ANALYSIS OF DMD TRANSLOCATIONS

David James Cockburn

A Thesis Submitted for the Degree of

Doctor of Philosophy

in the University of Oxford

St Catherine's College

Trinity Term 1991

To my family,

especially Rachel and Rebecca

ABSTRACT

David James Cockburn
St. Catherine's College

D.Phil.
Trinity term 1991

Analysis of DMD translocations

Duchenne and Becker muscular dystrophies (DMD, BMD) are allelic X-linked diseases which affect approximately one in 3500 male newborns. They are caused by mutations in a gene positioned on the short arm of the X chromosome at Xp21. The first indication of the location of this gene was the description of rare females expressing DMD and who were found to have constitutional X;autosome translocations with an X chromosome breakpoint at this site. There are now 24 such females known worldwide. They express DMD as a consequence of preferential inactivation of the normal X chromosome. In order to contribute to the understanding of the aetiology of mutations causing DMD and the aetiology of constitutional translocations, two types of study have been performed here. Firstly, the detailed mapping of the X chromosome breakpoints of DMD-associated X;autosome translocations has been investigated. The results of this study have been compared with data on the physical distribution of mutations causing DMD in male patients. Secondly, one translocation, an X;1 translocation with an autosomal breakpoint at 1p34, has been selected for more detailed investigation and the DNA sequence has been determined at the site of the rearrangement.

Translocation breakpoint mapping studies were performed by somatic cell hybrid analysis. Hybrids were karyotyped and this information was used to construct a hybrid panel for the purpose of determining the autosomal localisations of anonymous DNA probes. The mapping of seven probes using this panel is described.

The work described in this thesis revealed that the distribution of translocation breakpoints within the DMD gene appears to be random and may differ from the distribution of mutations in male patients. The X;1 translocation whose breakpoints are cloned and sequenced was found to involve two expressed loci, one coding for dystrophin on the X chromosome and one for the leukocyte antigen related protein on chromosome 1. Sequence data revealed that a deletion of four to seven nucleotides from the X chromosome and a duplication of two to five nucleotides are associated with the translocation. The possible involvement of trinucleotides adjacent to the breakpoints, and of a LINE, a SINE and a stretch of potential Z-DNA within 1 kb of the X chromosome or the chromosome 1 breakpoint, is discussed.

ACKNOWLEDGEMENTS

The main part of the work described in this thesis was carried out in the Genetics Laboratory, Department of Biochemistry in the University of Oxford. I should like to thank Professor J.H. Edwards for his interest and encouragement and I am grateful to the Muscular Dystrophy Group of Great Britain and Northern Ireland for financial support.

I have been extremely fortunate to be supervised throughout this project by Yvonne Boyd and Ian Craig. Despite Yvonne's work taking her to Harwell her support was undiminished and the joint supervision provided by her and Ian gave me every possible encouragement and assistance.

Those who collaborated with me on specific projects are acknowledged in the text but particular thanks must be given to Sheila Holt and Chris Porter with whom I have worked very closely. Special thanks must also be made to Elin Munro for her part in establishing the somatic cell genetics which provided the foundation of this work.

I wish to record my sincere thanks to all colleagues in the Genetics Laboratory and those at the MRC Radiobiology Unit, Harwell for their practical help and their friendship.

My family has provided understanding and help throughout. I thank them all, especially my wife, Rachel, who had our first baby in May of this year and yet has given generously of her time and skill.

CONTENTS

Chapter 1 General introduction

1.1 Historical background

The second half of the twentieth century has witnessed a huge expansion in the study and understanding of human genetics. This began with widespread analysis of the human karyotype which followed the correct identification of the diploid chromosome number as 46 (Tjio and Levan, 1956; Ford and Hamerton, 1956). For decades this had been thought to be 48 (Painter, 1923) although, as early as 1921, Painter reported on his study of spermatogonial metaphases from individuals 'castrated because of self abuse' that 'in the clearest equatorial plates so far studied only 46 chromosomes have been found' (Painter, 1921). Human chromosomes were classified into seven groups based on size and centromere position by Patau (1960). But the significant breakthrough was the development of banding techniques, first described for plant chromosomes using quinacrine dihydrochloride or quinacrine mustard, but quickly applied to human chromosomes, which allowed unambiguous identification of each chromosome (Caspersson et al., 1968; Caspersson et al., 1971). Simple reliable banding protocols are now readily available which have allowed human cytogenetics to become firmly established in clinical diagnosis (see Penn and Perle, 1986). Numerical chromosome abnormalities, also previously detected by unbanded analysis can be identified, such as Down syndrome which is associated with trisomy chromosome 21 (Lejeune et al., 1959), and the chromosomes involved in structural rearrangements such as translocations can be identified. In the last few years a role for the practice of molecular genetics in clinical diagnosis has been created, brought about by progress in the development of DNA technology and the mapping and cloning of the genes for important single-locus diseases.

Present-day DNA technology is largely dependent on the the ability to cut DNA using restriction enzymes isolated from microorganisms which recognise specific

sequence sites. The first were described by Lederberg and Meselson (1964), but particularly significant was the discovery of restriction endonucleases which always cut between the same nucleotides at the recognition site (type II restriction endonucleases; first described by Kelly and Smith, 1970; and Smith and Wilcox, 1970). DNA cloning procedures exploit the properties of the cohesive ends generated by many restriction enzymes. Cloning vehicles such as the plasmid pBR322 have been constructed which include the single cloning sites for several enzymes (Bolivar et al., 1977). Mapping of DNA regions is performed by analysis of the sizes of fragments generated by endonuclease digestion (first described for SV40 by Danna et al., 1973). Southern analysis allows the sizes of restriction fragments which hybridise to a DNA probe to be determined (Southern, 1975). This technology has been extended by using infrequently cutting enzymes and electrophoretic systems which resolve DNA fragments which are several hundred kb in length, to allow long range maps of the genome to be constructed (for instance around the DMD locus, van Ommen et al., 1987; Burmeister et al., 1988; Meitinger et al., 1988). The majority of polymorphisms described in the human genome and used in linkage studies are restriction fragment length polymorphisms (RFLPs) detected by Southern analysis. The polymorphism is caused by the presence/absence of a restriction site or the presence of a deletion/insertion polymorphism or a variable number of tandem repeats element (VNTR) within the restriction fragment.

1.2 Translocations

The frequency of constitutional translocations in the newborn is approximately 0.2%, and in spontaneous abortions 2-4% (Harper, 1988). The carriers of balanced translocations (where there is no significant gain or loss of chromosome material) are usually phenotypically normal but are at risk of producing unbalanced gametes. In approximately 5% of cases of recurrent miscarriage the woman or her partner carries a translocation. Carriers of apparently balanced translocations are occasionally

phenotypically abnormal. This might be because there is a cytologically undetectable deletion or duplication, or because the translocation disrupts a gene with a dominant disease expression. One special class of patients with disease expression and a balanced translocation is that comprising female carriers of X;autosome translocations where the X chromosome breakpoint disrupts a gene locus (Frezal and Schinzel, 1990). The expression of the disease is a consequence of a skewed X chromosome inactivation pattern, where the normal X chromosome is predominantly inactive. This feature is common to all X;autosome translocations, and in these cases leads to expression of the disruption caused by the translocation (Mattei et al., 1981; discussed further in section 1.4).

As well as constitutional translocations which affect every cell, translocations may arise in somatic tissue and are therefore only present in a proportion of cells. In particular, a number of somatic translocations have been identified and studied which are associated with human neoplasms, mainly those associated with haematological malignancies (reviewed by Croce, 1987; Cleary, 1991). Some associations are highly specific, for instance a 15;17 translocation is seen in virtually every case of acute promyelocytic leukemia and in no other malignancy (de Thé et al., 1990). Many translocations involve immunoglobulin genes and errors in the V(D)J recombinase mechanism are implicated in their aetiology (Croce, 1987).

The nomenclature used to describe translocations in this thesis complies with the report of the ISCN (1985). To clarify this terminology, a derivative chromosome is one rearranged by the translocation. The derivative X or der(X) chromosome retains the X chromosome centromere while the derivative autosome or der(autosome) retains the autosomal centromere (see figure 1.1). A translocation junction is referred to as the point on the der(X) or der(autosome) between DNA sequence from the reciprocal chromosomes, and a junction fragment as a fragment of DNA from the der(X) or der(autosome) spanning the translocation junction. A translocation breakpoint is referred to as the position on a normal chromosome corresponding to the junction

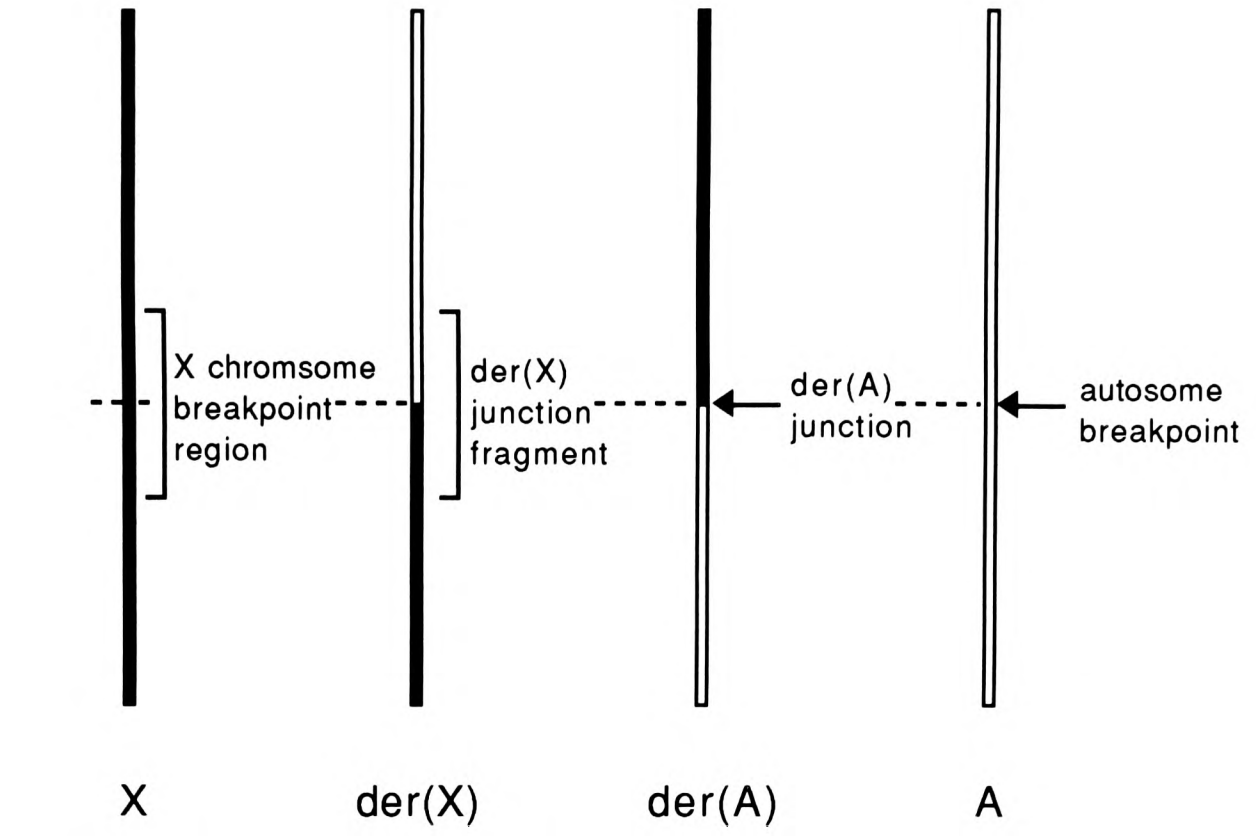


Figure 1.1 Diagram of an X;autosome translocation, illustrating the nomenclature used to describe the rearrangement. X chromosome DNA is represented by dark lines, and autosome (A) DNA by pale lines.

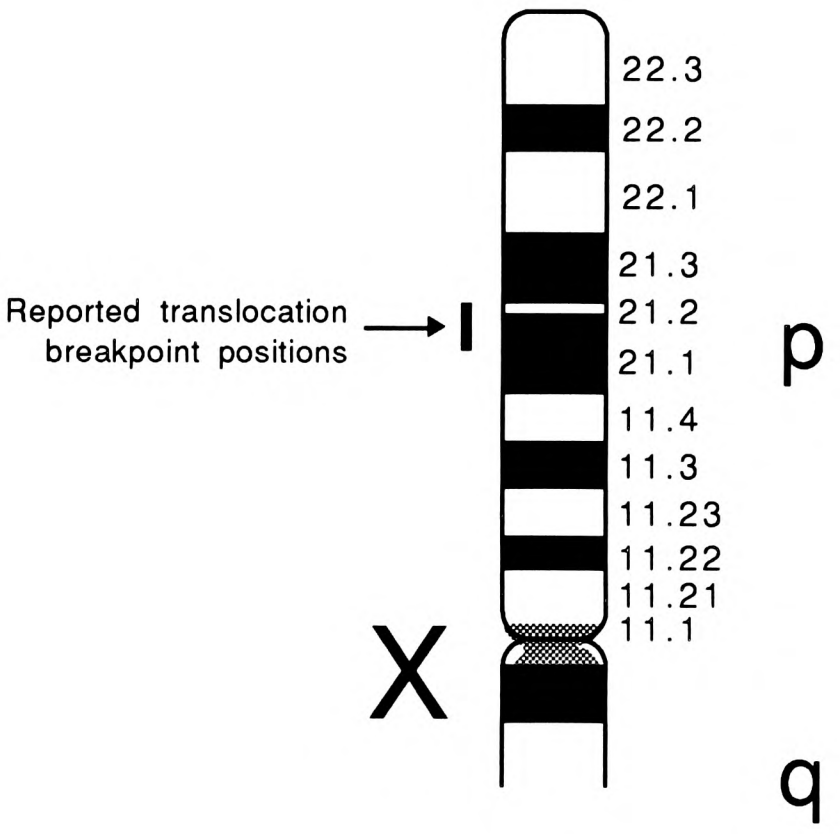


Figure 1.2 Ideogram of the short arm of the X chromosome illustrating the approximate positions of the X chromosome breakpoints of X;autosome translocations associated with DMD, and therefore indicating the likely location of the disease gene. The breakpoints are reported to lie in Xp21.1 and Xp21.2.

position, and a breakpoint region as a DNA region on a normal chromosome spanning the breakpoint.

1.3 Investigation of the DMD locus

1.3.1 Historical perspective

There is every reason to suppose that Duchenne muscular dystrophy has always existed in human populations. In his review of clinical and scientific aspects and the management of DMD, Emery (1987) includes a historical introduction presenting evidence that individuals with characteristics consistent with muscular dystrophy have been portrayed in paintings and drawings which date as far back as around 2500 BC. It was in 1868, however, that Duchenne de Boulogne described 13 cases, two of which were girls, and defined the characteristic clinical features (Duchenne, 1868). The incidence of Duchenne and the milder Becker muscular dystrophies is 1 in 3500 males, making it the most common lethal neuromuscular genetic disease (Engel and Banker, 1986). The disease is not restricted to humans although the disease pathogenesis is not identical in other animals. Dogs, cats and mice have been identified which are deficient in the protein product from the intact DMD gene, dystrophin (Bulfield et al., 1984; Cooper et al. 1988; Hoffman, 1990). There is unfortunately no reason to suppose that DMD will be eliminated in the future since a high proportion, approximately one third, of DMD patients are sporadic and have 'new' mutations (Haldane, 1935; Davie and Emery, 1978).

1.3.2 Cloning and genomic organisation of the DMD locus

In the past few years significant improvements in diagnosis, carrier detection and understanding of the disease pathogenesis have been achieved. These stem directly from the cloning of the gene which, when defective, causes DMD. The classical method for cloning disease genes has been by identifying and sequencing the protein product of the gene, and then using this knowledge to clone the cDNA and genomic sequences.

Examples include the genes for factor VIII and the globin genes (Gitschier et al., 1984; Fritsch et al., 1980). However, more recently, where the underlying biochemical defect is unknown, a reversed route has been taken. In these instances genes have been cloned from a knowledge only of their map locations, and the protein product has subsequently been determined and identified from the cDNA sequence. This method of cloning has been termed 'reverse genetics' or 'targetted cloning' (Orkin, 1986; Goodfellow, 1986). Genes for several important single-locus genetic diseases have been isolated in this manner in the last few years, including those associated with Duchenne muscular dystrophy, cystic fibrosis and fragile X-linked mental retardation (Kunkel et al., 1985; Koenig et al., 1987; Rommens et al., 1989;Verkerk et al., 1991). The isolation of the gene associated with DMD was the first successful application of targetted gene cloning. Due to the pattern of inheritance the gene was already known to be X-linked, but strong evidence for the localisation within band Xp21 was provided by the identification of cytogenetic rearrangements involving this band in affected individuals. Female patients expressing DMD and who had an X;autosome translocation were described first (Greenstein et al., 1977; Verellen et al., 1977; Lindenbaum et al., 1979). Later, a male patient expressing DMD in addition to other X-linked diseases and who had a deletion was described (Francke et al., 1985). Data supporting a localisation of DMD on the X chromosome short arm were provided by studies which showed linkage of the disease with RFLPs detected by the probes RC8 and L1.28 (Murray et al., 1982; Davies et al., 1983). The cytogenetic rearrangements involving the Xp21 region provided powerful physical tools for mapping probes which detected RFLPs or candidate gene sequences and were exploited in the cloning of DNA sequences from the gene. Kunkel et al. (1985) selectively cloned DNA absent from the male deletion patient mentioned above using a process of deletion enrichment. Two clones turned out to lie in the DMD gene. One, pERT87, was found to be deleted in several male DMD patients with normal karyotypes, and was later shown to be in an intron of the DMD gene (Monaco et al., 1985; Monaco et al., 1986). At the same time,

Ray et al. (1985) cloned the breakpoints of one of the X;autosome translocations. This was performed by exploitation of the finding that the autosomal breakpoint was within a cluster of repeated rRNA genes on the short arm of chromosome 21. These had already been cloned and characterised, and enabled the der(X) junction fragment to be detected and cloned. X chromosome clones adjacent to the breakpoint were found to be deleted in some male patients. Thus two independent strategies led to the cloning of sequences from within the DMD gene, both exploiting cytogenetic rearrangements.

There are now 24 known females with DMD/BMD and an X;autosome translocation involving Xp21 (Boyd, 1991). This figure is much higher than the number of X;autosome translocations associated with any other X-linked disorder (see Frezal and Schinzel, 1990). Six cases are known of patients with Incontinentia pigmenti and a breakpoint at Xp11.21-cen (Gorski et al., 1991). Other disorders include Anhidrotic ectodermal dysplasia with a breakpoint at Xq12.2-q13.1 (Gerald and Brown, 1974); Aicardi's syndrome with a breakpoint at Xp22 (Ropers et al., 1982); Hunter's syndrome with a breakpoint at Xq27.3-q28 (Mossman et al., 1983); Aarskog syndrome with a breakpoint at Xq13 (Bawle et al., 1984); Menke's syndrome with a breakpoint at Xq13.3 (Verga et al., 1991) and Lowe's syndrome with a breakpoint at Xq25 (Hodgson et al., 1986). The reason why so many cases associated with DMD have been identified may be due to the size of the gene locus (although the sizes of the genes associated with the above disorders are still unknown). There were signs even before the DMD gene was cloned that the locus might be very large. Wilcox et al. (1986) described a male DMD patient with a microscopically visible deletion but who did not express any other X-linked disorder. This indicated that the deletion was not sufficiently extensive to include adjacent disease gene loci and therefore that the DMD gene might be large or that adjacent disease gene loci were physically distant from the DMD locus. Boyd and Buckle (1986) studied extended chromosome preparations from nine of the DMD-associated X;autosome translocations and observed heterogeneity of X chromosome breakpoint position within Xp21. This finding was supported by others

who reported breakpoints either in Xp21.1 or Xp21.2 (see figure 1.2). Molecular analysis, using the first probes isolated from the DMD region, confirmed that the breakpoints were spread over a minimum of 176 kb and possibly a much larger region (Boyd et al., 1986). RFLP studies using probes from the pERT87 locus indicated frequent recombination between the probes and disease loci on both the proximal and distal side (Kunkel and co-authors, 1986). The cloning of the complete 14 kb cDNA revealed a multi-exonic genomic organisation. Southern analysis using cDNA fragments as probes revealed that 65 HindIII restriction fragments contain exons (Koenig et al., 1987). The physical size of the DMD locus has now been determined by long range mapping to be over 2,000 kb in length (Meitinger et al., 1988; Burmeister et al., 1988; see Chapter 4, figure 4.1). The genomic size is unprecedented, representing approximately 1% of the X chromosome. The factor VIII gene for comparison has a 9 kb cDNA and the exons are spread over 200 kb of genomic DNA (Gitschier et al., 1984).

1.3.3 The nature of DMD mutations in boys

As more genomic and cDNA probes from the DMD locus were isolated and used to screen male DMD patients the proportion of patients with detectable deletions increased. A worldwide study in 1986 indicated that 6.5% of male patients are deleted at the pERT87 locus (Kunkel and co-authors, 1986). Following the isolation of further genomic probes, a detectable deletion percentage of 17% was reported (Monaco et al., 1987), and a study by field inversion gel electrophoresis detected deletions in 50% of male patients (den Dunnen et al., 1987). Studies performed by Southern analysis and using cDNA fragments as probes indicated deletion frequencies of between 50% and 70% (Koenig et al., 1987; Forrest et al., 1987; Forrest et al., 1988). A small proportion of patients was found with duplications, as deduced from the intensity of bands in Southern analysis (Hu et al., 1988; den Dunnen et al., 1989; Koenig et al., 1989). The nature of the mutations in the remaining approximately 35% of male DMD/BMD patients remains unknown. They may mainly represent point mutations, which are

difficult to identify given the size of the gene, and only one has so far been identified (Bulman et al., 1991).

An intriguing finding from deletion mapping was that the size and position of the deletion does not appear to correlate with the disease severity (Monaco et al., 1987; Hart et al., 1987). An explanation for this proposed by Monaco et al. (1988) was that milder BMD patients have deletions removing exons but maintaining the reading frame, so that, after splicing of the mRNA, a shorter but functional protein is produced; while DMD deletions involve exons resulting in a disruption to the reading frame, and subsequently the translation of a truncated non-functional protein. This explanation has been confirmed, with a few exceptions, by more recent studies (Malhotra et al., 1988; den Dunnen et al., 1989; Koenig et al., 1989; Gillard et al., 1989); and by characterisation of dystrophin, the protein product of the DMD locus, which was generally found to be present but shorter in BMD patients and completely missing in DMD patients (Hoffman et al., 1988). A possible explanation is now available even for cases where the reading frame hypothesis was thought not to fit. Roberts et al. (1991) have recently observed, by amplification of mRNA from peripheral blood lymphocytes, that unexpected splicing of exons can occur in deletion patients so that an exon known to be present in the genomic DNA may be missing in the mRNA. Therefore the maintenance of the reading frame over a deletion cannot always be predicted from a knowledge of the exons which are deleted.

1.3.4 Dystrophin

A detailed description of the dystrophin protein is not directly relevant to the work described in this thesis and is available elsewhere (Hoffman and Kunkel, 1989). Its essential features are that it is a 3685 amino acid protein organised into four domains. The amino-terminal domain is related to α-actinin; a large, central, triple-helical domain resembles analogous domains of α-actinin and spectrin; at the carboxy-terminus are two domains, a cystein-rich domain similar to the carboxy-terminus of α-

actinin and a domain bearing no resemblance to any other known protein. Dystrophin is localised in the sarcolemma of human skeletal muscle and is thought to have a structural role (Zubrzycka-Gaarn et al., 1988). Deletions which are found in mild BMD patients are generally located within the portion of the gene encoding the central triple-helical domain of dystrophin (Hoffman and Kunkel, 1989).

Although dystrophin is mainly expressed in muscle, other tissues, most notably the brain, are also sites of expression (Hoffman et al., 1987). Approximately one third of affected males have some degree of mental retardation (reviewed in Emery, 1987). Interestingly, it has been found that alternative splicing at the carboxy-terminus generates isoforms of dystrophin in brain which differ from the predominant muscle isoforms (Feener et al., 1989); and that the brain uses an alternative promoter located more than 90 kb upstream of the muscle promoter and 400 kb from exon 2 to which it is spliced (Boyce et al., 1991).

1.4 The expression of X-linked disease in females

There are several possible explanations as to why a female might suffer from an X-linked disease such as DMD, usually only seen in males, assuming that an autosomally inherited disease with similar phenotype is not responsible. Many of these possibilities probably occur only rarely. Firstly, a female with Turner's syndrome and a 45,X karyotype may express a recessive gene on her only X chromosome. The frequency of an X-linked disease in Turner's syndrome patients might be expected to be similar to the frequency of the disease in boys. For DMD this would be 1 in 3500 Turner's syndrome patients. A girl with DMD and a 45,X karyotype was identified as long ago as 1957 (Walton, 1957). The second possibility is that of females with a normal karyotype, carrying a disease gene on each chromosome. This might appear to be very rare as it suggests the inheritance of defective genes from both parents, presumably a new mutation from the father unless he also is affected. However, this situation could also arise by uniparental disomy for the X chromosome

which would result in the expression of any recessive disease on the X chromosome. If it was paternal isodisomy, then the disomic chromosomes would have to carry a new mutation (unless the father was affected too). If it was maternal isodisomy and not due to a new mutation, then the mother would be a carrier of the X-linked disease. There is one reported case of uniparental sex chromosome disomy and this caused male to male transmission of haemophilia A (Vidaud et al., 1989). No estimates are available for the incidence of uniparental disomy for any chromosome (Schinzel, 1991). Maternal isodisomy of chromosome 7, however, has been attributed as the cause of cystic fibrosis in two individuals (Spence et al., 1988; Voss et al., 1988). Isodisomy of chromosome 7 in individuals without cystic fibrosis may also be expected to occur. The frequency of isodisomy may vary between different human chromosomes. Certainly where aneuploidy is concerned, different rates are observed for the different chromosomes whether one studies the frequencies in the newborn or in spontaneous abortions (Hamerton et al., 1975; Hassold et al., 1980; reviewed in Bond and Chandley, 1983). The differences in observed frequencies of the human chromosome aneuploidies may reflect the tolerance/spontaneous abortion rates of the different conditions and the differing susceptibilities of the chromosomes to non-disjunction (Bond and Chandley, 1983). Chromosomes which are most prone to non-disjunction might be expected to be the most prevalent uniparental disomies.

Other possible explanations for X-linked disease expression in females involve non-random X-inactivation permitting the expression of recessive genes on the active chromosome. Two pairs of twins have been fully reported where one is phenotypically normal while the other expresses DMD (Burn et al., 1986; Richards et al., 1990). A skewed X inactivation pattern was observed in both instances, so that in one twin one X chromosome was predominantly active and in the other twin the other X chromosome was predominantly active. The expression of DMD could be attributed to the inheritance of a defective maternal gene. Recent studies on X-inactivation patterns in phenotypically normal females have shown that significant deviation from a 50:50

proportion is not uncommon and the proportion varies in different tissues. This may cause variable expression in heterozygous females of X-linked disease genes which show incomplete dominance, such as pyruvate dehydrogenase E1α deficiency (Brown et al., 1990). Deviations from a 50:50 pattern of X chromsome inactivation are even more marked in individuals who have an X chromosome rearrangement. For example, in cases of X chromosome deletion in females with mild or partial Turner's syndrome, the abnormal X chromosome is usually inactive in most cells examined (Mattei et al., 1981). Where the normal X chromosome carries a recessive disease gene, this will be expressed. A girl with a structurally abnormal X chromosome and DMD has been described by Berg and Conte (1974). The reason for biased X-inactivation is probably due to a growth advantage in the early embryo of cells with an active normal X chromosome over those where the normal X is inactive. The growth disadvantage of cells with active deleted Xs can be explained by absence of products of genes from the deleted segment in these cells. The situation in the case of an X;autosome translocation is more complex. Here it has been observed in man that the normal X is inactive in almost all cells (Mattei et al. 1981). Several factors may contribute to the growth disadvantage of cells with an active normal X chromosome. Assuming an X chromosome breakpoint at Xp21 such as is observed in DMD-associated X;autosome translocations, the der(X) rather than the der(autosome) will become inactivated as it contains the inactivation centre located at Xq13 (Brown et al. 1991). This could produce two unusual circumstances. Firstly, inactivation spread into the autosomal segment of the der(X) would cause effective monosomy for that region, although this would depend upon the extent of inactivation spread. Secondly, the X chromosome region Xpter-Xp21 on the der(autosome) will be active in addition to the entire intact X chromosome. These effects on a cell of the developing embryo are unknown but are likely to be deleterious. A point worth noting is that in all DMD associated translocations the X chromosome segment on the der(autosome) is approximately equivalent (Xpter-Xp21) and will be active in two copies in such cells,

while the autosomal segments on the der(X) chromosomes, for which effective monosomy may exist in such cells, varies in size and composition depending on the autosome breakpoint position of the particular translocation. The consequence is that the position of the autosome breakpoint might influence the bias in the inactivation pattern. DMD associated X;autosome translocations are associated with variable phenotypes and it is thought likely that the reason for this is due to a less pronounced bias in the X-inactivation pattern in the patients with milder disease, so that a minority of cells with active intact X chromosomes are able to produce dystrophin (Boyd et al., 1988). Hence it might be expected that translocations with autosomal breakpoints nearest to the telomere will be associated with the mildest disease phenotypes. The correlation of autosome breakpoint position and disease severity is discussed further in section 1.5.

1.5 Females with DMD and an X;autosome translocation

The identification of females with an X;autosome translocation involving the Xp21 band and expressing DMD, played a crucial role in the mapping of the disease locus and subsequently in cloning the disease gene (section 1.3.2). The first cases were described in the late 1970s (see section 1.3.2) and 24 cases are now known (Boyd, 1991; table 1.1). All translocations share an X chromosome breakpoint position within Xp21, although the sub-band localisations and the autosomal breakpoint positions are heterogeneous. There is evidence that the severity of disease phenotype varies in the patients, although the severity cannot be accurately assessed in the majority of cases due to the age of the patient at reporting. A severe phenotype shown in table 1.1 was indicated in patients wheelchair-bound by the age of 13 or with severe muscle weakness above the age of 8. A mild BMD-like phenotype was indicated in patients still ambulant past the age of 15 or with only mild muscle weakness at the age of 13. Based on this clinical classification, six girls appear to have a severe classical DMD phenotype, while for three the disease is distinctly milder (table 1.1). As discussed in

Translocation	Breakpoints	Phenotype	Near telomere	Reference
X;1 (WLS)	p21.2;p34	severe	-	Lindenbaum et al., 1979
X;1	p21;q23		-	personal communication
X;2 (TM)	p21.2;q37.3	severe	+	Holden et al., 1986
X;2	p21;q14		-	personal communication
X;3 (VSN)	p21.2;q13.3		-	Canki et al., 1979
X;3	p21.2;q25.3		-	personal communication
X;3	p21;q27		+	personal communication
X;4	p21;q35		+	Bodrug et al., 1990
X;4	p21.1;q26		-	personal communication
X;4	p21;q31.2		-	Giacalone and Franke, 1988
X;5 (HEM)	p21.2;q31.1	severe	-	Nevin et al., 1986
X;5 (LUM)	p21.1;q35.3	severe	+	Jacobs et al., 1981
X;6 (EDN)	p21.2;q21	severe	-	Zatz et al., 1981
X;6	p21;q16		-	Perez-Vidal et al., 1983
X;8 (KIY)	p21.2;q24.3	mild	+	Narazaki et al., 1985
X;9 (GM6007)	p21;p22		-	Emanuel et al., 1983
X;9 (LC)	p21.2;q21.3		-	Robinson et al., 1990
X;9	p21;p21		-	Nielsen and Nielsen, 1984
X;11 (LAR)	p21.2;q23.3	mild	-	Nielsen et al., 1983
X;11	p21;q13		-	Greenstein et al., 1977
X;15	p21;q26	severe	+	Ribeiro et al., 1986
X;19 (ORI)	p21.2;q12 or q13.1		-	Boyd et al., 1988
X;21 (FRA)	p21.1;p12	mild	+	Verellen-Dumoulin et al., 1984
X;22 (DEB)	p21;q13		+	Boyd et al., 1988

Table 1.1 List of the 24 known females with DMD and an X;autosome translocation involving Xp21 (see Boyd, 1991). The severity of the disease is indicated where this is known or can be assessed (see text), and translocations with autosomal breakpoints adjacent to the telomere are indicated.

section 1.4, it might be expected that when the autosomal breakpoint is close to the telomere a less pronounced bias in X inactivation is produced causing a milder phenotype. Of the three mild cases, two have autosome breakpoints very close to the telomere (table 1.1). In the case of the t(X;21) the breakpoint is in the short arm of this acrocentric chromosome, and in the case of the t(X;8) the breakpoint is so close to the long arm telomere that chromosome 8 material is barely visible on G-banded der(X) chromosomes from this patient (not shown and Holt, personal communication). The other mild case has an autosome breakpoint at 11q23.3 which is not directly adjacent to the telomere. Of the six cases known to be severe, three have breakpoints adjacent to the telomere. Therefore, there is little evidence to support the prediction that the phenotype of mild cases is caused by the autosomal breakpoints being close to the telomere. An alternative explanation for the differing severity of disease phenotypes in this group of patients predicts that it depends upon the position of the X chromosome breakpoint within the DMD locus. This explanation is that unless the carboxy terminus of dystrophin is critical for function of the protein, then translocation breakpoints positioned in the 3' end of the gene will produce truncated protein products with partial function and cause only mild disease. The mapping of the X chromosome breakpoints of the translocations which is investigated and discussed in Chapter 4 does not support this explanation. There are three further reasons which might explain phenotypic differences in the translocation patients. Firstly, some of the autosome breakpoints, which vary in the patients, may disrupt autosomal gene loci causing dominant expression of phenotypes in addition to DMD or might cause partial expression of the dystrophin protein. Secondly, the background genetic heterogeneity of the patients leads to differing expression of the same biochemical defect. Finally, there is the possibility that environmental factors have played a significant role in the disease expression.

One of the cases with severe phenotype is the translocation selected for detailed molecular investigation in the work described in this thesis (Chapters 5, 6 and 7). This

is the X;1 translocation which was originally described by Lindenbaum et al. (1979). Cytogenetic characterisation had suggested that the translocation is not a simple two-break event, and the interpretation of the banding pattern was that a paracentric inversion on Xp was associated with the translocation (figure 1.3; see Chapter 5, sections 5.1 and 5.3 for fuller presentation and discussion of the cytogenetic characterisation). The autosomal breakpoint of the translocation was determined to be at 1p34 (see figure 1.4). Chromosome 1 has been extensively mapped and a continuous linkage map of markers from this chromosome has been constructed (O'Connell et al., 1989; Dracopoli et al., 1991). Two loci which have been mapped close to 1p34 by in situ hybridisation include the Rhesus blood group gene (1p36.1-p34.3; Chérif-Zahar et al., 1991) and the nonhistone chromosomal protein HMG-17 (1p36.1; Popescu et al., 1990). There are two fragile sites close to 1p34, located at 1p32 and 1p36 (Bruns and Sherman, 1989). Chromosome breakpoints associated with neoplasia which have been identified in this region include those associated with acute lymphoblastic leukemia (ALL) at 1p32; T-cell acute lymphoblastic leukemia (T-ALL) at 1p34-p32; and glioma (GL), malignant lymphoma (ML) and neuroblastoma (NB) all at 1p36-p32 (Trent et al., 1989). The locations of these loci with respect to the t(X;1) breakpoint position are illustrated in figure 1.4. Cloning of the t(X;1) breakpoints enables access to this autosomal point.

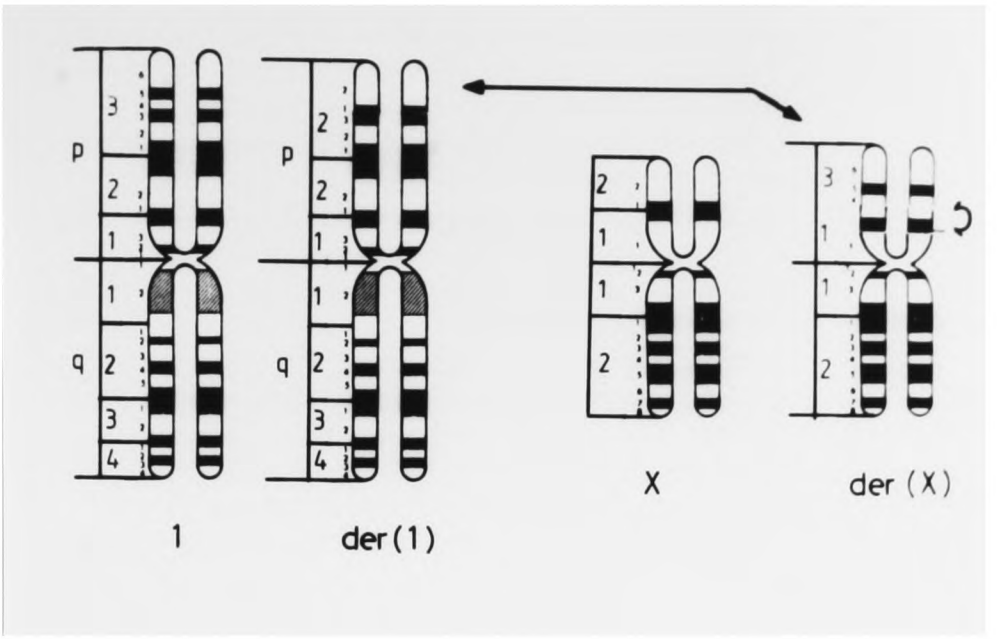


Figure 1.3 (reproduced from Lindenbaum et al., 1979) Diagrammatical representation of the rearrangement of the X;1 translocation (WLS) as originally interpreted. A break on chromosome 1 at 1p34 and a paracentric inversion between Xp2107 and Xp1106 were reported.

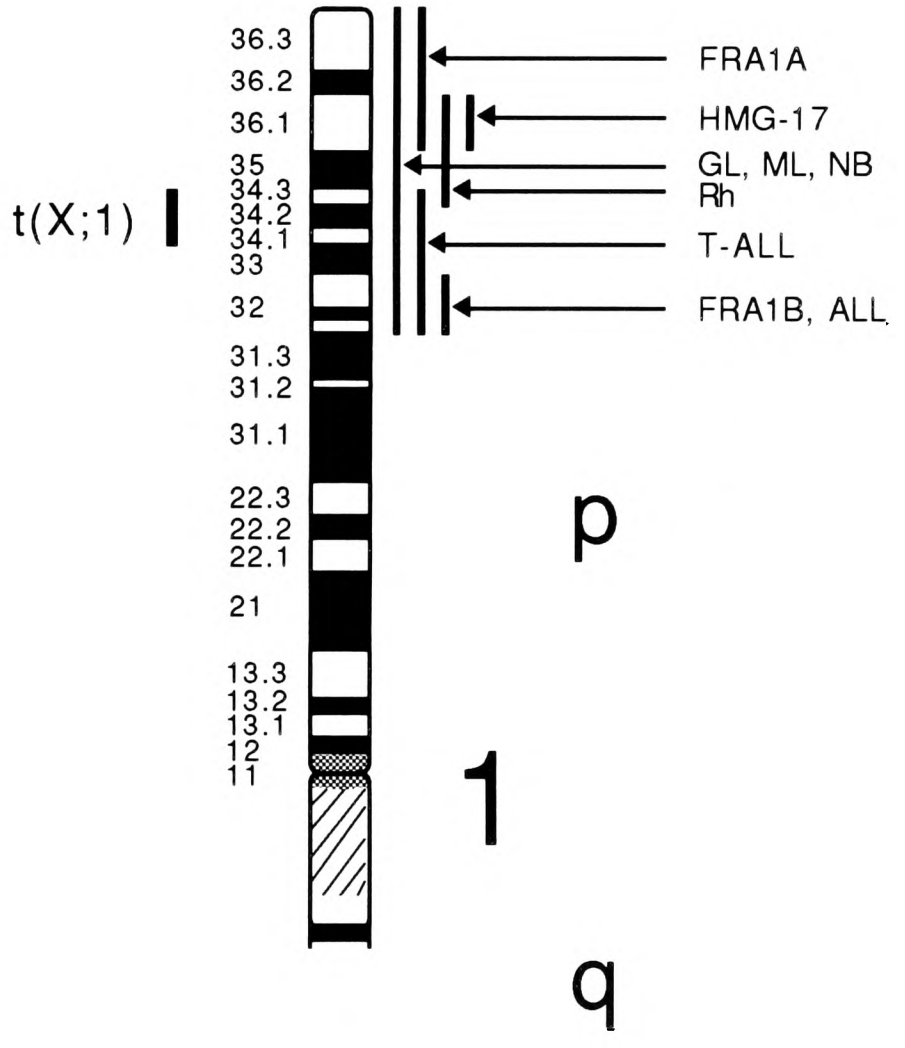


Figure 1.4 Ideogram of the short arm of chromosome 1 illustrating the position of the t(X;1) autosomal breakpoint at 1p34 and the positions of loci and chromosome abberations mapped in the same region. Abbreviations are: fragile sites (FRA1A and FRA1B); Rhesus (Rh); acute lymphoblastic leukemia (ALL); T-cell acute lymphoblastic leukemia (T-ALL); glioma (GL); malignant lymphoma (ML); and neuroblastoma (NB).

Chapter 2 Materials and methods

Many of the methods employed in the work described in this thesis are comprehensively described in 'Molecular Cloning, a laboratory manual' Sambrook et al., (1989). Other sources of methodology are indicated where appropriate.

General solutions:

SET: 150 mM NaCl, 5 mM Na_2EDTA, 50 mM Tris-HCl (pH 7.8).

TE: 1 mM Na_2EDTA, 10 mM Tris-HCl, pH 8.0.

2xSSC: 1.75% (w/v) NaCl, 0.88% (w/v) sodium citrate.

PBS: 0.8% (w/v) NaCl, 0.23% (w/v) Na_2HPO_4, 0.04% (w/v) KH_2PO_4, 0.04% (w/v) KCl.

Other solutions are given at the ends of each section, as appropriate, and % compositions are in v/v unless otherwise indicated.

2.1 Tissue culture

2.1.1 Cells and hybrids

Brief descriptions of the somatic cell hybrids used are given in table 2.1. Many of the hybrids were karyotyped and these details are provided in Chapter 3, table 3.2. TWDAM 1-2 and TWDAM 4-2 were isolated after fusion of blood lymphocytes from a patient with an abnormal banding pattern at Xq26-q28 with the HPRT- Chinese hamster cell line Wg3h. KAG 2-3 was isolated after fusion of a lymphoblastoid cell line from a patient with an Xp21 translocation but without DMD with the HPRT- mouse cell line RAG. All other hybrids were derived by fusion of lymphoblastoid cell lines from female DMD patients with a translocation involving Xp21 (see Chapter 1, table 1.1). Cell lines were either fused with RAG or Wg3h as indicated in table 2.1.

Hybrid	Human parent	Rodent parent	X content	Fully karyotyped	Use	Comments/reference
WAG 8-1	WLS	RAG	Der(X)	+	X mapping, autosome mapping	Monochromosome hybrid (Robinson et al., 1990).
WAG 21R-11	WLS	RAG	Der(1)	-	X mapping	G-banded analysis revealed the presence of normal chromosome 1 (not shown).
EDAG 2-9-8	EDN	RAG	Der(X)	+	X mapping, autosome mapping	Subcloned from EDAG 2-9 (Boyd et al., 1988).
EDAG 2-9-5	EDN	RAG	Der(X)	+	Autosome mapping	Subcloned from EDAG 2-9 (Boyd et al., 1988).
EDAG 2-9-3	EDN	RAG	Der(X), der(6)	+	Autosome mapping	Subcloned from EDAG 2-9 (Boyd et al., 1988).
EDAG 3R	EDN	RAG	Der(6)	-	X mapping	Boyd et al., (1988).
KIYAG 3-1R11	KIY	RAG	Der(8)	-	X mapping	Also contains low level background contamination of normal X; determined by PCR amplification of loci proximal to the X chromosome breakpoint (van Bakel, personal communication).
KIYAG 3-1-3	KIY	RAG	Der(X), der(8)	+	Autosome mapping	Unpublished.
LARAG 3R	LAR	RAG	Der(11)	-	X mapping	Boyd et al., (1988).
DAG 5-2	DEB	RAG	Der(X)	+	X mapping, autosome mapping	Boyd et al., (1988).
HEMAG N2	HEM	RAG	Der(X)	-	X mapping	Boyd et al., (1988).
ORIM 7-1	ORI	Wg3h	Der(X)	+	X mapping, autosome mapping	Boyd et al., (1988).
KAG 2-3	See comments	RAG	Der(X)	+	X mapping, autosome mapping	Derived from a female X;autosome translocation patient with a breakpoint in Xp21 but distal to the DMD locus and not associated with DMD expression (Boyd et al., 1988).
5E-F9	TM	RAG	Der(X)	+	Autosome mapping	Bodrug et al., (1991). Hybrid kindly supplied by S Bodrug and R Worton.
TWDAM 1-2	See comments	Wg3h	See comments	+	Autosome mapping	Derived from a male patient with an abnormal banding pattern at Xq26-q28. This hybrid retains this abnormal chromosome (unpublished).
TWDAM 4-2	See comments	Wg3h	See comments	+	Autosome mapping	Derived from a male patient with an abnormal banding pattern at Xq26-q28. This hybrid retains this abnormal chromosome (unpublished).

Table 2.1 Details of somatic cell hybrids used in studies described in this thesis. Translocation details of hybrids derived from DMD-associated translocations are given in Chapter 1, table 1.1. Karyotype analysis of the hybrids, where performed, is summarised in Chapter 3, table 3.2. The use of hybrids in the work described in this thesis, either for mapping the X chromosome breakpoints of the translocations from which the hybrids were derived, or for determining the autosomal localisations of DNA probes, is indicated.

2.1.2 Cell culture

All hybrids retaining a der(X) chromosome and TWDAM 1-2 and TWDAM 4-2 (table 2.1) were grown in HAT medium (Dulbecco's Modified Eagle's Medium, DME, supplemented with 10% fetal calf serum, 20mM glutamine, 10^{-5}M methotrexate, and 1.6×10^{-5}M thymidine) which selects for the presence of an active HPRT locus at Xq26 (see Choy et al. 1982). All other hybrids and the rodent cell lines RAG and Wg3h, were grown in 6-thioguanine medium (DME, supplemented with 10% fetal calf serum, 20mM glutamine and 10μg/ml 6-thioguanine) which selects against the presence of an active HPRT locus.

All manipulations were carried out in a Class II laminar flow cabinet using autoclaved glassware or sterilised disposable plasticware.

2.1.3 Harvesting

Medium was removed from the flask and T+V solution was added (2 ml to 25 cm^2 flask; 5 ml to 80 cm^2 flask; 10 ml to 200 cm^2 flask). The flask was stood at room temperature for 2 minutes. Cells were detached from the flask bottom with gentle banging and an equal volume of medium was added to eliminate further trypsin activity.

T+V: 0.02% (w/v) EDTA, 0.25% (w/v) trypsin

2.1.4 Frozen storage and resuscitation

Harvested cells were pelleted by centrifugation at 1000 rpm in a bench centrifuge for 5 minutes at room temperature and resuspended in freezing mix (dimethyl sulphoxide:fetal calf serum, 1:9) at a concentration of approxiately 10^7 cells/ml. Aliquots (1 ml) were frozen at -70^{0}C in polypropylene freezing vials for 24 hours and stored in liquid nitrogen tanks indefinitely. For resuscitation, frozen vials were dropped in a beaker of water at 37^{0}C for 5 minutes and the contents transferred using a plugged pasteur to an 80 cm^2 flask containing 20 ml medium. The medium was

replaced after 24 hours.

2.2 Cytogenetics

2.2.1 Standard cytogenetic harvesting method (see Penn and Perle, 1986)

Slide preparation: Slides were soaked in detergent (5% Decon 90) overnight, thoroughly rinsed in tap water, then stored in 70% ethanol. Just before chromosome spreading slides were individually hand polished using a paper hand towel.

Cell preparation: Colcemid (final concentration 0.1µg/ml) was added to a semi-confluent 80 cm^2 culture flask of rapidly growing hybrid cells for 1 hour. Dividing cells not firmly attached to the culture flask surface were detached by knocking the flask by hand. Where necessary the cell layer was treated with 5 ml T+V solution (see section 2.1.3) for 2 minutes, which releases cells at all stages of the cell cycle from the flask surface, and 5 ml culture medium was added. Cells were pelleted by centrifugation at 1000 rpm for 5 minutes in a bench centrifuge at room temperature and resuspended gently in 5 ml hypotonic solution prewarmed at 37^0C. Cell swelling was allowed for 5-10 minutes, following which cells were pelleted as before. The pellet was resuspended in approximately 200 µl hypotonic solution and 5 ml fixative precooled to -10^0C was added. The first few drops were added slowly and the remainder was added rapidly. The first fixation was left at room temperature for 30 minutes, and the fixative was changed twice more. The cell suspension was dropped at an appropriate concentration from a pasteur pipette at a height of between 1 and 60 cm onto prepared glass slides. Gentle breathing onto the slide immediately after dropping aided the dispersal of chromosomes on the slide.

Hypotonic solution: 0.56% (w/v) KCl

Fixative: 3:1 methanol:glacial acetic acid

2.2.2 Thymidine block harvesting method (based on Dutrillaux and Pequignot, 1979)

This modification of the standard protocol above was found to give improved G-

banding results in several instances. The method synchronises cells with respect to the cell cycle enabling the harvesting of increased numbers of metaphase cells and permitting choice of the stage of chromosome condensation at harvesting. The day before harvesting, thymidine was added to the culture flask to 0.3 mg/ml final concentration which was incubated for 18 hours to accumulate mid S-phase cells. The block was released by washing with fresh medium then incubating the cells further for $4^1/_2$-6 hours. Colcemid was added (final concentration 0.1µg/ml) for 10 minutes before harvesting which was as described above (section 2.2.1).

2.2.3 G-banding

Two methods of banding are described. Method 1 was generally used for slides prepared by the standard harvesting protocol (section 2.2.1), and method 2 for slides prepared by the thymidine block protocol (section 2.2.2).

Method 1: Slides aged for 2-5 days were dipped in trypsin solution (2.8% w/v in PBS; Gibco BRL) for 30-60 seconds, rinsed in Sorrensen's buffer, stained in Giemsa solution (6% in Sorrensen's buffer) for 70 seconds and rinsed in tap-water.

Method 2: Slides aged for 2-5 days were placed on a 60^{0}C hotplate for 10 minutes, cooled to room temperature and immersed in 5xPBS for 10 minutes, then rinsed in 1xPBS and dipped in trypsin solution for 40-70 seconds. Slides were stained as in method 1.

Sorrensen's buffer (pH 6.8): 0.95% (w/v) Na_2HPO_4, 0.9% (w/v) KH_2PO_4

2.2.4 In situ hybridisation (modified from Buckle and Craig, 1986)

Chromosome preparation slides were treated with 150 µl RNase solution under a coverslip at 37^{0}C in a moist chamber for 1 hour and then washed four times in 2xSSC. DNA denaturation was by immersion in denaturing solution at 65^{0}C for 4 minutes, after which slides were washed rapidly in 2xSSC, dehydrated by successive immersion in 10, 50, 75, 95 and 100% ethanol and transfered to a desiccator. Labelled DNA (10

ng/slide; specific activity $2x10^7$ dpm/μg; see section 2.10) dissolved in hybridisation solution (35 μl/slide) was denatured by boiling for 2 minutes and plunging on ice for 5 minutes. Hybridisation was performed under a coverslip, whose edges were sealed with rubber solution, at 42^{0}C in a moist chamber overnight. Coverslips were removed under 5xSSC and slides were washed in 2xSSC for 1 hour at room temperature, then in 2xSSC at 60^{0}C for a further hour before being dehydrated as described above and stored in a desiccator. Slides were dipped in photographic emulsion, drained and left to harden for 30 minutes before being stored for exposure at 4^{0}C for 2-7 days in a light-proof box containing desiccant. Development of photographic emulsion film was by immersion of slides (prewarmed to room temperature) in D19 solution at 20^{0}C for 5 minutes. A rinse in 1% glacial acetic acid preceded a 5 minute fixation in Hypam solution. Slides were finally rinsed for 1 hour in gently running tap water, stained in 6% Giemsa in Sorrensen's buffer for 5-10 minutes, rinsed briefly in tap water and air-dried.

RNase solution: 100 μg/ml in 2xSSC, boiled

Denaturing solution: 70% formamide, 0.1 mM Na_2EDTA, in 2xSSC pH 7.0

Hybridisation solution: 50% formamide, 5x Denhardts solution, 5xSSPE, 10% (w/v) dextran sulphate, 200 μg/ml salmon sperm DNA

SSPE: 900 mM NaCl, 50 mM NaH_2PO_4, 5 mM Na_2EDTA, pH 7.2

50xDenhardt's solution: 1% (w/v) BSA, 1% (w/v) polyvinylpyrrolidone, 1% (w/v) ficoll

Photographic emulsion: 1:1 Ilford nuclear emulsion L4:water at 55^{0}C

D19: 0.2% (w/v) metol, 9% (w/v) sodium sulphite, 0.8% (w/v) hydroquinone, 4.5% (w/v) sodium carbonate, 0.5% (w/v) KBr

Hypam solution: 20% Hypam (Ilford), 2.5% hardener (Ilford)

2.3 DNA preparation from mammalian cells

Approximately 10^8 cells were washed twice in SET and resuspended to 10 ml in a

polypropylene tube. Proteinase K was added to 100 μg/ml, SDS to 0.5% (w/v) and the solution incubated overnight at 37^{0}C. Two phenol extractions were performed using equal volumes of solution A and mixing the two phases gently by hand for 5-10 minutes until they had formed an emulsion. The phenol phase was removed following centrifugation of the emulsion at 3000 rpm in a bench centrifuge for 10 minutes at room temperature. Phenol extractions were followed by a phenol/chloroform extraction (solution C) and a chloroform extraction (solution B). RNase (boiled to destroy any DNase contamination) was added to the aqueous phase to 75 μg/ml and was incubated for 3 hours at 37^{0}C with gentle mixing. Following further phenol, phenol/chloroform, and chloroform extractions DNA was precipitated with 2 volumes of ethanol at room temperature, washed in 70% ethanol, allowed to air-dry for 10-30 minutes and redissolved in TE. The DNA concentration was determined by measuring the optical density of a diluted sample at 260 nm (OD_{260} = 1 for 50 μg/ml). Additional readings were taken at 235 nm and 280 nm. OD values at these wavelengths of approximately half the value at 280 nm indicated acceptable purity of the DNA. Typical yields of DNA approached the theoretical maximum yield of 600 μg from 10^8 diploid mammalian cells. Yields from hybrid cells with more than two sets of chromosomes often exceeded this value.

Solution A (phenol): 1000:1 phenol:8-hydroxyquinoline, extracted with an equal volume of 1M Tris-HCl (pH 8.0) and equilibrated in SET

Solution B (chloroform): 24:1 chloroform:isoamylalcohol

Solution C (phenol/chloroform): 1:1 solution A:solution B

2.4 Endonuclease restriction of DNA

DNA was digested according to the enzyme manufacturers' recommendations using the buffers supplied and up to a 10 fold excess of enzyme. Genomic digests were usually incubated overnight and other DNAs (plasmid, fragment or PCR product) were usually incubated for 1 hour. An aliquot of DNA was checked for complete digestion by

electrophoresis.

2.5 Gel electrophoresis

The standard conditions described below were for a resolution of genomic DNA fragments in the size range 1 kb to 15 kb. Conditions such as agarose concentration and voltage of electrophoresis were modified to suit specific size separation requirements following the guidelines of Sambrook et al., (1989). 250 ml of molten 0.8% (w/v) agarose (type II, A-9918, Sigma) in 1xTBE was poured to set into the 20 x 20 cm tray of a Pharmacia gel box (GNA 200) with a 22 tooth comb (well volume 20 µl). DNA samples in 5% glycerol, 0.025% (w/v) bromophenol blue were loaded and electrophoresed using TBE running buffer at approximately 50V overnight or until the bromophenol blue tracking dye had travelled 15 cm. The gel was stained in 0.5 µg/ml ethidium bromide for 15 minutes, destained in TBE for 15 minutes before being examined and photographed over a UV 254 nm transilluminator using a red filter and polaroid 665 film.

Electrophoresis of DNA fragment digests and PCR products was generally performed in a 10 x 8 cm gel (Pharmacia, GNA 100) using an agarose concentration of between 0.8 and 2% (w/v).

λ DNA fragments generated by digestion with HindIII or BstEII were used as size markers. Sizes in bp were:

HindIII: 23130, 9416, 6557, 4361, 2322, 2027, 564, 125

BstEII: 8454, 7242, 6369, 5686, 4822, 4324, 3675, 2323, 1929, 1371, 1264, 702, 224, 117

TBE electrophoresis buffer: 1.1% (w/v) Tris, 0.55% (w/v) boric acid, 0.002 M Na_2EDTA

2.6 Transfer of DNA to membranes by Southern blotting

DNA was transferred to membranes according to the manufacturer's instructions (Amersham International) adapted from Southern (1975). Agarose gels were gently

washed in 0.25M HCl for 30 minutes to fragment DNA molecules by depurination and to facilitate the transfer of restriction fragments larger than about 10 kb. DNA was denatured by equilibrating the gels for 30 minutes in alkali transfer buffer. Transfer of DNA to positively charged nylon membranes (Hybond N+; Amersham) was achieved by placing the gel on a pad of Whatmann 3mm filter paper over a reservoir of alkali transfer buffer, lying a sheet of transfer membrane over the gel, and over this, two sheets of moist 3mm filter paper, a 5cm high stack of absorbent paper towels, a flat glass plate and a 500g weight. Four hours were allowed for transfer, after which the transfer membrane was removed, gently rinsed in 2xSSC to remove adhering agarose and air dried.

Alkali transfer buffer: 0.4 M NaOH, 1.6 M NaCl

2.7 Membrane hybridisation

Hybridisation of probes to membranes was performed in bottles rotating in a 64^{0}C hybridisation oven (Hybaid). Membranes were prehybridised in Church hybridisation buffer at 64°C for approximately 30 minutes. Radiolabelled probe (see section 2.10) was added and allowed to hybridise to DNA on the membrane for 16 hours. Membranes were washed at 64^{0}C in sodium phosphate buffer solutions of decreasing molarity and containing 1% (w/v) SDS. The first wash was in 400 mM buffer (equivalent in molarity to 2xSSC) and final washes were generally in 200 or 100 mM buffer for 20 minutes. Filters were then briefly air-dried, wrapped in polythene cling film and autoradiographed using Fuji RX X-ray or Kodak AR film in X-ray cassettes with Kodak X-omatic intensifying screens at -70^{0}C. A satisfactory signal was usually obtained following an exposure of between 1 and 14 days.

Sodium Phosphate buffer: 1 M $Na_2H\ PO_4$, pH 7.2 adjusted by addition of orthophosphoric acid

Church hybridisation buffer (Church and Gilbert, 1984): 0.5 M sodium phosphate buffer, 1 mM Na_2EDTA, 1% (w/v) BSA (Sigma A-4503), 7% (w/v) SDS (Sigma L-4380),

stored frozen and filtered at 64^{0}C through a 0.45 μm membrane just before use

2.8 Preparation of plasmid or cosmid DNA by alkaline lysis

Bacterial cultures were grown overnight in 10ml L-broth containing the appropriate antibiotic. Cultures were centrifuged for 10 minutes at 3000 rpm in a bench centrifuge and the pellets resuspended in 150μl GTE in a 2000μl microfuge tube prior to incubation at room temperature for 5 minutes. Chilled NaOH/SDS (300μl) was added, the tubes shaken briefly, and then transferred to ice for 5 minutes. KAc (225μl) was added, the tubes again mixed by hand and returned to ice for a further 20 minutes. The tubes were microfuged at 4^{0}C for 15 minutes, the supernatants transferred to fresh tubes and extracted with phenol, phenol/chloroform, then chloroform as previously described (section 2.3). The aqueous DNA solution was precipitated with two volumes of ethanol at room temperature for 15 minutes, washed in 70% ethanol, lyophilised and dissolved in approximately 100μl TE. RNA was removed by addition of 50μg boiled RNase and incubation at 37^{0}C for 60 minutes. Typical yields of plasmid DNA were up to approximately 50 μg/10 ml culture.

GTE: 50 mM glucose, 10 mM Na_2EDTA, 25 mM Tris-HCl (pH 8.0)

NaOH/SDS: 0.2 M NaOH, 1% (w/v) SDS

KAc: 3M potassium acetate, 11.5% acetic acid

2.9 Gel-purification of DNA fragments

The following protocol was used to isolate inserts from plasmids or cosmids and to purify amplified products of PCR. Where necessary, DNA samples had been digested previously with appropriate restriction endonucleases. The sample volume was adjusted not to exceed 100ng/μl, glycerol was added to 5% final concentration and bromophenol blue to 0.025% (w/v) final concentration. Electrophoretic separation was generally performed in a Pharmacia GNA100 apparatus (10 x 8 cm gel) using low gelling temperature agarose (type VII, A-4018, Sigma) at a concentration between 0.5

and 1% (w/v) in TAE electrophoresis buffer for a time and at a voltage suitable to the separation required. Comb teeth were taped together, where required, to form wells of sufficient capacity for larger DNA samples. Gels were stained with ethidium bromide as previously described (section 2.5) and gel slices containing DNA fragments were cut using a scalpel blade with the aid of a UV (254 nm) transilluminator.

DNA was extracted from the agarose slice using a 'Geneclean' kit (Bio 101 Inc) following the manufacturer's instructions. The agarose was dissolved at room temperature in 2.5 volumes of 6M NaI. Silica bead suspension (5µl; sufficient for binding of up to 5µg DNA) was added and the tube kept on ice for 15 minutes with periodic gentle mixing. The tube was then centrifuged for 20 seconds and the DNA-silica bead pellet washed three times with "new" wash as supplied. Finally the pellet was resuspended in 10-50µl TE, incubated at 50^{0}C for 2 minutes, centrifuged and the supernatant collected. A further wash of the pellet with TE produced a DNA solution whose recovery was consistently between 50 and 100% for DNA fragments of between 100 bp and 5 kb in length.

TAE eletrophoresis buffer: 0.97% (w/v) Tris, 0.23% glacial acetic acid, 0.002 M Na_2EDTA

2.10 DNA labelling

All DNA radio-labelling was performed by the random hexanucleotide priming method (Feinberg and Vogelstein, 1983) using a kit supplied by Amersham International. DNA (10-50 ng) was denatured by boiling in a 1500µl microfuge tube for 3 minutes with a volume of water to make a final reaction volume of 50 µl. The solution was placed on ice for 5 minutes. 10 µl buffer solution and 5 µl primer solution were added followed by an appropriate quantity of α-^{32}P or α-^{3}H labelled d-CTP and finally 2 µl (2 units) Klenow fragment. The labelling reaction was performed at room temperature for 3-16 hours. Unincorporated nucleotides were removed by spin dialysis. The reaction mixture was added to a column of G50 coarse sephadex

(Pharmacia) washed with TE and centrifuged at 3000 rpm in a bench centrifuge for 3 minutes. The labelled DNA was collected from under the column and the label incorporation was estimated using a hand held monitor (for ^{32}P), or more accurately determined by use of a scintillation counter (^{3}H or ^{32}P). Typical specific activities of probes were in the order of 10^8 dpm/μg (^{32}P) or $2x10^7$ dpm/μg (^{3}H).

2.11 DNA ligation

The plasmid pUC18 was used in all double-stranded vector cloning experiments (Messing, 1983; Norrander et al., 1983). The multiple cloning site of pUC18 is positioned within the gene for the α-peptide of β-galactosidase and transformed colonies containing recombinant plasmids where the inserted DNA disrupts this gene can be identified by a colour test if X-Gal is included in agar plates. Recombinant colonies are white and non-recombinant colonies stain blue. The multiple cloning site of pUC18 includes the recognition sites of several restriction enzymes. A further advantage of using a pUC vector for cloning is that a good yield is obtained when DNA preparation is performed on bacterial cultures (see section 2.8) because of the high copy number of the plasmid in bacterial cells.

The method described here was used for sub-cloning of gel-purified DNA fragments, for shotgun sub-cloning DNA digests and for cloning gel purified PCR products. This DNA, where necessary, and vector DNA were restricted using appropriate endonucleases. Enzyme activity was destroyed by heat treatment where appropriate (e.g. 60^{0}C for 10 minutes) and/or phenol extraction followed by ethanol precipitation. Standard conditions for ligation were to incubate 50ng of vector DNA and a threefold molar excess of insert DNA with 1 unit of T4 ligase (Gibco BRL) at 16-18^{0}C for 16 hours using the manufacturer's reaction buffer. Generally 5μl of the ligated mixture was used to transform competent bacteria (see section 2.11). An obligatory control reaction was routinely performed where the insert DNA was replaced with an equal volume of TE, and the products of this reaction were also used in the

transformation process.

2.11.1 Infilling of 5' overhangs

Infilling of the cohesive ends of DNA molecules generated by restriction enzymes was performed in some experiments, in order to make the ends blunt-ended and therefore compatible in ligation reactions with other blunt-ended molecules. This can be performed for 5' overhangs using Klenow, since the 3' DNA end will prime DNA replication of the overhang.

Template DNA (approximately 100 ng) was incubated for 10 minutes at room temperature in a 25 µl volume, including 1.6 µl of each dNTP (from a nick-translation kit; N.5500, Amersham International) and 1 unit Klenow. The reaction was stopped by addition of 5 µl 0.5M EDTA, and the DNA was purified using a 'Geneclean' kit (see section 2.9).

2.12 Transformation of E.Coli (Method adapted from Hanahan, 1985)

In order to prepare competent cells, bacteria (JM83) from a frozen stock were streaked on a SOB agar plate and incubated overnight at 37^{0}C. Two colonies (2.3mm diameter) were picked and used to inoculate 20ml SOB medium in a 250ml conical flask which was incubated with moderate agitation at 37^{0}C. Once a density of 5x10 cells/ml (OD_{550} approximately 0.5) was reached, the culture was chilled on ice for 5 minutes and centrifuged at 1000g for 15 minutes at 4^{0}C in a polypropylene tube. The well drained cell pellet was resuspended in 5ml TFB with moderate vortexing and kept on ice for 15 minutes. The cells were pelleted as before and this time resuspended in 1600µl TFB. DnD solution (56µl) was added and mixed by swirling, and after 10 minutes incubation on ice, a second equal volume was added giving a 7% final concentration. The tube was further incubated on ice for 15 minutes after which the bacterial cells were ready for transformation. Aliquots (200µl) of competent cells were mixed with DNA solution in polypropylene tubes (Falcon 2059) and incubated on ice for

30 minutes prior to a 90 second heat shock in a 42^{0}C water bath. After 2 minutes cooling on ice, 800µl SOC medium was added to all tubes which were then incubated at 37^{0}C with moderate agitation for 60 minutes. Cells were then spread on L-broth agar plates containing appropriate antibiotics and other additives to select for and identify transformants.

JM83 genotype: F- φ80d*lacZ*ΔM15 Δ(*lac*-*pro*AB) ara *rps*L

Transformation buffer (TFB): 100 mM KCl, 45 mM MnCl, 10 mM $CaCl_2$, 2 mM $HACoCl_3$, 10 mM K-MES

DnD solution: 1 mM DTT, 90% DMSO, 10 mM Potassium acetate

L-broth: 1% (w/v) NaCl, 1% (w/v) Bacto-tryptone, 0.5% (w/v) Bacto-yeast extract

SOB medium: 2% (w/v) Bacto-tryptone, 0.5% (w/v) Bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$

SOC medium: SOB medium containing 20 mM glucose

Agar concentration in plates: 1.5% (w/v)

Antibiotics were added at the following final concentrations where appropriate: ampicillin (100 µg/ml), kanamycin (50 µg/ml)

X-gal concentration where used: 40 µg/ml, final concentration

2.13 Frozen storage of bacterial stocks

Autoclaved glycerol was added to exponential liquid bacterial cultures to give a final concentration of 15% and 1 ml aliquots were frozen and stored at -70^{0}C.

2.14 Colony blots

Nylon DNA transfer membrane (Hybond N, Amersham) was carefully laid over the surface of an agar plate containing colonies at suitable dilution, and a sterile needle was used to mark its orientation. After 2 minutes, the membrane was removed with forceps and placed colony side up on an absorbent paper pad (Whatmann 3mm)

soaked in denaturing solution. After 5 minutes the membrane was transferred for a further 5 minutes to another pad soaked with neutralising solution. The membrane was then washed gently in 2xSSC for 60 seconds and air-dried. DNA was covalently cross-linked to the membrane by placing it colony side down on a UV (254 nm) transilluminator for 60 seconds.

2.15 DNA sequencing

The DNA sequencing described in this thesis was performed on DNA fragments cloned into either M13mp18, M13mp19 or pUC18 vectors. The advantage of using a double-stranded pUC vector is that both strands of a clone's insert DNA can be sequenced by performing two sequencing reactions using sequencing primers which flank the insert. M13 is a single-stranded vector and therefore, in order to sequence both strands of a DNA fragment they must be cloned separately. To facilitate this, M13mp18 and M13mp19 can be used, as they contain the same multiple cloning site in opposite orientations. The sequencing results obtained by using M13 are superior to those from double stranded clones (Sambrook et al., 1989). It is possible to read sequence at a greater distance from the primer, and fewer band compressions and other sequence reading difficulties are produced.

Stocks of M13 bacteriophages are usually grown in liquid culture. The infected bacteria do not lyse, but grow slowly, releasing bacteriophage particles. The bacteria contain a double-stranded replicative form (RF) of M13 DNA, and may be separated from the bacteriophage particles by centrifugation, allowing single-stranded and double-stranded DNA preparations to be performed on the same culture.

Foreign DNA is cloned into M13 by ligation with double-stranded RF DNA and competent bacteria are transfected. The transfection process requires a bacterial strain carrying an F'episome. Bacterial strains such as JM101, used for transfection, have a chromosomal mutation causing defficiency in proline biosynthesis, so that only those bacteria with the F' plasmid (also including the genes required to complement

the proline defficiency) are able to grow on minimal medium. The spontaneous loss of F' plasmids from bacteria grown for limited periods in rich media is not a serious problem if the bacteria are previously grown on minimal agar plates.

2.15.1 Transfection

Ligation of DNA with RF M13 DNA was performed as previously described for pUC18 (section 2.11).

The transfection procedure was essentially the same as that described for transformation (section 2.12) except that the bacterial strain used was JM101 (prepared by streaking a frozen stock on a minimal medium agar plate and incubating at 37^0C for 40 hours). Colonies were picked, grown in SOB medium, made competent, incubated with DNA and heat shocked as previously described. Next, 175 µl of SOB medium lacking magnesium was added, and the tube gently vortexed. Aliquots of 1 µl, 10 µl and 100 µl were then added to prepared tubes containing top agar mix, followed by 200 µl of overnight liquid culture of JM101 (grown in SOB). The tubes were vortexed and the contents poured over L-broth agar plates. Plates were briefly left to harden and incubated at 37^0C for 10 hours.

JM101 genotype: Δ(lac-proAB) supE thi/F'-lacI^{q}ZΔM15 traD36 proAB$^+$

Minimal medium: 0.13% (w/v) $Na_2HPO_4.7H_2O$, 0.03% (w/v) KH_2PO_4, 0.05% (w/v) NaCl, 0.4% (w/v) glucose

Top agar mix: 3 ml SOB top agar (0.7% w/v;molten), 800 µg X-gal, 800 µg IPTG

2.15.2 Isolation of M13 plaques

Well separated (>1.5 cm) plaques were picked with the tip of a sterile pasteur pipette and the M13 particles allowed to diffuse out of the small plug of agar into 1 ml L-broth for 1 hour at room temperature.

2.15.3 Preparation of infected culture

A single colony of JM101 grown from a minimal agar plate (section 2.15.1) was used to inoculate 50 ml of LB medium in a 250 ml flask. The culture was incubated with shaking for 7 hours at 37^{0}C. 50 µl of this plating bacteria culture was added to 2 ml of LB medium in a 15 ml culture tube. One tenth of the M13 particle suspension derived from a single plaque (section 2.15.2) was added and the infected culture was incubated for 5-6 hours at 37^{0}C with constant shaking. The 2 ml culture was centrifuged in a 2000 µl microfuge tube at 12,000 g for 5 minutes at 4^{0}C and the supernatant was divided. Half was stored at -20^{0}C as a master stock of bacteriophage particles. The remainder was reserved for single-stranded DNA preparation. The bacterial pellet was used to prepare a small quantity of double-stranded RF DNA by the alkaline lysis method (see section 2.8) enabling recombinant M13 clones containing the relevant sized inserts to be identified by restriction enzyme digestion.

