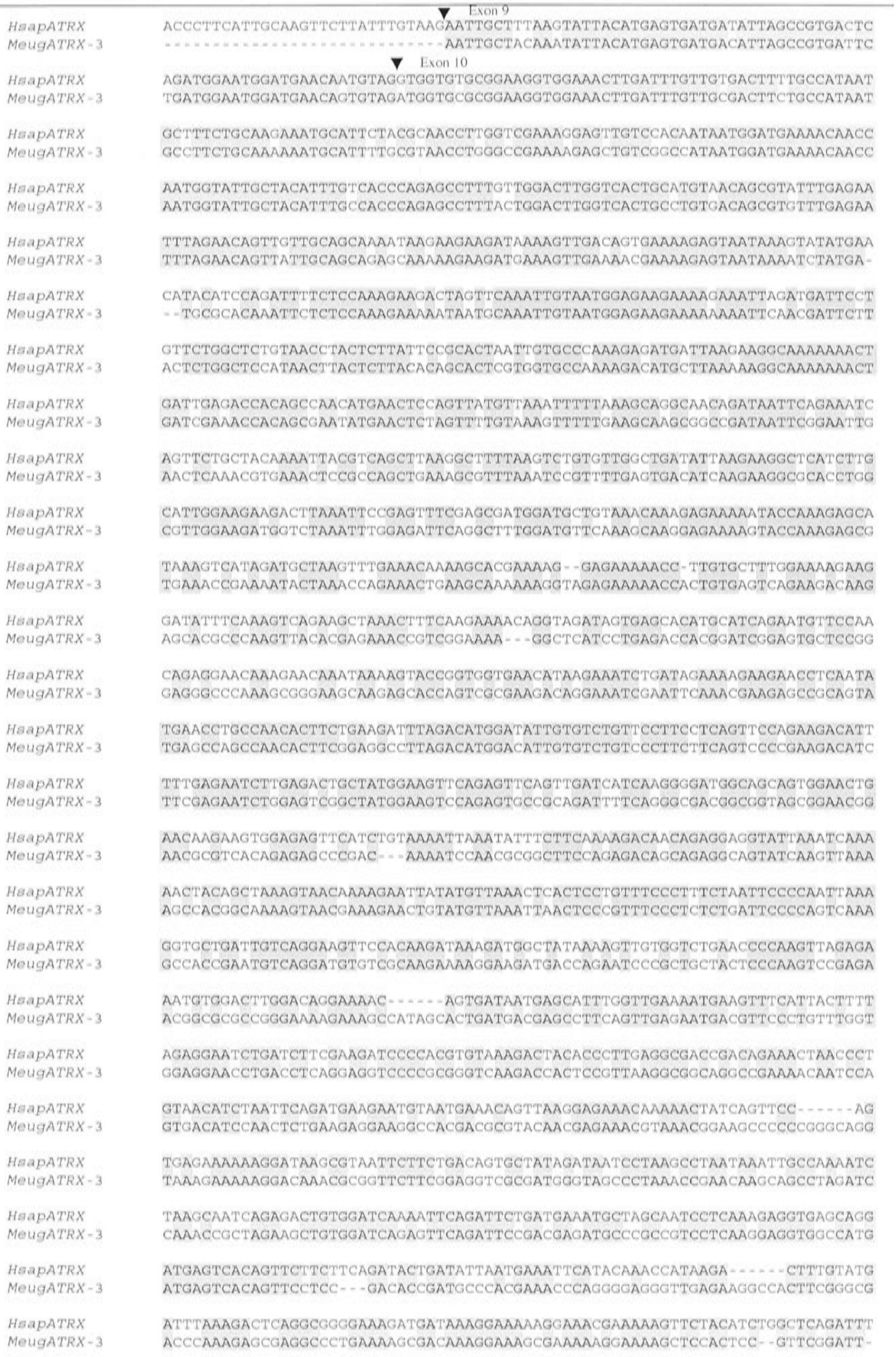
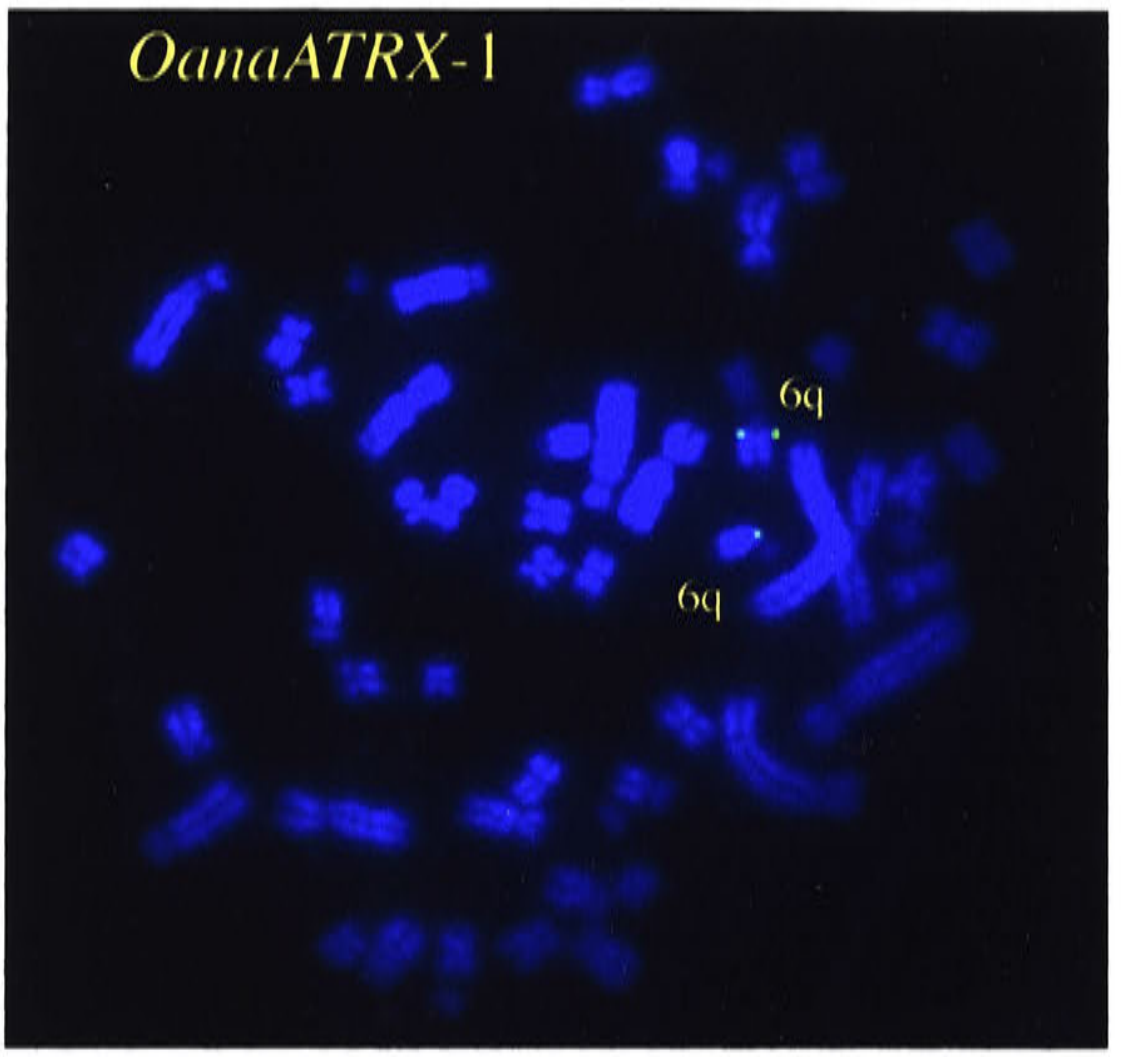


Figure4.2: Sequence alignment of human and tammar wallaby *ATRX* exon sequence. An open reading frame was observed across the region sequenced. The intron between exons 9 and 10 is 284bp. The position of this intron is conserved between species and

```
HsapATRX        CKNCFKYYMSDDISRDSDGMDEQCRWCAEGGNLICCDFCHNAFCKKCILRNLGRKELSTI
OanaATRX/Y      --NCYKYYMSDDISRDSDGMDEQCRWCAEGGNLICCDFCHNAFCKKCILRNLGRKELSAI

HsapATRX        MDENNQWYCYICHPEPLLDLVTACNSVFENLEQLLQQNKKKIKVDSEKSNKVYEHTSRFS
OanaATRX/Y      MDENNQWYCYICHPEPLLDLVTACDSVFENLEQLLQQSKKKMKVENEKSNKIYDAH-KFS

HsapATRX        PKKTSSNCNGEEKKLDDSCSGSVTYSYSALIVPKEMIKKAKKLIETTANMNSSYVKFLKQ
OanaATRX/Y      PKKNNANCNGEEKKFNDSYSGSITYSYTALVVPKDMLKKAKKLIETTANMNSSFVKFLKQ

HsapATRX        ATDNSEISSATKLRQLKAFKSVLADIKKAHLALEEDLNSEFRAMDAVNKEKNTKEHKVID
OanaATRX/Y      AADNSELNSNVKLRQLKAFKSVLSDIKKAHLALEDGLNLEIQALDVQSKEKSTKERETEN

HsapATRX        AKFETKARKGEK-PCALEKKDISKSEAKLSRKQVDSEHMHQNVPTEEQRTNKSTGGEHKK
OanaATRX/Y      TKPETEAKKVEKNHCESEDKSTPKLHEKPSEK-AHPETTDRSAPGGPKAGSKSTSREDRK

HsapATRX        SDRKEEPQYEPANTSEDLDMDIVSVPSSVPEDIFENLETAMEVQSSVDHQGDGSSGTEQE
OanaATRX/Y      SNSNEEPQYEPANTSEALDMDIVSVPSSVPEDIFENLESAMEVQSAADFQGDGGSGTERV

HsapATRX        VESSSVKLNISSKDNRGGIKSKTTAKVTKELYVKLTPVSLSNSPIKGADCQEVPQDKDGY
OanaATRX/Y      TESP-TKSNAASRDSRGSIKLKATAKVTKELYVKLTPVSLSDSPVKATECQDVSQEKEDD

HsapATRX        KSCGLNPKLEKCGLGQE--NSDNEHLVENEVSLLLEESDLRRSPRVKTTPLRRPTETNPV
OanaATRX/Y      QNPAATPKSENGAPGKESHSTDDEPSVENDVPCLVEEPDLRRSPRVKTTPLRRQAENNPV

HsapATRX        TSNSDEECNETVKEKQKLSVP--VRKKDKRNSSDSAIDNPKPNKLPKSKQSETVDQNSDS
OanaATRX/Y      TSNSEEEGHDAYNEKRKRKPPGQVKKKDKRGSSEVAMGSPKPNKQPRSKPLEAVDQSSDS

HsapATRX        DEMLAILKEVSRMSHSSSSDTDINEIH--TNHKTLYDLKTQAGKDDKGKRKRKSSTSGSD
OanaATRX/Y      DEMPAVLKEVAMMSHSSS-DTDAHETQGRVEKATSGDPKSEALKSDKGKRKRKSSTP---

HsapATRX        FDTKKGKSAKSSIISKKKRQTQSES
OanaATRX/Y      FG-----------------------
```

Figure 4.3: Amino acid alignment of platypus and human *ATRX*. Conservative changes are highlighted in yellow. Semi-conservative changes are highlighted in blue.



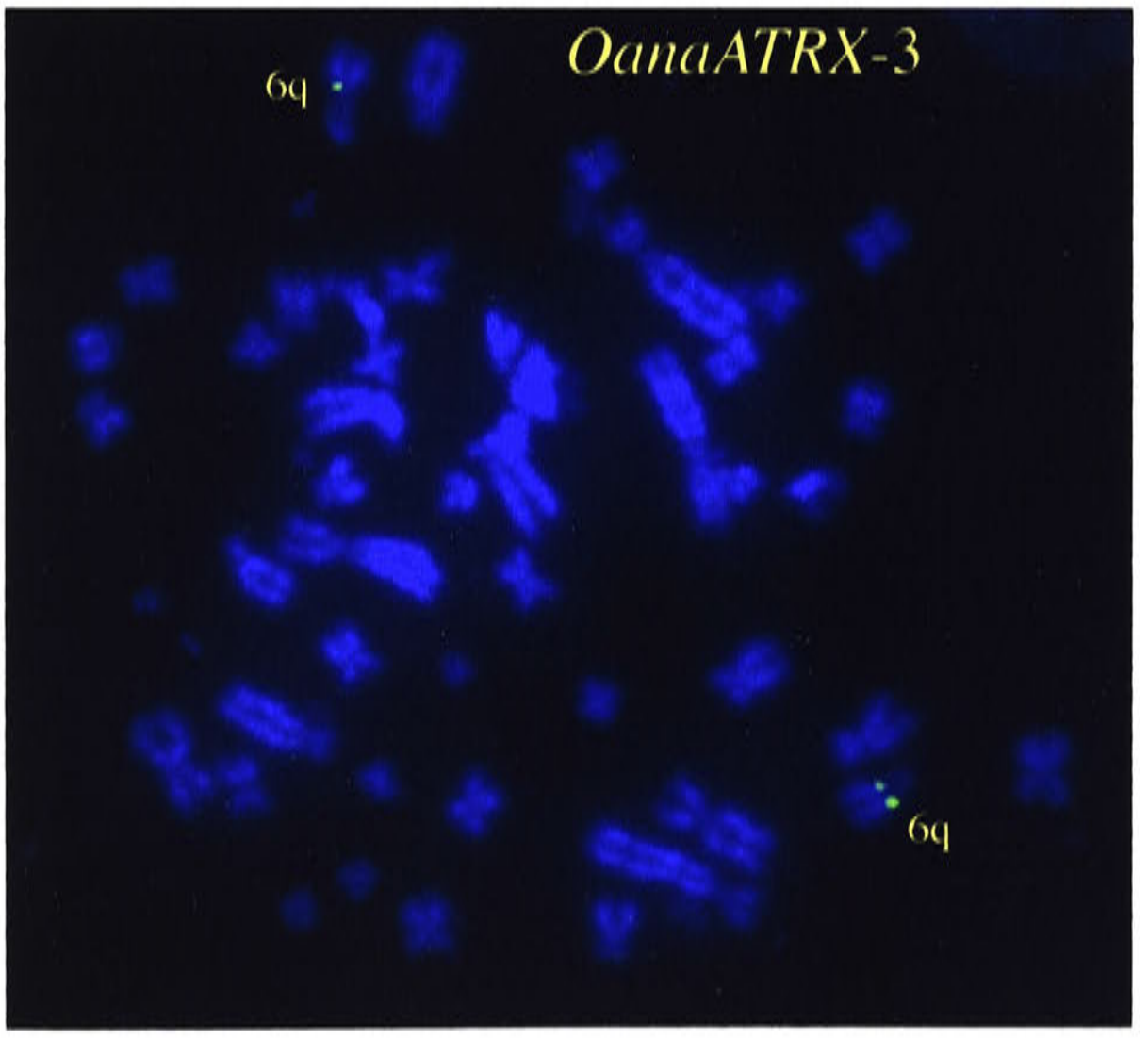


Figure 4.4: FISH mapping of platypus ATRX clones *OanaATRX*-1 and *OanaATRX*-3. Signals were observed on the long arm of chromosome six. The chromosomes were counterstained with DAPI.

There were no signals on the platypus X chromosome, or on platypus chromosome 1, the chromosome to which XAR genes map.

4.1.3 Summary

Screening of a male platypus genomic DNA library, with a *M. domestica ATRX* PCR product, yielded two clones that contained *ATRX/Y* orthologues. The two clones displayed 70-90% sequence homology to human *ATRX*. Both clones were mapped by FISH to the long arm of platypus chromosome 6. *ATRX/Y* was therefore not conserved on the sex chromosomes of all three mammalian groups, and constituted the first exception to the rule that genes conserved on the eutherian and marsupial sex chromosomes are also conserved on the monotreme sex chromosomes.

4.2 Platypus *USP9X/Y* and *DBX/Y*

Two Class I genes shared by the human X and Y chromosomes are *USP9X/Y* and *DBX/Y*. Their position on the human X (Xp11.4 and Xp11.3 respectively) and their location on tammar wallaby chromosome 5p leads to the expectation that they should be located on an autosome in platypus.

USP9Y has been shown to play a role in the AZFa phenotype (Sun *et al.*, 1999). *USP9X* spans 342kb of genomic DNA with 50 exons and 49 introns. Its transcript is 7.2kb and codes for 1367 amino acid protein (NCBI, www.ncbi.nih.gov/IEB/Research /Acembly).

DBY has also been shown to play a critical role in the AZFa phenotype. *DBY* consists of 17 exon and 16 introns. It produces a long transcript (4.4kb) that is ubiquitously expressed, as well as a shorter transcript (2.3kb) that is expressed specifically in the testis (Foresta *et al.*, 2000).

USP9Y and *DBY*, along with *UTY*, form a syntenic homology region shared between the human, cat and mouse Y chromosome (Murphy *et al.*, 1999; Mazeyrat *et al.*, 1998), all of which fall within both the human AZFa and the mouse ΔSxr^b intervals. The tammar wallaby orthologues of these three genes also cluster together on chromosome 5p (Waters *et al.*, 2001).

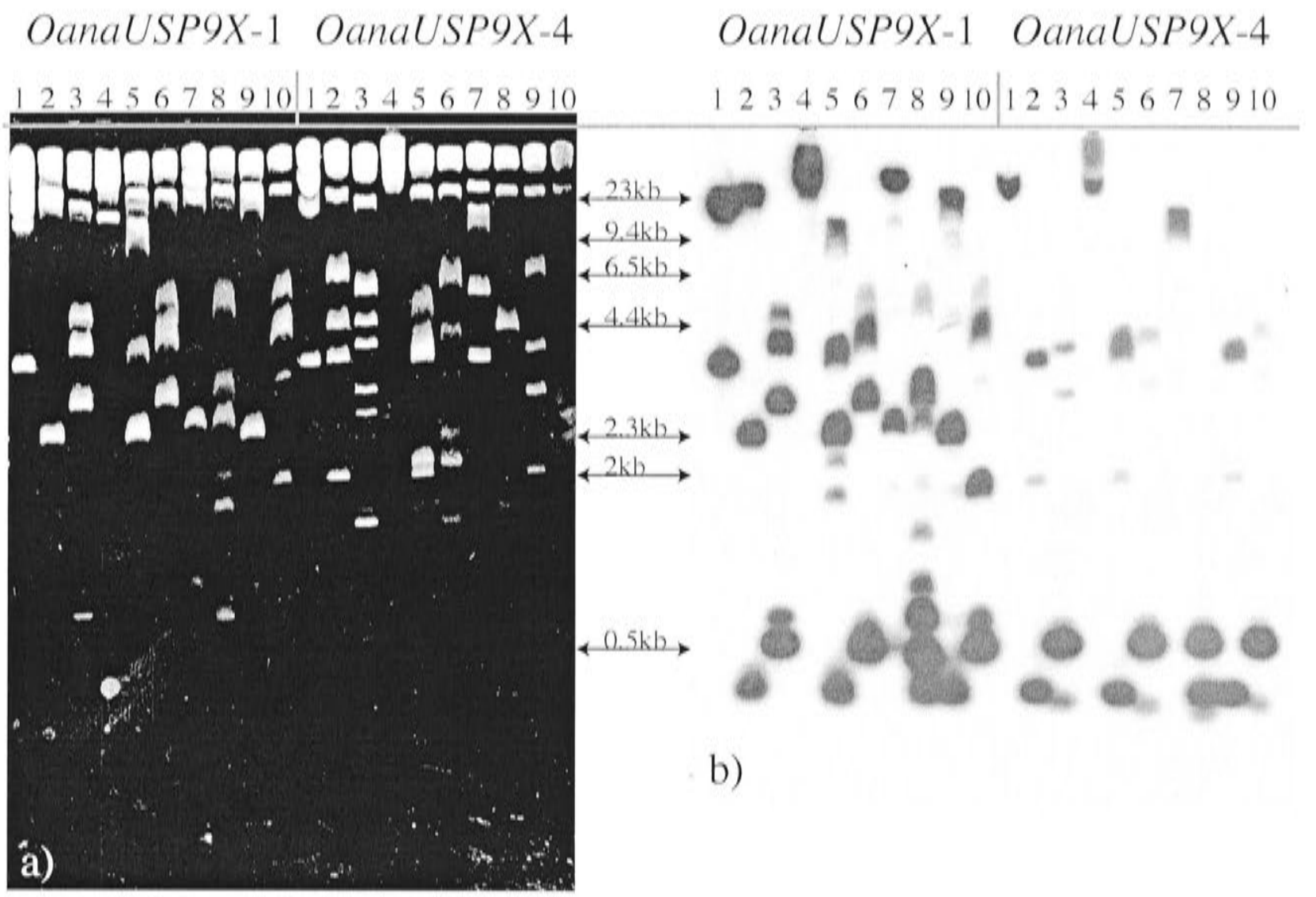
4.2.1 Attempts to clone platypus *DBX/Y*

A full length mouse *Dbx* cDNA, kindly supplied by Dr. M. Mitchell (INSERM Unite 406, Marseille, France; GenBank accession number NM010028), was used as a probe in an attempt to isolate the platypus orthologue from a male genomic DNA library. The mouse cDNA was radioactively labelled with ^{32}P dCTP and hybridised to a platypus genomic DNA library at 55°C. The membranes were washed at 55°C and exposed to X-ray film overnight. Under these conditions six primary clones were isolated from the library; however, after a further three rounds of screening no final clones were obtained.

4.2.2 Cloning and characterisation platypus *USP9X/Y*

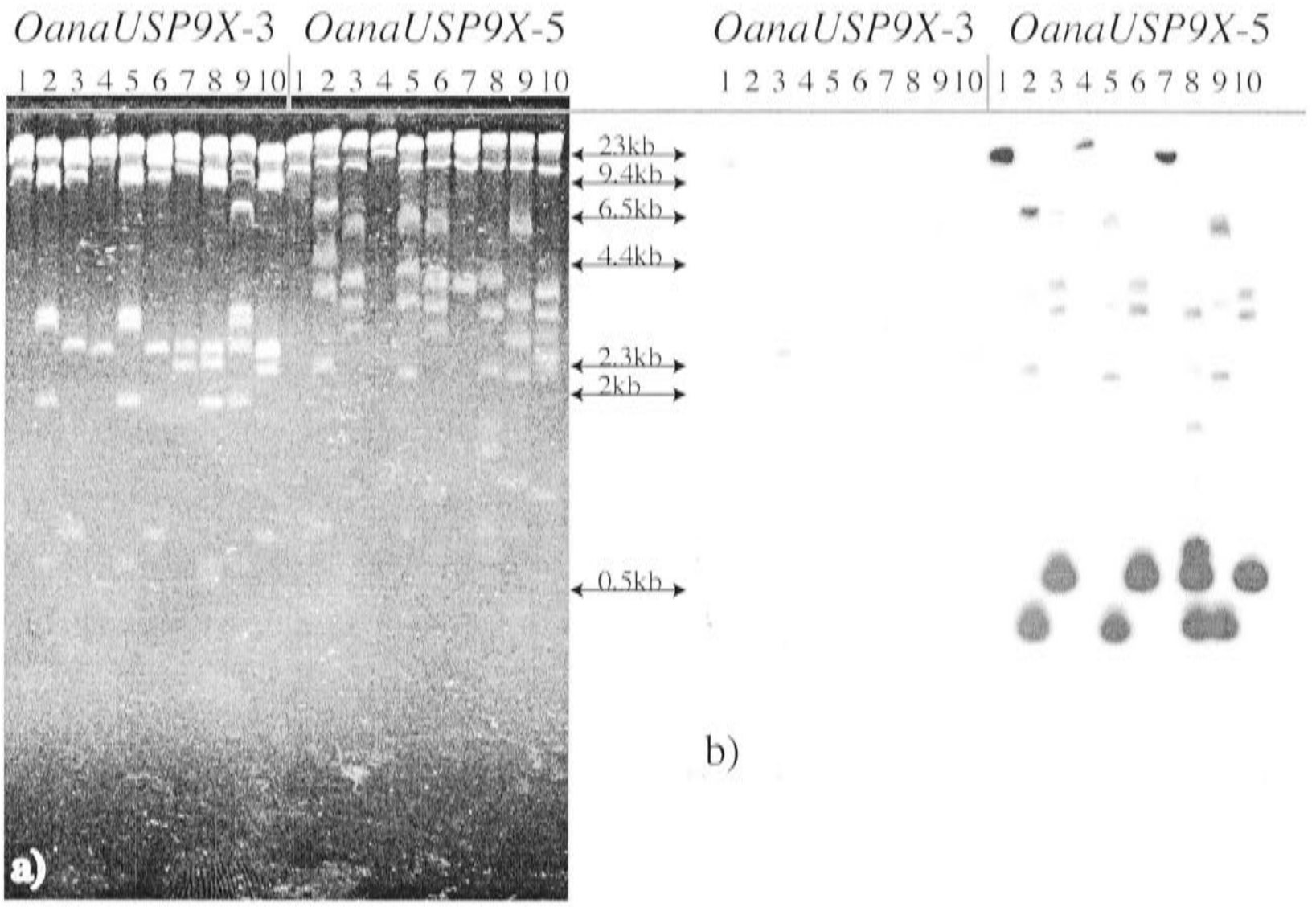
A full-length mouse *Usp9x* cDNA kindly supplied by Dr. M. Mitchell (INSERM, Marseille, France; GenBank accession number NM009481) was used as a probe in an attempt to isolate the platypus orthologue from a male genomic DNA library. The mouse *Usp9x* cDNA was radioactively labelled with ^{32}P dCTP and hybridised to a platypus genomic DNA library under low stringency (55°C). The membranes were washed at 55°C and exposed to X-ray film overnight. Under these conditions six primary clones were picked from the library. After a further three rounds of screening four final clones were obtained.

DNA from these final clones was amplified and digested with a variety of restriction endonucleases. The digestions were conducted at 37°C for 4 hours. The fragments generated were separated by electrophoresis on a 0.8% agarose gel, upon which a Southern blot was performed. The *Usp9x* cDNA probe was hybridised to this Southern blot under the same conditions used to isolate the clones from the library. The Southern blot was exposed to X-ray film for 2 hours and 30 minutes. The restriction analysis revealed that *OanaUSP9X*-1, *OanaUSP9X*-4 and *OanaUSP9X*-5 shared hybridising and non-hybridising bands, indicating that they were three different but overlapping clones (Figure 4.5). The remaining clone showed poor hybridisation of the cDNA probe (Figure 4.6).



Lane	Restriction enzymes		Lane	Restriction enzymes	
1	*Sal*I		6	*Sal*I	*Xba*I
2	*Eco*RI		7	*Sal*I	*Xho*I
3	*Xba*I		8	*Eco*RI	*Xba*I
4	*Xho*I		9	*Eco*RI	*Xho*I
5	*Sal*I	*Eco*RI	10	*Xba*I	*Xho*I

Figure4.5: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of platypus *USP9X* clones #1 and #4. b) Southern blot of restriction endonuclease gel with the *USP9X* probe hybridised. The autoradiograph was exposed for 2hrs and 30mins.



Lane	Restriction enzymes		Lane	Restriction enzymes	
1	*Sal*I		6	*Sal*I	*Xba*I
2	*Eco*RI		7	*Sal*I	*Xho*I
3	*Xba*I		8	*Eco*RI	*Xba*I
4	*Xho*I		9	*Eco*RI	*Xho*I
5	*Sal*I	*Eco*RI	10	*Xba*I	*Xho*I

Figure4.6: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of platypus *USP9X* clones #3 and #5. b) Southern blot of restriction endonuclease gel with the *USP9X* probe hybridised. The autoradiograph was exposed overnight.

```
HsapUSP9X       TGGAAGAGAGCAAAGAGCCAGATGACCAAGATGCTCCAGATGAACATGAGTCGCCTCCAC
OanaUsp9x       --------------GAACCSGATGACCAAGAGGCCCCAGATGAACACGAATCCTCTCCAC

HsapUSP9X       CTGAAGATGCCCCATTGTACCCCCATTCACCTGGATCTCAGTA-TCAACAGAATAACCAT
OanaUsp9x       CAGAAGATGCGCCAATTATATCCCATTC-CCTGGATCCCAGTAATCAGCAG---------
```

Figure4.7: Sequence alignment of human and platypus *USP9X*, exon 46 . This region is translated in human and an open reading frame is present in platypus.

A ~400bp hybridising fragment, generated by *Eco*RI from *OanaUSP9X*-4 was subcloned into pBluescriptSK$^+$ and sequenced using the plasmid primers. Sequencing revealed approximately 90% homology to exon 46 of human *USP9X*, which was approximately 100bp. *OanaUSP9X*-4 showed homology to the entire exon with homology falling away in the introns on either side. Exon 46 is part of the 3' untranslated region (Figure 4.7). *OanaUSP9X*-4 was therefore confirmed as containing platypus *USP9X/Y*. The overlapping clones *OanaUSP9X*-1 and *OanaUSP9X*-5 shared the hybridising fragment from *OanaUSP9X*-4 that was sequenced to confirm its identity. Therefore, all three clones contained exon 47. In addition, *OanaUSP9X*-1 contained larger hybridising fragments that were not subcloned. These fragments may have contained sequence from a different exon of *USP9X*. However, attempts to subclone and sequence these fragments were unsuccessful.

Attempts to map platypus *USP9X/Y* by FISH were unsuccessful. No signals were consistently observed on any platypus chromosome.

4.2.3 Summary

Attempts to clone *DBX/Y* were unsuccessful, but attempts to clone platypus *USP9X/Y* yielded three overlapping clones that all contained *USP9X/Y*. Fish mapping was unsuccessful and no conclusion could be made about its evolution in light of its human, tammar wallaby and platypus locations.

4.3 Conclusions

Cloning genes in platypus has traditionally proven to be difficult. The monotreme genome contains many repeats of MON-1, a monotreme specific Core-SINE element. Previous attempts to clones platypus genes by myself and others have resulted in isolation of these elements (P. Kirby, N. El-Mogharbel, personal communication).

ATRX/Y lies on the sex chromosomes of eutherians and marsupials. It has a copy that lies on the long arm of the human X, among other genes that are shared by the X in all mammals. *ATRX/Y* was therefore expected to be located on the monotreme X. It has a copy on the X as well as the Y in marsupials. The gonad-specific expression of

marsupial *ARTY*, as well as the sex-reversed phenotype of humans with mutant *ATRX*, suggested that it might have once served as the mammalian testis-determining gene, before the differentiation of *SRY*. Since platypus lacks a male-specific *SRY*, it was suggested that *ATRY* serves a testis determining function in platypus.

It was therefore surprising to find that platypus *ATRX* was located, not on the sex chromosomes, but rather on chromosome 6. This immediately contradicted the hypothesis that *ATRX* determines sex in monotremes. It also implied that part of the X conserved region is not conserved on the monotreme X.

This finding was supported by the recent mapping of *RBMX* and *CDX4* (conserved on the sex chromosomes of therian mammals) to platypus chromosome 6 (M. Delbridge; J. Deakin, personal communications). In addition, the *SOX3* gene, which was the ancestor of the sex-determining *SRY*, was also determined to be autosomal in platypus (Pask, 1999; Western, 1999), although there are now doubts about the correct identification of this member of the large SOX gene family (Kirby, 2002).

These result implied that the *ATRX-RBMX-CDX* region was either not part of the ancestral mammalian sex chromosomes, or that it was originally part of the ancestral mammalian sex chromosomes, but was subsequently moved to an autosome in the monotreme lineage. The question of where this cluster of genes was located in the mammalian ancestor was particularly interesting because two of the genes in the region have important male-specific functions in human. *RBMX* has a Y-linked homologue in eutherians and marsupials that is crucial to male spermatogenesis. *ATRX* has a Y-linked homologue in marsupials that may play a role in the sex-determining pathway. Did these genes arrive more recently on the sex chromosomes of the therian ancestor and acquire their roles in sex and reproduction, perhaps replacing more ancestral mechanisms, after the divergence of monotremes 170MyBP? If so, this would subdivide Stratum 1 (Lahn and Page, 1999) into two evolutionarily distinct regions. Alternatively, could a region of the X, carrying these genes, have been moved to an autosome in monotremes? A more likely scenario is that the region was moved to an autosome at a time when the proto-X and -Y were still homologous over this region, so that loss from the monotreme X represented a PAR to autosome translocation.

Gene mapping in a vertebrate outgroup was necessary to answer this question. Chicken represents the reptile-bird lineage from which the mammalian ancestor

diverged about 310MyBP. Thus attempts were made to clone and map the chicken orthologues of *ATRX*, *SOX3* and *RBMX*, as well as *G6PD*, a gene conserved on the X chromosome of all mammals.

Chapter 5: Comparative Mapping of Human X- and Y-linked Genes in Chicken

Twelve genes from the human X chromosome have been located on the X in marsupials and monotremes (Wilcox *et al.*, 1996; Wrigley and Graves, 1988b; Watson *et al.*, 1990). They defined a conserved region encompassing the entire marsupial X, and the long arm and proximal short arm of the human X. This region was hypothesized to represent the original autosomes from which mammal sex chromosomes evolved. However, there are a number of genes on human Xp that are autosomal in marsupials and monotremes. This region was therefore added to the eutherian sex chromosomes after the divergence of marsupials 130MyBP. It was therefore concluded that the human X chromosome consisted of a conserved region (XCR) shared by the X in all mammals, and recently added region (XAR). However, it now appears that at last a part of the region shared by human Xq and the marsupial X is not on the X in monotremes.

