A MOLECULAR GENETIC ANALYSIS OF TESTIS DETERMINATION IN THE MOUSE

John Ernest Gubbay

A thesis submitted for the Degree of Doctor of Philosophy 1992

Department of Eukaryotic Molecular Genetics,
National Institute for Medical Research,
The Ridgeway, Mill Hill,
London, NW7 1AA.

Department of Genetics,
University College London,
Gower Street,
London, WC1E 6BT.

ProQuest Number: 10609157

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10609157

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To my parents

Abstract

The testis determining gene is the Y-linked gene responsible for initiating the developmental pathway leading to testis formation in males. A strategy based upon determining the precise chromosomal location of this locus has been used to clone candidate genes. In this study members of the *ZFY* gene family have been analysed in terms of their expression pattern and their functional status in XY̶ female mice known to be defective in the testis determining gene. The failure to detect any molecular abnormality in *ZFY* genes in these mice, their inappropriate expression pattern and other data show that these genes cannot be involved in testis determination.

A second candidate gene known as *SRY* was subsequently isolated in humans, also by positional cloning. This study describes the cloning of the homologous gene in mouse (*Sry*) and presents data showing that this is indeed the testis determining gene. Its chromosomal location and putative protein structure are consistent with a role in sex determination. *Sry* seems to be the only gene affected in XY̶ female mice, indicating that it is normally necessary for testis determination. Models for how the deletion event which gave rise to these XY̶ female mice occurred are considered in the light of the unusual genomic organisation found at the *Sry* locus. Finally, the finding that a genomic fragment carrying *Sry* can cause male development in XX mice has proved that *Sry* is the only gene from the Y chromosome necessary for testis determination.

Contents

Abstract 3
Contents 4
Figures and Tables 8
Abbreviations 10
Acknowledgements 11

Chapter 1: Introduction 12
1.1 Non-mammalian sex-determining mechanisms 12
1.2 Mammalian sex determination 16
1.3 The biology of gonadal and secondary sexual differentiation 18
1.4 Genetic definition of *TDF* in man 21
1.5 Genetic definition of *Tdy* in mouse 24
1.6 Mode of action of *Tdy* 29
1.7 XY female sex reversal 35
1.7.1 XY^{Tdym1} female mice 36
1.7.2 Incompletely penetrant XY sex reversal in the mouse 38
1.8 The identification of candidate genes for *TDF/Tdy* 41

Chapter 2: Materials and Methods 44
2.1 Molecular Biology techniques 44
2.1.1 Purification of DNA fragments from agarose gels 44
2.1.2 Preparation of vectors 44
2.1.3 Ligations 44
2.1.4 Preparation of competent bacteria 44
2.1.5 Transformation 45
2.1.6 Plasmid minipreps 45
2.1.7 Plasmid maxipreps 45
2.1.8 Sequencing 46
2.2 Mouse stocks and breeding 46
2.3 DNA isolation 47
2.4 RNA isolation 47

2.5 Southern blotting 48
2.6 Northern blotting 48
2.7 Hybridization 49
2.8 Polymerase Chain Reaction (PCR) 50
2.8.1 *Zfy* expression 50
2.8.2 XY̶ female analysis 51
2.8.3 *Sry* transgenics 51
2.9 Genomic library construction and screening 52
2.9.1 Size selected library 52
2.9.2 Total genomic library 52
2.9.3 Plating and screening 52
2.9.4 Isolation of phage DNA 53
2.10 Microinjection of fertilised mouse eggs 54
2.11 Sex chromatin staining in amnion cells 54
2.12 Histology 55
2.13 Karyotype analysis 55
2.14 Commonly used solutions 56
2.15 Bacterial strains 56

Chapter 3: Results and Discussion 57
Zfy* as a candidate for *Tdy
3.1 Analysis of *Zfy* gene structure in XY̶ female mice 57
3.2 Analysis of *Zfy* gene expression in XY̶ female mice 63
3.2.1 Expression of *Zfy* genes in adult testis from wild type mice 63
3.2.2 Expression of *Zfy* genes in adult testis from XY̶Y mice 65
3.2.3 Expression of *Zfy* genes in normal male embryos 69
3.2.4 Expression of *Zfy* genes in XY̶ female embryos 70
3.3 *Zfy* genes cannot be involved in testis determination 71
3.4 What is the true function of the *ZFY* gene family? 74

Chapter 4: Results and Discussion 78
Analysis of endogenous retroviral elements in XY̶ female mice

Chapter 5: Results and Discussion 83
Cloning and sequence analysis of the mouse *Sry* gene
5.1 Cloning a new candidate for *TDF/Tdy* 83
5.2 Cloning the Y-linked mouse homologue of pY53.3 86
5.3 Sequencing the Y-linked mouse homologue of pY53.3 86
5.4 *Sry* is a DNA binding protein 89

Chapter 6: Results and Discussion 96
Analysis of the genomic structure of *Sry*
6.1 Cloning a 14.5kb genomic fragment containing *Sry* 96
6.2 The *Sry* locus consists of an inverted duplication 96
6.3 *Sry* from a *musculus* and a *domesticus* Y chromosome has a similar structure 98
6.4 The *Sry* inverted repeat: Origin and evolution 100

Chapter 7: Results and Discussion 102
***Sry* is deleted in XɎ female mice**
7.1 Identifying a deletion breakpoint in the XɎ females 102
7.2 Analysis of the deletion in a number of XɎ female mice. 105
7.3 Full characterization of the *Sry* deletion in XɎ female mice . . . 105
7.4 Possible mechanisms for the deletion of *Sry* in XɎ females 113

Chapter 8: Results and Discussion 117
Sry* is equivalent to *Tdy
8.1 Analysis of transgenic embryos 119
8.2 Analysis of adult transgenic mice 123
8.3 *Sry* transgenic mice 126
8.4 *Sry* does not always give sex-reversal 127
8.5 Human *SRY* does not cause sex-reversal in XX transgenic mice 129

Chapter 9: General Discussion 130
9.1 Human XX males 130
9.2 Differences between *SRY* and *Sry* 131

9.3 How is *Sry* regulated? 132
9.4 What are the target genes of *Sry*? 132

Appendix: Sequence of L741 136
References 147

Figures and Tables

Figure 1. The human and mouse Y chromosomes 22

Figure 2. Normal and abnormal X-Y interchanges 25

Figure 3. Sex reversal in the mouse 27

Figure 4. Comparison of *Zfy* genes in XY̶ female versus XY male mice by Southern analysis 58

Figure 5. Assignment of bands to *Zfy-1* and *Zfy-2* by Southern analysis 60

Figure 6. *Zfy-1* genomic map 61

Figure 7. Revised map of the *Zfy-1* genomic locus 62

Figure 8. Northern blot analysis of *Zfy* gene expression 64

Figure 9. Identification of XY̶^{m}Y^{d} progeny by Southern analysis 67

Figure 10. Sequences obtained from RT-PCR analysis of *Zfy* transcripts from XY̶^{m}Y^{d} adult testes 68

Figure 11. PCR analysis of *Zfy-1* expression in XY̶ female, XY male and XX female 12.5*dpc* embryonic gonads 72

Figure 12. Analysis of endogenous retroviral insertions in XY̶ female mice 79

Figure 13. Mapping conserved sequences to the sex-determining region of the mouse Y chromosome 85

Figure 14. Nucleotide comparison of mouse and human Y-linked sequences 87

Figure 15. Conservation of homologous protein domains 90

Figure 16. Guide to *Sry* genomic clones and structure of the wild type locus 97

Figure 17. Analysis of a deletion breakpoint in the XY̶ female 103

Figure 18. Analysis of the deletion in a panel of XY̶ females 106

Figure 19. Mapping the genomic region corresponding to the deletion in XY̶ females 108

Figure 20. Sequences at the deletion breakpoints in XY̶ females 109

Figure 21. Two possible schemes for the deletion of *Sry* in XY̶ female mice 110

Figure 22. PCR analysis of the deletion in XY̶ female mice 112

Figure 23.	Schematic representation of recombination events that may occur at the *Sry* locus	115
Table 1.	Summary of embryonic transgenic data	120
Figure 24.	Analysis of sex-reversed transgenic embryos	121
Figure 25.	Analysis of adult sex-reversed transgenic mouse m33.13	124

Abbreviations

ATP	adenosine triphosphate
(A) dATP	2' deoxyadenosine triphosphate
bp	base pairs
cDNA	complementary deoxyribose nucleic acid
(C) dCTP	2' deoxycytidine triphosphate
DNA	deoxyribose nucleic acid
(G) dGTP	2' deoxyguanidine triphosphate
kb	kilobase pairs
LTR	long terminal repeat
M	molar
(N) dNTP	2' deoxynucleotide triphosphate
OD	optical density
RNA	ribonucleic acid
rpm	revolutions per minute
(T) dTTP	2' deoxythymidine triphosphate
UV	ultraviolet light
≥	greater than or equal to
≤	less than or equal to

Acknowledgements

This thesis would not have been possible without the help, advice, collaboration and patience of many friends and colleagues over the last three years. In particular I am grateful to my supervisor Robin Lovell-Badge whose gentle guidance was the inspiration for this work. Blanche Capel, Androulla Economou, Peter (Koops) Koopman, David Jackson and Nigel Vivian have all made huge contributions to this work, both intellectually and physically, and I owe them a big thanks. All the following people have been wonderful friends and always willing to help: Jérôme Collignon, Bob Braun, Daniel Guerrier, Adam Hacker, Andrea Münsterberg, Michaela Riddell, Barbara Skene and Shantini Sockanathan.

I would also like to thank David Wilkinson, Robb Krumlauf and everybody in their labs for making our wing of NIMR such an enjoyable place to work.

Special thanks to Peter Goodfellow and his group, whose collaboration has been so important.

I am also grateful to Paul Burgoyne for his collaboration, help and advice, and to Anne McLaren my university supervisor for help and discussion.

In addition, I would like to thank Alan Ashworth for supplying one of the probes, Anna-Marie Frischauf for making one of the libraries and Peter Little who supplied a magical strain of bugs.

Last but not least, thanks to the photography and graphics departments at NIMR for all their hard work and Debbie Duthie for lots of typing.

Chapter 1: Introduction

A pervasive theme in the animal kingdom is the existence of organisms in two sexual forms the result of which leads to sexual reproduction which facilitates genetic exchange, variation and diversification. Males and females differ in reproductive function, anatomy, physiology and behaviour. This plethora of differences is known to stem from a single decision during development, which dictates whether the male or female blueprint will be followed. An impressive range of sex-determining mechanisms exists in nature. In some cases these mechanisms have proved amenable to study due to the availability of developmental mutations whose effect is solely to perturb the resulting sexual phenotype. In the last few years progress has extended towards an analysis of mammalian sex determination at a molecular level.

1.1 Non-mammalian sex-determining mechanisms

A detailed genetic analysis has established the hierarchy of genes involved in sex determination in two organisms in particular, the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans* (*C.elegans*) (Kimble, 1988; Meyer, 1988; Wolfner, 1988; Baker, 1989; Cline, 1989; Hodgkin, 1990; Villeneuve & Meyer, 1990). In both cases the primary sex-determining signal is the ratio of X chromosomes to autosomes. This signal determines the functional status of a cascade of regulatory genes which govern three independent processes: somatic sexual development, germ-line sexual development and dosage compensation.

Somatic sexual development refers to sex specific differentiation of all tissues besides the germ line. Germ-line sexual development refers to the type of gametes generated by the organism. Dosage compensation is the process which results in equal levels of gene expression from the single X chromosome in one sex compared to the two X chromosomes of the other. In *Drosophila* this is achieved by up-regulation of gene expression from the single X chromosome in males whilst in *C.elegans* down-regulation of gene expression occurs from the two X chromosomes in hermaphrodites. Somatic and germ-line sexual differentiation and dosage

compensation are controlled by an initial common pathway of regulatory genes which subsequently diverges into three separate paths. In *Drosophila* these three "parameters" are set to either male or female status, while in *C.elegans* the natural states are male or self-fertilizing hermaphrodite.

In *Drosophila* the processes of somatic sexual development and dosage compensation are regulated through a key X-linked gene, *Sex-lethal* (*Sxl*), whose activity must be on in females and off in males. From mid-embryogenesis and throughout adult life *Sxl* is transcribed in both sexes, however *Sxl* transcripts in males contain a male specific exon carrying an in-frame stop codon, such that male transcripts cannot generate a functional protein (Bell.LR *et al.*, 1988). In contrast female specific transcripts give rise to a full length protein containing an RNA binding domain. This protein subsequently controls the splicing pattern of its own transcript (generating more of the female specific form) and of those from other key genes in the cascade.

Whereas *Sxl* maintenance is regulated at the level of RNA processing, its initiation appears to be transcriptionally regulated. Very early during embryogenesis *Sxl* transcripts, which differ in structure to those detected in later life, are found transiently only in female embryos. These early transcripts appear to initiate at a site distinct from that of later transcripts. This has led to a model of *Sxl* regulation. In females *Sxl* transcriptional activation may occur from a female specific promoter which allows a functional protein product to be generated (Salz *et al.*, 1989). Later identical primary transcripts from a constitutive promoter are produced in both males and females, but only the transcripts in females would be productively spliced due to the antecedent presence of active *Sxl* protein. Thus the transient regulation of *Sxl* at the level of transcription initiates a positive feedback loop which ensures that *Sxl* remains active in females even once the initiating signal has disappeared (Bell *et al.*, 1991).

Recent work has shed light on how the primary sex-determining signal in *Drosophila*, the X:autosome (X:A) ratio, operates at the molecular level to initiate *Sxl* transcription only in females (Parkhurst *et al.*, 1990). Two X-linked genes,

sisterless-a (*sis-a*) and *sisterless-b* (*sis-b*) are thought to be numerator elements responsible for measuring the X:A ratio (Cline, 1988). *sis-a* maps within a 20kb region containing the *achaete-scute T4* and *T7* (*AS-C T4*, *T7*) genes. Mutations at *AS-C T4* affect sex determination suggesting that this gene is equivalent to *sis-a* (Torres & Sanchez, 1989). *AS-C T4* encodes a protein containing a helix-loop-helix (HLH) domain, which has been implicated in protein-protein dimerization and DNA binding. Parkhurst *et al.* (1990) have proposed that *AS-C T4* forms heterodimers with another HLH containing protein, daughterless (*da*) (Cronmiller *et al.*, 1988). These heterodimers are responsible for the initiation of *Sxl* transcription. As *da* is a maternal product it is present in equal amounts in male and female embryos. However males having only one X chromosome produce only half as much *AS-C T4* protein as females, consequently not enough *da*: *AS-C T4* dimers are present in males to initiate *Sxl* transcription.

One explanation for the existence of a mechanism for *Sxl* maintenance independent of the initiating signal is the fact that once dosage compensation has been established this signal, which is comprised of X-linked elements such as *sis-a*, may alter. Thus a cell which continuously reads the X:A ratio may receive ambiguous signals.

In addition to *Sxl*, several of the other regulatory interactions involved in *Drosophila* sex determination entail differential splicing of primary transcripts that are common to both sexes (Hodgkin, 1989).

In *C.elegans* three different genes seem to play a similar pivotal role in sex determination to that of *Sxl* in *Drosophila*. *sdc-1* and *sdc-2* (Villeneuve & Meyer, 1987; Nusbaum & Meyer, 1989), control both sexual differentiation and dosage compensation, although the contribution of each gene to these processes differs, as mutations in the former have less extreme effects than mutations in the latter. A third gene, known as *xol-1* is a negative regulator of *sdc-1* and *sdc-2* (Miller,LM *et al.*, 1988). *xol-1* may act in one of two possible fashions. (i) It may respond directly to the X:A ratio, in turn controlling the expression of the *sdc* genes or (ii) it may be necessary for *sdc* gene down-regulation in the presence of other factors

which respond directly to this ratio. In the latter case its action would be comparable to that of *da* in *Drosophila.*

In contrast to *Drosophila,* sex determination in *C.elegans* seems to be controlled at the transcriptional level rather than by alternative splicing (Hodgkin, 1987). The involvement of an autoregulatory step in *C.elegans* sex determination, as is the case for *Drosophila,* is unknown.

Many invertebrates exhibit sexual states which appear to be determined by environmental factors. The females of the worm *Bonellia viradis* live in small burrows and have a long proboscis which protrudes for food collection. Male worms are considerably smaller than the females which they parasitise. If a larva comes into contact with the proboscis of an adult female it will become a male. Larvae that fail to reach a female develop into females themselves (Bacci, 1965). Experiments in which larvae are cultured with and without an exogenous proboscis indicate that this effect may be mediated by a substance secreted by the proboscis. However the finding that some males are produced even in cultures without a proboscis and some females in cultures with a proboscis, suggests that some intrinsic factors may exist such as a genetic predisposition to one sex or the other (Leutert, 1975).

In the aphid family where an asexual form exists in addition to males and females, at least three factors seem to be important for sex determination: daylength, crowding and nutritional status (Richards & Davis, 1977).

Environmental factors are also critical for sex determination in some vertebrate species. Temperature-dependent sex determination has been observed for reptiles such as lizards and crocodiles. Ferguson & Joanen (1982) have shown that the sex of *Alligator mississippiensis* depends on the temperature of egg incubation. At temperatures ≤30°C only females are produced while temperatures ≥34°C yield only males. The eggs exhibit a defined temperature-sensitive period after which the sex of the embryo is fixed (Deeming & Ferguson, 1988).

Some species of fish are also subject to temperature-sensitive sex determination. *Rivulus marmoratus* is particularly interesting, as apart from the primary sex-determining effect of temperature during embryonic development, young adult hermaphrodites can also undergo conversion to males under the influence of temperature (Harrington, 1975) thus differing in the fixed temperature-sensitive period exhibited by *Alligator mississippiensis*.

In birds, as might be expected for a homeothermic species, environmental sex-determining mechanisms have not been described. The female sex is heterogametic having a ZW karyotype while males are ZZ lacking the W chromosome (Bloom, 1974). Unfortunately, as diploid birds with ZO or ZZW karyotypes have not been observed, it is not known whether the W chromosome acts as a dominant female determinant, or if Z chromosome dosage determines sex. Abdel-Hameed (1971) has described 13 intersex chickens which were triploid/ZZW. Although their gonads were malformed they appeared to be composed only of testicular tissue. This argues against the W chromosome being a strong female determinant, although the triploid status of these chickens makes these results inconclusive.

1.2 Mammalian sex determination

In mammals, males and females differ in sex chromosome constitution: males normally have a single X and Y chromosome, while females normally have two X chromosomes. Initially it was assumed that sex in mammals would depend on X chromosome dosage as in *Drosophila* and *C.elegans*. However with improved techniques for karyotyping, the investigation of individuals with abnormal sex chromosome constitutions refuted this hypothesis. Individuals with a single X chromosome (XO) develop as females (Ford,CE *et al.*, 1959; Welshons & Russell, 1959), although in humans 97% of such conceptions are inviable and those that do survive to term suffer from a range of stigmata, collectively known as Turner syndrome (Turner, 1938). In contrast, individuals with a normal Y chromosome, regardless of the presence of supernumery X chromosomes (XY, XXY XXXY, XXXXY) develop as males (Jacobs & Strong, 1959; Cattanach, 1961). The XXY karyotype in humans is referred to as Klinefelter syndrome, in which patients suffer

from infertility and tend to be mildly educationally subnormal. From this evidence it was concluded that sex in mammals is determined by the presence or absence of the Y chromosome, which acts as a dominant male determinant.

One distinctive feature of mammals is that sexual differentiation at the level of secondary (i.e. extra-gonadal) sexual characteristics appears to be controlled not by direct gene action, but by hormones. This is in stark contrast to the development of *Drosophila* and *C.elegans* where the primary sex-determining signal seems to operate in each cell to control their sexual phenotype. The elegant experiments of Jost (1947) have shown that in the male these hormones are produced by the testis. Jost performed *in utero* castration of rabbit fetuses. Removal of the fetal testis resulted in female development, whilst removal of fetal ovaries did not alter the sexual fate of the animal. This evidence clearly demonstrated that the male phenotype derives from the presence of testes. In the absence of testes, female development occurs and this can therefore be thought of as the default developmental pathway. Thus the process of mammalian sex determination can be simplified to the developmental decision to produce testes (Jost *et al.*, 1973). The resultant hypothesis concerning testis formation is as follows: the Y chromosome directs the formation of testes by coding for a gene (or a number of genes) which has been called the testis-determining factor (*TDF*) in man and testis-determining gene-Y chromosome (*Tdy*) in mouse. It is assumed that this locus operates by regulating the expression of genes elsewhere in the genome to bring about testis differentiation. The testis then directs the acquisition of the mature male phenotype via its hormonal output.

One intriguing exception to this view of sexual differentiation has come from the study of marsupials. O *et al.* (1988) have reported that in the Tammar Wallaby, *Macropus Eugenii*, some sexual dimorphism precedes any morphological differentiation of the gonads. They conclude that the sexual differentiation of the scrotum, mammary glands, gubernaculum, process vaginalis and pouch cannot be controlled by the gonads in this species, suggesting a difference in at least part of the sex-determining process between eutherian and marsupial mammals. Recently, Sharman *et al.* (1990) have described a number of intersex marsupials. In three

mosaic animals carrying an XO component but no Y chromosome (i.e. XO/XX and XO/XX/XXX), ovaries and normal female reproductive tracts were found. Scrotal development was also observed in all cases. In one case no pouch or mammary glands were found, while in the others mammary tissue and a partial pouch were present in addition to the scrotum. Two XXY intersex marsupials and one XXY/XY/XX mosaic intersex marsupial were also described, which had male internal reproductive tracts and testes, complete pouches and mammary glands, but lacked a scrotum. This data confirms an earlier finding (Sharman *et al.*, 1970) that gonadal sex in marsupials is determined by the presence or absence of the Y chromosome. In addition, these authors have proposed that the differentiation of the pouch, mammary glands and scrotum depends on X chromosome dosage (or the X:A ratio). Thus the presence of two X chromosomes leads to development of mammary glands and the pouch, while a single X chromosome leads to development of the scrotum. Examination of a larger number of sex chromosome aneuploid marsupials will be necessary to confirm these data.

1.3 The biology of gonadal and secondary sexual differentiation

Conventionally sex determination refers to the events which fix the nature of the gonad (i.e. testis or ovary). All subsequent events influencing sexual form are referred to as sexual differentiation. The following section briefly describes the different cell types which differentiate in the developing gonads and the morphological changes which occur during this differentiation event. This is followed by a description of the hormonal control of secondary sexual differentiation.

The gonadal anlagen form in the tissue covering the coelomic side of the mesonephros (Everett, 1943; Witschi, 1948; Pelliniemi, 1975, 1976). In their undifferentiated (indifferent) form the gonads are thought to be comprised of the following three somatic cell lineages which can differentiate into either testis- or ovary-specific cell types. (i) The "supporting cell precursors" will give rise to either Sertoli cells in the testis, or follicle cells in the ovary. The idea of a common precursor cell lineage for these two cell types was suggested by Vigier *et al.* (1984).

Both types of cell carry a common lineage-specific surface antigen (Ciccarese & Ohno, 1978), form a diffusion barrier shielding germ cells from the somatic environment and release their secretion products into the fluid which bathes germ cells during normal maturation. In addition, both produce anti-Müllerian hormone (AMH), Sertoli cells during embryogenesis and follicle cells in the adult ovary (Tran *et al.*, 1977; Vigier *et al.*, 1984; Donahoe *et al.*, 1987; Münsterberg & Lovell-Badge, 1991). (ii) A "steroid cell precursor" lineage is thought to give rise to steroidogenic interstitial (Theca) cells in the ovary and testosterone secreting Leydig cells in the testis. (iii) The connective tissue gives rise to peritubular myoid cells and the *tunica albuginea* in males and stromal cells in females. The connective tissue framework of the fetal testis is more complex than that of the fetal ovary and provides a pathway for testosterone export.

The primordial germ cells do not originate in the gonadal anlagen but are first distinguishable at the posterior end of the primitive streak in a 7.25 days post *coitum* (*dpc*) embryo (Ginsburg *et al.*, 1990). Over the following three or four days they migrate along the hind gut into the developing genital ridge.

The origin of the somatic cells which comprise the indifferent gonad is uncertain. However ultrastructural studies, using electron microscopy, provide morphological evidence that both the coelomic epithelium and mesonephros, which are in intimate contact with the gonadal anlage, may be sources of such cells (Satoh, 1985; Smith & Mackay, 1991). In the fetal testis, Sertoli cells can be identified at 12.5*dpc* on the basis of their association with germ cells, and alignment into cord like structures (testis cords). They also begin to deposit a basement membrane at this stage (Taketo *et al.*, 1985). At the same time the connective tissue takes on a specific organisation; prominent blood vessels grow between developing cords (Pelliniemi, 1975) and peritubular myoid cells accumulate at the boundaries of the testis cords. Once the germ cells in the fetal testis become surrounded by pre-Sertoli cells in testis cords, they undergo a last prenatal mitotic division and then arrest in the G_1 (or G_0) phase of the cell cycle in the prospermatogonial stage (McLaren, 1984). After birth mitotic divisions are resumed as the germ cells begin to undergo spermatogenesis (Hilscher & Hilscher, 1976).

By contrast, developing ovaries initially lack any clear features which distinguish them from the indifferent gonads, apart from changes in overall shape. A few days later follicle cells aggregate around germ cells to form follicles (Torrey, 1945; Gillman, 1948; Jost, 1958, 1970). The germ cells enter meiotic prophase and become arrested at the dictyate stage (Borum, 1961).

The differentiation of the testes has a profound effect on the subsequent development of the genital tract, which is mediated by hormonal secretions. Initially the genital tract of both male and female embryos consists of unipotential Wolffian and Müllerian ducts and of bipotential external genitalia (Jirasek, 1971; Dyche, 1979). In females the Müllerian duct will give rise to the oviducts, uterus and cervix, however, in males, Müllerian duct regression occurs under the influence of the hormone AMH secreted by differentiating Sertoli cells (Jost, 1953). The Leydig cells of the testis produce testosterone, which is responsible for stabilizing the Wolffian ducts and inducing them to differentiate into the epididymis, *vas deferens* and seminal vesicles (Josso, 1970). In the absence of this hormonal influence (i.e. in the female) the Wolffian duct quickly regresses. Testosterone also induces the bipotential primordia of the external genitalia to develop into penis shaft, scrotum and glans penis. Again, in the absence of testosterone, female genitalia develop.

Sex differentiation has mainly been analysed by experiments designed to change the hormonal status of the embryo. However a number of mutations exist in which the gonadal and phenotypic sex of the individual have been dissociated. The molecular defect in *Testicular feminization* (*Tfm*) is a deficiency of androgen receptors (Attardi & Ohno, 1974; Ohno, 1979; Migeon *et al.*, 1981). In males, despite normal secretion of testosterone by the testes, the Wolffian ducts regress and female external genitalia develop as these tissues are not able to respond to the hormonal signal. The oviducts, uterus and cervix also fail to develop as AMH secreted by the testes still results in Müllerian duct regression. XY male individuals deficient for the enzyme 5α-reductase, which converts testosterone to dihydrotestosterone, have male internal genital tracts but female external genitalia. This indicates that dihydrotestosterone is the active androgen for masculinization of the external

genitalia, while testosterone itself seems to act on the Wolffian ducts (Imperato-McGinley *et al.*, 1974; Wilson, 1985; Andersson *et al.*, 1991). Another condition which affects a different aspect of male sexual differentiation is the persistent Müllerian duct syndrome (PMDS). This syndrome may be caused by a defect in the AMH gene or its receptor (Guerrier *et al.*, 1989). In males suffering from PMDS the Müllerian ducts fail to regress, giving rise to female internal structures. However, through the action of testosterone normal male external genitalia are formed although the testes sometimes fail to descend. Similarly, in females, masculinization of external genitalia can occur in cases such as 21-hydroxylase deficiency (adrenal hyperplasia) in which faulty steroid metabolism in the fetal adrenal cortex misdirects steroid production into an androgenic pathway (Lee & Gareis, 1975).

While the process of sex differentiation is fairly well understood, at least at the anatomical and physiological levels, until recently our knowledge of the biochemical basis of sex determination was sparse. Efforts to identify *TDF/Tdy* have been particularly protracted. Information regarding the chromosomal location of this locus in both man and mouse and the expected biological properties of *TDF/Tdy* is considered in the following sections.

1.4 Genetic definition of *TDF* in man

Advances in the genetic analysis of the human Y chromosome have led to a progressively more refined definition of the location of *TDF*. The human Y chromosome is divided into two parts, the pseudoautosomal region and the Y-specific region (see figure 1, human Y). The pseudoautosomal region lies at the tip of the short arm and is homologous with a region of the X chromosome (Pearson & Bobrow, 1970; Chen & Falek, 1971; Ellis & Goodfellow, 1989). It is this region which is responsible for the pairing of the X and Y chromosomes at the tips of their short arms, which can be observed cytogenetically, during male meiosis (Solari, 1980). This pairing event is critical to ensure proper segregation of the sex chromosomes during spermatogenesis and is also responsible for the 1:1 sex ratio found in mammals. Molecular evidence suggests that an obligatory recombination

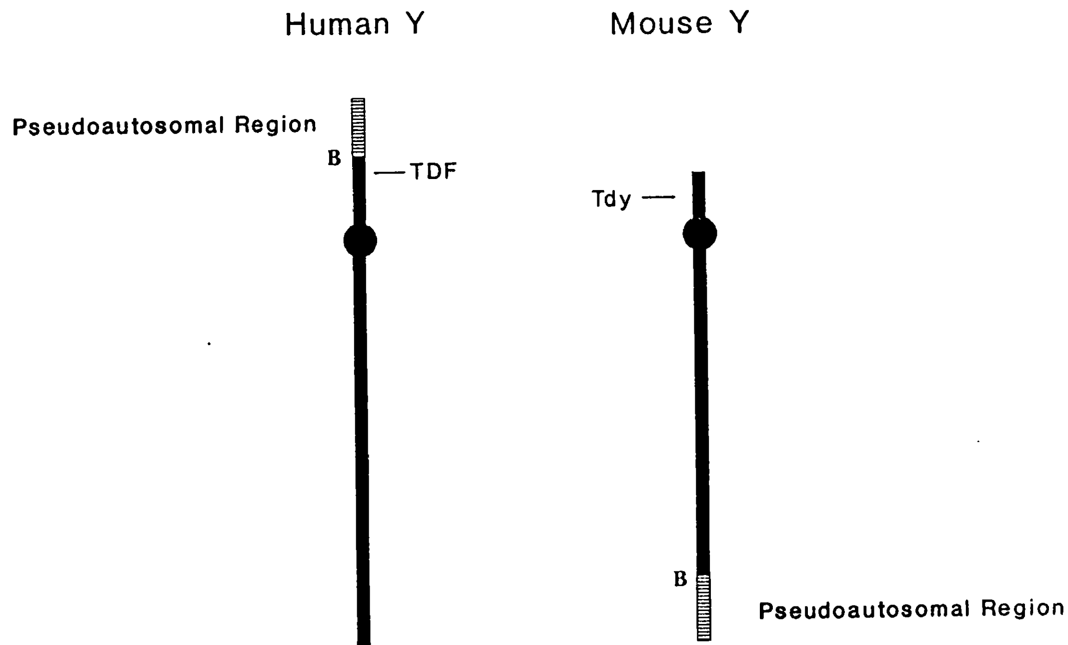


Figure 1. The human and mouse Y chromosomes. A schematic representation of the human and mouse Y chromosomes is shown in order to highlight differences in their overall structure. The Y-specific portions are shown in black. The pseudoautosomal regions (at the end of the short arm of the human Y chromosome and at the end of the long arm of the mouse Y chromosome) are hatched. The position of the pseudoautosomal boundary is shown (B). The positions of *TDF* and *Tdy* are indicated.

event occurs within the pseudoautosomal region during meiosis. Sequences at the short arm telomeres recombine with respect to X and Y specific sequences in 50% of meiotic products implying that telomeric sequences crossover in every XY pair (Cooke *et al.*, 1985). The position of this obligatory recombination event within the pseudoautosomal region is variable. This leads to sequences showing a gradient of sex linkage, depending on their position within the pseudoautosomal region. This has allowed the construction of a meiotic map of pseudoautosomal DNA markers (Goodfellow,PJ *et al.*, 1986; Rouyer *et al.*, 1986). Two genes are known to map to this region: the receptor for the granulocyte-macrophage colony stimulating factor (Gough *et al.*, 1990) and the structural gene for the cell surface antigen 12E7, known as *MIC2* (Darling *et al.*, 1986). These sequences exhibit the expected pattern of partially sex-linked inheritance.

The Y-specific region is separated from the pseudoautosomal region by the pseudoautosomal boundary. This is the normal limit to recombination and the locus at which Y-specific sequences begin. The boundary has been isolated from a human X and Y chromosome and analysed at the molecular level. An abrupt change in sequence homology between the sex chromosomes begins at the site of an *Alu* insertion element present on the Y chromosome. This site has therefore been postulated to be the boundary itself (Ellis *et al.*, 1989). The *Alu* insertion is thought to have been a relatively recent event, as it is absent from the pseudoautosomal boundary of Old World Monkeys, but yet is present in all great apes. Thus it must have occurred after the divergence of these two lineages. Nevertheless, the position of the boundary is the same in both Old World Monkeys and great apes, suggesting that the *Alu* repeat element was inserted at the pre-existing boundary (Ellis *et al.*, 1990).

Several individuals have been found who have visible deletions and translocations of the Y chromosome (Buhler, 1980; Davis,RM, 1981; Goodfellow,PN *et al.*, 1985). Correlation of sexual phenotype with these Y-chromosome structural abnormalities suggested that *TDF* was located in a region around the centromere. Further progress towards defining the location of *TDF* has been made from the study of "exceptional" males with an apparently normal female XX karyotype. Such

individuals occur with a frequency of about 1 in 20000 males (de la Chapelle, 1981). The majority have a normal phenotype, with the exception of the testes which are small. This size difference is probably due to an absence of spermatogenesis. In fact all XX males are sterile (the causes of sterility in XX males are discussed in section 1.5). The genetic basis of this form of sex reversal was correctly explained by the far-sighted hypotheses of Ferguson-Smith (1966). He suggested that XX males may be caused by the inheritance of undetected Y-chromosome sequences which include *TDF*. He also suggested that this could occur if the meiotic recombination that normally takes place within the pseudoautosomal region were to occur abnormally within Y-specific sequences, so transferring Y-derived material to an X chromosome (see figure 2). Subsequently, with the availability of Y-specific cloned DNA probes, it has been shown that about 80% of XX males have inherited Y-derived sequences (Guaellaen *et al.*, 1984; Page, 1986) and that these sequences are found at the tip of the X chromosome short arm as predicted by Ferguson-Smith's abnormal XY interchange model (Anderson,M *et al.*, 1986; Buckle *et al.*, 1987). The mechanism of this abnormal interchange is unknown, but the finding in one patient that the translocation breakpoint on both the X and Y chromosomes was an Alu repeat, suggests that some homology may be required for the recombination event (Rouyer *et al.*, 1987). The amount of Y-derived sequences transferred in different XX males is variable which has allowed the construction of a deletion map of the Y chromosome (Affara *et al.*, 1986; Vergnaud *et al.*, 1986). From this map *TDF* has been placed in the distal region of the Y-specific part of the short arm adjacent to the pseudoautosomal region.

1.5 Genetic definition of *Tdy* in mouse

In mouse the sex-reversed mutation (Sxr) has helped to define the location of *Tdy*. Sxr was first described by Cattanach *et al.* (1971) as an autosomally inherited form of sex reversal in which sterile XX males, XY carrier males and normal males and females are formed. Evidence that this interpretation was incorrect came from the work of Singh & Jones (1982). These workers had isolated a female specific satellite DNA sequence from the banded krait snake (*Bkm*, banded krait minor satellite) (Singh *et al.*, 1980). These snakes show female heterogamety, as also

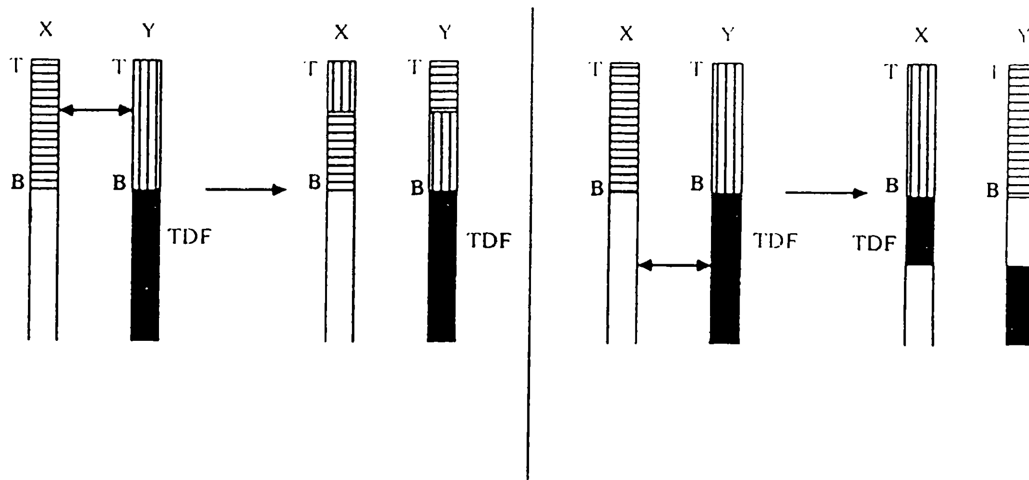


Figure 2. Normal and abnormal X-Y interchanges (from Weissenbach *et al.*, 1987). The ends of the human X and Y chromosome short arms are shown schematically. The X- and Y-specific regions are shown as solid white and solid black, respectively. The X and Y pseudoautosomal regions are indicated by horizontal and vertical hatching respectively. (T) denotes the telomere while (B) marks the position of the pseudoautosomal boundary. In the left panel a normal crossover event within the pseudoautosomal region is indicated by arrows. This leads to exchange of homologous blocks of chromosome. The right hand panel indicates an aberrant crossover occurring within the sequence specific parts of the chromosome. This can lead to an X chromosome carrying *TDF* on its short arm. The majority of XX males possess such an X chromosome. A Y chromosome which is deleted for *TDF* can also be generated by this mechanism. Some cases of XY females are due to inheritance of such a Y chromosome.

described for birds. Southern analysis indicated that *Bkm* homologous sequences in the mouse were Y linked. In addition, DNA from an XX sex-reversed male (XXSxr) showed the same pattern of *Bkm* hybridization as a normal XY male. This strongly suggested that Sxr was a fragment of the Y chromosome (presumably encompassing *Tdy*). By using the same sequence as a probe for *in situ* hybridization it was possible to demonstrate the presence of a concentration of *Bkm* related sequences at the centromere of the normal Y chromosome.

Roberts *et al.* (1988) and McLaren *et al.* (1988) have more accurately defined the region of the Y chromosome that gave rise to Sxr. Roberts *et al.* (1988) made use of a single copy sequence which recognises a Y-specific band on Southern blots also present in DNA from XXSxr male mice. *In situ* hybridization using this sequence as a probe has shown that the peak of hybridization on the normal Y chromosome in fact corresponds to the very small short arm of the Y (see figure 3a). An XY carrier male (XYSxr) showed two peaks of signal, one defining the Y short arm and the other at the telomere of the long arm of the Y chromosome. As the mouse pseudoautosomal region is at the end of the long arm (see figure 1, mouse Y), in contrast to its location on the human Y chromosome, this pattern is consistent with the presence of Sxr distal to the Y chromosome pseudoautosomal region in XYSxr carrier males. This suggests a simple explanation for the origin of Sxr involving the duplication and translocation of a large part (or all of) the Y chromosome short arm to the distal tip of the Y chromosome pseudoautosomal region (see figure 3b). During meiosis in a mouse carrying such a YSxr chromosome, the obligatory crossover within the X and Y pseudoautosomal regions will transfer Sxr to the distal tip of the X chromosome. As Sxr, and consequently *Tdy*, are now X-linked, mice inheriting such an XSxr chromosome along with a normal X chromosome will be sex reversed (see figure 3c). In addition to testis-determining information (*Tdy*), Sxr, also includes the structural gene for H-Y antigen or a gene controlling its expression (*Hya*) (Simpson,E *et al.*, 1981). H-Y antigen is a male specific minor histocompatibility antigen (Eichwald & Silmser, 1955; Billingham & Silvers, 1960), found also to be expressed in XXSxr male mice.

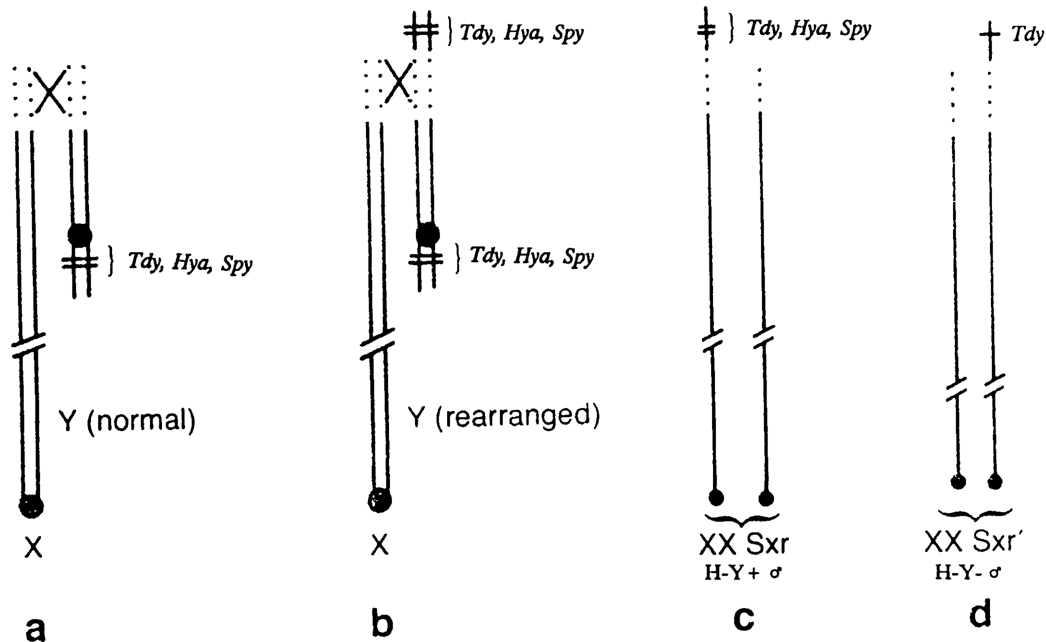


Figure 3. Sex reversal in the mouse (from McLaren, 1988a). a, In a normal male mouse the genes for testis determination (*Tdy*), H-Y antigen expression (*Hya*) and spermatogenesis (*Spy*) are located on the short arm of the Y chromosome. The pseudoautosomal region, which pairs with the X chromosome is at the distal end of the long arm (....). An obligatory crossover occurs within this region. b, In the rearranged Y chromosome the region containing *Tdy*, *Hya* and *Spy* (the Sxr region) has been copied and transposed to the distal end of the long arm. c, Sxr can now be transferred to an X chromosome by crossing-over, generating XXSxr males that express H-Y antigen. d, Sxr' is a deleted version of Sxr which has retained *Tdy*. XXSxr' males do not express H-Y antigen and lack *Spy*.

XXSxr males differ from normal males not only in that they lack most of the Y chromosome but also in having two X chromosomes. In order to fully assess the phenotypic effects of these Y-chromosome deficiencies in sex reversed males, it was necessary to eliminate the second variable. This is particularly important when looking at spermatogenesis as the presence of two X chromosomes is incompatible with male germ cell survival beyond the perinatal period (Lyon, 1970; McLaren, 1983). Sxr males with single X chromosomes (XOSxr) were first described by Cattanach *et al.* (1971). Viable XOSxr mice can be produced by mating XO females with XYSxr carrier males. The fate of XO germ cells has been studied in mosaic male mice where an XO component has arisen by mitotic non-disjunction. XO germ cells present in an XO/XY/XYY mosaic male mouse show a severe spermatogenic impairment with only a few germ cells reaching early meiotic prophase (Levy & Burgoyne, 1986). In contrast, all stages of spermatogenesis were represented in the testes of XOSxr mice, although the later stages are severely depleted and the few sperm produced are often diploid and/or abnormal, such that the mice are sterile. These findings led Burgoyne *et al.* (1986) to suggest that Sxr carries a gene involved in spermatogenesis which they have called *Spy*. The failure of XOSxr germ cells to fully progress to functional sperm indicates that other Y encoded sequences besides *Spy* are required for spermatogenesis (Burgoyne, 1991). In addition, Miklos (1974) and subsequently Burgoyne (for example see Sutcliffe *et al.*, 1991) have proposed that the presence of unpaired sex chromosomes is also deleterious to spermatogenesis (and oogenesis).

In 1984 McLaren *et al.* discovered a variant of Sxr termed Sxr' that had retained *Tdy* (XXSxr' mice were male) but had lost the *Hya* gene, as XXSxr' mice did not express H-Y antigen. When XOSxr' germ cells were analysed they were found to have the same severe spermatogenic block that XO germ cells suffer from. This indicated that Sxr' lacked *Spy* as well (Burgoyne *et al.*, 1986). Southern analysis using a molecular probe that detected several loci mapping to Sxr suggested that Sxr' was a deleted version of Sxr (Bishop,CE & Mitchell, 1991). *Tdy* is therefore the only gene function known to be encoded by Sxr' (see figure 3d).