2.15.4 Single-stranded DNA preparation

PEG solution (200 µl) was added to the bacteriophage substrate (1 ml) and the tube was mixed, gently vortexed and stood for 15 minutes at room temperature. Precipitated M13 particles were pelleted by centrifugation at 12,000 g for 5 minutes at 4^{0}C, the supernatant was removed and the tube briefly re-centrifuged to allow the removal of residual supernatant. The pellet was resuspended in 100 µl TE by vortexing vigorously. Phenol, phenol/chloroform and chloroform extractions were performed as previously described (section 2.3). The aqueous phase was then transferred to a tube containing 10 µl 3M sodium acetate and 300 µl ethanol and the tube was mixed and incubated at room temperature for 15 minutes. The single-stranded DNA precipitate was centrifuged at 12,000 g for 10 minutes at 4^{0}C, washed in 70% ethanol which was then removed very carefully and the pellet allowed to dry by standing the open tube at room temperature for 30 minutes. The DNA was dissolved in 50 µl TE and stored at -20^{0}C. DNA concentration was estimated by running 1 µl on a 0.8% (w/v) agarose gel

alongside standard M13 single-stranded DNA of known concentration. A typical yield was approximately 5 μg DNA from 1 ml infected culture.

PEG solution: 20% (w/v) polyethelene glycol (8000), 2.5 M NaCl, sterilised by filtration

2.15.5 Hybridisation of complementary DNA strands in M13 clones

Confirmation of cloned product in M13 was obtained by testing for hybridisation between single-stranded M13 clones containing opposite DNA strands of the same insert, to form a figure of 8 structure. This hybrid molecule migrates more slowly in gel electrophoresis than either clone alone. Hybridisation of 100ng of each clone was performed in 10 μl with a final concentration of 5xSSC at 65^{0}C for 1 hour. The mixed DNAs were examined after electrophoresis in 0.8% (w/v) agarose at 80V for 90 minutes.

2.15.6 Annealing of M13 single stranded DNA to sequencing primers

Single stranded M13 DNA (7μl; approximately 1 μg) was mixed with 2 μl 5x Sequenase reaction buffer (USB) and 1 μl (0.5 pmol/μl) forward sequencing primer. The mixture was heated to 68^{0}C for 2 minutes and allowed to cool to 30^{0}C over 30 minutes by placing in a plastic beaker containing water at 68^{0}C on the bench. Annealed template/primer was stored on ice and sequencing reactions were carried out within 2 hours.

forward sequencing primer: 5' GTTTTCCCAGTCACGAC 3'

2.15.7 Preparation of plasmid DNA for double stranded sequencing

Plasmid DNA prepared by the alkaline lysis method (section 2.8) was further purified by polyethylene glycol precipitation. DNA (10μg in 200μl) was mixed with 120 μl PEG solution (section 2.15.4) and placed on ice for 1 hour. The tube was centrifuged at 12,000 g for 10 minutes at 4^{0}C, the supernatant carefully removed and the DNA pellet

washed with 200 µl 70% ethanol before being briefly lyophilised and redissolved in 40 µl TE.

2.15.8 Annealing of plasmid DNA to sequencing primers

Denaturing solution (2 µl) was mixed with 2-3 µg plasmid DNA (10 µl; prepared as above) and the tube was incubated at 37^{0}C for 30 minutes. The DNA was precipitated by addition of 1.2 µl 3M sodium acetate (pH 5) and 26 µl ethanol, left at -70^{0}C for 15 minutes and then pelleted by centrifugation at 12,000 g for 15 minutes at 4^{0}C. The pellet was washed in 70% ethanol, briefly lyophilised, dissolved in 10 µl forward or reverse primer annealing solution and incubated for 30 minutes at 37^{0}C. Annealed primer/template was stored on ice and sequencing reactions were performed within 2 hours.

Denaturing solution: 0.2 M NaOH, 2 mM Na_2EDTA

Annealing solution: 2 µl Sequenase reaction buffer (USB), 1 µl (0.5 pmol) sequencing primer, 7 µl H_2O

Reverse sequencing primer: 5' AACAGCTATGACCATG 3'

2.15.9 DNA sequencing reactions

Sequencing reactions were performed by the di-deoxy chain termination method of Sanger et al. (1977) using the components of a Sequenase Version 2.0 kit (USB Corporation) and its accompanying protocol. The reaction was carried out in 2 stages. Firstly, a labelling reaction was performed by mixing the following with the 10 µl template/ primer: 1µl DTT (0.1 M); 2 µl labelling mixture; 0.5 µl α-^{35}S d-ATP (5 µCi); and 2 µl Sequenase version 2.0. The reaction was incubated at room temperature for 2-5 minutes then termination reactions were performed by transferring 3.5 µl labelling reaction product into each of 4 tubes, containing 2.5 µl termination mix for one of the four DNA nucleotides prewarmed to 37^{0}C. Following incubation at 37^{0}C for 3- 5 minutes, 4 µl stop solution was added. Tubes were stored at -20^{0}C until the sequencing

gel was ready to be loaded.

2.15.10 Polyacrylamide gel electrophoresis of sequencing reaction products

An LKB 2010 Macrophor apparatus was used for running sequencing gels with 40 x 10 cm glass plates separated by 0.2 mm spacers and with a 0.6 mm wedge former at the bottom of the gel. The front glass plate was treated with 5 ml Repelcote 2810 solution (BDH) and polished using 70% ethanol and paper towels to prevent gel binding, while the back plate was treated with 5 ml Bind-silane solution then polished using 70% ethanol and paper towels to promote gel adhesion. Acrylamide solution was carefully poured and left to polymerise for two hours with the gel plates resting on a horizontal platform. A 36 toothed well-former was used permitting the reading of up to 9 DNA sequences per gel. The gel was positioned vertically in the gel-running assembly, TBE electrophoresis running buffer was added, the well-former removed, and the wells were quickly and thoroughly rinsed using a pasteur pipette with a large teat to remove unpolymerised acrylamide and air bubbles. An aluminium plate was clamped onto the back glass plate of the gel to minimise temperature differences within the gel, and the gel was then pre-run at 2000 V for 20 minutes. Wells were rinsed again with electrophoresis buffer, to wash out diffusing urea, and sequencing reaction samples were individually heated to 80^0C for 2 minutes and 1 µl carefully loaded using a macrophor sample syringe (LKB 2010-150) with a glass fibre needle. The loaded gel was run at 500 V for 5 minutes then the voltage increased to 2000. Running times of between 1 .5 and 3 hours (4% w/v acrylamide) or 2.5 and 5 hours (6% w/v acrylamide) were used. 4% gels were found to give better resolution when reading DNA sequence far from the primer. The apparatus was dismantled and the two gel plates were gently prised apart leaving the gel attached to the "bind-silane" treated plate. This gel plate was rinsed in 10% acetic acid for 20 minutes twice allowing all urea to diffuse out of the gel, then rinsed in distilled water for 20 minutes. The gel was dried by placing in a 60^0C oven for 90 minutes and autoradiographed using Fuji RX X-ray film sandwiched

between the gel plate and another glass plate, clamped together in a light proof box at room temperature for 12-96 hours.

Typical sequencing gels enabled accurate reading of DNA sequence using single- or double-stranded templates up to approximately 300 bp from the primer. Longer runs of reactions using single-stranded templates on optimal quality gels enabled accurate reading of sequence 500 bp from the primer.

Bind-silane solution: 20 ml ethanol, 5 ml 10% acetic acid, 75 µl Bind-silane (LKB 1850-251)

Acrylamide solution: 50 ml of 4-6% (w/v) acrylamide (19:1, acrylamide:bisacrylamide), 10.5 % (w/v) urea, in 1xTBE; was filtered through a 0.45 µm membrane and degassed briefly by swirling in a vacuum. Just before pouring, 400 µl 10% (w/v) ammonium persulphate and 40 µl TEMED were added.

2.15.11 Sequence analysis

Individual nucleic acid sequences were analysed using a computer program 'nip' from the Staden package (Staden, 1984). This program was used to search for stretches of alternating purine/pyrimidine, secondary structure, restriction enzyme sites, splice sites and short specific sequences. DNA sequences were compared for similarity to sequences in the EMBL databases using a program, 'BLAST' (Karlin and Altschul, 1990; Altschul et al., 1990). Access to this program was through the computing service provided by the UK Human Genome Mapping Project Resource Centre.

2.16 Polymerase Chain Reaction (PCR)

Assistance in the experimental design of PCRs was obtained from Innis and Gelfand (1990).

2.16.1 Primer design

DNA sequence was examined and candidate 20mer primers were identified with

approximately 50% GC content and which were devoid of runs of consecutive purines, consecutive pyrimidines, alternating purine/pyrimidines or other unusual base sequences. The sequences of candidate primer pairs were compared to check that significant primer-primer hybridisation could not occur (homology longer than four bases in length was avoided).

2.16.2 Reaction conditions

The experimental conditions which varied between experiments, such as the times and temperatures of steps employed in the thermal cycle, are described separately for each experiment in the text. The following standard reaction conditions were constant in all PCR experiments: template DNA, 100ng; primers, 20 pmol each; Tris-HCl (pH 8.3), 20 mM; $MgCl_2$, 1.5 mM; KCl, 25 mM; gelatin, 100 µg/ml; dNTPs, 50 µM each; Taq DNA polymerase, 1 unit. A reaction volume of 100 µl was used in a 500 µl microfuge tube and the mixed components were overlaid with 100 µl mineral oil. Just prior to the addition of enzyme and the commencement of the first thermal cycle, all other components of the reaction were incubated at 94^0C for 5 minutes. Reactions were performed using a Techne PHC-1 programmable thermal cycler. Amplification products were analysed on 1-2% (w/v) agarose gels without further purification.

A detailed description is given in Chapter 6 of an adaptation of PCR, termed 'inverse PCR', whose use played a significant part in the work described in this thesis.

Chapter 3 Construction and use of a somatic cell hybrid panel to map autosomal DNA probes

3.1 Introduction

The construction and use of a somatic cell hybrid panel is described for the determination of the chromosomal localisations of autosomal DNA probes. This hybrid panel is a by-product of work to characterise as fully as possible the hybrid panel constructed to map DMD-associated translocation breakpoints (see Chapter 4). Characterisation of hybrids was by cytogenetic analysis using standard G-banding techniques and in the case of one hybrid, in situ hybridisation using total human DNA as a complex probe was additionally used. Eleven hybrids were karyotyped, and this information enabled the unambiguous localisation of a DNA fragment from any autosome.

Any modern laboratory engaged in molecular genetics research sporadically generates unmapped DNA fragments of interest for which chromosomal localisations would be valuable. There is a variety of methods available for obtaining chromosomal localisations. The principal three are somatic cell hybrid analysis, in situ hybridisation analysis, and linkage analysis. The advantages of in situ hybridisation analysis are that immediate sub-localisation to one or a few chromosomal bands is often possible and probe hybridisation to related sequences at other sites in the genome may be detected (Buckle and Craig, 1986). The main disadvantage is that each localisation is labour intensive. The determination of genetic map position by linkage analysis is only possible where a probe detects a polymorphism. It is not often used for de novo chromosomal assignments in man where the small family size presents an extra difficulty although this is made somewhat easier by the exploitation of large families (for example CEPH families, Dausset et al., 1990). A large number of experiments is generally required before significant co-segregation of alleles is observed. This method of analysis is usually reserved for mapping important disease loci, as it is the only way of mapping

phenotypes unless they are associated with chromosome rearrangements (Frezal and Schinzel, 1990). Alternatively it may be used for the regional localisation of a probe whose chromosomal location is already known. In many circumstances somatic cell hybrid analysis is the method of choice for the de novo chromosomal assignment of a DNA fragment since one or two simple experiments should be all that is required, using a well characterised hybrid panel. Prior to use of such a panel, it is usually worth checking the sizes of hybridising DNA fragments produced by different restriction enzymes, the strengths of hybridisation signals and whether the probe cross-hybridises to rodent DNA. A restriction enzyme digest can then be chosen with a strong human signal which does not co-migrate with any rodent signal.

The somatic cell hybrid panel described here was characterised principally by using standard G-banding protocols (Penn and Perle, 1986; Dutrillaux and Viegas-Péquignot, 1979). This permits human chromosomes to be recognised amongst mouse or hamster chromosomes based on their banding patterns. An aid to differentiating human and mouse chromosomes by this banding method is that mouse chromosomes are all telocentric and the high AT content of the mouse centromeres causes them to stain intensely (Penn and Perle, 1986). There are alternative protocols which permit human and rodent chromosomes to be distinguished. The G-11 staining technique produces a differential staining intensity of human and mouse chromosomes or human and hamster chromosomes (Bobrow and Cross; 1974). Although individual human chromosomes cannot be identified using this method alone, G-11 banding enables the number of human chromosomes in a cell to be counted quickly and can be combined with trypsin or quinacrine G-banding to allow the chromosomes to be identified subsequently. Small human marker chromosomes or interspecific rearrangements can also be detected using this method. Differentiation of human and rodent chromosomes can alternatively be made by in situ hybridisation of total human or rodent DNA to metaphase chromosomes and can be combined with BudR banding methods to identify specific chromosomes. Another method has been described for

simultaneously identifying and banding human chromosomes in somatic cell hybrids (Tucker et al., 1988). The procedure involves in situ hybridisation of biotinylated human DNA which is then bound with fluoresceinated avidin, and simultaneous staining with DAPI and actinomycin D which produces the banding. The fluorescein and DAPI dyes excite at different wavelengths, therefore the human material can be identified at one wavelength and the banding pattern visualised at another. New methods of somatic cell hybrid characterisation are now available based on PCR technology which do not require expertise in cytogenetic techniques and analysis (see discussion in section 3.3).

The hybrid mapping panel described here was used to map DNA fragments by Southern blot analysis, but it could also be used to map a segment of DNA which is amplified between a pair of PCR primers. To do this, each hybrid would be tested to determine whether the expected DNA segment amplified. A potential problem of this type of analysis is the very high sensitivity of the PCR. DNA might be amplified from a hybrid where the relevant chromosome was present only at a low level and therefore not seen in the original cytogenetic characterisation. This would not matter however if the hybrid panel was recharacterised by PCR. Theune et al. (1991) have described a somatic cell hybrid mapping panel which was characterised using 48 pairs of PCR primers representing loci from the long and short arms of every autosome and the X chromosome.

A possible source of error in the cytogenetic characterisation of the hybrid panel described in this chapter is that chromosomal rearrangements may have occurred, particularly interspecific translocations (Boyd, 1986). If this is the case then fragments of human chromosomes may be present which have not been detected by G-banded karyotype analysis. In some cases, in order to confirm the localisations described in this chapter, additional experiments were performed. For example, the panel was probed with a DNA segment of known location to check the expected hybridisation pattern; a hybrid retaining the assigned human chromosome as its only

human complement was obtained and hybridisation of the probe tested; or an in situ hybridisation experiment was performed.

Karyotype analysis (see section 3.2.1) was performed while hybrid cells were being grown up in bulk for DNA preparation. This was done to minimise differences between the chromosome complements of the hybrid cells karyotyped and those from which the DNA was prepared. The risk of change in the chromosomal complements is greatest when the population of cells goes through a bottle-neck. Large numbers of cells (confluent 80 cm^2 flasks per freezing vial) were frozen for storage at the same time so that it would be possible to grow up more cells at a later date with minimal risk of a change in the chromosomal complements.

3.2 Results

3.2.1 Karyotype analysis

Details regarding the derivation of the 11 hybrids used in this study are listed in Chapter 2, table 2.1. At least 10 G-banded metaphase spreads were karyotyped from each hybrid. Table 3.1 lists the mean number of total chromosomes and the mean number of human chromosomes observed in each hybrid. The human autosomal representation is summarised in table 3.2 and ranges from a fragment of a chromosome arm (1pter-1p34 in WAG 8-1; see figure 3.1) to 18 out of the 22 autosomes (KIYAG 3-1-3). An example of a trypsin-Giemsa banded metaphase cell from the hybrid EDAG 2-9-3 which retains 19 human chromosomes is illustrated in figure 3.2.

The results of analysis of all hybrids presented in table 3.2 predict that unambiguous chromosomal localisation can be made for any autosome. Since many of the hybrids in this mapping panel are derived from cell lines with X;autosome translocations, many contain derivative chromosomes which include only a part of the relevant autosome (see table 3.2). Therefore a regional localisation is possible for some chromosomes using this mapping panel.

The results of hybridisation of anonymous DNA probes to the hybrid panel

Hybrid	Rodent parent	Cells analysed	Total chromosome number	Human chromosome number
EDAG 2-9-8	RAG	12	114	10
EDAG 2-9-5	RAG	10	120	9
EDAG 2-9-3	RAG	10	124	17
ORIM 7-1	Wg3h	15	44	6
WAG 8-1	RAG	20	55	1
DAG 5-2	RAG	10	130	14
KAG 2-3	RAG	14	69	8
KIYAG 3-1-3	RAG	10	148	24
5E-F9	RAG	10	136	9
TWDAM 1-2	Wg3h	10	58	15
TWDAM 4-2	Wg3h	10	59	16

Table 3.1 Summary of the karyotypic analysis of 11 somatic cell hybrids, indicating the number of G-banded cells analysed, the mean total number of chromosomes counted per cell and the mean number of human chromosomes counted per cell. Full details of the human chromosome complements of the hybrids are provided in table 3.2.

	Human chromosome																							Probe									
	X	dX	dA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	A	B	C	D	E	F	G	H	I	J
EDAG 2-9-8		+				+		+	*		‡		‡				+							‡		-	+	-	-	-	-	+	+	-	-
EDAG 2-9-5		+		+		+			*				‡			+	+		+					+	+	n	+	+	-	n	n	n	-	+	-
EDAG 2-9-3		+	+	+		+		+	+	+			‡		+	+		‡	+	+	+			‡	(+)	+	+	n	n	+	+	-	n	n	n
ORIM 7-1		+							+		(+)	(+)		+	+			(+)			+	*		(+)	(+)	+	-	-	+	+	+	(+)	(+)	-	+
WAG 8-1		+		*																						-	-	-	-	-	-	-	-	-	-
DAG 5-2		+		+	+	(+)			+	‡			+	+			‡				+	+		+	*	-	-	+	-	-	-	-	-	+	+
KAG 2-3		+					+			+	+		+	+	+			(+)			+		(+)		*	(+)	-	-	+	+	+	(+)	+	-	-
KIYAG 3-1-3		‡	+	‡	+	‡	+		+	+	‡		+	+	+	+	+	‡		+		+	+	+	+	+	+	+	+	+	+	+	+	+	+
5E-F9		+			*	+		+	+	+	+		+	+			+									-	-	-	-	-	-	+	+	-	-
TWDAM 1-2	‡										‡				‡		+	‡	‡		‡			‡		+	-	-	-	+	+	+	+	-	-
TWDAM 4-2	+					‡			+			+	+	‡		+	+	+	+		+	+	+		+		-	-	-	-	-	-	-	-	-	+

Table 3.2 Summary of karyotype analysis of the autosome mapping panel, and of probe hybridisation results.

Karyotype analysis: The presence of a human chromosome in the majority of karyotyped cells is indicated by '+' and in the minority of karyotyped cells by '(+)'. '‡' indicates that more than one chromosome was observed in the majority of cells examined. 'dX' and 'dA' refer to the der(X) and der(autosome) chromosomes respectively of the translocation patient from which the hybrids were derived (see Chapter 2, table 2.1). The presence of an autosomal fragment on a derivative chromosome is indicated by '*'.
Probe results: The presence '+' or absence '-' of hybridisation signal of 10 probes to the mapping panel is indicated. A faint hybridisation is indicated by '(+)' and an untested hybridisation by 'n'. The probes are: A, amylin; B, β_2I; C, RSA; D, 2.2β-1; E, Quog 3.8; F, Quog 3.4; G, 6.3(100G); H, LPL; I, 4.4B (4.1 kb band on EcoRI digest); and J, 4.4B (0.5 kb band on BamHI digest).

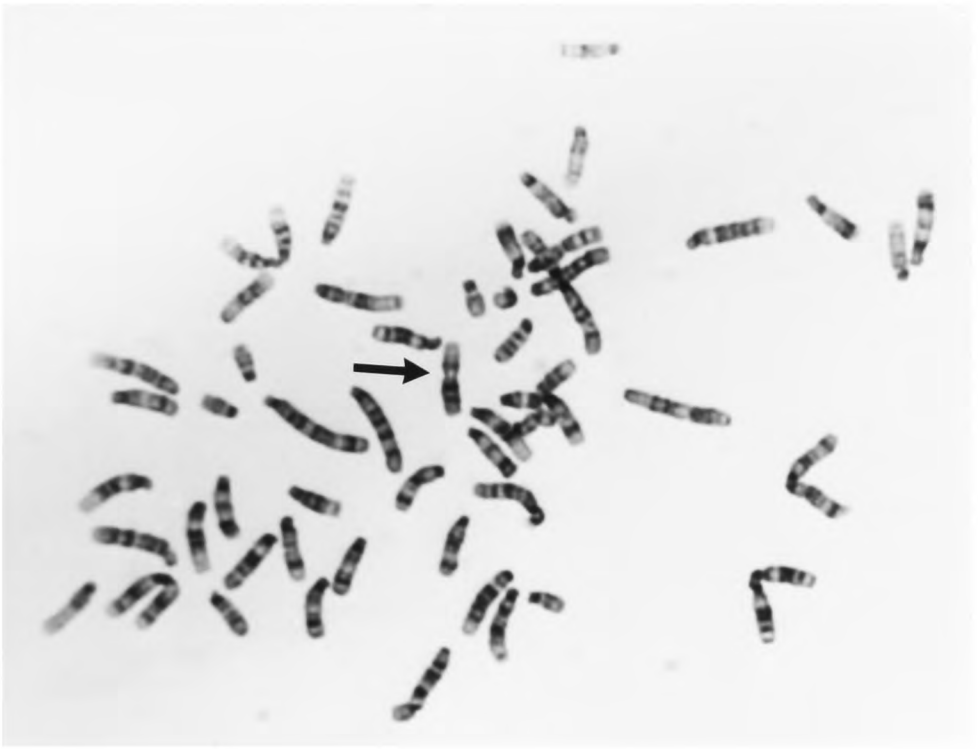


Figure 3.1 G-banded chromosome spread from the hybrid WAG 8-1 which retains the der(X) from an X;1 translocation as its only human complement (arrowed).

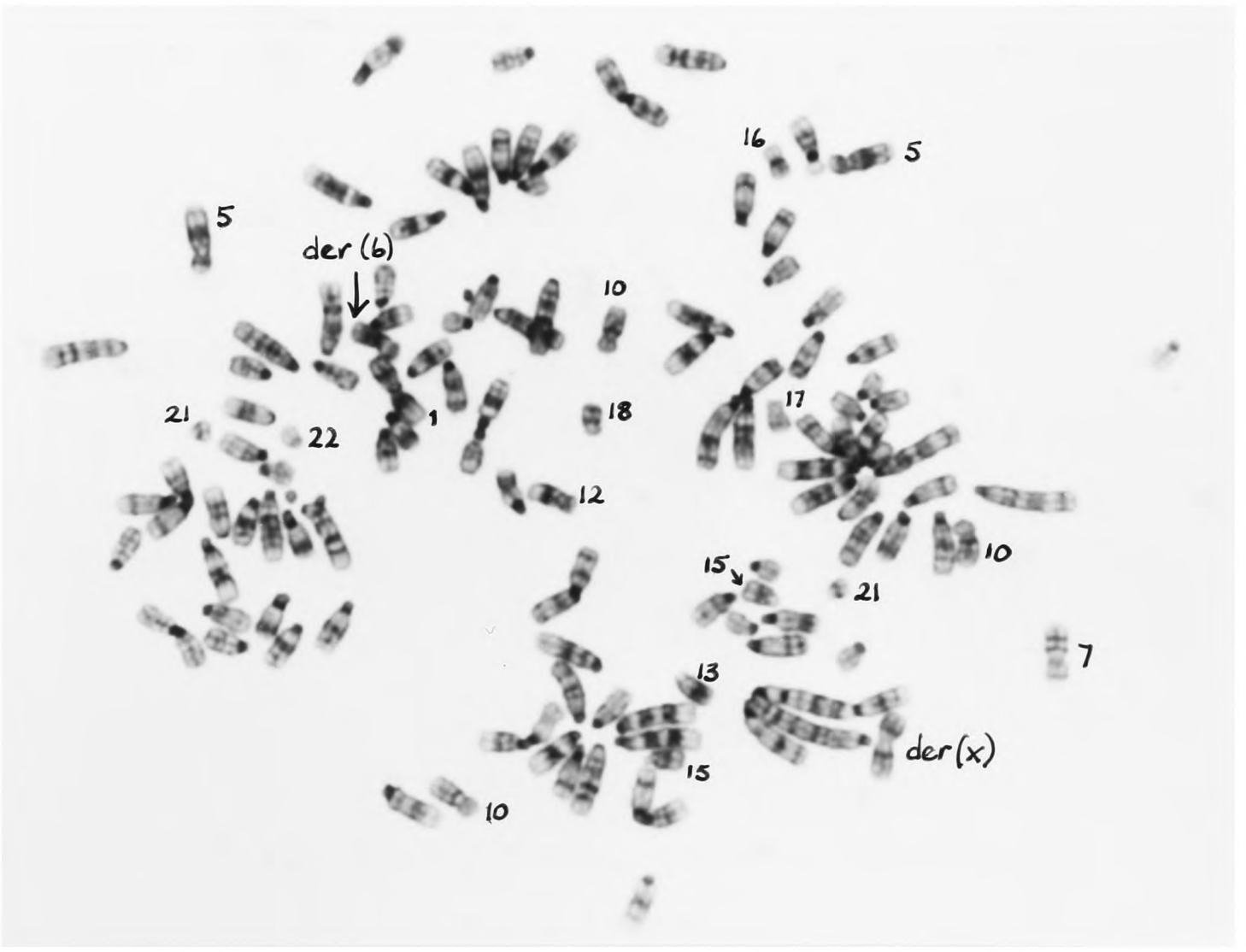


Figure 3.2 G-banded chromosome spread from the hybrid EDAG 2-9-3. This cell retains 19 human chromosomes. A summary of the full karyotypic analysis is included in tables 3.1 and 3.2.

described in this chapter, are presented in terms of concordance and discordance of the observed hybridisation pattern with the presence or absence of the autosomes in the karyotyped hybrids. 100% concordance indicates that hybridisation of the probe is detected in every hybrid where the chromosome was observed in karyotype analysis, and that no hybridisation is observed in hybrids not retaining this chromosome. Localisation of a probe to a chromosome is indicated by 100% concordance of the hybridisation pattern with the presence of this chromosome, and a lower concordance with the presence of every other chromosome.

3.2.2 In situ hybridisation of total human DNA to WAG 8-1

This experiment was performed in order to confirm the indication from trypsin-Giemsa banded analysis that the hybrid contained only a single human chromosome (section 3.2.1). Fifteen cells were examined for the presence of human chromosomes (indicated by dense cover with silver grains; see figure 3.3). One cell contained two human chromosomes of similar size and the remaining cells each contained one human chromosome of a size consistent with that expected for the der(X) observed in trypsin-Giemsa banded chromosome preparations. There was no evidence of inter-specific chromosomal rearrangements. In the cell where two human chromosomes were present in a single cell, it is possible that both were der(X) chromosomes. One cell with two der(X)s had also been identified out of 15 trypsin-Giemsa cells examined (not shown).

This experiment was performed before the majority of the other hybrids described in this chapter had been karyotyped. However it would be a valuable experiment to perform on all these hybrids. An attempt was made to do this using biotin labelled human DNA as a probe and a streptavidin-alkaline phosphatase detection system (BlueGENE detection kit; BRL). The chromosome slides were between 2 and 4 years old, and unfortunately the experiment was unsuccessful. The reason for this was probably the quality of the slides which appeared to have become coated in an oily film.

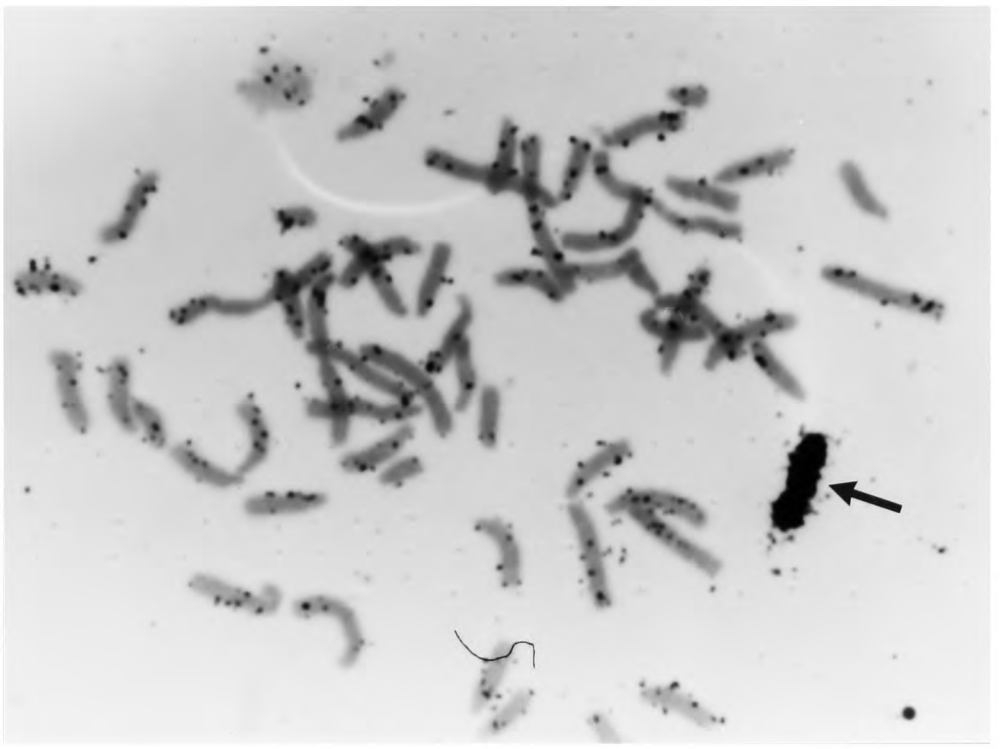


Figure 3.3 In situ hybridisation of tritium-labelled total human DNA to a chromosome spread from the hybrid WAG 8-1. A high concentration of silver grains over a single chromosome (arrowed) confirms that this hybrid retains a single human chromosome.

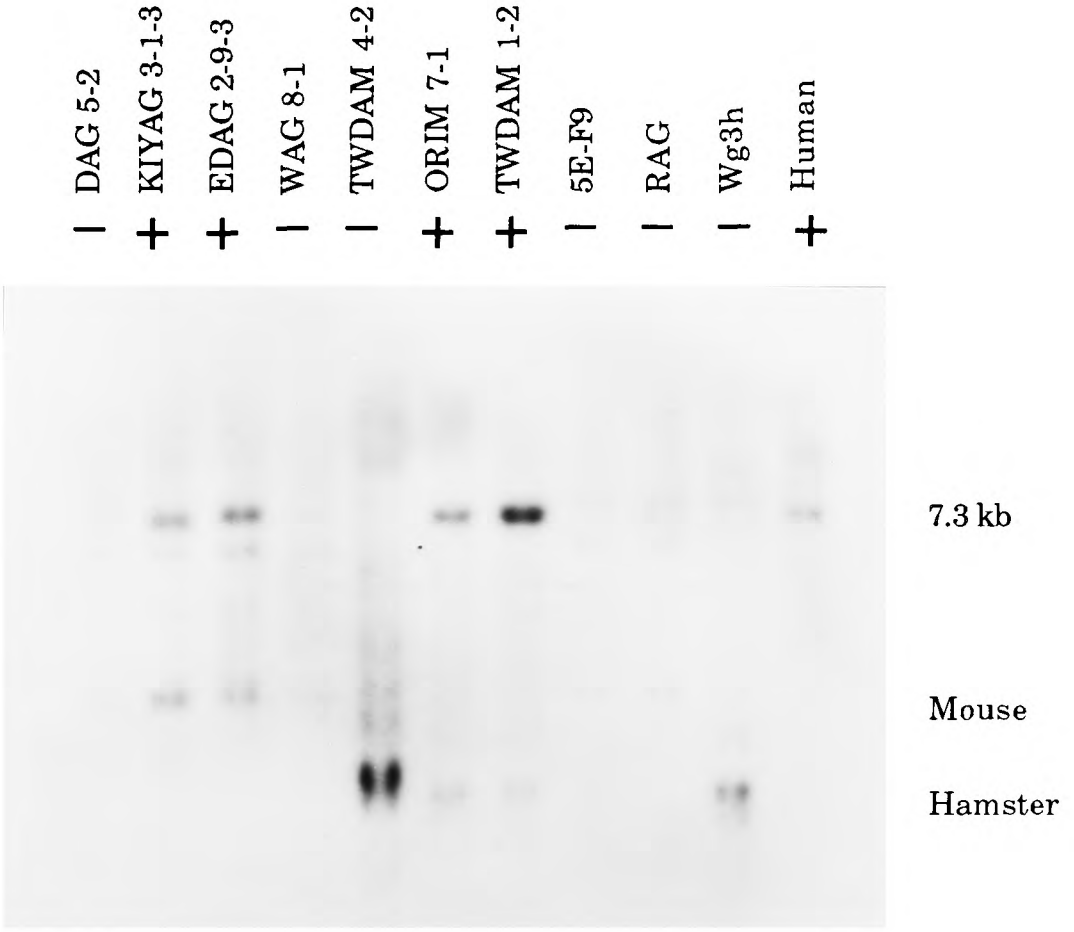


Figure 3.4 Hybridisation pattern produced by probing a Southern blot of HindIII digested DNA from eight hybrids with the amylin probe. Presence or absence (+/-) of a human-specific 7.3 kb band is indicated for each hybrid. Rodent cross-hybridising bands are also indicated.

Human and rodent chromosomes could be differentiated in control fresh slides from another hybrid.

3.2.3 Mapping results

A number of probes, including several cloned genes, were mapped on the hybrid panel in collaboration with donating laboratories. Each assignment is discussed separately below.

3.2.3.1 Amylin

Amylin or Islet amyloid polypeptide (IAPP) is a 37 amino acid peptide with biological activity which is found in the pancreases of normal individuals and is a major component of the pancreatic amyloid deposits found in patients with type 2 (non-insulin dependent) diabetes mellitus (Cooper et al., 1987; Leighton and Cooper, 1988). The amino acid sequence is over 40% identical with that of the calcitonin gene related peptide (CGRP) family members CGRP-1 AND CGRP-2 and shows weaker homology with insulin and insulin-like growth factors (IGF) (Cooper et al., 1987). With the exception of IGF-1, these have all been mapped to chromosome 11 (Kazazian and Junien, 1987).

A 110 bp DNA probe was supplied which had been cloned following in vitro DNA amplification of Islet of Langerhans derived cDNA by Taq polymerase, using nondegenerate oligonucleotide primers designed from the C and N termini of the peptide sequence and based on codon usage (Roberts et al., 1989b). The pattern of hybridisation of this probe to the autosome mapping panel is shown in figure 3.4 and the interpretation in table 3.2. This pattern is 100% concordant with the presence of human chromosome 12 (table 3.3). Only one hybrid is discordant for a localisation on chromosome 15 (table 3.3). This hybrid is TWDAM 4-2, which retains a chromosome 15 although no hybridisation signal was detected (figure 3.4; table 3.2).

Localisation to chromosome 12 was supported by in situ hybridisation of the same

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	5	5	50
1p34-qter	6	4	60
2pter-q27	5	5	50
2q27-2qter	4	6	40
3	4	6	40
4	7	3	70
5	4	6	40
6pter-q21	5	5	50
6q21-qter	4	6	40
7	6	4	60
8	7	3	70
9	5	5	50
10	4	6	40
11	5	5	50
12	10	0	100
13	6	4	60
14	3	7	30
15	9	1	90
16	6	4	60
17	7	3	70
18	7	3	70
19pter-q13	4	6	40
19q13-qter	5	5	50
20	6	4	60
21	7	3	70
22pter-q13	7	3	70
22q13-qter	7	3	70
Xpter-p21	7	3	70
Xp21-qter	5	5	50
Y	5	5	50

Table 3.3 Correlation of hybridisation signal from the amylin probe with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the amylin probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

probe to BudR replication banded metaphase chromosomes (Cockburn et al., 1989). Analysis detected 12% (23/187) of silver grains scored from 73 metaphases on chromosome 12 with 8.2% (15/187) assigned to the region 12cen-12q21.2. Data from other laboratories also confirm the assignment of amylin to chromosome 12 (Mosselman et al., 1988; Fan et al., 1989; Buckle et al., 1989).

3.2.3.2 Beta-2-glycoprotein I (β_2I)

Beta-2-glycoprotein I (β_2I) or apolipoprotein H (APOH) is a protein of uncertain physiological function which shows some association with plasma lipoproteins. Complete absence of β_2I has been described in several individuals with little consistent effect, and therefore it is suggested that this should be considered a 'nondisease' (OMIM 138700; Lozier et al., 1984; Hoeg et al., 1985). The cDNA nucleotide and deduced amino acid sequences indicate a mature protein of 326 amino acids consisting of five contiguous repeating units that belong to the Complement Control Protein superfamily (Lozier et al., 1984; Steinkasserer et al., 1991a). A 1.1 kb human cDNA clone was provided in order to determine the gene's chromosomal localisation. For comparison, the cDNA of a related gene, Regulator of Complement Activation (RCA), known to lie on chromosome 1 was also used to probe the autosomal mapping panel, because of the possible localisation of these genes on the same chromosome. The hybridisation of the control probe RCA was as expected for a localisation on chromosome 1p34-qter (figure 3.5 and tables 3.2 and 3.4). That of β_2I was different (figure 3.6). The signal obtained using this probe did not show 100% concordance with the presence of any single human chromosome (table 3.5). The highest concordance (82%) was with chromosomes 13 and 17. The discordant hybrids for a localisation on chromosome 13 are EDAG2-9-8 in which a signal was detected but no chromosome 13 had been observed, and TWDAM 4-2 in which no signal was detected but a chromosome had been observed (figure 3.6 and table 3.2). Therefore in order to explain a localisation on chromosome 13, EDAG 2-9-8

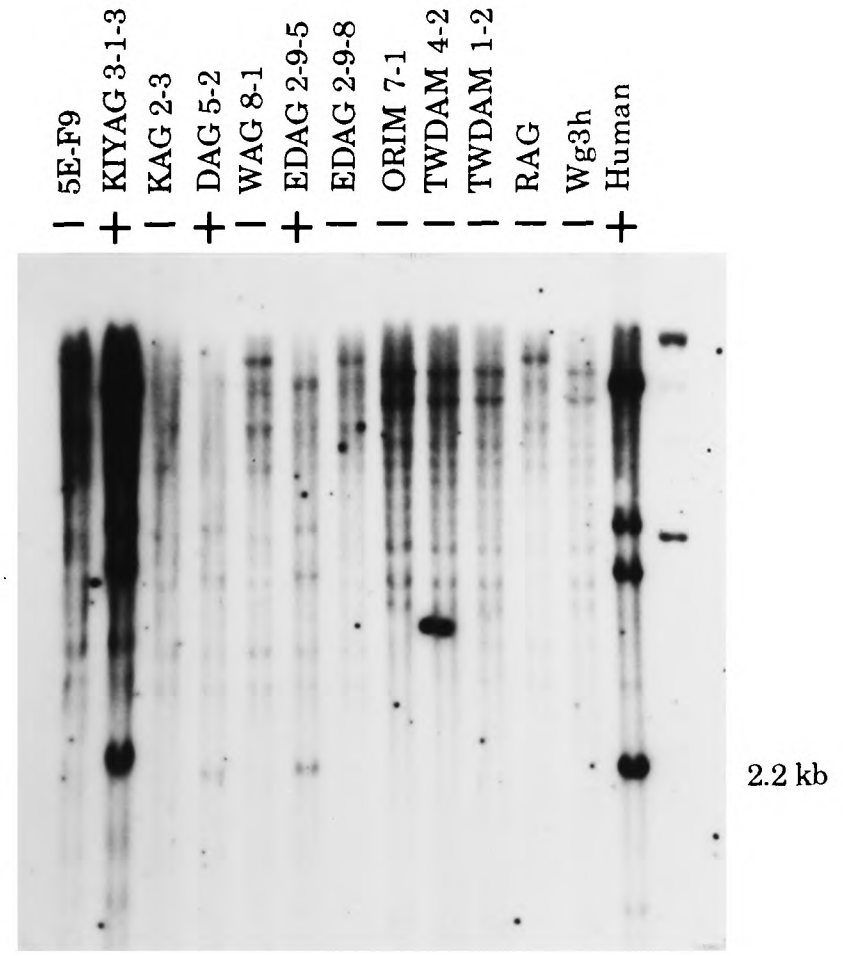


Figure 3.5 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from 10 hybrids with RCA. Presence or absence (+/-) of the scored 2.2 kb band is indicated for each hybrid.

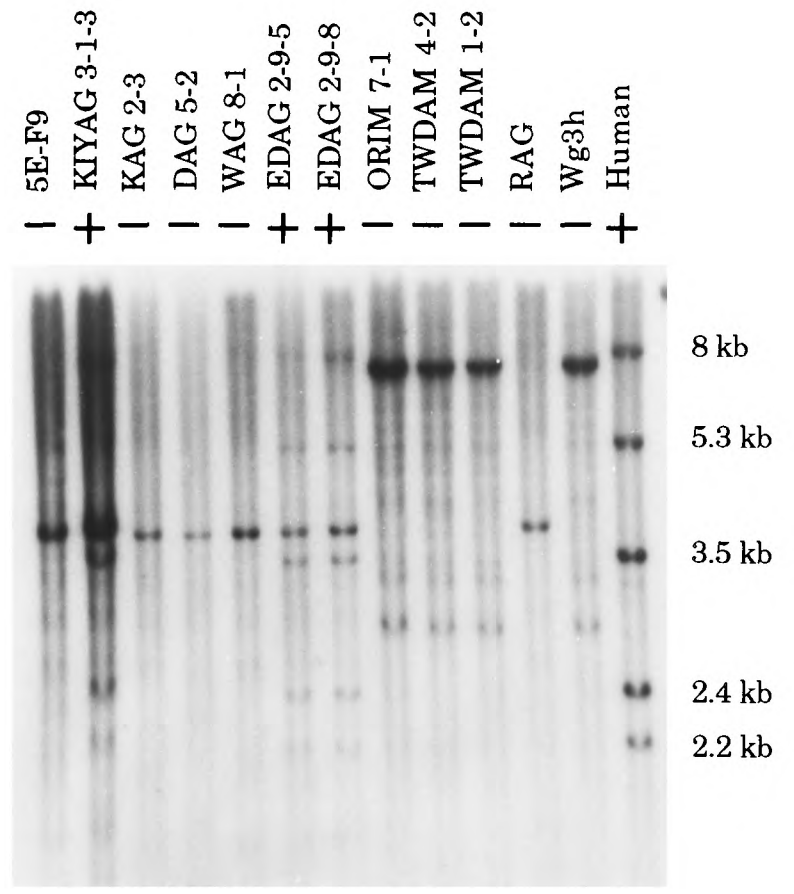


Figure 3.6 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from 10 hybrids with β2I. Presence or absence (+/-) of the scored 3.5, 2.4 and 2.2 kb bands is indicated for each hybrid.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	9	1	90
1p34-qter	10	0	100
2pter-q27	9	1	90
2q27-qter	8	2	80
3	7	3	70
4	7	3	70
5	5	5	50
6pter-q21	6	4	60
6q21-qter	6	4	60
7	7	3	70
8	3	7	30
9	5	5	50
10	6	4	60
11	5	5	50
12	5	5	50
13	8	2	80
14	6	4	60
15	4	6	40
16	6	4	60
17	8	2	80
18	4	6	40
19pter-q13	8	2	80
19q13-qter	7	3	70
20	6	4	60
21	7	3	70
22pter-q13	7	3	70
22q13-qter	7	3	70
Xpter-p21	6	4	60
Xp21-qter	3	7	30
Y	7	3	70

Table 3.4 Correlation of hybridisation signal from the RCA probe with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	8	3	73
1p34-qter	9	2	82
2pter-q27	7	4	64
2q27-qter	6	5	55
3	8	3	73
4	7	4	64
5	8	3	73
6pter-q21	5	6	45
6q21-qter	7	4	64
7	6	5	55
8	5	6	45
9	5	6	45
10	7	4	64
11	3	8	27
12	6	5	55
13	9	2	82
14	6	5	55
15	5	6	45
16	7	4	64
17	9	2	82
18	3	8	27
19pter-q13	6	5	55
19q13-qter	5	6	45
20	6	5	55
21	8	3	73
22pter-q13	8	3	73
22q13-qter	6	5	55
Xpter-p21	7	4	64
Xp21-qter	4	7	36
Y	7	4	64

Table 3.5 Correlation of hybridisation signal from the β_2I probe with human chromosomes present in 11 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the β_2I probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

must contain a fragment of chromosome 13 which includes the β_2I locus, and the chromosome 13 observed in TWDAM 4-2 must contain a deletion including this locus which is too small to be seen cytologically. There is a simpler explanation for a localisation on chromosome 17. Two hybrids are also discordant for this localisation, however, the hybrids are EDAG 2-9-8 and EDAG 2-9-5 which are subclones from the same parent hybrid EDAG 2-9. In both hybrids a signal was detected using the probe β_2I although no chromosome 17 had been observed in their characterisation (figure 3.6; table 3.2). This result can be explained by the presence of the same chromosome 17 fragment including the β_2I locus in both hybrids.

Assignment of β_2I to chromosome 17 was confirmed through observing positive hybridisation of the same probe to DNA from PCTBA1.8 (kindly supplied by E. Solomon), a somatic cell hybrid which retained chromosome 17 as its only human component (figure 3.7). This signal must be from the same locus as detected previously because the same sized fragments are detected. A regional localisation of the gene to 17q22-23 has now been made by hybridising the cDNA clone to a panel of chromosome 17 translocation somatic cell hybrids (Steinkasserer et al., 1991b).

The results of this experiment indicate that the hybrids EDAG 2-9-5 and EDAG 2-9-8 both contain a fragment of chromosome 17 which includes the β_2I locus. The size of this fragment is unknown but could be determined by hybridisation of chromosome 17 probes to Southern blots of hybrid DNA. These hybrids might then be useful additions to somatic cell hybrid panels for the regional localisation of chromosome 17 probes.

3.2.3.3 2.2β-1

The probe 2.2β-1 is a 600 bp EcoRI fragment subcloned from a phage (2.2) isolated from a library of a 'human X only' somatic cell hybrid with a mouse background (Laval et al., 1991; Chen unpublished data). The phage was picked on the basis of its hybridisation to a bovine monoamine oxidase probe (X-linked in the human) and the

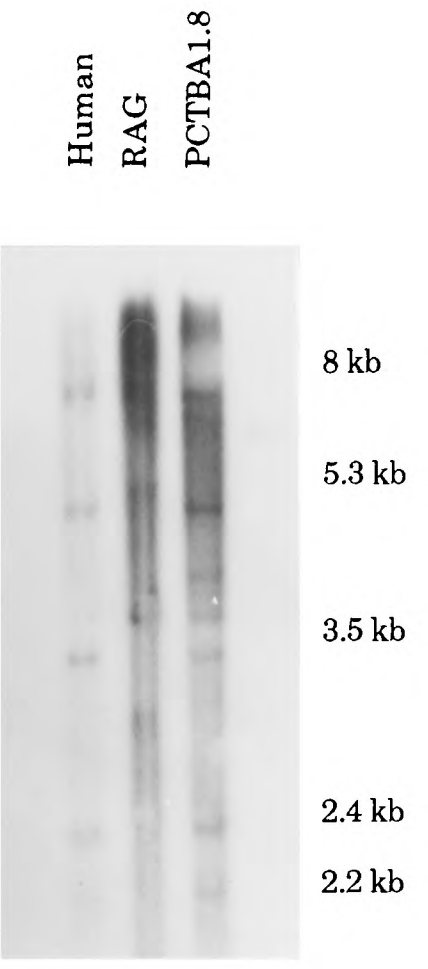


Figure 3.7 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA with β2I. The presence of the human-specific bands detected by this probe in the PCTBA1.8 track confirmed that the probe is from human chromosome 17, as this hybrid retains chromosome 17 as its only human complement.

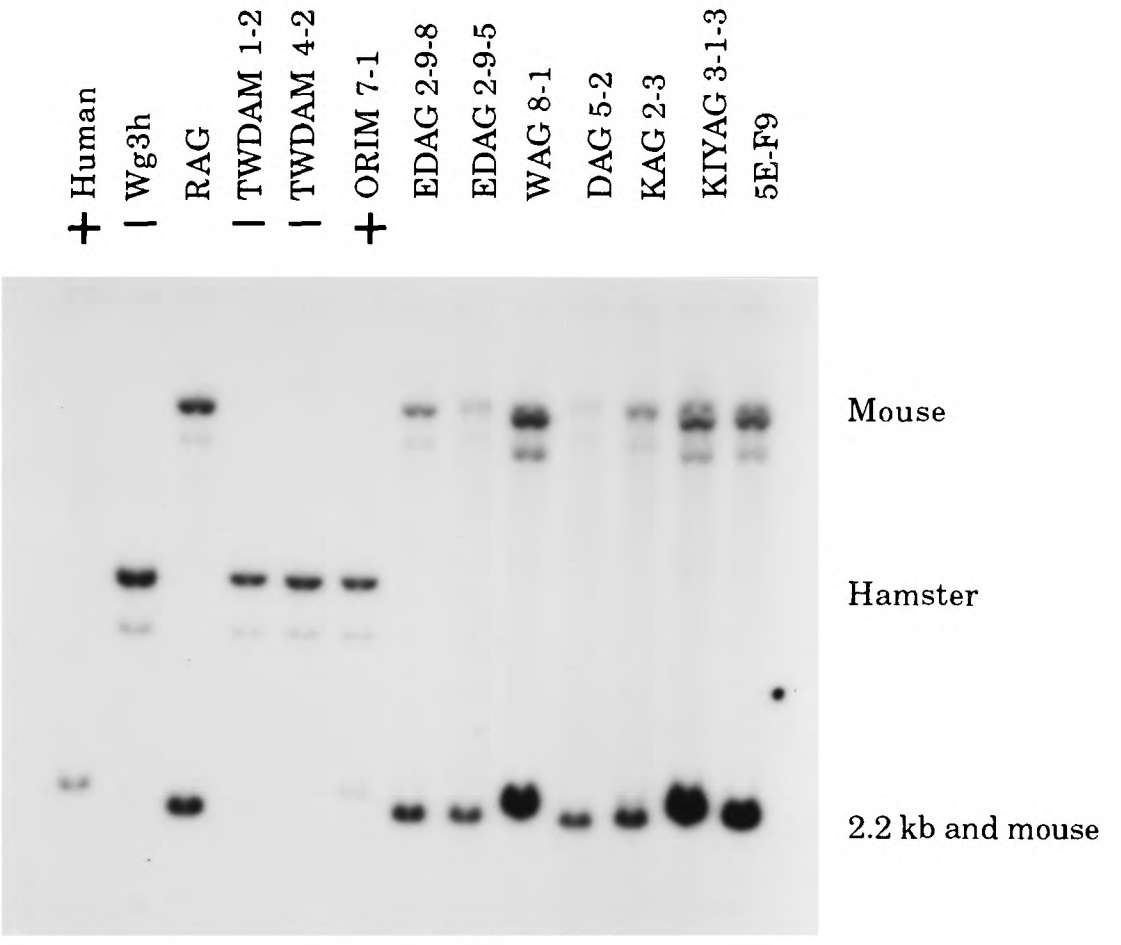


Figure 3.8 Hybridisation pattern produced by probing a Southern blot of PvuII digested DNA from 10 hybrids with 2.2β-1. The human signal (2.2 kb) co-migrates with a mouse signal under these conditions therefore only the human/hamster hybrids can be scored here.

cloned DNA was found to be murine in origin (as determined by strong hybridisation of mouse DNA to the clone's insert). Characterisation of 2.2β-1 showed that two restriction fragments, one X-linked and one autosomal, are detected in Southern blots of mouse DNA digested with several different restriction enzymes and that the autosomal band could be provisionally assigned to mouse chromosome 15 (Laval and Williamson, personal communication). It is thought that the clone was produced by a co-ligation of two DNA fragments during library construction, but this is not fully understood. In Southern analysis of human DNA a single autosomal hybridising band is generally detected. Cross species DNA hybridisation of this nature indicates DNA conservation and therefore possible functional significance of the DNA. In order to aid the characterisation of the DNA fragment I attempted to map this human cross hybridising signal.

Hybridisation of 2.2β-1 to the autosome mapping hybrid panel is illustrated in figures 3.8 and 3.9 (the human signal is separated from the mouse in a HindIII digest and from the hamster in a PvuII digest) and the interpretation included in table 3.2. The hybridisation signal is not 100% concordant with the presence of any human chromosome (table 3.6). The best scores (90% concordance) are for chromosomes 4 and 12. The discordant result for chromosome 12 is the absence of hybridisation signal from TWDAM 1-2 which was observed to retain the chromosome in karyotype analysis. Such a result could only be explained by a deletion within chromosome 12 in this hybrid, including 2.2β-1 but too small to be cytogenetically visible. The discordance for human chromosome 4 can be more easily explained. Here a signal was detected in the hybrid ORIM 7-1 where no chromosome 4 was observed in chromosome analysis. A chromosome rearrangement, most likely an inter-specific translocation, might be present so that the human chromosome fragment including the locus cross-hybridising to 2.2β-1 is not readily observed in routine G-banded analysis. Thus the most likely location of the DNA segment cross-hybridising to 2.2β-1 is on chromosome 4.

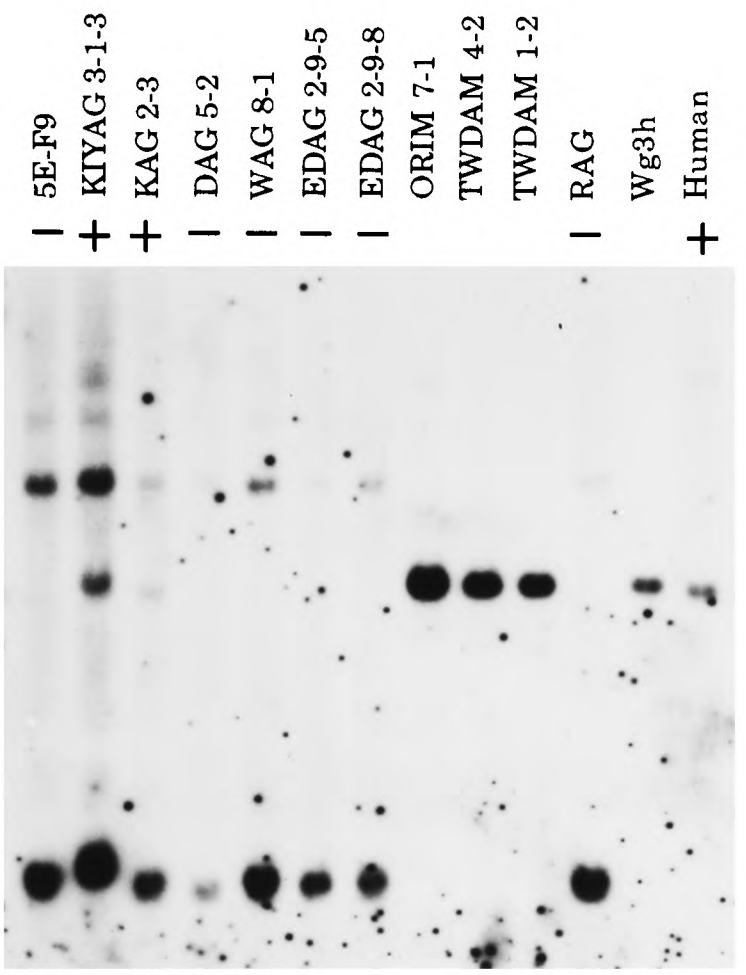


Figure 3.9 Hybridisation pattern produced by probing a Southern blot of HindIII digested DNA from 10 hybrids with 2.2β-1. The human signal (4.5 kb) co-migrates with the hamster signal under these conditions therefore only the human/mouse hybrids can be scored here.

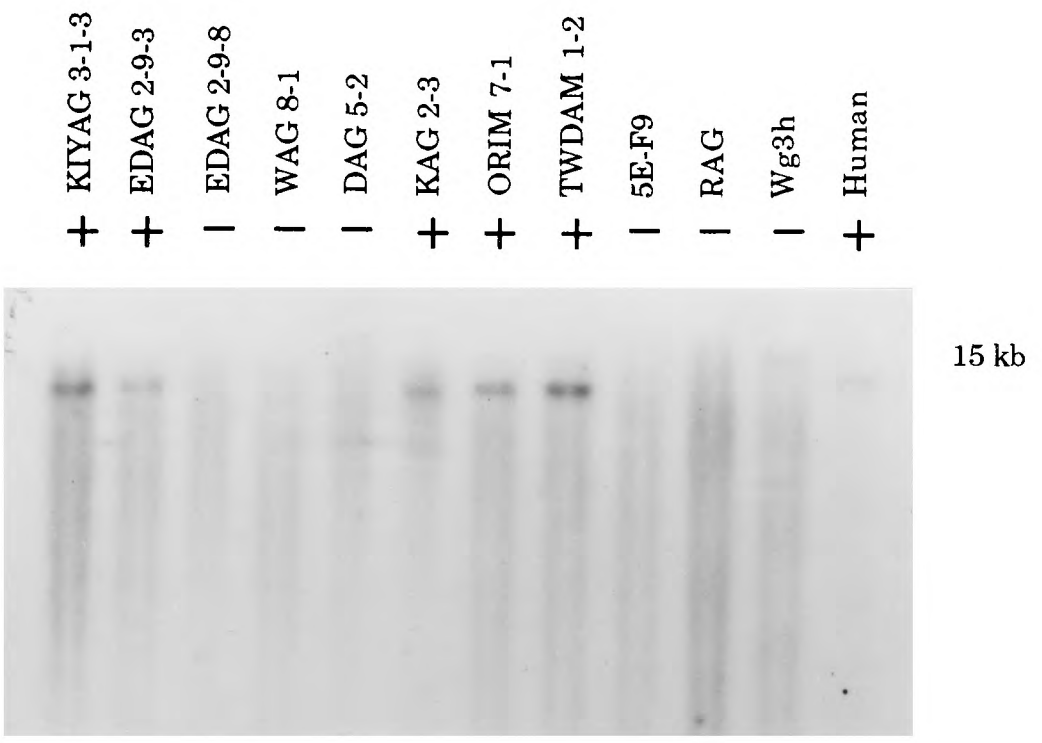


Figure 3.10 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from nine hybrids with Quog 3.8. The presence or absence (+/-) of the 15 kb human hybridisation signal is indicated for each hybrid.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	5	5	50
1p34-qter	6	4	60
2pter-q27	7	3	70
2q27-qter	6	4	60
3	4	6	40
4	9	1	90
5	5	5	50
6pter-q21	6	4	60
6q21-qter	4	6	40
7	7	3	70
8	7	3	70
9	7	3	70
10	4	6	40
11	7	3	70
12	9	1	90
13	6	4	60
14	2	8	20
15	8	2	80
16	4	6	40
17	8	2	80
18	6	4	60
19pter-q13	6	4	60
19q13-qter	7	3	70
20	8	2	80
21	5	5	50
22pter-q13	7	3	70
22q13-qter	7	3	70
Xpter-p21	6	4	60
Xp21-qter	3	7	30
Y	7	3	70

Table 3.6 Correlation of hybridisation signal from 2.2β-1 with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probe 2.2β-1 are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

If this localisation and that of the autosomal mouse signal on chromosome 15 are correct, then they may represent a hitherto unrecognised homologous block between these two species. A compilation of homologous DNA sequences mapped in both species has been constructed (Edwards, 1991). It shows homology of human chromosome 4 sequences, so far, only with those from mouse chromosomes 3 and 5 (five loci each). Homology of mouse chromosome 15 sequences have been observed with human sequences from chromosomes 5 (2 loci),8 (4 loci),11 (1 locus),12 (6 loci) and 22 (5 loci). Therefore if the human localisation of 2.2β-1 is on chromosome 12, it may add to the six loci so far known with homologous loci on mouse chromosome 15. Regions of homology between species provide insight into the course of chromosome evolution and suggest candidate mouse mutations for human genetic diseases which then may provide a useful model for the study of the natural course and treatment of the disease.

The Southern blot, to which the probe 2.2β-1 was hybridised, contained DNA from 10 of the 11 karyotyped somatic cell hybrids. The hybrid not included was EDAG 2-9-3 since no DNA was left from this hybrid. This hybrid retains chromosome 12 but does not retain chromosome 4 and therefore should be able to determine to which of the two chromosomes 2.2β-1 hybridises.

The results of this experiment suggest that either the hybrid ORIM 7-1 contains a fragment of human chromosome 4 which includes the locus detected by 2.2β-1, or that the chromosome 12 present in the hybrid TWDAM 1-2 contains a deletion including this locus. An in situ hybridisation experiment using total human DNA as a probe would be a straight forward way of looking for an inter-specific translocation in ORIM 7-1 (see section 3.3).

3.2.3.4 Quog 3.8 and Quog 3.4

Two probes (Quog 3.8 and Quog 3.4) which detect RFLPs were supplied for mapping. Both probes were isolated from a single cosmid derived from a total human cosmid library. They were both mapped in case they had been co-ligated into the cosmid and

represented different chromosomal loci. They produced identical patterns of hybridisation to the hybrid mapping panel as expected for their origin from a single cloning event (see figures 3.10 and 3.11). This pattern is the same as that produced by the amylin probe (section 3.2.3.1) and is 100% concordant with the localisation of these sequences to chromosome 12 (table 3.7; Newton et al., 1991). The addition of these RFLPs to the existing list of polymorphisms on chromosome 12 improves this valuable resource for linkage analysis and gene mapping (Craig and McBride, 1990).

3.2.3.5 6.3(100G) and 4.4B

The probe M27β detects a VNTR at Xp11.2 and the repeating DNA unit has been sequenced (Fraser et al., 1989). An artificial oligonucleotide (HXOX) homologous to part of the repeat sequence was used to isolate clones from a human phage library which may contain similar VNTR structures. VNTR polymorphisms have not yet been found using these clones although the inserts hybridise strongly when re-screened with HXOX (Riley, personal communication). The hybridisation of DNA fragments from two such clones, 6.3(100G) and 4.4B, to the autosome mapping panel is described below.

6.3(100G): The pattern of hybridisation of a 4 kb fragment from 6.3(100G) was 100% concordant with the presence of human chromosome 8 in the hybrid mapping panel (figure 3.12, tables 3.2 and 3.8). The gene for lipoprotein lipase (LPL) is known to lie on chromosome 8 (Sparkes et al., 1987) and a cDNA probe for this gene was used to check the expected pattern of hybridisation for a probe on this chromosome (figure 3.13). The pattern of hybridisation is consistent with that observed using 6.3(100G) and is 100% concordant with the presence of chromosome 8 in the hybrid panel (table 3.9).

4.4B: A 1.5 kb fragment from 4.4B was found to hybridise to several human genomic DNA fragments in Southern blot analysis using a variety of different restriction enzymes, therefore suggesting that the probe hybridised to several different loci (Riley, personal communication). Additionally, hybridisation to rodent DNA was found signifying conservation of the sequence and a possible functional role. On

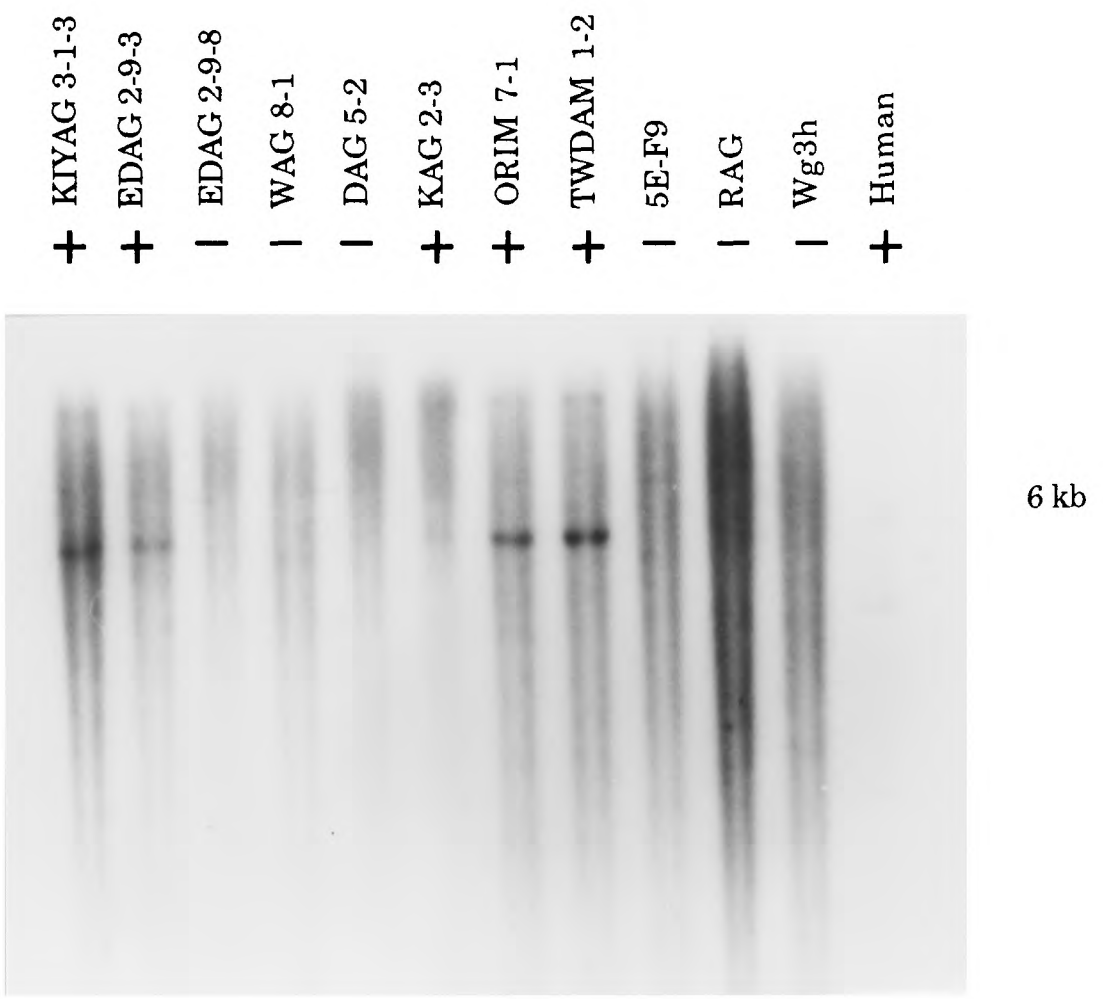


Figure 3.11 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from nine hybrids with Quog 3.4. The presence or absence (+/-) of the 6 kb human hybridisation signal is indicated for each hybrid.

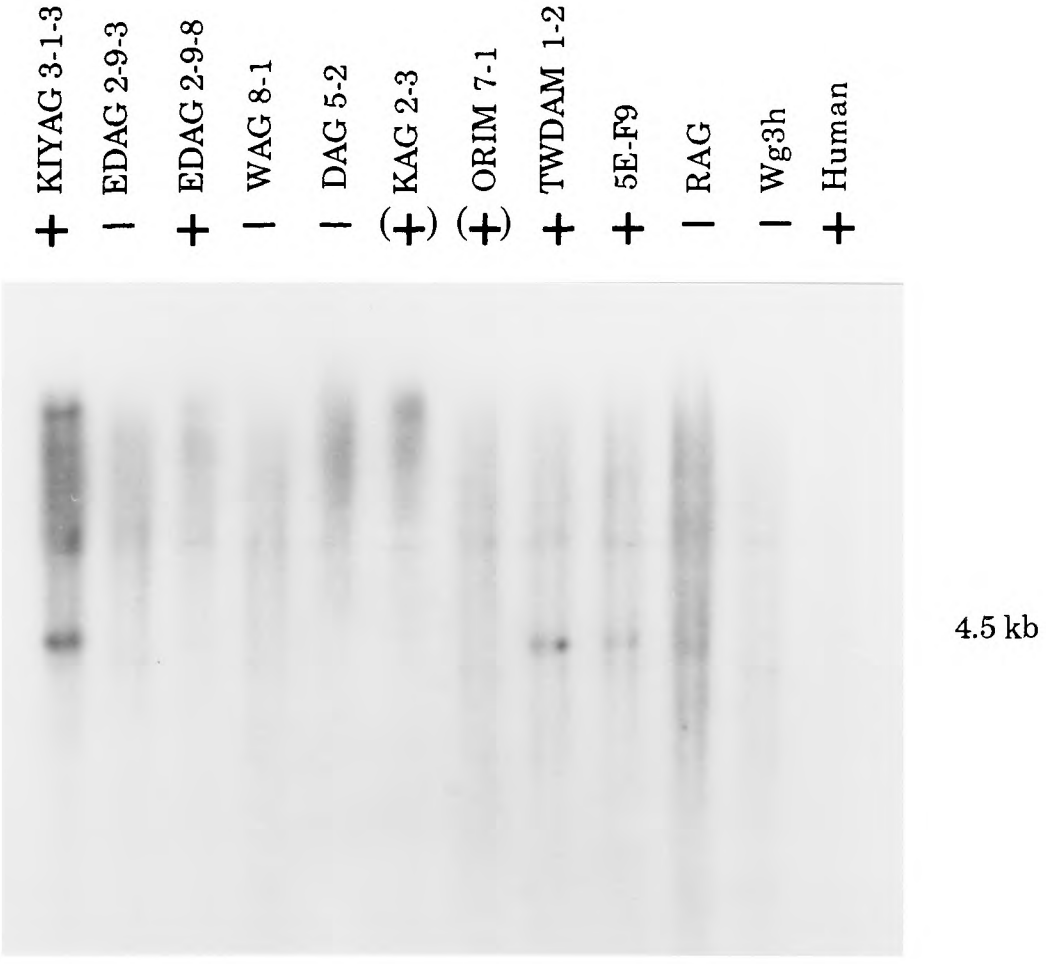


Figure 3.12 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from nine hybrids with 6.3(100G). The presence or absence (+/-) of the 4.5 kb human hybridisation signal is indicated for each hybrid. Faint signals in the hybrids KAG 2-3 and ORIM 7-1 are visible on the original autoradiograph.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	5	5	50
1p34-qter	6	4	60
2pter-q27	5	5	50
2q27-2qter	4	6	40
3	4	6	40
4	7	3	70
5	4	6	40
6pter-q21	5	5	50
6q21-qter	4	6	40
7	6	4	60
8	7	3	70
9	5	5	50
10	4	6	40
11	5	5	50
12	10	0	100
13	6	4	60
14	3	7	30
15	9	1	90
16	6	4	60
17	7	3	70
18	7	3	70
19pter-q13	4	6	40
19q13-qter	5	5	50
20	6	4	60
21	7	3	70
22pter-q13	7	3	70
22q13-qter	7	3	70
Xpter-p21	7	3	70
Xp21-qter	5	5	50
Y	5	5	50

Table 3.7 Correlation of hybridisation signal from Quog 3.8 and Quog 3.4 with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probes are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	3	7	30
1p34-qter	4	6	40
2pter-q27	4	6	40
2q27-qter	5	5	50
3	4	6	40
4	6	4	60
5	5	5	50
6pter-q21	4	6	40
6q21-qter	5	5	50
7	5	5	50
8	10	0	100
9	4	6	40
10	5	5	50
11	6	4	60
12	7	3	70
13	3	7	30
14	6	4	60
15	6	4	60
16	3	7	30
17	4	6	40
18	4	6	40
19pter-q13	3	7	30
19q13-qter	4	6	40
20	5	5	50
21	6	4	60
22pter-q13	4	6	40
22q13-qter	4	6	40
Xpter-p21	4	6	40
Xp21-qter	6	4	60
Y	4	6	40

Table 3.8 Correlation of hybridisation signal from 6.3(100G) with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probe 6.3(100G) are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

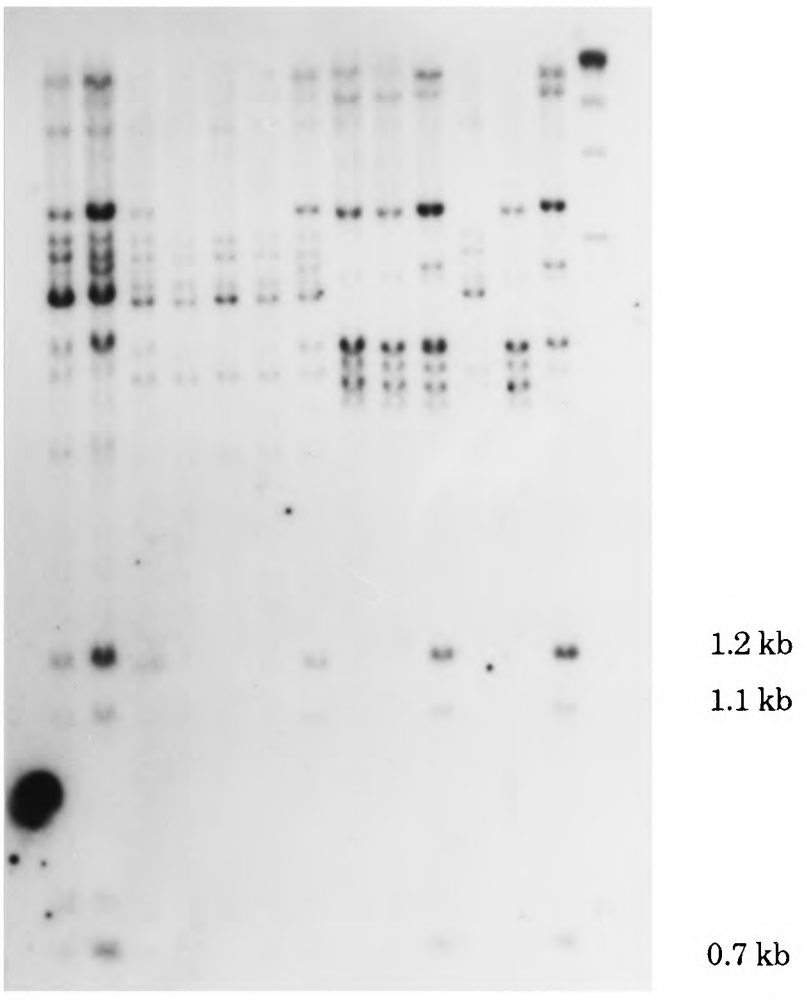


Figure 3.13 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from 10 hybrids with LPL. Presence or absence (+/-) of the scored 1.2, 1.1 and 0.7 kb bands is indicated for each hybrid. A faint hybridisation signal in ORIM 7-1 is visible on the original autoradiograph.