The finding that *ATRX*, *RBMX* and *CDX4* (all conserved between the human and marsupial sex chromosomes) resided on chromosome 6 in platypus, and that *FMR1*, *HPRT* and *PGK* were also autosomal platypus, suggested that this region may represent an addition to the X chromosome in the therian lineage after the divergence of monotremes. Or, it may represent a region lost from the X in the monotreme lineage, and retained by the X in the therian linage. Thus it was important to distinguish between these two hypotheses by mapping some or all of these genes in an outgroup.

The division of the human X chromosome into two regions (XCR and XAR) with an independent evolutionary origin has been confirmed by reference to an outgroup. Chicken was chosen as an outgroup because birds/reptiles are the closest vertebrate relatives of the mammals. The bird/reptile lineage diverged from mammals approximately 310MyBP, whereas mammals diverged from amphibians and fish 360MyBP and 400MyBP respectively. Chicken was the obvious choice as a model bird because they have the best studied genome of all birds and reptiles.

Three genes from the human XCR have been mapped in chicken and all map to chromosome 4p. Genes from the XAR are located on chicken chromosome 1 (Figure 5.1). This confirmed that XCR and XAR were both present intact, but separate, in a common ancestor of mammals and reptiles/birds 310 MyBP.

However, these results do answer the question of whether the XCR is composed of more than one evolutionary unit, because there are no human XCR genes (on the X in all mammals) that have been mapped in both platypus and chicken. None of the genes

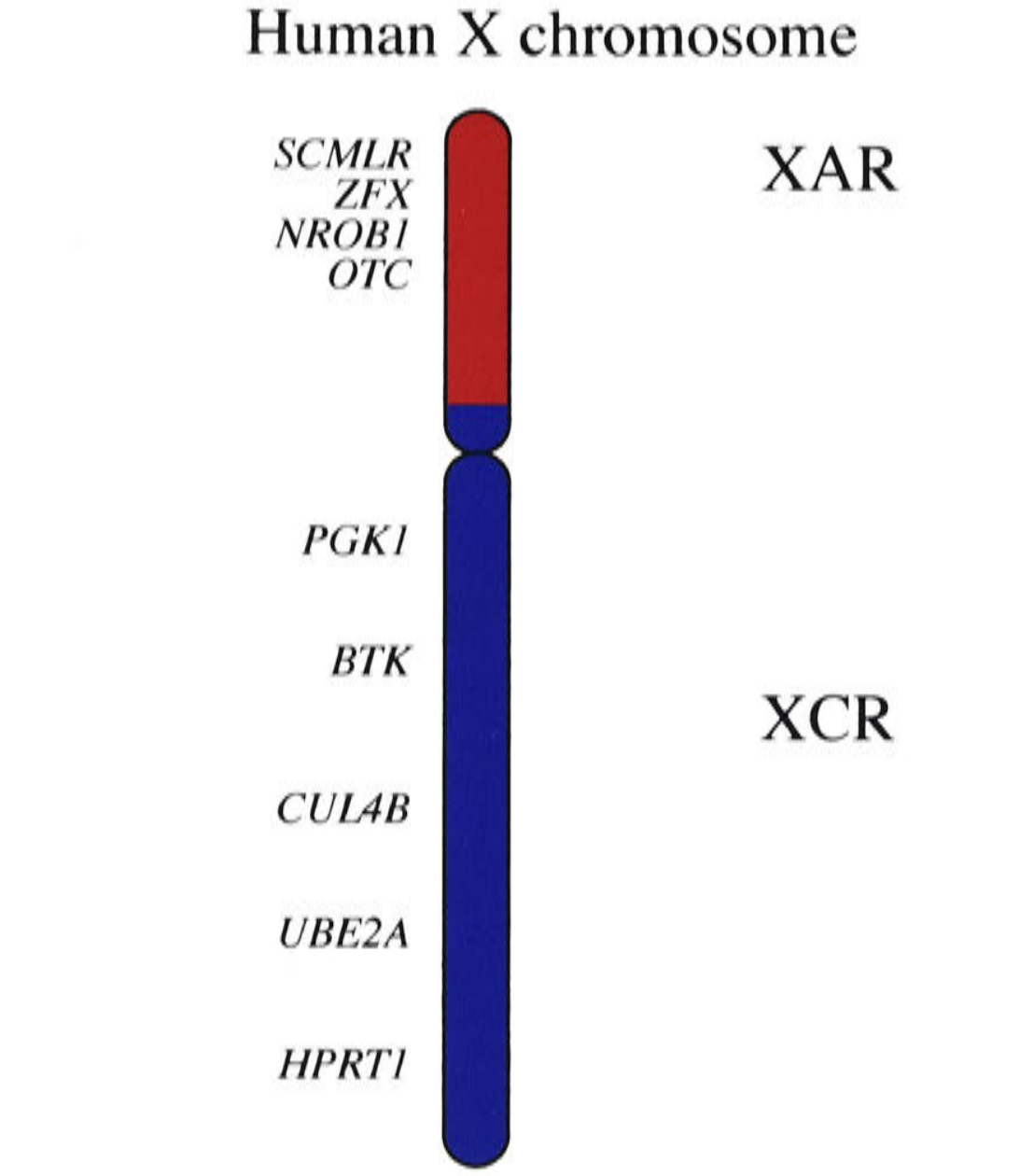


Figure 5.1: Location of human X gene in chicken. Human XCR genes (red) are located on chicken chromosome 1. XCR genes (blue) are located on chicken chromosome 4.

that have been mapped to the platypus X chromosome (*G6PD*, *RCP*, *AR*, *GDX*, *UBE1X*, *PLP*, *F9*, *F8*, *GATA1*, *ALAS2*, *P3* and *GLA*) have been mapped in chicken, and none of the XCR genes mapped in chicken (*BTK*, *CUL4B* and *UBE2A*) have been mapped in platypus. Since no genes have been mapped in both monotremes and birds, the ancestral origin of the genes missing from the monotreme X is unknown. A finding that the cluster of genes on platypus chromosome 6 lays on chicken 4p would favour the loss of this region from an original mammal XCR. A finding that these genes lay in a separate cluster in the chicken would imply that the region was independently added to the sex chromosomes later in the therian lineage.

To answer this question, attempts were made to clone and map orthologues of three human X-Y shared genes (*ATRX/Y*, *SOX3/SRY* and *RBMX/Y*) in chicken. *SOX3* was chosen because of its status as a gene conserved on the therian X and the importance of its Y-linked homologue (*SRY*) in the evolution of the therian Y chromosome. *RBMX* was chosen also because of its status as a conserved gene on the therian X and the importance of its Y-linked homologue (*RBMY*) in the evolution of the therian Y chromosome. In addition, *G6PD* was chosen for mapping in chicken because of its location on the platypus X, as well as marsupial and eutherian X (Watson *et al.*, 1990).

5.1 Chicken *G6PD*

Glucose-6-phosphate-dehydrogenase (*G6PD*) is an X-specific gene that maps to the long arm of the human X and lies on the X chromosomes of all eutherians, marsupials and monotremes (Watson *et al.*, 1990; Dawson and Graves, 1984). It spans 16kb of genomic DNA with 13 exons and 12 introns. The *G6PD* mRNA is 2631bp and codes for a 545 amino acid protein (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly).

5.1.1 Cloning and characterisation of chicken *G6PD*

A partial human *G6PD* cDNA (Persico *et al.*, 1986) was used as a probe to screen the chicken female genomic DNA library. The *G6PD* insert was liberated from its plasmid vector by digestion with the restriction endonuclease *Eco*RI for 1 hour at 37°C. The insert was separated from the vector by electrophoresis on a 1% agarose gel, and the band was excised and purified. 100ng of the *G6PD* probe was radioactively labelled with ^{32}P dCTP and hybridised to a female genomic DNA chicken library under low

stringency (55°C). The library membranes were washed at 55°C and exposed to X-ray film overnight.

Under these conditions five primary clones were isolated from the library. After a further 3 rounds of screening all five were purified to final clones. Three final clones were isolated from primary clone #1 and two final clones were isolated from primary clone #2. DNA was amplified from all of the final clones and digested with a variety of restriction endonucleases for 4 hours at 37°C. The fragments generated by the digests were separated on a 0.8% agarose gel and a Southern blot was performed. The radioactively labelled *G6PD* probe was hybridised to this Southern blot overnight at 55°C and then washed at 55°C. The Southern blot membrane was exposed to X-ray film overnight (Figure 5.2). All three final clones from primary clone #1 had identical restriction patterns, whereas the two final clones from primary clone #2 were different from each other and from the final clones isolated from primary clone #1. Good hybridisation to the G6PD probe was observed to all of the final clones that were isolated.

GgalG6PD-1.1 was selected for shotgun-cloning to confirm its identity. Sequencing of the relevant recombinant shotgun-clones and a nucleotide blast search revealed 90% sequence homology to part of the 159bp exon 6 of human *G6PD* (Figure 5.3). However, this exon in chicken was tandemly repeated at least 11 times, and showed good homology only at the end of the exon. The good hybridisation of the *G6PD* probe to *GgalG6PD*-1.1 was probably due to the tandem repetition of exon 6. However, the possibility that a complete, translatable copy of exon 6 was present in the clone could not be excluded. Unsuccessful attempts were therefore made to sequence this clone further to find a complete exon 6. Attempts to map this clone to chicken metaphase spreads were made, but these were unsuccessful. Time constraints restricted the characterisation of other chicken *G6PD* clones.

5.2 Chicken *SOX3/SRY*

SOX3 and *SRY* are located on the human X and Y chromosome respectively. Both are conserved on the sex chromosomes of eutherians and marsupials, demonstrating that they have existed on the mammalian sex chromosomes at least since the divergence of these two lineages 130MyBP.

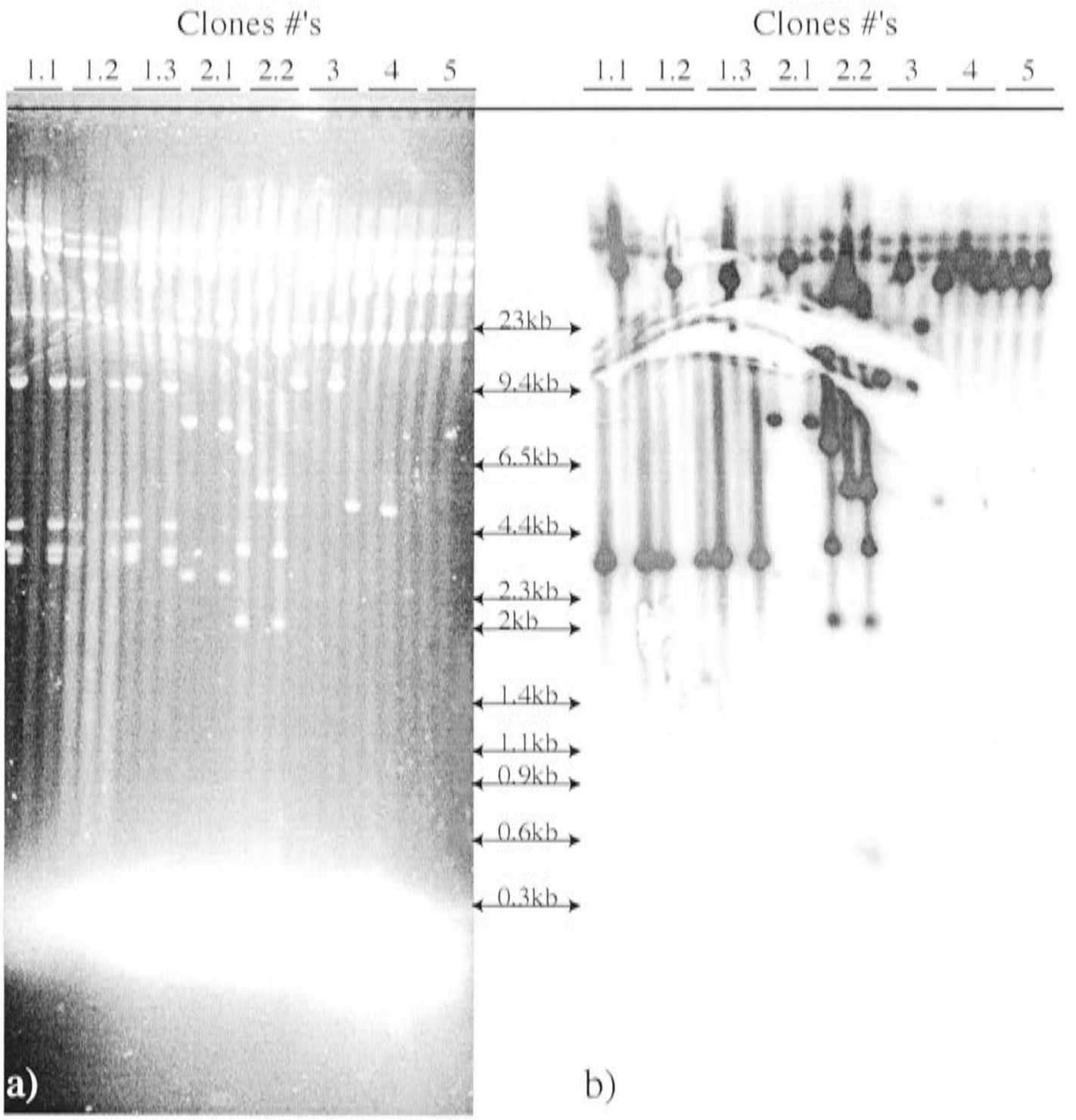


Figure5.2: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of chicken *G6PD* clones. Each clone was digested with *Eco*RI, *Sal*I and *Eco*RI/*Sal*I respectively. b) Southern blot of restriction endonuclease gel with the *G6PD* probe hybridised. The autoradiograph was exposed overnight.

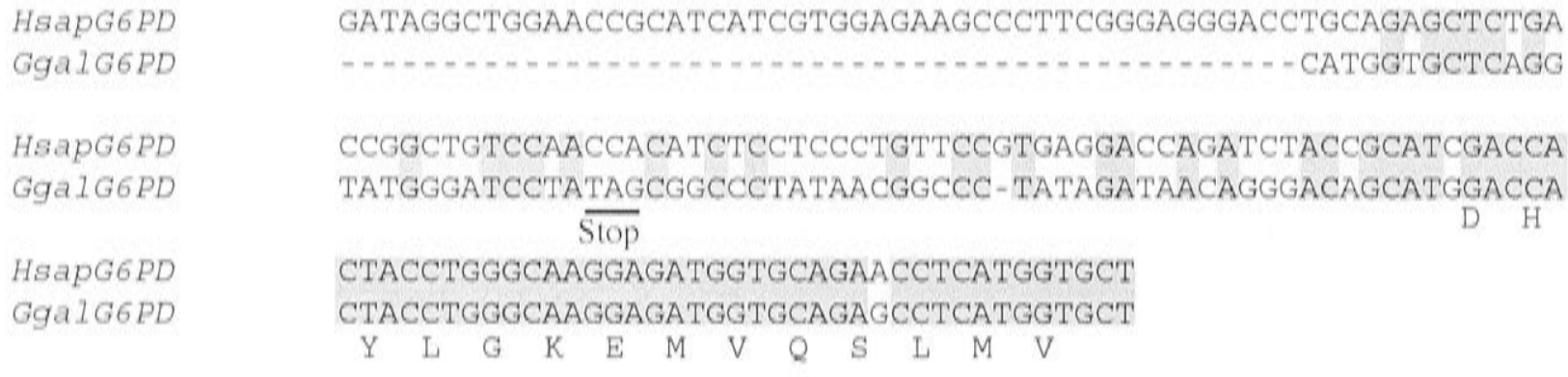


Figure 5.3: Exon 6 (159bp) of human *G6PD* aligned to the 112bp repeated unit of in *GgalG6PD*-1.1. This repeat ocurrs at least 11 times in *GgalG6PD*-1.1. Amino acid sequence is shown where there is homology. No amino acid homology was observed outside this region. An in-frame stop codon was observed and is indicated.

SOX3 is a member of the SOX group B family and in the mouse embryo it is expressed in the central nervous system, as well as briefly in the genital ridge (Collignon *et al.*, 1996). It is intronless, and its1.39kb transcript codes for a 446 amino acid protein. *SOX3* orthologues were cloned from platypus and mapped to a small autosome that is a member of the group of cytogenetically indistinguishable chromosomes 7-11 (Pask, 1999); however, the identity of this SOXB gene in now not certain (Kirby, 2002).

SRY, the dominant testis-determining factor (TDF) in eutherians, is thought to have evolved from *SOX3* by truncation outside the HMG box. As for *SOX3*, *SRY* is intronless, and its 897bp transcript codes for a 218 amino acid protein (NCBI, www.ncbi.nih.gov/IEB/Research/Acembly). Although it was originally classified as a Class II gene by Lahn and Page (1997), the discovery of an X-borne homologue (Foster and Graves, 1994) showed that it was, after all, a Class I gene. Despite many attempts, no male-specific *SRY* has been identified in platypus. Because the progenitor of *SRY* (*SOX3*) has not been identified and mapped in platypus, the origin of *SRY* is unclear.

There is no sex-specific *SOX/SRY* in chickens. *SOX3* has been cloned and sequenced (Kamachi *et al.*, 1998) but not mapped. To determine the location of the ancestral *SRY* in chicken, *SOX3* (rather than the very variable *SRY*) was cloned and mapped in chicken.

5.2.1 Cloning and characterisation of chicken *SOX3*

A *SOX3* probe was amplified from tammar wallaby male genomic DNA using the degenerate primers SOX3f and SOX3r. To amplify the expected product of 900bp, the cycling parameters used were 30 cycles of (94°C, 30'' / 51°C, 30'' / 72°C, 30'') following a 2 minute denaturation at 94°C. The product was subcloned into pGEM T Easy so it could be easily amplified and used as a probe to isolate *SOX3* from a female chicken genomic DNA library. A library was constructed using DNA from female chicken tissue (section 2.3.1). Because females are the heterogametic sex in birds a male chicken genomic DNA library would lack representation of the W chromosome.

The *SOX3* probe was radioactively labelled with ^{32}P dCTP and hybridised to a female chicken genomic DNA library overnight under low stringency (55°C). The library membranes were washed at 55°C and exposed to X-ray film overnight. Under

these conditions six primary clones were picked from the library. After a further three rounds of screening, at the same stringency, four final clones were isolated.

DNA from the final clones was amplified and PCR with two different primer sets was performed. The first primer set, SOX3-HMG(f) and SOX3-HMG(r) (Pask, 2000), amplified the HMG box from group B SOX genes. The second primer set, SOX3(f) and SOX3(r), amplified *SOX3* specifically. To amplify the expected products of 300bp (HMG box) and 900bp (*SOX3*) the cycling parameters of a 2' denaturation at 94°C followed by 30 cycles of (94°C, 30'' / 51°C, 30'' / 72°C, 30'') followed by 72°C for 10' were used. The reactions were subjected to electrophoresis on a 1% agarose gel. Under these conditions products of the expected size were amplified for *GgalSOX3*-4 and *GgalSOX3*-5. No products were amplified for *GgalSOX3*-1 and *GgalSOX3*-2 (Figure 5.4).

The products were subcloned into pGEM T Easy and the inserts sequenced using the vector primers. A nucleotide blast search revealed that *GgalSOX3*-4 and *GgalSOX3*-5 were both chicken *SOX3*. The 900bp PCR products displayed 99% homology to the chicken sequence in the NCBI database; the few differences are located in the primer-binding region at the degenerate nucleotide positions (Figure 5.5). An open reading frame across the entire sequence was observed and there was 100% homology at the amino acid level (Figure 5.6).

5.2.2 FISH mapping of chicken *SOX3*

GgalSOX3-5 was labelled for 4 hours and 30 minutes in a nick translation reaction with biotin. 200ng of the labelled probe was precipitated with 10µg of suppressor DNA and a further 200ng of the labelled probe was precipitated with 60µg of ssDNA. The two precipitation reactions containing the biotin labelled probe were hybridised to female chicken metaphase chromosomes prepared from fibroblast cells. The slides were pretreated with RNase and pepsin to remove excess RNA and protein. The hybridisation was conducted at 37°C for 48 hours. After hybridisation excess probe was washed off at 37°C. After washing, biotin labelled probe that hybridised to the target sequences was detected and visualised as described in section 2.5.4.

Under these conditions consistent signals were observed on the short arm of chromosome 4 (Figure 5.7), near to the location of other genes from the conserved

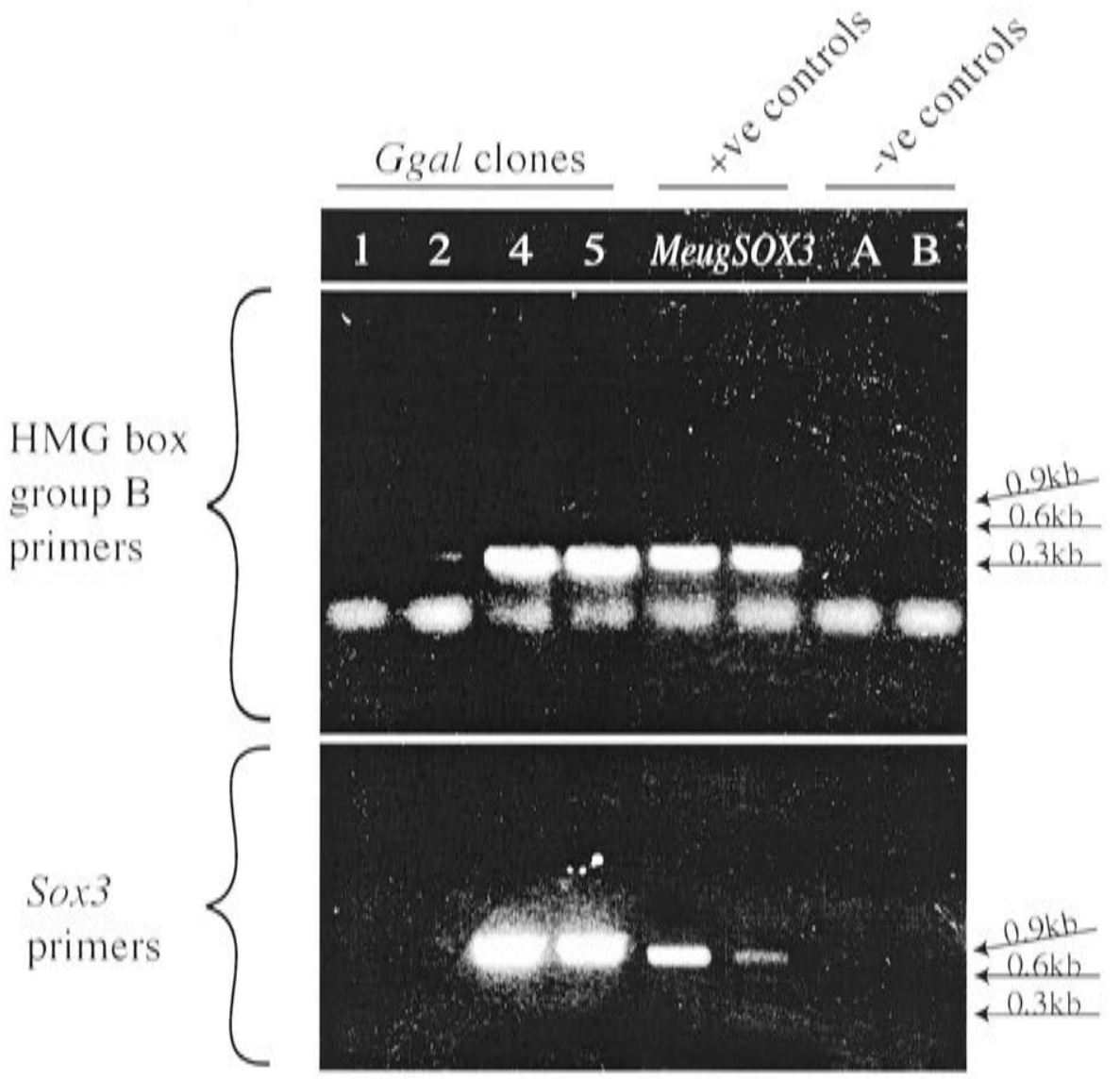


Figure5.4: PCR of *SOX3* chicken clones #1,2,4 and 5. Tammar wallaby *SOX3* genomic clones were used as positive controls. Negative control A was a *Monodelphis domestica RBMX* genomic clone and negative control B contained no DNA.

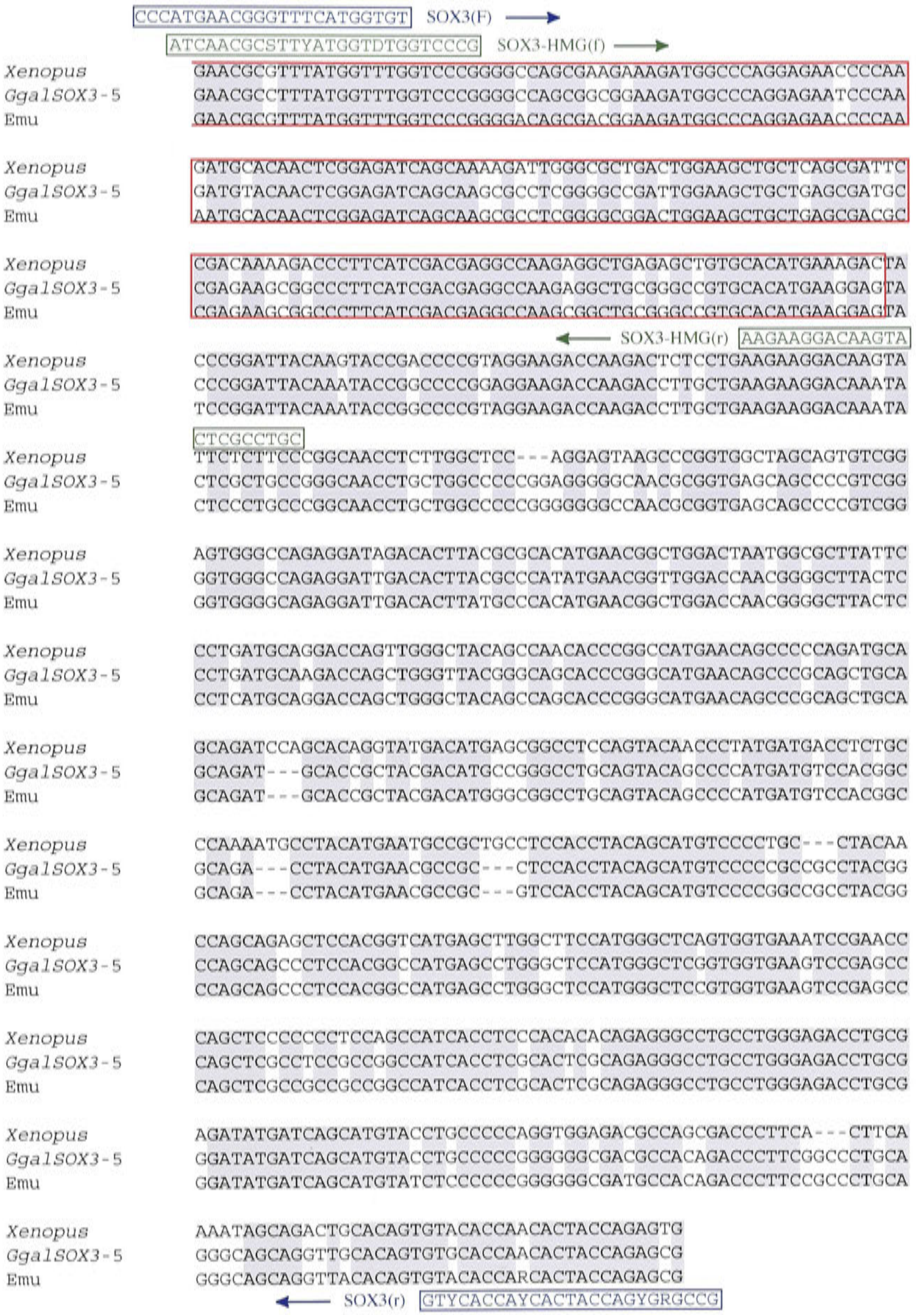


Figure 5.5: Alignment of chicken, emu (Shetty, 2001) and *Xenopus* (GenBank accession number BM262916) *SOX3*. The HMG Box is outlined in red, 15bp of the 5' region of the HMG box is missing. HMG box primers (green) and cross-species *SOX3* primers (blue) are indicated. An open reading frame was observed across the entire region sequenced.

```
EmuSOX3          NAFMVWSRGQRRKMAQENPKMHNSEISKRLGADWKLLSDAEKRPFIDEAKRLRAVHMKEY
GgalSOX3         --------GQRRKMAQENPKMYNSEISKRLGADWKLLSDAEKRPFIDEAKRLRAVHMKEY
XenopusSOX3      NAFMVWSRGQRRKMAQENPKMHNSEISKRLGADWKLLSDSDKRPFIDEAKRLRAVHMKDY

EmuSOX3          PDYKYRPRRKTKTLLKKDKYSLPGNLLAPGGANAVSSPVGVGQRIDTYAHMNGWTNGAYS
GgalSOX3         PDYKYRPRRKTKTLLKKDKYSLPGNLLAPGGGNAVSSPVGVGQRIDTYAHMNGWTNGAYS
XenopusSOX3      PDYKYRPRRKTKTLLKKDKYSLPGNLLAPG-VSPVASSVGVGQRIDTYAHMNGWTNGAYS

EmuSOX3          LMQDQLGYSQHPGMNSPQLQQM-HRYDMGGLQYSPMMSTAQ-TYMN-AASTYSMSPAAYG
GgalSOX3         LMQDQLGYGQHPGMNSPQLQQM-HRYDMPGLQYSPMMSTAQ-TYMN-AASTYSMSPAAYG
XenopusSOX3      LMQDQLGYSQHPAMNSPQMQQIQHRYDMSGLQYNPMMTSAQNAYMNAAASTYSMSP-AYN

EmuSOX3          QQPSTAMSLGSMGSVVKSEPSSPPPAITSHSQRACLGDLRDMISMYLPPGGDATDPSALQ
GgalSOX3         QQPSTAMSLGSMGSVVKSEPSSPPPAITSHSQRACLGDLRDMISMYLPPGGDATDPSALQ
XenopusSOX3      QQSSTVMSLASMGSVVKSEPSSPPPAITSHTQRACLGDLRDMISMYLPPGGDASDPS-LQ

EmuSOX3          GSRLHSVHQHYQS
GgalSOX3         GSRLHS-------
XenopusSOX3      NSRLHSVHQHYQS
```

Figure 5.6: Amino acid alignment of chicken, emu and *Xenopus SOX3*. Conservative changes are highlighted in yellow. Semi-conservative changes are highlighted in blue.

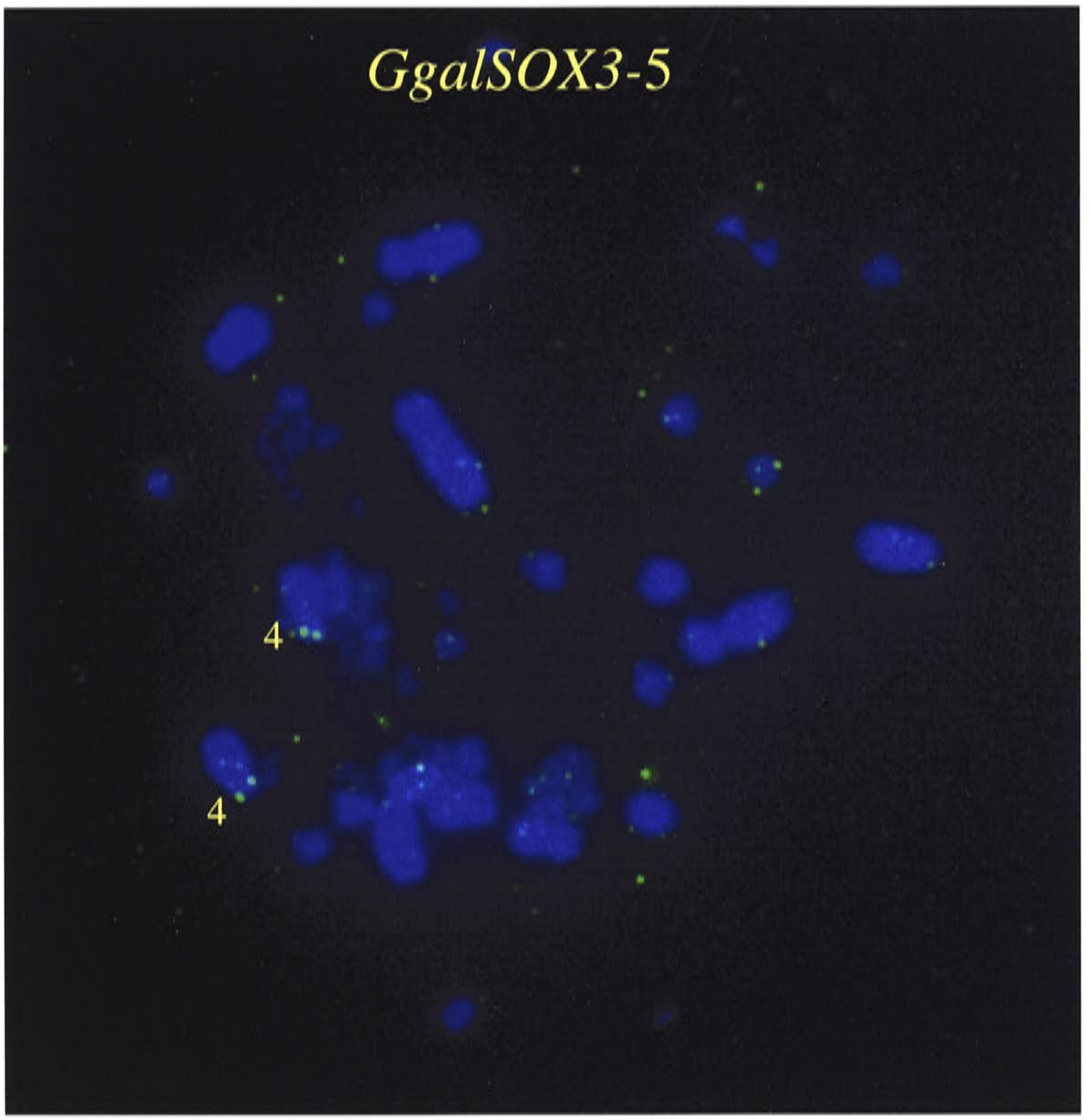


Figure 5.7: FISH mapping of the chicken *SOX3* clone *GgalSOX3*-5. Signals were observed on the short arm of chromosome four. The chromosomes were counterstained with DAPI.

region of the human X chromosome. This was the first localisation of a human X/Y shared gene, conserved on the therian X, in chicken. This directly confirmed that the origin of the conserved part of the Y chromosome was represented by chicken chromosome 4p.