1.6 Mode of action of *Tdy*

Other predictions can be made about *TDF/Tdy* besides those concerning its chromosomal location. The developmental time and site of action of *Tdy* and the cell types involved in its expression are considered in this section.

In a male mouse embryo, the series of events which result in the formation of a morphologically recognisable fetal testis begins at around 11.5*dpc*, when the genital ridge is still morphologically identical in males and females. 24 hours later (at 12.5*dpc*) the first signs of testis development are clearly visible (for a description of these events at the cellular level see section 1.3). As this differentiation event must be initiated by *Tdy*, it is reasonable to suppose that *Tdy* will be expressed in the genital ridge at around this time. However, this does not preclude the possibility that *Tdy* may continue to be expressed after 12.5*dpc* or that its action may occur before this time.

Various experiments have been done which address the question of which gonadal cell type is likely to be responsible for *Tdy* expression. Germ cells can be eliminated from rat fetuses by busulfan treatment (Merchant, 1975). However, testes still form normally in treated males, suggesting that germ cells are not required for testis determination. Mutations at two loci in the mouse have lent additional weight to this hypothesis: *steel* (*Sl*) and *white spotting* (*W*) are both pleiotropic mutations in which the development of neural crest derived melanocytes, primordial germ cells and haematopoietic stem cells is disrupted. This has effects on coat colour, fertility and causes anaemia. Recent work has shown that *W* encodes the *c-Kit* proto-oncogene, a tyrosine kinase membrane receptor and that *Sl* encodes the ligand for this receptor (Gillman, 1948; Mintz & Russell, 1957; Chabot *et al.*, 1988; Geissler *et al.*, 1988; Anderson,DM *et al.*, 1990; Copeland *et al.*, 1990; Flanagan & Leder, 1990; Huang *et al.*, 1990; Martin *et al.*, 1990; Williams *et al.*, 1990; Zsebo *et al.*, 1990a, b). In both cases in the homozygous state proliferation and perhaps also migration of primordial germ cells is impaired so that virtually none are able to colonise the genital ridge. Once again, this does not affect the differentiation of the testis (Cattanach, 1978). It now seems clear that the

developmental fate of germ cells is itself subject to environmental signals and does not determine testis differentiation. In fact, although the ability of a germ cell to complete gametogenesis does depend upon its sex chromosome constitution (McLaren, 1988b), the decision to embark upon oogenesis or spermatogenesis does not. For instance in a fetal testis XO germ cells can reach the early stages of spermatogenesis (see section 1.5), and in experimental situations such as female XX↔XY mouse chimaeras XY cells can undergo oogenesis (Ford,CE *et al.*, 1975). It has also been shown that in males, XY germ cells which fail to reach the gonad, and end up in the nearby adrenal primordium, differentiate as oocytes at their ectopic location (Zamboni & Upadhyay, 1983; McLaren, 1984). This suggests that the oocyte may be the default germ cell pathway and the male germ cell pathway only occurs in the presence of the appropriate signals provided by the testicular environment (where germ cells are enclosed by Sertoli cells in testis cords). It is interesting to note at this point that the concept of female development being the default developmental pathway has recurred in three separate contexts: (i) ovary determination (section 1.2), (ii) female sexual differentiation (sections 1.2 and 1.3) and (iii) germ cell differentiation (see above).

If germ cells are not required for testis determination, *Tdy* must be expressed in the somatic portion of the genital ridge. What somatic cell type is responsible for this expression? Various lines of evidence suggest that it may be the pre-Sertoli cells. The initial steps of testis differentiation have been studied in rat fetuses. The first event observed, which distinguishes the testis from the indifferent gonad, is the appearance of large clear cells which aggregate and encompass the germ cells into the forming seminiferous tubules. In the electron microscope characteristic interdigitations are seen between these cells (Jost, 1972; Jost *et al.*, 1973; Magre & Jost, 1980; Jost & Magre, 1988). The morphological relationship of these cells to the germ cells indicates that they are Sertoli cells. This has led to the suggestion that pre-Sertoli cells are the first male-specific cell type to differentiate in the fetal testis. This differentiation event may therefore be the first consequence of *Tdy* action. One interpretation of this is that *Tdy* is expressed within this lineage to cause its differentiation.

The analysis of sex chimaerism in mice has also provided some insight into the expression of *Tdy*. Burgoyne *et al.* (1988a) have examined the contribution of XX and XY cells to different gonadal lineages in adult XX↔XY male mouse chimaeras. Such chimaeras, which are made by aggregation of 8 cell embryos, will develop as males if the contribution of XY cells in the gonad is ≥ 25% (Eicher *et al.*, 1980; Burgoyne *et al.*, 1988a; Palmer,S & Burgoyne, 1991). Burgoyne *et al.* (1988a) used parental embryos carrying electrophoretic variants of glucose phosphate isomerase (GPI), the gene for which is X-linked. Chimaeric testes were fractionated into Leydig cells, Sertoli cells and *tunica albuginea*. By assaying these for GPI activity, the authors showed that both XX and XY cells were present amongst Leydig cells, peritubular myoid cells and the vascularized connective tissue. However the Sertoli cells were found to be exclusively XY. This led Burgoyne *et al.* (1988a) to postulate that *Tdy* acts cell-autonomously within the supporting-cell lineage, directing its differentiation to Sertoli cells, and that the commitment of the other components of the gonad to the male pathway is directed by Sertoli cells without further involvement of *Tdy*. Thus XX cells can contribute to all somatic cell types in the gonad with the exception of Sertoli cells. One criticism of the conclusions drawn by Burgoyne *et al.* (1988a) is that they are based on the study of postnatal testes which may not accurately reflect the situation in fetal XX↔XY chimaeric testes when Sertoli cells first form. Palmer,SJ & Burgoyne (1991a) have therefore undertaken a more extensive study of male XX↔XY chimaeras in which one component carried a transgene consisting of a single insertion of 1000 copies of ß-globin, allowing sectioned and air-dried material from fetal and adult mouse testes to be analysed. In agreement with the previous results for adult testis, no prejudice was found against XX cells contributing to peritubular myoid or Leydig cells. By contrast most Sertoli cells were found to be XY. Nevertheless, some XX Sertoli cells were found to be present in fetal testes. The mean XX contribution to Sertoli cells was 10% in four XX↔XY chimaeras in which the mean XX contribution to non-testicular cells was about 50%. The incidence of XX Sertoli cells in adult testes was even lower (less than 3%), which may explain how it was missed in the earlier study. The presence of XX Sertoli cells in an adult mouse carrying both XX and XY components was previously inferred by Singh *et al.* (1987) and has been reported for a series of adult XX↔XY chimaeras by Patek *et al.* (1991). The

existence of XX Sertoli cells in these studies was deduced by their morphology and location within testis cords. As this identification has not been verified by testing for Sertoli cell markers such as AMH, the possibility still remains that these cells may not in fact be Sertoli cells.

Despite these new findings there was still a strong XY bias for Sertoli cells not found in any of the other cell types in the fetal testes. This seems to strengthen the original conclusions of Burgoyne *et al.* (1988a) that *Tdy* acts cell-autonomously in this lineage. The presence of a small number of XX Sertoli cells can be explained by postulating that at some point between the expression of *Tdy* and the formation of Sertoli cell cords, there may be a step that can locally recruit XX cells. As these XX Sertoli cells are seen in the early embryonic testis, Palmer,SJ & Burgoyne (1991a) argue that they co-differentiate with XY Sertoli cells. Patek *et al.* (1991) have attempted to explain this codifferentiation by postulating that both XY and XX Sertoli cells may be formed concurrently by a single non-cell-autonomous mechanism, but that the commitment of the supporting cell lineage to form Sertoli cells depends on the presence of a threshold level of XY cells in the supporting cell progenitors in which *Tdy* must act. The fetal XX Sertoli cells described by Palmer,SJ & Burgoyne (1991a) were observed at 13.5*dpc*. In the mouse the first signs of Sertoli cell differentiation are observed at least a day earlier (when testis cords are already visible). Thus the secondary recruitment of XX Sertoli cells, as a consequence of the prior differentiation of XY cells cannot be excluded.

XX Sertoli cells have been observed in other situations, in which they are thought to have undergone transdifferentiation from the follicle or pre-follicle cell phenotype. For instance, fetal ovary grafting under the kidney capsule of male and female hosts (Taketo & Merchant-Larios, 1986) and treatment of cultured fetal rat ovaries with AMH (Vigier *et al.*, 1987) both lead to the formation of cord like structures containing cells which by morphological criteria are thought to be Sertoli cells. The freemartin effect (Jost *et al.*, 1973), occurs in cattle when a female calf is joined by vascular anastomosis of the placenta to a male twin. This leads to regression of the Müllerian ducts in the female. In addition the ovaries cease to grow, become depleted of germ cells and develop seminiferous tubules containing

Sertoli cells. Regression of the Müllerian ducts is caused by exposure of the female twin to AMH present in the bloodstream of the male. Because of the masculinizing effects of AMH *in vitro* (see above) it is thought that this may also be responsible for the ovarian freemartin effect (Vigier *et al.*, 1987).

In all the cases described above, masculinization of ovarian tissue is preceded by oocyte loss. In fact the presence of oocytes seems to be critical for maintenance of the follicle cell phenotype. Thus in contrast to testis development, which can occur in the absence of germ cells (see section 1.6), ovary differentiation is profoundly influenced by the presence or absence of germ cells (Taketo-Hosotani & Merchant-Larios, 1985). The finding that AMH induces atrophy of oocytes *in vitro* (Vigier *et al.*, 1987), suggests that its masculinizing properties may be a consequence of its effect on oocytes. Behringer *et al.* (1990) have recently extended the functional study of AMH by constructing transgenic mice chronically expressing human AMH under the control of the mouse metallothionein promoter. As expected, the females lacked Müllerian duct derivatives. The germ cell component of transgenic ovaries was severely depleted after birth and testis cords containing Sertoli cells were seen. However, unlike the grafted ovaries described by Taketo-Hosotani & Merchant-Larios (1985) and freemartins, no Leydig cell differentiation was induced. In addition, the testis cords were transient, such that in most adult transgenic females no gonads could be found. The failure of Sertoli cells to survive or to induce Leydig cell differentiation suggests that Sertoli cells formed in female mice transgenic for the AMH gene, may not be fully competent. Nevertheless, these results reinforce a direct or indirect role for AMH in Sertoli cell differentiation in certain "abnormal" situations. However, AMH cannot be necessary for Sertoli cell differentiation as PMDS patients who lack AMH have apparently normal testes. The fact that masculinization of the gonads can occur in the cases described above where *Tdy* is absent suggests that redundant mechanisms of testis differentiation may exist in addition to that initiated by *Tdy*. Some of these redundant mechanisms may involve AMH.

The studies by Patek *et al.* (1991) on XX↔XY male mouse chimaeras have also provided information on Leydig cell function. Patek *et al.* (1991) have postulated

that while the Y chromosome may not be required for Leydig cell differentiation (in contrast to Sertoli cells), it may be necessary for normal Leydig cell function. They found a high incidence of seminiferous tubules with germinal failure in chimaeric testes. This correlated with a particularly high proportion of XX Leydig cells in these tubules, leading them to propose that the germinal failure is brought about by XX Leydig cell dysfunction. An alternative explanation proposed by Palmer,SJ & Burgoyne (1991a) is that the strain combination used in these experiments may not only favour XX Leydig cells but also XX germ cells. Germinal failure would then be attributable to loss of XX germ cells around birth.

A further prediction of the "cell-autonomous action of *Tdy*" model of sex determination is that in an XX↔XY female chimaera, follicle cells should be exclusively XX, as all supporting cell precursors containing a Y chromosome should be diverted to the Sertoli cell pathway. However, as many as 50% of the follicle cells in a number of XX↔XY female chimaeras have been shown to be XY (Ford,CE *et al.*, 1974; Burgoyne *et al.*, 1988b). Burgoyne *et al.* (1988b), suggested that these XY follicle cells could be formed by a "timing mismatch" mechanism in which *Tdy* action was pre-empted by the ovarian determination process initiated by the XX component of the chimaera. This model presupposes two properties of ovary determination. First, that there are active ovary-determining genes analogous to testis-determining genes and second that follicle cell development (unlike most cases of Sertoli cell development) is not cell autonomous but involves some form of inducing signal. Given the influence of germ cells on follicle cell maintenance and possibly differentiation (see above), this cell type is a good candidate as a source for the inducing signal. In chimaeras there is clearly the potential for a developmental mismatch between the two components, which may lead to ovary-determining signals occurring before *Tdy* has acted. In fact in some of the XX↔XY chimaeras studied, the XY component was derived from the AKR strain, which other evidence suggests may have a Y chromosome carrying a late acting allele of *Tdy* which induces testis differentiation later than many other strains (Eicher & Washburn, 1986 and see section 1.7.2). More recently Palmer,S & Burgoyne (1991) have analysed the chromosome constitution of follicle cells in XO/XY and XO/XY/XYY mosaic hermaphrodites. These mosaics arise through mitotic non-

disjunction of the Y, so there are no strain differences between the components. Some XY follicle cells are found in these mosaics. While the proportion of XY follicle cells was consistently low, their presence cannot be explained by the "timing-mismatch" theory. The authors speculate that in such a situation XY supporting cells are triggered to form fetal Sertoli cells by the action of *Tdy*, but if their numbers fall below a critical threshold, they may subsequently transdifferentiate into follicle cells under the influence of ovarian factors. McLaren (1991) has suggested that the requirement for a critical threshold of Sertoli cells may be a requirement for the presence of enough of these cells to be able to organise into testis cords. If this threshold is not reached cord formation will not occur. In the absence of this environment the germ cells will enter the female pathway and initiate meiosis (see section 1.3). The resulting oocytes may be responsible for the production of ovarian factors which cause Sertoli cell transdifferentiation to follicle cells. Once again transdifferentiation seems to be an intriguing property associated with the supporting cell lineage.

In summary, strong evidence exists to indicate that *Tdy* acts cell autonomously within the supporting cell precursors. In certain circumstances its action may be pre-empted by the ovary-determining mechanism or reversed, possibly by a transdifferentiation mechanism. Equally, both experimental and natural situations exist in which Sertoli cell and testicular differentiation occurs in the complete absence of *Tdy*.

1.7 XY female sex reversal

The study of sex determination and the search for *TDF/Tdy* has been facilitated by the existence of genetic mutations giving rise to a sexual phenotype inconsistent with karyotype. Sections 1.4 and 1.5 described cases of XX males in man and mouse. In the next sections the occurrence of XY females is described and the insight they have provided regarding *TDF/Tdy* is considered.

1.7.1 XY^{Tdym1} female mice

Lovell-Badge & Robertson (1990) devised a genetic scheme, designed to isolate new mutations in the sex-determining pathway in mice. They reasoned that mutations in *Tdy* or in "downstream" responder genes would result in a breakdown of normal differentiation events to give complete or partial phenotypic sex reversal. Their strategy made use of chimaeric male mice that had been constructed using an XY embryonic stem (ES) cell line that had been multiply infected in culture with the MPSV.mos^{-1} neo replication defective retroviral vector. Previous analysis of such chimaeric animals had shown that the proviral vector sequences, integrated as single copy events at many different chromosomal locations, were transmitted to their F_1 progeny (Robertson,E *et al.*, 1986). If a single contributing XY ES cell carries a mutation affecting testis determination, the germ cell descendants of this cell will be in the appropriate gonadal setting of the testis to contribute to functional sperm. It would therefore be possible to screen for mutations affecting sex determination simply by looking amongst the offspring of the chimaeras for the presence of XY females. This circumvents the problem that mutations affecting sex determination may be difficult to transmit, as gonadal environment and sex-chromosome complement affect a germ cells ability to complete gametogenesis. (McLaren, 1988b and see section 1.5). In addition, if such a mutation had been caused by the disruption of a cellular gene due to a specific proviral insertion event, this may allow the cloning and identification of the mutated gene. In such a screen the most useful chimaeras are those generated by injecting XY ES cells into XX blastocysts (reviewed by Robertson,EJ & Bradley, 1986). If the contribution of XY cells to the somatic portion of the gonad is high enough (see section 1.6) the chimaera will undergo male development. As host XX cells present in the testis of these chimaeras will fail to undergo spermatogenesis, these animals will only produce sperm derived from the injected ES cells, which simplifies the analysis. To identify XY phenotypically female offspring, chimaeras were mated to females carrying distinct X-chromosome markers. In one of the systems used, females homozygous for the X-linked *Pgk-1*^{a} allele were mated to founder males. The ES cell line carried a *Pgk-1*^{b} allele, so normal XX female progeny would type as PGK-1A/1B, while any anomalous females would be PGK-1A only. Lovell-Badge & Robertson

(1990) identified a single founder germ-line chimaera who sired phenotypically female F_1 progeny lacking paternally inherited X chromosome markers. Karyotypic analysis of these females indicated that they were indeed XY and not XO females. The latter could have arisen by meiotic non-disjunction during spermatogenesis and would type similarly for PGK. These XY females were found to be fertile, although they have a limited reproductive lifespan and produce only small litters. Mahadevaiah *et al.* (1991) have shown that the small number of oocytes produced by these mice is the result of an extensive failure of the X and Y chromosomes to pair during oogenesis. The progeny of XY females included other XY females showing that the phenotype was heritable. To test whether this mutation segregated with the Y chromosome, XY females were mated to males carrying a very small Y chromosome (referred to here as lower case "y") that lacks approximately two-thirds of the long arm but which otherwise appears to be functionally normal (Conway *et al.*, 1991). It is readily distinguished cytologically from the Y chromosome carried by the XY females, which is derived from the 129 mouse strain. In over 200 informative cases XY females always carried the normal sized maternal Y chromosome, not the small y chromosome. Thus the mutation segregates with the Y chromosome from the founder XY females. In addition XYy progeny (which are commonly produced as a result of meiotic non-disjunction in the XY females) were always male, suggesting that this small y chromosome can complement the mutation. The Y chromosome carrying the mutation was given the symbol Ɏ. The availability of F_2 progeny allowed Lovell-Badge & Robertson (1990) to test whether inheritance of the mutation is associated with a specific proviral insertion site. No provirus or viral LTR (which could represent an internally deleted provirus) was found to segregate with the mutation. They concluded that the mutation was not a direct result of integration of MPSV vector sequences.

When XɎ females were mated to XYSxr' males, a number of XSxr'Ɏ progeny were obtained which were fertile males. Complementation of the mutation in this case, places it within the region of the Y chromosome delineated by Sxr', the minimal portion known to contain *Tdy* (see section 1.5). The sex-reversing phenotype and deduced location of this mutation indicate that it must have occurred in *Tdy*. The mutated gene is therefore referred to as Tdy^{m1}. Karyotypic analysis of XɎ females

indicated that the Y chromosome was unaffected at the gross morphological level and that the short arm region was visible. Southern analyses using two repetitive probes which detect sequences mapping to Sxr' did not reveal any difference in hybridization pattern between XɎ females and normal male siblings suggesting that there was no large deletion or rearrangement affecting *Tdy* expression. Lovell-Badge & Robertson (1990) also showed that other genes mapping to the Y chromosome are normal in Ɏ. Thus the spermatogenesis gene *Spy* must be active as XSxr'Ɏ males are fertile and the *Hya* gene must be functional, as XɎ females are positive for the H-Y transplantation antigen.

In summary, Lovell-Badge & Robertson (1990) demonstrated that the fertile XɎ female mice suffer from a mutation in *Tdy*. Unfortunately, one of the original aims of the scheme was to be able to make use of the retrovirus as a tag for the affected gene. The finding that insertional mutagenesis was not a direct cause of the mutation means this was not possible. However the availability of mice known to posses a defective *Tdy* is an invaluable resource. The work described in this thesis concerns the XɎ female mice in the following ways: (i) the possibility that the mutation may have resulted from the insertion of an endogenous retroviral sequence at or near the *Tdy* locus is explored, (ii) candidate genes for *Tdy* are analysed in XɎ female mice for the presence of a molecular defect and (iii) such an analysis has led to an understanding of the mutation which has occurred in this line of mice.

1.7.2 Incompletely penetrant XY sex reversal in the mouse

Eicher *et al.* (1982) have reported a case of XY sex reversal which involves both a locus on the Y chromosome (presumably *Tdy*) and autosomal loci. These authors had crossed a wild strain of mice descended from mice captured in Val Poschiavo, Switzerland with the C57BL/6 inbred strain. When the F_1 progeny of this cross, which were normal, were backcrossed onto C57BL/6 it was found that some of the XY progeny developed as females. The XY female trait was found to depend on the presence of the *poschiavinus* derived Y chromosome (referred to here as Y^{POS}). Y^{POS} is *Mus musculus domesticus* derived, in contrast to the Y chromosome carried

by C57BL/6 which is *Mus musculus musculus* derived. Adult C57BL/6-Y^{POS} mice show a range of sexual phenotypes including fully sex-reversed XY females, true hermaphrodites (i.e. both ovarian and testicular tissue in the same individual) and males. The males have small testes and almost certainly originate from hermaphrodites which during fetal life had a sufficient testicular component to masculinize the external genitalia and reproductive tract. Transmission of the trait is possible through these males (Eicher & Washburn, 1983; Nagamine *et al.*, 1987). Eicher & Washburn (1986) have proposed that this form of sex reversal is due to a specific incompatibility between Y^{POS} and the C57BL/6 genetic background as the Y^{POS} chromosome functions normally on its "own" *Mus musculus domesticus* background and when C57BL/6-Y^{POS} males are mated to females of unrelated inbred strains the F_1 hybrid XY^{POS} mice all develop as males. An explanation for the requirement for a C57BL/6 background for sex reversal to occur was proposed by Eicher & Washburn (1986) when they noticed that in the first backcross generation of an F_1 C57BL/6-Y^{POS} male to a C57BL/6 female, exactly half of the XY progeny show some ovarian tissue in their gonads. The 1:1 ratio suggested the involvement of a single autosomal locus. Eicher & Washburn (1986) have proposed that this locus is involved in testis determination and have called it *testis determination autosomal-1* (*Tda-1*). In this model the C57BL/6 allele, *Tda-1*^{b}, does not interact properly with the *Tdy* locus of the *domesticus* Y chromosome. Because the first generation of the cross was normal, *Tda-1*^{b} is considered recessive to the *domesticus* allele at this locus, *Tda-1*^{a}, which allows normal interaction with the *domesticus* Y chromosome. Eicher & Washburn (1986) have also reported differences in the proportion of XY females with exclusively ovarian tissue, as opposed to hermaphrodites. These differences depend on the number of backcross generations of strains carrying Y^{POS} to C57BL/6. On the basis of these variations they have proposed the involvement of at least one more autosomal gene, *testis determination autosomal-2* (*Tda-2*). While homozygosity for *Tda-1*^{b} causes the development of ovarian tissue in conjunction with Y^{POS}, homozygosity for the C57BL/6 allele of genes like *Tda-2* leads to an increase in the amount of ovarian tissue developing in XY female individuals.

Further studies have shown that a number of (but not all) *domesticus*-derived Y chromosomes generate XY females when placed on the C57BL/6 background (Biddle & Nishioka, 1988). Eicher & Washburn (1986) have suggested that the *domesticus* Y carries a *Tdy* allele which is later-acting than that on the *musculus*-derived Y chromosome of the C57BL/6 inbred strain, and that this delay sometimes enables ovary determination to pre-empt Y action (see section 1.6). Burgoyne (1988) has proposed that *Tda-1* would fit well into this scheme if it was an ovary-determining gene, with the C57BL/6 allele acting earlier than those of other inbred strains. When this is brought together with the late-acting *domesticus Tdy* allele, a "timing-mismatch" may occur such that ovary determination may pre-empt testis determination (Burgoyne & Palmer, 1991). It is not difficult to imagine how minor differences in the timing of expression of these two key genes could lead to the range of gonadal phenotypes observed in these mice. Palmer,SJ & Burgoyne (1991b) have tested this hypothesis by showing that in the presence of the Y^{POS} chromosome, testis development is delayed by 14 hours, suggesting that timing-mismatch is indeed involved in this form of sex reversal.

XY sex reversal can also occur in mice heterozygous for *T-hairpin-tail* (T^{hp}) or *T-Orleans* (T^{Orl}), both deletion alleles at the *Brachyury* (*T*) locus (Washburn & Eicher, 1983, 1989). Washburn & Eicher (1983) have postulated that this dominant autosomal sex-reversal trait is associated with a locus located in the region of chromosome 17 common to T^{hp} and T^{orl}. This locus has been called *T-associated sex reversal* (*Tas*). In this case sex reversal has been shown to occur only on a C57BL/6 background in combination with a Y chromosome from the AKR strain. The AKR Y is *Mus musculus domesticus* derived like Y^{POS}, and although it is associated with delayed testicular development, it does not give sex reversal on its own in a C57BL/6 background. It is therefore possible that the AKR Y may be on the threshold of causing XY sex reversal (Darling *et al.*, 1986). Burgoyne & Palmer (1991) have postulated that the *Tas* deletion, present on a C57BL/6 background, may further accelerate the onset of ovary determination so that the testis-determining effect of the AKR-Y is now pre-empted. When a deletion allele of the *W* locus, W^{19}, is introduced into the C57BL/6 background along with the AKR-Y, it has an equivalent effect to *Tas*, giving rise to XY females (B. Cattanach,

discussion to Eicher, 1988). Cattanach (1987) has therefore proposed that it is the general developmental abnormalities caused by mutations like T^{hp}, T^{orl} (*Tas*) and W^{19} which delay the timing of *Tdy* action relative to ovary determination, giving rise to sex-reversal. Nevertheless the possibility still remains that genes which are directly involved in the testis- or ovary-determining pathways map to these deletions.

The timing mismatch model differs from the ideas of Mittwoch (Mittwoch, 1969, 1989), who has proposed that the sex-determining effect of the Y chromosome can be equated with an accelerating effect of the Y chromosome on the growth of the genital ridge. This supposes that if this growth rate is above a certain threshold, male development occurs. In this model there is no positive ovary-determination process to be pre-empted by the action of the Y chromosome. These arguments cannot distinguish between an accelerated growth rate of the genital ridge causing testis determination or simply being a consequence of testis determination.

1.8 The identification of candidate genes for *TDF/Tdy*

The mouse is considered by many to be the organism of choice for studying mammalian developmental genetics. Thousands of mouse mutants have been described, providing a resource for those interested in cloning genes involved in developmental processes. In the case of sex determination however, the exceptional situation exists in which humans, not mice, have provided the most direct route to cloning a candidate gene. This is due to the relatively common occurrence of XX males that have arisen by abnormal X-Y interchange (see section 1.4) and to the fact that such individuals, being sterile, are generally diagnosed in fertility clinics. Needless to say, such spontaneous cases would invariably go unnoticed in the case of mouse.

In 1987 Page *et al.* described a chromosomal walk along the human Y chromosome designed to clone *TDF* on the basis of its chromosomal position. The authors made use of DNA isolated from two particularly informative sex-reversed individuals. LGL203 was an XX male in whom no Y specific sequences had been found at that

time and WHT1013 was a female with a Y:22 translocation. All known Y-specific sequences were present in this female, except for locus DXYS42 which was deleted. From previous deletion mapping of XX males and XY females who had arisen by abnormal X-Y interchange, Page *et al.* (1987) had shown that DXYS42 was the most closely linked locus to *TDF*. Consequently they initiated a chromosomal walk from this locus. Sequences obtained from the walk were tested for their presence or absence from LGL203 and WHT1013. In this way it was established that individual LGL203 carried only between 235-335kb of the distal short arm of the Y chromosome. A 230kb block of this region had been spanned by the walk. In addition, the entire deletion in WHT1013 seemed to be within this cloned region. This study defined a 140kb interval of the cloned region which was both present in male LGL203 and absent from female WHT1013. It was reasoned that *TDF* (or at least an essential portion of *TDF*) must reside in this 140kb interval. As testis determination is a conserved function of the mammalian Y chromosome, Page *et al.* (1987), attempted to identify *TDF* by isolating sequences from this region showing strong evolutionary conservation amongst a range of mammals. In this way a restriction fragment was discovered that detected highly conserved sequences in many other mammals. Sequencing of this fragment revealed the presence of an open reading frame. The most striking feature of the predicted amino acid sequence of this open reading frame is that it encodes a protein containing 13 zinc fingers of the Cysteine-Cysteine/Histidine-Histidine type. The zinc-finger domain is a highly conserved sequence motif indicative of eukaryotic, sequence specific, DNA binding-proteins (Miller,J *et al.*, 1985). This gene has been called *ZFY* (zinc finger gene on the Y chromosome). *ZFY* has two additional domains, a putative nuclear localization signal and an acidic domain (Ashworth *et al.*, 1989; Mardon & Page, 1989). Acidic domains have been shown to mediate transcriptional activation functions when associated with DNA binding domains (Ptashne, 1988). A single *ZFY* homologue is present on the Y chromosome of most placental mammals, in addition to a highly homologous gene, *ZFX*, which resides on the X chromosome (Page *et al.*, 1987; Schneider-Gaedicke *et al.*, 1989a). In mice the situation is somewhat more complex. There are two Y-linked *ZFY* homologues (*Zfy-1* and *Zfy-2*), an X-linked gene (*Zfx*) and an autosomal homologue (*Zfa*) (Mardon *et al.*, 1989; Nagamine *et al.*, 1989; Ashworth *et al.*, 1990). Both *Zfy-1* and *Zfy-2* have

been found to map to the Sxr region of the mouse Y chromosome. Initially it was thought that only *Zfy-1* mapped to Sxr', *Zfy-2* having been deleted. However, Simpson,EM & Page (1991) have now shown that the deletion which gave rise to Sxr' from Sxr was in fact the result of unequal crossing over between the tandemly arranged *Zfy-1* and *Zfy-2* loci. Thus Sxr' in fact contains a fusion between the 5' portion of *Zfy-2* and the 3' portion of *Zfy-1* and this fusion gene can be expressed. *ZFY* was considered an attractive candidate for *TDF/Tdy* for the following reasons. (i) *ZFY* or its homologues map to the sex determining regions of the human and mouse Y chromosomes. (ii) *ZFY* homologues are generally conserved on mammalian Y chromosomes. (iii) The structural properties of *ZFY* strongly suggested a role as a transcriptional regulator, this would fit in well with the proposed cell autonomous action of *TDF/Tdy* in triggering Sertoli cell differentiation.

The experiments reported in this thesis together with a body of additional evidence (see section 3.3) now show that *ZFY* genes play no role in sex determination. This thesis also describes the cloning and characterization of a new candidate for *Tdy*, known as *Sry*. Both genetic and functional evidence are presented that prove that *Sry* is indeed the mammalian testis-determining gene.

Chapter 2: Materials and Methods

2.1 Molecular Biology techniques

2.1.1 Purification of DNA fragments from agarose gels

DNA fragments contained in gel slices were isolated in two ways. Either by centrifugation through a plug of glass wool (Sigma), or by adsorption to glass beads (Geneclean, Bio 101 Inc.) in 4M NaI. Bound DNA was washed in 0.01M Tris (2-amino-2-(hydroxymethyl)propane-1,3-diol,(tris)) pH 7.5, 0.01M NaCl, 1mM EDTA (disodium ethylenediaminetetraacetate) in 70% ethanol and then eluted in distilled water at 65°C for 5 minutes. Only gels run in 1xTAE (see section 2.14) are suitable for extraction with glass beads. Fragments for microinjection were further purified by passage through a "spin-x" filter unit (Costar).

2.1.2 Preparation of vectors

Vectors (pBluescript, Stratagene) were digested with the appropriate restriction enzyme(s). Singly digested vectors were dephosphorylated in 50mM NaCl, 10mM Tris pH 7.5, 10mM $MgCl_2$, 1mM DTT (Dithiothreitol) and 0.01 units of Calf Intestinal Phosphatase (Boehringer Mannheim) per 1pmole of DNA, for 30 minutes at 37°C. Reactions were phenol extracted and ethanol precipitated.

2.1.3 Ligations

Ligations were carried out using 50ng of vector and a three fold, molar, excess of insert in 50mM Tris pH 7.5, 7mM $MgCl_2$, 1mM DTT, 1mM ATP and 5 units T4 DNA ligase (Northumbria Biologicals), at 4°C overnight.

2.1.4 Preparation of competent bacteria

An overnight culture of bacterial strain DH5αF' (see section 2.15) was diluted into LB medium (see section 2.14) and grown to OD_{600}=0.5-0.6. The cells were pelleted

and resuspended in 1/3 volume of 100mM RbCl, 50mM $MnCl_2$, 30mM K Acetate, 10mM $CaCl_2$ and 15% (volume/volume (v/v)) glycerol and left on ice for 15 minutes. Bacteria were again pelleted and resuspended in 1/12.5 volumes of MOPS (10mM 3-[N-Morpholino]propane-sulfonic acid), 10mM RbCl, 75mM $CaCl_2$ and 15% glycerol. This was incubated on ice for a further 15 minutes and then frozen as aliqouts in liquid nitrogen. 100μl aliquots were stored at -70°C.

2.1.5 Transformation

10μl ligation reactions (see above) were added to 100μl competent bacteria (see above) and incubated on ice for 10 minutes. This was then heat shocked at 42°C for 45 seconds and incubated at 37°C for 10 minutes after the addition of 900μl LB. Bacteria were plated on LB agar (0.7% (weight/volume (w/v)) agar in LB) plates containing 100μg/ml Ampicillin, 0.5mM IPTG (isopropylthio-ß-D-galactosidase) and 40μg/ml X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside). Plates were incubated overnight at 37°C and white colonies were picked.

2.1.6 Plasmid minipreps

2ml of an overnight bacterial culture were pelleted and resuspended in 100μl 50mM glucose, 25mM Tris pH 8.0 and 10mM EDTA. 200μl of 0.2M NaOH, 1% (w/v) SDS (sodium dodecyl sulphate) was added and mixed by gentle inversion. 150μl of 5M Na Acetate pH 5.2 was added and mixed by shaking. Debris was removed by centrifugation and the supernatant was ethanol precipitated. DNA pellets were resuspended in TE (see section 2.14) containing 20μg/ml ribonuclease A (RNAase A).

2.1.7 Plasmid maxipreps

Bacteria were harvested from a 500ml overnight culture by centrifugation and resuspended in 10ml of 50mM glucose, 10mM EDTA, 25mM Tris pH 8.0 and 5mg/ml lysozyme. This was left for 5 minutes at room temperature. 20ml of 0.2M NaOH, 0.1% SDS was added and mixed carefully. This was left on ice for 10

minutes after which 15ml of 5M K Acetate pH 5.2 was added and incubated for a further 10 minutes. Bacterial debris was removed by centrifugation and the plasmid was recovered by precipitation with 0.6 volumes isopropanol. The pellet obtained was resuspended in 10mM Tris pH 8.0, 10mM EDTA containing 1g/ml $CsCl_2$ and 0.5mg/ml ethidium bromide. This was centrifuged in 5ml quick seal tubes (Beckman) at 75000rpm in a vti80 rotor for 2.5 hours. The plasmid band obtained was extracted with isopropanol ($CsCl_2$ and water saturated) until all traces of ethidium bromide had been removed. Two volumes of water were added and the plasmid was ethanol precipitated.

2.1.8 Sequencing

Miniprep DNA (50μl volume obtained from 2ml of culture) was precipitated by the addition of 30μl 20% (w/v) Polyethyleneglycol (PEG), 2.5M NaCl. This was left on ice for 1 hour and then centrifuged. The pellet was washed in 70% ethanol and resuspended in 20μl of water. Plasmid was denatured by the addition of 2μl 2M NaOH, 2mM EDTA. After a 5 minute incubation at room temperature, Na Acetate pH 5.2 was added to 0.3M and the DNA was ethanol precipitated. The pellet was resuspended in water in sufficient volume for a single sequencing reaction. Sequencing, by the dideoxy method (Sanger *et al.*, 1977) was performed using the "T7 sequencing" kit (Pharmacia), according to the manufacturers instructions. Sequence analysis was performed using software designed by the genetics computer group at the University of Wisconsin (Devereux *et al.*, 1984).

2.2 Mouse stocks and breeding

All mice were from stocks maintained at the animal facilities at NIMR except where otherwise stated. For RNA isolation Parkes outbred mice were used except where otherwise stated. Midday on the day of vaginal plug was taken to be 0.5*dpc*. The stage of embryos was confirmed by examination of limb morphology and comparison to the schedule outlined by Hogan *et al.* (1986). Embryos at 11.5*dpc* were sexed by staining for sex chromatin in amniotic cell nuclei (see section 2.11). Older embryos could be sexed by gonadal morphology. XXSxr and XXSxr' males

were from stocks maintained at the Mammalian Development Unit (London). All normal male and female mice used for DNA isolation were of the 129 strain, except CD1 males which carry the *Mus musculus domesticus* Y chromosome (referred to in the text as Y^d). The origin of the Tdy^{m1} mutation and some of the properties of the mice carrying it have been described by Lovell-Badge & Robertson (1990). The ¥ chromosome is of 129 origin and is of the *Mus musculus musculus* type and so where appropriate is referred to as ¥^{m}. The ¥ is maintained most readily by breeding XX¥ females, usually to outbred MF1 males carrying the RIII del 'small Y' chromosome (referred to here as a lower case 'y') (Conway *et al.*, 1991). These crosses normally give rise to roughly equal proportions of XX females, XX¥ females, X¥y males and Xy males. Karyotype analysis (see section 2.13) was routinely used to distinguish offspring (Lovell-Badge & Robertson, 1990). X¥ females used in this study were either founder X¥ females or their offspring. X¥ female embryos were from matings of an X¥y male with CA females and were identified at 12.5*dpc* by gonad morphology and staining of sex chromatin in amnion cells.

2.3 DNA isolation

Genomic DNA was isolated from adult spleens as described by Lovell-Badge (1987). Spleens were homogenised briefly in 100 mM EDTA, 50mM Tris pH 8.0 and incubated overnight at 37°C after the addition of NaCl, SDS and proteinase K to final concentrations of 200mM, 2.5% and 50μg/ml respectively. The lysates were extracted twice with phenol and twice with phenol:chloroform and then dialysed against TE. NaCl and RNAase A were added to final concentrations of 200mM and 100μg/ml respectively and samples incubated for 1 hour at 37°C. After two further phenol extractions DNA was ethanol precipitated and resuspended in water.

2.4 RNA isolation

Total RNA was isolated by the method of Auffray & Rougeon (1980). Tissues were homogenised in 6M urea, 3M LiCl. After overnight incubation at 4°C the pellet

obtained was washed twice in the same solution. The pellet was then resuspended in 10mM Tris pH 7.5 , 0.5% SDS containing 50μg/ml proteinase K and incubated for 30-120 minutes at 37°C. After three phenol extractions, total RNA was ethanol precipitated and resuspended in water. For small quantities of tissue, such as single pairs of genital ridges or fetal gonads an alternative procedure was used which has been described by Chomczynski & Sacchi (1987). The tissue was resuspended in 6M guanidinium thiocyanate, 40mM Na Citrate pH 7.0, 0.8% (w/v) Sarkosyl, 15mM β-mercaptoethanol and 0.2M Na Acetate pH 4.0 up to a total volume of 100μl. This was extracted once with chloroform:isoamylalcohol (49:1) and ethanol precipitated. The pellet was resuspended in water. For cases where removal of DNA contamination was critical (eg for PCR reactions involving primers within a single exon) an optional deoxyribonuclease (DNAase) treatment was performed. The RNA was resuspended in 200mM Tris pH 8.0, 50mM NaCl, 30mM $MgCl_2$ containing 1 unit/μl DNAase and 1 unit/μl RNAase inhibitor (RNasin, Promega Biotec). After a 2 hour incubation at 37°C the reaction was extracted once with phenol and the RNA was ethanol precipitated.

2.5 Southern blotting

10 μg samples of genomic DNA were digested with the appropriate restriction enzymes and electrophoresed on 0.7% TBE/agarose gels (see section 2.14). The gels were capillary blotted onto Hybond-N nylon filters (Amersham) for 4 hours in 1.5M NaCl and 0.5M NaOH. After blotting, filters were neutralized briefly in 2xSSC (see section 2.14). Nucleic acid was crosslinked to dry filters by UV irradiation in a Stratalinker (Stratagene).

2.6 Northern blotting

10μg samples of total RNA were electrophoresed in 1% agarose gels containing 7% (v/v) formaldehyde in 1xMOPS buffer (see section 2.14). RNA samples were prepared by resuspending in sample buffer (1xMOPS, 50% (v/v) formamide, 7% formaldehyde and 40μg/ml ethidium bromide) and heating at 65°C for 10 minutes

before loading. Gels were transferred overnight in 20xSSC to Hybond-N filters, which were subsequently treated as for Southern blots.

2.7 Hybridization

DNA probes, which had been denatured by boiling for 2 minutes, were labelled with ^{32}P by incubation for at least 30 minutes at 37°C in a total volume of 50μl containing 200mM Tris pH 8.0, 40mM $MgCl_2$, 40mM β-mercaptoethanol, 0.1mM dATP/dTTP/dGTP, 1M HEPES (N-2-Hydoxyethylpiperazine-N'-2-ethanesulfonic acid) pH 6.6, 27 OD units/ml random hexanucleotides (Pharmacia), 50μCi α-^{32}P dCTP (3000 Ci/mmol, DuPont) and 1 unit of klenow (the large fragment of DNA polymerase I). Unincorporated nucleotides were removed by spinning through a Sephadex G-50 (Pharmacia) column, that had been equilibrated in TE containing 0.1% SDS. After boiling for 10 minutes, probes were added to the filter in 3xSSC, 0.1% (w/v) sodium pyrophosphate, 5xDenhardt's solution (see section 2.14), 0.1% SDS, 10% (w/v) dextran sulphate and 50μg/ml denatured herring testis DNA, and hybridized for 16 hours at 65°C in a rotisserie oven (Hybaid). Filters were washed in 1xSSC, 0.1% SDS at 65°C for 1 hour or, at higher stringency where stated, 0.1xSSC, 0.1% SDS at 65°C for 1 hour and then exposed to X-ray film (XAR-5, Kodak) for 1-3 days. Filters were stripped in 0.1% SDS at 100°C for 1 hour. For genomic probes thought to contain repetitive sequences probes were prehybridized at 65°C for 2 hours, after 5 minutes boiling, with 2.5mg/ml mouse genomic DNA (sonicated to give an average length of 500bp). 0.25mg/ml mouse genomic DNA was also added to the filter during prehybridization. Oligonucleotides were end-labelled by incubating 20ng of oligonucleotide at 37°C for 1 hour in 50mM Tris pH 7.6, 10mM $MgCl_2$, 5mM DTT, 0.1mM spermidine, 0.1mM EDTA, 100μCi γ-^{32}P dATP (3000Ci/mmol, DuPont) and 5 units of polynucleotide kinase. Filters were hybridized in 25mM NaH_2PO_4, 6xSSC, 10xDenhardt's solution, 0.1% SDS and 10μg/ml yeast tRNA for 16 hours at 55°C and washed in 6 x SSC at 65°C for 1 hour. The following oligonucleotides (Stoye *et al.*, 1988) were used as probes: JS4; 5'-GCAGC CTCTA TACAA CCTGG GACGG GAG-3', JS5; 5'-GCAGC CTCTA TAGTC CCTGA GACTG CCC-3', JS6; 5'-ACGGT CTCTA TGGTA CCTGG GGCTC CCC-3', JS10; 5'-ACGGT CTCTA TGGTG CCTGG GGCTC CCC-3'.

2.8 Polymerase Chain Reaction (PCR)

PCR reactions were carried out in a 50μl volume containing 50mM Tris pH 9.0, 15mM ammonium sulphate, 7mM $MgCl_2$, 0.17mg/ml Bovine serum albumin (BSA), 0.05% (v/v) Nonidect-P40 (NP40), 3.75mM dATP/dCTP/dGTP/dTTP, 500ng of each primer and 1 unit of *Thermus aquaticus (Taq)* DNA polymerase (Anglian). 1μg of genomic DNA was used per reaction. Reactions were overlayed with paraffin oil and amplified in a Techne PHC-2 thermocycler (see individual primer sets for conditions). Electrophoresis of 5 μl of each reaction confirmed the amplification of only PCR products of the predicted size. For reverse transcriptase-PCR (RT-PCR) 1μg total RNA was reverse transcribed using 200 units of Moloney murine leukaemia virus (MoMuLV) reverse transcriptase (BRL) in 30 μl of the supplied buffer containing 500 ng oligo-dT (Pharmacia) at 42°C for 30 minutes. 5μl of this cDNA was then added to a standard PCR reaction. For isolated genital ridges, the entire yield of RNA was reversed transcribed in 7.5μl and amplified as described above.

Oligonucleotide primers were prepared by a service at NIMR using an automated DNA synthesizer (Applied Biosystems 380B DNA synthesizer). Primers were purified from their ammonium hydroxide solvent by lyophilization in an evacuated centrifuge (Speedvac). The pellet obtained was washed in ethanol and resuspended in water. The primer was then ethanol precipitated and resuspended in water at 500ng/μl.