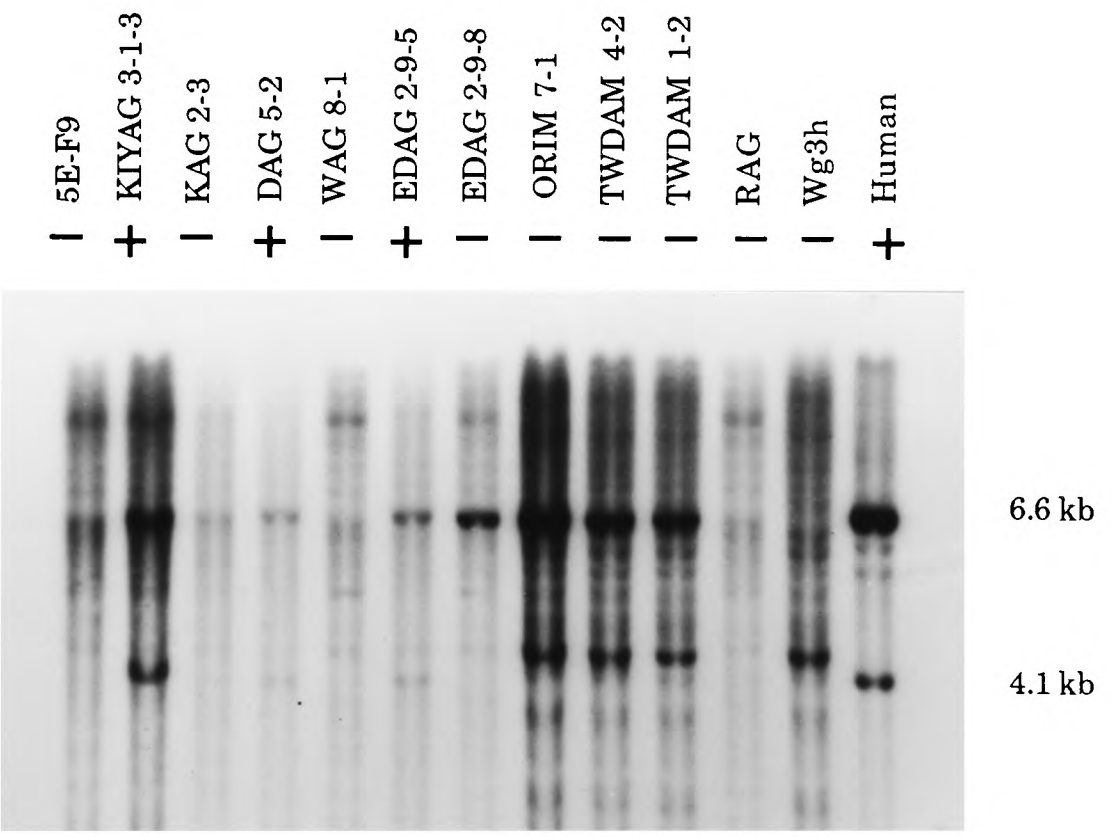


Figure 3.14 Hybridisation pattern produced by probing a Southern blot of EcoRI digested DNA from 10 hybrids with 4.4B. Presence or absence (+/-) of the scored 4.1 kb band is indicated for each hybrid. The 6.6 kb band was not scored (see text).

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	3	7	30
1p34-qter	4	6	40
2pter-q27	4	6	40
2q27-qter	5	5	50
3	4	6	40
4	6	4	60
5	6	4	60
6pter-q21	5	5	50
6q21-qter	5	5	50
7	6	4	60
8	10	0	100
9	4	6	40
10	5	5	50
11	6	4	60
12	8	2	80
13	3	7	30
14	5	5	50
15	7	3	70
16	3	7	30
17	5	5	50
18	5	5	50
19pter-q13	3	7	30
19q13-qter	4	6	40
20	5	5	50
21	6	4	60
22pter-q13	4	6	40
22q13-qter	4	6	40
Xpter-p21	5	5	50
Xp21-qter	6	4	60
Y	4	6	40

Table 3.9 Correlation of hybridisation signal from the LPL cDNA probe with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the LPL cDNA probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

hybridising the fragment to an EcoRI Southern blot of the autosome mapping panel two principal human bands were observed (figure 3.14). The probe did not contain an EcoRI site (Riley, personal communication). The stronger band of 6.6 kb was not 100% concordant with the presence of any single human chromosome. The other (4.1 kb) showed a pattern of hybridisation 100% concordant with a chromosome 1p34-qter localisation (table 3.10) and with band intensities the same as observed for RSA (figure 3.5). When hybridised to a BamHI Southern blot of the same panel three principal bands were observed (figure 3.15). The probe did not contain a BamHI site (Riley, personal communication). The 4.7 kb fragment co-migrated with a strong hamster band but in the mouse hybrids produced the same pattern of hybridisation as the 6.6 kb fragment in the EcoRI digest. This was also the hybridisation pattern of the 1.4 BamHI band. The 0.5 kb band however produced a new pattern of hybridisation which was 100% concordant with a localisation on chromosome 19q12-qter (tables 3.2 and 3.11).

Given that the hybridisation signals produced by this probe were stronger than those of other single copy probes used with the same filters, it is possible that the stronger signals here represent more than one locus from different chromosomes and therefore cannot be mapped by somatic cell hybrid analysis. It is particularly intriguing that the probe also cross-hybridises to rodent DNA indicating a possible functional role of the sequence. The logical experiment to follow this investigation is an <u>in situ</u> hybridisation. This should confirm signals on 1p34-qter and 19q12-qter which I have predicted and identify any other loci but has not yet been performed.

3.2.3.6 C1QA, C1QB and C1QC

The genes C1QA, C1QB and C1QC code for the A-, B- and C- chains of human complement subcomponent C1q. C1QA and C1QB have been localised on the short arm of chromosome 1 (Hedge et al., 1987; Solomon et al., 1985). I have further positioned these genes relative to the chromosome 1 breakpoint of an X;1 translocation (p21;p34) (see Chapter 4). A cDNA probe for C1QB was hybridised to a Southern blot including

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	9	1	90
1p34-qter	10	0	100
2pter-q27	9	1	90
2q27-qter	8	2	80
3	7	3	70
4	7	3	70
5	5	5	50
6pter-q21	6	4	60
6q21-qter	6	4	60
7	7	3	70
8	3	7	30
9	5	5	50
10	6	4	60
11	5	5	50
12	5	5	50
13	8	2	80
14	6	4	60
15	4	6	40
16	6	4	60
17	8	2	80
18	4	6	40
19pter-q13	8	2	80
19q13-qter	7	3	70
20	6	4	60
21	7	3	70
22pter-q13	7	3	70
22q13-qter	7	3	70
Xpter-p21	6	4	60
Xp21-qter	3	7	30
Y	7	3	70

Table 3.10 Correlation of 4.1 kb EcoRI hybridisation signal from 4.4B with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

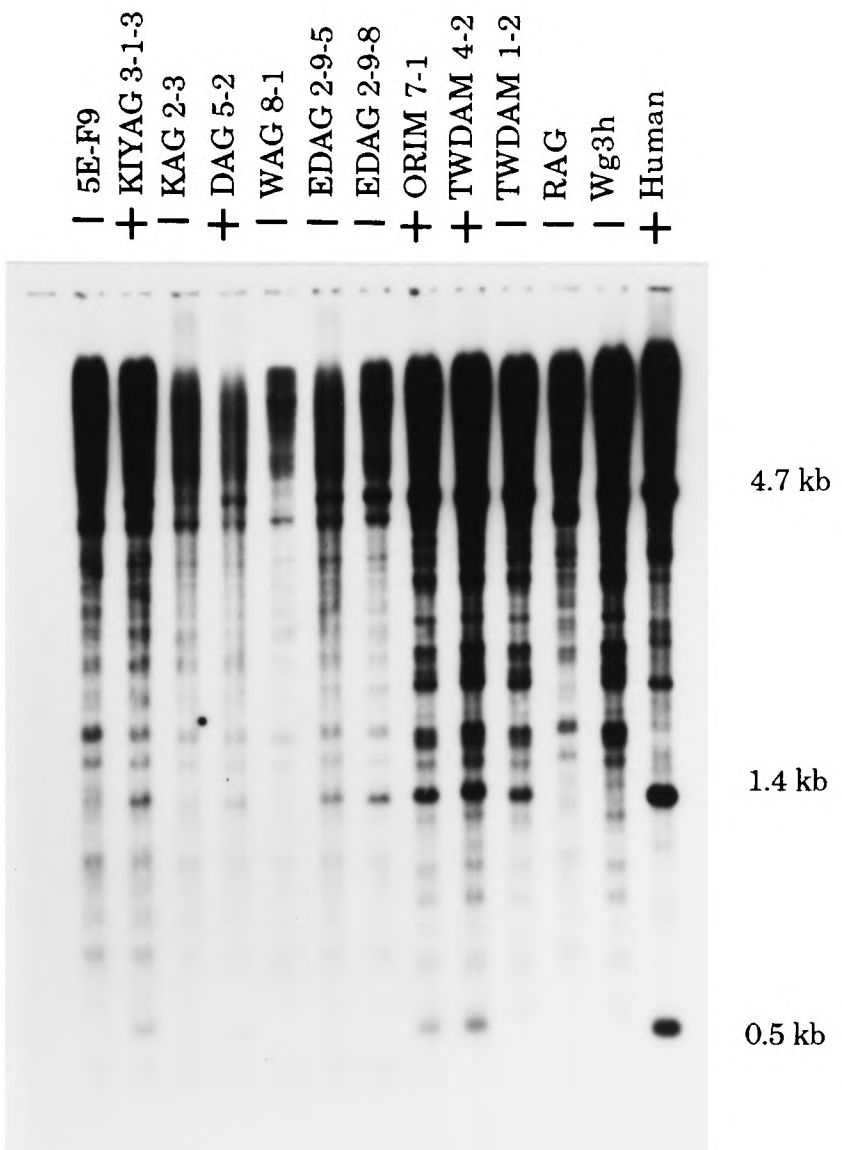


Figure 3.15 Hybridisation pattern produced by probing a Southern blot of BamHI digested DNA from 10 hybrids with 4.4B. Presence or absence (+/-) of the scored 0.5 kb band is indicated for each hybrid. The 4.7 and 1.4 kb bands were not scored (see text).

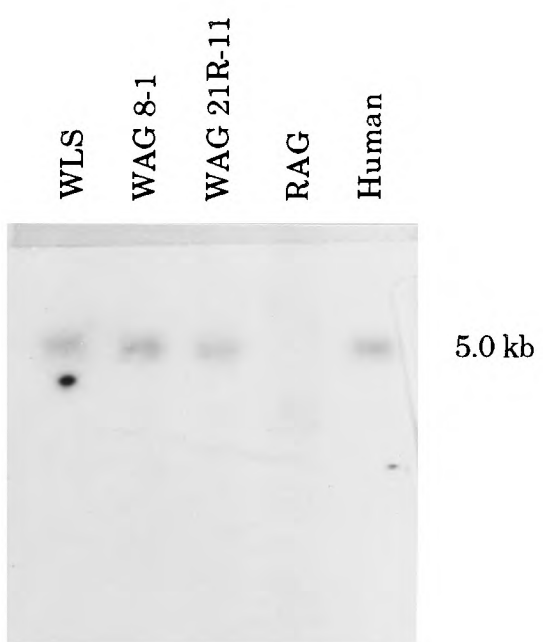


Figure 3.16 Hybridisation of a cDNA probe for C1QB to a Southern blot of PvuII digested DNA. The presence of a 5.0 kb signal in the hybrid WAG 8-1 indicates a regional localisation of this locus to 1pter-p34.

Chromosome	Concordant hybrids	Discordant hybrids	% Concordance
1pter-p34	6	4	60
1p34-qter	7	3	70
2pter-q27	8	2	80
2q27-qter	7	3	70
3	6	4	60
4	6	4	60
5	4	6	40
6pter-q21	9	1	90
6q21-qter	7	3	70
7	6	4	60
8	4	6	40
9	8	2	80
10	5	5	50
11	8	2	80
12	6	4	60
13	7	3	70
14	5	5	50
15	7	3	70
16	5	5	50
17	7	3	70
18	7	3	70
19pter-q13	9	1	90
19q13-qter	10	0	100
20	7	3	70
21	6	4	60
22pter-q13	8	2	80
22q13-qter	8	2	80
Xpter-p21	7	3	70
Xp21-qter	4	6	40
Y	6	4	60

Table 3.11 Correlation of 0.5 kb BamHI hybridisation signal from 4.4B with human chromosomes present in 10 somatic cell hybrids.

The numbers of somatic cell hybrids which are concordant and discordant with the probe are shown for each chromosome. Some hybrids contain translocated chromosomes, hence concordance figures are shown separately for the chromosomal regions either side of the breakpoint.

DNA from the somatic cell hybrid WAG 8-1 which retains the der(X) chromosome from this translocation as its only human chromosome and therefore only includes the region 1pter-p34 from a normal chromosome 1 (see section 3.2.1). The presence of a band corresponding to C1QB in the WAG 8-1 track indicates that this gene is located distal to the chromosome 1 breakpoint 1p34 (figure 3.16). The presence of a band in the reciprocal translocation hybrid WAG 21R-11 can be attributed to the presence of a normal chromosome 1 in this hybrid (see Chapter 2, table 2.1). The three loci, C1QA, C1QB and C1QC have been cloned and found to lie within a 24 kb DNA genomic interval (Sellar et al., 1991). Thus all three loci map to 1pter-1p34.

3.3 Discussion

The characterisation of a panel of 11 somatic cell hybrids has permitted the chromosomal localisation of seven new loci. One hybrid retains a single human chromosome, one retains 18 out of the 22 autosomes and the others retain an intermediate number. The number of hybrids karyotyped was probably about the minimum required to construct a useful panel. If only monochromosomal hybrids were used in a panel capable of localising a DNA fragment from any autosome, then 22 hybrids, each retaining one of the autosomes, would be required to construct this panel. The most informative hybrids in an autosome mapping panel are those which retain approximately half of the human chromosomes. In some of the mapping experiments described in this chapter, 90% concordance was found between the hybridisation pattern and the presence of more than one autosome (for example in the mapping of Amylin, section 3.2.3.1 and in the mapping of 2.2β-1, section 3.2.3.3). The confidence in the mapping results would be improved if a few additional hybrids were added to this panel. In two recent publications of autosomal localisations by somatic cell hybrid analysis, the numbers of hybrids in the panels were 18 and 25 (Dodge et al., 1991; Tait et al., 1991).

Use of this panel indicated that errors of characterisation were present (such as

fragments of chromosome 17 in EDAG 2-9-5 and EDAG 2-9-8, see section 3.2.3.2). Therefore, although this indicates that results from this type of analysis must be treated with caution, it also indicates that the more the hybrid panel is used to localise probes the better characterised it will become. One very worthwhile further study in the characterisation of these hybrids would be in situ hybridisation of total human and/or rodent DNA to metaphase chromosome preparations. This experiment has only been performed on the hybrid WAG 8-1, but was attempted for the other hybrids and unfortunately failed (see section 3.2.2). The value of this experiment is that it would identify hybrids which contain interspecific chromosome rearrangements and would allow the sizes of human chromosome fragments involved to be estimated.

PCR based methods have now been described for the rapid characterisation of the autosomal contents of somatic cell hybrids. For example, hybrids may be screened using PCR primer pairs from all the human autosomes (Theune et al., 1991). An alternative approach is based on Alu-PCR or L1-PCR whereby specific amplification of human DNA fragments from hybrids is achieved (Nelson et al., 1989). Using a dot-blot method, amplified DNA from a hybrid can be hybridised to similarly amplified DNA from a reference panel of monochromosomal or highly reduced hybrids to produce a 'PCR-karyotype' (Ledbetter et al., 1990; Bicknell et al., 1991). This approach would enable the identification of human marker chromosomes or translocated fragments of human chromosomes that are detected but not readily identified by G-banded cytogenetic analysis. An alternative method for identification of small marker chromosomes is chromosomal in situ suppression (CISS-) hybridisation, where libraries of flow-sorted chromosomes are used complex probes (Jauch et al., 1990).

This chapter has described the characterisation and application of a somatic cell hybrid panel in the mapping of autosomal DNA fragments. The panel was constructed after working on the characterisation of somatic cell hybrids which contain derivative chromosomes from X;autosome translocations. The principal reason for the

characterisation was in order to map the X chromosome breakpoints of these translocations with respect to DNA markers from the DMD locus. The mapping of translocation breakpoints within the DMD locus using cDNA fragments as probes is described in the following chapter.

3.4 Acknowledgements

I am grateful to the many collaborators who mainly supplied me with probes or who probed filters which I supplied to them. The collaborators were: Ann Roberts and Sheila Holt (Amylin); Alexander Steinkasserer (β_2I); Steve Laval (2.2β-1); Robert Newton (Quog 3.8 and Quog 3.4); Sue Riley (6.3(100G) and 4.4B); and Grant Sellar (C1QA and C1QB). I am also grateful to Penny Williamson for assistance with an experiment to characterise the hybrids by in situ hybridisation using biotin labelled total human DNA. This experiment failed due to deterioration of my chromosome slides.

Chapter 4 Refined mapping of DMD-associated translocations using cDNA probes

4.1 Introduction

In this chapter, the detailed mapping of DMD-associated X;autosome translocation breakpoints with respect to exons of the DMD gene is described. The mapping of seven breakpoints is given here, and data from seven additional translocations whose breakpoints have been determined elsewhere is included in an examination of the distribution of the breakpoints over the physical DMD gene map. This distribution is compared with that of deletions and duplications which have been reported in male patients.

There are 24 females known worldwide with DMD/BMD and a constitutional X;autosome translocation involving the band Xp21 (Boyd 1991; see Chapter 1, table 1.1). The translocations are associated with preferential inactivation of the normal X, a feature common to other X;autosome translocations, and the effect is that only loci on the active, translocated X chromosome fragments are expressed (Mattei et al., 1981; further discussed in Chapter 1, section 1.4). Therefore the expectation is that the translocation produces the disease by direct disruption to dystrophin gene expression. This prediction has been confirmed by mapping studies which have shown that the X chromosome breakpoints of 14 translocations all lie within the coding region of the dystrophin locus (Boyd et al.,1988; Meitinger et al., 1988; Giacalone and Franke, 1988; Bodrug et al., 1989; Bodrug et al.,1991).

Mapping studies have been performed by three principal methods of analysis. These are somatic cell hybrid analysis, pulsed field gel electrophoresis (PFGE), and in situ hybridisation. Of these, somatic cell hybrid analysis has established itself as the most versatile method. In situ hybridisation has proved most useful for obtaining key results not immediately obtainable by other methods. All the methods rely on the hybridisation of DNA probes from the DMD region and on recognising whether the signals detected derive from the normal X, the der(X) or the der(autosome). In in situ

hybridisation the three chromosomes are identified at the microscope and the hybridisation signals for each chromosome are scored after examination of many metaphase spreads (Buckle and Craig, 1986; Boyd et al., 1986). Somatic cell hybrid construction physically isolates one of the translocation derivative chromosomes from the other and from the normal X chromosome in a rodent background, thus DNA prepared from these hybrids contains X chromosome DNA from only one side of the translocation breakpoint whose position can then be determined by Southern blot hybridisation analysis. The flexibility of this method of mapping is that a panel of characterised hybrids is always available for future experiments such as when new probes become available. PFGE analysis of translocation breakpoints is dependent on knowledge of the order of probes and of the long range map. When these are available, a translocation breakpoint position can determined by detecting a junction fragment produced by a rare cutting restriction enzyme, which is altered in size compared to the fragments detected in normal individuals (see Chapter 5, section 5.1 for a description of translocation junction fragments and their detection). This approach was adopted in the study of 10 DMD-associated X;autosome translocations by Meitinger et al. (1988). The mapping of the same 10 translocations was also performed by somatic cell hybrid analysis using genomic probes from the DMD region (Boyd et al., 1988). The diagrams summarising the results from these two studies are reproduced in figures 4.1 and 4.2. Together with information about the topography of the DMD gene from other groups (for example den Dunnen et al., 1987; Burmeister et al., 1988), the results show that the translocation breakpoints occur throughout the DMD locus and that none lies outside the region of the gene encompassing the coding sequences. The studies also showed that two translocations not associated with DMD have breakpoints lying outside the DMD gene. Some of the DMD-associated translocation breakpoints were not resolved in these studies. For example, it was found that the breakpoints of KIY and EDN or of LAR and DEB may either coincide or be separated by up to around 100 kb (figures 4.1 and 4.2). The accuracy of the breakpoint positions obtained by these

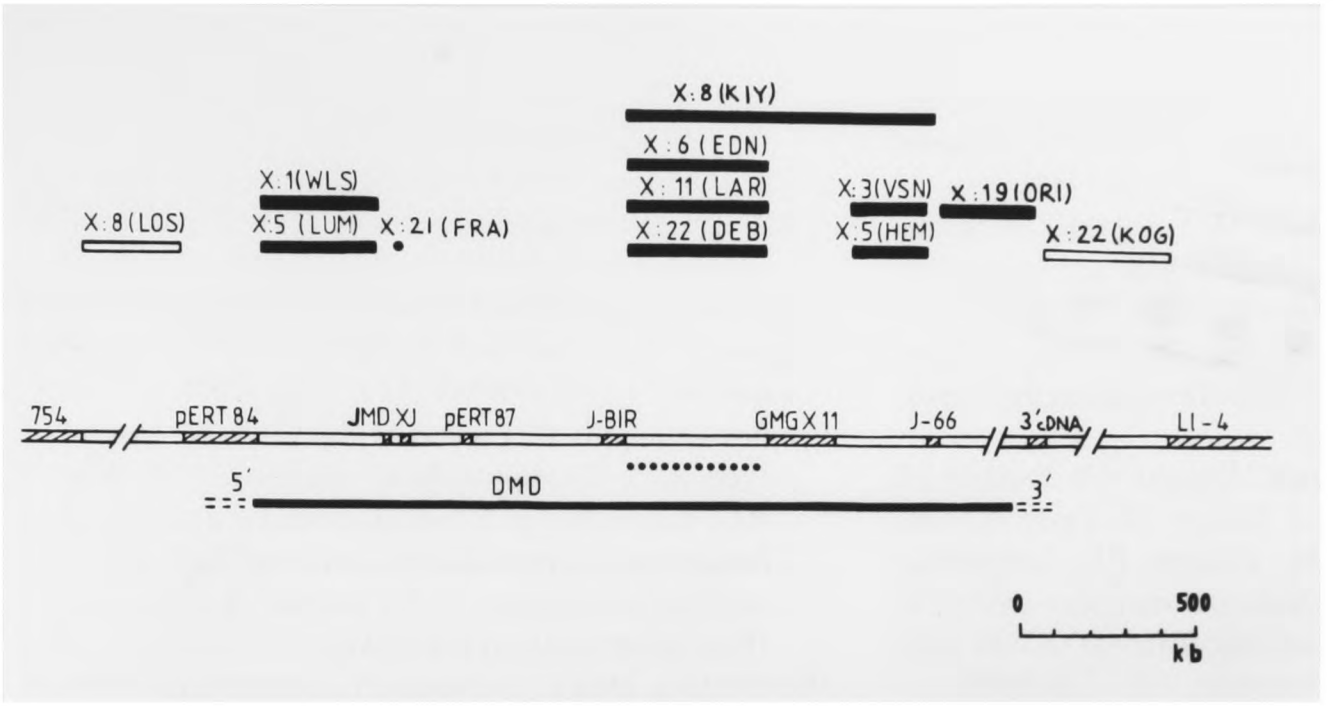


Figure 4.1 Summary diagram of the results of mapping DMD-associated translocation breakpoints within the DMD locus by somatic cell hybrid analysis (diagram reproduced from Boyd et al., 1988). a) Dark lines represent the locations of DMD-associated breakpoints and pale lines represent the locations of breakpoints from translocations not associated with DMD. b) The positions of probes used with respect to the DMD gene are indicated.

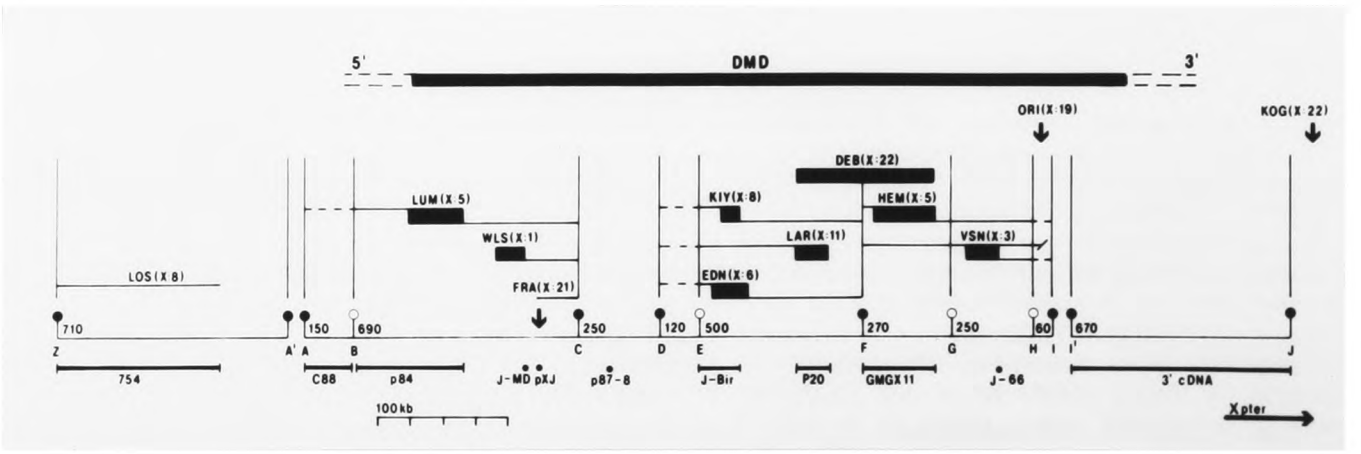


Figure 4.2 Summary diagram from Meitinger et al. (1988) showing the mapping positions of 12 translocations breakpoints (b) with respect to probes (c) from within and around the DMD locus (a) determined by PFGE analysis. Two breakpoints lie outside the gene locus (LOS and KOG) and are from translocations not associated with DMD expression.

methods is typically in the range of a few hundred kb. In the case of somatic cell analysis, the accuracy is determined by the number and positions of probes used in the study. The availability of the dystrophin cDNA has effectively increased the number of probes in this type of analysis since it hybridises to over 65 exons from the gene (Koenig et al., 1987). In this chapter, an improved map of translocation breakpoints within the DMD locus is described, by comparing somatic cell analysis results using cDNA probes with the most accurate physical map available of the exon positions.

Seven translocations are studied here and are listed in table 4.1. The breakpoint positions of these and seven additional breakpoints from the literature are compared with the physical DMD map determined by den Dunnen et al. (1989). The study of den Dunnen et al. included a definition of 115 deletions and 13 duplications from male DMD patients. In this chapter, the endpoints of these deletions and duplications are plotted and compared with the distribution of translocation breakpoints. This is probably the most appropriate method for comparing these types of mutation since it is more likely that sequences in the vicinity of the endpoints of deletions/duplications are involved in producing the mutations than sequences within the deleted/duplicated segment. The data of Koenig et al. (1989) from 258 deletions have not been added to those of den Dunnen et al. (1989) in case the same patients are represented in both sets of data, and because of minor discrepancies in the orders of exon containing restriction fragments.

If there are differences in the distributions of deletions/duplications and translocations, then this may indicate that the mechanisms producing them differ. Further information relevant to the mechanisms may be obtained by determining the parental origins of the mutations. A bias in the parental origin of the mutation may reflect the fundamental differences which exist in male and female gametogenesis (Chandley, 1991). In this chapter, the determination of the parental origin of the X;1 translocation (WLS; see Chapter 1, table 1.1) is described and discussed in relation to other investigations on the parental origins of X;autosome translocations and DMD-

Translocation	breakpoints	Cell line	Hybrid	Content
X;1	p21.2;p34	WLS	WAG 8-1	der(X)
			WAG 21R-11	der(1)
X;6	p21.2;q21	EDN	EDAG 2-9-8	der(X)
			EDAG 3R	der(6)
X;8	p21.2;q24.3	KIY	KIYAG 3-1R1	der(8)
X;11	p21.2;q23.3	LAR	LARAG 3R	der(11)
X;22	p21;q13	DEB	DAG 5-2	der(X)
X;5	p21.2;q31.1	HEM	HEMAG N2	der(X)
X;19	p21.2;q13	ORI	ORIM 7-1	der(X)

Table 4.1 List of the DMD-associated translocation hybrids used in this study. Fuller details on the derivation of the hybrids are provided in Chapter 2, section 2.1, and karyotype analysis of several of the hybrids is described in Chapter 3.

associated deletions.

Appropriate sections of the dystrophin cDNA clones were used in this investigation (kindly provided by L. Kunkel). These were, probe 1b+2a, BamHI (nucleotide 406) to nucleotide 1538; probe 7b, BglII (nucleotide 6509) to EcoRI (nucleotide 7002); probe 8, EcoRI (nucleotide 7002) to EcoRI (nucleotide 7866); and probe 10, BamHI (nucleotide 9079) to BamHI (nucleotide 9786) (see Koenig et al., 1987; Koenig et al., 1988). All hybridisations were performed to HindIII digested DNA, and the order and nomenclature of exon containing HindIII (Hd) fragments used is as described by den Dunnen et al. (1989). The equivalent exon numbers are as described by Koenig et al. (1989).

4.2 Results

4.2.1 Mapping of the t(X;1) between Hd7 and Hd8 (intron 7)

The cDNA probe 1b+2a which detects Hd fragments 5 to 9 (exons 5 to 11) was hybridised to a Southern blot of HindIII digested DNA from the hybrids WAG 8-1 and WAG 21R-11 (figure 4.3). Exons 5 to 7 were detected in WAG 8-1, and exons 8 to 11 in WAG 21R-11, indicating that the breakpoint of the X;1 translocation must lie between exons 7 and 8 (table 4.2). This confirmed the result obtained using the der(X) retaining hybrid alone by Bodrug et al. (1989) and demonstrated that there was no deletion or duplication of DMD exons in the vicinity of the X chromosome breakpoint. Detailed molecular investigation of the X;1 translocation which is described later in this thesis revealed that this breakpoint is 10kb distal to exon 7 (Chapter 5, section 5.2.4). The position in relation to the physical DMD map shows that it lies approximately 50kb proximal to the cloned X;21 translocation breakpoint which is in the same intron (figure 4.4; Ray et al., 1985).

4.2.2 Mapping of three breakpoints between Hd31 and Hd34

The breakpoint positions of three translocations were determined from

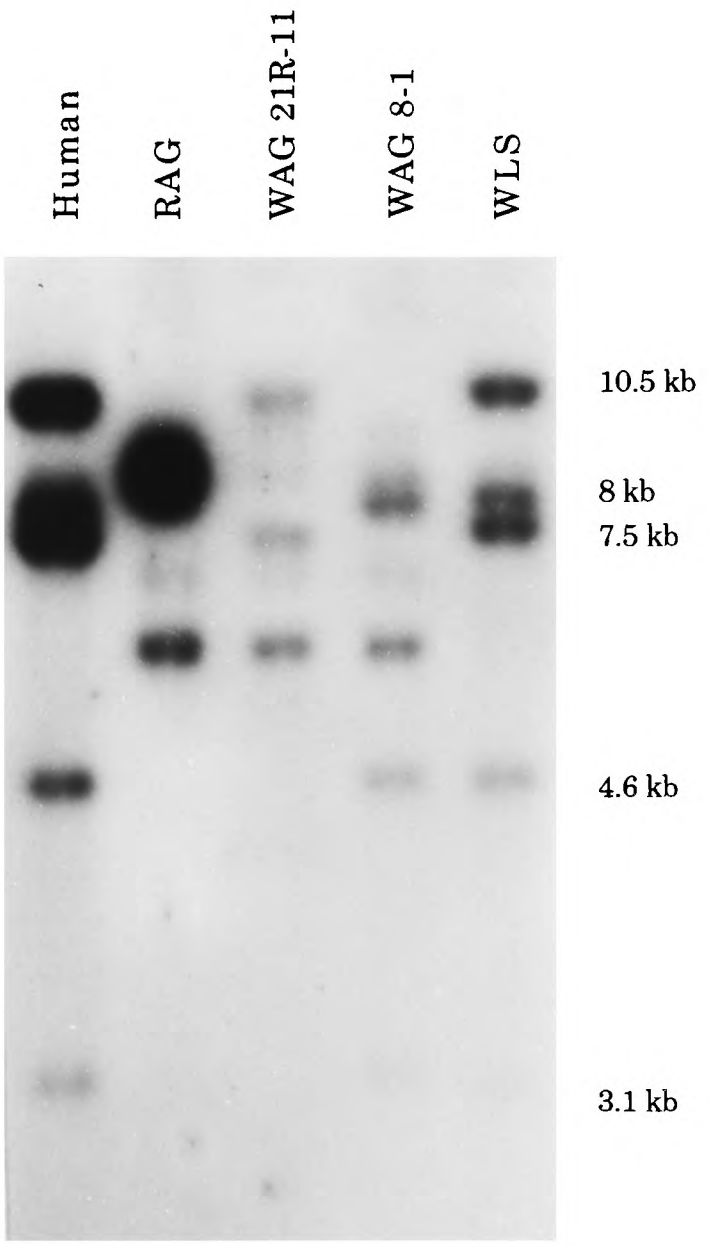


Figure 4.3 Mapping of the X;1 translocation between Hd7 and Hd8. Hybridisation of probe 1b+2a to a Southern blot of HindIII digested DNA revealed that Hd5, Hd6 and Hd7 (3.1, 8 and 4.6 kb) are present in the der(X) retaining hybrid WAG 8-1, and Hd8 and Hd9 (7.5 and 10.5 kb) are in the der(1) retaining hybrid WAG 21R-11.

Probe[a]	1b + 2a					7b					8					
Hd fragment[b]	5	6	7	8	9	31	32	33	34	35	36	37	38	39	40	41
Size[b]	3.1	8.0	4.6	7.5	10.5	11.0	4.1	0.5	1.5	10.0	1.2	3.8	1.6	3.7	3.1	7.0
Exon[c]	5	6	7	8/9	10/11	43	44	45	46	47	48	48	49	50	51	52
ORIM 7-1								•		+				+	+	+
HEMAG N2														+	+	-
DAG 5-2						+	+		+	+		-		-	-	-
LARAG 3R						-	-	-	+	+		+		+	+	+
KIYAG 3-1R1						-	+	+	+	+				+	+	+
EDAG 2-9-8						+	-	-	-	-	-	-		-	-	-
EDAG 3R						-	+	+	+	+	+	+		+	+	+
WAG 8-1	+	+	+	-	-											
WAG 21R-11	-	-	-	+	+											

Probe[a]	9						10						
Hd fragment[b]	42	43	44	45	46	47	48	49	50	51	52	53	54
Size[b]	7.8	1.0	8.3	2.3	1.0	8.8	6.0	3.5	2.4	10.5	6.6	2.8	2.5
Exon[c]	53	53	54	55	56	57	58/9	60					
ORIM 7-1	+		+	+		+	+	+	-		-	-	-
HEMAG N2	-		-	-		-	-	-	-		-	-	-
DAG 5-2													
LARAG 3R													
KIYAG 3-1R1													
EDAG 2-9-8													
EDAG 3R													
WAG 8-1													
WAG 21R-11													

Table 4.2 Summary of Southern blot hybridisation results. Presence (+) or absence (-) of hybridisation of cDNA probes to Hd fragments is indicated for nine somatic cell hybrids. [a]Koenig et al. (1987); [b]den Dunnen et al. (1989); [c]Koenig et al. (1989).

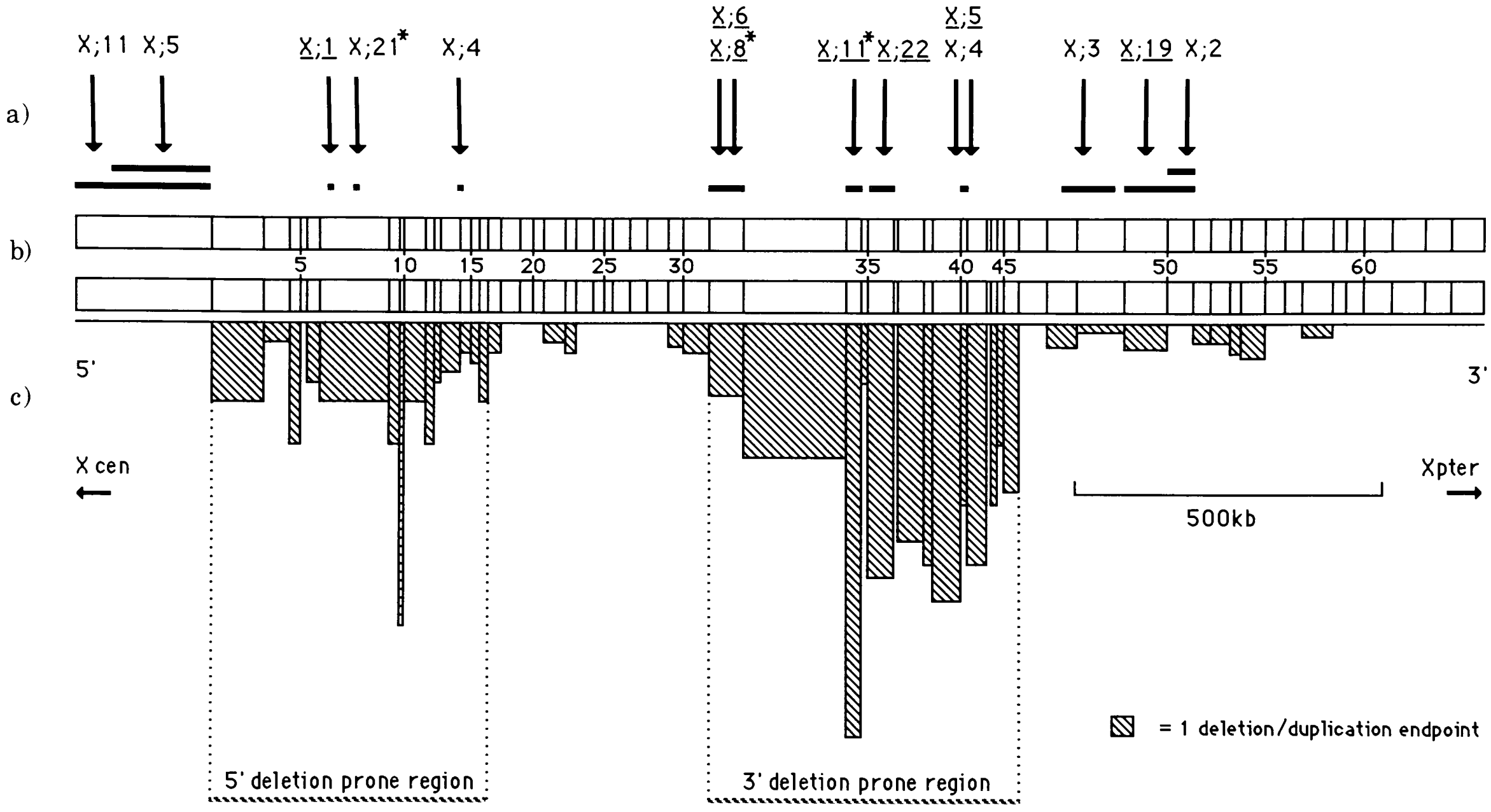


Figure 4.4 Physical map of the dystrophin locus. a - X chromosome breakpoint positions of 14 DMD-associated X;autosome translocations. Those determined here are underlined and those associated with mild phenotypes are marked with an asterisk. b - Exon containing Hd fragment map (from den Dunnen et al. 1989). c - Distribution of deletion/duplication endpoints from 117 male patients. The area of each bar represents the number of deletion/duplication endpoints mapped between the Hd fragments (data from den Dunnen et al. 1989).

hybridisation of the cDNA probe 7b to HindIII digested DNA of the relevant somatic cell hybrids (table 4.1). This probe detects Hd fragments 31 to 34, and encompasses the 160-180 kb intron containing the P20 locus (Blonden et al., 1991). The breakpoints of EDN (X;6) and KIY (X;8) were found to map to the intron between Hd31 and Hd32 (table 4.2). This intron contains the genomic locus J-Bir which is already known to lie on the proximal side of these breakpoints (Boyd et al., 1988; Meitinger et al., 1988). The interval between J-Bir and Hd32 (exon 44) is estimated to be around 100kb (from den Dunnen et al., 1989). Figure 4.5 illustrates that Hd34 is present in DNA of LARAG 3R while Hd32 and Hd33 are absent. This indicated that the translocation breakpoint of LAR (X;11) is in the intron between Hd33 and Hd34 (exons 45 and 46; table 4.2). This intron is thought to be only 10-20 kb in size, and taking the size of the intron into consideration, it appears to contain more deletion/duplication endpoints than any other intron (see figure 4.4; den Dunnen et al., 1989). None of these three translocation breakpoints lies in the large intron which contains the P20 locus (Hd32-Hd33).

4.2.3 Mapping of two breakpoints between Hd35 and Hd41

Southern analysis using the cDNA probe 8 which detects Hd35 to Hd41 determined the breakpoint positions of two translocations in this region of the DMD gene. Figure 4.6 illustrates that Hd35 is present in DNA from the hybrid DAG 5-2 and Hd37 to Hd41 are absent. Hd36 could not be scored on this Southern blot, however, Hd36 and Hd37 both represent exon 48 since this exon contains a HindIII site (den Dunnen et al., 1989). Therefore this result indicates that the translocation breakpoint of DEB (X;22) maps between Hd35 and Hd37 (exons 47 and 48; table 4.2; figure 4.4). From the physical map of the dystrophin gene reported by den Dunnen et al. (1989), it can be estimated that the X;22 breakpoint is probably between 10kb and 100kb distal to the breakpoint of LAR (X;11; figure 4.4). Analysis revealed that the breakpoint of HEM (X;5) maps into the small (approximately 20kb) intron between Hd40 and Hd41 (exons 51 and 52; table 4.2; figure 4.4).

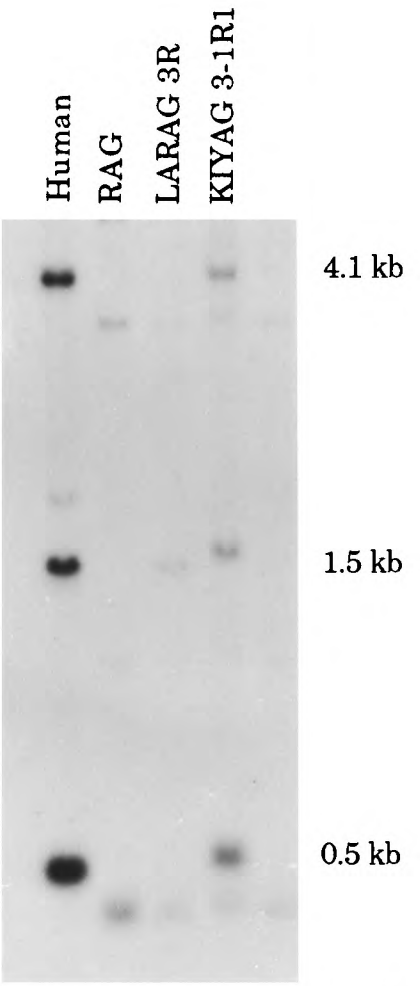


Figure 4.5 Mapping of the X;11 (LAR) translocation breakpoint between Hd33 and Hd34. Hybridisation of probe 7b to a Southern blot of HindIII digested DNA revealed the absence of Hd32 and Hd33 (4.1 and 0.5 kb) and the presence of Hd34 (1.5 kb) in the der(11) retaining hybrid LARAG 3R.

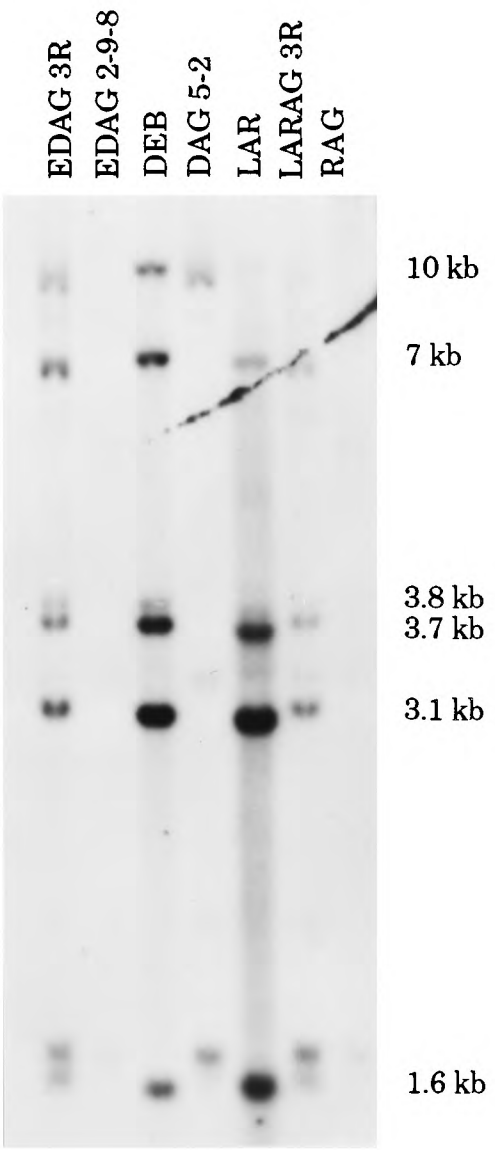


Figure 4.6 Mapping of the X;22 (DEB) translocation breakpoint between Hd35 and Hd37. Hybridisation of probe 8 to a Southern blot of HindIII digested DNA revealed the presence only of Hd35 (10 kb) in the der(X) retaining hybrid DAG 5-2. The absence of the 10 kb band in LAR is due to high molecular weight DNA degradation.

4.2.4 Mapping of the t(X;19) in the distal region of the gene

Hybridisation of cDNA probe 10 to HindIII digested DNA of the somatic cell hybrid ORIM 7-1 detected fragments Hd48 and Hd49 and not Hd50 and Hd52 to Hd 54 (figure 4.7; table 4.2). The fragment Hd51 could not be scored due to co-migration of a hamster cross-hybridising band. This result indicates that the breakpoint of the X;19 translocation lies distal to Hd49, however, since the order of Hd50 and Hd51 is unknown, its position relative to Hd51 is uncertain (see figure 4.8).

4.2.5 Determination of the parental origin of the X;1 translocation

The probe 782 (DXS85) which maps to Xp22.3-p22.2 detects an RFLP in Southern blots of EcoRI digested DNA. The t(X;1) patient is heterozygous for this polymorphism, her mother homozygous for the 14 kb allele, and her father hemizygous for the 7kb allele (Boyd, personal communication). Probing of a Southern blot of der(1) retaining somatic cell hybrid DNA with DXS85 detected the 7kb allele (figure 4.9), indicating the paternal origin of the translocation. This result was confirmed independently through the use of the hypervariable probe M27β which can detect methylation differences at the MspI/HpaII sites adjacent to this locus (Robinson et al., 1990). The recognition sites of these enzymes are the same but HpaII will only cut when it is unmethylated, and the methylation status at this site has been demonstrated to be associated with the inactivation status of the X chromosome (Boyd and Fraser, 1990). As the intact X chromosome of X;autosome translocations is invariably inactivated in preference for the der(X) (Mattei et al., 1981), analysis of the methylation pattern of the patient and her parents demonstrated that the chromosome carrying the methylated MspI site resistant to HpaII digestion, hence the active der(X), was paternally inherited.

Strictly, the above data show that the X chromosome involved in the X;1 translocation is paternally inherited but do not indicate the parental origin of the chromosome 1. If the translocation event occured prior to zygote formation, then both

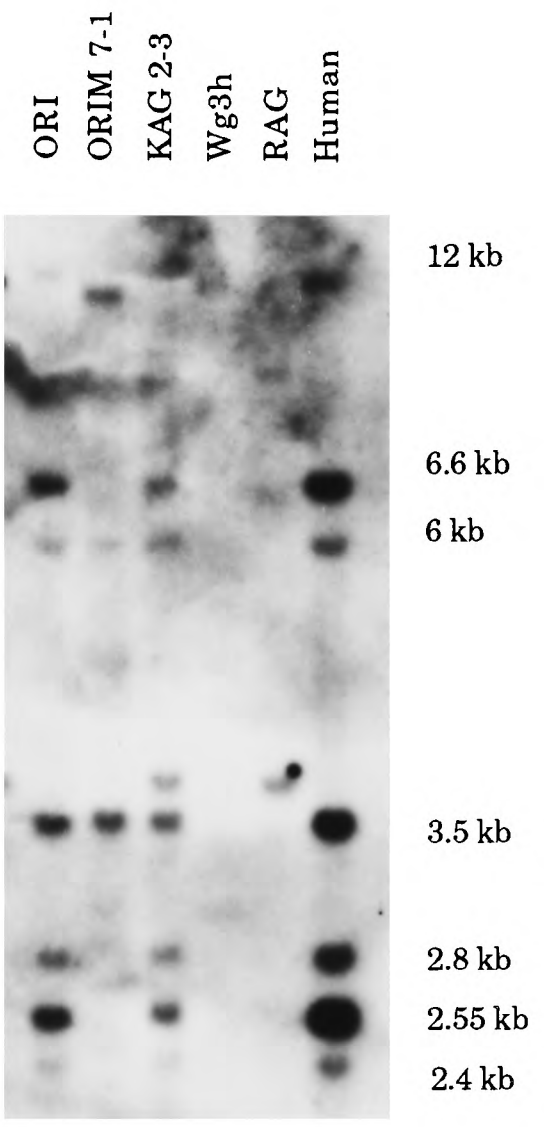


Figure 4.7 Mapping of the X;19 (ORI) translocation between Hd49 and Hd50/Hd51. Hybridisation of probe 10 to a Southern blot of HindIII digested DNA revealed that Hd48 and Hd49 (6 and 3.5 kb) are present in the der(X) retaining hybrid ORIM 7-1. Hd50 (12 kb) could not be scored due to co-migration of a cross-hybridising hamster band.

	48	49	50	51	52	53	54
Order 1	+	+	-	n	-	-	-
		t(X;19)					
Order 2	48	49	51	50	52	53	54
	+	+	n	-	-	-	-
		t(X;19)					

Figure 4.8 Alternative mapping positions of the X;19 (ORI) translocation depending upon the order of Hd49 and Hd50. Presence (+) or absence (-) of hybridisation of probe 10 to Hd fragments in DNA from the der(X) retaining hybrid ORIM 7-1 is indicated (n - not determined).

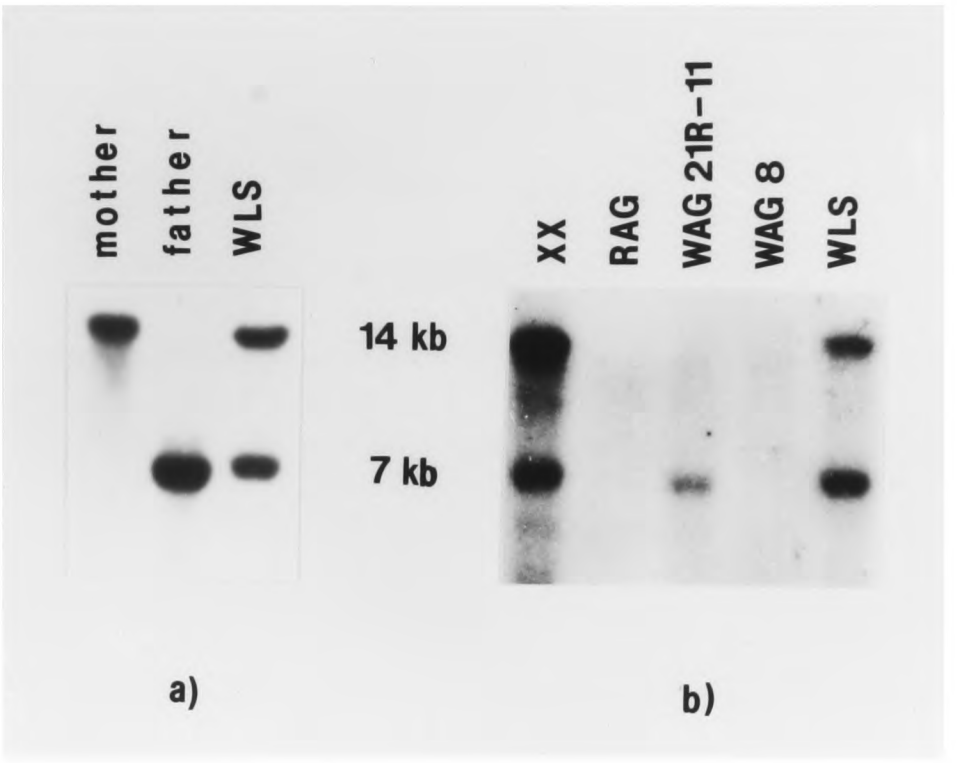


Figure 4.9 Demonstration of the parental origin of the X;1 translocation. The 7 kb allele which is detected by the probe 782 was detected in DNA of the der(1) retaining hybrid WAG 21R-11 and of the patient's father but the patient's mother was found to be homozygous for the 14 kb allele.

chromosomes involved would necessarily derive from the same parent. If, however, the rearrangement occurred after zygote formation, then it is possible that the chromosomes involved derived from the different parents (see Chapter 8, section 8.1 for discussion). The availability of the somatic cell hybrid WAG 8-1 which retains the der(X) but does not contain a normal chromosome 1 (see Chapter 3, section 3.2.2) would allow the parental origin of the chromosome 1 involved in the translocation to be determined assuming that RFLPs which map between 1pter and 1p34 are informative. The der(1) retaining hybrid WAG 21R-11 would not be useful in this study as it also contains the normal chromosome 1 from the t(X;1) patient (see Chapter 2, table 2.1).

4.3 Discussion

The DMD locus offers an unprecedented opportunity to characterise the distribution of mutations over an extensive genomic segment. The significance that can be drawn from the distribution of translocation breakpoints is somewhat restricted by the small number of cases so far studied. Therefore analysis of further cases remains critical in the context of a full understanding of the nature of this rare class of mutation in the DMD locus. Data are presented here on the refined mapping of seven translocation breakpoints with respect to exon containing HindIII fragments from the dystrophin locus. The positions of these breakpoints on the physical map of the DMD locus reproduced from den Dunnen et al. (1989) are illustrated in figure 4.4. In this figure, the distribution of the endpoints of 115 deletions and 13 duplications is also shown (data from den Dunnen et al., 1989). There are two regions where deletion/duplication endpoints are predominantly located, one in the 5' region of the gene between Hd fragments 2 and 17 (containing the XJ and pERT87 loci) and the other between Hd fragments 31 and 46 (around the P20 locus). These deletion prone regions have also been described by others (Koenig et al., 1987; den Dunnen et al., 1987; Forrest et al., 1987). Deletion/duplication breakpoints appear to be fairly evenly distributed within these regions and it is reported that there is no evidence of a deletion hot spot within the

160-180kb P20 intron (Hd32-Hd33; Blonden et al., 1991). Although this intron is often referred to as the deletion prone intron, when its size is taken into consideration it can be seen that deletion endpoints occur less frequently here than within the adjacent genomic region on the distal side from Hd33 to Hd46 (see figure 4.4).

The presence of deletion prone regions could partly reflect ascertainment bias. Many DMD deletions were originally detected with genomic probes located in these regions and it is possible that such patients are over-represented in the cases selected for study with cDNA probes. The distribution of mutations from approximately 35% of male DMD patients is unknown, as presently their mutations have not been detected. These mutations may include point mutations and small deletions. Their distribution in the dystrophin gene could differ from that of the majority of male DMD patients with deletions. Only one point mutation has so far been described (Bulman et al., 1991). It is a G to T transversion at nucleotide 3677 within exon 26 (Hd20) producing an Amber stop codon. Additionally, it has been proposed that a small deletion/duplication in the portion of the gene coding for the triple-helical domain of dystrophin and which maintains the reading frame may produce an insignificant or very mild phenotype (Koenig et al., 1990). Such mutations may be predicted to occur in the interval between the two deletion prone regions.

By contrast, the distribution of translocation breakpoints would not be expected to reflect these biases. One possible ascertainment bias of translocations might be in the 3' end of the gene, where translocations in the untranslated portion of the gene or causing small truncations of the C-terminus of dystrophin might produce milder phenotypes.

Data are available from the literature concerning the locations of seven translocation breakpoints additional to those reported here and are included in figure 4.4. Two translocation breakpoints, an X;11 and an X;5 lie within the first DMD intron (Bodrug et al., 1989). The PFGE data of Meitinger et al. (1988) place the X;5 translocation in the distal portion of the intron. The location of the cloned X;21

translocation breakpoint within intron 7 is precisely known (Ray et al., 1985). Giacolone and Francke (1987) have identified an X;4 translocation whose X chromosome breakpoint is 10 kb proximal to the genomic clone pERT87-15 and therefore just proximal to Hd14. Bodrug et al. (1990) have mapped a different X;4 translocation to the intron between Hd40 and Hd41, the same small intron to which the X;5 translocation (HEM) has been mapped in this study. The best information regarding the location of the X;3 translocation (VSN) is from PFGE studies (van Ommen et al., 1987; Meitinger et al., 1988). Thus DMD-associated translocation breakpoints are distributed throughout the dystrophin locus from the first intron (X;5 and X;11) up to the X;19 and X;2 breakpoints in the vicinity of Hd50 (figure 4.4).

Three of the 14 translocations whose mapping positions are illustrated in figure 4.4 are associated with a distinctly milder phenotype than classical DMD. They are the X;21, X;8 and X;11 (LAR) translocations (Verellen et al., 1984; Narazaki et al., 1985; Nielsen et al., 1983). There appears to be no correlation between the breakpoint position within the DMD locus and the severity of the phenotype (see figure 4.4). Several translocations positioned 3' of these produce severe phenotypes and therefore severity cannot be correlated with the amount of the locus potentially transcribed. In keeping with this observation, the X;6 and X;8 translocation breakpoints which lie in the same intron (Hd31-Hd32) are associated with contrasting phenotypes (Zatz et al., 1981; Narazaki et al., 1985). Therefore a more likely explanation for the differing phenotypes of these patients is that the milder phenotypes are due to transcription of dystrophin loci from a small number of active intact X chromosomes in these patients (Boyd et al., 1986). The possible effect that the autosomal breakpoint position could have on the inactivation bias and hence on the severity of DMD expression is discussed in Chapter 1, section 1.5.

There are noteworthy similarities between the distribution of translocation breakpoints with that of deletion/duplication endpoints and noteworthy differences. Firstly, nine of the fourteen translocation breakpoints are within the deletion prone

regions while these regions represent less than half of the DMD locus, around 1000/2300kb, although this association is not statistically significant ($P < 0.01$). Additionally the observation that six are within the 3' deletion prone region and three within the 5' region mirrors the distribution of deletions/duplications. Secondly, it is apparent that while no translocation breakpoint lies in the interval between the deletion prone regions, five translocation breakpoints do lie outside the deletion prone regions. Two are within the first intron which is around 200kb in length. Interestingly, no deletion or duplication endpoint has been localised to this intron despite its size and the number of cases studied (Koenig et al., 1989, studied 273 deletions). Three translocation breakpoints lie distal to the 5' deletion prone region in a part of the gene where few deletion endpoints have been observed.

At present it is unclear whether ascertainment bias or a difference in the aetiology of mutation is the main reason for the different distributions. A clue however is suggested from the study of the parental origins of the mutations. Eight DMD-associated translocations have been studied in this respect including the t(X;1) described here (Robinson et al., 1990; Bodrug et al., 1990) and all are paternal in origin (table 4.3). Five additional X;autosome translocations are paternal in origin (Robinson et al.,1990; Verga et al., 1991). The probability that this is a chance deviation from a 50:50 ratio is less than 1 in 4000. All of these X;autosome translocations exist in females. Males with X;autosome translocations do exist (Schinzel, 1983). Mendelian inheritance would predict that they inherited the translocation from their mothers although this has not been confirmed by linkage analysis. If X;autosome translocations are rarely produced in oogenesis then males with these rearrangements would be rare also. In the records of the Department of Medical Genetics in Oxford, there are nine females and two male that have been identified with an X;autosome translocation. Worldwide, there are 24 females with DMD and an X;autosome translocation but no male has been identified (Boyd, 1991). This might further indicate that males with X;autosome translocations are rare, since if the breakpoint disrupted

Translocation	Breakpoints	Parental origin	Reference
X;1 (WLS)	p21.2;p34	Paternal	Lindenbaum et al., 1979
X;1	p21;q23		personal communication
X;2 (TM)	p21.2;q37.3	Paternal	Holden et al., 1986
X;2	p21;q14		personal communication
X;3 (VSN)	p21.2;q13.3		Canki et al., 1979
X;3	p21.2;q25.3		personal communication
X;3	p21;q27		personal communication
X;4	p21;q35	Paternal	Bodrug et al., 1990
X;4	p21.1;q26		personal communication
X;4	p21;q31.2		Giacalone and Franke, 1988
X;5 (HEM)	p21.2;q31.1		Nevin et al., 1986
X;5 (LUM)	p21.1;q35.3	Paternal	Jacobs et al., 1981
X;6 (EDN)	p21.2;q21		Zatz et al., 1981
X;6	p21;q16		Perez-Vidal et al., 1983
X;8 (KIY)	p21.2;q24.3		Narazaki et al., 1985
X;9 (GM6007)	p21;p22		Emanuel et al., 1983
X;9 (LC)	p21.2;q21.3	Paternal	Robinson et al., 1990
X;9	p21;p21	Paternal	Nielsen and Nielsen, 1984
X;11 (LAR)	p21.2;q23.3		Nielsen et al., 1983
X;11	p21;q13		Greenstein et al., 1977
X;15	p21;q26	Paternal	Ribeiro et al., 1986
X;19 (ORI)	p21.2;q12 or q13.1		Boyd et al., 1988
X;21 (FRA)	p21.1;p12	Paternal	Verellen-Dumoulin et al., 1984
X;22 (DEB)	p21;q13		Boyd et al., 1988

Table 4.3 Table of the 24 known cases of X;autosome translocation (Boyd, 1991) indicating the parental origins of the translocations where this has been studied.

the DMD locus then a male would be expected to express the disease. It is possible, however, that such males have not been identified due to ascertainment bias. There is a case of Menke's syndrome in a male patient who has an X chromosome rearrangement. The rearrangement is an insertion of a block from the X chromosome long arm (q13-q21.2) into the short arm at Xp11.4 (Horn and Tonneson, personal communication). By contrast, evidence from the study of deletion/duplication mutations, does not suggest that there is a biased parental origin. Bakker et al. (1989) studied the origin of 41 new mutations (deletions and duplications) in boys and found that in 27 cases the mutation was not present in the mother, indicating that the mutations had occurred in her germline. In nine of the remaining cases the mother carried the mutation on the grandmaternal chromosome, and in 5 cases she carried the mutation on the grandpaternal chromosome. Therefore, by comparison of the numbers of mutations originating in grandparents, Bakker et al. concluded from the small set of data that 'the mutation rates for DMD in grandparental male and female X chromosomes are equal for each chromosome'. It would appear from these data that differences in the parental origins of deletions/duplications and of translocations exist which may reflect differences in the mechanisms responsible for producing them or differences in the processes of male and female gametogenesis.

The clearest indications as to what mechanism might cause a translocation must be obtained by examining the DNA sequence in the vicinity of the breakpoints. For this reason, the work described in the following three chapters involves the detailed analysis and sequencing of a single translocation.

Chapter 5 Characterisation of the X;1 translocation breakpoints

5.1 Introduction

The following three chapters (5, 6 and 7) concern the detailed investigation of a single translocation. Each chapter describes a similar type of analysis performed during the course of the investigation, therefore the work is not described in the same order as it was performed. The main part of the work described in this chapter is the characterisation of the X;1 translocation by restriction mapping and Southern blot analysis. The derivation of many of the cloned DNA segments and probes characterised in this chapter is described in detail in Chapter 6. Therefore only brief descriptions with cross references are included in the text here.

The principal objective of cloning and sequencing the breakpoints of the t(X;1) was to elucidate the mechanisms which were involved in producing the translocation. The scope of this thesis permitted only one translocation to be studied at this resolution although it is likely that the study of several will be necessary in order to determine the aetiology of this class of chromosome rearrangement.

The t(X;1) patient lived in the Oxford area and the diagnosis of muscular dystrophy at 5 years of age, followed by the identification and characterisation of the chromosomal rearrangement, were all made locally (Lindenbaum et al., 1979). She manifested classical DMD symptoms and died just before her 16th birthday (Lindenbaum, personal communication). This case was among the first described of 24 females now known with DMD and an X;autosome translocation involving Xp21 (Boyd, 1991). A detailed cytogenetic characterisation of the X;1 translocation is not included in this thesis since it has been performed previously by others (Lindenbaum et al., 1979; Boyd and Buckle, 1986; Clarke, unpublished observations). Figure 5.1 illustrates reproduced partial G-banded karyotypes prepared from peripheral blood lymphocytes of the patient. The translocation breakpoints were identified at Xp21 and 1p34. An abnormal cytogenetic observation accompanies the identification of the

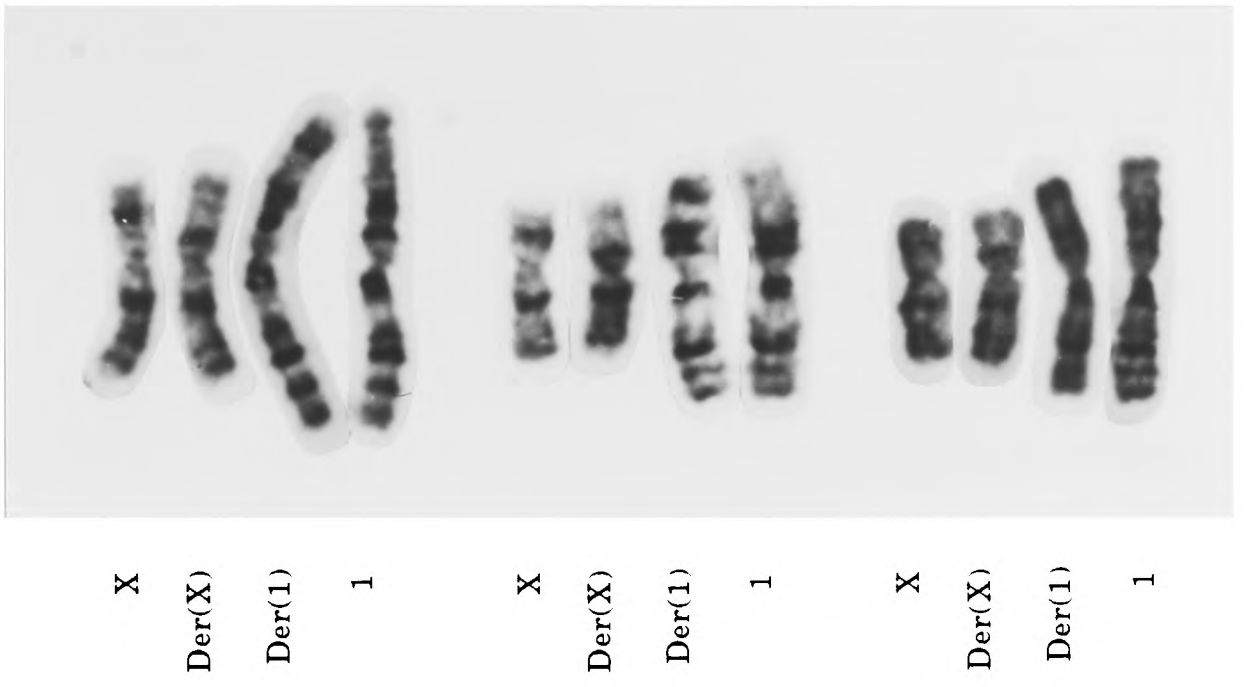


Figure 5.1 Partial G-banded karyotypes of the X;1 translocation patient (WLS). Photographs were kindly supplied by G Clarke.

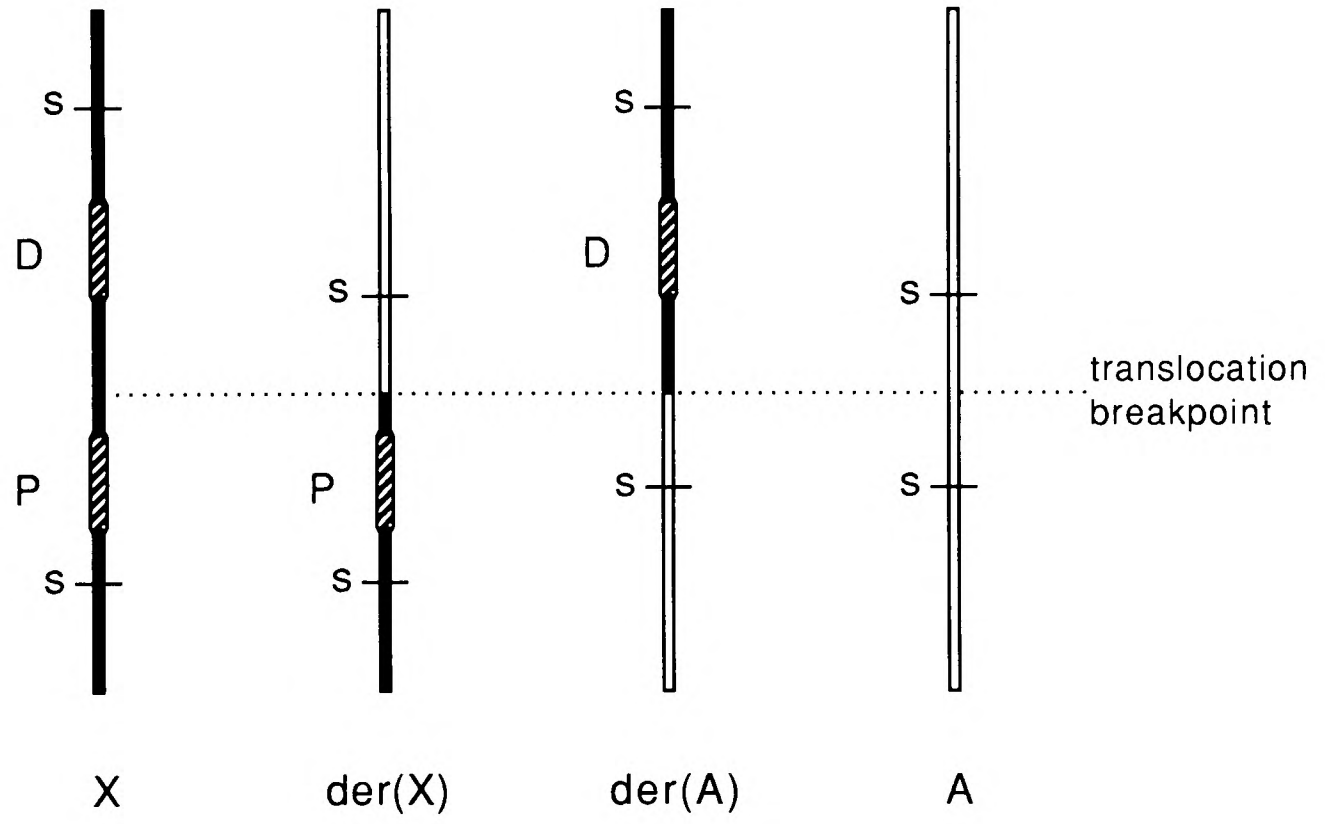


Figure 5.2 The detection of the junctions fragments of an X;autosome translocation. Dark lines represent X chromosome DNA and pale lines represent autosome (A) DNA. The sites of hybridisation of X chromosome probes proximal (P) and distal (D) to the breakpoint position; and the restriction fragments generated by digestion of the DNA are indicated.

translocation. This is that the remnant Xp21 band on the der(X) is closer than expected to the centromere as judged by comparing the banding patterns of the X and der(X) chromosomes. The simplest interpretation to account for the banding pattern, a three-break rearrangement involving a small paracentric inversion on Xp associated with the translocation, was the structure of the rearrangement suggested by the original investigators (see Chapter 1, section 1.5, figure 1.3). By cloning the breakpoints of this translocation, I have been able to test this interpretation. The results presented in this chapter indicate that a three-break rearrangement is not responsible. Alternative structures for the rearrangement are discussed in section 5.3.

