Genetic Analysis of the X Chromosomes of Man and Mouse

Thesis

UNRESTRICTED

Genetic Analysis of the X Chromosomes of Man and Mouse

Helen Jane Blair B.Sc. (Hons.)

A thesis submitted in partial fulfilment of the requirements of
the Open University for the degree of Master of Philosophy

Sponsoring establishment

MRC Mammalian Genetics Unit
(Formally the MRC Radiobiology Unit),

Harwell, Didcot,
Oxon. OX11 ORD

July 1996

Date of submission: July 1996

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

Acknowledgements

Most of all, I would like to thank my supervisor, Dr Yvonne Boyd, for her advice and encouragement throughout my M.Phil studies and thesis preparation.

I would also like to thank my second supervisor, Dr Bruce Cattanach, for helpful comments about the thesis. I thank all members of my laboratory at Harwell, past and present, for their help, including Viv Reed, Ify Uwechue, Steve Laval, Sheila Holt, Penny Williamson and, in particular, Wal Masson for proof reading.

I am very grateful to Kevin Glover for producing most of the prints in this thesis and Stuart Townsend for producing the FISH photographs. I am also grateful to Adrian Ford for help with 'Designer'.

I would like to thank M. Ho, I.W. Craig, S. Fisher and A.P. Monaco for providing probes and N. Quaderi, A.P. Monaco and D. Burke for providing YACs. I thank the HGMP-RC for the synthesis of primers and Mike Rhodes and the EUCIB group for supplying backcross DNAs.

Lastly, I would like to thank my parents, Sue, Craig and all my friends for their encouragement to complete these studies.

Table of Contents

Chapter 1 - Introduction

Chapter 2 - Materials and Methods

Chapter 3 - The Construction of Genetic Maps of Regions of the Mouse X Chromosome which Share Homology with Human Proximal Xp

Chapter 4 - Physical Characterisation of the Regions on the Proximal Mouse X Chromosome Homologous to Proximal Xp in Man

Chapter 5 - Discussion

List of Figures

List of Tables

Abbreviations

bp	Base pair(s)
DNA	Deoxyribonucleic acid
d(A, T, C, G, N)TP	Deoxy(adenosine, thymidine, cytidine, guanosine, nucleoside) triphosphate.
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetracetic acid
(m, μ, n)g	(milli, micro, nano) gram
kb	Kilo base pair(s)
(m, μ)l	(milli, micro)litre
min	Minute(s)
OD	Optical density
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
RPM	Revolutions per minute
sec	Second(s)
SDS	Sodium dodecyl sulphate
SDW	Sterile distilled water
Tris	Tris-hydroxymethane methylamine
X-gal	5-bromo, 4-choro, 3-indoyl, β-D-galactopyranoside
YAC	Yeast artifical chromosome
uv	Ultra-violet

Abstract

A detailed man-mouse comparative map aids the identification of mouse models for human inherited diseases and provides insight into the chromosomal rearrangements which have occurred during evolution. The man-mouse comparative map of the X chromosome is known to comprise at least eight conserved segments. Four of these lie in the human proximal short arm (Xp) in the CYBB-centromere region. Three share homology with the proximal, and one with the distal, region of the mouse X chromosome.

To improve the man-mouse comparative map of proximal Xp detailed recombination mapping has been performed. Panels of mice with recombination events in regions of the mouse X chromosome which share homology with human proximal Xp have been identified from laboratory *M. musculus* x *M. spretus* backcrosses. A high resolution map of the *DXWas70* to *Cybb* region of the mouse X chromosome was created by analysis of additional recombination from the European Collaborative Interspecific Backcross. The order of loci and the position of evolutionary breakpoints (EBs) in human proximal Xp was established as: DMD - EB - XK - PFC - EB - GATA1 - NPHL - EB - DXS674/DXS679 - ALAS2 - EB - DXF34.

Detailed physical analysis of the *DXWas70* to *Cybb* interval of the mouse X chromosome was carried out using YAC clones. YACs were analysed using PCR, PFGE and Southern and *in situ* hybridisation and as a result of these analyses, three YAC contigs were constructed. These include twelve YACs and may contains sequences at 2-3 evolutionary breakpoints. They may also contain the genes involved in the mouse mutant phenotypes tattered and scurfy.

Chapter 1

Introduction.

1.1 Human Gene Mapping.

Mapping of genes in the human genome began over 50 years ago in the hope that the knowledge gained would provide information useful for the therapy and prevention of disease. Genes responsible for disease phenotypes were originally mapped using classical Mendelian techniques. However, this approach was limited by the paucity of polymorphic markers. Possibilities for gene mapping have since been expanded by the ability to clone gene sequences and the identification of restriction fragment length polymorphisms (RFLPs) by Southern blot analysis. Also, the high level of allelic variation found in repeat sequences, both variable number tandem repeats (VNTRs, minisatellites) and simple sequence tandem repeats (microsatellites), coupled with the development of polymerase chain reaction (PCR) technology, has greatly increased the number of markers on the human genetic map (Gyapay *et al.*, 1994; Hearne *et al.*, 1992). In addition, many genes have been assigned physically to human chromosomes and specific chromosome bands using interspecies somatic cell hybrids and *in situ* hybridisation (Ruddle and Creagan, 1975; Buckle and Craig, 1986). Physical maps of human genomic regions have been constructed using pulsed field gel electrophoretic analysis (PFGE) of large DNA fragments defined by restriction endonuclease cut sites (Burmeister and Lehrach, 1986). A large proportion of the human genome has now been cloned, for example in yeast artificial chromosome and cosmid vectors, and much work is now concentrated on the completion of a map of overlapping clones (the contig map) (Olson *et al.*, 1989; Schlessinger, 1990). Of particular interest is a

complete cDNA (transcript) map of the human genome since it is the expressed sequences that are most likely to be involved in human genetic disease and biological processes.

1.2 The Mouse as a Model

Once identified, genes involved in disease phenotypes can be studied to understand disease etiology. There are however technical and ethical difficulties in the identification and the study of many genes in the human system. To fully gain the benefit of molecular technologies for the study and identification of genes and their involvement in disease, it is necessary to use a model organism. The mouse provides a very good model system, as it is a mammal with a genetic composition very similar to that of humans and has similar physiological and metabolic processes. It is small in size, resistant to infection, has a large litter size and a relatively rapid generation time, all of which mean it is relatively cheap and easy to maintain and study.

1.3 Laboratory Mice.

The majority of mouse strains used in the laboratory are derived from the wild mouse species *Mus musculus* and the analysis of mitochondrial DNA indicates that these mice are from the sub species *Mus musculus domesticus* (Ferris *et al.*, 1983) found in west and southern Europe (Bonhomme *et al.*, 1984). There is, however, evidence that some strains have a Y chromosome from *Mus musculus musculus* (Bishop *et al.,* 1985), a sub species found in central and eastern Europe, Russia and China. Most laboratory stocks are inbred, i.e. they have been maintained for more than 20 generations by brother-sister matings, and are therefore homozygous at all loci with the exception of

any new spontaneous mutations. There are over 200 inbred strains used in laboratories all over the world (Festing, 1996).

1.4 Mouse Mutants.

Genetically the mouse is one of the best studied of all mammalian species. The first gene linkage between two mouse mutant phenotypes, albino and pink eye dilution, was reported in the early part of this century (Haldane *et al.*, 1915). There are now a wealth of mutations available in the mouse for the study of many biological phenomena, for example, tissue interactions can be studied in agouti mice and cell migration in the mutant extreme white spotting (Hogan *et al.*, 1986). Many mouse mutants, in particular those with visible differences in coat colour, hair morphology, pigmentation and behavioural abnormalities were first discovered by mouse breeders. Some mutants resulting from recessive gene mutations have been identified during successive inbreeding of wild mouse strains in the laboratory. Exposure of mice to X-ray irradiation or chemical mutagens has also given rise to various mouse mutants (Russell *et al.,* 1979; Rinchik, 1991). Many of the loci implicated in the mutant phenotypes have been localised to specific chromosomes or chromosomal regions by linkage to existing mutant phenotypes or to biochemical markers (Nielsen and Chapman, 1977; Doolittle *et al.*, 1990). Targeted mutagenesis, the introduction of a mutant gene in place of the normal copy using homologous recombination, can also be used to create mice with mutations at a specific locus (Capecchi, 1989; Frohman and Martin, 1989). This technique, for example, has been successfully used to create a mouse model for cystic fibrosis (Snouweart *et al.*, 1992; Dorin *et al.*, 1992).

1.5 Mouse Mutants as Models for Human Inherited Diseases.

Mouse mutants can be used to provide information about human inherited diseases in a number of ways: (1) for the identification of the genes responsible; (2) for investigation into the nature of mutations; (3) for studying effects in development and (4) for investigations into possible therapy.

To provide a good model for a human genetic disease, a mouse mutant requires a comparable phenotype and a defect in an homologous locus. A mouse mutant which possesses one of these characteristics could prove useful for a particular aspect of disease study but the most relevant models would possess both characteristics. Gene homologies between man and mouse can be confirmed by satisfying one of a number of standards outlined by Davisson *et al.* (1991). These include similarities in the molecular structure and biological function of human and mouse genes; however, these may be difficult to determine if the gene responsible for the human disease or mouse mutation has not been isolated. The establishment of a homologous map position can help to confirm if the defects are occurring in the same gene. To allow homologous genes and sequences to be identified, a comparative map of the human and mouse genomes is required.

1.6 Man - Mouse Comparative Mapping.

An estimated 70 million years have elapsed since human and mouse divergered from a common ancestor (Nadeau and Taylor, 1984). During this time changes in the genomic organisation of each species will have occurred. Cytogenetically, considerable rearrangement must have occurred since the human karyotype has 22 pairs of

autosomes of differing morphology (metacentric, submetacentric and acrocentric chromosomes) whereas the mouse karyotype has 19 pairs of acrocentric autosomes. Changes in sequences which have a biological function, such as those which are expressed in a gene, are likely to have been selected against during evolution; therefore the position of these homologous sequences (or loci) can be compared in different species and rearrangements can be identified. When the gene maps of human and mouse are compared there are regions of conserved synteny where two or more homologous loci have been mapped to the same chromosome in each species. Regions of conserved synteny where the gene order of homologous loci has remained the same are known as conserved linkages[1] (Nadeau, 1989). The Oxford grid shows the distribution of homologous loci on the autosomes of man and mouse (Edwards, 1991). The most recent version of the Oxford grid (courtesy of J. Peters) is given in Figure 1.1. Some regions contain a large number of loci showing conserved linkage, as on human chromosome 1 and mouse chromosome 4 which, at present, share 53 loci. This region is estimated to cover approximately 40cM on mouse chromosome 4 (Abbott *et al.*, 1993; Dracopoli *et al.*, 1994) and using high resolution G-banding, it appears to have maintained a similar chromosome banding pattern in the human and mouse karyotype (Sawyer and Hozier, 1986). Nadeau and collaborators have estimated that there are approximately 100 conserved segments on the autosomes of human and mouse with a mean length of 8.8cM and that there have been 144 rearrangements disrupting linkage since the divergence of the two species (unpublished data; 9th International Mouse Genome Conference, Ann Arbor, USA, 1995).

[1] Conserved linkages are referred to as conserved segments thoughout this thesis.

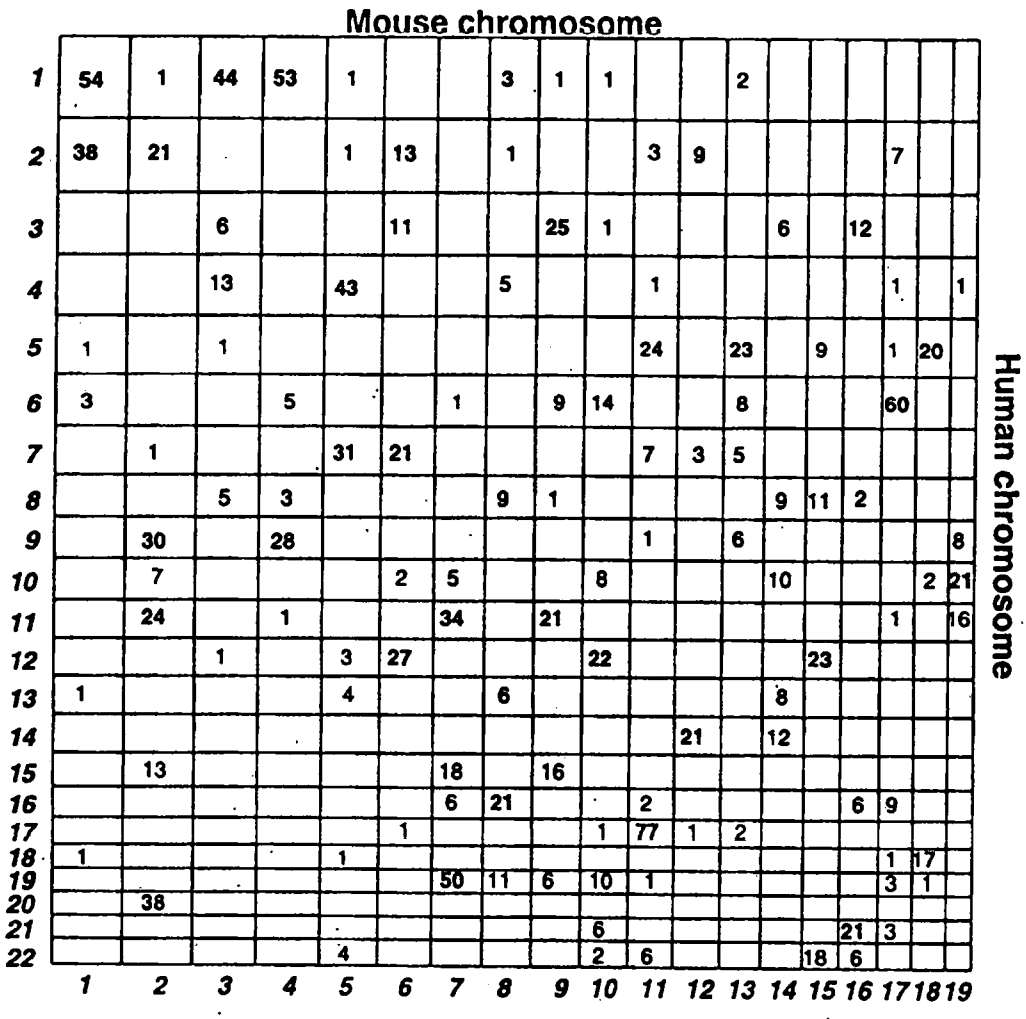


Figure 1.1

The Oxford Grid: An array which displays the number and distribution of homologous loci on the autosomes of man and mouse. (courtesy of J. Peters, MRC Radiobiology Unit).

Chromosomal rearrangement can occur in several different ways such as transposition, translocation and inversion. Study of the comparative maps of man and mouse and the identification of evolutionary breakpoints, alongside information from other mammalian species, will allow insight into genome organisation and the mechanisms of chromosomal rearrangement.

1.7 Linkage Mapping in the Mouse.

Somatic cell genetics for genome mapping is less well developed in the mouse than in man and there are very few somatic cell hybrids which have a retained single mouse chromosome. *In situ* hybridisation has also been less frequently employed for mouse mapping studies as mouse chromosomes are similar in size and morphology, making them difficult to distinguish from each other. Mouse crosses, however, can allow mapping of genes and mouse mutant phenotypes to specific chromosomal regions at high resolution. The most frequently employed have been interspecific backcrosses between *Mus musculus* laboratory strains and *Mus spretus*, a wild derived mouse species from Spain and North Africa. These are sympatric species in Spain (Britton and Thaler, 1978); however in the laboratory they will breed, with *M. musculus* female and *M. spretus* male crosses being far more fruitful than the reciprocal cross. F1 males from *M. musculus* x *M. spretus* crosses are sterile but the females can be backcrossed to either *M. musculus* or *M. spretus* males to produce backcross offspring (Figure 1.2). Using these backcross offspring, genes and markers are positioned with respect to recombination events that occur at meiosis in F1 females (Robert *et al.*, 1985), by taking advantage of the large amount of allelic variation between the two species. For almost any marker a restriction fragment length variation

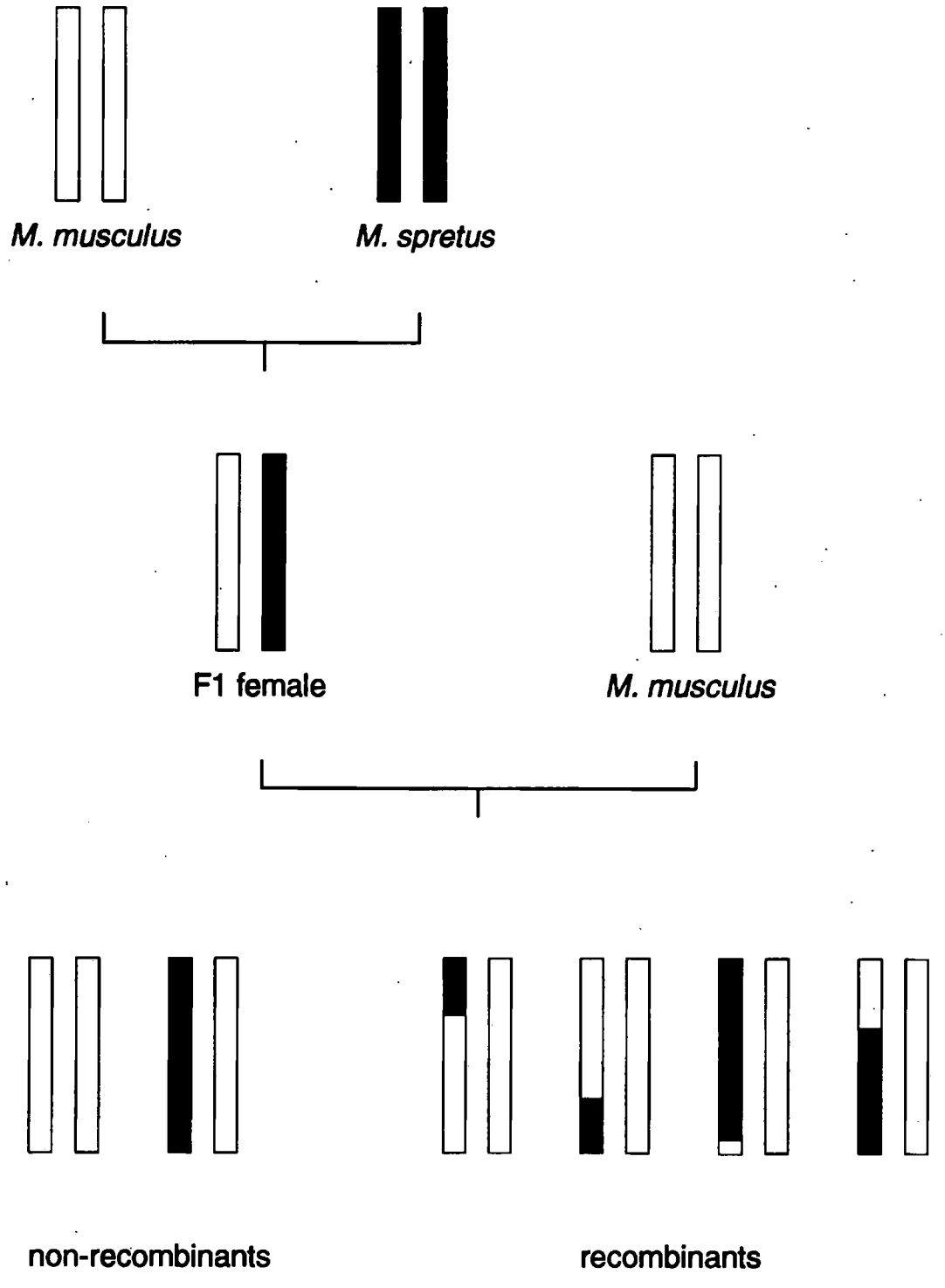


Figure 1.2
Schematic representation of a *M. musculus* x *M. spretus* interspecific backcross. In this example case, F1 hybrid females are backcrossed to *M. musculus* males but they can also be backcrossed to *M. spretus* males. The rectangles represent one chromosome of an homologous pair. Open rectangles are *M. musculus* and black rectangles are *M. spretus* type chromosomes. Note: recombination at female meiosis results in backcross offspring with recombinant chromosomes (part *M. musculus* and part *M. spretus* type).

(RFLV), using Southern blotting or an amplification product variant (APV), using the polymerase chain reaction (PCR), can be found between the two species. This allows the position of recombination events in a panel of backcross offspring to be identified. An order and a genetic distance (in cM) between typed markers can then be calculated. Offspring with recombination events in a particular genomic region can be identified, these can then be used as a subset of the panel to position markers within that region. This is known as pedigree analysis and allows fine positioning of markers without testing a large number of animals (Avner *et al.*, 1988).

One of the largest backcrosses available, the European Collaborative interspecific backcross (EUCIB) consists of almost 1000 offspring and has been used for some of the studies presented in Chapter 3 of this thesis. Offspring are derived from F1 interspecific females backcrossed to *M. musculus* (C57BL/6) or from backcrosses to *M. spretus* male mice. The large number of offspring from this backcross should allow positioning of markers to 0.1cM resolution (Breen *et al.*, 1994). All the information from the scoring of the EUCIB animals is stored on a database called MBx which is available on the World Wide Web (WWW). The production of high resolution maps has also been aided by the large number of recently developed microsatellite markers (Dietrich *et al.*, 1996). These are highly variant between mouse strains and species and can be analysed easily by PCR. Microsatellites are not usually conserved so provide little information for comparative mapping studies. However, they provide useful tools for the identification of backcross offspring for pedigree analysis and for the isolation of DNA clones for physical mapping experiments.

1.8 The Man - Mouse X Chromosome Comparative Map.

The X chromosome is unique in that X-linked genes in man are usually located on the X chromosome in the mouse (and the X chromosomes of other eutherian mammals). This phenomenon is known as Ohno's law (Ohno, 1969; 1973) and can be explained by the evolutionary constraints brought about by X inactivation in females. Inactivation of one of the X chromosomes is thought to occur in females in order to maintain equivalent dosage of X-linked genes in both sexes and rearrangements involving the X chromosome and autosomes would be unfavourable. On the X chromosomes of man and mouse although there are conserved blocks within which the order of loci has been retained, the position of these blocks, with respect to each other, has changed.

In 1988, Amar *et al.* and in 1989, Searle *et al.* reported that X-linked genes could be divided into 5 conserved segments shared between man and mouse. Two of the conserved segments were located on the human X chromosome long arm (Xq) and three on the short arm (Xp) (Figure 1.3). The difference between the position of these segments in man and mouse could be explained by two inversion events occuring during the evolutionary time which separates the two species (Lyon, 1988). However, the extent of the segments and the exact position of evolutionary breakpoints was not well defined as only a small number of loci were mapped in both species. Many more loci have now been positioned on both the X chromosome maps of man and mouse (Nelson *et al.*, 1995; Herman *et al.*, 1996); these include genes positioned using molecular probes, or by biochemical analysis, as well as a number of cross-hybridising conserved sequences that are from unknown or uncharacterised genes (e.g. Laval and

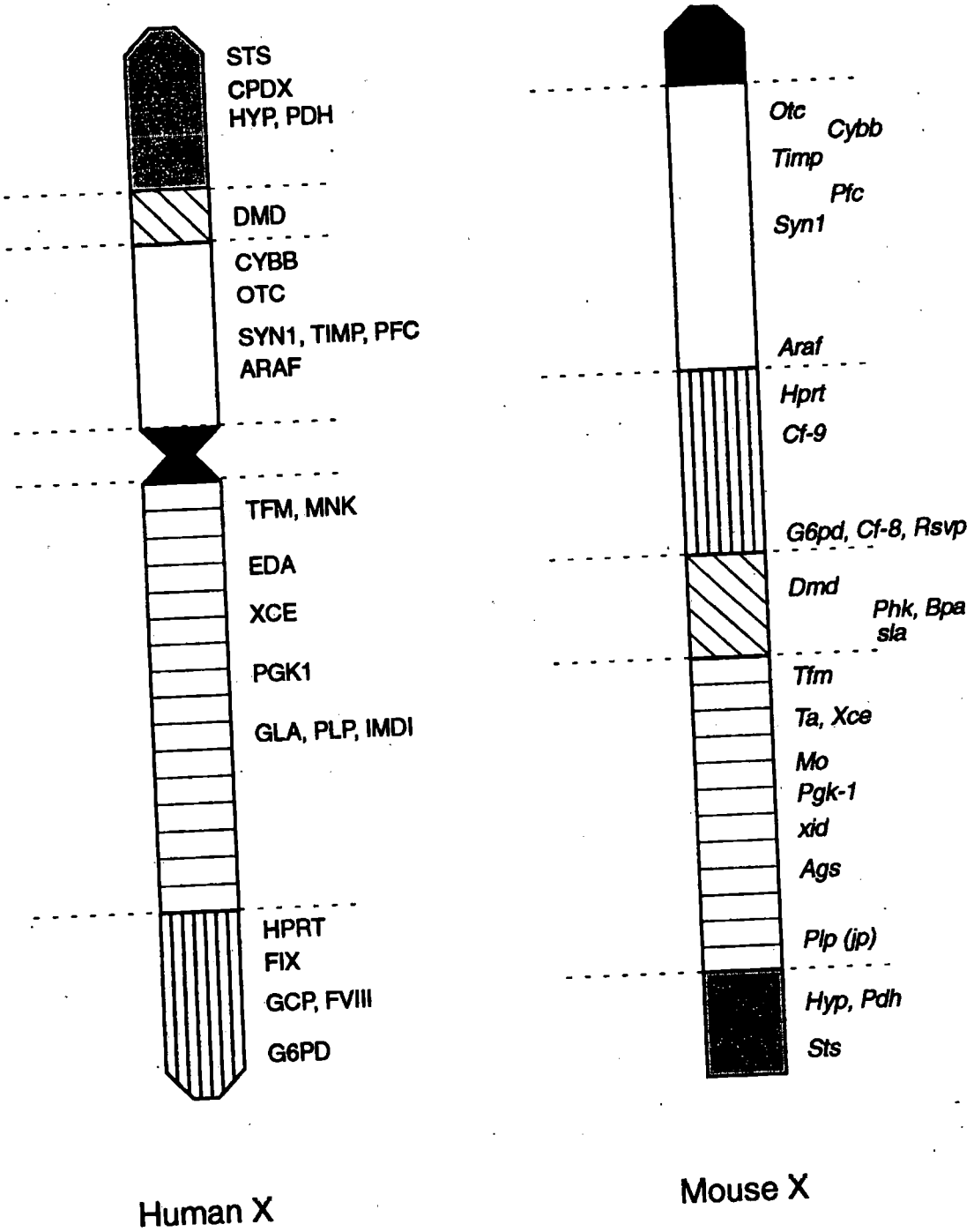


Figure 1.3
Human and mouse X chromosomes showing five conserved segments. Corresponding segments have been given the same shading pattern and homologous loci are shown adjacent to each chromosome.
(Adapted from Searle *et al.*, 1989)

Boyd, 1993a). It now appears that the man-mouse comparative map consists of eight segments (Laval and Boyd, 1993b). The region on the human X chromosome from CYBB[2] (the locus encoding cytochrome b-245 β peptide) to the centromere, which previously was thought to be one complete segment, has been shown to contain at least four conserved segments. The organisation of conserved segments on the human and mouse X chromosome can no longer be explained by two inversion events as previously hypothesised. The work in this thesis focuses on the man-mouse comparative map of human proximal Xp from CYBB to the centromere.

1.9 Conserved Segments Composing Human Proximal Xp[3].

In 1989, the loci on human proximal Xp from CYBB to the centromere, were believed to have remained in the same order but had been inverted with respect to the centromere on the proximal mouse X chromosome (Searle *et al.,* 1989; Figure 1.3). Further mapping studies have shown that this segment is composed of a minimum of four distinct conserved segments on the mouse X chromosome (Laval and Boyd, 1993a; Chapman *et al.*, 1994; Blair, H.J. *et al.*, 1994 and Laval and Boyd, 1993b). Each of these segments can be defined by either a single locus (segments 2 and 4) or the most proximal and distal loci they contain (segments 1 and 3) as follows: 1) **CYBB-PFC**; 2) **GATA1**; 3) **DXS674-ALAS2** and 4) **DXF34**[4]. Details of each, beginning with the DXF34 conserved segment and continuing to CYBB-PFC

[2] Loci expressed in upper case are from the human X chromosome map and those in lower case and italics are from the mouse X chromosome map in accordance with standard nomenclature.

[3] Details of loci referred to in the following sections are given in Figure 3.1

[4] The most proximal and distal loci which define the conserved segments are expressed in upper case throughout this thesis.

conserved segment are given below.

In 1993a, Laval and Boyd isolated the human genomic clones pH3-7 and pH3-68. The clone pH3-7 detected a family of sequences known as DXF34 in the Xp11.21 - centromere region of the human X chromosome, all of which were found to be in the same YAC clone (Reed *et al.*, 1994). pH3-7 also detected several fragments in digested mouse genomic DNA (Figure 1.4), all of which were assigned to the mouse X chromosome using a panel of somatic cell hybrids (Williamson *et al.*, 1995). One of these sequences, known as *DXHXF34* (previously known as *DXF34h*-rs1) had been assigned to the proximal region of the mouse X chromosome close to *Otc* (Laval and Boyd, 1993a). *DXHXF34* was later shown to cosegregate with *DXWas70,* a non-conserved repeat sequence locus (Disteche *et al.*, 1985), and to lie proximal to *Gata1* and *Cybb* on the mouse X chromosome (Laval and Boyd, 1993b). The position of this sequence indicated that there was a conserved segment near the centromere on the comparative map of the X chromosomes of man and mouse (Figure 1.5). At the beginning of this work it was not known whether all the sequences detected by pH3-7 on the mouse X chromosome lay in the same region.

The human conserved single-copy clone, pH3-68 detects the human locus DXS674 which lies in the Xp11.23-Xp11.21 region of the human X chromosome (Laval and Boyd, 1993a). Unexpectedly, the murine homologue of DXS674, *DXHXS674*, was found to lie in the distal region of the mouse X chromosome between *Plp* and *Pdha1* and to define a novel region of homology. In addition, the murine homologue of ALAS2, the gene which encodes erythroid δ-aminolevulinate synthase, was assigned

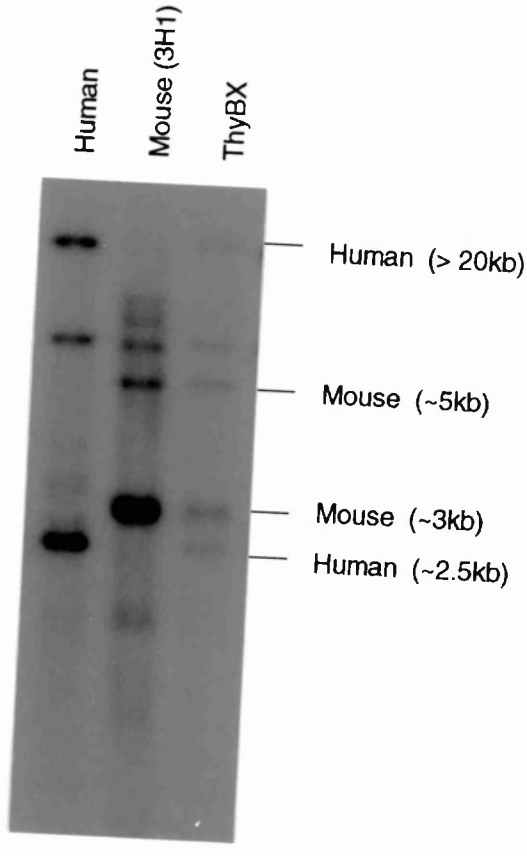


Figure 1.4

Hybridisation of pH3-7, which detects the locus DXF34, to DNA digested with TaqI. pH3-7 detects several fragments in both human and mouse DNA. ThyBX is a somatic cell hybrid which contains the human X chromosome on a mouse background (Goodfellow *et al.*, 1980). The approximate sizes of the stronger human and mouse TaqI fragments are shown. All of the strong human specific fragments are present in ThyBX DNA indicating that they are from the X chromosome.

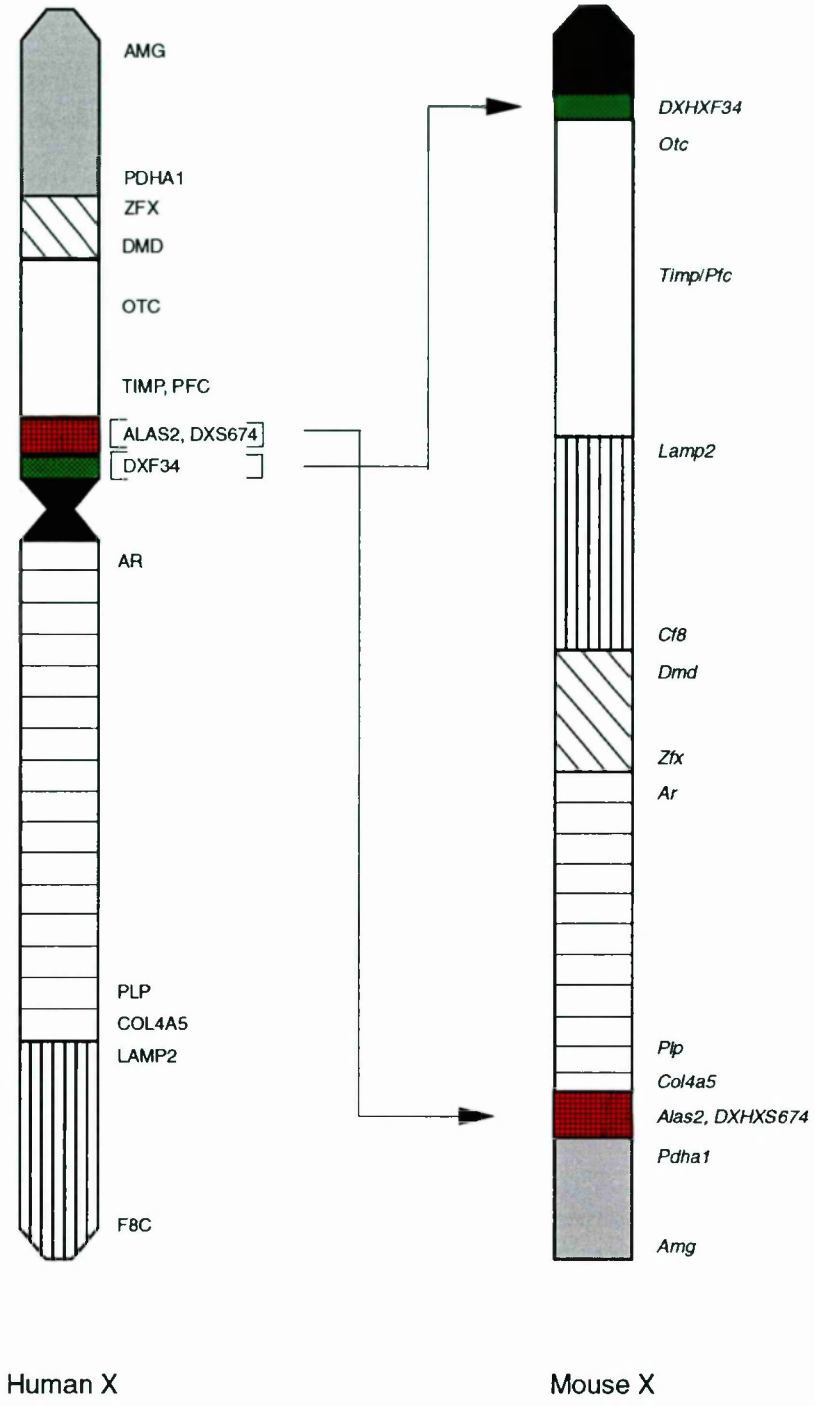


Figure 1.5

The comparative map of the human and mouse X chromosomes showing the five conserved segments as in Figure 1.3 and in addition the position of segments defined by DXF34 as described by Laval and Boyd, 1993a (shown in green) and ALAS2/DXS674 as decribed by Chapman, *et al.*, 1994 and Laval and Boyd, 1993a (shown in red). The pseudoautosomal regions are not shown.

to the distal region of the mouse X chromosome (Chapman *et al.*, 1994). In man, ALAS2 was found to lie within 400kb of DXS674 (Boyd *et al.*, 1994) which indicated that together they defined a conserved segment in the distal region of the mouse X chromosome (Figure 1.5). At the beginning of this work the size, orientation and precise position of this segment on the mouse X chromosome was not known.

Also in 1993(b), Laval and Boyd positioned the murine homologue of GATA1 between *DXHXF34* and *Cybb* in the proximal region of the mouse X chromosome. GATA1 lies between DXS674 and PFC in man (Nelson *et al.,* 1995) therefore it was expected from the comparative map that the murine homologue of GATA1 would lie close to *Pfc*. It was proposed that there had been an inversion during the evolution of man and mouse, within the segment from CYBB to GATA1, and that an evolutionary breakpoint lies between GATA1 and PFC in man (Figure 1.6). At the beginning of this work it was not known if any other loci which mapped close to GATA1 in man would also lie close to *Gata1* on the mouse X chromosome.

The CYBB-PFC conserved segment appears to span the Xp11.3-Xp21.1 region in man and a region of approximately 5cM on the proximal mouse X chromosome (Figure 1.6). Several other loci from this region in man have been positioned on the mouse X chromosome (e.g. MAOA; NDP, UBE1, DXHX676) and it appears that the gene order within this segment has been retained (Nelson *et al.*, 1995; Herman *et al.*, 1996).

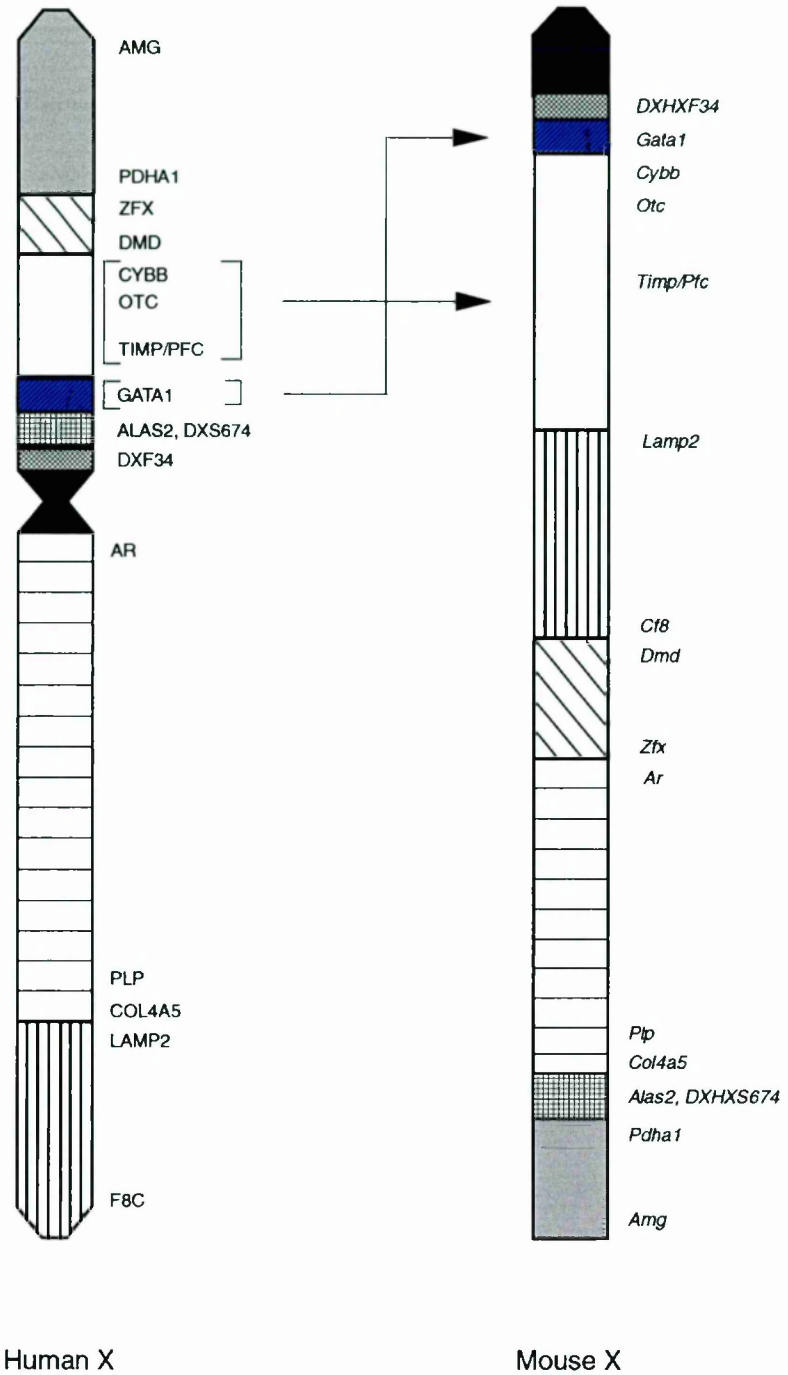


Figure 1.6

The comparative map of the human and mouse X chromosomes showing the seven conserved segments as in Figure 1.5 and in addition the position of the segment defined by GATA1 as described by Laval and Boyd, 1993b (shown in blue). The position of the CYBB-PFC segment is also indicated.

1.10 Prediction of Murine Map Position of Sequences from Proximal Xp using the Comparative Map.

The human X chromosome has been extensively studied and many X-linked genes and conserved sequences have been cloned and mapped. A large proportion of the chromosome has also been isolated in YAC clones which has enabled genes and other conserved markers to be ordered in detail on the human map within YAC contigs (Nelson *et al.*, 1995). It can be particularly difficult to predict the murine map position of disease genes and conserved sequences from the proximal Xp region of the human X chromosome due to the uncertainty about the extent and orientation of conserved segments in this region. The known extent of the conserved segments and possible positions of the evolutionary breakpoints in proximal Xp are shown in relation to the human physical map in Figure 1.7. TFE3 (the transcription factor enhancer 3) lies very close to GATA1 in man (Nelson *et al.*, 1995) and the murine homologue of TFE3, *Tcfe3,* has been positioned cytogenetically between the centromere and band XA3 on the mouse X chromosome (Roman, *et al.*, 1992). It could not be predicted from the present comparative map whether *Tfe3* lies close to *Gata1* on the mouse X chromosome or in a more distal location, close to *Pfc* (Figure 1.7).