2.8.1 *Zfy* expression

For analysis of RNA isolated from testes of adult $X\not{Y}^{m}Y^{d}$ mice, the following primers were used: 5'-CCTAT TGCAT GGACT GCAGC TTATC-3' and 5'-CGTAA AGTTT GTCGA TCAGG AGCAA C-3'. Primer sequences matched both *Zfy-1* and *Zfy-2* cDNA sequences (Ashworth *et al.*, 1989; Mardon & Page, 1989) except for the single mismatches shown in bold type. These were designed to introduce restriction sites for *Pst*I and *Taq*I respectively (underlined) to facilitate subsequent cloning (see below). Amplification conditions were 30 cycles of 94°C,

5 sec, 65°C, 30 sec and 72°C, 30 sec. Reactions were extracted with phenol:chloroform, digested with *Pst*I and *Taq*I and subcloned into the *Pst*I and *Acc*I sites of pBluescript, as described above.

For analysis of *Zfy-1* expression in XY̶ female embryonic samples *Zfy-1* specific primers (5'-GTTAC TCATT TTCAG GTGTT CTGGG-3' and 5'-GTGTC AGCTG TTATA GGATC AGTGA-3') (Ashworth *et al.*, 1989; Koopman *et al.*, 1989) were used. The hypoxanthine phosphoribosyltransferase (*Hprt*)-specific primers (5'-CCTGC TGGAT TACAT TAAAG CACTG-3' and 5'-GTCAA GGGCA TATCC AACAA CAAAC-3') (Melton *et al.*, 1984) were included as a positive control for the presence of RNA in each sample. Amplification was for 20 cycles using the conditions above. 1 μl of the completed PCR reaction was transferred to a new reaction containing only "nested" primers (5'-TGAAG TCTGC AGTAC TTGTC GTCAT-3' and 5'-TCACT CATCA AGACA TGTTC AGGCA-3') which further amplify the *Zfy-1* PCR product. 20 μl of the second reaction was visualised by electrophoresis on a 2% agarose gel.

2.8.2 XY̶ female analysis

For confirmation of the structure of the deletion in XY̶ female mice, a single primer (5'-GTGTC TCAAA GCCTG CTCTT C-3') was used to amplify a 522bp fragment from XY̶ female genomic DNA, using the conditions described above.

2.8.3 *Sry* transgenics

For transgenic diagnosis the following primers were used: *Sry*, 5'-TCATG AGACT GCCAA CCACA G-3' and 5'-CATGA CCACC ACCAC CACCA A-3'; *Zfy-1*, 5'-CCTAT TGCAT GGACT GCAGC TTATG-3' and 5'-GACTA GACAT GTCTT AACAT CTGTC C-3'; myogenin primers corresponded to nucleotides 656-675 and 882-901 of the rat cDNA sequence (Wright,WE *et al.*, 1989). Amplification was for 30 cycles using the above conditions.

2.9 Genomic library construction and screening

2.9.1 Size selected library

DNA isolated from the spleen of a 129 male mouse was digested to completion with the restriction enzyme *Eco*RI. DNA in the size range 3-4kb was recovered from an agarose gel (see section 2.1.1). 0.1μg of this DNA was ligated into 1μg of *Eco*RI digested and phosphatased λ ZapII vector (Stratagene) by overnight incubation at 4°C in 50mM Tris pH 7.5, 7mM $MgCl_2$, 1mM DTT and 2 units of T4 DNA ligase. The reaction was packaged using Gigapack gold packaging extract (Stratagene) according to the manufacturer's instructions.

2.9.2 Total genomic library

Small scale pilot digests were performed on genomic DNA using serial dilutions of the restriction enzyme *Sau*3aI in order to establish conditions for obtaining fragments in the size range 18-23kb. DNA in this size range was isolated by gel purification of a suitable large scale partial digest. The ends of this DNA were partially filled in by incubating 50μg for 15 minutes at room temperature in 50mM NaCl, 10mM Tris pH 7.5, 1mM dGTP, 1mM dATP and 15 units of Klenow. 0.4μg of this DNA was ligated in a λ FixII vector (Stratagene) and packaged as described above.

2.9.3 Plating and screening

Plating bacteria were prepared by resuspending an overnight culture of bacterial strain DL652 in 0.5 volumes of ice cold 10mM $MgSO_4$. An appropriate volume of the library was plated on LB agar containing 10mM $MgSO_4$ in 22x22cm bioassay dishes (Nunc) to give $2.5x10^5$ plaques per plate. Phage were incubated without agitation at 37°C with 2ml of plating bacteria for 15 minutes and then poured onto the plates after the addition of 30ml of molten (55°C) top agarose (0.7% agarose in LB containing 10mM $MgSO_4$). Plates were incubated at 37°C overnight (not more than 12 hours). Phage DNA was transferred to duplicate Hybond-N filters

which were denatured in 0.5M NaOH, 1.5M NaCl, neutralised in 1M NH_4 Acetate and washed in 2xSSC. Subsequent filter treatment and hybridization was as described for Southern blotting and hybridization. Positive plaques were picked, and rescreened a number of times until single plaques could be isolated.

2.9.4 Isolation of phage DNA

High titre ($>10^{11}$ plaque forming units/ml) liquid lysates were prepared as follows: an agarose plug containing a single phage plaque was placed in 0.3ml of plating bacteria in a 50ml tube. After 15 minutes incubation at room temperature, 10ml of LB containing 10mM $MgSO_4$ and 5mM $CaCl_2$, were added. The culture was then incubated with vigorous agitation at 37°C until lysis occurred (5-6 hours). Bacterial debris was removed by centrifugation. The lysate could be kept for long periods in the presence of chloroform as a phage stock.

For phage minipreps, liquid lysates were centrifuged at 16000rpm (Sorval rotor HB4) for 2 hours. The pellet was resuspended in 0.3ml SM (see section 2.14), containing 0.1% SDS and 0.1% (w/v) EDTA and incubated for 15 minutes at 37°C. After one phenol and one phenol:chloroform extraction, phage DNA was ethanol precipitated and resuspended in water.

For phage maxipreps 50μl of liquid lysate was incubated with 16ml of plating bacteria for 15 minutes at 37°C. This was added to 500ml of LB containing 10mM $MgSO_4$ and 5mM $CaCl_2$ and then incubated with vigorous agitation at 37°C until lysis occurred (5-6 hours). NaCl was added to a final concentration of 1M and 10 ml of chloroform were added. Bacterial debris was removed by centrifugation in 250ml bottles at 6000rpm for 10 minutes. The supernatant was further filtered by passing through muslin. PEG (6000-8000) was added up to 10% and left for between 1 hour to overnight at 4°C. This was centrifuged at 11000rpm for 10 minutes and the pellet resuspended in SM. After addition of 0.75g/ml CsCl, the solution was centrifuged in 5ml quick seal tubes (Beckman) at 75000rpm in a vti80 rotor for 2.5 hours. The phage band obtained was isolated and dialysed against TE.

This was then phenol, phenol:chloroform and chloroform extracted. Phage DNA was ethanol precipitated and resuspended overnight in water.

2.10 Microinjection of fertilised mouse eggs

The procedures used are described by Hogan *et al.* (1986). Four week old (CBA x C57BL/10) F_1 female mice were superovulated by injecting intraperitoneally with 5 units (0.1ml) of Pregnant Mare Serum (PMS) 48 hrs before and 5 units (0.1ml) of Human Chorionic Gonadotrophin on the day of caging with (CBA x C57BL/10) F_1 male mice. The next day, oviducts from plugged females were removed and the fertilised eggs released by dissection under M2 medium (Hogan *et al.*, 1986). Hyaluronidase was added to a final concentration of 300 μg/ml. The eggs were then washed through at least four drops of M2 medium to remove all cumulus cells and traces of hyaluronidase. These 1 cell embryos were washed once more in M2 with hyaluronidase to remove the cumulus cells before being rinsed once in M2 medium alone. Purified DNA fragment (see section 2.1.1) was injected into either (usually the male) pronucleus of one cell embryos. Injected eggs were washed once in M16 medium (Hogan *et al.*, 1986) and transferred to M16 droplets under oil equilibrated in a humidified incubator at 37°C with 5% CO_2 for overnight culture. Fertilised eggs which survived the microinjection procedure were reimplanted into the oviducts of day of plug pseudopregnant female mice the following morning. Mice were anaesthetised by intraperitoneal injection of 0.1ml (for a 25g mouse) of 0.5mg/ml midazolam (Hypnovel, Roche), 2.5mg/ml fluanisone and 0.079 mg/ml fentanyl citrate (Hypnorm, Crown Chemical Company).

2.11 Sex chromatin staining in amnion cells

The method described by Monk & McLaren (1981) was used. In brief, amnions were dissected into freshly made fixative (3:1 methanol:glacial acetic acid) in a 1.5ml microfuge tube. The fixative was then aspirated and three drops of 60% acetic acid were added and vortexed. This was taken up in excess fixative and centrifuged (5000rpm, 1 minute). The supernatant was poured off and the cells were resuspended in the residual liquid. This was then applied to a microscope

slide and allowed to dry. Slides were stained by inversion on to a coverslip to which one drop of 1% (w/v) toluidine blue had been added. Cells from embryos with two X chromosomes show a densely staining granule at the periphery of the nucleus, when viewed under x200 magnification.

2.12 Histology

Gonads were fixed in 4% (w/v) paraformaldehyde, dehydrated in ethanol and embedded in paraffin. Sections (7μm) were stained in haematoxylin and eosin.

2.13 Karyotype analysis

Embryonic fibroblasts were cultured by disagregating embryonic tissue with a scalpel in a petri dish, in Dulbecco's modified Eagle's medium (DMEM). Cultures were maintained in a humidified incubator at 37°C, in an atmosphere of 5% CO_2, until they had reached semi-confluency. At this stage colcemid was added to a final concentration of 0.06μg/ml and the culture was further incubated for 60 minutes. For a single 60mm dish, cells were collected after trypsinization (using 0.25% (w/v) trypsin) and washed twice in phosphate-buffered saline (PBS see section 2.14) by pelleting in a benchtop centrifuge at 1000rpm for 5 minutes. The cells were then resuspended in 10ml 0.56% (w/v) KCl and incubated at room temperature for 10 minutes. The cells were then pelleted again (as above). The pellet was resuspended by the dropwise addition of 10ml of fixative (3:1 methanol:glacial acetic acid). After incubation at room temperature for 10 minutes the cells were pelleted once more and resuspended in the same volume of fixative. 1-5 drops of this cell suspension were dropped on to a microscope slide. Slides were immersed in a solution of 2% (v/v) Giemsa in 50mM phosphate buffer pH 7.0 for 5 minutes and allowed to dry before examination under the microscope. For adult and new born mice spleen cells or cells from tail tip cultures were used respectively.

2.14 Commonly used solutions

1xDenhardt's:	0.02% (w/v) Ficoll (Type 400, Pharmacia), 0.02% (w/v) polyvinylpyrrolidone and 0.02% (w/v) BSA (Fraction V, Sigma)
LB:	1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract and 0.17M NaCl
1xMOPS:	20mM 3-[N-Morpholino]propane-sulfonic acid, 5mM Na Acetate and 1mM EDTA
PBS:	140mM NaCl, 3mM KCl, 10mM Na_2HPO_4 and 2mM KH_2PO_4
SM:	100mM NaCl, 8mM $MgSO_4$, 50mM Tris and 0.01% (w/v) gelatin
1xSSC:	0.15M NaCl and 0.015M Na Citrate pH 7.0
1xTAE:	0.04M Tris-acetate pH 8.0 and 1mM EDTA
1xTBE:	0.089M Tris-borate, 0.089M boric acid and 0.002M EDTA
TE:	10mM Tris pH 7.5 and 1mM EDTA

2.15 Bacterial strains

DH5αF':	*supE44ΔlacU*169 (*φ80lacZ*ΔM15) *hsdR*17*recA*1*endA*1*gyrA*96*thi*-1*relA*1 F'[*traD*36*proAB*$^+$*lacI*^{q}*lacZ*ΔM15]
DL652:	K802[*supEhsdRgalmetB*]recC

Chapter 3: Results and Discussion

Zfy as a candidate for *Tdy*

XɎ female mice carry a mutation in *Tdy* (see section 1.7.1). Thus they represent an ideal system to test the hypothesis that either *Zfy-1* or *Zfy-2* is equivalent to *Tdy*. This chapter contains an analysis of the structure of these genes in XɎ female mice and their expression patterns both in wild type mice and XɎ females.

3.1 Analysis of *Zfy* gene structure in XɎ female mice

The physical nature of the mutation which had occurred at the *Tdy* locus in the XɎ female mice was unknown. However, on the assumption that the mutation may have created a restriction fragment length polymorphism (RFLP) an extensive Southern analysis of both the *Zfy-1* and *Zfy-2* genes was carried out. A set of three contiguous probes that together comprise a full length *Zfy-1* cDNA clone was used on Southern blots of XY male and XɎ female DNA cut with a range of restriction enzymes. Figure 4 shows the results of this comparison. No differences, in either the number or size of bands, between XY male and XɎ female DNA were detected with any of the probes used. Therefore, at this level of resolution (see section 3.3), no defect in either *Zfy* gene can be found which correlates with sex reversal in the XɎ females.

The pattern of hybridization seen in figure 4 consists of bands corresponding to all four *ZFY*-related genes in the mouse (*Zfy-1*, *Zfy-2*, *Zfx* and *Zfa*). cDNAs derived from *Zfy-1* and *Zfy-2* are almost identical in sequence along their entire length (Ashworth *et al.*, 1989; Mardon & Page, 1989), thus despite the fact that the probes used were derived from *Zfy-1*, bands corresponding to both *Zfy* genes are visible. Similarly, *Zfx* and *Zfa* cDNA sequences are almost identical (Ashworth *et al.*, 1990). However there are significant differences between the two sets of genes. At the amino acid level *Zfx* and *Zfy-1* are 79% homologous in the zinc finger region but only 55% homologous in the acidic domain. Consequently, while all the probes used in the study shown in figure 4 cross-hybridized weakly with *Zfx* and *Zfa*, probe 3, which contains the *Zfy-1* zinc finger domain, showed the strongest level of cross-

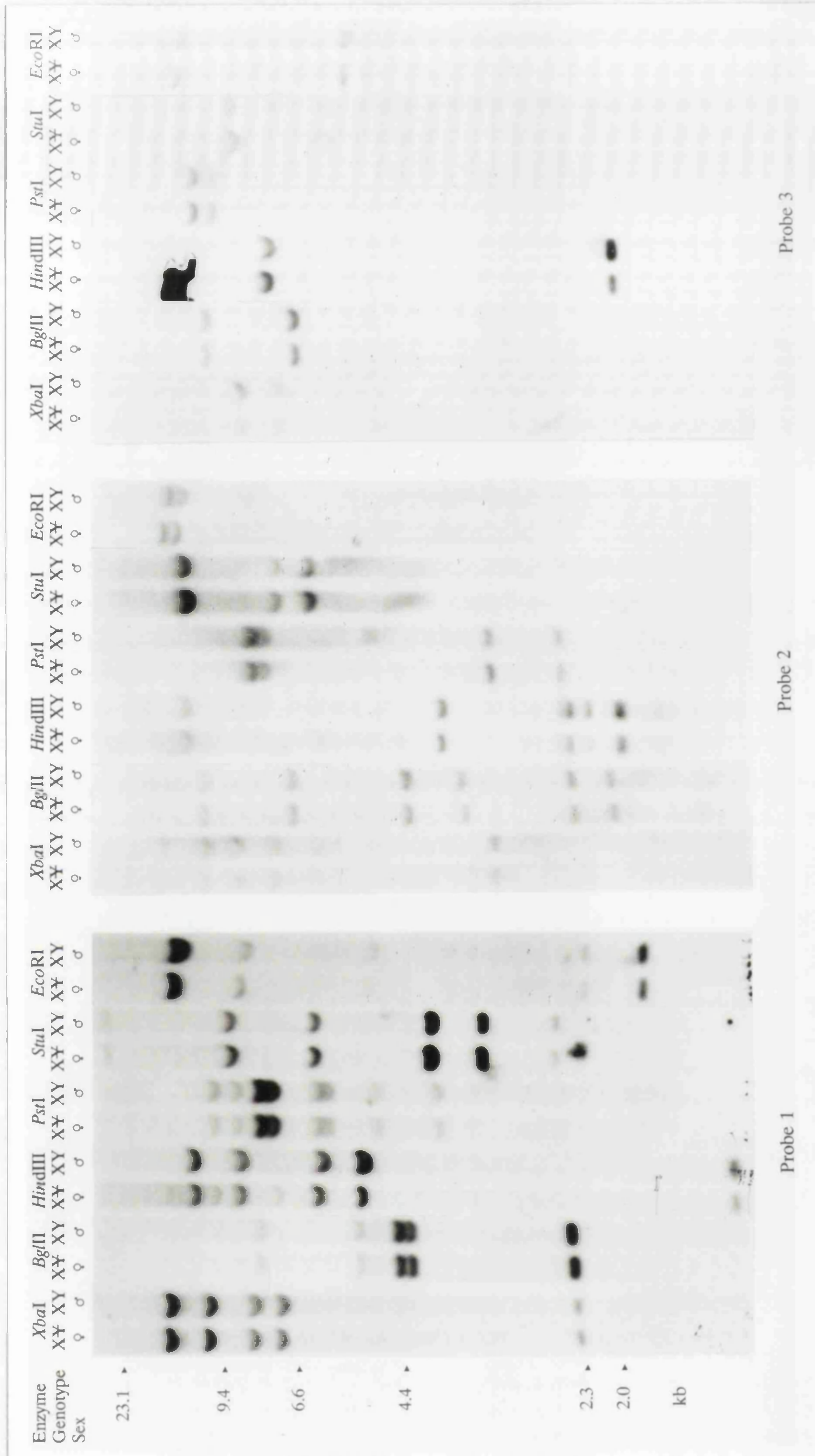


Figure 4. Comparison of *Zfy* genes in XY̶ female versus XY male mice by Southern analysis. Southern blots were hybridized to the following probes; Probe 1, nucleotides 1 to 788 (*Eco*RI-*Bgl*II restriction fragment) of a full length *Zfy-1* cDNA clone (Ashworth *et al.*, 1989). Probe 2, nucleotides 789 to 1496 (*Bgl*II-*Hin*dIII restriction fragment) of the *Zfy-1* cDNA. Probe 3, *Zfy-1* zinc finger probe, contiguous with probe 2 and including some 3' flanking sequences to *Zfy-1*.

hybridization to *Zfx* and *Zfa* sequences. The origin of the bands in figure 4 has been established by hybridizing the same set of probes to additional Southern blots in which DNA from XX females, XXSxr' males, XɎ females and XY males has been digested with various restriction enzymes (an example of such a blot is shown in figure 5). Bands corresponding to *Zfx* and *Zfa* are present in DNA from both XY males and XX females, while all Y-linked bands are absent from XX female DNA. In addition Mardon *et al.* (1989) and Nagamine *et al.* (1989) proposed that *Zfy-1* but not *Zfy-2* maps to Sxr'. Thus bands present in DNA from XXSxr' males but absent from XY male DNA were assumed to correspond to *Zfy-1*, while bands common to both XY male and XXSxr' male DNA were considered to be from *Zfy-2*. A restriction map of the *Zfy-1* genomic locus has been derived by combining three sets of data (see figure 6). (i) The assignment of bands, detected by probes 1,2 and 3 (figure 4), on Southern blots to *Zfy-1* by further hybridization of these probes to Southern blots of XXSxr' genomic DNA (as described above). (ii) Using an *Eco*RI-*Nco*I restriction fragment of probe 1 (see figure 6), as an additional probe on Southern blots. This probe represents the extreme 5' end of *Zfy-1* and detects a subset of the bands recognised by the larger probe 1, allowing these bands to be ordered on the map. (iii) A partial *Zfy-1* genomic clone, c21, isolated by B. Skene (unpublished results), was analysed to determine its restriction map which is shown in figure 6. The derivation of the probes used and their relationship to the restriction map is also depicted. This analysis gives a minimum size for *Zfy-1* of 24.5kb.

Recently Simpson,EM & Page (1991) have shown that Sxr' was generated by an interstitial deletion occurring between *Zfy-1* and *Zfy-2*. Thus Sxr' contains genomic restriction fragments corresponding to the 5' portion of *Zfy-2* and the 3' portion of *Zfy-1*. This new data means that in some cases the assignment of bands to *Zfy-1* or *Zfy-2* was incorrect. This does not compromise the conclusions regarding the structure of the *Zfy* genes in XɎ female mice, as in this study no differences were seen in any of the bands regardless of their gene of origin. Simpson,EM & Page (1991) have shown that the position of the breakpoint in *Zfy-1*, which gave rise to Sxr', is between the first and second exons. Thus the most 5' restriction fragments detected by probe 1 (figure 4), which correspond to the genomic region containing

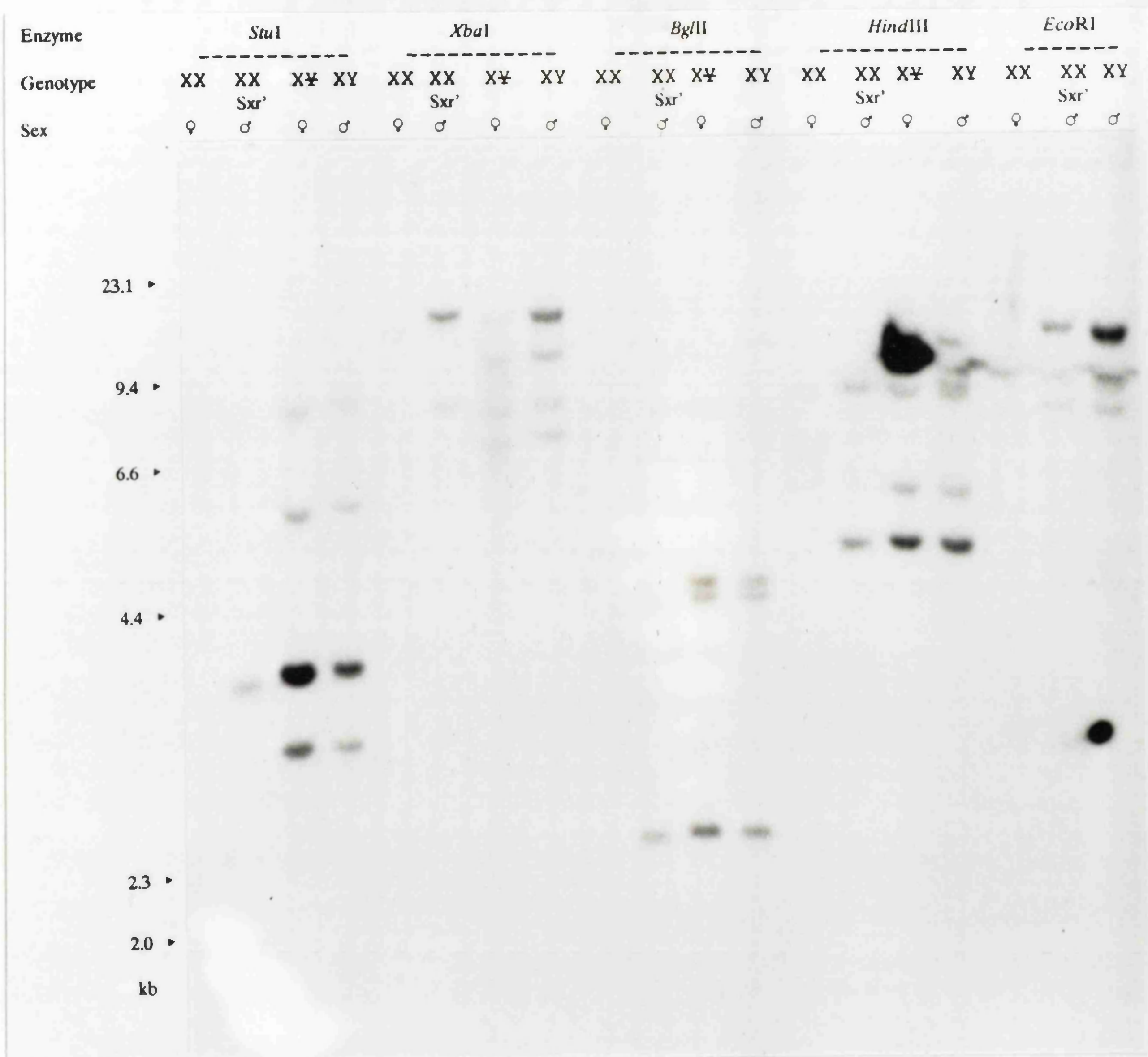


Figure 5. Assignment of bands to *Zfy-1* and *Zfy-2* by Southern analysis. A Southern blot hybridized to *Zfy-1* probe 1 (described in the legend to figure 4) is shown. DNA size markers shown are λ DNA digested with *Hind*III.

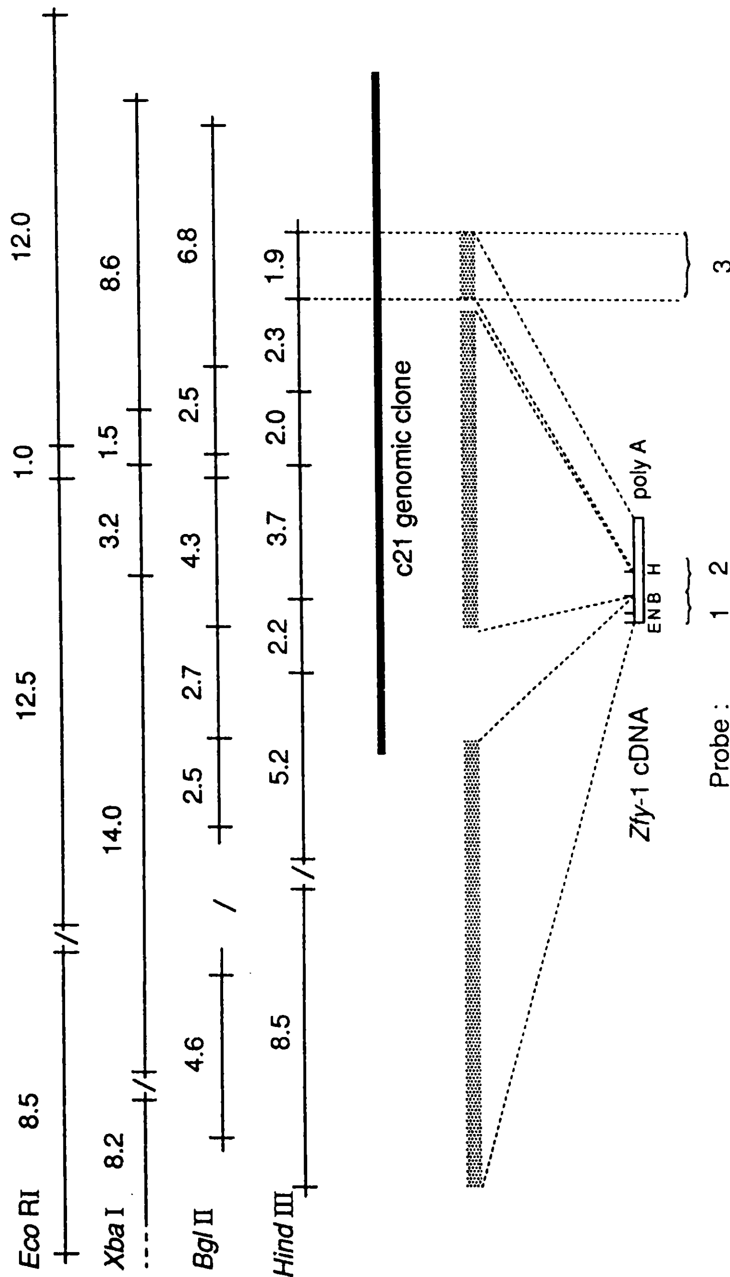


Figure 6. *Zfy-1* genomic map. A map showing the order of *Zfy-1* genomic restriction fragments and the probes to which they hybridize. *c21* is a genomic clone covering part of the *Zfy-1* gene. Bands seen on Southern blots, which are not overlapping with this clone, may not be contiguous with it and are separated by +/+. Stippled lines (▒) indicate which genomic fragments hybridize to a given probe. The relationship of the three probes used to the *Zfy-1* cDNA is also shown schematically. Sites for the following restriction enzymes are shown: B; *Bam*HI, H; *Hind*III, E; *Eco*RI and N; *Nco*I. These sites define the *Eco*RI-*Nco*I restriction fragment of probe 1, which has been used as an additional probe to map the *Zfy-1* locus. Later work (Simpson,EM & Page, 1991) has indicated that the extreme 5' restriction fragments shown, in fact correspond to *Zfy-2* (see figure 7).

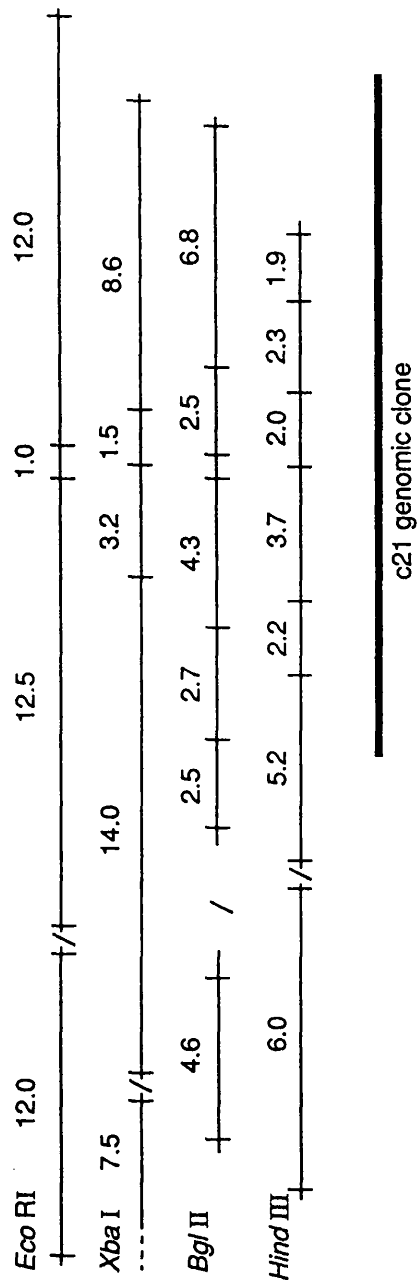


Figure 7. Revised map of the *Zfy-1* genomic locus. See legend to figure 6 for full description. This map incorporates data from Simpson,EM & Page (1991) which shows that Sxr' was created by a fusion event between the two *Zfy* genes.

the first exon, originate from *Zfy-2* rather than *Zfy-1* as assigned previously. Figure 7 shows a revised restriction map of the *Zfy-1* genomic locus incorporating this new data. The estimate for the minimum size of *Zfy-1* remains unaltered.

3.2 Analysis of *Zfy* gene expression in XY̶ female mice

It is conceivable that the mutation in the XY̶ females is not within the *Zfy-1* or *Zfy-2* coding regions but, instead, affects the expression of either of these genes. To study this it was necessary to look for *Zfy* transcripts from the mutant Y̶ chromosome. In order to do this the expression patterns of these genes in wild type XY male mice had to be determined. Such an expression analysis is critical in its own right for assessing the involvement of candidate genes in sex determination, as such genes need to satisfy various predictions concerning the expression pattern of *Tdy* (see section 1.6). Experiments concerning *Zfy* expression in wild type males and XY̶ females are described in the next four sections.

3.2.1 Expression of *Zfy* genes in adult testis from wild type mice

In order to determine the expression pattern of the *Zfy* genes and establish whether this pattern was consistent with that expected for *Tdy*, Northern blots were prepared using RNA isolated from male and female genital ridges, embryonic gonads at various stages of development, adult testis and adult ovary. Figure 8a (top panel) shows such a Northern blot hybridized to a probe containing the *Zfy-1* zinc finger region (probe 3, legend to figure 4). This probe cross-hybridizes to all of the *ZFY*-related genes in the mouse. Two bands corresponding to transcripts from the *Zfx* gene (Ashworth *et al.*, 1990) can be seen in all the samples. This and other data (Ashworth *et al.*, 1990) have shown that *Zfx* is ubiquitously expressed in mouse. The *Zfy* band is only visible in RNA from adult testis. This result was confirmed by subsequent hybridization of the same filter to a different *Zfy-1* probe (probe 1, legend to figure 4) which is more specific for *Zfy* genes at high stringency (figure 8a, bottom panel). In this case the only band visible is that in adult testis, the size of which corresponds to that predicted from the known cDNA sequence (Ashworth *et al.*, 1989; Mardon & Page, 1989). Neither probe distinguishes between *Zfy-1* and

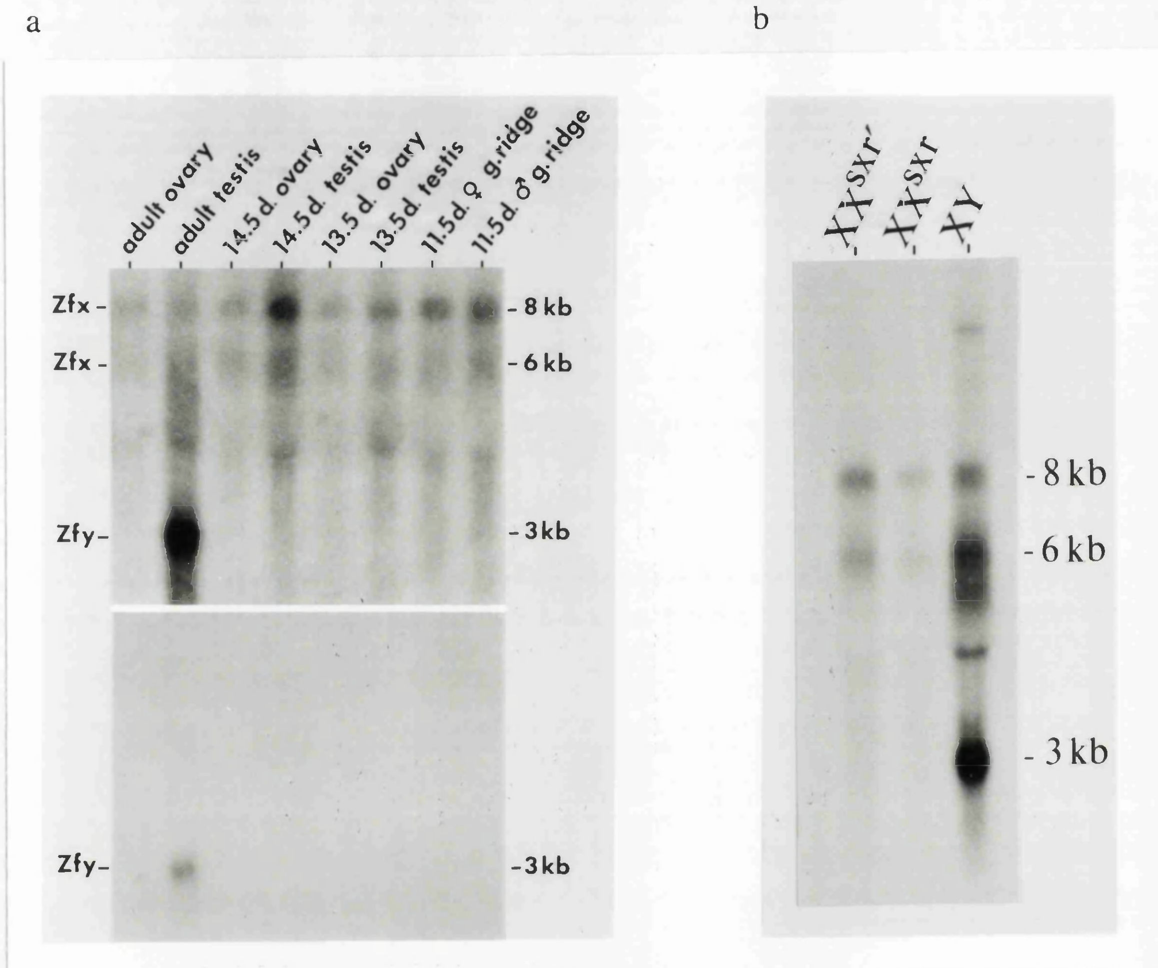


Figure 8. Northern blot analysis of *Zfy* gene expression. a, Analysis of RNA from adult and embryonic mouse gonads and urogenital ridges (g. ridge) shows expression of a 3kb *Zfy* transcript specific to adult testis. The upper panel shows hybridization to a 1.9kb *Hind*III *Zfy-1* genomic fragment containing the zinc-finger encoding exon (probe 3, legend to figure 4), which also detects 6 and 8kb *Zfx* transcripts. The lower panel shows the same filter rehybridized with the first 788bp of a *Zfy-1* cDNA (probe 1, legend to figure 4) which is more specific for *Zfy* sequences at high stringency. b, Absence of the *Zfy* transcript in XXSxr and XXSxr' adult testis which lack germ cells. The *Zfy* zinc-finger probe shows that *Zfx* transcripts are present in both testis types, as well as in normal XY testes.

Zfy-2. These results and data from Mardon & Page (1989) suggest that the only site of *Zfy* expression detectable by Northern analysis is adult testis. Koopman *et al.* (1989) and Nagamine *et al.* (1989) have demonstrated that *Zfy-1* and *Zfy-2* are expressed in adult testis using RT-PCR based assays that could distinguish *Zfy-1* and *Zfy-2* transcripts.

To establish whether *Zfy* expression was associated with the somatic or the germ cell component of the adult testis, RNA from the testes of adult male XXSxr and XXSxr' mice was compared to RNA from normal adult testis by Northern blot analysis. The testes of adult XXSxr and XXSxr' mice are somatically normal but lack germ cells as a result of a degeneration of spermatogonia prior to 10 days of age (Cattanach *et al.*, 1971; Burgoyne *et al.*, 1986 and see section 1.5). Figure 8b shows such a Northern blot hybridized to the *Zfy-1* zinc finger probe (probe 3, legend to figure 4). Although the two *Zfx* transcripts can be seen in all tracks, the characteristic 3kb *Zfy* band is undetectable in the XXSxr and XXSxr' samples. This indicates that *Zfy* expression in adult testis is germ cell dependent. Nagamine *et al.* (1990) have extended this study by analysing *Zfy* expression in RNA isolated from fractionated germ cells from adult testis. Their study showed first that germ cells are the source of *Zfy* expression in adult testis and second that this expression can be attributed to the maturing germ cell population, in particular the round spermatids.

3.2.2 Expression of *Zfy* genes in adult testis from XɎY mice

To test whether transcription of the *Zfy* genes could occur from the mutant Ɏ chromosome, in adult testis, it was necessary to obtain male mice carrying this chromosome. Lovell-Badge & Robertson (1990) have described that it is possible to obtain complementation of Tdy^{m1} with a normal Y chromosome in XɎY males. These occur spontaneously amongst the offspring of XɎ females through non-disjunction of the X and Ɏ, but are more readily obtained by breeding XXɎ females with normal XY males, where 25 % of offspring will be XɎY. However, it was necessary to do this in such a way that the products of the *Zfy* genes from the two Y chromosomes in the XɎY males could be distinguished.

Comparison of sequences obtained from the partial *Zfy-1* genomic clone, c21 (see figure 6), with the published sequences of *Zfy-1* and *Zfy-2* cDNAs revealed a number of nucleotide differences. The genomic library, from which c21 was cloned, was constructed from DNA of the 129 mouse strain which carries the *Mus musculus musculus* type Y chromosome (referred to here as Y^m), whereas both published sequences were derived from mice carrying a Y chromosome of the *Mus musculus domesticus* type (Y^d). As the mutant Ɏ is of 129 and therefore of *musculus* origin, this suggested that there may be sufficient differences between it and a *domesticus* type Y to allow all the various *Zfy* transcripts to be distinguished. XXɎ mice were therefore mated with males of the CD1 strain which are known to carry a *domesticus* type Y.

Karyotypes were determined from tail tip cultures to distinguish the XɎ^{m}Y^d sons (data not shown) and Southern blot analysis was used to confirm the presence of the two types of Y chromosome by taking advantage of *Zfy-1* and *Zfy-2* *Taq*1 RFLPs (Mardon *et al.*, 1989). Figure 9 shows the distinct patterns derived from Y^m and Y^d and the combined pattern seen in DNA from mice carrying both chromosomes.

Using RNA isolated from the testis of an XɎ^{m}Y^d adult mouse an RT-PCR based strategy was employed to amplify a specific region of the *Zfy* gene transcripts known to contain sequence polymorphisms between Y^m and Y^d. Two oligonucleotide primers were designed which fulfilled the following criteria. (i) Both recognised identical sequences in *Zfy-1* and *Zfy-2,* according to all available sequence data (Ashworth *et al.*, 1989; Mardon & Page, 1989 and J. Collignon, unpublished results). (ii) Both were located in exons adjacent to the exon to be amplified (J. Collignon, unpublished results) to ensure that amplification of contaminating genomic DNA would either not occur, or if it did, would give rise to PCR products of a different size to the predicted band. (iii) Both oligonucleotides contained single mismatches to the known sequences, generating convenient restriction sites for subsequent cloning.

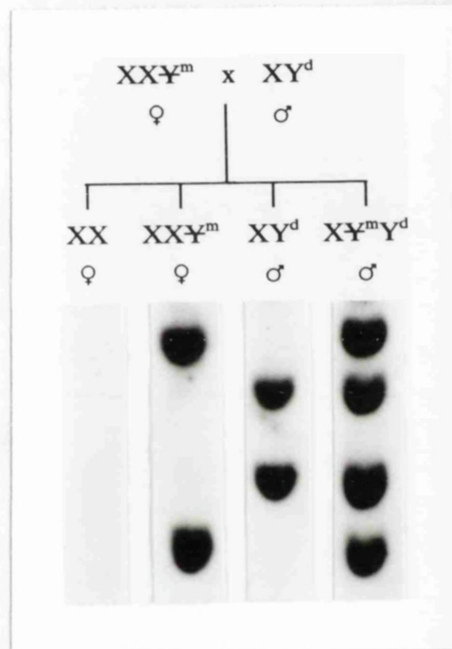


Figure 9. Identification of XɎ^{m}Y^d progeny by Southern analysis. Southern blots of *Taq*I digested DNA, were hybridized to the *Zfy* zinc finger probe (probe 3, legend to figure 4). All the bands seen are absent in females and are therefore Y-linked. Mice carrying Y^m show two bands of approximately 9kb and 4.5kb, mice carrying Y^d show bands of 7.5kb and 5.5kb. Mice with both Y chromosomes display all four bands.

```
Gene    Type    Sequence

        mus     GAAGACCAGAATGTCACTGCCAGTGCTCTCTTAAACCAAGATGAGTCTGG
Zfy-1           ||||| ||||||||||||||||||||||||||||||||||||||||||||
        dom     GAAGATCAGAATGTCACTGCCAGTGCTCTCTTAAACCAAGATGAGTCTGG
                ::::  ::::::: ::::::::::::: : ::: :::: ::::: :::::
        mus     GAAGTGCAGAATGCCACTGCCAGTGCTATGTTACACCATGATGAATCTGG
Zfy-2           |||||||||||||||||||||||||||||||||||||||||||| |||||
        dom     GAAGTGCAGAATGCCACTGCCAGTGCTATGTTACACCATGATGAGTCTGG
```

Number of clones of each type:

	mus	dom	Total
Zfy-1	8	4	12
Zfy-2	15	20	35
Total	23	24	47

Figure 10. Sequences obtained from RT-PCR analysis of *Zfy* transcripts from X$\cancel{Y}^{m}$$Y^{d}$ adult testes. Four different sequences were obtained. Critical polymorphisms between these sequences are shown in bold type. Homology between *Mus musculus musculus* (mus) and *Mus musculus domesticus* (dom) are indicated by | and between *Zfy-1* and *Zfy-2* by :. The number of clones of each sequence obtained indicates that the total number of *Zfy-1* and *Zfy-2* transcripts originating from each Y chromosome is equal and that the level of *Zfy-2* expression is three times that of *Zfy-1*.

RT-PCR produced a single band of the expected size which was subcloned into a pBluescript vector (see section 2.1). 47 subclones were sequenced and all could be assigned to one of four types, shown in figure 10. These corresponded to the sequences of both *Zfy-1* and *Zfy-2* of the *domesticus* type, to the *musculus* type *Zfy-1* sequence and to a fourth type which was almost identical to the Y^d *Zfy-2* sequence, except for one nucleotide difference. This last type is assumed to correspond to Y^m *Zfy-2* (the complete sequence of which is unknown) and that the single base difference is a *musculus/domesticus* polymorphism. This is a third base change and would not alter the amino acid sequence. The number of subclones corresponding to each type of sequence are also given in figure 10. Assuming that this reflects the relative proportions of the different types of transcript, then *Zfy-2* seems to be expressed at a three fold higher level than *Zfy-1*, in adult testis, which has also been observed by Nagamine *et al.* (1990). This result clearly shows that both *Zfy-1* and *Zfy-2* are transcribed from the mutant Y̶ at approximately the same rate as they are from the normal *domesticus* Y chromosome. It can be concluded from this that the *Tdy*^{m1} mutation has not affected the regulation of either gene in adult testis.