A practical reason for cloning the X;1 translocation in preference to one of the other translocations which are described in Chapter 4 was that the X breakpoint was localised to a relatively narrow 20 kb interval (described in section 5.2.1). Furthermore, a cosmid spanning the region was made available for the purpose.

The first stage in the strategy I adopted to clone the breakpoints of the X;1 translocation was to isolate an X chromosome probe close enough to the breakpoint to detect junction fragments. A junction fragment is a restriction fragment which spans the der(X) or der(1) translocation junction. It can be detected by Southern analysis, and is altered in size compared to the restriction fragment detected by the same probe in a normal individual. For example, figure 5.2 illustrates a hypothetical restriction map around the breakpoints of an X;autosome translocation. Probe D is an X chromosome probe which lies distal to the breakpoint and detects a 5 kb fragment in a normal individual. When hybridised to DNA from the translocation patient, probe D will detect two fragments. One is the 5 kb fragment from the normal X chromosome while the other is a fragment from the der(autosome) chromosome which spans the translocation junction and is 4 kb in size. Probe P is proximal to the translocation but detects the same 5 kb band in a normal individual as probe D. When hybridised to DNA from the translocation patient, probe P detects the 5 kb fragment from the normal X chromosome and a 3 kb band which is the der(X) junction fragment. If the

restriction map of the X chromosome region around the breakpoint is known, then the sizes of junction fragments can be used to predict the positions of the restriction sites which flank the autosomal breakpoint. Information derived in this way was used to design experimental conditions for the amplification and cloning of the t(X;1) junction fragments by inverse PCR (described in Chapter 6). Figure 5.3 summarises the procedure used although the sizes and positions of probes and restriction fragments are invented. Firstly, an X chromosome probe (1) was isolated distal to the breakpoint and which detected the der(1) junction fragment (2). This probe was sequenced and the der(1) junction fragment was amplified and cloned by inverse PCR. A subclone of this fragment proximal to the junction (3) was used to probe Southern blots and detected the same der(1) junction fragment but also detected a normal chromosome 1 restriction fragment (4). This probe was sequenced and the restriction fragment spanning the chromosome 1 breakpoint was amplified and cloned by inverse PCR. A probe distal to the chromosome 1 breakpoint was isolated (5) and found by Southern analysis to detect the der(X) junction fragment (6). Finally, sequence proximal to the X chromosome breakpoint (7), was used to amplify the der(X) junction fragment by inverse PCR (8).

The cloning and sequencing of the t(X;1) is described in three chapters. This chapter describes all of the restriction mapping, Southern blot characterisation and related studies necessary to provide the basis for the subsequent cloning of the breakpoints. Chapter 6 describes the inverse PCR approach adopted and the cloning of amplified DNA. Chapter 7 describes the detailed analysis of the sequence rearrangement involved in the translocation and of flanking DNA.

Two additional experiments are described in this chapter. The first is the localisation by in situ hybridisation of a probe from the autosomal breakpoint region of the t(X;1) to the expected site on the chromosome 1 short arm (1p34). The second is the characterisation of an RFLP detected by a probe adjacent to the X chromosome breakpoint of the X;1 translocation and which is useful in linkage analysis of families with Duchenne muscular dystrophy.

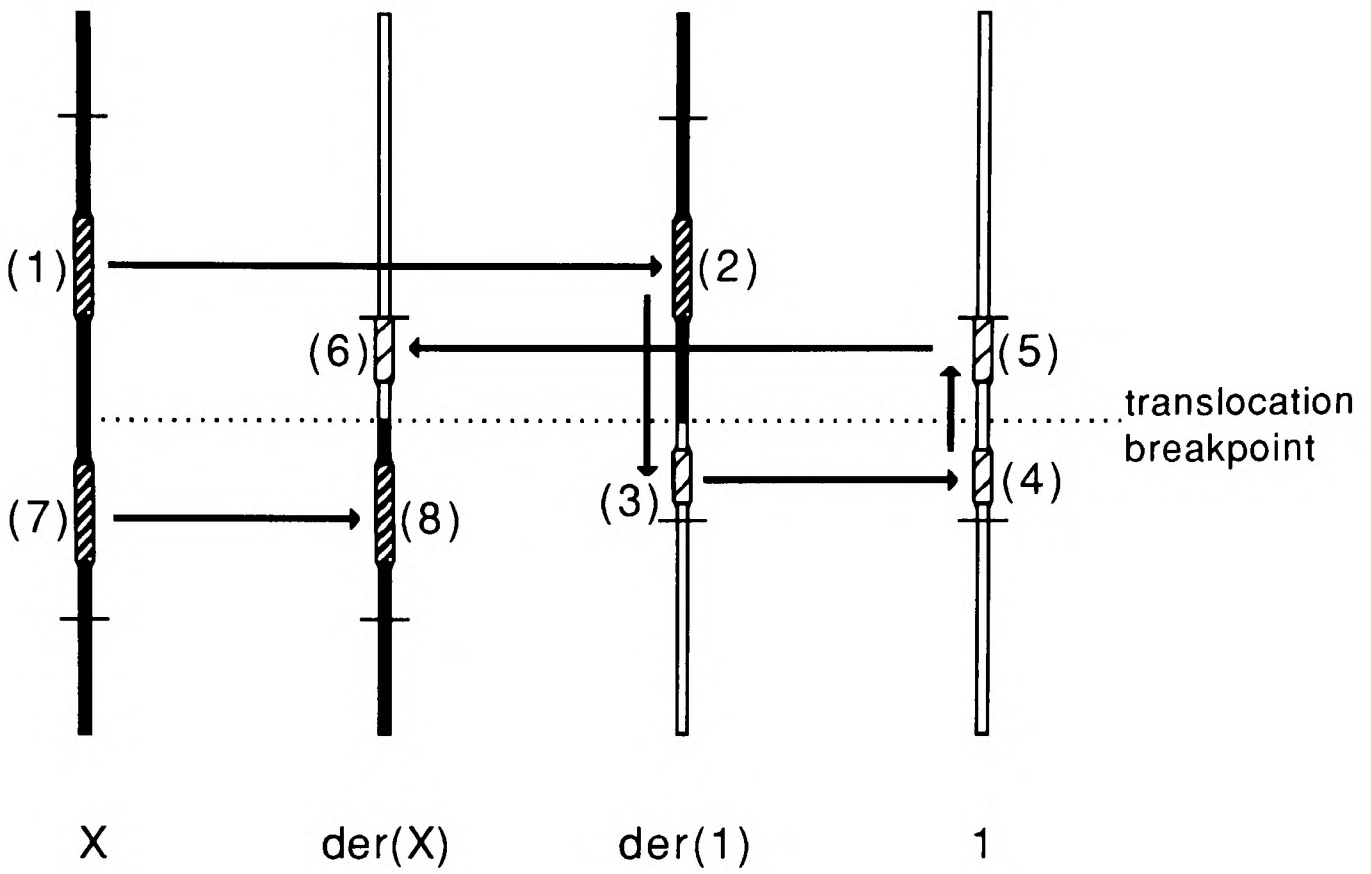


Figure 5.3 Summary of the proceedure used to clone the X;1 translocation. X chromosome DNA is represented by dark lines and chromosome 1 DNA is represented by pale lines. Details are given in the text (section 5.1). Restriction sites are invented for the purpose of the diagram.

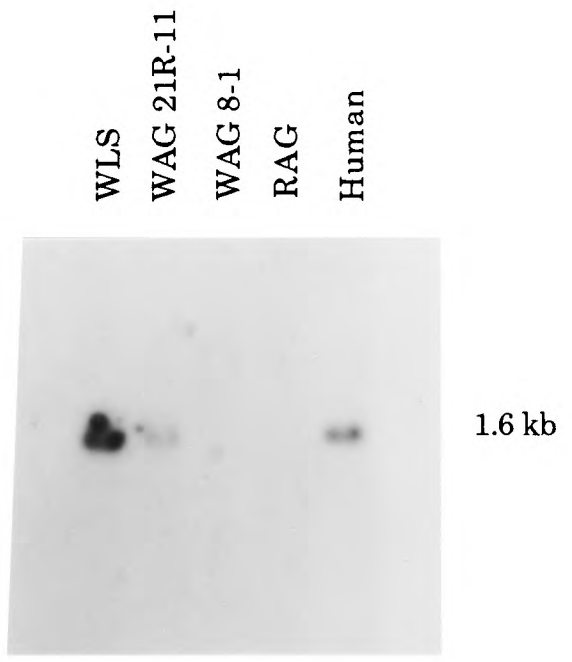


Figure 5.4 Mapping of the X;1 translocation breakpoint proximal to the genomic probe J-MD1. Probing of a Southern blot of HindIII digested DNA revealed a hybridisation signal (1.6 kb) in the der(1) retaining hybrid WAG 21R-11 and no signal in the der(X) retaining hybrid WAG 8-1.

5.2 Results

5.2.1 Hybrid mapping

The somatic cell hybrids WAG 8-1 and WAG 21R-11 (see Chapter 2, table 2.1) retaining the reciprocal der(X) and der(1) chromosomes respectively from the X;1 translocation were used to identify the X chromosome breakpoint position (see Chapter 4, section 4.1 on mapping by somatic cell hybrid analysis). Intragenic genomic and cDNA probes of known order from the DMD locus were hybridised to Southern blots of DNA from both hybrids. J-MD1 and J-MD2 are single-copy genomic DNA probes isolated from around the endpoint of a deletion from a male DMD patient (Monaco et al., 1987). Figure 5.4 illustrates that the probe J-MD1 hybridises to DNA from WAG 21R-11 but not to DNA from WAG 8-1. Therefore the X chromosome breakpoint is proximal to J-MD1. Figure 5.5 illustrates, however, that J-MD2 hybridises to DNA from WAG 8-1 and therefore the X chromosome breakpoint must lie distal to this probe. The hybridisation of a cDNA fragment representing exons 5 to 11 to DNA from both hybrids was illustrated in Chapter 4 (figure 4.3). Three exon-containing restriction fragments were detected in the hybrid WAG 8-1 and two in WAG 21R-11. This indicated that the cDNA fragment spanned the t(X;1) X chromosome breakpoint. Fragments representing exons 5 to 7 are present in WAG 8-1 and fragments representing exons 8 to 11 in WAG 21R-11. The results obtained from these experiments are summarised in figure 5.6. They predict that the X chromosome breakpoint is between DMD exon 7 and the genomic clone J-MD1. These results confirm and extend those of Bodrug et al. (1989) who used the der(X) retaining hybrid only, and Meitinger et al. (1988) who detected the der(X) and der(1) SfiI junction fragments by PFGE analysis (see Chapter 4, section 4.1 on translocation mapping by PFGE).

A cosmid, XJC-5, was obtained which spanned the genomic interval between exon 7 and J-MD1. This cosmid is a walk from the t(X;21) breakpoint which lies within the

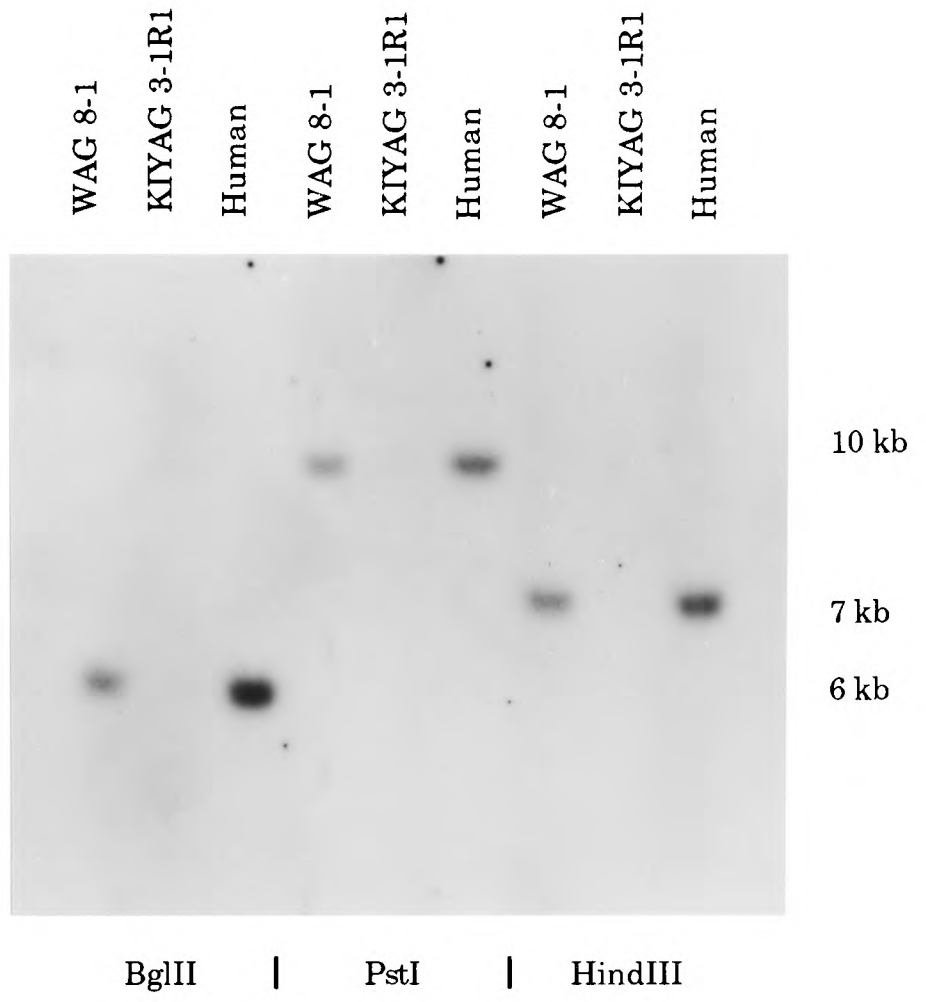


Figure 5.5 Mapping of the X;1 translocation breakpoint distal to the genomic probe J-MD2. Probing of a Southern blot of BglII, PstI and HindIII digested DNA revealed a hybridisation signals in the der(X) retaining hybrid WAG 8-1.

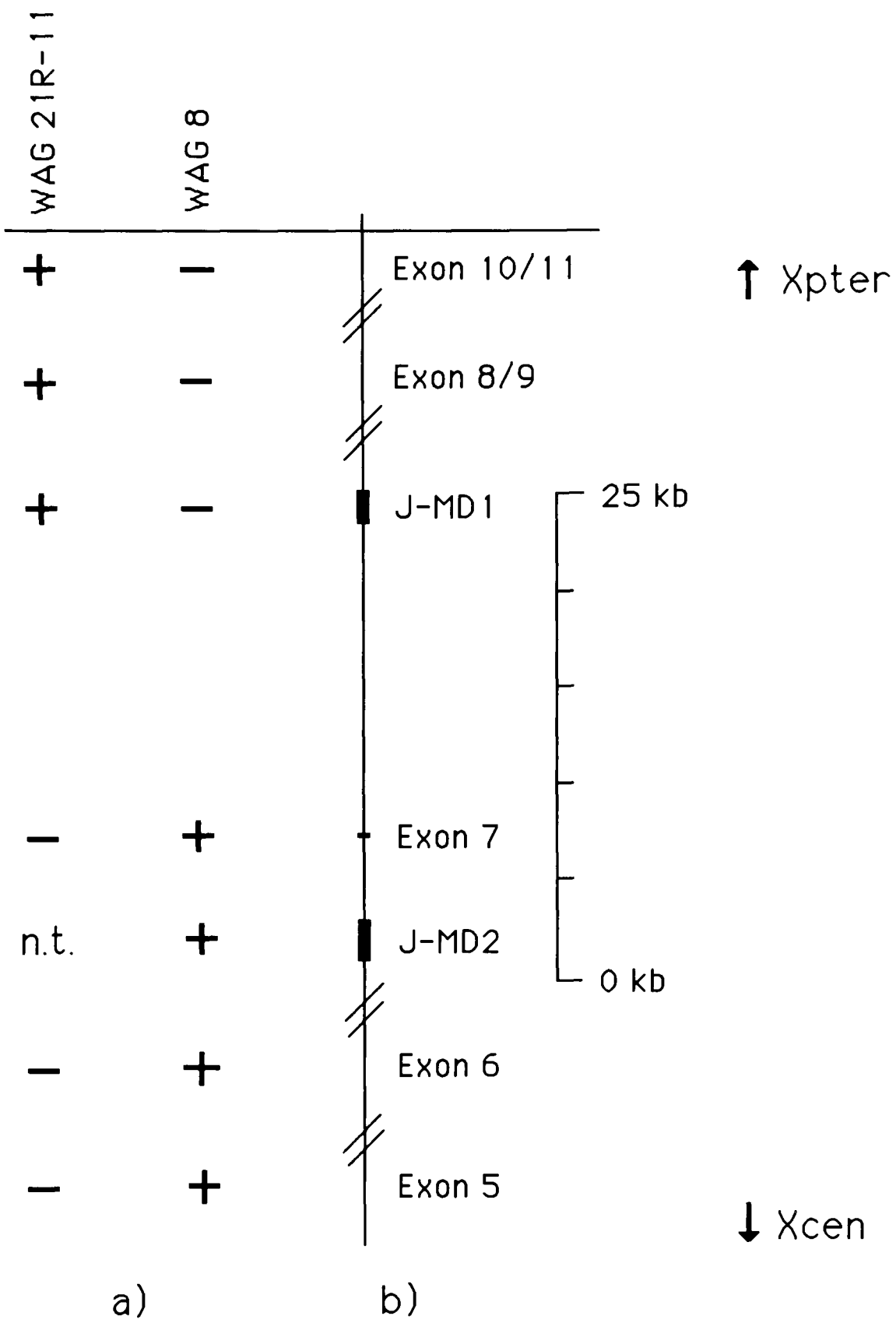


Figure 5.6 Summary of results on mapping the X chromosome breakpoint of the X;1 translocation by Southern analysis. a) Presence (+) or absence (-) of hybridisation (n.t. not tested) of probes (b) is indicated to the der(1) retaining hybrid WAG 21R-11 and the der(X) retaining hybrid WAG 8-1.

same intron as the t(X;1) (see Chapter 4) and was kindly supplied by Henry Klamut, Peter Ray and Ron Worton (Toronto) who had already partially mapped the insert of the cosmid with respect to EcoRI and HindIII restriction sites. This map is illustrated in figure 5.7 and indicated clearly that the breakpoint interval between exon 7 and J-MD1 was around 20 kb. This region was further characterised in order to identify the precise translocation breakpoint position on the X chromosome. This was performed by subcloning DNA fragments from the cosmid insert.

5.2.2 Characterisation of the X chromosome breakpoint region

The cosmid XJC-5 was digested with combinations of the restriction enzymes HindIII, EcoRI and PstI and the fragments generated were separated by electrophoresis in order to confirm the restriction map for this region [figure 5.8(a)]. The gel was blotted and probed with total human DNA to search for segments devoid of repetitive sequences and therefore likely to be useful as probes [figure 5.8(b)]. A 4.6 kb HindIII fragment (which gave a weak signal after hybridisation with total human DNA) was chosen in the first instance (see figure 5.7) and was purified by gel-elution. Southern analysis was performed using competitive DNA hybridisation to reduce the hybridisation of repetitive sequences from within the 4.6 kb fragment (Sealey et al., 1985). The 4.6 kb probe detected a 9.5 kb BamHI fragment in the hybrid WAG 8-1 (figure 5.9). This band was smaller than that detected in a normal individual (approximately 15 kb), suggesting that the der(X) junction fragment had been identified, although a possibility existed that the 9.5 kb and 15 kb BamHI fragments represented alleles of an RFLP and that the der(X) junction fragment had not been detected. However, if the assumption was made that the 9.5 kb fragment was the der(X) junction fragment, then the result refined the X chromosome breakpoint position to the interval of around 14 kb between this 4.6 kb HindIII fragment and J-MD1 (figure 5.7). The breakpoint could actually lie within the 4.6 kb HindIII fragment, although not in the proximal portion otherwise the probe would not have hybridised to DNA from the

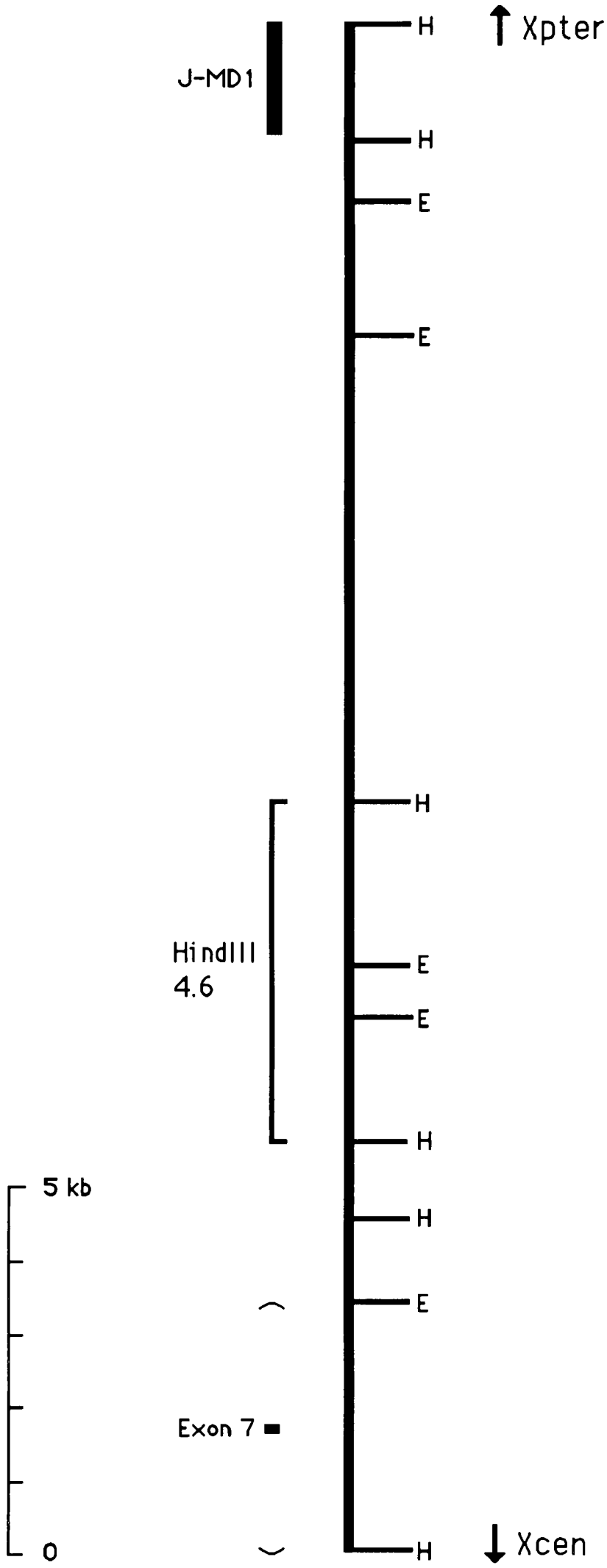


Figure 5.7 Preliminary restriction map supplied by H Klamut, P Ray and R Worton of the HindIII and EcoRI sites within the genomic DNA region surrounding the X chromosome breakpoint of the X;1 translocation.

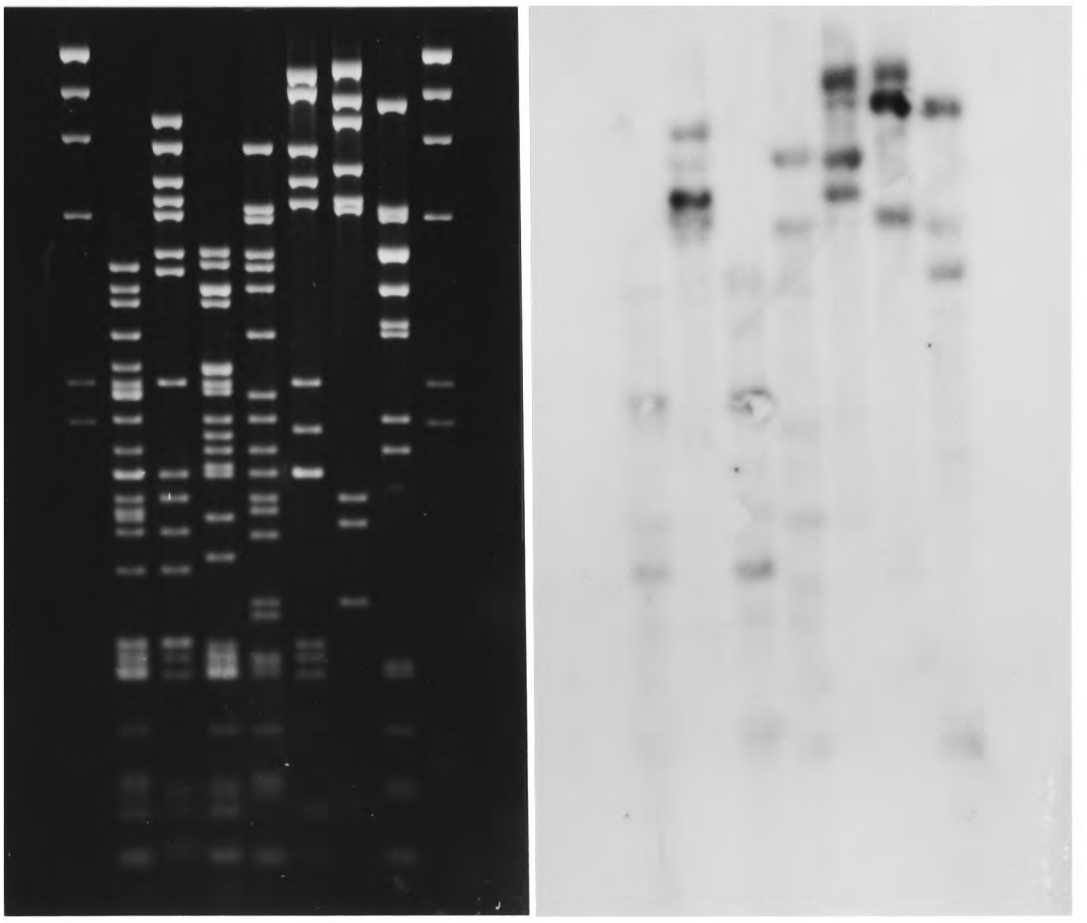


Figure 5.8 a) Restriction digests of the cosmid XJC-5 which spans the X chromosome breakpoint of the X;1 translocation. b) Southern blot of this gel probed with total human DNA in order to identify which fragments contain repetitive DNA sequences.

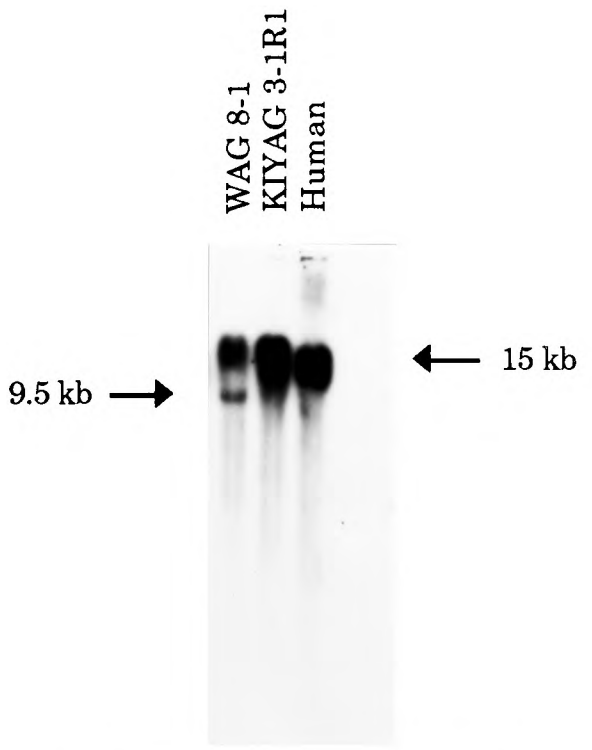


Figure 5.9 Detection of the BamHI der(X) junction fragment. A 9.5 kb fragment was detected in DNA from the der(X) retaining hybrid WAG 8-1 by Southern analysis using a 4.6 kb HindIII fragment gel-eluted from the cosmid XJC-5 as a probe. The 9.5 kb signal was absent in DNA from a human control and a mouse/human hybrid (KIYAG 3-1R1) not containing this portion of the X chromosome.

der(X) retaining hybrid. While performing these experiments, data from simultaneous subcloning and restriction mapping experiments of DNA fragments from the cosmid predicted that the 4.6 kb HindIII fragment and DMD exon 7 were within the same 15 kb BamHI fragment (see section 5.2.3, figure 5.11). The cDNA probe 1b+2a (see Chapter 4) which includes exon 7, would therefore be expected to detect the same 15 and 9.5 kb BamHI fragments (in addition to bands corresponding to other exons detected by this cDNA fragment). Southern blot hybridisation confirmed this expectation and is illustrated in figure 5.10.

5.2.3 Subcloning of the X chromosome breakpoint region

Cosmid fragments which spanned the region containing the X chromosome breakpoint were subcloned into pUC18. Two 'shotgun' cloning experiments were performed, where the cosmid DNA was digested using HindIII in the first experiment and using XbaI in the second. In the first experiment a clone was identified, C5-75, which contained two adjacent HindIII fragments of 1.6 kb (corresponding to the genomic clone J-MD1) and 9 kb in length (see figure 5.11). Thus this clone was either produced by double cloning of the two fragments which by chance retained their original orientation with respect to each other, or else the entire 10.6 kb was cloned in one piece and had not been completely digested prior to ligation. The insert of C5-29 which was isolated from the XbaI 'shotgun' cloning experiment contained a 4 kb insert which overlapped the insert of C5-75 (figure 5.11). Together, these clones span 13 kb and cover the expected breakpoint site between the 4.6 HindIII fragment and J-MD1 (section 5.2.2). The inserts of both clones were mapped with respect to the cutting sites of six restriction endonucleases: EcoRI, HindIII, PstI, XbaI, EcoRV and BamHI. For example, figure 5.12 shows a gel of the single and double digests in all combinations of six restriction enzymes of the 9 kb HindIII fragment of C5-75, together with the result of probing a blot of the gel with total human DNA to identify fragments containing repetitive sequences. Similar analyses were performed of C5-29 and other subcloned

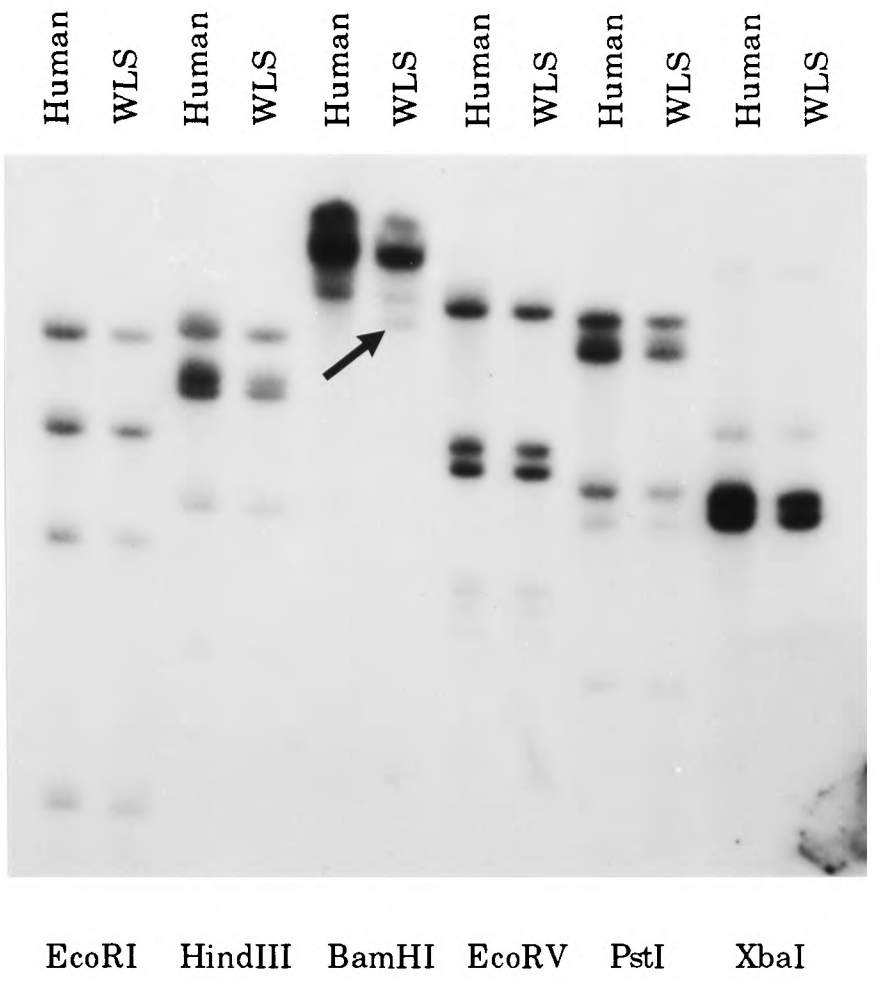


Figure 5.10 Detection of the der(X) junction fragment using cDNA probe 1b+2a (see Chapter 4, section 4.1). A 9.5 kb fragment was detected in BamHI digested DNA of the der(X) retaining hybrid WAG 8-1 and not in DNA of a control human.

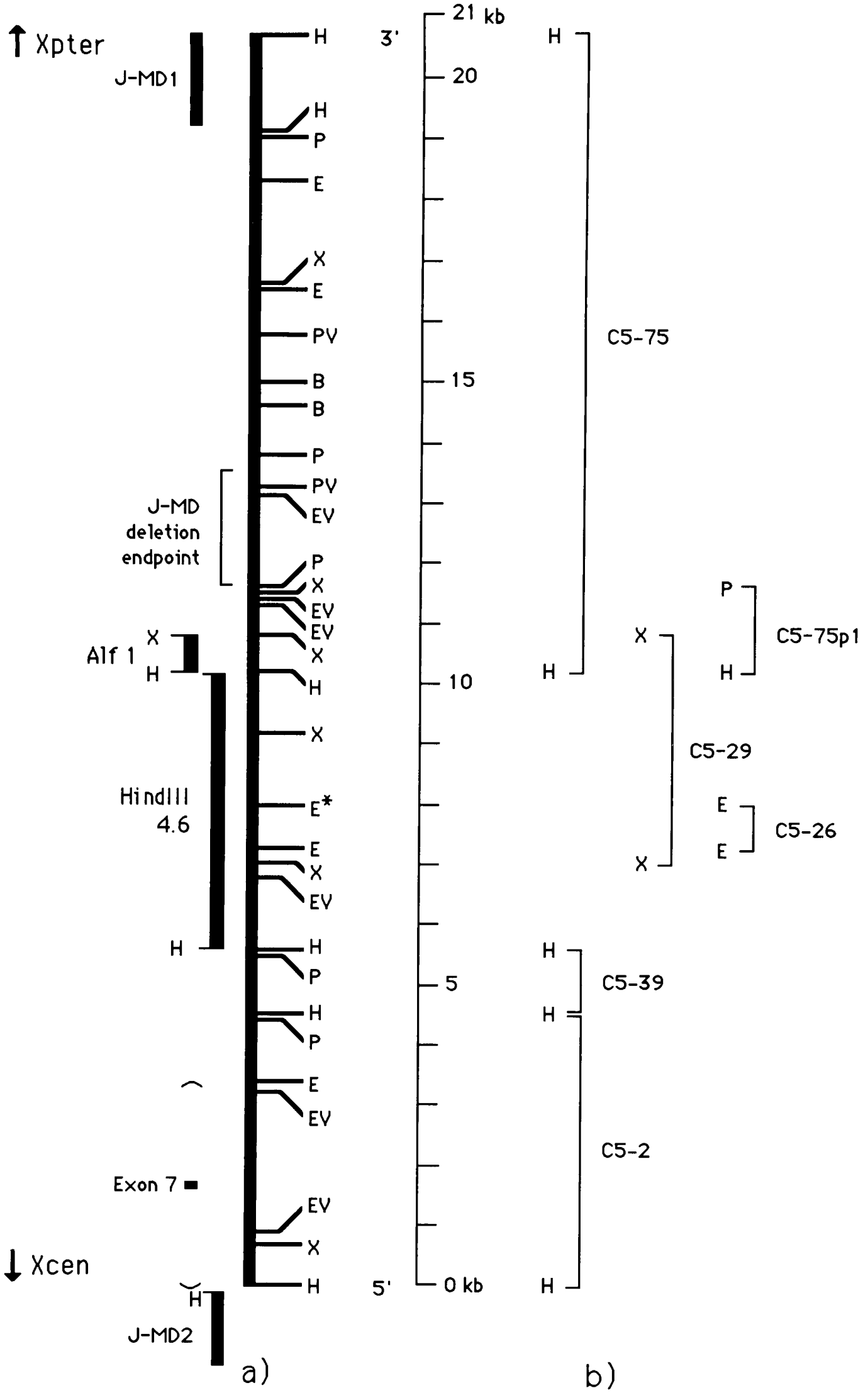


Figure 5.11 a) Detailed restriction map of the genomic interval between the genomic probes J-MD1 and J-MD2. Restriction sites are HindIII (H), XbaI (X), PstI (P), EcoRI (E), PvuII (PV), BamHI (B) and EcoRV (EV). An EcoRI site found to be polymorphic is marked with an asterisk. b) DNA fragments subcloned from the cosmid XJC-5.

a)

b)

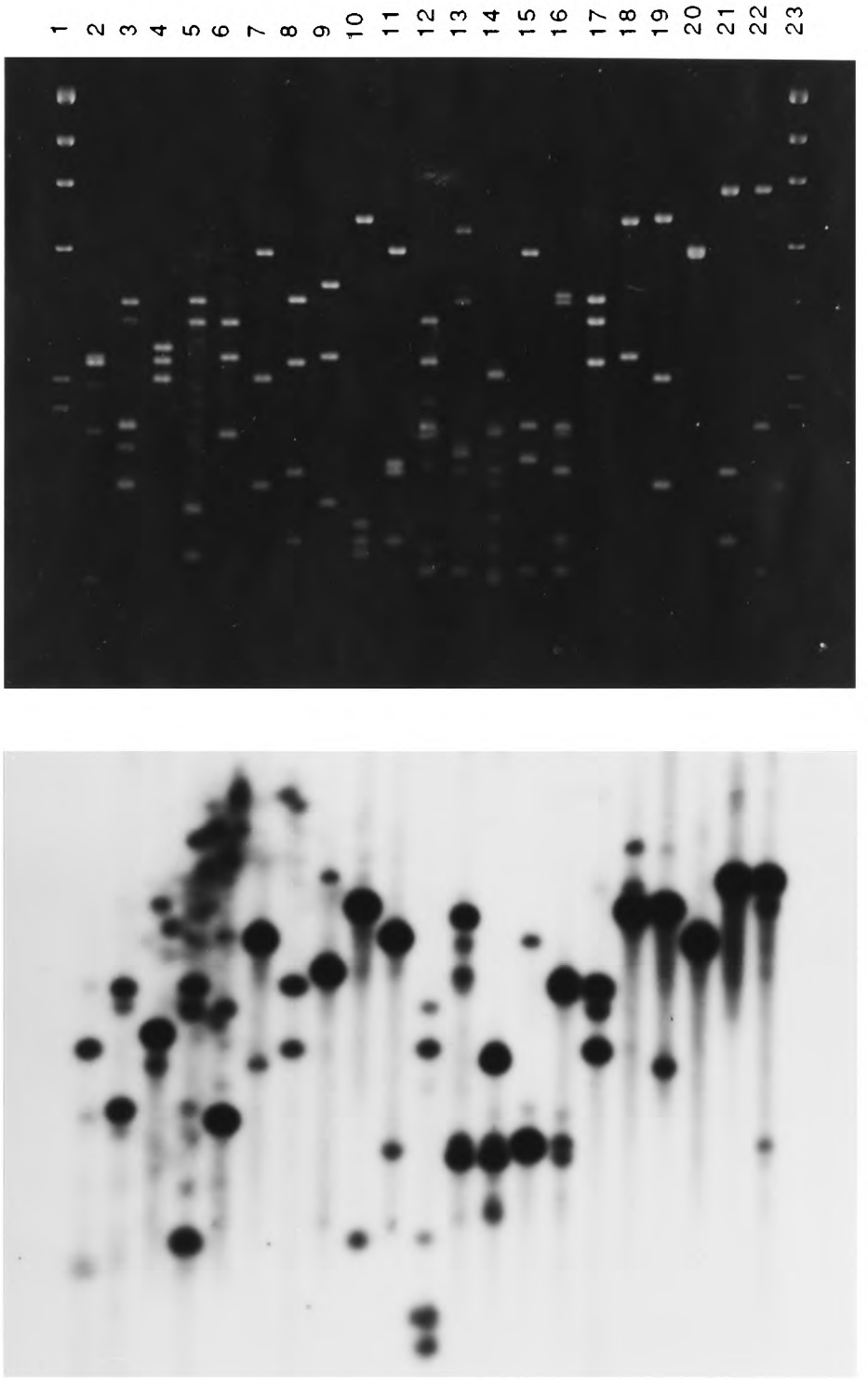


Figure 5.12 Restriction map analysis of the 9 kb insert of C5-75. Tracks 1 and 23 are HindIII digested λ. a) Ethidium bromide-stained fragments generated by digestion with 1, X/PV; 2, P/PV; 3, P/X; 4, B/PV; 5, B/X; 6, B/P; 7, EV/PV; 8, EV/X; 9, EV/P; 10, EV/B; 11, E/PV; 12, E/X; 13, E/P; 14, E/B; 15, E/EV; 16, P; 17, X; 18, P; 19, B; 20, EV; and 21, E. Restriction enzymes are EcoRI (E), EcoRV (EV), BamHI (B), PstI (P), XbaI (X) and PvuII (PV). b) Southern blot of the same gel probed with total human DNA to identify which fragments contain human repetitive sequences.

DNA fragments. The locations of C5-75, C5-29, and other subclones are shown in figure 5.11 with the full deduced restriction map between exon 7 and J-MD1. This map is consistent with the published restriction map for the enzymes EcoRI and HindIII of 50 kb cloned DNA surrounding the endpoint of the J-MD deletion (Monaco et al., 1987). The position of this deletion endpoint is within an approximately 2 kb region of DNA immediately distal to the insert of the subclone C5-75p1 (see figure 5.11).

5.2.4 Detection of the der(1) junction

A 600 bp HindIII, XbaI single copy fragment, Alf1 (the overlap of C5-75 and C5-29; see figure 5.11), was subcloned and found in Southern blot analysis to detect altered fragments (hence the name A l f 1) in the t(X;1) patient's DNA for five out of six restriction enzymes tested (see figure 5.13). An altered fragment was detected in XbaI digests but not HindIII digests, which indicated that the breakpoint must lie beyond the HindIII site of Alf1 (i.e. proximal to the probe). This was confirmed by detecting a 10 kb BamHI junction fragment in the der(1) retaining hybrid WAG 21R-11 and not in WAG 8-1 which retains the der(X) (figure 5.14). The same BamHI junction fragment was shown not to be present in eight additional unrelated individuals (seven of them female, therefore representing 15 X chromosomes; figure 5.15). As the detailed restriction map of the X chromosome was known, the sizes of junction fragments indicated the positions of restriction enzyme cutting sites beyond the translocation breakpoint (figure 5.16). These sites were presumed, but not yet proven, to lie on chromosome 1. This information was important for two reasons. Firstly, it allowed refined positioning of the X chromosome breakpoint. The proximal limit was now defined by the chromosome 1 PstI site estimated to be 300 bp from the distal limit of the breakpoint which was the HindIII site of Alf1 (see figures 5.13 and 5.16). Therefore the X;1 translocation junction was found to lie approximately 1.5 kb to 4 kb proximal of the J-MD deletion endpoint (Monaco et al., 1987; see figure 5.16). Secondly, this detailed information was required for the design of conditions required by the inverse PCR

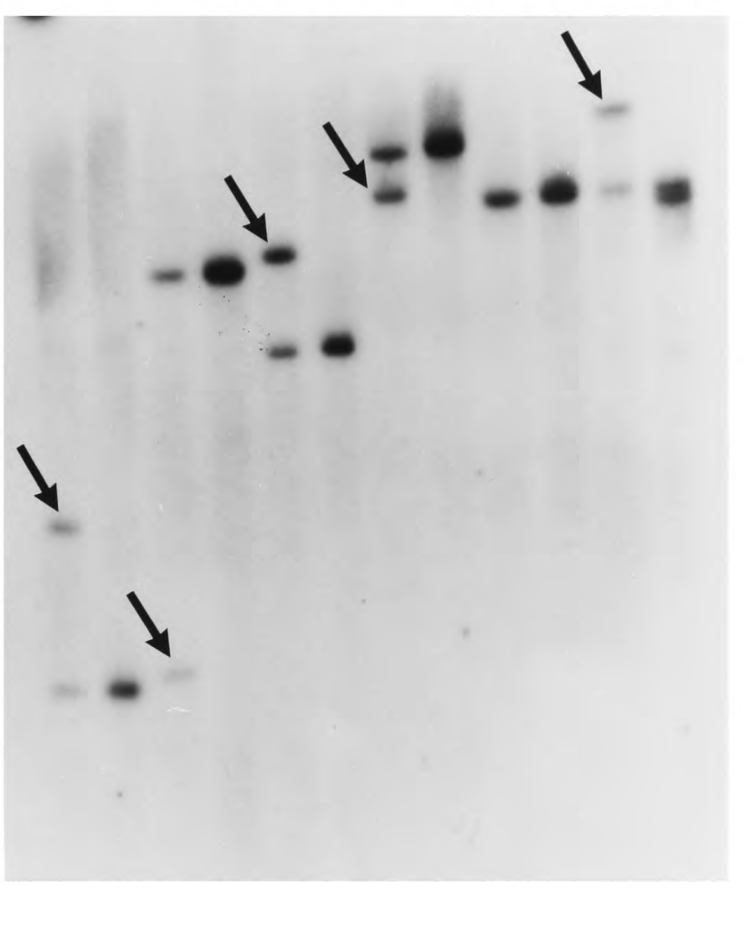


Figure 5.13 Detection of junction fragments using the probe Alf1. Restriction fragments which are altered in size compared to bands detected in a normal individual are seen using five out of six restriction enzymes (arrowed). The sizes of junction fragments are XbaI (2.7 kb), PstI (1.6 kb), EcoRV (6.5 kb), BamHI (10 kb) and EcoRI (15 kb). The EcoRI fragment in the control human (female) track appears to be a doublet. Further investigation revealed that Alf1 detects an EcoRI polymorphism (section 5.2.8).

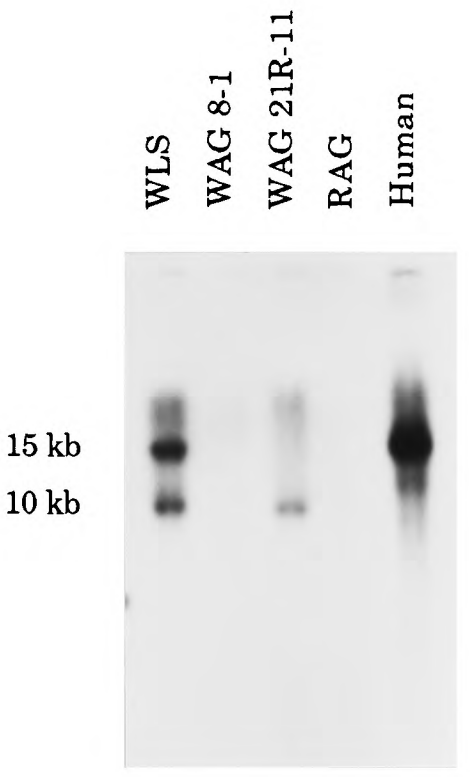


Figure 5.14 Mapping of the X;1 translocation breakpoint proximal to Alf1. The BamHI junction fragment (10 kb) was detected in the der(1) retaining hybrid WAG 21R-11 and not the der(X) retaining hybrid WAG 8-1.

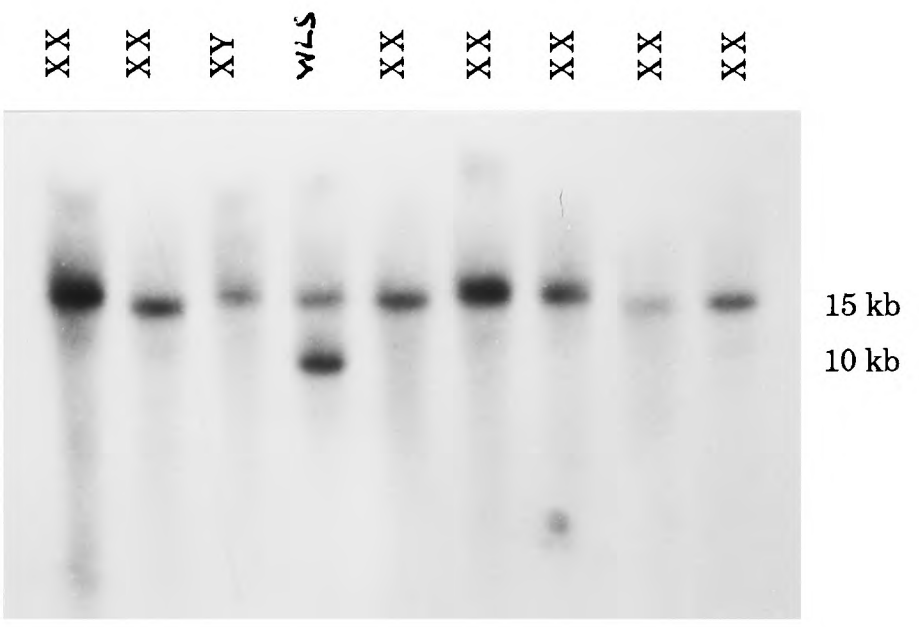


Figure 5.15 Confirmation that the altered BamHI fragment (10 kb) detected in the X;1 translocation patient (WLS) is not detected in seven normal females (XX) or in a normal male (XY).

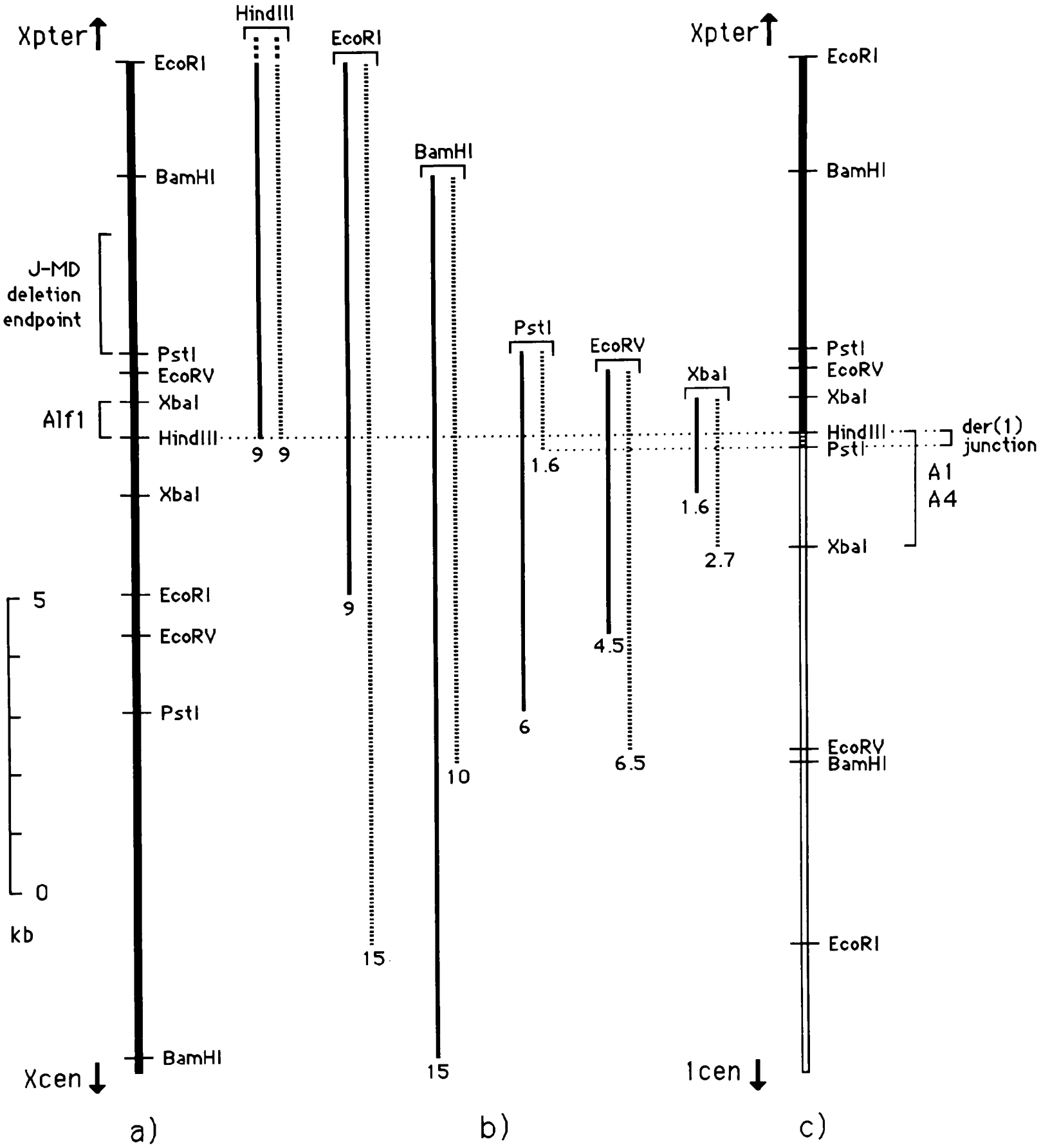


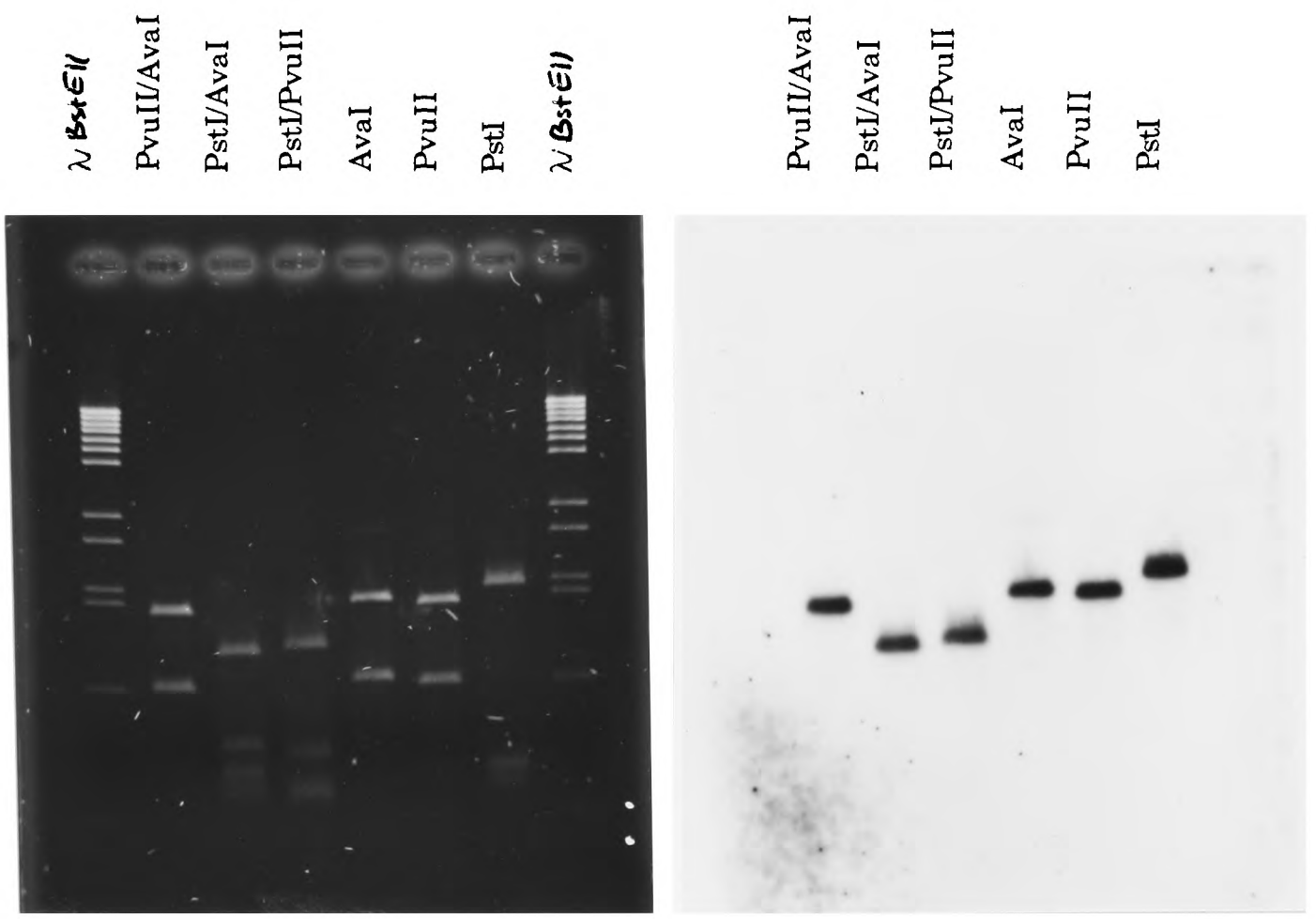
Figure 5.16 Determination of the restriction map around the der(1) junction by Southern analysis using the probe Alf1. The sizes of restriction fragments detected by Alf1 are indicated in b) (solid lines, normal X chromosome fragments and dotted lines, der(1) fragments). The positions of restriction sites on the der(1) (c) were determined by comparing the sizes of junction fragments in b) with the known normal X chromosome restriction map (a). The position of the endpoint of a deletion, J-MD, is indicated (Monaco et al. 1987). The 2.7 kb XbaI junction fragment was used as a template for an inverse PCR (Chapter 6). Two clones, A1 and A4, spanning the der(1) junction were isolated from the amplified DNA.

amplification strategy used in Chapter 6 to clone the der(1) junction fragment. The 2.7 kb XbaI junction fragment was chosen as a template for inverse PCR (see figure 5.16).

5.2.5 Preliminary characterisation of the cloned der(1) junction and detection of normal chromosome 1 restriction fragments

The cloning of the der(1) junction is described in Chapter 6, section 6.2.1. Two identical 2 kb clones (A1 and A4) containing the 2 kb HindIII/XbaI der(1) junction were isolated (see figure 5.16). The insert of A1 was gel-purified and digested using combinations of the restriction endonucleases PstI, PvuII and AvaI. A gel of the restriction fragments generated and the result of probing a blot of this gel with total human DNA are illustrated in figure 5.17. This experiment permitted the construction of the restriction map of the 2 kb A1 fragment which is illustrated in figure 5.18. The distal 320 bp HindIII/PstI portion of this DNA segment is the expected site of the der(1) junction (see figure 5.16). Repetitive human DNA was detected within the insert of A1 and was localised to the 880 bp AvaI/PstI restriction fragment (figure 5.17; figure 5.18). A 430 bp AvaI/PstI fragment (named AP430) which lies between the fragment containing the repetitive DNA and the expected site of the der(1) junction (figure 5.18) was gel-purified from the A1 insert and used as a probe in Southern blot analysis. Sequencing of the der(1) junction later revealed that the repetitive DNA within the 880 bp AvaI/PstI fragment is an Alu element (Chapter 7, section 7.2.2.2).

AP430 was found in Southern blot analysis to detect the same altered restriction fragments in the t(X;1) patient's DNA as Alf1. This is illustrated for the restriction enzymes PvuII, SstI and BglII in figure 5.19, and confirms that both probes are located within the der(1) junction fragments produced by these enzymes. However, the other fragments detected in the t(X;1) patient's DNA which are the same as in the control DNA, were different from the normal X chromosome bands detected by Alf1. These fragments must represent restriction fragments from the normal chromosome 1 breakpoint region (assuming that the translocation is a simple two-breakpoint



a) b)

Figure 5.17 a) Restriction digests of the insert of the clone A1 which spans the der(1) junction run on a gel and stained with ethidium bromide. b) Southern blot of the same gel probed with total human DNA to identify fragments containing human repetitive sequences. The smallest fragment containing repetitive DNA is an 880 bp PstI/AvaI fragment.

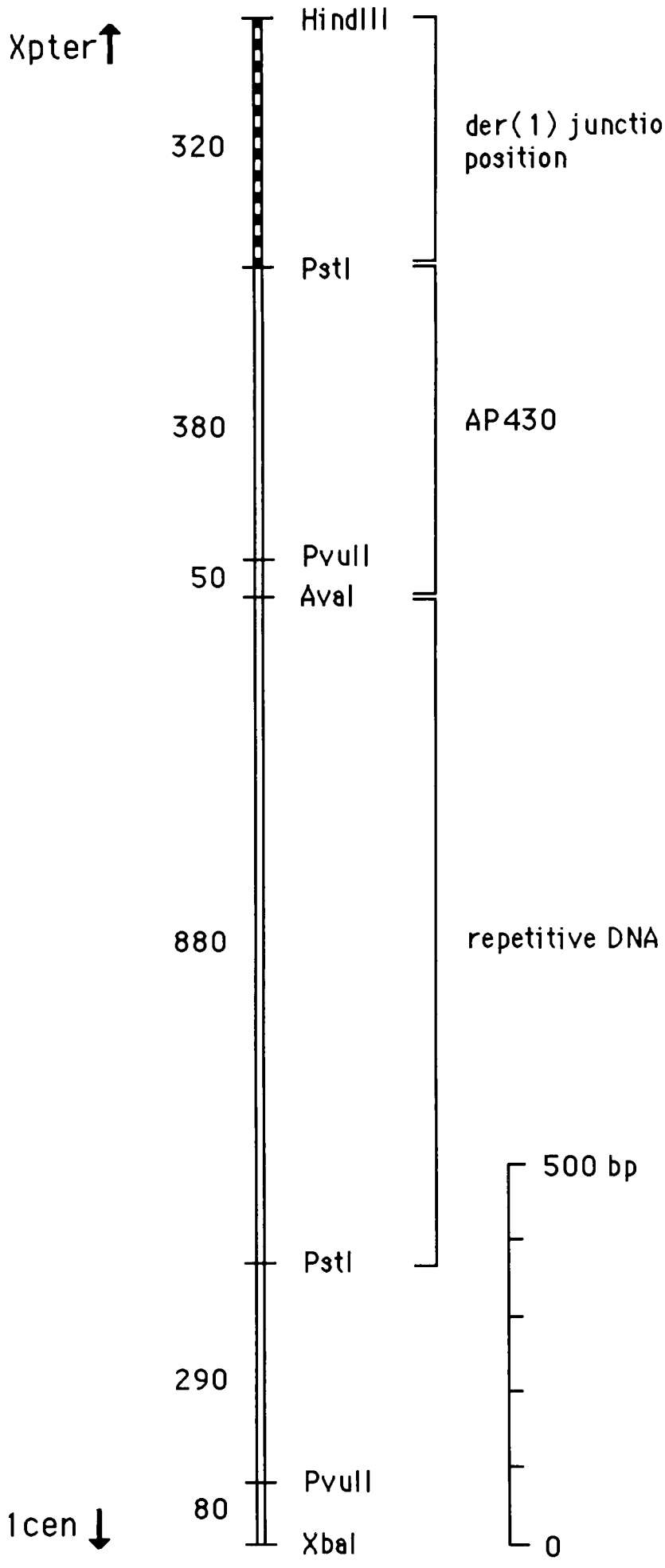


Figure 5.18 Restriction map of the insert of A1. The position of the der(1) junction is indicated, and the position of a PstI/AvaI 430 bp fragment which was gel-eluted and used in Southern analysis is also indicated.

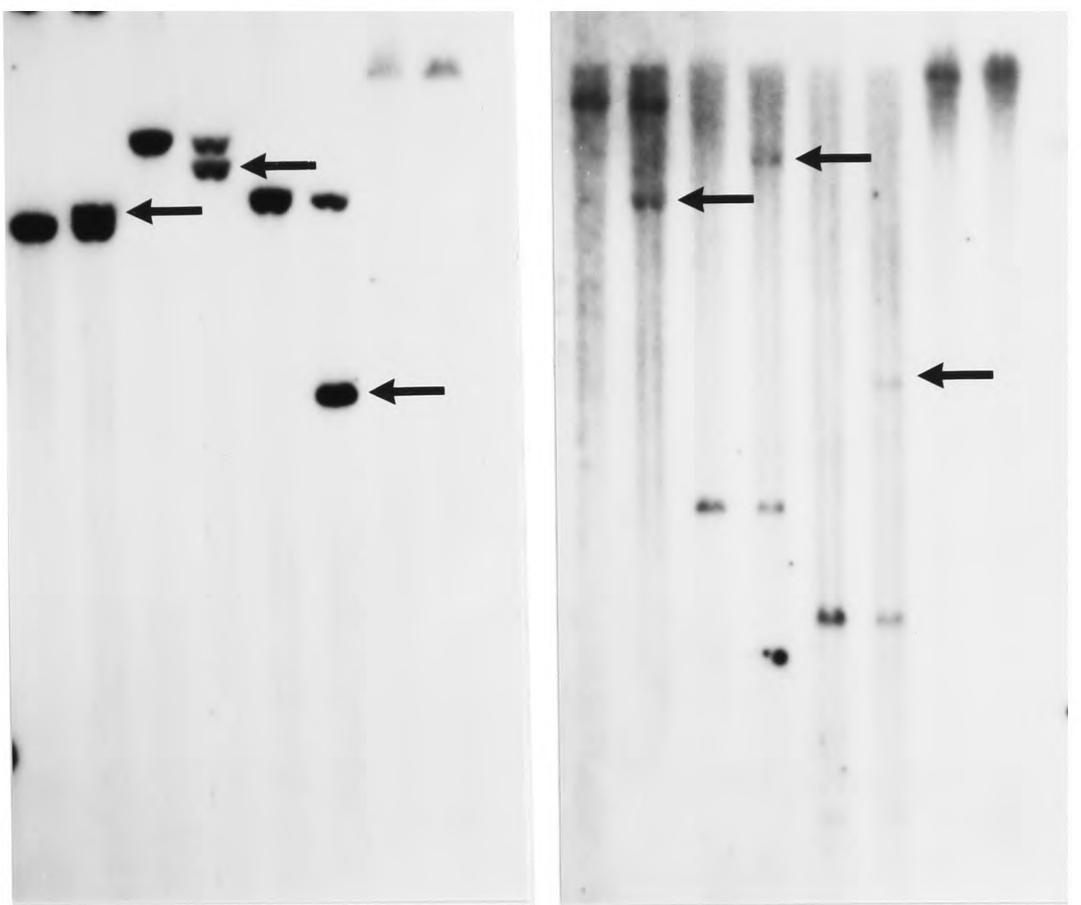


Figure 5.19 Alf1 and AP430 detect the same junction fragments. Southern blot of DNA from the t(X;1) patient and a normal individual sequentially probed with Alf1 and AP430 showing that the junction fragments (arrowed) produced by BglII (6.6 kb), SstI (8 kb) and PvuII (3.4 kb) are identical although the fragments detected by these probes in a normal human are different.

rearrangement) and therefore confirmed that Alf1 and AP430 must flank the der(1) junction. Thus Southern blot analysis obtained using probe AP430 provided data from which a preliminary restriction map around the chromosome 1 breakpoint was constructed [Chapter 6, figure 6.5(a)]. This information was used to amplify and clone the region by inverse PCR (Chapter 6, section 6.2.2).

5.2.6 Preliminary characterisation of the cloned chromosome 1 breakpoint region and detection of the der(X) junction

The cloning of DNA from around the chromosome 1 breakpoint region following two independent inverse PCR amplifications is described in Chapter 6, section 6.2.2. The location of clones and the restriction map of the cloned region which extends 2 kb distal of the expected breakpoint site was determined by digestion of these clones with EcoRI, HindIII, PstI, XbaI, EcoRV and BamHI (figure 5.20). The insert of one clone, PPii, proved to be a useful single copy DNA probe. In Southern blot analysis of DNA from normal individuals, the gel-purified insert of PPii generally detected the same restriction fragments in a normal individual as AP430 (for example KpnI, PvuII, SstI and BglII; figure 5.21), thus confirming that the expected DNA had been amplified and cloned. However the junction fragments which were detected in the t(X;1) patient's DNA differed in size from those detected by AP430, and thus confirmed that AP430 and PPii must flank the breakpoint site. Figure 5.22 illustrates the hybridisation pattern of PPii to a Southern blot of BamHI digested DNA from the somatic cell hybrids WAG 8-1 and WAG 21R-11 and control DNA. Two bands of 13 kb and 6 kb are detected in a normal individual (track 1) since PPii contains a BamHI restriction site (see figure 5.23). The 6 kb fragment must lie proximal to the 13 kb fragment as AP430 detects the 6 kb BamHI fragment (not shown). An altered band of 9.5 kb is detected in cell line DNA from the translocation patient in addition to both normal bands. In DNA from the der(X) retaining hybrid WAG 8-1, the normal 13 kb and the altered 9.5 kb bands are detected. Therefore both these bands represent the der(X). The altered band (9.5 kb) is

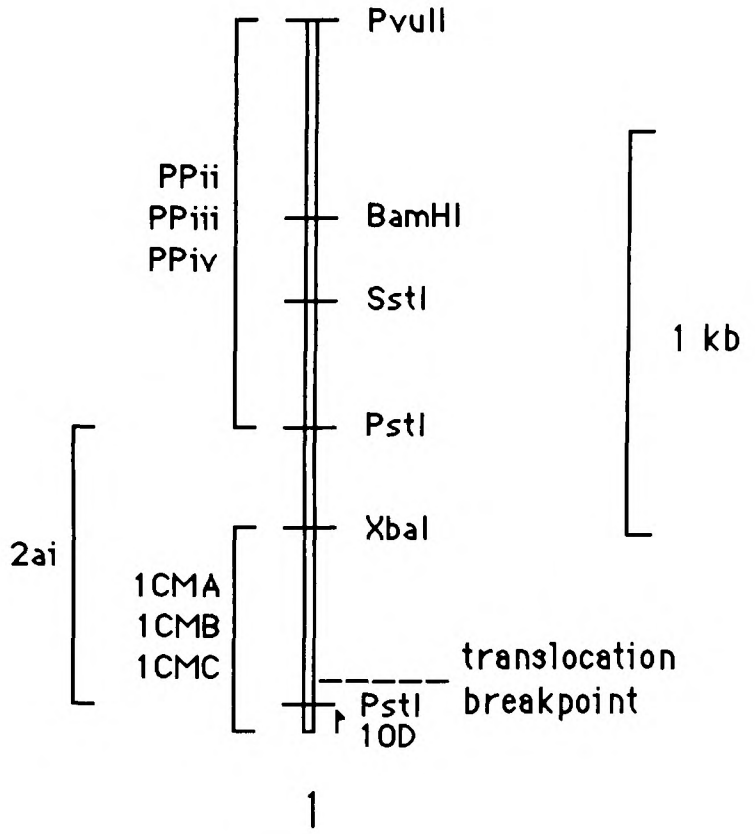


Figure 5.20 Restriction map of DNA cloned from the products of inverse PCR amplification of the normal chromosome 1 region which spans the X;1 translocation breakpoint. The locations of the cloned fragments is indicated (see Chapter 6, section 6.2.2).