5.2.3 Summary

A female chicken genomic DNA library was screened with a tammar wallaby *SOX3* probe. Two clones were isolated that contained chicken *SOX3*. One of these clones was mapped to the short arm of chromosome 4. *SOX3* therefore maps with other human XCR genes in chicken.

5.3 Chicken *RBMX/Y*

RBMY is present in multiple copies on the human Y chromosome. There are copies located in all of the AZF regions defined on the human Y. However, only the copies located in the AZFb region are expressed, and *RBMY* is the primary candidate for azoospermia resulting from deletion of this region (Elliott *et al.*, 1997). *RBMY* is conserved on the marsupial Y chromosome, indicating that it was part of the ancient Y, and has been maintained for at least 130 million years. This is consistent with a conserved role for *RBMY* in spermatogenesis (Delbridge *et al.*, 1997).

It was originally thought that *RBMY* arose on the human Y chromosome via translocation from an autosome, or retrotransposition from an unprocessed transcript in the mammalian ancestor (Vogt *et al.*, 1997). The *HNRPG* gene located on human chromosome 6 was found to share about 60% sequence homology with *RBMY* (Soulard *et al.*, 1993). However, it was shown that *RBMY* has a ubiquitously expressed X-linked homologue (*RBMX*) located on the long arm of the human X chromosome in a region (Xq26) responsible for several mental retardation syndromes (Delbridge *et al.*, 1999). The chromosome 6 copy is one of several pseudogenes, some of which are expressed (Lingenfelter *et al.*, 2001). Thus *RBMX* and *RBMY* are a conserved X/Y pair that have resided on the mammalian X and Y chromosomes for at least 130 million years. The

platypus *RBMX/Y* was mapped to the long arm of chromosome 6. Therefore, *RBMX/Y*, like *SOX3*, is not conserved on the sex chromosomes of all mammals.

RBMX/Y was mapped in chicken to determine whether it clusters with other XCR genes, including *SOX3/SRY*.

5.3.1 Cloning and characterisation of chicken *RBMX/Y*

A partial tammar wallaby *RBMX* cDNA (Delbridge *et al.*, 1998) of approximately 1kb was used as a probe to isolate the chicken orthologue. The insert was liberated from its plasmid vector by digestion with the restriction endonuclease *Eco*RI for 1 hour at 37°C. The insert was separated from the vector by electrophoresis on a 1% agarose gel, which was excised and purified. 100ng of the *RBMX* probe was radioactively labelled with ^{32}P dCTP and hybridised to a female chicken genomic DNA library under low stringency (55°C). The library membranes were washed at 55°C and exposed to X-ray film overnight.

Under these conditions seven primary clones were isolated from the library. After a further three rounds of screening all were purified through to final clones. DNA was amplified for all of the final clones and digested with a variety of restriction endonucleases at 37°C for 4 hours. The fragments generated were separated on a 0.8% agarose gel, on which a Southern blot was performed. Hybridisation of radioactively labelled *RBMX* probe and washing of the Southern blot was performed under the same conditions used to screen the library. The Southern blot membrane was exposed to X-ray film overnight, after which the film was developed. Good hybridisation was observed to two of the seven final clones, *GgalRBMX*-3 and *GgalRBMX*-5 (Figure 5.8).

GgalRBMX-5 was chosen for shotgun-cloning in order to confirm its identity as chicken *RBMX*. Sequencing of the relevant shotgun-clones revealed that *GgalRBMX*-5 showed 85-90% homology to human *RBMX* across exons 4 through 9. Introns were observed between all exons, confirming its identity as the chicken *RBMX* orthologue (Figure 5.9). An open reading frame was observed in the chicken sequence, displaying 90-95% amino acid homology to human *RBMX* (Figure 5.10).

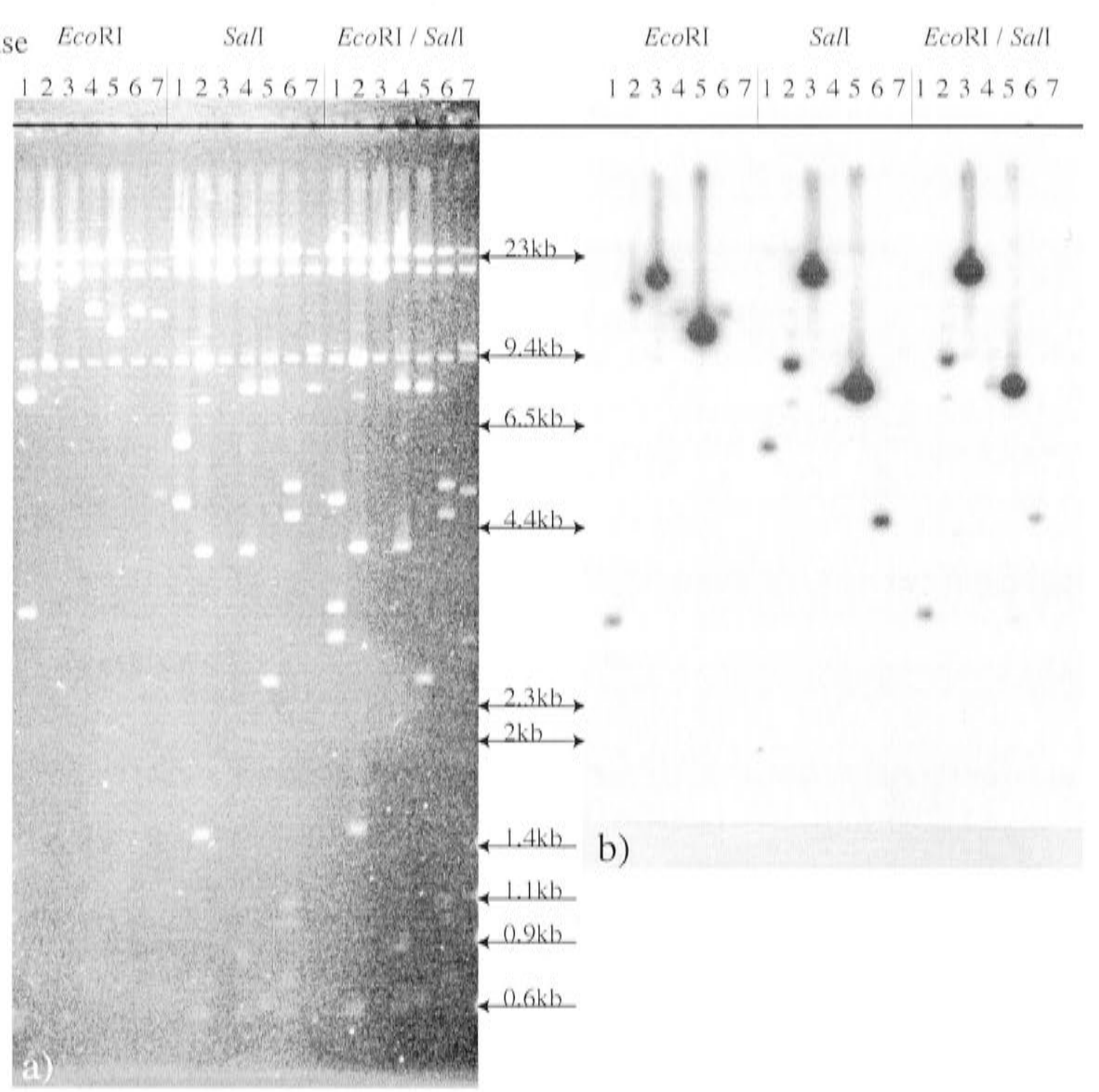


Figure5.8: a) Ethidium bromide stained agarose gel of a restriction endonuclease analysis of chicken *RBMX* clones #1 to #7. b) Southern blot of restriction endonuclease gel with the *RBMX* probe hybridised. The autoradiograph was exposed overnight.

Figure 5.9: Sequence alignment of human and chicken *RBMX*. Start and stop codons are indicated. The chicken *RMBX* displayed an open reading frame from the beginning of exon 4 until the stop codon in exon 9. Intron positions are conserved between species, their sizes are indicated for human and, where known, chicken.

```
HsapRBMX      DAKDAARDMNGKSLDGKAIKVEQATKPSFESG-RRGPPPPPRSRGPPRGLRGGRGGSGGT
GgalRBMX/Y    ------------SLDGKAIKVEQATKPSFESGGRRGPPPPPRSRGPPRGLRGGRGGSG-A

HsapRBMX      RGPPSRGGHMDDGGYSMNFNMSSSRGPLPVKRGPPPRSGGPPPKRSAPSGPVRSSSGMGG
GgalRBMX/Y    RGPPSRGSHLVTSWY-MAYLVGSSRGPLPMKRGPPPRSGGPPPKRSAPSGPVRSSS-MGG

HsapRBMX      RAPVSRGRDSYGGPPRREPLPSRRDVYLSPRDDGYSTKDSYSSRDYPSSRDTRDYAPPPR
GgalRBMX/Y    RAPVSRGRDSYGGPPRREPMPSRRDVYMSPRDDGYSTKDGYSSRDYPSSRDTRDYAPPPR

HsapRBMX      DYTYRDYGHSSSRDDYPSRGYSDRDGYGRDRDYSDHPSGGSYRDSYESYGNSRSAPPTRG
GgalRBMX/Y    DYAYRDYGHSSSRDEYPSRGYSDGYGGGRDRDYSDHPSGGSYRDSYESYGNSRSAPPARG

HsapRBMX      PPPSYGGSSRYDDYSSSRDGYGGSRDSYSSSRSDLYSSGRDRVGRQERGLPPSMERGYPP
GgalRBMX/Y    PPPSYGGSSRYDDYGSTRDGYG-SRESYSSSRSDVYSSGRDRVGRQDRGLPPSMERGYPP

HsapRBMX      PRDSYSSSSRGAPRGGGRGGSRSDRGGGRSRY
GgalRBMX/Y    PRDSYSSSSRGAPRGGGRGGSRSDRGGGRSRY
```

Figure 5.10: Amino acid alignment of human and chicken *RBMX*. Conservative changes are highlighted in yellow. Semi-conservative changes are highlighted in blue.

5.3.2 FISH mapping of chicken *RBMX*

GgalRBMX-5 was labelled for 3 hours in a nick translation reaction with biotin. 200ng of the labelled probe was precipitated with 10μg of suppressor DNA and a further 200ng of the labelled probe was precipitated with 60μg of ssDNA. The two precipitation reactions containing the biotin labelled probe were hybridised to female chicken metaphase chromosomes prepared from fibroblast cells. The slides were pretreated with RNase and pepsin to remove excess RNA and protein. The hybridisation was conducted at 37°C for 48 hours. After hybridisation excess probe was washed off at 37°C. The hybridised biotin labelled probe was detected and visualised as described in section 2.5.4.

Under these conditions, consistent signals were observed on the short arm of chicken chromosome 4 (Figure 5.11), near to the location of the chicken orthologues of other genes from the human XCR. Thus the two human X/Y shared genes *SOX3/SRY* and *RBMX/Y* map together to this location in chicken.

5.3.3 Summary

Screening a female chicken genomic DNA library yielded one clone that was confirmed as containing *RBMX/Y*. This *RBMX/Y* orthologue was mapped to the short arm of chicken chromosome 4. *RBMX/Y* therefore clusters with other XCR genes in chicken, including *SOX3*.

5.4 Attempts to clone chicken *ATRX/Y*

A *Monodelphis domestica ATRX* PCR product was used as a probe in an attempt to isolate the chicken orthologue from a female genomic DNA library. The *ATRX* probe was radioactively labelled with ^{32}P dCTP and hybridised to the genomic DNA library at 55°C. The membranes were washed at 55°C and exposed to X-ray film overnight. Under these conditions seven primary clones were isolated from the library; however, after a further three rounds of screening no final clones were obtained.

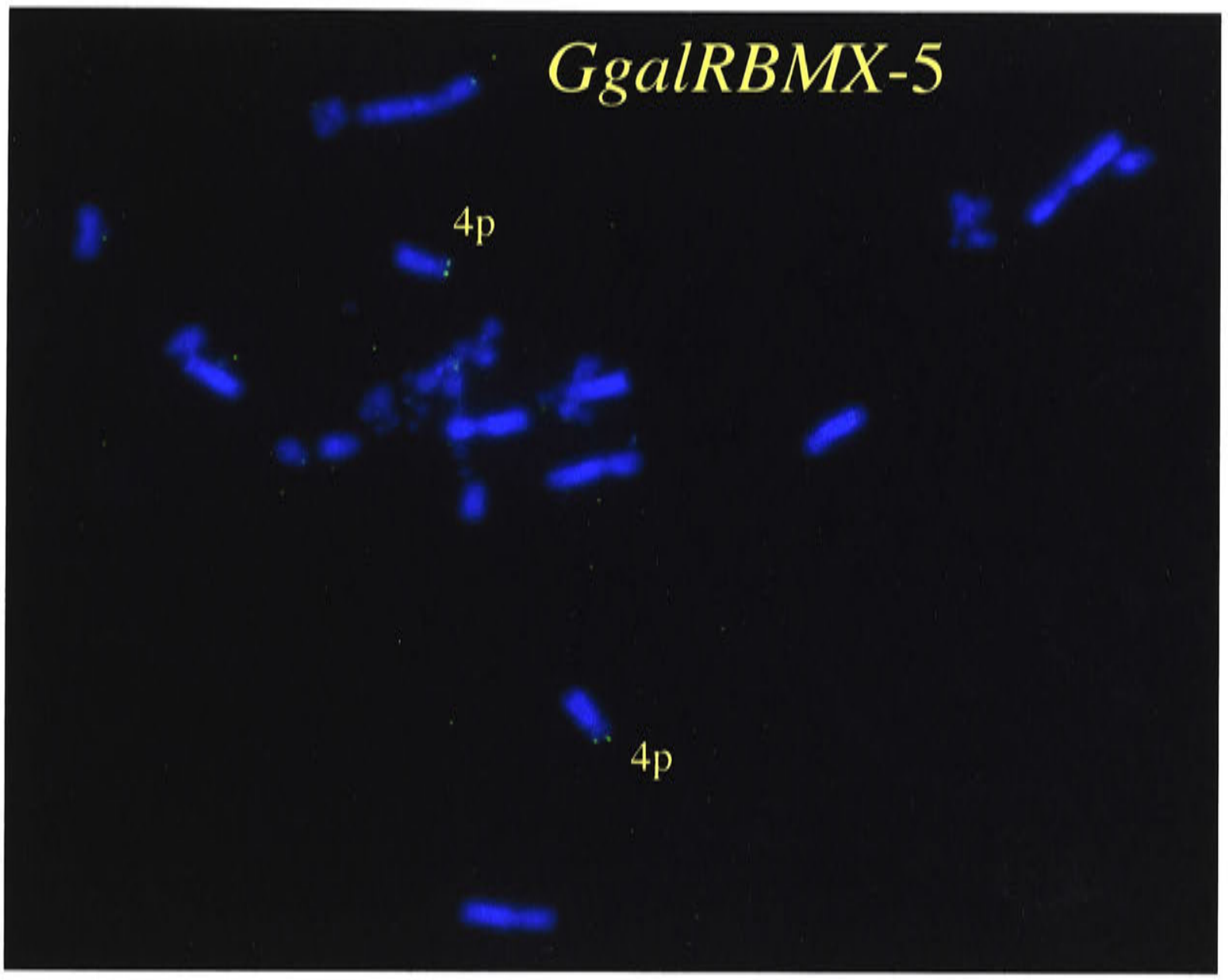


Figure 5.11: FISH mapping of the chicken *RBMX* clone *GgalRBMX*-5. Signals were observed on the short arm of chromosome four. The chromosomes were counterstained with DAPI.

The map location of chicken *ATRX/Y* is clearly important in understanding the evolution of the mammalian sex chromosomes. However, time constraints prevented further attempts to clone and map chicken *ATRX/Y*.

5.5 Conclusions

Three XCR genes that map to human Xq and the marsupial X, were unexpectedly found to map to chromosome 6 in platypus. This demonstrated that not all of the genes shared by the marsupial and eutherian X chromosomes were also located on the monotreme X chromosome. Human XCR genes were mapped in chicken to determine whether these genes were lost from the X chromosome in the monotreme lineage, or added to the X chromosome later in the therian lineage.

Cloning, characterisation and mapping of chicken *SOX3* and *RBMX* revealed that they were both located on the short arm of chicken chromosome 4. This is where other genes from the human XCR map. Chicken chromosome 4p contains *SOX3*, *RBMX*, *BTK*, *UBE2A*, *PGK1*, *CUL4B* and *HPRT1*, and therefore represents at least a significant part of, and possibly the entire, human XCR.

However, without a map location in chicken for genes that are on the X in all mammals, the question of whether these genes had an independent origin from genes that are on the platypus, as well as the therian X, cannot be answered. These results were consistent with either hypothesis; that a large region was lost from the monotreme X, or added to the therian X after the divergence of monotremes.

Chapter 6: The shrinking Y

The mammalian sex chromosomes are thought to have evolved from a pair of autosomes. The proto-Y chromosome acquired a sex-determining locus, and subsequently other male beneficial genes accumulated nearby. Recombination with the proto-X chromosome was suppressed across this region to preserve a male-advantage haplotype, and the Y began to waste away. In the eutherian lineage, autosomal regions were added to the X chromosomes, which then suffered the same degradation. The human X therefore consists of regions conserved from the proto-sex chromosomes, and regions that were added after the divergence of marsupials, but before the eutherian radiation.

Y-borne genes could have arrived on the human Y chromosome in one of two ways. First, genes carried on the original autosomes could have persisted on the Y since the proto-X and -Y diverged because they adopted a selectable (male-specific) function. Secondly, genes could be transposed or retrotransposed from an autosome to the Y chromosome, and have homology only to an autosomal copy. Lahn and Page (1997) placed all of the genes on the Y chromosome into one of two classes. Class I genes were defined as X/Y shared, single copy genes that were ubiquitously expressed (*e.g. SMCY* and *RPS4Y*). Class II genes were defined as multi-copy, Y specific and expressed only in the testis. Class II genes were thought to have been acquired by a "selfish" Y chromosome via transposition (*DAZ*) or retrotransposition (*CDY*). Originally, the testis-specific *SRY* and *RBMY* were thought to be Class II genes. However, both have been shown to possess X homologues. It therefore appears that most genes on the Y chromosome today have persisted from the original proto-Y, or additions to it.

The initial aim of this thesis was to determine whether parts of the human Y chromosomes were conserved from an original Y and shared with the marsupial Y chromosome, and whether parts were recently added and therefore autosomal in marsupials. Genes conserved between the eutherian and marsupial sex chromosomes were then investigated in monotremes and birds to determine the ultimate origin of the mammalian sex chromosomes of all mammals.

To answer these questions, this study involved the cloning, characterisation and mapping of orthologues to human Y-borne genes in marsupials, monotremes and birds. Attempts were made to clone and map the genes *PRY*, *XKRY*, *BPY2*, *VCX/Y*, *PCDHX/Y*, *EIF1AX/Y* and *TB4X/Y*, in addition to the two genes (*USP9X/Y* and *DBX/Y*) investigated

in tammar wallaby as part of an honours project (Waters *et al.*, 2001). The location of these genes in tammar wallaby would determine whether they were recent additions to the human sex chromosomes, or whether they had been conserved on the Y chromosome since before the divergence of marsupials and eutherians 130MyBP.

The ultimate origin of these genes was further explored by attempting to clone and map the same genes, and others conserved on the therian sex chromosomes, in monotremes and chicken. Genes conserved on sex chromosomes in all mammals, including the monotremes, must have resided there since before the prototherian and therian divergence 170MyBP, indicating that they were part of the original mammalian proto-X and -Y. Genes located on the therian, but not the prototherian, sex chromosomes were also mapped in a vertebrate outgroup (chicken), to determine whether they were lost from the monotreme sex chromosomes, or had an independent origin and were added later in the therian lineage.

Summarised in table 6.1 are the results of gene I attempted to clone and map in this study. Results previously obtained by myself and others in the Graves lab are indicated by blue print. Isolation of pseudogenes is indicated by ψ.

Table 6.1: Results of this study

Gene	Tammar wallaby		Platypus		Chicken	
	Clones isolated	Map loation	Clones isolated	Map loation	Clones isolated	Map loation
TB4X	+	5p				
EIF1AX	ψ	-				
USP9X	+	5p	+	-		
DBX	+	5p	-	-		
PCDHX	*PCDH13*	6q				
VCX	*SMARCF1*	5q				
PRY	-	-				
XKRY	-	-				
BPY2	-	-				
RPS4X	ψ	-				
ATRX	+	X/Y	+	6q	-	-
RBMX	+	X/Y	+	6q	+	4p
SOX3	+	X/Y	+	6q	+	4p
G6PD	-	X	-	X	ψ	-

6.1 Phylogenetic analysis of *RBMX/Y*

The relationship of X/Y shared genes to orthologues that are autosomal in other species can give insights into the evolution of genes on sex chromosomes. The autosomal orthologue represents the progenitor of both the X and Y copies. The *RBM* genes were chosen for this analysis because chicken *RBM* sequence was obtained during this study that overlapped with the human, mouse and wallaby *RBMX/Y* sequences. The sequence obtained for platypus *RBM* included only 3' untranslated region, so was not included in this analysis because there was little overlap with the chicken sequence.

RBMX is located on the X and Y chromosome of both eutherian and marsupial mammals. *RBMX/Y* has therefore been conserved on the mammalian sex chromosomes since their divergence 130MyBP. It clusters on chicken chromosome 4, co -localising with other XCR genes.

The chicken *RBMX* sequence obtained in this study was aligned to the tammar wallaby, mouse and human X and Y *RBM* genes. A DNA maximum likelihood analysis, DNA parsimony analysis and protein parsimony analysis was conducted on this alignment, which spanned from the beginning of exon 4 through to the stop codon in exon 9. The trees were constructed using PHYLIP 3.6. DNA maximum likelihood (DNAML) estimation involves finding the evolutionary tree that yields the highest probability of evolving the observed data. All sites are included in the analysis, not just the ones that have undergone change. This method does not assume rate consistency between lineages and can accommodate unequal rates of transversions and transitions. Protein parsimony (PROTPARS) analysis and DNA parsimony (DNAPARS) analysis estimates phylogenies on the fewest number of changes required to explain the observed data. The ancestral sites are estimated from the consensus.

In the DNAML and DNAPARS analyses, the human and mouse *RBMX* clustered. The wallaby *RBMX* was the next closet relative, followed by chicken *RBM*. The close clustering of the *RBMX* genes with chicken *RBM*, implied that there had been little change in this gene over 310MyBP (Figure 6.1).

As observed, the eutherian *RBMX* genes were expected to be more related to each other then either were to wallaby. The therian *RBMX* copies clustered together and were more closely related to each other than any were to chicken. However, the PROTPARS

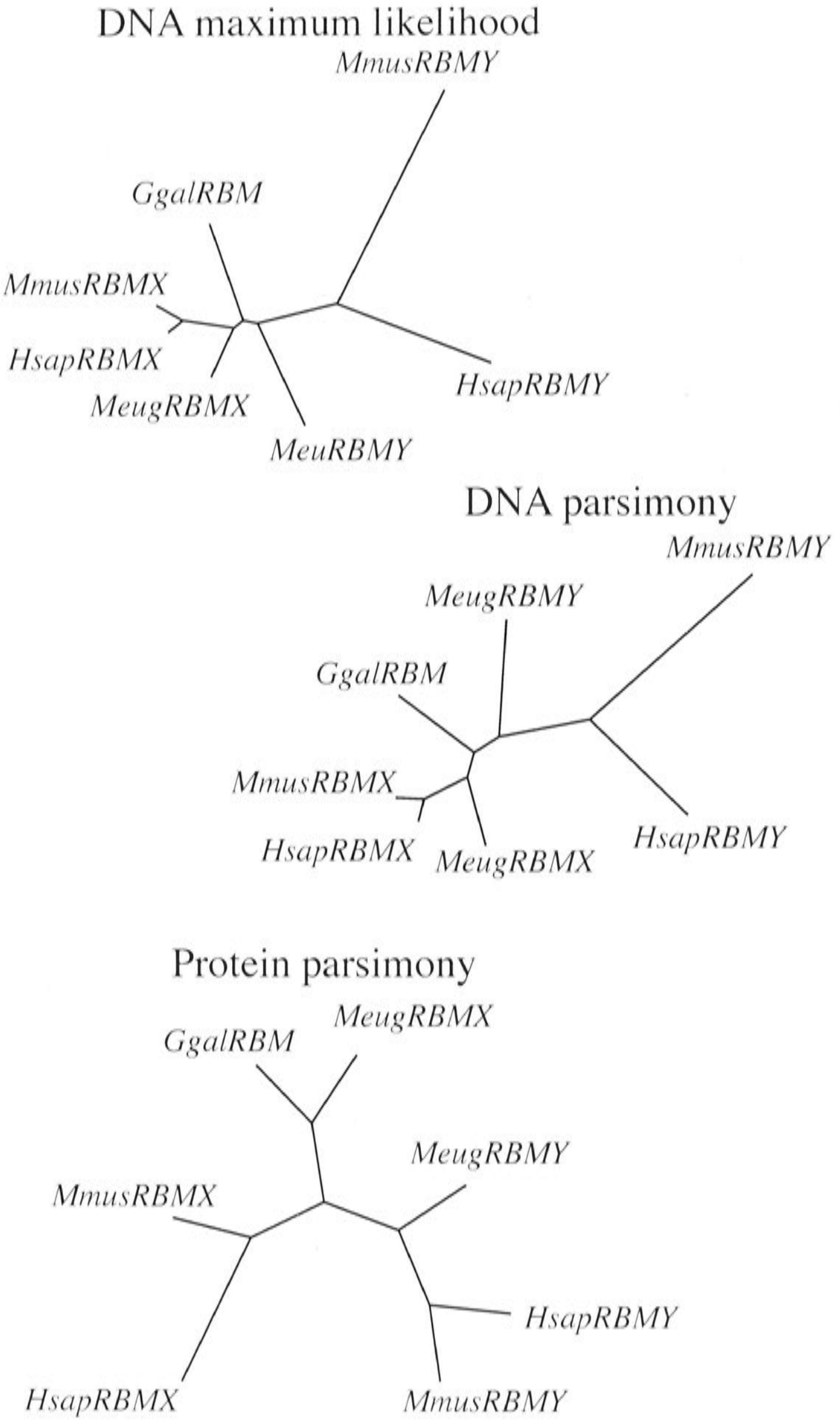


Figure 6.1: Results of phylogenetic studies of *RBM* genes. The unrooted trees were constructed using PHYLIP 3.6. The *RBMX* genes cluster with the chicken *RBMX* homologue. The *RBMY* genes diverged from the *RBMX* cluster.

analysis was not as clear. The human and mouse *RBMX* clustered together and the wallaby *RBMX* and chicken *RBM* clustered together, but separately from human and mouse *RBMX*.

The eutherian Y-borne copies were very divergent from the *RBMX* cluster. The DNAML and DNAPARS analyses indicated that the *RBMY* copies had diverged, both from each other, and from the *RMBX* cluster. The wallaby *RBMY* was also removed from the *RBMX* cluster, although not as far as the eutherian *RBMY* copies implying less rapid divergence. The eutherian *RBMY* copies were more closely related to each other than either were to chicken *RBM*, or to the *RBMX* cluster (human, mouse or wallaby *RBMX*). This indicated a common *RBMY* ancestor after the divergence of the X and Y linked copies, but before the marsupial and eutherian lineages diverged. The more rapid divergence of the *RBMY* genes, in relation to the therian X-borne copies and bird autosomal copy, supports the theory that the Y chromosome evolves more rapidly than the X chromosome and autosomes.

This analysis confirmed that mammalian X-borne genes remain similar to their autosomal orthologues in distantly related vertebrates, whereas, Y-borne genes change rapidly.

6.2 Ancient and added regions of the human sex chromosomes

Comparative mapping of genes on the human X has defined conserved (XCR) and added (XAR) regions. In the same way, in this study, comparative mapping of human Y genes in tammar wallaby defined regions of the human Y chromosome that were conserved on the marsupial sex chromosomes, and regions that were added after the marsupial divergence (130MyBP). Therefore, genes conserved on the human Y chromosome must have persisted on the Y for at least 130MyBP, whereas genes recently added to the human Y chromosome have persisted since before the eutherian radiation (80MyBP), but after the marsupial divergence.

6.2.1 Ancient and added regions of the human Y

The eutherian Y chromosome, like the X chromosome, is composed of a conserved ancient region (YCR) and added region (YAR). The surprising finding from this study was that the most of the human Y was derived from the added region. A total of 8 genes that have been mapped to chromosome 5p of tammar wallaby (*CSF2RA*, *ZFY*, *AMELY*, *USP9Y*, *DBY*, *UTY*, *TB4Y* and *STSP*) are distributed along the length of the human Y chromosome. The PAR gene *SLC25A6* is also autosomal in tammar wallaby, but it maps to chromosome 3/4q. Excluding PAR2 genes, no other genes from the human Y have been mapped to a marsupial autosome. These genes therefore represent a recently added region of the human Y chromosome. Since eight of the nine human Y genes mapped in tammar wallaby are located on 5p, almost all of the YAR arose as the result of a single addition.

What then has become of the ancient Y? There are only three genes on the human Y chromosome that are known to be located on the sex chromosome of tammar wallaby (*SRY*, *SMCY*, and *RBMY*). A fourth gene, *Ube1y*, on the mouse Y chromosome is also on the wallaby Y chromosome, as well as *ATRY*, which is X-specific in eutherians. *RPS4Y* was shown to be located on the sex chromosomes of opossum by Southern blot analysis, although this could not be confirmed in this study. The four genes conserved between human and marsupial sex chromosomes, and therefore representing the ancient Y, form two small conserved regions on the human Y chromosome. The larger of the conserved regions contains *SMCY* and *RMBY*. The maximum extent of this region is from the *DAZ* gene cluster to *STSP* (7Mb), and the minimum only from *RBMY* to *SMCY* (2.7Mb). A second smaller (YCR) region of conserved material resides next to the pseudoautosomal region boundary, and includes *SRY* and *RPS4Y*. It could extend from the PAR boundary to *ZFY* (80kb), or only from *SRY* to *RPS4Y* (20kb) (Figure 6.2). Therefore, only 2.7-7.1Mb remains of the original proto-Y chromosome (the smaller estimate from the minimum size possible for each region, and the larger estimate from the maximum size possible for each region). This amounts to only 8-20% of the euchromatic region. Assuming that the proto-Y was the same size as the conserved region of the modern human X chromosome (because the X and Y chromosomes evolved from an homologous pair of autosomes), the modern human Y chromosome has

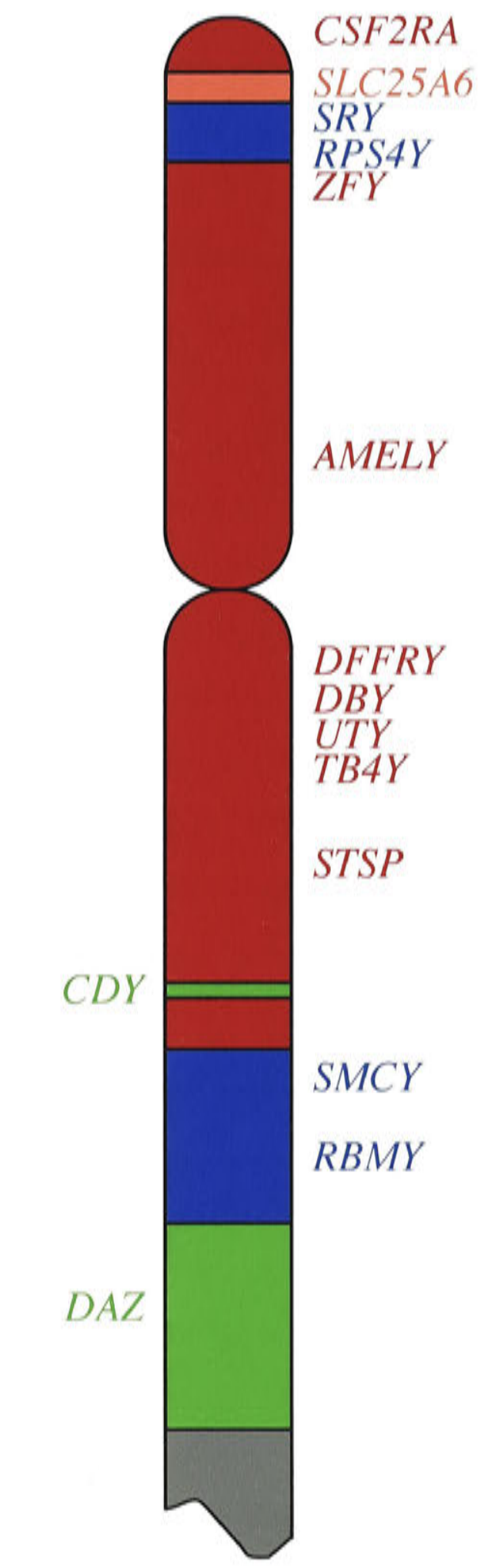


Figure 6.2: Genes on the human Y chromosome. The two small conserved regions are indicated in blue. The large recent addition is indicated by red. Class II genes that have jumped to the Y chromosome by transposition are listed on the left in green. Class I genes are listed on the right.

lost as much as ~110Mb of its original material; equating to as many as 1000 genes. Most of the human Y was derived from the single recent addition.