1.11 Mouse Mutants as Possible Models for Inherited Diseases Associated with Proximal Xp.

There are several disease loci which map to proximal Xp in man. Of particular relevance for this thesis are two diseases which have potential mouse mutant homologues. The first, Wiskott-Aldrich syndrome (WAS) is characterised by severe

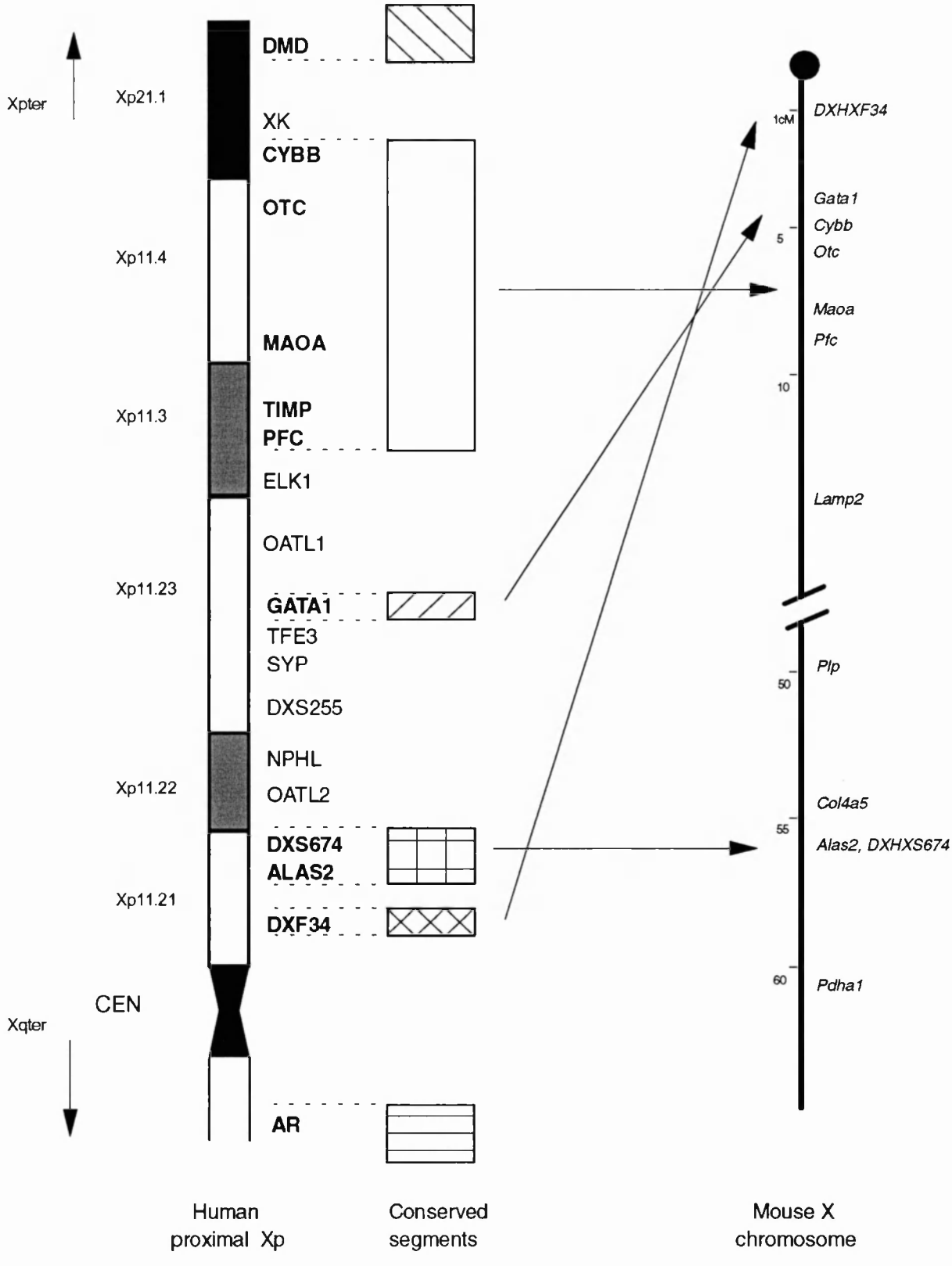


Figure 1.7

Map of the proximal region of the human X chromosome short arm showing the extent of conserved segments with the mouse X chromosome. The positions of loci on the human physical map are taken from Nelson *et al.*,1995 and the positions of loci on the mouse X chromosome genetic map are taken from Herman *et al.*, 1996.

immunodeficiency, thrombocytopenia, eczema and an increased risk of malignancies (Aldrich *et al.*, 1954). The gene responsible for WAS was localised by linkage analysis to a region flanked by the loci TIMP and DXS255 on Xp11.22-Xp11.3 (Kwan *et al.*, 1991); thus the position of the murine homologue of this gene would be difficult to predict (Figure 1.7). During this work, WASP, the gene mutated in Wiskott-Aldrich syndrome, was isolated and found to lie very close to GATA1 in man (Derry *et al.*, 1994; see Chapter 5). The mouse mutation scurfy (*sf*) has a similar phenotpye to WAS and has been proposed as its murine homologue (Lyon *et al.*, 1990). *sf* has been mapped close to *Gata1* and *Tfe3* on the mouse X chromosome (Blair, P.J. *et al.*, 1994), therefore, the gene responsible for this disorder could lie in the conserved segment identified by the *Gata1* locus. Analysis of the murine homologue of WASP in scurfy mice and further comparative mapping studies around GATA1 would help to confirm this proposed homology.

The other disorder, incontinentia pigmenti type 1 (IP1) is a rare form of X-linked dominant incontinentia pigmenti which is characterised by neuroectodermal abnormalities (McKusick, 1990). It has been associated with X chromosome translocations with breakpoints scattered in Xp11.2 - Xp11.1 (Gorski *et al.,* 1991). The mouse mutant tattered (*Td*) has been proposed as a model for IP1 (Davisson *et al.,* 1991) as it has skin lesions similar to IP1 and has been mapped close to *Otc* on the mouse X chromosome (Cattanach, 1982) a region which shares some homology with Xp11.2-Xp11.1. However, IP1 breakpoints also lie close to ALAS2 and DXS674 and as the murine homologues of these loci map to the distal end of the mouse X chromosome, the murine homologue of IP1 could lie in this region. The mouse mutant

lined (*Li*) has skin abnormalities and maps to the distal mouse X chromosome (Cattanach, 1985) so it too is a potential mouse model for IP1. Determination of the position of evolutionary breakpoints in Xp11.2-Xp11.1 with respect to the IP1 breakpoints and precise map positions for *Td* and *Li* would help to confirm which, if either, of these mouse mutants is the true homologue.

1.12 Aims of this Study.

The aim of this study is to improve the comparative map of the X chromosomes of man and mouse with emphasis on the definition of evolutionary breakpoints between the conserved segments which comprise proximal Xp in man. Two approaches were taken to achieve this aim:

1) To develop markers for linkage mapping in the mouse and create a detailed genetic map to provide further information about the extent and orientation of conserved segments which comprise proximal Xp in man (Chapter 3).

2) To initiate the development of a physical map of the proximal region of the mouse X chromosome using YAC clones in regions close to man/mouse evolutionary breakpoints (Chapter 4).

Chapter 2

Materials and Methods.

The solutions and media used in the methods described are listed in section 2.12.

2.1 Genomic DNA Preparation.

The spleen, liver and kidneys from adult mice were dissected and placed in liquid nitrogen. Generally, the liver and kidneys were stored at -70°C and DNA extracted from the spleen.

Spleens were ground to a fine powder using a pestle and mortar which was precooled on dry ice and filled with liquid nitrogen. The powder was added to 4ml of SET containing 1% SDS and 200μg/ml proteinase K (Boehringer Mannheim) and the samples incubated overnight at 60°C. The lysate was extracted twice with an equal volume of phenol, once with phenol/chloroform and once with chloroform. The DNA was precipitated with absolute ethanol and washed with 70% ethanol. The DNA pellet was resuspended in 1ml 1x TE containing 20μg/ml RNase A (Sigma) and incubated at 37°C. The purity and the yield were assessed by the absorbence at the wavelengths 235nm, 260nm and 280nm using a spectrophotometer (DU-64, Beckmann) (1OD unit = 50μg/ml at 260nm). Samples were diluted to 1μg/μl and aliquoted for storage at 4°C and -20°C.

2.2 Plasmid DNA Preparation.

Plasmid DNA was prepared by adaptations of the alkaline lysis methods of Birnboim

and Doly (1979), Ish-Horowicz and Burke (1981) and Sambrook *et al.* (1989).

Single *E.coli* colonies containing the plasmid of interest were used to inoculate 100ml of LB medium containing 100μg/ml antibiotic (usually ampicillin). Cultures were incubated overnight at 37°C with shaking (250 RPM). The cell suspension was centrifuged at 12000g for 10 min. The supernatant was decanted and the tube inverted on a paper towel to drain the remaining medium. 2.5ml of cold SET was added and the pelleted cells resuspended. The cells were spun again at 12000xg for 10 min and the supernatant removed. 1ml of ice cold GET solution containing 4mg/ml lysozyme (Sigma) was added, the bacteria resuspended and incubated at room temperature for 5 min. 2ml ice cold 1% SDS, 200mM NaOH was then added, mixed by inversion, and incubated on ice for 5 min. 1.5ml KAc was added, mixed by inversion and incubated on ice for 5-10 min. Samples were spun at 12000xg for 10 min, the supernatant being retained and extracted once with 4 ml of phenol/chloroform. RNA was removed by incubation at 37°C for 1-2 hours with 10μg/ml RNase A (DNase-free, Sigma). Samples were extracted once each with phenol, phenol/chloroform and chloroform. Two volumes of absolute ethanol were added to precipitate the DNA and the pellet washed with 70% ethanol and dried. The plasmid DNA was resuspended in 500μl of sterile 1x TE or SDW and the concentration assessed by comparison with known standards after agarose gel electrophoresis.

2.3 Plasmid Transformation of *E.coli* by Heat Shock.

Approximately 100ng of plasmid DNA was added to 200μl JM83 competent cells and incubated on ice for 30min. The cells were heat shocked at 42°C for 90sec and placed

back on ice for 5min. 800ml of LB medium was added to the cells and they were allowed to recover at 37°C for 1hour with shaking. Four 50µl aliquots of cells were spread on L-agar plates containing the appropriate supplements, usually (100µg/ml) ampicillin and X-gal(40µg/ml) and incubated overnight at 37°C.

2.4 Restriction Enzyme Digestion of Genomic DNA.

Restriction enzyme digestion was generally carried out as 20µl reactions with 10µg of genomic DNA, 2µl restriction enzyme (8-10 units/µl, Life Technologies, NBL Gene Sciences, NEB) and 2µl of the appropriate 10 x reaction buffer. Reactions were incubated overnight at the optimum temperature for the enzyme (usually 37°C, but 65°C for TaqI). 2µl of the reaction was removed for agarose gel electrophoresis to confirm complete digestion of the DNA. Digestion, if partial, could usually be completed by the addition of more restriction enzyme and further incubation at the appropriate temperature.

2.5 Submerged Horizontal Agarose Gel Electrophoresis.

Agarose (Sigma) was melted in 0.5 x TBE and allowed to cool before pouring into a gel casting tray with a well forming comb. The concentration of agarose was dependant on the size of the DNA to be fractionated; 0.8% agarose gels were generally used for digested genomic DNA, whilst 3% or greater gels were used for DNA fragments of 100-500bp in size e.g. most PCR products. Samples containing loading buffer were loaded onto the gel and fractionated under constant voltage in 0.5 x TBE. DNA markers were run adjacent to the samples. The DNA markers used were: λ Hind III fragments which range in size from approximately 550bp to 23kb; 1 kb

ladder which ranges in size from <500bp to 12kb and 123bp ladder which ranges in size from 123bp to >3kb each band increasing by 123bp (Life Technologies). Gels were stained in EtBr (0.5µg/ml) for 20-30 min and photographed on a uv transilluminator (high power) using a Land camera (Polaroid MP-4) for photographs requiring a negative or a video copy processor (Model K65HM, Mitsubishi) for direct prints on thermal paper.

2.6 Isolation of DNA Fragments from Agarose Gels.

DNA fragments were isolated from agarose in two ways depending upon their size:

2.6.1 Geneclean II Kit for fragments >500bp.

The slice of agarose gel containing the DNA fragment to be purified was excised under low power uv light and placed in a 1.5ml Eppendorf tube. 0.5 x volume TBE modifier and 4.5 x volume NaI solution were added from the Geneclean II kit (Bio 101 Inc; Statech Scientific). The gel slice was incubated at 55°C for 5 min to melt the agarose. 10-15µl of Glassmilk was added to the tube, the suspension mixed and placed at room temperature for 5-10 min. The tube was spun for 5 sec in a microcentrifuge and all the supernatant removed. The Glassmilk was washed with three changes of cold New Wash solution. 10-15µl of SDW was mixed with the Glassmilk and incubated at 55°C for 5min. The tube was spun, the supernatant collected and the Glassmilk reincubated with SDW to collect further supernatant.

2.6.2 DEAE Membrane for fragments >100bp.

The fragment of DNA was fractionated on an 1% agarose gel with 1 x TBE running

buffer and stained with EtBr, was viewed under low power uv light. A slit was made in the agarose just ahead of, but slightly wider than the fragment band to be purified. A piece of DEAE membrane (Schleicher and Schuell, NA45), cut to the same size as the slit, was presoaked in running buffer and placed in the slit using tweezers. The gel was run at high voltage until the fragment was observed on the membrane. The membrane was placed in an 1.5ml Eppendorf tube with 0.4ml 1M NaCl at 65°C for 15 min to elute the DNA. The membrane was removed from the tube. The NaCl solution was extracted with absolute ethanol at -20°C overnight. The tube was spun 10min at 15000xg and washed with 70% ethanol. The pellet was dried and resuspended in 20µl 1 x TE.

2.7 Southern Blotting.

Digested DNA was fractionated on an agarose gel and photographed as described above. A ruler was usually included next to the gel during photography to enable size calculation of bands on the final autoradiograph. After the photograph was taken the area containing the DNA markers was removed. Blots were assembled using 0.4M NaOH as the transfer buffer. Four pieces of buffer soaked 3MM filter paper (Whatman) were placed, as a wick, over a blotting stand in buffer. Gels were inverted and placed on the stand and air bubbles removed by rolling the gel with a glass rod. Hybond N+ positively charged nylon membrane (Amersham Int.), 1cm larger than the gel was labelled and placed face down on the gel. Air bubbles were removed as above. Four pieces of 3MM filter paper soaked in buffer, the same size as the agarose gel, were placed on top of the membrane and air bubbles again removed. Cling film was placed around the edge of the gel between the gel and membrane. Paper towels,

a perspex plate and a small weight (~200g) were placed on top of the gel and blotting allowed to proceed overnight. The membrane was washed briefly in 2 x SSC and used in a hybridisation reaction or stored in cling film at -20°C. Blotted gels were restained with EtBr and observed under uv light to ensure complete transfer of the DNA onto the membrane.

2.8 Preparation of Radiolabelled DNA Probes and Hybridisation.

The method of labelling was chosen according to the size of DNA probe:

2.8.1 Nick Translation Reactions: For probes > 1 kb.

Approximately 30ng of purified insert DNA was combined with 10µl of nucleotide mix and 5µl of enzyme from a Nick Translation kit (Amersham Int.). SDW was added to make the volume up to 47µl and the reaction kept on ice. 3µl (30µCi) of 5'α^{32}P dCTP (3000Ci/mM; ICN) was added and the reaction mixed with the pipette tip. The reaction was placed at 14°C for 2 hours.

2.8.2 Multiprime/Random prime Reactions: For probes < 1kb.

Approximately 30ng of purified insert DNA was suspended in a 10µl volume with SDW. The DNA was denatured at 94°C for 10 min and placed immediately on ice. 10µl of labelling mix, 5µl of primers and 2µl of enzyme were added from a Multiprime kit (Amersham Int.), the reaction made up to 47µl with SDW, and kept on ice. The reaction was maintained on ice. 3µl (30µCi) of 5'α^{32}P dCTP (3000Ci/mM; ICN) was added and the reaction incubated at room temperature for 3-5hours.

2.8.3 PCR Labelling: For probes < 500bp for which PCR primers were available.

Approximately 5ng of template DNA, either purified insert DNA, or purified product amplified from a complex source of DNA such as total genomic or YAC DNA, was amplified in a 50µl reaction overlayed with oil. The reaction mixture was composed of 50mM KCl; 10mM Tris HCl pH8.4; 1-2mM $MgCl_2$; 200µM dATP, dGTP and dTTP; 2µM dCTP; 0.5µM both forward and reverse primer; 1 unit Taq DNA polymerase (Perkin Elmer) and 30µCi 5'α^{32}P dCTP (3000Ci/mM; ICN). 10 cycles of 94°C, 1min; Ta, 1min; 72°C, 1min. were used, where Ta is the annealing temperature determined for the primer pair.

2.8.4 Hybridisation and washing.

After the radiolabelling reaction was complete, 50µl of 1 x TE was added and unincorporated nucleotides removed by filtration through a Sephadex G-50 (Sigma) column. At this stage, approximately 2µg of sonicated genomic DNA could be added for competition of repeat sequences in the probe. The probe was denatured at 94°C for 10min and placed immediately on ice. If competitor DNA was added the probe mixture could be incubated at 37°C to allow annealing of repeat sequence. The incubation time for this was varied according to the probe. Nylon filters were prepared for hybridisation by incubation at 63°C in a pre-hybridisation buffer (6 x SSC; 5 x Denhardts solution; 1% SDS; 200µg/ml herring sperm DNA degraded free acid, Sigma; 1mM EDTA pH 8.0) for at least 2 hours. The probe was added to hybridisation buffer (pre-hybridisation solution containing 10% dextran sulphate) and incubated with the filter, overnight, at 63°C. Filters were washed twice with 2 x SSC; 0.1% SDS for 20 min at 63°C. Further washes were carried out at higher stringency

when required. Filters were rinsed in 2 x SSC, placed on filter paper and sealed in plastic bags. Autoradiography was carried out using X ray film (Fuji RX) at -70°C.

2.8.5 Removal of probe from filters.

Place filter in 0.1% SDS; 0.1 x SSC at 90-95°C for 15 min. Wash in 2 x SSC. Store at -20°C.

2.9 Polymerase Chain Reaction.

Reactions were carried out in a 10µl volume using approximately 50ng of genomic or YAC DNA, or 5ng of plasmid or insert DNA. Alternatively, DNA from single yeast or bacterial colonies was amplified by picking a small number of cells from the colony directly into the reaction tube. Amplification was performed in the presence of 50mM KCl; 10mM Tris HCl pH 8.4; 1.5mM $MgCl_2$; 200µM each dNTP; 0.5µM forward and reverse primer and 0.5 units of Taq DNA polymerase (Perkin Elmer) in a tube overlayed with mineral oil. Normal PCR cycles were: 1 cycle of 94°C, 3min; 55°C, 1½min; 72°C, 1½min; 30 cycles of 94°C, 1min; 55°C, 1½min; 72°C, 1½min, and 1 cycle of 94°C, 1min; Ta (55°C), 2min; 72°C, 3min; in a thermal cycler (Applied Biosystems). To change the specificity of the reaction Ta, and/or the $MgCl_2$ concentration was varied.

2.10 YAC Liquid and Plug DNA Preparation.

Yeast colonies were inoculated into AHC medium containing 50µg/ml ampicillin and incubated for 24-48 hours at 30°C with shaking (250RPM). Yeast strain DNA not containing a YAC, was grown on YPD medium without supplements.

2.10.1 YAC Liquid DNA Preparation. Used for Southern and *in situ* hybridisation and PCR applications.

Cells from 50ml of culture were spun at 3K for 10min and resuspended in 5ml of 0.9M sorbitol; 20mM EDTA and 0.1% β-mercaptoethanol. 40μg/ml lyticase was added and the cells incubated at 37°C for 1 hour with shaking. The tube was spun gently (1K) for 10min and the supernatant removed. Cells were resuspended in 5ml of 4.5M guanidinium hydrochloride; 100mM EDTA; 150mM NaCl; and 50mM sarkosyl pH 8.0 and incubated 65°C for 10min. The sample was cooled and an equal volume of ethanol added. Precipitate was collected by centrifugation (1200xg) for 10 min and resuspended in 2ml of 1 x TE; 100μg/ml RNase A (Sigma) and 200μg/ml proteinase K ((Boehringer Mannheim). The sample was incubated at 65°C for 1hour. DNA was phenol/chloroform extracted three times, chloroform extracted once, then ethanol precipitated, washed in 70% ethanol and resuspended in 250μl of 1 x TE. The resultant YAC liquid DNA was stored at 4°C.

2.10.2 YAC Plug Preparation. Used for pulsed-field gel electrophoresis.

Cells from a 50ml culture were spun (3K) for 10min and resuspended in 500μl SEB. 1μl of suspension was added to 1ml of water and the cells counted using a haemocytometer. The cell concentration was adjusted to 1 x 10^9 cells/ml. The cell suspension was warmed to 37°C and 300μg lyticase was added. It was then combined with an equal volume of prewarmed 1% low geling temperature agarose (Type VII; Sigma) in SEB, mixed and pipetted into plug moulds (6x2x10mm) precooled on ice. Plugs were allowed to set for 15min, transferred from the plug moulds into approximately 10ml SEBT containing 200μg lyticase and incubated at 37°C for 2

hours. The SEBT solution was carefully removed and replaced with 10-20ml of yeast lysis solution for a further 30min incubation at 37°C. The yeast lysis solution was replaced before incubation at 37°C overnight. Plugs were stored in fresh yeast lysis solution at room temperature.

2.10.3 Pulsed Field Gel Electrophoresis.

YAC sizes were determined by fractionation of the yeast chromosomes on 'Waltzer' gel boxes (Southern *et al.*, 1987) or using a CHEF MAPPER XA pulsed field electrophoresis system (BioRad). Plugs were removed from the yeast lysis solution and washed twice in 1 x TE at 4°C for 30min each wash. The plugs were halved, loaded onto 1.5% agarose gels (Type I; Low EEO; Sigma) and run using 0.5 x TAE buffer at 15°C for 24 hours at 250volts, 65sec pulse time and an angle of 110°. EtBr stained gels were illuminated with uv light and photographed. DNA was transferred onto Hybond N+ nylon membrane (Amersham Int.) by Southern blotting. Approximately 100ng mouse genomic DNA was nick translated and hybridised to the Nylon membrane to detect the YAC insert DNA.

2.10.4 Biotin Labelling of YAC DNA for Fluorescent *in situ* Hybridisation (FISH).

YAC DNA (approximately 350ng in 10µl total hybridisation buffer/slide) was labelled with biotin by nick translation (using a BioNick kit, Life Technologies). The labelled YAC was separated from unincorporated nucleotides using a Nick column (Pharmacia) and 0.1 x volume of 3M NaAc and 2 x volume of absolute ethanol were added. The labelled YAC was stored in the ethanol at -20°C until required. Sonicated mouse genomic DNA (for competition of repeat sequences in the YAC) was added to the

ethanol. The probe and competitor DNA were spun 15000xg for 10min and the pellet washed with 70% ethanol then dried. 2µl of SDW was added, the tube mixed, spun and maintained at room temperature to allow the pellet to dissolve. Equal volumes of formamide and 2 x hybridisation buffer per slide were added to give a final concentration of approximately 50% formamide; 5 x Denhardts solution; 5 x SSPE pH 7.2; 10% dextran sulphate and 200µg/ml sonicated salmon sperm DNA. The hybridisation mix was denatured at 94°C for 10min, placed immediately on ice, and incubated at 37°C for 5-30min. The hybrisiation mix was pipetted onto a slide and sealed under a 22 x 22mm coverslip with rubber glue. Hybridisation was performed in a 2 x SSC moist chamber at 42°C for 4 days. The amount of competitor DNA required and the length of the preannealing time depended on the repeat sequences present in the YAC. Approximately 2µg of sonicated mouse genomic DNA and incubation at 37°C for 15min was sufficient for most YACs studied.

Post hybridisation washes were carried out to a stringency of 50% formamide, 1 x SSC at 42°C. YAC signals were detected using avidin conjugated to fluorescein (FITC) and amplified with biotinylated goat anti-avidin (Vector Laboratories). Photographs of *in situ* hybridised slides were taken using MRC 600 confocal microscope (BioRad).

2.11 Genetic Distance and Standard Error Calculations.

Genetic distances, in centimorgans (cM), between loci were calculated from backcross studies. One cM corresponds to one recombination among one hundred meiotic products (Fincham, 1983). The standard error of a genetic distance was calculated as

follows:-

$$SE=\sqrt{\frac{r(1-r)}{n}}$$

where **SE** is the standard error, **r** is the recombination fraction (the number of recombination events over the total number of animals tested) and **n** is the total number of animals tested (adapted from Bailey, 1959).

2.12 Solutions and Media.

AHC Medium

Casein hydrolysate-acid (to make 1% in final medium) and Bacto agar (15g/l if required) were autoclaved, cooled to less than 60°C and the following solutions (filter sterilised) added to give the appropriate concentrations:-

0.7% yeast nitrogen base

2% glucose

10μg/ml adenine hemi-sulphate

50μg/ml ampicillin

50 x Denhardts solution

1% Ficoll (Sigma)

1% Polyvinylpyrrolidine

1% Bovine serum albumin (Fraction V; Sigma)

This solution was stored at -20°C.

GET

50mM glucose

25mM Tris HCl (pH 8.0)

10mM EDTA (pH 8.0)

This solution was autoclaved and stored 4°C.

LB Medium

1% Bacto tryptone

1% NaCl

0.5% Bacto-yeast extract

0.1% D-glucose

1.5% Bacto agar was added when preparing plates.

This solution was autoclaved.

10 x Loading buffer

0.4% Bromophenol blue

25% Ficoll (Sigma).

KAc (Potassium acetate for plasmid preparations)

3M Potassium

5M Acetate

(Add 11.5ml glacial acetic acid and 28.5ml of H_2O to 60ml

5M Potassium acetate).

SEB

1.2M Sorbitol

20mM EDTA

15mM β-mercaptoethanol

SEBT As for SEB, with the addition of 10mM Tris HCl (pH 7.5).

SET

0.15M NaCl

5mM EDTA

50mM Tris

The pH of the solution was ajusted to 7.8. This solution was autoclaved and stored at 4°C.

2 x SSC

300mM NaCl

30mM Tri NaCi (Tri sodium citrate)

The pH was adjusted to 7.0 with NaOH.

20 x SSPE

0.2M $NaH_2PO_4.H_2O$

3M NaCl

20mM EDTA

The pH was ajusted to 7.4 with NaOH

1 x TAE (pH 8.0)

40mM Tris-acetate

1mM EDTA (24.2g Tris, 5.71ml glacial acetic acid and 10ml 0.5ml EDTA in 1l).

1 x TE

10mM Tris

1mM EDTA

The pH was adjusted to 7.8

0.5 x TBE (pH 8.0)

45mM Tris-borate

1mM EDTA (5.4g Tris, 2.75g orthoboric acid and 2ml 0.5M EDTA in 1l).

Yeast Lysis solution

1% Lithium Dodecyl sulphate

100mM EDTA

10mM Tris-HCl; pH8.0

YPD Medium

1% bacto-yeast extract

2% peptone

2% glucose

The yeast extract and peptone were added before autoclaving and the filter sterilised glucose to the medium at less than 50°C, after autoclaving.

Chapter 3

The Construction of Genetic Maps of Regions of the Mouse X Chromosome which Share Homology with Human Proximal Xp.

3.1 Introduction.

Detailed maps of human and mouse X chromosomes will allow accurate prediction and identification of mouse models for human genetic disease. The region on the human X chromosome short arm from CYBB to the centromere (Figure 1.5) was thought to be composed of one complete segment of homology shared between the human and mouse X chromosomes (Searle *et al.*, 1989); however, more recently this region has been shown to contain at least four conserved segments (Laval and Boyd, 1993 a and b). These conserved segments can be defined by a single locus (for conserved segments 2 and 4), or by the most proximal and distal loci they contain (for conserved segments 1 and 3) as follows: 1) CYBB-PFC; 2) GATA1; 3) DXS674-ALAS2 and 4) DXF34. Three segments (1, 2 and 4) lie in the proximal region of the mouse X chromosome and one, the DXS674-ALAS2 segment, lies in the distal region (Figure 1.7). To improve the comparative map of the proximal region of the human X chromosome short arm, murine homologues of human genes and conserved sequences were positioned on the genetic map of the mouse X chromosome. In particular, the map positions have been determined for loci which lie close to evolutionary breakpoints between human and mouse. Several simple sequence repeat/microsatellite loci have been mapped with respect to genic loci on the mouse X chromosome; although these provide no information for comparative mapping studies they are useful tools for the identification of backcross offspring for pedigree

analysis and for the isolation of DNA clones for physical mapping studies (see Chapter 4).

3.1.1 General Strategy.

The following sequential approach was taken to position markers on the genetic map of the mouse X chromosome:-

(1) Amplification product variants (APVs) defined by PCR, or restriction fragment length variants (RFLVs) detected by hybridisation of DNA probes to Southern blots, were identified between the *M. musculus* strain 3H1[1] and *M. spretus* at each locus.

(2) The X-linked inheritance pattern of the *M. spretus* variant was confirmed using a 3H1 female x *M. spretus* male cross, where the *M. spretus* allele of an X-linked locus is inherited by the female but not the male offspring. As initial mapping was carried out using backcrosses to 3H1, it was of lesser importance to confirm X-linkage of the *M. musculus* variant.

(3) The locus was positioned into a ~10-15cM interval using a small set of backcross offspring with single recombination events dividing the mouse X chromosome into eight intervals (referred to as a mouse X chromosome mapping panel, MXMP; Laval and Boyd, 1993a). The haplotypes of the backcross offspring which compose the MXMP are shown in Figure 3.1.

(4) Detailed localisation of the locus was carried out by analysis of a panel of mice with recombination events covering the interval of interest. Panels of backcross offspring were constructed (see section 3.2 below) with recombination

[1] 3H1 is an F_1 hybrid between C3H/HeH female and 101/H male *M. musculus* inbred strains.

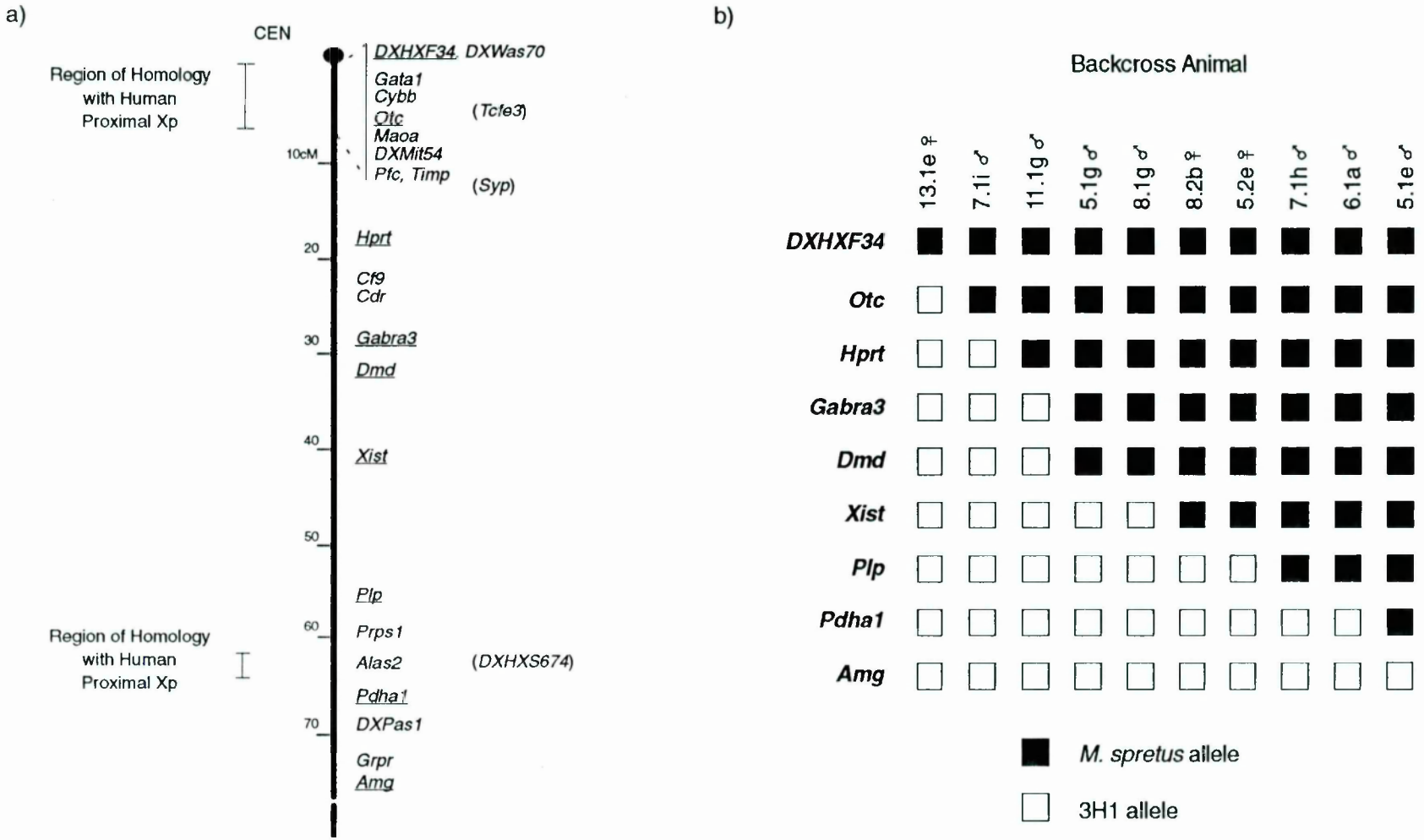


Figure 3.1

(a) Genetic map of the mouse X chromosome. The positions of loci are taken from Herman *et al.*, 1994. Loci in parentheses have not been positioned accurately on the map. Regions of homology with human proximal Xp are indicated on the left. Details of relevant loci are given in Table 3.1.

(b) The haplotypes of backcross offspring which constitute the mouse X chromosome mapping panel, MXMP (Laval and Boyd, 1993a), at the eight loci underlined in (a). The offspring have single recombination events which divide the X chromosome into eight intervals of approximately 10-15cM.

events in the proximal and distal regions of the mouse X chromosome which share homology with human proximal Xp. (Figure 3.1).

(5) By minimising the number of double recombinants the order of loci was determined and genetic distances calculated to produce a genetic map (genetic distance and standard error calculations are explained in 2.11).

Details of all loci discussed in this study are listed in Table 3.1, hybridisation probe details and the X-linked RFLVs they detect in Table 3.2, and primer details and product sizes of X-linked APVs in Table 3.3. The approximate positions of loci on the genetic map of the mouse X chromosome are shown in Figure 3.1.

3.2 Construction of Mouse Panels for Pedigree Analysis.

Offspring from two *M. musculus* x *M. spretus* interspecific backcrosses have been analysed and panels of mice identified with recombination events in:

(1) the proximal region of the mouse X chromosome, between the loci *DXWas70* and *Pfc*. This region shares homology with three conserved segments from human proximal Xp which are defined by CYBB-PFC, GATA1 and DXF34.

(2) the distal region of the mouse X chromosome, between the loci *Plp* and *Pdha1*. This region shares partial homology with one conserved segment from human proximal Xp which is defined by DXS674-ALAS2.

The identification of relevant recombination events in each backcross is described separately below.

Table 3.1 - Locus Details.

Human Locus Symbol	Mouse Locus Symbol	Gene Name	Mouse Location	Human Location	Reference to mapping of locus in man and/or mouse
ALAS2	*Alas2*	Erythroid δ-aminoevulinate synthase	XF	Xp11.21	Cotter *et al.*, 1992; Chapman *et al.*, 1994.
AMG	*Amg*	Amelogenin	XF3/4	Xp22.31-Xp22.1	Lau *et al.*, 1989; Chapman *et al.*, 1991a.
CDR1	*Cdr*	Cerebellar degeneration related protein	XA6/7	Xp27.1	Hirst *et al.*, 1991; Laval *et al.*, 1992.
F9	*Cf9*	Coagulation factor IX	XA6	Xq26.3	Purrello *et al.*, 1985; Avner *et al.*, 1987.
CYBB	*Cybb*	Cytochrome b-245 β peptide	XA	Xp21.1	Royer-Pokora *et al.*, 1986; Chapman *et al.*, 1991b.
DMD	*Dmd*	Dystrophin	XC	Xp21.3-Xp21.2	Brockdorff *et al*; 1987.
DXF34	*DXHXF34*	Sequence family conserved human genomic clone	XA1	Xp11.22-cen	Laval and Boyd, 1993a.
DXS674	*DXHXS674*	Single copy conserved human genomic clone		Xp11.2-Xp11.21	Laval and Boyd, 1993a.
DXS679	*DXHXS679*	Single copy conserved human genonic clone		Xp11.2-Xp11.21	Laval and Boyd, 1993a.
	DXWas70	Mouse region specific repeat (non-conserved)	XA		Disteche *et al.*, 1987.
GABRA3	*Gabra3*	GABA receptor α3 subunit	XA/D	Xq28	Buckle *et al.*, 1989; Faust *et al.*, 1991.
GATA1	*Gata1* (*Gf1*)	GATA binding protein 1	XA	Xp11.23	Zon *et al.*, 1990; Chapman *et al.*, 1991b.
GRPR	*Grpr*	Gastrin releasing peptide receptor	XF	Xp22.2	Schantz *et al.*, 1991; Maslen and Boyd, 1993.
HPRT	*Hprt*	Hypoxanthine phosphoribosyl transferase	XA6	Xq26.1	Avner, *et al.*, 1987.
MAOA	*Maoa*	Monoamine oxidase A	XA	Xp11.4-Xp11.3	Lan *et al.*, 1989; Derry *et al.*, 1989.
NPHL	*Nphl* (*Clcn5* or *Clc5*)	X-linked nephrolithiasis (Dents disease) Voltage-gated chloride channel N5		Xp11.23	Fisher *et al.*, 1994.
OTC	*Otc*	Ornithine transcarbamylase	XA1/2	Xp21.1	Lindgren *et al.*, 1984; Amar *et al.*, 1988.

PDHA1	*Pdha1*	Pyruvate dehydrogenase E1α subunit	XF3/4	Xp22.1-Xp22.2	Brown *et al.*, 1989; 1990.
PFC	*Pfc*	Properdin factor complement	XA2/3	Xp11.3	Goundis *et al.*, 1988; Laval, *et al.*, 1991.
PLP	*Plp*	Proteolipid protein	XF	Xq21.33-Xq22	Dautigny *et al.*, 1986
SYP	*Syp*	Synaptophysin		Xp11.23-Xp11.22	Özçelik *et al.*, 1990
TFE3	*Tcfe3*	Transcription factor enhancer 3	XA2	Xp11.23	Roman *et al.*, 1992
TIMP1	*Timp*	Tissue inhibitor of metalloproteases	XA	Xp11.3	Huebner *et al.*, 1986; Mullins *et al.*, 1988.
XIST	*Xist*	Xi-specific transcript	XD	Xq13.2	Brockdorff *et al.*, 1991
XK	*Xkh*	McLeod syndrome		Xp21.1	Ho *et al.*, 1994

Table 3.2 - Hybridisation Probe Details and X-linked RFLVs Detected.

Mouse Locus Symbol	Probe Name	Probe Origin	Restriction Enzyme for RFLV	X-Linked RFLV band sizes (kb)		Mouse Mapping Reference
				M. musculus (3H1)	*M. spretus*	
Cdr	pCDR-9	Human cDNA	TaqI	5.0, 0.8	6.0	Laval *et al.*, 1992
Cybb	CYBB	Human cDNA	TaqI	4.0	2.3, 1.8	Laval and Boyd, 1993b
DXHXF34	pH3-7 (and cDNA clone pHFB7.2A2)	Human genomic (and human cDNA)	EcoRI	4.0, 1.3	3.0	Laval and Boyd, 1993a (Laval *et al.*, unpublished data)
DXHXS674	pH3-68	Human genomic	RsaI	1.4	1.6	Laval and Boyd, 1993a
DXHXS679	pH3-145	Human genomic	BglII	6.5	2.5	This study
DXWas70	70-68	Mouse genomic	TaqI	3.5, 1.8	4.0, 3.0	Disteche *et al.*, 1987
Gata1	K14	Human cDNA	EcoRI	3.2	3.0	Laval and Boyd, 1993b
Nphl	RL-6	Human cDNA	PvuII	9.4	3.5	This study
Pfc	mP-830	Mouse cDNA	TaqI	3.1	5.0	Laval *et al.*, 1991
Syp	mSYPE4	Mouse genomic	TaqI	5.0	3.0	This study
Xkh	XK	Human cDNA	TaqI	0.8	2.2	This study

Table 3.3 - Details of Amplification Product Variants used in this Study

Mouse Locus Symbol	Primer Sequence (5'→3') Forward Reverse	Type of Sequence	Product sizes (bp) *M. musculus* (3H1)	 *M. spretus*	Reference
DXMit26	TTGGCAAGCATGCTTTACTG AGGAACATGGAAACACCTGC	Simple Sequence Repeat (SSR)	220	206	Dietrich *et al.*, 1996
DXMit28	ACTGCAGGTACCTCTATGTGAGC TGTTATGGCAACAGAAACAAGC	SSR	178	192	"
DXMit34	GCTTGATCCTTGAATTTTAACAA TTTTTCTCTGTAGTGTTTCCTTGA	SSR	248	196	"
DXMit35	GGGACAATTGTCAAAGGCAT CTGCTGACACTTTTTCTGAAACA	SSR	158	122	"
DXMit54	ACATCTGATGTGGATGTTCAGCAGC CTCCGATGGAAACTGTTGGT	SSR	144	212	"
DXMit55	CTGCTTCCAGAATATTATCACTACTCC AAAACATCCATTTATGTTAACACACA	SSR	116	100	"
Tcfe3	TAAGGGTATGCCCCTGGCCAC AAGGTCAGCACAGAGTCCTCA	3' UTR (untranslated region)	267, 188[a]	455[a]	This study

[a] The amplification product variant for *Tcfe3* is observed after digestion of the PCR product with HhaI.

3.2.1 Pedigree analysis panel for the *DXWas70-Pfc* interval of the mouse X chromosome.

3.2.1.1 Backcross 1 - (3H1♀ x *M. spretus* ♂) F_1 ♀ x 3H1 ♂.

3H1 female animals were mated with *M. spretus* males (wild-derived stock maintained by random mating at the MRC Radiobiology Unit) and the F_1 hybrid female offspring from this cross were mated with 3H1 males. DNAs from 108 offspring prepared by S.H. Laval (MRC Radiobiology Unit) had been analysed previously using DNA markers which detected the loci *DXWas70* and *DXHXF34* (Laval and Boyd, 1993a); *Otc, Maoa, Pfc* and *Hprt* (Laval *et al.*, 1991); *Cf9*, *Cdr, Gabra3* and *Dmd* (Laval *et al.*, 1992); *Xist, Plp* and *Pdha1* (Blair *et al.*,1993); *Grpr* (Maslen and Boyd, 1993) and *Amg* (Blair *et al.*, 1993). The scorings of the 108 backcross offspring at these loci are shown in Appendix A. There were 7 animals which had recombination events between *DXWas70* and *Pfc,* from 103 animals scored for both loci. Pedigree analysis of these animals gave an order of *DXWas70/DXHXF34 - Gata1 - Cybb - Otc - Maoa - Pfc* (Laval *et al.*, 1991; Laval and Boyd, 1993b).