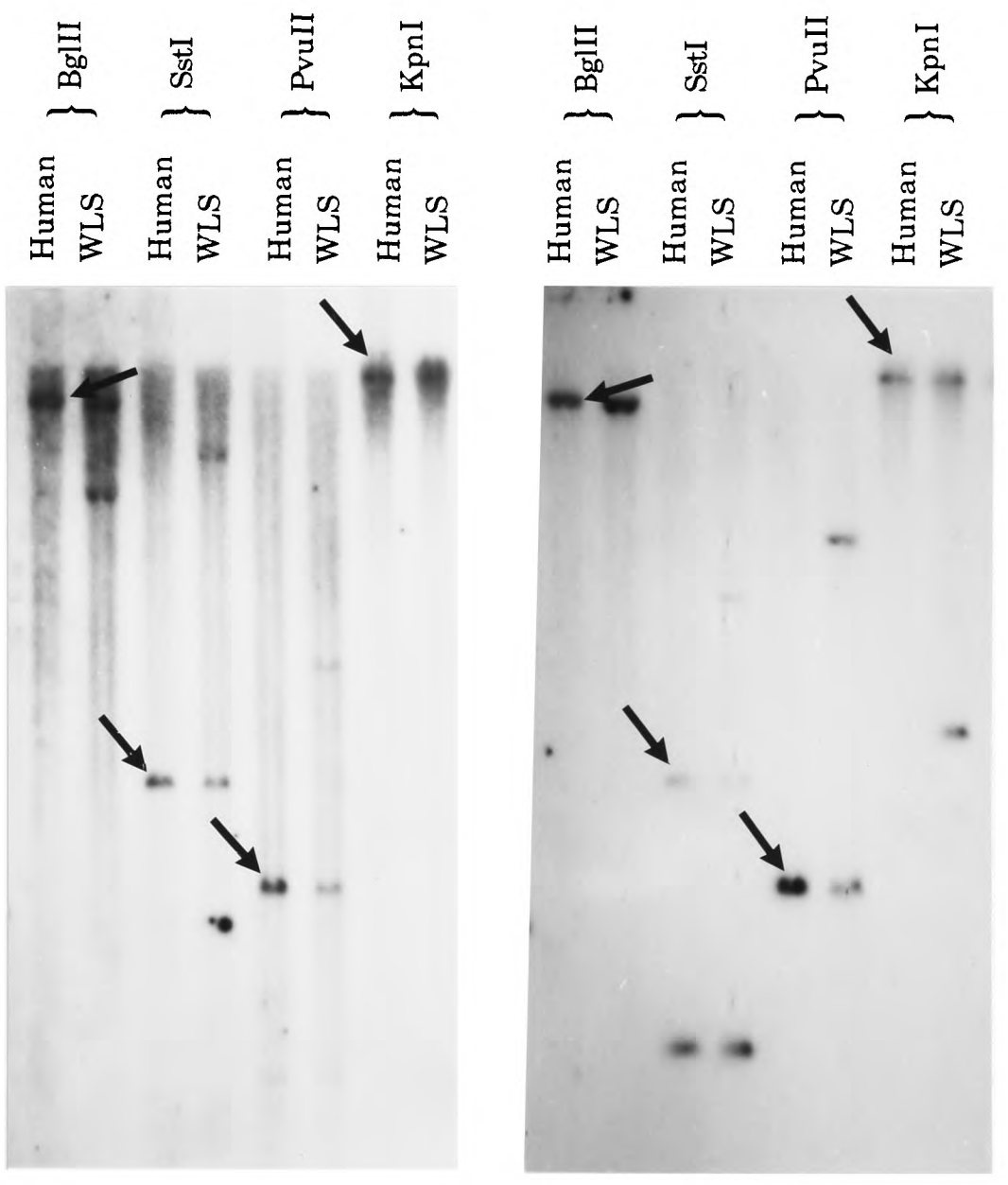


Figure 5.21 AP430 and PPii detect the same 'normal' fragments. Southern blot of DNA from the t(X;1) patient and a normal individual sequentially probed with AP430 and PPii showing that the fragments detected in a normal individual (arrowed) produced by BglII (14 kb), SstI (2.5 kb), PvuII (1.9 kb) and KpnI (>20 kb) are the same, although the junction fragments detected by these probes in the t(X;1) patient are different.

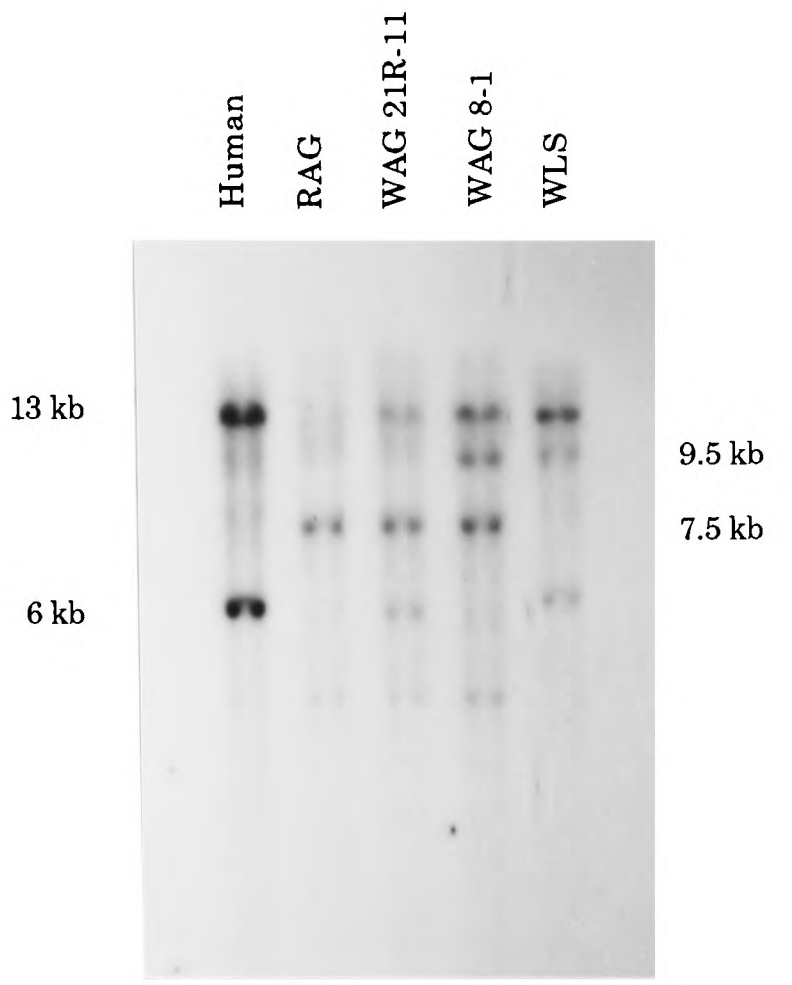


Figure 5.22 Confirmation that the BamHI junction fragment detected by PPii is on the der(X). Southern analysis indicated that the BamHI 9.5 kb junction fragment is present in the der(X) retaining hybrid WAG 8-1 and not in the der(1) retaining hybrid WAG 21R-11. PPii was observed to cross-hybridise to a 7.5 kb mouse BamHI fragment. Two DNA fragments (13 and 6 kb) are detected on a normal chromosome 1 since PPii contains a BamHI site (see text and figure 5.23 for details).

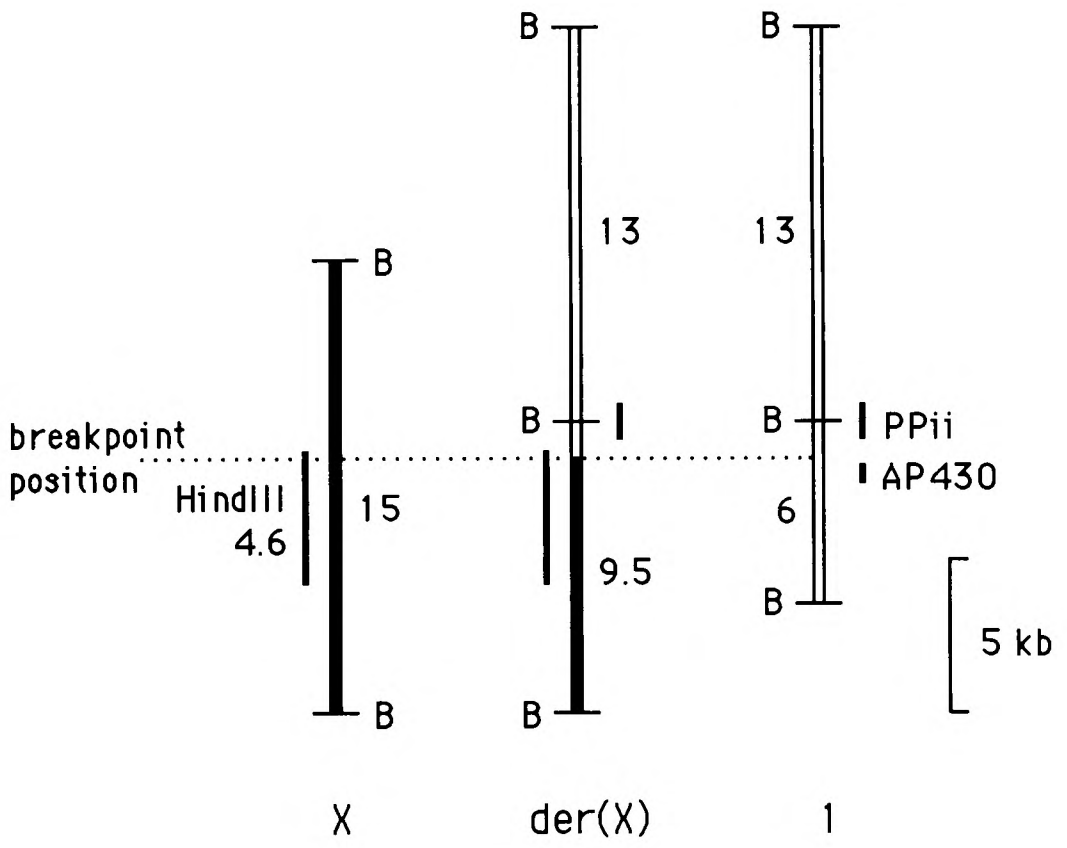


Figure 5.23 Diagram of BamHI (B) restriction sites around the X chromosome (dark line) and chromosome 1 (pale line) breakpoint positions, and on the der(X). The sizes of restriction fragments detected by the probes, PPii, AP430 and HindIII 4.6 are indicated. PPii spans a BamHI site and therefore detects both the 13 kb and the 6 kb restriction fragments from chromosome 1.

not detected in WAG 21R-11, however both normal fragments are present (13 kb and 6 kb). This can be attributed to the presence of normal chromosome 1 in this hybrid (Chapter 2, table 2.1). A further finding of this hybridisation was that cross hybridisation of PPii was observed to a mouse 7.5 kb fragment indicating DNA sequence conservation between mouse and the human. The significance of this is explored further in Chapter 7, section 7.2.3.

The sizes of der(X) junction fragments detected by PPii were significant in two respects. Firstly, they were as would be expected for a two breakpoint translocation, i.e. they were consistent with known X chromosome restriction enzyme sites proximal to the breakpoint position. For example, PPii detects a 9.5 kb BamHI der(X) junction fragment (figure 5.22). This is the same size as the BamHI junction fragment that was detected earlier by X chromosome probes proximal to the breakpoint (see section 5.2.2). This therefore was the first evidence that the translocation being studied was a two breakpoint rather than a three breakpoint rearrangement involving an inversion of X chromosome material as originally described (Lindenbaum et al., 1979). Secondly, the sizes indicated that no significant (i.e. not detectable by Southern analysis; not larger than about 100 bp) deletion, duplication or insertion was associated with the translocation in the vicinity of the breakpoints. A full restriction map of the X;1 translocation regions and of the junction fragments could finally be constructed and is illustrated in figure 5.24.

5.2.7 Localisation of PPii to 1p34 by in situ hybridisation

In order to confirm that the rearrangement cloned in the work described here was indeed the X;1 translocation and not another rearrangement coincidentally present in the patient (for instance the breakpoints may have been cloned of an inversion on Xp, present in addition to the translocation), an in situ hybridisation was performed of PPii to BudR-replication banded male metaphase chromosomes (in collaboration with S. Holt, MRC Radiobiology Unit, Harwell). Full karyotype analysis of 37 metaphases

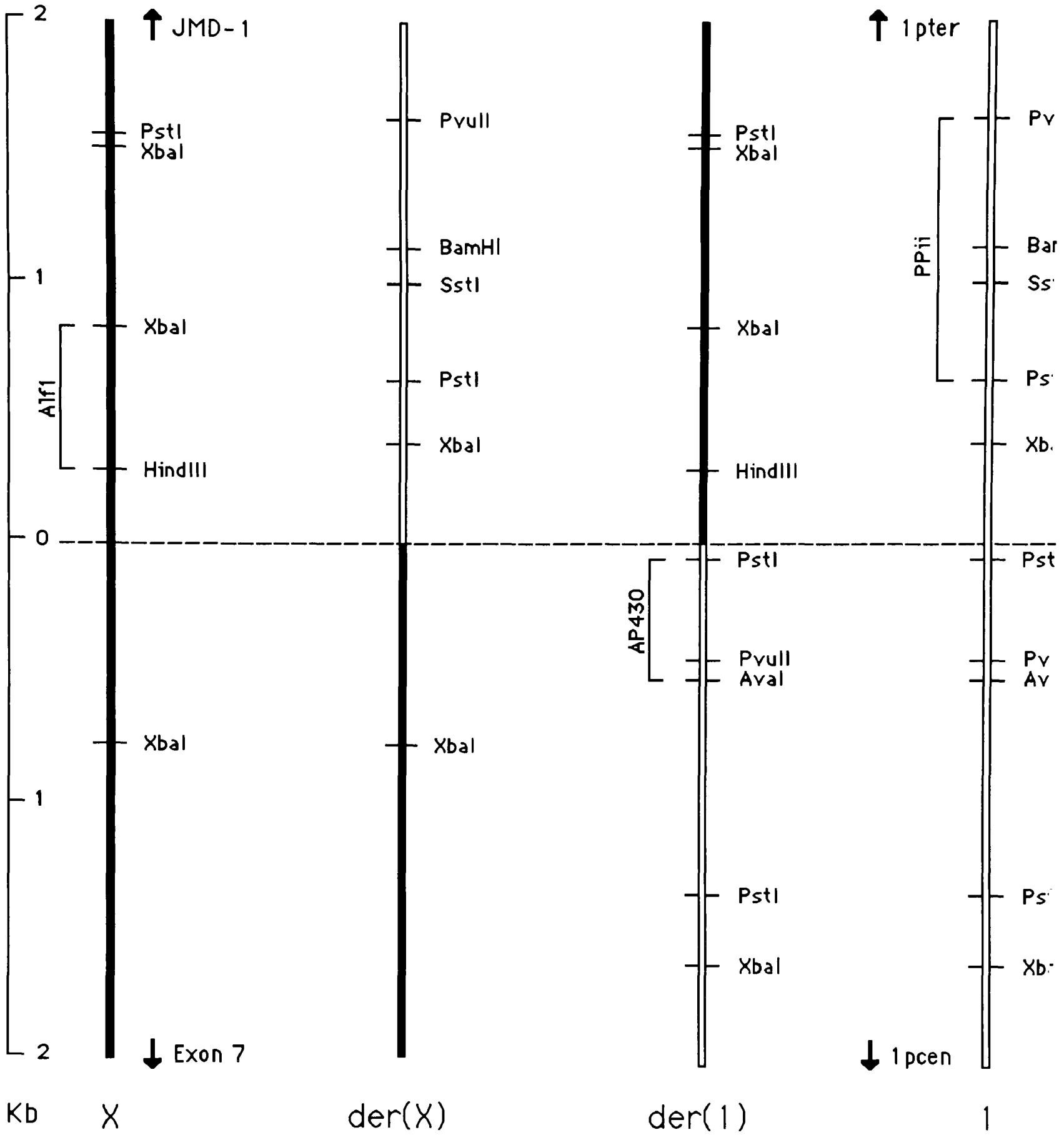


Figure 5.24 Complete restriction map around the X;1 translocation junctions and breakpoint regions from the normal chromosomes. The positions of probes isolated in the course of this investigation are indicated.

detected 26.9% (46/171) of scored grains on chromosome 1 and no other peak of hybridisation (figure 5.25). A further 53 cells were partially analysed and 41 grains were scored from 32 chromosome 1s. Out of a total of 87 grains scored on chromosome 1, 41 (48%) were in the region 1p35-p33 and 28 of these (32%) were in 1p34 (figure 5.26). Seven chromosomes with a grain on 1p34 are illustrated in figure 5.27. These results confirm that the X;1 translocation junctions have been cloned and also confirm the original cytogenetic assignment of the chromosome 1 breakpoint to 1p34 (Lindenbaum et al., 1979; Boyd and Buckle, 1986). Sequencing the X;1 translocation breakpoints revealed that PPii lies in an intron of the gene for the leukocyte antigen related protein (LAR; see Chapter 7, section 7.2.2.2). Therefore the localisation of PPii to 1p34 refines the localisation of the gene for LAR (Cockburn et al., 1991).

5.2.8 Detection of an EcoRI polymorphism in intron 7 of the DMD gene

Probing of a Southern blot of EcoRI digested DNA from a normal female using the probe Alf1 revealed that the 9 kb band detected might be a doublet (see figure 5.13). On examination of the X chromosome restriction map (figure 5.11), an EcoRI recognition site 2 kb proximal to Alf1 was identified whose absence/presence would produce two restriction fragments of 9kb and 8.2 kb respectively. If this was the cause of the apparent polymorphism, then double digestion using EcoRI in combination with BamHI, PstI or EcoRV would result in improved resolution of alleles (see figure 5.11). A Southern blot of EcoRI/EcoRV digested DNA from 20 unrelated females was probed with Alf1 (figure 5.28). Two alleles of 4 kb and 3.2 kb were detected. This confirms that the presence/absence of the EcoRI site 2 kb proximal to Alf1 is responsible for the polymorphism. Thirteen alleles of 4 kb and 27 of 3.2 kb were detected. Nine out of the 20 females tested were heterozygous (45%). The alleles of 11 additional X chromosomes from unrelated individuals were also determined (not shown) bringing the total number of the rarer large alleles to 16/51 (31%) and of the commoner small alleles to 35/51 (69%). The expected heterozygote frequency in females is therefore 43%. This is

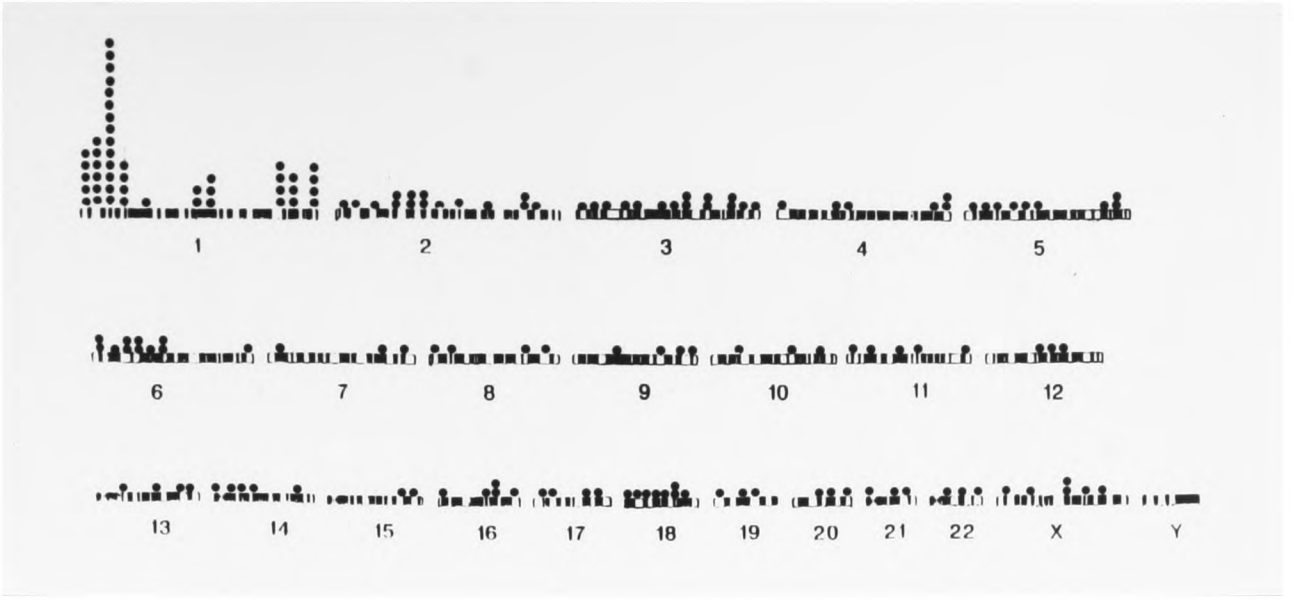


Figure 5.25 The distribution of 171 grains scored over the total karyotype in 37 cells after <u>in situ</u> hybridisation to PPii. All grains in these metaphases could be assigned unequivocolly to a chromosome region. Karyotype analysis was performed by S Holt and this figure prepared by Y Boyd.

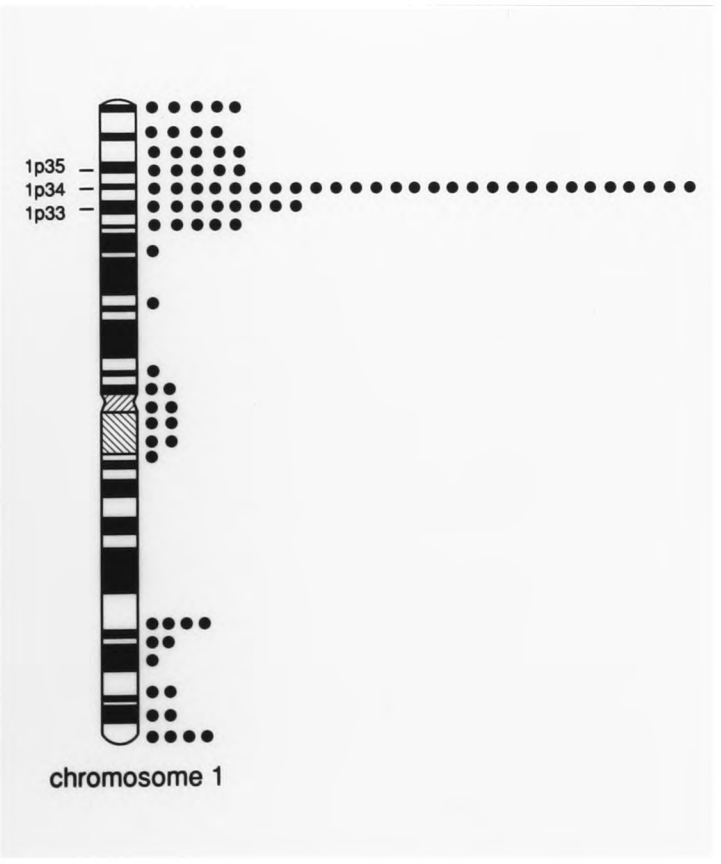


Figure 5.26 Distribution of 87 grains scored on chromosome 1 after <u>in situ</u> hybridisation to PPii. The peak of hybridisation is on band 1p34. Karyotype analysis was performed by S Holt and this figure prepared by Y Boyd.

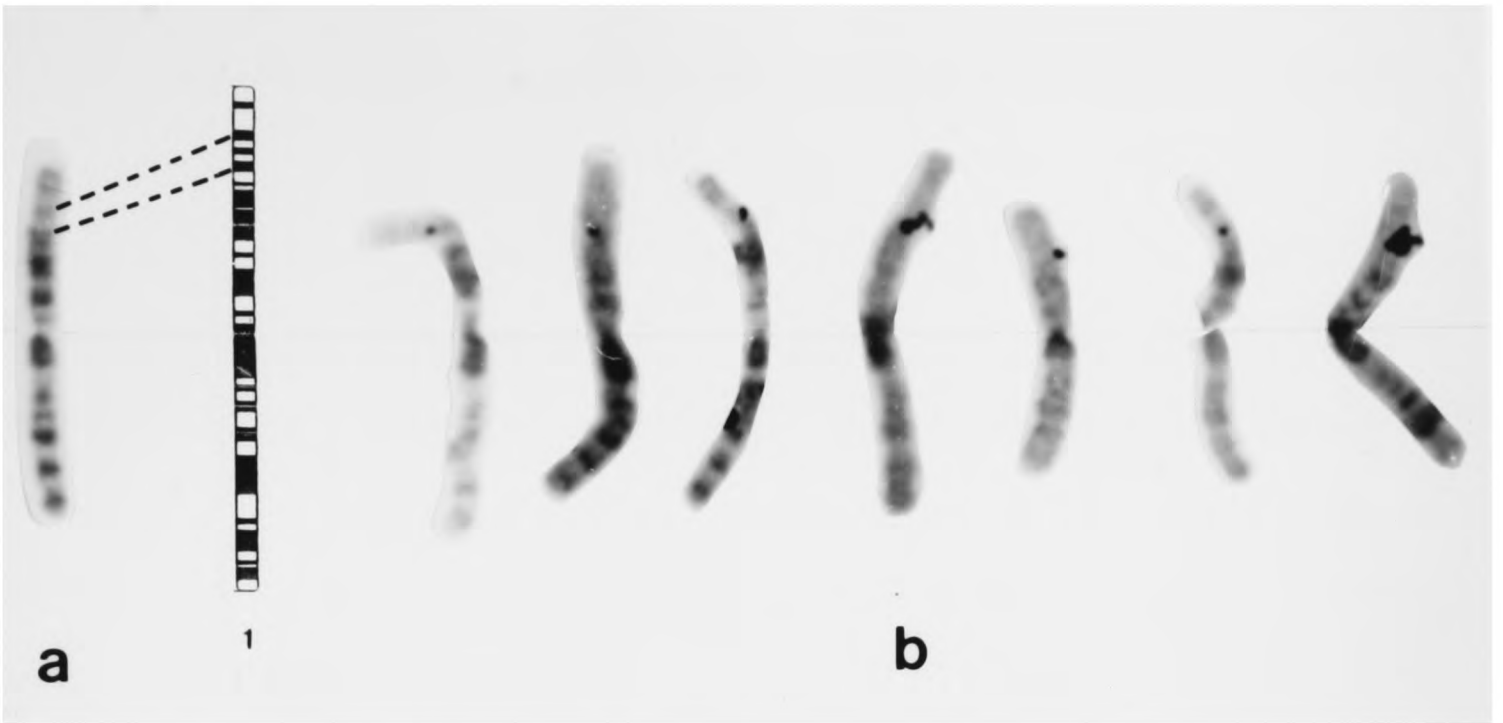


Figure 5.27 a) A typical replication-banded chromosome 1 compared with the standard ideogram for this chromosome (ISCN, 1985). b) Examples of grains scored on different chromosome 1s in the region 1p35-p33. Karyotype analysis and figure prepared by S Holt.

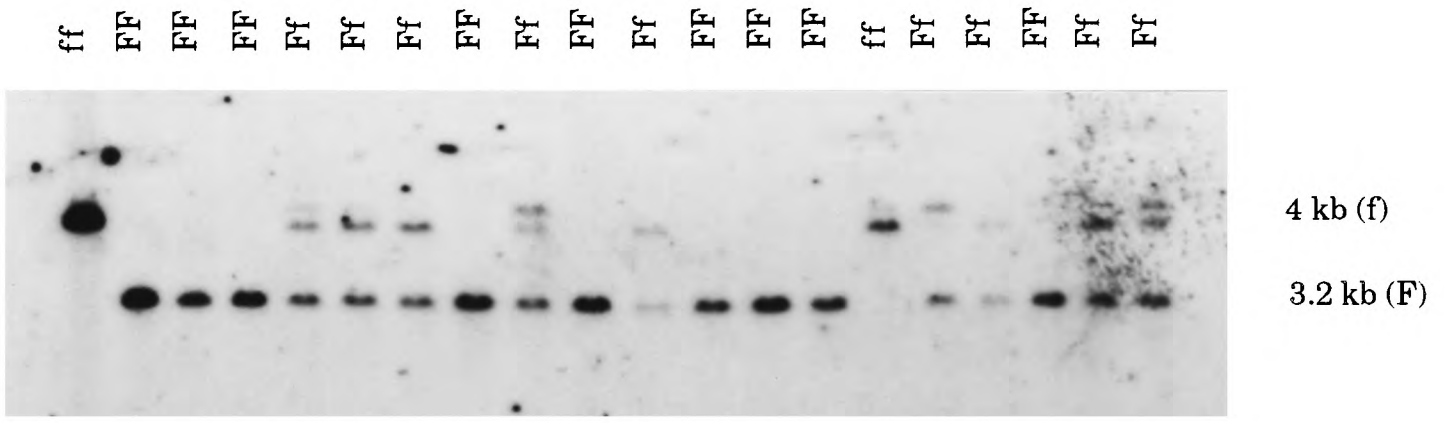


Figure 5.28 Investigation into the heterozygosity of the EcoRI polymorphism detected by Alf1. Southern blot of EcoRI/EcoRV digested DNA from 20 unrelated females. Two individuals are homozygous for the 4 kb allele (ff), nine are homozygous for the 3.2 kb allele (F) and nine individuals are heterozygotes. The presence of a 4.2 kb band in some tracks may be due to partial digestion of the EcoRV site distal to Alf1 but this is somewhat unclear.

in close agreement with the observed female heterozygote frequency (45%). This level of polymorphism is suitable for clinical use. Inheritance of alleles of this polymorphism was used to predict the carrier status of three sisters with respect to DMD. Figure 5.29 illustrates the pedigree and the result of probing EcoRI/PstI digested DNA from the living family members with Alf1. The allele sizes detected by this double-digest are 4.3 and 3.5 kb. The 4.3 kb allele was detected in the three living unaffected brothers, which indicates that their mother's 3.5 kb allele is linked to the disease expression (excluding the possibility of germline mosaicism; Bakker et al., 1989). One sister is homozygous for the 3.5 kb allele and has therefore inherited her mother's high risk 3.5 kb allele and is likely to be a carrier. Her genotype also indicates that the dead father must have had the 3.5 kb allele. This in turn indicates that her two heterozygous sisters have inherited their mother's low risk 4.3 kb allele and are unlikely to be carriers.

Polymorphic restriction sites can be detected by PCR. Several such polymorphisms have been described within the DMD gene (Roberts et al., 1989a; Roberts et al., 1990). The advantage of detecting polymorphisms by PCR is that the analysis is rapid and that only a small quantity of template DNA is required. Primers are designed which amplify a segment of DNA including the polymorphic site and the PCR product is digested using the restriction enzyme which recognises the site. The DNA is then run on an agarose or polyacrylamide gel to identify the alleles.

DNA in the vicinity of the polymorphic EcoRI site detected by Alf1 is currently being sequenced so that primers can be designed to amplify a DNA segment of the region.

It is interesting to note that the endpoint of the J-MD deletion lies in the middle of the polymorphic 9/8.2 kb EcoRI fragments (see figure 5.11), and therefore that this polymorphism might very easily have been detected several years ago and become widely used in clinical diagnosis (Monaco et al., 1987). The clone J-MD1 which was isolated after walking from this cloned deletion endpoint flanks the polymorphic

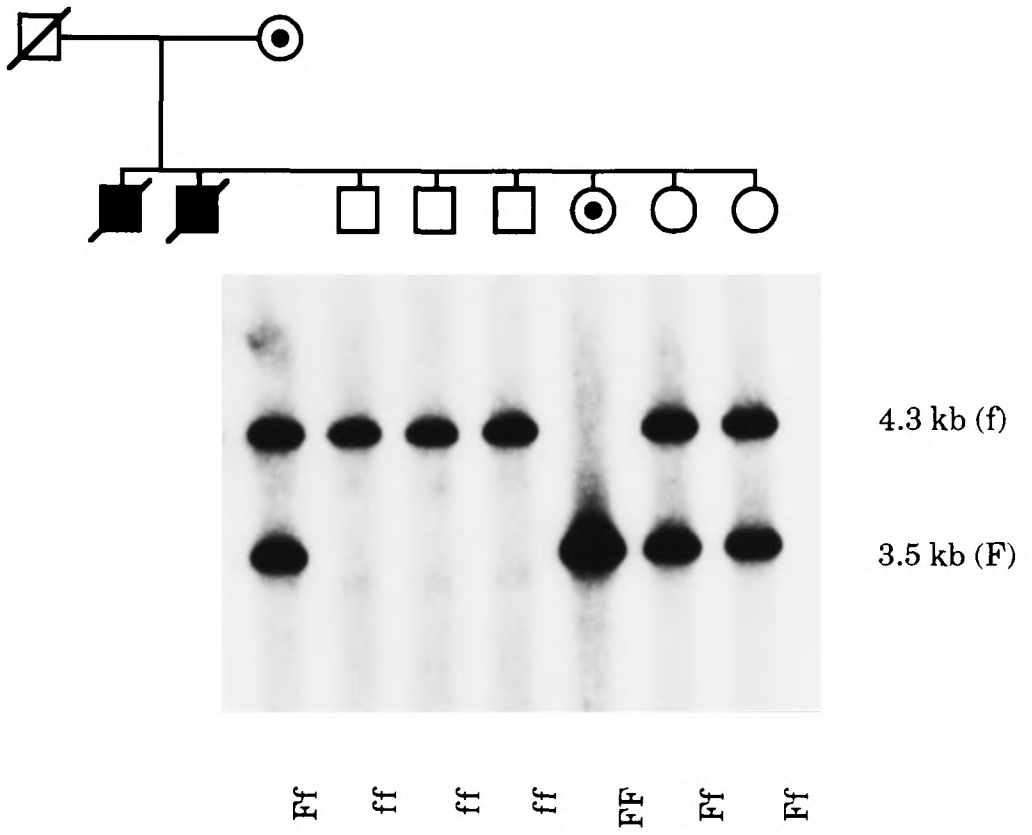


Figure 5.29 Prediction of carrier status in a family with DMD by linkage analysis using Alf1. Explanation in text (section 5.2.8).

EcoRI fragment on its distal side, while another clone, J-MD-2, lies approximately 8 kb proximal to the polymorphic site (see figure 5.11).

5.3 Discussion

This chapter has described the characterisation of the breakpoint regions of the X;1 translocation and of the translocation junction fragments. As already explained in the introduction (section 5.1), this information was required in order to perform inverse PCR to clone the translocation junctions.

The restriction map of the translocation generated by this analysis is presented in figure 5.24 and indicates that no deletion, duplication or insertion larger than about 100 bp is associated with the translocation in the vicinity of the breakpoints. This was later confirmed by sequencing the translocation junctions (Chapter 7). The rearrangement cloned is a two-break event. This result is inconsistent with the original cytogenetic interpretation of the translocation (Lindenbaum et al., 1979; Boyd and Buckle, 1986; Clarke, unpublished observations) which predicted a three-break rearrangement by involving a paracentric inversion between bands Xp11.4 and Xp21.2. This interpretation was suggested as it involved the smallest number of breakpoints to explain the observed banding pattern (see section 5.1, figure 5.1). The results of cloning the X;1 translocation do not preclude the presence of an additional rearrangement, as long as it involves at least two further breakpoints. Four physical structures for possible rearrangement structures are illustrated in figure 5.30. The first (b) is a two breakpoint translocation and is consistent with the cloning and restriction mapping data presented in this chapter although it does not accomodate the cytogenetic observations which have been described. The second (c) is the structure suggested by Lindenbaum et al. (1979) and can now be excluded as it involves three breakpoints. On presenting the results of cloning the X;1 translocation to the cytogeneticists who first characterised the rearrangement, the cytogenetic evidence of an abnormal der(X) banding pattern was judged to be strong enough to involve four

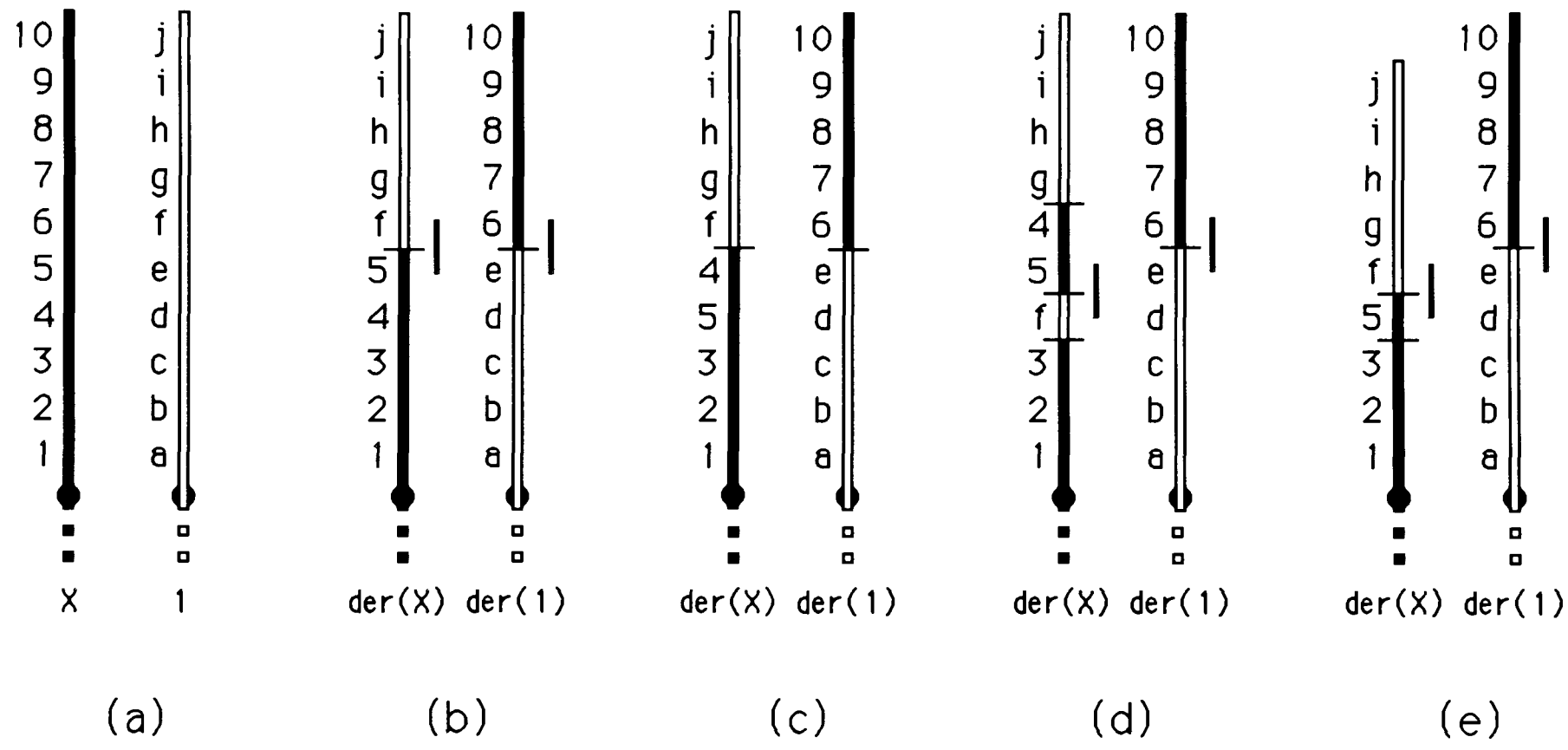


Figure 5.30 Diagram of possible rearrangement structures which may represent the X;1 translocation. Dark lines represent X chromosome DNA and pale lines represent chromosome 1 DNA. Numbers indicate the order of DNA sequence on an intact X chromosome and letters indicate that on an intact chromosome 1. Bars in (b), (c) and (d) indicate the junction fragments which have been cloned in this study. A full explanation is given in the text (section 5.3).

breakpoints in the rearrangement (Clarke, personal communication). The third and fourth structures in figure 5.30 are four breakpoint rearrangements which could explain the reported cytogenetic observations and which are consistent with the cloning and restriction mapping data presented in this chapter. One (d) involves a paracentric inversion occurring after the translocation event and around the translocation breakpoint. The other (e) involves a deletion proximal to the translocation site. There are two ways in which structures (b) and (d) could be distinguished, however they could become prolonged studies. One is by searching for the junction fragments predicted in the four-break structure (3-f and 4-g; figure 5.30). This would be best done by using PFGE methods but would be dependent on a full length long range map of Xp21 to Xp11.3. Alternatively, in situ hybridisation could be used to determine the position of probes from around Xp11.4 (equivalent to segment '4', figure 5.22) relative to the Xp21.1 band on the patient's der(X) chromosome. This experiment would probably be the better way to distinguish the two possible structures. Identification of a deletion such as in structure (e) could be performed by hybridisation of probes proximal to Xp21 to Southern blots of DNA from the der(X) retaining somatic cell hybrid WAG 8-1. Alternatively, PFGE could be used to identify junction fragments. If a deletion as in structure (e) exists, then it is unlikely to include loci of clinical importance since the X;1 translocation patient did not manifest clinical symptoms apart from DMD and mild short stature (Lindenbaum, personal communication). There is, however, a growing list of genes which escape X inactivation, and these might be included in the putative deletion without deleterious effects (Davies, 1991).

A further possibility to explain the observed banding of the t(X;1) translocation chromosomes is that a simple two-breakpoint translocation has occurred, but that the breakpoints have altered the staining/banding nature of the chromatin in the vicinity of the rearrangement (Raimondi et al. 1983).

In this chapter, the characterisation of a DMD-associated X;autosome translocation was described by restriction mapping of DNA segments around the breakpoint sites

and of junction fragments cloned from the der(X) and der(1) chromosomes. Southern blot analysis using probes from these regions was also described. The next chapter explains how these DNA segments were cloned.

5.4 Acknowledgements

I am grateful to Chris Porter for collaboration on the subcloning and restriction mapping of the cosmid XJC-5 and to Sheila Holt for collaboration on the in situ localisation of PPii.

Chapter 6 Inverse PCR amplification and cloning of the t(X;1) breakpoint regions and junction fragments

6.1 Introduction

This chapter describes the application of the polymerase chain reaction in the cloning of the junction fragments from the t(X;1) and of the corresponding regions from the normal chromosomes. The use of this method of cloning requires DNA sequence information in the vicinity of the breakpoints so that oligonucleotide primers can be designed. In this chapter, relevant stretches of DNA sequence from which primers were designed are illustrated, their locations with respect to the restriction map are indicated, and cross references are included to the full sequence presented in Chapter 7.

6.1.1 Strategies for cloning translocation breakpoints

Inverse PCR was adopted as a cloning strategy in preference to more orthodox strategies which involve DNA library construction and identification and isolation of the desired clones. A comprehensive practical guide to library construction is provided by Sambrook et al. (1989). Here, the features of DNA libraries relevant to the choice of strategy for cloning translocation breakpoints are described. Libraries can be constructed from complete genomic DNA or alternatively enrichment for the desired fragment can be achieved by digesting the DNA using a restriction enzyme and using a size selected fraction. This was the approach used by Bodrug et al. (1987) in cloning the breakpoints and junctions of an X;21 translocation. An alternative enrichment procedure is to use somatic cell hybrid DNA as a source for library construction. The best hybrids for use are those containing little human material, for instance hybrids containing a single human chromosome or radiation-fragmented hybrids which only contain a human chromosome fragment (Burk et al., 1985; Benham et al. 1989). A choice of cloning vectors is available for library construction. This includes

plasmids, bacteriophages and cosmids. They vary in the size of foreign DNA fragment which can be efficiently cloned. For plasmids this is a few kb, for bacteriophages this is up to approximately 20 kb and for cosmids this is approximately 50 kb (Sambrook et al., 1989). In order to clone translocation breakpoints and junctions using a library construction strategy, it would be necessary to construct a library specifically for this purpose, as only DNA from the translocation patient contains the junction fragments. Although such a library may be a useful resource for other future experiments, its construction could prove time consuming for someone with little experience in the relevant techniques. The advantage of inverse PCR on the other hand is that it is fast and relatively straightforward. There are, however, disadvantages in its use in a cloning/sequencing strategy. One disadvantage is the practical limit that PCR technology places on the size of DNA fragment which can be efficiently amplified. Although amplification of fragments up to 10 kb in length has been reported, shorter fragments amplify more readily (Jeffreys et al.,1988). This limits the size of the DNA fragment which can subsequently be cloned. The application of inverse PCR involves the preparation of specially modified PCR templates. This step might lower further the efficiency of the reaction and therefore reduce the length of fragment which can be amplified and cloned. A small cloned junction fragment or breakpoint region might be disadvantageous for several reasons. Firstly, a smaller fragment is directly available for sequencing and restriction mapping, although standard chromosome walking methods could be used to enlarge the cloned region. Secondly, if one is unlucky, it might not be possible to identify a single copy DNA probe within the fragment. This could hamper Southern analysis if competitive hybridisation methods are not found to be effective (Sealey et al., 1985). This could also prevent enlargement of the cloned region by chromosome walking which relies on the identification of single copy sequences. PCR primers designed from within repetitive DNA sequences may amplify several products from several loci. Another disadvantage of using the inverse PCR approach to clone translocation breakpoint

regions and junction fragments is that cloned fragments may include sequence errors produced by misincorporation of nucleotides by Taq polymerase, errors which would not be produced by other cloning methods (Karlovsky, 1990). PCR errors which I have identified by sequencing the X;1 translocation and how the problem may be overcome are discussed in Chapter 7, section 7.2.4.

6.1.2 Inverse PCR

The polymerase chain reaction (PCR) is one of the most recent methods of analysis in molecular biology. Saiki et al. (1985) first described how the thermostability of DNA polymerase isolated from Thermus aquaticus could be exploited to allow automated repeated cycles of DNA denaturation and DNA replication. This cycling process causes an exponential increase in the production of a DNA segment designated by artificially synthesised oligonucleotide primers. The main advantages of PCR techniques are in their ease of use (due to semi-automation), the speed of reaction (one can amplify and visualise a specified DNA segment in a day), and in its flexibility to suit specific needs (see Innis et al., eds., 1990). Examples of the diversity of PCR applications include multiplex PCR for simultaneous analysis of several loci, for example within the Duchenne muscular dystrophy gene (Chamberlain et al., 1988; Beggs et al., 1990; Abbs et al., 1991); asymmetric PCR for generating single-stranded DNA sequencing templates (Gyllensten and Erlich, 1988); amplification of microdissected chromosome bands in DNA cloning experiments (Han et al., 1991; Kao et al., 1991); Alu-mediated PCR for specific amplification of human DNA fragments from a complex DNA source (Nelson et al., 1989); and inverse PCR as described in the work presented in this chapter in the context of the amplification of translocation junction fragments.

Inverse PCR was independently developed by three groups (Ochman et al., 1988; Triglia et al., 1988; Silver and Keerikatte, 1989) and is essentially a method of taking a short chromosome walk. The size of walk is limited by the length of DNA which can be

efficiently amplified (see section 6.1.1). Inverse PCR allows the amplification of an uncloned segment of DNA which is adjacent to cloned DNA. There are two prerequisites for the method's use. Firstly, the sequence of the cloned DNA must be determined so that oligonucleotide primers may be designed. Secondly, determination of the positions of restriction enzyme sites within the uncloned DNA segment is desirable, so that an appropriate DNA fragment can be chosen as a template for PCR amplification. The strategy is illustrated in figure 6.1 (showing the der(1) fragment amplification). A pair of oligonucleotides is designed whose 3' ends will prime DNA replication in opposite directions. Instead of the directions of replication being towards the other primer, they are the reverse (hence inverse PCR). The DNA template is cut with a restriction enzyme. It is preferable that the cutting sites of this restriction enzyme have been already determined so that the expected fragment size of the amplification product is known. Restricted DNA is circularised by treatment with T4 DNA ligase at a DNA concentration low enough to favour circularisation over concatamerisation (see the following section 6.1.3 for theoretical calculations used in this study). The circularised DNA is used as a template in a PCR using the designed primers. DNA between the 3' ends of the oligonucleotide primers which will be amplified in such an experiment now includes the segment of uncloned DNA. One or both ends of the amplified DNA product can then be trimmed with a restriction enzyme to facilitate its cloning (see figure 6.1).

Two modifications to the basic strategy described above were employed in some experiments. The first of these was the linearisation of the circular DNA template, which is an unusual PCR template, by using a restriction enzyme which has a cutting site between the 5' ends of the primers (for example see the position of the NdeI restriction site in figure 6.1). Such restriction sites were identified by computer search of the sequence using the program 'nip' (Staden, 1984). Restriction enzymes with a 6 bp recognition site were used since they were unlikely to have a second recognition site within the circular molecule. The other modification was to perform a second PCR

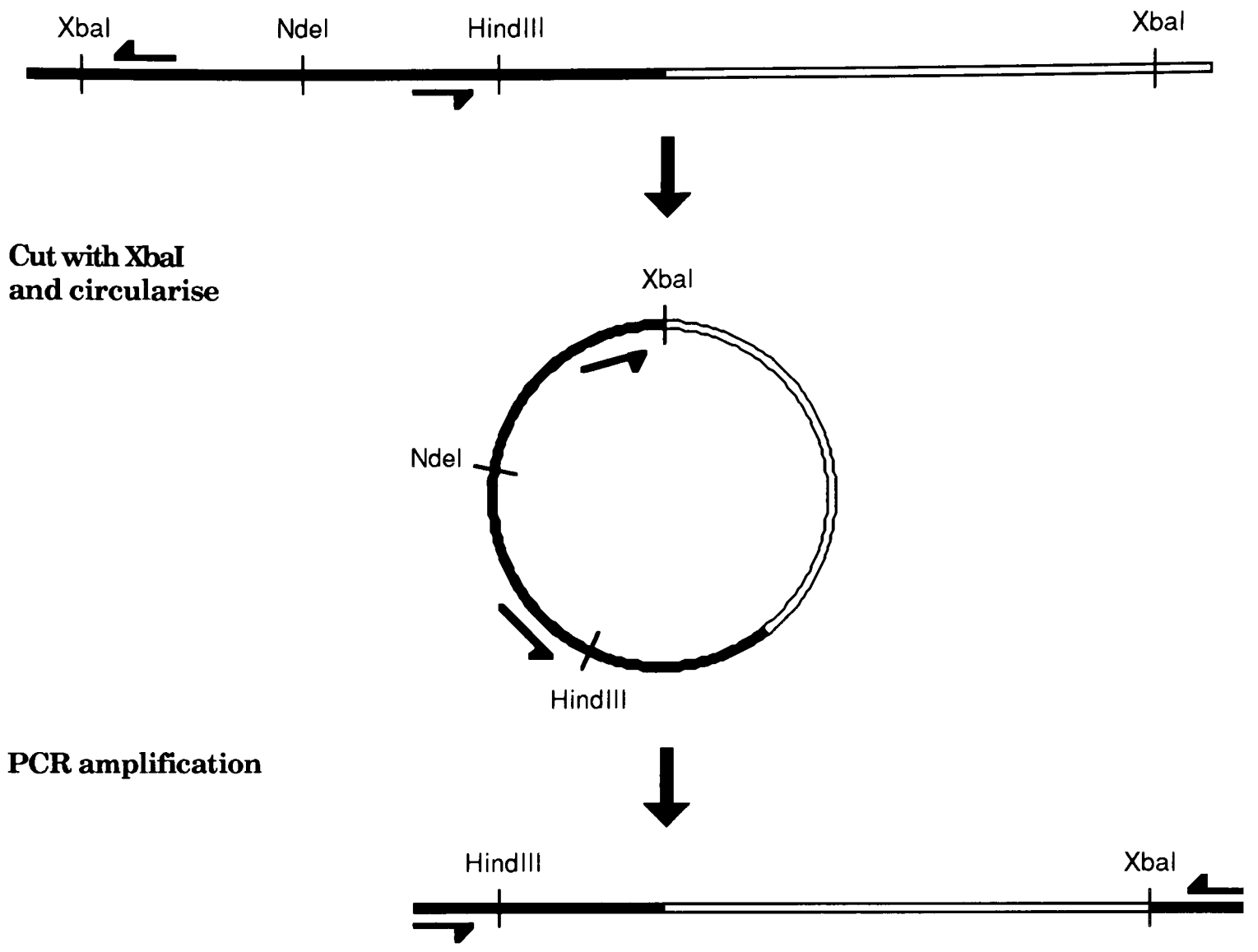


Figure 6.1 Strategy of inverse PCR, illustrated here for the der(1) XbaI junction fragment (not to scale). X chromosome DNA (cloned) is represented by dark lines and chromosome 1 DNA (uncloned) by pale lines. Genomic DNA is cut with a restriction enzyme (XbaI) and ligated to produce circular molecules. PCR amplification using this circularised DNA as template results in the amplification of a fragment including the uncloned DNA segment (in this case from chromosome 1). The position of a restriction site (NdeI) which may be cut to linearise the circular template prior to amplification is indicated.

using the gel-eluted initial PCR product as template and using nested primers. The success of these modifications is discussed in section 6.3.

6.1.3. Circularisation of DNA templates for inverse PCR

The application of inverse PCR is dependent upon successful circularisation of target restriction fragments by treatment with DNA ligase. Ligation reactions were performed using suitably restricted total human genomic DNA as substrate. In order that target restriction fragments circularised, conditions of reaction were chosen so that one end of the molecule was more likely to ligate with its other end than with the end of another DNA molecule. Circularisation is favoured when the concentration of DNA molecules in solution is low and when the restriction fragment is short. Two parameters were used to determine the required conditions (Collins and Weissman, 1984; Sambrook et al., 1989). These are *i*, the concentration of all molecule ends in solution, and *j*, the effective concentration of one end of a molecule in the vicinity of the other. The value of *i* can be varied, but that of *j* is constant for a molecule of given length.

Polyethylene glycol (PEG) is often included in the buffers supplied by manufacturers for use with T4 ligase. Its purpose is to increase the effective concentration of DNA molecule ends in solution and thereby decrease the time required for a ligation reaction. PEG should have a proportionally equal effect on the values of *i* and *j* and therefore should not interfere with theoretical considerations of circularisation. Its use is advantageous as it increases the efficiency of the ligation reaction which for a circularisation reaction is performed at an unusually low DNA concentration.

When *i* is equal to *j*, 50% of molecules of the given length will circularise. In order that more than 50% molecules circularise, *i* must be lower in value than *j*. The percentage of molecules that circularise is predicted by the equation (Collins and

Weissman, 1984):

$$\% \text{ circles} = [j / (i + j)] \times 100$$

It follows that:

$$i = j\,[(100 / \% \text{ circles}) - 1]$$

and that for 90% circles:

$$i = j / 9$$

The value of j for a molecule of length 'a' can be determined from the equation (Sambrook et al., 1989; where MW is the molecular weight, and j_λ is 3.6×10^{11} ends/ml):

$$j_a = j_\lambda\,[MW_\lambda / MW_a]^{3/2}.$$

In order to convert values of i from ends/ml into μg/μl, it is necessary to consider the average size of restriction fragment in a particular enzyme digest. This was estimated by observing an ethidium bromide-stained gel of digested DNA, judging the size of the most intensely staining portion of the smear and halving this figure (Sambrook et al., 1989).

The following equivalent values were also assumed:

1 kb ≈ 660 kdal

1 kdal ≈ 1.7×10^{-15} μg double-stranded DNA

6.2. Results

6.2.1. Amplification and cloning of the der(1) junction fragment and of the normal X chromosome breakpoint region

Through the analysis of der(1) junction fragment sizes using the X chromosome probe Alf1, a 2.7 kb XbaI junction fragment was identified as a suitable target for the application of inverse PCR (Chapter 5; see figures 5.8 and 5.11). The conditions for circularisation of this DNA fragment prior to inverse PCR were calculated as follows. The value of j for a 2.7 kb DNA fragment is 2.8×10^{13} ends/ml. The value of i for 90%

circularisation (*j*/9) is 3.1 x 10^{12} ends/ml. This is equivalent to approximately 1 µg / 150 µl assuming that the average length of DNA fragment generated by XbaI digestion of human genomic DNA is 4 kb. The template DNA for this inverse PCR was prepared by ligating 1 µg XbaI digested genomic DNA from the X;1 translocation patient in a 200 µl volume. Under these conditions, over 90% of 2.7 kb fragments would be expected to circularise.

The 2.7 kb XbaI junction fragment includes the probe Alf1. Alf1 was sequenced and the primers XO11 and XO12 and the nested primers XO17 and XO18 were designed for inverse PCR from this sequence (figure 6.2; full sequencing described in Chapter 7, section 7.2.1.1). The locations of the primers with respect to the sequence are illustrated in figure 6.2. The expected length of DNA amplified from the der(1) junction using these primers is approximately 2.5 kb. This length is suitable for efficient PCR amplification and approximately 2 kb of this product would be expected to be the uncloned region beyond the translocation breakpoint which is of presumed chromosome 1 origin (figure 6.3). Both modifications to the basic inverse PCR method outlined in section 6.1.2 were tested. Firstly, the linearisation of the template DNA with NdeI was tried. This enzyme was predicted from Alf1 sequence to cut between the oligonucleotide primers XO11 and XO12 (figure 6.2). Secondly, the gel-purified product of this PCR amplification was used in a secondary amplification using the nested primer pair XO17 and XO18 (figure 6.2).

The products of the three amplification experiments are shown in figure 6.4. The experimental conditions used for amplification were denaturation at 94^{0}C for 1 minute, primer annealing at 58^{0}C for 40 seconds, and primer extension at 72^{0}C for 4 minutes for a total of 30 cycles. These conditions had been determined by performing several amplifications using various conditions until adequate yield of product with low non-specific amplification was obtained. Two principal DNA fragments were amplified in these reactions. The sizes of these fragments conform to the expected products amplified from the normal X chromosome and der(1) XbaI fragments, which had been

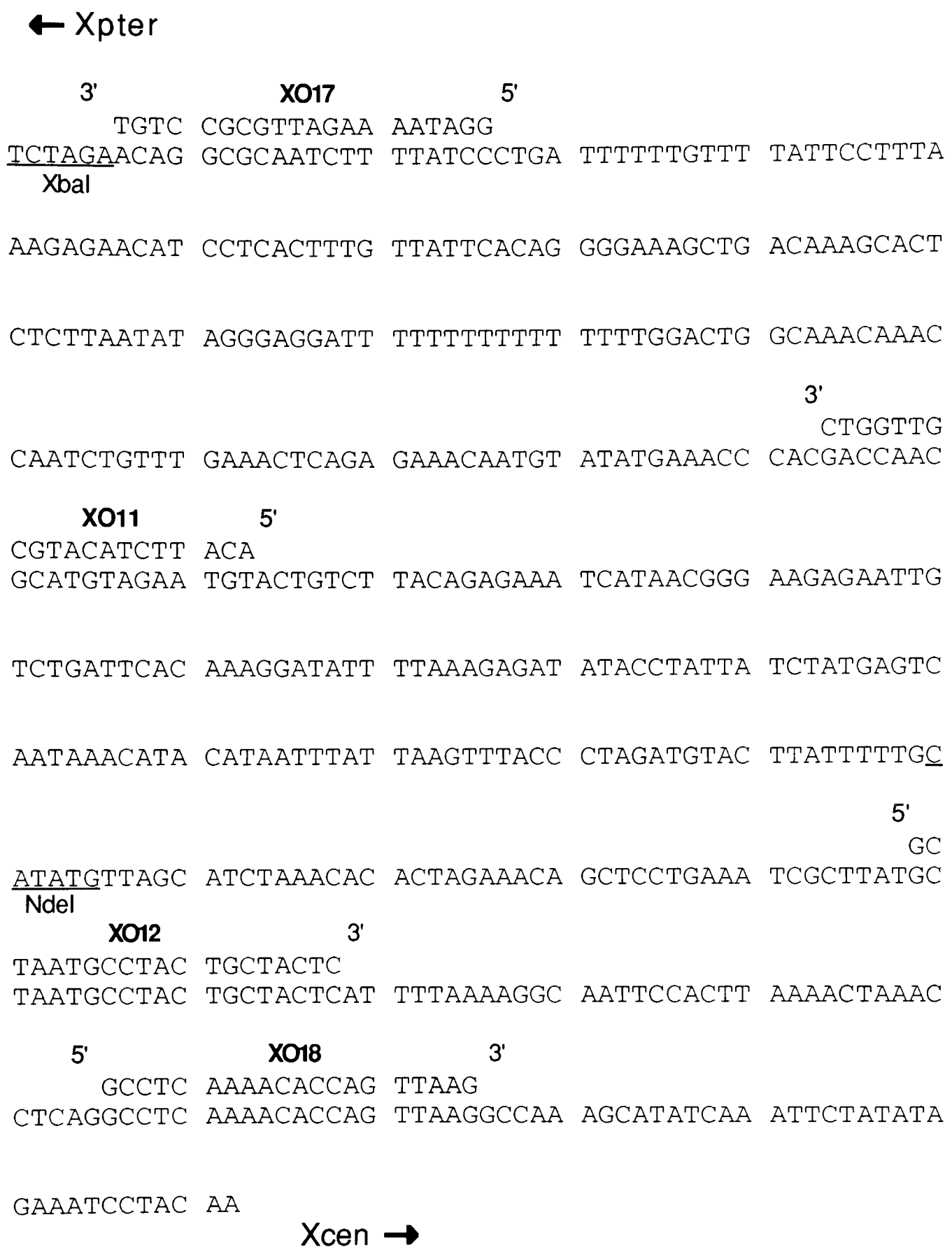


Figure 6.2 DNA sequence from within Alf1 showing the positions of the primers XO11 and XO12 and of the nested primers XO17 and XO18 used in the inverse PCR amplification of the der(1) junction fragment and of the normal X chromosome breakpoint region. The position of a NdeI restriction site is indicated which was cut in some experiments to linearise the inverse PCR template prior to amplification.

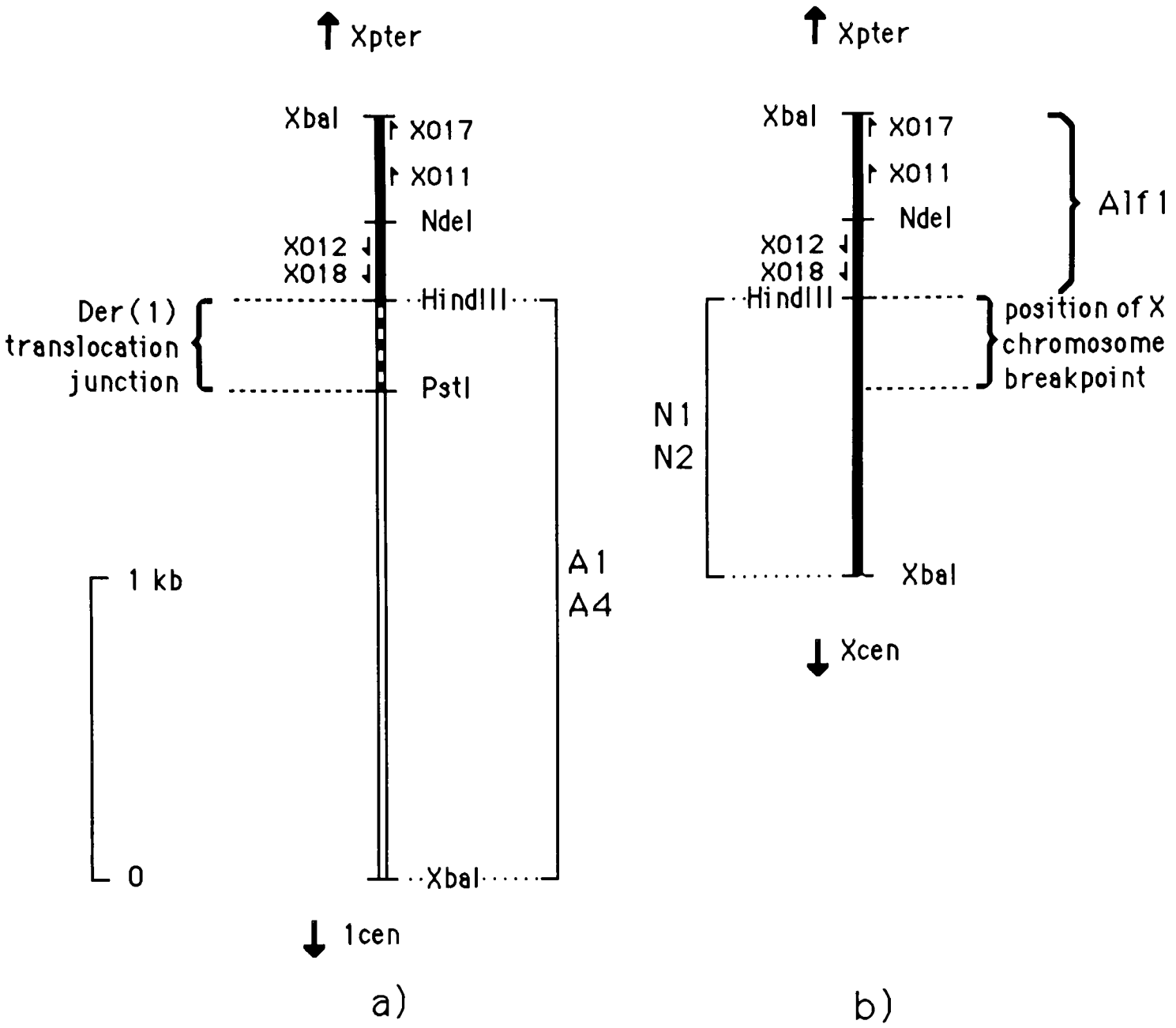


Figure 6.3 Details of a) the der(1) XbaI junction fragment and b) the XbaI restriction fragment spanning the translocation breakpoint on the normal X chromosome, which were circularised to be used as templates for inverse PCRs using the primers XO11 and XO12. The positions of the nested primers XO17 and XO18 are also indicated as is the site of a NdeI restriction site which was cut in some experiments to linearise the PCR template prior to amplification. The locations of DNA segments cloned following inverse PCR amplification are shown. A1 and A4 span the der(1) translocation junction, and N1 and N2 span the normal X chromosome breakpoint region.

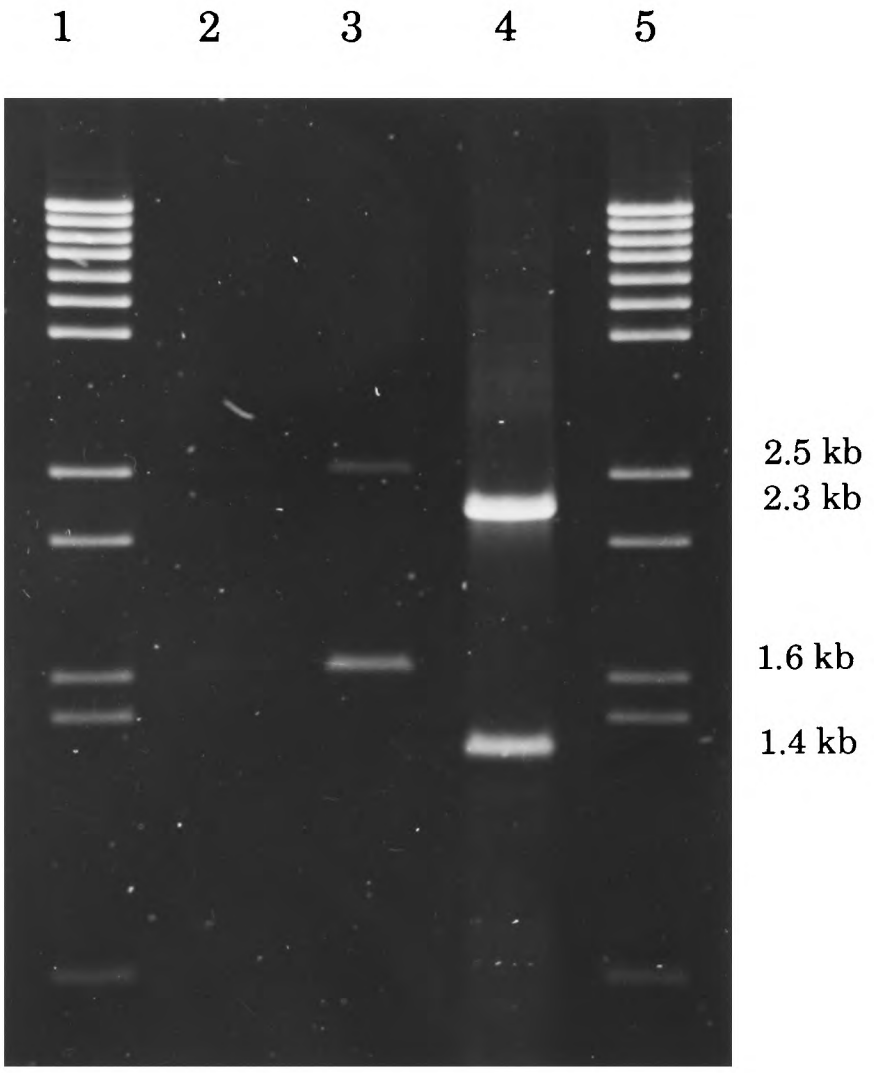


Figure 6.4 Ethidium bromide-stained products of three PCRs. Tracks 1 and 5, BstEII digested λ. The products of a reaction using the primers XO11 and XO12 with circular template derived from the X;1 translocation are shown in track 2 (barely visible). Track 3 illustrates the products from the same primers but using a linearised template. The gel-purified 2.5 kb product from this reaction was used as template in a further PCR using the nested primers XO17 and XO18 (track 4). Further explanation is provided in the text. The sizes of PCR product are indicated to the side of the figure.

determined by Southern blot analysis using probe Alf1 (see Chapter 5, figures 5.8 and 5.11). Major improvements in yield (more than 10 fold) were obtained by employing each of the two modifications to the basic strategy described above. The yield from the basic strategy was barely visible when run on an agarose gel and stained with ethidium bromide (track 2, figure 6.4), although readily detectable when a similar gel was blotted and probed with Alf1 (not shown). The amplification products of the NdeI linearised DNA template are illustrated in track 3. The yield of the normal X chromosome fragment (1.6 kb) was approximately 40 ng per 100 µl reaction and that of the der(1) junction fragment (2.5 kb) was approximately 20 ng per 100 µl reaction. The junction fragment was gel-purified and 5 ng was used as template in a PCR using the nested primers XO17 and XO18. The products of amplification are illustrated in track 4. The yield of der(1) junction fragment was approximately 1 µg per 100 µl reaction. The size of fragment is approximately 2.3 kb. The size difference between this fragment and that produced using the primers XO11 and XO12 is consistent with the positions of the primers within the DNA sequence (figure 6.2). The presence of some normal X chromosome amplified product (1.4 kb) in track 4 indicates contamination of the gel-eluted 2.5 kb PCR template with some of the 1.6 kb band illustrated in track 3.

The amplification of the normal X fragment did not prove to be a serious handicap to the amplification of the der(1) despite its smaller size. If difficulty had been experienced in this respect, then DNA from the der(1) retaining hybrid WAG 21R-11, which does not contain a normal X chromosome (see Chapter 2, table 2.1), could have been used as the inverse PCR template in place of cell line DNA from the translocation patient. However, the estimated yield of PCR amplified der(1) junction fragment after the sequential use of both modifications was approximately 1 µg per 100µl reaction, representing a 10^6 fold amplification of the initial template.

The 2.3 kb der(1) PCR product described above was trimmed using HindIII and XbaI and cloned into pUC18 (see figure 6.3). In the cloning experiment, six out of ten picked clones contained the expected 2 kb insert (see figure 6.3; the other four clones

appeared not to contain inserts, not shown). Two clones, A1 and A4, were fully characterised. The inserts of these plasmids were observed to contain recognition sites for the restriction enzymes PstI, PvuII, and AvaI but not for BamHI, EcoRV, EcoRI, KpnI, SacI and MluI (not shown). These observations were consistent with expectations from Southern blot data derived by probing WLS DNA with Alf1 (see Chapter 5, section 5.2.4). The restriction mapping of the PstI, PvuII and AvaI sites within the cloned der(1) junction fragment, A1, was described in Chapter 5, section 5.2.5. A 430 bp PstI/AvaI DNA fragment, AP430, presumed to be from chromosome 1, was gel-eluted from the insert of A1 and used as a probe in Southern analysis of DNA from the t(X;1) patient and from a normal individual (described in Chapter 5, section 5.2.5). This analysis was compared with the restriction map of the der(1) junction and used to construct a preliminary restriction map of the normal chromosome 1 region around the translocation breakpoint [figure 6.5(a)].

The amplified DNA product which represented the normal X chromosome from the patient and which was produced in the above experiment was also cloned. The target for this PCR amplification is illustrated in figure 6.3(b). The 1.4 kb fragment illustrated in track 4 of figure 6.4 was gel-eluted, trimmed with HindIII and XbaI and cloned into pUC18. Four clones were tested and all contained the expected sized insert (1.1 kb; see figure 6.3). Two of these, named N1 and N2, were further analysed. The inserts did not contain the recognition sites for BamHI, EcoRV, EcoRI, PstI or PvuII (not shown), as expected from restriction mapping data of the normal X chromosome region (Chapter 5, section 5.2.3). However, a single HaeIII site was detected approximately 400 bp from one end of the insert (not shown). The orientation of this site with respect to the 1.1 kb fragment was not determined until the XbaI/HaeIII 400 bp and HindIII/HaeIII 700 bp fragments were later subcloned from N1 and N2 and sequenced (see Chapter 7, section 7.2.1.1). This revealed that the 700 bp fragment is distal to the 400 bp fragment and contains the X chromosome breakpoint.