3.2.3 Expression of *Zfy* genes in normal male embryos

The Northern analysis (figure 8) failed to detect any *Zfy* expression in embryonic testes or genital ridge. This clearly does not satisfy the prediction that *Tdy* is most likely to be expressed in the genital ridge at or just before the first signs of testis differentiation. As Northern analysis is relatively insensitive it was still possible that a low level of *Zfy* expression in embryonic tissues could have gone undetected. RT-PCR is highly sensitive, can distinguish between genes with similar sequences and requires only small amounts of material. Thus Koopman *et al.* (1989) used such an assay to analyse *Zfy* expression in individual genital ridges and fetal testes. Koopman *et al.* (1989) were unable to detect *Zfy-2* transcripts at any of the embryonic stages tested, indicating that *Zfy-2* is not *Tdy*. Nagamine *et al.* (1990) have reported finding a very low level of *Zfy-2* expression in 12*dpc* fetal testes. A similar level of expression was seen in other fetal tissues (indicating that expression in the genital ridge is not much above the background expression level of this gene). These conflicting results may be due to strain specific differences or differences in

the sensitivity of the assay. Koopman *et al.* (1989) did find that *Zfy-1* is expressed in fetal gonads from 11.5*dpc* until at least 17.5*dpc*. They also demonstrated the existence of three different *Zfy-1* transcripts which are generated by alternative splicing within the 5' untranslated region. Thus *Zfy-1* is expressed, in the genital ridge, just before and during testis differentiation. However the finding that germ cells are responsible for *Zfy* expression in adult testis led Koopman *et al.* (1989) to test whether the same is true in the embryo. *Tdy* must be expressed in the somatic portion of the genital ridge as germ cells are not required for testis determination (see section 1.6). Koopman *et al.* (1989) analysed *Zfy-1* expression in fetal testes dissected from embryos homozygous for the extreme allele at the *W* locus (W^{e}). In such embryos, testes develop normally despite a complete lack of germ cells (see section 1.6). These testes were found to lack *Zfy-1* transcripts, demonstrating that *Zfy-1* expression in the embryo is also due to the presence of germ cells as in adult testis. In fact, the level of *Zfy-1* expression detected by Koopman *et al.* (1989) correlates well with the relative proportion of germ cells in the testis. The number of germ cells in the testis increases up to about 13*dpc* due to migration of germ cells into the genital ridge and cell division. *Zfy-1* expression is maximal at this stage. However at 13*dpc* the germ cells become arrested in the cell cycle (see section 1.3). The somatic portion of the gonad proliferates and it undergoes a considerable increase in size. The subsequent decrease in the *Zfy-1* signal is consistent with this decrease in the effective concentration of germ cells. The formation of testes in the absence of *Zfy* expression and the association of *Zfy-1* expression with germ cells, argues that *Zfy-1* is not *Tdy*.

3.2.4 Expression of *Zfy* genes in XY̶ female embryos

Having shown that the adult expression of *Zfy-1* and *Zfy-2* was unaltered in XY̶ female mice, it was necessary to test whether an alteration in *Zfy-1* expression in XY̶ female fetal gonads could be associated with their failure to develop as testes.

12.5*dpc* embryos were obtained from matings of an XY̶y male with CA strain females (see section 2.2). RNA was extracted from single gonads and analysed by an RT-PCR method which was specific for *Zfy-1*. Figure 11 shows the results of

this analysis on gonads from an X¥ female, XX female and Xy male from two mothers. *Zfy-1* expression in X¥ ovaries was indistinguishable from that in sibling Xy testes. The three bands seen in the Xy tracks are due to alternative splicing in the 5' non-coding region of *Zfy-1* (Koopman *et al.*, 1989). These three bands are also apparent in the X¥ tracks. This indicates that the embryonic regulation of *Zfy* genes is also unaltered in X¥ female mice. The finding that ovarian development occurs in X¥ female embryos despite expression of *Zfy-1* in the embryonic gonad, shows that *Zfy-1* expression is not sufficient for testis determination.

3.3 *Zfy* genes cannot be involved in testis determination

Any candidate for *Tdy* should display some form of abnormality in X¥ female mice, known to be mutant in *Tdy*. However, in this study no evidence was found to indicate that either the genomic organisation of the *Zfy* genes or their regulation is abnormal in X¥ females. The analysis of these genes was based on a consideration of the type of mutation that may have occurred in Tdy^{m1}. Initially *Zfy-1* and *Zfy-2* were tested for the presence of major rearrangements (insertions, translocations, inversions, duplications or deletions) by comparing X¥ female DNA against normal XY male DNA on Southern blots, using a variety of restriction enzymes. No differences were seen in *Zfy-1* and *Zfy-2* bands when normal males and X¥ females were compared. The use of probes representing a full length cDNA means that bands corresponding to all the exons of this transcript are represented on the blots. This implies that no gross alteration has occurred in the coding regions or adjacent intron sequences of *Zfy-1* and *Zfy-2*. The limit of resolution of this approach is estimated to be approximately 100bp. Any smaller mutation would not be detected unless it happened to alter the site for one of the restriction enzymes used. However, previous work using retroviral mutagens (the mutagen used during the generation of Tdy^{m1}) has implied that the most likely form of gene disruption to occur in these cases is a gross alteration, which would be visible on a Southern blot, rather than a point mutation. In most cases the disruption is caused by insertional mutagenesis involving whole virus or just viral LTRs. However, this is not always the case. In a similar mutagenesis protocol to that which gave rise to the Tdy^{m1} mutation, embryonic stem cells were selected for

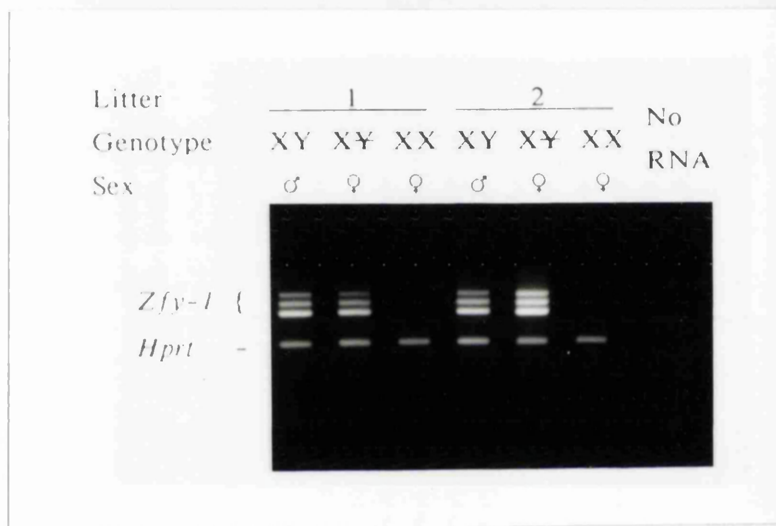


Figure 11. PCR analysis of *Zfy-1* expression in XɎ female, XY male and XX female 12.5*dpc* embryonic gonads. The three bands corresponding to *Zfy-1* expression were seen in RNA extracted from both normal XY male and mutant XɎ female embryonic gonads but not in normal XX females. The control *Hprt* band was present in all samples. Littermates from two separate litters were analysed.

mutations in the hypoxanthine phosphoribosyl transferase (*Hprt*) gene and a number of lines derived. As expected most of these lines carried inactivating viral insertions, but one line was found not to have viral sequences associated with the mutation and instead to have a 2kb sequence inserted into the fifth intron of the *Hprt* gene. The origin of this insertion is unknown (M. Evans, personal communication). In another study (Stocking *et al.*, 1988), growth factor independent variants of a myeloid precursor cell line were isolated after retroviral mutagenesis. In many cases activation of the Granulocyte Macrophage-Colony stimulating factor (GM-CSF) gene was found to have resulted from retroviral insertion into the locus, however in one case an intracisternal A particle (Dalton *et al.*, 1961) had integrated into the last exon of GM-CSF, producing a distinguishable band pattern compared to the wild type allele on Southern blots.

Alternatively Tdy^{m1} could be a mutation affecting the regulation of *Tdy*. While the Southern analysis would have detected gross alterations in 5' and 3' control sequences relatively near the exon sequences of the *Zfy* genes, it would not detect alterations in regulatory regions more distant from the genes. However the analysis of an adult X¥^{m}Y^{d} testis has shown that both *Zfy-1* and *Zfy-2* are capable of being transcribed from ¥ at apparently normal levels in adult testis, both in absolute terms and relative to each other, implying that the elements controlling the regulation of *Zfy* genes in the adult are unaltered in X¥ female mice.

Similarly, *Zfy-1* but not *Zfy-2* is expressed in 12.5*dpc* X¥ female embryonic gonads at the same level and displays an identical pattern of alternative splicing as normal male siblings. Thus the overall regulation of these genes seems to have been unaffected by the Tdy^{m1} mutation. These data strongly suggest that a defect in *Zfy-1* or *Zfy-2* is not the cause of the Tdy^{m1} mutation and consequently that neither of these genes can be *Tdy* (Gubbay *et al.*, 1990b).

The possibility still remains that there is a point mutation affecting the *Zfy-1* protein product. To exclude this, it would be necessary to determine full length cDNA sequences of *Zfy-1* from both the mutant ¥ chromosome and from a normal Y chromosome of *musculus* type (the only available full length sequences are from a

domesticus type Y chromosome). However additional overwhelming evidence supports the conclusion that *Zfy* genes are not involved in testis determination. (i) Theories of sex determination involving *ZFY* like genes must take into account the presence of an X-linked homologue in many species and an additional autosomal homologue in mice. There are also differences in the expression patterns and X-inactivation status of members of the family between man and mouse (see section 3.4). (ii) *ZFY* like sequences are not found on the sex chromosomes of marsupials (Sinclair *et al.*, 1988), which also seem to have a Y-dependent sex determining mechanism (see section 1.2). (iii) Efforts to find molecular abnormalities in *ZFY*, in human XY females, not deleted for *ZFY*, have been unsuccessful. In an analysis of nine such individuals Schneider-Gädicke *et al.* (1989b) showed that *ZFY* was transcribed in cultured cells isolated from these patients. These transcripts were indistinguishable in size from those of normal controls. However, this analysis does not exclude the possibility that missence or nonsense mutations are present in these transcripts or that these females may suffer from mutations in downstream sex determining genes rather than *TDF* itself. (iv) Koopman *et al.* (1989) (see section 3.2.3) have shown that *Zfy-2* is not expressed in the mouse genital ridge or fetal testis and that the expression of *Zfy-1* in the embryonic gonad is dependent upon germ cells, which are not required for testis determination. (v) Four XX male patients have been described that possess Y-linked sequences that do not include *ZFY*, indicating that the chromosomal location of *ZFY* is distinct from that of *TDF* (Palmer,MS *et al.*, 1989 and see chapter 5).

3.4 What is the true function of the *ZFY* gene family?

The actual functions of this gene family still remain a mystery. Some clues may be obtained by considering their expression patterns, especially in terms of differences between humans and mice. Both *ZFY* and *ZFX* are ubiquitously expressed in essentially all cells, suggesting a general cellular function rather than a role in a specific developmental event (Lau & Chan, 1989; Schneider-Gädicke *et al.*, 1989b). *ZFX* has been found not to be subject to X inactivation (Schneider- Gädicke *et al.*, 1989b) i.e. it escapes the process whereby one X chromosome becomes transcriptionally inactive in females in order to equalize the dose of X-linked genes

between the two sexes (Lyon, 1961). Recently Ashworth *et al.* (1991) have shown that in the mouse *Zfx* does undergo X inactivation.

In contrast to *ZFY* and the X-linked genes, mouse *Zfy-1* and *Zfy-2* have a very restricted pattern of expression. *Zfy-1* seems to be expressed in male germ cells in both the adult and the embryo, while *Zfy-2* is also expressed in male germ cells but only becomes abundant in the testis between 7-14 days postnatally (Koopman *et al.*, 1989; Nagamine *et al.*, 1989, 1990). In addition Koopman *et al.* (1989) have shown that *Zfy-1* but not *Zfy-2* is expressed in ES cells. Mardon & Page (1989) have argued that *Zfy-1* and *Zfy-2* may be functionally interchangeable, on the basis of their similar sequences. The differences in the expression patterns of the two genes can be interpreted in different ways. The two genes may have diverged sufficiently from a single ancestor, so that they have distinct functions associated with their respective sites of expression. Alternately, they may play identical roles but their regulatory sequences have diverged resulting in overlapping but slightly different patterns of expression.

Although the pattern of expression of the *ZFY* gene family is different in man and mouse, if all these genes are considered to be interchangeable, then equivalent dosage of the family as a whole is maintained between the sexes in both cases, albeit in different ways. In humans where *ZFX* does not undergo X inactivation, *ZFY* is ubiquitously expressed. Thus ensuring that two *ZFY* like genes are functional in every cell in both sexes. While in mouse where *Zfx* does undergo X inactivation, the *Zfy* genes (and *Zfa*) are not expressed in the vast majority of cells. In fact *Zfy-1* expression seems to correlate with cell types that in the female would not show X inactivation, such as ES cells and germ cells within the fetal gonad. It is possible that expression occurs in these sites simply to compensate dosage between males and females.

Alternatively, selective pressure to maintain these genes on the Y chromosome during mammalian evolution, suggests a distinct male specific function. The expression of *Zfy* genes in male germ cells may reflect a role in spermatogenesis. The deletion event which gave rise to the Sxr' portion of the Y chromosome from

Sxr, also resulted in the loss of *Spy*, a gene involved in spermatogenesis (see section 1.5). As this deletion involved a fusion of the *Zfy-1* structural gene to the *Zfy-2* promoter which may interfere with the normal function of both genes, and as these genes are expressed during spermatogenesis, Koopman *et al.* (1991a) have suggested that one or both *Zfy* genes may be *Spy*. Recently a new candidate gene for *Spy* has been described. *A1S9* (Zacksenhaus & Sheinin, 1988, 1990; Brown *et al.*, 1989) is a human X-linked gene which complements a temperature sensitive cell cycle mutation in mouse L cells. It encodes the ubiquitin activating enzyme, E1 (Hatfield *et al.*, 1990; McGrath *et al.*, 1990; Handley *et al.*, 1991), which catalyses the first step in ubiquitin-mediated protein degradation in eukaryotic cells (Reichsteiner, 1988). Kay *et al.* (1991) and C. Bishop (personal communication) have shown that two copies of this gene map to the Sxr portion of the mouse Y chromosome. One of these, *A1S9Y-2*, is retained by Sxr' and appears to be a non-functional pseudogene. The other, *A1S9Y-1*, is lost from Sxr' and is expressed in adult testis. The mapping and expression data regarding *A1S9Y-1* is consistent with it being *Spy* (as is the case for the *Zfy* genes). In addition, its role in the cell cycle (see above), makes it an attractive candidate for a gene whose deficiency seems to block the proliferation of germ cells.

Another gene which maps to Sxr and is deleted in Sxr' is *Hya*. Simpson,E *et al.* (1987) have typed a panel of sex-reversed XX male human individuals for H-Y antigen. All of the individuals had arisen through aberrant X-Y interchange. In this way the gene for H-Y antigen was found to map to the long arm or centromeric region of the Y chromosome. This mapping separated the H-Y antigen gene from *TDF*, as XX males were found who lacked H-Y antigen, and from *ZFY* which maps much nearer to *TDF*. In the mouse, the Sxr' mutation has shown that *Tdy* is distinct from *Hya*. Although, in this case, the genetic data is not inconsistent with either *Zfy-1* or *Zfy-2* being *Hya*. However this seems unlikely, given the restricted pattern of *Zfy* gene expression compared to the widespread detection of H-Y antigen (Johnson,LL *et al.*, 1981), which presumably reflects a similarly widespread expression pattern for *Hya*.

Tsunoda *et al.* (1985) have reported a growth rate difference between XX and XY preimplantation embryos. Burgoyne (1991) have shown that this effect depends on the strain of origin of the Y chromosome and have therefore postulated the existence of a Y-linked gene responsible for accelerating the early development of XY blastocysts. The location of this gene on the Y chromosome is unknown, however the finding that *Zfy-1* is expressed in ES cells which are derived from, and approximate to, cells of the inner cell mass of the blastocyst would be consistent with it being responsible for this effect. This hypothesis could be pursued by correlating the growth rate variations associated with different Y chromosomes with differences in *Zfy-1*.

The presence of two *Zfy* genes in the mouse is due to a duplication event which involves at least one other gene, *A1S9Y* (Bishop,CE & Mitchell, 1991; Kay *et al.*, 1991; Simpson,EM & Page, 1991). The finding that only *A1S9Y-1* is functional and that *A1S9Y-2* is a pseudogene, suggests that a second copy of this gene is superfluous and is now in the process of accumulating mutations. The same could be true of the *Zfy* genes. *Zfy-2* from the *musculus* Y chromosome (Ashworth *et al.*, 1990) has been found to have a six amino acid deletion, which may affect the coordination of the zinc ion in one of the fingers. Whether the protein produced is functional is unknown. However one interpretation is that one *Zfy* gene is redundant in the mouse and can therefore accumulate mutations. By this argument, the fact that *Zfy-2* is only expressed in a subset of the sites where *Zfy-1* is expressed may reflect a gradual loss of the appropriate *Zfy-2* regulatory elements by mutation. This hypothesis could be tested by inactivation of one of the *Zfy* genes by gene targeting. Whether any of the sequences that map to this duplication conferred a selective evolutionary advantage by being present in two copies, over ancestral mice which did not carry the duplication, is not clear.

Chapter 4: Results and Discussion

Analysis of endogenous retroviral elements in X¥ female mice

The strategy from which the Tdy^{m1} mutation arose was based on using a retroviral vector as an insertional mutagen. However attempts to correlate a retroviral or retroviral LTR insertion with this mutation have failed (Lovell-Badge & Robertson, 1990 and see section 1.7.1). The mutation could therefore have been a completely random event or may have arisen as a result of the retroviral infection strategy, for example by the mobilisation of an endogenous proviral element. Such an element would be a convenient marker for cloning *Tdy*. There is evidence to suggest that some naturally occurring mutations in mouse have been caused by such a mechanism (Jenkins *et al.*, 1981; Stoye *et al.*, 1988) and the retroviral infection strategy used during the generation of Tdy^{m1} may have promoted retroviral mobilisation (Lovell-Badge & Robertson, 1990). To test this possibility the pattern of proviral LTRs in DNA from a panel of X¥ females was analysed by Southern blotting. As there are on average 40-60 non-ecotropic C-type proviral insertions in an inbred mouse genome, a set of four oligonucleotides, designed by Stoye *et al.* (1988), each of which only recognises a subset of these proviral LTRs, was used as probes in order to simplify the analysis. Figure 12 shows Southern blots with DNA from a range of mice, many of which carry the mutant ¥ chromosome, hybridized to these oligonucleotides. Each band corresponds to a single LTR insertion. Each oligonucleotide seems to recognise 15-20 insertion sites. No band can be seen to segregate exclusively with the ¥ chromosome, neither can the absence of a band be correlated with this chromosome. This result suggests that the Tdy^{m1} mutation was not caused by mobilisation of an endogenous proviral LTR.

Figure 12. Analysis of endogenous retroviral insertions in X¥ female mice. Southern blots of *Eco*RI digested DNA isolated from a number of different normal female, normal male, X¥ female, XX¥ female, X¥y male and Xy male mice were hybridized to the oligonucleotides JS4, JS5, JS6 and JS10 each of which recognise a subset of nonecotropic C-type endogenous retroviruses (see section 2.7 for sequences). DNA size markers shown are λ DNA digested with *Hind*III.

23.0
9.4
6.6
4.3
2.3
2.0

Probe JS4

23.0
9.4
6.6
4.3
2.3
2.0

Probe JS5

XX*♀ XX*♀ XX♀ XX*♀ Xy♂ XX*y♂ XX*y♂ XX*♀ XX*♀ XX*♀ XX*♀ XX*♀ XX*♀ XX♀ XX♀ XX♀ XX♀ XY♂

23.0
9.4
6.6
4.3
2.3
2.0

Probe JS6

XX*♀ XX*♀ XX♀ XX*♀ Xy♂ XX*y♂ XX*y♂ XX*♀ XX*♀ XX*♀ XX*♀ XX*♀ XX*♀ XX♀ XX♀ XX♀ XX♀ XY♂

23.0
9.4
6.6
4.3
2.3
2.0

Probe JS10

The failure to find any proviral insertion that correlates with Ɏ suggests that the mutation must have arisen by another mechanism or that an endogenous provirus from another family may be responsible. However, the total number of endogenous retroviruses in the mouse genome may be up to 1000 (Stoye & Coffin, 1985) and cannot be easily distinguished. Retroviral like intracisternal A particles (IAPs) are also present in 1000-2000 copies per haploid genome (Kuff & Lueders, 1988). Brilliant *et al.* (1991) have used this sequence to identify DNA sequences associated with the mouse Pink-Eyed Unstable (p^{un}) mutation by "genome scanning". This method is based on Southern hybridization using an IAP probe. Mutations caused by large deletions or rearrangements may be detected as band differences on such Southern blots. Preliminary analysis using an IAP probe (obtained from M. Brilliant) have not revealed any differences in the hybridization patterns of DNA from XY males compared to XɎ females (data not shown).

Chapter 5: Results and Discussion

Cloning and sequence analysis of the mouse *Sry* gene

5.1 Cloning a new candidate for *TDF/Tdy*

In 1989 Palmer,MS *et al.* described four human XX male individuals who did not possess the *ZFY* gene but were positive for the Y chromosome-derived pseudoautosomal boundary. These individuals must therefore have inherited *TDF* by aberrant X-Y interchange but yet lacked *ZFY.* This data provides strong genetic evidence that *ZFY* is not *TDF.* In addition it suggests that *TDF* actually lies somewhere between *ZFY* and the pseudoautosomal boundary. This was further confirmed by the finding that the XY female patient that had been used in the previous mapping of *TDF* (Page *et al.*, 1987), which was deleted for the region of the Y chromosome containing *ZFY*, possessed additional deletions nearer to and including the pseudoautosomal boundary. Ellis *et al.* (1989) had previously used chromosome walking to clone the Y chromosome-derived pseudoautosomal boundary and 110kb of the flanking Y-specific region which is between *ZFY* and the boundary. Palmer,MS *et al.* (1989) used a series of probes, isolated from this chromosomal walk, to map the region of the Y chromosome present in these XX males. All the breakpoints were found to cluster about 60kb away from the pseudoautosomal boundary. Sinclair *et al.* (1990) subsequently used additional probes from this region to show that these XX males, in fact, possessed no more than 35kb of Y-specific sequences. *TDF* had to lie within this short region.

Various strategies are available for the isolation of transcribed (genic) sequences from cloned genomic DNA (reviewed by Hochgeschwender & Brennan, 1991). These include cDNA library screening, Northern blot hybridization and exon trapping (Duyk *et al.*, 1990). In this particular case, a search for sequences exhibiting cross-species evolutionary conservation was considered most appropriate for two main reasons: (i) although most transcribed sequences are likely to show some conservation amongst closely related organisms, the testis-determining gene was expected to be both highly conserved and Y-linked in all eutherian mammals, as all these species are thought to show Y-dependent sex determination and (ii) the

mouse system can be exploited by simultaneously testing if sequences map to Sxr' where *Tdy* is known to map.

Sinclair *et al.* (1990) performed this search by subcloning DNA from the 35kb region into fragments of about 4kb in size and then using frequent cutting restriction enzymes to generate 50 smaller fragments (0.5-2kb) covering the entire region. These were used as probes on Southern blots of DNA from human males and females and bovine males and females. Most of the probes used failed to detect single copy sequences in the human genome. In fact only seven probes detected single-copy Y-specific bands in human DNA. Of these a single 2.1kb *Hind*III restriction fragment known as pY53.3 hybridized to a Y-specific band in the bovine as well as the human genome. The conserved sequences responsible for this hybridization were found to map to a 0.9kb *Hinc*II fragment within pY53.3. To assess the involvement of this sequence in sex determination, the 0.9kb *Hinc*II fragment from pY53.3 was used to probe a Southern blot of mouse genomic DNA digested with the restriction enzyme *Eco*RI. A 3.5kb band is present in DNA from the normal male but absent in DNA from the female (see figure 13b). This indicates that a homologue of pY53.3 exists on the Y chromosome of the mouse. The 3.5kb band is also present in DNA from XXSxr' males showing that it maps to Sxr'. Figure 13a shows the same filter rehybridized to a *Zfy* probe. The two male specific bands correspond to *Zfy-1* and *Zfy-2*. Using this probe, only the band corresponding to *Zfy-1* is visible in XXSxr' DNA. This mapping data indicates that pY53.3 defines a single copy sequence present in the minimal portion of the Y chromosome known to be sex determining in both humans and mice.

Figure 13b also shows that the 3.5kb male specific band hybridizing to pY53.3 is absent in DNA from XɎ female mice. Because there are no additional bands which would be expected if this region were polymorphic, this result shows that the hybridizing sequences have been deleted in the XɎ females. This is in contrast to the bands corresponding to *Zfy-1* and *Zfy-2* which are both present in XɎ female DNA (figure 13a). pY53.3 is the first probe to detect a molecular abnormality in the mutant Ɏ chromosome. This indicates that the sequence on the mouse Y

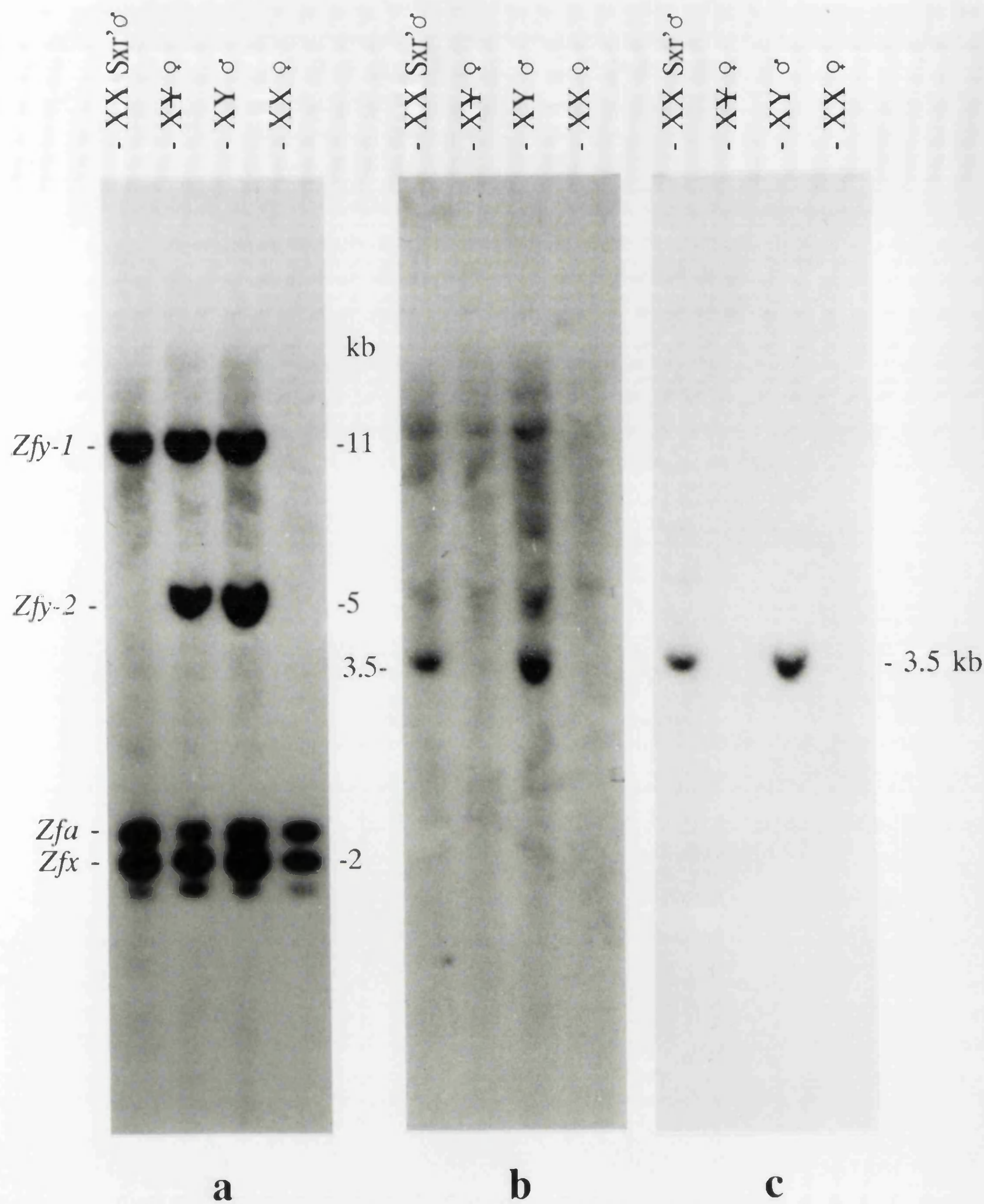


Figure 13. Mapping conserved sequences to the sex-determining region of the mouse Y chromosome. A Southern blot of *Eco*RI-digested genomic DNA from an XXSxr' male, an XɎ female, a normal XY male and an XX female mouse was hybridized consecutively to three separate probes. a, The *Zfy-1* zinc finger probe (probe 3, legend to figure 4), which reveals bands corresponding to each of the four *Zfy*-related genes in XY male DNA. All of these genes are also seen in the XɎ female track. *Zfy-1*, but not *Zfy-2*, is shown to map to Sxr' in the XX Sxr' track, while only *Zfa* and *Zfx* are present in a normal XX female. b, A 0.9kb *Hinc*II fragment of pY53.3 which hybridizes strongly to a 3.5kb band present in DNA from XY male and XX Sxr' male mice but absent from XX female and XɎ female mice. A number of additional, weakly hybridizing bands are present in all tracks, so cannot be Y-linked. c, A mouse Y chromosome derived clone, p422.04, containing sequences homologous to pY53.3, which hybridizes at high stringency only to the 3.5kb Y-linked band.

chromosome homologous to pY53.3 is at least the closest marker to *Tdy* and is consistent with it being *Tdy*.

Also visible in figure 13b are a number of faintly hybridizing bands present in all tracks. These loci cannot be Y-linked, as they are present in DNA from normal females. This indicates that the mouse genome contains additional loci, which are not Y-linked, that share sequence similarity with pY53.3. Although the possibility of X-linkage has not been formerly excluded, these loci will be subsequently referred to as autosomal.

5.2 Cloning the Y-linked mouse homologue of pY53.3

Two mouse genomic phage libraries and one cosmid library were screened with pY53.3. Although several clones were isolated, none of them contained the expected 3.5kb *Eco*RI fragment or mapped to the Y chromosome (data not shown). They seemed instead to correspond to some of the autosomal loci seen on Southern blots (figure 13b). The difficulty in cloning the Y sequence may have been due to a general under-representation of Y-chromosomal DNA in genomic libraries or to a specific instability of the clone of interest in amplified libraries. Size-selected DNA was therefore used in a forced cloning strategy to isolate the 3.5kb *Eco*RI Y-specific fragment (see section 2.9.1). Three clones, out of a total of $5x10^5$ screened, were obtained which were 3.5kb *Eco*RI fragments as expected and that hybridized strongly to pY53.3. One of these clones, p422, was analysed further and shown to contain a 374 base pair (bp) *Bgl*II-*Pst*I restriction fragment (clone p422.04), that retained homology with pY53.3. When this 374bp fragment was used as a probe at high stringency on genomic Southerns it hybridized only to the Y-specific 3.5kb *Eco*RI band (figure 13c). Clone p422 therefore contains the Y-linked mouse homologue of pY53.3.

5.3 Sequencing the Y-linked mouse homologue of pY53.3

Clone p422.04 was fully sequenced. Figure 14 shows a nucleotide comparison of the mouse and human sequences over the entire length (374bp) of this clone. The

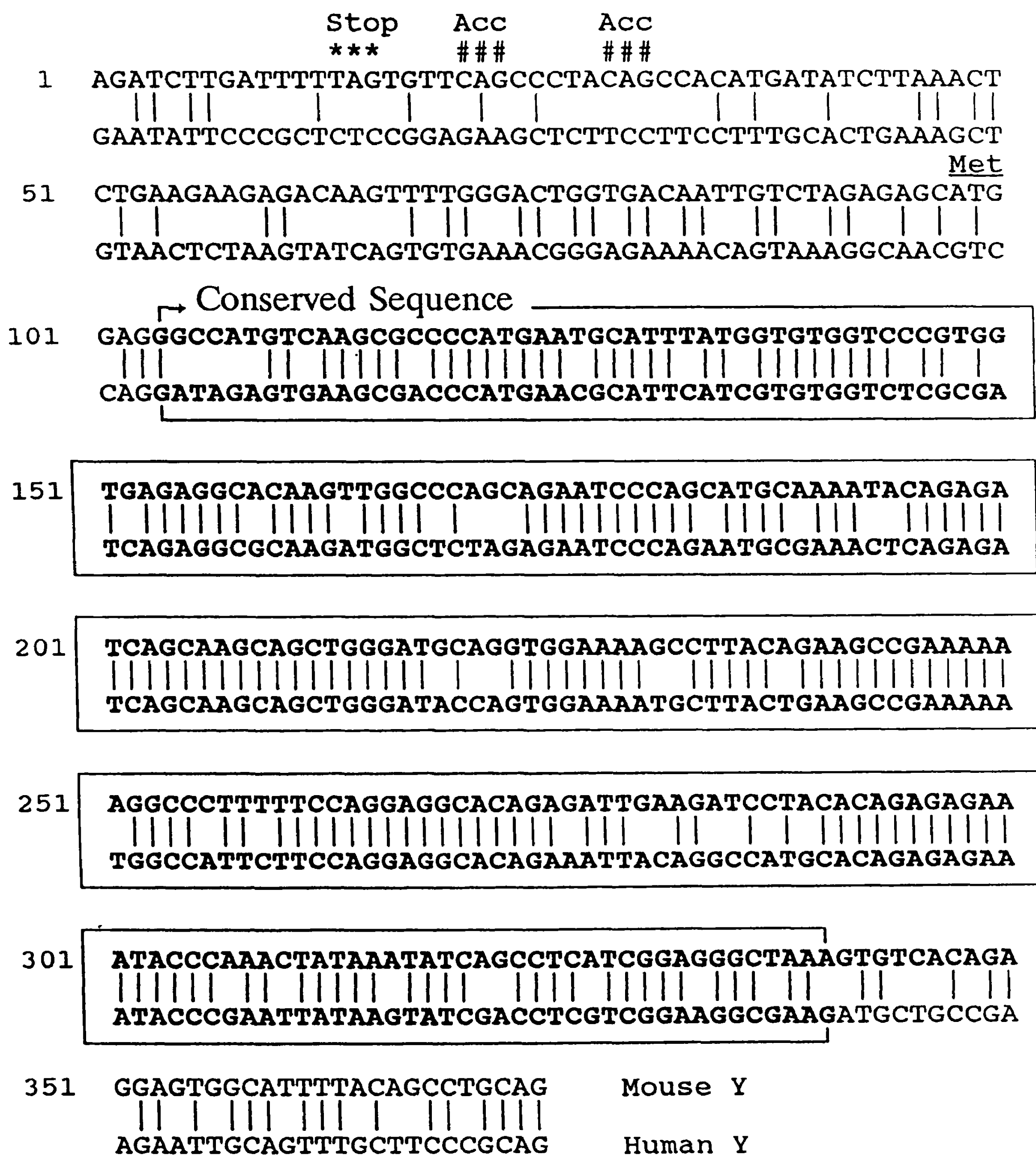


Figure 14. Nucleotide comparison of mouse and human Y-linked sequences. The sequences of p422.04 (Mouse Y) and pY53.3 (Human Y) are shown on the upper and lower lines respectively. Nucleotide homology is indicated by vertical bars. The single open reading frame in the mouse sequence is defined at the 5' end by a stop codon at nucleotides 14-16 (***). Within this open reading frame are two possible splice acceptor sites (###) and an in-frame translational start codon (Met). The region of greatest amino acid homology between the two sequences is boxed and marked Conserved Sequence.

overall degree of homology is 62%. Of the six possible translation frames in p422.04, only one contained a long open reading frame consisting of 119 amino acids. Part of this open reading frame corresponds to one of the two open reading frames in pY53.3, and to an open reading frame in the corresponding Y-linked sequence from rabbit DNA (Sinclair *et al.*, 1990).

The nucleotide homology between the mouse and human open reading frames rises to 80% over a 237bp sequence (corresponding to 79 amino acids). This sequence is indicated in figure 14 (Conserved Sequence). The appropriate codon usage has been found for both the human and mouse open reading frames. This and the high degree of evolutionary conservation of the open reading frame are strong indications that it is part of a functional gene. This gene has been called *Sry* (*SRY* in humans) (a gene from the Sex-determining Region of the Y-chromosome) (Gubbay *et al.*, 1990a; Sinclair *et al.*, 1990).

The human open reading frame consists of 223 amino acids (202 amino acids from the first ATG codon to a stop codon). The sequence of cDNAs generated by 3' rapid amplification of cDNA ends (RACE) -PCR experiments has been found to be colinear with the genomic sequence of pY53.3 except for the presence of a poly(A) tract, after the 3' stop codon of the open reading frame. This suggests that the open reading frame in pY53.3 corresponds to the last exon of *SRY* (Sinclair *et al.*, 1990). Preliminary results of cDNA cloning, 5' RACE-PCR and RNAase protection experiments (P. Goodfellow, personal communication) suggests that *SRY* may consist of a single exon encompassing the entire open reading frame in pY53.3.

Sry cDNA clones have not yet been isolated. Therefore the complete structure of this gene remains unknown. The open reading frame in p422.04 contains an ATG codon, at its 5' end, in a reasonable context for translational initiation (Kozak, 1987) which is preceded by an in frame stop codon, and two potential intron splice acceptor sites (see figure 14). If either of these splice acceptor sites are utilized *in vivo* then the methionine residue at the 5' end of p422.04 may not represent the amino terminus of the *Sry* protein, as additional coding exons may be spliced onto this end. However if the in frame stop codon is not spliced out then this

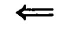
methionine must initiate the protein. The stop codon, potential splice acceptor sites and methionine at the 5' end of p422.04 have no cognates in pY53.3. In fact the two sequences diverge considerably 5' of this methionine. As these sequences are probably present within the human *SRY* transcript (see above), this loss of conservation can be interpreted in at least two ways: (i) that the sequences at the 5' end of p422.04 belong to an intron or (ii) that the mouse and human transcripts, and possibly proteins, differ 5' to the conserved region. From a comparison of sequence flanking p422.04 (in clone p422) to the sequence of pY53.3 it has been found that nucleotide homology between the two sequences also decreases considerably 3' to the conserved domain (see figure 14 and data not shown), which can be interpreted in a similar manner.

5.4 *Sry* is a DNA binding protein

One way of assessing the function of a gene is to compare its sequence to that of other cloned genes. When the predicted amino acid sequence of *Sry* was used to search a protein sequence database, homology was found to a number of other proteins. These include HMG (high mobility group) proteins , such as HMG1 (Wen *et al.*, 1989), the Mc mating-type protein of *Schizosaccharomyces pombe* (*S. pombe*) (Kelly *et al.*, 1988), human upstream binding factor (hUBF) (Jantzen *et al.*, 1990), T-cell factor 1 (TCF-1) (van de Wetering *et al.*, 1991), mouse lymphoid enhancer binding factor 1 (LEF-1, and its equivalent in humans TCF-1α) (Travis *et al.*, 1991; Waterman *et al.*, 1991) and human mitochondrial transcription factor 1 (mtTF1) (Parisi & Clayton, 1991). These homologies are shown in figure 15. The homologous domain within all these proteins was first recognised in the HMG proteins. Jantzen *et al.* (1990) have referred to this domain as the "HMG-box". This 79 amino acid domain also corresponds to the region of greatest conservation between mouse and human *SRY*.

Figure 15. Conservation of homologous protein domains. Amino acid sequences are compared between the homologous regions of the following sequences: *Sry*; *SRY*; TCF-1 (van de Wetering *et al.*, 1991); LEF-1 (Travis *et al.*, 1991); Mc, mating type protein Mc from *S.pombe* (Kelly *et al.*, 1988); UBF, the third HMG box of hUBF (Jantzen *et al.*, 1990) which shares the highest homology to the other proteins; mtTF1 (Parisi & Clayton, 1991) and huHMG1, DNA binding domain of human HMG1. Residues which are absolutely or highly conserved between 5 or more of the 8 sequences are shaded. The following amino acid pairs were considered highly conserved: P/G, K/R, D/E and F/Y. The beginning and end of the 79 amino acid domain, known as the HMG box, is marked by arrows above the sequence.

```
                 ⇒
Sry     TGDNCLESME GHVKRPMNAF MVWSRGERHK LAQQNPSMQN TEISKQLGCR
SRY     TGENSKGNVQ DRVKRPMNAF IVWSRDQRRK MALENPRMRN SEISKQLGYQ
TCF-1   .......... ..IKKPLNAF MLYMKEMRAK VIAECTLKES AAINQILGRR
LEF-1   ..RKEQEPKR PNIKKPLNAF MLYMKEMRAN VVAECTLKES AAINQILGRR
Mc      KSLRKDTTST ERTPRPPNAF ILYRKEKHAT LLKSNPSINN SQVSKLVGEM
UBF     ......KGGS EKPKRPVSAM FIFSEEKRRQ LQEERPELSE SELTRLLARM
mtTF1   .......... ..PKKPVSSY LRFSKEQLPI FKAQNPDAKT TELIRRIAQR
huHMG1  .......... NAPKRPPSAF FLFCSEYRPK IKGEHPGLSI GDVAKKGEMW
```



```
Sry     WKSLTEAEKR PFFQEAQRLK ILHREKYPNY KYQPHR.RAK VSQRSGILQP
SRY     WKMLTEAEKW PFFQEAQKLQ AMHREKYPNY KYRPRR.KAK MLPKNCSLLP
TCF-1   WHALSREEQA KYYELARKER QLHMQLYPGW SARDNY.GKK
LEF-1   WHALSREEQA KYYELARKER QLHMQLYPGW SARDNY.GKK KKRK
Mc      WRNESKEVRM RYPKMSEFYK AQHQKMYPGY KYQPRKNKVK R
UBF     WNDLSEKKKA KYKAREAALK AQSERKPGGE REERGK
mtTF1   WREIPDSKKK IYQDAYRAEW QVYKEEISRF KEQLT
huHMG1  NNTAADDKQ. PYEKKAAKLK EKYEKDIAAY RAKGKPD
```

Although the role of all these genes is not yet fully understood it seems that the proteins which they encode may all bind DNA and that a subset of these proteins bind in a sequence specific manner.

HMG proteins are non-histone components of chromatin, which interact with both DNA and histones. While they are not sequence-specific DNA-binding factors, some display enhanced DNA-binding interactions with certain sequences such as AT-rich stretches (Wright,JM & Dixon, 1988).

The Mc protein is involved in mating type determination in fission yeast (Kelly *et al.*, 1988). Cell type in *S.pombe* is determined by the presence of one of two alleles (M or P) at the *mat1* locus. Two other loci, *mat2-P* and *mat3-M* serve as donors of information that are transposed to *mat1* during a switch of mating type. This pattern of mating type interconversion is formally similar to that in the budding yeast *Saccharomyces cerevisiae* (*S.cerevisiae*) and has been referred to as a cassette mechanism (Strathern *et al.*, 1981). There are two genes at each mating-type locus *Mc*, *Mi* and *Pc*, *Pi*. *Mc* and *Pc* have been found to be necessary and sufficient to confer the appropriate mating phenotype and to allow conjugation of haploid cells but all four genes are required for meiosis of diploid cells. *S.cerevisiae* has three mating-type genes *α1*, *α2* and *a1* (Astell *et al.*, 1981). The only sequence homology that could be detected between any of these genes was a putative diverged homeobox sequence shared by *Pi* and *α2*. The homeobox is a highly conserved protein domain involved in DNA binding (Gehring, 1987). In *S.cerevisiae* mating type genes act to regulate the expression of unlinked genes (Bender & Sprague, 1987; Hall & Johnson, 1987). By analogy it seems likely that *S.pombe* mating type genes (including *Mc*), also act as transcriptional regulators, although this has not been formally proved. Interestingly the *mt a1* gene from the a mating type region of *Neurospora crassa* (*N.crassa*), also contains an HMG box (Staben & Yanofsky, 1987). Unlike yeast this organism does not undergo mating type switching and does not carry silent copies of mating type information. Thus in *N.crassa*, a single meiotic product (ascophore) cannot become operationally "self-fertile" by switch of mating type as occurs in yeast (Glass *et al.*, 1988). The remarkable conservation of this domain, given the evolutionary divergence between yeast, fungi and mammals

suggests that it is an important functional domain. Its occurrence in genes involved in diverse mating type and sex determination processes is not thought to imply any evolutionary conservation of mechanism, but merely to reflect common use of a specific functional domain.

hUBF contains four HMG domains (Jantzen *et al.*, 1990). hUBF is a cofactor of RNA polymerase I, essential for the transcription of tandemly arranged genes encoding the large ribosomal RNA (rRNA) precursors. hUBF has been shown to bind to elements within the rRNA promoter. This sequence-specific binding is retained in deletion mutants containing only a single HMG box, but is abolished in mutants lacking all four HMG boxes, indicating that DNA binding by hUBF involves the HMG domain. hUBF shares greater homology with mtTF1 than any of the other HMG box containing proteins. Both proteins seem to play similar roles as cofactors of transcriptional machinery. mtTF1 binds to the two promoters of human mitochondrial DNA in a sequence specific manner. It can then stimulate accurate transcriptional initiation by the mitochondrial RNA polymerase.