To identify more animals with recombination events in the *DXWas - Pfc* interval, DNAs from a further 55 offspring were prepared and analysed for Taq I RFLVs between 3H1 and *M. spretus* at *DXWas70* and *Pfc.* (Figure 3.2 a and b). As a result a further 5 animals with recombination events between *DXWas70* and *Pfc* were identified (Table 3.4). From this backcross, therefore, a total of 12 animals carried recombination events between *DXWas70* and *Pfc*, from 158 animals scored

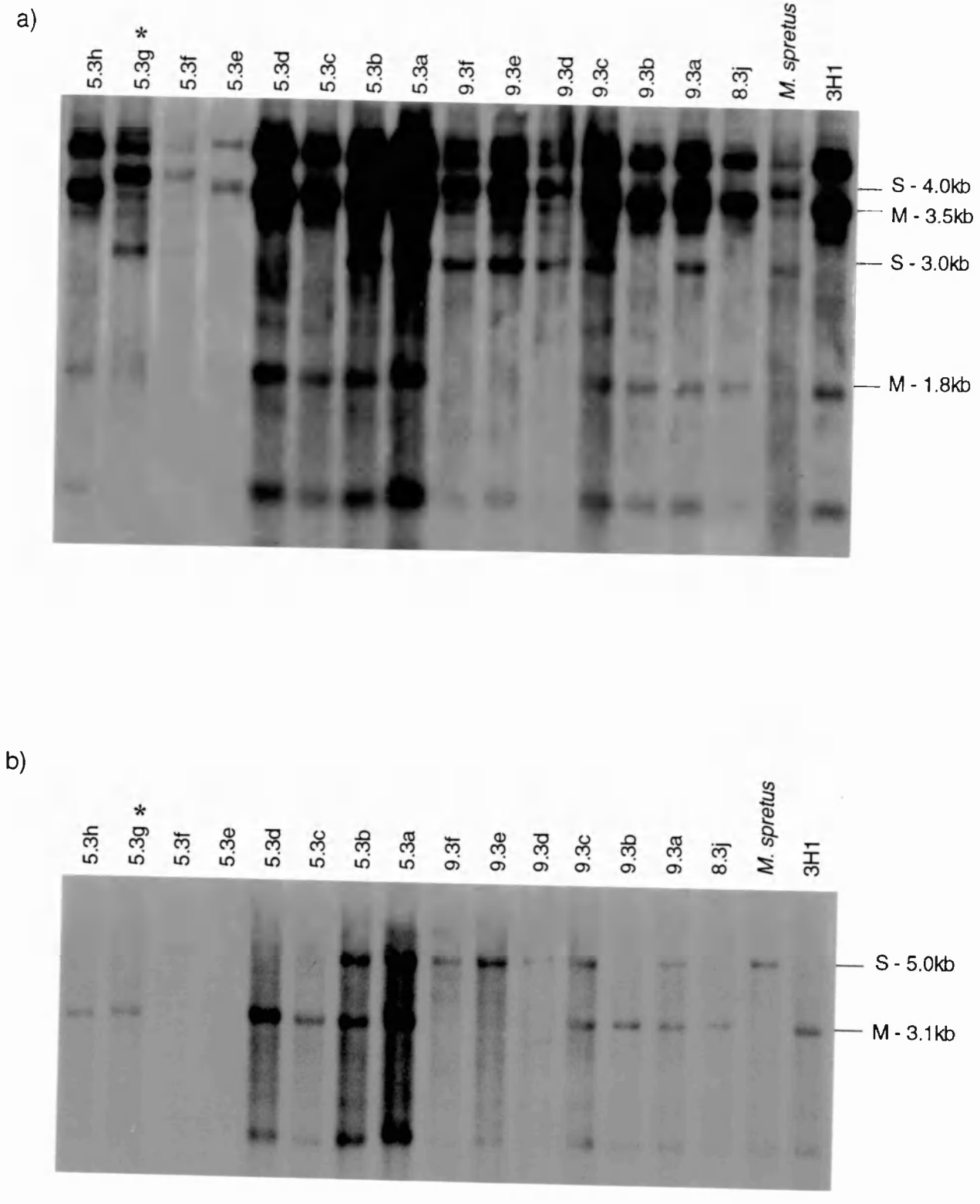


Figure 3.2
Hybridisation of probes which detect the loci (a) *DXWas70* and (b) *Pfc* to Taq I digested DNA from representative extra backcross 1 offspring. Details of probes are listed in Table 3.2. S and M indicate the *M. spretus* and *M. musculus* (3H1) X-linked bands respectively. Animal 5.3g, indicated by * has a recombination event between *DXWas70* and *Pfc*.

Table 3.4 Scorings of Extra Offspring from Backcross 1 for RFLVs at *DXWas70* and *Pfc*.

Animal	Sex[a]	*DXWas70*	*Pfc*
5.3a	f	S[b]	S
5.3b	f	S	S
5.3c	f	M	M
5.3d	f	M	M
5.3e	m	M	M
5.3f	m	S	S
5.3g	m	S X	M
5.3h	m	M	M
5.3i	m	S X	M
5.3j	m	M	M
6.3a	f	S	S
6.3b	f	S	S
6.3c	f	S	S
6.3d	f	S	S
6.3e	f	M X	S
6.3f	f	M	M
6.3g	f	S	S
6.3h	f	S	S
6.3i	m	M	M
6.3j	m	M X	S
6.3k	m	M	M
7.3a	f	S	S
7.3b	f	S	S
7.3c	f	M	M
7.3d	f	M	M
7.3e	m	M	M
7.3f	m	S	S
7.3g	m	M	M

Animal	Sex	*DXWas70*	*Pfc*
7.3h	m	S	S
7.3i	m	S	S
8.3b	f	M	M
8.3c	f	M	M
8.3d	f	S	S
8.3e	f	S	S
8.3f	f	M	M
8.3g	m	S	S
8.3h	m	M	M
8.3i	m	M	M
8.3j	m	M	M
9.3a	f	S	S
9.3b	f	M	M
9.3c	f	S	S
9.3d	m	S	S
9.3e	m	S	S
9.3f	m	S	S
10.3h	m	M	M
10.3i	m	M	M
10.3j	m	S	S
11.3a	f	M	M
11.3c	f	S X	M
11.3d	f	M	M
11.3f	m	S	S
11.3g	m	M	M
11.3h	m	S	S
11.3i	m	S	S

[a] f = female; m = male.

[b] M or S for the *M.musculus* or *M.spretus* allele inherited from the F_1 female parent of the backcross offspring respectively.

X= a recombination event between *DXWas70* and *Pfc*.

for both loci. The genetic distance between these two loci was calculated as 7.6 ± 2.1cM.

3.2.1.2 Backcross 2 - (An♀ x *Mus spretus* ♂) F_1 ♀ x An ♂.

An[2] females were mated with *M. spretus* males and the F_1 females produced were mated with An males to produce the backcross progeny. DNAs were prepared by V. Reed (MRC Radiobiology Unit) from 84 backcross offspring which had previously been analysed for RFLVs at the X-linked loci *Xist, Plp, Pdha1* and *Amg* (Blair *et al.*, 1993). In order to identify animals with recombination events between *DXWas70* and *Pfc* these DNAs were analysed for RFLVs at *DXWas70, DXHXF34, Gata1* and *Pfc*. RFLVs at the *DXHXF34* locus were detected by a human cDNA clone, pHFB7.2A2, isolated using the pH3-7 probe (S.H. Laval, personal communication). Representative scorings of animals at the four loci are shown in Figure 3.3, whilst the complete scorings of this backcross, at all loci tested, are shown in Appendix B. 5 animals were found to have recombination events between *DXWas70* and *Pfc*, from a total of 66 scored for both loci. The genetic distance between *DXWas70* and *Pfc* was calculated as 7.6 ± 3.3cM, and is the same as that calculated from the analysis of backcross 1 mice (see above). No recombination events were detected between the loci *DXWas70* and *DXHXF34* in the 60 mice tested for both loci. There were 3 mice with recombination events between *DXWas70/DXHXF34* and *Gata1* (60 animals tested) and 2 mice with recombination events between the loci *Gata1* and *Pfc* (64 animals tested). Therefore, the order and genetic distances between loci from the proximal region

[2] An is a steroid sulphatase (STS)-deficient *M. musculus* C3H substrain (Jones *et al.*, 1989).

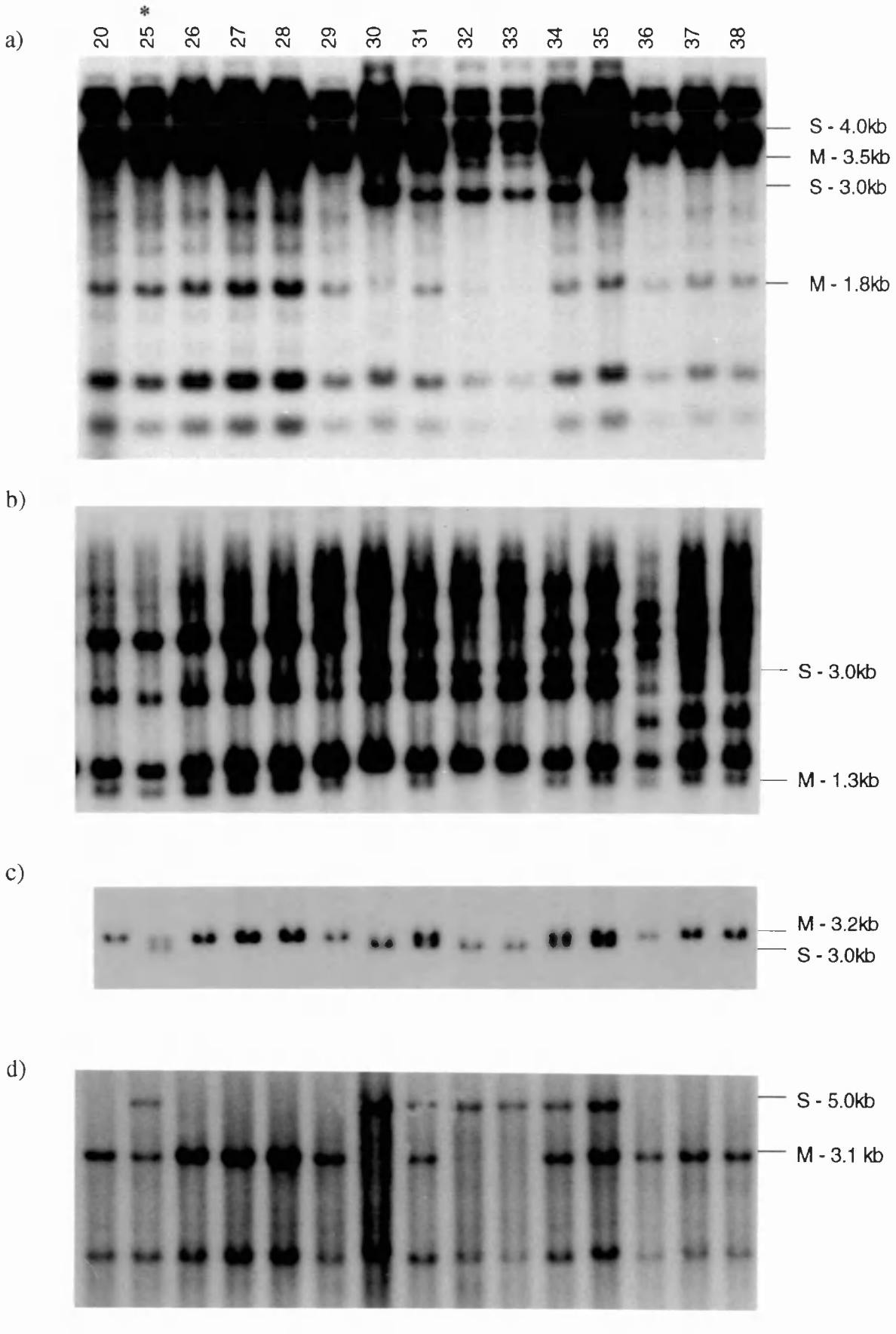


Figure 3.3

Analysis of representative offspring from backcross 2 for RFLVs at (a) *DXWas70*, (b) *DXHXF34*, (c) *Gata1* and (d) *Pfc*. DNAs from the offspring were digested with the restriction enzymes TaqI for *DXWas70* and *Pfc* analyses (a and d), and EcoRI for *DXHXF34* and *Gata1* analysis (b and c). Details of probes are listed in Table 3.2. M and S indicate the *M. musculus* and *M. spretus* X-linked bands respectively. Animal 25, indicated by * has a recombination event between *DXHXF34* and *Gata1*.

of the mouse X chromosome were calculated as: *DXWas70/DXHXF34* - (5.0 ± 2.8cM) - *Gata1* - (3.1 ± 2.2cM) - *Pfc*.

3.2.2 Pedigree analysis panel for the *Plp-Pdha1* interval of the mouse X chromosome.

There were 15 mice with recombination events in the distal region of the mouse X chromosome between the loci *Plp* and *Pdha1* from 100 mice tested from backcross 1 (see 3.2.1.1). There were 7 mice with recombination events between *Plp* and *Pdha1* in the distal region of the mouse X chromosome in backcross 2 (84 animals analysed for both loci). The genetic distance, calculated from the combined data of backcrosses 1 and 2, between *Plp* and *Pdha1* was 12.0 ± 2.4cM (Blair *et al.*, 1993).

3.3 Mapping of the Conserved Sequences *DXHXS674* and *DXHXS679*.

The human conserved single-copy clone, pH3-68, which detects the locus DXS674, lies in the Xp11.23-Xp11.21 region of the human X chromosome within 400kb of ALAS2, the erythroid-specific δ-aminolevulinate synthase locus (Laval and Boyd, 1993a; Boyd *et al.*, 1994). The murine homologue of DXS674, *DXHXS674*, was positioned in the distal region of the mouse X chromosome in the interval between *Plp* and *Pdha1* using the MXMP described in Section 3.1 (Laval and Boyd, 1993a). This locus identified a novel region of homology between the human and mouse X chromosomes. *Alas2* has now also been positioned in the distal region of the mouse X chromosome, between the loci *Prps1* and *DXPas1* (Chapman *et al.*, 1994; Figure 3.1). These data indicated that *Alas2* also lies in the novel conserved segment on the mouse X chromosome identified by *DXHXS674*.

The cross-hybridising human single-copy clone, pH3-145, detects the locus DXS679, which lies within 40kb of DXS674 in man (Boyd *et al.*, 1994, Maslen *et al.*, 1995). The mapping of the murine homologue of this locus and the refinement of the position of *Alas2* and *DXHXS674* could help to determine the size, orientation and precise position of the conserved segment in the distal region of the mouse X chromosome.

An X-linked BglII RFLV between 3H1 and *M. spretus* had previously been identified with the probe pH3-145 (V. Reed, personal communication). To position the murine homologue of DXS679 on the mouse X chromosome, pH3-145 was hybridised to BglII digested DNAs of the MXMP (Figure 3.4a). Haplotype analysis of the panel indicated that *DXHXS679* also lies in the interval between *Plp* and *Pdha1* on the mouse X chromosome (Figure 3.4b). As *DXHXS674, DXHXS679* and *Alas2* all lie in the distal region of the mouse X chromosome and are physically linked in man, they were likely to belong to the same conserved segment. To confirm this, *DXHXS674* and *DXHXS679* were positioned more accurately using the 22 backcross offspring with recombination events in the *Plp* - *Pdha1* interval (see 3.2.2). DNAs prepared from these mice were digested with RsaI and BglII and hybridised with the probes pH3-68 and pH3-145 respectively (Figure 3.5a). No recombination events were observed between *DXHXS674* and *DXHX679*. However, 12 mice had recombination events between *Plp* and *DXHXS674* and *DXHX679*, and 10 mice had recombination events between these two loci and *Pdha1* (Figure 3.5b). This indicates that *DXHXS674* and *DXHX679* are in the same conserved segment which must lie approximately half way between *Plp* and *Pdha1* on the genetic map of the mouse X chromosome.

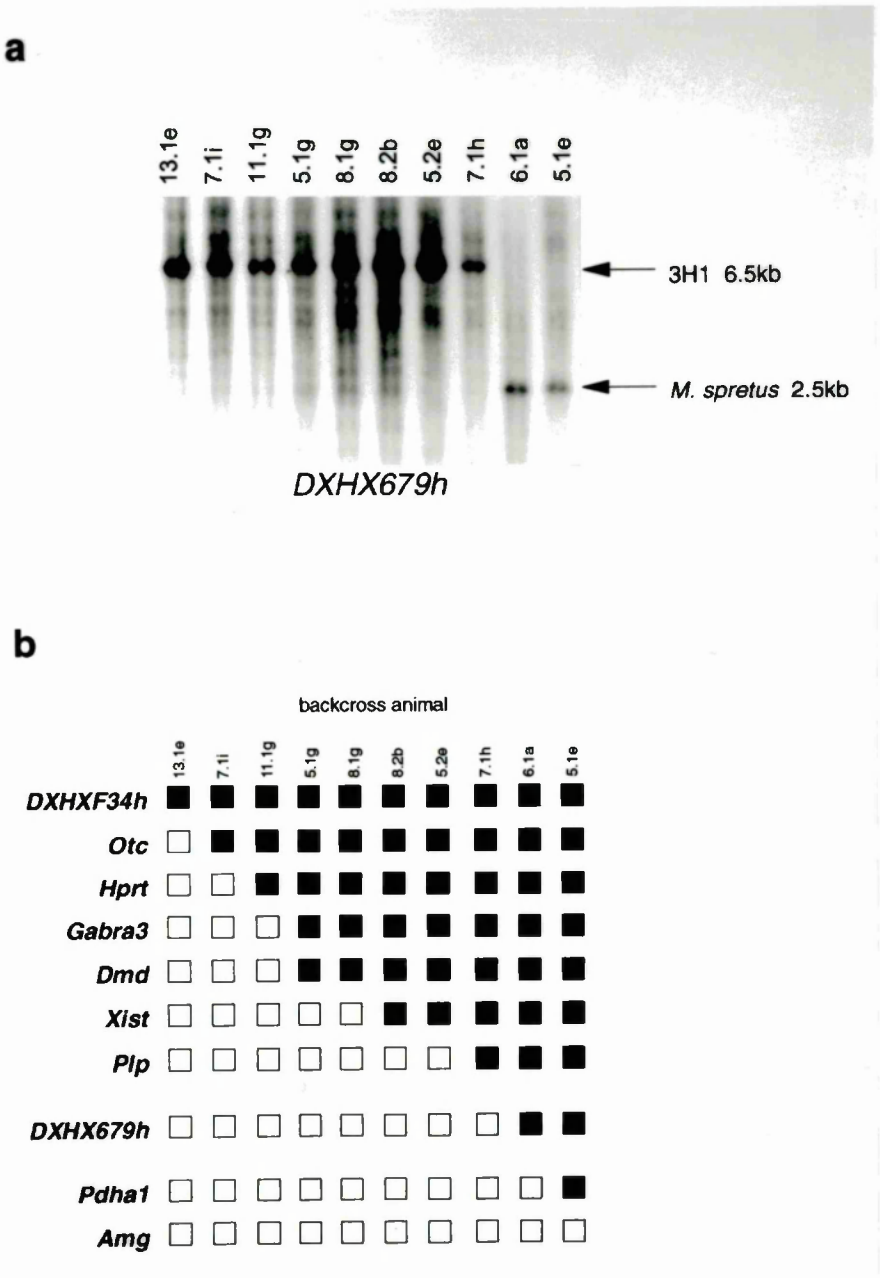
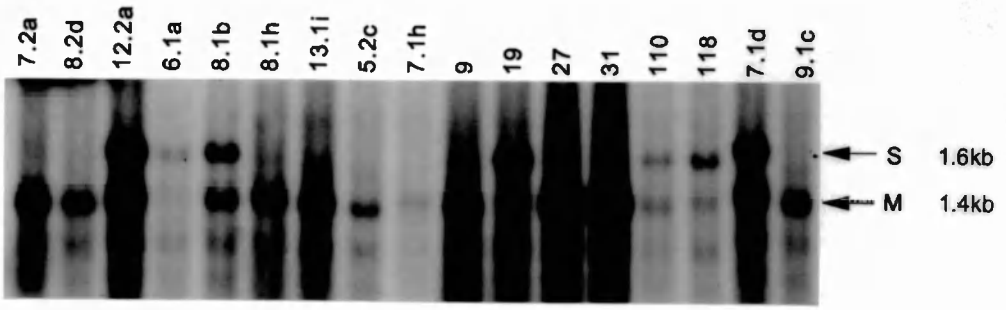


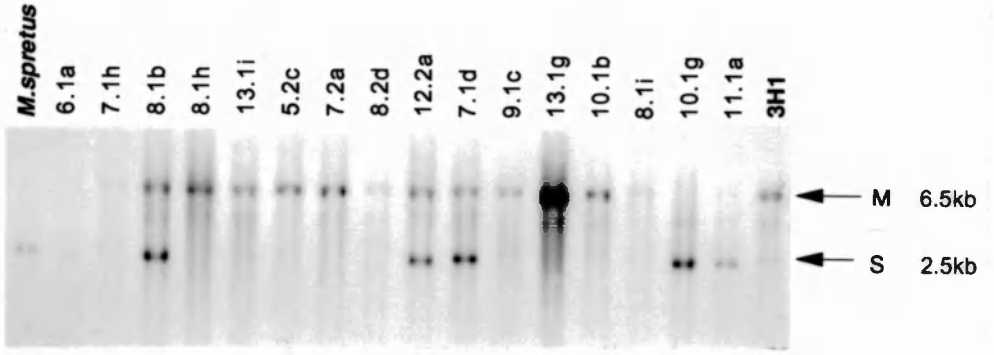
Figure 3.4

a) Hybridisation of the probe pH3-145, which detects the locus *DXHXS679* (also now known as *DXHX679h*), to BglII digested DNA from the MXMP. The sizes and positions of X-linked bands are indicated. (b) Haplotype analysis of the MXMP. Open squares represent the presence of the *M. musculus* (3H1) allele and black squares the presence of the *M. spretus* allele. *DXHXS679* (*DXHX679h*) lies between *Plp* and *Pdha1* on the mouse X chromosome.

a)



DXHX674h



DXHX679h

b)

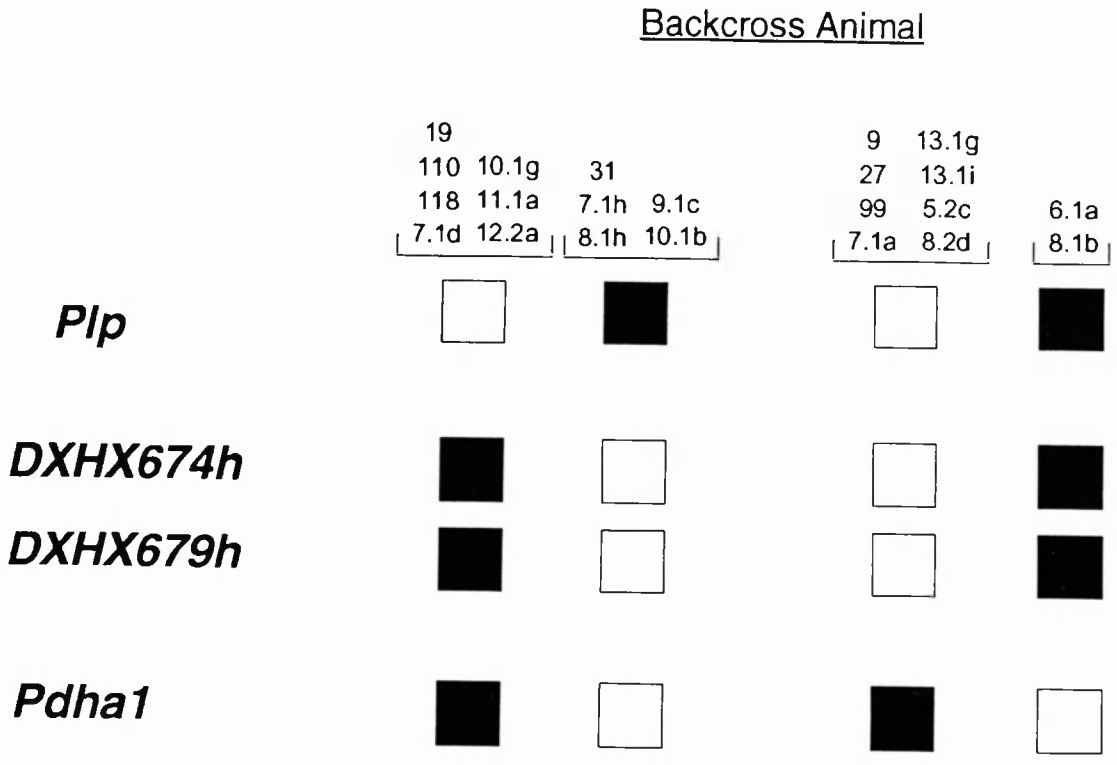


Figure 3.5

a) Hybridisation of probes which detect *DXHXS674* (*DXHX674*h) and *DXHXS679* (*DXHX679h*) to representative offspring from backcrosses 1 and 2 with recombination events in the *Plp* to *Pdha1* region of the mouse X chromosome. DNAs were digested with RsaI for *DXHXS674* analysis and BglII for *DXHXS679* analysis. M and S indicate the *M. musculus* and *M.spretus* X-linked bands respectively (b) Haplotype analysis of backcross offspring with recombination events between *Plp* and *Pdha1*. Open squares represent the presence of the *M. musculus* allele and black squares the presence of the *M. spretus* allele.

The number of markers positioned in the *Plp-Pdha1* interval was increased by PCR analysis of APVs between 3H1 and *M. spretus* at the *DXMit34, DXMit35* and *DXMit28* loci (Figure 3.6, Table 3.5). The order of, and genetic distances between, loci in this interval were calculated to be: *Plp* - (4.9 ± 1.6cM) - *DXMit35* - (1.1 ± 0.4cM) - *DXMit34* - (0.5 ± 0.5cM) - (*DXHXS674, DXHXS679*) - (4.4 ± 1.4cM) - *DXMit28* - (1.1 ± 0.7cM) - *Pdha1*.

In an attempt to position the murine *Alas2* locus with respect to *DXHXS674* and *DXHX679,* and because the *Alas2* cDNA was not available, PCR primers were designed from the sequence and intron/exon boundary details of Scheonhaut and Curtis (1986, 1989). A product of 187bp assumed to be from exon 2 of the *Alas2* gene (nucleotides 33 to 220, Scheonhaut and Curtis, 1986) was obtained after amplification of 3H1 DNA. Amplification was carried out using standard PCR conditions (see 2.9) from a forward primer, 5' - TGG GCT CAG GAT GGT GGC - 3', and a reverse primer, 5' - CTC CAG CCT TGG TTG CCT TA - 3'. The product was radiolabelled using PCR and hybridised to 3H1 and *M. spretus* DNA digested with several different restriction enzymes. Although clear bands were observed, no scorable RFLVs between 3H1 and *M. spretus* were detected using the PCR labelled probe (data not shown). Therefore *Alas2* could not be positioned on the mouse genetic map with respect to *DXHXS674* and *DXHXS679.*

Alas2 has since been shown to cosegregate with *DXMit34* by Pasteris *et al.* (1995) using recombination events from 126 interspecific backcross offspring. *DXHXS674* and *DXHXS679* were positioned only 0.5 ± 0.5cM distal to *DXMit34* in this study. The

Table 3.5 Scorings of Offspring from Backcrosses 1 and 2 with Recombination Events between *Plp* and *Pdha1*.

Animal	6.1a	7.1d	7.1h	8.1b	8.1h	9.1c	10.1b	10.1g	11.1a	13.1g	13.1i	5.2c	7.2a	8.2d	12.2a	9	19	27	31	99	110	118
Sex	♂	♀	♂	♀	♂	♀	♀	♂	♀	♂	♂	♀	♀	♀	♀	♂	♂	♀	♀	♀	♀	♂
Plp	S	M	S	S	S	S	S	M	M	M	M	M	M	M	M	M	M	M	S	M	M	M
		X				X	X	X	X						X		X		X			X
DXMit35	nt	S	S	nt	S	M	M	S	S	nt	nt	nt	nt	nt	S	nt	S	nt	M	nt	M	S
					X																X	
DXMit34	nt	S	S	nt	M	M	M	S	S	nt	nt	nt	nt	nt	S	nt	S	nt	M	nt	S	S
			X																			
DXHXS674	S	S	M	S	M	M	M	S	S	M	M	M	M	M	S	M	S	M	M	M	S	S
DXHXS679	S	S	M	S	M	M	M	S	S	M	M	M	M	M	S	M	S	M	M	M	S	S
	X									X		X	X	X		X		X		X		
DXMit28	M	S	M	S	M	M	M	S	S	S	M	S	S	S	S	S	S	S	M	S	S	S
				X							X											
Pdha1	M	S	M	M	M	M	M	S	S	S	S	S	S	S	S	S	S	S	M	S	S	S

M indicates the *M. musculus* (3H1) allele and S the *M. spretus* allele inherited from the F1 female parent of the backcross offspring; nt indicates animal not tested; X indicates the position of a recombination event.

murine homologues of several other loci which lie in proximal Xp in man have now also been positioned in the distal region of the mouse X chromosome. The implications of these mapping data for the comparative map of proximal Xp in man and the identification of mouse models for human inherited disease are discussed in Chapter 5.

3.4 Mapping of *Tcfe3*, the Transcription Factor Enhancer 3 Locus on the Mouse X Chromosome.

At the start of this study, one of the four conserved segments which compose proximal Xp in man was defined by a single locus, GATA1. It was not known if any other conserved markers which lie close to GATA1 in man also lie close to *Gata1* in the mouse. TFE3, the gene encoding the transcription factor enhancer 3, was known to lie proximal and very close to GATA1 in man (Nelson *et al.*, 1995) but its murine homologue, *Tcfe3,* had not been positioned on the mouse genetic map. However, *Tcfe3* had been placed in band A2 on the mouse X chromosome by *in situ* hybridisation (Roman *et al.*, 1992). If it is assumed that all the rearrangements between the human and mouse X chromosome have been discovered, then *Tcfe3* would be expected to lie either distal to *Pfc* or close to *Gata1* on the mouse X chromosome. To position *Tcfe3* with respect to other markers on the mouse genetic map and determine to which conserved segment it belongs, an RFLV at the locus was required.

Initially, a ~1.8kb human TFE3 cDNA clone which contained the region from nucleotide 655 to the 3' end of the gene, including the 3' untranslated region (Beckmann *et al.*, 1990), was hybridised to mouse genomic DNA. The hybridisation

signal was non-specific and repetitive and subclones isolated from the cDNA clone also failed to detect clear hybridisation signals. RFLVs between 3H1 and *M. spretus* could therefore not be found (data not shown).

A variant between 3H1 and *M. spretus* was identified through the differential ability of the restriction enzyme HhaI to digest a PCR product amplified from the murine *Tcfe3* gene. Primers, 5' - TAA GGG TAT GCC CCT GGC CAC - 3', and 5' - AAG GTC AGC ACA GAG TCC TCA - 3', were designed to amplify a 455bp product corresponding to nucleotides 1618 - 2073 in the 3' untranslated region of the mouse cDNA sequence published by Roman *et al.* (1992). PCR was carried out using standard conditions and an annealing temperature of 62°C. The 455bp product amplified from *M. spretus* DNA was not cleaved by HhaI whereas, that from 3H1 yielded 267bp and 188bp HhaI digestion products. The HhaI cut site in the amplification product is found at nucleotide 1885 in the *Tcfe3* cDNA sequence (Roman *et al.*, 1992). X-linkage of this variant was confirmed by amplification, and subsequent HhaI digestion, of DNA from the male and female offspring of a 3H1 x *M. spretus* cross (Figure 3.7a). In addition, products were also amplified from the *M. musculus* strains; C3H/HeH, 101/H, C57BL/6J, CBA/H, 129, JU, MTX, JPX, PT, $Hprt^-$ and the subspecies *M. musculus castaneus* all of which were found to digest with HhaI to yield the 267bp and 188bp products (Figure 3.7b).

Tcfe3 was first positioned in the interval between the centromere and *Otc* in the proximal region of the mouse X chromosome using the MXMP (Figure 3.8a and b). The PCR analysis was found to be quick and easy to perform therefore the entire

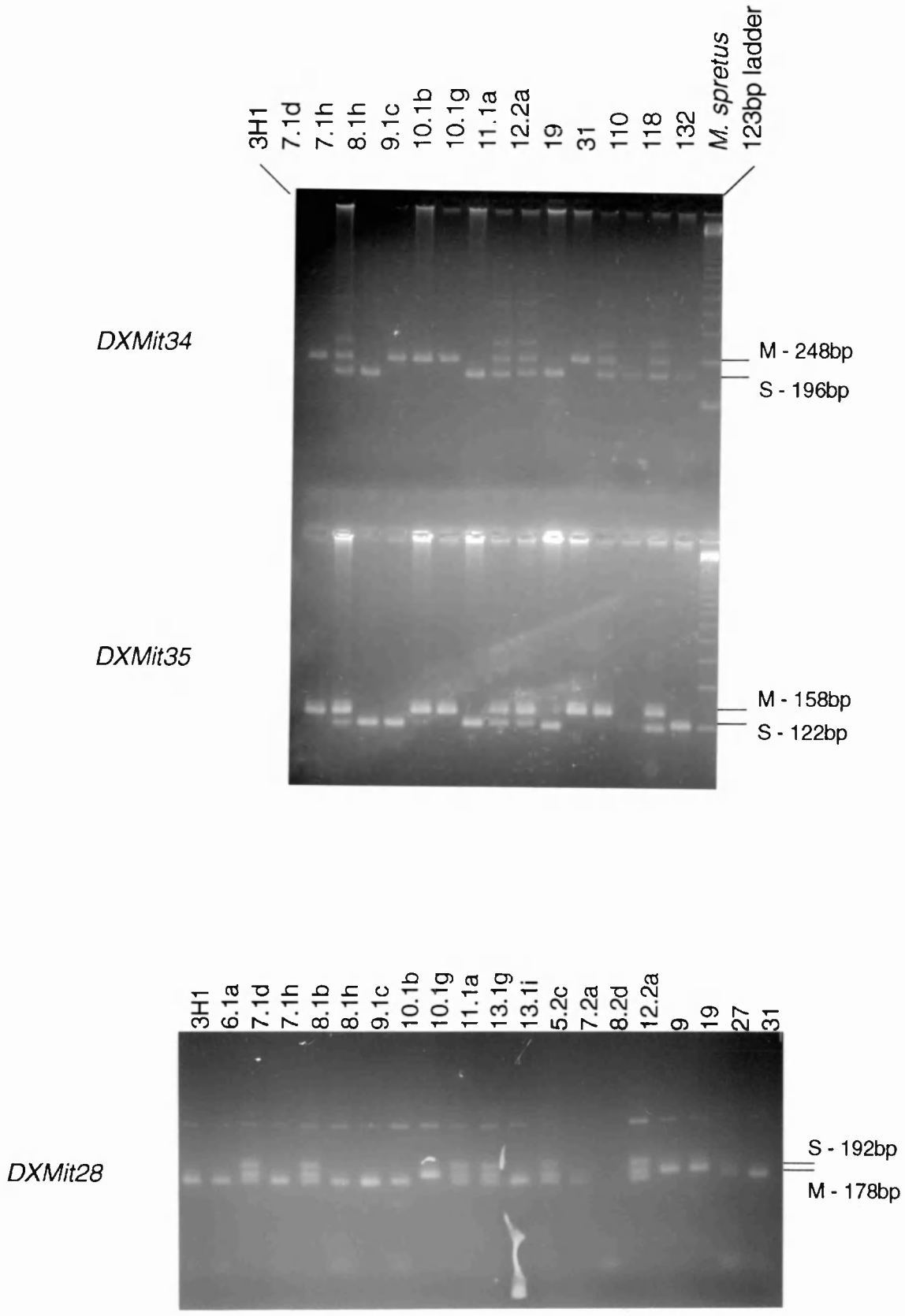


Figure 3.6

Representative scorings of APVs in offspring from backcrosses 1 and 2 at the microsatellite loci *DXMit34, 35* and *28*. Amplification products and their sizes for 3H1 (M) and *M.spretus* (S) are indicated for each microsatellite.

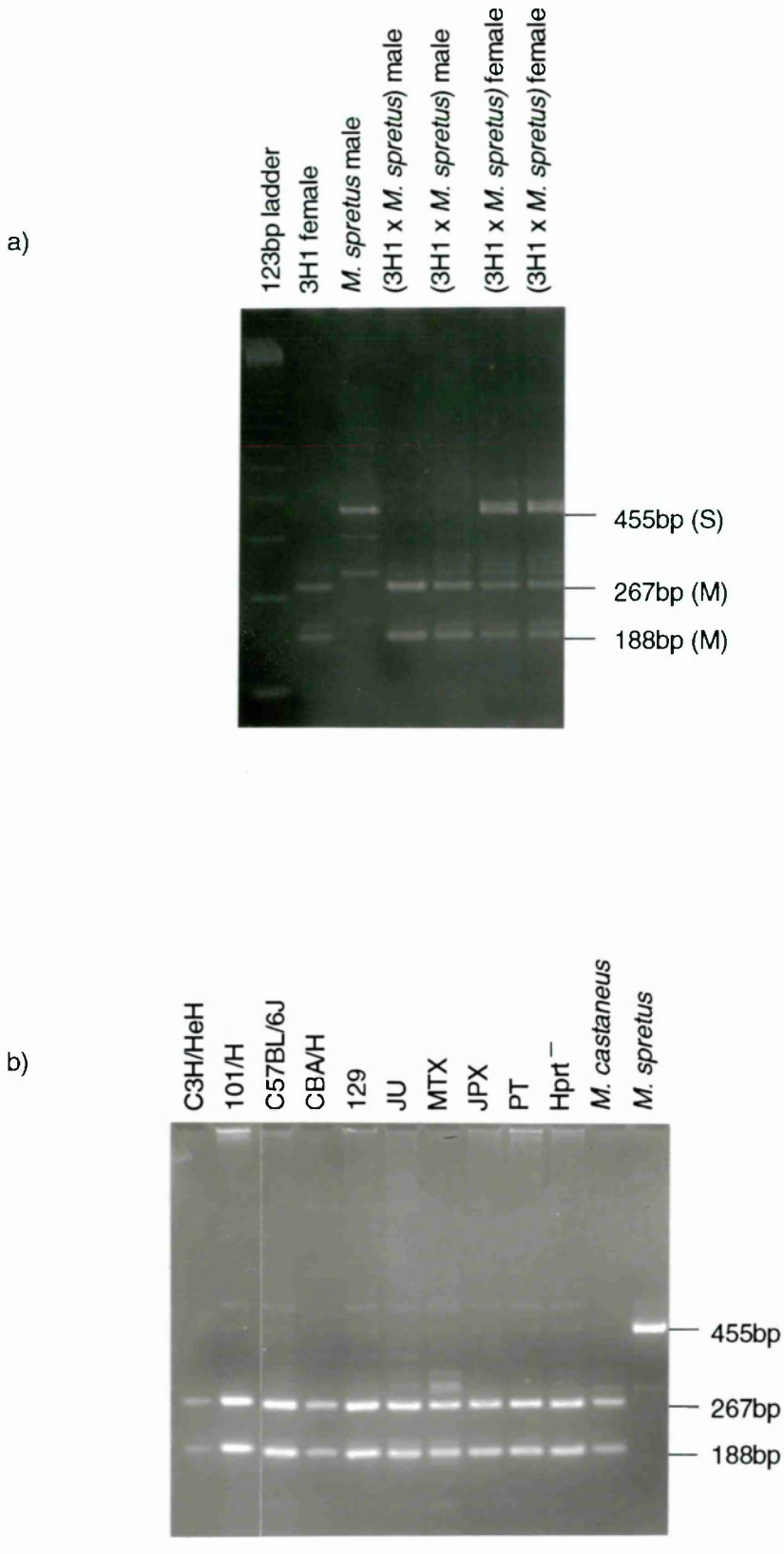
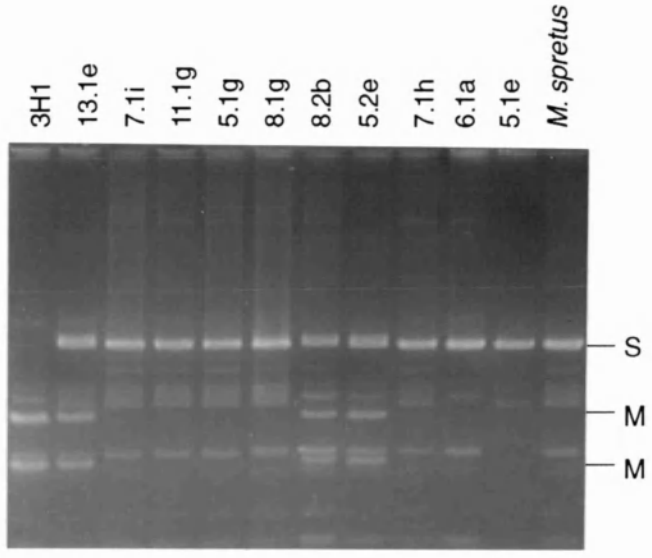


Figure 3.7

Digested amplification products at *Tcfe3* from (a) 3H1 female and *M. spretus* male mice and the offspring from a 3H1 female x *M. spretus* male cross. The digested amplification product variant exhibits X-linked inheritance and (b) *M. musculus* strains and *M. musculus castaneus* digested to yield the 267bp and 188bp products. Details of inbred strains can be found in Lyon *et al.*, 1996. MTX, JPX and Hprt⁻ are partially inbred strains maintained at the MRC Harwell.

a)



b)

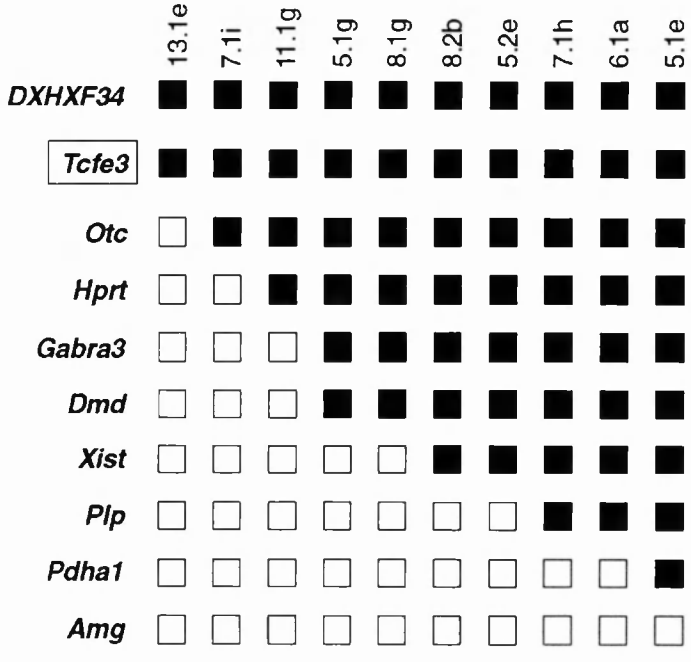


Figure 3.8

a) HhaI digested amplification products at *Tcfe3* from MXMP DNAs. M and S indicate the *M. spretus* and 3H1 APVs. (b) Summary of pedigree analysis data for *Tcfe3* on the MXMP. Open squares represent 3H1 alleles and black squares *M. spretus* alleles. *Tcfe3* lies proximal to *Otc* on the mouse X chromosome.

backcrosses were scored for *Tcfe3*. (Figure 3.9, Appendices A and B). There were a total of 6 animals with recombination events between *DXWas70* and *Tcfe3* (221 mice analysed for both loci) including a mouse from backcross 1 (8.1f♀) not previously identified as having a recombination event in the *DXWas70* to *Pfc* interval (see section 3.2.1). There were 12 animals with recombination events between *Tcfe3* and *Pfc* (234 analysed for both loci) and the genetic distances were therefore calculated as: *DXWas70* - (2.7 ± 1.1cM) - *Tcfe3* - (5.1 ± 1.4cM) - *Pfc*. Of the 18 mice with recombination events between *DXWas70* and *Pfc* only 11 had been analysed for RFLVs at *Gata1* (Laval and Boyd, 1993b; Appendix B); none of these had recombination events between *Tcfe3* and *Gata1*. Therefore, TFE3 must lie in the conserved segment defined by GATA1. TFE3 lies immediately proximal to GATA1 in man (Nelson *et al.* 1995) however, the order of the mouse homologues was unknown. To determine the size and orientation of this conserved segment the mouse homologues of genes which lie close to GATA1 and TFE3, in human proximal Xp, had to be positioned (see 3.5).