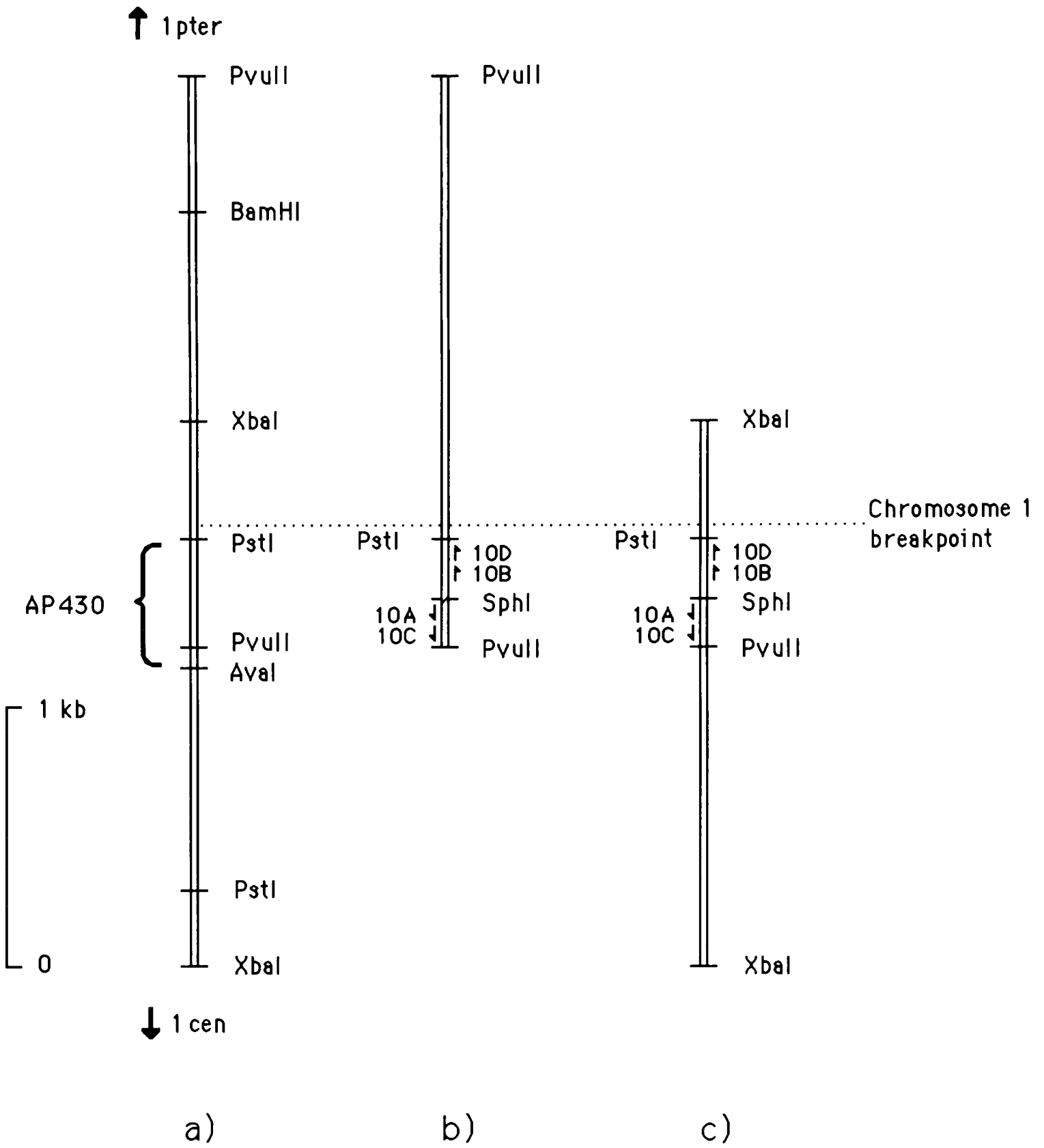


Figure 6.5 a) Preliminary restriction map of the chromosome 1 breakpoint region. Data proximal to the breakpoint position was obtained by restriction enzyme digestion of the der(1) clone A1, and data distal to the breakpoint position was determined by Southern analysis using the probe AP430. b) PvuII and c) XbaI restriction fragments spanning the chromosome 1 breakoint which were circularised and used as templates in inverse PCR reactions designed to allow the amplification and cloning of the normal chromosome 1 breakpoint region and DNA segments distal to the breakpoint position.

6.2.2 Amplification and cloning of the normal chromosome 1 breakpoint region

Examination of the preliminary restriction map of the genomic region around the chromosome 1 breakpoint obtained by probing genomic DNA with AP430 [figure 6.5(a)] indicated two suitable restriction fragments for circularisation and use as templates in inverse PCR experiments. These are the PvuII and XbaI restriction fragments. Although both fragments are approximately 2 kb in length, the PvuII fragment is the preferred template since it permits the amplification of a larger segment of chromosome 1 DNA distal to the chromosome breakpoint, approximately 1.5 kb compared to approximately 0.4 kb from the XbaI fragment (figure 6.5). However, PvuII generates blunt-ended DNA fragments which are known to ligate less efficiently than fragments with compatible single-stranded tails (Sambrook et al., 1989). Therefore, in case inverse PCR using the PvuII fragment as template was unsuccessful, the XbaI fragment was also used. The PvuII and XbaI restriction fragments both include a PstI/PvuII segment of approximately 410 bp which is the distal portion of the probe AP430 and lies just proximal to the chromosome 1 breakpoint (see figure 6.5). The PstI/PvuII segment was sequenced (described in Chapter 7, section 7.2.1.2). A pair of primers, 1OA and 1OB, and a nested pair of primers, 1OC and 1OD, were designed from this sequence (figures 6.5 and 6.6). An SphI restriction site was identified from sequence data between the primers 1OA and 1OB (figs 6.5 and 6.6). The nested primers and SphI site enabled the same modifications as used in the amplification of the der(1) junction fragment to be used here (see sections 6.1.2 and 6.2.1).

The conditions for circularisation of the PvuII and XbaI DNA fragments prior to inverse PCR were calculated as follows. Since the PvuII and XbaI fragments are both approximately 2 kb in length, the value of j is the same for each. This is approximately 4.7×10^{13} ends/ml. Therefore the value of i for 90% circularisation ($j/9$) is approximately 5.2×10^{12} ends/ml. This value for XbaI digested DNA is equivalent to approximately 1 µg / 90 µl assuming that the average XbaI fragment of human genomic DNA is 4 kb. Assuming a size of 2 kb for the average PvuII fragment in human

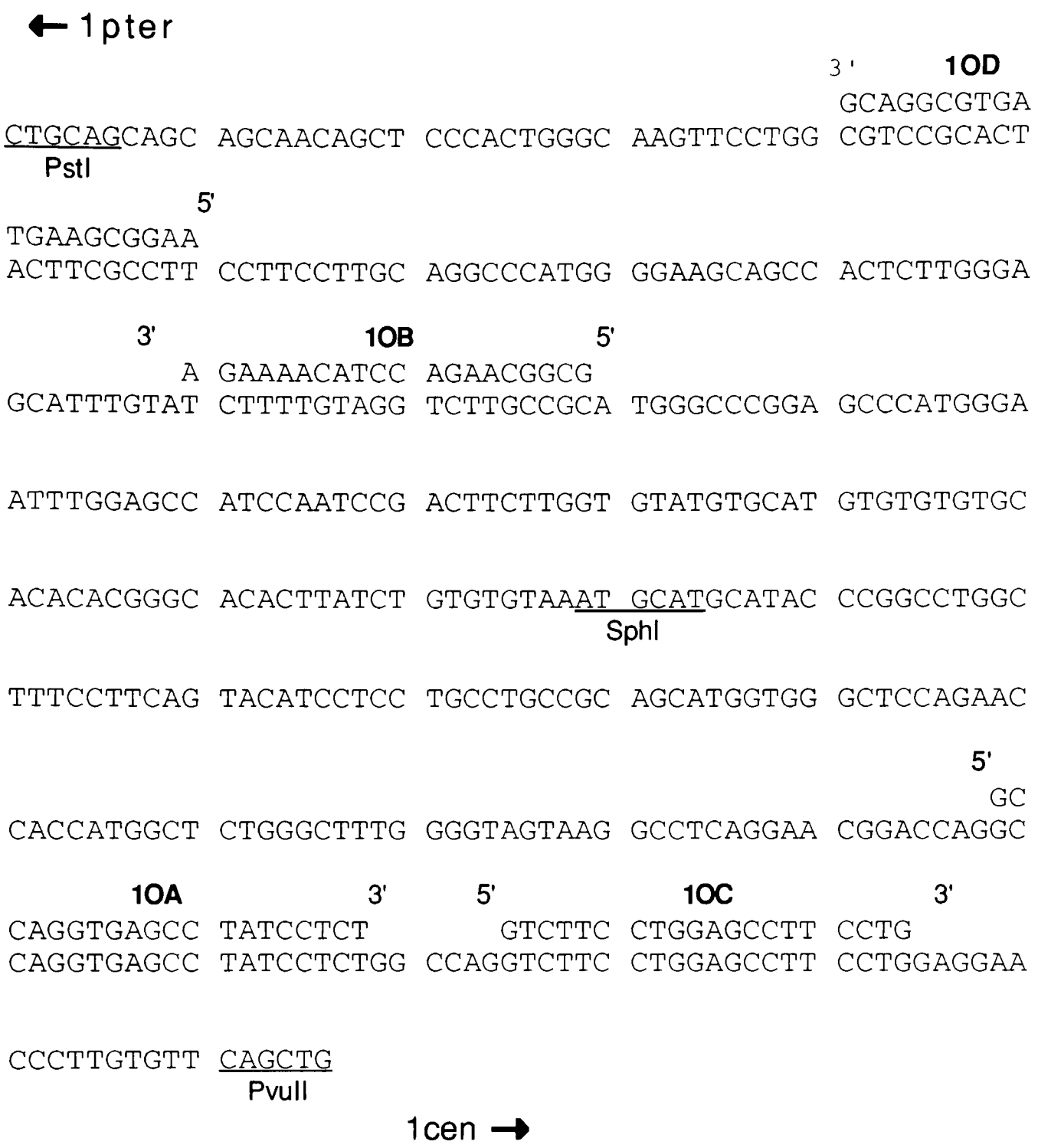


Figure 6.6 DNA sequence from within AP430 showing the positions and sequences of the primers 1OA and 1OB and of the nested primers 1OC and 1OD used in the inverse PCR amplification of the normal chromosome 1 breakpoint region. An SphI restriction site is indicated which was cut in some experiments to linearise the inverse PCR template prior to amplification.

genomic DNA, this value approximates to 1 μg / 175 μl. Both ligation reactions were performed using 1 μg restricted DNA in a total volume of 200 μl, therefore over 90% of target XbaI and PvuII restriction fragments would be expected to circularise. The source of DNA for these reactions was placental DNA from an anonymous normal female rather than from the t(X;1) patient, since a normal individual has two normal chromosome 1s and therefore has twice as many templates for PCR amplification.

The PCR amplification conditions used with these primers were DNA denaturation at 94^{0}C for 1 minute, primer annealing at 61^{0}C for 30 seconds, and primer extension at 72^{0}C for 3 minutes for a total of 30 cycles. These conditions are those which were found to produce an adequate yield of expected product and a low background of non-specific amplification, after testing several PCRs using different conditions. The products of several different experiments were run on an agarose gel and stained with ethidium bromide (figure 6.7). Track 2 contains the products of amplification from the XbaI cut and circularised template using the primers 1OA and 1OB. A band of approximately 1.9 kb is visible, which is the expected size for the desired amplification product (see figure 6.5). The yield was around 100 ng in a 100 μl reaction. The gel-eluted product of this reaction, named Pβ, was later used to screen bacterial colonies from different cloning experiments by colony hybridisation (see below and section 6.2.3). Pβ was also used as template in a further amplification using the nested primers 1OC and 1OD, and the yield of this PCR was over 2 μg of approximately 1.8 kb product in a 100μl reaction (>10^6 fold amplification of original template; tracks 4 and 5, figure 6.7). The size of PCR product using the nested primers is expected to be 95 bp smaller than using the primers 1OA and 1OB (see figure 6.6). The yield of amplified PCR product from the PvuII cut and circularised template was lower. Approximately 50 ng of a 1.8 kb fragment was amplified in a 100 μl reaction using the primers 1OA and 1OB (10^5 fold amplification; track 3, figure 6.7). This is the expected size (see figure 6.5). The fragment was gel-eluted and 5 ng was used as template in a further reaction using the nested primers 1OC and 1OD. A 50 ng yield of approximately 1.7 kb DNA fragment

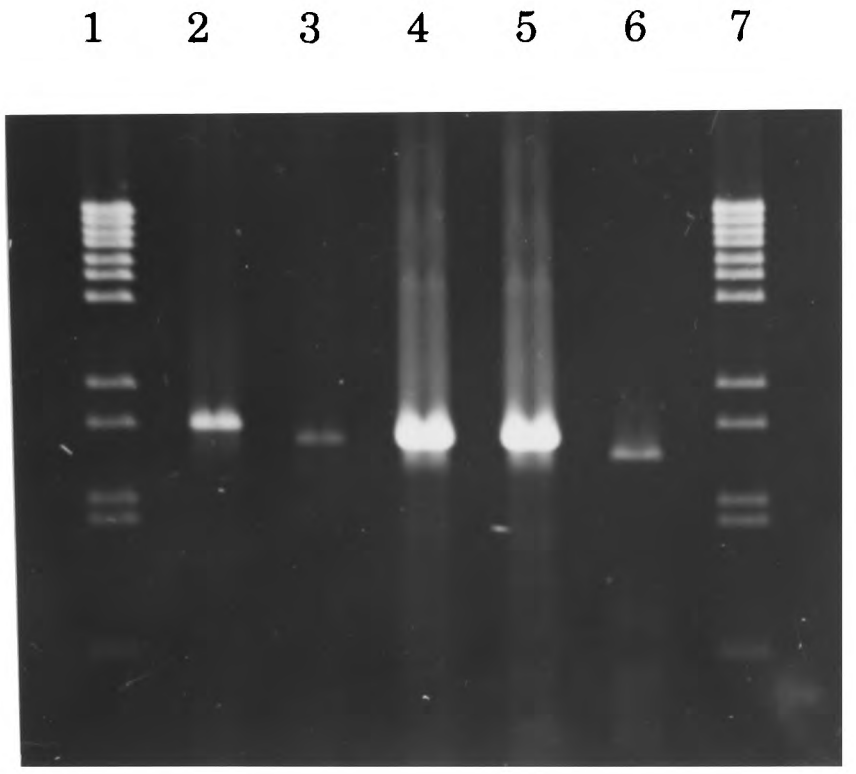


Figure 6.7 Ethidium bromide-stained products of five inverse PCRs whose products include the chromosome 1 breakpoint region. Tracks 1 and 7 are BstEII digested λ. A full description of the reactions is given in the text. The sizes of the products are between 1.7 kb (track 6) and 1.9 kb (track 1).

was produced (track 6, figure 6.7). Thus although a 10-fold amplification was achieved in this reaction, use of the nested primers failed to increase the final yield of PCR product per 100 µl reaction from the PvuII cut and circularised template.

The 1.7 kb DNA fragment generated by inverse PCR using PvuII cut and circularised genomic DNA as template and using the nested primers 1OC and 1OD (see track 6; figure 6.7) was gel-eluted, and restricted with the endonucleases PvuII and PstI in order to trim the fragment ends ready for cloning. However, when the digested DNA was run on a gel, two bands of approximately 1 kb and 0.7 kb were observed instead of a single band of 1.7 kb (not shown). This indicated that another PstI site must be present within this fragment, and therefore that two cloning experiments would be necessary in order to clone the entire fragment [see figure 6.8(a)]. Two ligation experiments were performed using this prepared DNA. Firstly, in order to clone the PstI fragment containing the chromosome 1 breakpoint, the DNA was ligated into pUC18 DNA treated with PstI. Of six clones tested following transformation, one named 2ai contained a 700 bp insert (see figure 6.8). Digestion of 2ai plasmid DNA with XbaI demonstrated that a recognition site for this enzyme is present within the plasmid's insert (not shown) as had been predicted from Southern analysis using the probe AP430 (Chapter 5, section 5.2.5; see figure 6.5). In the second experiment, in order to clone the PstI/PvuII fragment, pUC18 DNA was treated with PstI and SmaI. SmaI and PvuII both generate blunt-ended DNA fragments therefore the ends are compatible in cloning experiments. The PvuII end of the insert can be excised using EcoRI which cuts within the multiple cloning site but not within the insert. Following transformation, 97 white colonies were picked and a colony-blot was probed with Pβ (see above). Seventeen positively hybridising colonies were identified (not shown). Plasmid DNA was prepared from four of these and three (PPii, PPiii and PPiv) contained the expected insert of approximately 1 kb, following digestion with PstI and EcoRI (figure 6.8; the fourth contained a smaller insert and was discarded). The restriction mapping of the insert of PPii is described in Chapter 5, section 5.2.6. The

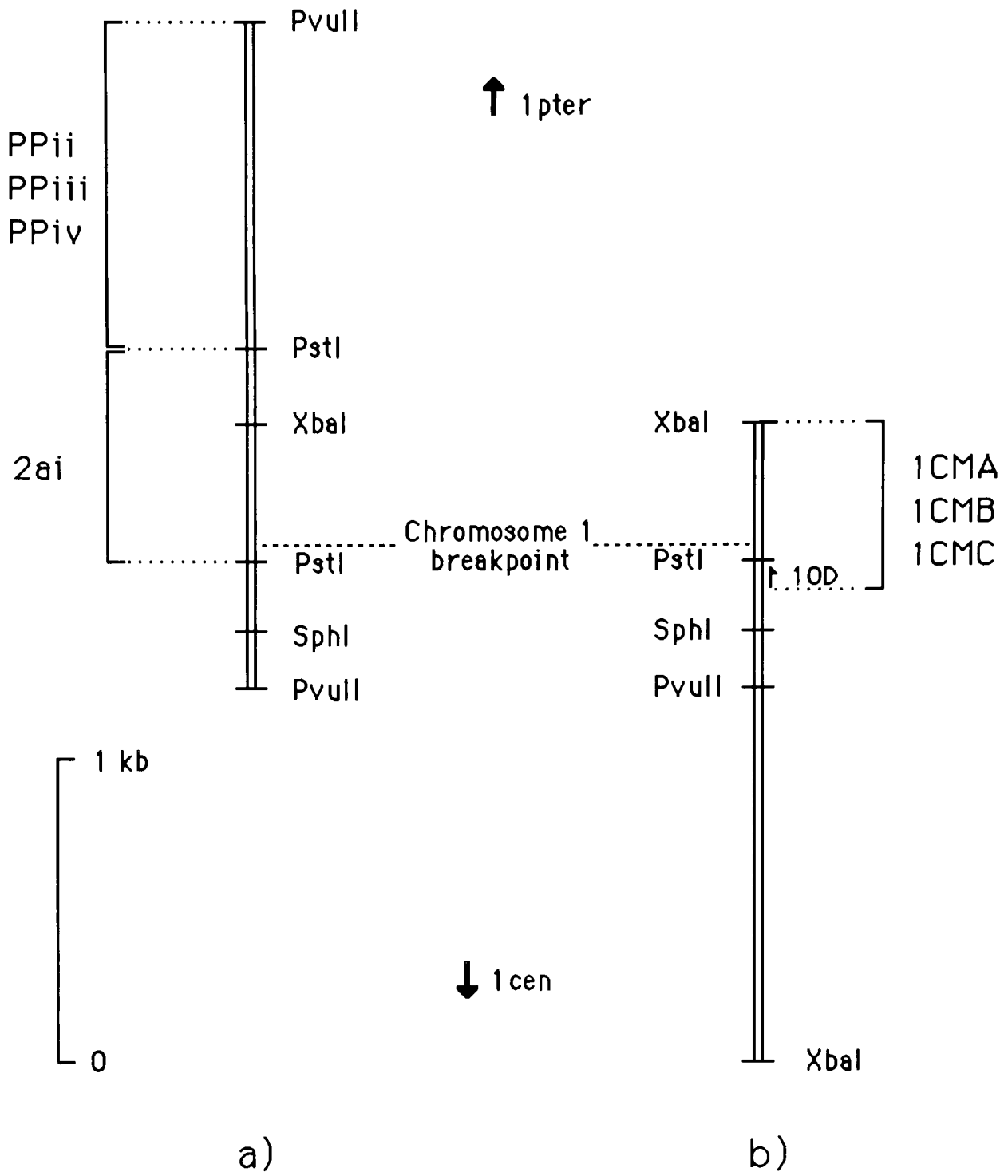


Figure 6.8 Locations of cloned DNA segments from the inverse PCR amplified normal chromosome 1 breakpoint region. a) Clones derived from the PvuII cut and circularised template. b) Clones derived from the XbaI cut and circularised template.

sequencing of subclones of PPii and PPiii is described in Chapter 7, section 7.2.1.3.

The 1.8 kb DNA fragment generated by inverse PCR using XbaI cut and circularised genomic DNA as template and using the nested primers 1OC and 1OD (see tracks 4 and 5; figure 6.7) was gel-eluted, and restricted with XbaI. This generated two DNA fragments, an approximately 0.5 kb fragment extending from the primer 1OD to the closest XbaI site distal to the chromosome 1 breakpoint position, and the other, an approximately 1.5 kb fragment extending from the primer 1OC to the closest XbaI site proximal to the chromosome 1 breakpoint position (see figure 6.5). This DNA was ligated with pUC18 DNA restricted with XbaI and SmaI. Some of the ends of the PCR product were assumed to be blunt-ended and therefore compatible with SmaI (see section 6.3 for discussion on the cloning of PCR products. This end of the insert can be excised using EcoRI which cuts within the multiple cloning site but not in the insert. Following transformation, 90 white colonies were picked and a colony blot was probed with Pβ (see above) and three positively hybridising clones, 1CMA, 1CMB and 1CMC were identified (not shown). They were found to contain 500 bp inserts excised by XbaI and EcoRI. This is the size of fragment expected to contain the chromosome 1 breakpoint (see figure 6.8).

The DNA cloning experiments described above led to the isolation of four independent clones derived from the PvuII cut and circularised template (PPii; PPiii; PPiv; 2ai) and three from the XbaI cut and circularised template (1CMA; 1CMB; 1CMC). The locations of these clones are indicated in figure 6.8 and the sequencing of the clones is described in Chapter 7, section 7.2.1.3.

Southern analysis of the X;1 translocation was performed using the insert of PPii as a probe and is described in chapter 5, section 5.2.6. This analysis confirmed that the expected normal chromosome 1 fragment had been amplified and provided data on the preliminary characterisation of the der(X) junction fragments. PPii was additionally the probe used in an in situ hybridisation experiment which confirmed the genomic localisation of this DNA segment at the expected site, 1p34 (Chapter 5, section 5.2.7).

The sequence appears by Southern analysis to be conserved in vertebrate species (Chapter 7, section 7.2.3).

6.2.3 Amplification and cloning of the der(X) junction fragment

So far in this chapter, the cloning of the X chromosome breakpoint region, the der(1) junction fragment and the chromosome 1 breakpoint regions have been described. The restriction mapping of these DNA segments is described in Chapter 5. Southern analysis using the probe PPii (described in Chapter 5, section 5.2.6) indicated that the translocation was probably a two breakpoint reciprocal rearrangement without significant deletion or insertion of sequence at the breakpoints. The XbaI junction fragment on the der(X) chromosome would therefore be a suitable segment for circularisation and inverse PCR (figure 6.9). A choice of PCR strategies was available for the amplification of the junction. Since the normal X chromosome and chromosome 1 segments either side of this breakpoint were already cloned, it would be possible to design inverse PCR primers from the DNA sequence of either chromosome. Alternatively, an orthodox PCR could be performed using a primer from the sequence of each chromosome. The strategy I chose was inverse PCR using primers from X chromosome sequence, because at the time of the experiment, this sequence was already available, and that from the chromosome 1 segments distal to the breakpoint had not yet been obtained.

The conditions for circularisation of the der(X) XbaI junction fragment were calculated as follows. The value of j for a 1.2 kb DNA fragment is 9.4×10^{13} ends/ml. Therefore the value of i giving 90% circularisation ($j/9$) is 1.0×10^{13} ends/ml. This is equivalent to a DNA concentration of approximately 1µg / 45 µl assuming that the average length of XbaI fragment in human genomic DNA is 4 kb. Template for the inverse PCR of the der(X) was prepared by ligating 1 µg XbaI digested human genomic DNA in a 200 µl volume. This dilution is significantly greater than that calculated to produce 90% circularisation of the 1.2 kb junction fragment.

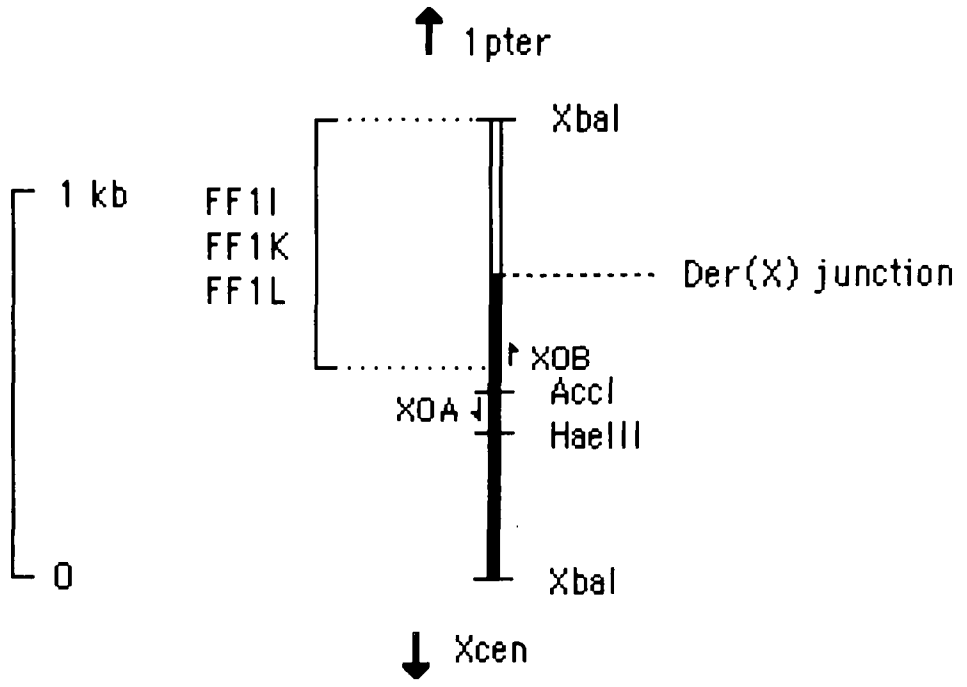


Figure 6.9 Details of the der(X) XbaI junction fragment which was circularised for use as a template in an inverse PCR using the primers XOA and XOB. The locations of FF1I, FF1K and FF1L are indicated which are the cloned DNA segments from the amplified PCR product containing the der(X) translocation junction.

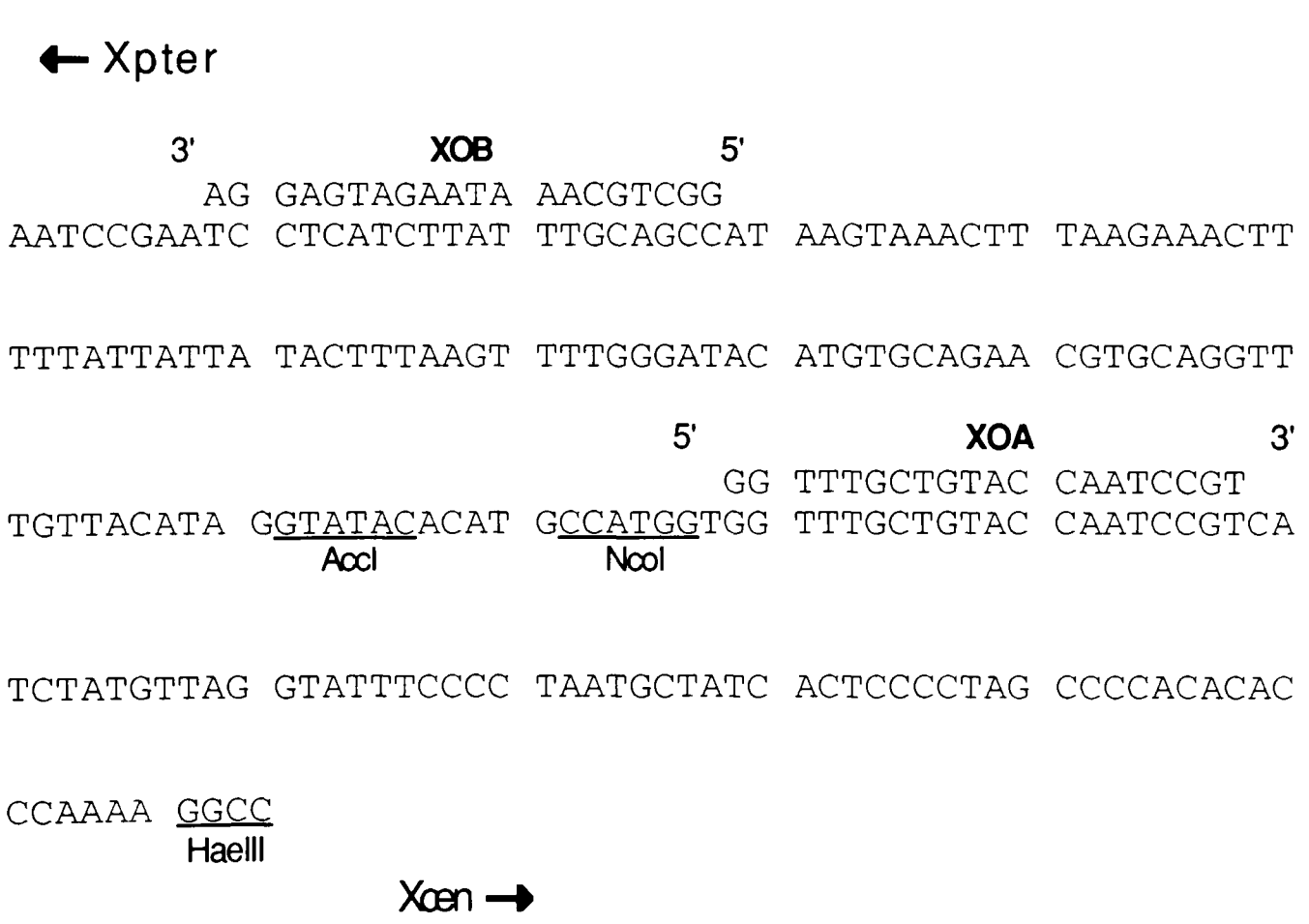


Figure 6.10 Normal X chromosome DNA sequence, approximately 200 to 400 bp proximal of the breakpoint position. The locations and sequences are shown of the primers XOA and XOB which were used in the inverse PCR amplification of the der(X) translocation junction. AccI and NcoI restriction sites are indicated which were used in some experiments to linearise the PCR template prior to amplification.

A pair of primers, XOA and XOB, were designed from X chromosome sequence approximately 300 bp proximal to the breakpoint (see figs 6.9 and 6.10). Figure 6.10 also illustrates the positions of AccI and NcoI restriction sites that lie between the primers. Initial PCR experiments were performed using NcoI linearised template. These experiments were unsuccessful. No bands were visible when the products of PCR were run on ethidium bromide stained gels (not shown). This result might be caused by the primers not working. Alternatively, it might be explained if there is a second NcoI site within the XbaI junction fragment, as it would cut between the 3' ends of the primers preventing amplification. The presence of such a site was not tested for, however, since inverse PCRs using AccI linearised template were more successful and indicated that there were no problems with the primers. The following conditions were used for PCR: denaturation at 94^{O}C for 1 minute, primer annealing at 59^{O}C for 40 seconds, and primer extension at 72^{O}C for 1.5 minutes for a total of 30 cycles. These conditions are those which were found to produce an adequate yield of expected product and a low background of non-specific amplification, after testing several PCRs using different conditions. The products of two identical PCRs which were run on an agarose gel and stained with ethidium bromide are shown in figure 6.11. Two strongly staining DNA fragments were observed in addition to more weakly staining smaller fragments. The larger fragment is approximately 1.1 kb, the size expected for the der(X) junction. The amplification of the smaller fragment (approximately 0.9 kb) was not expected. Additionally, no band of approximately 1.5 kb corresponding to the expected amplification product of the normal X chromosome fragment was observed. In order to further clarify the identification of the products, the gel of PCR products was blotted and probed (figure 6.11) with the insert of N1 (see section 6.2.1; an X chromosome clone containing the breakpoint region). A strong hybridisation was observed to the 1.1 kb band suggesting that it was indeed the amplified junction fragment. A weaker hybridisation to the 0.9 kb fragment was also observed. The origin of this amplification product could not be explained at this stage and was not pursued further.

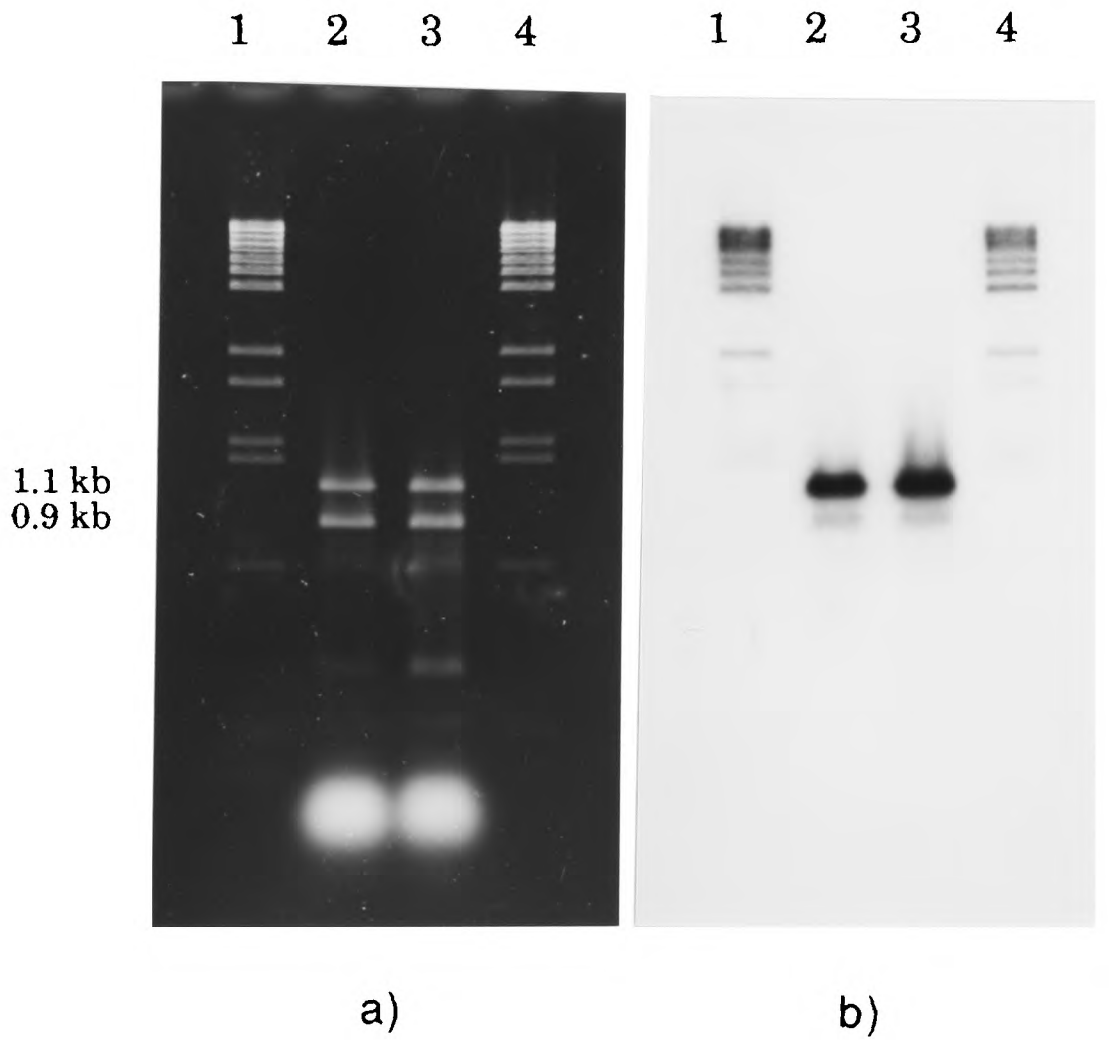


Figure 6.11 a) Ethidium bromide-stained products of two identical inverse PCRs (tracks 2 and 3) which include the der(X) junction from the X;1 translocation; and b) result of probing a Southern blot of the gel with the insert of N1. Tracks 1 and 4 are BstEII digested λ. A full description of the reactions is given in the text. The sizes of the two main amplified products are indicated.

However, sequence analysis of the X chromosome which was performed later and is described in Chapter 7 revealed the likely origin. A KpnI LINE repetitive element was identified, the distal end of which lies between the primers XOA and XOB (see Chapter 7, section 7.2.2.1). Thus a sequence very similar to XOA may be present on average every 100 kb in the human genome (Deininger and Daniel, 1986). For instance XOA is homologous at 17/20 nucleotides to the KpnI LINE element which is 1 kb upstream from exon 1 of the human factor IX gene, if an adjustment to the alignment of the sequences is made in their comparison within the primer sequence (see Chapter 7, figure 7.10; Yoshitake et al., 1985). This finding clearly predicts that a higher than average background of non-specific PCR products is expected using the primers XOA and XOB. It also indicates the reason why some hybridisation was observed of the insert of N1 to the 0.9 kb PCR product (figure 6.11). N1 includes 521 bp of the Kpn1 LINE element (see Chapter 7, section 7.2.2.1) and would therefore be expected to hybridise to any PCR product derived from another Kpn1 LINE element.

The 1.1 kb fragment amplified as described above was gel-eluted and digested with XbaI. Two fragments were observed of approximately 0.7 kb and 0.4 kb (not shown). The 0.7 kb fragment was expected to represent the segment spanning the translocation junction from XOB to the XbaI site approximately 400 bp distal to the junction; and the 0.4 kb fragment was expected to represent the segment from XOA to the XbaI site approximately 800 bp proximal to the junction (see figure 6.9). This DNA was ligated with XbaI/SmaI cut pUC18 DNA. The ends of the PCR product were treated as if blunt-ended and therefore compatible for cloning with SmaI (see section 6.3 for discussion on the cloning of PCR products). This end of the insert can then be excised using EcoRI which cuts within the multiple cloning site of pUC18. Following transformation, 100 white colonies were picked, and three were positive when a colony-blot was probed with Pβ (see section 6.2). This probe represents chromosome 1 DNA and therefore distinguishes clones containing the junction fragment from those containing the 0.4 kb X chromosome fragment. These clones were named FF1I, FF1K and FF1L (figure

6.9), and their inserts were excised using XbaI and EcoRI. All were approximately 700 bp in length (not shown). The sequencing of FF1I and FF1K is described in Chapter 7, section 7.2.1.4.

6.3 Discussion

This chapter describes the successful application of inverse PCR with respect to short chromosome walks and to cloning translocation junction fragments. The method is ideally suited to the task for the following reasons. The application of inverse PCR requires some detailed knowledge of the cutting sites of restriction endonucleases and requires DNA sequence information so that oligonucleotide primers can be designed (Ochman et al., 1988; Silver and Keerikatte, 1989; Triglia et al., 1988). Since the ultimate goal of this study was a full characterisation of the translocation structure at both these resolutions, the application of inverse PCR did not require any information further than was already needed in this respect. One limitation that it placed on the study was that such a characterisation was required for each amplification in turn so that to clone all the relevant DNA fragments, a cycle of restriction mapping, sequencing, PCR amplification and cloning was necessary. A further limitation was that only a relatively small region of DNA around the breakpoints was made available for cloning because of the size of amplification product (see section 6.1.2). This might make the identification of a restriction fragment suitable for an inverse PCR template difficult. Fortunately, in this study suitable restriction sites for a single enzyme were found. If they had not been, then different restriction sites at either end of the fragment could have been identified as sites for circularisation. The single-stranded ends could be infilled using Klenow for 5' overhangs and the fragment circularised through blunt ended ligation. This strategy of ligating 'non-compatible' DNA ends in subcloning experiments is described in Chapter 2, section 2.11.1 and was used in Chapter 7, section 7.2.1.3. In the case of 3' overhangs, these can be trimmed using an enzyme with 3' to 5' exonuclease

activity (Hemsley et al., 1989). It has been demonstrated here that inverse PCR can be performed successfully from a circularised blunt-ended PvuII fragment (section 6.2.2).

Two modifications to the basic inverse PCR strategy were used in this study as described earlier in this chapter. Both produced significant improvements in yield of amplified product in certain circumstances. The first was the linearisation of the circular PCR template. This method was essentially as originally described by Triglia et al. (1988). However, Ochman et al. (1990) stated that this procedure is not necessary to promote efficient amplifications. In my hands, opening of the circular template prior to amplification consistently produced an increase in yield. The reason why this effect is produced is probably concerned with competition between primer-template and template-template annealing. When a linear DNA template is denatured, the complementary strands separate, and their concentration in solution is very low compared to the primer concentrations, thus primer-template hybridisation is favoured over that between complementary template strands. When a covalently closed circular DNA template is denatured, the complementary strands are unable to physically separate unless a single-stranded DNA nick is present, thus their effective concentration in solution will be significantly higher leading to an increased probability that they will reanneal, thus reducing the incidence of template-primer hybridisation and template amplification.

The other modification of the basic inverse PCR strategy was the use of nested primers. The use of such primers has the benefit of specific reamplification of the desired DNA segment only, from a possibly complex product of an initial PCR. Generally, the initial template concentration will be much higher in such an experiment too, thereby maximising the chances of a successful amplification of the desired DNA segment. The use of nested primers produced more than a tenfold increase in amplification yield in two experiments (amplification of the der(1) junction fragment and of the normal chromosome 1 fragment from the XbaI cut and

circularised template; see figures 6.4 and. 6.7); however it did not significantly increase the yield in another (amplification of the normal chromosome 1 fragment from the PvuII cut and circularised template; see figure 6.7). Many factors probably influence the stage at which the exponential accumulation of PCR product ceases, the so called 'plateau effect' (Innis and Gelfand, 1990). These include utilisation of substrates, stability of reactants, end-product inhibition, competition for reactants by non-specific products or primer-dimer, and reannealing or incomplete denaturation of specific product at high concentration. It is evident that the final yield of product is highly dependent on the reaction conditions and the reactants used, such as the primers; therefore the use of some nested primers might not always increase the product yield. If an inverse PCR experiment failed to produce an adequate yield of the desired fragment following the usual investigation of conditions of reaction such as buffer composition and thermal cycling conditions, then the use of alternative primers or the linearisation of the circular template with alternative restriction enzymes would be recommended.

The cloning of PCR products described in this work was facilitated in many cases by trimming the fragment ends using restriction enzymes. An alternative approach is to include a restriction site within the primer sequence (Scharf et al., 1986). This method has the advantage that it is flexible and does not require restriction sites to be present within the amplified DNA. In the cloning of some PCR products described in this chapter, only one end of the DNA segment was trimmed and the other end was treated as if it was blunt-ended (for example the cloning of the der(X) junction fragment; section 6.2.3). However, these cloning experiments may have been somewhat fortuitous, since Taq polymerase has template-independent terminal transferase activity resulting in the addition of a single nucleotide at the 3' end of most fragments (Clarke, 1988; Mole et al., 1989). This makes the cloning of PCR products as blunt-ended fragments inefficient. Alternative means of overcoming this problem are to use a 3' to 5' exonuclease to remove the 3' overhang at the ends of the PCR product

(Hemsley et al., 1989) or to exploit the presence of this overhang which is usually an adenosine by ligating with vector with a 3' thymidine overhang (Marchuk et al., 1990).

In this chapter, the cloning of the junctions of an X;1 translocation and of the DNA segments around the breakpoints from the normal X chromosome and chromosome 1 is described. This was achieved using a particular application of PCR, inverse PCR. The following chapter describes how these clones were sequenced, revealing features which may have played a role in producing the translocation.

Chapter 7 DNA sequencing of the X;1 translocation breakpoint regions and of the junction fragments

7.1 Introduction

In this chapter, the structure of the X;1 translocation is described at the finest level of resolution - at the level of the DNA sequence. The sequencing of approximately 1.5 kb DNA from around the X chromosome breakpoint and of approximately 3 kb DNA from around the chromosome 1 breakpoint is also described. These sequences are compared with each other and with sequences held in the EMBL database. The results of this analysis are discussed in relation to the possible mechanisms causing the translocation.

The cloning of DNA whose sequence is presented in this chapter was described in Chapter 6. These clones were comprehensively subcloned in order to assist sequencing (the reading of DNA sequence was found to be reliable up to approximately 400 bp from the sequencing primer).

In Chapter 4, the X;1 translocation was shown to be paternally inherited (section 4.2.5). It follows that the most appropriate normal chromosome sequences with which to compare the sequences of the translocation junction fragments would be those of the patient's father which rearranged to produce the translocation. This would mean that any differences found between the sequences of the normal and rearranged chromosomes could not be confused with normal polymorphic sequence variation. Such normal variation is thought to occurs on average every 1 kb on the X chromosome (Hofker et al.,1986). In practice, identifying which of the father's chromosome 1s had rearranged and selectively cloning the relevant chromosome fragment would not be a straightforward procedure, except in the unlikely event of identifying an informative polymorphism within the fragment to be cloned. As a male has only one X chromosome one could clone the relevant portion of the father's X chromosome with confidence, however, at the time of the relevant cloning experiment, DNA from the father was not available therefore DNA from another source was used.

The sequence described in this chapter is derived from the following sources. Both junction fragments were cloned and sequenced from lymphoblastoid cell line DNA of the t(X;1) patient (WLS). Sequence containing the breakpoint on the normal X was derived from two sources. The X chromosome clone Alf1 is a subclone of the cosmid XJC-5 (see Chapter 5, section 5.2) which in turn was derived from a 49,XXXXY library (Bodrug et al., 1989). The remaining X chromosome clones sequenced were derived by inverse PCR of the normal X chromosome of the t(X;1) patient (Chapter 6, section 6.2.1). The normal chromosome 1 clones sequenced were derived by inverse PCR of DNA from an anonymous normal female (Chapter 6, section 6.2.2) therefore these clones may be derived from either of her chromosomes.

7.2 Results

7.2.1 Subcloning and sequencing

7.2.1.1 Subcloning and sequencing of the normal X chromosome fragments

The subcloning of the 560 bp XbaI/HindIII DNA segment, Alf1, into pUC18 (figure 7.1), was described in Chapter 5, section 5.2.3. Alf1 was also subcloned into pBluescribe and was sequenced using single and double stranded sequencing protocols (Chapter 2, section 2.15).

The 1 kb HindIII/XbaI fragment adjacent to Alf1 on its proximal side, and which contains the X chromosome translocation breakpoint (figure 7.1) was cloned from DNA whose amplification by inverse PCR was described in Chapter 6 (section 6.2.1). The template of this PCR was DNA from the t(X;1) patient and thus the sequence obtained represents her normal X chromosome. Two clones were isolated and named N1 and N2. The gel-eluted inserts of these clones were found to contain a single HaeIII site and the fragments of 700 bp and 400 bp generated by restriction of the inserts of N1 and N2 were subcloned into M13mp18 and M13mp19 single-stranded sequencing vectors (see table 7.1 for full details of clones). Sequencing reactions were performed using the forward sequencing primer. The sequence obtained from all these subclones

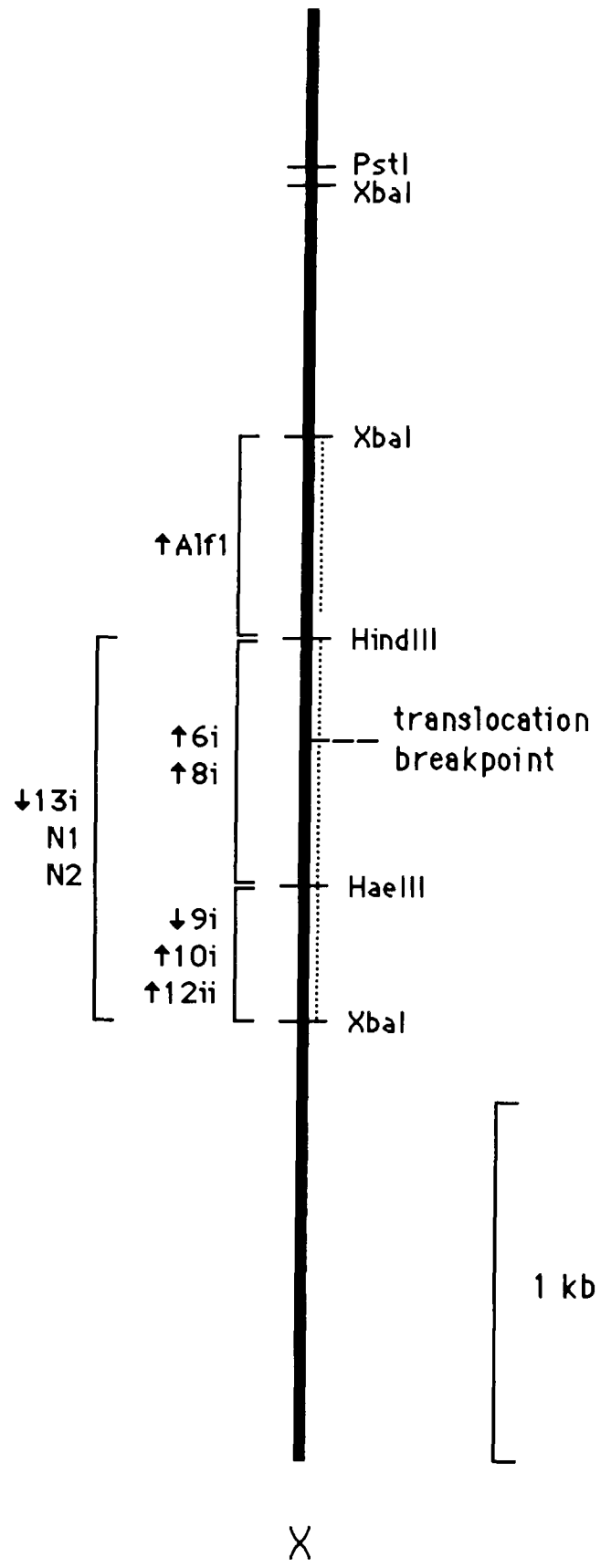


Figure 7.1 Locations of cloned DNA segments around the breakpoint of the normal X chromosome. Arrows on single-stranded M13 clones indicate the direction of sequencing possible in these clones, and the arrow on the Alf1 clone indicates the direction of sequencing this DNA segment using the forward sequencing primer. The dotted line indicates the extent of sequenced DNA.

X chromosome subclones

Subclone	Insert size	Insert sites	Vector	Vector sites	Parent
6i	700 bp	HindIII/HaeIII	M13mp19	HindIII/HindII	N1
8i	700 bp	HindIII/HaeIII	M13mp19	HindIII/HindII	N2
9i	400 bp	XbaI/HaeIII	M13mp18	XbaI/HindII	N1
10i	400 bp	XbaI/HaeIII	M13mp19	XbaI/HindII	N1
12ii	400 bp	XbaI/HaeIII	M13mp19	XbaI/HindII	N2
13i	1100 bp	HindIII/XbaI	M13mp18	HindIII/XbaI	N1

Der(1) subclones

Subclone	Insert size	Insert sites	Vector	Vector sites	Parent
2i	320 bp	HindIII/PstI	M13mp19	HindIII/PstI	A1
3i	320 bp	HindIII/PstI	M13mp18	HindIII/PstI	A4
4i	320 bp	HindIII/PstI	M13mp19	HindIII/PstI	A4
15i	420 bp	PstI/PvuII	M13mp18	PstI/SmaI	A1
15ii	300 bp	PstI/PvuII	M13mp18	PstI/SmaI	A1
15iii	950 bp	PstI/PvuII	M13mp18	PstI/SmaI	A1
16i	300 bp	PstI/PvuII	M13mp19	PstI/SmaI	A1
18i	740 bp	HindIII/PvuII	M13mp18	HindIII/HindII	A1
18vi	1990 bp	HindIII/PvuII	M13mp18	HindIII/HindII	A1
19iii	740 bp	HindIII/PvuII	M13mp19	HindIII/HindII	A1

Chromosome 1 subclones

Subclone	Insert size	Insert sites	Vector	Vector sites	Parent
IIIPa	500 bp	BamHI*/PvuII	pUC18	PstI*/SmaI	PPiii
IIEa	550 bp	PstI/BamHI*	pUC18	PstI/EcoRI*	PPii
IIIEa	550 bp	PstI/BamHI*	pUC18	PstI/EcoRI*	PPiii
F5A	300 bp	PstI/SstI	pUC18	PstI/SstI	PPiii
F3A	220 bp	PstI/XbaI	pUC18	PstI/XbaI	2ai

Table 7.1 Details of the subclones generated from the X chromosome, the der(1) chromosome and the normal chromosome 1, in order to assist DNA sequencing. An asterisk indicates that a cohesive end produced by the restriction enzyme was made blunt-ended.

is listed in the appendix and the consensus X chromosome sequence is listed in figure 7.2. This consensus includes sequence derived from the cloned der(X) and der(1) junction fragments.

7.2.1.2 Subcloning and sequencing of the der(1) junction fragment

The amplification and cloning of the der(1) junction fragment was described in Chapter 6, section 6.2.1. Two clones, A1 and A4, were isolated. These contain HindIII/XbaI inserts of 2 kb including over 1.7 kb DNA of chromosome 1 origin (the restriction mapping of the insert of A1 was described in Chapter 5, section 5.2.5; see figure 5.18). Several subclones of A1 and A4 were generated by ligation of suitably restricted gel-eluted DNA fragments into M13mp18 and M13mp19 single-stranded sequencing vectors. Full details regarding the clones are provided in table 7.1. Their locations are illustrated in figure 7.3. Sequencing reactions were performed using the forward sequencing primer. The PvuII/PstI insert of clone 15iii is approximately 1 kb in length and therefore too large to be completely sequenced using the forward sequencing primer (figure 7.3). Restriction sites had not been identified which would allow the segment to be subcloned. Sequence at the distal end of the DNA fragment represented by clone 15iii was obtained by sequencing the clone 18vi using the primer 1OA (see figure 7.3; 1OA was designed for inverse PCR, see Chapter 6, section 6.2.3). The insert of clone 18vi is 1.8 kb in length and includes the DNA fragment cloned in 15iii, although these clones contain the opposite DNA strands (figure 7.3). Sequence from this reaction was used to design a 16mer primer, 1OE (figures 7.3 and 7.4). This primer was used to obtain further sequence of the insert of clone 18vi. The appendix includes a list of the sequence obtained from all these clones. The der(1) sequence in the immediate vicinity of the translocation junction is presented in section 7.2.1.5 (figure 7.8), and the sequence at the junction is illustrated in figure 7.5.

Xcen

TCTAGAGCCA GAAATACCAT TTGACCAAGC AATCCCATCA CTGGGTATAT ACCCAAAGGA 60
XbaI
TTATAAATTA TTCTGTGCTT ATTTCAGCAC TATTCACAAT AGCAAAGACT TGAAACCAAC 120
CCAAATGCCC ATCAATGATA GACTGGATAA AGAAAATTTG GCACATATAC ACCATGGAAT 180
ACTATGCAGC AACAAAAAAG GATGAGTTTC ATGTCCTTTG CAGGGATGTG GATGAAGCTG 240
GAAAACATCA TTCTCAGCAA ACTAACATAG GAACAGAAAA TCAAACTCGG CATGTTCTCA 300
CTCATAAATC GGAGTTGAAC AACCAGAAGA CATGGATGCA GTGGAGCATC ACACACTGGG 360
GCCTTTTGGG TGTGTGGGGC TAGGGGAGTG ATAGCATTAG GGGAAATACC TAACATAGAT 420
GACGGATTGG TACAGCAAAC CACCATGGCA TGTGTATACC TATGTAACAA ACCTGCACGT 480
LINE
TCTGCACATG TATCCCAAAA CTTAAAGTAT AATAATAAAA AGTTTCTTAA AGTTTACTTA 540
TGGCTGCAAA TAAGATGAGG ATTCGGATTC AGCATAAGTC ATATATTTTT AAAATATGTC 600
AGTTCTTGCC ATTGTGTTAG GGAAACCACC AAGTACACAT TTATACATTT TTGCTCAGCA 660
TTTTTCAGGA GATAATCTCA ATGATGATTC AAAGAAAAAA ATAATAACTC ATAAGAATTT 720
TTATTTACTG ACTGCTTAGT CACAAAAGAC AGAACACCAT CTCTTGAAGT CTCAGACTTA 780
Breakpoint
GAACACTCTC TTCGATAGAT TTACCAGCAA AATTCGTGAT TCATGCCTGT ATAGTAATTG 840
CTTTTGTGAT GTTCATATTT TTATTTAACC AGATATTGGT TCTGTAGCTT ATGCATGATG 900
TTGGCTACTC TCGAAAGAAA TTTCATTTTA AATTAGTTAT GCCCCTTAGG AGAATAATGA 960
AACATTTGGG GAGAATTTTA TTATTGGGGG ATTGAAGCTA TATTTGGTTT AAGAAATACC 1020
TGCCAAATAA TGCTTTGACA ATTTTAAACA TAAGCTTTTG TAGGATTTCT ATATAGAATT 1080
HindIII
TGATATGCTT TGGCCTTAAC TGGTGTTTTG AGGCCTGAGG TTTAGTTTTA AGTGGAATTG 1140
CCTTTTAAAA TGAGTAGCAG TAGGCATTAG CATAAGCGAT TTCAGGAGCT GTTTCTAGTG 1200
TGTTTAGATG CTAACATATG CAAAAATAAG TACATCTAGG GTAAACTTAA TAAATTATGT 1260
ATGTTTATTG ACTCATAGAT AATAGGTATA TCTCTTTAAA ATATCCTTTG TGAATCAGAC 1320
AATTCTCTTC CCGTTATGAT TTCTCTGTAA GACAGTACAT TCTACATGCG TTGGTCGTGG 1380
GTTTCATATA CATTGTTTCT CTGAGTTTCA AACAGATTGG TTTGTTTGCC AGTCCAAAAA 1440
AAAAAAAAAA ATCCTCCCTA TATTAAGAGA GTGCTTTGTC AGCTTTCCCC TGTGAATAAC 1500
AAAGTGAGGA TGTTCTCTTT AAAGGAATAA AACAAAAAAT CAGGGATAAA AGATTGCGCC 1560
TGTTCTAGA 1569
XbaI
Xpter

Figure 7.2 Consensus sequence from the X chromosome surrounding the t(X;1) breakpoint position. The position of the three nucleotides (GAA) of junctional homology at the breakpoint position and the region of sequence with homology to a LINE (nucleotides 1 to 521) are indicated. The asterisk adjacent to the HindIII site indicates a gap in the DNA sequence estimated to represent fewer than 20 nucleotides.

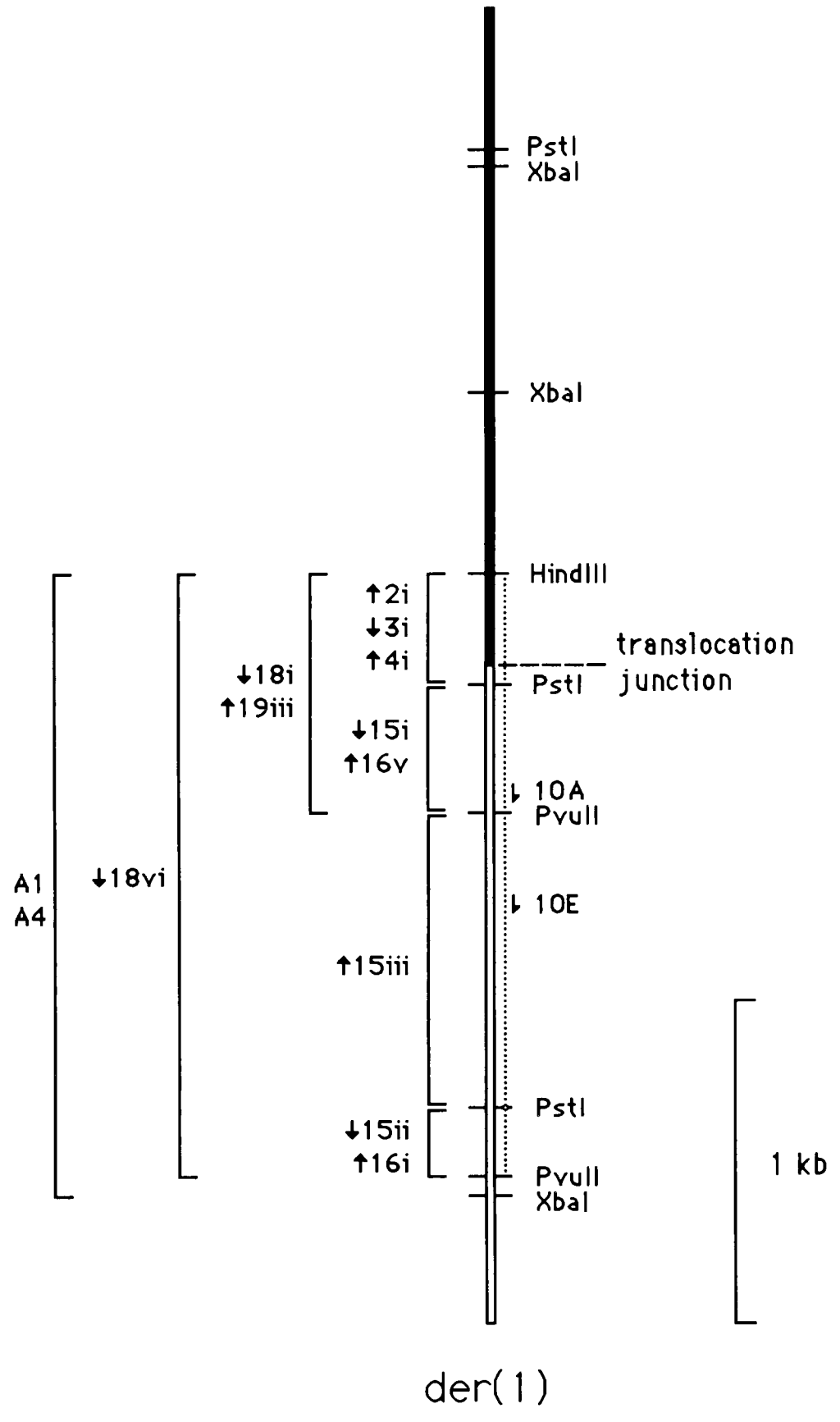


Figure 7.3 Locations of cloned DNA segments from the der(1) junction. The dark line represents X chromosome DNA and the pale line represents chromosome 1 DNA. Arrows indicate the direction of sequencing permitted in these single-stranded M13 clones. The positions of two oligonucleotides, 1OA and 1OE, are indicated which were used to prime DNA sequencing of clone 18vi. The extent of sequenced DNA is indicated by a dotted line.

```
        3'GTCATACTCCCATCTT 5'                    1OE
5'  AGGGGACAGTATGAGGGTAGAAGGGCTCTCA  3'    Chromosome 1
    |                               |        sequence
    940                             970
```

Figure 7.4 Sequence of oligonucleotide primer 1OE and position with respect to the consensus chromosome 1 sequence (see figure 7.7).

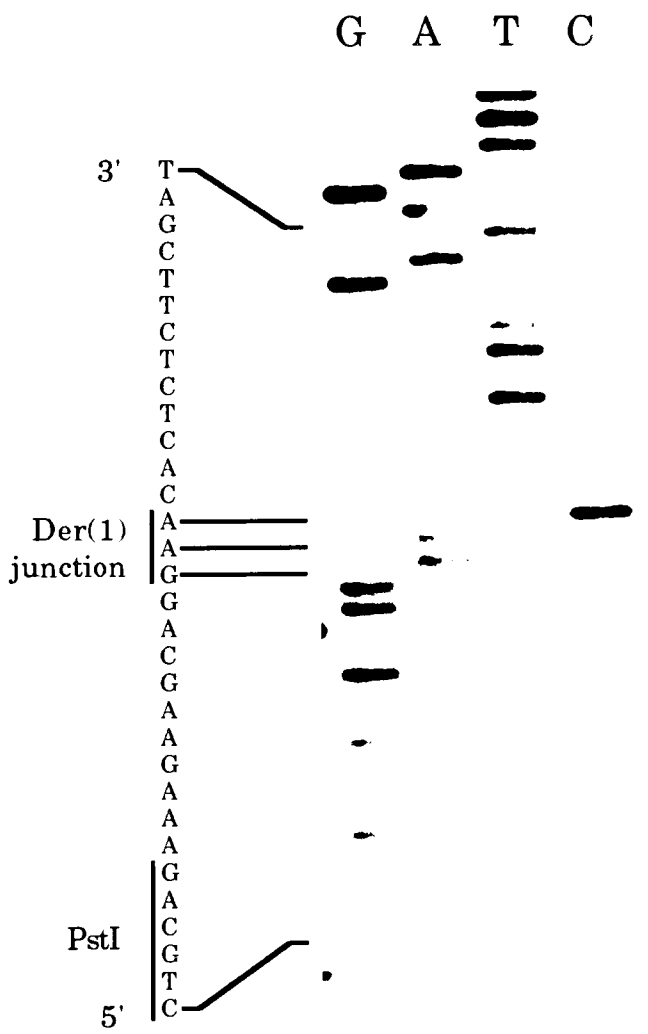


Figure 7.5 DNA sequence across the der(1) junction determined by sequencing the clone single-stranded clone 4i. The trinucleotide (GAA) at the junction site is indicated, as is the PstI site which lies just proximal to the junction.

7.2.1.3 Subcloning and sequencing of DNA fragments from normal chromosome 1

The cloning of DNA segments from around the breakpoint position on chromosome 1 was described in Chapter 6, section 6.2.2. This cloned region extends from 70 bp proximal to approximately 1.5 kb distal of the breakpoint position (figure 7.6). The clones 2ai, PPii, PPiii and PPiv are derived from one PCR reaction while the clones 1CMA, 1CMB and 1CMC are derived from another. The positions of these cloned DNA segments are indicated in figure 7.6. Further subcloning of these clones facilitated sequencing which was performed using the double stranded protocol and using the forward and reverse sequencing primers. A summary of the details of all subclones is included in table 7.1 and their positions are illustrated in figure 7.6. The subclone IIIPa was produced by digesting PPiii plasmid DNA with BamHI and PstI, thereby cutting the proximal BamHI/PstI portion from the clone. The cohesive ends were in-filled using Klenow (Chapter 2, section 2.11.1) and the blunt ended molecule was circularised and used to transform competent bacteria. The subcloning of IIEa and IIIEa from PPii and PPiii respectively was similarly performed, this time removing the distal BamHI/PvuII fragments from the parental clones by digesting the plasmids with BamHI and EcoRI (EcoRI cuts within the multiple cloning site of pUC18 just outside the PvuII end of the insert - PvuII could not be used as the clones PPii and PPiii were produced by ligating the PvuII end of the insert into the SmaI site of the pUC18 multiple cloning site; see Chapter 6, section 6.2.2). F5A is a PstI/SstI subclone of PPiii generated by digesting PPiii plasmid DNA with SstI, circularising the fragment and using it to transform competent bacteria. F3A was subcloned from 2ai by the same method using XbaI.

The sequence obtained from the above clones is listed in the appendix. A complete list of the chromosome 1 consensus DNA sequence determined is provided in figure 7.7. This includes sequence derived from the cloned der(X) and der(1) junction fragments.

Since the clones 2ai, PPii, PPiii, PPiv,1CMA, 1CMB and 1CMC are derived from

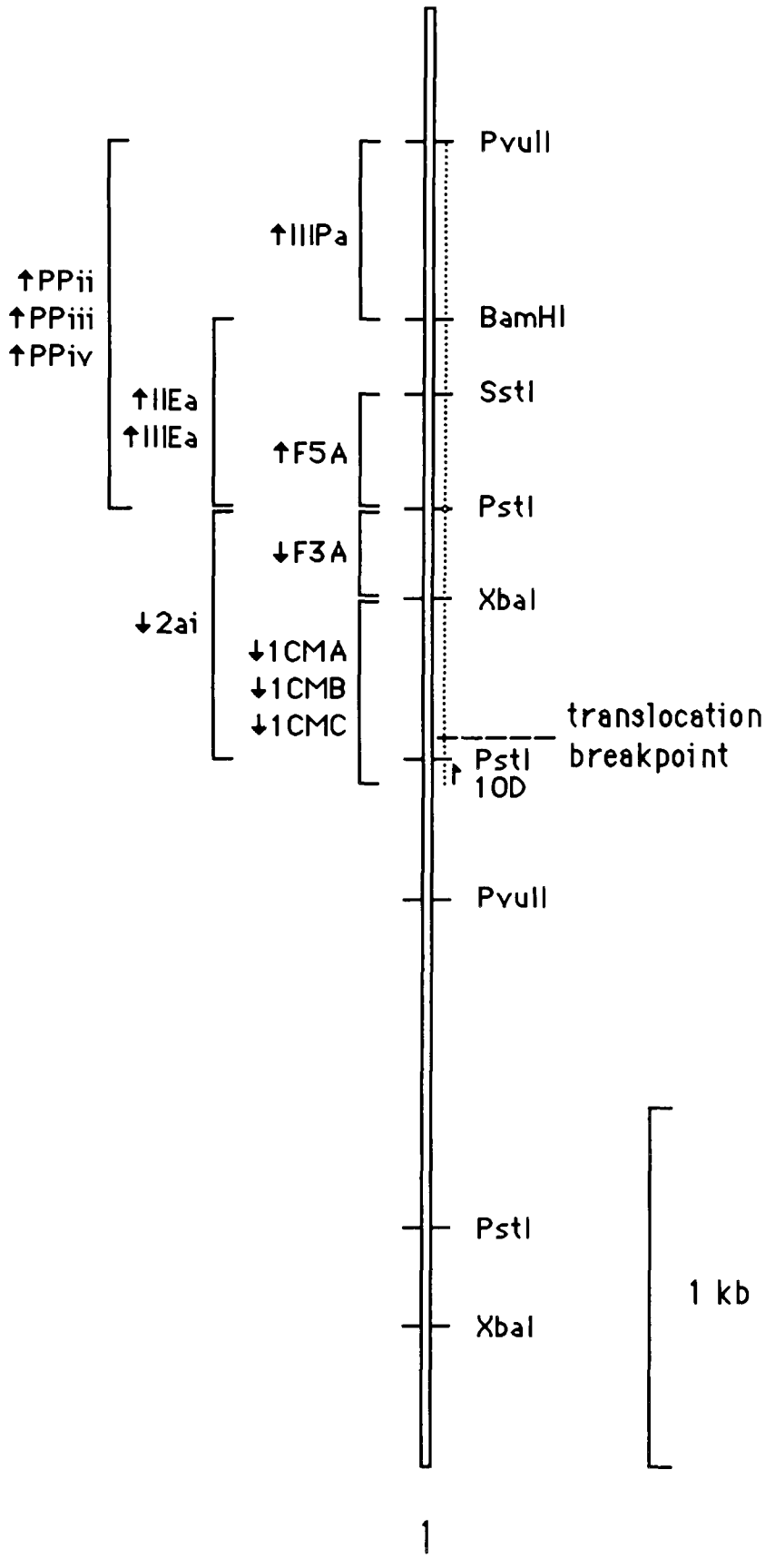


Figure 7.6 Locations of DNA segments cloned from inverse PCR amplified normal chromosome 1 around the t(X;1) chromosome 1 breakpoint position. All DNA segments are cloned into pUC18, and arrows indicate the direction of sequencing using the forward sequencing primer. The dotted line indicates the extent of sequenced DNA.

genomic DNA from a normal individual, some clones may represent one chromosome 1 and some the other. No differences were observed in the sequences obtained from these clones. However, if differences had been observed, then they could either represent PCR errors (see section 7.2.4) or normal polymorphic differences in DNA sequence.

7.2.1.4 Sequencing of DNA at the der(X) junction

The description of the cloning of the der(X) junction fragment is in Chapter 6, (section 6.2.3). Three independent clones, FF1I, FF1K and FF1L, were isolated from a single inverse PCR amplification. The inserts of these clones are approximately 650 bp in length and extend from 250 bp proximal of the translocation breakpoint to 400 bp distal of the breakpoint (figure 7.8). These clones were sequenced by double-stranded sequencing protocols using forward and reverse sequencing primers. A list of the sequences obtained from all these clones is included in the appendix. The der(X) sequence in the immediate vicinity of the translocation junction is presented in section 7.2.1.5 (figure 7.9).