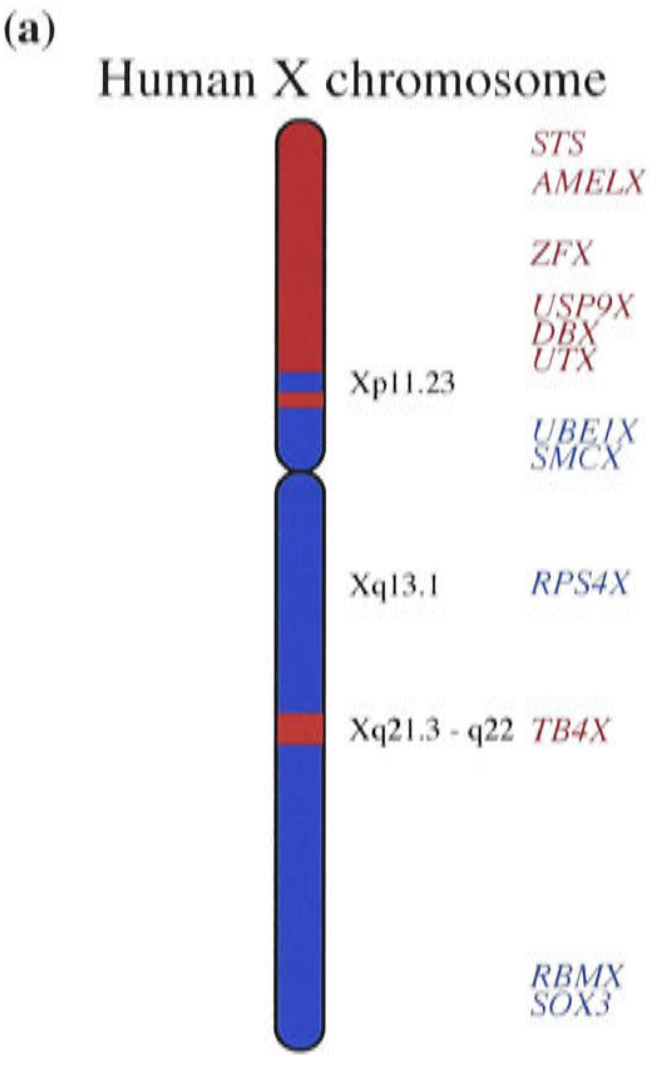
6.2.2 Rearrangement of the ancient and added regions of the human X

The mapping of eight X/Y shared human genes, as well as 15 X-specific genes, to wallaby 5p confirmed a single addition. The genes previously mapped to the ancient and added regions of the human X chromosome have fallen into a distinct XCR region (the long and pericentric region of the X) and XAR region (the short arm of the X, distal to Xp 11.23). Only a local inversion around the fusion point rearranges these regions.

However, the finding that *TB4X/Y* also mapped to wallaby 5p did not accord with this simple picture, because *TB4X* maps on the human X chromosome within a group of XCR genes at Xq21.3-q22. It was therefore the first gene from the long arm of the human X chromosome to map to tammar wallaby chromosome 5p. This indicated that a block of recently added material lies in the heart of the conserved human Xq (Figure 6.3a). These results implied that there was an inversion, which positioned *TB4X* on the long arm of the human X chromosome, perhaps along with other genes that were part of the same recently added material to the human X chromosome (Figure 6.3b).

However, there is a large block of material on the human Xp that has been duplicated on Xq. Thus a number of recently added genes located on Xp have copies located on Xq. *TB4X* could be one of these genes duplicated from Xp (Sargent *et al.*, 1993) (Figure 6.3c). Either of these mechanisms, for mixing of conserved and recently added regions on the human X chromosome, was in addition to the small local rearrangement near the border of the conserved and added regions at Xp11.23, described previously by Wilcox *et al.* (1996).

Thus comparative mapping of human X genes identifies a conserved region (17 genes on the wallaby X) and an added region (ten genes on wallaby 5p and three genes on 1p). Largely the regions have remained separate, except for a small inversion around the fusion point, and an inversion or duplication Xp gene(s) onto Xq. These conserved and added regions have parallels on the human Y chromosome, where mapping human



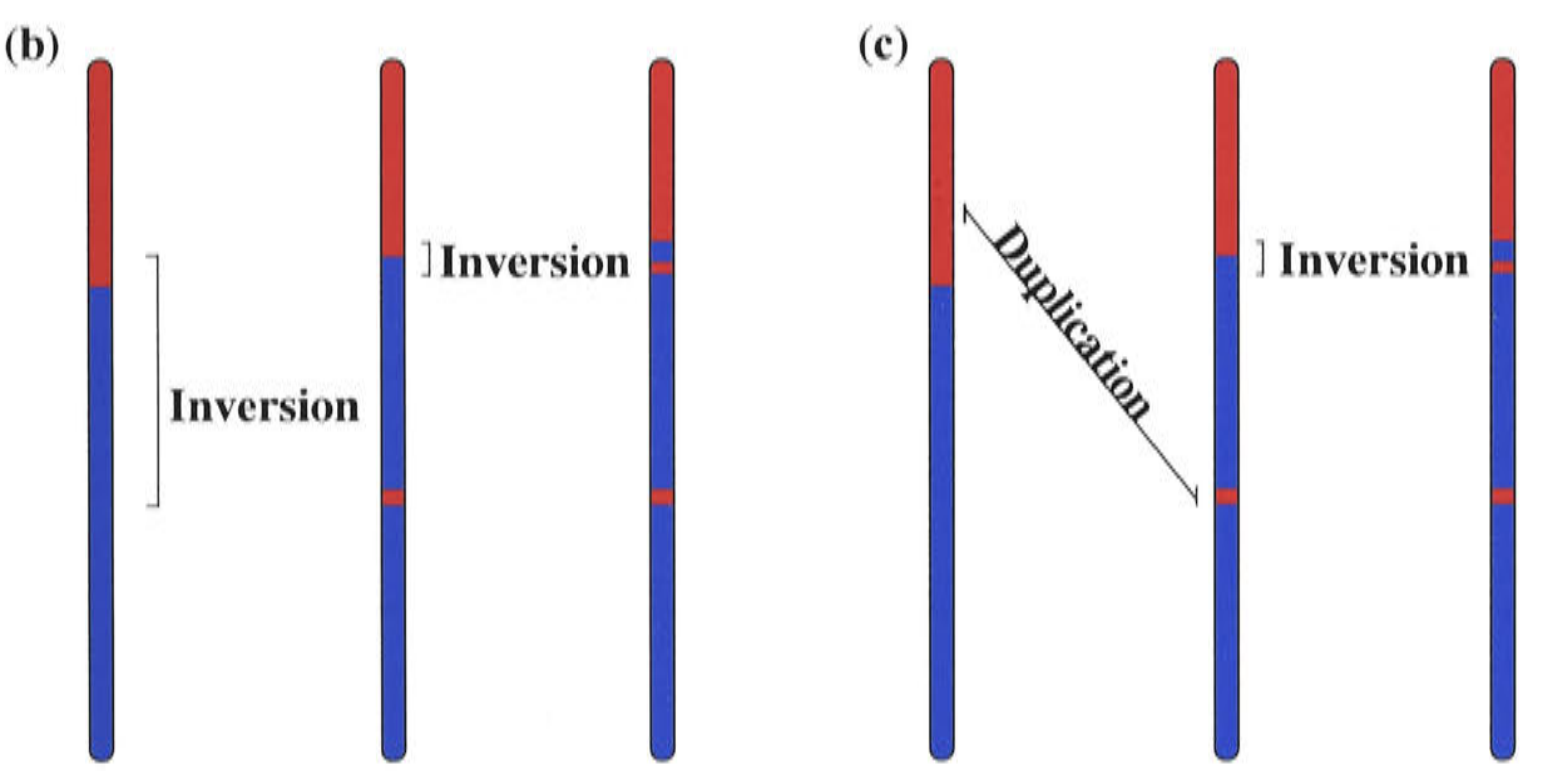


Figure 6.3: a) Recently added regions of the human X chromosome (red) includes *TB4X* on the long arm in the middle of conserved material (blue). b) An inversion may have positioned recently added genes on the long arm of human X chromosome. c) Alternatively, a duplication of genes from human Xp to Xq may have positioned XAR in XCR. A small local inversion occurred for both b) and c) across the XCR/XAR border.

X/Y shared genes identified a conserved region (four genes on the wallaby Y) and an added region (seven genes on wallaby 5p).

6.3 A new geological stratum of the X and Y?

The autosomal location of (previously defined) XCR genes in platypus was an unexpected result, which demonstrated that not all of human Xq derived from the same ancient autosome.

6.3.1 The hole in the platypus X

Somatic cell genetic analysis showed that *HPRT* and *PGK*, genes conserved on the X chromosome of both marsupials (Dawson and Graves, 1984) and eutherians, were autosomal in platypus (Watson and Graves, 1988). Their co-expression in cell hybrids showed that these genes in platypus were syntenic and did not correlate with retention or loss of the platypus X, but they could not be assigned to an autosome. Therefore, there were a total of six genes located on the therian X, but autosomal in platypus. Of these six genes, one had a partner on the eutherian and marsupial Y chromosome *RBMY*, whereas *ATRY* was confined to the marsupial Y chromosome.

The mapping of *ATRX/Y* with *RBMX/Y* and *CDX4* on platypus chromosomes 6 (along with the exclusion of *HPRT*, *PGK* and *FMR1* from the platypus X), demonstrated that a large block of human Xq genes, conserved on the therian sex chromosomes, were missing from on the monotreme sex chromosomes. The regions that contained these genes will be termed X therian conserved regions (XtCR). Two scenarios were possible: XtCR genes, located on platypus chromosome 6q, were added to the therian sex chromosomes after the monotreme divergence. Alternatively, they were part of the original proto-X/Y, but were lost from the monotreme sex chromosomes after the divergence of the monotreme and therian lineages.

Addition to the therian X of an autosomal region containing *CDX4*, *HPRT*, *PGX*, *SOX3*, *RBMX* and *ATRX*, or loss of this region from the mammalian X/Y, is likely to have occurred before recombinational isolation of the X and Y homologues, after the divergence of monotremes from therians. Under both scenarios, *RBMY*, *SRY* and *ATRY*

acquired male specific functions, after monotreme/therian divergence, and therefore probably do not play a role in platypus male specific functions.

To distinguish between these possibilities, XCR genes were mapped in a vertebrate outgroup.

6.3.2 Mapping of XCR genes in an outgroup

Genes on the human X chromosome that were previously mapped in chicken cluster at two locations. XCR genes, present on the therian X, were located on the short arm of chicken chromosome 4. XAR genes, located on the sex chromosomes in eutherians but not marsupials, were located on chicken chromosome 1. One gene from the added region of the human X chromosome, *DMD*, is located on chicken chromosome 10.

My finding that the XtCR genes *RBMX* and *SOX3* clustered with the other XCR genes on the short arm of chicken chromosome 4 (Table 6.2), indicated that chicken chromosome 4p was equivalent to platypus chromosome 6.

Table 6.2: Locations of human X-linked genes that have been mapped in chicken, tammar wallaby, platypus and zebrafish.

	Gene	Human	Tammar wallaby	Platypus	Chicken	Zebrafish
Human Xp genes	*SCMLR*	Xp22			1	
	ZFX	Xp21.3-22.2	5p	2/1	1	
	DMD	Xp21.2-21.3	5p	1	10	
	NROB1	Xp21.2-21.3			1q21	
	OTC	Xp21.1	1p	2	1q13-14	
Human Xq genes	*AR*	Xq11.2-12	X	X		
	CDX4	Xq13.3		6q		LG14
	PGK1	Xq13.3	X	A	4	
	ATRX	Xq13.1-21.1	X	6q		
	BTK	Xq21.2-22			4p14	
	CUL4B	Xq23			4	
	UBE2A	Xq24-25			4p11-14	
	HPRT1	Xq26.1	X	A	4p14	LG14
	RBMX	Xq26	X	6q	4p	LG14
	SOX3	Xq26-27	X		4p	LG14
	FMR1	Xq27.3		A		
	RCP	Xq28	X	X		LG11
	G6PD	Xq28	X	X		

If this region conserved on the therian X was on the original proto-X and lost from the monotreme X, it would be expected that XCR gene conserved on the platypus, as well as the therian X chromosome, would also map to chicken chromosome 4. If the region represents an independent addition to the therian X, genes on the platypus X, as well as the marsupial X, would be expected to map elsewhere in chicken. Unfortunately, none were successfully mapped in chicken (Table 6.2).

Further attempts to map *G6PD*, and other genes conserved on the X chromosome of all mammals (*e.g. AR* and *RCP*), would answer this question.

Not withstanding lack of mapping data from chicken, mapping data from a different outgroup could give insights into the origin of human Xq.

6.3.3 Two independent origins for gene on human Xq

Fish are more distantly related to mammals than are chickens. The fish lineage diverged from other vertebrates 400MyBP.

In zebrafish, two XtCR genes (*CDX4* and *RBMX*) that were on the X chromosome in therians, but on chromosome 6 in platypus, were located in linkage group 14 (LG14). A third XtCR gene, *HPRT1*, (located on an unknown platypus autosome) also mapped to zebrafish LG14. *SOX3* is also located on zebrafish LG14, but its location is unknown in platypus. Thus LG14 contains four XtCR genes from the long arm of the human X chromosome that map to the marsupial, but not the monotreme X (*CDX4*, *HPRT1*, *RBMX* and *SOX3*). It therefore appears that LG14 in zebrafish, chromosome 4p in chicken and chromosome 6q in platypus are all equivalent to each other, representing a significant part of the conserved region of the human X chromosome

Only one gene that was on the X chromosome of all mammals (XCR) has been mapped in fish. Red Cone Pigment (*RCP*) was located in the conserved region of the human X chromosome (Xq28). It was located on the tammar wallaby X chromosome and also the platypus X chromosome. Significantly, *RCP* did not map to LG14, but instead was located on LG11 in zebrafish. *RCP* was therefore the only gene on the platypus X chromosome that has been mapped in a vertebrate outgroup. LG11 could therefore represent XCR genes from the platypus, as well as the therian X chromosome. In *Tetraodon nigroviridid* (pufferfish) *ATRX* and *RCP* were also located on different chromosomes (F, Grützner, personal communication), lending further support to the

idea that they were on different chromosomes in the vertebrate ancestor. This needs to be confirmed by mapping more genes located on the platypus X in chicken and zebrafish (*e.g. AR* and *HPRT1*).

I therefore propose that platypus X/ zebrafish LG11 and platypus 6/ chicken 4p/ zebrafish LG14 represent different ancestral blocks. The original proto-X is represented by platypus X/ zebrafish LG11. Platypus 6/ chicken 4p/ zebrafish LG14 was added after the monotreme and therian lineages diverged, but before eutherians and marsupials diverged (Figure 6.4). The conserved region of the human X chromosome therefore consists of two evolutionarily distinct regions. The first is the region conserved between the X chromosomes of all mammalian groups (proto-X). This region includes *UBE1X*, *AR*, *G6PD*, *RCP*, *GDX*, *PLP*, *GLA*, *F8*, *F9*, *GATA1*, *ALAS2* and *P3*. The second is the region that was added to the sex chromosomes of the therian ancestor, which includes *HPRT*, *SOX3*, *RBMX*, *CDX4* and *ATRX*. This region was added to the sex chromosomes 130-170MyBP.

Lahn and Page (1999) defined the entire long arm of the human X chromosome as Stratum 1. Stratum 2 included the proximal region of human Xp. Nine of the 12 genes mapped to the platypus X chromosome were located on human Xq (Stratum 1). The remaining three, *UBE1X*, *GATA1* and *ALAS2*, were located in Stratum 2 of the human X chromosome. Therefore Strata 1 and 2 did not correlate to XCR and XtCR. The strata were not defined by the time when they arrived on the sex chromosomes, but rather the time at which they were recombinationally isolated from the Y chromosome. Stratum 2 was estimated to be between 130-170 million years old, which was compatible with *UBE1X*, *GATA1* and *ALAS2* having been on the original proto-sex chromosomes.

The XtCR genes are all located in Stratum 1., which was estimated to be between 240-320 million years old. The arrival of the XtCR Stratum 1 genes *HPRT*, *SOX3*, *RBMX*, *CDX4* and *ATRX* on the therian sex chromosomes, between 130-170MyBP, completely contradicted the age of Stratum 1. The recombinational isolation of Stratum 1 from the Y could not have occurred before it arrived on the sex chromosomes (130-170MyBP).

Thus the long arm of the human X chromosome was comprised of two blocks with independent origins. The first remained from the original mammalian proto-X

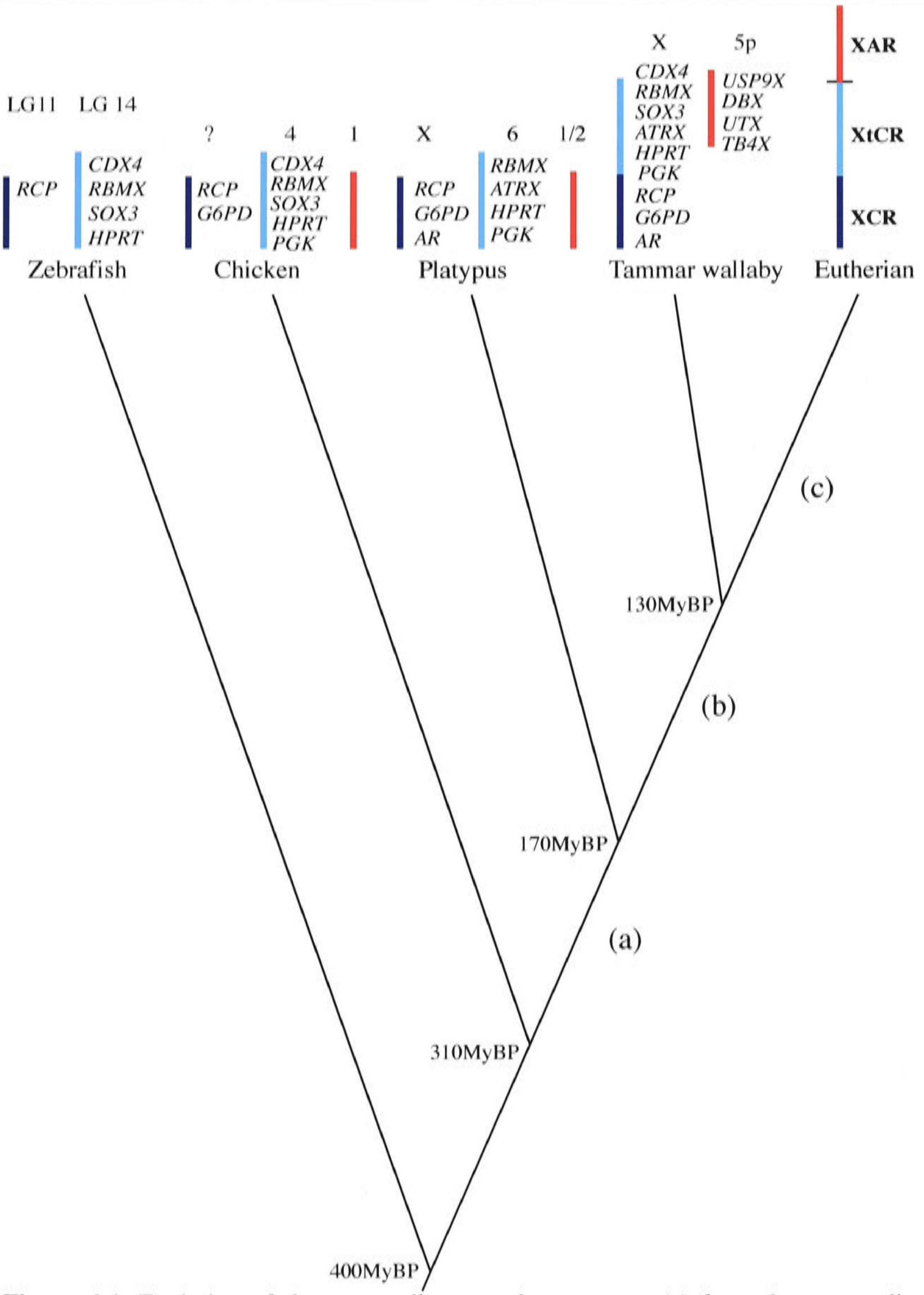


Figure 6.4: Evolution of the mammalian sex chromosomes (a) from the mammalian proto-X/Y XCR (blue) contained *RCP, G6PD, AR* and other genes located on the monotreme sex chromosomes. (b) Addition of XtCR region (light blue) containing *SOX3, RBMX, ATRX* and other genes located on platypus chromosome6 and zebrafish LG14 to the sex chromosomes of the therian ancestor.(c) XAR (red) containing genes on the human X and Y chromosomes that are located on an autosome in wallaby were added to the sex chromosomes of the eutherian ancestor.

chromosome (XCR). The second was added after monotremes diverged from the therian lineage, but before the marsupial and eutherian lineages diverged (XtCR).

The independent origin of genes on human Xq implied that genes previously classified as YCR genes could also have different origins.

6.3.4 Origins of the Y conserved region

Since the ancient mammalian X was equivalent to the ancient mammalian Y, the oldest part of the Y was homologous to the platypus X/ zebrafish LG11. The only Y gene that remains from this most ancient Y is *UBE1Y*. It is conserved on the monotreme Y chromosome in the region that pairs with the X chromosome. *UBE1Y* was therefore a member of the mammalian proto-Y. It is on the mouse Y but absent from the human Y, and is the only eutherian Y-borne gene known to have been part of the proto-Y.

Potentially, the modern human Y chromosome contains no genes that have survived from the original mammalian proto-Y chromosome. However, *SMCY* and *RPS4Y* have not been mapped in platypus, chicken or zebrafish and could therefore have been members of the proto-Y chromosome with *Ube1y*. The origin of *SMCY* and *RPS4Y* could be resolved by mapping them in platypus and chicken.

The next oldest region of the Y was equivalent to platypus chromosome 6/ chicken 4p/ zebrafish LG14 (*SRY*, *RBMY* and *ATRY*). Thus, of four (previously defined) YCR genes on the human Y chromosome (*SRY*, *RBMY*, *RPS4Y* and *SMCY*) and a fifth on the mouse Y (*Ube1y*), only *Ube1y* was present on the original proto-Y. Two genes (*RBMY* and *ATRY*) were added in the therian lineage (YtCR). *RBMY* and *ATRY* were located on chromosome 6 in platypus. *RBMY* and *SOX3* (X-borne homologue of *SRY*) were located on LG14 in zebrafish, and were probably part of the same ancient cluster of genes (which also included *ATRY*) added to the therian sex chromosomes (Figure 6.5). These genes were therefore not part of the mammalian proto-Y.

RBMY and *SRY* acquired their roles in male sex and reproduction after they were added to the therian sex chromosomes. *SRY* replaced an ancient mammalian sex-determining gene, and *RBMY* acquired a role in therian spermatogenesis, ensuring their survival.

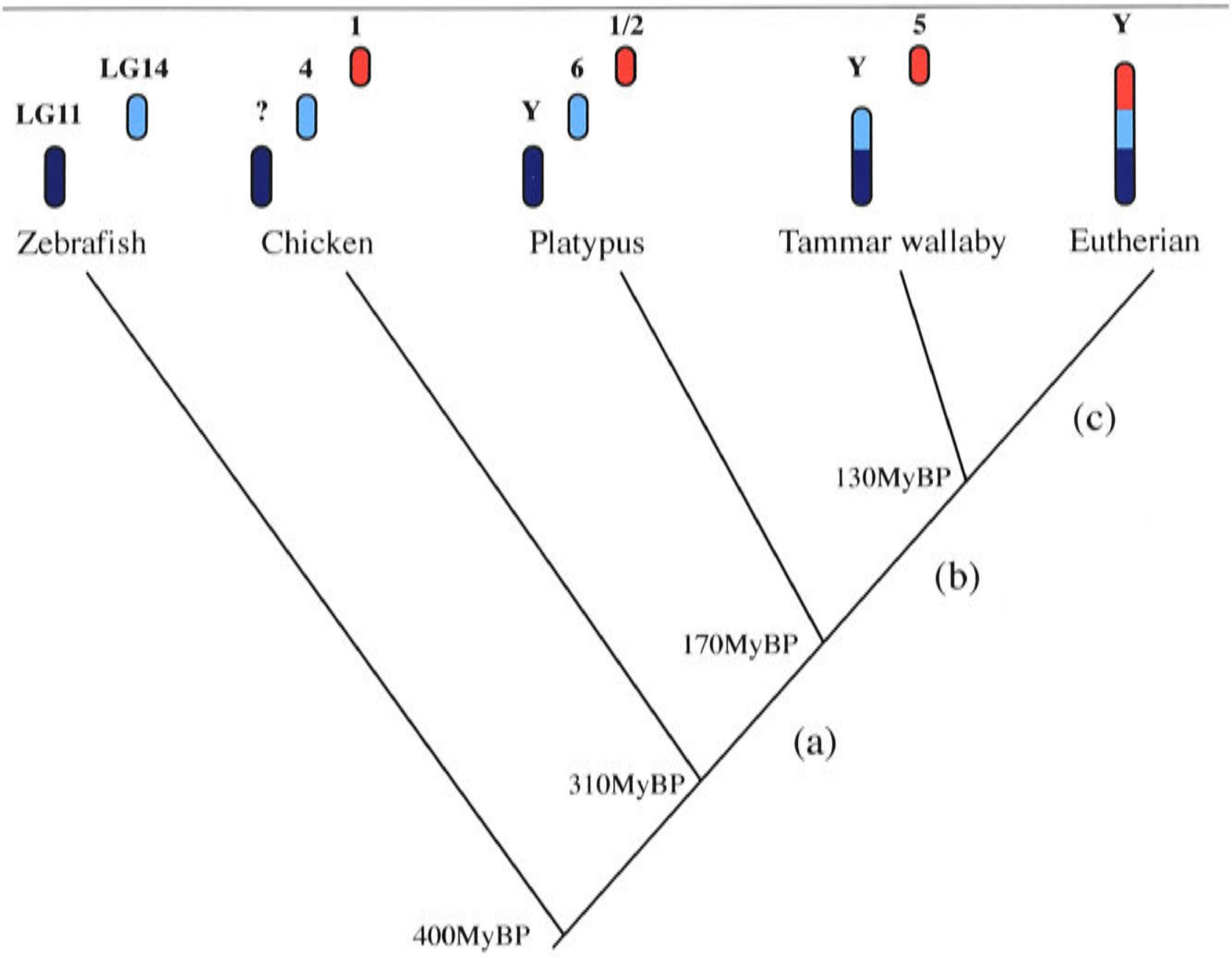


Figure 6.5: The eutherian Y chromosome consists of YCR, YtCR and YAR. (a) YCR has persisted from the proto-Y chromosome (dark blue). (b) YtCR was added to the Y chromosome in the therian ancestor 170-130MyBP (light blue). (c) YAR was added after marsupials diverged from eutherians but before the eutherian radiation (130-80MyBP). YCR and YtCR are separate in distantly related vertebrates (fish) and were therefore seperate in the mammalian ancestor.

Human	Wallaby	Platypus	Chicken	Zebrafish
X/Y	5p	1/2	1	?
X	1p	?	?	?
X/Y	X/Y	X	4p	LG14
X	X/ *UBE1Y*	X/Y	?	LG11

Figure 6.6: Known location of regions added to the eutherian Y chromosome in wallaby, platypus, chicken and zebrafish. The original Y (dark blue) had three independent additions (light blue, orange and red). Only genes from two additions persist on the human Y.

The mammalian Y chromosome has been losing genes since it became recombinationally isolated from the X chromosome, and perhaps none of the original proto-Y chromosome remains on the modern human Y and only a few remain from the YtCR addition. Therefore, most of the genes borne by the eutherian Y have persisted from a recent addition in the eutherian lineage. These additions were located on different autosomes in distantly related mammalian, and vertebrate, species (Figure 6.6).

The critical functions that the Y chromosome plays, ensures that it has survived for as long as it has. The loss of these functions could permit the complete loss of the Y.

6.4 Decline and fall of the Y

6.4.1 Loss of genes from the Y chromosome

Clearly, most of the 1000 genes in the original YCR, and the 400 genes in the original added region have been lost from the human Y. Genes that have survived on the Y chromosome for a significant length of time must therefore have acquired a function that is selectable. *SRY* (sex determination) and *RBMY* (spermatogenesis) are examples of genes that acquired male specific functions and persisted on the Y chromosome. These two genes lie on the Y chromosome of all therian mammals studied. *SRY* evolved from *SOX3*, and is the dominant sex-determining gene that acts as a switch to set off a cascade of events leading to male determination. *RBMY* evolved from *RMBX*, gaining testis specific expression and a role in mammalian spermatogenesis.

However, even a selectable function is not always enough to assure the maintenance of a Y-borne gene. There are several genes that which are retained on the Y only in one lineage or another. *Eif2s3x/y* is an X/Y shared gene residing on the mouse Y but not the human Y chromosome (Ehrmann *et al.*, 1998). It also maps to 5p in tammar wallaby (M. Delbridge, personal communication) and is therefore part of the recent addition to the sex chromosomes of the eutherian ancestor. *Eif2s3y* is located in the ΔSxr^b deletion interval of the mouse Y, and introduction of an *Eif2s3y* transgene rescues spermatogenic failure in mice with the ΔSxr^b deletion (Mazeyrat *et al.*, 2001), demonstrating that it plays an important role in mouse spermatogenesis.

Even a critical Y-borne gene may be lost if a backup gene evolves. In human, along with an X borne copy, there is a retroposed copy of *EIF2S3* on chromosome 12 that is expressed in the testis (Ehrmann *et al.*, 1998) that may have compensated for the loss of *EIF2S3Y*. Similarly, in mouse there is a retroposed copy of *Ddx3* (the X homologue of *Dby*) on chromosome 1, which is transcribed in the testis (Leroy *et al.*, 1989). This could explain why mouse spermatogenesis can proceed without *Dby*, whereas *DBY* is critical to human spermatogenesis.

UBE1Y is a Class I gene with an X-borne homologue, *UBE1X*, in mice and marsupials. It is also present on the short arm of the monotreme X chromosome. This region of the monotreme X chromosome pairs with the first element in the meiotic chain and is therefore part of the platypus PAR. Its location on the sex chromosomes of all three mammalian groups indicated that it has persisted there since before the divergence of the prototherian and therian lineages 170MyBP. *UBE1Y* is present on the Y chromosome in marsupial, mouse and new world monkey, but not in human, chimpanzee and old world monkey (Mitchell *et al.*, 1998). The location of *UBE1X/Y* on the tammar wallaby sex chromosomes means that it was on the original mammalian X and Y. To exist for so long on the therian Y chromosome *UBE1Y* must have acquired a selectable male specific function. This was lost, and its function replaced by another gene in the lineages that led to hominids and old world monkeys (Figure 6.7).
Thus, the male specific role of a gene on the Y chromosome could be lost if another gene can fulfil that same role.

Clearly, YAR genes with critical male specific functions (*e.g. USP9Y* and *DBY*) acquired their roles after arriving on the Y, perhaps replacing the roles of other genes. The YtCR genes, with critical male-specific functions (*e.g. SRY* and *RBMY*), arrived on the sex chromosomes of the therian ancestor and may have replace the roles of genes that no longer exist on the therian Y chromosome.

Marsupials possess an *ATRY*, whereas eutherian mammals do not. Both eutherians and marsupials have an X-borne *ATRX*, which demonstrated that *ATRY* was on the sex chromosomes in the therian ancestor. It was hypothesised that *ATRY* could play a conserved role in marsupial sex-determination that was lost in the eutherian lineage (Pask *et al.*, 2000). Perhaps *ATRY* was the original TDF that was replaced by *SRY* (a

Figure 6.7: *UBE1Y* evolution in mammals. *UBE1Y* was on the ancestral Y chromosome. It was lost independently from the Y in the ancestor to old world monkeys, hominids and marmosets. In all other eutherians tested it remains as an X/Y shared gene, and in monotremes it remains as a PAR gene. Adapted from Mitchell *et al.*, 1998.

later arrival), thus making *ATRY* redundant in eutherians. To test this hypothesis *ATRY* was mapped in the most distantly related mammals, the monotremes. To be the ancestral sex-determining gene, *ATRX/Y* should be located on the monotreme sex chromosomes.