3.5 The Production of a Refined Comparative Map of Human Proximal Xp.

At the start of this study, the proximal region of the mouse X chromosome, from the centromere to *Pfc,* was thought to be composed of three segments of homology with human proximal Xp. One of these segments in man was defined by the most distal and proximal loci it contained; CYBB-PFC. The other two conserved segments were defined by single loci; DXF34 and GATA1 (Laval and Boyd, 1993 a and b). The GATA1 segment now appeared to contain TFE3 (see section 3.4) but the size and

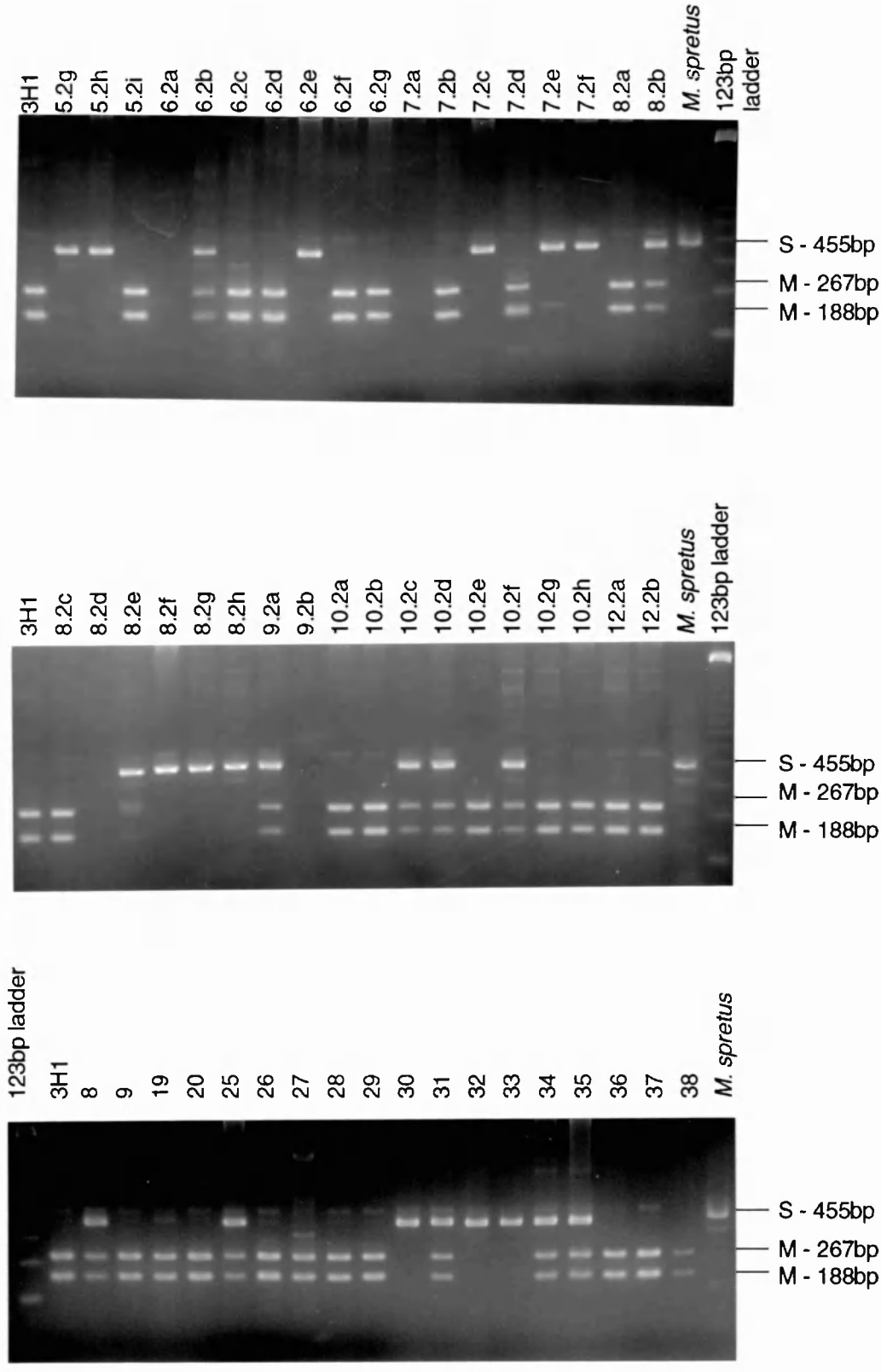


Figure 3.9

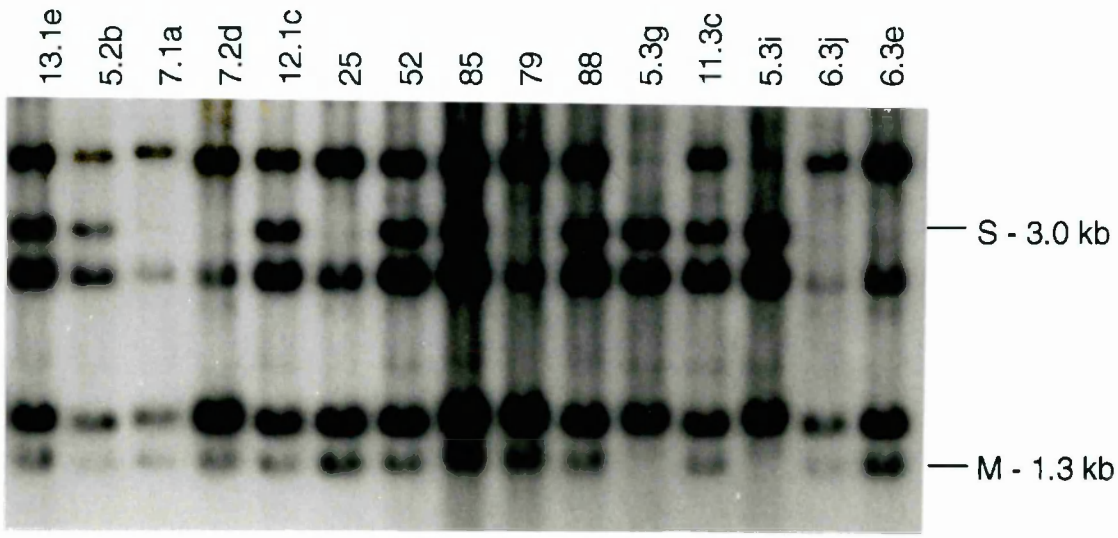
Analysis of the amplification product variant at *Tcfe3* in representative animals from backcrosses 1 and 2. M and S indicate the 3H1 and *M. spretus* variants respectively.

orientation of this segment was not known. To allow accurate prediction of the position of murine homologues of human disease loci and identify possible mouse models for human disease, the comparative map of this region required improvement. With this in mind, an attempt was made to position the murine homologues of genes which lay close to evolutionary breakpoints in human proximal Xp on the mouse genetic map. To map markers in the proximal region of the mouse X chromosome, pedigree analysis was performed on the panel of 18 backcross offspring (13 from backcross 1 and 5 from backcross 2) with recombination events between *DXWas70* and *Pfc* (section 3.2). This panel was fully characterised (see 3.5.1) before being used to position three human genes on the mouse genetic map (see 3.5.2 - 4).

3.5.1 Further characterisation of the *DXWas70-Pfc* pedigree analysis panel.

Although the markers *DXHXF34*, *Gata1* and *Cybb* had previously been positioned in the proximal region of the mouse X chromosome by scoring backcross offspring used in this study (Laval and Boyd, 1993b; Appendix A and B), several of the 18 offspring with recombination events between *DXWas70* and *Pfc* had not been analysed for all three loci. Therefore, they were all scored for RFLVs at *DXHXF34*, *Gata1* and *Cybb* (Figure 3.10; Table 3.6). As no recombination events were identified between *DXWas70* and *DXHXF34*, in a total of 161 backcross offspring (Appendixes A and B), their order could not be determined from this analysis. *Tcfe3* and *Gata1* also cosegregated in all 18 backcross offspring. There were 6 recombination events between *DXWas70/DXHXF34* and *Tcfe3/Gata1* and 12 between *Tcfe3/Gata1* and *Pfc*; one between *Tcfe3/Gata1* and *Cybb* and 10 between *Cybb* and *Pfc*. Only one animal (11.3c) was not scored at *Cybb* (Table 3.6). The order of loci and the positions of

a)



b)

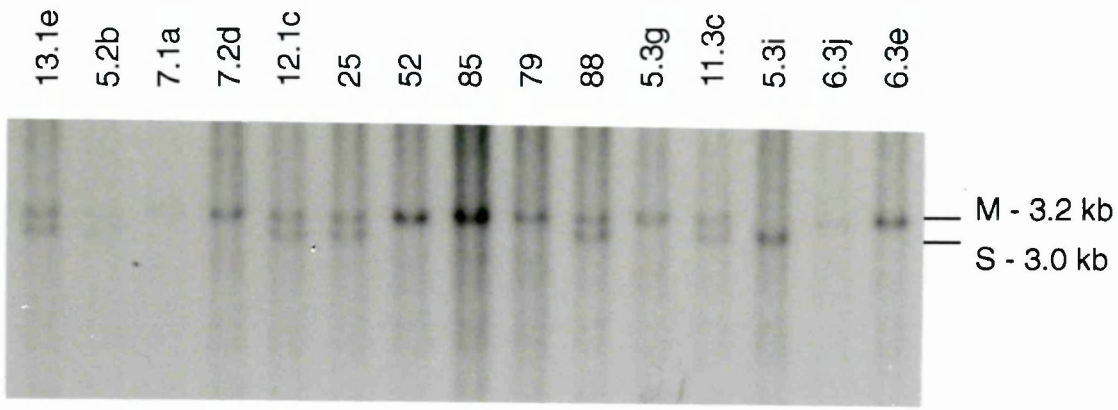


Figure 3.10

Hybridisation of probes which detect the loci (a) *DXHXF34* and (b) *Gata1* to EcoRI digested DNA from representative mice with recombination events between *DXWas70* and *Pfc*. Several of these mice had previously been scored for these loci. The approximate sizes of 3H1 (M) and *Mus spretus* (S) X-linked bands are indicated on the right of each photograph. *Cybb* analysis is not shown.

Table 3.6 Scorings of Offspring from Backcrosses 1 and 2 with Recombination Events between *DXWas70* and *Pfc*.

Animal	8.1f	10.1b	5.2i	13.1e	7.1a	5.2b	7.2d	12.1c	5.3g	5.3i	6.3e	6.3j	11.3c	25	52	85	79	88
Sex	♂	♀	♂	♀	♀	♀	♂	♀	♂	♂	♀	♂	♀	♀	♀	♀	♂	♀
DXWas70	S	M	M	S	M	S	M	S	S	S	M	M	S	M	S	S	M	S
DXHXF34	S	M	M	S	M	S	M	S	S	S	M	M	S	M	S	S	M	S
	X	X												X	X	X		
Nphl	M	S	M	S	M	S	M	S	S	S	nt	nt	S	S	M	M	M	S
Syp	M	S	M	S	M	S	M	S	S	S	M	M	S	S	M	M	M	S
DXMit55	M	S	M	S	M	nt	nt	nt	S	nt	nt	nt	nt	S	nt	M	nt	nt
DXMit26	M	S	nt	S	M	nt	nt	nt	S	S	M	M	S	S	M	M	M	S
									X									
Tcfe3	M	S	M	S	M	S	M	S	M	S	M	M	S	S	M	M	M	S
Gata1	M	S	M	S	M	S	M	S	M	S	M	M	S	S	M	M	M	S
Xkh	M	S	M	S	M	S	M	S	M	S	M	M	S	S	M	M	M	S
			X										(X)					
Cybb	M	S	S	S	M	S	M	S	M	S	M	nt	nt	S	M	M	M	S
				X	X	X							(X)				X	X
DXMit54	nt	nt	S	M	S	M	M	S	M	S	M	M	M	nt	nt	nt	S	M
							X	X		X	X	X						
Pfc	nt	S	S	M	S	M	S	M	M	M	S	S	M	S	M	M	S	M

M indicates the *M.musculus* (3H1) allele and S the *M spretus* allele inherited from the F1 female parent of the backcross offspring; nt indicates animal not tested; X indicates the position of a recombination event; (X) indicates the possible position of a recombination event.

evolutionary breakpoints (EB) in this region were determined as: (*DXWas70/DXHXF34*) - EB - (*Tcfe3/Gata1*) - EB - *Cybb* - *Pfc.*

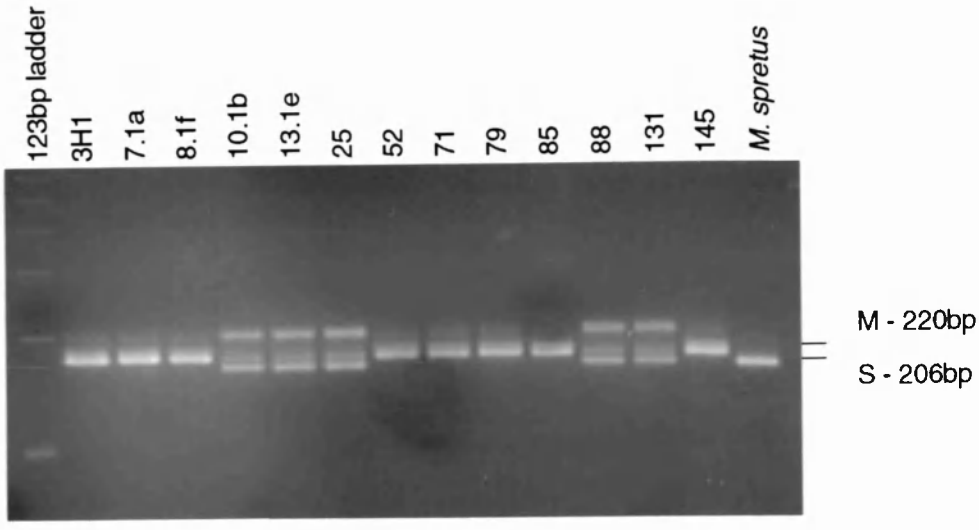
To increase the number of markers positioned in the *DXWas70-Pfc* interval and produce a framework map for physical mapping, simple sequence repeat/microsatellite loci which appeared to lie in the region, were selected from the map of Dietrich *et al.* (1994). APVs at *DXMit26, DXMit55* and *DXMit54* were analysed in the panel of 18 backcross offspring (Figure 3.11; Table 3.6). The number of recombination events and the recombination frequencies between loci are shown in Table 3.7. The order of loci in the proximal region of the mouse X chromosome is: (*DXWas70, DXHXF34*) - (*DXMit26, DXMit55*) - (*Tcfe3, Gata1*) - *Cybb -DXMit54* - *Pfc.*

In order to further improve the human-mouse X chromosome comparative map, the mouse homologues of the human genes, SYP, NPHL and XK which lie near evolutionary breakpoints in human proximal Xp were positioned in the *DXWas70* - *Pfc* interval on the mouse X chromosome.

3.5.2 The human synaptophysin gene. SYP.

Synaptophysin is found in the membrane of small synaptic vesicles in brain and endrocrine cells and is thought to play a role in the transmission of nerve impulses across the synapse (Özçelik *et al.,* 1990). The human synaptophysin gene (SYP) had been isolated and positioned on the X chromosome in Xp11.22-Xp11.23 close to TFE3 (Nelson *et al.*, 1995). The mouse *Syp* gene had previously been localised to region A-D of the mouse X chromosome by somatic cell hybrid analysis (Özçelik *et al.*,

a) *DXMit26*



b) *DXMit54*

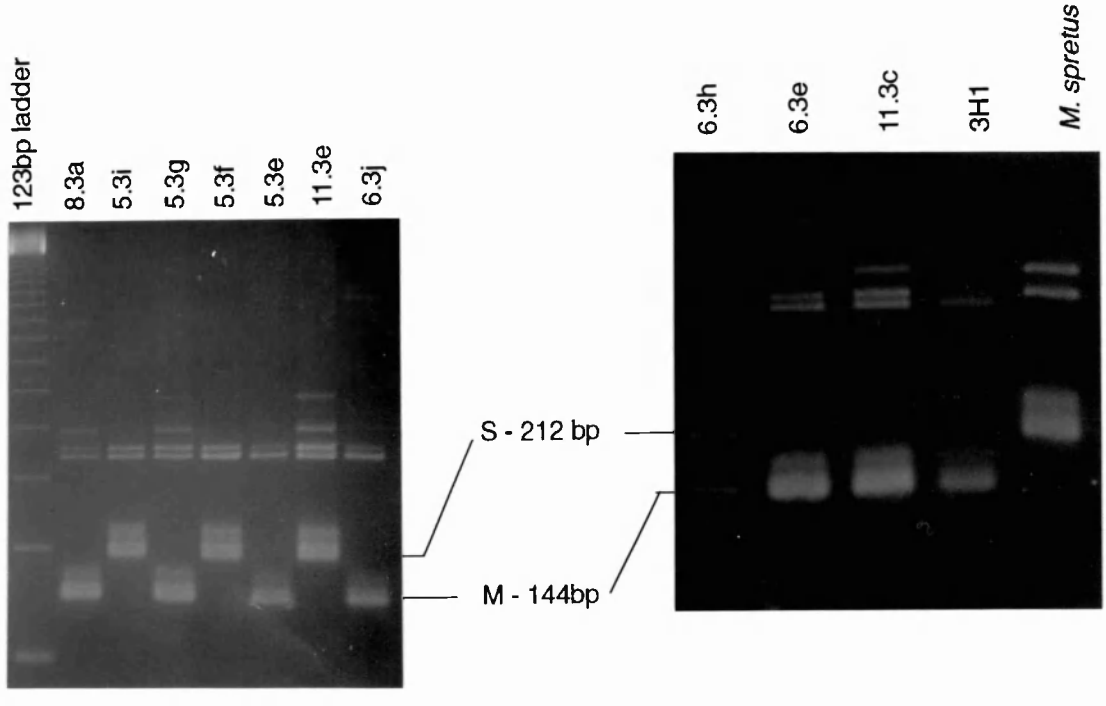


Figure 3.11

Representative scorings of (a) *DXMit26* and (b) *DXMit54* of offspring from backcrosses 1 and 2. M and S indicate the approximate sizes of the 3H1 and *M. spretus* amplification products respectively. *DXMit55* scorings are not shown.

Table 3.7 Pairwise Recombination Frequencies Between Loci in the Proximal Region of the Mouse X Chromosome.

	DXWas70/ DXHXF34	*DXMit55/ DXMit26*	*Tcfe3/ Gata1*	*Cybb*[a]	*DXMit54*	*Pfc*
DXWas70/ DXHXF34	-	5/225 2.2±1.0	6/225 2.7±1.0	7 or 8/225 3.1±1.1 or 3.5±1.2	13/225 5.8±1.6	17/224 7.6±1.8
DXMit55/ DXMit26		-	1/225 0.4±0.4	2 or 3/225 0.9±0.6 or 1.3±0.8	8/225 3.5±1.2	13/225 5.8±1.5
Tcfe3/ Gata1			-	1 or 2/225 0.4±0.4 or 0.9±0.6	7/225 3.1±1.1	12/225 5.3±1.5
Cybb[a]				-	5 or 6/225 2.2±1.0 or 2.7±1.0	10 or 11/225 4.5±1.4 or 4.9±1.4
DXMit54					-	5/225 2.2±1.0
Pfc						-

a: Two recombination frequencies have been calculated for any distances involving *Cybb* because the recombination event in animal 11.3c (not scored at *Cybb*) could be located between *Tcfe3*/*Gata1* and *Cybb* or between *Cybb* and *DXMit54*.

1990), but had not been positioned on the genetic map of the mouse. *Syp* would be expected to belong to the *Tcfe3/Gata1* conserved segment and a refined *Syp* localisation on the genetic map of the mouse X chromosome could provide further information about the size and orientation of this segment.

In order to position *Syp* on the mouse genetic map, primers were designed to amplify a product from exon 4 of the human SYP gene. A forward primer, 5' - GAC CCA AGT GTA CTT TGA TGC - 3', and a reverse primer, 5' - GCC CTT TGT TAT TCT CTC - 3', which represent nucleotides 1570-1591 and 1735-1752 of the human exon 4 SYP gene (Özçelik *et al.*, 1990), were used to amplify a product of approximately 180bp from 3H1 DNA. The human product was slightly smaller than the product amplified from mouse and a product of the same size could be amplified from a human SYP cosmid supplied by S. E. Fisher (University of Oxford), (results not shown). Amplification was carried out using standard conditions except for a buffer containing a 1mM concentration of $MgCl_2$.

Because no difference in size was observed between the products amplified from 3H1 and *M. spretus* DNA using the exon 4 primer pair, PCR could not be used to analyse backcross mice. In order to identify a RFLV, the mouse exon 4 product was radiolabelled using PCR and hybridised to mouse DNA. The PCR exon 4 probe (mSYPE4) detected variant bands of 5 kb and 3 kb in TaqI digests of DNA prepared from 3H1 and *M. spretus*. The *M. spretus* 3kb band was inherited in a X-linked manner in the male and female offspring of a 3H1 x *M. spretus* cross (results not shown). DNA from the MXMP was digested with TaqI and probed with mSYPE4

(Figure 3.12a). Analysis of these results positioned *Syp* proximal to *Otc* in the proximal region of the mouse X chromosome (Figure 3.12b). Pedigree analysis of the 18 mice with recombination events between *DXWas70* and *Pfc* (Figure 3.13b; Table 3.6) revealed that *Syp* cosegregated with *DXMit26* and *DXMit55* and lay proximal to *Tcfe3/Gata1*.

As predicted, SYP lies in the same conserved segment as GATA1 and TFE3. At the beginning of this study, it was not clear whether SYP was located distal or proximal to TFE3, therefore these data helped to clarify the order in man. The most proximal and distal loci which compose this segment are now SYP and GATA1 in man and *Syp* and *Tcfe3/Gata1* in the mouse.

3.5.3 NPHL: The voltage-gated chloride channel gene implicated in Dent's disease. (Previously known as ClCK2 and ClCN5).

NPHL encodes a member of the voltage-gated chloride channel family and has been found to be mutated in patients with the X-linked nephrolithiasis, Dent's disease (Fisher *et al.*, 1994). In man, NPHL lies between the conserved loci SYP and DXS674 in Xp11.22 (Nelson *et al.*, 1995); therefore the mouse homologue of NPHL could lie in the *Syp - Gata1/Tcfe3* conserved segment, or in the *Alas2-DXHXS674/DXHXS679* conserved segment, or could identify a new region of homology between the human and mouse X chromosomes (see Figure 1.7).

The human NPHL cDNA probe, RL-6, (Fisher *et al.*, 1994) was hybridised to DNAs from 3H1 and *M. spretus* mice digested with several different restriction enzymes

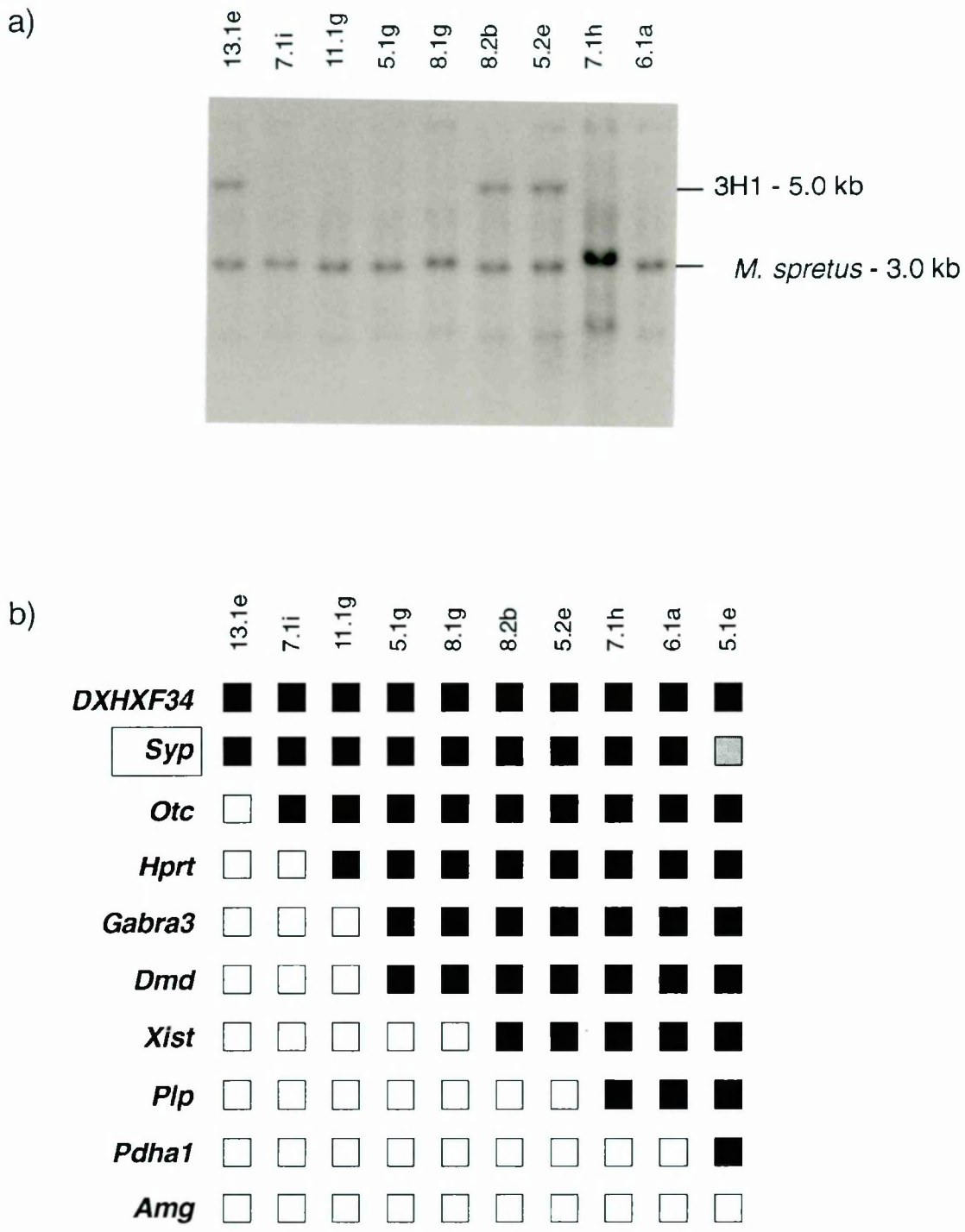


Figure 3.12

a) Pedigree analysis of *Syp* on TaqI digested genomic DNA from the MXMP. 3H1 (M) and *Mus spretus* (S) X-linked RFLVs are shown next to the bands. (b) Summary of pedigree analysis data for *Syp* on the MXMP. 3H1 alleles are represented as open squares, *Mus spretus* alleles as black squares and alleles not scored as grey squares. *Syp* lies proximal to *Otc* on the mouse X chromosome.

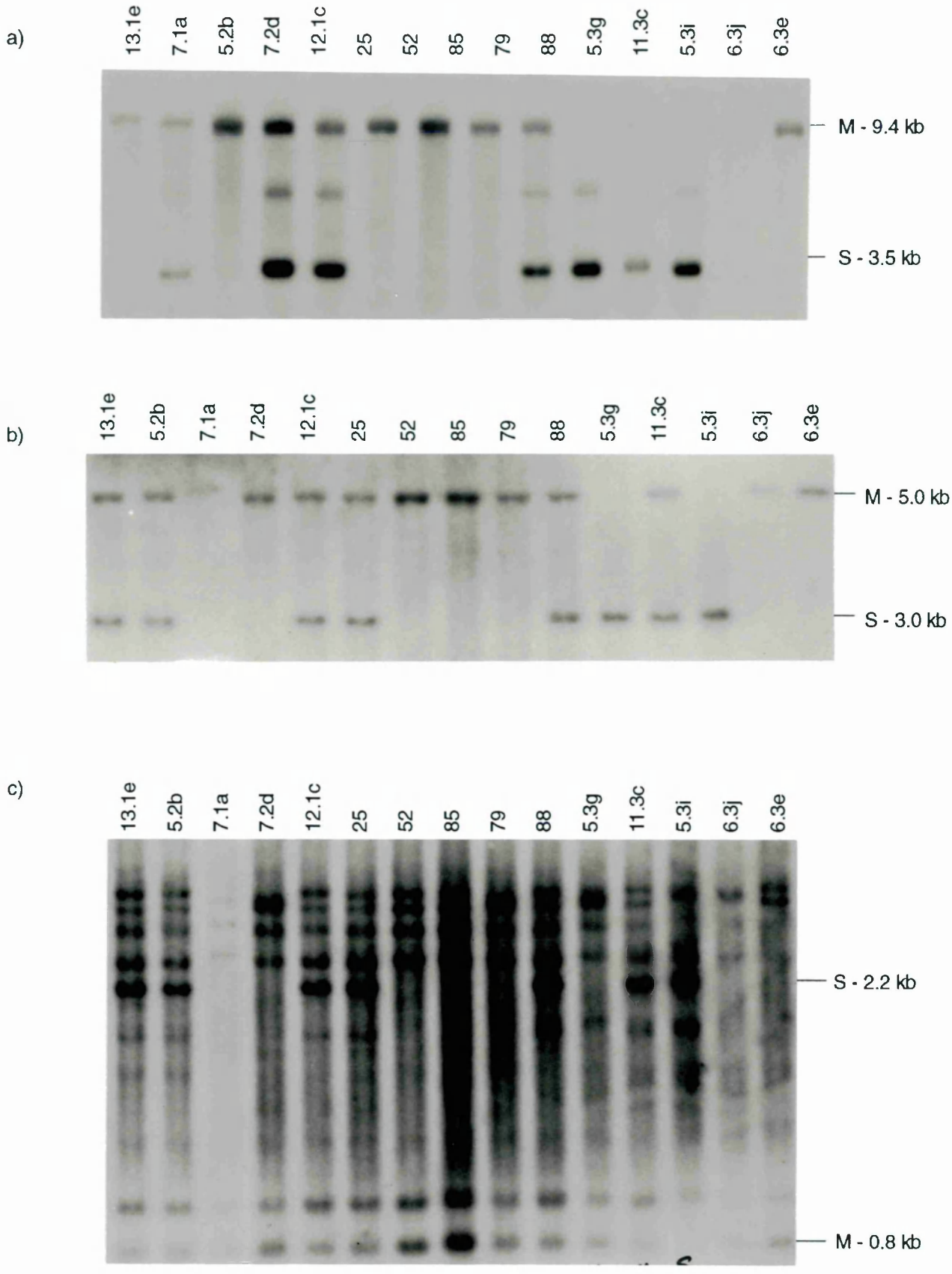


Figure 3.13

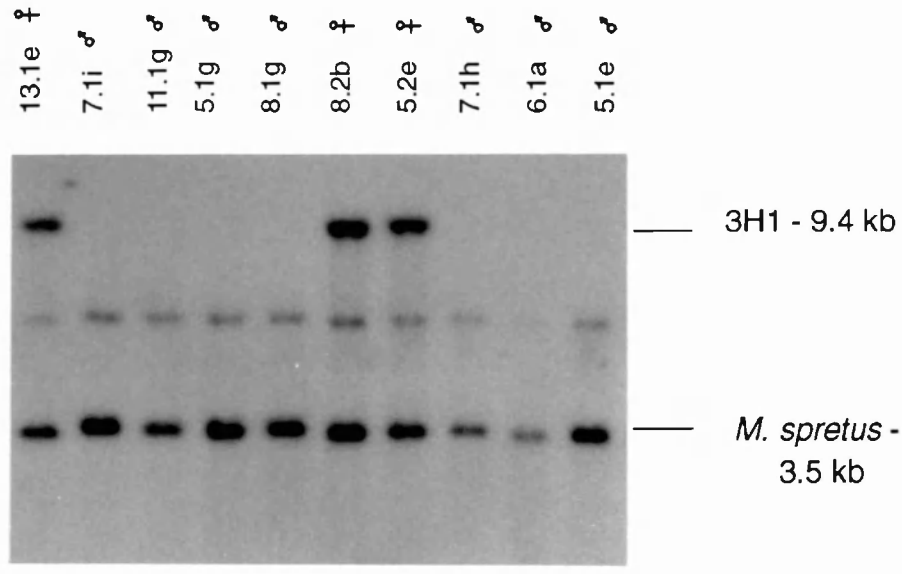
Hybridisation of probes which detect (a) *Nphl,* (b) *Syp* and (c) *Xkh* to representative offspring from backcrosses 1 and 2 which have recombination events between *DXWas70* and *Pfc*. The sizes of X-linked 3H1(M) and *Mus spretus* (S) bands are indicated. DNA were digested with the restiction enzymes PvuII for (a) *Nphl* analysis and TaqI for (b) *Syp* and (c) *Xkh* analyses.

(results not shown). Variant bands of 9.4kb and 3.5kb were detected in PvuII digested 3H1 and *M. spretus* DNA respectively. The *M. spretus* 3.5kb band was inherited in an X linked manner in male and female offspring of a 3H1 x *M. spretus* cross (results not shown). The mouse homologue of NPHL was found to lie proximal to *Otc* from analysis of the segregation pattern of the PvuII RFLV in the MXMP (Figure 3.14a and b). This indicated that *Nphl* lay in the *Syp* - *Tcfe3/Gata1* conserved segment. To confirm this, and the orientation of the segment discovered from the mapping *Syp*, RL-6 was hybridised to PvuII digested DNA prepared from the panel of 18 mice with recombination events between *DXWas70* and *Pfc*. No recombination events were identified between *Nphl*, *Syp* and *DXMit55/DXMit26* (Figure 3.13a and Table 3.6). Therefore, from this analysis, the order of *Nphl* and *Syp* could not be determined in the mouse, although *Nphl* would be expected to lie proximal to *Syp* from the order in man (Figure 1.7). The most proximal and distal loci which compose the GATA1 segment are now NPHL and GATA1. Evolutionary breakpoints must lie between DXS674/DXS679 and NPHL and between GATA1 and PFC on the human X chromosome.

3.5.4 The gene mutated in McLeod syndrome. XK.

The human XK gene encodes a membrane transport protein and is mutated in patients with McLeod syndrome, a disorder characterised by abnormalities in the neuromuscular and haematopoietic systems (Ho *et al.*, 1994). XK was known to lie on the human X chromosome approximately 200kb distal to CYBB, however, the location of the murine homologue, *Xkh*, was unknown. It was of particular interest to determine the position of *Xkh* on the comparative map since it could lie either

a)



b)

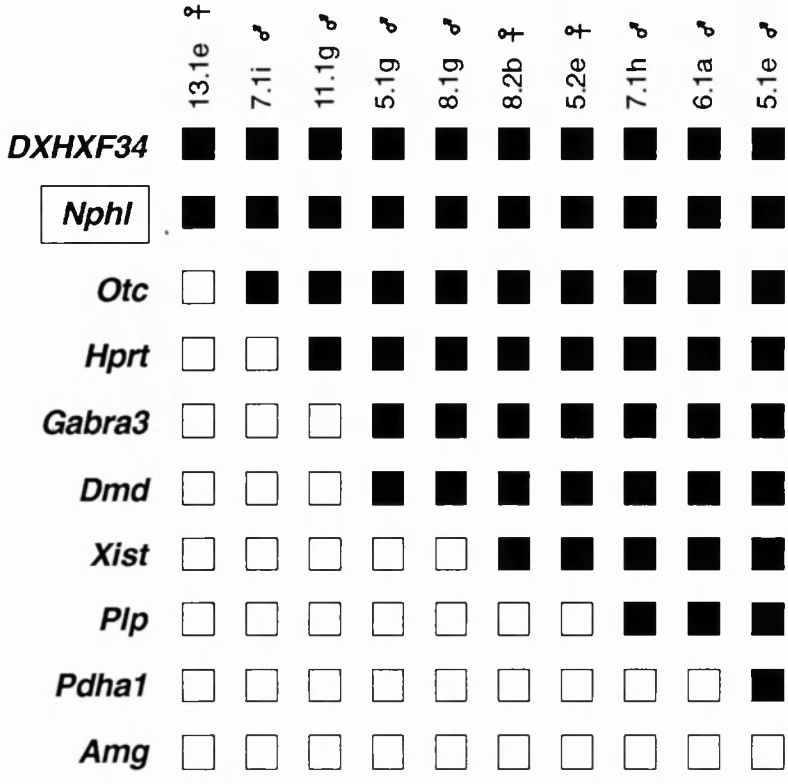


Figure 3.14

a) Pedigree analysis of probe RL-6, which detects the locus *Nphl,* on PvuII digested genomic DNA from the MXMP. M and S indicate the 3H1 and *Mus spretus* X-linked RFLVs.
(b) Summary of pedigree analysis data for *Nphl* on the MXMP. 3H1 alleles as represented as open sqaures and *Mus spretus* alleles as black squares. *Nphl* lies proximal to *Otc* on the mouse X chromosome.

proximal to *Cybb* or close to *Dmd* in the central region of the mouse X chromosome (Figure 1.5).

A 2 kb partial cDNA clone, from the human XK clone FSM5122 (Ho *et al.*, 1994), was hybridised to TaqI digested 3H1 and *M. spretus* DNAs and RFLVs of 0.8kb and 2.2kb respectively were observed. X-linkage of these bands was confirmed using the male and female offspring of a 3H1 x *M. spretus* cross (results not shown). Analysis of mice from the MXMP revealed that *Xkh* lies proximal to *Otc* on the mouse X chromosome (Figure 3.15a and b). Pedigree analysis of the 18 mice with recombination events between *DXWas70* and *Pfc* showed that *Xkh* cosegregated with the loci *Tcfe3* and *Gata1* on the mouse X chromosome (Figure 3.13c; Table 3.6). There was one recombinant (mouse 5.2i) which positioned *Xkh, Tcfe3* and *Gata1* proximal to *Cybb* and one recombinant (animal 5.3g) which positioned the three loci distal to *Nphl, Syp, DXMit26* and *DXMit55*. Therefore, *Xkh* did not lie in the central region of the mouse X chromosome close to *Dmd* but was proximal to *Cybb*. The evolutionary breakpoint must therefore lie between XK and DMD in man and the segment which was previously defined by CYBB-PFC can now be defined by XK-PFC. From these data the deduced order of loci and the position of evolutionary breakpoints (EBs) in the proximal region of the mouse X chromosome was (*DXWas70, DXHXF34*) - EB - (*Nphl, Syp, DXMit26, DXMit55*) - (*Tcfe3, Gata1,* EB, *Xkh*) - *Cybb* - *DXMit54* -*Pfc*.

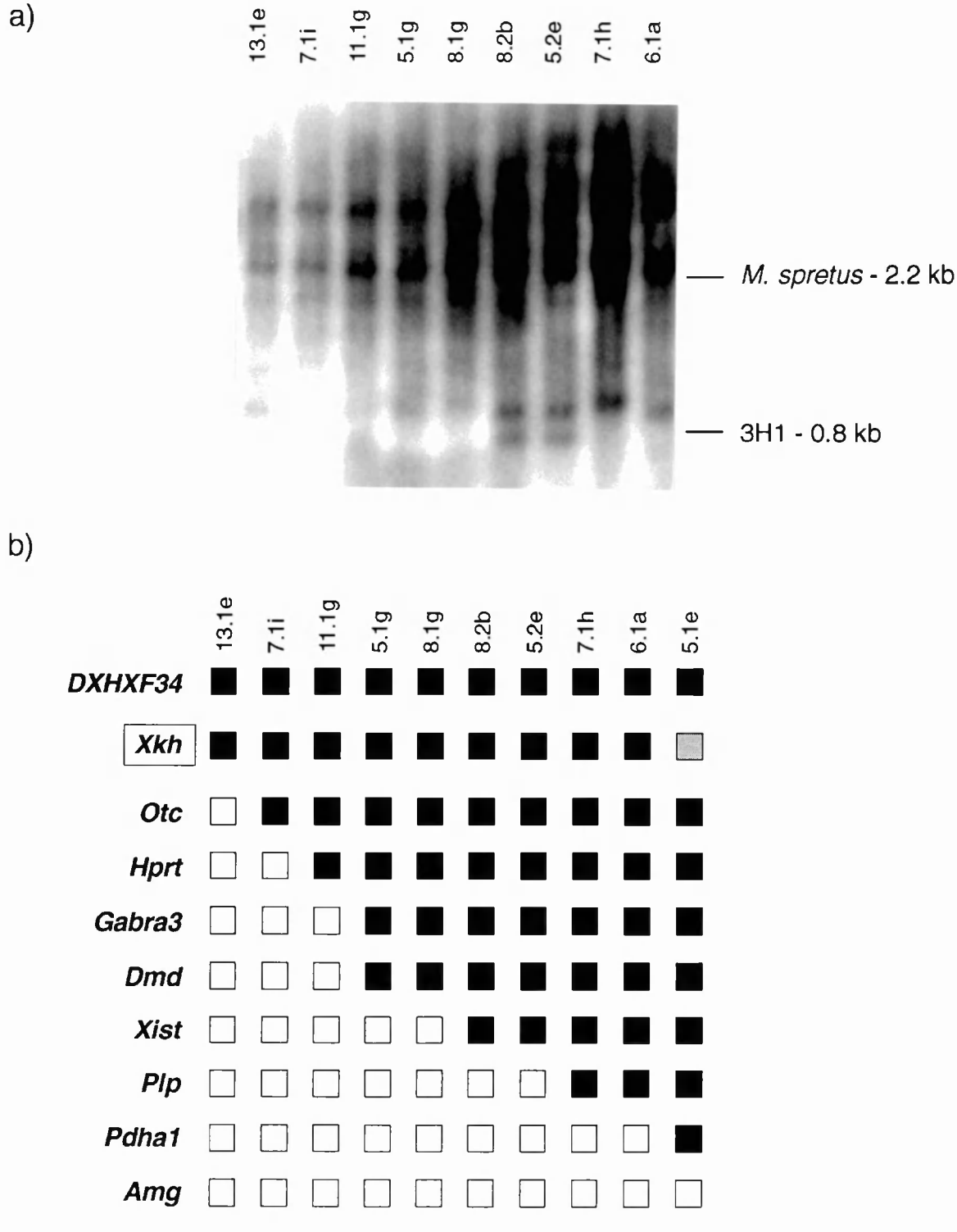


Figure 3.15

a) Pedigree analysis of *Xkh* on TaqI digested genomic DNA from the MXMP. 3H1 (M) and *Mus spretus* (S) X-linked bands are indicated. (b) Summary of pedigree analysis data for *Xkh* on the MXMP. 3H1 alleles as represented as open squares, *Mus spretus* alleles as black squares and alleles not scored as grey squares. *Xkh* lies proximal to *Otc* on the mouse X chromosome.

3.6 High Resolution Mapping in the Proximal Region of the Mouse X Chromosome using the European Collaborative Interspecific Backcross.

In an attempt to improve the resolution of the proximal mouse X map, additional recombination events in the region were identified in offspring from the large European Collaborative Interspecific backcross (EUCIB), (section 3.6.2). These were analysed for key loci. (section 3.6.3).

3.6.1 The European Collaborative Interspecific Backcross (EUCIB)

(C57BL/6 x *Mus spretus*) F_1 ♀ x C57BL/6 ♂ or *M. spretus* ♂.