The remaining sequences belong to an increasingly large class of more *bona fide* transcription factors. TCF-1 is a T-cell specific factor that has been shown to bind to enhancer sites for other T-cell specific genes such as *CD3-ε*, T-cell receptor α (*TCRα*), *TCRβ* and *TCRδ* (Oosterwegel *et al.*, 1991; Waterman *et al.*, 1991; van de Wetering *et al.*, 1991). TCF-1 binds specifically to a seven base pair motif represented by the consensus AACAAAG.

Another HMG box containing protein which may be involved in haematopoiesis is known as LEF-1, which is presumed to be the murine equivalent of TCF-1α as their protein sequences are 97% identical. LEF-1 is expressed in pre-T and -B cells (Travis *et al.*, 1991), while Waterman *et al.* (1991) have reported that TCF-1α is T-cell specific and not expressed in mature B-cells. However the latter authors have not analysed pre B-cells, which may explain this discrepancy. The HMG box of LEF-1/TCF-1α differs from that of TCF-1 by only a single amino acid residue. However these proteins do differ outside the HMG box. TCF-1α interacts with the canonical sequence CANAG, which includes the TCF-1 binding site. Transfection

studies (Waterman *et al.*, 1991) have shown that TCF-1α causes up to 20 fold stimulation of transcription of a reporter gene with a single copy of the *TCRα* enhancer positioned upstream.

The homology of the HMG box in *Sry* to other HMG box containing proteins is relatively low (56% for the Mc protein and 47% for the third (most homologous) HMG box of hUBF, at the amino acid level). An additional group of homologous sequences exist which are much more strongly related to *Sry* in their HMG domains. As shown in this study, bands corresponding to these genes are visible on genomic Southern blots probed with *Sry* (see figure 13b). Four of these genes were initially cloned when an 8.5*dpc* whole embryo cDNA library was screened with *Sry* in attempts to obtain *Sry* cDNAs (Gubbay *et al.*, 1990a). These genes are known to be part of a large gene family in the mouse and to have closely related homologues in other organisms (A. Ashworth, J. Collignon, personal communication and Griffiths, 1991). As the highly conserved HMG box of the human and mouse *SRY* genes is referred to as the *Sry* box, these *Sry*-related genes have been termed *Sox* (*Sry* box) genes. The expression patterns of three of these genes during mouse embryogenesis have been studied in some detail (J. Collignon and R. Lovell-Badge, manuscript in preparation). *Sox*-1, -2 and -3 are early markers of neuroectoderm, which can be detected at 8*dpc*. Between 8*dpc* and 13.5*dpc* *Sox*-1 is expressed in the central nervous system (CNS), with some patches of non-expression in the mid- and hindbrain. *Sox*-2 is expressed throughout the CNS and in the trigeminal, geniculate and spinal ganglia, which are all derived from the neural crest, the endodermal lining of the gut and all the sensory placodes. These genes are also expressed in the developing eye. *Sox*-2 is initially expressed all over the developing eye. Subsequently expression in the lens vesicle is down-regulated and *Sox*-2 expression becomes restricted to what will become the neuronal layer of the retina. Simultaneously *Sox*-1 expression becomes upregulated in the lens vesicle. These expression patterns suggest a role for at least some of these genes in developmental processes. The strong homology of *Sox* genes to *Sry* (80-85% at the protein level) suggests that the latter is in fact a member of this gene family.

From the cases described above where functional data is available, it seems that HMG box containing proteins have diverse modes of action, although DNA binding seems to be a common factor in each case. This variability is presumably influenced by the exact sequence of the HMG box, which differs considerably in some cases (see figure 15) and the sequence of protein regions outside the HMG box. Harley *et al.* (1991) and S. Sockanathan (unpublished results) have tested whether *SRY* function is analogous to HMG box containing proteins which are transcription factors. They have shown that both mouse and human *SRY* can bind to DNA containing a TCF-1 binding motif *in vitro*. Nasrin *et al.* (1991) have shown that both *Sry* and a rat *Sox* gene which is most similar to mouse *Sox-4* can bind *in vitro* to the insulin response element (IRE-A) of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene. The protein product of the rat *Sox* gene has therefore been called IRE-A DNA binding protein (IRE-ABP). The sequence motif, within IRE-A, to which *Sry* and IRE-ABP bind contains the consensus binding site for TCF-1α which is also present within the TCF-1 consensus binding site. The finding that four different HMG box containing proteins (TCF-1, TCF-1α, IRE-ABP and *Sry*) can bind *in vitro* to enhancer elements belonging to disparate genes, but sharing a common sequence motif, indicates that binding to this motif *in vitro* may be a general property of these proteins. This sequence motif may not represent the *in vivo* binding site for *Sry* as the specificity of *Sry in vivo* will depend on the cellular environment, for instance the presence of other proteins which may cooperate with *Sry* in DNA binding or the accessibility of target sequences in the cells where *Sry* acts. Nevertheless, the possibility that *Sry* may be a sequence specific DNA binding protein is attractive for two reasons. First, such a role is consistent with *Sry* acting cell autonomously, which has previously been proposed for *Tdy* (see section 1.6). Second, a molecule responsible for a developmental switch, such as testis determination, is likely to regulate other genes in the developmental cascade. One way that this could occur is through the action of a transcription factor. Distinguishing the *in vivo* target sites of *Sry* from those with which it interacts *in vitro* (the former will presumably be a subset of the latter) will require additional data, such as expression analysis and targeted inactivation of candidate genes.

Chapter 6: Results and Discussion

Analysis of the genomic structure of *Sry*

6.1 Cloning a 14.5kb genomic fragment containing *Sry*

In order to extend the analysis of *Sry* it was considered important to obtain an *Sry* clone spanning a greater region of the genome than p422. Therefore clone p422.04 was used to screen an unamplified genomic library constructed from male mouse DNA (see section 2.9.2) and plated using bacterial strain DL652. This strain stabilises end to end repeat sequences, and was used in order to overcome the previous problems experienced in cloning this locus. From 10^6 clones, one was obtained which showed very strong homology with p422.04. Restriction mapping confirmed that this clone, L741, contained the 3.5kb *Eco*RI band corresponding to p422 and to the mouse Y-linked locus. Clone L741 has been used for three separate studies. (i) To investigate the deletion of *Sry* in XY̶ female mice (chapter 7). (ii) To analyse *Sry* function by the construction of transgenic mice (chapter 8). (iii) To study the structure of the *Sry* genomic locus.

6.2 The *Sry* locus consists of an inverted duplication

Initial characterization of L741 by restriction mapping revealed the existence of identically sized fragments for a number of restriction endonucleases. Figure 16 shows an example of this using the enzyme *Eco*RI. In order to further characterise the locus, probe p2.3/A, figure 16) which is from one end of L741, was used to isolate additional phage clones from a mouse genomic library. p2.3/A was chosen as a probe, as it did not hybridize to repetitive sequences in the mouse genome as assayed by Southern blot analysis (data not shown). The *Eco*RI restriction map of two of the overlapping phage inserts obtained, L832 and L823 is also shown in figure 16. The presence of identically sized restriction fragments and their symmetrical arrangement on either side of the *Sry* box strongly suggested that the sequence surrounding the gene was an inverted repeat. Combining the phage mapping data it has been possible to extend the restriction map to cover approximately 33kb of the *Sry* locus (figure 16). This map suggests that the repeat

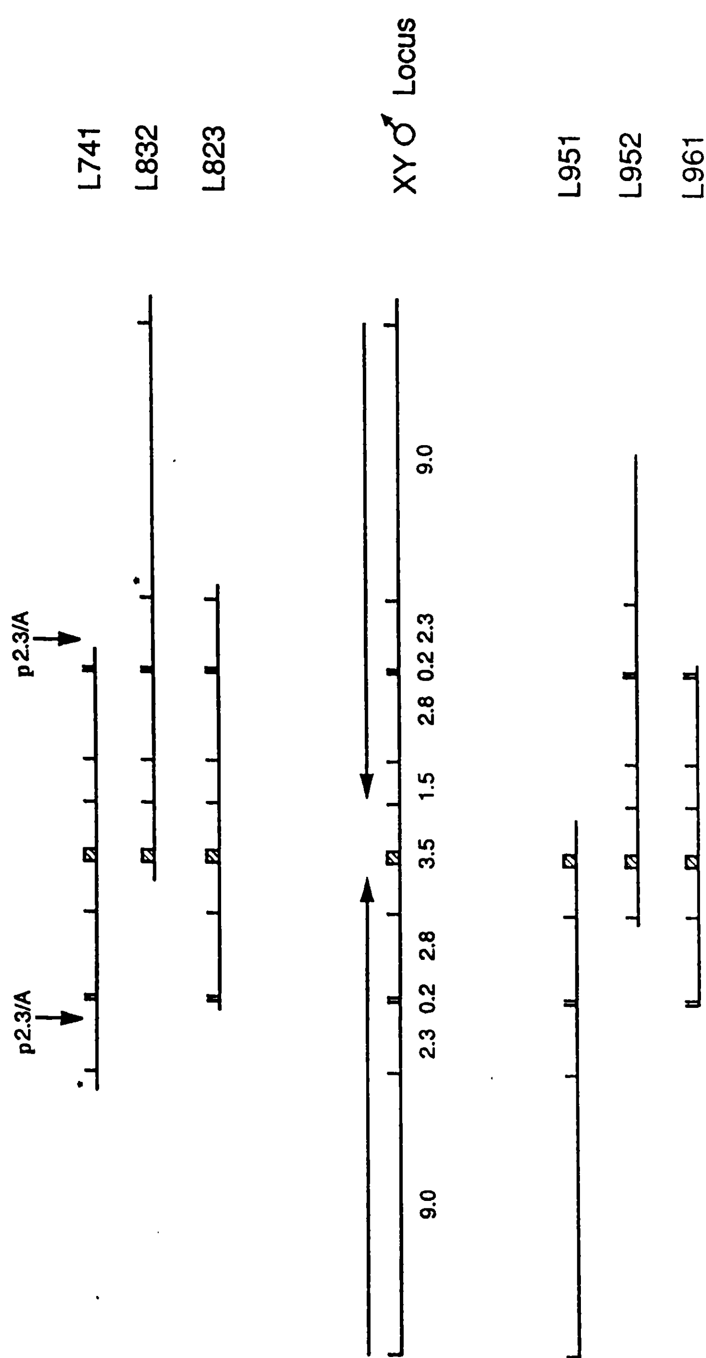


Figure 16. Guide to *Sry* genomic clones and structure of the wild type locus. *Eco*RI restriction maps of the following genomic phage inserts are shown: L741, L832, L823, *Sry* locus from a 129 mouse (*Mus musculus musculus* Y chromosome) and L951, L952, L961, *Sry* locus from a *poschiavinus* mouse (*Mus musculus domesticus* Y chromosome). Restriction sites are indicated by vertical dashes. The position of the *Sry* box is indicated by a hatched box. The positions of probe p2.3/A are shown by arrows. XY♂ Locus indicates the derived restriction map of the wild type locus. Restriction fragment sizes are indicated (kb). The extent and end points of the duplicated sequences are indicated by horizontal arrows. The presence of a second 9.0kb *Eco*RI fragment on the left hand side of the *Sry* box was inferred in the following way: a probe from the extreme left hand end of L741 is known to hybridize in two places (shown by asterisks), but only recognises a single 9.0kb *Eco*RI fragment on a genomic Southern blot (data not shown).

is a minimum of 15.5kb. The entire sequence of L741 has now been obtained (see appendix). A 2.8kb region of unique sequence, including the *Sry* box, can be defined within the inverted repeat as shown in figure 16 (Gubbay *et al.*, 1991).

Various lines of evidence suggest that this is the genuine structure of the region rather than an artifact of the cloning procedure used. (i) A number of overlapping genomic clones (some of which are shown in figure 16) have been isolated from three different unamplified genomic phage libraries made from two different mouse strains (see below) and the structure of all of these agree with the proposed restriction map. (ii) There is agreement between the phage maps shown and fragment sizes seen in genomic Southern blots (data not shown and see figure 17b). (iii) The *Eco*RI site immediately to the right of the *Sry* box, being in the unique region, is the only asymmetric *Eco*RI site present (see figure 16). Thus a probe covering one end of the inverted repeat such as the 1.5kb *Eco*RI fragment from L741 when used on a genomic Southern blot of DNA digested with *Eco*RI should recognise both a 1.5kb fragment (itself) and a 3.5kb fragment which contains the same sequence as part of the other copy of the repeat. This result is shown in figure 17b (probe p1.5). (iv) Although the duplicated region on either side of the *Sry* box is almost identical, there are at least 7 single base differences (see appendix). These differences have been verified by sequencing from both strands. If the duplication was introduced during cloning, such mismatches would be unlikely. In addition, a number of blocks of repeats of the dinucleotide CA, also known as microsatellites, are present within the repeated region. Some of these microsatellites have been found to be polymorphic in length on either side of the duplication (see appendix). (v) Finally, as shown in chapter 8, clone L741 which contains the inverted duplication is functionally active, as XX individuals transgenic for L741 can be sex reversed.

6.3 *Sry* from a *musculus* and a *domesticus* Y chromosome has a similar structure

The *Sry* genomic locus was initially cloned from mouse strain 129, which carries a Y chromosome of *Mus musculus musculus* origin. In order to approach the question of the evolutionary age of this inverted repeat, the *Sry* genomic locus from a Y

chromosome of *Mus musculus domesticus* origin has been cloned. As the *poschiavinus* substrain has been proposed to carry a late-acting allele of *Tdy* (Eicher & Washburn, 1986 and see section 1.7.2), this was chosen as an example of a *domesticus* type Y chromosome in order to also eventually define the molecular basis of this effect. A *poschiavinus* genomic library was therfore screened with the *Sry* box. Figure 16 shows the *Eco*RI restriction maps of three of the phage inserts obtained L951, L952 and L961. The inverted duplication is clearly present on this Y chromosome as shown by the symmetrical arrangement of fragment lengths around *Sry*. The *Eco*RI restriction pattern seen is identical to that seen in the *musculus* Y chromosome. These data suggest that the inverted repeat at the *Sry* locus arose before the divergence of the *musculus* and *domesticus* Y chromosomes. To confirm that the DNA used to construct this genomic library was from a mouse carrying a *domesticus* Y chromosome, a Southern blot of this DNA, digested with *Taq*I, was hybridized to the *Zfy-1* zinc finger probe (probe 3, legend to figure 4). This method can be used to distinguish *musculus* and *domesticus* Y chromosomes (see section 3.2.2). The appropriate band pattern for a *domesticus* Y chromosome was obtained (data not shown).

Given the similarity in the restriction maps of the *musculus* and *domesticus Sry* loci it is clear that a much higher resolution of analysis will be required to correlate *Sry* structure with functional differences between alleles. It is important to note that a late acting *Tdy* could be the result of a change in the regulation of this gene, or a change in protein structure, such that for instance, more protein needs to be accumulated before it reaches a threshold of activity. In fact, the sequence of the *Sry* box isolated from a number of *domesticus* Y chromosomes, including the *poschiavinus* Y chromosome has recently been obtained (E. Eicher, personal communication and T. Kunieda, EMBL database: X60687). In all cases, the only difference seen is the substitution of a threonine residue in the *domesticus* Y chromosome where an isoleucine is found in the *musculus* Y chromosome. Not all *domesticus* Y chromosomes are associated with sex-reversal on a C57BL/6 background, therefore this substitution may simply be a *musculus/domesticus* polymorphism. Alternately it may contribute to the sex-reversing effect, but only

when accompanied by other alterations in *Sry* only present in the *poschiavinus* Y chromosome.

6.4 The *Sry* inverted repeat: Origin and evolution

The mouse *Sry* box is contained within 2.8kb of unique sequence flanked by inverted repeats of at least 15.5kb. Such structures are a known feature of eukaryotic genomes, although the duplication generally involves sequences coding for genes rather than the gene itself being present in a unique inter-repeat region (Kant *et al.*, 1985; Russnak & Candido, 1985; Wang *et al.*, 1985). In fact, it appears from the literature that the arrangement found at the *Sry* locus has not been previously recorded. Inverted repeats usually involve large gene families. For example, the mouse major urinary proteins are organised in 45kb repeat units of 12-15 pairs of genes containing large internal palindromes (Bishop,O *et al.*, 1985). In the case of the α-amylase multi-gene family two members of the family are organised in the form of a large inverted repeat with a unique central core about 2kb in length (Groot *et al.*, 1990). Interestingly inverted repeats also occur in *Drosophila melanogaster* α-amylase genes, suggesting that this arrangement is important (Schwartz & Doane, 1989). The longest inverted repeats recorded are associated with gene amplification. Ford,M & Fried (1986) have reported inverted duplications in amplified DNA containing the *Myc* gene in four human tumour cell lines. These authors speculate upon the existence of inverted duplications found in amplified gene arrays and in a number of eukaryotic multi-gene families which presumably arose from a single ancestral copy. They suggest that gene amplification and stable gene duplication during evolution may share common features, such as repeated initiation from a particular location on the chromosome followed by multiple recombination events. Whether such mechanisms have been involved in generating the *Sry* locus is unclear. The centromeric regions of chromosomes from the fission yeast *S. pombe* are also organised as an inverted duplication flanking a small unique core (Polizzi & Clarke, 1991). Many direct repeats reside within this entire region. Mammalian centromeres have also been shown to be major sites for various satellite sequences, however whether these are also present as inverted repeats is not known (Rattner, 1991).

The duplicated arms at the *Sry* locus show almost perfect conservation at the sequence level. There are at least four possible explanations for this phenomenon. (i) The duplication may have been a recent event allowing little time for divergence. Although it has been shown that both a *musculus* and a *domesticus* Y chromosome carry this duplication, these chromosomes diverged relatively recently. Studies of more distantly related *Mus* species and of other species will be necessary to determine the evolutionary history of this event. (ii) Alternatively gene conversion events may be responsible for maintaining conservation between the two arms of the repeat. (iii) Sequence conservation can also result from a functional role, for instance *Sry* regulatory sequences, such as enhancers, which may need to be present in duplicate, may reside within the repeated area. However these are generally small motifs which would not explain the large area of conservation. Studies on the regulation of *Sry* should shed light on this matter. (iv) Sequences in the duplication (or the duplication itself) may be required to maintain a specific chromatin configuration necessary for the correct regulation of *Sry* expression. It is possible that the inverted repeats identified at yeast centromeres play such a structural role.

Human *SRY* is not surrounded by an inverted duplication (Ellis *et al.*, 1989; Sinclair *et al.*, 1990). Perhaps this is not surprising as there are many dramatic differences between the organisation of the human and mouse Y chromosomes (see figure 1). Human *SRY* is located only 5kb away from the pseudoautosomal boundary (Palmer,MS *et al.*, 1989), while mouse *Sry* is located at the opposite end of the chromosome from the pseudoautosomal region (McLaren *et al.*, 1988; Roberts *et al.*, 1988). It is known that the short arm of the mouse Y chromosome (where *Sry* resides) has undergone a number of duplications and rearrangements. For instance mouse *Zfy* genes are present in two copies as a large tandem repeat, while only one *ZFY* gene is found in most other species (Mardon *et al.*, 1989; Nagamine *et al.*, 1989; Simpson,EM & Page, 1991).

Chapter 7: Results and Discussion

Sry is deleted in XɎ female mice

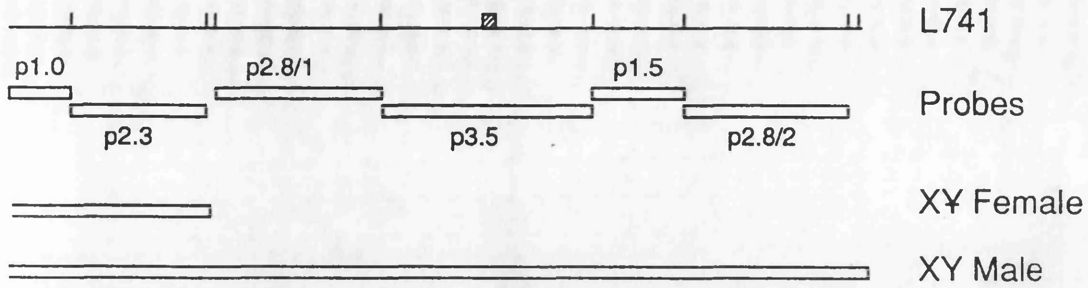
XɎ female mice can be used to determine whether a candidate sequence is involved in testis determination. The finding that *Sry* is deleted from the Ɏ chromosome which has lost testis-determining activity is consistent with *Sry* being *Tdy*. An analysis of the *Sry* deletion is described in this chapter.

7.1 Identifying a deletion breakpoint in the XɎ females

Since the DNA band hybridizing to pY53.3 (*SRY*), p422.04 (*Sry* box) and p422 (a 3.5kb *Sry* genomic fragment) is absent in XɎ females (see figure 13b, 13c and data not shown for p422), it was necessary to use the larger genomic fragment from this region (L741) in order to search for the boundaries of this deletion. Radiolabelled probes were made from each of the *Eco*RI fragments contained within L741 (see figure 17a), and used to screen genomic Southern blots of DNA from XY male, XɎ female and XX female mice. Figure 17b shows that three of the six probes, p2.8/1, p2.8/2 (which are essentially identical, as discussed in section 6.2) and p1.5, failed to detect a Y specific band in XɎ female DNA. Note that probe p1.5 detects a 3.5kb band in the normal XY male in addition to the expected 1.5kb cognate band. This is due to the presence of one end of the inverted repeat sequence at the *Sry* locus being in each of these fragments (see section 6.2). Probe p3.5 (p422) also failed to detect either the 3.5kb or the 1.5kb band in XɎ female DNA (data not shown). Unlike probes p2.8/1, p2.8/2, p1.5 and p3.5, probe p2.3 detects an *Eco*RI band of the same size in both XY male and XɎ female DNA. The adjacent probe, p1.0, which corresponds to one end of the phage insert, similarly detects an *Eco*RI band of the same size in both XY male and XɎ female DNA (data not shown). These results suggest that a breakpoint of this deletion is within the 200bp *Eco*RI fragment between probes p2.3 and p2.8/1. This was confirmed by hybridizing *Sac*I digested DNA with probe p2.3. This reveals a band of altered size in XɎ females compared with XY males (also shown in figure 17b). However, given the palindromic nature of the *Sry* locus (see figure 16), this breakpoint could be on either side of the *Sry* box.

Figure 17. Analysis of a deletion breakpoint in the X¥ female. a, An *Eco*RI restriction map of phage L741 is shown. The open boxes indicate the location of *Eco*RI fragments (p1.0, p2.3, p2.8/1, p3.5, p1.5, p2.8/2) used as probes. Note that p3.5 is equivalent to p422 and that p2.8/1 and 2 are identical, being within the repeated region. The *Sry* box is indicated by a hatched box. The limit of the region detectable in the X¥ female and the colinear genomic region in XY males is indicated by open boxes, labelled X¥ Female and XY Male respectively. b, Southern blots of *Eco*RI- or *Sac*I-digested genomic DNA from XY male, X¥ female, and XX female mice. Band sizes are indicated (kb). Probe p1.5 and a combination of probes p2.8/1 and 2 detect Y-specific bands in an XY male which are deleted in an X¥ female. Probe p2.3 detects a Y-specific *Eco*RI band of the same size in an XY male and an X¥ female. However, a *Sac*I digest reveals a difference in the size of the band detected by probe p2.3, indicating a breakpoint within this genomic region in the X¥ female. The size of the *Sac*I restriction fragment in XY male DNA was as expected from the derived restriction map of the wild type *Sry* locus (data not shown). Probe p1.0, although found to be highly repetitive when used to probe a Southern blot of *Sac*I-digested DNA, nevertheless gave the same result as probe p2.3 (data not shown). Probes p2.3, p2.8/1 and p2.8/2 were precompeted with mouse genomic DNA before hybridization.

a



b

-XY ♂ -XY̶ ♀ -XY ♂ -XY̶ ♀ -XX ♀ -XY ♂ -XY̶ ♀ -XX ♀ -XY ♂ -XY̶ ♀ -XX ♀

15
9.5
2.3
2.8
3.5
1.5

Enzyme:	*Sac*I	*Eco*RI	*Eco*RI	*Eco*RI
Probe:	p 2.3	p 2.3	p 2.8/1,2	p 1.5

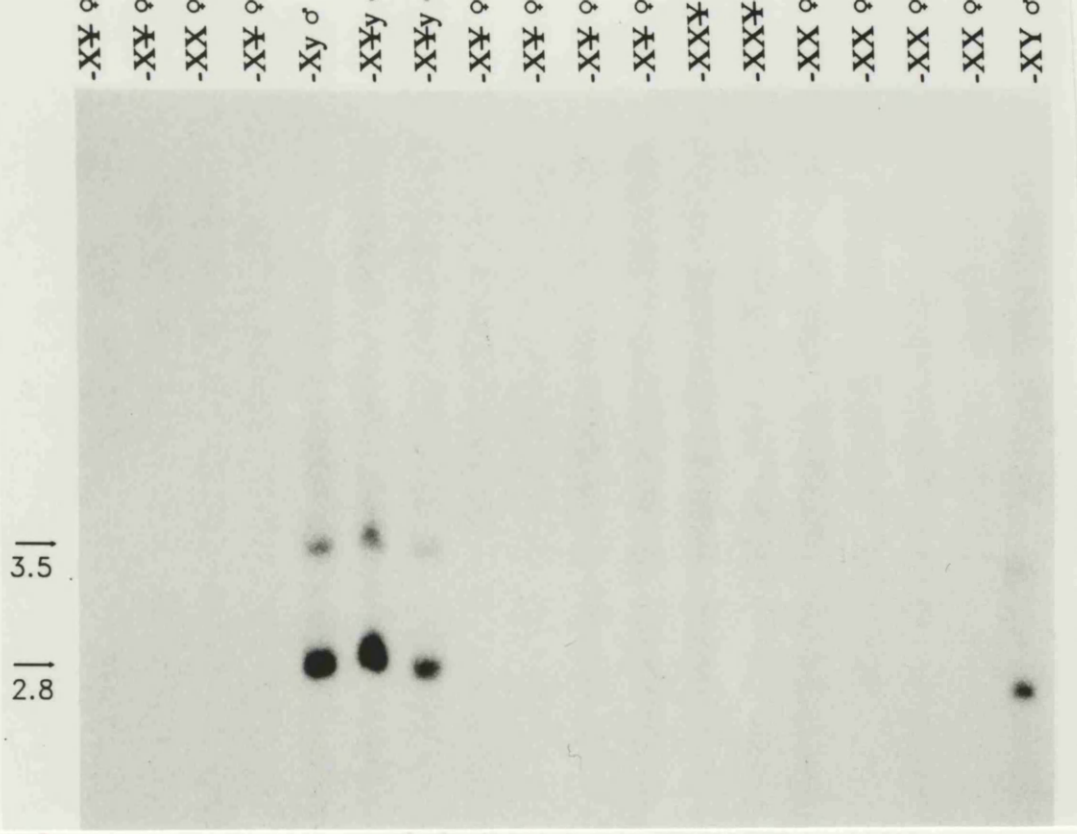
7.2 Analysis of the deletion in a number of X¥ female mice.

It was necessary to confirm that the deletion correlated with the mutant ¥ chromosome and was not just a secondary mutation, unconnected with sex determination, in the particular X¥ female used for the analyses shown in figures 13 and 17. A Southern blot of DNA from several X¥ females and their progeny was therefore hybridized to probes p2.8/1, p2.8/2 and p422.04. Figure 18 (upper panel) shows that the bands corresponding to these probes are absent from all the X¥ female samples. Similarly, figure 18 (lower panel) shows that probe p2.3 is present in all the mutant ¥ chromosomes analysed. These results confirm that the presence of this deletion correlates with the loss of *Tdy* activity.

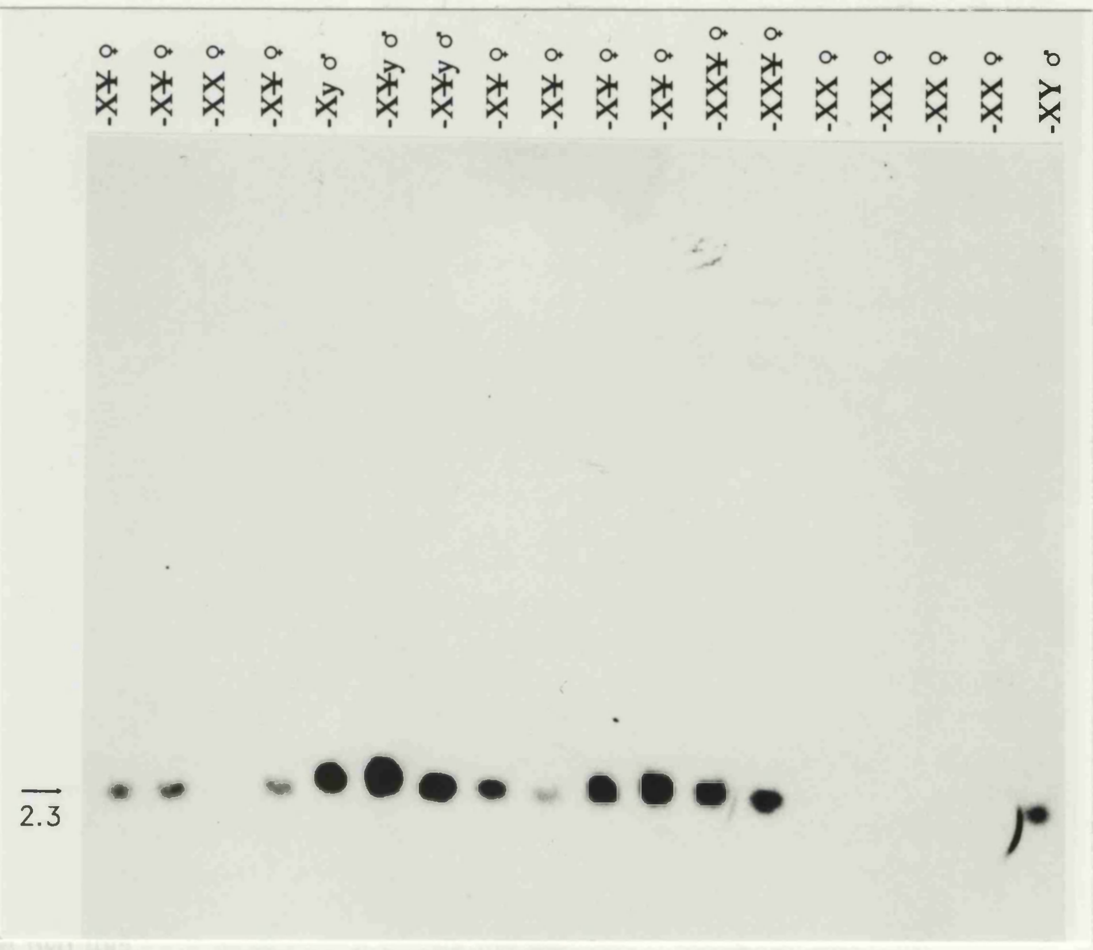
7.3 Full characterization of the *Sry* deletion in X¥ female mice

In order to fully define the deletion, a non-repetitive probe, p2.3/A (see figure 16 and section 6.2), that was not deleted from the ¥ chromosome (data not shown) was used to screen a phage genomic library of DNA from an X¥ female mouse. This library was made by A. M. Frischauf, using the phage vector λ EMBL3a. One positive clone was obtained out of $1x10^6$ plaques screened. An *Eco*RI restriction map of the phage insert obtained, L1041, is shown in figure 19. This indicates that the entire deletion covers just under 11kb of sequence including the *Sry* box. Upon digestion with *Eco*RI the only novel fragment produced from L1041 not seen in any phage inserts from wild type DNA was a 400bp fragment. The position of this is shown in figure 19. This fragment was sequenced as it was presumed to contain the deletion breakpoints themselves. Figure 20 shows the sequence of this fragment and one possible alignment of this sequence to the two arms of the inverted duplication at the *Sry* locus. The alignment shown in figure 20 represents the mutation as two deletions, one being only 239bp, while the other spans from one arm of the inverted duplication to the other. This is shown schematically as scheme 1 in figure 21. However, as the sequences on either side of the *Sry* box are palindromic, an alternative interpretation is that a single deletion has occurred accompanied by a small inversion of 264bp. This is shown in scheme 2 figure 21. To confirm that L1041 accurately represents the genome in X¥ female mice, an

Figure 18. Analysis of the deletion in a panel of X¥ females. Southern blots described previously (see legend to figure 12) were hybridized to probes p2.8/1, p2.8/2 and p422.04 together (upper panel) and to p2.3 (lower panel). These probes are described above (see legend to figure 17). Numbers indicate the size of the hybridizing bands (kb). The bands corresponding to probes deleted in X¥ female mice (upper panel) are absent from normal females and all samples whose only Y chromosome is the mutant ¥. While the band detected by probe p2.3 is absent only in DNA from normal females.



Probes p2.8/1,2 and p422.04



Probe p2.3

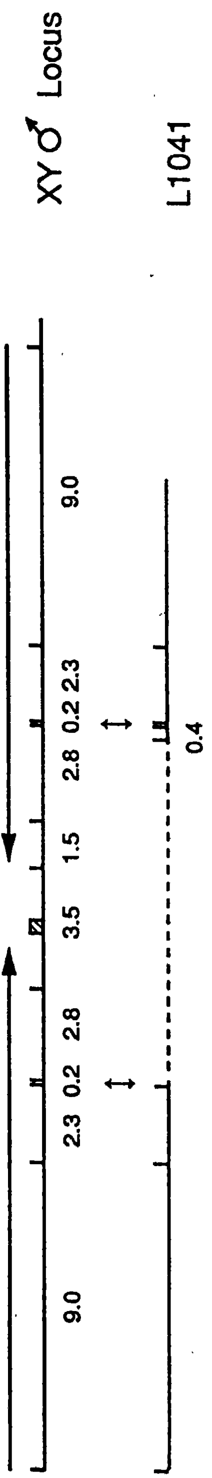
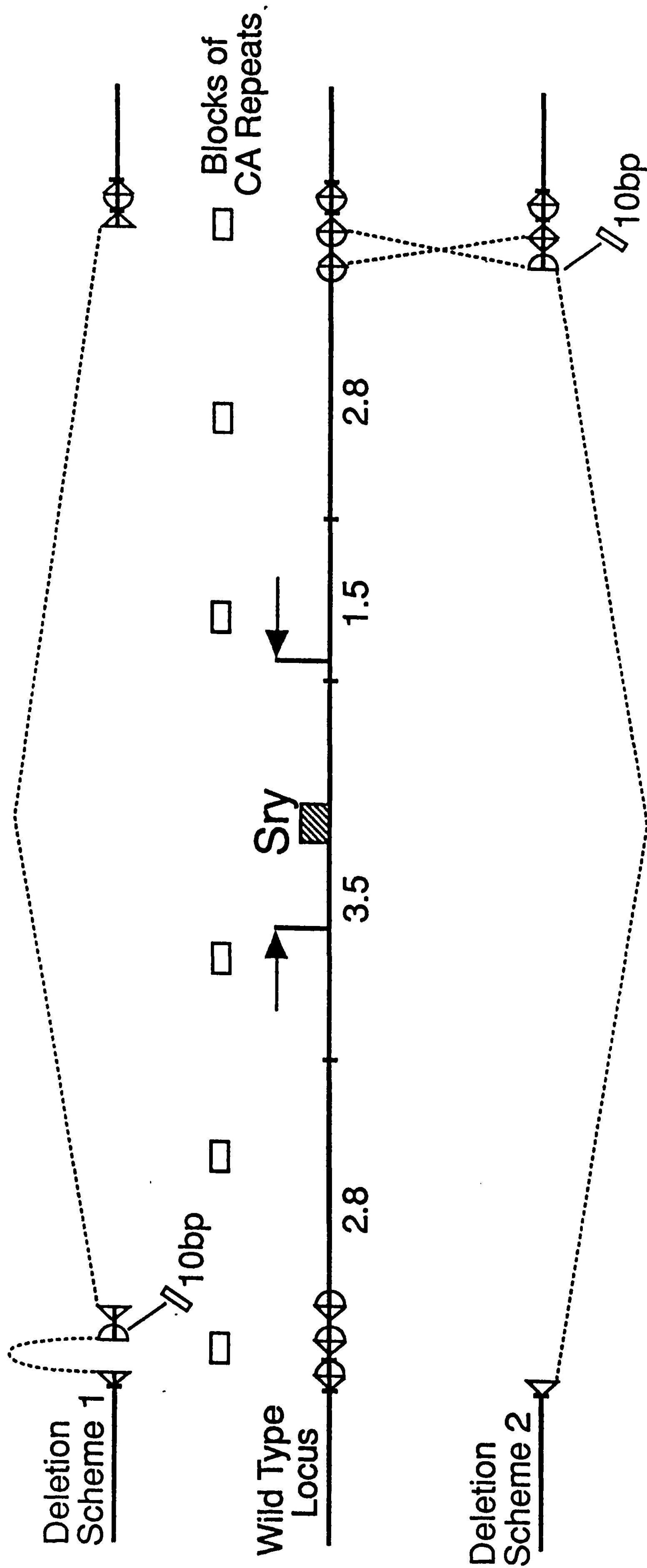


Figure 19. Mapping the genomic region corresponding to the deletion in XY̶ females. The insert of phage L1041 which corresponds to the *Sry* locus from an XY̶ female mouse carrying the Tdy^{m1} mutation is compared to the wild type *Sry* locus (XY♂ locus). The wild type locus is as described in the legend to figure 16. Restriction fragment sizes are indicated (kb). In L1041 the dotted line indicates deleted sequences. The orientation of this clone with respect to the *Sry* box is not known. The approximate positions of a single oligonucleotide primer used for PCR amplification of the deletion breakpoints are indicated by vertical arrows between the two loci.

```
       1                                                                               80
Left.  GAATTCCACACTGAGGCCTAATGCCTGCTGTTTCTTCTGTGTCTCAAAGCCTGCTCTTCCTACCTGAAATAGGATAGTCC
Del.   GAATTCCACACTGAGGCCTAATGCCTGCTGTTTCTTCTGTGTCTCAAAGCCTGCTCTTCC....................

       81                                                                             160
Left.  TTTTTTTCCCCCTTGGTACTTCTTTACCTTATTACCACACCAAGACTTAATACTCTTGCTTTTATTTCTCCTTTCAGTTA
Del.   ................................................................................

       161                                                                            240
Left.  CCTGTTATGCCTTTCAGATTGTCACCCTCTATCTGAATTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Del.   ................................................................................

       241                                                                            320
Left.  GTGTGTGTGTGTGTGTGTGTGTATGTGTGTGTGTGTGTGTTTGTGTGTGTCTTTTTGTGTGTGATGTGTGATTGGCGATT
Del.   ...........................................................TGTGATGTGTGATTGGCGATT

       321                                                                            400
Left.  GTTGTCTTTGCTTTGTACACACACTTTACTTTGCACACACACACACACACACACACACACACA........GAGTG
Del.   GT...CTTTGCTTTGTACACACACTTTACTTTGCACACACACACACACACACACACACACACACACACACACACAGAGTG

       401                                                                            480
Left.  TTTATTGATATGTAAATCACTGGATTCAGTTTACAGAGATTCAGATGTCGTTTGGAATATGATAGCTGCTAGCAGTACAT
Del.   TTTATTGATATGTAAATCACTGGATTCAGTTTACAGAGATTCAGATGTCGTTTGGAATATGATAGCTGCTAGCAGTACAT

       481                                                                            560
Left.  GTTGCTTGTTCAGCTTTGCTTCCTGATTCAGATGTGTGCCCACTGTTGCTTTAGGTTGCTTCAGGCCTGCCTAACCACAC
Del.   GTTGCTTGTTCAGCTTTGCTTCCTGATTCAGATGTGTGCCCACTGTTGCTTTAGGTTGCTTCAGGCCTGCCTAACCACAC

       561                                                                            640
Left.  CACCTC| -->
Del.   CACCTC|aacacacacacacacacatacacacacacacacacacacacacacacacacacacacacacacacacacacac
Right.   <-- |aacacacacacacacacatacacacacacacacacacacacacacacacacacacacacacacacacacacacac

       641           654
Del.   acacacatgaattc
Right. acacacatgaattc
```

Figure 20. Sequences at the deletion breakpoints in XY̶ females. The sequence of the 0.4kb *Eco*RI fragment from L1041 (Del.), which contains the XY̶ female deletion breakpoints, compared to sequences from either arm of the duplication. Sequences from the left arm of the duplication (Left.) are shown in upper case while those from the right arm (Right.) are shown in lower case. The deletion sequence is shown in upper case where it corresponds to the left arm and lower case when it corresponds to the right arm. The arrangement shown corresponds to scheme 1 (Figure 21) i.e. one short deletion (shown by dots) and a major deletion indicated by a vertical line across the sequence. The horizontal arrows at this point represent a continuation of sequence from the left and right hand sides of the locus in wild type DNA which is deleted in the XY̶ female. The deletion sequence corresponds to sequences at the wild type locus except at two points: (1) a TGT sequence is missing from the deletion sequence at position 323, (2) The length of one of the CA repeat motifs differs in the deletion sequence from that of the wild type sequence. This particular motif is polymorphic in length on either side of *Sry* in wild type mice, and thus may represent a polymorphism that already existed in the Y chromosome that gave rise to the XY̶ females. Alternatively additional CA residues may have been added during the deletion process.

Figure 21. Two possible schemes for the deletion of *Sry* in X~~Y~~ female mice. The wild type locus is shown with restriction sites for *Eco*RI indicated by vertical dashes and fragment sizes in kb. The end points of the duplicated regions are marked by horizontal arrows above the locus. Rectangular blocks represent the approximate position of blocks of CA/GT repeats. Three points along the sequence, marked on both arms of the repeat, are illustrated by triangle/semicircle symbols. DNA breakage points have occurred at the centre of these symbols. As these symbols have polarity, the presence of either a triangle or a semi-circle in the deleted locus (schemes 1 and 2) indicates which side of a particular breakpoint sequences originate from. In both schemes there are three breakpoints. Two of these breakpoints are at almost identical positions on either side of *Sry* but differ by 10bp. A 10bp block of sequence is therefore shown being lost to compensate for this slight difference. Dotted lines indicate the position of deletion and inversion events.


Deletion Scheme 1
10bp
Blocks of CA Repeats
Wild Type Locus
2.8
3.5
Sry
1.5
2.8
Deletion Scheme 2
10bp

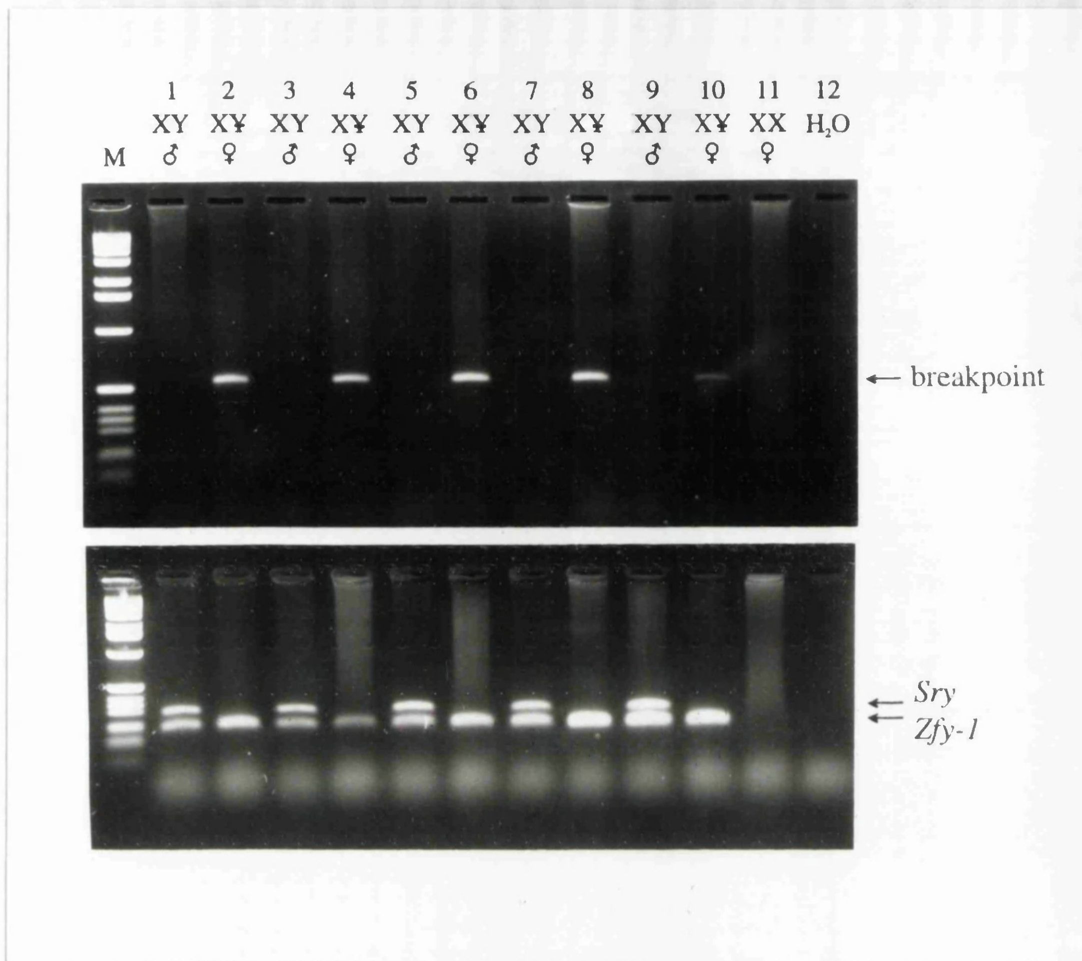


Figure 22. PCR analysis of the deletion in X¥ female mice. Analysis was performed on 11 different genomic DNA samples (from 5 XY♂ mice, 5 X¥♀ mice and one XX♀ mouse). Upper panel, A 522bp fragment (breakpoint) was only amplified in X¥ female samples using a single PCR primer adjacent to the deletion breakpoints. The strict correlation between amplification of the breakpoint fragment and the ¥ chromosome indicates that plasmid DNA contamination is not present. Lower panel, the same samples were tested for PCR amplification of the *Sry* box (*Sry*), for which a band is only present in normal males, and *Zfy-1* for which a band is present in all samples except XX♀. Marker bands (M) are: 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118, 72bp.

assay to amplify a fragment at the breakpoint in DNA from XɎ females using PCR was designed. Because the deletion juxtaposes sequences which are palindromic, a single PCR primer should be sufficient to amplify a 522bp fragment around the breakpoint (figure 19). It was possible to amplify such a fragment from XɎ female DNA but not wild type DNA (figure 22). The fragment obtained was of the appropriate size and its sequence identical to that of the breakpoint fragment from the XɎ female phage clone L1041.