6.4.2 Evolution of *ATRX/Y* and sex determination

ATRY is located on the marsupial Y chromosome and could potentially play a role in the marsupial sex-determining pathway (Pask *et al.*, 2000) (Figure 6.8). However, the eutherian Y chromosome does not appear to contain *ATRY* related sequences. It has therefore been conserved on the marsupial Y chromosome and lost from the eutherian Y chromosome. The location of *ATRX/Y* in monotremes is very interesting, as it could play a role in an ancient mammalian sex-determining pathway.

Surprisingly, cloning and mapping of *ATRX/Y* in platypus revealed that it was located on the proximal long arm of chromosome 6, co-locating with *RMBX/Y* and *CDX4*.

The unexpected result that *ATRX/Y* was autosomal in platypus immediately contradicted the hypothesis that *ATRY* represented an ancestral mammalian sex-determining gene. *ATRX/Y* was therefore part of YtCR that was added to the mammalian sex chromosomes after the monotreme and therian lineages diverged. *ATRY* could, therefore, not have been the ancestral sex determining gene. *ATRY* must have gained a function in marsupials to persist on the Y. However, in the eutherian lineage either no function for *ATRY* was obtained, or, a function was obtained and then lost. *ATRY* was subsequently lost from the eutherian Y. There must have been another gene responsible for ancient mammalian sex determination, which could still determine sex in platypus (Figure 6.9).

6.4.3 Fall of the Y

With the mammalian Y chromosome rapidly degrading and getting smaller, how much of it can disappear before it is no longer stable? Could it disappear completely? It seems that the answer is "yes".

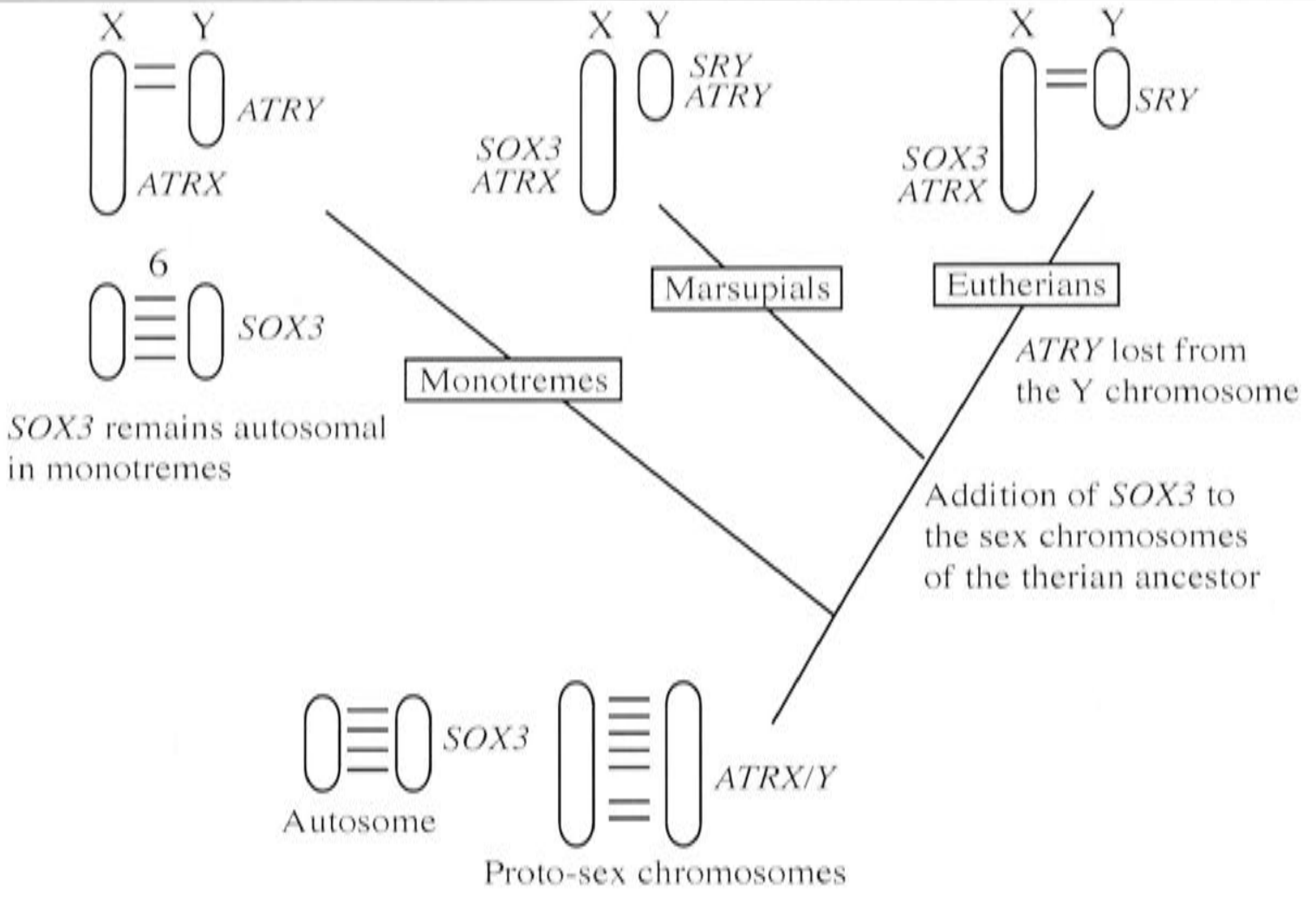


Figure 6.8: Evolution of the mammalian *ATRX/Y* proposed by Pask *et al.* (1999). *ATRY* was the original TDF whereas *SOX3* was autosomal. *SOX3* was added to the sex chromosomes in the therian lineage. On the Y chromosome it evolved into *SRY*, replacing the function of *ATRY*, which was lost in eutherians.

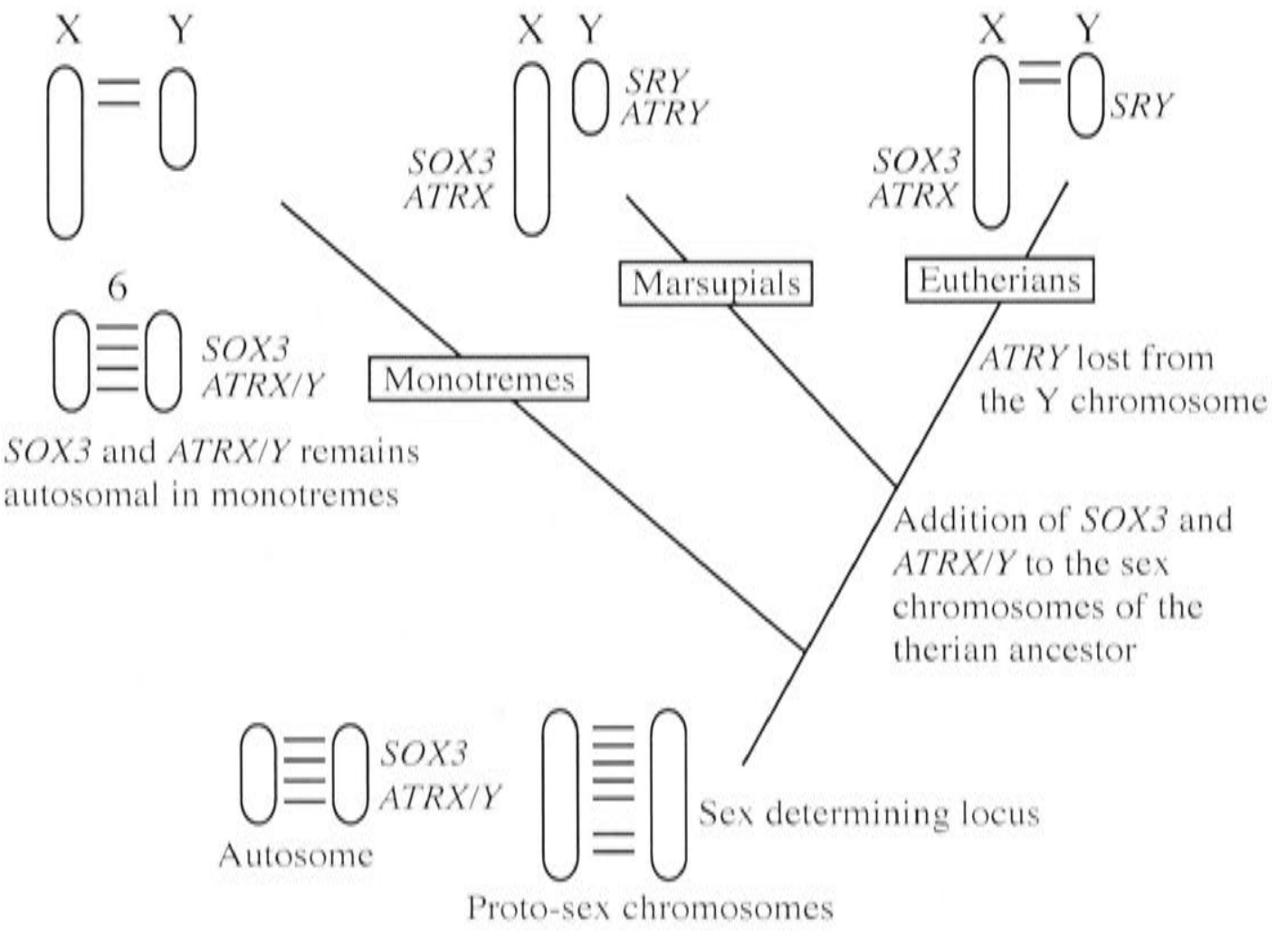


Figure 6.9: Revised evolution of *ATRX/Y*. *ATRY* was not the ancestral sex determining locus. A different genes must have been the ancestral sex determining locus. It was added to the sex chromosomes of the therian ancestor. *ARTY* was lost in eutherians and retained in marsupials, where it could still play a role in sex determination.

There are two species of mole vole (genus *Ellobius*) that have no Y chromosome and no *SRY*. A third *Ellobius* species has a standard XY male: XX female system, in which the males possess an *SRY* (Section 1.2.1.4). How it is possible for these two species of mole vole to determine sex without a Y chromosome is still not clear. The male determining function of *SRY* must have been replaced by another gene in *E. lutescens* and *E. tancrei*. This made *SRY*, and ultimately the Y chromosome, redundant. There could be a new sex-determining locus on a different pair of what were autosomes, which are still homomorphic.

The loss of the Y chromosome meant that other genes on the mammalian Y chromosome were also lost. This included the conserved gene *Rbmy* that plays a role in spermatogenesis, as well as genes (*Usp9y* and *Eif2s3y*) recently added to the eutherian Y chromosome that are important in rodent spermatogenesis. There are two ways this could have occurred. The important functions of genes on the missing Y chromosomes could have been replaced before the Y was lost. Alternatively, these genes could have been translocated to an autosome, and still retain their male specific functions. A search for these genes in the *Ellobius* species would answer these questions. *ZFY* is absent, but the presence of other Y genes in unknown. To know what is on the Y chromosome of *E. fuscocapillus* would also help in understanding how the *E. lutescens* and *E. tancrei* Y chromosome met its doom. Once the sex-determining locus on a Y chromosome becomes non-functional, and genes critical to male specific functions have been replaced or moved from the Y chromosome, the end is near for that Y chromosome (Figure 6.10).

The fall of the Y chromosome has therefore occurred in *Ellobius*. And could be close to its fall in species that commonly have XY females (*e.g.* akodont rodents and lemmings; discussed in section 1.5). *Ellobius* demonstrates that the Y chromosome is not immortal, and the retention of its male determining function is crucial to its survival. The process of Y degradation is continual and it may only be a matter of time before other mammalian species lose their Y chromosome.

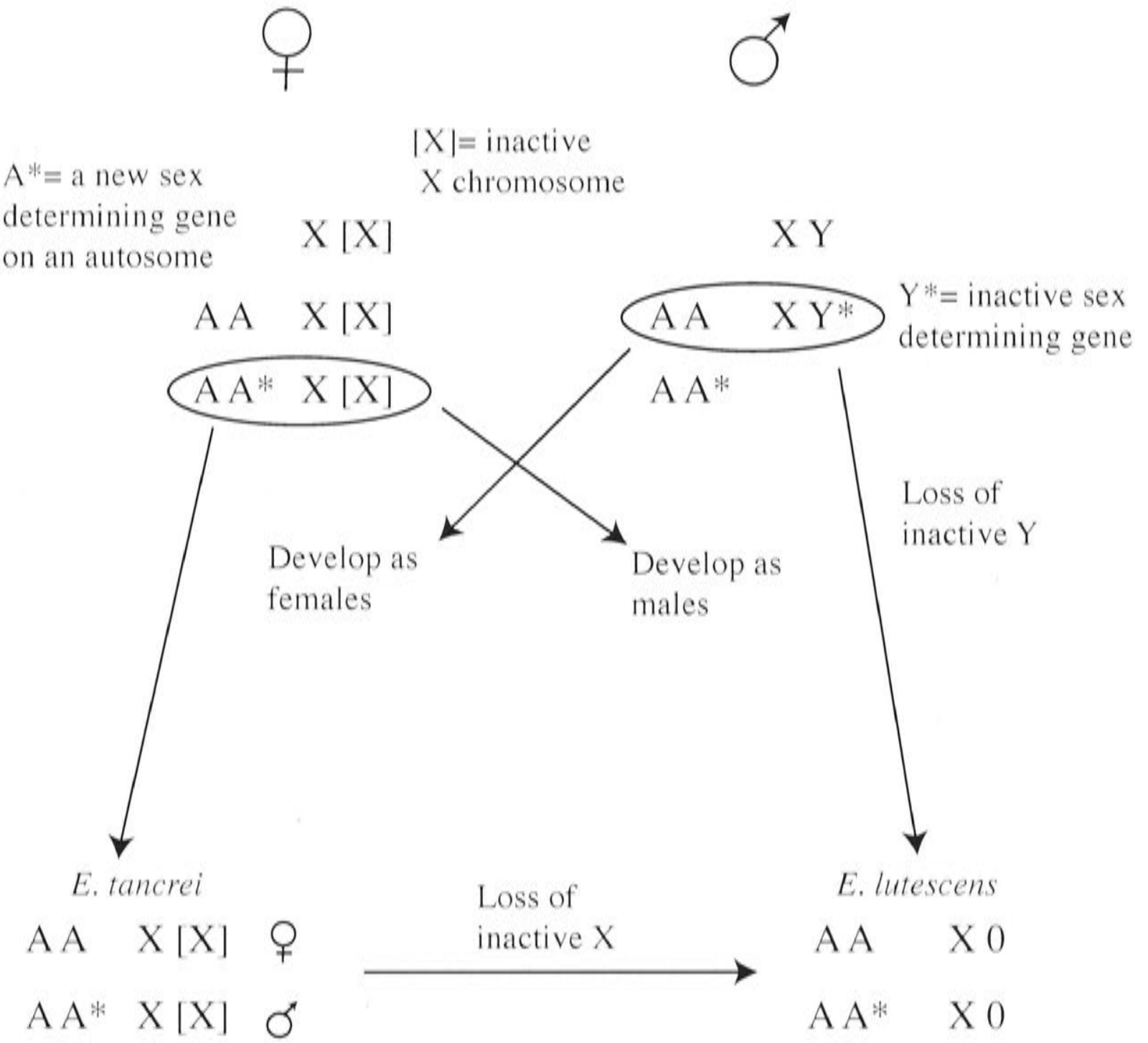


Figure 6.10: Loss of the Y chromosome in *Ellobius*. A new sex determining gene arrises on an autosome, and SRY becomes redundent. The inactive Y chromosome is lost. *E. lutescens* lost the Y and has X0 males and females. *E. lutescens* could have gone through an intermediate time when they had X [X] males and females, as is seen with *E. tancrei*. The inactive X chromosome was then lost.

6.5 Conclusions

The mammalian X and Y chromosomes evolved from an homologous pair of autosomes. The original proto-X and -Y chromosomes contained the same set of genes, but throughout its history the Y chromosome has been subjected to selective forces that have caused the loss of most of the genes it once contained. Only genes that are beneficial to males have survived on the Y for a long evolutionary period.

Previous comparative mapping of human X/Y shared genes in marsupials defined conserved (XCR) and recently added (XAR) regions of the human X chromosome. My mapping of human XY shared genes in the tammar wallaby showed that the human Y chromosome, too, is composed of an ancient region (YCR, represented on the marsupial Y), and a Y added region (YAR, represented on marsupial autosomes). The position of genes on tammar chromosome 5p implies that the human Y chromosome is derived primarily from a single recent addition. The ancient YCR constitutes only 8-20% of the euchromatic human Y, so has been virtually obliterated.

Unexpectedly, the locations of XCR genes in platypus, chicken and fish defined two evolutionarily distinct origins for the long arm of the human X chromosome. One region, on the X also in platypus, represents the original proto-X (XCR). A second region, mapping to chromosome 6 in platypus, was added to the therian X chromosome (XtCR) after the divergence of monotremes 170MyBP.

Mapping of these genes also defined a region on the human Y that has been conserved only on the therian Y (YtCR), but is autosomal in platypus. This region was therefore not a part of the proto-Y, but was added after monotremes diverged from marsupials 170MyBP, but before marsupials diverged from eutherians 130MyBP. This finding contradicts Lahn and Page's influential "Geological layers" of the sex chromosomes hypothesis, which states that the entire long arm of the X, and corresponding regions of the Y, represent the original proto-XY chromosome, and genes in this region diverged on the X and Y 180-240MyBP.

Thus of the 35 genes located on the human Y chromosome, none were present on the original proto-Y and only four survive from the region added 130-170MyBP. Most of the human Y chromosome is therefore derived from the recent (~80MyBP) addition to eutherian sex chromosomes (Figure 6.11). Thus of the 1400 genes on the original Y and the additions, the mammalian Y chromosome has lost all but 35. With time, more will be lost, until the human Y disappears entirely.

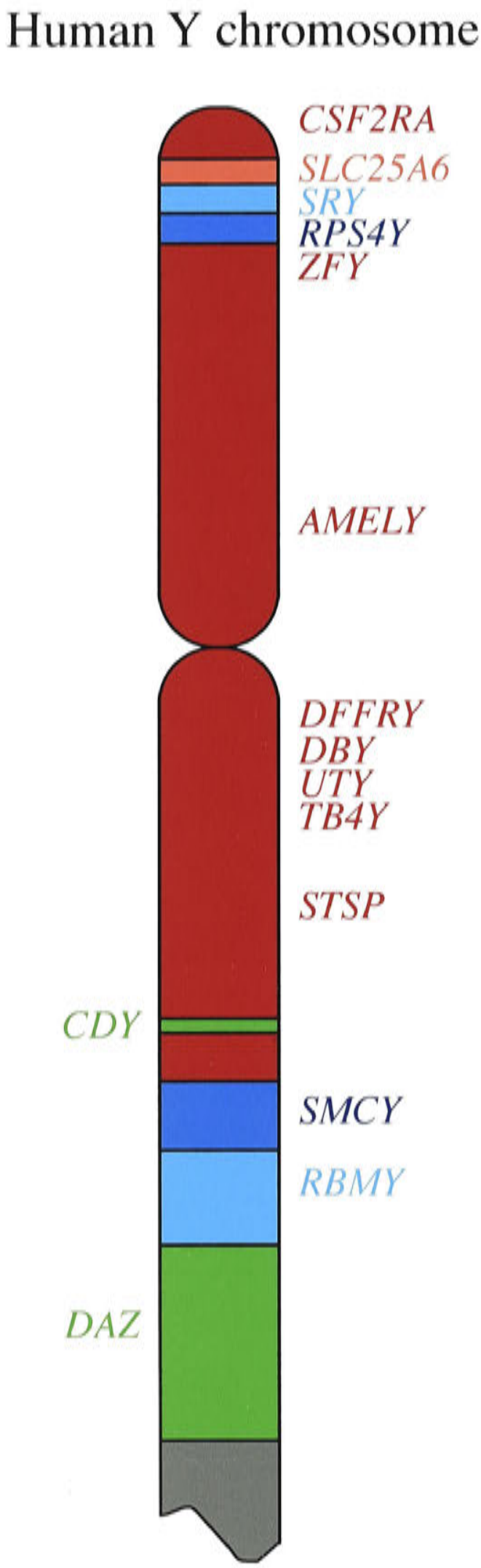


Figure 6.11: Genes on the human Y chromosome. The two small YtCR regions are indicated in light blue. Potentially, *RPS4Y* and *SMCY* could be the only genes remainng from the original proto-Y chromosome. The large recent addition is indicated by red. Class II genes that have jumped to the Y chromosome by transposition are listed on the left in green. Class I genes are listed on the right.

References

Abassi, V. (1998) Growth and normal puberty. *Pediatrics,* **102,** 507-11.

Achermann, J. C., Ito, M., Hindmarsh, P. C. and L., J. J. (1999) A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal. *Nat Genet,* **22,** 125-126.

Adler, D. A., Rugarli, E. I., Lingenfelter, P. A., Tsuchiya, K., Poslinski, D., Liggitt, H. D., Chapman, V. M., Elliott, R. W., Ballabio, A. and Disteche, C. M. (1997) Evidence of evolutionary up-regulation of the single active X chromosome in mammals based on Clc4 expression levels in Mus spretus and Mus musculus. *Proc Natl Acad Sci USA,* **94,** 9244-8.

Aitken, R. J. and Marshall Graves, J. A. (2002) The future of sex. *Nature,* **415,** 963.

Arango, N. A., Lovell-Badge, R. and Behringer, R. R. (1999) Targeted mutagenesis of the endogenous mouse Mis gene promoter: in vivo definition of genetic pathways of vertebrate sexual development. *Cell,* **99,** 409-19.

Arnermann, J., Epplen, J. T., Cook, H. J., Sauermann, U., Engel, W. and Schmidtke, J. (1987) A human Y chromosome sequence expressed in testicular tissue. *Nucleic Acids Res*, **15**, 8713-24.

Ashley, T. and Fredga, K. (1994) The curious normality of the synaptic association between the sex chromosomes of two arvicoline rodents: Microtus oeconomus and Clethrionomys glareolus. *Hereditas,* **120,** 105-11.

Auer, H., Mayr, B., Lambrou, M. and Schleger, W. (1987) An extended chicken karyotype, including the NOR chromosome. *Cytogenet Cell Genet,* **45,** 218-21.

Bachtrog, D. and Charlesworth, B. (2002) Reduced adaptation of a non-recombining neo-Y chromosome. *Nature,* **416,** 323-6.

Banting, G. S., Pym, B., Darling, S. M. and Goodfellow, P. N. (1989) The MIC2 gene product: epitope mapping and structural prediction analysis define an integral membrane protein. *Mol Immunol*, **26,** 181-8.

Baverstock, P. M., Adams, R., Plolkinghorne and Gelder, M. (1982) Sex linked enzyme in birds: Z-chromosome conservation but not dosage conservation. *Nature*, **296**, 763-66.

Belote, J. M. and Baker, B. S. (1982) Sex determination in Drosophila melanogaster: analysis of transformer-2, a sextransforming locus. *Proc Natl Acad Sci USA,* **79,** 1568-72.

Benton, M. J. (1990) Phylogeny of the major tetrapod groups: morphological data and

divergence dates. *J Mol Evol,* **30,** 409-24.

Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N. and Fellous, M. (1990) Genetic evidence equating SRY and the testis-determining factor. *Nature,* **348,** 448-50.

Bick, Y. A. and Jackson, W. D. (1967) DNA content of monotremes. *Nature,* **215,** 192-3.

Bick, Y. A., Murtagh, C. and Sharman, G. B. (1973) The chromosomes of an egg-laying mammal Tachyglossus aculeatus (the echidna). *Cytobios,* **7,** 233-43.

Bishop, C. E., Whitworth, D. J., Qin, Y., Agoulnik, A. I., Agoulnik, I. U., Harrison, W. R., Behringer, R. R. and Overbeek, P. A. (2000) A transgenic insertion upstream of Sox9 is associated with dominant XX sex reversal in the mouse. *Nat Genet,* **26,** 490-494.

Blair, H. J., Reed, V., Gormally, É., Wilson, J. B., Novak, J., McInnes, R. R., Phillips, S. J., Taylor, B. A. and Boyd, Y. (2000) Positioning of five genes (CASK, ARX, SAT, IMAGE, cDNAs 248928 and 253949) from the human X chromosome short arm with respect to evolutionary breakpoints on the mouse Y chromosome. *Mamm Genome*, **11**, 710-2.

Blanco, P., Sargent, C. A., Boucher, C. A., Mitchell, M. and Affara, N. A. (2000) Conservation of PCDHX in mammals; expression of human X/Y genes predominantly in brain. *Mamm Genome,* **11,** 906-14.

Blaschke, R. J. and Rappold, G. A. (1997) Man to mouse--lessons learned from the distal end of the human X chromosome. *Genome Res,* **7,** 1114-7.

Boumil, R. M. and Lee, J. T. (2001) Forty years of decoding the silence in X-chromosome inactivation. *Hum Mol Genet,* **10,** 2225-32.

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. and Willard, H. F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature,* **349,** 38-44.

Brown, G. M., Furlong, R. A., Sargent, C. A., Erickson, R. P., Longepied, G., Mitchell, M., Jones, M. H., Hargreave, T. B., Cooke, H. J. and Affara, N. A. (1998) Characterisation of the coding sequence and fine mapping of the human DFFRY gene and comparative expression analysis and mapping to the Sxrb interval of the mouse Y chromosome of the Dffry gene. *Hum Mol Genet,* **7,** 97-107.

Bull, J. J. (1983) *Evolution of Sex Determining Mechanisms*, The Benjamin/Cummings Publishing Company, California.

Burgoyne, P. S., Mahadevaiah, S. K., Sutcliffe, M. J. and Palmer, S. J. (1992) Fertility in mice requires X-Y pairing and a Y-cromosomal 'spermatogenesis' gene mapping to the long arm. *Cell*, **71,** 391-98.

Burrows, W. and Ryder, O. A. (1997) Y-chromosome variation in great apes. *Nature*, **385,** 125-6.

Caron, H., Peter, M., van Sluis, P., Speleman, F., de Kraker, J., Laureys, G., Michon, J., Brugieres, L., Voute, P. A., Westerveld, A. and et al. (1995) Evidence for two tumour suppressor loci on chromosomal bands 1p35-36 involved in neuroblastoma: one probably imprinted, another associated with N-myc amplification. *Hum Mol Genet*, **4,** 535-9.

Carrel, L., Cottle, A. A., Goglin, K. C. and Willard, H. F. (1999) A first-generation X-inactivation profile of the human X chromosome. *Proc Natl Acad Sci USA*, **96,** 14440-44.

Carrel, L. and Willard, H. F. (1999) Heterogeneous gene expression from the inactive X chromosome: An X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. *Proc Natl Acad Sci USA*, **96,** 7364-69.

Chai, N. N., Salido, E. C. and Yen, P. H. (1997) Multiple functional copies of the RBM gene family, a spermatogenesis candidate on the human Y chromosome. *Genomics*, **45,** 355-61.

Chapman, T., Liddle, L. F., Kalb, J. M., Wolfner, M. F. and Partridge, L. (1995) Cost of mating in Drosophila melanogaster females is mediated by male accessory gland products. *Nature*, **373,** 241-4.

Charchar, F. J., Svartman, M., El-Mogharbel, N., Kirby, P. J., Ciccodicola, A., D'Esposito, M. and Graves, J. A. M. (Submitted) Complex events in the evolution of the huaman pseudoautosomal region 2 (PAR2).

Charlesworth, B. (1991) The evolution of sex chromosomes. *Science*, **251**, 1030-33.

Charlesworth, B. (1996a) Background selection and patterns of genetic diversity in Drosophila melanogaster. *Genet Res*, **68,** 131-49.

Charlesworth, B. (1996b) The evolution of chromosomal sex determination and dosage compensation. *Curr Biol*, **6,** 149-62.

Charlesworth, B. (2002) Evolutionary genetics: the evils of abstinence from sex. *Curr*

Biol, **12,** R56-8.

Charlesworth, B. and Charlesworth, D. (1997) Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genet Res,* **70,** 63-73.

Charlesworth, B. and Charlesworth, D. (2000) The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci,* **355,** 1563-72.

Chiurazzi, P., Hamel, B. C. and Neri, G. (2001) XLMR genes: update 2000. *Eur J Hum Genet,* **9,** 71-81.

Ciccodicola, A., D'Esposito, M., Esposito, T., Gianfrancesco, F., Migliaccio, C., Miano, M. G., Matarazzo, M. R., Vacca, M., Franze, A., Cuccurese, M., Cocchia, M., Curci, A., Terracciano, A., Torino, A., Cocchia, S., Mercadante, G., Pannone, E., Archidiacono, N., Rocchi, M., Schlessinger, D. and D'Urso, M. (2000) Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region. *Hum Mol Genet,* **9,** 395-401.

Clark, M. S., Edwards, Y. J., McQueen, H. A., Meek, S. E., Smith, S., Umrania, Y., Warner, S., Williams, G. and Elgar, G. (1999) Sequence scanning chicken cosmids: a methodology for genome screening. *Gene,* **227,** 223-30.

Cleaveland, L. R. (1947) The origin and evolution of meiosis. *Science,* **105,** 287-88.

Clemson, C. M., Chow, J. C., Brown, C. J. and Lawrence, J. B. (1998) Stabilization and localization of Xist RNA are controlled by separate mechanisms and are not sufficient for X inactivation. *J Cell Biol,* **142,** 13-23.

Clinton, M. and Haines, L. C. (1999) An overview of factors influencing sex determination and gonadal development in birds. *Cell. Mol. Life Sci.,* **55,** 876-86.

Cockwell, A., MacKenzie, M., Youings, S. and Jacobs, P. (1991) A cytogenetic and molecular study of a series of 45,X fetuses and their parents. *J Med Genet,* **28,** 152-5.

Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. and Lovell-Badge, R. (1996) A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development,* **122,** 509-20.

Conrad, C., Hierl, T., Glaser, B., Taylor, K., Zeitler, S., Chandley, A. C. and Schempp, W. (1996) High-resolution fluorescence in situ hybridization of RBM- and TSPY- related cosmids on released Y chromatin in humans and pygmy

chimpanzees. *Chromosome Res,* **4,** 201-6.

Cooper, D. W., Johnston, P. G., Watson, J. M. and Graves, J. A. M. (1993) X inactivation in marsupials and monotremes. *Semin. Dev. Biol.,* **4,** 117-28.

Cooper, D. W., VandeBerg, J. L., Sharman, G. B. and Poole, W. E. (1971) Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. *Nat New Biol,* **230,** 155-7.

Costanzi, C. and Pehrson, J. R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature,* **393,** 599-601.

Crow, J. F. (2000) The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet,* **1,** 40-7.

D'Esposito, M., Ciccodicola, A., Gianfrancesco, F., Esposito, T., Flagiello, L., Mazzarella, R., Schlessinger, D. and D'Urso, M. (1996) A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X inactivation. *Nat Genet,* **13,** 227-9.

Dal Zotto, L., Quaderi, N. A., Elliott, R., Lingerfelter, P. A., Carrel, L., Valsecchi, V., Montini, E., Yen, C. H., Chapman, V., Kalcheva, I., Arrigo, G., Zuffardi, O., Thomas, S., Willard, H. F., Ballabio, A., Disteche, C. M. and Rugarli, E. I. (1998) The mouse Mid1 gene: implications for the pathogenesis of Opitz syndrome and the evolution of the mammalian pseudoautosomal region. *Hum Mol Genet,* **7,** 489-99.

Dallas, P. B., Pacchione, S., Wilsker, D., Bowrin, V., Kobayashi, R. and Moran, E. (2000) The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. *Mol Cell Biol,* **20,** 3137-46.

Dawson, G. W. and Graves, J. A. (1984) Gene mapping in marsupials and monotremes. I. The chromosomes of rodent- marsupial (Macropus) cell hybrids, and gene assignments to the X chromosome of the grey kangaroo. *Chromosoma,* **91,** 20-7.

de la Chapelle, A. (1972) Nature and origin of males with XX sex chromosomes. *Am J Hum Genet,* **24,** 71-105.

De Rosa, M., De Brasi, D., Zarrilli, S., Paesano, L., Pivonello, R., D'Agostino, A. and LongoBardi, S. (1997) Short stature and azoospermia in a patient with Y chromosome long arm deletion. *J Endocrinol Invest,* **20,** 623-8.

Debrand, E., Chureau, C., Arnaud, D., Avner, P. and Heard, E. (1999) Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol,*

19, 8513-25.

Dechend, F., Williams, G., Skawran, B., Schubert, S., Krawczak, M., Tyler-Smith, C. and Schmidtke, J. (2000) TSPY variants in six loci on the human Y chromosome. *Cytogenet Cell Genet,* **91,** 67-71.

Delbridge, M. L., Harry, J. L., Toder, R., O'Neill, R. J., Ma, K., Chandley, A. C. and Graves, J. A. (1997) A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome. *Nat Genet,* **15,** 131-6.

Delbridge, M. L., Lingenfelter, P. A., Disteche, C. M. and Graves, J. A. (1999) The candidate spermatogenesis gene RBMY has a homologue on the human X chromosome. *Nat Genet,* **22,** 223-4.