(C57BL/6 x *M. spretus*) F_1 females were backcrossed to either C57BL/6 or *M. spretus* males. Backcross offspring were produced at the Institut Pasteur, France and the MRC Clinical Research Centre, London, UK. *M. spretus* mice used in the cross were from two different colonies; SPR for the UK bred mice and SEG/Pas for the French mice (Breen *et al.*, 1994). Backcross offspring are named according to where they were derived (L for London and P for Paris) and whether they were produced from a backcross to C57BL/6 (B) or *M. spretus* (S). For example, LS718 is mouse 718 bred at the MRC Clinical Research Centre, London, UK to *M. spretus* males (SPR). 982 offspring, 501 from the backcross to *Mus spretus* and 481 from the backcross to C57BL/6, were analysed for anchor loci spaced approximately 20cM apart on each chromosome and panels of mice with recombination events in the intervals between two anchor markers were identified. Four anchor markers were used to identify recombination events on the X chromosome: *DXWas70, DXMit8, Plp* and *DXHar2* (Breen *et al.,* 1994).

3.6.2 The identification of EUCIB offspring with recombination events in the proximal region of the mouse X chromosome.

To identify offspring with recombination events useful for high resolution mapping in the proximal region of the mouse X chromosome, all backcross offspring were analysed for *DXMit54*, a simple sequence repeat marker thought to lie approximately 4cM distal to *DXWas70* (Herman *et al.*, 1994; Dietrich *et al.*, 1994). PCR analysis of *DXMit54* was carried out with the assistance of P. Williamson (MRC Radiobiology Unit). 38 offspring were found to have recombination events in the interval between *DXWas70* and *DXMit54* from a total of 723 scored for both loci (Table 3.8; data not shown). Therefore, the genetic distance between the two loci was calculated as 5.3 ± 0.8 cM. Data from analysis of all the backcross animals was deposited in the *Mbx* database which is available through the UK Human Genome Mapping Project (HGMP) Resource Centre on the World Wide Web: http://www.hgmp.mrc.ac.uk/local-data/mbx/Mbx-Homepage.html (Brown, 1995).

3.6.3 Pedigree analysis of EUCIB offspring with recombination events between *DXWas70* and *DXMit54*.

Aliquots of DNA, for RFLV and APV analysis, from the 38 EUCIB offspring with recombination events between *DXWas70* and *DXMit54* were obtained from the HGMP Resource Centre, Cambridge. The *DXWas70* and *DXMit54* scorings were repeated to ensure they were from the same backcross animals and had recombination events between the two loci (Figure 3.16; *DXMit54* scorings not shown). Several discrepancies were observed between the previous scorings and the scorings on these new aliquots of DNA for both *DXWas70* and *DXMit54* (shown in brackets in Table

Table 3.8 Scorings of EUCIB Offspring with Recombination Events between *DXWas70* and *DXMit54*.

Animal	Sex	*DXWas70*	*DXMit54*
LB061	f	S[b]	M
LB202	m	M	S
LB304	m	S	M
LB390	f	M	S
LB395	f	S	M
LB396	f	S	M
LB397	f	M (S)	S
LB426	f	S	M
LB522	f	S	M
LB604	m	M	S
LS033	m	M	S
LS095	f	S (M)	M
LS151	f	M	S
LS214	f	M	S
LS224	f	M	S
LS228	m	M	S (M)
LS288	f	M	S (M)
LS324	f	S	M (S)
LS327	f	M	S

Animal	Sex	*DXWas70*	*DXMit54*
LS548	f	S	M
LS562	m	S	M
LS568	m	M	S
LS579	m	M	S
LS589	f	M	S
LS718	f	S (M)	M
LS719	m	M (S)	S
LS760F	f	M	S
LS808F	f	M	S
LS815F	f	M	S
LS855F	m	M	S
LS867F	m	M	S
PB003	f	S	M
PB020	f	S	M
PB123	f	M	S (M)
PB214	f	M	S (M)
PB225	f	M	S
PS069	f	S	M
PS070	f	S	M

[a] Sex: f = female; m = male.
[b] M or S for the *M. musculus* (C57BL/6) and *M.spretus* allele respectively inherited from the F_1 female parent of the backcross offspring.
Scorings shown in parentheses indicate the allele observed after retyping the new DNA sample.

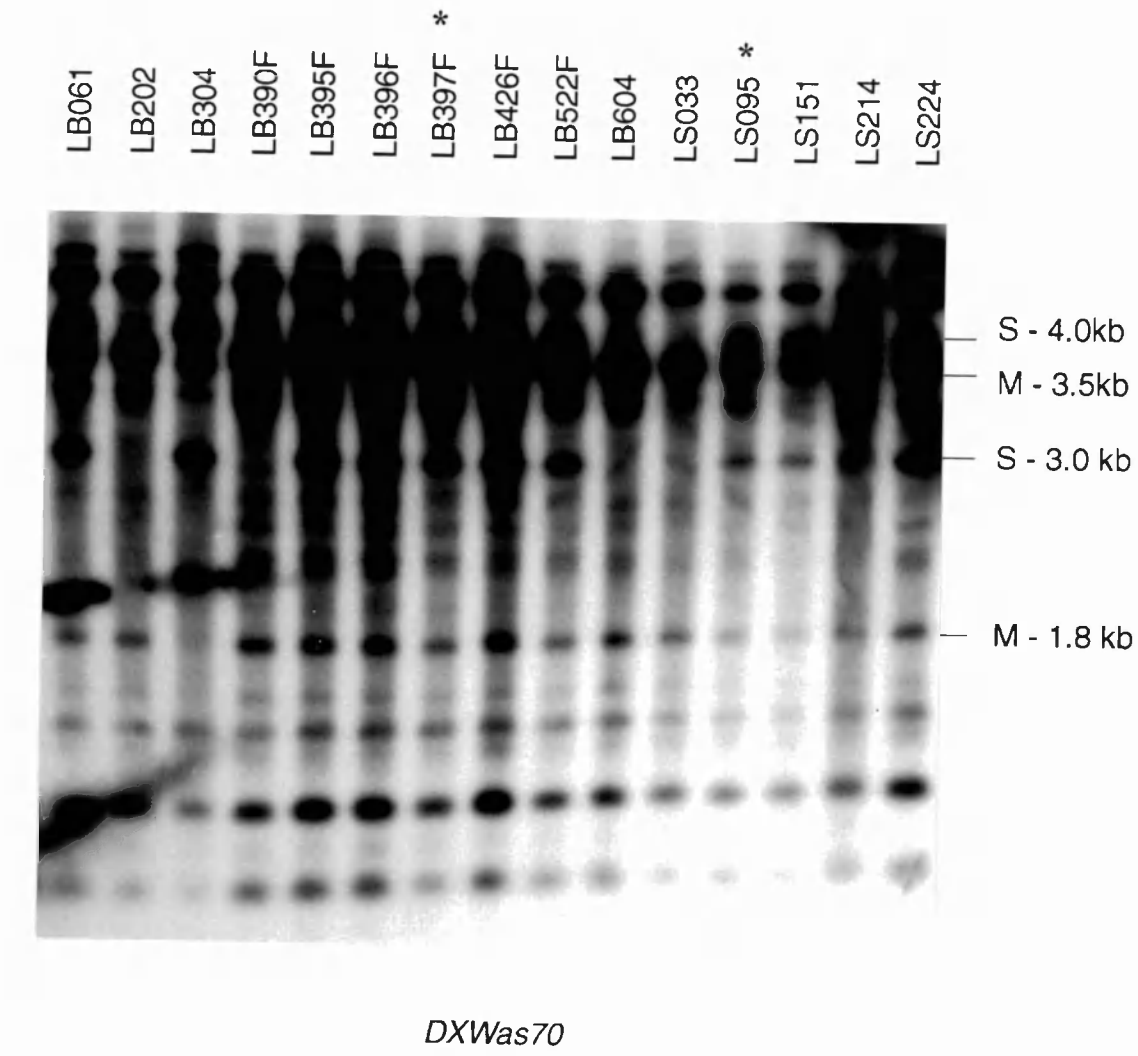
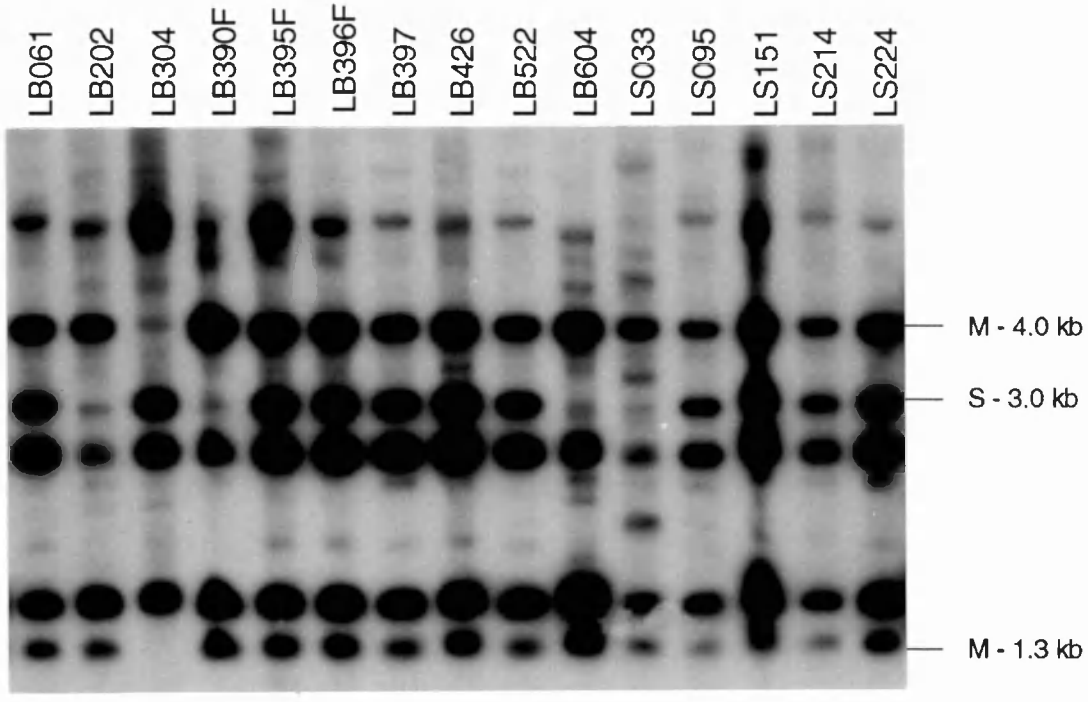


Figure 3.16

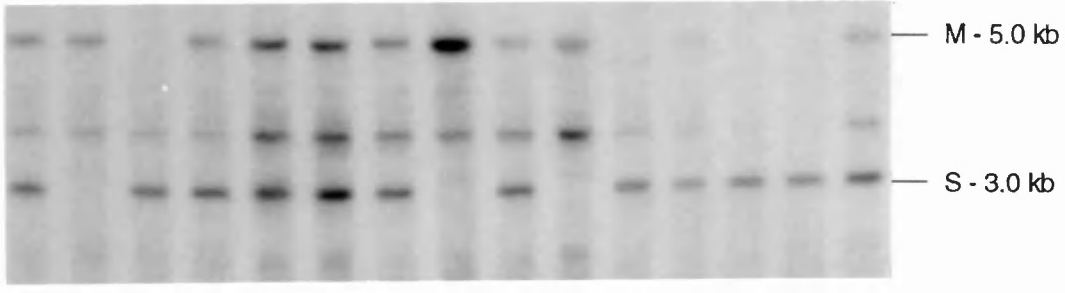
Representative *DXWas70* scorings of DNAs from EUCIB offspring after digestion with TaqI. M and S indicate the sizes of 3H1 and *M. spretus* X-linked bands respectively. Animals LB397F and LS095 marked with * were removed from the analysis as they were found not to have recombination events between *DXWas70* and *DXMit54*. *DXMit54* scorings are not shown.

3.8). These discrepancies could have been due to errors in the typing of the samples, or errors in DNA distribution. The *DXMit54* analysis was repeated on both the old and the new five DNA samples with discordant scorings at this locus and the scorings confirmed (results not shown). Therefore, in each case the pair of DNA samples (animals LS228, LS288, LS324, PB123 and PB214) could not have originated from the same mouse. The samples of the DNA from 4 mice with discordant *DXWas70* scorings (animals LB397, LS095, LS718, LS719) could not be reanalysed as the initial scorings were achieved by hybridisation of the *DXWas70* probe onto Southern blots prepared by Généthon (Paris). These hybridisations were carried out in our laboratory, therefore the original autoradiographs were reanalysed and the scorings confirmed. As a result of rescoring, 9 from the 38 original mice did not appear to have recombination events in the interval between *DXWas70* and *DXMit54* and were not useful in this analysis. Variants at *DXHXF34, Nphl, Syp, DXMit55, DXMit26, Tcfe3, Gata1* and *Xk* were analysed in the remaining 29 animals with recombination events between *DXWas70* and *DXMit54* (Figure 3.17 and Table 3.9). The order of several loci which co-segregated in the previous backcross, were obtained from studies of EUCIB offspring. *Nphl* was positioned proximal to *Syp, DXMit26* and *DXMit55* by 4 recombination events, *Tcfe3* was positioned proximal to *Gata1* by one recombination event and *Xkh* distal to *Gata1* by 3 recombination events. The order of loci known to belong to the conserved segment which previously was only defined by *Gata1* is: *Nphl - Syp - Tcfe3 - Gata1*. This is in agreement with the order of these loci on the human X chromosome (Figure 1.7; Nelson *et al.*, 1995). *DXWas70* cosegregated with *DXHXF34* in all offspring analysed which indicated that *DXHXF34* could lie either proximal, or distal but very close to *DXWas70.*

a) *DXHXF34*



b) *Syp*



c) *Gata1*

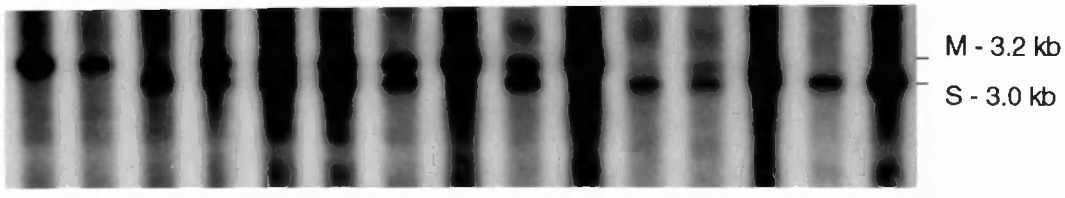
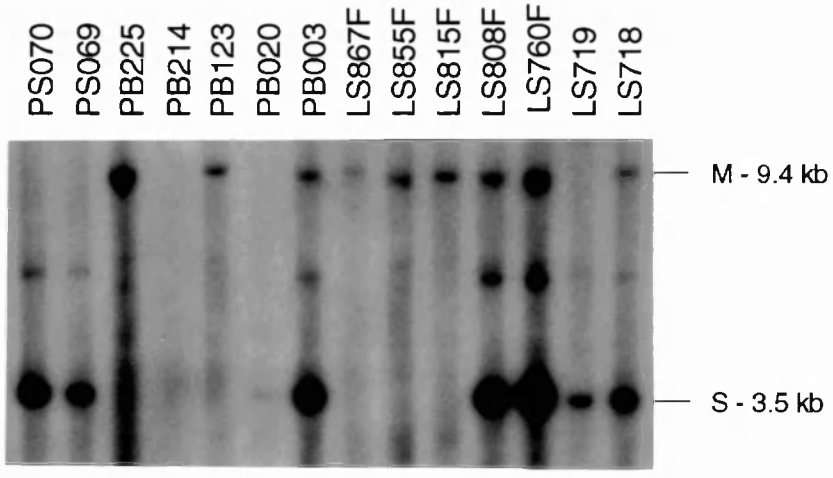


Figure 3.17

Representative EUCIB animals with recombination events between *DXWas70* and *DXMit54* typed for (a) *DXHXF34,* (b) *Syp,* (c) *Gata1,* (d) *Nphl,* (e) *Xkh,* (f) *DXMit55* (g) *DXMit26* and (h) *Tcfe3.*Samples were digested with EcoRI for *DXHXF34* and *Gata1* analysis (a and c), TaqI for *Syp* and *Xkh* analysis (b and e) and PvuII for *Nphl* analysis (d). M and S indicate the X-linked RFLVs or APVs at each locus.

d) *NphI*



e) *XkH*

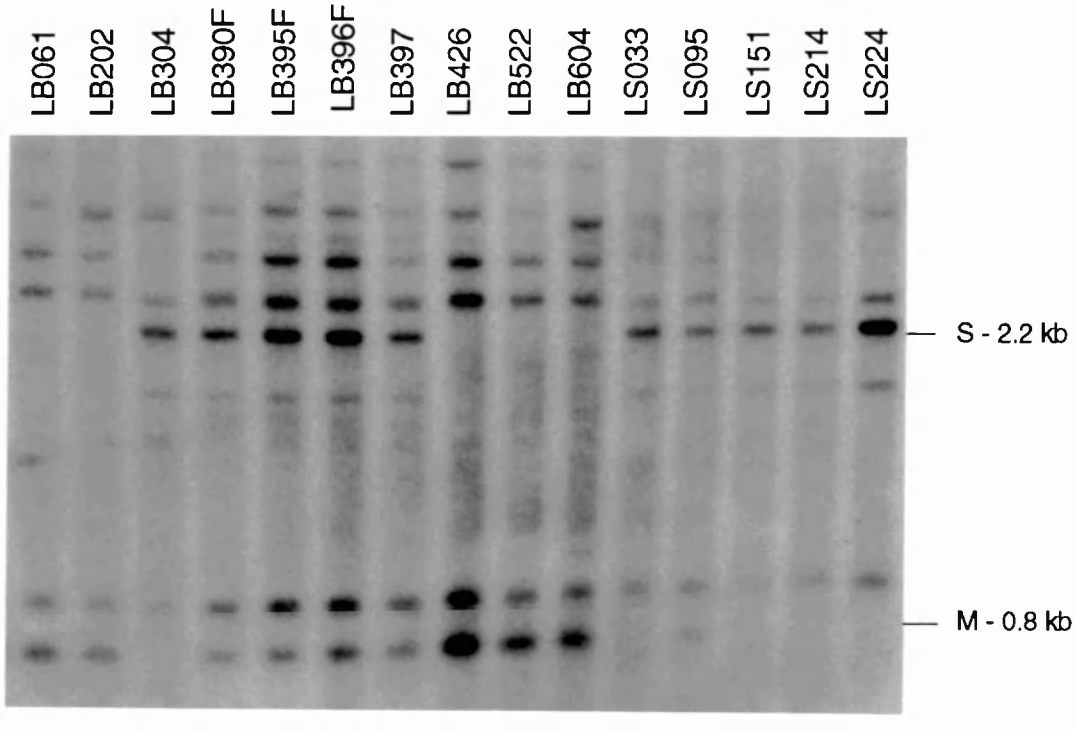
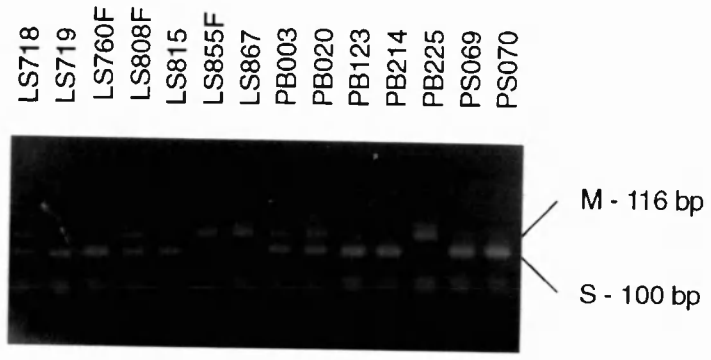
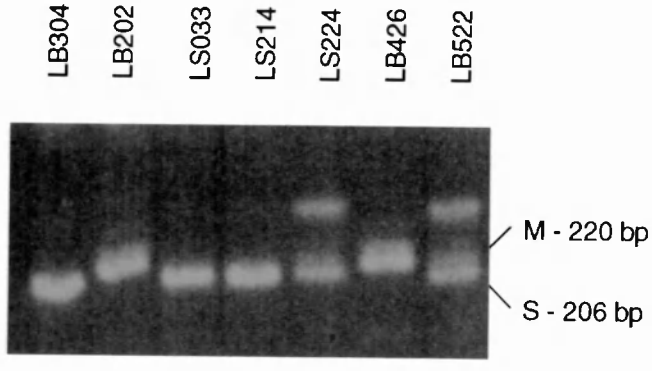


Figure 3.17 - continued

f) *DXMit55*



g) *DXMit26*



h) *Tcfe3*

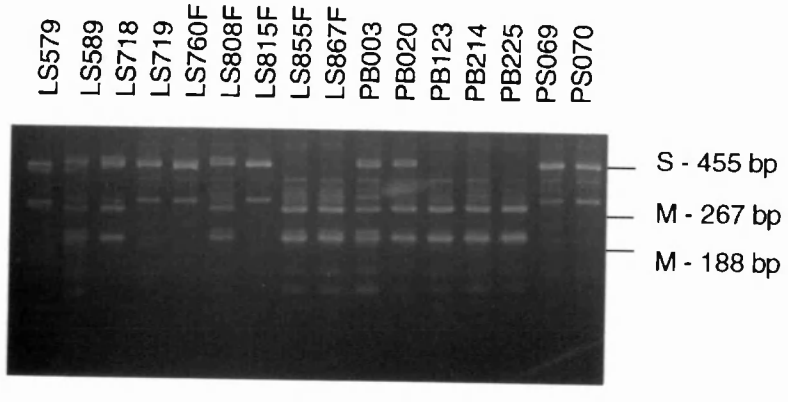


Figure 3.17 continued

Table 3.9 Scorings of Offspring from EUCIB with Recombination Events between *DXWas70* and *DXMit54*.

Animal	LB061	LB202	LB304	LB390F	LB395F	LB396F	LB426	LB522F	LB604F	LS033	LS151	LS214	LS224	LS327	LS548	LS562	LS568	LS579	LS589	LS760F	LS808F	LS815F	LS855F	LS867F	PB003	PB020	PB225	PS069	PS070
Sex	f	m	m	f	f	f	f	f	m	m	f	f	f	f	m	m	m	m	f	f	f	m	m	m	f	f	f	f	f
DXWas70	S	M	S	M	S	S	S	S	M	M	M	M	M	M	S	S	M	M	M	M	M	M	M	M	S	S	M	S	S
DXHXF34	S	M	S	M	S	S	S	S	M	M	M	M	M	M	S	S	M	M	nt	M	M	M	M	M	S	S	M	S	S
				x						x		x				x		x											
Nphl	S	nt	S	S	S	S	S	S	nt	S	M	nt	M	nt	nt	nt	nt	S	nt	M	M	M	M	M	S	nt	M	S	S
							x				x									x		x							
DXMit55	S	M	S	S	S	S	M	S	M	S	S	S	M	M	S	M	M	S	M	S	M	S	M	M	S	S	M	S	S
Syp	S	M	S	S	S	S	M	S	M	S	S	S	M	nt	nt	nt	nt	nt	nt	S	M	S	M	M	nt	S	M	S	S
DXMit26	S	M	S	S	S	S	M	S	M	S	S	S	M	M	S	M	M	S	M	S	M	S	M	M	S	S	M	nt	S
	x												x	x															
Tcfe3	M	M	S	S	S	S	M	S	M	S	S	S	S	S	S	M	M	S	M	S	M	S	M	M	S	S	M	S	S
																								x					
Gata1	M	M	S	S	S	S	M	S	M	S	S	S	S	S	S	nt	nt	S	nt	S	M	S	M	S	S	S	M	S	S
								x											x								x	x	
Xkh	M	M	S	S	S	S	M	M	M	S	S	S	S	nt	nt	M	M	nt	nt	S	M	S	M	S	nt	S	S	M	S
		x	x		x	x			x						x		x				x		x		x	x			x
DXMit54	M	S	M	S	M	M	M	M	S	S	S	S	S	S	M	M	S	S	S	S	S	S	S	S	M	M	S	M	M

M indicates the *M. musculus* (3H1) allele and S the *M. spretus* allele inherited from the F1 female parent of the backcross offspring; nt indicates animal not tested; x indicates the position of a recombination event; f indicates the sex of the animal is female and m indicates the sex of the animal is male.

There were two factors concerning the EUCIB offspring in this study which complicated the calculation of genetic distances between markers:

(1) Only 29 DNA samples from offspring with recombination events between *DXWas70* and *DXMit54* were available for pedigree analysis from the total of 38 identified by scoring the whole backcross.

(2) The amount of DNA was limited therefore, Southern blot scorings which were ambiguous could not be repeated, Thus, many mice were not scored for all loci.

As 9 offspring were eventually found to be non-recombinant, it could be assumed that another 9 animals were recombinant but not identified from the whole backcross, therefore, the genetic distance of 5.3cM between *DXWas70* and *DXMit54* was justified. As all 29 offspring were scored for *Tcfe3* by PCR analysis, distances between *Tcfe3* and the anchor markers *DXWas70* and *DXMit54* were estimated by calculating the proportion of the backcross actually scored. For example, 12 offspring carried recombination events between *DXWas70* and *Tcfe3* from the 29 scored, therefore, the expected number of recombinants, if all of the 38 offspring had been scored, would be approximately 12/29 x 38. The genetic distance between *DXWas70* and *Tcfe3* was therefore calculated as 2.2 cM (12/29 x 38/723) and in the same way the distance between *Tcfe3* and *DXMit54* was calculated as 3.1cM (17/29 x 38/723).

Not all 29 available recombinants were scored at *Nphl, Syp, Gata1* and *Xkh* (see factor 2 above). Distances were estimated between these loci by calculating the proportion of the offspring scored which had recombination events in the *DXWas70/Tcfe3* or the *Tcfe3/DXMit54* intervals. For example, 3 offspring carried recombination events between *DXWas70* and *Nphl* from 9 scored for both loci in the *DXWas70-Tcfe3*

interval, therefore the expected number of recombination events would be approximately 3/9 x 12. The distance was then calculated as 0.7cM in the same way as for the *DXWas70-Tcfe3* distance above (3/9 x 12/29 x 38/723).

From pedigree analysis of EUCIB offspring the order and approximate genetic distances, in centimorgans, between loci in the proximal region of the mouse X chromosome was calculated as (*DXWas70, DXHXF34*) - (0.7) - *Nphl* - (1.0) - (*Syp, DXMit55, DXMit26*) - (0.5) - *Tcfe3* - (0.2) - *Gata1* - (0.7) - *Xkh* - (2.2) - *DXMit54*. The refined order of conserved loci and the position of evolutionary breakpoints (EBs) from analysis of recombination events in EUCIB offspring was *DXHXF34* - EB - *Nphl* - *Syp* - *Tcfe3* - *Gata1* - EB - *Xkh*.

3.7 Summary and Conclusions.

To improve the comparative map of the human and mouse X chromosomes, the mouse map position of several genes and conserved sequences from human proximal Xp have been determined. In particular, an attempt was made to position those loci which would provide maximum information for the comparative map; i.e. those which lie close to evolutionary breakpoints between conserved segments. Towards this aim several novel genes and conserved sequences were positioned on the genetic map of the mouse with respect to previously mapped key loci. These include the mouse homologues of XK, the gene involved in McLeod syndrome; NPHL, the gene mutated in patients with Dent's disease; SYP, the synaptophysin gene and TFE3, the gene which encodes the transcription factor enhancer 3. Offspring from *M. musculus* x *M. spretus* backcrosses were analysed to identify two panels of mice, with recombination

events between *DXWas70* and *Pfc* and between *Plp* and *Pdha1*, in the proximal and distal regions of the mouse X chromosome respectively. RFLVs or APVs between *M. musculus* and *M. spretus* were identified at each locus and pedigree analysis of the appropriate panel carried out. Further offspring, recombinant in the *DXWas70-Pfc* interval, were identified from the large European Collaborative interspecific backcross (EUCIB) for high resolution mapping of loci in the proximal region of the mouse X chromosome. In addition, several microsatellites, from each of the two regions, were also positioned on the genetic map. These proved useful for the initiation of a clone contig of one of the regions and provide a resource for the fine mapping of mouse mutant phenotypes.

The order of loci in the proximal region of the mouse X chromosome determined from backcross mapping studies described in this Chapter is:
centromere - (*DXWas70, DXHXF34*) - *Nphl* - (*Syp, DXMit26, DXMit55*) - *Tcfe3* - *Gata1* - *Xkh* - *Cybb* - *DXMit54* - *Pfc* - telomere.

The order of loci in the distal region of the mouse X chromosome determined from these studies is centromere - *Plp* - *DXMit35* - *DXMit34* - (*DXHXS674, DXHXS679*) - *DXMit28* - *Pdha1*- telomere.

The conserved segments which compose human proximal Xp can be redefined by the most proximal and distal loci they contain (conserved segments 1, 2 and 3) or by a single locus (conserved segment 4) as follows: 1) XK-PFC; 2) GATA1-NPHL; 3) DXS674/DXS679-ALAS2 and 4) DXF34. These data improve the comparative map

of human proximal Xp and their implications to the identification of mouse models for human inherited disease are discussed in Chapter 5.

The construction of a high resolution genetic map of the proximal region of the mouse X chromosome proved invaluable in the next stage of the project, the initiation of a YAC contig covering the region, which is described in Chapter 4.

Genetic Analysis of the X Chromosomes of Man and Mouse

Helen J. Blair

M.Phil

July 1996

Volume 4 of 5

Chapter 4

Physical Characterisation of the Regions on the Proximal Mouse X Chromosome Homologous to Proximal Xp in Man.

4.1 Introduction

The order of loci (*DXWas70, DXHXF34*) -*Nphl* - (*Syp, DXMit55, DXMit26*) - *Tcfe3* - *Gata1* - *Xkh* - *Cybb*, in a region of approximately 3-4 cM in the proximal mouse X chromosome, was established from the genetic mapping studies described in Chapter 3. This region of the mouse X chromosome shares homology with three conserved segments on the human X chromosome proximal short arm. By studying the human X chromosome map (Nelson *et al.,* 1995) and comparing the positions of genes and conserved sequences in both species, it can be concluded that one evolutionary breakpoint must lie between *DXHXF34* and *Nphl* and another between *Gata1* and *Xkh* (see Chapter 3). Each of these intervals spans approximately 0.7cM on the mouse genetic map. Comparison of genetic and physical distances in other regions of the mouse X chromosome has indicated that 1cM is equivalent to approximately 2Mb (Chatterjee *et al.*, 1994; Hamvas *et al.*, 1993); therefore, 0.7cM should represent a physical distance of less than 2Mb. Although this distance is small, there may have been other, as yet undetected, rearrangements within this region since the human and mouse X chromosomes evolved from an ancestral X chromosome.

The *DXHXF34-Cybb* interval on the mouse X chromosome contains the gene responsible for the tattered (*Td*) phenotype and is likely to contain the gene mutated

in scurfy (*sf*) mice (Merrell *et al.*, 1995; Blair, P.J. *et al.*, 1994). Isolation and study of these genes would provide information about the underlying defects which cause the mutant phenotypes and indicate if they are relevant models for human inherited disorders associated with human proximal Xp.

An attempt was therefore made to construct a physical map of the *DXHXF34-Cybb* interval on the mouse X chromosome. Such a map would help to determine if there have been further rearrangements in the human-mouse X chromosome comparative map, by providing accurate sizes for the regions which include evolutionary breakpoints. A complete cloned contig would allow homologues of additional human genes, e.g. those which lie close to the evolutionary breakpoints, to be readily positioned in the mouse. It would also provide a resource for the cloning of sequences at the evolutionary breakpoints themselves, and for the isolation of the genes responsible for the *Td* and *sf* phenotypes.

DNA fragments of up to 50kb in cosmid, and 100kb in P1 bacteriophage vectors can be cloned in *E.coli*-based systems and have been used to build physical maps of several regions of the mouse genome (Wicking and Williamson, 1991). Because the *DXHXF34-Cybb* interval spans several megabases many P1 clones would be required to build a complete contig. Instead, YAC clones, which are able to carry DNA inserts of up to, and greater than 1Mb (Schlessinger, 1990), were chosen because it was assumed that relatively few clones would be required to build a complete contig of the region.

4.1.1 Yeast artificial chromosomes (YACs).

YAC vectors, developed by Burke, Carle and Olson (1987), allow the cloning of large DNA inserts between two vector arms which contain all the sequences necessary for the stable maintenance of the YAC as an artificial chromosome within yeast cells. These sequences include ARS1, an autonomous replication sequence, CEN4, which functions as a centromere, and TEL, telomere sequences at the end of each arm. The genes TRP1 (tryptophan) and URA3 (uracil) are also present to allow selection of yeast cells containing the YAC (Burke *et al.*, 1987). YAC libraries have been constructed from several different eukaryotic DNA sources including the mouse (Larin *et al.*, 1991); general cloning procedures are represented in Figure 4.1. In addition to the large insert size, the YAC cloning system offers other features which facilitate the construction of physical maps:

(1) The presence of selectable marker sequences and a bacterial origin of replication in one arm which allows the isolation of end sequences by plasmid rescue (Schlessinger, 1990). End sequences can be used for YAC library screening and to identify overlapping clones.

(2) The absence of rare cutter enzyme sites and the presence of unique sequences in each arm which allows a restriction map of the insert DNA to be constructed easily by indirect end-label mapping (Burke *et al.*, 1987).

(3) The evidence that the YAC system, which employs the replication machinery of yeast, may allow the cloning of some sequences which are unclonable in bacterial systems (Coulson *et al.*, 1988).

Despite the advantages of the YAC mapping system problems can occur from clones containing more than one YAC or from the co-ligation, in the same YAC, of two or

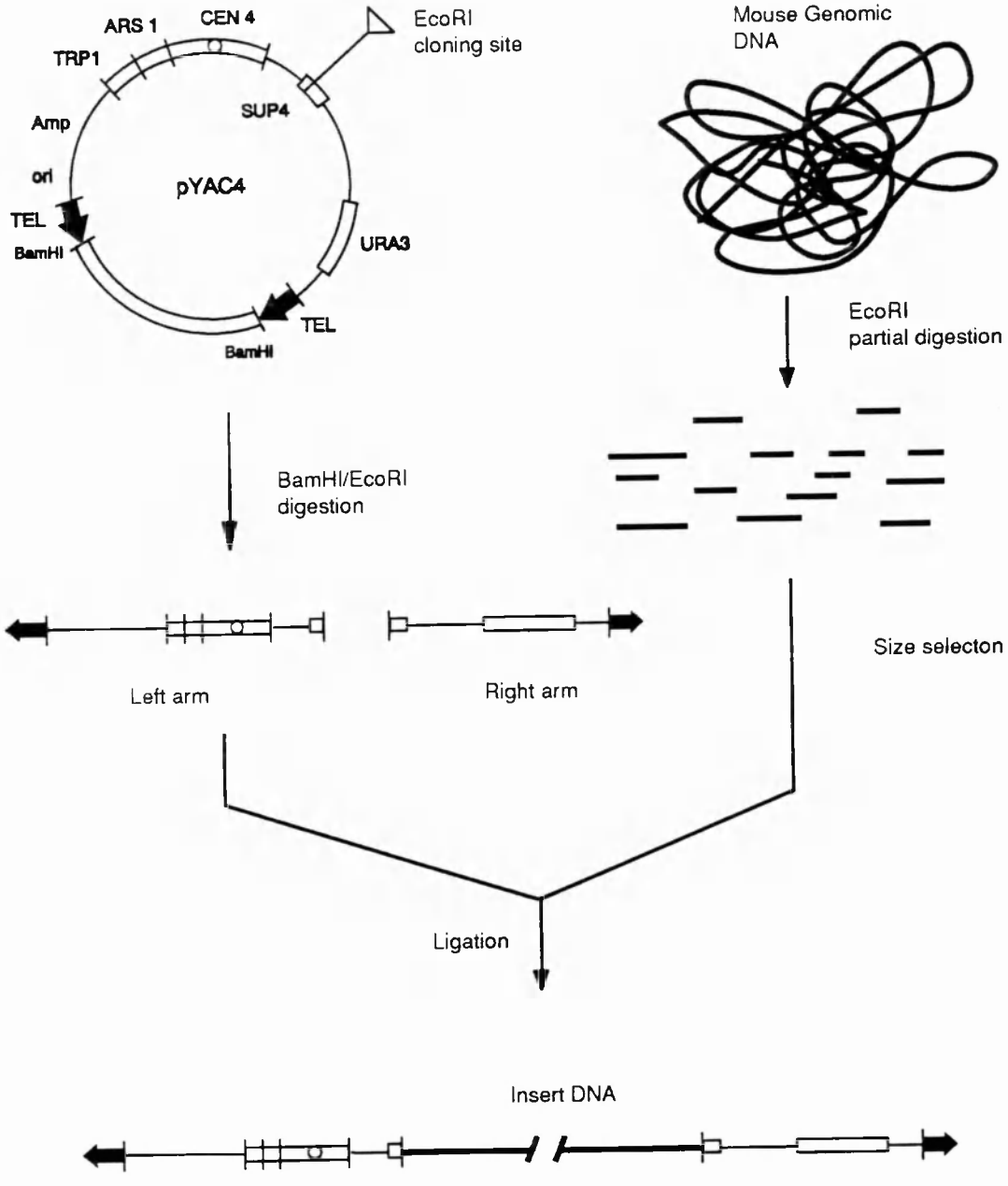


Figure 4.1
Yeast artificial chromosome cloning system. Adapted from Burke, *et al.*, 1987 and Larin, *et al.*, 1991. Restriction sites used for the construction of YACs, structures of the restricted vector as well as the structure of the final YAC are shown. Single lines in the pYAC4 vector represent pBR322-derived sequences whereas boxed regions are yeast-derived sequences.

more different DNA segments. The frequency of such events appears to vary according to the YAC library and the region of the genome. Another disadvantage of the YAC system is the inability to readily isolate large quantities of pure YAC DNA from the yeast genomic DNA, therefore many of the techniques involved in the analysis of YACs are carried out in the presence of yeast sequences.

4.1.2 General strategy.

The following sequential approach was adopted for the construction of the YAC contig:

1) YACs isolated with markers in the *DXHXF34-Cybb* interval were obtained.

2) YAC DNAs were prepared and PCR or Southern blot analysis performed to confirm the presence of the marker with which they were isolated (4.2)

3) The approximate size of each YAC was determined using pulsed-field gel electrophoresis (4.3.1).

4) The proximal X chromosome content of each YAC was determined by fluorescent *in situ* hybridisation of YAC DNA to mouse metaphase spreads (4.3.2).

5) The presence of other markers which were known to lie in the *DXHXF34-Cybb* interval was determined by PCR and/or Southern blot analysis of YAC DNA (4.3.3).

6) Using the information from 2-5 above a YAC contig of the region was constructed (4.5).

4.2 Details of YACs from the Proximal Region of the Mouse X Chromosome.

YACs positive for *DXHXF34, DXMit26, Gata1* and *Cybb* were made available for these studies. Details of all the YACs are given in Table 4.1.

4.2.1 *DXHXF34* YAC clones.

Six *DXHXF34* positive YAC clones were isolated from the Imperial Cancer Research Fund (ICRF) Reference Library 902 mouse (C3H) YAC library (Larin *et al.*, 1991). The human cDNA clone, pHFB7.2A2, which detects the locus *DXHXF34* (S.H. Laval, personal communication), and four other probes not relevant to this study, were co-hybridised to YAC colony filters (V. Reed, MRC Radiobiology Unit). The six *DXHXF34* clones were identified by hybridisation of pHFB7.2A2 and the genomic probe (pH3-7) to YAC DNA digested with EcoRI (Figure 4.2). Each detects four bands of approximately 4.0, 2.5, 1.8 and 1.3kb in 3H1 DNA digested with EcoRI. Previously only the 4.0kb and 1.3kb bands had been shown to lie in the proximal region of the mouse X chromosome (Laval and Boyd, 1993a). The presence of the 2.5kb EcoRI band in the YAC DNAs could not be established due to the frequent presence of a similar size band in AB1380 yeast strain DNA. As ICRFy902F0751 contained the 4.0 and 1.3kb EcoRI bands and ICRFy902C069 contained the 4.0 and 1.8kb EcoRI bands (Table 4.2) all three bands must lie within the combined region covered by these two YACs. The genomic organisation of *DXHXF34* in the mouse therefore appears to be similar to that of DXF34 on the human X chromosome because it has previously been shown that the family of sequences detected by pH3-7 in man is found in the same 450kb YAC (Reed *et al.*, 1994).

Table 4.1 Background Details of YAC Clones Characterised in this Study

YAC Reference	Library[a]	Locus Screened	Method of Isolation[b]	Laboratory of Isolation[c]
ICRFy902F0651	ICRF	*DXHXF34*	F	RBU
ICRFy902F0751	ICRF	*DXHXF34*	F	RBU
ICRFy902E0547	ICRF	*DXHXF34*	F	RBU
ICRFy902C069	ICRF	*DXHXF34*	F	RBU
ICRFy902E1216	ICRF	*DXHXF34*	F	RBU
ICRFy902H0127	ICRF	*DXHXF34*	F	RBU
I114H3	ICRF	*DXMit26*	P	SMH
FDDB10	MIT	*Gata1*	P	UM
FDBE3	MIT	*Gata1*	P	UM
F192	ICRF	*Cybb*	F	IMM
C8127	ICRF	*Cybb*	F	IMM
E840	ICRF	*Cybb*	F	IMM

Note:
a : ICRF -Imperial Cancer Research Fund; MIT- Massachusetts Institute of Technology.
b : F - Filter hybridisation; P - PCR
c : RBU - Radiobiology Unit; SMH - St. Mary's Hospital Medical School; UM - University of Michigan Medical School; IMM - Institute of Molecular Medicine.

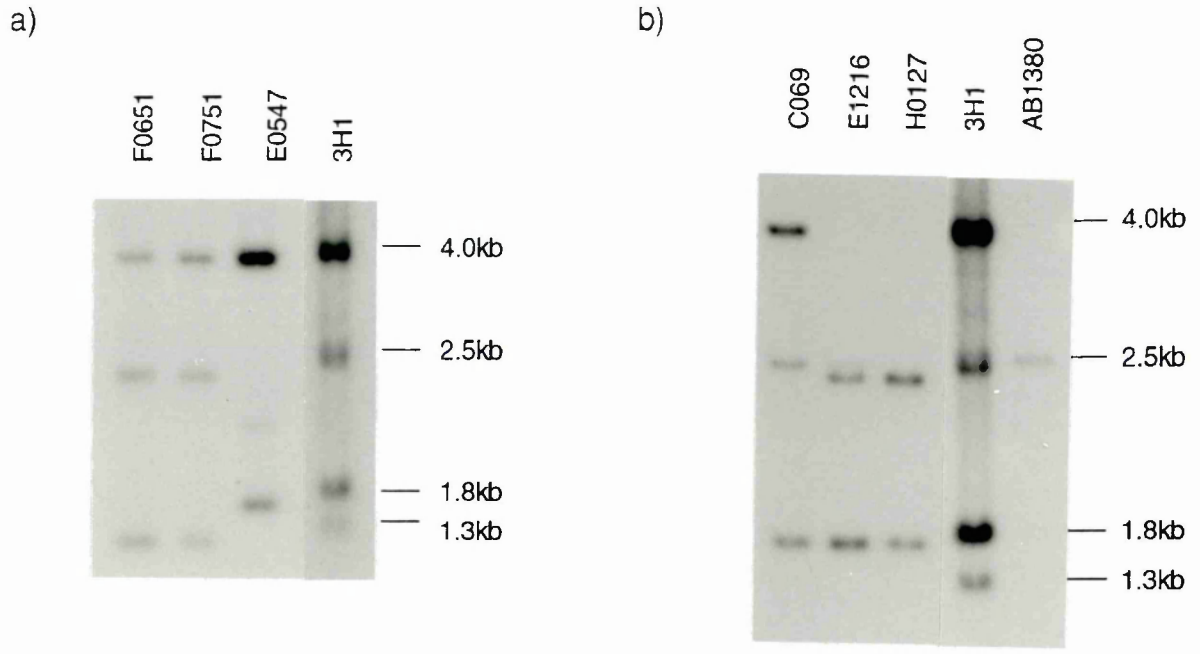


Figure 4.2
Hybridisation of (a) pH3-7, and (b) pHFB7.2A2, which detect the locus *DXHXF34,* to EcoRI digests of YAC DNA. Both probes detect fragments of approximately 4.0, 2.5, 1.8 and 1.3kb in EcoRI digested 3H1 DNA. As there was a fragment of approximately 2.5kb in AB1380 yeast strain DNA (yeast control for (a) not shown) it was unclear which of the YACs contained this fragment. None of the YACs possessed all of the other three EcoRI fragments and a summary of these results are given in Table 4.2.