The finding that the *Sry* box is deleted in XɎ female mice represents the first molecular difference to be detected between Ɏ and a normal Y chromosome. Cloning the deletion breakpoints has shown that only 11kb at the *Sry* locus has been deleted. Sequence comparison data (see section 8.3) reveals no other conserved genes residing within the portion deleted, so unless the mutation results in cis-effects being exerted on a gene that is not deleted, *Sry* must be the only gene affected in these mice. This finding strongly argues that *Sry* is normally necessary for testis determination, although it cannot provide proof, because of the possibility of other mutations existing on the Ɏ chromosome.

7.4 Possible mechanisms for the deletion of *Sry* in XɎ females

The *Tdy*^{m1} mutation arose during insertional mutagenesis of ES cells using retroviral vectors. Lovell-Badge & Robertson (1990), proposed that the mutation in *Tdy* must have occurred after the infection protocol and may therefore have been an entirely unconnected event. Retroviral excision events can give rise to deletions. However a single LTR is usually left behind as excision occurs by recombination between LTRs (Mager & Goodchild, 1989). As all of the sequences at the breakpoint of this deletion are from the endogenous locus it seems unlikely that insertion and subsequent excision of a retroviral or retroviral like element was its cause.

Clearly the mutation itself was a complex event, as shown in figure 21, involving either 2 individual deletions or one deletion accompanied by a short inversion. Either scheme would involve three points of DNA breakage, two of these being at almost identical points (±10bp) in each arm of the inverted duplication. Rather

than this being a completely random event, the unusual properties of the locus may have contributed to the mutation in one or more of the following ways. (i) The role of direct repeats in the generation of deletions is well documented. Such deletions occur by unequal sister chromatid exchange, unequal crossing over or replication slippage (Crow & Dove, 1988). In humans these mechanisms have been proposed to explain deletions giving rise to diseases such as the thalassemias (Henthorn *et al.*, 1990) and the generation of new length alleles at tandem repetitive hypervariable loci (Jeffreys *et al.*, 1988). The *Sry* deletion has occurred near blocks of CA repeats present either side of the *Sry* box. Such repeats are thought to be randomly distributed about the genome but are not uncommon near genes (Jongeneel *et al.*, 1990). The proximity of the deletion breakpoints to microsatellites, suggests that they may have acted as tandem repetitive elements to promote the deletion by one of the mechanisms mentioned above (see figure 23). (ii) An alternative possibility is that the palindromic nature of the locus may have played a part in juxtaposing the deletion breakpoints by forming a stem loop structure. Lehrman *et al.* (1986) have proposed that short palindromic sequences may have played such a role in a 5kb deletion in the human low density lipoprotein receptor. Their model invokes the formation of a stem loop structure possibly involving single stranded DNA during replication. Nicking and subsequent religation within the stem leads to the looped out unique sequences being deleted (see figure 23).

One other expected consequence of the inverted duplication at this locus may be a capacity to invert the unique sequences at the centre of the duplication, including *Sry*. This could occur by homologous pairing between sequences in the inverted repeat followed by intramolecular recombination (see figure 23). Flip-flop inversion of specific DNA sequences flanked by inverted repeats is known to occur in bacteria (Simon & Herskowitz, 1985; Johnson,RC *et al.*, 1986). It is also thought to be the mechanism for an inversion event at the *Drosophila melanogaster* α-amylase locus, resulting in a null allele of this gene (Schwartz & Doane, 1989). Due to the extreme sequence conservation of the *Sry* inverted repeat, inversion events not accompanied by additional mutations, may not affect *Sry* function at all. Once a DNA sequence outside the limits of the duplication has been isolated it should be

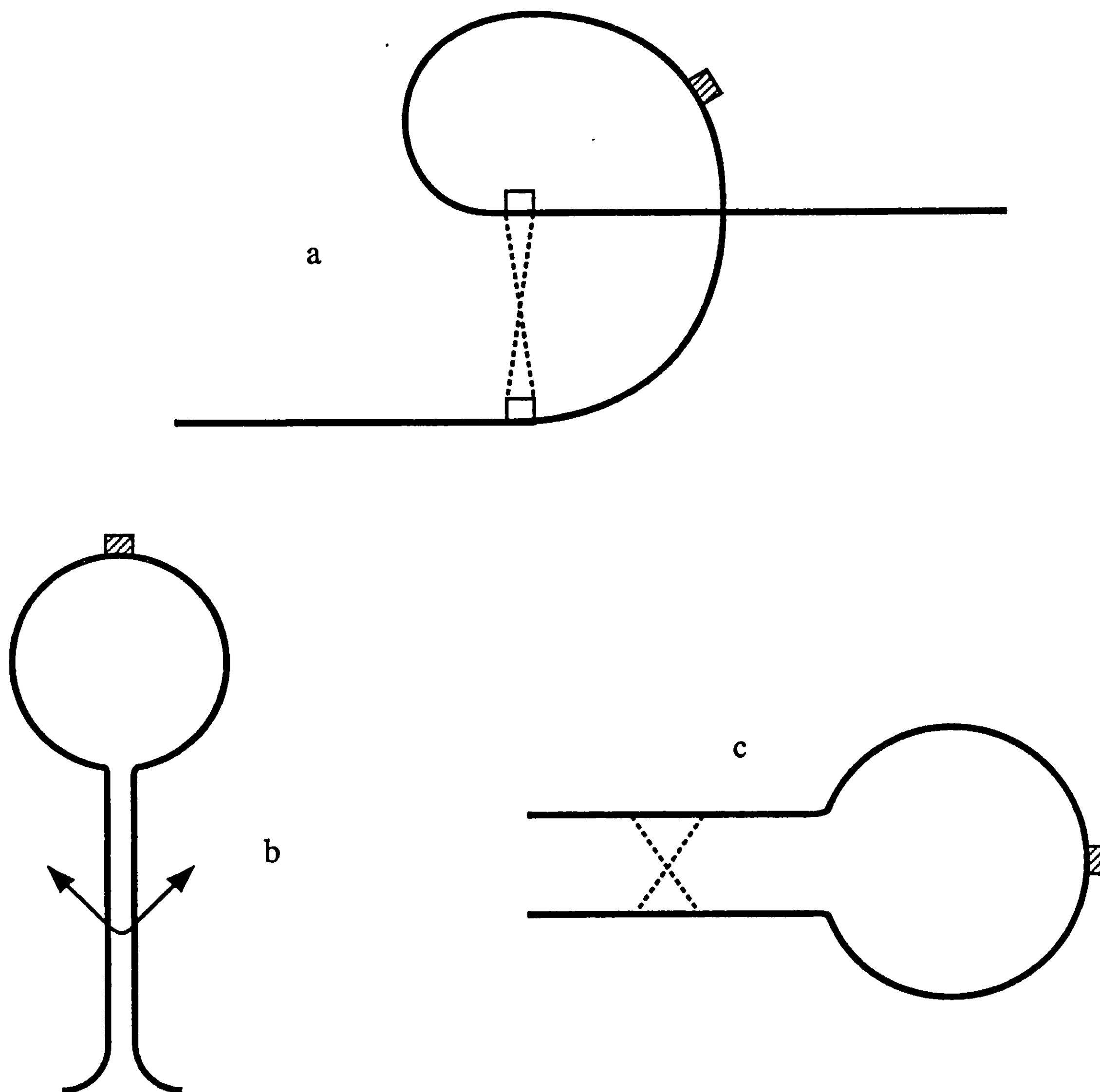


Figure 23. Schematic representation of recombination events that may occur at the *Sry* locus. The hatched box indicates the position of the *Sry* box. Blocks of CA/GT repeats are indicated by open boxes. Dotted lines indicate the position of double stranded recombination events. a, Tandemly repeated elements can promote deletions by intramolecular recombination. b, DNA sequences (possibly single stranded) can be juxtaposed due to their palindromic nature, subsequent strand breakage and religation can result in the generation of a deletion (indicated by a double headed arrow). c, Intramolecular recombination promoted by palindromic sequences can lead to "flip-flop" inversion of intervening sequences.

possible to test for changes in the orientation of *Sry* with respect to this new marker.

Chapter 8: Results and Discussion

Sry is equivalent to *Tdy*

The data presented here regarding the chromosomal location of *Sry* (chapter 5), its possible mode of action (chapter 5) and its deletion in X~~Y~~ female mice (chapter 7) indicated that *Sry* was a strong candidate to be *Tdy*. This data has been complemented by two further studies. Koopman *et al.* (1990) have used RT-PCR to study the expression of *Sry* in mouse embryos. They found a strong correlation between *Sry* expression and testis differentiation. *Sry* was found to be expressed only in the genital ridge in a narrow window of time beginning at 10.5*dpc* and ending at approximately 12.5*dpc*. This corresponds exactly with the first morphological signs of testis differentiation. These authors also examined which cells in the genital ridge were responsible for this expression. As testis development occurs in W^e/W^e XY embryos, which completely lack germ cells, *Tdy* must be expressed in the remaining somatic portion of the genital ridge. By analysing *Sry* expression in genital ridges isolated from such embryos, Koopman *et al.* (1990) showed that *Sry* is expressed from one of the somatic cell lineages in the developing gonad. The only other site of *Sry* expression reported was adult testis where *Sry* transcripts were found to be germ cell dependent. If this expression is a property of the germ cells themselves, rather than a result of somatic-germ cell interactions, then it probably has no critical cell-autonomous function in male germ cell differentiation, as functional sperm lacking *Sry* can be made. This occurred during the generation of X~~Y~~ female mice, deleted for *Sry* (Lovell-Badge & Robertson, 1990 and see section 1.7.1), via a male chimaera. A number of genes suspected of having a role in developmental events in the embryo show reactivation in adult testis, but in no case has their function been defined (Willison & Ashworth, 1987).

Berta *et al.* (1990) and Jäger *et al.* (1990a) have analysed a panel of XY female human individuals that do not possess Y chromosome deletions or rearrangements. They reasoned that sex reversal may have occurred in at least some of these individuals as a result of subtle mutations in *TDF*. By sequencing the *SRY* box region from these individuals, three were found to possess point mutations in *SRY*. In one case this would result in the production of a truncated protein whereas the

other two cases would cause single amino acid substitutions at positions which are highly conserved in all *SRY* and *SRY*-related genes. Two of these mutations were shown to be *de novo*, in that the father possessed a wild type *SRY* sequence. This provides strong genetic evidence that *SRY* is normally necessary for testis determination.

One of the XY females described had a variant in her *SRY* sequence also present in her father. Clearly this *SRY* variant is not always associated with sex reversal. Berta *et al.* (1990) have suggested three possible reasons for this effect. (i) The variant may be constitutively sex reversing but the father is mosaic for wild type and variant sequences. However this possibility was subsequently excluded as not only the father, of this XY female, but the father's XY sister all possessed the same *SRY* variant, (McElreavey *et al.*, 1991). (ii) The variant may have been fortuitously found in a family segregating for a downstream, autosomal or X-linked, sex determining gene. (iii) The variant could cause conditional sex-reversal depending on other genetic or environmental factors. The latter possibility would be particularly interesting as it may mirror cases of XY sex reversal in the mouse, involving the *poschiavinus* Y chromosome (see section 1.7.2).

Further studies (R. Hawkins, personal communication) suggest that approximately 15% of *SRY* positive XY females possess point mutations in the *SRY* box. It is possible that in the remaining cases *SRY* mutations may have fallen outside the region tested, for instance within an element responsible for *SRY* regulation. Alternatively these individuals may possess a mutation in another part of the sex determining pathway. In some cases such mutations may be in genes which are themselves regulated by *SRY*. Such patients are therefore a valuable resource, as they can be used to test for mutations in genes thought to be downstream targets of *SRY*.

These data, show that *Sry* is normally necessary for testis determination. The experiments described below show that *Sry* is the only Y-linked gene required for testis determination.

8.1 Analysis of transgenic embryos

The best way to test the function of *Sry* is to introduce the gene into XX embryos, and to determine if they develop as males. The pattern of *Sry* expression during fetal gonad development in the mouse suggests that precise regulation of the gene may be critical for its action (Koopman *et al.*, 1990 and see above). Murine *Sry* was therefore introduced as part of a genomic fragment, in the expectation that this would provide the correct regulatory sequences.

Fertilized eggs were microinjected with purified insert from the *Sry* containing phage, L741 (see section 2.10). Eggs were transferred to pseudopregnant recipients, and some of the resulting embryos were analysed 14 days later, rather than allowing them to develop to term. At this stage phenotypic sex can be assayed rapidly by examination of the fetal gonads. The first visible signs of testis development from the genital ridge occur at about 12.5*dpc*. By 14*dpc* a number of morphological differences distinguish fetal testes from ovaries. Thus Sertoli cell cord formation leads to a characteristic stripy appearance, rapid growth has occurred and a prominent vasculature is present. At this stage even partial sex reversal should be clear (Eicher *et al.*, 1980; Eicher, 1988). An indication of chromosomal sex (XX or XY/XO) was obtained by staining for sex chromatin in amnion cells.

After injection with the L741 insert (subsequently referred to as f741), 158 embryos were obtained. Most of these were XY males or XX females in roughly equal proportions (see Table 1). However, in two cases, testes were observed in embryos whose sex chromatin indicated an XX rather than an XY sex chromosome constitution. Southern blot analysis of genomic DNA prepared from these embryos, using probes to both the *Zfy* and *Sry* genes, showed that both of these males lacked *Zfy* sequences, confirming the absence of a Y chromosome, and were transgenic, with many copies of *Sry* (see figure 24a). Histological examination showed that their testis-cord formation was normal and that their gonads were indistinguishable from testes of normal XY sibling embryos (figure 24b and c).

TABLE 1. Summary of embryonic transgenic data

No. of embryos	Sex Chromatin	*Sry*	*Zfy*	Deduced Karyotype	Transgenic	Phenotypic sex
63	+	-	27-/40 ND	XX	-	♀
27	-	+	+	XY	ND	♂
58	-	ND	ND	XY	ND	♂
2	-	-	-	XO	-	♀
6	+	+	-	XX	+*	♀
2	+	+	-	XX	+	♂

Table 1. Summary of embryonic transgenic data. Embryos injected with f741 were examined 14 days after transfer. Embryos were analysed in the following way: chromosomal sex (XX or XY/XO) was determined by staining for sex chromatin in amnion cells. Transgenesis was assayed either by Southern blot or PCR detection of *Sry*, and the presence or absence of a Y chromosome was judged from similar assays for *Zfy* gene sequences. Phenotypic sex was determined by scoring for testis or ovary development. The frequency of XO progeny was consistent with previous studies (Russell, 1976). Asterisk, indicates that in four cases of XX transgenesis, comparison of the *Sry* signal to that of a control male indicated mosaicism for the transgene. Shading, highlights the transgenic embryos obtained. ND, not determined.

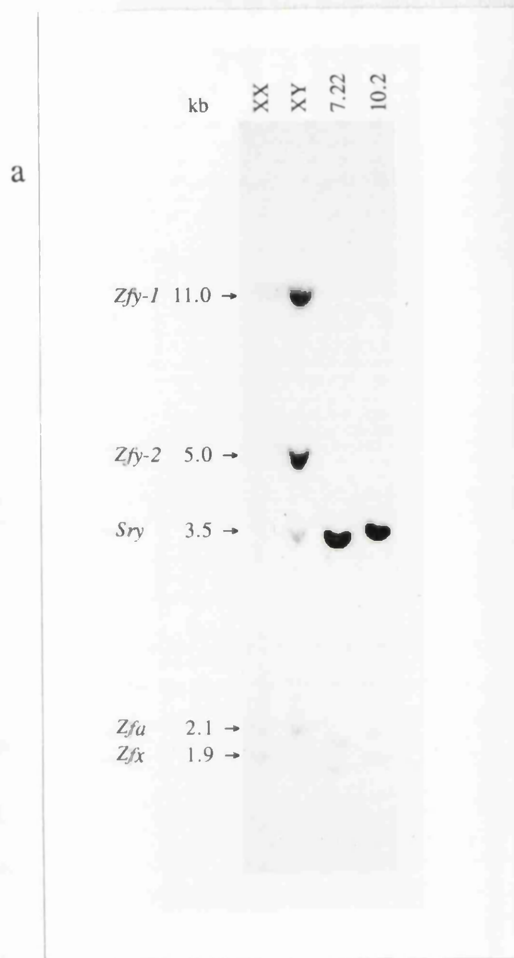
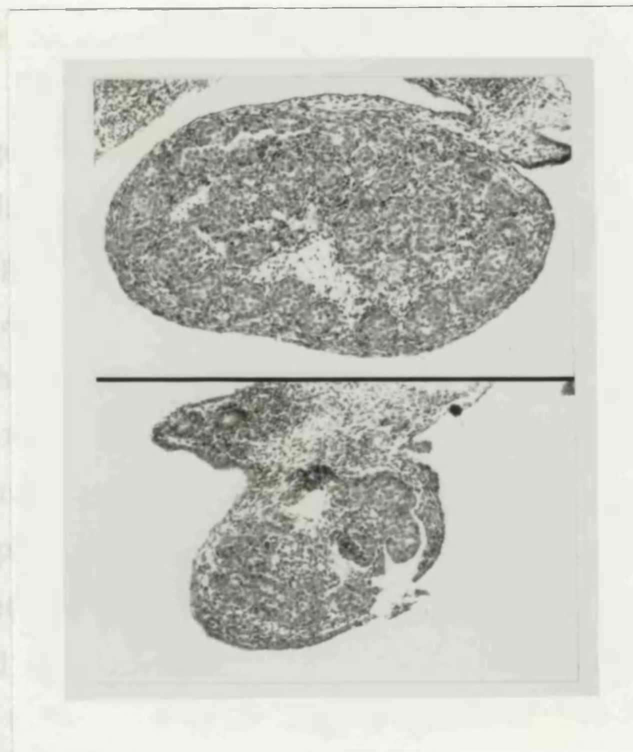


Figure 24. Analysis of sex-reversed transgenic embryos. a, Southern blot analysis of *Eco*RI-digested DNA from two phenotypic male embryos m7.22 and m10.2. Lack of hybridization to a probe recognising *Zfy-1* and *Zfy-2* (*Zfy* zinc finger probe, probe 3, legend to figure 4), shows the absence of a Y chromosome, whereas the intensity of hybridization to the *Sry* probe (p422.04) demonstrates that they carry multiple copies of the transgene. XY male and XX female samples are included for comparison. The *Zfy* zinc finger probe hybridizes to *Zfx* and *Zfa* as well, providing a control for the amount of DNA in each lane. b, Gonad morphology. Pairs of gonads, dissected from embryo m7.22 (upper panel, centre) and m10.2 (lower panel, centre), are shown between single testes (left) and ovaries (right) of nontransgenic sibs. The gonads of the transgenic embryos show the characteristic stripes associated with testis-cord formation. c, Histology of m7.22 (upper panel) and m10.2 (lower panel) testis sections. The apparent difference in size is due to plane of section. Cord morphology was similar to that of littermates (not shown).

b
♂
Embryo 7.22
♀
♂
Embryo 10.2
♀


c

These experiments indicate that a genomic fragment carrying *Sry* is sufficient to initiate testis development in mice. To determine the frequency with which f741 gives sex reversal, all the embryos scored as females were examined for the presence of *Sry* sequences by PCR (data not shown). Two were unequivocally identified as transgenic for *Sry*. Four more were probably mosaics possessing the transgene in a low proportion of cells, as only weak signals were detected for *Sry* (Table 1). This indicates that sex reversal does not occur in all embryos transgenic for f741.

8.2 Analysis of adult transgenic mice

To test the adult phenotype of *Sry* transgenic mice some of the embryos injected with f741 were allowed to develop to term. A total of 93 animals were born (49 males and 44 females). Genomic DNA was prepared from tail tip biopsies, and analysed by Southern blotting, using both *Zfy* and *Sry* sequences as probes (data not shown). This indicated that five of these mice were definitely transgenic (because of the presence of the endogenous *Sry* gene, XY mice possessing less than a single copy of the transgene per cell may not be detected). Two of the five transgenics were XY males that did not transmit the transgene and so were uninformative with respect to sex reversal.

One of the transgenics, m33.13, had no Y chromosome as determined by PCR analysis using primers directed against *Zfy* sequences (figure 25a), but was externally male (figure 25b). He was similar in size and weight to his normal XY male littermates. At about six weeks *post partum*, m33.13 was caged with females. His copulatory behaviour was normal, mating four times in six days. As expected for a sex-reversed male, m33.13 was found to be sterile. None of the four females with which he mated became pregnant. In three cases the vaginal plugs were examined for the presence of sperm, but none were found (data not shown). The only difference between m33.13 and a normal XY sibling was in the size of the testes: m33.13 had a testis weight of 17mg (in the range expected for an XXSxr' male), as opposed to 76mg for an XY control littermate. Sex-reversed males, such as XXSxr or XXSxr' male mice and human XX males, are always sterile for two

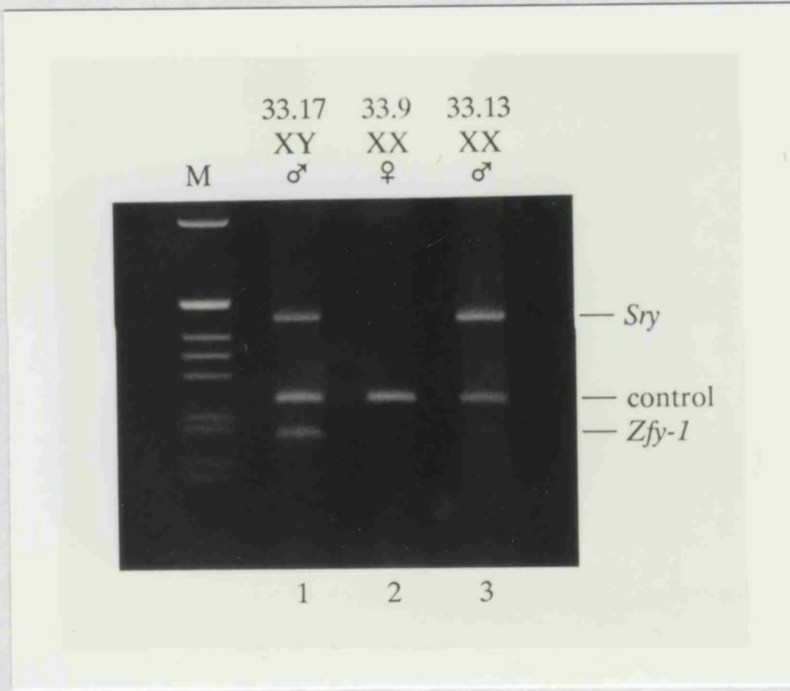
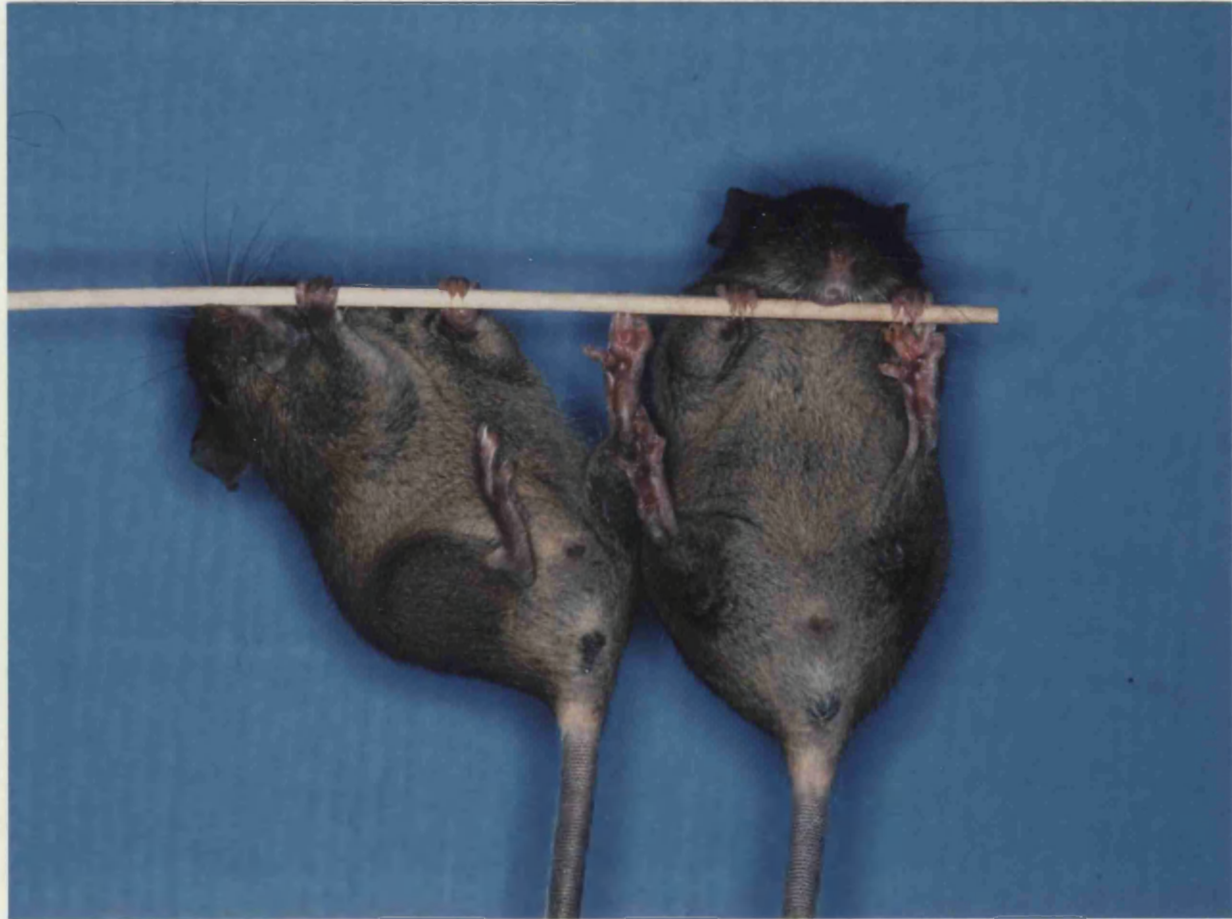
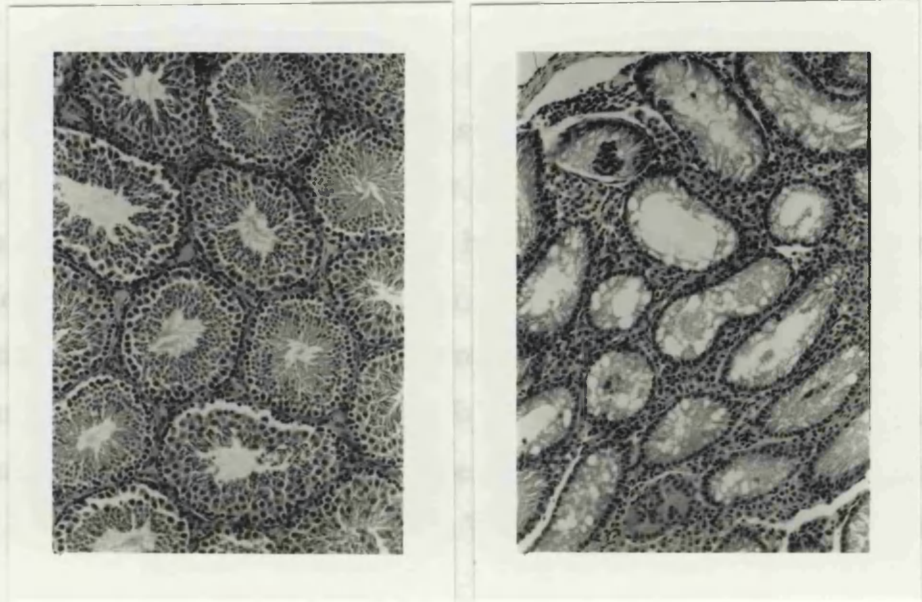


Figure 25. Analysis of adult sex-reversed transgenic mouse m33.13. a, PCR analysis of genomic DNA from m33.13 (lane 3), showing *Sry* and control (myogenin) bands. No band corresponding to *Zfy-1* was seen, demonstrating the lack of a Y chromosome; This result was confirmed by Southern blotting using Y-chromosome probes Y353B (Bishop,CE *et al.*, 1985) and Sx1 (Roberts *et al.*, 1988) (data not shown). Normal XY male and XX female littermates (33.17, lane 1 and 33.9, lane 2) are shown for comparison. M, marker bands (1018, 510, 396, 344, 298, 220, 201, 154 and 134 base pairs). b, External genitalia of XX male m33.13 (right) and a control nontransgenic XY male littermate, m33.17, (left). No abnormalities of male development could be seen. c, Histology of testis sections from XX male m33.13 (right) and the control XY male (left). m33.13 shows a complete lack of spermatogenesis, however other cell types and testicular organisation appear normal.

b



c



reasons: the presence of two X chromosomes is incompatible with spermatogenesis and sequences on the Y chromosome are required for spermatogenesis to occur (see section 1.5). Histological examination of sections of the testes revealed the presence of tubules, with clearly defined and apparently normal populations of Leydig cells, peritubular myoid cells and Sertoli cells, but no cells undergoing spermatogenesis (figure 25c). Internal examination of m33.13 revealed a normal male reproductive tract with no sign of hermaphroditism (data not shown). This indicates that Sertoli and Leydig cells must have functioned normally in producing AMH and testosterone respectively, as these determine the fate of the reproductive tract.

Two further transgenics, m32.10 and m33.2, showed an external female phenotype, yet both carried many copies of *Sry*. These mice have produced offspring and so have functional reproductive tracts and ovaries. They also provide further evidence, along with the XX female fetuses, that f741 does not always cause sex reversal.

8.3 *Sry* transgenic mice

Final proof that *Sry* is both necessary and sufficient to cause testis determination on an XX chromosomal background has come from the transgenic work described in this study. These experiments demonstrate that a 14.5kb mouse genomic fragment containing *Sry* is sufficient to direct the formation of testes in XX transgenic embryos and subsequently to give rise to full phenotypic sex reversal in an XX transgenic adult.

These conclusions are based on the assumption that the genomic fragment used does not encode an additional gene (or genes) that may also be necessary for testis determination. This formal possibility has been excluded on the basis of two lines of evidence (D. Jackson and A. Sinclair, unpublished results). (i) Hybridization of restriction fragments covering the entire length of f741 to genomic Southern blots of human DNA, indicates that besides *Sry*, no other sequences in f741 are conserved between mouse and man. (ii) Y-linked genes necessary for testis determination should map to the 35kb of the human Y chromosome which can give sex-reversal

in XX males. Cross-hybridization experiments between this region and f741, have shown that *Sry* is the only conserved sequence present. This has been confirmed by comparing the sequence of f741 to 25kb of sequence obtained at the human *SRY* locus. No significant homology besides that between the human and mouse *SRY* box was found. These data strengthen the conclusion that *Sry* is the only Y-linked gene required to give rise to male development. In other words that *Sry* is functionally equivalent to *Tdy*.

8.4 *Sry* does not always give sex-reversal

Although *Sry* alone can promote testicular development in the absence of other Y-linked genes, sex reversal does not always occur. There are a number of possible explanations for this. (i) Subtle rearrangements of the transgene may have occurred, rendering *Sry* non-functional. However, as four phenotypically female XX transgenics carrying one or more copies of *Sry* per cell have been observed, it seems unlikely that inactivating mutations have occurred in *Sry* during the generation of each of these transgenics. (ii) These females could be mosaic for the transgene. Mouse XX↔XY chimaeras develop as females or hermaphrodites if the XY component of the genital ridge is less than 25% (see section 1.6). Such chimaeras may be an appropriate model for *Sry* mosaicism. However mouse m32.10 (a non sex-reversed XX transgenic) has transmitted the transgene to female offspring, suggesting that mosaicism is not relevant in this case (Koopman *et al.*, 1991b). (iii) Expression of the transgene (and consequently phenotypic affect) may be influenced by the position at which it integrates. Except for a few cases where locus-controlling regions are present, expression of transgenes almost always depends on their chromosomal location (Grosveld *et al.*, 1987). This hypothesis can be tested by analysing *Sry* expression in embryos from lines carrying non sex-reversing integrations of the *Sry* transgene. XX transgenic progeny from m32.10 were found to express *Sry* in the genital ridge at 12.5*dpc* (P. Koopman, unpublished results). The simplest interpretation of this result is that lack of expression of the transgene in this case is not the cause of its failure to give sex-reversal. Nevertheless this assay cannot exclude the possibility that the absolute level of *Sry* expression from the transgene or its temporal regulation have been affected by the position of

chromosomal integration. Comparison of *Sry* expression from transgenes to that from the endogenous gene in XY embryos is not satisfactory as minor differences in embryo stage cannot be controlled for. In order to be able to compare *Sry* expression from transgenes to that from the endogenous gene, in a single embryo, an *Sry* transgene has now been constructed that carries a point mutation in the *Sry* box (data not shown). This alteration does not affect the amino acid sequence of *Sry* but will allow transcripts originating from the transgene to be distinguished from endogenous *Sry* transcripts by RT-PCR.

Recently (N. Vivian, unpublished results), heterozygous transgenic progeny from m32.10 have been mated in an effort to obtain mice which were homozygous for the transgene. Three mice out of a total of 15 progeny from a single cross, were phenotypically male XX transgenics. There are various possible explanations for why the transgene is giving sex-reversal in these cases. (i) If the level of *Sry* expression in embryos heterozygous for this particular transgene integration is insufficient for sex-reversal, then doubling this dose, which may occur in homozygous transgenics, may overcome this insufficiency. The correlation between sex-reversal and homozygosity for the transgene is now being studied. (ii) Alternatively, in this transgenic line, the temporal expression pattern of *Sry*, in XX transgenic embryos, may differ slightly from that of the endogenous gene in XY embryos. For instance *Sry* expression from the transgene may occur slightly later than from the endogenous gene. If this were the case then the genetic background may influence whether sex-reversal occurs or not. This model is identical to that proposed to explain the failure of the *poschiavinus* Y chromosome to cause male development on a C57BL/6 background (see section 1.7.2). The mice used in these transgenic experiments are F_1 hybrids between C57BL/10 and CBA. Extending the analogy with the *poschiavinus* system, it would be expected that the C57BL background, being associated with early ovary determination, would pre-empt a late-acting *Sry* transgene. This hypothesis is being tested by breeding the transgene onto a pure CBA or pure C57BL background.

8.5 Human *SRY* does not cause sex-reversal in XX transgenic mice

Parallel experiments have been carried out using a 25kb genomic fragment carrying human *SRY* (Koopman *et al.*, 1991b). In this case transgenic mice were not sex-reversed. Human *SRY* is expressed in transgenic embryos, in the genital ridge at 11.5*dpc*, suggesting that the mechanisms of *SRY* regulation are conserved between man and mouse. However as discussed above, subtle deviations from the wild type pattern of expression, may nevertheless be responsible for a failure of the transgene to cause sex-reversal. There are various other explanations for this finding. (i) The human *SRY* box differs from that of the mouse at 23 out of a total of 79 amino acids. These differences may result in the DNA binding properties of *SRY* being sufficiently different from *Sry*, that the human gene cannot interact with target sequences in the mouse. Single amino acid substitutions in *SRY*, in XY females (Berta *et al.*, 1990; Jäger *et al.*, 1990a), indicate that the function of the *SRY* box may be exquisitely sensitive to changes in its primary structure. Thus one gene may not be able to substitute for the other. This can be tested by exchanging the *SRY* box between the mouse and human genes. (ii) The mouse and human *SRY* genes may differ outside the *SRY* box (see section 5.3). This may prevent the human gene interacting with both other proteins and DNA targets in the mouse. (iii) It is possible that *SRY* transcripts are not translated in the mouse or that if a protein product is made it may not undergo correct post-translational modifications or may be unstable.

Note. All the work described in this chapter was the result of a collaboration between P. Koopman, N. Vivian, R. Lovell-Badge and myself, except for the PCR analysis of transgenic embryos which was carried out by P. Koopman alone.

The decision of a group of cells to differentiate into the components of a testis rather than an ovary during mammalian embryogenesis, ultimately decides the sexual phenotype of that organism. Certain features of sex determination have facilitated the molecular study described here. In particular, mutations affecting *TDF* or *Tdy* in both humans and mice affect only sexual fate and fertility. In contrast, mutations in genes controlling other developmental decisions are likely to have more serious if not lethal effects upon the embryo. The cloning and analysis of *Sry*, described in this thesis, has shown that the dominant male-determining affect of the mammalian Y chromosome is due to this single gene. Various aspects of the *Sry* gene and sex determination are discussed below.

9.1 Human XX males

The four XX males described by Palmer,MS *et al.* (1989), which only carry 35kb of Y derived sequences including *SRY*, vary in phenotype from normal male (although lacking germ cells) to hermaphrodite with bilateral ovotestes and a uterus. Jäger *et al.* (1990b) have also described an XX hermaphrodite with a similar small portion of the Y chromosome. One possible explanation for the incomplete male differentiation seen in some of these cases is that in addition to *SRY* another gene (or genes) on the Y chromosome is required for complete rather than partial testis determination. Such an argument would not explain the variability in phenotype seen in these XX males. In particular of the four patients described by Palmer,MS *et al.* (1989) two "brothers" carrying identical Y-chromosome translocations showed very different degrees of sex reversal. In addition, the work described here shows that *Sry* alone causes complete sex reversal in mice.

A more likely explanation for these phenotypes is that the proximity of *TDF* to the breakpoints in the XX males, makes its activity highly sensitive to position effects, at its new position on the X chromosome. Such a model was proposed to explain the failure of *Sry* to cause sex-reversal in some transgenic mice (see section 8.4). Another factor that may be contributing to the variable phenotypes in these cases

is X-inactivation. As *SRY* becomes X-linked in XX males, its level of expression may be affected by a spread of X-inactivation. Such a random process may account for the variability in phenotype seen even in familial cases of XX maleness. There are several precedents for this in the mouse; for example females or hermaphrodites frequently develop instead of males when the Sxr fragment is present only on the inactive X chromosome (McLaren & Monk, 1982).

9.2 Differences between *SRY* and *Sry*

The work described in this thesis suggests that human and mouse *SRY* differ in a number of ways. The inverted repeat structure found at the mouse *Sry* locus has no equivalent in human *SRY* (see chapter 6). In addition, as described in chapter 5, the human and mouse *SRY* transcripts and possibly proteins may differ outside the *SRY* box region. The absence of any homology between the mouse and human genomic *SRY* loci, besides the *SRY* box (see chapter 8) can be interpreted in a number of ways. Possibly the *SRY* box represents the whole *SRY* protein and is thus the only region conserved between man and mouse. This cannot be true for the human gene, for which more information is available on the transcript than in the mouse. However it may be that *SRY* sequences outside the *SRY* box have diverged between man and mouse because these parts of the protein are non-functional or at least have lesser constraints on their divergence than the *SRY* box itself. The latter might be true of domains such as acidic activating regions (Ptashne, 1988). If the proteins do differ significantly outside the *SRY* box it does not necessarily mean that the sex-determining pathway is different in man and mouse. If a functional domain is present in, for instance, mouse *Sry* but not human *SRY* it is possible that in humans the function of this domain is provided by a separate protein encoded by a distinct gene elsewhere in the genome. One could therefore imagine that in this case mouse *Sry* may have evolved by recruitment of such proteins as new domains (presumably at the level of exons being copied and transposed).

9.3 How is *Sry* regulated?

Sex determination in mammals is controlled by the presence or absence of *Sry*. However, in order for *Sry* to act a number of conditions must be met. (i) The gonadal anlagen must have formed in the appropriate part of the embryo. (ii) The appropriate cell types in which *Sry* will be expressed must be present in these anlagen. (iii) A mechanism must exist for *Sry* expression to occur at the appropriate time and at a suitable level. As morphologically identical gonadal anlagen arise in both male and female embryos, *Sry* cannot be involved in this process. In addition, the precursors of ovarian and testicular cell types are thought to be common to both the male and female indifferent gonad (see section 1.3), once more suggesting that the formation of these precursors is an *Sry*-independent process. The simplest models for how *Sry* transcription is controlled are as follows. *Sry* may be a leaky gene that does not need any specific mechanism of transcriptional activation. If this were the case it would be expressed in all cells like a housekeeping gene, even though only the appropriate cells in the genital ridge could respond to this expression. Clearly this is not the case, as Koopman *et al.* (1989) have shown that *Sry* has an extremely specific spatial and temporal expression pattern in the mouse embryo. The mechanism controlling this expression is presumably in place, in the indifferent genital ridge, in both male and female embryos, but the gene itself is only present in the former. By studying the regulation of *Sry* in the embryo, it may be possible to determine the factors and hence the genes involved in this process. These molecules will be the earliest markers of cells in the gonadal anlage. The isolation of such markers or indeed the genes involved in their activation may provide insight into the temporal and positional signals that control the formation of the gonadal anlagen which is competent to respond to *Sry*.

9.4 What are the target genes of *Sry*?

Sry initiates a cascade of events which lead to the differentiation of a specific organ from an undifferentiated primordium. Understanding these events should reveal much about both organogenesis and cellular differentiation. Assuming that *Sry* is

a transcription factor, then the question of how it controls cellular fate becomes one of what gene (or genes) is either activated or repressed by *Sry*. From this point of view *Sry* can be thought of as a master regulatory gene controlling first cell type and through this organ type. The *MyoD* family of genes are thought to play such a role in the determination of muscle cells. Transfection of these genes can convert fibroblasts into muscle (Davis,RL *et al.*, 1987). Preliminary results (M. Waters, personal communication) suggests that *Sry* does not have the same dramatic affects as *MyoD* on the phenotype of cells where it is not normally expressed. This could either be because the appropriate cofactors for *Sry* are absent or because the target genes are hidden (for instance due to their chromatin configuration).

Sry downstream genes can probably be classified in two general groups. The first of these are genes likely to be specifically involved in testis differentiation. Mutations in some of these genes could explain the occurrence of XX males lacking *SRY* (Palmer,MS *et al.*, 1989) and XY females where *SRY* is intact (Fredga, 1988; Scherer *et al.*, 1989; Berta *et al.*, 1990; Jäger *et al.*, 1990a). The second group would be genes that are involved in the development of many organs in the body, such as structural components of the extracellular matrix. Mutations in such genes would therefore have wide ranging, if not lethal affects. As *Sry* acts over a short period of time it seems likely that the latter group may be indirect targets of *Sry*, with other testis-specific genes acting as intermediates.

While there are a number of molecular genetic approaches that can be used to isolate *Sry* target genes, an alternative approach is to assess the interaction of *Sry* with genes which are candidates for being *Sry* targets by virtue of their function or expression pattern. The first molecular marker of Sertoli cell differentiation is the hormone AMH. Analysis of AMH expression in the developing mouse embryo, has shown that AMH transcripts are first detected 48 hours after *Sry* transcription begins. Although this does not exclude a role for *Sry* in the control of AMH expression, it suggests that this may be an indirect rather than a direct interaction (Münsterberg & Lovell-Badge, 1991). Steroid metabolism is another area of difference between the early ovary and testis. Dorizzi *et al.* (1991) have shown that in turtles, which show temperature-sensitive sex determination, temperature sex-

reversal occurs when oestrogens are injected into eggs. This results in ovarian development occurring even at the masculinizing temperature. This suggests that oestrogen levels play a key role in gonadal differentiation. The enzyme P-450 aromatase is responsible for the conversion of testosterone to oestrogen. Desvages & Pieau (1991) have shown that aromatase activity is higher at the feminizing temperature than at the masculinizing temperature. This suggests that, in turtles, temperature may determine the activity of this enzyme which in turn controls the levels of oestrogen in the egg. Aromatase is responsible for the conversion of testosterone to oestrogen in mammals as well as amphibians. Thus aromatase may play a similar role in the differentiation of the mammalian gonad. In this case its activity would not be controlled by temperature, but possibly by the presence or absence of *Sry*.