Delbridge, M. L., Subbarao, M. N., Cook, H. J., Bhasin, S. and Graves, J. A. (1998) Evolution of mammalian HNRPG and its relationship with the putative azoospermia factor RBM. *Mamm Genome,* **9,** 168-70.

Dimitratos, S. D., Stathakis, D. G., Nelson, C. A., Woods, D. F. and Bryant, P. J. (1998) The location of human CASK at Xp11.4 identifies this gene as a candidate for X-linked optic atrophy. *Genomics*, **51**, 308-9.

Disteche, C. M. (1995) Escape from X inactivation in human and mouse. *Trends Genet,* **11,** 17-22.

Disteche, C. M., Brannan, C. I., Larsen, A., Adler, D. A., Schorderet, D. F., Gearing, D., Copeland, N. G., Jenkins, N. A. and Park, L. S. (1992) The human pseudoautosomal GM-CSF receptor alpha subunit gene is autosomal in mouse. *Nat Genet,* **1,** 333-6.

Dorfman, D. M., Genest, D. R. and Reijo Pera, R. A. (1999) Human DAZL1 encodes a candidate fertility factor in women that localizes to the prenatal and postnatal germ cells. *Hum Reprod,* **14,** 2531-6.

Dorit, R. L., Akashi, H. and Gilbert, W. (1995) Absence of polymorphism at the ZFY locus on the human Y chromosome. *Science,* **268,** 1183-5.

Eberhart, C. G., Maines, J. Z. and Wasserman, S. A. (1996) Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. *Nature,* **381,** 783-5.

Ehrmann, I. E., Ellis, P. S., Mazeyrat, S., Duthie, S., Brockdorff, N., Mattei, M. G., Gavin, M. A., Affara, N. A., Brown, G. M., Simpson, E., Mitchell, M. J. and Scott, D. M. (1998) Characterization of genes encoding translation initiation

factor eIF-2γ in mouse and human: sex chromosome localization, escape from X-inactivation and evolution. *Hum Mol Genet,* **7,** 1725-37.

Eichwald, E. J. and Silmser, C. R. (1955) Untitled communication. *Transplant-Bull,* **2,** 148-49.

Ellegren, H. (2002) Human mutation--blame (mostly) men. *Nat Genet,* **31,** 9-10.

Ellegren, H. and Fridolfsson, A. K. (1997) Male-driven evolution of DNA sequences in birds. *Nat Genet,* **17,** 182-4.

Elliott, D. J., Millar, M. R., Oghene, K., Ross, A., Kiesewetter, F., Pryor, J., McIntyre, M., Hargreave, T. B., Saunders, P. T., Vogt, P. H., Chandley, A. C. and Cooke, H. (1997) Expression of RBM in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm. *Proc Natl Acad Sci USA,* **94,** 3848-53.

Ellison, J. W., Wardak, Z., Young, M. F., Gehron Robey, P., Laig-Webster, M. and Chiong, W. (1997) PHOG, a candidate gene for involvement in the short stature of Turner syndrome. *Hum Mol Genet,* **6,** 1341-7.

Ellison, J. W., Ramos, C., Yen, P. H. and Shapiro, L. J. (1992) Structure and expression of the human pseudoautosomal gene XE7. *Hum Mol Genet,* **1,** 691-6.

Esposito, T., Gianfrancesco, F., Ciccodicola, A., Montanini, L., Mumm, S., D'Urso, M. and Forabosco, A. (1999) A novel pseudoautosomal human gene encodes a putative protein similar to Ac-like transposases. *Hum Mol Genet,* **8,** 61-7.

Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. and Bianchi, M. E. (1992) SRY, like HMG1, recognizes sharp angles in DNA. *Embo J,* **11,** 4497-506.

Fisher, E. M., Beer-Romero, P., Brown, L. G., Ridley, A., McNeil, J. A., Lawrence, J. B., Willard, H. F., Bieber, F. R. and Page, D. C. (1990) Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. *Cell,* **63,** 1205-18.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection,* Oxford University Press.

Ford, C. E., Jones, K. W., Polani, P. E., Almida, J. C. and Briggs, J. H. (1959) A sex chromosome anomaly in the case of gonadal dysgenesis. *Lancet,* **1,** 711-13.

Foresta, C., Ferlin, A. and Moro, E. (2000) Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. *Hum Mol Genet,* **9,** 1161-9.

Foster, J. W., Brennan, F. E., Hampikian, G. K., Goodfellow, P. N., Sinclair, A. H., Lovell-Badge, R., Selwood, L., Renfree, M. B., Cooper, D. W. and Graves, J. A. (1992) Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. *Nature*, **359,** 531-3.

Foster, J. W. and Graves, J. A. (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. *Proc Natl Acad Sci USA*, **91,** 1927-31.

Franco, B., Guioli, S., Pragliola, A., Incerti, B., Bardoni, B., Tonlorenzi, R., Carrozzo, R., Maestrini, E., Pieretti, M., Taillon-Miller, P. and et al. (1991) A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature*, **353**, 529-36.

Fredga, K. (1988) Abberant chromosomal sex-determining mechanisms in mammals, with special reference to species with XY females. *Philos Trans R Soc Lond B Biol Sci*, **322,** 83-95.

Freije, D., Helms, C., Watson, M. S. and Donis-Keller, H. (1992) Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science*, **258,** 1784-7.

Fridolfsson, A. K., Cheng, H., Copeland, N. G., Jenkins, N. A., Liu, H. C., Raudsepp, T., Woodage, T., Chowdhary, B., Halverson, J. and Ellegren, H. (1998) Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proc Natl Acad Sci USA*, **95,** 8147-52.

Gianfrancesco, F., Esposito, T., Montanini, L., Ciccodicola, A., Mumm, S., Mazzarella, R., Rao, E., Giglio, S., Rappold, G. and Forabosco, A. (1998) A novel pseudoautosomal gene encoding a putative GTP-binding protein resides in the vicinity of the Xp/Yp telomere. *Hum Mol Genet*, **7,** 407-14.

Gianfrancesco, F., Sanges, R., Esposito, T., Tempesta, S., Rao, E., Rappold, G., Archidiacono, N., Graves, J. A., Forabosco, A. and D'Urso, M. (2001) Differential divergence of three human pseudoautosomal genes and their mouse homologs: implications for sex chromosome evolution. *Genome Res*, **11,** 2095-100.

Gibbons, R. J., Picketts, D. J., Villard, L. and Higgs, D. R. (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell*, **80,** 837-45.

Glas, R., Graves, J. A. M., Toder, R., Ferguson-Smith, M. A. and O'Brien, P. C. (1999) Cross-species chromosome painting between human and marsupial directly demonstrates the ancient region of the mammalian Y chromosome. *Mamm Genome*, **10**, 1115-16.

Glaser, B., Grutzner, F., Taylor, K., Schiebel, K., Meroni, G., Tsioupra, K., Pasantes, J., Rietschel, W., Toder, R., Willmann, U., Zeitler, S., Yen, P., Ballabio, A., Rappold, G. and Schempp, W. (1997) Comparative mapping of Xp22 genes in hominoids--evolutionary linear instability of their Y homologues. *Chromosome Res*, **5,** 167-76.

Glaser, B., Grutzner, F., Willmann, U., Stanyon, R., Arnold, N., Taylor, K., Rietschel, W., Zeitler, S., Toder, R. and Schempp, W. (1998a) Simian Y chromosomes: species-specific rearrangements of DAZ, RBM, and TSPY versus contiguity of PAR and SRY. *Mamm Genome*, **9,** 226-31.

Glaser, B., Myrtek, D., Rumpler, Y., Schiebel, K., Hauwy, M., Rappold, G. A. and Schempp, W. (1999) Transposition of SRY into the ancestral pseudoautosomal region creates a new pseudoautosomal boundary in a progenitor of simian primates. *Hum Mol Genet*, **8,** 2071-8.

Glaser, B., Yen, P. H. and Schempp, W. (1998b) Fibre-fluorescence in situ hybridization unravels apparently seven DAZ genes or pseudogenes clustered within a Y-chromosome region frequently deleted in azoospermic males. *Chromosome Res*, **6,** 481-6.

Gordo, I. and Charlesworth, B. (2000) The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics*, **154,** 1379-87.

Graves, J. A. (1995) The origin and function of the mammalian Y chromosome and Y-borne genes- -an evolving understanding. *Bioessays*, **17,** 311-20.

Graves, J. A. (1998) Interactions between SRY and SOX genes in mammalian sex determination. *Bioessays*, **20,** 264-9.

Graves, J. A. and Delbridge, M. L. (2001) The X--a sexy chromosome. *Bioessays*, **23,** 1091-4.

Graves, J. A. M. (1982) 5-azacytidine-induced re-expressionof alleles on the inactive X chromosome in a hybrid mouse cell line. *Exp Cell Res*, **141**, 99-105.

Graves, J. A. M., Disteche, C. M. and Toder, R. (1998) Gene dosage in the evolution and function of mammalian sex chromosomes. *Cytogenet Cell Genet*, **80,** 94-

103.

Greenfield, A., Scott, D., Pennisi, D., Ehrmann, I., Ellis, P., Cooper, L., Simpson, E. and Koopman, P. (1996) An H-YDb epitope is encoded by a novel mouse Y chromosome gene. *Nat Genet,* **14,** 474-8.

Harley, V. R., Jackson, D. I., Hextall, P. J., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R. and Goodfellow, P. N. (1992) DNA binding activity of recombinant SRY from normal males and XY females. *Science,* **255,** 453-6.

Hayman, D. L. and Martin, P. G. (1974) *Animal Cytogenetics,* Stuttgart: Gebruder Borntraeger, Berlin.

Herzing, L. B., Romer, J. T., Horn, J. M. and Ashworth, A. (1997) Xist has properties of the X-chromosome inactivation centre. *Nature,* **386,** 272-5.

Hickey, D. A. (1982) Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics,* **101,** 519-531.

Hirano, S., Ono, T., Yan, Q., Wang, X., Sonta, S. and Suzuki, S. T. (1999) Protocadherin 2C: a new member of the protocadherin 2 subfamily expressed in a redundant manner with OL-protocadherin in the developing brain. *Biochem Biophys Res Commun,* **260,** 641-5.

Hope, R. M., Cooper, S. and Wainwright, B. (1990) Globin Macromolecular Sequences in Marsupials and Monotremes. In *Mammal from Pouches and Eggs: Genetics and Evolution of Marsupials and Monotremes,* (Eds, Graves, J. A. M., Hope, R. M. and Cooper, D. W.) Commonwealth Scientific and Industrial Research, Melbourne, pp. 147-173.

Huang, B., Wang, S., Ning, Y., Lamb, A. N. and Bartley, J. (1999) Autosomal XX sex reversal caused by duplication of SOX9. *Am J Med Genet,* **87,** 349-53.

Huber, R., Hansen, R. S., Strazzullo, M., Pengue, G., Mazzarella, R., D'Urso, M., Schlessinger, D., Pilia, G., Gartler, S. M. and D'Esposito, M. (1999) DNA methylation in transcriptional repression of two differentially expressed X-linked genes, GPC3 and SYBL1. *Proc Natl Acad Sci USA,* **96,** 616-21.

Hurst, L. D. and Randerson, J. P. (1999) An eXceptional chromosome. *Trends Genet,* **15,** 383-5.

Jacobs, P. A. and Strong, J. A. (1959) A case of human intersexuallity having a possible XXY sex-determining mechanism. *Nature,* **183,** 302-303.

Janke, A., Xu, X. and Arnason, U. (1997) The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia and Eutheria. *Proc Natl Acad Sci USA*, **94**, 1276-81.

Janke, A., Magnell, O., Wieczorek, G., Westerman, M. and Arnason, U. (2002) Phylogenetic analysis of 18S rRNA and the mitochondrial genomes of the wombat, *Vombaus ursinus*, and the spiny anteater, *Tachyglossus aculeatus*: increased support for the Marsupionta hypotesis. *J Mol Evol*, **54**, 71-80.

Jegalian, K. and Page, D. C. (1998) A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature*, **394,** 776-80.

Jeppesen, P. and Turner, B. M. (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell*, **74,** 281-9.

Jones, M. H., Furlong, R. A., Burkin, H., Chalmers, I. J., Brown, G. M., Khwaja, O. and Affara, N. A. (1996) The Drosophila developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. *Hum Mol Genet*, **5,** 1695-701.

Jost, A. (1970) Hormonal factors in the sex differentiation of the mammalian foetus. *Philos Trans R Soc Lond B Biol Sci*, **259,** 119-30.

Jost, A., Vigier, B. and Prepin, J. (1972) Freemartins in cattle: the first steps of sexual organogenesis. *J Reprod Fertil*, **29,** 349-79.

Just, W., Rau, W., Vogel, W., Akhverdian, M., Fredga, K., Graves, J. A. and Lyapunova, E. (1995) Absence of Sry in species of the vole Ellobius. *Nat Genet*, **11,** 117-8.

Kadonaga, J. T. (1998) Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell*, **92,** 307-13.

Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R. and Kondoh, H. (1998) Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction. *Development*, **125,** 2521-32.

Kamp, C., Hirschmann, P., Voss, H., Huellen, K. and Vogt, P. H. (2000) Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. *Hum Mol Genet*, **9,** 2563-72.

Karashima, T., Sugimoto, A. and Yamamoto, M. (2000) Caenorhabditis elegans homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis. *Development*, **127,** 1069-79.

Kardong, K. V. (1995) *Vertebrates*, Wm. C. Brown Publishers, Dubuque.

Kent-First, M. (2000) The Y chromosome and its role in testis differentiation and spermatogenesis. *Semin Reprod Med*, **18,** 67-80.

Kent-First, M., Muallem, A., Shultz, J., Pryor, J., Roberts, K., Nolten, W., Meisner, L., Chandley, A., Gouchy, G., Jorgensen, L., Havighurst, T. and Grosch, J. (1999) Defining regions of the Y-chromosome responsible for male infertility and identification of a fourth AZF region (AZFd) by Y-chromosome microdeletion detection. *Mol Reprod Dev*, **53,** 27-41.

Kermouni, A., Van Roost, E., Arden, K. C., Vermeesch, J. R., Weiss, S., Godelaine, D., Flint, J., Lurquin, C., Szikora, J. P., Higgs, D. R. and et al. (1995) The IL-9 receptor gene (IL9R): genomic structure, chromosomal localization in the pseudoautosomal region of the long arm of the sex chromosomes, and identification of IL9R pseudogenes at 9qter, 10pter, 16pter, and 18pter. *Genomics*, **29,** 371-82.

Kidwell, M. G. (1993) Lateral transfer in natural populations of eukaryotes. *Annu Rev Genet*, **27,** 235-56.

Killian, J. K., Buckley, T. R., Stewart, N., Munday, B. L. and Jirtle, R. L. (2001) Marsupials and eutherians reunited: genetic evidence for the Theria hypothesis of mammalian evolution. *Mamm Genome*, **12**, 513-17.

Kingston, R. E. and Narlikar, G. J. (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev*, **13,** 2339-52.

Kirby, P. J. (2002) In *Department of Genetics and Human Evolution*La Trobe University, Melbourne.

Kirsch, J. A. W., Laponite, F. J. and Sprinder, M. S. (1997) DNA-hybridisation studies of marsupials and their implications for metatherian classification. *Aust J Zool*, **45,** 211-80.

Kirsch, S., Weiss, B., De Rosa, M., Ogata, T., Lombardi, G. and Rappold, G. A. (2000) FISH deletion mapping defines a single location for the Y chromosome stature gene, GCY. *J Med Genet*, **37,** 593-9.

Kondrashov, A. S. (1993) Classification of hypotheses on the advantage of amphimixis.

J Hered, **84,** 372-87.

Koopman, P., Gubbay, J., Collignon, J. and Lovell-Badge, R. (1989) Zfy gene expression patterns are not compatible with a primary role in mouse sex determination. *Nature*, **342,** 940-2.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1991) Male development of chromosomally female mice transgenic for Sry. *Nature*, **351,** 117-21.

Kuroda, Y., Arai, N., Arita, M., Teranishi, M., Hori, T., Harata, M. and Minzuno, S. (2001) Absence of Z-chromosome inactivation for five genes in male chickens. *Chromosome Res*, **9**, 457-68.

Kvaloy, K., Galvagni, F. and Brown, W. R. (1994) The sequence organization of the long arm pseudoautosomal region of the human sex chromosomes. *Hum Mol Genet*, **3,** 771-8.

Lahn, B. T. and Page, D. C. (1997) Functional coherence of the human Y chromosome. *Science*, **278,** 675-80.

Lahn, B. T. and Page, D. C. (1999) Four evolutionary strata on the human X chromosome. *Science*, **286,** 964-7.

Lahn, B. T. and Page, D. C. (2000) A human sex-chromosomal gene family expressed in male germ cells and encoding variably charged proteins. *Hum Mol Genet*, **9,** 311-9.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R.,

Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409,** 860-921.

Lang, J. W. and Andrews, H. V. (1994) Temperature-dependent sex determination in crocodilians. *J Exp Biol*, **270,** 28-44.

Lau, E. C., Mohandas, T. K., Shapiro, L. J., Slavkin, H. C. and Snead, M. L. (1989) Human and mouse amelogenin gene loci are on the sex chromosomes. *Genomics*, **4,** 162-8.

Lau, Y., Chou, P., Iezzoni, J., Alonzo, J. and Komuves, L. (2000) Expression of a candidate gene for the gonadoblastoma locus in gonadoblastoma and testicular seminoma. *Cytogenet Cell Genet*, **91,** 160-4.

Lau, Y. F. (1999) Gonadoblastoma, testicular and prostate cancers, and the TSPY gene. *Am J Hum Genet*, **64,** 921-7.

Lee, J. T., Davidow, L. S. and Warshawsky, D. (1999a) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet*, **21,** 400-4.

Lee, J. T. and Lu, N. (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell*, **99,** 47-57.

Lee, J. T., Lu, N. and Han, Y. (1999b) Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. *Proc Natl Acad Sci USA*, **96,** 3836-41.

Leroy, P., Alzari, P., Sassoon, D., Wolgemuth, D. and Fellous, M. (1989) The protein encoded by a murine male germ cell-specific transcript is a putative ATP-dependent RNA helicase. *Cell*, **57,** 549-59.

Li, L. and Hamer, D. H. (1995) Recombination and allelic association in the Xq/Yq homology region. *Hum Mol Genet*, **4,** 2013-6.

Li, X., Zimmerman, A., Copeland, N. G., Gilbert, G. J., Jenkins, N. A. and Yin, H. L. (1996) The mouse thymosin β4 gene: structure, promoter identification, and chromosome localization. *Genomics*, **32,** 388-94.

Lingenfelter, P. A., Delbridge, M. L., Thomas, S., Hoekstra, H. E., Mitchell, M. J., Graves, J. A. and Disteche, C. M. (2001) Expression and conservation of

processed copies of the *RBMX* gene. *Mamm Genome*, **12,** 538-45.

Little, M. and Wells, C. (1997) A clinical overview of WT1 gene mutations. *Hum Mutat*, **9,** 209-25.

Lock, L. F., Takagi, N. and Martin, G. R. (1987) Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell*, **48,** 39-46.

Lyon, M. F. (1961) Gene action in the X-chromosome of the mouse. *Nature*, **190,** 372-3.

Lyon, M. F. (1974) Evolution of X-chromosome inactivation in mammals. *Nature*, **250,** 651-3.

Ma, K., Inglis, J. D., Sharkey, A., Bickmore, W. A., Hill, R. E., Prosser, E. J., Speed, R. M., Thomson, E. J., Jobling, M., Taylor, K. and et al. (1993) A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis. *Cell*, **75,** 1287-95.

Makova, K. D. and Li, W. H. (2002) Strong male-driven evolution of DNA sequences in humans and apes. *Nature*, **416,** 624-6.

Margulis, L. and Sagan, D. (1986) *Origins of Sex*, Halliday Lithograph, West Hanover, Massachusetts.

Marin, I., Siegal, M. L. and Baker, B. S. (2000) The evolution of dosage-compensation mechanisms. *Bioessays*, **22,** 1106-14.

Mazeyrat, S., Saut, N., Sargent, C. A., Grimmond, S., Longepied, G., Ehrmann, I. E., Ellis, P. S., Greenfield, A., Affara, N. A. and Mitchell, M. J. (1998) The mouse Y chromosome interval necessary for spermatogonial proliferation is gene dense with syntenic homology to the human AZFa region. *Hum Mol Genet*, **7,** 1713-24.

Mazeyrat, S., Saut, N., Grigoriev, V., Mahadevaiah, S. K., Ojarikre, O. A., Rattigan, A., Bishop, C., Eicher, E. M., Mitchell, M. J. and Burgoyne, P. S. (2001) A Y-encoded subunit of the translation initiation factor Eif2 is essential for mouse spermatogenesis. *Nat Genet*, **29,** 49-53.

McElreavey, K., Vilain, E., Abbas, N., Herskowitz, I. and Fellous, M. (1993) A regulatory cascade hypothesis for mammalian sex determination: SRY represses a negative regulator of male development. *Proc Natl Acad Sci USA*, **90,** 3368-72.

McLaren, A., Simpson, E., Epplen, J. T., Studer, R., Koopman, P., Evans, E. P. and

Burgoyne, P. S. (1988) Location of the genes controlling H-Y antigen expression and testis determination on the mouse Y chromosome. *Proc Natl Acad Sci USA*, **85,** 6442-5.

McQueen, H. A., Fantes, J., Cross, S. H., Clark, V. H., Archibald, A. L. and Bird, A. P. (1996) CpG islands of chicken are concentrated on microchromosomes. *Nat Genet*, **12,** 321-4.

McQueen, H. A., McBride, D., Miele, G., Bird, A. P. and Clinton, M. (2001) Dosage compensation in birds. *Curr Biol*, **11,** 253-7.

Meadows, L., Wang, W., den Haan, J. M., Blokland, E., Reinhardus, C., Drijfhout, J. W., Shabanowitz, J., Pierce, R., Agulnik, A. I., Bishop, C. E., Hunt, D. F., Goulmy, E. and Engelhard, V. H. (1997) The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity*, **6,** 273-81.

Mermoud, J. E., Costanzi, C., Pehrson, J. R. and Brockdorff, N. (1999) Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J Cell Biol*, **147,** 1399-408.

Meroni, G., Franco, B., Archidiacono, N., Messali, S., Andolfi, G., Rocchi, M. and Ballabio, A. (1996) Characterization of a cluster of sulfatase genes on Xp22.3 suggests gene duplications in an ancestral pseudoautosomal region. *Hum Mol Genet*, **5**, 423-31.

Milatovich, A., Kitamura, T., Miyajima, A. and Francke, U. (1993) Gene for the alpha-subunit of the human interleukin-3 receptor (IL3RA) localized to the X-Y pseudoautosomal region. *Am J Hum Genet*, **53,** 1146-53.

Mitchell, M. J., Woods, D. R., Wilcox, S. A., Graves, J. A. and Bishop, C. E. (1992) Marsupial Y chromosome encodes a homologue of the mouse Y-linked candidate spermatogenesis gene Ube1y. *Nature*, **359,** 528-31.

Mitchell, M. J., Wilcox, S. A., Watson, J. M., Lerner, J. L., Woods, D. R., Scheffler, J., Hearn, J. P., Bishop, C. E. and Graves, J. A. M. (1998) The origin and loss of the ubiquitin activating enzyme gene on the mammalian Y chromosome. *Hum Mol Genet*, **7,** 429-34.

Mohandas, T. K., Speed, R. M., Passage, M. B., Yen, P. H., Chandley, A. C. and Shapiro, L. J. (1992) Role of the pseudoautosomal region in sex-chromosome pairing during male meiosis: meiotic studies in a man with a deletion of distal

Xp. *Am J Hum Genet,* **51,** 526-33.

Moore, S. S., Byrne, K., Johnson, S. E., Kata, S. and Womack, J. E. (2001) Physical mapping of CSF2RA, ANT3 and STS on the pseudoautosomal region of bovine chromosome X. *Anim Genet,* **32,** 102-4.

Muller, H. J. (1964) The relation of recombination to mutational advance. *Mutation Research,* **1,** 2-9.

Mumm, S., Molini, B., Terrell, J., Srivastava, A. and Schlessinger, D. (1997) Evolutionary features of the 4-Mb Xq21.3 XY homology region revealed by a map at 60-kb resolution. *Genome Res,* **7,** 307-14.

Murphy, W. J., Sun, S., Chen, Z. Q., Pecon-Slattery, J. and O'Brien, S. J. (1999) Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. *Genome Res,* **9,** 1223-30.

Murtagh, C. E. (1977) A unique cytogenetic system in monotremes. *Chromosoma,* **65,** 37-57.

Nachtigal, M. W., Hirokawa, Y., Enyeart-VanHouten, D. L., Flanagan, J. N., Hammer, G. D. and Ingraham, H. A. (1998) Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. *Cell,* **93,** 445-54.

Nanda, I., Shan, Z., Schartl, M., Burt, D. W., Koehler, M., Nothwang, H., Grutzner, F., Paton, I. R., Windsor, D., Dunn, I., Engel, W., Staeheli, P., Mizuno, S., Haaf, T. and Schmid, M. (1999) 300 million years of conserved synteny between chicken Z and human chromosome 9. *Nat Genet,* **21,** 258-9.

Nanda, I., Zend-Ajusch, E., Shan, Z., Grutzner, F., Schartl, M., Burt, D. W., Koehler, M., Fowler, V. M., Goodwin, G., Schneider, W. J., Mizuno, S., Dechant, G., Haaf, T. and Schmid, M. (2000) Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene DMRT1: a comparative (re)view on avian sex determination. *Cytogenet Cell Genet,* **89,** 67-78.

Nelson, G. A., Lew, K. K. and Ward, S. (1978) Intersex, a temperature-sensitive mutant of the nematode, *Caenorhabditis elegans. Developmental Biology,* **66,** 386-409.

Nygard, O. and Nika, H. (1982) Identification by RNA-protein cross-linking of ribosomal proteins located at the interface between the small and the large subunits of mammalian ribosomes. *Embo J,* **1,** 357-62.

O, W. S., Short, R. V., Renfree, M. B. and Shaw, G. (1988) Primary genetic control of

somatic sexual differentiation in a mammal. *Science,* **331,** 716-17.

Ogata, T., Tomita, K., Hida, A., Matsuo, N., Nakahori, Y. and Nakagome, Y. (1995) Chromosomal localisation of a Y specific growth gene(s). *J Med Genet,* **32,** 572-5.

Ogawa, A., Murata, K. and Mizuno, S. (1998) The location of Z- and W-linked marker genes and sequence on the homomorphic sex chromosomes of the ostrich and the emu. *Proc Natl Acad Sci USA,* **95,** 4415-8.

Ohno, S. (1967) *Sex Chromosomes and Sex-linked Genes,* Springer-Verlag, New York.

Ono, T., Hirano, S., Yonezawa, S., Aono, S., Osaki, M., Masaki, S., Yamashita, S., Tsukasaki, T., Oohira, A., Suzuki, S. T. and Sonta, S. (2000) Comparative mapping of seven genes in mouse, rat and Chinese hamster chromosomes by fluorescence in situ hybridization. *Cytogenet Cell Genet,* **89,** 209-13.

Otto, S. P. and Lenormand, T. (2002) Resolving the paradox of sex and recombination. *Nat Rev Genet,* **3,** 252-61.

Page, D. C. (1987) Hypothesis: a Y-chromosomal gene causes gonadoblastoma in dysgenetic gonads. *Development,* **101,** 151-5.

Page, D. C., Disteche, C. M., Simpson, E. M., de la Chapelle, A., Andersson, M., Alitalo, T., Brown, L. G., Green, P. and Akots, G. (1990) Chromosomal localization of ZFX--a human gene that escapes X inactivation--and its murine homologs. *Genomics,* **7,** 37-46.

Page, D. C., Harper, M. E., Love, J. and Botstein, D. (1984) Occurrence of a transposition from the X-chromosome long arm to the Y- chromosome short arm during human evolution. *Nature,* **311,** 119-23.

Page, D. C., Mosher, R., Simpson, E. M., Fisher, E. M., Mardon, G., Pollack, J., McGillivray, B., de la Chapelle, A. and Brown, L. G. (1987) The sex-determining region of the human Y chromosome encodes a finger protein. *Cell,* **51,** 1091-104.

Page, J., Berrios, S., Rufas, J. S., Parra, M. T., Sija, J. A., Heyting, C. and Fernandez-Donosoo, R. (2002) The meiotic pairing of X and Y chromossomes in the marsupial species *Thylamys elegans* is maintained by a dense plate developed from their axial elements.

Pallavicini, A., Zimbello, R., Tiso, N., Muraro, T., Rampoldi, L., Bortoluzzi, S., Valle, G., Lanfranchi, G. and Danieli, G. A. (1997) The preliminary transcript map of a

human skeletal muscle. *Hum Mol Genet*, **6,** 1445-50.

Palmer, M. S., Sinclair, A. H., Berta, P., Ellis, N. A., Goodfellow, P. N., Abbas, N. E. and Fellous, M. (1989) Genetic evidence that ZFY is not the testis-determining factor. *Nature*, **342,** 937-9.

Palmer, S., Perry, J., Kipling, D. and Ashworth, A. (1997) A gene spans the pseudoautosomal boundary in mice. *Proc Natl Acad Sci USA*, **94,** 12030-5.

Pask, A. and Renfree, M. B. (2001) Sex determining genes and sexual differentiation in a marsupial. *J Exp Zool*, **290,** 586-96.

Pask, A., Renfree, M. B. and Marshall Graves, J. A. (2000) The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. *Proc Natl Acad Sci USA*, **97,** 13198-202.

Pask, A. J. (1999) In *School of Biochemistry and Genetics*La Trobe University, Melbourne.

Perry, J., Palmer, S., Gabriel, A. and Ashworth, A. (2001) A short pseudoautosomal region in laboratory mice. *Genome Res*, **11,** 1826-32.

Persico, M. G., Viglietto, G., Martini, G., Toniolo, D., Paonessa, G., Moscatelli, C., Dono, R., Vulliamy, T., Luzzatto, L. and D'Urso, M. (1986) Isolation of human glucose-6-phoshate dehydrogenase (G6PD) cDNA clones: primary structure of the protein and unusual 5' non-coding region. *Nucleic Acids Res*, **14,** 2511-22.

Pfeifer, G. P., Tanguay, R. L., Steigerwald, S. D. and Riggs, A. D. (1990) In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev*, **4,** 1277-87.

Picketts, D. J., Tastan, A. O., Higgs, D. R. and Gibbons, R. J. (1998) Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. *Mamm Genome*, **9,** 400-3.

Pierce, R. A., Field, E. D., den Haan, J. M., Caldwell, J. A., White, F. M., Marto, J. A., Wang, W., Frost, L. M., Blokland, E., Reinhardus, C., Shabanowitz, J., Hunt, D. F., Goulmy, E. and Engelhard, V. H. (1999) Cutting edge: the HLA-A*0101-restricted HY minor histocompatibility antigen originates from DFFRY and contains a cysteinylated cysteine residue as identified by a novel mass spectrometric technique. *J Immunol*, **163,** 6360-4.

Raines, M. A., Liu, L., Quan, S. G., Joe, V., DiPersio, J. F. and Golde, D. W. (1991) Identification and molecular cloning of a soluble human granulocyte-macrophage colony-stimulating factor receptor. *Proc Natl Acad Sci U S A*, **88,** 8203-7.

Ratti, A., Stuppia, L., Gatta, V., Fogh, I., Calabrese, G., Pizzuti, A. and Palka, G. (2000) Characterization of a new TSPY gene family member in Yq (TSPYq1). *Cytogenet Cell Genet*, **88,** 159-62.

Raymond, C. S., Murphy, M. W., O'Sullivan, M. G., Bardwell, V. J. and Zarkower, D. (2000) Dmrt1, a gene related to worm and fly sex regulators, is required for mammalian testis differentiation.*Genes and Development*, **14**, 2587-95.

Reddy, J. C. and Licht, J. D. (1996) The WT1 Wilms' tumor suppressor gene: how much do we really know? *Biochim Biophys Acta*, **1287,** 1-28.

Redfield, R. J. (2001) Do bacteria have sex? *Nat Rev Genet*, **2,** 634-9.

Reijo, R., Lee, T. Y., Salo, P., Alagappan, R., Brown, L. G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., Hovatta, O. and et al. (1995) Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet*, **10,** 383-93.

Rice, W. R. (1987) Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics*, **116,** 161-7.

Ried, K., Rao, E., Schiebel, K. and Rappold, G. A. (1998) Gene duplications as a recurrent theme in the evolution of the human pseudoautosomal region 1: isolation of the gene ASMTL. *Hum Mol Genet*, **7,** 1771-8.

Rodriguez, I. R., Mazuruk, K., Schoen, T. J. and Chader, G. J. (1994) Structural analysis of the human hydroxyindole-O-methyltransferase gene. Presence of two distinct promoters. *J Biol Chem*, **269,** 31969-77.

Rota, A., Ballarin, C., Vigier, B., Cozzi, B. and Rey, R. (2002) Age dependent changes in plasma anti-Mullerian hormone concentrations in the bovine male, female, and freemartin from birth to puberty: relationship between testosterone production and influence on sex differentiation. *Gen Comp Endocrinol*, **15,** 39-44.

Rousseaux-Prevost, R., Rigot, J. M., Delobel, B., Lesur, P., Collier, F., Croquette, M. F., Gauthier, A., Mazeman, E. and Rousseaux, J. (1996) Molecular mapping of a Yq deletion in a patient with normal stature. *Hum Genet*, **98,** 505-7.