Table 4.2 Characterisation of *DXHXF34* YACs.

YAC Reference	*DXHXF34* EcoRI Fragments (kb)		
	4.0	1.8	1.3
F0651	+	-	+
F0751	+	-	+
C069	+	+	-
E0547	+	+	-
E1216	-	+	-
H0127	-	+	-

+ present in YAC; - absent in YAC

4.2.2 *DXMit26* and *Gata1* YAC clones.

The YAC I114H3 was isolated by Nandita Quaderi (St Mary's Hospital) by PCR amplification of ICRF mouse YAC library pools using *DXMit26* microsatellite primers. The presence of *DXMit26* in this YAC was confimed by PCR amplification of a *DXMit26* product from YAC DNA (Figure 4.3a). I114H3 was also known to contain *Gata1*.

The *Gata1* positive YACs FDDB10 and FDBE3 were isolated by D. T. Burke (University of Michigan) by PCR amplification of Massachusetts Institute of Technology (MIT) YAC library pools from GATA1 primers. The presence of *Gata1* in the YACs I114H3, FDDB10 and FDBE3 was confirmed by the hybridisation of the human GATA1 cDNA clone, K14, to EcoRI digested YAC DNA (Figure 4.3b). Bands of approximately 3.2kb were detected by K14 in all three *Gata1* positive YAC clones. A 3.2kb band detected in EcoRI digested 3H1 DNA by K14 is located on the mouse X chromosome (Laval and Boyd, 1993b; Table 3.2).

4.2.3 *Cybb* YAC clones.

YAC clones F192, C8127 and E840 were isolated from the ICRF mouse YAC library by A.P. Monaco (Wellcome Trust Centre for Human Genetics, Oxford) after hybridisation of the human CYBB cDNA to library colony filters. The YAC references F192, C8127 and E840 are laboratory names and do not refer to the ICRF mouse YAC library coordinates. The presence of *Cybb* in the YACs was confirmed by hybridisation of the human CYBB cDNA to HindIII digested YAC DNA (Figure 4.3c).

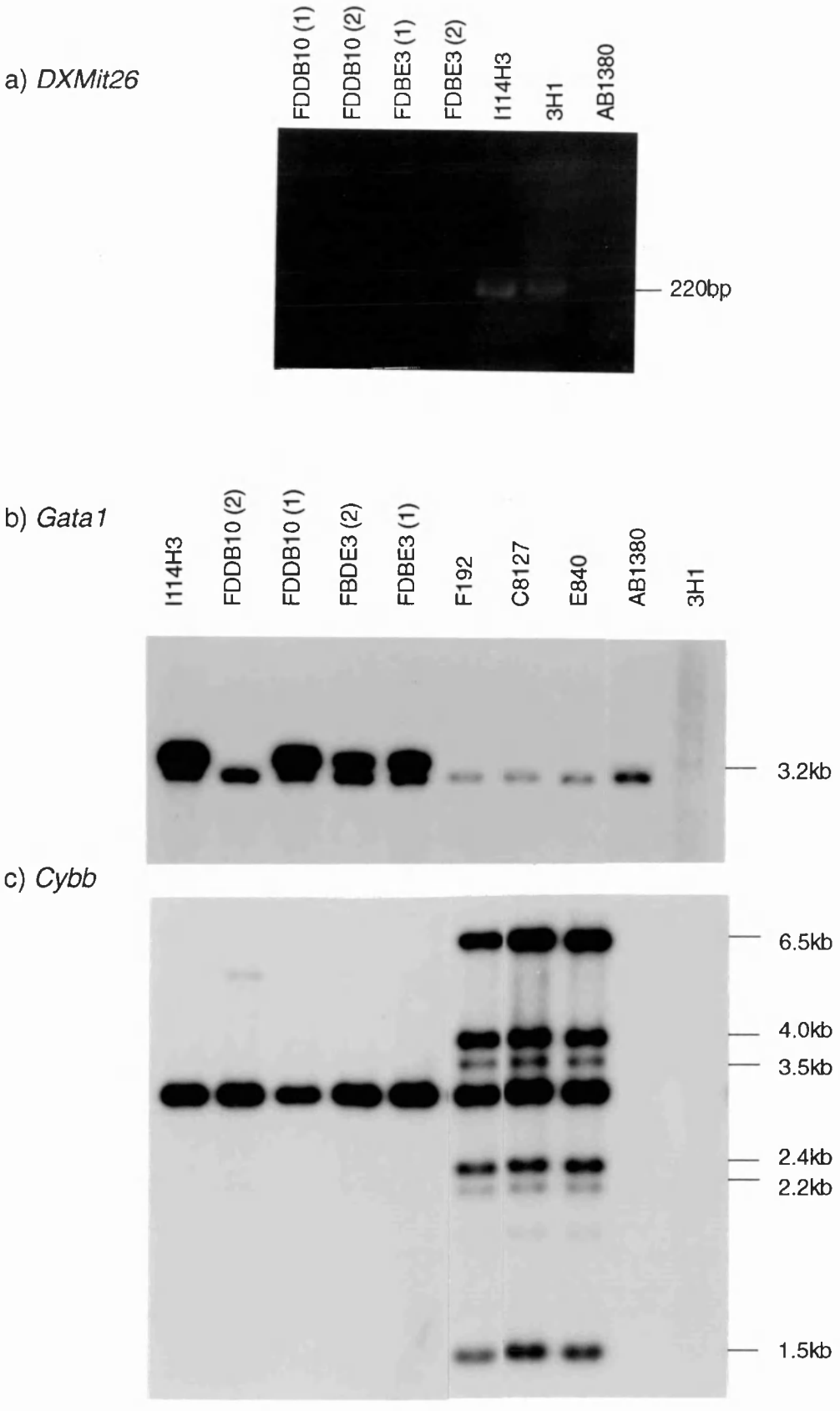


Figure 4.3

Analysis of YAC DNA for the presence of loci with which they were isolated. (a) Analysis of the *DXMit26* PCR amplification products from YAC DNA samples, 3H1 and AB1380 yeast strain controls. The 220bp *DXMit26* product is only present in YAC I114H3 and the 3H1 control. Hybridisation of (b) K14, which detects *Gata1,* to EcoRI digested YAC and control DNAs. The band present in all of the YACs is a yeast sequences, and (c) the human CYBB cDNA toHindIII digested YAC and control DNAs. The ~3.0kb band present in all the YACs is a vector sequence (analysis not shown). The *Cybb* cDNA detects a similar pattern of fragments in 3H1 DNA digested with HindIII as in the YACs, F192, C8127 and E840, however the signal in the 3H1 track in (c) is too weak to be observed. The FDDB10 and FDDBE3 yeast host cells were cultured twice and DNA prepared indicated by (1) and (2). FDDB10 (2) does not appear to contains *Gata1* (this is discussed in section 4.5.1).

4.3 Characterisation of YACs from the Proximal Region of the Mouse X Chromosome.

All twelve YAC clones were analysed to determine their size, X chromosome content and the presence/absence of markers from the proximal region of the mouse X chromosome.

4.3.1 Determination of YAC size by pulsed-field gel electrophoresis.

Because conventional unidirectional-field agarose gel electrophoresis can only resolve DNA fragments of up to approximately 20kb, larger DNA molecules, such as YACs, require alternative techniques. Pulsed-field gel electrophoresis, the application of bidirectional, alternating electric fields to DNA fragments in an agarose gel, can be used to separate yeast chromosome-sized DNA fragments (Schwartz and Cantor, 1984).

Intact YAC DNA was prepared in agarose blocks and separated by pulsed-field gel electrophoresis (see section 2.10.2 and 2.10.3). DNA samples are fractionated according to their size, therefore the approximate size of each of the YACs could be determined by comparison of the distance migrated by each YAC to the distance migrated by yeast chromosomes of known size. YAC DNA of a size different to any yeast chromosomes can be observed on an ethidium bromide stained pulsed-field gel e.g. FDBE3 (3), Figure 4.4b; however YACs often co-migrate with one of the yeast chromosomes so can not be discriminated. Therefore, all gels were blotted and hybridised with mouse genomic DNA, which hybridises to the mouse YAC but not the yeast DNA. The presence of several YAC bands of different sizes within the same

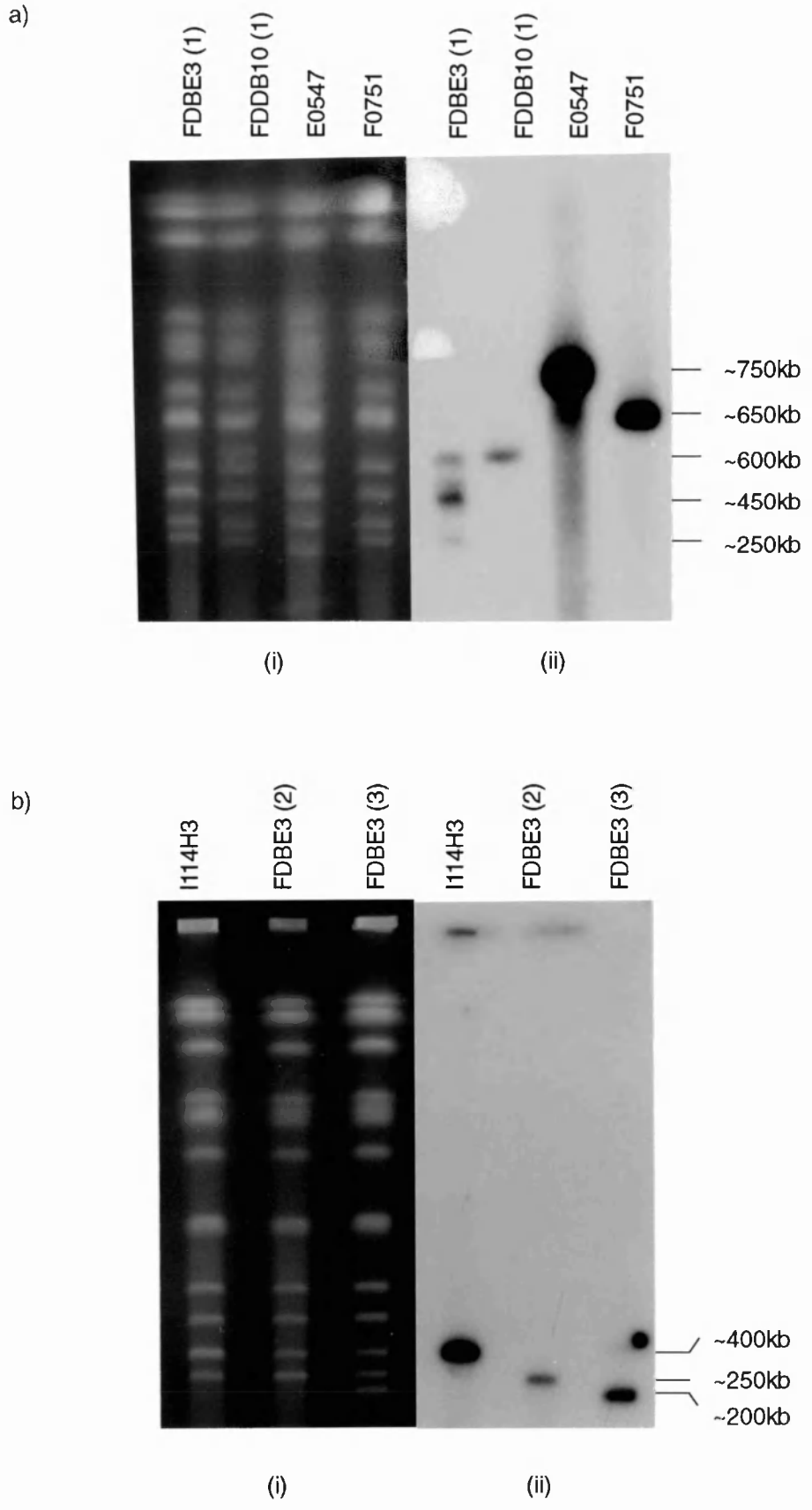


Figure 4.4

Determination of YAC clone size by pulsed-field gel electrophoresis. Pulsed-field gel electrophoresis was performed on (a) a Waltzer apparatus as described by Southern, *et al.*, 1987 and (b) a CHEF mapper system (BioRad). Pulsed-field gels were, (i) stained with ethidium bromide and, (ii) blotted and hybridised with labelled mouse genomic DNA. FDBE3 yeast host cells were cultured separately three times indicated by 1 (a) and 2 and 3 (b) and each time the YAC was a different size (see section 4.5.1). Approximate YAC sizes shown on the right were estimated from the distance migrated through the gel by yeast chromosomes of known size (i).

yeast cell can sometimes be observed by this procedure. The ethidium bromide stained pulsed-field gels and corresponding mouse genomic DNA hybridisation signals of several of the YACs analysed in this study are shown in Figure 4.4 and the deduced YAC sizes are given in Table 4.3.

Three of the YACs analysed, F0751, FDBE3 and C8127 contained more than one band when hybridised with mouse genomic DNA. The extra bands could have been observed, either because the yeast cells which contained the YAC of interest also contained other different size YACs, or because the YAC had rearranged to produce different sized products. The presence of the locus used to isolate the YAC in all the bands was tested by hybridisation (e.g. *DXHXF34, Gata1* and *Cybb* in F0751, FDBE3 and C8127 respectively). For each YAC, all of the bands detected by mouse genomic DNA also appeared to contain their isolation locus (results not shown) indicating that the extra bands were rearranged products of the YACs. Yeast cells, containing the YAC FDBE3, were cultured separately three times and pulsed field gels prepared and blotted (Figure 4.4). Each time the cells were cultured, the YAC appeared to be of a different size, indicating that this YAC is unstable and prone to deletion (see 4.5.1).

The average size of the YAC inserts, excluding those with multiple bands, was approximately 700kb which is similar to the average mouse genomic DNA insert size of the ICRF mouse YAC library (Larin *et al.*, 1991).

Table 4.3 Determination of YAC Size and Chimaerism.

YAC Reference	Isolation Locus	Approx. Size (kb)	Chimaeric[a]	No. of Autosomal Signals
ICRFy902F0651	*DXHXF34*	>1000, 650, 150	Yes	1
ICRFy902F0751	*DXHXF34*	650	No	0
ICRFy902E0547	*DXHXF34*	750	No	0
ICRFy902C069	*DXHXF34*	600	nt	-
ICRFy902E1216	*DXHXF34*	600	nt	-
ICRFy902H0127	*DXHXF34*	900	Yes	1
I114H3	*DXMit26*	400	No	0
FDDB10	*Gata1*	600	Yes	2
FDBE3	*Gata1*	600, 450, 250	Yes	2
F192	*Cybb*	>1000	Yes	2
C18127	*Cybb*	900, 750	Yes	2
E840	*Cybb*	900	Yes	1

a: YACs were classified as chimaeric if autosomal signals were detected following *in situ* hybridisation. The number of autosomal signals is shown in the next column. nt : not tested.

4.3.2 Confirmation of X chromosome location and determination of YAC chimaerism using FISH.

In order to confirm the X chromosome location of the YACs and to determine if they were chimaeric, liquid DNA was labelled and hybridised to metaphase spreads prepared from male and/or female mice carrying either the T(X;4)37H or T(X;11)38H translocations (Searle *et al.*, 1983). Both translocations have resulted in the production of marker chromosomes which can be distinguished by their size from other mouse chromosomes on a metaphase spread. The T38H translocation breakpoint has been localised proximal to *DXWas70* and the T37H breakpoint distal to *Otc* (Fisher and Tease, 1992; Searle *et al.,* 1983; Figure 4.5). YACs which contain DNA from the *DXHXF34-Cybb* interval of the mouse X chromosome should hybridise to the short marker chromosome of T37H mice and the long marker chromosome of T38H mice. YACs which hybridised to autosome pairs, in addition to the marker chromosome, must contain a substantial amount of genomic DNA from other regions of the mouse genome and were classified as chimaeric (Table 4.3). Metaphase spread preparation, signal detection and spread analysis was carried out by S.Holt (MRC Radiobiology Unit). Examples of YACs hybridised to metaphase spreads prepared from T37H and T38H mice are shown in Figure 4.6. Eight of the ten YAC clones tested appeared to be chimaeric. The proportion of X chromosomal DNA in a chimaeric YAC could not be deduced from the relative intensity of the X chromosome to autosomal signal(s) since signal intensity is related not to the size of the probe, but to the repeat sequences it contains. Also, a difference in the intensity of signal from an active and inactive X chromosome has been observed (Y. Boyd, personal communication), therefore signal intensity may be related to availability for hybridisation of the chromosomal DNA on

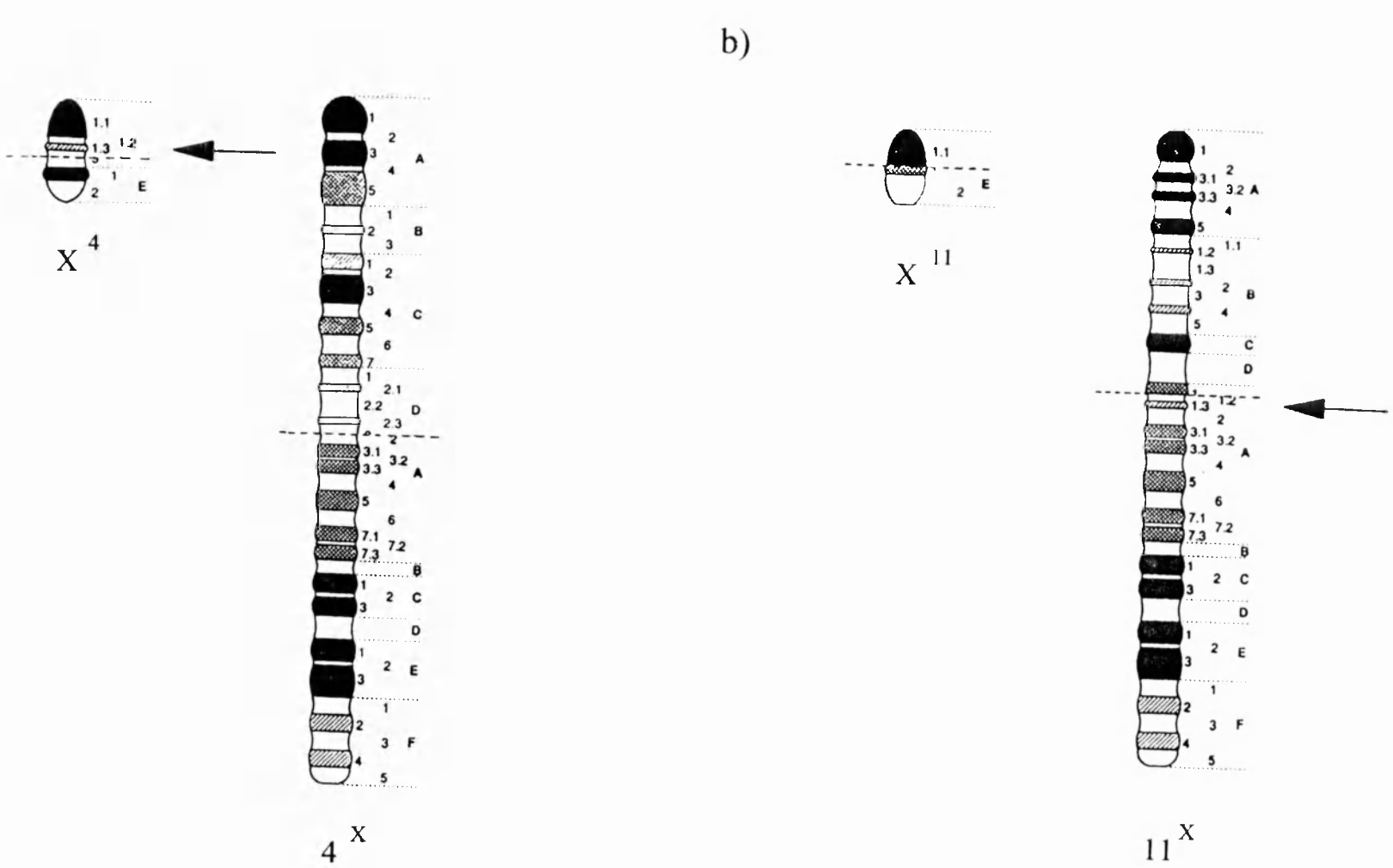
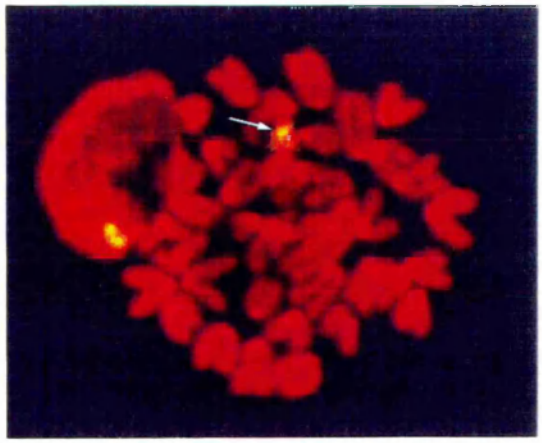
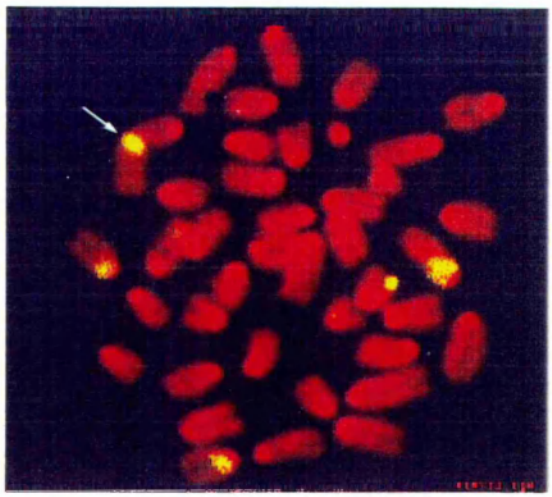


Figure 4.5
Idiograms of banded, long and short, marker chromosomes resulting from (a) T(X;4)37H and (b) T(X;11)38H reciprocal translocations (Searle *et al.*, 1983). The *DXWas70-Cybb* region is indicated by an arrow, on the T37H short, and the T38H long, marker chromosomes.

a) E0547
T(X;4)37H male



b) FDBE3
T(X;11)38H female



c) E840
T(X;11)38H female

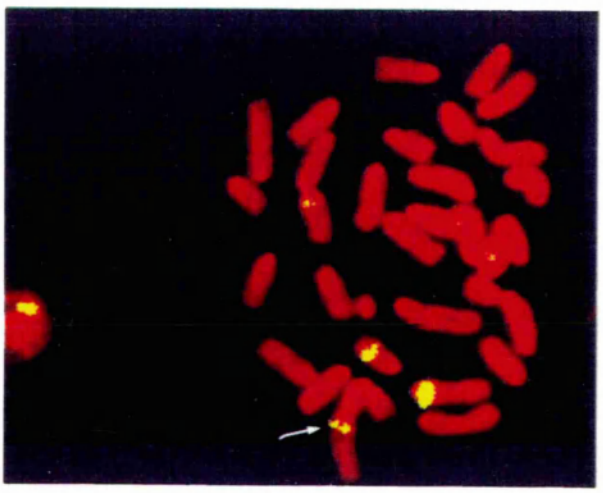


Figure 4.6
Hybridisation of (a) E0547 to T(X;4)37H male metaphase spreads, (b) FDBE3 and (c) E840 to T(X;11)38H female metaphase spreads. Approximately 30ng/µl of biotin labelled YAC DNA in the presence of 200ng/µl sonicated mouse DNA was hybridised for 4 nights to spreads prepared from bone marrow cells. Signals were detected with avidin conjugated to fluorescein (FITC) and amplified with biotinylated goat anti-avidin. The marker chromosomes are indicated by an arrow. Results from the analysis of more than ten spreads are given in Table 4.3.

the metaphase spread.

4.3.3 Determination of the presence/absence in the YACs of markers from the proximal region of the mouse X chromosome.

The presence of loci from the proximal region of the mouse X chromosome in the YACs was determined by the hybridisation of labelled probes to digested YAC DNA on a Southern blot, or by the amplification of X-specific products by PCR (Figure 4.7 and data not shown). The presence of loci which lay close to the isolation locus on the genetic map of the region was tested in the relevant YAC clones. The results of these analyses are shown in Table 4.4.

YAC clones were analysed for the presence of *Syp* by both Southern blot and PCR techniques. Results were consistent with these techniques for all YACs tested except I114H3. Labelled *Syp* exon 4 probe appeared to hybridise to a band in I114H3 digested DNA, but no product was amplified from the same I114H3 DNA using the *Syp* exon 4 primers (data not shown). YAC I114H3 was isolated using primers which amplify the *DXMit26* locus and, as *DXMit26* cosegregated with *Syp* in the studies described in chapter 3, it is possible that both loci are present in this YAC. The inability of the *Syp* exon 4 primers to amplify a product from the I114H3 DNA may be due to impurities in the DNA solution which inhibited the PCR reaction, or because the *Syp* exon 4 sequence is very close to one end of this YAC and one of the primer sites is absent in the clone. Although this could have been tested by hybridisation of a *Syp* cDNA clone to I114H3 DNA, no further studies were possible because a cDNA was not available.

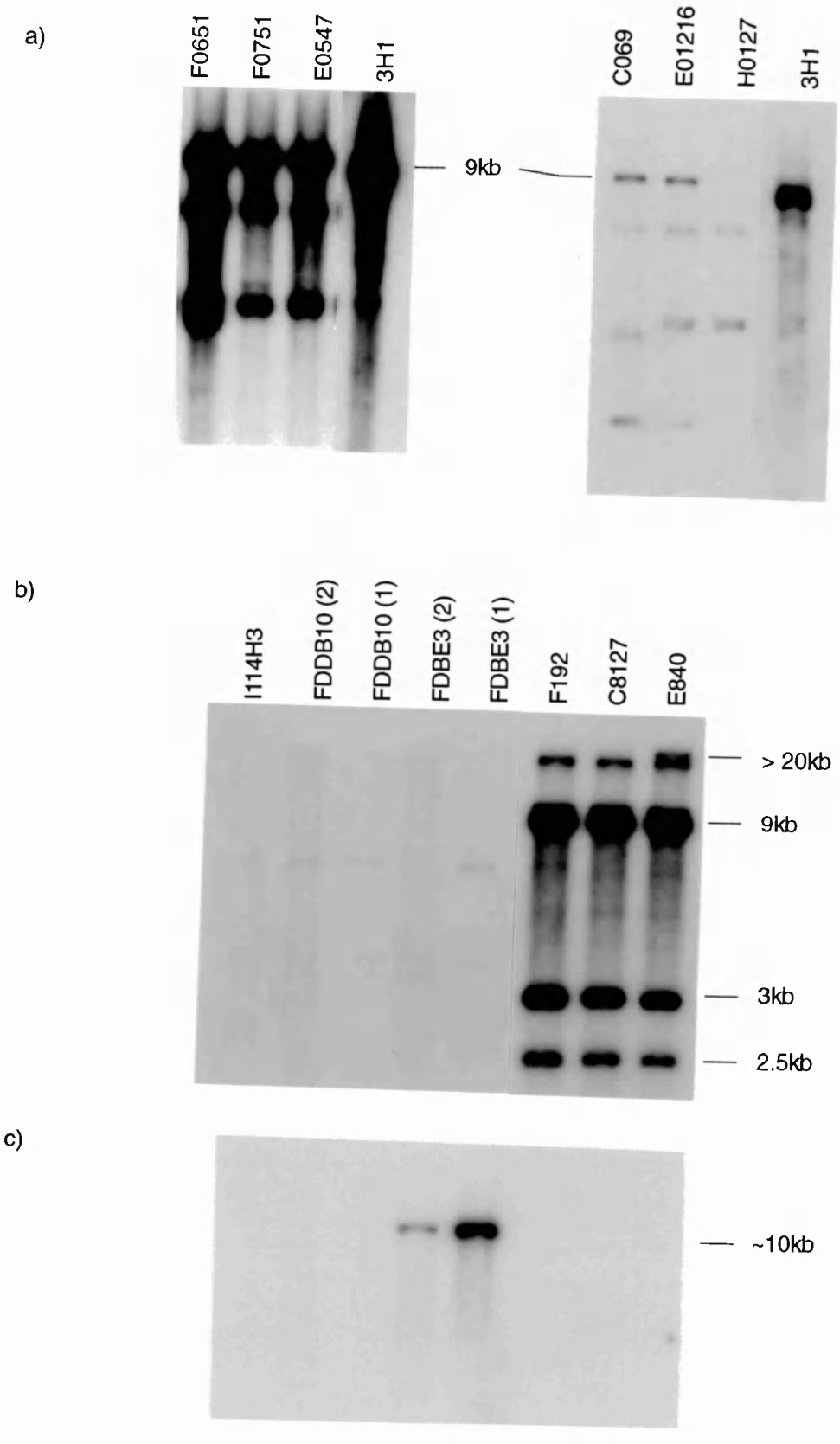


Figure 4.7

Analysis of YACs for the presence of markers from the proximal region of the mouse X chromosome. Hybridisation of (a) 70-68 which detects *DXWas70*; (b) XK and (c) radiolabelled *Tcfe3* PCR product to EcoRI (a and b) and HindIII (c) digested YAC DNA. *DXWas70* is only present in five of the six *DXHXF34* YACs (a); *Xkh* is present in YACs F192, C8127 and E840 (see b) and *Tcfe3* was only present in FDBE3, a result confirmed by PCR analysis (not shown).

Table 4.4 Analysis of YACs for the Presence of Markers from the Proximal Region of the Mouse X Chromosome.

YAC Ref.	Locus Content[a]										
	DXWas70	*DXHXF34*	*Nphl*	*Syp*	*DXMit55*	*DXMit26*	*Tcfe3*	*Gata1*	LAR3 end clone	*Xkh*	*Cybb*
ICRFy902F0651	+	+	-	-	-	-	-	-	nt	nt	nt
ICRFy902F0751	+	+	-	-	-	-	-	-	nt	nt	nt
ICRFy902E0547	+	+	-	-	-	-	-	-	nt	nt	nt
ICRFy902C069	+	+	-	-	-	-	-	-	nt	nt	nt
ICRFy902E1216	+	+	-	-	-	-	-	-	nt	nt	nt
ICRFy902H0127	-	+	-	-	-	-	-	-	nt	nt	nt
I114H3	-	-	-	+[b]	-	+	-	+	-	-	-
FDDB10	-	-	-	-	-	-	-	+	-	-	-
FDBE3	-	-	-	-	-	-	+	+	+	-	-
F192	nt	nt	-	-	-	-	-	-	-	+	+
C8127	nt	nt	-	-	-	-	-	-	-	+	+
E840	nt	nt	-	-	-	-	-	-	-	+	+

a : + = locus present in YAC; - = locus not present in YAC; nt = YAC not tested.
Presence/absence of loci in each YAC was tested by PCR and/or Southern hybridisation.
b : A Positive *Syp* signal was observed in YAC 1114H3 DNA on a Southern blot however PCR results were negative.

From the genetic mapping studies (Chapter 3) the order of loci in the proximal region of the mouse X chromosome was established as *DXMit26* - *Tcfe3* - *Gata1*. YAC I114H3 appeared to contain the loci *DXMit26* and *Gata1* but not *Tcfe3* indicating a deletion of *Tcfe3*. *DXMit26* and *Gata1* lie approximately 0.7cM apart on the genetic map of the mouse X chromosome, representing a possible physical distance greater than 1Mb. As I114H3 was only 400kb, a deletion of several hundred kb, including *Tcfe3* has probably occurred in this YAC.

4.5 Construction of a Physical Map of the Proximal Region of the Mouse X Chromosome.

A physical map of the 3-4cM interval between *DXHXF34* and *Cybb* was constructed based on the genetic order of loci established in Chapter 3. The twelve YACs analysed were organised into three contigs with two gaps (Figure 4.8). As many of the clones were found to be chimaeric, the amount of X-linked material present in these clones could not be determined. The extent of the deletion in clone I114H3 was also not known.

A possible organisation of the three *DXHXF34* and the one *DXWas70* EcoRI fragment determined from the presence of these fragments in the *DXHXF34* YACs (section 4.2.1) is illustrated in Figure 4.8. The orientation with respect to the centromere of this contig is not known.

4.5.1 The instability of FDBE3 and FDDB10.

YACs FDBE3 and FDDB10 both appear to be unstable. Yeast cells containing YAC

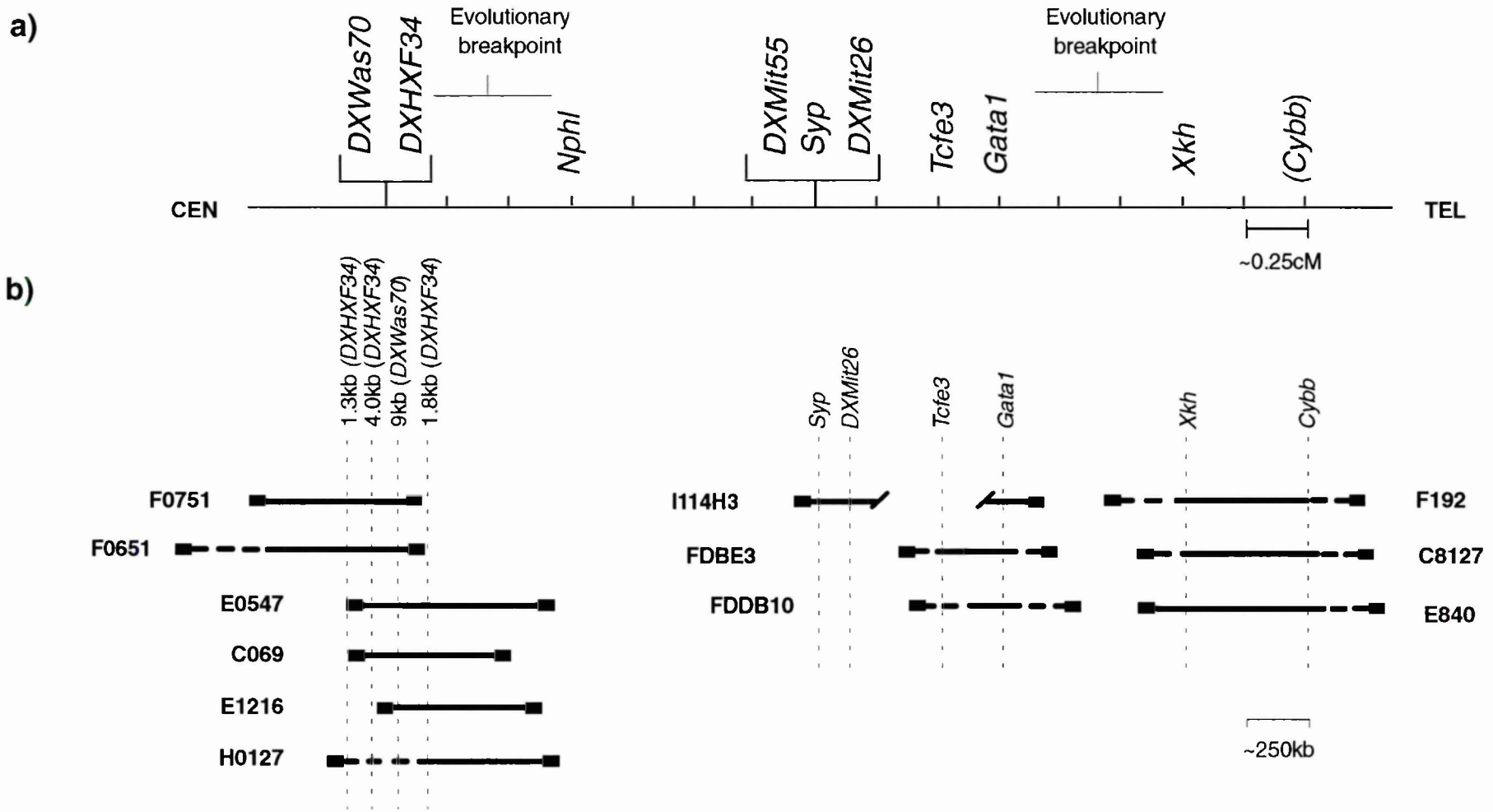


Figure 4.8

Diagrammatic representation of (a) the map of *DXWas70-Cybb* interval of the mouse X chromosome based on the genetic distances calculated in Chapter 3. The position of *Cybb* (shown in parentheses) was estimated using data from lab backcrosses (see section 3.5.1). The approximate position of the regions which contain evolutionary breakpoints is shown; (b) the three YAC contigs from the same region. Possible chimaeric regions of the YACs are indicated by discontinuous lines and the deletion in YAC I114H3 indicated by a gap. The possible positions of the three *DXHXF34* EcoRI fragments and the 9kb *DXWas70* EcoRI fragment are shown in the six *DXHXF34* YACs.

FDBE3 were cultured three times, and those containing FDDB10 twice and DNA prepared for analysis. The *Gata1* hybridisation signal of 3.2kb was present in only one of the FDDB10 DNA preparations (FDDB10 (1); Figure 4.3b) indicating that upon culturing of the yeast cells YAC FDDB10 (2) had lost this *Gata1* sequence. Results were therefore based on the analysis of FDDB10 (1) DNA.

Although both FDBE3 (1) and (2) DNAs appeared to contain the *Tcfe3* hybridisation signal of ~10kb, the hybridisation signal was much stronger in FDBE3 (1) than in FDBE3 (2) (Figure 4.7). The quantity of DNA in each of the tracks was similar therefore a proportion of the FDBE3 (2) YAC clone DNA had lost the *Tcfe3* sequence. Using pulsed-field gel electrophoresis, a band of smaller size was observed in FDBE3 (2) compared with FDBE3 (1) DNA, also indicating that sequences had been deleted (Figure 4.4). Results for the presence of other loci were based on the analysis of FDBE3 (1) DNA.

4.5.2 Estimation of intermarker physical distances from YAC studies.

The 9kb EcoRI fragment detected at *DXWas70* was present in five of the *DXHXF34* YACs (Figure 4.7). This sequence must therefore lie between the 1.8 and 4.0kb *DXHXF34* fragments (Figure 4.8). This is consistent with the cosegregation of *DXWas70* and *DXHXF34* in backcross studies. None of the *DXHXF34* YACs contained *Nphl*, therefore they did not cross the evolutionary breakpoint between the two loci.

The loci *Tcfe3* and *Gata1* were both located in YAC FDBE3 which had a maximum

size of 600kb. Therefore, *Tcfe3* and *Gata1* probably lie less than 600kb apart on the mouse X chromosome, a physical distance supported by a correspondingly small genetic distance of 0.2cM, established from backcross studies in Chapter 3. The human homologues of *Tcfe3* and *Gata1* lie less than 1Mb apart in the proximal region human X chromosome short arm (Nelson *et al.*, 1995). These results indicate that the physical distance between *Tcfe3* and *Gata1* is similar in both mouse and human.

YACs F192, C8127 and E840, which all contained *Xkh* and *Cybb.* were chimaeric, the smallest being 900kb. Therefore, *Xkh* and *Cybb* must lie less than 900kb apart on the mouse X chromosome. On the human X chromosome, XK and CYBB lie less than 400kb apart (Ho *et al.*, 1994). Again, the distances between two loci appear to be similar in man and mouse.

4.5.3 Looking for YAC clone overlaps.

None of the YAC clones studied contained both *Gata1* and *Xkh*, or both *DXHXF34* and *Nphl*, the loci which flank the human-mouse evolutionary breakpoints. In order to link the two sets of loci further markers and/or YACs were required. As these were not readily available, an attempt was made to isolate end sequences from YACs I114H3, FDDB10, FDBE3, F192, C8127 and E840. These YACs were selected as they contained either *Gata1* or *Xkh*, the loci which flank one of the gaps in the physical map. As *Nphl* was not present in any of the YACs no attempt was made to isolate end sequences from the *DXHXF34* YACs to close the other gap in the contig.

Five of the selected YACs were chimaeric, therefore at least one of their end

sequences would not be of proximal X chromosome origin and not useful to these studies. Also, because only one of the end clones from each YAC would be close to the gap in the contig it was important to choose a method which could be used to obtain both left and right arm clones. Only the left arm of pYAC4, the vector used to clone the YACs studied here, contains sequences for plasmid rescue. Several other methods, including inverse PCR, vectorette PCR and vector-repeat sequence PCR have been developed to isolate end sequences next to both YAC vector arms (Silverman *et al.*, 1991; Riley *et al.*, 1990 and Breukel *et al.*,1990). With the limited amount of time available for the completion of these studies, only the inverse PCR method was attempted.