The candidate gene approach can be extended to any gene known to be expressed in the genital ridge or early embryonic testis. Genes expressed earlier than *Sry* and/or in both males and females are more likely to be involved in *Sry* regulation than *vice versa* . A candidate Wilms' tumour gene *WT1* has recently been cloned (Call *et al.*, 1990; Gessler *et al.*, 1990). The gene encodes a zinc finger protein and therefore may be a transcription factor. Besides its role in Wilms' tumour, which is an embryonal malignancy of the kidney, there is evidence to suggest that *WT1* is also involved in genital development (van Heyningen *et al.*, 1990; Pelletier *et al.*, 1991a). *WT1* has been found to be expressed in developing human gonads (Pritchard-Jones *et al.*, 1990) and in 12.5*dpc* mouse embryonic testes (Pelletier *et al.*, 1991b). Whether it is expressed earlier than 12.5*dpc* in the mouse is unknown. The expression pattern of *WT1* and its proposed role in genital development, have led Pelletier *et al.* (1991b) to speculate that *WT1* may interact with *Sry* in some way.

Haematopoietic growth factor (mast cell factor, *Sl* factor), which is encoded by the mouse *Sl* locus is expressed in the somatic portion of the genital ridge in 12.5*dpc* mouse embryos (Matsui *et al.*, 1990). Mice carrying mutations at this locus show abnormalities in stem cell systems, such as primordial germ cells (see section 1.6). *Sl* expression in the embryonic gonad may therefore function only in the processes of primordial germ cell migration or proliferation. However, a role in the

development of the somatic part of the gonad cannot be ruled out. Thus the possibility remains, that *Sl* and *Sry* interact in some way. Recently Tajima *et al.* (1991) have shown that *Sl* is expressed in adult Sertoli cells, if the same is true in the embryo then this would be consistent with an interaction between *Sl* and *Sry*. In contrast the *c-kit* gene which encodes a receptor for *Sl* factor is only expressed in germ cells of the genital ridge (Manova & Bachvarova, 1991), which are not required for either testis determination or differentiation. Other genes expressed in the genital ridge include a number of homeobox containing genes (Wolgemuth *et al.*, 1987; Izpisua-Belmonte *et al.*, 1990) and a member of the paired box family of genes, *Pax-2*, (G. Dressler, personal communication).

In conclusion, the cloning of *Sry* has provided a molecular tool to begin to genetically dissect both the earliest signals which govern the formation of the indifferent genital ridge and the cascade of genes which must be involved in testicular differentiation. This may be the first developmental system in mammals that has become amenable to such an analysis.

Appendix: Sequence of L741

```
   1 GATCTATGAG TCTGCACACA CATCATACAT GTCCCAAAAA AATAGAGTCG
  51 AATATGAACT ACTCTGGCCT CACTGTCTCC TACACTGACA TATGAGGATG
 101 CTATAAACAG TGGAACCTGG TACCTAAGCC TACCCCTGGA CTGGTATGGG
 151 ACCTGCATGC TAGTTCAACT GACCATGCCC TCTTCTCTGC CTTTCAGTTC
 201 TTCATTTTAA TAAAAGCTAA TTTTATGGTT CACACCTGTC TTAGTTAGGG
 251 TTTGACTGCT GTGAACAGAC ACTACGACCC AGGCAGGTCT TATAAAGGAC
 301 AACATTAAAT AGGGGCTGGC ACACTGGTTC AGTCTATTAT CATCAAGGCA
 351 GGAACATGGC AGCATGCATG TTTTGGTGGT AGGGGAGCAG CAAGCTGGGA
 401 AAATGCGTGC AAGCATCTAA CTGGAAATCT ACTGTGGTCT GTAACATAAT
 451 TGCCTGGTCC TGGGAGTTAT ATCATAAGCT TAACTTCTGC TCTCCTCTGC
 501 ATTGGTGGTC TTTGGTCAAG GGCTGGGCTT GTAACCTAGG ACTGCAGGTT
 551 TTTTGGGGAA ATAACCTAGA GATGCTGGTT TTCTTGAGGG TGTAACCTGG
 601 AAACTCTGGT CTTGGTGGGG ATTAGCCTAG AGACTGGAGC TTGGCTCAGG
 651 TTTTGTTGGG AGGCAACTTG GAAACTAATG CTAGGTATCA GCCTGTTAGT
 701 TTACCTGAGT TCAAACTTAG GTCAGGTTCT CTAAAATGGA GTCTGAACTT
 751 TAAAAAAATA ATGGCATCTC AAGTCAACAT GGTAGAATCT AGAAGACCTG
 801 TGAGAGGTGA TGGTTTATCT AAACATGTGG GTGGAGGTTC TGCAACTTCC
 851 TGACAAATGG CACATATGTC ATGAACTGGG GATCCCGTAG AACAGAGGGA
 901 TAAGGCCTTA GTAGAATCAG CCCTCTGGAG AGTATGGGCC AGAATGTGAG
 951 GTGAATTCCA AGTGAAGTCA GAGGAGTCAG GGGGCTGTTC TTCCCTCTGC
1001 CTGTGCTGGT TGGGCAGGGG CCAAGGCTGA CCCTCCTTTA GAGGAAGAAG
1051 GGTGTGCGGT GGTGAGGCCT GTAAGAAAAA AGTTCTCATC CAAACAGAGG
1101 GCCTCTCCAC CAGGTATTGC CACATGGTGA TGTATGGCAC TTGGTCCAGG
1151 TGGCCCGAAG TATCACCCAA AAAATGATGG GGAGATGGAA AGTCCCTCCT
1201 CAGATCATAG GTGGGCCACT CATTCCTGCA AAAGATGGTT AGTTTGTAGG
1251 GACACCAAGC AGACCCACGT CTTCTATGAT CTCTGTCACA AAGTCCTTAA
1301 AATGGTTCAG GATCAGGTCC AGTGCAGTGC TTTATGCCTT TCCATCCTAG
```

1351 GAATTAAAAA AGAAAAAAAT TGTACAAGAG GACAAAAAGC ACAAACAAAA

1401 TACACACAAA CACAAAAGGA CATTCCCAGA AAACCAAAAC CAAAACCAAA

1451 AGGCTATCTC ACATGGTCTT CCTCCCTGCT CATGGGAGGC CCATAGACCC

1501 ACAGCCAGTT CAGGATTCAG ATGAGAACTA GCAAGTTCTC CTGCTTCTGG

1551 TAGGGAATCT GGGAGAGTCT TCCATTACCC CTGGCCAGGT AACTATGAGT

1601 CGGCCTTGAT CCTCTTGGCC TTCTGATTAA AACGGAAACA GAATTAACGG

1651 AGAAATAGAA ACATAGATAT GGAAAAGATG AAACAAAACC AGGCACCTGG

1701 ACTTACCAAT CAGACAAAGC CTCCACTCTC GGGGGTTGGG GGGGGGTACT

1751 AGAGTGAATC ACAGAATGAG CACCCAAAAG TAAGGGTCCA TGATTCACTG

1801 AAGAATGATA CCCCAGACTC AAACAGTATG TAAACGCAAA GGTTGTTTCA

1851 TTCTCCCAAA GTCCTGCATA CTGGGATCTC CCATTAAGAT AGAGAGACAC

1901 CAGAGTGAGT TAGGCCTGAT TTAATGTGCA CTCAGGGGAT TCCTGGCTAG

1951 GTGACCTTAT CCTACTCTAT CTCTCAGGGA CATTCCAGAA CATTACCTGG

2001 GTGTGAGGGC TGGAAACTGT TGCTGACCCA CTGTCCTTAC TTTAGGCCAA

2051 GTGGAGGAAC ATTTTGTGAT TAGTTGCTTA AGGCCTGGGT TTTTAGGCAT

2101 AATCCCCTGA ACTGCCTATT TGAAGCCTGT CATGGAGTCA GCCTGTCTCA

2151 TCACATTTAA GTACACTTAA GCTTCATGCT GCCAAAATAT TGACACCTCT

2201 CACTGCTGTT CCAATAAACC ACTGTCAGCC TGATGAATTT GAGTCTGTCT

2251 GCCTCTCTGC CTCATTTTTC TTATGTTCCT CTGCAATCAT TCAAATGGAA

2301 GCACACACAC TCTCTGAGCA CATTACATAG GCTAATGGTT ACTTAAGTGA

2351 CCTAGATCTC TGGCCTCTTG TAGCCTCTGC TTTCTTATGA AGGCAATTTC

2401 AGCTCCTGGG ACTTCATTAG TACTCTCTGT CAGCCCAGCT TCATCACAGT

2451 TTGCTCAGCT GTCCTTTGAG TCCTCCCCAT CCATCTTCCT AAGACTCTTC

2501 CTCGGACCTC TTCTTCAAGA GTCTCTGATC CTCAGGATAA GAGAGCTTTT

2551 TCCATTGGAT TGCTATCATC ACACACCCAA ATTGTGCAGG ACAAGGTAAC

2601 TGTCCATCCT GGTCTGACTG AGAGGTGCCC TGACCAGATC CCAGAAGCTG

2651 TGGGAGCTTC TCATTCCTTC TCTGGATCAC AATGTCCTGG AGGCTTCTAG

2701 GCACCTCCCA GTTCTTGACT GACATGACAA AGTCCCACCA TCGTCTTATG

2751 AGTATTGGGG TGCCACATCG GGCTCATTTT CATCTTTGCC AAAAACTTAG

2801 GAGGCAGTGG GATAAAACAC CCTACCAGTC CTGCCTGATA TGTCAAATTT

2851 AAGCCAGACT TAAAGATGGA CTCTAAAACT CTGTCTTAGT CCACCAAACA

2901 CAGTTATCCA CTCGAAGTAG CCACATTCTC AGATAATGGC CACATTTTTA

2951 AACCTTTCAA TTTAAGGTCT TCAAAACCTA TGATAAGTTT CTGAAAGAAA

3001 TGACAAGCCT GGGCCCACCT TCCTGATCCA CCCACCTGAG **G**CAATCATTC

3051 CCTCACCCAC TAACCACTAA CCCCTCCCTA TCTTTTATCC ACCTACCCCC

3101 CCCCC**CCC**AA CCCCAGTTCA TTTTAAGCCT TCTCAACCTG GCCACTCAGC

3151 TCCAAGTTGT AAGATTATAA GTAGGCCCAG AACCCAAGCT TAACAGAGCT

3201 GGGAGGGAAA TTGCAGATGA ATTCCACACT GAGGCCTAAT GCCTGCTGTT

3251 TCTTCTGTGT CTCAAAGCCT GCTCTTCCTA CCTGAAATAG GATAGTCCTT

3301 TTTTCCCCCT TGGTACTTCT TTACCTTATT ACCACACCAA GACTTAATAC

3351 TCTTGCTTTT ATTTCTCCTT TCAGTTACCT GTTATGCCTT TCAGATTGTC

3401 ACCCTCTATC TGAATTCAT**G TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG**

3451 **TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG T**ATGTGTGTG TGTGTGTGTT

3501 TGTGTGTGTC TTTTTGTGTG TGATGTGTGA TTGGCGATTG TTGTCTTTGC

3551 TTTGTACACA CACTTTACTT TG**CACACACA CACACACACA CACACACACA**

3601 **CACACA**GAGT GTTTATTGAT ATGTAAATCA CTGGATTCAG TTTACAGAGA

3651 TTCAGATGTC GTTTGGAATA TGATAGCTGC TAGCAGTACA TGTTGCTTGT

3701 TCAGCTTTGC TTCCTGATTC AGATGTGTGC CCACTGTTGC TTTAGGTTGC

3751 TTCAGGCCTG CCTAACCACA CCACCTCCTA CATCCAAAGT CTCAGCAGGC

3801 ACCCCAGGCC TGGCATCTCT CACGAATGGT ACAGTTCTCA ATTCCAGAAT

3851 CCCTCACCAC TAACCCTACA TGGATTTGGA GGAGTTAGCA AGAAAGGCCT

3901 CCTACATGGG GCAAAGGACT GGCCAGTGAT TTGTCCTAGA CAGAGGCAGG

3951 CTCTAGGTCA ACATACTTCT CACACCTGCA TTCCACACCT CCTGAACCTT

4001 GGACAGCTAA CACTGATCTT TTCTCAGAAT GTCAGGGAAA GCAGAAGTCT

4051 GGAATCTCCC ACGGATGGGA AGTGCAAAAG TCCCCTGCAA AATTGAAGGT

4101 CACATAGGCT AATAGAACTG TCCAGGTCCT GGGCTACATA GACCTAGGGA

4151 CCCTGGGTGT TTCCTCTATG CCCATCAGCC TCCCTCTTCC CAAAATGGTC

4201 ACAAATGTCA ATCATGCTCT TGATTGCCTT CTCAGCACAG AACAAGTGCC

4251 AGCCTCATTT GGGTGACTGG TGGCCCCTCC TCCTTGCCCT GTTTCCCCAC

4301 AGCAGCTCTC AGACAAGAGG TTCATGGGTT TTGTCCTGAG GACACCTGGC

4351 TGTTCTCTAT CTTGCCAACT CTCAACCTGG CATGACCCCC TCCTGCTCTT

4401 AATATCAAAG CTGTGCATAG AGCAAGCACC CAGGCTTTCC TAGAAATGCA

4451 TCCTGTGTCA CAATGGAATG AAGCTGTCAG GGCCTTAAAT GTCTTGACTT

4501 GGTTGAGCAG AGGTTATATG CAGA**A**CTTTC TCGCGGGAAT ACACACTCCT

4551 TCTGTTTTGT TGACCTCCAT AAATACTTTA TGTTCCAATT TTAAATGCTC

4601 ACTGGCCATC ACCTGAACAG ATTTCCTCTT TCCCTTTCCA CATAATGAGT

4651 CTCATGATTT TCTTTCCCAA CCTCACTGTA TTTGGAAACT AAAGTTTGCC

4701 AACACTAATG TTTTAGAAAT CCTTTATTAA TGGCATTTGA ATTATTCCTA

4751 AAGATGAACA ACTGAGAATG ACATAATCAC AAGCTATGAT TCTGGCTCAG

4801 TAAGTAAGAG CACTTGGACT TCATTTAGAG CTTTCTATTA TCTTTGTAAC

4851 AAGCAGGCAT TGAAGCCAGG TGCTATGTAA GCTTTCCAGA AGAACAGTAT

4901 CCCAGGGAAT ACTTTTTGCT AATG**GTGTGT GTGTGTGTGT GTGTGTGTGT**

4951 **GTGTGTGTGT GTGTGTGTGT GTGTGT**GATG GCAATTGTCT TTGCTTTGCT

5001 TTCAATTGCA AGTATCAGAG AGCAAACAAC CTCATAGTCC CTGGTCAGTA

5051 GGGAGTTAGC ATCCAGGGAA CAACAACTTC ATATTTCAGA GATAAAAGAA

5101 AGTTCTTTAG TTAAACCAAA ATCAACTAAT ACCAGTTTAA TTTTATAGTT

5151 TGTGTCATTA AATGTTTTGA AACCCTGGAC ATATTTTTGT TCAAAAGGCA

5201 AAAATAATGC TCCTTGTCAC AAATCCAGAG TTTCAATATA CTCTTTGTGG

5251 GAATACAAAA CAAAACAAAA ACAAACAAAC AAACAAACAA AAAAACAAGG

5301 ATACTTCGTA TTTACATTGT CCTTTATCCT GTCTAGTTTA GATTTTCAGT

5351 GAAATCTCAT GCAATTTGCT TGCAACTCTC CTCCTTAATA TGTTTGTTTG

5401 CTCTAAAACA TTAAAAGTTG GGTGGTAGGA CAAAGTATTT CAGGACTCCT

5451 TGAGCTGTCC TGTTCTGTAA GCCCCGTCTT CACCATTGTT CTAGAAGTTT

5501 CTCCCAGATA TTGGTGCTGT GGTCAGTGCC TTGGTTTGGA TTTCTTGGTG

5551 TGGCCCATGA AAAGTGCATA GCAGGAAGTA TCAGCAGAGG TATGTGTGTA

5601 GGGGGATGGC ATCTAACAAA AGAACAGCAT CTCATGAGCT CTCCAAGAAT

5651 AGCCACCAAT CCCAGCACTC ACCTCTAAGC AGATAGTCGA GTGAACATTC

5701 TGGAGAATGC TAAAGGTATA TAAATCTTAG GCAGAACACT GCTTGCAAAA

5751 TTCAAGCCAG AATTAACTCG AAAGCTCTCT TGCACAAGGA TACTCTGTGC

5801 ACAAAGATGA AGGAGAGAGA GGAGGAGGAA GAAGAAAGAG GAGGAGGAGG

5851 AGGAGGAGGA AGAAGAGGAG AAAGAAAGGA GAGGAAGAAA GAGGAGGAGG

5901 AGGAGGAGGA GACAGAGGAA GATATGTAGC AAGTAGGTAC AGGTGAAAAA

5951 GAGGAAGCCT GTGGAATGCC AAAGGATCAA TGAGGTCCTT GTCCTGTATC

6001 TATGGCTCTC TGCATACAAT TCACACAGAG GAGAATAAAA TCCACAAATG

6051 ACATGGAGAA ACAATAGCAG AAACTCTCCA GAGAAAGAGA CATAAGCACC

6101 CATTGGGATG GATATATTTG TATTCATTTG TGTTAGAAAT CTATCCAAAG

6151 GCTTATAAAG TCTTATTGAT TGTTGGCATT GAAATGGTGC TGAAACATAC

6201 CCTTAAAAGA AAAACTAACC ACTTAGAAGT CCAGTAGTGC CAGCTAACTG

6251 TTCCATTGAA TTCTAGAGAA CAGTCTTTTA ACTGATTTAC AGAAAACAAT

6301 ACATGTAAGT TACTGTTCAC AATACATAAT TTTGAAAATC AACTAAAACT

6351 CAAAGGTCTT AAGACCATGT AGAACAGTAG TTTAATTGTT TCCAAGAGTT

6401 CAATACATCA GATAATACAC CATGTTGTTG TATAGTATTT AGTGTATAAG

6451 ATGTCACAGT TGAGGCTCCT TGGCTGCTAT TTTCTTCCAA CCTATGTTTT

6501 AGAGTTGACA GGGCTACATA ATCAATTCAA AAGTTCATGG GTAACAGAGT

6551 TATTATGAGG GAAGCATCGA GGGAGATTAA TGTGATGAAA TTGGAACAGG

6601 CTAACTGAAG GTCTGGGTAT GTAACATGGT GTTAACAAGC CTATGTGTAG

6651 TTCCTTGGTC TTTCTTAAAG AAACTGTTGC AGGGGTTGCT GTGTACACTC

6701 TTTATGTCTC TTTAAACTTG TTTGTAAGAA CAAGAGTGGG TAACTTTGAA

6751 CCATTCCTTA GCTTAACAAT CTGGCAGTTG AGTTAATGTG CAGATGGCCA

6801 TTCATTCATC CCACATATAC TTGCCCTCCC CTTCACCTTT TGTTTTATTC

6851 CAATCATTTT GTATATGTGG TGTCTTAAGT GTTGGTACAT CCACTTAGCT

6901 TGACTATTTG TCCATTTTGT TTCTGATTCT TAGCTAGCCC TTGAAAACAC

6951 AGCTTTGCTG TATTGTCAAT AAAACAGTCT ATTCTTTAGA GATAGATAAT

7001 AAAAGAGAAG CAAAATGTGA CTCCTTATGC ATAAAGGCTT ATTTTACCAA

7051 AAGATATAGA GAGCTCATAG CATAGTTCAT AGAACAACTG GGCTTTGCAC

7101 ATTGTGGAGG AGAACTAAAA TGTCATACAC ACACACACAC ACACACACAC

```
7151 ACACACACAC ACACACACAC ACACACACAC ACACACACTC ACAAGACTTT
7201 ACATACTCAT GATGTAGTTT TGTGTGCTAG AGAGAAACCC TGATAAACAG
7251 AAAGATGCAT TTGATCAACA AAGGAGTGTT TTGCATACTA AAGAGACACA
7301 GCCTGTGTTT TCACAGTATT AAGTTTGGTA TGTTTTTCAA AGATGAGTAT
7351 TGAAGACTTA CAGATGGTTT GCATGGTGAT GCTGTAAGGA AATGACAGCC
7401 AGCCCTGACC CGATGGCACT GTGTTTTTGC CTGCTCAGAA ATGAACTACT
                                          ←|
7451 GCATCCCAGT CATTAGAATG CTCAACCTGG ATGATTTTTA CCATGAAAGA
7501 ATTTTTGCTA TTATGTAAGA AATTTGTATA CATTACTTTT ATGATGCATT
7551 GATATGACCA AGAAAGGAAG CAAAAGTTCT TTTATGGTGT AGATATTTAT
7601 ACTGGGCTAT AGAGCCAGAA CAGACTATTT CTTAGTTTTT TCCATGCTGC
7651 CTGCCGTAGT AGACTATGAT ACTAATGATA GGTTCCATGC TTATCATTTT
7701 TAACCCTTTC CACTACTTTT GCAAGCTCGT TAATTTCTCA CGTTAGTCCA
7751 TCCTAAACAT TAGGTACGTT ACAGGGCAAA AGAACTCATC AGTATGTTCT
7801 TTGTATCTTA ATACTTGTTG TGACAGAATG ATCAAAGAAA AAGCGATATA
7851 TTTTTAAGGA GTGTGAAAGT GCAGTCTCTA ACATTCACAA ACTTGAGGGG
7901 TTTTGCTTGA CAGTATCTAG GTTCATGAGA TGGCTTCAAA GTAGATTAGT
7951 TGGGCTGGAC TAGGGAGGTC CTGAAGGAGG AGGGATAAAT ATTTTCTTAC
8001 ACACGTTAAA TATTAAAATC TTTGTTACAA GCACATTTTG GTCAGTGGCT
8051 TTTAGCTCTT ACACTTTAAG TTTTGACTTC TGTATCTGTC TGTCTCTGTC
8101 TGTCTCTGTC TGTCTGTCTG TCTTTGTCTG TCTGTCTGTT TCTCTTTGCA
8151 TCACATCTCT GTCTTTTTGG TTGCAATCAT AATTCTTCCA TTATGTTCAG
8201 CATTTCAGAT CTTGATTTTT AGTGTTCAGC CCTACAGCCA CATGATATCT
8251 TAAACTCTGA AGAAGAGACA AGTTTTGGGA CTGGTGACAA TTGTCTAGAG
8301 AGCATGGAGG GCCATGTCAA GCGCCCCATG AATGCATTTA TGGTGTGGTC
8351 CCGTGGTGAG AGGCACAAGT TGGCCCAGCA GAATCCCAGC ATGCAAAATA
8401 CAGAGATCAG CAAGCAGCTG GGATGCAGGT GGAAAAGCCT TACAGAAGCC
8451 GAAAAAAGGC CCTTTTTCCA GGAGGCACAG AGATTGAAGA TCCTACACAG
8501 AGAGAAATAC CCAAACTATA AATATCAGCC TCATCGGAGG GCTAAAGTGT
```

8551 CACAGAGGAG TGGCATTTTA CAGCCTGCAG TTGCCTCAAC AAAACTGTAC

8601 AACCTTCTGC AGTGGGACAG GAACCCACAT GCCATCACAT ACAGGCAAGA

8651 CTGGAGTAGA GCTGCACACC TGTACTCCAA AAACCAGCAA AGCTTTTATT

8701 GGCAGCCTGT TGATATCCCC ACTGGGCACC TGCAGCAGCA GCAGCAGCAG

8751 CAGCAGCAGC AGCAGTTCCA TAACCACCAC CAGCAGCAAC AGCAGTTCTA

8801 TGACCACCAC CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGTTCC

8851 ATGACCACCA CCAGCAGAAG CAGCAGTTTC ATGACCACCA CCAGCAGCAA

8901 CAGCAGTTCC ATGACCACCA CCACCACCAC CAGGAGCAGC AGTTCCATGA

8951 CCACCACCAG CAGCAACAGC AGTTCCATGA CCACCAGCAG CAGCAGCAGC

9001 AGCAGCAGCA GCAGCAGTTC CATGACCACC ACCAGCAGAA GCAGCAGTTC

9051 CATGACCACC ACCACCACCA ACAGCAGCAG CAGTTCCATG ACCACCAGCA

9101 GCAGCAGCAG CAGTTCCATG ACCACCAGCA GCAGCAGCAT CAGTTCCATG

9151 ACCACCCCCA GCAGAAGCAG CAGTTCCATG ACCACCCCCA GCAGCAACAG

9201 CAGTTCCATG ACCACCACCA CCAGCAGCAG CAGAAGCAGC AGTTCCATGA

9251 CCACCACCAG CAGAAGCAGC AGTTCCATGA CCACCACCAG CAGAAGCAGC

9301 AGTTCCATGA CCACCACCAG CAGCAACAGC AGTTCCATGA CCACCACCAG

9351 CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGTTCC ACGACCAGCA

9401 GCTTACCTAC TTACTAACAG CTGACATCAC TGGTGAGCAT ACACCATACC

9451 AGGAGCACCT CAGCACAGCC CTGTGGTTGG CAGTCTCATG ACACTGGCCT

9501 TTTCTCCTAC CTATGCCAAC ACTCCCCCTT GCTATGATTT TTAAGTCTGA

9551 GATTCCTAGT TGGTTTAGAT ACTGACTTTT CTCCTGTCAT GAAAAAGGTC

9601 CTGCTCCTTC CTTTTGGGTT CTAAAATGTG TGTAAAGATG TGTAGTGTGT

9651 AAGGAAGGAG AAAGCTTAGA AATGCTTCTA CATAGTTTTT GTTTAACATA

9701 GTACTAGTTG CATTGTTGTG GCATTTAATC CACAGAAGAA ACCAGAGATT

9751 TACACTCAGA TCCAATTTAG ATGAAATGCA TTTATTCTTA GGCTAGCACA

9801 AAGACAGCAA ACAGCTTTGT GGTCAATACA GCATGCCATG CTAAAGTGGG

9851 GGCTAGGGGT GAGTAAGCTG GAGTTTTAAA GGAAACATCA GAATGACACA

9901 TTGTAATGAC AAATATCCAG CTGGGCAAGA AAATTCTTTT CATTCCTGTT

9951 ATGTAGTAAT AAAAGACCAT TTCATCCCCT TTTCCACAGA TAATCAGGTT

10001 TTACAGGAAC AAAGACAGCT GTGAATTCCT CACAGCCTAT TAGCCTATTC

10051 TTATCTTACT CTTTGGCCAA TAGGATCTCC CAGGCATCCC AAGGCAAAGG

10101 TCAAGGACTC TTACTAGATG CCTGGCTGCA TGTTAAATTT CTTTAGCCAC

10151 CACCGGAGAC ATATATATAT ATATATATGT GTATATATAT ATATAATATG

|→

10201 TAGGTATGTA TATAATCAT**T** CAGGTTGAGC ATTCTAATGA CTGGGATGCA

10251 GTAGTTCATT TCTGAGCAGG CAAAAACACA GTGCCATCGG GTCAGGGCTG

10301 GCTGTCATTT CCTTACAGCA TCACCATGCA AACCATCTGT AAGTCTTCAA

10351 TACTCATCTT TGAAAAACAT ACCAAACTTA ATACTGTGAA AACACAGGCT

10401 GTGTCTCTTT AGTATGCAAA ACACTCCTTT GTTGATCAAA TGCATCTTTC

10451 TGTTTATCAG GGTTTCTCTC TAGCACACAA AACTACATCA TGAGTATGTA

10501 AAGTCTTGTG A**GTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG**

10551 **TGT**ATGACAT TTTAGTTCTC CTCCACAATG TGCAAAGCCC AGTTGTTCTA

10601 TGAACTATGC TATGAGCTCT CTATATCTTT TGGTAAAATA AGCCTTTATG

10651 CATAAGGAGT CACATTTTGC TTCTCTTTTA TTATCTATCT CTAAAGAATA

*

10701 GACTGTTTTA TTGACATACA GCAAAGCTGT GTTTTCAAGG GCTAGCTAAG

10751 AATCAGAAAC AAAATGGACA AATAGTCAAG CTAAGTGGAT GTACCAACAC

10801 TTAAGACACC ACATATACAA AATGATTGGA ATAAAACAAA AGGTGAAGGG

10851 GAGGGCAAGT ATATGTGGGA TGAATGAATG GCCATCTGCA CATTAACTCA

10901 ACTGCCAGAT TGTTAAGCTA AGGAATGGTT CAAAGTTACC CACTCTTGTT

10951 CTTACAAACA AGTTTAAAGA GACATAAAGA GTGTACACAG CAACCCCTGC

11001 AACAGTTTCT TTAAGAAAGA CCAAGGAACT ACACATAGGC TTGTTAACAC

11051 CATGTTACAT ACCCAGACCT TCAGTTAGCC TGTTCCAATT TCATCACATT

11101 AATCTCCCTC GATGCTTCCC TCATAATAAC TCTGTTACCC ATGAACTTTT

11151 GAATTGATTA TGTAGCCCTG TCAACTCTAA AACATAGGTT GGAAGAAAAT

11201 AGCAGCCAAG GAGCCTCAAC TGTGACATCT TATACACTAA ATACTATACA

11251 ACAACATGGT GTATTATCTG ATGTATTGAA CTCTTGGAAA CAATTAAACT

11301 ACTGTTCTAC ATGGTCTTAA GACCTTTGAG TTTTAGTTGA TTTTCAAAAT

11351 TATGTATTGT GAACAGTAAC TTACATGTAT TGTTTTCTGT AAATCAGTTA

11401 AAAGACTGTT CTCTAGAATT CAATGGAACA GTTAGCTGGC ACTACTGGAC

11451 TTCTAAGTGG TTAGTTTTTC TTTTAAGGGT ATGTTTCAGC ACCATTTCAA

11501 TGCCAACAAT CAATAAGACT TTATAAGCCT TTGGATAGAT TTCTAACACA

11551 AATGAATACA AATATATCCA TCCCAATGGG TGCTTATGTC TCTTTCTCTG

11601 GAGAGTTTCT GCTATTGTTT CTCCATGTCA TTTGTGGATT TTATTCTCCT

11651 CTGTGTGAAT TGTATGCAGA GAGCCATAGA TACAGGACAA GGACCTCATT

11701 GATCCTTTGG CATTCCACAG GCTTCCTCTT TTTCACCTGT ACCTACTTGC

11751 TACATATCTT CCTCTGTCTC CTCCTCCTCC TCCTCCTCTT TCTTCCTCTC

11801 CTTTCTTTCT CCTCTTCTTC CTCCTCCTCC TCCTCCTCCT CTTTCTTCTT

11851 CCTCCTCCTC TCTCTCCTTC ATCTTTGTGC ACAGAGTATC CTTGTGCAAG

11901 AGAGCTTTCG AGTTAATTCT GGCTTGAATT TTGCAAGCAG TGTTCTGCCT

11951 AAGATTTATA TACCTTTAGC ATTCTCCAGA ATGTTCACTC GACTATCTGC

12001 TTAGAGGTGA GTGCTGGGAT TGGTGGCTAT TCTTGGAGAG CTCATGAGAT

12051 GCTGTTCTTT TGTTAGATGC CATCCCCCTA CACACATACC TCTGCTGATA

12101 CTTCCTGCTA TGCACTTTTC ATGGGCCACA CCAAGAAATC CAAACCAAGG

12151 CACTGACCAC AGCACCAATA TCTGGGAGAA ACTTCTAGAA CAATGGTGAA

12201 GACGGGGCTT ACAGAACAGG ACAGCTCAAG GAGTCCTGAA ATACTTTGTC

12251 CTACCACCCA ACTTTTAATG TTTTAGAGCA AACAAACATA TTAAGGAGGA

12301 GAGTTGCAAG CAAATTGCAT GAGATTTCAC TGAAAATCTA AACTAGACAG

12351 GATAAAGGAC AATGTAAATA CGAAGTATCC TTGTTTTTTT GTTTGTTTGT

12401 TTGTTTGTTT TTGTTTTGTT TTGTATTCCC ACAAAGAGTA TATTGAAACT

12451 CTGGATTTGT GACAAGGAGC ATTATTTTTG CCTTTTGAAC AAAAATATGT

12501 CCAGGGTTTC AAAACATTTA ATGACACAAA CTATAAAATT AAACTGGTAT

12551 TAGTTGATTT TGGTTTAACT AAAGAACTTT CTTTTATCTC TGAAATATGA

12601 AGTTGTTGTT CCCTGGATGC TAACTCCCTA CTGACCAGGG ACTATGAGGT

12651 TGTTTGCTCT CTGATACTTG CAATTGAAAG CAAAGCAAAG ACAATTGCCA

12701 TCACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACACAC

12751 CATTAGCAAA AAGTATTCCC TGGGATACTG TTCTTCTGGA AAGCTTACAT

12801 AGCACCTGGC TTCAATGCCT GCTTGTTACA AAGATAATAG AAAGCTCTAA

12851 ATGAAGTCCA AGTGCTCTTA CTTACTGAGC CAGAATCATA GCTTGTGATT

12901 ATGTCATTCT CAGTTGTTCA TCTTTAGGAA TAATTCAAAT GCCATTAATA

12951 AAGGATTTCT AAAACATTAG TGTTGGCAAA CTTTAGTTTC CAAATACAGT

13001 GAGGTTGGGA AAGAAAATCA TGAGACTCAT TATGTGGAAA GGGAAAGAGG

13051 AAATCTGTTC AGGTGATGGC CAGTGAGCAT TTAAAATTGG AACATAAAGT

13101 ATTTATGGAG GTCAACAAAA CAGAAGGAGT GTGTATTCCC GCGAGAAAG**A**

13151 TCTGCATATA ACCTCTGCTC AACCAAGTCA AGACATTTAA GGCCCTGACA

13201 GCTTCATTCC ATTGTGACAC AGGATGCATT TCTAGGAAAG CCTGGGTGCT

13251 TGCTCTATGC ACAGCTTTGA TATTAAGAGC AGGAGGGGGT CATGCCAGGT

13301 TGAGAGTTGG CAAGATAGAG AACAGCCAGG TGTCCTCAGG ACAAAACCCA

13351 TGAACCTCTT GTCTGAGAGC TGCTGTGGGG AAACAGGGCA AGGAGGAGGG

13401 GCCACCAGTC ACCCAAATGA GGCTGGCACT TGTTCTGTGC TGAGAAGGCA

13451 ATCAAGAGCA TGATTGACAT TTGTGACCAT TTTGGGAAGA GGGAGGCTGA

13501 TGGGCATAGA GGAAACACCC AGGGTCCCTA GGTCTATGTA GCCCAGGACC

13551 TGGACAGTTC TATTAGCCTA TGTGACCTTC AATTTTGCAG GGGACTTTTG

13601 CACTTCCCAT CCGTGGGAGA TTCCAGACTT CTGCTTTCCC TGACATTCTG

13651 AGAAAAGATC AGTGTTAGCT GTCCAAGGTT CAGGAGGTGT GGAATGCAGG

13701 TGTGAGAAGT ATGTTGACCT AGAGCCTGCC TCTGTCTAGG ACAAATCACT

13751 GGCCAGTCCT TTGCCCCATG TAGGAGGCCT TTCTTGCTAA CTCCTCCAAA

13801 TCCATGTAGG GTTAGTGGTG AGGGATTCTG GAATTGAGAA CTGTACCATT

13851 CGTGAGAGAT GCCAGGCCTG GGGTGCCTGC TGAGACTTTG GATGTAGGAG

13901 GTGGTGTGGT TAGGCAGGCC TGAAGCAACC TAAAGCAACA GTGGGCACAC

13951 ATCTGAATCA GGAAGCAAAG CTGAACAAGC AACATGTACT GCTAGCAGCT

14001 ATCATATTCC AAACGACATC TGAATCTCTG TAAACTGAAT CCAGTGATTT

14051 ACATATCAAT AAACACTC**TG TGTGTGTGTG TGTGTGTGTG TGTGTG**CAAA

14101 GTAAAGTGTG TGTACAAAGC AAAGACAACA ATCGCCAATC ACACATCACA

14151 CACAAAAAGA CACACACAAA CACACACACA CACACAT**ACA CACACACACA**

14201 **CACACACACA CACACACACA CACACACACA CAC**ATGAATT CAGATAGAGG

14251 GTGACAATCT GAAAGGCATA ACAGGTAACT GAAAGGAGAA ATAAAAGCAA

```
14301  GAGTATTAAG TCTTGGTGTG GTAATAAGGT AAAGAAGTAC CAAGGGGGAA

14351  AAAAGGACTA TCCTATTTCA GGTAGGAAGA GCAGGCTTTG AGACACAGAA

14401  GAAACAGCAG GCATTAGGCC TCAGTGTGGA ATTCATCTGC AATTTCCCTC

14451  CCAGCTCTGT TAAGCTTGGG TTCTGGGCCT ACTTATAATC TTACAACTTG

14501  GAGCTGAGTG GCCAGGTTGA GAAGGCTTAA AATGAACTGG GGTTGGGGGG
           *
14551  GGGGTAGGTG GATAAAAGAT AGGGAGGGGT TAGTGGTTAG TGGGTGAGGG

14601  AATGATTGTC TCAGGTGGGT GGATC
```

Full sequence of *Sry* genomic clone, L741. This sequence was obtained by D. Jackson and A. Economou. A number of features are marked on the sequence. (i) The end points of the inverted repeat are marked by arrows pointing away from the unique region of L741 (between positions 7486-7487 and 10213-10214). These arrows define the two arms of the inverted repeat, each of which is the reverse complement of the other. (ii) Nucleotides which are not complementary in the two arms of the repeat are shown underlined and in bold. Nucleotides present in only one arm of the repeat, are also shown in bold and underlined and the position where these bases are missing in the other arm of the repeat is marked with an asterisk above the sequence. Nucleotides at the following pairs of positions are non-complementary or missing from one arm of the repeat: G 3041, T 14609; CCC 3106-3108, * 14554-14555; A 4525, A 13150; T 6964, * 10715-10716; G 7480, T 10220. (iii) CA/GT microsatellites which are polymorphic in length on either arm of the repeat are also shown in bold and underlined. These occur at the following pairs of positions: 3420-3481, 14188-14233; 3573-3606, 14069-14096; 4925-4976, 12702-12749; 7127-7188, 10512-10553. (iv) Nucleotides corresponding to the *Sry* box are indicated by double underlining. The open reading frame runs from left to right, 5'-3'.

References

Abdel-Hameed, F. (1971). Intersexes and sex determination in chickens. *Science* **172**, 962-964.

Affara, N. A., Ferguson-Smith, M. A., Tolmie, J., Kwok, K., Mitchell, M., Jamieson, D., Cooke, A. and Florentin, L. (1986). Variable transfer of Y-specific sequences in XX males. *Nucleic Acids Res.* **14**, 5375-5387.

Anderson, D. M., Lyman, S. D., Baird, A., Wignall, J. M., Eisenman, J., Rauch, C., March, C. J., Boswell, H. S., Gimpel, S. D., Cosman, D. and Williams, D. E. (1990). Molecular cloning of mast cell growth factor, a hematopoietin that is active in both menbrane bound and soluble forms. *Cell* **63**, 235-243.

Anderson, M., Page, D. C. and De La Chapelle, A. (1986). Chromosome Y-specific DNA is transferred to the short arm of the X chromosome in human XX males. *Science* **233**, 786-788.

Andersson, S., Berman, D. M., Jenkins, E. P. and Russell, D. W. (1991). Deletion of steroid 5α-reductase 2 gene in male pseudohermaphroditism. *Nature* **354**, 159-161.

Ashworth, A., Swift, S. and Affara, N. (1989). Sequence of cDNA for murine *Zfy-1*, a candidate for *Tdy*. *Nucleic Acids Res.* **17**, 2864.

Ashworth, A., Skene, B., Swift, S. and Lovell-Badge, R. (1990). *Zfa* is an expressed retroposon derived from an alternative transcript of the *Zfx* gene. *EMBO J.* **9**, 1529-1534.

Ashworth, A., Rastan, S., Lovell-Badge, R. and Kay, G. (1991). X-chromosome inactivation may explain the difference in viability of XO humans and mice. *Nature* **351**, 406-408.

Astell, C. R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K. A. and Hall, B. D. (1981). The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae*. *Cell* **27**, 15-23.

Attardi, B. and Ohno, S. (1974). Cytosol and androgen receptor from kidney of normal and *testicular feminized* (*Tfm*) mice. *Cell* **2**, 205-212.

Auffray, C. and Rougeon, F. (1980). Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* **107**, 303-324.

Bacci, G., ed. (1965). *Sex determination*. Oxford: Pergamon.

Baker, B. S. (1989). Sex in flies: the splice of life. *Nature* **340**, 521-524.

Behringer, R. R., Cate, R. L., Froelick, G. J., Palmiter, R. D. and Brinster, R. L. (1990). Abnormal sexual development in transgenic mice chronically expressing Müllerian inhibiting substance. *Nature* **345**, 167-170.

Bell, L. R., Horabin, J. L., Schedl, P. and Cline, T. W. (1991). Positive autoregulation of *sex-lethal* by alternative splicing maintains the female determined state in *Drosophila*. *Cell* **65**, 229-239.

Bell.LR, Maine, E. M., Schedl, P. and Cline, T. W. (1988). *Sex-lethal*, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* **55**, 1037-1046.

Bender, A. and Sprague, G. J., Jr. (1987). MATα1 protein, a yeast transcriptional activator, binds synergistically with a second protein to a set of cell-type-specific genes. *Cell* **50**, 681-691.

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 male sex determining gene. *Nature* **348**, 248-250.

Biddle, F. G. and Nishioka, Y. (1988). Assays of testis development in the mouse distinguish three classes of *domesticus*-type Y chromosome. *Genome* **30**, 870-878.

Billingham, R. E. and Silvers, W. K. (1960). Studies on tolerance of the Y chromosome antigen in mice. *J. Immunol.* **85**, 14-26.

Bishop, C. E. and Mitchell, M. J. (1991). The molecular biology of the mouse Y chromosome. *Semn. Devl. Biol.* **2**, 241-249.

Bishop, C. E., Boursot, P., Baron, B., Bonhomme, F. and Hatat, D. (1985). Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. *Nature* **315**, 70-72.

Bishop, O., Selman, G. G., Hickman, J., Black, L., Saunders, R. D. P. and Clark, A. J. (1985). The 45kb unit of Major Urinary Protein gene organisation is a gigantic imperfect palindrome. *Mol. Cell. Biol.* **5**, 1591-1600.

Bloom, S. E. (1974). Current knowledge about the avian W chromosome. *Bioscience* **24**, 340-344.

Borum, K. (1961). Oogenesis in the mouse, a study of the meiotic prophase. *Exp. Cell Res.* **24**, 495-507.

Brilliant, M. H., Gondo, Y. and Eicher, E. M. (1991). Direct molecular identification of the mouse pink-eyed unstable mutation by genome scanning. *Science* **252**, 566-569.

Brown, C. J., Powers, V. E., Munroe, D. L., Sheinin, R. and Willard, H. F. (1989). A gene on the short arm of the human X chromosome complements murine tsA1S9 DNA synthesis mutation. *Somat. Cell Mol. Genet.* **15**, 173-178.

Buckle, V. J., Boyd, Y., Fraser, N., Goodfellow, P. N., Goodfellow, P. J., Wolfe, J. and Craig, I. W. (1987). Localisation of Y chromosome sequemces in normal and XX males. *J. med. Genet.* **24**, 197-203.

Buhler, E. M. (1980). A synopsis of the human Y chromosome. *Hum. Genet.* **55**, 145-175.

Burgoyne, P. S. (1988). Role of mammalian Y chromosome in sex determination. *Phil. Trans. R. Soc. Lond.* **322**, 63-72.

Burgoyne, P. S. (1991). Y chromosome function in mammalian development. In *Advances in Developmental Biology*, vol. 1 (ed. P. M. Wassarman). Greenwich, Connecticut: JAI Press, Inc. (in press)

Burgoyne, P. S. and Palmer, S. J. (1991). The genetics of XY sex reversal in the mouse and other animals. *Semn. Devl. Biol.* **2**, 277-284.

Burgoyne, P. S., Levy, E. R. and McLaren, A. (1986). Spermatogenic failure in male mice lacking H-Y antigen. *Nature* **320**, 170-172.

Burgoyne, P. S., Buehr, M., Koopman, P., Rossant, J. and McLaren, A. (1988a). Cell autonomous action of the testis-determining gene : Sertoli cells are exclusively XY in XX↔XY chimaeric mouse testes. *Development* **102**, 443-450.

Burgoyne, P. S., Buehr, M. and McLaren, A. (1988b). XY follicle cells in ovaries of XX-XY female mouse chimaeras. *Development* **104**, 683-688.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C. and Housman, D. E. (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumour locus. *Cell* **60**, 509-520.

Cattanach, B. M. (1961). XXY mice. *Genet. Res.* **2**, 156-158.

Cattanach, B. M. (1978). Private Communication. *Mouse News Letters* **59**, 18.

Cattanach, B. M. (1987). Sex-reversed mice and sex determination. *Ann. N.Y. Acad. Sci.* **513**, 27-29.

Cattanach, B. M., Pollard, C. E. and Hawkes, S. G. (1971). Sex reversed mice : XX and XO males. *Cytogenetics* **10**, 318-337.