7.2.1.5 DNA sequence at the translocation breakpoints

The DNA sequence at the translocation breakpoints is illustrated in figure 7.9. Comparison with the normal sequence reveals the structure of the rearrangement. Slight uncertainty in the precise structure is presented by three nucleotides, GAA, which are present at the junction of the der(1) and are present in the normal sequences from both chromosome 1 and the X chromosome. The finding of short nucleotide sequences of junctional homology has previously been observed at sequenced deletion breakpoints (Henthorn et al., 1990; Love et al., 1991; see discussion, section 7.3.2). The four nucleotides, CTTA, immediately proximal to the nucleotides of junctional homology on the X chromosome are deleted in the rearrangement. They are not present in the sequences of either derivative chromosome. The pair of nucleotides, GA,

1cen

CAGCTGCTCC TCTGTTGGAG TGGACTCGCC AGCCCCAATG CCTGCCAGAC ACTCTGCCTG 60
PvuII

GGGCAGGGTG GGGCCAAGGC TGTGCCGCTT ACAACTCCAC TCCTCTAGGC CATTTGGGAA 120

GTTTTCAGGC TCTTGTTCCT GAAGACAGAC AAGCTGACCT GAGGCCTGAG CCTTGAGGAG 180

TAAGACACCT GACAGTAAGC TAGGGGACCT CCCGTCCCAT GCCCCCAACT CTCATGAAGC 240

ATCTTTAATA GTCTCTGAAG CACAATTTGC CCACCACAAA TAGTTCCTCT GCAGCTCAAC 300
PstI

TGGTCACCGA ATGGGCTGGC AGCTGTGCCT CTCCAGAACC ACCCTCTCCC ATGACAGACA 360

GCGACTCCTT GATAAGTGTA TTTTGGTTGA CAGCCTTGAC AGCCTAATTC TCTCTGCTCT 420

GAGTAAGCAA GTCTGACCAC ATGGCCCCAA AGAGTCGGAC AGGCCACCGC GATCACCCAC 480

CACGTGCACT GGGTGGAGAG TCAGACGACG CCCCTCAGGA AGCCGCGCTA AGCACTTGGC 540

GTGATGGAAT CATCACTGCA AGGAGCCATG TCCGCTTTCC TTTACAATCC CCTGGGTGGA 600

Alu

GACAGTTGTA CTTTTTTTTT TTGAGATGGA GTCTTGCTCT TTCACCAAGC TGGAGTGCAG 660

TGGCGCGACC TTGGCTCACT GCAACCTCTG GTTCAAGCGA TTTTCCTACC TCAGCCTCCT 720

GAGTAGCTGG GATTACAGGC ACGCGCCACC ACGCCCAGCT AATTTTTGTA GTTTTAGTAG 780

AGACGGGGTT TCACCATGTT GGCCAGGATG GTCTCGATCT CCTGACCTCG TGATCTGCCT 840

Alu

GCCTCAGCCT CCCAAAGTGC TGGGATTACA GGCGTGGCCA CCACACCCGG TCAACAGTTT 900

CCCTTTTAAT GACATCCATC TGTACAAGAA GGCTAGGTGA GGGGACAGTA TGAGGGTAGA 960

AGGGCTCTCA GGGCCCCTTC TGGGCACATG TATACCAGAC CCAGTGTGTT CTCAGGAAAG 1020

GGCAAGGTGG GGAGCTTTAG TCACCCCTAG GCTGTGAGCA TAGCAGAGAT TGGAAACCCC 1080

AAAACAGTCC CCAGGTTCCT CTCTGCTGGG AACTTGGACT TCCAGAAAGC ACACTGTCCA 1140

CTCGCTCCCA GCAAAGAACA TCTGGAACAC ACCAGGTGCT CACCCGAGGC AACATAACCA 1200

CAGGGGCTCA CAGCTGAACA CAAGGGTTCC TCCAGGAAGG CTCCAGGAAG ACCTGGCCAG 1260
PvuII

AGGATAGGCT CACCTGGCCT GGTCCGTTCC TGAGGCCTTA CTACCCCAAA GCCCAGAGCC 1320

ATGGTGGTTC TGGAGCCCAC CATGCTGCGG CAGGCAGGAG GATGTACTGA AGGAAAGCCA 1380

Z-DNA

GGCCGGGTAT GCATGCATTT ACACACAGAT AAGTGTGCCC GTGTGTGCAC ACACACATGC 1440

Z-DNA

ACATACACCA AGAAGTCGGA TTGGATGGCT CCAAATTCCC ATGGGCTCCG GGCCCATGCG 1500

GCAAGACCTA CAAAAGATAC AAATGCTCCC AAGAGTGGCT GCTTCCCCAT GGGCCTGCAA 1560

GGAAGGAAGG CGAAGTAGTG CGGACGCCAG GAACTTGCCC AGTGGGAGCT GTTGCTGCTG 1620

Breakpoint

CTGCAGAAAG AAGCAGGAAG GAGAAGAGAC TGGGCCCATG GCAGTCCCTT GCGTGTCTGC 1680
PstI

ACCAAGGGCT CTGGGCAGAG AGCCCGGCAG GTGCAGCACA GCTCTGTGGT GACAGGGGCC 1740

LAR

AGGCACTGCC TGAGTCCTTA CCTCGCACAT ACAGGTTCGC AGGGGCTGAG TAACGTGTGC 1800

Figure 7.7 Consensus chromosome 1 sequence indicating the positions of restriction sites and features of note within the sequence. The nucleotides (GAA) of junctional homology at the t(X;1) breakpoint position are indicated. The positions of an Alu element, a stretch of potential Z-DNA and an exon of the LAR gene are also indicated. Asterisks indicate the presence of blocks in the sequencing gel and therefore represent a small but unknown number of nucleotides.

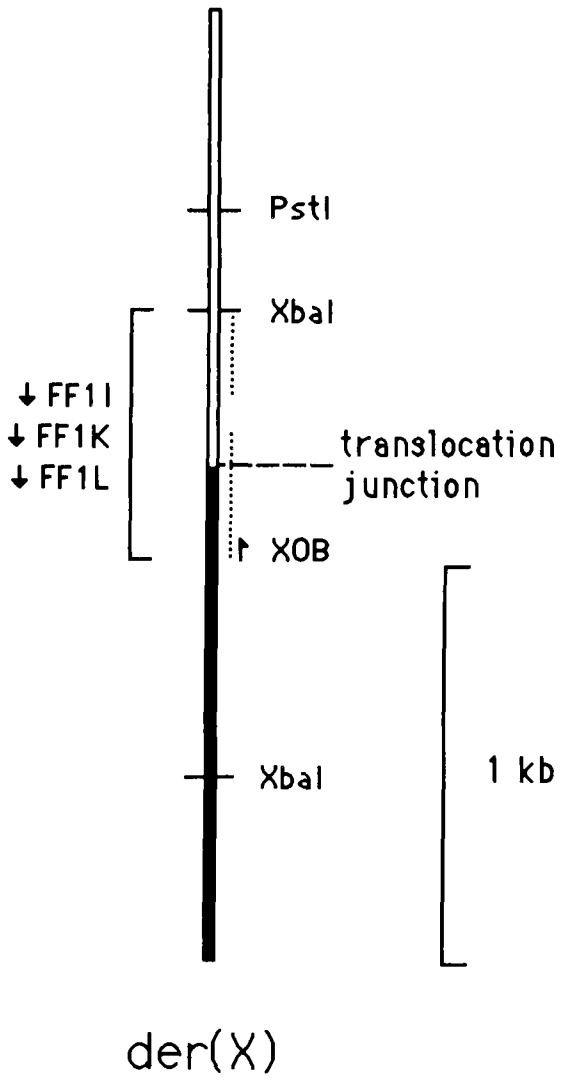


Figure 7.8 Locations of der(X) junction containing clones FF1I, FF1K and FF1L which extend from an XbaI site distal to the der(X) junction to an oligonucleotide primer, 1OB, proximal to the junction. The dark line represents X chromosome DNA and the pale line represents chromosome 1 DNA. The dotted line indicates the extent of sequenced DNA.

Figure 7.9 Nucleotide sequence across and immediately surrounding the t(X;1) junctions and of the corresponding sequences from a normal X chromosome and chromosome 1. Three nucleotides (GAA) of junctional homology are boxed. The presence of this trinucleotide at three additional sites on the normal chromosome 1 sequence and at two additional sites on the normal X chromosome sequence is indicated by underlining. Another trinucleotide (TGC) present four times on the normal chromosome 1 sequence is also underlined.

immediately proximal to the nucleotides of junctional homology on the normal chromosome 1 appear to have been duplicated. They are present in the sequences of both derivative chromosomes. Therefore, given the uncertainty of three nucleotides representing the nucleotides of junctional homology, a deletion of 4 to 7 nucleotides from the X chromosome and a duplication of 2 to 5 nucleotides from chromosome 1 are associated with the X;1 translocation at the breakpoint site. The chromosome 1 DNA strand illustrated in figure 7.9 in the vicinity of the breakpoint is purine rich. The 25 nucleotide region from 12 nucleotides proximal to the nucleotides of junctional homology to 10 nucleotides distal contains a single pyrimidine. This feature is not shared with the X chromosome sequence in the vicinity of the breakpoint. Within this purine rich region, the nucleotide sequence GAA, which represents the nucleotides of junctional homology, is present three more times. The same sequence is also present three times on the normal X chromosome sequence (including the site of junction homology) within the 60 bp DNA sequence illustrated in figure 7.9 surrounding the breakpoint. Immediately proximal to the purine rich region on the chromosome 1 sequence three nucleotides, TGC, are repeated four times (figure 7.9).

7.2.2 Sequence analysis

In order to determine any possible significance of the DNA sequences surrounding the t(X;1) breakpoints, computer-aided sequence analysis was performed.

7.2.2.1 Analysis of the X chromosome sequence

The complete X chromosome sequence (1569 bp; figure 7.2) was compared with all sequences held in the EMBL database in June 1991. Highly significant similarity was found to several sequences which contain KpnI L1 repetitive elements. For example, 86% homology (446 nucleotides out of 521) was observed between the proximal 521 bp of the X chromosome sequence and the equivalent portion of the KpnI L1 repetitive element which is approximately 1 kb upstream of exon 1 of the factor IX gene (nucleotides 1375-

1926; Yoshitake et al., 1985; see figure 7.10). This finding indicated that a KpnI L1 repetitive element is present in the X chromosome test sequence. The region of the X chromosome sequence within which similarity was found is indicated in figure 7.11. The analysis suggests that the KpnI L1 repetitive element extends from beyond the proximal end of the sequenced DNA to within 210 bp of the translocation breakpoint. The proximal end of the KpnI L1 repetitive element may lie up to 5.5 kb proximal to the end of the sequenced DNA segment depending on whether or not it is a full length element (Fanning and Singer, 1987). Within the distal 1048 bp portion of the X chromosome sequence, only the consecutive run of 16 adenines (nucleotides 1436 to 1451) was found to have significant homology to other known sequences which contain similar consecutive adenines. Homology was not observed to extend into the surrounding sequences.

No significant direct or inverted repeats were detected in the X chromosome sequence.

7.2.2.2 Analysis of the chromosome 1 sequence

The complete chromosome 1 sequence (3154 bp; figure 7.7) was compared with nucleic acid sequences contained in the EMBL database in June 1991. An Alu repetitive sequence was identified from nucleotide 623 to 892 which lies 750 bp proximal to the chromosome 1 breakpoint position (figure 7.7). Analysis also revealed a 100% match between a 113 bp segment of the cDNA sequence of the human leukocyte antigen related protein (LAR) and the chromosome 1 test sequence (nucleotides 1762 to 1868, figure 7.12; a compression was present in the sequencing gel within the sequence corresponding to this region, however the sequence matched perfectly on either side). Examination of the chromosome 1 test sequence revealed that sequences matching the consensus donor and acceptor splice sites are present at either end of the homologous region of DNA sequence (figure 7.13; Mount 1982). This finding together with the match to the LAR cDNA sequence indicates that the region represents an exon of the

```
1 (X chromosome test sequence)
|
TCTAGAGCCAGAAATACCATTTGACCAAGCAATCCCATCACTGGGTATATACCCAAA
|||||| | |||||||||||| ||||| ||| ||||||| | |||||||||||| |||
TCTAGACCTAGAAATACCATCTGACCCAGCCATCCCATTATTGGGTATATACC AAA
|
1375 (F IX)

58 (X chromosome test sequence)
|
GGATTATAAATTATTCTG.....................TGCTTATTTCAGCACT
| ||||||||| || |||                        || ||||| | |||||
GTATTATAAATCATGCTGCTATAAAGACACATGCACACGTATGTTTATTGCGGCACT
|
1431 (F IX)

92 (X chromosome test sequence)
|
ATTCACAATAGCAAAGACTTGAAACCAACCCAAATGCCCATCAATGATAGACTGGAT
 |||||||||||||| |||||| |||||||||||||| ||| ||||||||||||||||
TTTCACAATAGCAATGACTTGGAACCAACCCAAATGTCCAACAATGATAGACTGGAT
|
1488 (F IX)

149 (X chromosome test sequence)
|
AAAGAAAATTTGGCACATATACACCATGGAATACTATGCAGCAACAAAAAAGG.ATG
 |||||||| |||||||||||||||  ||||||||| ||||| | |||||    |||
TAAGAAAATGTGGCACATATACACCTAGGAATACTAGGCAGCCATAAAAAGAAAATG
|
1545 (F IX)

205 (X chromosome test sequence)
|
AGTTTCATGTCCTTTGCAGGGATGTGGATGAAGCTGGAAAACATCATTCTCAGCAAA
|||| |||||||||||| ||||   |||||||||| |||| ||||||||||||||||
AGTT.CATGTCCTTTGTAGGGCA.TGGATGAAGCTAGAAACCATCATTCTCAGCAAA
|
1602 (F IX)

262 (X chromosome test sequence)
|
CTAACATAGGAACAGAAAATCAAACTCGGCATGTTCTCACTCATAAATCGGAGTTGA
||| |  | | ||| |||| ||||| | ||||||||||||||||  | |||  |||
CTATCGCAAGGACAAAAAACCAAACACCGCATGTTCTCACTCATAGGTGGGAACTGA
|
1657 (F IX)

319 (X chromosome test sequence)
|
ACAACCAGAAGACATGGATGCAG....TGGAGCATCACACACTGGGGCCT TTGGGG
||||  |||| || ||||  |||     ||| |||||||||| ||||||| ||| ||
ACAATGAGAACACTTGGACACAGGAAGGGGAACATCACACACCGGGGCCTGTTGTGG
|
1714 (F IX)
```

Figure 7.10 Alignment of X chromosome consensus sequence (nucleotides 1 to 588) with region of F IX gene sequence containing a KpnI L1 element. The positions of oligonucleotide primers, XOA and XOB, used in the inverse PCR amplification of the der(X) junction fragment are indicated (see Chapter 6, section 6.2.3). Primer XOA is homologous to the F IX sequence at 17 out of 20 nucleotides if an adjustment of alignment of the sequences is made within the primer sequence.

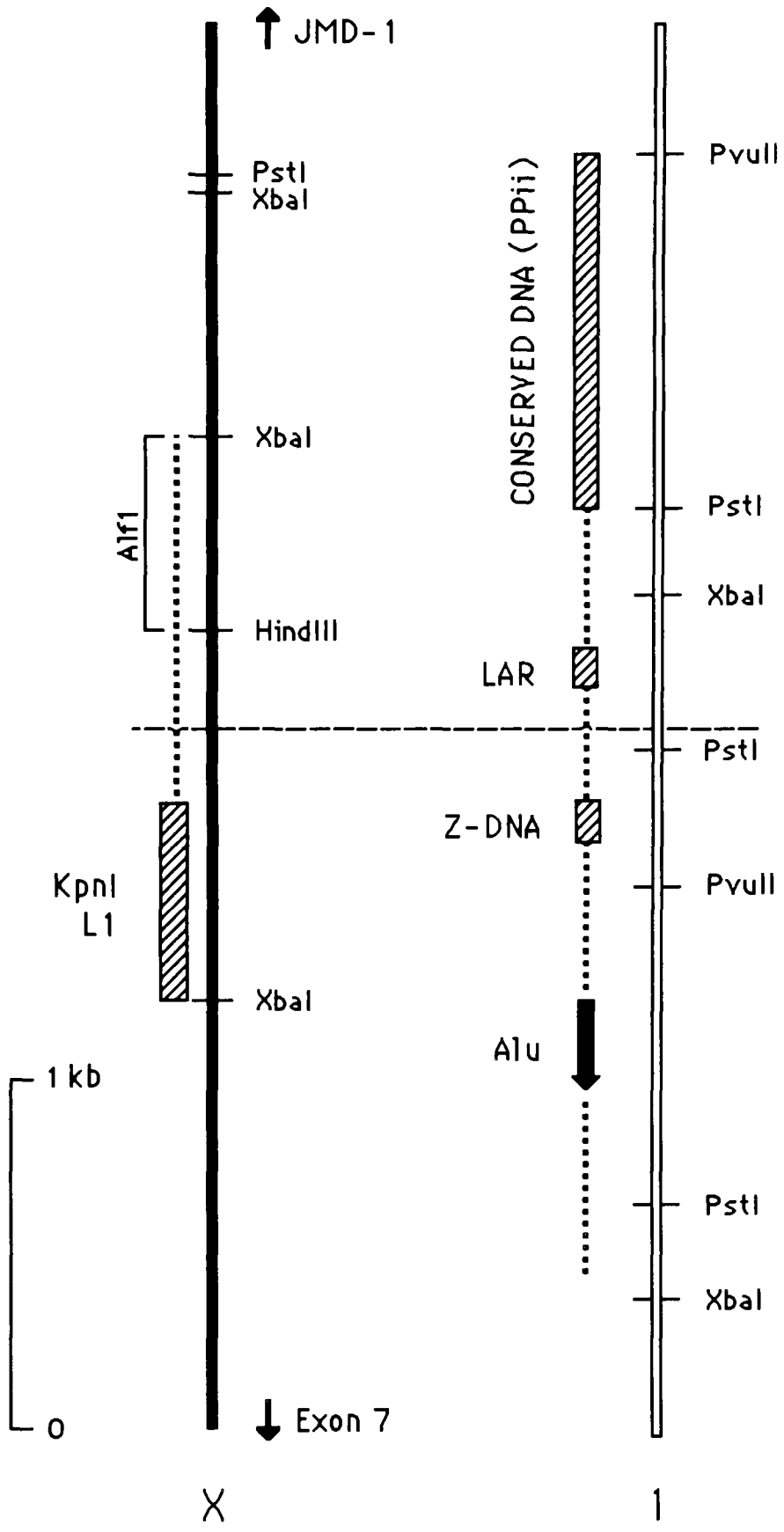


Figure 7.11 Positions of features within the DNA sequence surrounding the t(X;1) chromosome 1 and X chromosome breakpoint positions revealed by sequence analysis. Other DNA sequenced between these regions is indicated by dotted lines.

```
1895 (chromosome 1 test sequence)
|
CTGACCTCTCGCTTCGTGTACCCACAGGTGCCTTGCAGATAGAGAGCAGTGAGGAATCCG
  | || |       |   | |   ||||||||||||||||||||||||||||||||||||
ACGGCCGCATCAAGCAGCTGCGTTCAGGTGCCTTGCAGATAGAGAGCAGTGAGGAATCCG
|
882 (LAR)

1835 (chromosome 1 test sequence)
|
ACCAAGGCAAGTACGAGTGTGT****ACCAACTCGGCAGGCACACGTTACTCAGCCCCTG
||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
ACCAAGGCAAGTACGAGTGTGTGGCGACCAACTCGGCAGGCACACGTTACTCAGCCCCTG
|
942 (LAR)

1779 (chromosome 1 test sequence)
|
CGAACCTGTATGTGCGAGGTAAGGACTCAGGCAGTGCCTGGCCCCTGTCACCACAGAGCT
|||||||||||||||||       |   |    |   |   || |    ||    |||
CGAACCTGTATGTGCGAGTGCGCCGCGTGGCTCCTCGTTTCTCCATCCCTCCCAGCAGCC
|
1002 (LAR)
```

Figure 7.12 Comparison of a portion of the chromosome 1 consensus DNA sequence with a portion of the cDNA sequence of the leukocyte antigen related protein (LAR). A perfect match was seen between nucleotides 906 to 1019 of the LAR sequence with nucleotides 1871 to 1762 of the consensus chromosome 1 sequence. Asterisks indicate where a block in the sequencing gel prevented reading of the chromosome 1 sequence. Comparison with the LAR sequence indicated that four nucleotides were unreadable.

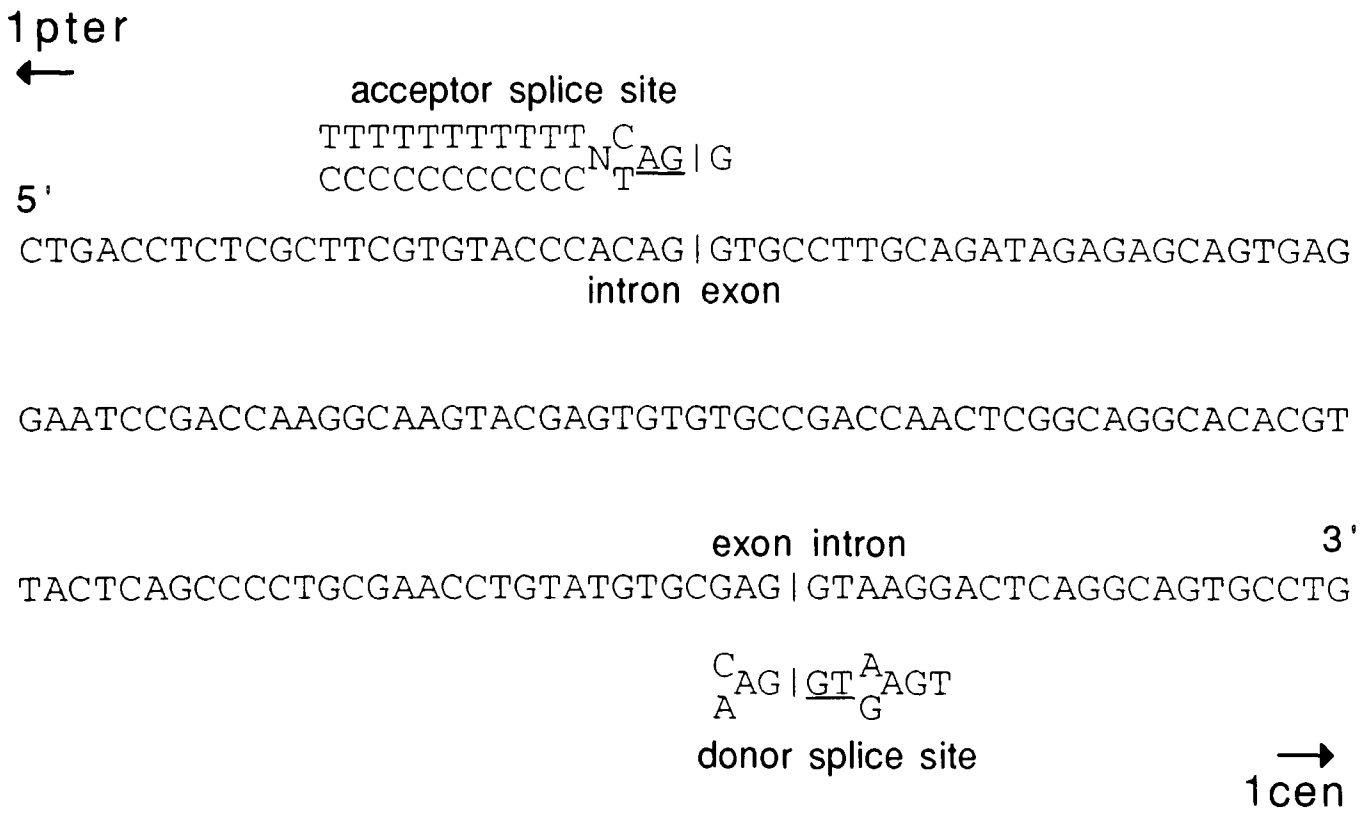


Figure 7.13 Genomic sequence around the LAR exon identified within the chromosome 1 consensus sequence (see figure 7.12). Potential exon/intron borders were identified by searching for homology of sequence with the consensus acceptor and donor splice sites (Mount, 1982; invariant nucleotides within these sites are underlined). The exon/intron borders indicate that the exon is between nucleotides 909 and 1019 of the LAR sequence (see figure 7.12).

LAR gene. The exon is 122 bp distal to the t(X;1) chromosome 1 breakpoint (figure 7.11).

A 62 bp stretch of alternating purine/pyrimidine nucleotides was identified which is 188 bp proximal to the chromosome 1 breakpoint (nucleotides 1387 to 1448; figure 7.7). 58 out of 62 bases match the alternating sequence (figure 7.14).

No significant direct or inverted repeats were identified in the chromosome 1 sequence.

7.2.2.3 Search for recombination signals in the sequenced DNA from the X chromosome and chromosome 1

A search was made for specific short DNA sequences which are known to be signals for recombination. Heptamer and nonamer signals specify V(D)J recombination in immunoglobulin genes (consensus sequences CACAGTG and ACAAAAACC; Sakano et al., 1979; Max et al., 1979). Studies show that the nonamer signal is not always required for recombination especially if the process is illegitimate (Baer et al., 1988; Boehm et al., 1988). Searching of the complete DNA sequences from the X chromosome and chromosome 1, revealed that perfect matches to the heptamer and nonamer signals are not present. Searching the complete chromosome 1 sequence revealed that there are 17 sites where 6 out of 7 nucleotides match the heptamer consensus. The closest of these to the chromosome 1 breakpoint is 32 bp proximal (nucleotides 1598-1604). A search of the complete X chromosome sequence detected six sites where 6 out of 7 nucleotides match the heptamer consensus. The closest of these to the X chromosome breakpoint is 164 bp proximal (nucleotides 610-616). Heptamer signals lie directly adjacent to the recombination sites in V(D)J recombination (Max et al., 1979). Therefore, the involvement of illegitimate V(D)J recombination can be ruled out in producing the X;1 translocation.

A search was made for the presence of the octamer GCTGGTGG which is the Chi site around which Rec-promoted recombination in bacteriophageλ is elevated (Smith et al., 1981). No perfect match was found in the complete X chromosome or chromosome 1

```
1387 (chromosome 1 sequence)                                        1448
|                                                                     |
GTATGCATGCATTTACACACAGATAAGTGTGCCCGTGTGTGCACACACACATGCACATACAC
pypypypypypypypypypypypypypypypypypypypypypypypypypypypypypypypy
```

Figure 7.14 Alternating purine (GA; p) / pyrimidine (CT; y) DNA segment located 188 bp proximal to the chromosome 1 breakpoint. Four bases which do not match the alternating sequence shown underneath are underlined.

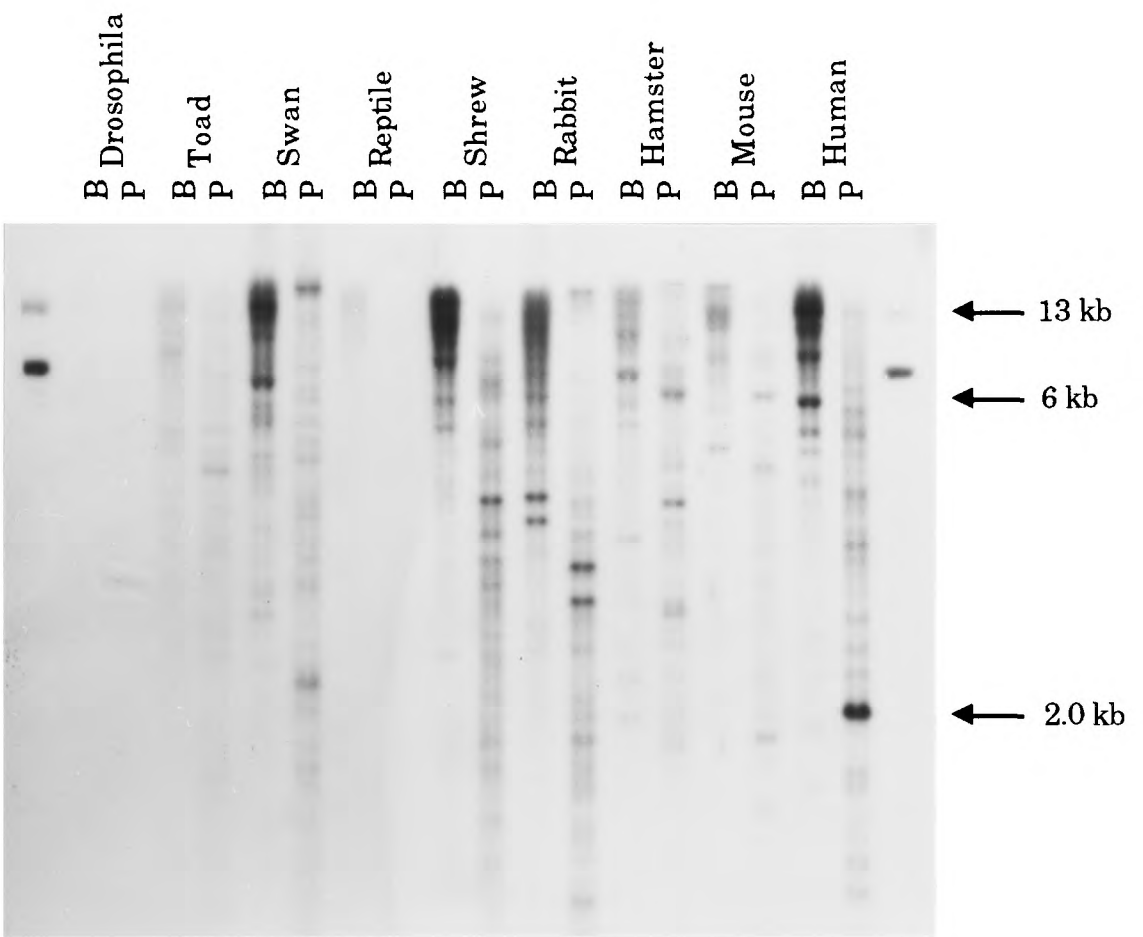


Figure 7.15 Hybridisation of PPii to a 'zoo blot' of BamHI (B) and PstI (P) DNA from 8 vertebrate and 1 invertebrate species. Hybridisation signals are visible in all tracks except those of the reptile (containing a small quantity of DNA) and of Drosophila. The final wash prior to autoradiography was performed using 200 mM phosphate buffer at 60°C. The principal bands in the human tracks (PstI, 2.0 kb; BamHI, 13 and 6 kb) are arrowed.

sequences. It is not known whether the Chi sites in λ are the same sequence as sites that stimulate recombination in the human, however Chi sites enhance recombination up to 10 kb away and therefore it is possible that an equivalent human signal has stimulated the X;1 translocation but lies outside the sequenced chromosomal regions. Alternatively an unknown recombination signal may lie in the vicinity of the t(X;1) breakpoints.

A search for the presence of the topoisomerase I and II consensus sequences (5' A/C/T, C/G, A/T, T 3' and 5' A/G, N, T/C, N, N, C, N, N, G, T/C, N, G, G/T, T, N, T/C, N, T/C 3' respectively; Spitzner and Muller, 1988) did not reveal significant homology.

7.2.2.4 Comparison of DNA sequence from around the chromosome 1 breakpoint region with that from the X chromosome region

A computer comparison of the complete X chromosome sequence with the complete chromosome 1 sequence detected no significant homology.

The X chromosome breakpoint of the t(X;1) is within the DMD locus which is within a dark staining G-band, Xp21, while the translocation breakpoint on chromosome 1 falls within a pale region of the chromosome following G-banding, at band 1p34 (see Chapter 5, figure 5.1). It is possible, however, that the DMD locus is within the pale sub-band Xp21.2 or that the chromosome 1 breakpoint is within the dark sub-band 1p34.2 (see Chapter 1, figures 1.2 and 1.4). A glance at the DNA sequences from the two breakpoint regions shows that they are very different in composition (figures 7.2 and 7.7). The GC content of the 3154 bp chromosome 1 sequence is 60% while that from the 1569 bp X chromosome sequence is 34%. The differences in composition of DNA from dark and pale G bands are discussed in section 7.3.1.

7.2.3 Investigation of evolutionary conservation of DNA from chromosome 1 around the breakpoint position

Sequence analysis indicated that the chromosome 1 breakpoint is within an intron

of the LAR gene (section7.2.2.2), however, the first indication that the breakpoint might be within an expressed locus was through the hybridisation of the chromosome 1 probe PPii to a Southern blot which included mouse DNA (Chapter 5, figure 5.22). The probe was observed to cross-hybridise to a 7.5 kb BamHI restriction fragment. Cross-species hybridisation indicates evolutionary conservation of the sequence and implies that the DNA has functional significance and may represent part of a structural gene or a pseudogene. In order to investigate whether PPii might also hybridise to species which diverged from the human earlier in evolution than the mouse, PPii was hybridised to a 'zoo-blot' containing DNA from several mammalian species, and representative species of reptile, bird, amphibian and insect (figure 7.15). Hybridisation signals were observed in all tracks except those containing DNA from the fruit fly, Drosophila melanogaster. Since sequence analysis revealed that PPii is within an intron of the LAR gene, the reason for conservation of DNA within PPii is unclear. One possibility could be that alternative splicing of the LAR transcript occurs, and that PPii does include an exon which is not included in the cDNA sequence held in the EMBL database. The 'zoo-blot' has not been probed with a DNA clone including the LAR exon. The clone 1CMA would be suitable for this purpose (see figures 7.6 and 7.11). This experiment would be expected to detect cross-species hybridisation similar to that detected using the probe PPii described above.

7.2.4 Estimation of the frequency of sequence errors caused by Taq Polymerase

The major disadvantage in using inverse PCR to clone the translocation breakpoints over traditional cloning techniques (see Chapter 6, section 6.1.1) is that the sequence data come from the products of amplification. It is recognised that during PCR, Taq Polymerase occasionally makes errors in DNA replication, usually A.T to G.C transitions, and that these errors are copied in subsequent rounds of amplification (Keohavong and Thilly, 1989; Belyavsky et al., 1989). The frequency of erroneous DNA replication is such that any cloned PCR product of around 1 kb is likely to include

a change in DNA sequence compared to that of the DNA template. Thus sequence derived from a single cloned PCR product may include such errors (Karlovsky, 1990). Thermostable DNA polymerases of higher fidelity have now been isolated from Thermococcus litoralis and Pyrococcus furiosus (available from Stratagene). The use of these enzymes would therefore be preferred to that of Taq Polymerase when the PCR product is to be cloned and sequenced. The problem of sequencing PCR errors can be overcome in two ways. Firstly, methods have been described which allow the direct sequencing of PCR products without cloning. This can been performed using double-stranded PCR products (Wrischnik et al., 1987; Newton et al., 1988). Alternatively single-stranded PCR products can be generated, either by asymmetric PCR where primers are used at dissimilar concentrations or by performing a secondary amplification using only a single nested primer (Gyllensten and Erlich, 1988; McCabe, 1990). Since the number of template molecules for DNA amplification provided by a typical initial DNA template concentration of 100 ng per reaction is high (approximately 20,000), errors which are made even in early PCR cycles and clonally copied will be undetectable against a consensus DNA sequence. Secondly, DNA sequence from more than one cloned PCR product can be obtained and a consensus sequence determined. This was the method I used to ensure against sequence errors of both translocation junction fragments and of both normal chromosome fragments around the breakpoints. The full sequence determined in this work is given in the appendix which shows where sequence was confirmed from more than one clone. The chance that two clones derived from a single PCR amplification are clonally related may be as low as 1 in 20,000 assuming that all 20,000 template molecules represented in 100 ng DNA are proportionally amplified. Thus two such clones are unlikely to contain the same errors. The sequencing of different clones of PCR products, permitted the frequency that PCR errors are present in the clones isolated in this work to be estimated. A total of six such errors were identified (one illustrated in figure 7.16) where a total of 5740 bp of sequence was available for comparison. Thus a PCR error is

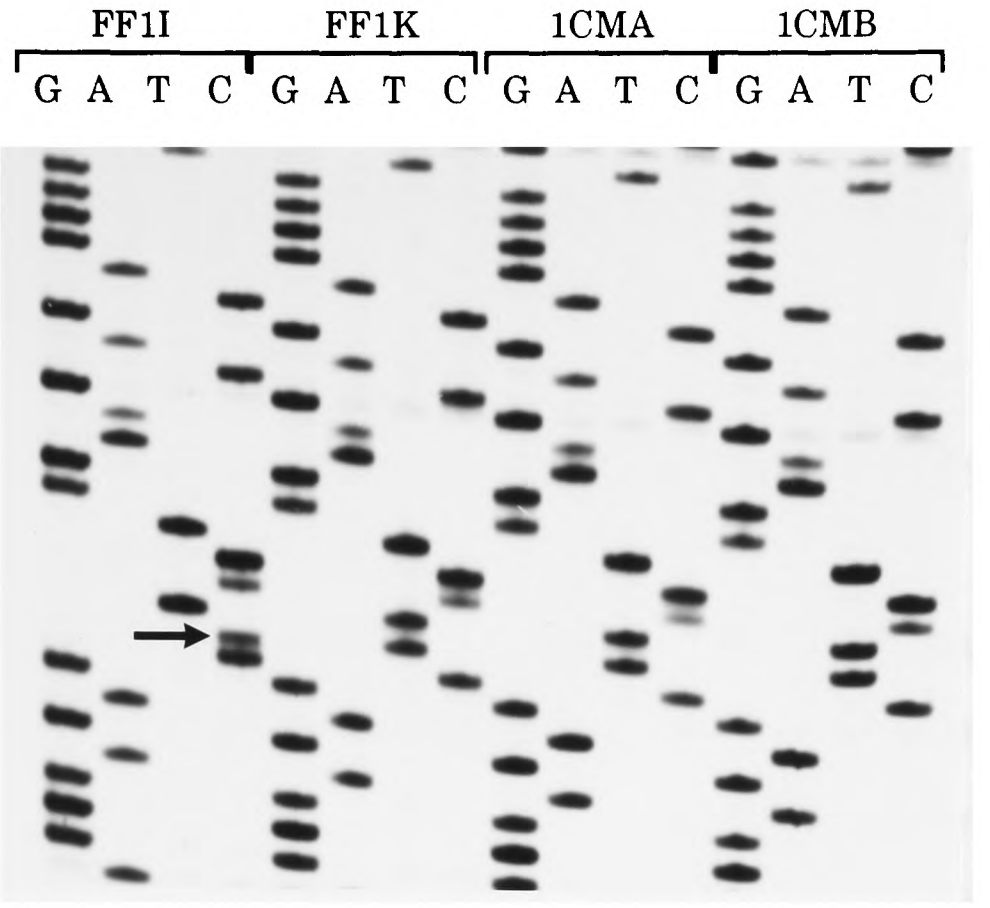


Figure 7.16 DNA sequence from four clones illustrating a PCR error which is a T to C transition (arrowed) within the sequence of FF1I determined by comparison of the sequence with those of FF1K, 1CMA and 1CMB.

```
5' C T A T C G A A G A G A G T G T T C T A A G T C T G A G A C T T C A A  3'  X
   | | | | | | | | | | | | | | | | |         |
  5' T A T C G A A G A G A G T G T T C C T G  3'                          JDER1
5' C T A T C G A A G A G A G T G T T C C T G C T T C T T T C T G C A G C  3'  Der(1)
   |                             | | | | | | | | | | | | | | | | | | |
5' C C C A G T C T C T T C T C C T T C C T G C T T C T T T C T G C A G C  3'  1
```

```
5' T C T C T T G A A G T C T C A G A C T T A G A A C A C T C T C T T C G  3'  X
   | | | | | | | | | | | | | | | | |       |                       |
  5' C T C T T G A A G T C T C A G A A G G A  3'                          JDERX
5' T C T C T T G A A G T C T C A G A A G G A A G G A G A A G A G A C T G  3'  Der(X)
      | |   |           | |  | | | | | | | | | | | | | | | | | | |
5' C T G C T G C A G A A A G A A G C A G G A A G G A G A A G A G A C T G  3'  1
```

Figure 7.17 Sequences of oligonucleotide primers JDER1 and JDERX and positions with respect to the junctions of the der(1) and der(X) chromosomes respectively. The nucleotides deleted from JDER1 and JDERX in the sequences of the primers DELJDER1 and DELJDERX respectively are underlined (see figure 7.19).

present in the cloned products on average every 957 bp or approximately every 1 kb which is in agreement with expectations. It should be noted that the DNA sequence which is derived from single cloned PCR products is likely to contain PCR errors at this frequency.

7.2.5 Verification of the sequence at the translocation junctions by PCR using primers homologous to the junctions

This experiment was performed in order to confirm the DNA sequence determined at the translocation breakpoints. The principle of the experiment was to perform a PCR amplification using one primer complementary to the DNA sequence determined at one of the translocation junctions and another primer at a known distance from the junction. If the predicted sequence at the junction was the true sequence then the primer would be able to anneal to template DNA derived from the t(X;1) patient and a product of the expected size would be amplified. A minor difference between the predicted and true sequences would be assumed to be sufficient to prevent DNA amplification. Separate reactions were performed for the der(X) and der(1) junctions.

A 20mer oligonucleotide PCR primer, JDERX, was designed from sequence determined at the der(X) junction (see figure 7.17). Four nucleotides at the 3' end of this primer are complementary to chromosome 1 sequence and the 16 nucleotides at the 5' end are complementary to X chromosome sequence. Another 20mer oligonucleotide PCR primer, JDER1, was similarly designed from sequence determined at the der(1) junction (see figure 7.17). Six nucleotides at the 3' end of this primer are complementary to chromosome 1 sequence while the 17 most 5' nucleotides are complementary to X chromosome sequence (three nucleotides, TTC, are of junctional homology and are complementary to DNA from both chromosomes; see section 7.2.1.5 and figure 7.9).

The locations of primers with respect to the translocation junctions are illustrated in figure 7.18. Primers JDERX and 1OF should amplify a 360 bp DNA segment from

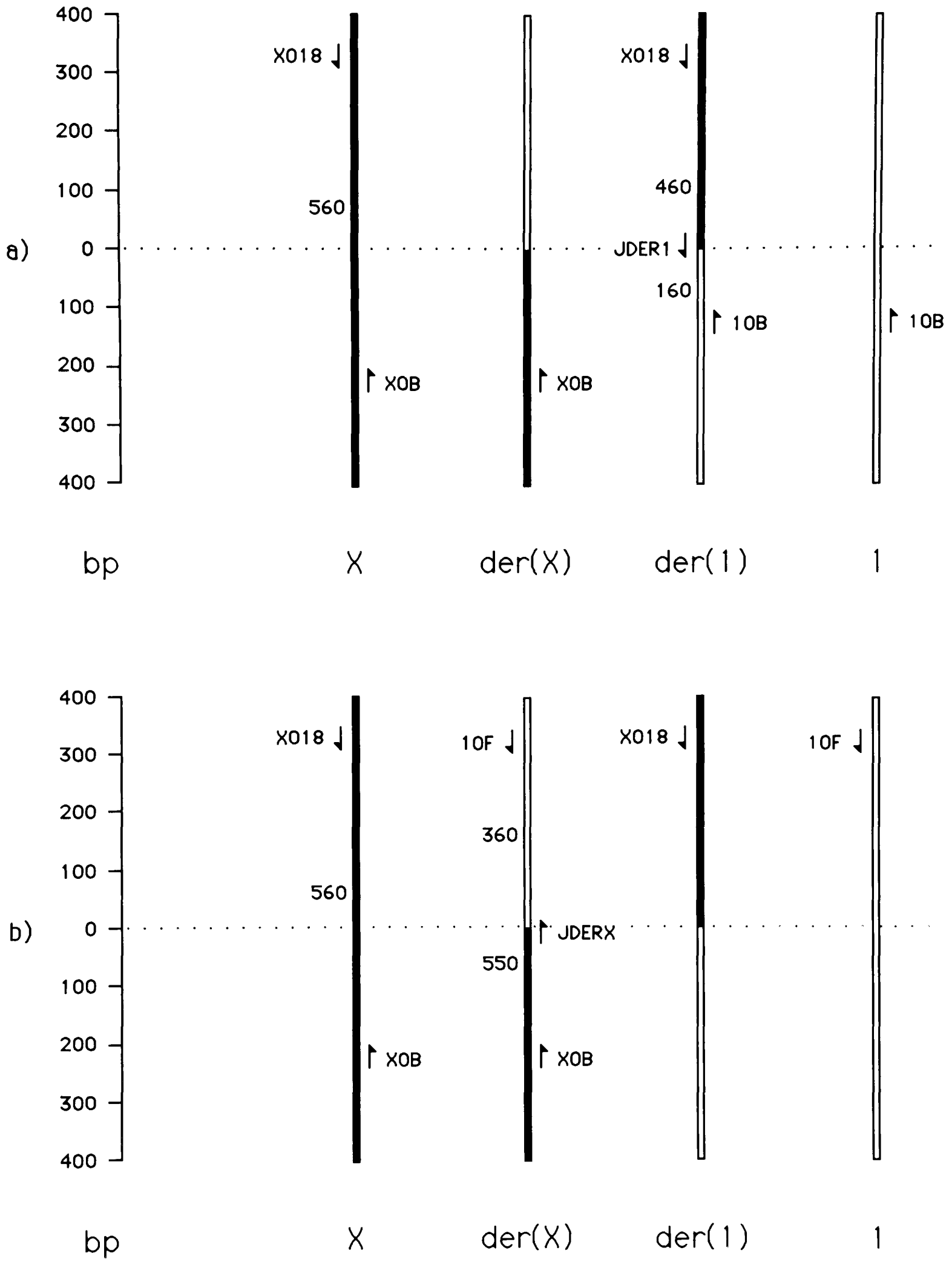


Figure 7.18 Strategy of PCR experiments to verify the DNA sequence at the der(1) (a) and der(X) (b) junctions. Dark lines represent X chromosome DNA and pale lines represent chromosome 1 DNA. The primers JDER1 and JDERX are homologous to DNA sequence determined at the der(1) and der(X) junctions respectively. The expected sizes of amplified PCR product from each chromosome are indicated.

the t(X;1) patient, and primers JDER1 and 1OB a 160 bp product from the patient. Another pair of primers, XO18 and XOB, were included in reactions as a positive control, predicted to amplify a 560 bp DNA segment from the normal X chromosome. This product would be expected whether the template DNA was from the t(X;1) patient or from a normal individual. The sites that XO18 and XOB anneal to on either side of the X chromosome breakpoint are indicated in figure 7.18. These locations predict that further amplification products would be expected from the t(X;1) patient's DNA. Primers 1OF and XOB in one experiment would be expected to amplify a 550 bp DNA segment representing the der(X) junction; while 1OB and XO18 in the other experiment would be expected to amplify a 460 bp DNA segment representing the der(1) junction (see figure 7.18). Thus the amplification of these fragments would indicate that the junction fragments are of the sizes predicted by sequence analysis.

Negative control reactions were included in the experimental design. In one reaction, a 19mer oligonucleotide primer DELJDERX was substituted for the primer JDERX. This primer is identical to JDERX except that it includes a deletion of the guanine which is six nucleotides from the 3' end of JDERX (see figure 7.17). The consequent change in homology to the predicted sequence of the der(X) is illustrated in figure 7.19. In another reaction, a 19mer oligonucleotide primer, DELJDER1, was substituted JDER1. This primer was identical to JDER1 except that it included a deletion of the thymine which is six nucleotides from the 3' end of JDER1 (see figure 7.17). The consequent change in homology of this primer to the predicted sequence of the der(1) is illustrated in figure 7.19. The intention of including these deletions was to destabilise the annealing of the 3' ends of DELJDERX and DELJDER1 to the DNA template derived from the t(X;1) patient and thus prevent Taq DNA polymerase extension from these primers. Therefore if no product was observed in these reactions, this would indicate that although the sequences of DELJDERX and DELJDER1 are very similar to the predicted sequences of the translocation junctions, they are sufficiently dissimilar to the true sequence to prevent primer annealing and DNA amplification.

```
          5' T A T C G A A G A G A G T G T C C T G  3'          DELJDER1
             | | | | | | | | | | | | | | | | |  |
5' A A A T C T A T C G A A G A G A G T G T T C C T G C T T C  3'   Der(1)


          5' C T C T T G A A G T C T C A A A G G A  3'          DELJDERX
             | | | | | | | | | | | | | | | |  |  |
5' A C C A T C T C T T G A A G T C T C A G A A G G A A G G A  3'   Der(X)
```

Figure 7.19 Sequences of the oligonucleotide primers DELJDER1 and DELJDERX and comparison with the sequences at the junctions of the der(1) and der(X) chromosomes respectively.

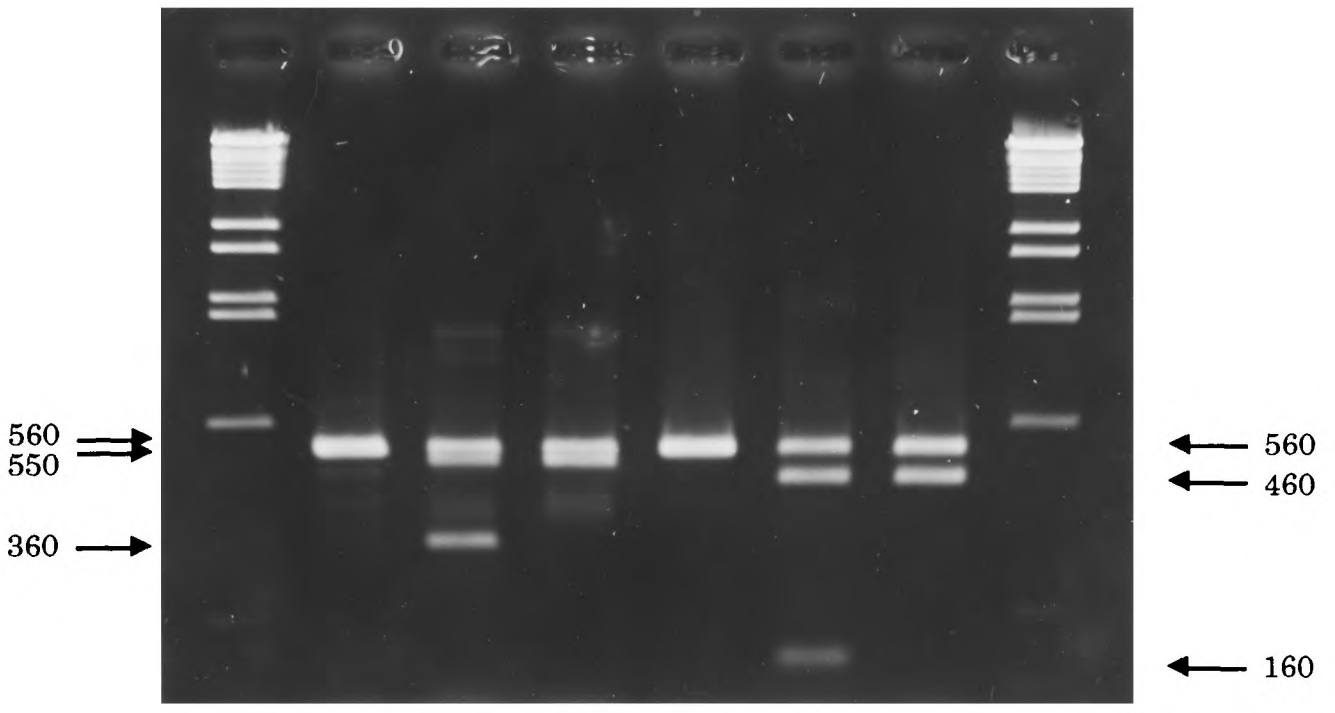


Figure 7.20 Products of amplification from six PCRs designed to verify the sequence determined at the junctions of the der(1) and der(X) chromosomes. Tracks 1 and 8 are BstEII digested λ DNA. The description of the PCRs whose products are illustrated in tracks 2 to 7 is given in the text. The sizes of the principal products of amplification in tracks 2 to 4 are shown on the left of the figure and those of the principal products of amplification in tracks 5 to 7 are shown on the right.

A gel of the ethidium bromide-stained products of six PCR amplifications is illustrated in figure 7.20. An intense 560 bp band was observed in each track. This was the expected size of DNA segment from the normal X amplified between the control primers, XO18 and XOB, which were included in each reaction. Tracks 2 and 3 contain DNA products of amplification using primers XO18, XOB, 1OF and JDERX. In track 2 the template DNA was from a normal individual. No strongly staining band apart from the control 560 bp fragment is evident. In track 3 the template DNA was derived from the t(X;1) patient. Two strongly staining bands are visible, in addition to the 560 bp product derived from the patient's normal X chromosome. These correspond in size to the expected approximately 550 bp and 360 bp fragments which represent the der(X) junction amplified between 1OF and XOB, and the product amplified between 1OF and JDERX respectively (figure 7.18). The amplification of the 550 bp fragment confirms the size of this DNA fragment as predicted by DNA sequencing. The amplification of the 360 bp fragment indicates that the DNA sequence determined at the der(X) translocation junction must be close enough to the true sequence (if it is different at all) to permit stable binding of the primer JDERX and amplification of the fragment. The reaction illustrated in track 4 was identical to that of track 3 except that JDERX was replaced with DELJDERX. The 360 bp fragment did not amplify in this reaction which indicates that a very small change to the 3' sequence of the primer was sufficient to prevent stable primer binding and DNA amplification. This is despite the amplification of the 550 bp fragment which could additionally serve as a template for amplification by 1OF and DELJDERX (see figure 7.18). Thus this result further indicates that DNA sequence determined for the der(X) junction is probably the true sequence. Tracks 5 and 6 contain DNA products of amplification using primers XO18, XOB, JDER1 and 1OB. In track 5 the template DNA was from a normal individual. The 560 bp control product is the only strongly staining band. The reaction in track 6 used the t(X;1) patient's DNA as the template. Two strongly staining bands are visible in addition to the control 560 bp band. These correspond in size to the expected

fragments of approximately 460 bp and 160 bp which represent the der(1) junction amplified between the primers XO18 and 1OB and the product amplified between JDER1 and 1OB respectively. The presence of the amplified 460 bp band confirms that this junction fragment is the size predicted from DNA sequencing. The presence of the amplified 160 bp band indicates that the determined DNA sequence at the der(1) translocation junction site is close enough to the true sequence to allow stable annealing of the oligonucleotide JDER1 and amplification of the 160 bp fragment. The reaction illustrated in track 7 is identical to that in track 6 except that JDER1 was replaced with DELJDER1. The effect was that the 160 bp fragment did not amplify. Thus the sequence of DELJDER1 must be sufficiently different from the true der(1) junction sequence to prevent stable annealing of this oligonucleotide and amplification of the 160 bp fragment. This is despite the amplification of the 460 bp DNA fragment which could serve as a template for amplification by the primers DELJDERX and 1OB (see figure 7.18). Thus this result further indicates that DNA sequence determined for the der(1) junction is probably the true sequence.

Amplification of the 550 bp der(X) and 460 bp der(1) junction fragments provided a large quantity of template for amplification of the 360 bp fragment by the primers JDERX and 1OF, and of the 160 bp fragment by the primers JDER1 and 1OB respectively (see figure 7.18). However, it was also confirmed that the 360 bp and 160 bp amplification products were present after PCR using these primer pairs alone in the absence of the control primers XO18 and XOB (figure 7.21).

7.3 Discussion

7.3.1 Discussion of the findings from sequencing the X;1 translocation

The main part of the work described in this thesis concerns the detailed analysis of a single constitutional translocation, and this chapter has described the investigation of this translocation at the finest level, that of the DNA sequence. Examination of the sequence at the breakpoints of this translocation and comparison with that at the

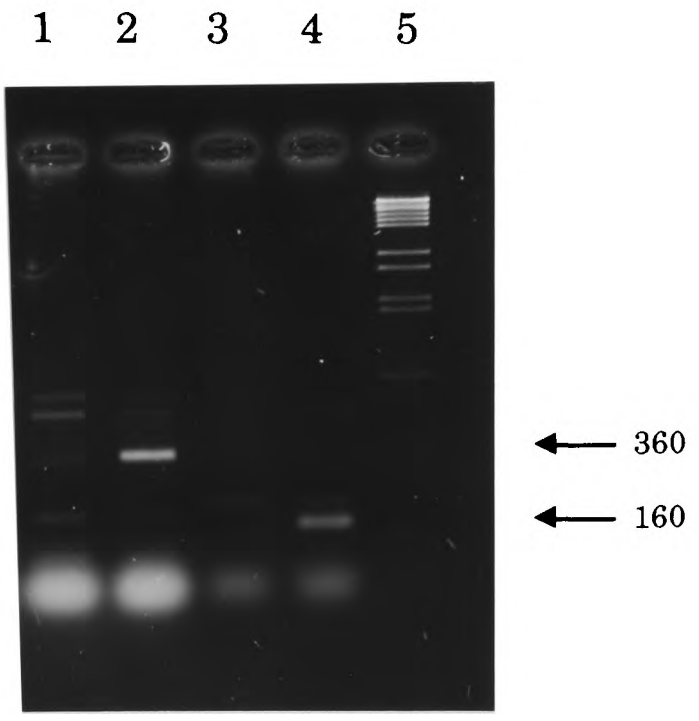


Figure 7.21 Products of PCR amplification using the primers JDERX and 1OF (tracks 1 and 2) and JDER1 and 1OB (tracks 3 and 4). The template DNA was from a normal female in tracks 1 and 3 and from the t(X;1) patient in tracks 2 and 4. A DNA fragment of 360 bp is visible in track 2, and a fragment of 160 in track 4. These fragments are of the expected size (see figure 7.18). Track 5 is BstEII digested λ DNA.

breakpoints of others is the most likely course to an understanding of the mechanisms producing these chromosomal rearrangements. However there are very few comparable studies in the literature. In this discussion, the results of the analysis of the t(X;1) are explored, the findings of similar investigations are discussed and the mechanisms involved in producing the rearrangements are considered. Although features of structural chromosome rearrangements studied at DNA sequence resolution can be compared, it is often not possible to draw firm conclusions owing to the small number of studies so far performed.

Figure 7.11 illustrates the positions of the following segments of DNA in relation to the breakpoints of the X;1 translocation. The X chromosome breakpoint is approximately 10 kb distal to exon 7 of the dystrophin locus (see Chapter 5, section 5.2.4). A LINE sequence has been identified which is 210 bp proximal to the X chromosome breakpoint. The chromosome 1 breakpoint is also within an intron of an expressed locus, that which codes for LAR. The breakpoint is 122 bp proximal to a 110 bp exon. A repetitive DNA element is also close (750 bp proximal) to the chromosome 1 breakpoint, but it is an Alu sequence which is a SINE. A 62 bp stretch of potential Z-DNA is 188 bp proximal to the same breakpoint. There is a 1 kb single-copy DNA segment, PPii, which is 630 bp distal to the chromosome 1 breakpoint and which appears to be conserved in evolution although apparently does not contain an exon of the LAR gene.

The localisation of the gene for LAR has already been made by in situ hybridisation of a cDNA probe and is at 1p34-p32 (Disteche et al., 1989). This is consistent with the cytogenetic characterisation of the X;1 translocation by Lindenbaum et al. (1979) and by Boyd and Buckle (1986) who determined the chromosome 1 breakpoint was at band 1p34. It is also consistent with the localisation by in situ hybridisation of a cloned DNA segment, PPii, which is adjacent to the X;1 translocation breakpoint to the same band (described in Chapter 5, section 5.2.7). Thus all available evidence suggests that the chromosome 1 breakpoint of the X;1

translocation is adjacent to an exon of the LAR gene. The cloning of genomic segments from the LAR gene and of its cDNA is described by Streuli et al. (1988). The cDNA sequence held in the EMBL database is 7762 bp in length and the putative exon adjacent to the t(X;1) breakpoint is from nucleotide 909 to 1019. Therefore the t(X;1) breakpoint is within the 5' portion of the gene. Since the orientation of the chromosome 1 sequence around the t(X;1) breakpoint is known with respect to the centromere, the orientation of the LAR gene can be determined to be with the 5' end distal and the 3' end proximal to the chromosome 1 breakpoint (see figure 7.13). The mRNA of LAR codes for a protein of 1,897 amino acids which is approximately half the number found in dystrophin (3,685; Koenig et al., 1988). The protein encoded by LAR is thought to be a transmembrane protein with Immunoglobulin (Ig)-like and neural cell adhesion molecule (N-CAM)-like domains in the extracellular portion while the cytoplasmic region is homologous to the leukocyte common antigen (LCA). The putative exon adjacent to the t(X;1) chromosome 1 breakpoint corresponds to amino acids 164-200 as determined from Streuli et al. (1988). This is within the Ig-like domain of LAR.

The finding that the chromosome 1 breakpoint is within an expressed gene was unexpected given that genes are thought to represent approximately 5% of human genomic DNA (Singer and Berg, 1991). Therefore the possibility that DNA within expressed loci is prone to chromosomal rearrangement deserves consideration. This question will only be answered when further translocations have been sequenced. It is interesting that the autosomal breakpoint of the X;21 translocation, which is the only published human constitutional translocation that has been sequenced, is within the rRNA gene cluster on the short arm of this chromosome. However, this translocation was selected for study because of the autosomal breakpoint position being within this defined locus. The X;1 and X;21 translocations were ascertained owing to their X chromosome breakpoint positions within the dystrophin locus, therefore, of the four breakpoints which are all within expressed loci, only the autosomal breakpoint position of the t(X;1) was unexpected. If expressed loci are a preferred target for translocations,

then the reason for this is unclear given that most genes are not transcriptionally active during the later stages of spermatogenesis, which appears to be the most likely timing of the rearrangement (Erickson, 1990; see Chapter 8, section 8.1). On this point it is interesting to note that the rRNA genes which represent the nucleolar organising regions (NORs) on the short arms of human acrocentric chromosomes are active during pachytene, which is the stage of homologous chromosome pairing and chiasma formation (Evans et al., 1974). The association of NORs at pachytene has been suggested to play a role in the prevalence of Robertsonian translocations in the human (Ferguson-Smith, 1967; Krystal et al., 1981; Cheung et al., 1990).

The presence of an Alu sequence 750 bp proximal to the t(X;1) chromosome 1 breakpoint does not appear to have significance with respect to the rearrangement, given that no Alu sequence is close to the X chromosome breakpoint and that Alu sequences are found on average in every 4 kb human DNA (Britten et al., 1988; approximately 4.5 kb DNA around the breakpoint regions has been sequenced in the study described here). There are several reported instances, however, where unequal recombination between Alu sequences has been identified as the cause of deletions (Ariga et al., 1990; Ottolenghi and Giglioni, 1982; Jagadeeswaran et al., 1982). A similar mechanism has been proposed for deletions causing growth hormone-1 deficiency through unequal recombination between 594 bp and 454 bp homologous DNA segments that flank the disease locus (Vnencak-Jones and Phillips, 1990). Unequal recombination between highly similar sequence elements of at least 10 kb in length and which lie 1900 kb apart appears to be the major mechanism responsible for deletions of the STS locus resulting in X-linked ichthyosis (Yen et al., 1990). The identification of a LINE element 210 bp from the X chromosome breakpoint is unlikely by chance. LINEs are present approximately every 50-100 kb and represent approximately 4% of the human genome (Deininger and Daniels, 1986). A LINE was identified close to one of the breakpoints of a ring chromosome 21 (discussed in section 7.3.2) but was not seen near the breakpoints of the X;21 translocation. The sequencing

of further translocations and other chromosome rearrangements will establish whether the presence of LINEs at the breakpoints is a consistent feature.

The identification of a 62 bp stretch of alternating purine/pyrimidine (potential Z-DNA) 188 bp from the chromosome 1 breakpoint is one further finding whose significance is unclear. It has been suggested that alternating purine/pyrimidine tracts promote lymphoid tumour chromosome translocations which result from erroneous V(D)J recombinase activity (Boehm et al., 1989). Bacterial plasmids containing alternating purine/pyrimidine sequences have been found to be prone to deletion (Freund et al., 1989). Antibodies to Z-DNA bind to pale G-bands in the human, indicating a non-random distribution of these sequences, however their functional role is still unknown (Viegas-Péquignot et al., 1983). Proteins have been detected which bind to such regions (Fishel et al., 1988; Leith et al., 1988). These might cause local chromatin disruption and allow recombinase access.

The alternating purine/pyrimidine sequence detected close to the t(X;1) breakpoint includes a number of repeated CA dinucleotides (see figure 7.14). Five CA repeats is the longest uninterrupted run (nucleotides 1428-1437). The number of dinucleotide CA repeats in such runs has frequently been found to be polymorphic, especially when the number of repeats is large (for example Powell et al., 1991; Feener et al., 1991). Although the number of repeats present in the chromosome 1 sequence described here is smaller than in the above reports (the number of repeats reported by Powell et al. was 17 on the chromosome sequenced), primers are currently being designed to test for the possibility of polymorphism in the number of CA repeats at this chromosome 1 locus.

Several differences are known to exist between the composition of DNA within dark and pale G-bands (Bernardi, 1989). These include GC content, the distribution of repetitive DNA elements and Z-DNA sequences, and the distribution of genes. For example, the GC content of dark G-bands is usually around 40-41% and that of pale G-bands in the range 46-52% (Gardiner et al., 1990). These values are calculated over compositionally homogenous segments of DNA, called isochores, which are around 500

kb in length (Bernardi, 1989). Therefore the GC content of the relatively short segments of DNA sequenced in this study from the X chromosome and chromosome 1 may be different from that of the isochores within which they lie. The GC content of the DNA sequenced around the X chromosome breakpoint is 34% and that around the chromosome 1 breakpoint is 60% (section 7.2.2.4). The finding that the GC content is low in the X chromosome DNA and high in the chromosome 1 DNA is consistent with the assumption that the DMD locus is within a dark G-band and that the LAR locus is within a pale G-band (section 7.2.2.4). The majority of genes are located in isochores of the highest GC content (53% GC in man) although these represent only 3-5% of the genome (Bernardi and Bernardi, 1985). However the GC content of DNA from within a gene is not usually the same as that of the isochore within which the gene lies. Typically, a gene with a GC content of around 60% would be expected to lie within an isochore whose GC content is around 50% (Bernardi, 1989), therefore the high GC content (60%) observed around the chromosome 1 breakpoint of the t(X;1) would appear mainly to be due to the presence of the LAR gene at this site.

The distribution of repetitive DNA elements within the human genome is non-random. LINES are predominantly located in dark G-bands or isochores of low GC content and SINES within pale G-bands or isochores of high GC content (Singer, 1982; Soriano et al., 1983; Holmquist and Caston, 1986). The finding of a KpnI L1 element (a LINE) close to the X chromosome breakpoint and of an Alu sequence (a SINE) close to the chromosome 1 breakpoint is consistent with the assignment of these breakpoints to a dark and pale G-band respectively.

7.3.2 Comparison of the sequence of the t(X;1) with other chromosome rearrangements

There is one published description of the sequence at the breakpoints of a constitutional translocation in man (Bodrug et al., 1987). This is another X;autosome translocation whose X chromosome breakpoint is within the DMD locus and which is associated with disease expression (Verellen-Dumoulin et al., 1984). The breakpoints

were cloned before other DNA segments from the DMD locus had been isolated, following the observation that the chromosome 21 breakpoint was within a cluster of ribosomal RNA genes on the short arm of this chromosome (Ray et al., 1985). Sequence analysis revealed that deletions at both the X chromosome breakpoint (71 to 72 bp) and the chromosome 21 breakpoint (16 to 23 bp) were associated with the translocation. These deletions are larger than the 4 to 7 bp deletion and 2 to 5 bp duplication which have been found at the t(X;1) breakpoints described here. The most significant finding of sequence analysis of the t(X;21) was the presence of a trinucleotide (GGC) repeat at the chromosome 21 breakpoint. The number of repeats is polymorphic in the population and is between four and six. Within the deleted DNA from the X chromosome region, a tetranucleotide sequence CGGC was found which is homologous to the trinucleotide repeat sequence from chromosome 21. It was suggested that these sequences could be the recognition sites for an enzyme involved in the translocation process. Four repeated trinucleotides (TGC; see figure 7.9) have been found in this study close to the chromosome 1 breakpoint of the t(X;1), although no similar sequence is present on the X chromosome. Two of the nucleotides in this trinucleotide are homologous to the trinucleotide repeat found at the X;21 breakpoint. Another trinucleotide (GAA; see figure 7.9) has also been found to be present four times within the chromosome 1 sequence and three times within the X chromosome sequence close to the t(X;1) breakpoints. It will be interesting to note whether this feature is present at the breakpoints of other translocations when they are sequenced. Alternating purine/pyrimidine sequences (potential Z-DNA) which often consist of dinucleotide CA repeats, have been observed near the breakpoints of translocations associated with lymphoid tumours and these translocations are thought to be caused by erroneous V(D)J recombinase activity (Boehm et al., 1989). The short trinucleotide repeats close to the X;21 and X;1 breakpoints might promote the activity of another recombinase or represent the recognition signal of such an enzyme.

Apart from the X;21 translocation, the only other published sequences from

mutations within the DMD locus are from two deletions and one point mutation (Love et al., 1991; Bulman et al., 1991). One of the deletions described by Love et al. (1991, patient GP, whose deletion encompasses exons 45 to 50) shares two features with the t(X;1). Three nucleotides of junctional homology are present at the deletion junction, and a 28 bp stretch of pyrimidine nucleotides is present 10 bp from the junction (the nucleotides of the opposite DNA strand will be purines). Henthorn et al. (1990) have reviewed studies on the sequence of 21 deletions within the β-globin gene cluster. Thirteen contain from 1 to 6 nucleotides of junctional homology which they report is more than would be expected by chance. Henthorn et al. devised a potential scheme for generating deletion junction sequences with junctional homology which is reproduced in figure 7.22(a). In figure 7.22(b), an analogous process that could have generated the structure of the X;1 translocation is illustrated. The process involves chromosome breakage to produce single-stranded overhangs, degradation/deletion of some nucleotides, association of bases within the single-stranded overhangs that are of junctional homology and primed DNA synthesis and ligation to fill the single-stranded gaps and repair the derivative chromosome junctions. As well as providing an explanation for the presence of bases of junctional homology, this model accommodates the deletion/duplication of nucleotides which is associated with the translocation.

One further constitutional structural chromosome rearrangement that has been sequenced is a de novo ring chromosome 21 (r(21) (Wong et al., 1989). It is a complex rearrangement which involves a duplication of the centromere, proximal short arm and proximal long arm. The sequenced junction is formed by the fusion of two long arm breakpoints. The most significant finding of this investigation was that at one breakpoint a 17mer sequence and at the other breakpoint a 13mer sequence were identified which matched perfectly sequences 335 bp and 557 bp from the opposite breakpoints respectively. One feature of this rearrangement in common with the t(X;1) is that homology to a LINE repeat was observed close to one of the breakpoints. In the

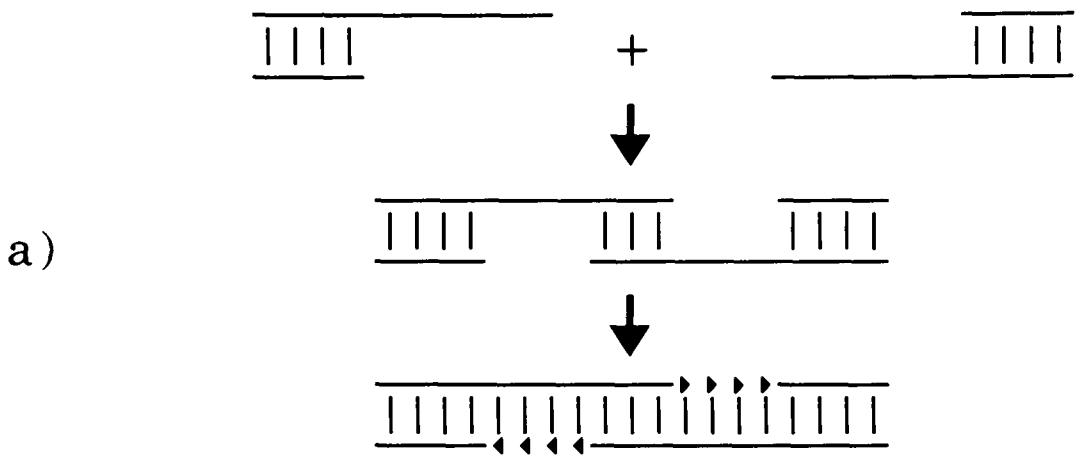


Figure 7.22 a) Scheme for generating deletions with nucleotides of junctional homology described by and reproduced from Henthorn et al. (1990). b) Analogous process which may explain the observed sequence at the junctions of the X;1 translocation described here. Further details are given in the text.

case of the r(21) this was 1 kb from one of the breakpoints while the distance from the LINE repeat to the X chromosome breakpoint of the t(X;1) was 210 bp.

The normal chromosome sequences compared with those of the derivative chromosomes in this chapter were not those of the father's chromosomes which rearranged to form the translocation (see section 7.1). Therefore the possibility that the deletion and/or duplication of a few nucleotides observed at the breakpoints of the derivative chromosomes were already present in the father's chromosomes before the translocation occurred cannot be ruled out. Such a deletion or duplication, if present, would represent a deletion/insertion polymorphism if it were also present in other individuals in the population. If it were not, then it might be considered to be a premutation in so far as the translocation is concerned. It would not be easy to test for the presence of the deletion and duplication in the father's chromosomes without cloning and sequencing the relevant DNA fragments. An additional difficulty would be encountered in identifying the relevant chromosome 1. The breakpoints of the X;21 translocation sequenced by Bodrug et al. (1987) were found to be associated with deletions of nucleotides from both the X chromosome and chromosome 21. In this case the sequences of the derivative chromosomes were compared with the sequence of the father's X chromosome and RFLP analysis had shown that the origin of the translocation was paternal (Kean et al., 1986). Therefore the deletion of X chromosome nucleotides in this translocation was shown to be associated with the translocation event.

In this chapter, the DNA sequence of a constitutional X;autosome translocation has been described. Features around the breakpoints have been identified which may have played a role in promoting the rearrangement. This information, along with that from a few similar investigations, provides a basis for the aetiology of mammalian constitutional chromosome rearrangement.

7.4 Acknowledgements

I am grateful to Glen Gaughan for technical assistance in performing DNA sequencing. I am also extremely grateful to Chris Porter for sequencing the X chromosome clone Alf1, and for time consuming and invaluable help in conducting DNA sequence analysis.

Chapter 8 General discussion and summary

The work described in Chapter 7 revealed features of the DNA sequence in the vicinity of the breakpoints of the t(X;1) which may have played a role in promoting the generation of the translocation. In this final chapter, the types of mechanism which may have caused the rearrangement are considered and the possible timing of the event is discussed.