The inverse PCR procedure followed is represented schematically in Figure 4.9. YAC DNA was digested with the restriction enzyme XhoI which cuts only once in each arm to remove the TEL sequences but does not cut between the vector and the genomic DNA insert. The technique relies on the presence of a XhoI cut site within the insert DNA, sufficiently close to the vector arm, to allow PCR amplification to be carried out on a circularised vector/insert XhoI fragment with primers designed from the vector sequence. Primers were designed from the GenBank database pYAC4 sequence (Accession number U01086); their positions on each arm are shown in Figure 4.9. For left end sequence isolation, XhoI cut and religated YAC DNA was amplified by PCR from the primers A (5'-TTT AAG GCG CAA GAC TTT-3') and A' (5'-AGG TTT TCA CCG TCA TCA CCG-3'). For right end sequence isolation the same DNA was amplified by PCR from the primers B (5'-TCC CTT GGC GAT CGA ACG CCC-3') and B' (5'-CGC TTC ACG ACC ACG CTG ATC-3'). Products of approximately 0.8kb

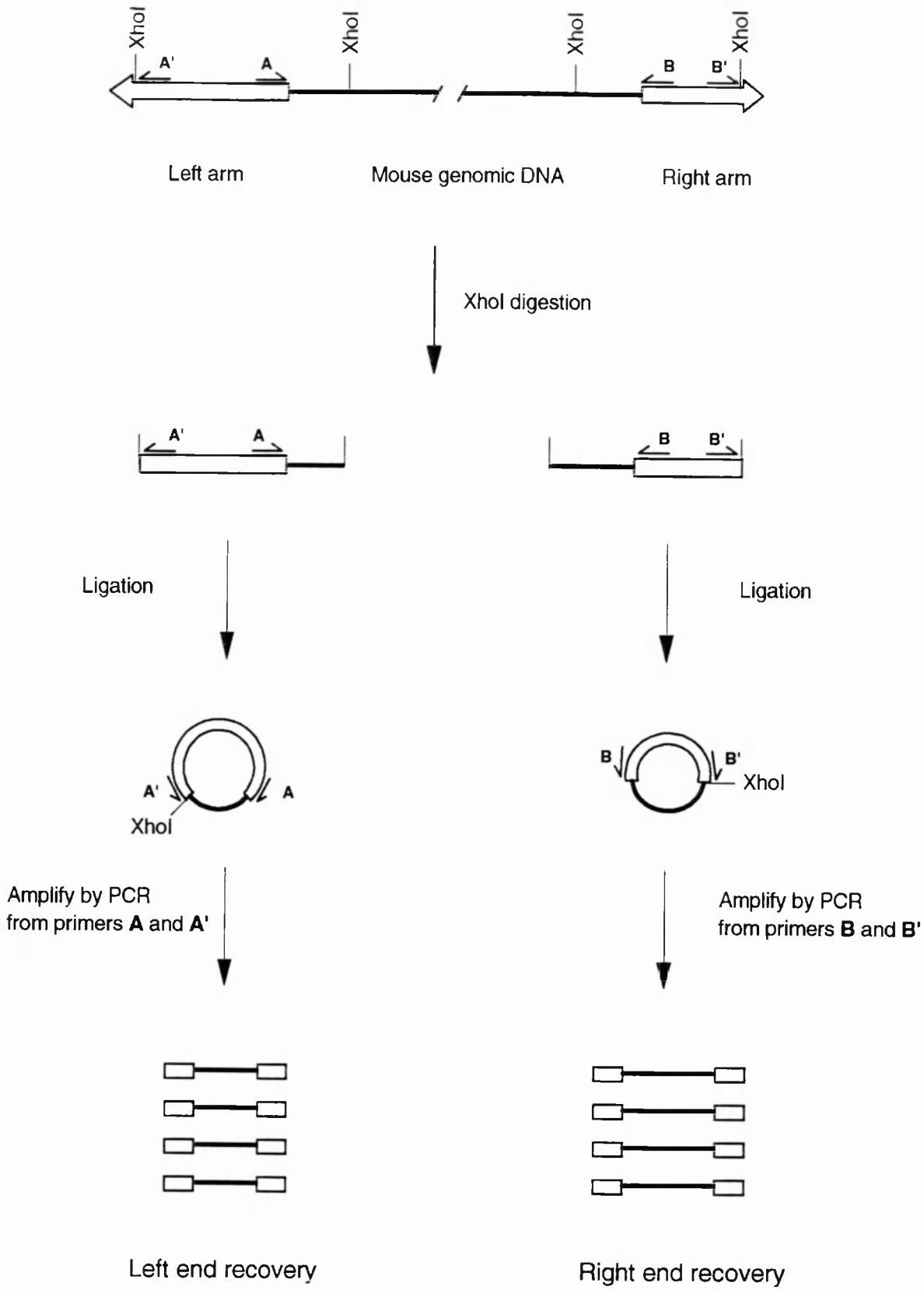


Figure 4.9

Schematic representation of the recovery of YAC end sequences by inverse PCR (Silverman, *et al.*, 1991). YAC DNA is digested with the restriction enzyme XhoI which cuts only once in each vector arm. The technique relies on the presence of a XhoI cut site in the mouse genomic DNA close enough to the vector arm to allow PCR amplification. The digested YAC DNA is diluted and ligated to allow the vector/insert fragment to circularise. End sequence can then be amplified from vector primers. The primers A and A' are employed for left end recovery and B and B' for right end recovery. The sequence of these primers is given in the text.

and 1.3kb were amplified from the left end of YACs FDBE3 and C8127 respectively; however no distinct end products were amplified from I114H3, FDDB10, F192 and E840 (results not shown). It was assumed that the distances between the XhoI cut sites and the vector in YACs I114H3, FDDB10, F192 and E840 were not optimal for PCR amplification of the end sequences. There was no products from the right arm of any of the YACs and subsequently the sequence from GenBank was found to contain several errors in the sequence of primer B, which would have prevented proper annealing of the primer.

4.5.4 Characterisation of left end clones from FDBE3 and C8127.

The left end products from FDBE3 and C8127 (LAR3 and LAR6 respectively) were isolated and hybridised to digested DNA from all six *Gata1* or *Xkh* YACs (Figure 4.10) Both probes were pre-reassociated with 2μg of sonicated mouse genomic DNA before hybridisation to reduce signal from repetitive sequences within the probe (see Chapter 2: section 2.8.4). LAR3 hybridised to ~1.5kb and ~2.2kb bands in FDBE3 and 3H1 DNA respectively, digested with HindIII. LAR3 also hybridised to several bands in yeast strain AB1380 DNA which were common to all the YAC DNA samples. LAR6 hybridised to a ~0.8kb band in EcoRI digested C8127 DNA; however the 3H1 signal was highly repetitive even after pre-reassociation. LAR6 must contain repeat sequences which occur frequently in the mouse genome, supported by the multiple bands observed in all the YAC DNA samples.

Unfortunately the left end clones gave strong signals only in the YACs from which they were isolated and therefore failed to identify overlaps between the *Gata1* and *Xkh*

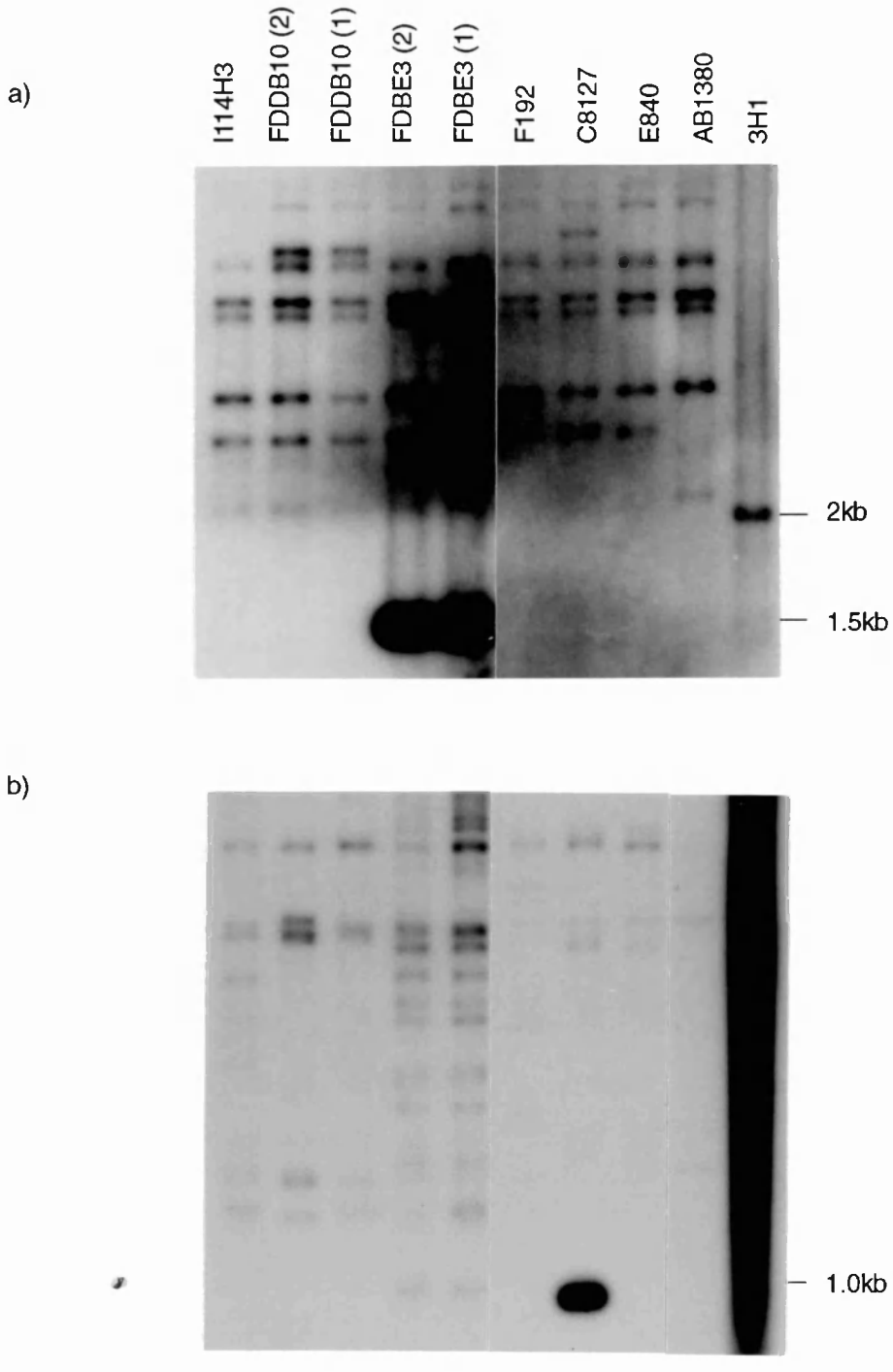


Figure 4.10

Hybridisation of (a) LAR3, end PCR product from FDBE3, to HindIII digested YACs, AB1380 yeast strain and 3H1 DNA and (b) LAR6, end PCR product from C8127, to the same DNAs, but digested with EcoRI. Probes were pre-reassociated with sonicated mouse genomic DNA to reduce signal from any repetitive sequences in the probe. Both end products hybridised to the YACs from which they were isolated but neither product hybridised to sequences in any of the other YACs.

contigs. Because LAR6 detected a highly repetitive signal in 3H1 DNA and was isolated from YAC C8127, which comprised two segments of autosomal DNA in addition to the X chromosome sequences, it was no longer studied. To establish the origin of LAR3, an attempt was made to position it on the X chromosome using recombination events from interspecific backcross offspring. HindIII RFLVs between 3H1 and *M. spretus* of ~2.0kb and ~1.8kb respectively were identified. Hybridisation of LAR3 to HindIII digested DNA from male and female offspring of a 3H1 female x *M .spretus* male cross confirmed an X-linked pattern of inheritance for the *M. spretus* variant (results not shown). LAR3 was positioned proximal to *Otc* using the MXMP described in 3.1.1, because the recombinant chromosome of backcross animal 13.1e carries the *M. spretus* LAR3 allele. (Figure 4.11). The position of LAR3 on the map could not be established from YAC studies; it could lie either between *Gata1* and *Xkh*, or in the region deleted in YAC I114H3 (Figure 4.8). The position of LAR3 could have been determined with respect to the recombination breakpoint between *Gata1* and *Tcfe3* in EUCIB offspring LS867F (Table 3.9); but there was insufficient DNA from this mouse available for analysis.

Surprisingly, LAR3 also detected a ~2.0kb band in TaqI digested human placental DNA. This band was shown to be X-linked using the somatic cell hybrid ThyBX which contains a human X chromosome on a mouse background (Goodfellow *et al.*, 1980). LAR3 detected a ~0.8kb band in TaqI digested 3H1 DNA therefore the ~2.0kb band observed in ThyBX DNA must be the human X chromosome signal (Figure 4.11c). As LAR3 is from the mouse X chromosome, it is a useful clone for the isolation of further YACs from the region. In addition, it will also be useful for

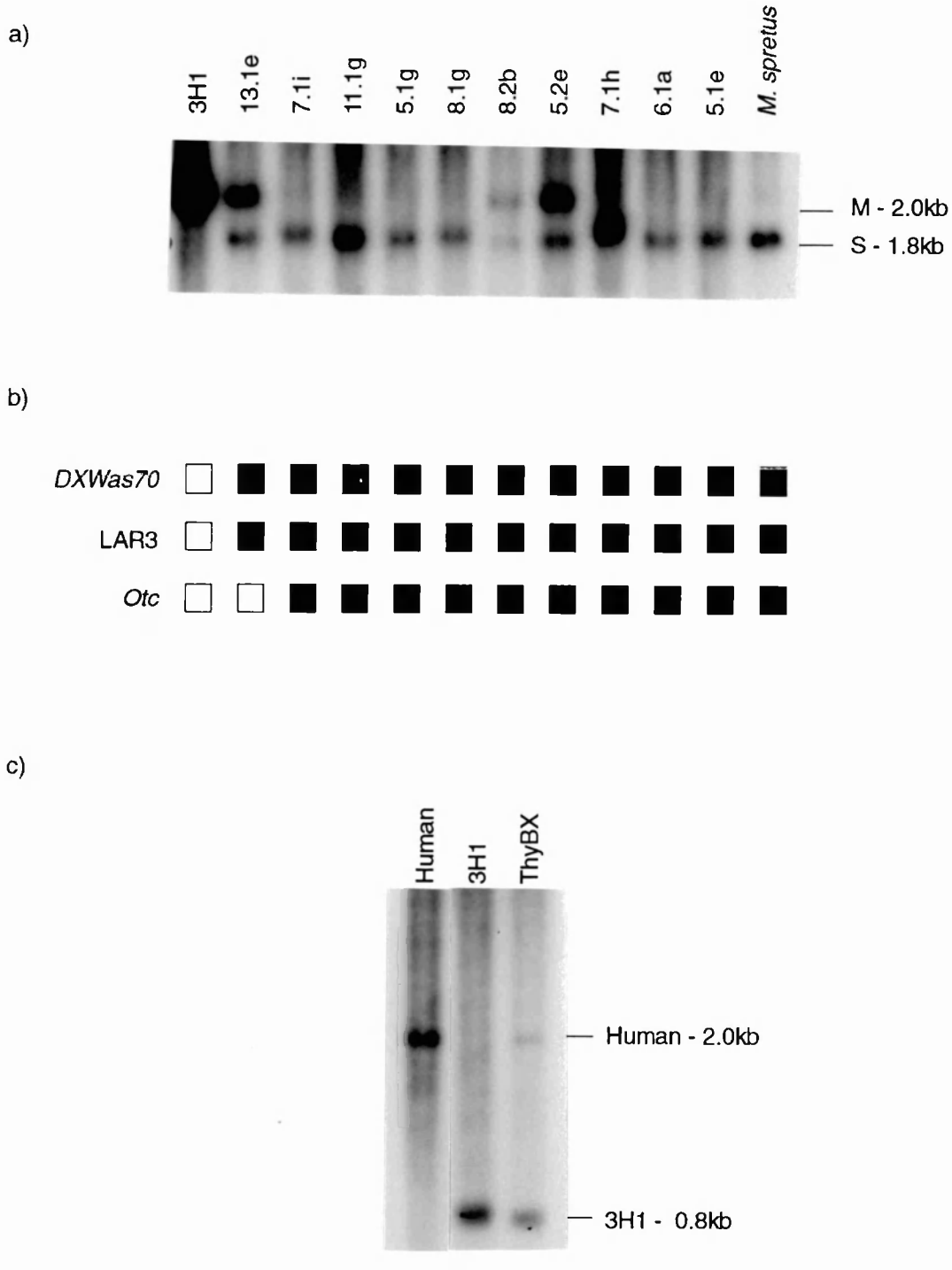


Figure 4.11

(a) Hybridisation of LAR3 PCR product to HindIII digested MXMP backcross DNAs (see section 3.1.1; Fig. 3.1). The position and size of the 3H1 and *M. spretus* bands are indicated by M and S respectively. (b) Haplotype analysis of the MXMP for LAR3; backcross animal 13.1e has a 3H1 *Otc* allele (open square) but *M .spretus* alleles (black squares) at LAR3 and *DXWas70,* indicating that LAR3 lies proximal to *Otc.* (c) Hybridisation of LAR3 to TaqI digested human, 3H1 and ThyBX DNAs. ThyBX is a somatic cell hybrid which contains the human X chromosome on a mouse background (Goodfellow, *et al.*, 1980). The 2.0kb (human) TaqI band is present in ThyBX therefore the sequence detected by LAR3 in man lies on the X chromosome.

comparative mapping studies as it appears to be conserved.

4.5 Summary and Conclusions.

Twelve YACs were analysed from the *DXWas70-Cybb* interval on the proximal mouse X chromosome. Three YAC contigs were constructed, the two resultant gaps coinciding with the position of human-mouse evolutionary breakpoints. The contigs are illustrated in Figure 4.8. The average size of the YACs was 700kb and eight of the ten tested were chimaeric.

One of the contigs was composed of six *DXHXF34* positive YACs. Study of the *DXHXF34* EcoRI restriction fragments present in these YACs established that the family of sequences at *DXHXF34* was present in two overlapping YAC clones. The close proximity of *DXWas70* to *DXHXF34* suggested from genetic mapping studies was confirmed by the presence of *DXWas70* repeat sequences in five of the *DXHXF34* positive YACs.

The second contig contained three *Gata1* YACs: FDDB10, which was positive only for *Gata1*, I114H3 which was also positive for *DXMit26*, and FDBE3 which was also positive for *Tcfe3*. FDDB10 and FDBE3 were both found to be chimaeric and unstable upon successive culturing. I114H3 was not chimaeric but had deleted *Tcfe3*. Physical distances between these markers could therefore not be determined. It is of interest that YACs from the homologous region in man, which contain TFE3 and GATA1, are also unstable and prone to deletion (Derry *et al.*, 1994). This suggests that there may be sequences conserved in man and mouse in this region which are unstable in YACs.

The third contig consists of three YACs positive for *Cybb,* all of which contained *Xkh.* These loci must lie less than 900kb apart on the mouse X chromosome.

In an attempt to identify regions of YAC overlap between *Gata1* and *Xkh,* end sequences from a *Gata1* and a *Xkh* YAC were isolated by inverse PCR. No overlaps were detected however a conserved clone, LAR3, was isolated from the left end of FDBE3 which will be useful for future YAC isolation and comparative mapping studies.

The YACs described in this chapter provide a framework for the construction of a complete contig of the *DXWas70-Cybb* interval of the mouse X chromosome. They provide an alternative resource to backcross mapping for the positioning of murine homologues of human genes to improve the human-mouse X chromosome comparative map. In addition, they provide cloned material from the region which can be used as a starting point of a search for the genes responsible for tattered and scurfy.

Genetic Analysis of the X Chromosomes of Man and Mouse

Helen J. Blair

M.Phil

July 1996

Volume 5 of 5

Chapter 5

Discussion

5.1 An Improved Man-Mouse Comparative Map of Human Proximal Xp.

A detailed comparative map of the human and mouse X chromosomes will facilitate the identification of mouse models for human inherited diseases and provide insight into the chromosomal rearrangements which have occurred since the divergence of human and mouse from a common ancestor. The aim of this study was to improve the comparative map of the human and mouse X chromosomes and, in particular, to define the evolutionary breakpoints between the conserved segments which compose human proximal Xp (Figure 1.7).

To fulfil this aim, detailed maps of the two relevant regions of the mouse X chromosome were constructed using offspring from *M. musculus* x *M. spretus* interspecific backcrosses (Chapter 3). Panels of mice with recombination events between *DXWas70* and *Pfc* and between *Plp* and *Pdha1* were analysed for RFLVs and APVs at loci known to lie in these intervals. To improve the map, new markers were developed for linkage mapping in the mouse. These were chosen to provide critical information about the extent and orientation of conserved segments composing human proximal Xp.

TFE3, which encodes the transcription factor enhancer 3, is located proximal and very close to GATA1 in Xp (Nelson *et al.*, 1995). *Tcfe3* was shown to lie only 0.2cM proximal to *Gata1* in the mouse and therefore to belong to the same conserved segment. The murine homologues of SYP, the synaptophysin gene, and NPHL, the gene mutated in Dent's disease, were also positioned on the mouse genetic map and found to lie in this conserved segment (Figure 5.1). The most distal and proximal loci in this conserved segment are now GATA1 and NPHL.

The human XK gene encodes a membrane transport protein and is mutated in patients with McLeod syndrome (Ho *et al.*, 1994). It is located on the X chromosome approximately 200kb distal to CYBB. The murine homologue of XK was found to lie between *Gata1* and *Cybb* in the proximal region of the mouse X chromosome therefore the conserved segment previously defined by CYBB-PFC can now be defined by XK-PFC (Figure 5.1).

The loci DXS679 and DXS674 lie within 40kb of each other in man and less than 400kb distal to ALAS2 (Nelson *et al.*, 1995). In the mouse, their homologues cosegregated in the laboratory backcross and were positioned with respect to *Plp*, *Pdha1* and several microsatellite loci. These data confirmed that *DXHXS674* and *DXHXS679* belonged to the same conserved segment. This segment is likely to contain *Alas2* and lie between the conserved segments defined by *Ar-Col4a5* and *Pdha1-Amg* in the distal region of the mouse X chromosome (Figure 1.5). As the murine homologue of ALAS2 could not be positioned with respect to *DXHXS674* and *DXHXS679*, the orientation of this conserved segment in the mouse was not

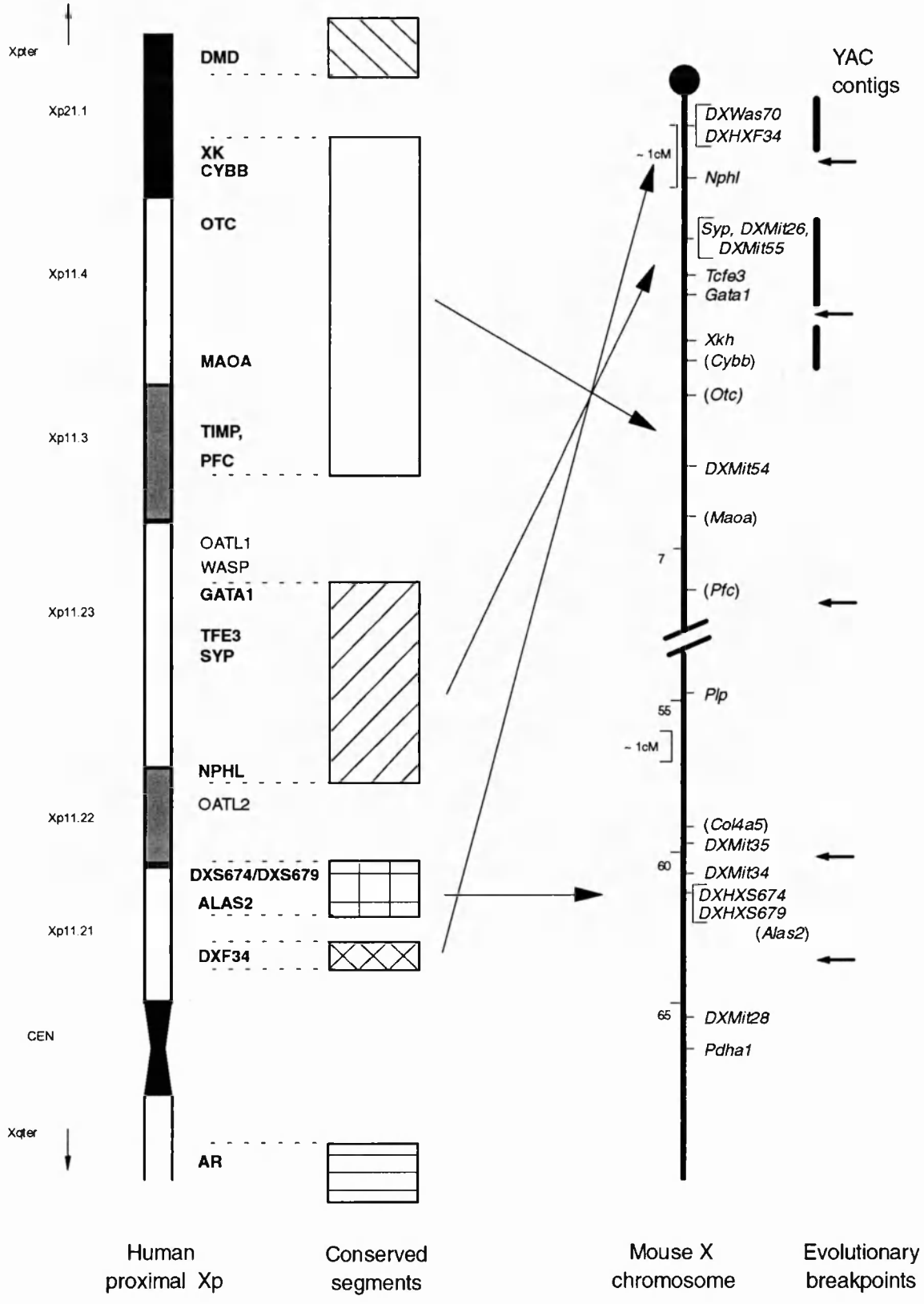


Figure 5.1

Improved comparative map of the proximal region of the human X chromosome short arm incorporating data from Chapters 3 and 4. The extent of each of the conserved segments with the mouse X chromosome is shown in the centre. The positions of loci on the human physical map are taken from Nelson *et al.*, 1995. The map of the proximal mouse X chromosome is based on the genetic distances determined from EUCIB offspring (see 3.6.3). The position of loci in parentheses have been determined from laboratory backcrosses (see 3.2.1 and 3.5.1). The map of the distal mouse X chromosome from *Plp* to *Pdha1*, is based on genetic distances determined from laboratory backcrosses (see 3.3) and the positions of loci in parentheses are taken from Herman *et al.*, 1996. YAC contigs constructed in this study (see 4.5) are represented alongside the mouse X chromosome and the approximate position of evolutionary breakpoints are indicated by arrows.

determined.

In summary, the order of loci on the mouse X chromosome and the position of evolutionary breakpoints (EBs) established from backcross mapping studies was: centromere - (*DXWas70, DXHXF34*) - EB - *Nphl* - (*Syp, DXMit26, DXMit55*) - *Tcfe3* - *Gata1* - EB - *Xkh* - *Cybb* - telomere and, centromere - *Plp* - EB - (*DXHXS674/DXHXS679*) - EB - *Pdha1* - telomere. By combining these data the order of loci and the position of evolutionary breakpoints (**EBs**) in human proximal Xp was established as: DMD - **EB** - XK - PFC - **EB** - GATA1 - NPHL - **EB** - DXS674/DXS679 - ALAS2 - **EB** - DXF34 (Figure 5.1).

Information from the detailed map of the proximal mouse X chromosome was used to initiate the construction of a YAC contig between *DXHXF34* and *Cybb* (Chapter 5). This region spans 3cM and contains two human-mouse evolutionary breakpoints. It may also contain the genes responsible for the mouse mutant phenotypes *Td* and *sf*. Three contigs were constructed consisting of twelve YACs (Figure 5.1). An attempt was made to isolate end clones from some of these YACs in order to link the *Gata1* and *Xkh* contigs and therefore span the region which contained the evolutionary breakpoint between these two loci. One end clone, LAR3, was found to be conserved and therefore will be useful for comparative mapping as well as further contig construction. To complete the contig between *DXHXF34* and *Cybb* an increased number of markers from the region would be required.

5.2 Additional Data Relevant to the Comparative Map of Human Proximal Xp.

Since the completion of these studies, the mouse homologues of several other human genes that map to proximal Xp have been positioned in the mouse. Two of the *Gata1* positive YACs, from the contigs described above, have been found to contain the murine homologue of WASP (I.C. Uwechue, personal communication). WASP is mutated in patients with Wiskott-Aldrich syndrome (Derry *et al.*, 1994), and is known to lie distal to GATA1 on the human X chromosome (Kwan *et al.*, 1995). Although, the order of *Gata1* and *Wasp* on the mouse X chromosome is not known, these data indicate that WASP is located in the same conserved segment as GATA1, therefore, this segment can be redefined by NPHL-WASP (Figure 5.2).

ELK1, a member of the Ets-like family of oncogenes, is located on the human X chromosome proximal to PFC (Rao *et al.*, 1989; Nelson, *et al.*, 1995). The expected position of the mouse homologue of ELK1 on the mouse X chromosome would therefore be distal to *Gata1* or distal to *Pfc*. Recently *Elk* was positioned between *DXHXS32* and *DXPas3* in the proximal region of the mouse X chromosome (Tamai *et al.*, 1995). It can be deduced from the X chromosome consensus map of Herman *et al.* (1996) that *DXHXS32* and *DXPas3* lie approximately 5cM apart and flank *Pfc*. Therefore, *Elk* probably lies proximal to *Pfc* and the XK-PFC conserved segment on human proximal Xp could be extended to include ELK1 (Figure 5.2).

The mouse homologues of two additional human genes have been positioned in the distal region of the mouse X chromosome; these map data provide information about

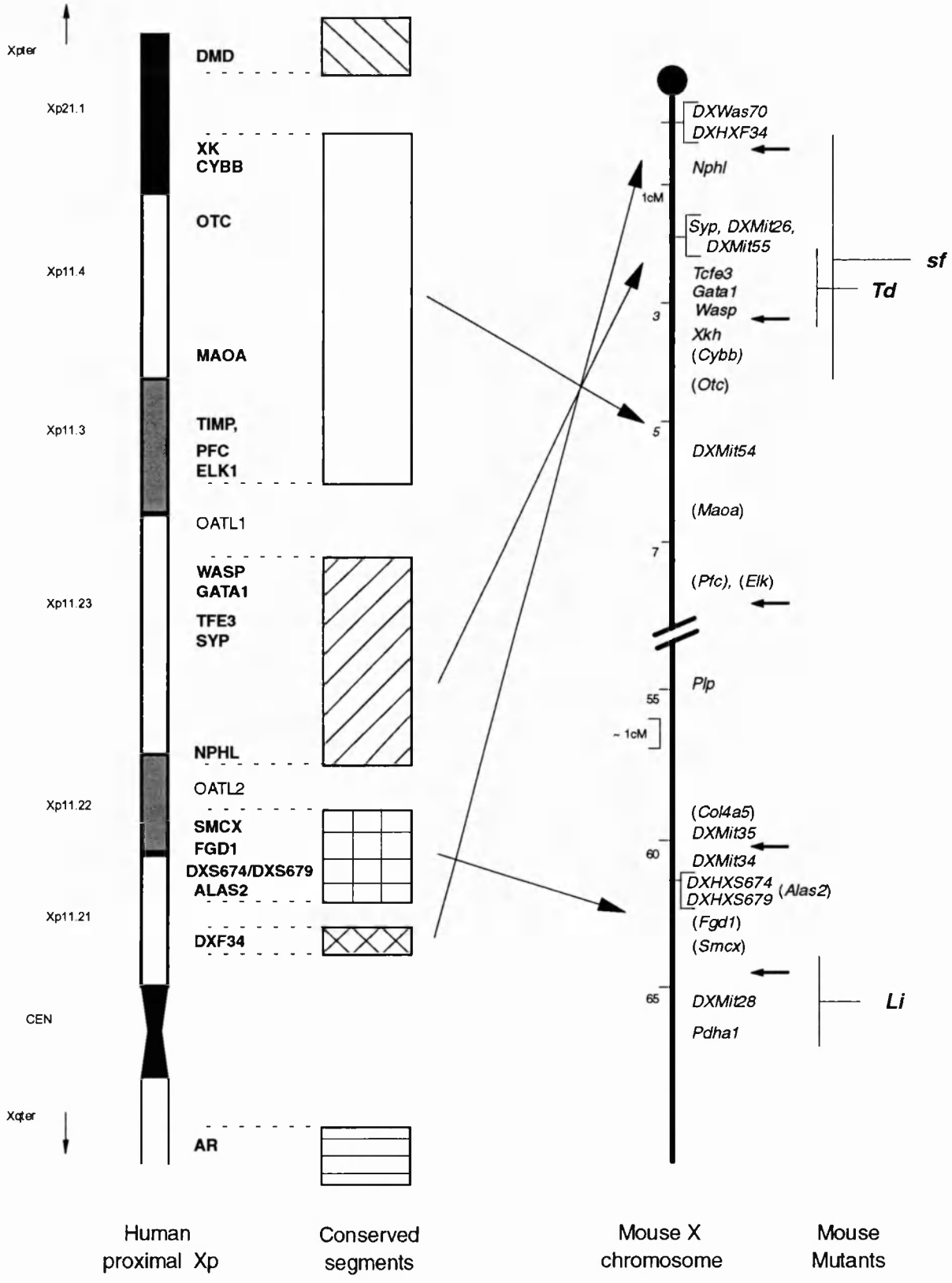


Figure 5.2

Comparative map of the proximal region of the human X chromosome short arm incorporating data from Figure 5.1 and data from other sources (see text). The approximate position of mouse mutations is indicated on the right of the mouse genetic map and is discussed in the text. Note that for all of the mouse mutants an homologous human phenotype could lie in one of two conserved segments using the present information.

the conserved segment defined by DXS674/DXS679-ALAS2. The first locus, FGD1 is located in human proximal Xp within 1Mb of and distal to, ALAS2 (Nelson *et al.*, 1995). Defects in the FGD1 gene are responsible for faciogenital dysplasia or Aarskog syndrome, a disease characterised by abnormal embryonic development involving facial, skeletal and urogenital malformations (Pasteris *et al.*, 1994). *Fgd1*, has been positioned distal to *Alas2* and therefore must lie in the distal conserved segment (Pasteris *et al.*, 1995). The second locus, SMCX (previously known as XE169 and DXS1272E), is a human expressed sequence with a homologue on the Y chromosome. In man, it lies between ALAS2 and NPHL in proximal Xp (Nelson *et al.*,1995). The mouse homologue, *Smcx*, has recently been shown to lie between *DXHXS674/DXHXS679* and *Pdha1* on the mouse X chromosome (Laval *et al.*, 1996). Furthermore, *Smcx*, has been shown to lie distal to *Fgd1* (I.C. Uwechue, personal communication). These data indicate that the orientation of the distal conserved segment in the mouse is: centromere - (*Alas2/DXHXS674/DXHXS679*) - *Fgd1* - *Smcx* - telomere. ALAS2 is proximal to DXS674/DXS679 in man and, although *Alas2* and *DXHXS674/DXHXS679* have not been positioned with respect to each other on the mouse genetic map, *Alas2* is expected to be the more proximal on the mouse X chromosome. These data suggest that the segment previously defined by DXS674/DXS679-ALAS2 in human proximal Xp can now be defined by SMCX-ALAS2 (Figure 5.2).

It is of interest to note that the position of the evolutionary breakpoints which define the WASP-NPHL conserved segment appear to coincide with the position of two OAT-'like' (OATL) repeated regions, OATL1 and OATL2, on the human X

chromosome (Figure 5.2). Both OATL clusters appear to contain several copies of a processed pseudogene of OAT, the ornithine amino transferase gene which is located on human chromosome 10q (Looney *et al.*, 1987; Lafreniere *et al.*, 1991). No evidence for the presence of OATL sequences on the mouse X chromosome has been described. Further studies to investigate the organisation of the OATL1 cluster on the human X chromosome have revealed that it contains several non-OATL sequences which cross-hybridise with sequences at OATL2. Because these sequences are not present at the OAT locus on chromosome 10, it has been proposed that the two clusters have arisen from duplication of one OATL cluster and its insertion into a different region on the X chromosome (Chand *et al.*, 1995). Whether these sequences have played a role in the rearrangements which have occurred since the human and mouse X chromosomes diverged is not known. The duplication of the original OATL cluster could have been followed by an inversion event resulting in distinct OATL1 and OATL2 clusters and the conserved segment defined by WASP-NPHL. It would be interesting to determine if evolutionary breakpoints coincide with the OATL1 and OATL2 clusters and whether sequences similar to OATL are present at any of the corresponding evolutionary breakpoints on the mouse X chromosome.

5.3 Mouse Mutants as Possible Models for Inherited Diseases Associated with Proximal Xp.

To provide a good model for a human inherited disease, a mouse mutant must have a comparable phenotype and a mutation in a homologous locus. Homologous map positions can indicate if defects in mutant mice and human inherited diseases are occurring in homologous loci. It was hoped that the improved comparative map of

human proximal Xp, described in this thesis, would confirm homologies between human diseases and mouse mutations.

There are two diseases which map to human proximal Xp which are particularly relevant to this study (see 1.11). The first, Wiskott-Aldrich syndrome (WAS), is characterised by immunodeficiency, thrombocytopenia, eczema and an increased risk of malignancies (Aldrich *et al.*, 1954). The mouse mutant *sf* resembles WAS and has been proposed as the mouse homologue (Lyon *et al.*, 1990). Mapping has revealed that the *sf* gene lies between *DXWas70* and *Otc* and cosegregates with *Gata1* and *Tcfe3* (Blair, P.J. *et al.*, 1994). As discussed in section 5.2, *Wasp* lies very close to *Gata1* and is therefore, a good candidate for *sf*. To date, as no mutations have been found in the *Wasp* gene in *sf* mice (J. Derry, personal communication), it is not known if *sf* is the mouse homologue of Wiskott-Aldrich syndrome.

The second disorder, incontinentia pigmenti type 1 (IP1) is a sporadic form of IP characterised by ectodermal abnormalities (McKusick, 1990). It is associated with X:autosome translocations, with breakpoints scattered in Xp11, between the locus NPHL and the centromere (Hatchwell, 1996). The mouse mutant *Td,* which has skin abnormalities, has been proposed as the homologue of IP1 based on a map position close to *Otc* (Davisson *et al.*, 1991). Recent mapping data indicate that *Td* lies in the interval between *DXMit55* and *Xkh* (Uwechue *et al.*, 1996). If IP1 results from the disruption of a gene by an IP1 breakpoint, the mouse homologue of this gene would not be expected to lie in the *DXMit55* and *Xkh* interval (Figure 5.2). Therefore, *Td* does not appear to be a good model for IP1.

As the IP1 translocation breakpoints all lie proximal to NPHL, the murine homologue of a gene disrupted by one of them could lie in the distal, *Alas2-Smcx*, conserved segment on the mouse X chromosome. The mouse mutant *Li*, like IP1, has skin abnormalities (Cattanach, 1985) and has been positioned between *Smcx* and *DXMit121*, a microsatellite marker which lies immediately distal to *Pdha1* (I.C. Uwechue, personal communication). Further characterisation of *Li* mice and mapping of the mutation with respect to additional conserved loci would help to determine if *Li* is the homologue of IP1.

5.4 Comparative Map Surprises.

The determination of the map position of mouse homologues of loci from human proximal Xp has improved the comparative map of the human and mouse X chromosomes. There continue to be four conserved segments which compose the region from CYBB to the centromere and further conserved segments have not been identified. The comparative map can now be used to predict the position of the murine homologues of human disease loci with some accuracy.

However, recent comparative mapping studies with genes from distal human Xp have yielded unexpected results. The mouse homologues of OA1 and APXL, which lie less than 450kb apart in distal Xp near the pseudoautosomal region (PAR), (Ferrero *et al.*, 1995), were found to lie proximal to *Amg* and not in the predicted mouse X chromosomal region (Dinulos *et al.*, 1996). Therefore, these two loci must belong to a new, human-mouse X chromosome conserved segment which had not previously been identified. The mouse homologue of CLCN4, which lies close to OA1 and APXL

in man, was found to lie in the predicted position in *M. spretus* but, surprisingly, on chromosome 7 in several *M. musculus* inbred strains (Rugarli *et al.*, 1995; Palmer *et al.*, 1995). These data indicate that there may be other undetected rearrangements involving the human and mouse X chromosomes since their divergence from an ancestral chromosome. It would therefore not be prudent to assume that homologous map predictions from the comparative map of proximal Xp are always correct.

Appendix A - Scorings of Offspring from Backcross 1.

Animal	5.1a	5.1b	5.1c	5.1d	5.1e	5.1f	5.1g	5.1h	6.1a	6.1b	7.1a	7.1b	7.1c	7.1d	7.1e	7.1f	7.1g	7.1h	7.1i	7.1j	8.1a	8.1b	8.1c	8.1d	8.1e	8.1f	8.1g	8.1h	8.1i	9.1a	9.1b	9.1c	9.1d
Sex	f	f	f	f	m	m	m	m	m	m	f	f	f	f	f	f	m	m	m	m	f	f	f	f	m	m	m	m	m	f	f	f	m
DXWas70	MS	MS	M	MS	S	M	S	S	S	M	M	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	S	S	S	M	MS	M	M	M
DXHXF34	MS	MS	M	MS	S	M	S	S	S	M	nt	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	S	S	S	M	MS	M	M	M
Tfe3	MS	MS	M	nt	S	M	S	S	S	M	M	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	M	S	S	M	MS	M	M	M
Otc	MS	MS	M	MS	S	M	S	S	S	M	M	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	nt	S	S	M	MS	M	M	M
											x																						
Maoa	MS	MS	M	MS	S	M	S	S	S	M	MS	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	M	S	S	M	MS	M	M	M
Pfc	MS	MS	M	MS	S	M	S	S	S	M	MS	MS	MS	MS	M	M	S	S	S	M	M	MS	MS	MS	S	nt	S	S	M	MS	M	M	M
			x											x					x					x							x	x
Hprt	MS	MS	MS	MS	S	M	S	S	S	M	MS	MS	MS	M	M	M	S	S	M	M	M	MS	MS	M	S	nt	S	S	M	MS	M	MS	S
										x			x																				
Cf9	nt	MS	MS	MS	S	M	S	S	S	S	MS	MS	M	M	M	M	S	S	M	M	M	MS	MS	M	S	nt	S	S	M	MS	M	MS	S
Cdr	MS	MS	MS	MS	S	M	S	S	S	S	MS	MS	M	M	M	M	S	S	M	M	M	MS	MS	M	S	M	S	S	M	MS	M	MS	S
																x																	
Gabra3	MS	MS	MS	MS	S	M	S	S	S	S	MS	MS	M	M	M	MS	S	S	M	M	M	MS	MS	M	S	M	S	S	M	MS	M	MS	S
																	x							x								
Dmd	MS	MS	MS	MS	S	M	S	S	S	S	MS	MS	M	M	M	MS	M	S	M	M	M	MS	MS	MS	S	M	S	S	M	MS	M	MS	S
							x																				x						
Xlst	MS	MS	MS	MS	S	M	M	S	S	S	MS	MS	M	M	M	MS	M	S	M	M	M	MS	MS	MS	S	M	M	S	M	MS	M	MS	S
															x	x							x										
Plp	MS	MS	MS	MS	S	M	M	S	S	S	MS	MS	M	M	MS	M	M	S	M	M	M	MS	M	MS	S	M	M	S	M	MS	M	MS	S
									x					x				x				x						x				x	
Pdha1	MS	MS	MS	MS	S	M	M	S	M	S	MS	MS	M	MS	MS	M	M	M	M	M	M	M	M	MS	S	M	M	M	M	MS	M	M	S
					x																								x				
Grpr	nt	MS	MS	MS	M	M	M	S	M	S	MS	nt	M	MS	MS	M	M	M	M	M	M	M	M	nt	S	M	M	nt	S	MS	M	nt	nt
																	x																
Amg	MS	MS	MS	MS	M	M	M	S	M	S	MS	MS	M	MS	MS	M	S	M	M	M	M	M	M	MS	S	M	M	M	S	MS	M	M	S

M indicates the *M. musculus* (3H1) allele and S the *M. spretus* allele; f indicates female and m indicates male mice; nt indicates animal not tested; X indicates the position of a recombination event.

Appendix A - Scorings of Offspring 1 continued.