Saifi, G. M. and Chandra, H. S. (1999) An apparent excess of sex- and reproduction-

related genes on the human X chromosome. *Proc R Soc Lond B Biol Sci,* **266,** 203-9.

Salido, E. C., Li, X. M., Yen, P. H., Martin, N., Mohandas, T. K. and Shapiro, L. J. (1996) Cloning and expression of the mouse pseudoautosomal steroid sulphatase gene (Sts). *Nat Genet,* **13,** 83-6.

Salido, E. C., Yen, P. H., Koprivnikar, K., Yu, L. C. and Shapiro, L. J. (1992) The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *Am J Hum Genet,* **50,** 303-16.

Salo, P., Kaariainen, H., Page, D. C. and de la Chapelle, A. (1995a) Deletion mapping of stature determinants on the long arm of the Y chromosome. *Hum Genet,* **95,** 283-6.

Salo, P., Kaariainen, H., Petrovic, V., Peltomaki, P., Page, D. C. and de la Chapelle, A. (1995b) Molecular mapping of the putative gonadoblastoma locus on the Y chromosome. *Genes Chromosomes Cancer,* **14,** 210-4.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual,* Cold Spring Harbor Laboratory Press, New York.

Sargent, C. A., Briggs, H., Chalmers, I. J., Lambson, B., Walker, E. and Affara, N. A. (1996) The sequence organization of Yp/proximal Xq homologous regions of the human sex chromosomes is highly conserved. *Genomics,* **32,** 200-9.

Sargent, C. A., Affara, N., Bentley, E., Pelmear, A., Bailey, D. M. D., Davey, P., Dow, D., Leversha, M., Aplin, H., Besley, G. T. N. and Ferguson-Smith, M. A. (1993) Isolation and characterisation of human glycerol kinase transcripts: indentification of loci at Xp22.1-p21.2 and Xq22-q23. *Cytogenet Cell Genet,* **64,** 186-87.

Saxena, R., Brown, L. G., Hawkins, T., Alagappan, R. K., Skaletsky, H., Reeve, M. P., Reijo, R., Rozen, S., Dinulos, M. B., Disteche, C. M. and Page, D. C. (1996) The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nat Genet,* **14,** 292-9.

Schiebel, K., Meder, J., Rump, A., Rosenthal, A., Winkelmann, M., Fischer, C., Bonk, T., Humeny, A. and Rappold, G. (2000) Elevated DNA sequence diversity in the genomic region of the phosphatase PPP2R3L gene in the human pseudoautosomal region. *Cytogenet Cell Genet,* **91,** 224-30.

Schiebel, K., Mertz, A., Winkelmann, M., Glaser, B., Schempp, W. and Rappold, G. (1997) FISH localization of the human Y-homolog of protein kinase PRKX (PRKY) to Yp11.2 and two pseudogenes to 15q26 and Xq12-->q13. *Cytogenet Cell Genet,* **76,** 49-52.

Schmid, M., Nanda, I., Guttenbach, M., Steinlein, C., Hoehn, M., Schartl, M., Haaf, T., Weigend, S., Fries, R., Buerstedde, J. M., Wimmers, K., Burt, D. W., Smith, J., A'Hara, S., Law, A., Griffin, D. K., Bumstead, N., Kaufman, J., Thomson, P. A., Burke, T., Groenen, M. A., Crooijmans, R. P., Vignal, A., Fillon, V., Morisson, M., Pitel, F., Tixier-Boichard, M., Ladjali-Mohammedi, K., Hillel, J., Maki-Tanila, A., Cheng, H. H., Delany, M. E., Burnside, J. and Mizuno, S. (2000) First report on chicken genes and chromosomes 2000. *Cytogenet Cell Genet,* **90,** 169-218.

Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Mardon, G., Luoh, S. W. and Page, D. C. (1989) Putative transcription activator with alternative isoforms encoded by human ZFX gene. *Nature,* **342,** 708-11.

Schnieders, F., Dork, T., Arnemann, J., Vogel, T., Werner, M. and Schmidtke, J. (1996) Testis-specific protein, Y-encoded (TSPY) expression in testicular tissues. *Hum Mol Genet,* **5,** 1801-7.

Schwartz, A., Chan, D. C., Brown, L. G., Alagappan, R., Pettay, D., Disteche, C., McGillivray, B., de la Chapelle, A. and Page, D. C. (1998) Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum Mol Genet,* **7,** 1-11.

Scott, D. M., Ehrmann, I. E., Ellis, P. S., Bishop, C. E., Agulnik, A. I., Simpson, E. and Mitchell, M. J. (1995) Identification of a mouse male-specific transplantation antigen, H-Y. *Nature,* **376,** 695-8.

Scully, R. E. (1970) Gonadoblastoma. A review of 74 cases. *Cancer,* **25,** 1340-56.

Sharman, G. B. (1971) Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature,* **230,** 231-2.

Sharman, G. B., Hughes, R. L. and Cooper, D. W. (1990) The chromosomal basis of sex differentiation in marsupials. In *Mammal from Pouches and Eggs: Genetics and Evolution of Marsupials and Monotremes*, (Eds, Graves, J. A. M., Hope, R. M. and Cooper, D. W.) Commonwealth Scientific and Industrial Research,

Melbourne.

Sharp, P. (1982) Sex chromosome pairing during male meiosis in marsupials. *Chromosoma,* **86,** 27-47.

Shetty, S., Griffin, D. K. and Graves, J. A. (1999) Comparative painting reveals strong chromosome homology over 80 million years of bird evolution. *Chromosome Res,* **7,** 289-95.

Shetty, S. (2001) In *Evolution of Sex Chromosomes in Birds and Reptiles.* Department of Genetics and Human Evolution, La Trobe University, Melbourne.

Shetty, S., Kirby, P. J., Zarkower, D. and Graves, J. A. M. (2002) DMRT1 in a ratite bird: evidence for a role in sex determination and identification of a putative regulatory element. *Cytogenet Gen Res*, In press.

Shimmin, L. C., Chang, B. H. and Li, W. H. (1993) Male-driven evolution of DNA sequences. *Nature,* **362,** 745-7.

Shimmin, L. C., Chang, B. H. and Li, W. H. (1994) Contrasting rates of nucleotide substitution in the X-linked and Y- linked zinc finger genes. *J Mol Evol,* **39,** 569-78.

Simpson, E., Chandler, P., Goulmy, E., Disteche, C. M., Ferguson-Smith, M. A. and Page, D. C. (1987) Separation of the genetic loci for the H-Y antigen and for testis determination on human Y chromosome. *Nature,* **326,** 876-8.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R. and Goodfellow, P. N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature,* **346,** 240-4.

Sinclair, A. H., Foster, J. W., Spencer, J. A., Page, D. C., Palmer, M., Goodfellow, P. N. and Graves, J. A. (1988) Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. *Nature,* **336,** 780-3.

Slee, R., Grimes, B., Speed, R. M., Taggart, M., Maguire, S. M., Ross, A., McGill, N. I., Saunders, P. T. and Cooke, H. J. (1999) A human DAZ transgene confers partial rescue of the mouse Dazl null phenotype. *Proc Natl Acad Sci USA,* **96,** 8040-5.

Slim, R., Levilliers, J., Ludecke, H. J., Claussen, U., Nguyen, V. C., Gough, N. M., Horsthemke, B. and Petit, C. (1993) A human pseudoautosomal gene encodes the ANT3 ADP/ATP translocase and escapes X-inactivation. *Genomics*, **16,** 26-

33.

Smith, M. J., Goodfellow, P. J. and Goodfellow, P. N. (1993) The genomic organisation of the human pseudoautosomal gene MIC2 and the detection of a related locus. *Hum Mol Genet,* **2,** 417-22.

Smith, J., Bruley, C. K., Paton, I. R., Dunn, I., Jones, C. T., Windsor, D., Morrice, D. R., Law, A. S., Masabanda, J., Sazanov, A., Waddington, D., Fries, R. and Burt, D. W. (2000) Differences in gene density on chicken macrochromosomes and microchromosomes. *Anim Genet,* **31,** 96-103.

Smith, J. and Burt, D. W. (1998) Parameters of the chicken genome (Gallus gallus). *Anim Genet,* **29,** 290-4.

Soulard, M., Della Valle, V., Siomi, M. C., Pinol-Roma, S., Codogno, P., Bauvy, C., Bellini, M., Lacroix, J. C., Monod, G., Dreyfuss, G. and et al. (1993) hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. *Nucleic Acids Res,* **21,** 4210-7.

Spencer, J. A. (1991) In *Mapping Human X-Linked Genes in Marsupials and Monotremes.* School of Genetics and Human Variation, La Trobe University, Melbourne.

Spencer, J. A., Sinclair, A. H., Watson, J. M. and Graves, J. A. (1991) Genes on the short arm of the human X chromosome are not shared with the marsupial X. *Genomics,* **11,** 339-45.

Stern, C. (1957) The problen of complete Y-linkage in man. *Am J Hum Genet,* **9,** 147-66.

Stouffs, K., Lissens, W., Van Landuyt, L., Tournaye, H., van Steirteghem, A. and Liebaers, I. (2001) Characterization of the genomic orgaization, localization and expression of four *PRY* genes (*PRY1*, *PRY2*, *PRY3* and *PRY4*). *Mol Hum Reprod,* **7,** 603-10.

Sun, C., Skaletsky, H., Birren, B., Devon, K., Tang, Z., Silber, S., Oates, R. and Page, D. C. (1999) An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat Genet,* **23,** 429-32.

Sun, C., Skaletsky, H., Rozen, S., Gromoll, J., Nieschlag, E., Oates, R. and Page, D. C. (2000) Deletion of azoospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses. *Hum Mol Genet,* **9,** 2291-6.

Sutcliffe, M. J. and Burgoyne, P. S. (1989) Analysis of the testes of H-Y negative XOSxrb mice suggests that the spermatogenesis gene (Spy) acts during the differentiation of the A spermatogonia. *Development*, **107,** 373-80.

Suzuki, S. T. (1996) Structural and functional diversity of cadherin superfamily: are new members of cadherin superfamily involved in signal transduction pathway? *J Cell Biochem*, **61,** 531-42.

Swain, A. and Lovell-Badge, R. (1999) Mammalian sex determination: a molecular drama. *genes Dev*, **13,** 755-67.

Swain, A., Narvaez, V., Burgoyne, P., Camerino, G. and Lovell-Badge, R. (1998) Dax1 antagonizes Sry action in mammalian sex determination. *Nature*, **391,** 761-67.

Takeuchi, T., Furihata, M., Heng, H. H., Sonobe, H. and Ohtsuki, Y. (1998) Chromosomal mapping and expression of the human B120 gene. *Gene*, **213,** 189-93.

Thorne, M. and Sheldon, B. L. (1993) *Triploid intersex and chimeric chickens: useful models for stuies of avian sex determination*, Harwood Academic, Chur, Switzerland.

Tiepolo, L. and Zuffardi, O. (1976) Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Hum Genet*, **34,** 119-24.

Toder, R., Glaser, B., Schiebel, K., Wilcox, S. A., Rappold, G., Graves, J. A. and Schempp, W. (1997a) Genes located in and near the human pseudoautosomal region are located in the X-Y pairing region in dog and sheep. *Chromosome Res*, **5,** 301-6.

Toder, R. and Graves, J. A. (1998) CSF2RA, ANT3, and STS are autosomal in marsupials: implications for the origin of the pseudoautosomal region of mammalian sex chromosomes. *Mamm Genome*, **9,** 373-6.

Toder, R., Wakefield, M. J. and Graves, J. A. (2000) The minimal mammalian Y chromosome - the marsupial Y as a model system. *Cytogenet Cell Genet*, **91,** 285-92.

Toder, R., Wienberg, J., Voullaire, L., O'Brien, P. C., Maccarone, P. and Graves, J. A. (1997b) Shared DNA sequences between the X and Y chromosomes in the tammar wallaby - evidence for independent additions to eutherian and marsupial sex chromosomes. *Chromosoma*, **106,** 94-8.

Tribioli, C., Tamanini, F., Patrosso, C., Milanesi, L., Villa, A., Pergolizzi, R., Maestrini, E., Rivella, S., Bione, S., Mancini, M. and et al. (1992) Methylation and sequence analysis around EagI sites: identification of 28 new CpG islands in XQ24-XQ28. *Nucleic Acids Res,* **20,** 727-33.

Tsuchiya, K., Reijo, R., Page, D. C. and Disteche, C. M. (1995) Gonadoblastoma: molecular definition of the susceptibility region on the Y chromosome. *Am J Hum Genet,* **57,** 1400-7.

Vacca, M., Matarazzo, M. R., Jones, J., Spalluto, C., Archidiacono, N., Ma, P., Rocchi, M., D'Urso, M., Chen, E. Y., D'Esposito, M. and Mumm, S. (1999) Evolution of the X-specific block embedded in the human Xq21.3/Yp11.1 homology region. *Genomics,* **62,** 293-6.

Vaiman, D. and Pailhoux, E. (2000) Mammalian sex reversal and intersexuality: deciphering the sex- determination cascade. *Trends Genet,* **16,** 488-94.

Veitia, R. A., Salas-Cortes, L., Ottolenghi, C., Pailhoux, E., Cotinot, C. and Fellous, M. (2001) Testis determination in mammals: more questions than answers. *Mol Cell Endocrinol,* **179,** 3-16.

Vermeesch, J. R., Petit, P., Kermouni, A., Renauld, J. C., Van Den Berghe, H. and Marynen, P. (1997) The IL-9 receptor gene, located in the Xq/Yq pseudoautosomal region, has an autosomal origin, escapes X inactivation and is expressed from the Y. *Hum Mol Genet,* **6,** 1-8.

Viets, B. E., Ewert, M. A., Talent, L. G. and Nelson, C. E. (1994) Sex determining mechanisms in squamate rptiles. *J Exp Zool,* **270,** 45-56.

Vogt, M. H., de Paus, R. A., Voogt, P. J., Willemze, R. and Falkenburg, J. H. (2000) DFFRY codes for a new human male-specific minor transplantation antigen involved in bone marrow graft rejection. *Blood,* **95,** 1100-5.

Vogt, P. H., Affara, N., Davey, P., Hammer, M., Jobling, M. A., Lau, Y. F., Mitchell, M., Schempp, W., Tyler-Smith, C., Williams, G., Yen, P. and Rappold, G. A. (1997) Report of the Third International Workshop on Y Chromosome Mapping 1997. Heidelberg, Germany, April 13-16, 1997. *Cytogenet Cell Genet,* **79,** 1-20.

Vogt, P. H., Edelmann, A., Kirsch, S., Henegariu, O., Hirschmann, P., Kiesewetter, F., Kohn, F. M., Schill, W. B., Farah, S., Ramos, C., Hartmann, M., Hartschuh, W., Meschede, D., Behre, H. M., Castel, A., Nieschlag, E., Weidner, W., Grone, H. J., Jung, A., Engel, W. and Haidl, G. (1996) Human Y chromosome

azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet*, **5,** 933-43.

Vollrath, P. H., Foote, S., Hilton, A., Brown, L. G., Romero, P. G., Bogan, J. S. and Page, D. C. (1992) The human Y chromosome: A 43-interval map based on naturally occuring deletions. *Science*, **258**, 52-9.

Wachtel, S. S., Ono, S., Koo, G. C. and Boyse, E. A. (1975) Possible role for H--Y antigen in the primary determination of sex. *Nature*, **257,** 235-6.

Wakefield, M. J. and Graves, J. A. (1996) Comparative maps of vertebrates. *Mamm Genome*, **7,** 715-6.

Wakefield, M. J., Keohane, A. M., Turner, B. M. and Graves, J. A. (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. *Proc Natl Acad Sci USA*, **94,** 9665-8.

Wang, P. J., McCarrey, J. R., Yang, F. and Page, D. C. (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet*, **27,** 422-6.

Wang, W., Meadows, L. R., den Haan, J. M., Sherman, N. E., Chen, Y., Blokland, E., Shabanowitz, J., Agulnik, A. I., Hendrickson, R. C., Bishop, C. E. and et al. (1995) Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein. *Science*, **269,** 1588-90.

Waters, P. D., Duffy, B., Frost, C. J., Delbridge, M. L. and Graves, J. A. (2001) The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80-130 million years ago. *Cytogenet Cell Genet*, **92,** 74-9.

Watson, J. M., Frost, C., Spencer, J. A. and Graves, J. A. (1993) Sequences homologous to the human X- and Y-borne zinc finger protein genes (ZFX/Y) are autosomal in monotreme mammals. *Genomics*, **15,** 317-22.

Watson, J. M., Riggs, A. and Graves, J. A. (1992a) Gene mapping studies confirm the homology between the platypus X and echidna X1 chromosomes and identify a conserved ancestral monotreme X chromosome. *Chromosoma*, **101,** 596-601.

Watson, J. M., Spencer, J. A., Graves, J. A., Snead, M. L. and Lau, E. C. (1992b) Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. *Genomics*, **14,** 785-9.

Watson, J. M., Spencer, J. A., Riggs, A. D. and Graves, J. A. (1990) The X chromosome of monotremes shares a highly conserved region with the eutherian and marsupial X chromosomes despite the absence of X chromosome

inactivation. *Proc Natl Acad Sci USA*, **87,** 7125-9.

Watson, J. M., Spencer, J. A., Riggs, A. D. and Graves, J. A. (1991) Sex chromosome evolution: platypus gene mapping suggests that part of the human X chromosome was originally autosomal. *Proc Natl Acad Sci USA*, **88,** 11256-60.

Watson, J. M. and Graves, J. A. (1988) Gene mapping in marsupials and monotremes, V. Synteny between hypoxanthine phosphoribosyltransferase and phosphoglycerate kinase in the platypus. *Aust J Biol Sci*, **41,** 231-7.

Weiss, L. (1983) *Histology: cell and tissue biology*, Elsevier Biomedical, New York.

Weller, P. A., Critcher, R., Goodfellow, P. N., German, J. and Ellis, N. A. (1995) The human Y chromosome homologue of XG: transcription of a naturally truncated gene. *Hum Mol Genet*, **4,** 859-68.

Welshons, W. J. and Russell, L. B. (1959) The Y-chromosome as the bearer of the male determining factors in the mouse. *Proc Natl Acad Sci USA*, **45,** 560-66.

Western, P. S. (1999) In *School of Biochemistry and Genetics*La Trobe University, Melbourne.

Wilcox, S. A., Watson, J. M., Spencer, J. A. and Graves, J. A. (1996) Comparative mapping identifies the fusion point of an ancient mammalian X-autosomal rearrangement. *Genomics*, **35,** 66-70.

Willard, H. F. and Latt, S. A. (1976) Analysis of deoxyribonucleic acid replication in human X chromosomes by fluorescence microscopy. *Am J Hum Genet*, **28,** 213-27.

Wolf, Y. I., Kondrashov, A. S. and Koonin, E. V. (2000) Interkingdom gene fusions. *Genome Biol*, **1,** research0013.1-0013.13.

Wolffe, A. P. (1998) Packaging principle: how DNA methylation and histone acetylation control the transcriptional activity of chromatin. *J Exp Zool*, **282,** 239-44.

Wrigley, J. M. and Graves, J. A. (1988a) Karyotypic conservation in the mammalian order monotremata (subclass Prototheria). *Chromosoma*, **96,** 231-47.

Wrigley, J. M. and Graves, J. A. (1988b) Sex chromosome homology and incomplete, tissue-specific X-inactivation suggest that monotremes represent an intermediate stage of mammalian sex chromosome evolution. *J Hered*, **79,** 115-8.

Wunderle, V. M., Critcher, R., Hastie, N., Goodfellow, P. N. and Schedl, A. (1998) Deletion of long-range regulatory elements upstream of SOX9 causes

campomelic dysplasia. *Proc Natl Acad Sci USA*, **95,** 10649-54.

Xu, J., Burgoyne, P. S. and Arnold, A. P. (2002) Sex differences in sex chromosome gene expression in mouse brain. *Hum Mol Genet*, **11,** 1409-19.

Yagi, T. and Takeichi, M. (2000) Cadherin superfamily genes: functions, genomic organisation, and neurologic diversity. *Genes Dev*, **14,** 1169-80.

Yen, P. H., Marsh, B., Allen, E., Tsai, S. P., Ellison, J., Connolly, L., Neiswanger, K. and Shapiro, L. J. (1988) The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution. *Cell*, **55**, 1123-35.

Yen, P. H., Chai, N. N. and Salido, E. C. (1997) The human DAZ genes, a putative male infertility factor on the Y chromosome, are highly polymorphic in the DAZ repeat regions. *Mamm Genome*, **8,** 756-9.

Yoshida, K. and Sugano, S. (1999) Identification of a novel protocadherin gene (PCDH11) on the human XY homology region in Xq21.3. *Genomics*, **62,** 540-3.

Zanaria, E., Bardoni, B., Dabovic, B., Calvari, V., Fraccaro, M., Zuffardi, O. and Camerino, G. (1995) Xp duplications and sex reversal. *Philos Trans R Soc Lond B Biol Sci*, **350,** 291-6.

Zechner, U., Wilda, M., Kehrer-Sawatzki, H., Vogel, W., Fundele, R. and Hameister, H. (2001) A high density of X-linked genes for general cognitive ability: a run-away process shaping human evolution? *Trends Genet*, **17,** 697-701.

Zhai, L., Mu, J., Zong, H., DePaoli-Roach, A. A. and Roach, P. J. (2000) Structure and chromosomal localization of the human glycogenin-2 gene GYG2. *Gene*, **242**, 229-35.

Zinn, A. R., Alagappan, R. K., Brown, L. G., Wool, I. and Page, D. C. (1994) Structure and function of ribosomal protein S4 genes on the human and mouse sex chromosomes. *Mol Cell Biol*, **14,** 2485-92.

Appendix I: Primer list

Primer name	Sequence 5'-3'
T3	AATTAACCCTCACTCCCGGG
T7	CGGGATATCACTCAGCATAATG
SP6	ATTTAGGTGACACTATAGAAA
BPY1#4t3(345)	CCTTCGTCAGTTTTGGGAGTG
BPY1#4t7(470)	GCCCTTCTGCTTCTGTCTTTG
PCDHX-2020f	GAATAACTCTCCTGGCATCC
PCDHX-2735r	TCTCTTGTGTTTCCTCCTAC
RPS4X-142f	CTCCGAAACCGCCTCAAGTA
RPS4X-585r	GGTGTTTCTCCCTGTTGGTG
RPS4X-12f	GCCCAAGAAGCACCTGAAGC
RPS4X-700r	CGGGGCAGAGAAATCCAAGC
PRY-418f	ATGGGAAGGTTGGCTGTATCT
PRY-1138r	AAGTGTTGTTGGAGGTTGTGA
XKRY-695f	TATTCCCTCTTATCAGTTGTG
XKRY-1481r	ACTCAGCCTCTTTACTCTTTG
ATRX(f)	CAATAATGGATGAAAACAGCC
ATRX(r)	TGCCTGCTTCAAAAATCTTAC
SOX3(f)	ATCAACGCSTTYATGGTDTGGTCCCG
SOX3(r)	CGGCRCYCTGGTAGTGYTGGTGYAC
SOX3-HMG(f)	CCCATGAACGGGTTTCATGGTGT
SOX3-HMG(r)	GCAGGCGAGTACTTGTCCTTCTT
PCDHX4(508)	CTGTTGATGGAGGTGAACC
PCDHX4(566)	AGACAGGGCGACTAAAGTGG
TB4X286-t3	TGCTAACCAAGGATAAAAAG
TB4X354-t3	TTACCCAACATAGAAAAACC
TB4X392-t3	ACCCAAAGGCATCAGTAAGG
TB4X81-t7	GGAGAGGGGAGAAGGAGGAG
TB4X397-t7	GAGGCTGTGAATGGTTTTGG
TB4X102-t7	ATTGTTGAGCCAGAGAAAGC

Appendix II: Publications

Cytogenet Cell Genet 92:74–79 (2001)

Cytogenetics and Cell Genetics

The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80–130 million years ago

P.D. Waters, B. Duffy, C.J. Frost, M.L. Delbridge, and J.A.M. Graves

Department of Genetics, La Trobe University, Melbourne, Victoria (Australia)

Abstract. Mapping of human X-borne genes in distantly related mammals has defined a conserved region shared by the X chromosome in all three extant mammalian groups, plus a region that was recently added to the eutherian X but is still autosomal in marsupials and monotremes. Using comparative mapping of human Y-borne genes, we now directly show that the eutherian Y is also composed of a conserved and an added region which contains most of the ubiquitously expressed Y-borne genes. Little of the ancient conserved region remains, and the human Y chromosome is largely derived from the added region.

In all mammals, females have two X chromosomes, while males have a single X and a male-determining Y. The Y is much smaller than the X and contains few active genes. The mammalian X and Y chromosomes are thought to have evolved from a homologous pair of autosomes by a process of Y degradation (Ohno, 1967) which left only a small pseudoautosomal region (PAR) and a few genes shared between the differential regions of the X and Y chromosomes.

The origin of mammalian sex chromosomes has been explored by comparing their gene content in humans and their most distant mammalian relatives, marsupials and monotremes, which diverged from the eutherian ("placental") lineage about 130 and 170 million years ago, respectively (Hope et al., 1990). Our previous work showed that genes on the long arm and pericentric region of the human X were also on the X chromosome in marsupials and monotremes. However, genes distal to human Xp11.23 were located in two autosomal clusters in other mammal groups, the larger on tammar wallaby *(Macropus eugenii)* chromosome 5p. Cross-species chromosome painting between human and tammar wallaby confirm these regions of homology and nonhomology (Glas et al., 1999). Thus comparative mapping and painting defined an X conserved region (XCR), which represents the ancestral mammalian X, and an X added region (XAR), which was originally autosomal but was translocated in one or more events to the eutherian X after the divergence of eutherians from marsupials (Graves, 1995).

Since several human Y chromosome-borne genes share homology with sequences on the X, it is likely that the Y, too, consists of conserved and added regions. The evolutionary origin of genes on the human Y may be explored by mapping eutherian Y-borne genes or their X-borne relatives in marsupials. Limited observations on the content of the marsupial Y chromosome supported the hypothesis that the eutherian Y is also bipartite. We have now investigated the origin of five more genes on the human Y which enable us to draw at least a rough map of the ancient and added regions of the human Y chromosome. Surprisingly, this map reveals that most of the human Y was not original, but was derived from a recently added region.

Materials and methods

Male and female *M. eugenii* genomic DNA (10 µg) was digested with *Eco*RI or *Hin*dIII, run in 0.8% agarose, and blotted on Hybond-N+ membranes (Amersham). These membranes were subsequently hybridised with [α-^{32}P]dCTP-labelled probes (Megaprime Labelling Kit, Amersham).

An EMBL 3 genomic library was constructed from DNA extracted from the liver of a male tammar wallaby. Genomic DNA was partially digested with *Sau*3A and size-selected through a 10% to 40% glycerol gradient. Fractions were taken, and partially digested DNA ranging in size from 15 to 20 kb was ligated to λEMBL3 arms. The ligation reaction was packaged with Gigapack Gold (Stratagene). The library was titred and plated to a density of 120,000 plaque-forming units on four 22 × 22cm Nunc plates.

Small hybridising fragments of wallaby genomic clones were subcloned into pBluescript. The plasmid primers T3 and T7 were used to obtain sequence from either end of the cloned fragment with an Amplicycle kit (Promega) according to the manufacturer's instructions.

Supported by an Australian Research Council grant to J.A.M.G.

Received 16 June 2000; revision accepted 18 August 2000.

Request reprints from Dr. Paul D. Waters, Department of Genetics, La Trobe University, Bundoora, 3083, Victoria (Australia); telephone: +61-3-9479-2770; fax: +61-3-9479-2480; e-mail: waters@gen.latrobe.edu.au.

KARGER

Fax + 41 61 306 12 34
E-mail karger@karger.ch
www.karger.com

0301–0171/01/0922–0074$17.50/0

Accessible online at:
www.karger.com/journals/ccg

Lymphocytes were cultured from *M. eugenii* blood, supplied by Prof. M.B. Renfree (University of Melbourne), and held under Animal Ethics and Experimentation Permit No. LSB 96/4. Air-dried chromosome preparations were made from cultured peripheral blood according to standard methods (Schempp and Meer, 1983), with minor modifications. Chromosome in situ suppression (CISS) hybridisation of the genomic probes was performed (Lichter et al., 1988). Probes were labelled with digoxigenin and detected with anti-digoxigenin mouse antibodies (Sigma), followed by binding of tetramethylrhodamine isothiocyanate (TRITC)-conjugated goat anti-mouse antibodies (Sigma) and then TRITC-conjugated rabbit anti-goat antibodies (Toder et al., 1996). Chromosomes were then stained with 4′,6-diamidino-2-phenylindole (DAPI) dihydrochloride, and images were collected using a Zeiss Axioplan microscope with a liquid-cooled Photometrics charge-coupled device camera.

Results

We cloned and mapped the marsupial homologues of DBY, USP9Y, UTY, ZFY, and SMCY (characteristics described in Table 1) in the model marsupial *M. eugenii* (tammar wallaby) in order to determine whether these genes were part of the conserved or recently added regions of the human Y chromosome.

Southern blots containing digested DNA from male and female wallabies were probed with cDNA from mouse *Usp9x, Dbx, Uty,* and *Smcy* (Agulnik et al., 1994a) and human ZFX and ZFY (Page et al., 1987) (Fig. 1). All probes detected a few clear bands, showing that the tammar wallaby genome contained sequences homologous to all genes. The *Smcy* probe detected male-specific bands that confirmed a Y-chromosome location for the marsupial homologue, as well as a dosed band, implying an X-borne homologue. The *Usp9x, Dbx, Uty,* and ZFX/ZFY probes detected no male-specific bands and no consistent male:female gene dosage differences, implying autosomal locations for their wallaby homologues.

Size-selected genomic DNA libraries were screened with the same cDNA probes. Clones were isolated for *Usp9x/y, Dbx/y, and Utx/y.* Restriction analysis was performed, and fragments hybridising strongly to the relevant probe were subcloned and partially sequenced to confirm clone identity. The marsupial clones displayed greater than 87% homology to their human homologues over coding regions. ZFX/Y clones were of two types, one of which was highly homologous to human ZFX/Y throughout the gene and the other only in the zinc finger region; the latter proved to be more similar to ZNF6 (Frost, unpublished). Two wallaby clones were isolated with *Smcy* cDNA; one was most similar to human SMCX and detected dosed bands on Southern blots, whereas the other was most closely related to human SMCY and detected male-specific bands.

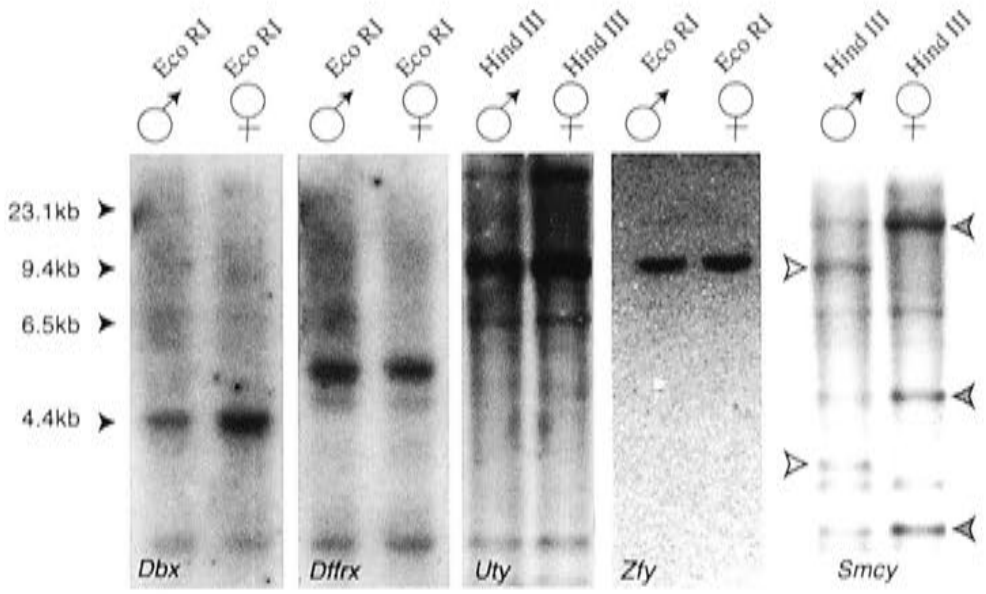


Fig. 1. Southern blots containing fully digested total genomic tammar wallaby DNA hybridized with either mouse *Dbx, Usp9y, Uty,* or *Smcy* or human ZFY. Open arrows denote male-specific bands. Grey arrows denote dosed bands.

Fluorescence in situ hybridisation (FISH) onto wallaby chromosomes was carried out with the full-length clones containing wallaby SMCX/Y, USP9X/Y, DBX/Y, and UTX/Y (Fig. 2). Digoxigenin-dUTP-labelled probes were hybridised onto DAPI-stained tammar wallaby metaphase chromosomes. The low diploid number (2n = 16) and the distinctive morphology made it easy to identify each chromosome except 3 and 4. SMCX mapped to the wallaby X, and SMCY localised to two regions on the wallaby Y chromosome, suggesting multiple Y-borne copies. Wallaby USP9Y/X, DBY/X, and UTY/X homologous clones all mapped to the same location on wallaby 5p. In situ hybridisation with a radioactively labelled wallaby ZFX/Y homologue produced a single peak over wallaby chromosome 5p (Fig. 2).