Chabot, B., Stephenson, D. A., Chapman, V. M., Besmer, P. and Bernstein, A. (1988). The proto-oncogene *c-kit* encoding a transmembrane tyrosine kinase receptor maps to the mouse *W* locus. *Nature* **335**, 88-89.

de la Chapelle, A. (1981). The etiology of maleness in XX men. *Hum. Genet.* **58**, 105-116.

Chen, A. and Falek, A. (1971). Cytological evidence for the association of the short arms of the X and the Y in the human male. *Nature* **232**, 555-556.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.* **162**, 156-159.

Ciccarese, S. and Ohno, S. (1978). Two plasma membrane antigens of testicular Sertoli cells and H-2 restricted versus unrestricted lysis by female T cells. *Cell* **13**, 643-650.

Cline, T. W. (1988). Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X:A sex determination signal in Drosophila that switch *Sxl* between two alternative stable expression states. *Genetics* **119**, 829-862.

Cline, T. W. (1989). The affairs of *daughterless* and the promiscuity of developmental regulators. *Cell* **59**, 231-234.

Conway, S. J., Darling, S. M. and Burgoyne, P. S. (1991). Spermatid specific expression of a multiple copy Y chromosomal DNA sequence which may be required for development of the sperm head. *Genes Dev.* (in preperation)

Cooke, H. J., Brown, W. R. A. and Rappold, G. (1985). Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. *Nature* **317**, 688-692.

Copeland, N. G., Gilbert, D. J., Cho, B. C., Donovan, P. J., Jenkins, N. A., Cosman, D., Anderson, D. and Lyman, S. D. (1990). Mast cell growth factor maps near the *steel* locus on mouse chromosome 10 and is deleted in a number of *steel* alleles. *Cell* **63**, 175-183.

Cronmiller, C., Schedl, P. and Cline, T. W. (1988). Molecular characterization of *daughterless*, a *Drosophila* sex determination gene with multiple roles in development. *Genes Dev.* **2**, 1666-1676.

Crow, J. F. and Dove, W. F. (1988). Unequal crossing over then and now. *Genetics* **120**, 1-6.

Dalton, A. J., Potter, M. and Merwin, R. M. (1961). Some ultrastructural characteristics of a series of transplanted plasma cell tumors in mouse. *J. Nat. Cancer Inst.* **26**, 1221-1267.

Darling, S. M., Banting, G. S., Pym, B., Wolfe, J. and Goodfellow, P. N. (1986). Cloning an expressed gene shared by the human sex chromosomes. *Proc. natn. Acad. Sci. U.S.A.* **83**, 135-139.

Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000.

Davis, R. M. (1981). Localization of male determining factors in man: A thorough review of structural anomolies of the Y chromosome. *J. med. Genet.* **18**, 161-195.

Deeming, D. C. and Ferguson, M. W. J. (1988). Environmental regulation of sex determination in reptiles. *Phil. Trans. R. Soc. Lond.* **B 322**, 19-39.

Desvages, G. and Pieau, C. (1991). Aromatase activity in gonads of turtle embryos as a function of the incubation temperature of eggs. *J. Steroid Biochem. Molec. Biol.* (in press)

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.

Donahoe, P. K., Cate, R. L., MacLaughlin, D. T., Epstein, J., Fuller, A. F., Takahashi, M., Coughlin, J. P., Ninfa, E. G. and Taylor, L. A. (1987). Müllerian inhibiting substance: gene structure and mechanism of action of a fetal regressor. *Recent Prog. Horm. Res.* **43**, 431-467.

Dorizzi, M., Mignot, T.-M., Guichard, A., Desvages, G. and Pieau, C. (1991). Involvement of oestrogens in sexual differentiation of gonads as a function of temperature in turtles. *Differentiation* **47**, 9-17.

Duyk, G. M., Kim, S., Myers, R. M. and Cox, D. R. (1990). Exon trapping: A genetic screen to identify candidate sequences in cloned mammalian genomic DNA. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8995-8999.

Dyche, W. J. (1979). A comparative study of the differentiation and involution of the Müllerian duct and Wolffian duct in the male and female fetal mouse. *J. Morph* **162**, 175-210.

Eicher, E. M. (1988). Autosomal genes involved in mammalian primary sex determination. *Phil. Trans. R. Soc. Lond.* **322**, 109-118.

Eicher, E. M. and Washburn, L. L. (1983). Inherited sex reversal in mice: identification of a new primary sex-determining gene. *J. exp. Zool* **228**, 297-304.

Eicher, E. M. and Washburn, L. L. (1986). Genetic control of primary sex determination in mice. Ann. *Rev. Genet.* **20**, 327-360.

Eicher, E. M., Beamer, W. G., Washburn, L. L. and Whitten, W. K. (1980). A cytogenetic investigation of inherited true hermaphroditism in BALB/cWt mice. *Cytogenet. Cell Genet.* **28**, 104-115.

Eicher, E. M., Washburn, L. L., Whitney III, J. B. and Morrow, K. E. (1982). *Mus poschiavinus* Y chromosome in the C57BL/6J murine genome causes sex reversal. *Science* **217**, 535-537.

Eichwald, E. J. and Silmser, C. R. (1955). Communication. *Transplantn. Bull.* **2**, 146-149.

Ellis, N. and Goodfellow, P. N. (1989). The mammalian pseudoautosomal region. *Trends Genet.* **5**, 406-410.

Ellis, N., Goodfellow, P. J., Pym, B., Smith, M., Palmer, M., Frischauf, A.-M. and Goodfellow, P. N. (1989). The pseudoautosomal boundary in man is defined by an Alu repeat sequence inserted on the Y chromosome. *Nature* **337**, 81-84.

Ellis, N., Yen, P., Neiswanger, K., Shapiro, L. J. and Goodfellow, P. N. (1990). Evolution of the pseudoautosomal boundary in old world monkeys and great apes. *Cell* **63**, 977-986.

Everett, N. B. (1943). Observational and experimental evidences relating to the origin and differentiation of the definite germ cells in mice. *J. exp. Zool* **92**, 49-91.

Ferguson, M. W. J. and Joanen, T. (1982). Temperature of egg incubation determines sex in *Alligator mississippiensis*. *Nature* **296**, 850-853.

Ferguson-Smith, M. A. (1966). X-Y chromosomal interchange in the aetiology of true hermaphroditism and of XX Klinefelter's syndrome. *Lancet* **ii**, 475-476.

Flanagan, J. G. and Leder, P. (1990). The *kit* ligand: A cell surface molecule altered in *steel* mutant fibroblasts. *Cell* **63**, 185-194.

Ford, C. E., Jones, K. W., Polani, P. E., de Almeida, J. C. and Briggs, J. H. (1959). A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* **i**, 711-713.

Ford, C. E., Evans, E. P., Burtenshaw, M. D., Clegg, H., Barnes, R. D. and Tuffrey, M. (1974). Marker chromosome analysis of tetraparental AKR-CBA-T6 mouse chimeras. *Differentiation* **2**, 321-333.

Ford, C. E., Evans, E. P., Burtenshaw, M. D., Clegg, H. M., Tuffrey, M. and Barnes, R. D. (1975). A functional sex-reversed oocyte in the mouse. *Proc. R. Soc. Lond.* **B.190**, 187-197.

Ford, M. and Fried, M. (1986). Large inverted duplications are associated with gene amplification. *Cell* **45**, 425-430.

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

Gehring, W. J. (1987). Homeoboxes in the study of development. *Science* **236**, 1245-1252.

Geissler, E. N., Ryan, M. A. and Housman, D. E. (1988). The Dominant-White Spotting (*W*) Locus of the Mouse Encodes the *c-kit* Proto-Oncogene. *Cell* **55**, 185-192.

Gessler, M. A., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H. and Bruns, G. A. P. (1990). Homozygous deletion in Wilms' tumours of a zinc finger gene identified by chromosome jumping. *Nature* **343**, 774-778.

Gillman, J. (1948). Development of gonads in man, with consideration of the role of fetal endocrines and histogenesis of ovarian tumors. *Carnegie Inst. Contrib. Embryol.* **32**, 81-131.

Ginsburg, M., Snow, M. and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. *Development* **110**, 521-528.

Glass, N. L., Vollmer, S. J., Staben, C., Grotelueschen, J., Metzenberg, R. L. and Yanofsky, C. (1988). DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. *Science* **241**, 570-573.

Goodfellow, P. J., Darling, S. M., Thomas, N. S. and Goodfellow, P. N. (1986). A pseudoautosomal gene in man. *Science* **234**, 740-743.

Goodfellow, P. N., Darling, S. M. and Wolfe, J. (1985). The human Y chromosome. *J. med. Genet.* **22**, 332-344.

Gough, N. M., Gearing, D. P., Nicola, N. A., Baker, E., Pritchard, M., Callen, D. F. and Sutherland, G. R. (1990). Localization of the human GM-CSF receptor gene to the X-Y pseudoautosomal region. *Nature* **345**, 734-736.

Griffiths, R. (1991). The isolation of conserved DNA sequences related to the Human Sex-determining region Y gene from the lesser black-backed gull (*Larus Fuscus*). *Proc. R. Soc. Lond.* **B 244**, 123-128.

Groot, P. C., Mager, W. H., Henriquez, N. V., Pronk, J. C., Arwert, F., Planta, R. J., Eriksson, A. W. and Frants, R. R. (1990). Evolution of the human α-amylase multigene family through unequal, homologous, and inter- and intrachromosomal crossovers. *Genomics* **8**, 97-105.

Grosveld, F., Blom van Assendelft, G., Greaves, D. R. and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975-985.

Guaellaen, G., Casanova, M., Bishop, C., Geldwerth, D., Andre, G., Fellous, M. and Weissenbach, J. (1984). Human XX males with single copy DNA fragments. *Nature* **307**, 172-173.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990a). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346**, 245-250.

Gubbay, J., Koopman, P., Collignon, J., Burgoyne, P. and Lovell-Badge, R. (1990b). Normal structure and expression of *Zfy* genes in XY female mice mutant in *Tdy*. *Development* **109**, 647-653.

Gubbay, J., Vivian, N., Jackson, D., Economou, A., Goodfellow, P. and Lovell-Badge, R. (1991). Inverted repeat structure of the mouse *Sry* locus. . (in preparation)

Guerrier, D., Tran, D., Vanderwinden, J. M., Hideux, S., van Outryve, L., Legeai, L., Bouchard, M., van Vliet, G., de Laet, M. H., Picard, J. Y., Kahn, A. and Josso, N. (1989). The persistent Müllerian duct syndrome: a molecular approach. *J. Clin. Endo. Metab.* **68**, 46-52.

Hall, M. N. and Johnson, A. D. (1987). Homeo domain of the yeast repressor $\alpha 2$ is a sequence-specific DNA-binding domain but is not sufficient for repression. *Science* **237**, 1007-1012.

Handley, P. M., Mueckler, M., Siegel, N. R., Ciechanover, A. and Schwartz, A. L. (1991). Molecular cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme E1. *Proc. natn. Acad. Sci. U.S.A.* **88**, 273-279.

Harley, V. R., Jackson, D., Hextall, P., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R. and Goodfellow, P. N. (1991). Sequence-specific DNA binding by the mammalian sex determining gene product, *SRY*, is abolished in some XY females. *Science*. (in press)

Harrington, R. W. J. (1975). Sex determination and differentiation among uniparental homogotes of the hermaphroditic fish, *Rivulus Marmoratus* (*Cyprinodontidae: Atheriniformes*). In *Intersexuality in the Animal Kingdom* (ed. R. Reinboth), pp. 249-262. Heidelberg: Springer Verlag.

Hatfield, P. M., Callis, J. and Vierstra, R. D. (1990). Cloning of ubiquitin-activating enzyme from wheat and expression of a functional protein in Escherichia coli. *J. biol. Chem.* **265**, 15813-15817.

Henthorn, P. S., Smithies, O. and Mager, D. L. (1990). Molecular analysis of deletions in the human beta-globin gene cluster:Deletion junctions and locations of breakpoints. *Genomics* **6**, 226-237.

van Heyningen, V., Bickmore, W. A., Seawright, A., Fletcher, J. M., Maule, J., Fekete, G., Gessler, M., Bruns, G. A. P., Huerre-Jeanpierre, C., Junien, C., Williams, B. R. G. and Hastie, N. D. (1990). Role for the Wilms' tumour gene in genital development? *Proc. natn. Acad. Sci. U.S.A.* **87**, 5383-5386.

Hilscher, W. and Hilscher, B. (1976). Kinetics of the male gametogenesis. *Androlog.* **8**, 105-106.

Hochgeschwender, U. and Brennan, M. B. (1991). Identifying genes within the genome: New ways for finding the needle in a haystack. *Bioessays* **13**, 139-144.

Hodgkin, J. (1987). Sex determination and dosage compensation in *Caenorhabditis elegans*. *Annu. Rev. Genet.* **21**, 133-154.

Hodgkin, J. (1989). *Drosophila* sex determination: a cascade of regulated splicing. *Cell* **56**, 905-906.

Hodgkin, J. (1990). Sex determination compared in *Drosophila* and *Caenorhabditis*. *Nature* **344**, 721-728.

Hogan, B., Costantini, F. and Lacy, E. (1986). *Manipulating the Mouse Embryo*. New York: Cold Spring Harbor Laboratory.

Huang, E., Nocka, K., Beier, D. R., Chu, T. Y., Buck, J., Lahm, H. N., Wellner, D., Leder, P. and Besmer, P. (1990). The hematopoietic growth factor KL is encoded by the *Sl* locus and is the ligand of the *c-kit* receptor, the gene product of the *W* locus. *Cell* **63**, 225-233.

Imperato-McGinley, J., Guerrero, L., Gautier, T. and Peterson, R. E. (1974). Steroid 5α-reductase deficiency in man; an inherited form of male pseudohermaphroditism. *Science* **186**, 1213-1215.

Izpisua-Belmonte, J.-C., Dolle, P., Renucci, A., Zappavigna, V., Falkenstein, H. and Duboule, D. (1990). Primary structure and embryonic expression of the mouse *Hox-4.3* homeobox gene. *Development* **110**, 733-745.

Jacobs, P. A. and Strong, J. A. (1959). A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature* **183**, 302-303.

Jäger, R. J., Anvret, M., Hall, K. and Scherer, G. (1990a). A human XY female with a frame shift mutation in *SRY*, a candidate testis determining gene. *Nature* **348**, 452-454.

Jäger, R. J., Ebensperger, C., Fraccaro, M. and Scherer, G. (1990b). A *ZFY*-negative 46, XX true hermaphrodite is positive for the Y pseudoautosomal boundary. *Hum. Genet.* **85**, 666-668.

Jantzen, H.-M., Admon, A., Bell, S. P. and Tjian, R. (1990). Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344**, 830-836.

Jeffreys, A. J., Royle, N. J., Wilson, V. and Wong, Z. (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* **332**, 278-281.

Jenkins, N. A., Copeland, N. G., Taylor, B. A. and Lee, B. K. (1981). *Dilute* (*d*) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. *Nature* **293**, 370-374.

Jirasek, J. E., ed. (1971). *Development of the genital system and male pseudo-hermaphroditism*. Baltimore: Johns Hopkins press.

Johnson, L. L., Bailey, D. W. and Mobraaten, L. E. (1981). Genetics of histocompatibility in mice IV. Detection of certain minor (Non-H-2) H antigens in selected organs by the popliteal node test. *Immunogenetics* **14**, 63-71.

Johnson, R. C., Bruist, M. F. and Simon, M. I. (1986). Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* **46**, 531-539.

Jongeneel, C. V., Acha-Orbea, H. and Blankenstein, T. (1990). A polymorphic microsatellite in the tumor necrosis factor alpha promoter identifies an allele unique to the NZW mouse strain. *J. Exp. Med* **171**, 2141-2146.

Josso, N. (1970). Action de la testosterone sur le canal de Wolff du foetus de rat en culture organotypique. *Arch. Anat. Micr.* **59**, 37-50.

Jost, A. (1947). Recherches sur la différenciation sexuelle de l'embryon de lapin. *Archs Anat. microsc. Morph. Exp.* **36**, 271-315.

Jost, A. (1953). Problems of fetal endocrinology: the gonadal and hypophyseal hormones. *Recent Prog. Horm. Res.* **8**, 379-418.

Jost, A. (1958). Embryonic sexual differentiation (morphology, physiology, abnormalities). In *Hermaphroditism, genital anomalies and related endocrine disorders* (ed. H. W. Jones and W. W. Scott), pp. 15-45. Baltimore: Williams and Wilkins.

Jost, A. (1970). Hormonal factors in the sex differentiation of the mammalian foetus. *Phil. Trans. R. Soc. Lond.* **B 259**, 119-130.

Jost, A. (1972). Données preliminaires sur les stades initiaux de la differenciation du testicule chez le rat. *Archs Anat. microsc. Morph. Exp.* **61**, 415-438.

Jost, A. and Magre, S. (1988). Control mechanisms of testicular differentiation. *Phil. Trans. R. Soc. Lond.* **B 322**, 55-61.

Jost, A., Vigier, B., Prépin, J. and Perchellet, J. P. (1973). Studies on sex differentiation in mammals. *Recent Prog. Horm. Res.* **29**, 1-41.

Kant, J. A., Fornace, A. J., Saxe, D., Simon, M. I., McBride, O. W. and Crabtree, G. R. (1985). Evolution and organisation of the fibrinogen locus on chromosome 4: gene duplication accompanied by transposition and inversion. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2344-2348.

Kay, G. F., Ashworth, A., Penny, G. D., Dunlop, M., Swift, S., Brockdorff, N. and Rastan, S. (1991). The mouse ubiquitin-activating enzyme E1: a candidate for the *Spy* gene, with an X-inactivated homologue. *Nature*. (in press)

Kelly, M., Burke, J., Smith, M., Klar, A. and Beach, D. (1988). Four mating-type genes control sexual differentiation in the fission yeast. *EMBO J.* **7**, 1537-1547.

Kimble, J. (1988). Genetic control of sex determination in the germ line of *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond.* **B 322**, 11-18.

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**, 940-942.

Koopman, P., Münsterberg, A., Capel, B., Vivian, N. and Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* **348**, 450-452.

Koopman, P., Ashworth, A. and Lovell-Badge, R. (1991a). The *ZFY* gene family in humans and mice. *Trends Genet.* **7**, 132-136.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1991b). Male development of chromosomally female mice transgenic for *Sry*. *Nature* **351**, 117-121.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 veterbrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8148.

Kuff, E. L. and Lueders, K. K. (1988). The intracisternal A-particle gene family structure and functional aspects. *Adv. Cancer Res.* **51**, 183-276.

Lau, Y.-F. C. and Chan, K. (1989). The putative testis-determining factor and related genes are expressed as discrete-sized transcripts in adult gonadal and somatic tissues. *Am. J. Hum. Genet.* **45**, 942-952.

Lee, P. A. and Gareis, J. F. (1975). Evidence for partial 21-hydroxylase deficiency among heterozygotes carriers of congenital adrenal hyperplasia. *J. Clin. Endocr.* **41**, 415-418.

Lehrman, M. A., Russel, D. W., Goldstein, J. L. and Brown, M. S. (1986). Exon-Alu recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3679-3683.

Leutert, R. (1975). Sex determination in *Bonellia*. In *Intersexuality in the Animal Kingdom* (ed. R. Reinboth), pp. 84-90. Heidelberg: Springer Verlag.

Levy, E. and Burgoyne, P. S. (1986). The fate of XO germ cells in the testes of XO/XY and XO/XY/XYY mouse mosaics: evidence for a spermatogenesis gene on the mouse Y chromosome. *Cytogenet. Cell Genet.* **42**, 208-213.

Lovell-Badge, R. H. (1987). Introduction of DNA into embryonic stem cells. In *Teratocarcinomas and Embryonic Stem Cells, a Practical Approach,; ed. E.J.* (ed. Robertson), pp. 153-182. Oxford: IRL Press.

Lovell-Badge, R. H. and Robertson, E. (1990). XY female mice resulting from a heritable mutation in the murine primary testis determining gene, *Tdy*. *Development* **109**, 635-646.

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus L.*). *Nature* **190**, 372-373.

Lyon, M. F. (1970). The activity of the sex chromosomes in mammals. *Sci. Prog.* **58**, 117-130.

Mager, D. L. and Goodchild, N. L. (1989). Homologous recombination between the LTRs of a human retrovirus-like element causes a 5kb deletion in two siblings. *Am. J. Hum. Genet.* **45**, 848-854.

Magre, S. and Jost, A. (1980). The initial phases of testicular organogenesis in the rat. An electron microscope study. *Archs Anat. microsc. Morph. Exp.* **69**, 297-318.

Mahadevaiah, S. K., Lovell-Badge, R. and Burgoyne, P. S. (1991). *Tdy*-negative XY, XXY and XYY female mice breeding data and synaptonemal complex analysis. *Chromosoma.* (submitted)

Manova, K. and Bachvarova, R. F. (1991). Expression of *c-kit* encoded at the *W* locus of mice in developing embryonic germ cells and the presumptive melanoblasts. *Dev. Biol.* **146**, 312-324.

Mardon, G. and Page, D. C. (1989). 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.

Mardon, G., Mosher, R., Disteche, C. M., Nishioka, Y., McLaren, A. and Page, D. C. (1989). Duplication, deletion, and polymorphism in the sex-determining region of the mouse Y chromosome. *Science* **243**, 78-80.

Martin, F. H., Suggs, S. V., Langley, K. E., Lu, H. S., Ting, J., Okino, K. H., Morris, C. F., McNiece, I. K., Jacobsen, F. W., Mendiaz, E. A., Birkett, N. C., Smith, K. A., Johnson, M. J., Parker, V. P., Flores, J. C., Patel, A. C., Fischer, E. F., Erjavec, H. O., Herrera, C. J., Wypych, J., Sachdev, R. K., Pope, J. A., Leslie, I., Wen, D., Lin, C.-H., Cupples, R. L. and Zsebo, K. M. (1990). Primary structure and functional expression of rat and human stem cell factor DNAs. *Cell* **63**, 203-211.

Matsui, Y., Zsebo, K. M. and Hogan, B. L. M. (1990). Embryonic expression of a haematopoietic growth factor encoded by the *Sl* locus and the ligand for *c-kit*. *Nature* **347**, 667-669.

McElreavey, K., Vilain, E., Abbas, N., Brauner, R., Nihoul-Fekete, C., Richaud, F., Rappoporte, R., Raymond, J. R., Toublanc, J. E. and Fellous, M. (1991). Conditional mutations in the sex-determining region suggests a 2nd sex-determining gene in humans. *Int. Cong. Hum. Genet.* **8**, 413.

McGrath, J. P., Jentsch, S. and Varshavsky, A. (1990). UBA1: An essential yeast gene encoding ubiquitin-activating enzyme. *EMBO J.* **10**, 227-236.

McLaren, A. (1983). Does the chromosomal sex of a mouse germ cell effect its development. In *Current problems in germ cell differentiation* (ed. A. McLaren and C. C. Wylie), pp. 225-240. Cambridge: Cambridge University Press.

McLaren, A. (1984). Meiosis and differentiation of mouse germ cells. In *Controlling Events in Meiosis. Soc. Exp. Biol. Symp*, vol. 38 (ed. C. W. Evans and H. G. Dickinson), pp. 7-23. Cambridge: Cambridge University Press.

McLaren, A. (1988a). Sex determination in mammals. *Trends Genet.* **4**, 153-157.

McLaren, A. (1988b). Somatic and germ-cell sex in mammals. *Phil. Trans. R. Soc. Lond.* **B 322**, 3-9.

McLaren, A. (1991). Development of the mammalian gonad: the fate of the supporting cell lineage. *Bioessays* **13**, 151-156.

McLaren, A. and Monk, M. (1982). Fertile females produced by inactivation from X chromosome of '*sex-reversed*' mice. *Nature* **300**, 446-448.

McLaren, A., Simpson, E., Tomonari, K., Chandler, P. and Hogg, H. (1984). Male sexual differentiation in mice lacking H-Y antigen. *Nature* **312**, 552-555.

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. natn. Acad. Sci. U.S.A.* **85**, 6442-6445.

Melton, D. W., Konecki, D. S., Brennand, J. and Caskey, C. T. (1984). Structure, expression, and mutation of the hypoxanthine phosphoribosyltransferase gene. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2147-2151.

Merchant, H. (1975). Rat gonadal and ovarian organogenesis with and without germ cells. An ultrastructural study. *Dev. Biol.* **44**, 1-21.

Meyer, B. J. (1988). Primary events in *C.elegans* sex determination and dosage compensation. *Trends Genet.* **4**, 337-342.

Migeon, B. R., Brown, T. R., Axelman, J. and Migeon, C. J. (1981). Studies of the locus for androgen receptor: Localization on the human X chromosome and evidence for homology with the *Tfm* locus in the mouse. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6339-6343.

Miklos, G. L. G. (1974). Sex chromosome pairing and male fertility. *Cytogenet. Cell Genet.* **13**, 558-577.

Miller, J., McLachlan, A. D. and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO J.* **4**, 1609-1614.

Miller, L. M., Plenefisch, J. D., Casson, L. P. and Meyer, B. J. (1988). *xol-1*: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in *C.elegans*. *Cell* **55**, 167-183.

Mintz, B. and Russell, E. S. (1957). Gene-Induced embryological modifications of primordial germ cells in the mouse. *J. exp. Zool* **134**, 207-237.

Mittwoch, U. (1969). Do genes determine sex? *Nature* **221**, 446-448.

Mittwoch, U. (1989). Sex differentiation in mammals and tempo of growth: Probabilities vs. switches. *J. theor. Biol.* **137**, 445-455.

Monk, M. and McLaren, A. (1981). X-chromosome activity in foetal germ cells of the mouse. *J. Embryol. Exp. Morph.* **63**, 75-84.

Münsterberg, A. and Lovell-Badge, R. (1991). Expression of the mouse anti-mullerian hormone gene suggests a role in both male amd female sexual differentiation. *Development* **113**, 613-624.

Nagamine, C. M., Taketo, T. and Koo, G. C. (1987). Morphological development of the mouse gonad in *tda-1* XY sex-reversal. *Differentiation* **33**, 212-214.

Nagamine, C. M., Chan, K., Kozak, C. A. and Lau, Y.-F. (1989). Chromosome mapping and expression of a putative testis-determining gene in mouse. *Science* **243**, 80-83.

Nagamine, C. M., Chan, K., Hake, L. E. and Lau, Y.-F. C. (1990). The two candidate testis-determining Y genes (*Zfy-1* and *Zfy-2*) are differentially expressed in fetal and adult mouse tissues. *Genes Dev.* **4**, 63-74.

Nasrin, N., Buggs, C., Kong, X. F., Carnazza, J., Goebl, M. and Alexander-Bridges, M. (1991). DNA-binding properties of the product of the testis-determining gene and a related protein. *Nature* **354**, 317-320.

Nusbaum, C. and Meyer, B. J. (1989). The *Caenorhabditis elegans* gene *sdc-2* controls sex determination and dosage compensation in XX animals. *Genetics* **122**, 579-593.

O, W.-S., Short, R. V., Renfree, M. B. and Shaw, G. (1988). Primary genetic control of somatic sexual differentiation in a mammal. *Nature* **331**, 716-717.

Ohno, S., ed. (1979). *Major sex determining genes. Monographs Endocrinology*, vol. II. Berlin: Springer.

Oosterwegel, M. A., Holstege, F. C. P., Prosser, H. M., Owen, M. J. and Clevers, H. C. (1991). TCF-1 binds to the *CD3-ε*, *TCR-α*, *TCR-β* and *TCR-δ* enhancers. *Eur. J. Immunol.* (in press)

Page, D. C. (1986). Sex reversal: deletion mapping the male-determining function of the human Y chromosome. In *Molecular Biology of Homo Sapiens. Cold Spring*

Harbor Symposium on Quantitative Biology, vol. 51, pp. 229-235. New York: Cold Spring Harbor.

Page, D. C., Mosher, R., Simpson, E. M., Fisher, E. M. C., 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**, 1091-1104.

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**, 937-939.

Palmer, S. and Burgoyne, P. S. (1991). XY follicle cells in the ovaries of XO/XY and XO/XY/XYY mosaic mice. *Development* **111**, 1017-1020.

Palmer, S. J. and Burgoyne, P. S. (1991a). In situ analysis of fetal, prepuberal and adult XX-XY chimaeric mouse testes: Sertoli cells are predominantly, but not exclusively, XY. *Development* **112**, 265-268.

Palmer, S. J. and Burgoyne, P. S. (1991b). The *Mus musculus domesticus Tdy* allele acts later than the *Mus musculus musculus Tdy* allele: A basis for XY sex reversal in C57Bl/6-Ypos mice. *Development* **113**, 709-714.

Parisi, M. A. and Clayton, D. A. (1991). Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science* **252**, 965-969.

Parkhurst, S. M., Bopp, D. and Ish-Horowicz, D. (1990). X:A ratio, the primary sex-determining signal in *Drosophila*, is transduced by helix-loop-helix proteins. *Cell* **63**, 1179-1191.

Patek, C. E., Kerr, J. B., Gosden, R. G., Jones, K. W., Hardy, K., Muggleton-Harris, A. L., Handyside, A. H., Whittingham, D. G. and Hooper, M. L. (1991). Sex chimaerism, fertility and sex determination in the mouse. *Development* **113**, 311-325.

Pearson, P. L. and Bobrow, M. (1970). Definitive evidence for the short arm of the Y chromosome associating with the X during meiosis in the human male. *Nature* **226**, 959-961.

Pelletier, J., Bruening, W., Li, F. P., Haber, D. A., Glaser, T. and Housman, D. E. (1991a). *WT1* mutations contribute to abnormal genital system development and hereditary Wilms' tumour. *Nature* **353**, 431-433.

Pelletier, J., Schalling, M., Buckler, A. J., Rogers, A., Haber, D. A. and Housman, D. (1991b). Expression of the Wilms' tumour gene *WT1* in the murine urogenital system. *Genes Dev.* **5**, 1345-1356.

Pelliniemi, L. J. (1975). Ultrastructure of the early ovary and testis in pig embryos. *Am. J. Anat* **144**, 89-112.

Pelliniemi, L. J. (1976). Ultrastructure of the indifferent gonad in male and female pig embryos. *Tissue and Cell* **8**, 163-175.

Polizzi, C. and Clarke, L. (1991). The chromatin structure of centromeres from the fission yeast:Differentiation of the central core that correlates with function. *J. Cell. Biol.* **112**, 191-201.

Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteus, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., van Heyningen, V. and Hastie, N. (1990). The candidate Wilms' tumour gene is involved in genitourinary development. *Nature* **346**, 194-197.

Ptashne, M. (1988). How eukaryotic transcription factors work. *Nature* **335**, 683-689.

Rattner, J. B. (1991). The structure of the mammalian chromosome. *Bioessays* **13**, 51-56.

Reichsteiner, M., ed. (1988). *Ubiquitin*. New York: Plenum.

Richards, O. W. and Davis, R. G., eds. (1977). *Imm's general textbook of entomology, classification and biology*, 10th edn, vol. II. London: Chapman and Hall.

Roberts, C., Weith, A., Passage, E., Michot, J. L., Mattei, M. G. and Bishop, C. E. (1988). Molecular and cytogenetic evidence for the location of *Tdy* and *Hya* on the mouse Y chromosome short arm. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6446-6449.

Robertson, E., Bradley, A., Kuehn, M. and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* **323**, 445-448.

Robertson, E. J. and Bradley, A. (1986). Production of permanent cell lines from early embryos and their use in studying developmental problems. In *Experimental Approaches to Mammalian Embryonic Development* (ed. J. Rossant and R. A. Pedersen), pp. 475-508. Cambridge: Cambridge University Press.

Rouyer, F., Simmler, M.-C., Johnson, C., Vergnaud, G., Cooke, H. J. and Weissenbach, J. (1986). A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. *Nature* **319**, 291-295.

Rouyer, F., Simmler, M.-C., Page, D. C. and Weissenbach, J. (1987). A sex chromosome rearrangement in a human XX male caused by Alu-Alu recombination. *Cell* **51**, 417-425.

Russell, L. B. (1976). Numerical sex chromosome anomalies in mammals: their spontaneous occurrence and use in mutagenesis studies. In *Chemical mutagens. Principles and methods for their detection*, vol. 4 (ed. A. Hollaender), pp. 55-91. New York, London: Plenum Press.

Russnak, R. H. and Candido, E. P. M. (1985). Locus encoding a family of small heat shock genes in *Caenorhabditis elegans*: two genes duplicated to form a 3.8kb inverted repeat. *Mol. Cell. Biol.* **5**, 1268-1278.

Salz, H. K., Maine, E. M., Keyes, L. N., Samuels, M. E., Cline, T. W. and Schedl, P. (1989). The *Drosophila* female-specific sex-determination gene, *Sex-lethal*, has stage- tissue- and sex-specific RNAs suggesting multiple modes of regulation. *Genes Dev.* **3**, 708-719.

Sanger, F. S., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-.

Satoh, M. (1985). The histogenesis of the gonad in rat embryos. *J. Anat.* **143**, 17-37.

Scherer, G., Schempp, W., Baccichetti, C., Lenzini, E., Bricarelli, F. D., Carbone, L. D. L. and Wolf, U. (1989). Duplication of an Xp segment that includes the *ZFX* locus causes sex inversion in man. *Hum. Genet.* **81**, 291-294.

Schneider-Gaedicke, A., Beer-Romero, P., Brown, L. G., Mardon, G., Luoh, S.-W. and Page, D. C. (1989a). Putative transcription activator with alternative isoforms encoded by human *ZFX* gene. *Nature* **342**, 708-711.

Schneider-Gädicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. and Page, D. C. (1989b). The *ZFX* gene on the human X chromosome escapes X-inactivation and is closely related to *ZFY*, the putative sex determinant on the Y chromosome. *Cell* **57**, 1247-1258.

Schwartz, P. E. and Doane, W. W. (1989). Molecular cloning of α-Amylase genes from *Drosophila Melanogaster*.III.An inversion at the *Amy* locus in an Amylase-null strain. *Biochem. Genet.* **27**, 31-46.

Sharman, G. B., Robinson, E. S., Walton, S. M. and Berger, P. J. (1970). Sex chromosomes and reproductive anatomy of some intersexual marsupials. *J. Reprod. Fert.* **21**, 57-68.

Sharman, G. B., Hughes, R. L. and Cooper, D. W. (1990). The chromosomal basis of sex differentiation in marsupials. In *Mammals from pouches and eggs: genetics beeding and evolution of marsupials and monotremes* (ed. J. A. M. Graves, R. M. Hope and D. W. Cooper), pp. 309-324. Melbourne: CSIRO.

Simon, M. and Herskowitz, I., eds. (1985). *Genome rearrangements*. New York: Alan R. Liss.

Simpson, E., Edwards, P., Wachtel, S., McLaren, A. and Chandler, P. (1981). H-Y antigen in Sxr mice detected by H-2 restricted cytotoxic T cells. *Immunogenetics* **13**, 355-358.

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 the human Y chromosome. *Nature* **326**, 876-878.

Simpson, E. M. and Page, D. C. (1991). An interstitial deletion in mouse Y chromosomal DNA created a transcribed *Zfy* fusion gene. *Genomics* **11**, 609-620.

Sinclair, A. H., Foster, J. W., Spencer, J. A., Page, D. C., Palmer, M., Goodfellow, P. N. and Graves, J. A. M. (1988). Sequences homologous to *ZFY*, a candidate human sex-determining gene, are autosomal in marsupials. *Nature* **336**, 780-783.

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**, 240-244.

Singh, L. and Jones, K. W. (1982). Sex reversal in the mouse (*Mus musculus*) is caused by a recurrent nonreciprocal crossover involving the X and an aberrant Y chromosome. *Cell* **28**, 205-216.

Singh, L., Purdom, I. F. and Jones, K. W. (1980). Sex chromosome associated satellite DNA: evolution and conservation. *Chromosoma* **79**, 137-157.

Singh, L., Matsukuma, S. and Jones, W. (1987). Testis development in a mouse with 10% of XY cells. *Dev. Biol.* **122**, 287-290.

Smith, C. and Mackay, S. (1991). Morphological development and fate of the mouse mesonephros. *J. Anat.* **174**, 171-184.

Solari, A. J. (1980). Synaptonemal complexes and associated structures in microspread human spermatocytes. *Chromosoma* **81**.

Staben, C. and Yanofsky, C. (1987). *Neurospora crassa* a mating type region. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4917-4921.

Stocking, C., Löliger, C., Kawai, M., Suciu, S., Gough, N. and Ostertag, W. (1988). Identification of genes involved in growth factor autonomy of hematopoietic cells by analysis of factor-independent mutants. *Cell* **53**, 869-879.

Stoye, J. P. and Coffin, J. M. (1985). Endogenous retroviruses. In *Molecular biology of tumour viruses, PartIII. RNA tumour viruses*, 2nd edn, vol. 2 (ed. R. Weiss, N. Teich, H. E. Varmus and J. M. Coffin), pp. 357-404. New York: Cold Spring Harbor Laboratory.

Stoye, J. P., Fenner, S., Greenoak, G. E., Moran, C. and Coffin, J. M. (1988). Role of endogenous retroviruses as mutagens: The Hairless mutation of mice. *Cell* **54**, 383-391.

Strathern, J., Hicks, J. and Herskowitz, I. (1981). Control of cell type in yeast by the Mating Type locus. *J. Mol. Biol.* **147**, 357-372.

Sutcliffe, M. J., Darling, M. S. and Burgoyne, P. S. (1991). Spermatogenesis in XY, $XYSxr^a$ and $XOSxr^a$ mice: A quantitative analysis of spermatogenesis throughout puberty. *Mol. Reproduct. Devel.* **30**, 81-89.

Tajima, Y., Onoue, H., Kitamura, Y. and Nishimune, Y. (1991). Biologically active kit ligand growth factor is produced by mouse Sertoli cells and is defective in Sl^d mutant mice. *Development* **113**, 1031-1035.

Taketo, T. and Merchant-Larios, H. (1986). Gonadal sex reversal of fetal mouse ovaries following transplatation into adult mice. In *Progress in developmental biology, part A* (ed. H. C. Slavkin), pp. 171-174. New York: Alan R. Liss.

Taketo, T., Koide, S. S. and Merchant-Larios, H. (1985). Gonadal sex differentiation in mammals. In *The origin and evolution of sex* (ed. H. O. Halvorson and A. Monroy), pp. 271-287. New York: Alan R. Liss.

Taketo-Hosotani, T. and Merchant-Larios, H. (1985). Testicular cell differentiation in fetal mouse ovaries following transplantation into adult male mice. *J. exp. Zool* **236**, 229-237.

Torres, M. and Sanchez, L. (1989). The *scute (T4)* gene acts as a numerator element of the X:A signal that determines the state of activity of *Sex-lethal* in *Drosophila*. *EMBO J.* **8**, 3079-3086.

Torrey, T. W. (1945). Development of the urogenital system of albino rat gonads. *Am. J. Anat* **76**, 375-397.

Tran, D., Meusy-Dessole, N. and Josso, N. (1977). Anti-Müllerian hormone is a functional marker of foetal Sertoli cells. *Nature* **269**, 411-412.

Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991). LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function. *Genes Dev.* **5**, 880-894.

Tsunoda, Y., Tokunaga, T. and Sugie, T. (1985). Altered sex-ratio of live young after transfer of fast-developing and slow-developing mouse embryos. *Gamete Res.* **12**, 301-304.

Turner, H. H. (1938). A syndrome of infantilism, congenital webbed neck and cutibus vulgus. *Endocrinol.* **23**, 566-574.

Vergnaud, G., Page, D. C., Simmler, M. C., Brown, L., Rouyer, F., Noel, B., Botstein, D., de la Chapelle, A. and Weissenbach, J. (1986). A deletion map of the human Y chromosome based on DNA hybridisation. *Am. J. Hum. Genet.* **38**, 109-124.

Vigier, B., Picard, J.-Y., Tran, D., Legeai, L. and Josso, N. (1984). Production of anti-Müllerian hormone: another homology between Sertoli and granulosa cells. *Endocrinol.* **114**, 1315-1320.

Vigier, B., Watrin, F., Magre, S., Tran, D. and Josso, N. (1987). Purified bovine AMH induces a characteristic freemartin effect in fetal rat prospective ovaries exposed to it in vitro. *Development* **100**, 43-55.

Villeneuve, A. M. and Meyer, B. J. (1987). *sdc-1*: A link between sex determination and dosage compensation in *C.elegans*. *Cell* **48**, 25-37.

Villeneuve, A. M. and Meyer, B. J. (1990). The regulatory hierachy controlling sex determination and dosage compensation in *Caenorhabditis elegans*. *J. Adv. Genet.* **27**, 117-188.

Wang, S. W., Robins, A. J., d'Andrea, R. and Wells, J. R. E. (1985). Inverted duplication of histone genes in chicken and disposition of regulatory sequences. *Nucleic Acids Res.* **13**, 1369-1387.

Washburn, L. L. and Eicher, E. M. (1983). Sex reversal in XY mice caused by dominant mutation on chromosome 17. *Nature* **303**, 338-340.

Washburn, L. L. and Eicher, E. M. (1989). Normal testis differentiation in the mouse depends on genetic interaction of a locus on chromosome 17 and the Y chromosome. *Genetics* **123**, 173-179.

Waterman, M. L., Fischer, W. H. and Jones, K. A. (1991). A thymus-specific member of the HMG protein family regulates the human T cell receptor Cα enhancer. *Genes Dev.* **5**, 656-669.

Weissenbach, J., Levilliers, J., Petit, C., Rouyer, F. and Simmler, M.-C. (1987). Normal and abnormal interchanges between the human X and Y chromosomes. *Development* **101**(Suppl.), 67-74.

Welshons, W. J. and Russell, L. B. (1959). The Y chromosome as the bearer of male determining factors in the mouse. *Proc. natn. Acad. Sci. U.S.A.* **45**, 560-566.

Wen, L., Huang, J. K., Johnson, B. H. and Reeck, G. R. (1989). A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1. *Nucleic Acids Res.* **17**, 1197-1214.

van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H. (1991). Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J.* **10**, 123-132.

Williams, D. E., Eisenman, J., Baird, A., Rauch, C., Van Ness, K., March, C. J., Park, L. S., Martin, U., Mochizuki, D. Y., Boswell, H. S., Burgess, G. S., Cosman, D. and Lyman, S. D. (1990). Identification of a ligand for the *c-kit* proto-oncogene. *Cell* **63**, 167-174.

Willison, W. and Ashworth, A. (1987). Mammalian spermatogenic gene expression. *Trends Genet.* **3**, 351-355.

Wilson, J. D. (1985). The endocrine control of sexual differentiation. *Harvey Lect.* **79**, 145-172.

Witschi, E. (1948). Migration of the germ cells of human embryos from the yolk sac to the primitive gonadal folds. *Carnegie Inst. Contrib. Embryol.* **32**, 67-80.

Wolfner, M. F. (1988). Sex-specific gene expression in somatic tissues of *Drosophila melanogaster*. *Trends Genet.* **4**, 333-337.

Wolgemuth, D. J., Viviano, C. M., Ginzang-Ginsberg, E., Frohman, M. A., Joyner, A. L. and Martin, G. R. (1987). Differential expression of the mouse homeobox-containing gene *Hox-1.4* during male germ cell differentiation and embryonic development. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5813-5817.

Wright, J. M. and Dixon, G. H. (1988). Induction by torsional stress of an altered DNA conformation 5' upstream of the gene for a high mobility group protein from trout and specific binding to flanking sequences by the gene product HMG-T. *Biochemistry* **27**, 576-581.

Wright, W. E., Sassoon, D. A. and Lin, V. K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to *MyoD*. *Cell* **56**, 607-617.

Zacksenhaus, E. and Sheinin, R. (1988). Identification of a human gene complementing tsA1S9 mouse L-cell defect in DNA replication following DNA-mediated gene transfer. *Somat. Cell Mol. Genet.* **14**, 371-379.

Zacksenhaus, E. and Sheinin, R. (1990). Molecular cloning of the human A1S9 locus: An X-linked gene essential for progression through S phase of the cell cycle. *Somat. Cell Mol. Genet.* **15**, 545-553.

Zamboni, L. and Upadhyay, S. (1983). Germ cell differentiation in mouse adrenal glands. *J. exp. Zool* **228**, 173-193.

Zsebo, K. M., Williams, D. A., Geissler, E. N., Broudy, V. C., Martin, F. H., Atkins, H. L., Hsu, R.-Y., Birkett, N. C., Okino, K. H., Murdock, D. C., Jacobsen, F. W., Langley, K. E., Smith, K. A., Takeishi, T., Cattanach, B. M., Galli, S. J. and Suggs, S. V. (1990a). Stem cell factor is encoded at the *Sl* locus of the mouse and is the ligand for the *c-kit* tyrosine kinase receptor. *Cell* **63**, 213-224.

Zsebo, K. M., Wypych, J., McNiece, I. K., Lu, H. S., Smith, K. A., Karkare, B., Sachdev, R. K., Yuschenkoff, V. N., Birkett, N. C., Williams, L. R., Satyagal, V. N., Tung, W., Bosselman, R. A., Mendiaz, E. A. and Langley, K. E. (1990b). Identification, purification and biological characterization of hematopoietic stem cell factor from buffalo rat liver-conditioned medium. *Cell* **63**, 195-201.