Animal	9.1e	9.1f	10.1a	10.1b	10.1c	10.1d	10.1e	10.1f	10.1g	10.1h	11.1a	11.1b	11.1c	11.1d	11.1e	11.1f	11.1g	11.1h	12.1a	12.1b	12.1c	12.1d	12.1e	12.1f	13.1a	13.1b	13.1c	13.1d	13.1e	13.1f	13.1g	13.1h	13.1i
Sex	m	m	f	f	m	m	m	m	m	m	f	f	f	f	f	m	m	m	f	f	f	f	m	m	f	f	f	f	f	m		f	m
DXWas70	M	M	M	M	S	M	S	nt	M	S	M	M	M	MS	MS	S	S	M	MS	MS	MS	MS	S	M	MS	M	M	M	MS	S	S	M	M
DXHXF34	M	M	M	M	S	M	S	M	M	S	M	M	M	MS	MS	S	S	M	MS	MS	MS	MS	S	M	MS	M	M	M	MS	S	S	M	nt
				x																												
Tfe3	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	S	M	MS	MS	MS	MS	S	M	MS	M	M	M	MS	S	S	M	M
																													x				
Otc	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	S	M	MS	MS	MS	MS	S	M	MS	M	M	M	M	S	S	M	M
Maoa	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	S	M	MS	MS	MS	MS	S	M	MS	M	M	M	M	S	S	M	M
																					x												
Pfc	nt	M	M	MS	S	M	S	M	M	nt	M	M	M	MS	MS	S	S	M	MS	MS	M	MS	S	M	MS	M	M	M	M	S	S	M	M
					x																					x				x			
Hprt	M	M	M	MS	M	M	S	M	M	S	M	M	M	MS	MS	S	S	M	MS	MS	M	MS	S	M	MS	MS	M	M	M	M	S	M	M
					x												x																
Cf9	M	M	M	MS	S	M	S	M	M	nt	M	M	M	MS	MS	S	M	M	MS	MS	M	MS	S	M	MS	MS	M	M	M	M	S	M	M
Cdr	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	M	M	MS	MS	M	MS	S	M	MS	MS	M	M	M	M	S	M	M
Gabra3	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	M	M	MS	MS	M	MS	S	M	MS	MS	M	M	M	M	S	M	M
																								x									
Dmd	M	M	M	MS	S	M	S	M	M	S	M	M	M	MS	MS	S	M	M	MS	MS	M	MS	S	S	MS	MS	M	M	M	M	S	M	M
	x		x									x	x						x			x									x		
Xist	S	M	MS	MS	S	M	S	M	M	S	M	MS	MS	MS	MS	S	M	M	M	MS	M	M	S	S	MS	MS	M	M	M	M	M	M	M
																																x	
Plp	S	M	MS	MS	S	M	S	M	M	S	M	MS	MS	MS	MS	S	M	M	M	MS	M	M	S	S	MS	MS	M	M	M	M	M	MS	M
				x					x		x																				x		x
Pdha1	S	M	MS	M	S	M	S	nt	S	S	MS	MS	MS	MS	MS	S	M	M	M	MS	M	M	nt	nt	nt	nt	nt	nt	M	M	S	MS	S
Grpr	S	nt	nt	M	S	M	S	M	S	S	MS	MS	MS	nt	MS	S	M	M	M	nt	M	nt	nt	S	MS	nt	M	nt	M	nt	S	nt	S
																														x			
Amg	S	M	MS	M	S	M	S	M	S	S	MS	MS	MS	MS	MS	M	M	M	M	MS	M	M	S	S	MS	MS	M	M	M	S	nt	MS	S

Appendix A - Scorings of Offspring 1 continued.

Animal	5.2a	5.2b	5.2c	5.2d	5.2e	5.2f	5.2g	5.2h	5.2i	6.2a	6.2b	6.2c	6.2d	6.2e	6.2f	6.2g	7.2a	7.2b	7.2c	7.2d	7.2e	7.2f	8.2a	8.2b	8.2c	8.2d	8.2e	8.2f	8.2g	8.2h	9.2a	9.2b
Sex	f	f	f	f	f	f	m	m	m	f	f	f	m	m	m	m	f	f	m	m	m	m	f	f	f	f	f	m	m	m	f	f
DXWas70	MS	MS	M	M	MS	M	S	S	M	nt	MS	M	M	S	M	M	M	M	S	M	S	S	M	MS	M	M	M	S	S	S	MS	M
DXHXF34	MS	MS	M	M	MS	M	S	S	M	nt	MS	M	M	S	M	M	M	M	nt	nt	S	S	M	MS	M	M	M	S	S	S	MS	M
									x																							
Tfe3	MS	MS	nt	M	MS	M	S	S	M	nt	MS	M	M	S	M	M	nt	M	S	M	S	S	M	MS	M	M	nt	S	S	S	MS	nt
Otc	MS	MS	M	M	MS	M	S	S	S	nt	MS	M	M	S	M	M	M	M	S	M	S	S	M	MS	M	M	nt	nt	S	S	nt	M
		x																		x												
Maoa	MS	M	M	M	MS	M	S	S	S	nt	MS	M	M	S	M	M	M	M	S	S	S	S	M	MS	M	M	M	S	S	S	MS	M
Ptc	MS	M	M	M	MS	M	S	S	S	nt	MS	M	M	S	M	M	M	M	S	S	S	S	M	MS	M	M	M	S	S	S	MS	M
															x																x	
Hprt	MS	M	M	M	MS	M	S	S	S	nt	MS	M	M	S	S	M	M	M	S	S	S	S	M	MS	M	M	M	S	S	S	M	M
																									x							
Cf9	MS	M	M	M	MS	M	S	S	S	nt	MS	M	nt	S	nt	M	M	M	nt	nt	nt	nt	M	nt	nt	nt	nt	nt	nt	nt	nt	nt
							x																									
Cdr	MS	M	M	M	MS	M	M	S	S	nt	MS	M	M	S	S	M	M	M	S	S	S	S	M	MS	MS	M	M	S	S	S	M	M
				x																												
Gabra3	MS	M	M	MS	MS	M	M	S	S	nt	MS	M	M	S	S	M	M	M	S	S	S	S	M	MS	MS	M	M	S	S	S	M	M
																													x			
Dmd	MS	M	M	MS	MS	M	M	S	S	nt	MS	M	M	S	S	M	M	M	S	S	S	S	nt	MS	MS	M	M	S	M	S	M	M
																											x					
Xist	MS	M	M	MS	MS	M	M	S	S	nt	MS	M	M	S	S	M	M	M	S	S	S	S	M	MS	MS	M	MS	S	M	S	M	M
					x											x							x	x								
Plp	MS	M	M	MS	M	M	M	S	S	nt	MS	M	M	S	S	S	M	M	S	S	S	S	MS	M	MS	M	MS	S	M	S	M	M
			x														x									x						
Pdha1	MS	M	MS	MS	M	M	M	S	S	nt	MS	M	M	S	S	S	MS	M	S	S	S	S	MS	M	MS	MS	MS	S	M	S	M	M
Grpr	MS	M	MS	MS	M	M	M	nt	S	nt	MS	M	M	S	S	nt	MS	M	nt	S	S	nt	nt	nt	nt	MS	MS	nt	nt	nt	nt	M
																													x			
Amg	MS	M	MS	MS	M	M	M	S	S	nt	MS	M	M	S	S	S	MS	M	S	S	S	S	MS	M	MS	MS	nt	S	S	S	M	M

Appendix A - Scorings of Offspring 1 continued.

Animal	10.2a	10.2b	10.2c	10.2d	10.2e	10.2f	10.2g	10.2h	12.1a	12.2b
Sex	f	f	f	f	f	f	m	m	f	f
DXWas70	M	M	MS	MS	M	MS	M	M	M	M
DXHXF34	M	M	nt	MS	M	MS	M	M	M	M
Tfe3	M	M	MS	MS	M	MS	M	M	M	M
Otc	M	M	nt	nt	M	MS	M	M	M	M
Maoa	M	M	MS	MS	M	MS	M	M	M	M
Pfc	M	M	MS	MS	M	MS	M	M	M	M
	x					x				x
Hprt	MS	M	MS	MS	M	M	M	M	M	MS
Cf9	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt
Cdr	MS	M	MS	MS	M	M	M	M	M	MS
Gabra3	MS	M	MS	MS	M	M	M	M	M	MS
Dmd	MS	M	MS	MS	M	M	M	M	M	MS
							x			
Xist	MS	M	MS	MS	M	M	S	M	M	MS
Plp	MS	M	MS	MS	M	M	S	M	M	MS
									x	
Pdha1	MS	M	MS	MS	M	M	S	M	MS	MS
Grpr	MS	M	nt	MS	M	M	S	M	MS	MS
Amg	MS	M	nt	MS	nt	nt	S	M	MS	MS

Appendix B - Scorings of Offspring from Backcross 2.

Animal	8	9	19	20	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
Sex	f	m	m		f		f		f	m	f	m	m	f	f									m		m
DXWas70	MS	M	M	M	M	M	M	M	M	S	MS	S	S	MS	MS	M	M	M	M	M	M	M	M	S	M	S
DXHXF34	MS	M	M	M	M	M	M	M	M	S	MS	S	S	MS	MS	M	M	M	M	M	M	M	M	S	M	S
					x																					
Tfe3	MS	M	M	M	MS	M	M	M	M	S	MS	S	S	MS	MS	M	M	M	M	M	M	M	M	S	M	S
Gata1	MS	M	M	M	MS	M	M	M	M	S	MS	S	S	MS	MS	M	M	M	M	M	M	M	M	S	M	S
Pfc	MS	M	M	M	MS	M	M	M	M	S	MS	S	S	MS	MS	M	M	M	M	M	M	M	M	S	M	S
	x												x													
Cdr	M	M	M	M	MS	M	M	M	M	S	MS	S	M	MS	MS	M	M	M	M	M	M	M	M	S	M	S
									x	x		x														x
Xist	M	M	M	M	MS	M	M	M	S	M	MS	M	M	MS	MS	M	M	M	M	M	M	M	M	S	nt	nt
Plp	M	M	M	M	MS	M	M	M	S	M	MS	M	M	MS	MS	M	M	M	M	M	M	M	M	S	M	M
		x	x				x				x															
Pdha1	M	S	S	M	MS	M	MS	M	S	M	M	M	M	MS	MS	M	M	M	M	M	M	M	M	S	M	M
Amg	M	S	S	M	MS	M	MS	M	S	M	M	M	M	MS	MS	M	M	M	M	M	M	M	M	S	M	M

M indicates the *M. musculus* (An or 3H1) allele and S the *M. spretus* allele; f indicates female and m indicates male mice; nt indicates animal not tested; X indicates the position of a recombination event.

Appendix B - Scorings of Offspring from Backcross 2 continued.

Animal	47	48	49	50	52	64	66	68	69	70	71	72	74	75	77	78	79	80	82	83	85	87	88	89	90	92
Sex	m	m	m	f	f	f		f	m		f	f	f	m		m	m		f	f	f	m	f	m	m	f
DXWas70	S	S	S	MS	MS	MS	M	M	S	M	nt	MS	MS	S	M	M	M	M	MS	MS	MS	M	MS	M	S	MS
DXHXF34	S	S	S	MS	MS	MS	M	M	S	M	M	MS	MS	S	M	M	M	M	MS	MS	MS	M	MS	M	S	MS
					x																x					
Tfe3	S	S	S	MS	M	MS	M	M	S	M	M	MS	MS	S	M	M	M	M	MS	MS	M	M	MS	M	S	MS
Gata1	S	S	S	MS	M	MS	M	M	S	M	M	MS	MS	S	M	M	M	M	MS	MS	M	M	MS	M	S	MS
																	x						x			
Pfc	S	S	S	MS	M	MS	M	M	S	M	M	MS	MS	S	M	M	S	M	MS	MS	M	M	M	M	S	MS
									x			x										x		x		x
Cdr	S	S	S	MS	M	MS	M	M	M	M	M	M	MS	S	M	M	S	M	MS	MS	M	S	M	S	S	M
								x						x						x						
Xist	S	S	S	MS	M	MS	M	MS	M	M	M	M	MS	M	M	M	S	M	MS	M	M	S	M	S	S	M
	x																				x					
Plp	M	S	S	MS	M	MS	M	MS	M	M	M	M	MS	M	M	M	S	M	MS	M	MS	S	M	S	S	M
Pdha1	M	S	S	MS	M	MS	M	MS	M	M	M	M	MS	M	M	M	S	M	MS	M	MS	S	M	S	S	M
																x										
Amg	M	S	S	MS	M	MS	M	MS	M	M	M	M	MS	M	M	S	S	M	MS	M	MS	S	M	S	S	M

Appendix B - Scorings of Offspring from Backcross 2 continued.

Animal	93	99	109	110	117	118	119	120	121	122	123	124	127	128	131	132	133	134	135	136	137	138	139	140	141	142
Sex		f	f	f	m				m		f		m		f	f	m	m		m	f		m	f	m	m
DXWas70	M	MS	M	M	M	M	M	M	M	M	M	nt	nt	M	nt	MS	S	M	M	nt	M	M	S	M	S	S
DXHXF34	M	MS	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	M	M	S	M	S	S
Tfe3	M	MS	M	M	M	M	M	M	M	M	M	M	S	M	MS	MS	MS	M	M	S	M	M	S	M	S	S
Gata1	M	MS	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	M	M	S	M	S	S
Pfc	M	MS	M	M	M	M	M	M	M	M	M	M	S	M	MS	MS	S	M	M	S	M	M	S	M	S	S
					x						x					x										
Cdr	M	MS	M	M	S	nt	nt	nt	nt	nt	nt	nt	S	nt	nt	nt	S	nt	nt	nt	nt	nt	S	nt	S	S
		x	x										x				x							x		x
Xist	M	M	MS	M	S	M	M	M	M	M	MS	M	M	M	MS	M	M	M	M	S	M	M	S	S	S	M
																		x			x					
Plp	M	M	MS	M	S	M	M	M	M	M	MS	M	M	M	MS	MS	M	S	M	S	MS	M	S	S	S	M
		x		x		x																				
Pdha1	M	MS	MS	MS	S	S	M	M	M	M	MS	M	M	M	MS	MS	M	S	M	S	MS	M	S	S	S	M
									x															x		
Amg	M	MS	MS	MS	S	S	M	M	S	M	nt	M	nt	nt	nt	nt	nt	nt	nt	nt	MS	M	S	M	S	M

Appendix B - Scorings of Offspring from Backcross 2 continued.

Animal	143	144	145	146	147	148
Sex	m	f		f	f	m
DXWas70	S	nt	M	MS	MS	S
DXHXF34	S	nt	nt	nt	nt	nt
Tfe3	nt	MS	nt	MS	MS	S
Gata1	S	MS	M	MS	MS	S
Pfc	nt	MS	nt	MS	MS	S
Cdr	S	nt	nt	nt	nt	S
Xist	S	MS	M	MS	MS	S
	x			x	x	x
Plp	M	MS	M	M	M	M
Pdha1	M	MS	M	M	M	M
Amg	M	nt	M	M	M	M

References

Abbott, C.M., Blank, R., Eppig, J.T., Fiedorek, Jr., F.T., Frankel, W., Friedman, J.M., Huppi, K.E., Jackson, I., Steel, K., Mock, B.A., Stoye, J. and Seman, W. (1993). Mouse Chromosome 4. *Mamm. Genome* 4:S58-S71.

Aldrich, R.A., Steinberg, A.G. and Campbell, D.C. (1954). Pedigree demonstrating a sex-linked recessive condition characterized by draining ears, eczematoid dermatitis and bloody diarrhea. *Pediatrics* 147:133-139.

Amar, L.C., Dandolo, L., Hanauer, A., Ryder Cook, A., Arnauld, D., Mandel, J-L. and Avner, P. (1988). Conservation and reorganization of loci on the mammalian X chromosome: A molecular framework for the identification of homologous subchromosomal regions in man and mouse. *Genomics* 2:220-230.

Avner, P., Amar, L., Arnaud, D., Hanauer, A. and Cambrou, J. (1987). Detailed ordering of markers localizing to the Xq26-Xqter region of the human X chromosome by the use of an interspecific *Mus musculus* mouse cross. *Proc. Natl. Acad. Sci. USA* 84:1629-1633.

Avner, P., Amar, L., Dandolo, L. and Guénet, J-L. (1988). Genetic analysis of the mouse using interspecific crosses. *Trends. Genet.* 4:18-23.

Bailey, N.T.J. (1959). 'Statistical methods in biology.' The English Universities Press Ltd, London.

Beckmann, H., Su, L-K. and Kadesch, T. (1990). TFE3:A helix-loop-helix protein that activates transcription through the immunoglobin enhancer μE3 motif. *Genes and Development* 4:167-179.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.

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.

Blair, H.J., Reed, V., Laval, S.H. and Boyd, Y. (1993). The locus for pyruvate dehydrogenase E1 α subunit (*Pdha1*) lies between *Plp* and *Amg* on the mouse X chromsome. *Mamm. Genome* 4:230-233.

Blair, H.J., Reed, V., Laval, S.H. and Boyd, Y. (1994). New insights into the man-mouse comparative map of the X chromosome. *Genomics* 19:215-220.

Blair, P.J., Carpenter, D.A., Godfrey, V.L., Russell, L.B., Wilkinson, J.E. and Rinchik, E.M. (1994). The mouse scurfy (*sf*) mutation is tightly linked to *Gata1* and *Tfe3* on the mouse proximal X chromosome. *Mamm. Genome* 5:652-654.

Bonhomme, F., Catalan, J., Britton-Davidian, J., Chapman, V.M., Moriwaki, K., Nevo, E. and Thaler, L. (1984). Biochemical diversity and evolution in the genus *Mus*. *Biochem. Genet.* 22:275-303.

Boyd, Y., Blair, H., Holt, S., Laval, S.H., Maslen, G.L., Reed, V. and Uwechue, I.C. (1994). Detailed comparative mapping, in man and mouse, of the conserved segments which compose proximal Xp. *Cytogenet. Cell Genet.* 67:331-332 (Abstract).

Breen, M., Deakin, L., MacDonald, B., Miller, S., Sibson, R., Tattelin, E., Avner, P., Bourgade, F., Guénet, J-L., Montagutelli, X., Poirier, C., Simon, D., Tailor, D., Bishop, M., Kelly, M., Rysavy, F., Rasten, S., Norris, D., Shepherd, D., Abbott, C., Pilz, A., Hodge, S., Jackson, I., Boyd, Y., Blair, H., Maslen, G., Todd, J.A., Reed, P.W., Stoye, J., Ashworth, A., McCarthy, L., Cox, R., Schalkwyk, L., Lehrach, H., Klose, J., Gangadharan, U. and Brown, S. (1994). Towards high resolution maps of the mouse and human genomes - a facility for ordering markers to 0.1cM resolution. *Hum. Mol. Genet.* 3:621-627.

Breukel, C., Wijnen, J., Tops, C., v/d Klift, H., Dauwerse, H. and Meera Khan, P. (1990). Vector-Alu PCR: a rapid step in mapping cosmids and YACs. *Nucleic Acids*

Res. 18: 3097.

Britton, J. and Thaler, L. (1978). Evidence for the presence of two sympatric species of mice (genus *Mus* L.) in Southern France based on biochemical genetics. *Biochem. Genet.* 16:213-225.

Brockdorff, N., Cross, G.S., Cavanna, J.S., Fisher, E.M.C., Lyon, M.F., Davies, K.E. and Brown, S.D.M. (1987). The mapping of a cDNA from the human X-linked Duchenne muscular dystrophy gene to the mouse X chromosome. *Nature* 328:166-168.

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D. and Rastan, S. (1991). Conservation of position and exclusive expression of the mouse Xist from the inactive X chromosome. *Nature* 351:329-332.

Brown, R.M., Dahl, H-H. and Brown, G.K. (1989). X chromosome localization of the functional gene for the E1α subunit of the human pyruvate dehydrogenase complex. *Genomics* 4:174-181.

Brown, R.M., Dahl, H-H. and Brown, G.K. (1990). Pyruvate dehydrogenase E1α subunit genes in the mouse: mapping and comparison with human homologs. (1990). *Som. Cell Mol. Genet.* 16:487-492.

Brown, S. (1995). The European Collaborative Backcross Database - *MBx* - is now available on the World Wide Web. *Mamm. Genome* 6:223.

Buckle, V.J., and Craig, I.W. (1986). in Human Genetic Diseases - a practical approach, Davies K.E. (eds) IRL Press, Oxford. p 85 - 100.

Buckle, V.J., Fujita, N., Ryder-Cook, A.S., Derry, J.M.J., Barnard, P.J., Lebo, R.V., Schofield, P.R., Seeburg, P.H., Bateson, A.N., Darlison, M.G. and Barnard, E.A. (1989). Chromosomal location of GABAA receptor subunit genes: relationship to human genetic disease. *Neuron* 3:647-654.

Burke, D.T., Carle, G.F. and Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236:806-812.

Burmeister, M. and Lehrach, H. (1986). Long-range restriction map around the Duchenne muscular dystrophy gene. *Nature* 324:582-585.

Capecchhi, M.R. (1989). The new mouse genetics: Altering the genome by gene targeting. *Trends Genet.* 5:70-76.

Cattanach, B.M. (1982). Location of tattered (Td). *Mouse News Lett.* 67:19.

Cattanach, B.M. (1985). Close linkage between Li and Hyp. *Mouse News Lett.* 73:17.

Chand, A., Clark, J., Cooper, C.S. and Craig, I.W. (1995). Long-range organisation of reiterated sequences, including the SSX1 cDNA, at the OATL1 cluster in Xp11.23. *Genomics* 30:545-552.

Chapman, V.M., Keitz, B.T., Disteche, C.M., Lau, E.C. and Snead, M.L. (1991a). Linkage of amelogenin (Amel) to the distal portion of the mouse X chromosome. *Genomics* 10:23-28.

Chapman, V.M., Stephenson, D.A., Mullins, L.J., Keitz, B.T., Disteche, C. and Orkin, S.H. (1991b). Linkage of the erythroid transcription factor gene (*Gf-1*) to the proximal region of the mouse X chromosome of mice. *Genomics* 9:309-313.

Chapman, V.M., Keitz, B.T. and Bishop, D.F. (1994). Genetic linkage of the erythroid-specific δ-aminolevulinate synthase gene (*Alas2*) to the distal region of the the mouse X chromosome. *Mamm. Genome* 5:741.

Chatterjee, A., Faust, C.J., Molinari-Storey, L., Kiochis, P., Poustka, A. and Herman, G.E. (1994). A 2.3Mb yeast artificial chromosome contig spanning from *Gabra3* to *G6pd* on the mouse X chromosome. *Genomics* 21:49-57.

Cotter, P.D., Willard, H.F., Gorski, J.L.and Bishop, D.F. (1992). Assignment of the human erythroid δ-aminolevulinate synthase (ALAS2) to the distal subregion of band Xp11.21 by PCR analysis of somatic cell hybrids containing X;autosome translocations. *Genomics* 13:211-212.

Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. *Nature* 335:184-186.

Dautigny, A., Mattei, M-G., Morello, D., Alliel, P.M., Pham-Dinh, D., Amar, L., Arnaud, D., Guénet, J-L., Jollès, P. and Avner, P. (1986). Structural gene coding for myelin-associated proteolipid protein is mutated in jimpy mice. *Nature* 321:867-869.

Davisson, M.T., Lalley, P.A., Peters, J., Doolittle, D.P., Hillyard, A.L. and Searle, A.G. (1991) Report of the comparative committee for human mouse and other rodents. *Cytogenet. Cell Genet.* 58:1152-1189.

Derry, J.M.J., Lan, N.C., Shih, J.C., Barnard, E.A. and Barnard, P.J. (1989). Localization of the monoamine oxidase A and B genes on the mouse X chromosome. *Nucleic Acids Res.* 17:8403

Derry, J.M., Ochs, H.D. and Francke, U. (1994). Isolation of a novel gene mutated in Wiskott-Aldrich Syndrome. *Cell* 78:635-644.

Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D.C., Wessel, M., Dredge, R.D., Marquis, A., Stein, L.D., Goodman, N., Page, D.C. and Lander, E.S. (1994). A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genet.* 7:220-245.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M., Damron-Boles, Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., O'Connor, T.J., Evans, C.A., DeAngelis, M.M., Levinson, D.M., Kruglyak, L, Goodman, N., Copeland, N.G., Jenkins, N.A., Hawkins, T.L., Stein, L., Page, D.C. and Lander, E.S. (1996). A comprehensive genetic map of the mouse genome. *Nature* 380:149-152.

Dinulos *et al.*, 1996 see end of list.

Disteche, C.M., Gandy, S.L. and Adler, D.A. (1987). Translocation and amplification of an X-chromosome repeat in inbred strains of mice *Nucleic Acids Res.* 15:4393-4401.

Doolittle, D.P., Davisson, M.T., Roderick, T.H. and Hillyard, A.L. (1990). The locus map of the mouse. *Mouse Genome* 87:14-27.

Dorin, J.R., Dickinson, P., Alton, W.F.W., Smith, S.N., Geddes, D.M., Stevenson, B.J., Kimber, W.L., Fleming, S., Clarke, A.R., Hooper, M.L., Anderson, L., Beddington, R.S.P. and Porteous, D.J. (1992). Cystic fibrosis in the mouse by targeted insertional mutagenesis. *Nature* 359:211-215.

Dracopoli, N.C., Bruns, G.A.P., Brodeur, G.M., Landes, G.M., Matise, T.C., Seldin, M.F., Vance, J.M. and Weith, A. (1994). Report of the first international workshop on human chromosome 1 mapping. *Cytogenet. Cell Genet.* 67:144-174.

Edwards, J.H. (1991). The Oxford grid. *Ann. Hum. Genet.* 55:17-31.

Faust, C.J. and Herman, G.E. (1991). Physical mapping of the loci *Gabra3*, *DXPas8*, *CamL1* and *Rsvp* in a region of the mouse X chromosome homologous to human Xq28. *Genomics* 11:154-164.

Ferrero, G.B., Franco, B., Roth, E.J., Firulli, B.A., Borsani, G., Delmas-Mata, J., Weissenbach, J., Halley, G., Schlessinger, D., Chinault, A.C., Zoghbi, H.Y., Nelson, D.L. and Ballabio, A. (1995). An integrated physical and genetic map of a 35Mb region on chromosome Xp22.3-Xp21.3. *Hum. Mol. Genet.* 4:1821-1827.

Ferris, S.D., Sage, R.D., Prager, E.M., Ritte, U. and Wilson, A.C. (1983). Mitochondrial evolution in mice. *Genetics* 105: 681-721.

Festing, M.F.W., (1996). In 'Genetic Variants and Strains of the Laboratory Mouse, Volume 2.' Eds. Lyon, M.F., Rastan, S. and Brown, S.D.M. Oxford University Press. pp1537-1576.

Fincham, J.R.S. (1983). 'Genetics.' J. Wright and Sons Ltd, Bristol, UK.

Fisher, G. and Tease, C. (1992). The position of the locus *DXWas70* in relation to 3 X chromosome rearrangements. *Mouse Genome* 90:430-432.

Fisher, S.E., Black, G.C.M., Lloyd, S.E., Hatchwell, E., Wrong, O., Thakker, R.V. amd Craig, I.W. (1994). Isolation and partial characterization of a chloride channel gene which is expressed in kidney and is a candidate for Dent's disease (an X-linked hereditary nephrolithiasis). *Hum. Mol. Genet.* 3:2053-2059.

Frohman, M.A. and Martin, G.R. (1989). Cut, paste, and save: New approaches to altering specific genes in mice. *Cell* 56:145-147.

Goodfellow, P., Banting, G., Levy, R., Povey, S., McMichael, A. (1980). A human X-linked antigen defined by monoclonal antibody. *Somat. Cell Genet.* 6:777-787.

Gorski, J.L., Burright, E.N., Harnden, C.E., Stein, C.K., Glover, T.W. and Reyner, E.L. (1991). Localization of DNA sequences to within a region within Xp11.21 between incontinentia pigmenti (IP1) X-linked chromosomal breakpoints. *Am. J. Hum. Genet.* 48:53-64.

Goundis, D., Holt, S., Boyd, Y. and Reid, K.B.M. (1989). Localization of the properdin structural locus to Xp11.23-Xp21.1. *Genomics* 5:56-60.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994). The 1993-94 Généthon human genetic linkage map. *Nature Genet.* 7: 246-339.

Haldane, J.B.S., Sprunt, A.D. and Haldane, N.M. (1915). Reduplication in mice. *J. Genet.* 5:133-135.

Hamvas, R.M.J., Larin, Z., Brockdorff, N., Rasten, S., Lehrach, H., Chartier, F.L. and Brown, S.D.M. (1993). YAC clone contigs surrounding the *Zfx* and *Pola* loci on the

mouse X chromosome. *Genomics* 17:52-58.

Hatchwell, E. (1994). D.Phil. thesis, University of Oxford.

Hearne, C.M., Ghosh, S. and Todd, J.A. (1992). Microsatellites for linkage analysis of genetic traits. *Trends Genet.* 8: 288-294.

Herman, G.E., Boyd, Y., Chapman, V., Chatterjee, A. and Brown, S.D.M. (1994). The mouse X chromosome. *Mamm. Genome* 5:5276-5288.

Herman, G.E., Blair, H.J., Boyd, Y., Brown, S.D.M., de Gouyon, B., Haynes, A. and Quaderi, N. (1996). The mouse X chromosome *Mamm. Genome* in press.

Hirst, M.C., Bell, M.V., MacKinnon, R.N., Watson, J.E.V., Callen, D., Sutherland, G., Dahl, N., Patterson, M.N., Schartz, C., Ledbetter, D. and Davies, K.E. (1991). Mapping of a cerebellar degeneration related protein and DXS304 around the fragile site. *Am. J. Hum. Genet.* 38:354-356.

Ho, M., Chelly, J., Carter, N., Danek, A., Crocker, P. and Monaco, A.P. (1994). Isolation of the gene for McLeod syndrome that encodes a novel membrane transport protein. *Cell* 77:869-880.

Hogan, B., Costantini, F. and Lacy, E. (1986). 'Manipulating the mouse embryo: A laboratory manual.' Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Huebner, K., Isobe, M., Gasson, J.C., Golde, D.W. and Croce, C.M. (1986). Localization of the gene encoding human erythroid-potentiating activity to chromosome region Xp11.1-Xp11.4. *Am. J. Hum. Genet.* 38:819-826.

Ish-Horowicz, D. and Burke, J.F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Res.* 9:2989-2998.

Jones, J., Peters, J., Rasberry, C. and Cattanach, B.M. (1989). X-inactivation of the *Sts*

locus in the mouse: an anamaly of the dosage compensation mechanism. *Genet. Res.* 53:193-199.

Kwan, S.P., Lehner, T., Hagemann, T., Lu, B., Blaese, M., Ochs, H., Wedgwood, R., Craig, I.W. and Rosen, F.S. (1991). Localization of the gene for Wiskott-Aldrich syndrome between two flanking markers, TIMP and DXS255, on Xp11.22-Xp11.3. *Genomics* 10:29-33.

Kwan, S.P., Hagemann, T., Blaese, R.M. and Roen, F.S. (1995). A high-resolution map of genes, microsatellite markers, and new dinucleotide repeats from UBE1 to the GATA locus in the region Xp11.23. *Genomics* 29:247-252.

Lafreniere, R.G., Geraghty, M.T., Valle, D., Shows, T.B. and Willard, H.F. (1991). Ornithine aminotransferase-related sequences map to two nonadjacent intervals on the human X chromosome short arm. *Genomics* 10:276-279.

Lan, N.C., Heinzmann, C., Gal, A., Klisak, I., Orth, U., Lai, E., Grimsby, J., Sparkers, R.S., Mohandas, T. and Shih, J.C. (1989). Human monoamine oxidase A and B genes map to Xp11.23 and are deleted in a patient with Norrie disease. *Genomics* 4:552-559.

Larin, Z., Monaco, A.P. and Lehrach, H. (1991). Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl. Acad. Sci. USA* 88:4123-4127.

Lau, E.C., Mohandas, T.K., Shapiro, L.J., Slavkin, H.C. and Snead, M.L. (1989). Human and mouse amelogenin gene loci are on the sex chromosomes. *Genomics* 4:162-168.

Laval, S.H., Chen, Z-Y. and Boyd, Y. (1991). The properdin structural locus (*Pfc*) lies close to the locus for tissue inhibitor of metallothione proteases (*Timp*) on the mouse X chromosome. *Genomics* 10:1030-1034.

Laval, S.H., Blair, H.J., Hirst, M.C., Davies, K.E. and Boyd, Y. (1992). Mapping of

FMR1, the gene implicated in fragile X-linked mental retardation, on the mouse X chromosome. *Genomics* 12:818-821.

Laval, S.H. and Boyd, Y. (1993a). Novel sequences conserved on the human and mouse X chromosome. *Genomics* 15:483-491.

Laval, S.H. and Boyd, Y. (1993b). Partial inversion of gene order within a homologous segment on the X chromosome. *Mamm. Genome* 4:119-123.

Laval, S.H., Blair, H.J., Mitchell, M.J. and Boyd, Y. (1996). *Smcx* lies distal to *DXHXS679* and *DXHXS679* on the mouse X chromosome. *Mamm. Genome* in press.

Lindgren, V., de Martinville, B., Horwich, A.L., Rosenberg, L.E. and Francke, U. (1984). Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. *Science* 226:698-700.

Looney, J., Mitchell, G., Brody, L., Suchanek, M., Steel, G., Willard, H. and Valle, D. (1987). Ornithine aminotransferase (OAT) hybridizing regions on the X chromosome: Integration of at least one processed gene on X and subsequent genomic duplication has generated multiple OAT-hybridizing fragments. *Am. J. Hum. Genet.* 41:A226 (abstract).

Lyon, M.F. (1988) The William Allan memorial award address: X-chromosome inactivation and the location and expression of X-linked genes. *Am. J. Hum. Genet.* 42:8-16.

Lyon, M.F., Peters, J., Glenister, P.H., Ball, S. and Wright, E. (1990) The scurfy mouse mutant has previously unrecognized hematological abnormalities and resembles Wiskott-Aldrich syndrome. *Proc. Natl. Acad. Sci. USA* 87:2433-2437.

Lyon, M. F., Rastan, S. and Brown, S.D.M. eds. (1996). 'Genetic Variants and Strains of the Laboratory Mouse.' Oxford University Press.

McKusick,V.A. (1990) Mendelian inheritance in man. John Hopkins University Press. 9th edition. pp1656-1657.

Maslen, G. L. and Boyd, Y. (1993) Comparative mapping of the *Grpr* locus on the mouse X chromosomes of man and mouse. *Genomics* 17:106-109.

Maslen, G.L. (1995). D.Phil. thesis, University of Oxford.

Merrell, K., Gonzales, J.C., Wells, S., Calame, K. and Herman, G.E. (1995). Genetic analyses of Tattered, an X-linked dominant, developmental mouse mutation. *Mamm. Genome* 6:291-294.

Mullins, L.J., Grant, S.G., Stephenson, D.A. and Chapman, V.M. (1988). Multilocus molecular map of the mouse X chromosome. *Genomics* 3:187-194.

Nadeau, J.H. (1989). Maps of linkage and synteny homologies between mouse and man. *Trends Genet.* 5: 82-86.

Nadeau, J.H. and Taylor, B.A. (1984). Lengths of chromosomal segments conserved since divergence of mouse and man. *Proc. Natl. Acad. Sci. USA* 81:814-818.

Nelson, D.L., Ballabio, A., Cremers, F., Monaco, A.P. and Schlessinger, D. (1995). Report of the sixth international workshop on X chromosome mapping 1995. *Cytogenet. Cell Genet.* 71:308-342.

Nielsen, J.T. and Chapman, V.M. (1977). Electrophoretic variation for X-chromosome-linked phosphoglycerate kinase (PGK-1) in the mouse. *Genetics* 87: 319-325.

Ohno, S. (1969). Evolution of sex chromosomes in mammals. *Ann. Rev. Genet.* 3: 495-524.

Ohno, S. (1973). Ancient linkage groups and frozen accidents. *Nature* 244: 259-262.

Olson, M., Hood, L., Cantor, C. and Botstein, D. (1989). A common language for physical mapping of the human genome. *Science* 245:1434-1435.

Özçelik, T., Lafreniere, R.G., Archer, B.T., Johnston, P.A., Willard, H.F., Francke, U. and Südhof, T.C. (1990). Synaptophysin: Structure of the human gene and assignment to the X chromosome in man and mouse. *Am. J. Hum. Genet.* 47:551-561.

Palmer, S., Perry, J. and Ashworth, A. (1995). A contravention of Ohno's law in mice. *Nature Genet.* 10:472-476.

Pasteris, N.G., Cadle, A., Logie, L.J., Porteous, M.E.M., Schwartz, C.E., Stevenson, R.E., Glover, T.W., Wilroy, R.S. and Gorski, J.L. (1994). Isolation and characterization of the faciogenital dysplasia (Aarskog-Scott syndrome) gene: A putitive Rho/Rac guanine nucleotide exchanger factor. *Cell* 79:669-678.

Pasteris, N.G., de Gouyon, B., Cadle, A.B., Cambell, K., Herman, G.E. and Gorski, J.L. (1995). Cloning and regional localization of the mouse faciogenital dysplasia (*Fgd1*) gene. *Mamm. Genome* 6:658-661.

Purrello, M., Alhadeff, B., Esposito, D., Szabo, P., Rocchi, M., Truett, M., Masiarz, F. and Siniscalco, M. (1985). The human genes for hemophilia A and hemophilia B flank the X chromosome fragile site at Xq27.3. *EMBO J.* 4:725-729.

Rao, V.N., Huebner, K., Isobe, M., Ar-Rushdi, A., Croce, C.M., Shyam, E. and Reddy, P. (1989). *Elk*, tissue-specific *ets*-related genes on chromosome X and 14 near translocation breakpoints. *Science* 244:66-70.

Reed, V., Rider, S., Maslen, G.L., Hatchwell, E., Blair, H.J., Uwechue, I.C., Craig, I.W., Laval, S.H., Monaco, A.P. and Boyd, Y. (1994). A 2-Mb contig encompassing three loci (DXF34, DXS14, and DXS390) that lie between Xp11.2 translocation breakpoints associated with incontinentia pigmenti type 1. *Genomics* 20:341-346.

Riley, J., Butler, R., Ogilive, D., Finniear, R., Jenner, D., Powell, S., Anand, R.,

Smith, J.C. and Markham, A.F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* 18:2887-2890.

Rinchik, E.M. (1991). Chemical mutagenesis and fine-structure functional analysis of the mouse genome. *Trends Genet.* 7:15-21.

Robert, B., Barton, P., Minty, A., Daubas, P., Weydert, A., Bonhomme, F., Catalan, J., Chazottes, D., Guenet, J-L. and Buckingham, M. (1985). Investigation of genetic linkage between myosin and actin genes using an interspecific mouse back-cross. *Nature* 314:181-183.

Roman, C., Matera, A.G., Cooper, C., Artandi, S., Blain, S., Ward, D.C. and Calame, K. (1992). mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix and zipper domains, utilizes the zipper to stabilize both DNA binding and multimerization. *Mol. Cell. Biol.* 12: 817-827.

Royer-Pokora, B., Kunkel, L.M., Monaco, A.P., Newburger, P.E., Baehner, R.L., Cole, F.S., Lurnette, J.T. and Orkin, S.H. (1986). Cloning the gene for an inherited disorder-chronic granulomatous disease - on the basis of its chromosomal location. *Nature* 322:32-38.

Ruddle, F.H. and Creagan, R.P. (1975). Parasexual approaches to the genetics of man. *Ann. Rev. Genet.* 9:407-486.

Rugarli, E.I., Adler, D.A., Borsani, G., Tsuchiya, K., Franco, B., Hauge, X., Disteche, C., Chapman, V. and Ballabio, A. (1995). Different chromosomal localization of the *Clcn4* gene in *Mus spretus* and C57BL/6J mice. *Nature Genet.* 10:466-471.

Russell,W.L., Kelly, E.M. Kelly, Hunsicker, P.R., Bangham, J.W., Maddux, S.C. and Phipps, E.L. (1979). Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc. Natl. Acad. Sci. USA* 76:5818-5819.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). 'Molecular cloning - A laboratory manual,' Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sawyer, J.R. and Hozier, J.C. (1986). High resolution of mouse chromosomes: Banding conservation between man and mouse. *Science* 232:1632-1635.

Schantz, L.J., Naylor, S.L., Giladi, E. and Spindel, E.R. (1991). Assignment of the GRP receptor gene to the human X chromosome. *Cytogenet. Cell Genet.* 58:2085-2086.

Schlessinger, D. (1990). Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet.* 6: 248-258.

Schoenhaut, D.S. and Curtis, P.J. (1986). Nucleotide sequence of the mouse 5-aminolevulinic acid synthase cDNA and expression of its gene in hepatic and erythroid tissues. *Gene* 48:55-63.

Schoenhaut, D.S. and Curtis, P.J. (1989). Structure of the mouse erythroid 5-aminolevulinate synthase gene and mapping of erythroid-specific DNAse I hypersensitive sites. *Nucleic Acids Res.* 17:7013-7025.

Schwartz, D.C. and Cantor, C.R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* 37:67-75.

Searle, A.G., Beechey, C.V., Evans, E.P. and Kirk, M. (1983). Two new X-autosome translocations in the mouse. *Cytogenet. Cell Genet.* 35:279-292.

Searle, A.G., Peters, J., Lyon, M.F., Hall, J.G., Evans, E.P., Edwards, J.H. and Buckle, V.J. (1989). Chromosome maps of man and mouse. IV. *Ann. Hum. Genet.* 53: 89-140

Silverman, G.A., Jockel, J.I., Domer, P.H., Mohr, R.M., Taillon-Miller, P. and Korsmeyer, S.J. (1991). Yeast artificial chromosome cloning of a two-megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2

and plasminogen activator inhibitor type-2. *Genomics* 9:219-228.

Snouwaert, J.N., Brigman, K.K., Latour, A.M., Malouf, N.N., Boucher, R.C., Smithies, O. and Koller, B.H. (1992). An animal model for cystic fibrosis made by gene targeting. *Science* 257:1083-1088.

Southern, E.M., Anand, R., Brown, W.R.A. and Fletcher, D.S. (1987) A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* 15:5925-5664.

Tamai, Y., Taketo, M., Nozaki, M. and Seldin, M.F. (1995). Mouse *Elk* oncogene maps to chromosome X and a novel *Elk* oncogene (*Elk3*) maps to chromosome 10. *Genomics* 26:414-416.

Uwechue, I.C., Goble, R., Blair, H.J., Hacker, T., Burke, D., Herman, G. and Boyd, Y. (1996). The X-linked developmental mutant, tattered, lies between *DXMit55* and *Xkh* and is associated with hyperkeratinisation. *Genomics* submitted.

Wicking, C. and Williamson, R. (1991). From linked marker to gene. *Trends. Genet.* 7: 288-293.

Williamson, P., Holt, S., Townsend, S. and Boyd, Y. (1995). A somatic cell hybrid panel for mouse characterized by PCR and FISH. *Mamm. Genome* 6:429-432.

Zon, L.I., Tsai, S-F., Burgess, S., Matsudaira, P., Bruns, G.A.P. and Orkin, S.H. (1990). The major human erythroid DNA-binding protein (GF-1): Primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci. USA.* 87:668-672.

Dinulos, M.B., Bassi, M.T., Rugarli, E.I., Chapman, V., Ballabio, A. and Disteche, C.M. (1996). A new region of conservation is defined between human and mouse X chromosome. *Genomics* 35:244-247.

Three publications have arisen from the material presented in this thesis. These are:

Blair, H.J. and Boyd, Y. (1994). *Tfe3* maps close to *Gata1* on the mouse X chromosome. *Mouse Genome* 92:511-512.

Blair, H.J., Reed, V., Laval, S.H. and Boyd, Y. (1994). New insights into the man-mouse comparative map of the X chromosome. *Genomics* 19:215-220.

Blair, H.J., Ho, M., Monaco, A.P., Fisher, S., Craig, I.W. and Boyd, Y. (1995). High-resolution comparative mapping of the proximal region of the mouse X chromosome. *Genomics* 28:305-310.