Table 1. Position and expression of X- and Y-borne genes in human and mouse

Gene	Human				Mouse		Putative function	References
	Location on		Expression of locus on		Expression of locus on			
	X chromosome	Y chromosome	X chromosome	Y chromosome	X chromosome	Y chromosome		
ZFX/Y	Xp22.1	Yp11.3	Ubiquitous, not inactivated	Ubiquitous	Ubiquitous, inactivated	Testis-specific	Spermatogenesis factor?	Page et al. (1987); Mardon and Page (1989); Jegalian and Page (1998)
USP9X/Y	Xp11.4	Yq11.2	Ubiquitous, not inactivated	Ubiquitous	Ubiquitous, inactivated	Testis-specific	Spermatogenesis factor?	Jones et al. (1996); Lahn and Page (1997a); Brown et al. (1998)
DBX/Y	?	Yq11.2	Ubiquitous, not inactivated	Ubiquitous	?	?	?	Lahn and Page (1997a)
UTX/Y	Xp11.23	Yq11.2	Ubiquitous, not inactivated	Ubiquitous	Ubiquitous, not inactivated	Ubiquitous	Hya determinant	Greenfield et al. (1996, 1998)
SMCX/Y	Xp11.2	Yq11.2	Ubiquitous, not inactivated	Ubiquitous	Ubiquitous, not inactivated	Ubiquitous	Hya determinant	Agulnik et al. (1994a, b); Wang et al. (1995)

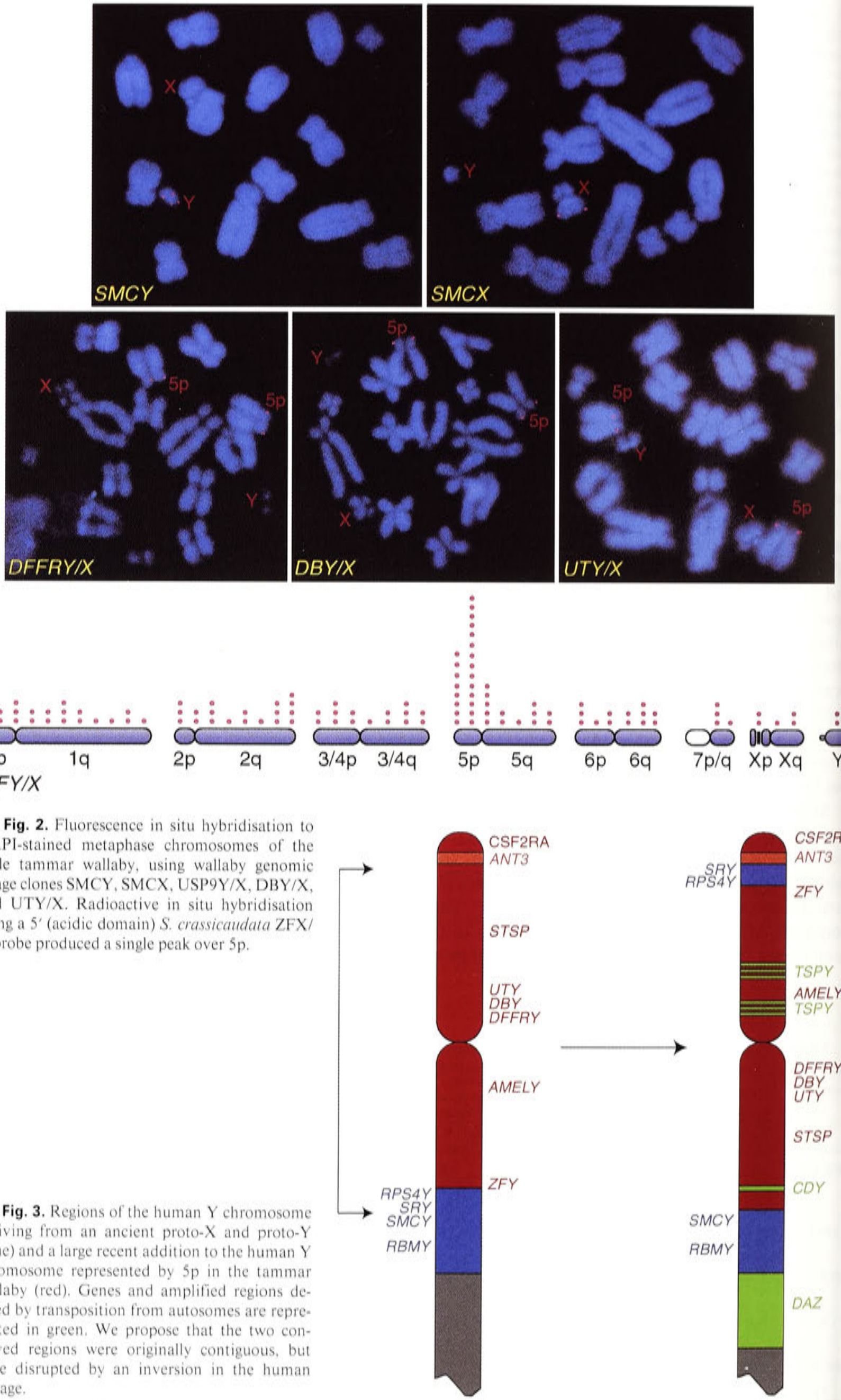


Fig. 2. Fluorescence in situ hybridisation to DAPI-stained metaphase chromosomes of the male tammar wallaby, using wallaby genomic phage clones SMCY, SMCX, USP9Y/X, DBY/X, and UTY/X. Radioactive in situ hybridisation using a 5′ (acidic domain) *S. crassicaudata* ZFX/Y probe produced a single peak over 5p.

Fig. 3. Regions of the human Y chromosome deriving from an ancient proto-X and proto-Y (blue) and a large recent addition to the human Y chromosome represented by 5p in the tammar wallaby (red). Genes and amplified regions derived by transposition from autosomes are represented in green. We propose that the two conserved regions were originally contiguous, but were disrupted by an inversion in the human lineage.

Table 2. Characteristics of genes on the human Y chromosome

Gene	Y location (Mb from Yp telomere)	Phenotype	X location[a]	Location in tammar wallaby
Yp telomere	0			
CSF2RA	1.23		Xp22.3	5p
SLC25A6	1.3		Xp22.3	3/4p
SRY/SOX3	2.56	Sex determination	Xq26→q27	X/Y
RPS4Y/X	2.58		Xq13.1	X
ZFY/X	2.64		Xp22.1	5p
TSPY major	8.31		–	Autosomal
AMELY/X	9.51	GBY[b] 9.00–11.49	Xp22.31→p22.1	5p
TSPY minor	10.53		–	Autosomal
Centromere	11.40	GCY[c] 11.31–14.69		
USP9Y/X	12.68	AZFa 11.92–13.37	Xp11.4	5p
DBY/X	12.84		?	5p
UTY/X	13.12		Xp11.32	5p
STS/P	15.86		Xp22.32	5p/2p
CDY	18.31–18.79		–	?
SMCY/X	20.11		Xp11.2	X/Y
RBMY/X	21.95–22.93	AZFb 20.64–22.93	Xq26→q27	X/Y
DAZ	22.93–27.28		–	?
Heterochromatin border	27.28			

[a] The lack of an X-chromosome homologue is denoted by a dash (–).
[b] GBY is the region associated with gonadoblastoma.
[c] GCY is the Y-specific growth control region.

Genes shown in italics represent recent additions to the Y chromosome. Genes shown in boldface are conserved. Genes shown in plain type were transposed to the Y chromosome from autosomes.

Discussion

These results consolidate the fragmentary evolutionary picture of the human Y. It was previously shown by somatic hybrid analysis that marsupial homologues of ZFY were autosomal (Sinclair et al., 1988), but localization was uncertain and complicated by cross-hybridisation with ZNF6, since shown to map to chromosome 1 (Frost, unpublished). AMELX/Y was previously shown to be autosomal in marsupials and monotremes, mapping near the centromere of wallaby chromosome 5 (Watson et al., 1992). STS/STSP and the PAR gene CSF2RA were recently mapped to wallaby 5p (Toder and Graves, 1998), while SLC25A6 mapped to wallaby chromosome 3 (Toder and Graves, 1998). The marsupial homologues of two human Y-borne genes, SRY and RBMY, have also been mapped to the Y chromosome in the tammar wallaby, and their X homologues have been mapped to the wallaby X (Foster et al., 1992; Foster and Graves 1994; Delbridge et al., 1999). A report of a male-specific RPS4Y (Jegalian and Page, 1998) band on Southern blots of the opossum suggests that this gene also has a homologue on the Y in some marsupial species, although no male-specific band is apparent in the tammar wallaby (Duffy, unpublished). The present results substantiate an earlier report of a male-specific band in opossum DNA detected with human SMCY (Agulnik et al., 1994a).

Our present results, together with previous localizations, define a Y conserved region (YCR) of the human Y containing four genes (SRY, RBMY, RPS4Y, and SMCY) and a Y added region (YAR) containing eight genes (USP9Y, DBY, UTY, ZFY, AMELY, STSP, CFS2RA, and SLC25A6). All but one of these YAR genes co-localise on wallaby chromosome 5p and, therefore, are part of a single recent addition to the eutherian Y.

Significantly, the wallaby homologues of 10 X-specific human genes located distal to human Xp11.23 also map in the same region of chromosome 5p in the wallaby. They are also grouped on chromosome 3q in a distantly related marsupial species *Sminthopsis crassicaudata,* and on chromosome 1 in monotremes (Graves,1995), implying that this is an ancient gene cluster. This cluster of X and Y genes on wallaby chromosome 5p therefore represents an ancient region which was originally autosomal in the ancestors of monotremes, marsupials, and eutherians, but was added to both sex chromosomes in the eutherian linage, then subsequently degraded on the Y.

The arrangement of YCR and YAR genes on the human Y chromosome (Table 2) reveals the evolutionary origin of parts of the human Y (Fig. 3). There are two small conserved regions on the human Y chromosome. One (adjacent to the Yp PAR) contains SRY and RPS4Y, and the other (near the Yq heterochromatin) contains SMCY and the major cluster of expressed RBMY copies. These regions amount to only 8% to 24% of the euchromatic portion of the Y chromosome (the lower figure is calculated from the distance separating the pairs of YCR genes, and the higher figure from the distance separating the YAR genes on either side [Affara et al., 1996], excluding the highly repetitive region below RBMY, which contains only amplified sequences with no X homologues). The minimum size (about 10 Mb) accords with the very small size of the con-

served marsupial Y, with which it shares homology. Since the region of the X with which it once shared homology is at least 100 Mb, we conclude that the original Y chromosome has been much reduced; indeed, it has all but disappeared.

Remarkably, most of the euchromatic region of the human Y, including the Yp PAR, contains only genes which are autosomal in marsupials and map together on chromosome 5p in the wallaby. Exceptions are the amplified non-expressed copies of RBMY, which are distributed with TSPY sequences in two additional blocks in Yp and proximal Yq and may have originated via tandem duplication. In addition, there are genes (e.g., DAZ and CDY) with no X homologues, which are likely to have been transposed into the Y from an autosomal location. Most of the human Y therefore represents the relic of an autosomal region added to the proto-X and proto-Y between 80 and 130 million years ago.

The presence of these wallaby chromosome 5 genes on both the human X and Y implies that a large autosomal block was added to both sex chromosomes after the divergence of eutherians from marsupials 130 million years ago, but before the eutherian radiation 80 million years ago, as proposed previously (Graves, 1995). A single inversion is all that was required to redistribute the YCR and YAR genes into two smaller blocks (Fig. 3), although there must have been rearrangements within the YCR before the inversion to produce the gene order shown in Fig. 3 from the original order retained by the X, as well as other rearrangements within the YAR after it was added (Gläser et al., 1997). These rearrangements must all be relatively recent in the hominid lineage, since in lower primates the Y-chromosome gene order SHOX–IL3RA–SLC25A6P–SRY still reflects that of the X chromosome (Gläser et al., 1999), from which the Y derived.

Our direct demonstration that genes on the human Y are arranged in two small conserved blocks amid a majority of recently added material confirms and extends the hypothesis put forward several years ago (Graves, 1995) that the human Y chromosome consists of a small conserved region and a large added region. This demonstration of a separate origin by comparative mapping constitutes much stronger evidence than sequence comparison, which is subject to many vagaries, especially for genes on the Y chromosome; however, our hypothesis is quite compatible with Lahn and Page's division of the human Y chromosome into five "geological strata" claimed, on the basis of sequence homology between X- and Y-borne homologues, to have been isolated at different times (Lahn and Page, 1999a). Their strata I and II represent the YAR, while strata III, IV, and V represent the ancient YCR. Both regions were probably further subdivided by major rearrangements which eliminated homologous pairing and permitted rapid degradation within a large region.

Some human Y-specific genes with no X homologue appear to have a more recent origin than the X-Y shared genes within the YCR and YAR. DAZ lies on the human Y but is not on the mouse or marsupial Y (Delbridge et al., 1997), and the intronless CDY is present only in simians (Lahn and Page, 1999b). TSPY also has no X homologue in any species: it is present as a single copy on the mouse Y and as multiple copies on human Yp, but is absent from the Y in marsupials (Delbridge et al., 1997). These genes all have autosomal homologues from which they were evidently transposed or retrotransposed into the primate Y chromosome 50–80 million years ago, inserting into a part of either the conserved or added region (Fig. 3).

The evolutionary origin of genes on the human Y may be reflected by their function, or lack of it. Genes that persisted on the Y for the 130 million years since eutherians and marsupials diverged are likely to have been selected for a critical function in sex determination (SRY) or spermatogenesis (e.g., RBMY). Genes that were added more recently may still be present simply because they have not yet had time to degrade (e.g., human STSP, which appears to have recently suffered an exon deletion [Yen et al., 1988]). However, some genes within the YAR have evidently acquired a male-specific function; for instance, DFFRY is likely to be involved in spermatogenesis (Sargent et al., 1999; Sun et al., 1999). Evidently, the Y copies of many genes which were originally added to the X and Y (e.g., the X-specific genes DMD and PDHA1, which also map to wallaby chromosome 5p) have already been degraded and deleted (Sinclair et al., 1988; Fitzgerald et al., 1993). Another recent addition to the sex chromosomes, represented by a human PAR gene (SLC25A6) which maps to wallaby chromosome 3, may be the sole survivor of an independently added region, and a region of wallaby chromosome 1, which maps to the human XAR but not the YAR (Graves et al., 1995), may have been added earlier, for it has evidently left no trace on the human Y.

New genes found within the two small regions of the human Y which have been conserved from the original Y of an ancient mammalian ancestor are likely to have a critical male-specific function. The tiny marsupial Y may serve as a model of this ancient mammalian Y.

Acknowledgements

We thank Mrs. Iole Barbieri for her expert assistance with cell culture, Dr. M. Mitchell for supplying the mouse probes, Dr. David Page for supplying the ZFX/ZFY probes, and Prof. M.B. Renfree for supplying *M. eugenii* tissue. We are grateful to Dr. David Page for his critical comments on an earlier draft.

served marsupial Y, with which it shares homology. Since the region of the X with which it once shared homology is at least 100 Mb, we conclude that the original Y chromosome has been much reduced; indeed, it has all but disappeared.

Remarkably, most of the euchromatic region of the human Y, including the Yp PAR, contains only genes which are autosomal in marsupials and map together on chromosome 5p in the wallaby. Exceptions are the amplified non-expressed copies of RBMY, which are distributed with TSPY sequences in two additional blocks in Yp and proximal Yq and may have originated via tandem duplication. In addition, there are genes (e.g., DAZ and CDY) with no X homologues, which are likely to have been transposed into the Y from an autosomal location. Most of the human Y therefore represents the relic of an autosomal region added to the proto-X and proto-Y between 80 and 130 million years ago.

The presence of these wallaby chromosome 5 genes on both the human X and Y implies that a large autosomal block was added to both sex chromosomes after the divergence of eutherians from marsupials 130 million years ago, but before the eutherian radiation 80 million years ago, as proposed previously (Graves, 1995). A single inversion is all that was required to redistribute the YCR and YAR genes into two smaller blocks (Fig. 3), although there must have been rearrangements within the YCR before the inversion to produce the gene order shown in Fig. 3 from the original order retained by the X, as well as other rearrangements within the YAR after it was added (Gläser et al., 1997). These rearrangements must all be relatively recent in the hominid lineage, since in lower primates the Y-chromosome gene order SHOX–IL3RA–SLC25A6P–SRY still reflects that of the X chromosome (Gläser et al., 1999), from which the Y derived.

Our direct demonstration that genes on the human Y are arranged in two small conserved blocks amid a majority of recently added material confirms and extends the hypothesis put forward several years ago (Graves, 1995) that the human Y chromosome consists of a small conserved region and a large added region. This demonstration of a separate origin by comparative mapping constitutes much stronger evidence than sequence comparison, which is subject to many vagaries, especially for genes on the Y chromosome; however, our hypothesis is quite compatible with Lahn and Page's division of the human Y chromosome into five "geological strata" claimed, on the basis of sequence homology between X- and Y-borne homologues, to have been isolated at different times (Lahn and Page, 1999a). Their strata I and II represent the YAR, while strata III, IV, and V represent the ancient YCR. Both regions were probably further subdivided by major rearrangements which eliminated homologous pairing and permitted rapid degradation within a large region.

Some human Y-specific genes with no X homologue appear to have a more recent origin than the X-Y shared genes within the YCR and YAR. DAZ lies on the human Y but is not on the mouse or marsupial Y (Delbridge et al., 1997), and the intronless CDY is present only in simians (Lahn and Page, 1999b). TSPY also has no X homologue in any species: it is present as a single copy on the mouse Y and as multiple copies on human Yp, but is absent from the Y in marsupials (Delbridge et al., 1997). These genes all have autosomal homologues from which they were evidently transposed or retrotransposed into the primate Y chromosome 50–80 million years ago, inserting into a part of either the conserved or added region (Fig. 3).

The evolutionary origin of genes on the human Y may be reflected by their function, or lack of it. Genes that persisted on the Y for the 130 million years since eutherians and marsupials diverged are likely to have been selected for a critical function in sex determination (SRY) or spermatogenesis (e.g., RBMY). Genes that were added more recently may still be present simply because they have not yet had time to degrade (e.g., human STSP, which appears to have recently suffered an exon deletion [Yen et al., 1988]). However, some genes within the YAR have evidently acquired a male-specific function; for instance, DFFRY is likely to be involved in spermatogenesis (Sargent et al., 1999; Sun et al., 1999). Evidently, the Y copies of many genes which were originally added to the X and Y (e.g., the X-specific genes DMD and PDHA1, which also map to wallaby chromosome 5p) have already been degraded and deleted (Sinclair et al., 1988; Fitzgerald et al., 1993). Another recent addition to the sex chromosomes, represented by a human PAR gene (SLC25A6) which maps to wallaby chromosome 3, may be the sole survivor of an independently added region, and a region of wallaby chromosome 1, which maps to the human XAR but not the YAR (Graves et al., 1995), may have been added earlier, for it has evidently left no trace on the human Y.

New genes found within the two small regions of the human Y which have been conserved from the original Y of an ancient mammalian ancestor are likely to have a critical male-specific function. The tiny marsupial Y may serve as a model of this ancient mammalian Y.

Acknowledgements

We thank Mrs. Iole Barbieri for her expert assistance with cell culture, Dr. M. Mitchell for supplying the mouse probes, Dr. David Page for supplying the ZFX/ZFY probes, and Prof. M.B. Renfree for supplying *M. eugenii* tissue. We are grateful to Dr. David Page for his critical comments on an earlier draft.

References

Affara N, Bishop C, Brown W, Cooke H, Davey P, Ellis N, Graves JM, Jones M, Mitchell M, Rappold G, Tyler-Smith C, Yen P, Lau Y-FC: Report of the second international workshop on human Y chromosome mapping 1995. Cytogenet Cell Genet 73:33–76 (1996).

Agulnik A, Mitchell MJ, Lerner JL, Woods DR, Bishop CE: A mouse Y chromosome gene encoded by a region essential for spermatogenesis and expression of male-specific minor histocompatibility antigens. Hum molec Genet 3:873–878 (1994a).

Agulnik AI, Mitchell MJ, Mattel MG, Borsani G, Avner PA, Lerner JL, Bishop CE: A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. Hum molec Genet 3:879–884 (1994b).

Brown GM, Furlong RA, Sargent CA, Erickson RP, Longepied G, Mitchell, M, Jones MH, Hargreave TB, Cooke HJ, Affara NA: Characterisation of the coding sequence and fine mapping of the human DFFRY gene and comparative expression analysis and mapping to the *Srx (b)* interval of the mouse Y chromosome of the *Dffry* gene. Hum molec Genet 7:97–107 (1998).

Delbridge ML, Disteche CM, Lingenfelter PA, Graves JAM: The candidate spermatogenesis gene RBMY has a homologue on the human X chromosome. Nature Genet 22:223–224 (1999).

Delbridge ML, Harry JL, Toder R, O'Neill RJW, Ma K, Chandley AC, Graves JAM: A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome. Nature Genet 15:131–136 (1997).

Fitzgerald J, Wilcox SA, Graves JAM, Dahl HHM: A eutherian X-linked gene, PDHA1, is autosomal in marsupials: a model for the evolution of a second , testis-specific variant in eutherian mammals. Genomics 18:636–642 (1993).

Foster JW, Brennan FE, Hampikian GK, Goodfellow PN, Sinclair AH, Lovell-Badge R, Selwood L, Renfree MB, Cooper DW, Graves JAM: Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature 359:531–533 (1992).

Foster JW, Graves JAM: An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc natl Acad Sci, USA 91:1927–1931 (1994).

Glas R, Graves JMG, Toder R, Ferguson-Smith M, O'Brien PC: Cross-species chromosome painting between human and marsupial directly demonstrates the ancient region of the mammalian X. Mammal Genome 10:1115–1116 (1999).

Gläser B, Grutzner F, Taylor K, Schiebel K, Meroni G, Tsioupra K, Pasantes J, Rietschel W, Toder R, Willmann U, Zeitler S, Yen P, Ballabio A, Rappold G, Schempp W: Comparative mapping of Xp22 genes in hominoids: evolutionary linear instability of their Y homologues. Chrom Res 5:167–176 (1997).

Gläser B, Myrtek D, Rumpier Y, Schiebel K, Hauwy M, Rappold GA, Schempp W: Transposition of SRY into the ancestral pseudoautosomal region creates a new pseudoautosomal boundary in a progenitor of simian primates. Hum molec Genet 15:2071–2778 (1999).

Graves JAM: The origin and function of the mammalian Y chromosome and Y-borne genes: an evolving understanding. BioEssays 17:311–320 (1995).

Greenfield A, Carrel L, Pennisi D, Philippe C, Quaderi N, Siggers P, Steiner K, Tam PPL, Monaco AP, Willard HF, Koopman P: The UTX gene escapes X inactivation in mice and humans. Hum molec Genet 7:737–742 (1998).

Greenfield A, Scott D, Pennisi D, Ehrmann I, Ellis P, Cooper L, Simpson E, Koopman P: An H-YD^b epitope is encoded by a novel mouse Y chromosome gene. Nature Genet 14:474–478 (1996).

Hope RM, Cooper S, Wainwright B: in Graves JAM, Hope RM, Cooper DW (eds): Globin macromolecular sequences in marsupials and monotremes, in Mammals from Pouches and Eggs: Genetic Breeding and Evolution of Marsupials and Monotremes, pp 147–172 (CSIRO, Melbourne, 1990).

Jegalian K, Page DC: A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature 394:776–780 (1998).

Jones MH, Furlong RA, Burkin H, Chalmers IJ, Brown GM, Khwaja O, Affara NA: The *Drosophila* developmental gene *fat facets* has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. Hum molec Genet 5:1695–1701 (1996).

Lahn B, Page DC: Functional coherence of the human Y chromosome. Science 278:675-680 (1997).

Lahn B, Page DC: Four evolutionary strata on the human X chromosome Science 286:964–967 (1999a).

Lahn B, Page DC: Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome Nature Genet 21:429–433 (1999b).

Lichter P, Cremer T, Borden J, Manuelidis L, Ward D: Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet 80:224–234 (1988).

Mardon G, Page DC: The sex-determining region of the mouse Y chromosome encodes a protein with a highly acidic domain and 13 zinc fingers. Cell 56:765–770 (1989).

Ohno S: Sex Chromosomes and Sex Linked Genes (Springer-Verlag, Berlin, 1967).

Page DC, Mosher R, Simpson EM, Fisher EM, Mardon G, Pollack J, McGillivray B, de la Chapelle A, Brown LG: The sex determining region of the human Y chromosome encodes a finger protein. Cell 51:1091–1104 (1987).

Sargent CA, Boucher CA, Kirsch S, Brown G, Weiss B, Trundley A, Burgoyne P, Saut N, Durand C, Levy N, Terriou P, Hargreave T, Cooke H, Mitchell M, Rappold GA, Affara NA: The critical region of overlap defining the AZFa male infertility interval of proximal Yq contains three transcribed sequences. J med Genet 36:670–677 (1999).

Schempp W, Meer B: Cytological evidence for three human X-chromosomal segments escaping inactivation. Hum Genet 63:171–174 (1983).

Sinclair AH, Foster JW, Spencer JA, Page DC, Palmer M, Goodfellow PN, Graves JAM: Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. Nature 336:780–783 (1988).

Sun C, Sklaetsky H, Birren B, Devon K, Tang ZL, Silber S, Oates R, Page DC: An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. Nature Genet 23:429–432 (1999).

Toder R, Graves JAM: CSF2RA, ANT3, and STS are autosomal in marsupials: implications for the origin of the pseudoautosomal region of mammalian sex chromosomes. Mammal Genome 9:373–376 (1998).

Toder R, Wilcox SA, Smithwick M, Graves JAM: The human/mouse imprinted genes IGF2, H19, SNRPN and ZNF127 map to two conserved autosomal clusters in a marsupial. Chrom Res 4:295–300 (1996).

Wang W, Meadows LR, Denhaan JMM, Sherman NE, Chen Y, Blokland E, Shabanowitz J, Agulnik AI, Hendrickson RC, Bishop CE, Hunt DF, Goulmy E, Engelhard VH: Human H-y: a male-specific histocompatibility antigen derived from the SMCY protein. Science 269:1588–1590 (1995).

Watson JM, Spencer JA, Graves JAM, Snead ML, Lau EC: Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. Genomics 14:785–789 (1992).

Yen PH, Marsh B, Allen E, Tsai SP, Ellison J, Connolly L, Neiswanger K, Shapiro JL: The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution. Cell 55:1123–1135 (1988).

Cytogenet Cell Genet 93:315–316 (2001)

Cytogenetics and Cell Genetics

Assignment[1] of the SMARCF1 gene to tammar wallaby chromosome 5q by fluorescence in situ hybridisation

P.D. Waters, P.J. Kirby, and J.A.M. Graves

Comparative Genomics Group, Research School of Biological Science, The Australian National University, Canberra (Australia)

[1] To our knowledge this is the first time this gene has been mapped in tammar wallaby.

Rationale and significance

SWI/SNF-related, matrix associated, actin-dependent regulator of chromatin, subfamily f, member 1 (SMARCF1) is a member of a conserved family of proteins distinguished by a DNA binding motif called ARID (AT-rich interactive domain) (Dallas et al., 2000). First isolated in yeast, and followed by the characterization of similar SWI/SNF complexes in mammals and *Drosophila melanogaster*, these complexes have been suggested to play fundamental roles in the regulation of gene expression during cell growth and development via altering chromatin structure (Kingston and Narlikar, 1999; Kadonaga, 1998). SMARCF1 maps to human 1p36.1→p35 (Takeuchi et al., 1998) where it has been suggested that at least two tumor suppressor genes are located (Caron et al., 1995). In this study we cloned and mapped the genomic SMARCF1 in the model marsupial *Macropus eugenii* (tammar wallaby) of particular value for comparative genetics because of its early divergence (about 130 million years ago) from placental mammals. SMARCF1 was cloned and mapped by fluorescent in situ hybridization to the distal region of chromosome 5q. This is the first gene to be mapped to chromosome 5q in the tammar wallaby and may represent a region of homology between human chromosome 1 and wallaby chromosome 5.

Materials and methods

A tammar wallaby lambda genomic DNA library was screened with a labeled cDNA probe for VCY (kindly supplied by Dr. David Page, Howard Hughes Medical Institute, Whitehead Institute, Cambridge, MA, USA). Five positive clones were isolated and sequencing revealed that one contained the wallaby SMARCF1 gene. The SMARCF1 clone was biotin-14-dCTP labeled by nick translation and hybridised to tammar wallaby metaphase chromosomes. Hybridisation signals were detected with anti-biotin raised in goat and anti-goat coupled with FITC and counter stained with DAPI.

Probe name: MeSMARC1.1
Probe type: genomic DNA
Insert size: approximately 20 kb
Vector: Lambda EMBL3
Proof of authenticity: sequencing

Results

Mapping data
Location: 5q distal region (Fig. 1)
Number of cells examined: 20
Number of cells with specific signal: 16
Most precise assignment: 5q distal region
Location of background signals (sites with >2 signals): none observed

Supported by an Australian Research Council grant to J.A.M.G. We thank Mrs. Iole Barbieri for her assistance with cell culture and Professor M.B. Renfree for supplying *M. eugenii* tissue.

Received 8 May 2001; manuscript accepted 18 May 2001.

Request reprints from P.D. Waters, Comparative Genomics Group, Research School of Biological Science, The Australian National University, Actin, 2601, ACT (Australia); telephone: +16 2 6125 2371; fax: +16 2 6125 4891; e-mail: waters@rsbs.anu.edu.au

KARGER

Fax + 41 61 306 12 34
E-mail karger@karger.ch
www.karger.com

0301–0171/01/0934–0315$17.50/0

Accessible online at:
www.karger.com/journals/ccg

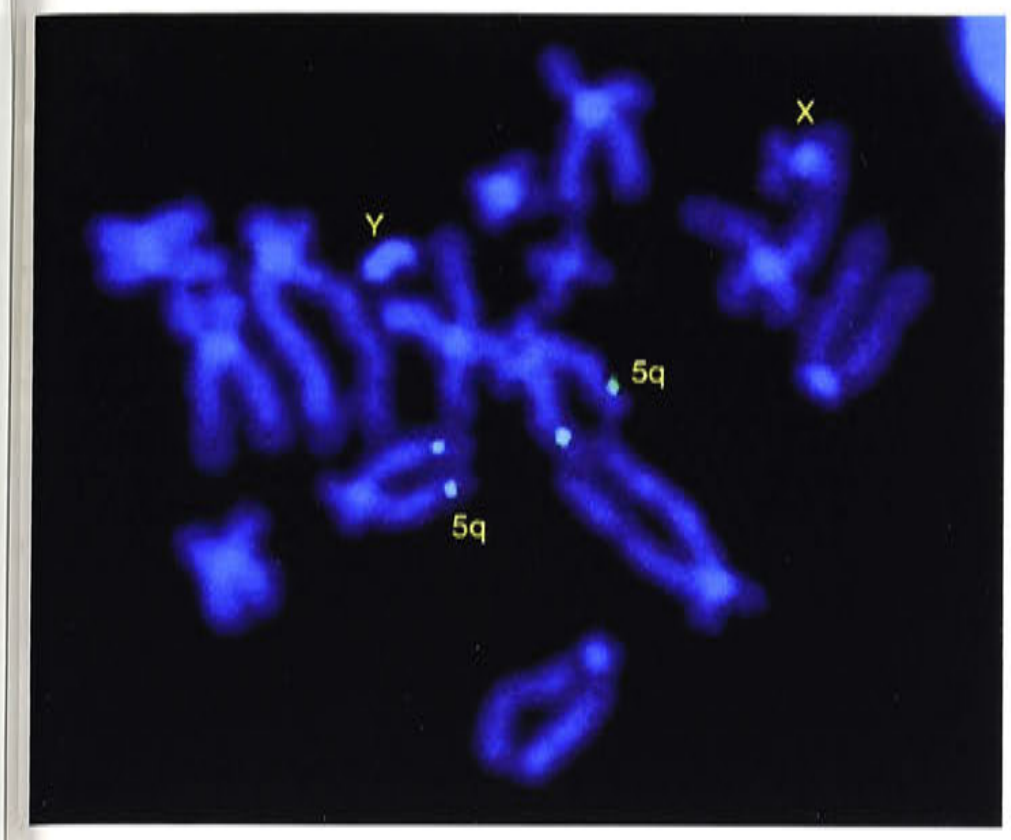


Fig. 1. The tammar wallaby SMARC1 homologue hybridised to 5q.

References

Caron H, Peter M, van Sluis P, Speleman F, de Kraker J, Laureys G, Michon J, Brugieres L, Voute PA, Westerveld A: Evidence for two suppressor loci on chromosome bands 1p36 → p35 involved in neuroblastoma: one probably imprinted, another associated with N-myc amplification. Hum molec Genet 4:535–539 (1995).

Dallas PB, Pacchione S, Wilsker D, Bowrin V, Kobayashi R, Moran E: The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. Mol Cell Biol 20:3137–3149 (2000).

Kadonaga JT: Eukaryotic transcription, an interlaced network of transcription factors and chromatin-modifying machines. Cell 92:307–313 (1998).

Kingston RE, Narlikar GJ: ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 13:2339–2352 (1999).

Takeuchi T, Furihata M, Heng HH, Sonobe H, Ohtsuki Y: Chromosomal mapping and expression of the human B120 gene. Gene 213:189–193 (1998).