8.1 The mechanisms involved in generating translocations

The process which produced the translocation described in this work must have involved two stages. Firstly, a breakage of the DNA strands and secondly, a rejoining or resolution stage. The first stage could be caused by damage or an enzymatic process while the second stage would have to be an enzymatic event. Therefore it would seem reasonable to consider two types of process. One is where the translocation is the result of erroneous repair to DNA damage. The other is where it is the result of a mistaken normal recombination mechanism. It would be difficult to examine the first of these possibilities, as damage may be expected to occur at any site and repair enzymes may be expected to repair damaged DNA of any sequence. However a recombinase may, and in some cases is known to, recognise specific signals in the DNA sequence. Two normal recombination processes that act on native chromosomes are meiotic recombination and V(D)J recombination which acts on the immunoglobulin genes of maturating lymphocytes (Whitehouse, 1973; Sakano et al., 1979). A possible timing for the translocation at meiosis would seem attractive, however there is evidence that cells with X;autosome translocations do not survive meiosis (see below). Meiotic recombination is known to be non-random as measured by chiasma distribution or comparison of physical and genetic chromosome maps (Laurie and Hultén, 1985; Steinmetz et al., 1987; Chakravarti, 1991), but it is not known whether specific DNA signals are involved such as the octomer Chi recombination hotspot signal in the bacteriophage λ and the 24 bp signal which stimulates homologous recombination in

yeast (Smith et al., 1981; Nickoloff et al., 1986). If the normal recombination process at meiosis frequenctly involves duplications and deletions as have been found at the X;1 and X;21 translocations, then this would appear to be a significant source of polymorphic variation and mutation in the genome and seems improbable.

The involvement of erroneous V(D)J recombinase in the production of many translocations associated with haematopoietic malignancies is well documented (Boehm et al., 1988; Croce, 1987; Haluska et al., 1987). The recombinase appears to join mistakenly the J segment of an immunoglobulin gene to a proto-oncogene. Sequence analysis has identified sequences similar to the heptamer and nonamer signals recognised by V(D)J recombinase at the breakpoints adjacent to the proto-oncogene. These events are brought to attention since they result in disease, however, the same mechanism would be presumed to join the J segments of immunoglobulin genes to other sites in the genome. A recent study has found that V(D)J recombinase also may cause deletions in the HPRT gene in human foetal T-cells, since sequences similar to the heptamer and nonamer signals recognised by V(D)J recombinase were identified at the breakpoints (Fuscoe et al., personal communication). Thus it appears that V(D)J recombinase occasionally produces chromosome rearrangements not involving the immunoglobulin genes. However no involvement has yet been identified which is in a cell type other than the lymphocyte and therefore there is no evidence as yet that V(D)J recombinase could cause constitutional chromosome rearrangements.

The parental origin of the X chromosome involved in the X;1 translocation is paternal (Chapter 4, section 4.2.5). A paternal origin is a feature common to all 13 X;autosome translocations so far studied in this respect (Robinson et al., 1990; Verga et al., 1991). This finding suggests that the timing of the majority of these translocations is unlikely to be after zygote formation since a mechanism occurring after this stage would be expected to involve maternal and paternal chromosomes equally. Therefore the predominant stage for inherited X;autosome translocation would appear to be in the male germline. Meiotic studies on males with constitutional X;autosome

translocations indicate that the cause of sterility in these individuals is spermatogenic arrest at meiosis I (Faed et al., 1982; Quack et al., 1988). This suggests that pre-meiotic cells carrying X;autosome translocations are unlikely to survive meiosis and therefore that the most likely timing of paternally inherited X;autosome translocations is after meiosis I. Chandley (1991) concludes that paternally inherited de novo X;autosome translocations almost certainly arise in the haploid spermatid or spermatozoan. The preferential paternal origin of X;autosome translocations that is observed may reflect the differences between the male and female meiotic processes (summarised in Chandley, 1991; Hultén et al., 1985). Alternatively, if the predominant cause of X;autosome translocations is radiation damage and mis-repair, then the position of the testis outside the body wall compared to that of the ovary which is shielded from the majority of external radiation could be the reason for the bias in parental origin. If this is the case, then since the commonest effects of radiation damage in mutagenesis experiments are deletions, unless the repair systems of translocations and deletions differ in efficiency, one might also expect deletions to have a biased paternal origin (Thacker, 1990). While there is some evidence that this may not hold for deletions causing DMD (see Chapter 4, section 4.3), the trend of relevant investigations is that point mutations and structural rearrangements are predominantly inherited from the father, and abnormalities of chromosome number from the mother (Chandley, 1991).

8.2 Summary and conclusions

The work described in this thesis concerns aspects of the analysis of DMD translocations, from the subcloning and characterisation of somatic cell hybrids retaining derivative chromosomes to the sequence analysis of an individual translocation. The characterisation of somatic cell hybrids constructed from DMD translocation patients enabled the mapping of the X chromosome breakpoints of seven translocations to be refined. This was performed by hybridisation of cDNA probes to

Southern blots of DNA from the hybrids (Chapter 4). The detailed cytogenetic information on the human complements of these hybrids was used to compile a hybrid panel for the de novo chromosomal localisation of anonymous DNA probes. The mapping of eight DNA fragments using this panel is described in Chapter 3. A large part of the work described in this thesis is the cloning and sequencing of the junction fragments and breakpoint regions from a single translocation (Chapters 5, 6 and 7). A cloning strategy was adopted involving inverse PCR amplification of these DNA fragments. This was found to be fast and reliable (Chapter 6). Sequencing of the cloned fragments revealed the structure of the translocation. Computer aided analysis of DNA sequence surrounding the translocation breakpoints revealed a number of interesting findings whose relevance to the translocation is discussed in Chapter 7.

The identification and study of DMD-associated X;autosome translocations has made a highly significant contribution towards the cloning of the DMD gene and to the consequent progress in understanding of the disease pathogenesis (see Chapter 1). Translocations are a rare and special class of mutation which can cause DMD. Approximately two thirds of males with DMD/BMD have deletions or duplications within the dystrophin locus which are large enough to be detected by Southern analysis using cDNA fragments as probes or by multiplex PCR (for example, den Dunnen et al., 1989; Abbs et al., 1991). The remaining mutations in males are presumed to be either deletions or duplications which are too small to be detected by these methods or point mutations. So far, however, only one point mutation has been described (Bulman et al., 1991). The locations of the majority of deletions and duplications within the DMD locus are well studied and are not random (den Dunnen et al., 1989; Koenig et al., 1989; see Chapter 4). It appears that the distribution of translocation breakpoints differs from these (Chapter 4). The distribution of uncharacterised mutations, such as point mutations, within the DMD locus may be different from either of these. There are two main reasons why the study of the distribution of mutations is important. Firstly, there are important implications to the understanding of the aetiology of the mutations.

Secondly, in clinical practice it is crucial to consider the likely location of the mutation (when it has not been identified) with respect to polymorphic markers. For instance, the presence of a point mutation hot-spot would have important consequences in genetic counselling. The way forward to a full understanding of how mutations affect the DMD locus is to study the distribution and DNA sequence of all classes of mutation. This ought to include the rare examples of X;autosome translocation. The mechanisms responsible for producing X;autosome translocations may be similar or the same as those which produce other chromosomal rearrangements. Therefore investigations into the causes of DMD-associated X;autosome translocations may have implications for the origins of these rearrangements too.

From the study of DMD-associated X;autosome translocations described in this thesis, several conclusions can be made. The distribution of translocation breakpoints within the DMD gene appears to be random and therefore differs from that of deletions and duplications in males which have been identified. The breakpoints of an X;1 translocation which has been cloned and which disrupts the DMD locus are from a two-break translocation, not a three-break rearrangement as originally predicted from cytogenetic characterisation. At least two further breakpoints, independent of those which have been cloned, are therefore required to explain the cytogenetic observations. The cloned translocation is between two expressed loci, those coding for dystrophin and the leukocyte antigen related protein (LAR). Mapping by in situ hybridisation of a DNA probe from adjacent to the chromosome 1 breakpoint confirms the localisation of the chromosome 1 breakpoint at 1p34. This experiment also refines the localisation of the gene for LAR to the same band. The origin of the X;1 translocation is paternal. This brings to 13 the number of X;autosome translocations with a paternal origin. The sequence at the translocation junctions indicates that a deletion of 4 to 7 nucleotides from the X chromosome and a duplication of 2 to 5 nucleotides from chromosome 1 are associated with the translocation. Three nucleotides of junctional homology (GAA) are present at the junctions of the der(X) and der(1) chromosomes. This trinucleotide is

also present twice more close to the X chromosome breakpoint and three times more close to the chromosome 1 breakpoint. Another trinucleotide (TGC) is repeated four times close to the chromosome 1 breakpoint. An Alu sequence is located near to the chromosome 1 breakpoint and a KpnI LINE element close to the X chromosome breakpoint. A stretch of alternating purine/pyrimidine nucleotides is also present close to the chromosome 1 breakpoint.

Due to the small number of studies of the sequence at constitutional translocation breakpoints, it is not yet possible to determine the full significance of the features described around the X;1 translocation breakpoints or whether unknown recombination signals may be present within the sequence. The study described in this thesis will provide a useful reference for similar future investigations, and once these are performed, a clearer understanding of the aetiology of constitutional chromosome rearrangements should emerge.

References

Abbs, S., Yan, S.C., Clarke, S., Mathew, C.G. and Bobrow, M. (1991). A convenient multiplex PCR system for the detection of dystrophin gene deletions: a comparative analysis with cDNA hybridisation shows mistypings by both methods. *J. Med. Genet.* **28**, 304-311.

Acsadi, G., Dickson, G., Love, D.R., Jani, A., Walsh, F.S., Gurusinghe, A., Wolff, J.A. and Davies, K.E. (1991). Human dystrophin expression in mdx mice after intramuscular injection of DNA constructs. *Nature* **352**, 815-818.

Alcalay, M., Zangrilli, D., Pandolfi, P.P., Longo, L., Mencarelli, A., Giacomucci, A., Rocchi, M., Biondi, A., Rambaldi, A., Coco, F.L., Diverio, D., Donti, E., Grignani, F. and Pelicci, P.G. (1991). Translocation breakpoint of acute promyelocytic leukemia lies within the retinoic acid receptor α locus. *Proc. Natl. Acad. Sci.* **88**, 1977-1981.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). A basic local alignment search tool. *J. Mol. Biol.* **214**,

Ariga, A., Carter, P.E. and Davis, A.E. III. (1990). Recombinations between *Alu* repeat sequences that result in partial deletions within the C1 inhibitor gene. *Genomics* **8**, 607-613.

Baer, R. Boehm, T., Yssel, H., Spits, H. and Rabbitts, T.H. (1988). Complex rearrangements within the human J_δ -C_δ /J_α -C_α locus and aberrant recombination between J_α segments. *The EMBO Journal* **7**, 1661-1668.

Bakker, E., Veenema, H., den Dunnen, J.T., van Broeckhoven, C., Grootscholten, P.M., Bonten, E.J., van Ommen, G.J.B. and Pearson, P.L. (1989). Germinal mosaicism increases the recurrence risk for 'new' Duchenne muscular dystrophy mutations. *J. Med. Genet.* **26**, 553-559.

Bawle, E., Tyrkus, M., Lipman, S. and Bozimoski, D. (1984). Aarskog syndrome: full male and female expression associated with an X-autosome translocation. *Am. J. Med. Genet.* **17**,595-602.

Beggs, A.H., Koenig, M., Boyce, F.M. and Kunkel, L.M. (1990). Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. *Hum. Genet.* **86**, 45-48.

Belyavsky, A., Vinogradova, T. and Rajewsky, K. (1989). PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. *Nuc. Acids Res.* **17**, 1919-2932.

Benham, F., Hart, K., Crolla, J., Bobrow, M., Francavilla, M. and Goodfellow, P.N. (1989). A method for generating hybrids containing nonselected fragments of human chromosomes. *Genomics* **4**, 509-517.

Berg, B.O. and Conte, F. (1974). Duchenne muscular dystrophy in a female with a structurally abnormal X chromosome. *Neurology* **24**, 356.

Bernard, L.E., Brooks-Wilson, A.R. and Wood, S. (1991). Isolation of DNA fragments from a human chromosomal subregion by *Alu* PCR differential hybridization. *Genomics* **9**, 241-246.

Bernardi, G. and Bernardi, G. (1985). Codon usage and genome composition. *J. Mol.*

Evol. **22**, 363-365.

Bernardi, G. (1989). The isochore organization of the human genome. *Annu. Rev. Genet.* **23**, 637-661.

Bhat, G.J., Lodes, M.J., Myler, P.J. and Stuart, K.D. (1991). A simple method for cloning blunt ended DNA fragments. *Nuc. Acids Res.* **19**, 398.

Bicknell, D.C., Markie, D., Spurr, N.K. and Bodmer, W.F. (1991). The human chromosome content in human x rodent somatic cell hybrids analyzed by a screening technique using *Alu* PCR. *Genomics* **10,** 186-192.

Blonden, L.A.J., Grootscholten, P.M., Den Dunnen, J.T., Bakker, E., Abbs, S., Bobrow, M., Boehm, C., van Broeckhoven, C., Baumbach, L., Chamberlain, J., Caskey, C.T., Denton, M., Felicetti, L., Galluzi, G., Fischbeck, K.H., Francke, U., Darras, B., Gilgenkrantz, H., Kaplan, J.-C., Herriman, F.H., Junien, C., Boileau, C., Liechti-Gallati, S., Lindlöf, M., Matsumoto, T., Niikawa, N., Müller, C.R., Poncin, J., Malcolm, S., Robertson, E., Romeo, G., Covone, A.E., Scheffer, H., Schröder, E., Schwartz, M., Verellen, C., Walker, A., Worton, R., Gillard, E., and van Ommen, G.J.B. (1991). 242 breakpoints in the 200-kb deletion-prone P20 region of the DMD gene are widely spread. *Genomics* **10,** 631-639.

Bobrow, M. and Cross, J. (1974). Differential staining of human and mouse chromosomes in interspecific cell hybrids. *Nature* **251**, 77-79.

Bodrug, S.E., Ray, P.N., Gonzalez, I.L., Schmickel, R.D., Sylvester, J.E., and Worton, R.G. (1987). Molecular analysis of a constitutional X-autosome translocation in a female with muscular dystrophy. *Science* **237**, 1620-1624.

Bodrug, S.E., Burghes, A.H.M., Ray, P.N., and Worton, R.G. (1989). Mapping of four translocation breakpoints within the Duchenne muscular dystrophy gene. *Genomics* **4**, 101-104.

Bodrug, S.E., Roberson, J.R., Weiss, L., Ray, P.N., Worton, R.G. and Van Dyke, D.L. (1990). Prenatal identification of a girl with a t(X;4)(p21;q35) translocation: molecular characterisation, paternal origin, and association with muscular dystrophy. *J. Med. Genet.* **27**, 426-432.

Bodrug, S.E., Holden, J.J.A., Ray, P.N. and Worton, R.G. (1991). Molecular analysis of X-autosome translocations in females with Duchenne muscular dystrophy. *The EMBO Journal* (in press).

Boehm, T., Baer, R., Lavenir, I., Forster, A., Waters, J.J., Nacheva, E. and Rabbitts, T.H. (1988). The mechanism of chromosomal translocation t(11;14) involving the T-cell receptor Cd locus on human chromosome 14q11 and a transcribed region of chromosome 11p15. *The EMBO Journal* **7**, 385-394.

Boehm, T., Mengle-Gaw, L., Kees, U.R., Spurr, N., Lavenir, I., Forster, A. and Rabbitts, T.H. (1989). Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. *The EMBO Journal* **8**, 2621-2631.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.V., Heynecker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95-113.

Bond, D.J. and Chandley, A.C. (1983). Aneuploidy. *Oxford Monographs on Medical Genetics No.11, Oxford University Press.*

Boyce, F.M., Beggs, A.H., Feener, C. and Kunkel, L.M. (1991). Dystrophin is transcribed in brain from a distant upstream promoter. *Proc. Natl. Acad. Sci.* **88**, 1276-1280.

Boyd, Y. (1986). Characterization and use of somatic cell hybrids with interspecific translocations involving the human X chromosome. *Ann. Hum. Genet.* **51**, 13-26.

Boyd, Y. and Buckle, V.J. (1986). Cytogenetic heterogeneity of translocations associated with Duchenne muscular dystrophy. *Clin. Genet.* **29**, 108-115.

Boyd, Y., Buckle, V.J., Holt, S.M., Munro, E.A., Hunter, D. and Craig, I.W. (1986). Muscular dystrophy in girls with X;autosome translocations. *J. Med. Genet.* **23**, 484-490.

Boyd, Y., Munro, E., Ray, P., Worton, R., Monaco, A., Kunkel, L.M. and Craig, I.W. (1987). Molecular heterogeneity of translocations associated with muscular dystrophy. *Clin. Genet.* **31**, 265-272.

Boyd, Y., Cockburn, D., Holt, S., Munro, E., Van Ommen, G.J., Gillard, B., Affara, N. and Ferguson-Smith, M. (1988). Mapping of 12 translocation breakpoints in the Xp21 region with respect to the locus for Duchenne muscular dystrophy. *Cytogenet. Cell. Genet.* **48**, 28-34.

Boyd, Y. and Fraser, N.J. (1990). Methylation patterns at the hypervariable X-chromosome locus DXS255 (M27β): Correlation with X-inactivation status. *Genomics* **7**, 182-187.

Boyd, Y. (1991). Females with Duchenne muscular dystrophy. *In press.* In *Molecular Genetics and Neurology* (ed. Roses, C.), Royal Society of Medicine.

Britten, R.J., Baron, W.F., Stout, D.B. and Davidson, E.H. (1988). Sources and evolution of human *Alu* repeated sequences. *Proc. Natl. Acad. Sci.* **85**, 4770-4774.

Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W. and Willard, H.F. (1991). Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* **349**, 82-84.

Brown, R.M., Fraser, N.J. and Brown, G.K. (1990). Differential methylation of the hypervariable locus DXS255 on active and inactive X chromosomes correlates with the expression of a human X-linked gene. *Genomics* **7**, 215-221.

Bruns, G.A.P. and Sherman, S.L. (1989). Report of the committee on the genetic constitution of chromosome 1, Human Gene Mapping 10 (1989): Tenth International Workshop on Human Gene Mapping. *Cytogenet. Cell Genet.* **51**, 1-1148.

Buckle, V.J. and Craig, I.W. (1986). In situ hybridisation. In *Human genetic diseases: A practical approach,*

Buckle, V.J., Marsland, A.M., Mosselman, S. and Höppener, J.W.M. (1989). Localization of the human islet amyloid polypeptide (IAPP) gene to 12p11.2-p12.3. *Cytogenet. Cell Genet.* **51**, 972.

Bulfield, G., Siller, W.G., Wight, P.A. and Moore, K.J. (1984). X chromosome-linked muscular dystrophy (*mdx*) in the mouse. *Proc. Natl. Acad. Sci.* **81**, 1189-1192.

Bulman, D.E., Gangopadhyay, S.B., Bebchuck, K.G., Worton, R.G., Ray, P.N. (1991). Point mutation in the human dystrophin gene: Identification through Western Blot analysis. *Genomics* **10**, 457-460.

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

Burmeister, M., Monaco, A., Gillard, E.F., van Ommen, G.-J.B., Affara, N.A., Ferguson-Smith, M.A., Kunkel, L.M. and Lehrach, H. (1988). A 10-megabase physical map of human Xp21, including the Duchenne muscular dystrophy gene. *Genomics* **2**, 189-202.

Burn, J., Povey, S., Boyd, Y. Munro, E.A., West, L., Harper, K. and Thomas, D. (1986). Duchenne muscular dystrophy in one of monozygotic twin girls. *J. Med. Genet.* **23**, 494-500.

Canki, N., Dutrillaux, B. and Tivadar, I. (1979). Dystrophie musculaire de Duchenne chez une petite fille porteuse d'une translocation t(X;3)(p21;q13) de novo. *Annls. Génét.* **22**, 35-39.

Caspersson,T., Farber, S., Foley, G.D., Kudynoski, J., Modest, E.J., Simmonson, E., Wagh, U. and Zech, L. (1968). Chemical differentiation along metaphase chromosomes. *Expl. Cell Res.* **49**, 219-222.

Caspersson, T., Lomakka, G. and Zech, L. (1971). The 24 fluorescence patterns of juman metaphase chromosomes - distinguishing characters and variability. *Hereditas* **67**, 89-102.

Chakravarti, A. (1991). A graphical representation of genetic and physical maps: the Marey map. *Genomics* **11**, 219-222.

Chamberlain, J.S., Gibbs, R.A., Ranier, J.E., Nguyen, P.N. and Caskey, C.T. (1988). Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nuc. Acids Res.* **16**, 11141-11156.

Chandley, A.C. (1991). On the parental origin of de novo mutation in man. *J. Med. Genet.* **28**, 217-223.

Chelly, J., Gilgenkrantz, H, Lambert, M., Hamard, G., Chafey, P., Récan, D., Katz, P., de la Chapelle, A., Koenig, M., Ginjaar, I.B., Fardeau, M., Tomé, F., Kahn, A. and Kaplan, J.-C. (1990). Effect of dystrophin gene deletions on mRNA levels and processing in Duchenne and Becker muscular dystrophies. *Cell* **63**, 1239-1248.

Chérif-Zahar, B., Mattéi, M.G., Le Van Kim, C., Bailly, P., Cartron, J.-P. and Colin, Y. (1991). Localization of the human Rh blood group gene structure to chromosome region 1p34.3-1p36.1 by in situ hybridization. *Hum. Genet.* **86**, 398-400.

Cheung, S.W., Sun, L. and Featherstone, T. (1990). Molecular cytogenetic evidence to characterize breakpoint regions in Robertsonian translocations. *Cytogenet. Cell Genet.* **54**, 97-102.

Choy, W.N., Gopalakrishnan, T.V. and Littlefield, J.W. (1982). Techniques for using HAT selection in somatic cell genetics. In *Techniques in somatic cell genetics* (ed. Shay, J.W.) Plenum Press, New York. 11-21.

Church, G.M. and Gilbert, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci.* **81**,1991-1995.

Clark, J.M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nuc. Acids Res.* **16**, 9677-9686.

Cleary, M.L. (1991). Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* **66**, 619-622.

Cockburn, D.J., Holt, S.M., Roberts, A.N., Cooper, G.J.S., Reid, K.B.M. and Boyd, Y. (1989). Localization of the amylin locus to chromosome 12. *Cytogenet. Cell Genet.* **51**, 977.

Cockburn, D.J., Holt, S.M., Craig, I.W. and Boyd, Y. (1991). Localisation of a DNA fragment isolated from the cloned breakpoints of an X;1 translocation to the expected site - 1p34 - by in situ hybridisation. *Submitted for publication.*

Collins, F.A. and Weissman, S.M. (1984). Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method. *Proc. Natl. Acad. Sci.* **81**, 6812-6816.

Cooper, B.J., Winand, N.J., Stedman, H., Valentine, B.A., Hoffman, E.P., Kunkel, L.M., Scott, M.-0., Fischbeck, K.H., Kornegay, J.N., Avery, R.J., Williams, J.R., Schmickel, R.D. and Sylvester, J.E. (1988). The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. *Nature* **334**, 154-156.

Cooper, G.J.S., Willis, A.C., Clark, A., Turner, R.C., Sim, R.B. and Reid, K.B.M. (1987). Purification and characterization of a peptide from amyloid-rich pancreases of type 2 diabetic patients. *Proc. Natl. Acad. Sci.* **84**, 8628-8632.

Craig, I.W. and McBride, W.O. (1990). Report of the committee on the genetic constitution of chromosome 12. *Cytogenet. Cell Genet.* **55**, 170-176.

Croce, C.M. (1987). Role of chromosome translocations in human neoplasia. *Cell* **49**, 155-156.

Danna, K.J., Sack, G.H. Jr. and Nathans, D. (1973). Studies of Simian virus 40 DNA VII. A cleavage map of the SV 40 genome. *J. Mol. Biol.* **78**, 363.

Davie, A.M. and Emery, A.E.H. (1978). Estimation of proportion of new mutants among cases of Duchenne muscular dystrophy. *J. Med. Genet.* **15**, 339-345.

Davies, K.E., Pearson, P.L., Harper, P.S., Murray, J.M., O'Brien, T., Sarfarazi, M. and Williamson, R. (1983). Linkage analysis of two cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. *Nuc. Acids. Res.* **11**, 2303-2312.

Davies, K. (1991). The essence of inactivity. *Nature* **349**, 15-16.

Deininger, P.L. and Daniels, G.R. (1986). The recent evolution of mammalian repetitive DNA elements. *Trends in Genet.* **2**, 76-80.

den Dunnen, J.T., Bakker, E., Klein Breteler, E.G., Pearson, P.L. and van Ommen, G.J.B. (1987). Direct detection of more than 50% of the Duchenne muscular dystrophy mutations by field inversion gels. *Nature* **329**, 640-642.

den Dunnen, J.T., Grootscholten, P.M., Bakker, E., Blonden, L.A.J., Ginjaar, H.B., Wapenaar, M.C., van Paassen, H.M.B., van Broeckhoven, C., Pearson, P.L. and van Ommen, G.J.B. (1989). Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am. J. Hum. Genet.* **45**, 835-847.

de Thé, H., Chomienne, C., Lanotte, M., Degos, L. and Dejean, A. (1990). The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor α gene to a novel transcribed locus. *Nature* **347**, 558-561.

Disteche, C.M., Adler, D.A., Tedder, T.F. and Saito, H. (1989). Mapping of the genes for LYAMI, a new lymphocyte adhesion molecule, and for LAR, a new receptor-linked protein tyrosine phosphatase, to human chromosome 1, Human Gene Mapping 10 (1989): Tenth International Workshop on Human Gene Mapping. *Cytogenet. Cell Genet.* **51**, 1-1148.

Dodge, G.R., Kovalsky, I., McBride, O.W., Yi, H.F., Chu, M.-L., Saitta, B., Stokes, D.G. and Iozzo, R.V. (1991). Human clathrin heavy chain (CLTC); Partial molecular cloning, expression and mapping of the gene to human chromosome 17q11-qter. *Genomics* **11**, 174--178.

Dracopoli, N.C., O'Connell, P., Elsner, T.I., Lalouel, J.-M., White, R.L., Buetow, K.H., Nishimura, D.Y., Murray, J.C., Helms, C., Mishra, S.K., Donis-Keller, H., Hall, J., Lee, M.K., King, M.-C., Attwood, J., Morton, N.E., Robson, E.B., Mahtani, M., Willard, H.F., Royle, N.J., Patel, I., Jeffreys, A.J., Verga, V., Jenkins, T., Weber, J.L., Mitchell, A.L. and Bale, A.E. (1991). The CEPH Consortium linkage map of human chromosome 1. *Genomics* **9**, 686-700.

Duchenne, G.B.A. (1868). Recherches sur la paralysie musculaire pseudohypertrophique ou paralysie myo-sclérosique. *Arch. Gen. Méd.* **11**, 5-25, 179-209, 305-321, 421-443, 552-588.

Dutrillaux, B. and Viegas-Péquignot, E. (1979). Thymidine synchronisation and chromosome labelling. In *Routine use of chromosome analysis in radiation protection. Proc. IAEA meeting, 1978, IAEA-TECDOC.* **224**, 123.

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

Emanuel, B.S., Zackai, E.H. and Tucker, S.H. (1983). Further evidence for Xp21 location of Duchenne muscular dystrophy (DMD) locus: X;9 translocation in a female with DMD. *J. Med. Genet.* **20**, 461-463.

Emery, A.E.H. (1987). Duchenne muscular dystrophy. *Oxford monographs on medical genetics No. 15, Oxford University Press.*

Engel, A.G. and Banker, B.Q. (1986). Myology: Basic and clinical. *New York: McGraw-Hill.*

Erickson, R.P. (1990). Post-meiotic gene expression. *Trends in Genet.* **6**, 264-269.

Evans, H.J., Buckland, R.A. and Pardue, M.L. (1974). Localization of the genes coding for 18S and 28S ribosomal RNA in the human genome. *Chromosoma* **48**,405-426.

Faed, M.J.W., Lamont, M.A. and Baxby, K. (1982). Cytogenetic and histological studies of testicular biopsies from subfertile men with chromosome anomaly. *J. Med. Genet.* **19**, 49-56.

Fan, Y.-S., Eddy, R.L., Byers, M.G., Haley, L.L., Henry, W.M., Sanke, T., Steiner, D.F., Shows, T.B. and Bell, G.I. (1989). Localization of the human islet amyloid polypeptide gene (IAP) to chromosome 12p12.3. *Cytogenet. Cell Genet.* **51**, 997.

Fanning, T.G. and Singer, M.F. (1987). LINE-1: a mammalian transposable element. *Biochimica et Biophysica Acta* **910**, 203-212.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.

Feener, C.A., Koenig, M. and Kunkel, L.M. (1989). Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy terminus. *Nature* **338**, 509-511.

Feener, C.A., Boyce, F.M. and Kunkel, L.M. (1991). Rapid detection of CA polymorphisms in cloned DNA: application to the 5' region of the dystrophin gene. *Am. J. Hum. Genet.* **48**, 621-627.

Ferguson- Smith, M.A. (1967). Chromosomal satellite association. *The Lancet* **I**, 1156-1157.

Fishel, R.A., Detmer, K. and Rich, A. (1988). Identification of homologous pairing and strand-exchange activity from a human tumor cell line based on Z-DNA affinity chromatography. *Proc. Natl. Acad. Sci.* **85**, 36-40.

Ford, C.E. and Hamerton, J.L. (1956). The chromosomes of man. *Nature* **178**, 1010-1023

Forrest, S.M., Cross, G.S., Speer, A., Gardner-Medwin, D., Burns, J. and Davies, K.E. (1987). Preferential deletion of exons in Duchenne and Becker muscular dystrophies. *Nature* **329**, 638-640.

Forrest, S., Cross, G.S., Flint, T., Speer, A., Robson, K.J.H. and Davies, K. (1988). Further studies of gene deletions that cause Duchenne and Becker muscular dystrophies. *Genomics* **2**, 109-114.

Francke, U., Ochs, H.D., de Martinville, B., Giacalone, J., Lindgren, V., Distèche, C., Pagon, R.A., Hofker, M.H., van Ommen, G.-J.B., Pearson, P.L. and Wedgwood, R.J. (1985). Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa, and Mcleod syndrome. *Am. J. Hum. Genet.* **37**, 250-267.

Fraser, N.J., Boyd, Y. and Craig, I.W. (1989). Isolation and characterisation of a human hypervariable copy number tandem repeat at Xcen-p11.22. *Genomics* **5**, 144-148.

Freund, A.-M., Bichara, M. and Fuchs, R.P.P. (1989). Z-DNA-forming sequences are spontaneous deletion hot spots. *Proc. Natl. Acad. Sci.* **86**, 7465-7469.

Frezal, J. and Schinzel, A., with the collaboration of Baule, M.S. (1990). Report of the committee on clinical disorders and chromosomal deletion syndromes. *Cytogenet. Cell Genet.* **55**, 321-357.

Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1980). Molecular cloning and characterization of the human β-like globin gene cluster. *Cell* **19**, 959-972.

Gardiner, K., Aissani, B. and Bernardi G. (1990). A compositional map of human chromosome 21. *The EMBO Journal* **9**, 1853-1858.

Gerald, P.S. and Brown, J.A. (1974). Report of the committee on the genetic constitution of the X chromosome. *Birth Defects Orig. Ser.* **10**, 29-34.

Giacalone, J. and Francke, U. (1988). Molecular analysis of a constitutional chromosomal translocation: a female DMD patient with a t(X;4) translocation and concomitant deletion of chromosomal material. *Am. J. Hum. Genet.* **43**, suppl. A184.

Gillard, E.F., Chamberlain, J.S., Murphy, E.G., Duff, C.L., Smith, B., Burghes, A.H.M., Thompson, M.W., Sutherland, J., Oss, I., Bodrug, S.E., Klamut, H.J., Ray, P.N. and Worton, R.G. (1989). Molecular and phenotypic analysis of patients with deletions within the deletion-rich region of the Duchenne muscular dystrophy (*DMD*) gene. *Am. J. Hum. Genet.* **45**, 507-520.

Gitschier, J., Wood, W.I., Goralka, T.M., Wion, K.L., Chen, E.Y., Eaton, D.H., Vehar, G.A., Capon, D.J. and Lawn, R.M. (1984). Characterization of the human factor VIII gene. *Nature* **312**, 326-330.

Goodfellow, P.N. (1986). Duchenne muscular dystrophy. Collaboration and progress. *Nature* **322**, 12-13.

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 a region within Xp11.21 between incontinentia pigmenti (IPI) X-chromosomal translocation breakpoints. *Am. J. Hum. Genet.* **48**, 53-64.

Greenstein, R.M., Reardon, R.M. and Chan, T.S. (1977). An X;autosome translocation in a girl with Duchenne muscular dystrophy (DMD): evidence for DMD gene localisation. *Pediatr. Res.* **11**, 475A.

Gyllensten, U.B. and Erlich, H.A. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci.* **85**, 7652-7656.

Haldane, J.B.S. (1935). The rate of spontaneous mutation of a human gene. *J. Genet.* **31**, 317-326.

Haluska, F.G., Tsujimoto, Y. and Croce, C.M. (1987). The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region. *Proc. Natl. Acad. Sci.* **84**, 6835-6839.

Hamerton, J.L., Canning, N., Ray, M. and Smith, S. (1975). A cytogenetic survey of 14,069 newborn infants. *Clin. Genet.* **8**, 223-43.

Han, J., Lu, C.-M., Brown, G.B. and Rado, T.A. (1991). Direct amplification of a single dissected chromosomal segment by polymerase chain reaction: A human brain sodium channel gene is on chromosome 2q22-q23. *Proc. Natl. Acad. Sci.* **88**, 335-339.

Hanahan, D. (1985). Techniques for transformation of *E. coli*. In *DNA cloning, Volume I : A practical approach* (ed. Glover, D.M.), IRL Press Ltd., 109-135.

Harper, P.S. (1988). Practical genetic counselling. Third edition. *Butterworth and Co. (publishers) Ltd.*

Hart, K.A., Hodgson, S., Walker, A., Cole, C.G., Johnson, L., Dubowitz, V. and Bobrow, M. (1987). DNA deletions in mild and severe Becker muscular dystrophy. *Hum. Genet.* **75**, 281-285.

Hassold, T.J., Chen, N., Funkhouser, J., Jooss, T., Manuel, B., Matsuura, J., Matsuyama, A., Wilson, C., Yamane, J.A. and Jacobs, P.A. (1980). A cytogenetic study of 1,000 spontaneous abortions. *Ann. Hum. Genet.* **44**,151-178.

Hedge, P.J., Sellar, G.C., Reid, K.B.M. and Solomon, E. (1987). Assignment of the A chain of C1q (C1QA) to the short arm of chromosome 1. *Cytogenet. Cell Genet.* **46**, 627.

Hemsley, A., Arnheim, N., Toney, M.D., Cortopassi, G. and Galas, D.J. (1989). A simple method for site-directed mutogenesis using the polymerase chain reaction. *Nuc. Acids Res.* **17**, 6545-6551.

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

Hodgson, S.V., Heckmatt, J.Z., Hughes, E., Crolla, J., Dubowitz, V. and Bobrow, M. (1986). A balanced de novo X/autosome translocation in a girl expressing many of the features of Lowe's syndrome. *Am. J. Med. Genet.* **23**, 837-847.

Hoeg, J.M., Segal, P., Gregg, R.E., Chang, Y.S., Lindgren, F.T., Adamson, G.L., Frank, M., Brickman, C. and Brewer, H.B.Jr. (1985). Characterization of plasma lipids and lipoproteins in patients with β-2-glycoprotein I (apolipoprotein H) deficiency. *Atherosclerosis* **55**, 25-34.

Hoffman, E.P., Brown, R.H.Jr. and Kunkel, L.M. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**, 919-928.

Hoffman, E.P., Fischbeck, K.H., Brown, R.H., Johnson, M., Medon, R., Loike, J.D., Harris, J.B., Waterston, R., Brooke, M., Specht, L., Kupsky, W., Chamberlain, J., Caskey, C.T., Shapiro, F. and Kunkel, L.M. (1988). Dystrophin characterization in muscle biopsies from Duchenne and Becker muscular dystrophy patients. *N. Engl. J. Med.* **318**, 1363-1368.

Hoffman, E.P. and Kunkel, L.M. (1989). Dystrophin abnormalities in Duchenne/ Becker muscular dystrophy. *Neuron* **2**, 1019-1029.

Hoffman, E.P. (1990). Dystrophin deficiency: Vignettes of cats and dogs, mice and men. *J. Neur. Sci.* **98** (suppl.), 12.

Hofker, M.H., Skraastad, M.I., Bergen, A.A.B., Wapenaar, M.C., Bakker, E., Millington-Ward, A., van Ommen, G.J.B. and Pearson, P.L. (1986). The X

chromosome shows less genetic variation at restriction sites than the autosomes. *Am. J. Hum. Genet.* **39**, 438-451.

Holden, J.A., Smith, A., MacLeod, P.M., Masotti, R. and Duncan, A.M.V. (1986). Xp21/autosome translocations. Case report and risk for Duchenne muscular dystrophy. *Clin. Genet.* **29**, 516-522.

Holmquist, G.P. and Caston, L.A. (1986). Replication time of interspersed repetitive DNA sequences in hamsters. *Biochimica et Biophysica Acta* **868**, 164-177.

Hu, X., Burghes, A.H.M., Ray, P.N., Thompson, M.W., Murphy, E.G. and Worton, R.G. (1988). Partial gene duplication in Duchenne and Becker muscular dystrophy. *J. Med. Genet.* **25**, 369-376.

Hulten, M., Saadallah, N., Wallace, B.M.N. and Cockburn, D.J. (1985). Meiotic investigations of aneuploidy in the human. *Proceedings of a symposium on Aneuploidy: Etiology and mechanisms, held March 25-29, 1985, in Washington D.C.* (Plenum Press, New York), 75-90.

Human Gene Mapping 10 (1989): Tenth International Workshop on Human Gene Mapping. *Cytogenet. Cell Genet.* **51**, 1-1148.

ISCN (1985). An international system for human cytogenetic nomenclature, (ed. Harnden, D.G. and Klinger, H.P.). Published in collaboration with *Cytogenet. Cell Genet. (Karger, Basel 1985);* also in *Birth Defects: Original Article Series, Vol 21, No1 (March of Dimes Birth Defects Foundation, New York 1985).*

Innis, M.A. and Gelfand, D.H. (1990). Optimization of PCRs. In *PCR Protocols: A Guide to Methods and Applications,* (ed. Innis et al.), Academic Press, Inc., 3-12.

Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (Eds.;1990). PCR protocols: A guide to methods and applications. (*Academic Press, Inc.*).

Jacobs, P.A., Hunt, P.A., Mayer, M. and Bart, R.D. (1981). Duchenne muscular dystrophy (DMD) in a female with an X;autosome translocation: further evidence that the DMD locus is at Xp21. *Am. J. Hum. Genet.* **33**, 513-518.

Jagadeeswaran, P., Tuan, D., Forget, B.G. and Weissman, S.M. (1982). A gene deletion ending at the midpoint of a repetitive DNA sequence in one form of hereditary persistence of fetal haemoglobin. *Nature* **296**, 469-470.

Jauch, A., Daumer, C., Lichter, P., Murken, J., Schroeder-Kurth, T. and Cremer, T. (1990). Chromosomal in situ suppression hybridization of human gonosomes and autosomes and its use in clinical cytogenetics. *Hum. Genet.* **85**, 145-150.

Jeffreys, A.J., Wilson, V., Neumann, R. and Keyte, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nuc. Acids Res.* **16**, 10953-10971.

Kao, F. and Yu, J. (1991). Chromosome microdissection and cloning in human genome and genetic disease analysis. *Proc. Natl. Acad. Sci.* **88**, 1844-1848.

Karlin, S. and Altschul, S.F. (1990). Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci.* **87**, 2264-2268.

Karlovsky, P. (1990). Misuse of PCR. *Trends Biochem. Sci.* **15**, 419.

Kazazian, H.H. and Junien, C. (1987). Report of the committee on the genetic constitution of chromosomes 10,11 and 12. *Cytogenet. Cell Genet.* **46**, 188-212.

Kean, V., Macleod, H.L., Thompson, M.W., Ray, P.N., Verellen-Dumoulin, C. and Worton, R.G. (1986). Paternal inheritance of translocation chromosomes in a t(X;21) patient with X-linked muscular dystrophy. *J. Med. Genet.* **23**, 491-493.

Kelly, T.J. and Smith, H.O. (1970). A restriction enzyme from *Hemophilus influenzae*. II. Base sequence of the recognition site. *J. Mol. Biol.* **51**, 393-409.

Kenwrick, S., Patterson, M., Speer, A., Fischbeck, K. and Davies, K. (1987). Molecular analysis of the Duchenne muscular dystrophy region using pulsed field gel electrophoresis. *Cell* **48**, 351-357.

Keohavong, P. and Thilly, W.G. (1989). Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci.* **86**, 9253-9257.

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**, 509-517.

Koenig, M., Monaco, A.P. and Kunkel, L.M. (1988). The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* **53**, 219-228.

Koenig, M., Beggs, A.H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Müller, C.R., Lindlöf, M., Kaarianinen, H., de la Chapelle, A., Kiuru, A., Savontaus, M.-L., Gilgenkrantz,H., Récan, D., Chelly, J., Kaplan, J.-C., Covone, A.E., Archidiacono, N., Romeo, G., Liechti-Gallati, S., Schneider, V., Braga, S., Moser, H., Darras, B.T., Murphy, P., Francke, U., Chen, D., Morgan, G., Denton, M., Greenberg, C.R., Wrongemann, K., Blonden, L.A.J., van Paassen, H.M.B., van Ommen, G.J.B. and Kunkel, L.M. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am. J. Hum. Genet.* **45**, 498-506.

Krystal, M., D'Eustachio, P., Ruddle, F.H. and Arnheim, N. (1981). Human nucleolus organizers on non-homologous chromosomes can share the same ribosomal gene variants. *Proc. Natl. Acad. Sci.* **78**, 5744-5748.

Kunkel, L.M., Monaco, A.P., Middlesworth, W., Ochs, H.D. and Latt, S.A. (1985). Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. *Proc. Natl. Acad. Sci.* **82**, 4778-4782.

Kunkel, L.M. and co-authors (1986). Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. *Nature* **322**, 73-77.

Laurie, D.A. and Hulten, M.A. (1985). Further studies on chiasma distribution and interference in the human male. *Ann. Hum. Genet.* **49**, 203-214.

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

Ledbetter, S.A., Garcia-Heras, J. and Ledbetter, D.A. (1990). "PCR-karyotype" of human chromosomes in somatic cell hybrids. *Genomics* **8**, 614-622.

Lederberg, S. and Meselson, M. (1964). Degradation of non-replicating bacteriophage DNA in non-accepting cells. *J. Mol. Biol.* **8**, 623-628.

Leighton, B. and Cooper, G. (1988). Pancreatic amylin and calcitonin gene-related peptide cause resistance to insulin in skeletal muscle *in vitro*. *Nature* **335**, 632-635.

Leith, I.R., Hay, R.T. and Russell, W.C. (1988). Detection of Z DNA binding proteins in tissue culture cells. *Nuc. Acids Res.* **16**, 8277-8289.

Lejeune, J., Gautier, M. and Turpin, R. (1959). Etude des chromosomes somatiques de neuf enfants mongoliens. *C.R. hebd. Séanc. Acad. Sci. Paris* **248**, 1721-1722.

Lew, A.M. and Kemp, D.J. (1989). Isolating DNA segments from cloned libraries without screening by affinity selection of PCR products. *Nuc. Acids Res.* **17**, 5859.

Liechti-Gallati, S., Platz, U., Janka, M., Moser, M., Meng, G., Müller, C.R. and Grimm, T. (1990). Predominance of Duchenne muscular dystrophy deletion mutations in oogenesis. *J. Neurol. Sci.* **98** (supplement), 163.

Lindenbaum, R. H., Clarke, G., Patel, C., Moncrieff, M., and Hughes, J. T. (1979). Muscular dystrophy in an X;1 translocation female suggests that Duchenne locus is on X chromosome short arm. *J. Med. Genet.* **16**, 389-392.

Lindlof, M., Kiuru, A., Kääriäinen, H., Kalimo, H., Lang, H., Pihko, H., Rapola, J., Somer, H., Somer, M., Savontaus, M-L. and de la Chapelle, A. (1989). Gene deletions in X-linked muscular dystrophy. *Am. J. Hum. Genet.* **44**, 496-503.

Love, D.R., England, S.B., Speer, A., Marsden, R.F., Bloomfield, J.F., Roche, A.L., Cross, G.S., Mountford, R.C., Smith, T.J., and Davies, K.E. (1991). Sequences of junction fragments in the deletion-prone region of the dystrophin gene. *Genomics* **10**, 57-67.

Lozier, J., Takahashi, N. and Putnam, F.W. (1984). Complete amino acid sequence of human plasma β-2-glycoprotein I. *Proc. Natl. Acad. Sci.* **81**, 3640-3644.

Lynch, J.R., Brown, J.M., Best, S., Jennings, M.W. and Weatherall, D.J. (1991). Characterization of the breakpoint of a 3.5-kb deletion of the β-globin gene. *Genomics* **10**, 509-511.

Malcolm, S., Cowell, J.K. and Young, B.D. (1986). Specialist techniques in research and diagnostic clinical cytogenetics. In *Human cytogenetics: a practical approach.* (ed. Rooney, D.E. and Czepulkowski, B.H.). IRL press Ltd.

Malhotra, S.B., Hart, K.A., Klamut, H.J., Thomas, N.S.T., Bodrug, S.E., Burghes, A.H.M., Bobrow, M., Harper, P.S., Thompson, M.W., Ray, P.N. and Worton, R.G. (1988). Frame-shift deletions in patients with Duchenne and Becker muscular dystrophy. *Science* **242**, 755-759.

Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. (1990). Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nuc. Acids Res.* **19**, 1154.

Mattei, M.G., Mattei, J.F., Vidal, I. and Giraud, F. (1981). Structural anomalies of the X chromosome and inactivation center. *Hum. Genet.* **56**, 401-408.

Max, E.E., Seidman, J.G. and Leder, P. (1979). Sequences of five potential recombination sites encoded close to an immunoglobulin K constant region gene. *Proc. Natl. Acad. Sci.* **76**, 3450-3454.

Meitinger, T., Boyd, Y., Anand, R. and Craig, I.W. (1988). Mapping of Xp21 translocation breakpoints in and around the DMD gene by pulsed field electrophoresis. *Genomics* **3**, 315-322.

Messing, J. (1983). *Methods in enzymology* **101**, 20-78.

Mole, S.E., Iggo, R.D. and Lane, D.P. (1989). Using the polymerase chain reaction to modify expression plasmids for epitope mapping. *Nuc. Acids Res.* **17**, 3319.

Monaco, A.P., Bertelson, C.J., Middlesworth, W., Colletti, C.-A., Aldridge, J., Fischbeck, K.H., Bartlett, R., Pericak-Vance, M.A., Roses, A.D. and Kunkel, L.M. (1985). Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. *Nature* **316**, 842-845.

Monaco, A.P., Neve, R.L., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M. and Kunkel, L.M. (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* **323**, 646-650.

Monaco, A.P., Bertelson, C.J., Colletti-Feener, C. and Kunkel, L.M. (1987). Localization and cloning of Xp21 deletion breakpoints involved in muscular dystrophy. *Hum. Genet.* **75**, 221-227.

Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H. and Kunkel, L.M. (1988). An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* **2**, 90-95.

Moser, H. (1984). Duchenne muscular dystrophy: pathogenic aspects and genetic prevention. *Hum. Genet.* **66**, 17-40.

Mosselman, S., Höppener, J.W.M., Zandberg, J., van Mansfeld, A.D.M., Geurts van kessel, A.H.M., Lips, C.J.M. and Jansz, H.S. (1988). Islet amyloid polypeptide: identification and chromosomal localization of the human gene. *FEBS Lett.* **239**, 227-232.

Mossman, J., Blunt, S., Stephens, R., Jones, E.E. and Pembrey, M. (1983). Hunter's disease in a girl: association with X;5 chromosome translocation disrupting the Hunter gene. *Arch. Dis. Child.* **58**, 911-915.

Mount, S.M. (1982). A catalogue of splice junction sequences. *Nuc. Acids Res.* **10**, 459-472.

Murray, J.M., Davies, K.E., Harper, P.S., Meredith, L., Mueller, C.R. and Williamson, R. (1982). Linkage relationship of a cloned DNA sequence on the short arm of the X chromosome to Duchenne muscular dystrophy. *Nature* **300**, 69-71.

Narazaki, O., Hanai, T., Ueki, Y. and Mitsudome, A. (1985). Duchenne muscular dystrophy in a female with an X;autosome translocation. *Clin. Neurol.* **25**, 432-436.

Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramírez-Solis, R., Webster, T., Ledbetter, D.H. and Caskey, C.T. (1989). *Alu* polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci.* **86**, 6686-6690.

Nevin, N.C., Hughes, A.E., Calwell, M. and Lim, J.H.K. (1986). Duchenne muscular dystrophy in a female with a translocation involving Xp21. *J. Med. Genet.* **23**, 171-173.

Newton, R., Henderson, D.J., Cockburn, D.J., Boyd, Y. and Moore, G.E. (1991). Two RFLPs at the D12S51 locus. *Nuc. Acids Res.* **19**, 194.

Nickoloff, J.A., Chen, E.Y. and Heffron, F. (1986). A 24-base-pair DNA sequence from the *MAT* locus stimulates intergenic recombination in yeast. *Proc. Natl. Acad. Sci.* **83**, 7831-7835.

Nielsen, L.B., Jacobsen, B.B., Nielsen, I.-M. and Tabor A. (1983). X;autosome translocation in a girl with muscular dystrophy. *Clin. Genet.* **23**, 242.

Nielsen, L.B. and Nielsen, I.-M. (1984). Turner's syndrome and Duchenne muscular dystrophy in a girl with an X;autosome translocation. *Ann. Genet.* **27**, 173-177.

Norrander, J., Kempe, T. and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**,101-106.

O'Connell, P., Lathrop, G.M., Nakamura, Y., Leppert, M.L., Ardinger, R.H., Murray, J.L., Lalouel, J.-M. and White, R. (1989). Twenty-eight loci form a continuous linkage map of markers for human chromosome 1. *Genomics* **4**, 12-20.

Ochman, H., Gerber, A.S., and Hartl, D.L. (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621-623.

Ochman, H., Medhora, M.M., Garza, D. and Hartl, D.L. (1990). Amplification of flanking sequences by inverse PCR. *PCR protocols: A guide to methods and applications* (ed. Innis et al.), Academic Press , Inc., 219-227.

Olson, S.B. and Magenis, R.E. (1988). The cytogenetics of mammalian autosomal rearrangements. *Progress and topics in cytogenetics* (ed. Sandberg, A.A.), 583-599.

Orkin, S.H. (1986). Reverse genetics and human disease. *Cell* **47**, 845-850.

Ottolenghi, S. and Giglione, B. (1982). The deletion in a type of α^0–β^0–thalassaemia begins in an inverted Alu1 repeat. *Nature* **300**, 770-771.

Page, D.C., Mosher, R., Simpson, E.M., Fisher, E.M.C., Mardon, G., Pollack, J., McGillivray, B., de la Chapelle, A. and Brown, L.G. (1987). The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* **51**, 1091-1104.

Painter, T.S. (1921). The Y-chromosome in mammals. *Science* **53**, 503-504.

Painter, T.S. (1923). Studies in mammalian spermatogenesis. II. The spermatogenesis of man. *J. Exptl. Zool.* **23**, 291-335.

Patau, K. (1960). The identification of individual chromosomes, especially in man. *Am. J. Hum. Genet.* **12**, 250-276.

Penn, P.A. and Perle, M.A. (1986). Chromosome staining and banding techniques. In *Human cytogenetics: a practical approach* (ed. Rooney, D.E. and Czepulkowski, B.H.), IRL Press Ltd.

Perez-Vidal, M.T., Sarret Grau, E., Prats Vinas, J., Vendrell Bayona, T. and Plaja Rustein (1983). Distrofia muscular tipo Duchenne en una hembra con una translocation equilibradfa X/6. *Rev. Neurol. (Barcelona)* **XI**, 51.

Popescu, N., Landsman, D. and Bustin, M. (1990). Mapping the human gene coding for chromosomal protein HMG-17. *Hum. Genet.* **85**, 376-378.

Powell, J.F., Fodor, F.H., Cockburn, D.J., Monaco, A.P., Craig, I.W. (1991). A dinucleotide repeat polymorphism at the DMD locus. *Nuc. Acids Res.* **19**, 1159.

Quack, B., Speed, R.M., Luciani, J.M., Noel, B., Guichaoua, M. and Chandley, A.C. (1988). Meiotic analysis of two human reciprocal X-autosome translocations. *Cytogenet. Cell Genet.* **48**, 43-47.

Raimondi, S.C., Luthardt, F.W., Summitt, R.L. and Martens, P.R. (1983). High resolution chromosome analysis of phenotypically abnormal patients with apparently balanced structural arrangements. *Hum. Genet.* **63**, 310-314.

Ray, P.N., Belfall, B., Duff, C., Logan, C., Kean, V., Thompson, M.W., Sylvester, J.E., Gorski, J.L., Schimickel, R.D. and Worton, R.G. (1985). Cloning of the breakpoint of an X;21 translocation associated with Duchenne muscular dystrophy. *Nature* **318**, 672-675.

Ribeiro, M.C.M., Melaragno, M.I., Schmidt, B., Brunoni, D., Gabbai, A.A. and Hackel, C. (1986). Duchenne muscular dystrophy in a girl with an (X;15) translocation. *Am. J. Med. Genet.* **25**, 231-236.

Richards, C.S., Watkins, S.C., Hoffman, E.P., Schneider, N.R., Milsark, I.W., Katz, K.S., Cook, J.D., Kunkel, L.M. and Cortada, J.M. (1990). Skewed X inactivation in a female MZ twin results in Duchenne muscular dystrophy. *Am. J. Hum. Genet.* **46**, 672-681.

Roberts, A.N., Leighton, B., Todd, J.A., Cockburn, D., Schofield, P.N., Sutton, R., Holt, S.M., Boyd, Y., Day, A.J., Foot, E.A., Willis, A.C., Reid, K.B.M. and Cooper, G.J.S. (1989b). Molecular and functional characterization of amylin, a peptide associated with type 2 diabetes mellitus. *Proc. Natl. Acad. Sci.* **86**, 9662-9666.

Roberts, R.G., Cole, C.G., Hart, K.A., Bobrow, M. and Bently, D.R. (1989a). Rapid carrier and prenatal diagnosis of Duchenne muscular dystrophy. *Nuc. Acids Res.* **17**, 811.

Roberts, R.G., Bobrow, M. and Bently, D.R. (1990). A Mael II polymorphism near the dystrophin gene promoter by restriction of amplified DNA. *Nuc. Acids Res.* **18**, 1315.

Roberts, R.G., Barby, T.F.M., Manners, E., Bobrow, M. and Bently, D.R. (1991). Direct detection of dystrophin gene rearrangements by analysis of dystrophin mRNA in peripheral blood lymphocytes. *Am. J. Hum. Genet.* **49**, 298-310.

Robinson, D.O., Boyd, Y., Cockburn, D., Collinson, M.N., Craig, I. and Jacobs, P.A. (1990). The parental origin of de novo X-autosome translocations in females with

Duchenne muscular dystrophy revealed by M27b methylation analysis. *Genet. Res.* **56**, 135-140.

Rommens, J.M., Iannuzzi, M.C., Kerem, B.-S., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.-C. and Collins, F.S. (1989). Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* **245**, 1059-1079.

Ropers, H.-H., Zuffardi, O., Bianchi, E. and Tiepolo, L. (1982). Agenesis of corpus callosum, ocular and skeletal anomalies (X-linked dominant Aicardi's syndrome) in a girl with balanced X/3 translocation. *Hum. Genet.* **61**, 364-368.

Saiki, R., Scharf, S.J., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985). Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354.

Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979). Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* **280**, 288-294.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: a laboratory manual, Second Edition. *Cold Spring Habor, New York: Cold Spring Habor Laboratory Press.*

Sanger, F., Niklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**, 5463-5467.

Scharf, S.J., Horn, G.T. and Erlich, H.A. (1986). Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233**, 1076-1078.

Schinzel, A.A. (1983). Catalogue of unbalanced chromosome aberrations in man. *Walter de Gruyter and Co., Berlin.*

Schinzel, A.A. (1991). Uniparental disomy and gene localization. *Am. J. Hum. Genet.* **48**, 424-425.

Sealey, P.G., Whittaker, P.A. and Southern, E.M. (1985). Removal of repeated sequences from hybridization probes. *Nuc. Acids Res.* **13**, 1905-1922.

Sellar, G.C., Blake, D.J. and Reid, K.B.M. (1991). Characterization and organization of the genes encoding the A-, B- and C-chains of human complement subcomponent C1q. *Biochem. J.* **274**, 481-490.

Silver, J. and Keerikatte, V. (1989). "Inside-out" PCR to amplify cellular DNA flanking proviruses. *J. Cell Biochem.* (suppl.) **13E**, 306.

Singer, M.F. (1982). SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* **28**, 433-434.

Singer, M. and Berg, P. (1991). Genes and genomes, a changing perspective. *University Science Books.*

Smith, G.R., Kunes, S.M., Schultz, D.W., Taylor, A. and Triman, K.L. (1981). Structure of Chi hotspots of generalized recombination. *Cell* **24**, 429-436.

Smith, H.O. and Wilcox, K.W. (1970). A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J. Mol. Biol.* **51**, 379-391.

Solomon, E., Skok, J., Griffin, J. and Reid, K.B.M. (1985). Human C1qB chain (C1QB) is on chromosome 1p. *Cytogenet. Cell Genet.* **40**, 749.

Soriano, P., Meunier-Rotival, M. and Bernardi, G. (1983). The distribution of interspersed repeats is non-uniform and conserved in the mouse and human genomes. *Proc. Natl. Acad. Sci.* **80**, 1816-1820.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.

Sparkes, R.S., Zollman, S., Klisak, I., Kirchgessner, T.G., Komaromy, M.C., Mohandas, T., Schotz, M.C. and Lusis, A.J. (1987). Human genes involved in lipolysis of plasma lipoproteins: mapping of loci for lipoprotein lipase to 8p22 and hepatic lipase to 15q21. *Genomics* **1**, 138-144.

Spence, J.E., Perciaccante, R.G., Greig, G.M., Willard, H.F., Ledbetter, D.H., Hejtmancik, J.F., Pollack, M.S. et al (1988). Uniparental disomy as a mechanism for human genetic disease. *Am. J. Hum. Genet.* **42**, 217-226.

Spitzner, J.R. and Muller, M.T. (1988). A consensus sequence for cleavage by vertebrate DNA topoisomerase II. *Nuc. Acids Res.* **16**, 5533-5556.

Streuli, M., Kreuger, N.X., Hall, L.R., Schlossman, S.F., Saito, H. (1988). A new member of the immunoglobulin superfamily that has a cytoplasmic region homologous to the leukocyte common antigen. *J. Exp. Med.* **168**, 1553-1562.

Staden, R. (1984). Graphic methods to determine the function of nucleic acid sequences. *Nuc. Acids Res.* **12**, 521-538.

Steinkasserer, A., Estaller, C., Weiss, E.H., Sim, R.B. and Day, A.J. (1991a). Complete nucleotide and deduced amino acid sequence of human beta 2-glycoprotein I. *Biochem. J.* **277**, 387-391.

Steinkasserer, A., Cockburn, D.J., Black, D.M., Boyd, Y., Solomon, E. and Sim, R.B. (1991b). Assignment of apolipoprotein H (APOH:beta 2-glycoprotein I) to human chromosome 17 band q23-25; determination of the major expression site. *Cytogenet. Cell Genet.* in press.

Steinmetz, M., Uematsu, Y. and Lindahl, K.F. (1987). Hotspots of homologous recombination in mammalian genomes. *Trends in Genet.* **3**, 7-10.

Tait, J.F., Frankenburg, D.A., Miao, C.H., Killary, A.M., Adler, D.A. and Disteche, C.M. (1991). Chromosomal localization of the human annexin III (ANX3) gene. *Genomics* **10**, 441-448.

Thacker, J. (1990). Molecular nature of ionising radiation-induced mutations of native and introduced genes in mammalian cells. In *Ionizing radiation damage to DNA: Molecular aspects* . Wiley-Liss, Inc., 221-229.

Theune, S., Fung, J., Todd, S., Sakaguchi, A.Y. and Naylor, S.L. (1991). PCR primers for human chromosomes: Reagents for the rapid analysis of somatic cell hybrids. *Genomics* **9**, 511-516.

Tjio, J.H. and Levan, A. (1956). The chromosome number of man. *Hereditas* **42**, 1-6.

Trent, J.M., Kaneko, Y. and Mitelman, F. (1989). Report of the committee on structural chromosome changes in neoplasia, Human Gene Mapping 10 (1989): Tenth International Workshop on Human Gene Mapping. *Cytogenet. Cell Genet.* **51**, 533-562.

Triglia, T., Peterson, M.G. and Kemp, D.J. (1988). A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nuc. Acids Res.* **16**, 8186.

Tucker, J.D., Christensen, M.L. and Carrano, A.V. (1988). Simultaneous identification and banding of human chromosome material in somatic cell hybrids. *Cytogenet. Cell Genet.* **48**, 103-106.

van Ommen, G.J.B., Bertelson, C., Ginjaar, H.B., den Dunnen, J.T., Bakker, E., Chelly, J., Matton, M., Van Essen, A.J., Bartley, J., Kunkel, L.M. and Pearson P.L. (1987). Long range genomic map of the Duchenne muscular dystrophy (DMD) gene: isolation and use of J66 (DXS268), a distal intragenic marker. *Genomics* **1**, 329-536.

Verellen, C., Meyer, R.de, Freund, M., Laterre, C. and Scholberg, B. (1977). Progressive muscular dystrophy of the Duchenne type in a young girl associated with an aberration of chromosome X, in Littlefield, S.W. (ed.) (1977). *Fifth International Congress of Birth Defects* (Excerpta Medica, Amsterdam), 42.

Verellen-Dumoulin, C., Freund, M., Meyer, R.de, Laterre, C., Frederic, J., Thompson, M.W., Markovic, V.D. and Worton, R.G. (1984). Expression of X-linked muscular dystrophy in a female due to translocation involving Xp21 and non-random inactivation of the normal X chromosome. *Hum. Genet.* **67**, 115-119.

Verga, V., Hall, B.K., Wang, S., Johnson, S., Higgins, J.V. and Glover, T.W. (1991). Localization of the translocation breakpoint in a female with Menkes syndrome to Xq13.2-q13.3 proximal to PGK1. *Am. J. Hum. Genet.* **48**, 1133-1138.

Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y.-H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., van Ommen, G.-J.B., Blonden, L.A.J., Riggins, G.J., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A. and Warren, S.T. (1991). Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905-914.

Vidaud, D., Vidaud, M., Plassa, F., Gazengel, C., Noel, B. and Goossens, M. (1989). Father-to-son transmission of hemophilia A due to uniparental disomy. *Am. J. Hum. Genet.* **45** (Suppl): A226.

Viegas-Péquignot, E., Derbin, C., Malfoy, B., Taillandier, E., Leng, M. and Dutrillaux, B. (1983). Z-DNA immunoreactivity in fixed metaphase chromosomes of primates. *Proc. Natl. Acad. Sci.* **80**, 5890-5894.

Vnencak-Jones, C. and Phillips J.A. III (1990). Hot spots for growth hormone gene deletions in homologous regions outside of Alu repeats. *Science* **250**, 1745-1748.

Voss, R., Ben-Simon, E., Avital, A., Godfrey, S., Zlotogora, J., Dagan, J., Tikochinski, Y. et al. (1988). Isodisomy of chromosome 7 in a patient with cystic fibrosis: could uniparental disomy be common in humans? *Am. J. Hum. Genet.* **45**, 373-380.

Walton, J.N. (1957). The inheritance of muscular dystrophy. *Acta Genet. Stat. Med. (Basel)* **7**, 318-320.

White, T.J., Arnheim, N. and Erlich, H.A. (1989). The polymerase chain reaction. *Trends in Genet.* **5**, 185-189.

Whitehouse, H.L.K. (1973). The chiasmatype theory. Chapter 8 in *Towards an understanding of the mechanism of heredity* (3rd edition). Edward Arnold (publishers) Ltd.

Wilcox, D.E., Cooke, A., Colgan, J., Boyd, E., Aitken, D.A., Sinclair, L., Glasgow, L., Stephenson, J.B.P. and Ferguson-Smith, M.A. (1986). Duchenne muscular dystrophy due to a familial Xp21 deletion detectable by DNA analysis and flow cytometry. *Hum. Genet.* **73**, 175-180.

Wong, C., Kazazian, H.H. Jr., Stetten, G., Earnshaw, W.C., Van Keuren, M.L. and Antonarakis, S.E. (1989). Molecular mechanism in the formation of a human ring chromosome 21. *Proc. Natl. Acad. Sci.* **86**, 1914-1918.

Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. (1987). *Nuc. Acids Res.* **15**, 529-542.

Yen, P.H., Li, X.-M., Tsai, S.-P., Johnson, C., Mohandas, T. and Shapiro, L.J. (1990). Frequent deletions of the human X chromosome distal short arm result from recombination between low copy repetitive elements. *Cell* **61**, 603-610.

Yoshitake, S., Schach, B.G., Foster, D.C., Davie, E.W. and Kurachi, K. (1985). Nucleotide sequence of the gene for human factor IX (antihemophilic factor B). *Biochem.* **24**, 3736-3750.

Zatz, M., Vianna-Morgante, A.M., Campos, P. and Diamant, A.J. (1981). Translocation (X;6) in a female with Duchenne muscular dystrophy: implications for the localisation of the DMD locus. *J. Med. Genet.* **18**, 442-447.

Zubrzycka-Gaarn, E.E., Bulman, D.E., Karpati, G., Burghes, A.H.M., Belfall, B., Klamut, H.J., Talbot, J., Hodges, R.S., Ray, P.N. and Worton, R.G. (1988). The Duchenne muscular dystrophy gene product is localized in sarcolemma of human skeletal muscle. *Nature* **333**, 466-469.

Appendix Full sequence of DNA around the X;1 translocation

The appendix contains the full basic sequence from all the DNA fragments cloned from the X;1 translocation patient and of the corresponding DNA fragments from normal chromosomes. Restriction sites are indicated which allow comparison of these sequences with the consensus X chromosome and consensus chromosome 1 sequences (Chapter 7, figures 7.2 and 7.7). The sequence across the junctions of the der(X) and der(1) chromosomes is also illustrated in Chapter 7, figure 7.9.

								C	9i(N1)
TCT	AGAGC	CAGAA	ATACC	ATTTG	ACCAA	GCAAT	CCCAT	CACTG	12ii(N2)
CCATA	TATGG	GTTTC	CTAAT	ATTTA	ATAAG	ACACG	AATAA	AGTCG	9i(N1)
GGTAT	ATACC	CAAAG	GATTA	TAAAT	TATTC	TGTGC	TTATT	TCAGC	12ii(N2)
TGATA	AGTGT	TATCG	TTTCT	GAACT	TTGGT	TGGGT	TTACG	GGTAG	9i(N1)
ACTAT	TCACA	ATAGC	AAAGA	CTTGA	AACCA	ACCCA	AATGC	CCATC	12ii(N2)
TTACT	ACCTG	ACCTA	TTTCT	TTTAA	ACCGT	GTATA	TGTGG	TACCT	9i(N1)
AATGA	TAGAC	TGGAT	AAAGA	AAATT	TGGCA	CATAT	ACACC	ATGGA	12ii(N2)
TATGA	TACGT	CGTTG	TTTTT	TCCTA	CTCAA	AGTAC	AGGAA	ACGTC	9i(N1)
ATACT	ATGCA	GCAAC	AAAAA	AGGAT	GAGTT	TCATG	TCCTT	TGCAG	12ii(N2)
CCTAC	ACCTA	CTTCG	ACCTT	TTGTA	GTAAG	AGTCG	TTTGA	TTGTA	9i(N1)
GGATG	TGGAT	GAAGC	TGGAA	AACAT	CATTC	TCAGC	AAACT	AACAT	12ii(N2)
TCCTT	GTCTT	TTAGT	TTGAG	CCGTA	CAAGA	GTGAG	TATTT	AGCCT	9i(N1)
AGGAA	CAGAA	AATCA	AACTC	GGCAT	GTTCT	CACTC	ATAAA		12ii(N2)
CAACT	TGTTG	GTCTT	CTGTA	CCTAC	GTCAC	CTCGT	AGTGT	GTGAC	9i(N1)
CCC									9i(N1)
CT	TTTGG	GGTGT	GTGGG	GCTAG	GGGAG	TGATA	GCATT	AGGGG	6i(N1)
CC	TTTTG	GGTGT	GTGGG	GCTAG	GGGAG	TGATA	GCATT	AGGGG	8i(N2)
AAATA	CCTAA	CATAG	ATGAC	GGATT	GGTAC	AGCAA	ACCAC	CATGG	6i(N1)
AAATA	CCTAA	CATAG	ATGAC	GGATT	GGTAC	AGCAA	ACCAC	CATGG	8i(N2)
CATGT	GTATA	CCTAT	GTAAC	AAACC	TGCAC	GTTCT	GCACA	TGTAT	6i(N1)
CATGT	GTATA	CCTAT	GTAAC	AAACC	TGCAC	GTTCT	GCACA	TGTAT	8i(N2)
CCCAA	AACTT	AAAGT	ATAAT	AATAA	AAAGT	TTCTT	AAAGT	TTACT	6i(N1)
CCCAA	AACTT	AAAGT	ATAAT	AATAA	AAAGT	TTCTT	AAAGT	TTACT	8i(N2)
TATGG	CTGCA	AATAA	GATGA	GGATT	C				6i(N1)
TATGG	CTGCA	AATAA	GATGA	GGATT	CGGAT	TCAGC	ATAAG	TCATA	8i(N2)
TATTT	TTAAA	ATATG	TCAGT	TCTTG	CCATT	GTGTT	AGGGA	AACCA	8i(N2)
CCAAG	TACAC	ATTTA	TACAT	TTTTG	CTCAG	CATTT	TTCAG	GAGAT	8i(N2)
						GTAAA	AAGTC	CTCTA	13i(N1)
AATCT	CAATG	ATGAT	TCAAA	GAAAA	AAATA	ATAAC	TCATA	AGAAT	8i(N2)
TTAGA	GTTAC	TACTA	AGTTT	CTTTT	TTTAT	TATTG	AGTAT	TCTTA	13i(N1)
TTTTA	TTTAC	TGACT	GCTTA	GTCAC	AAAAG	ACAGA	ACACC	ATCTC	8i(N2)
AAAAT	AAATG	ACTGA	CGAAT	CAGTG	TTTTC	TGTCT	TGTGG	TAGAG	13i(N1)
TTGAA	GTCTC	AGACT	TAGAA	CACTC	TCTTC	GATAG	ATTTA	CCAGC	8i(N2)
AACTT	CAGAG	TCTGA	ATCTT	GTGAG	AGAAG	CTATC	TAAAT	GGTCG	13i(N1)
AAAAT	TCGTG	ATTCA	TGCCT	GTATA	GTAAT	TGCTT	TTGTG	ATGTT	8i(N2)
TTTTA	AGCAC	TAAGT	ACGGA	CATAT	CATTA	ACGAA	AACAC	TACAA	13i(N1)

XbaI

HaeIII

CATAT TTT 8i(N2)
GTATA AAAAT AAATT GGTCT ATAAC CAAGA CATCG AATAC GTACT 13i(N1)

ACAAC CGATG AGAGG TTTCT TTAAA GTAAA ATTTA ATCAA TACGG 13i(N1)

GGAAT CCTCT TATTA CTTTG TAAAC CCCTC TTAAA ATAAT AACCC 13i(N1)

CCTAA CTTCG ATATA AACCA AATTC TTTAT GGACG GTTTA TTACG 13i(N1)

AAACT GTTAA AATTT GTATT CGAA 13i(N1)

HindIII

TTGTAGGATT TCTATATAGA ATTTGATATG CTTTGGCCTT AACTGGTGTT TTGAGGCCTG Alf1

AGGTTTAGTT TTAAGTGGAA TTGCCTTTTA AAATGAGTAG CAGTAGGCAT TAGCATAAGC Alf1

GATTTCAGGA GCTGTTTCTA GTGTGTTTAG ATGCTAACAT ATGCAAAAAT AAGTACATCT Alf1

AGGGTAAACT TAATAAATTA TGTATGTTTA TTGACTCATA GATAATAGGT ATATCTCTTT Alf1

AAAATATCCT TTGTGAATCA GACAATTCTC TTCCCGTTAT GATTTCTCTG TAAGACAGTA Alf1

CATTCTACAT GCGTTGGTCG TGGGTTTCAT ATACATTGTT TCTCTGAGTT TCAAACAGAT Alf1

TGGTTTGTTT GCCAGTCCAA AAAAAAAAAA AAAATCCTCC CTATATTAAG AGAGTGCTTT Alf1

GTCAGCTTTC CCCTGTGAAT AACAAAGTGA GGATGTTCTC TTTAAAGGAA TAAAACAAAA Alf1

AATCAGGGAT AAAAGATTGC GCCTGTTCTA GA Alf1

XbaI

Sequence of normal X chromosome clones.

				CAG	GAACT	TGCCC	AGTGG	GAGCT	GTTGC 1CMA+1CMB
TGCTG	CTGCA	GAAAG	AAGCA	GGAAG	GAGAA	GAGAC	TGGGC	CCATG	1CMA+1CMB
	PstI								
GCAGT	CCCTT	GCGTG	TCTGC	ACCAA	GGGCT	CTGGG	CAGAG	AGCCC	1CMA+1CMB
GGCAG	GTGCA	GCACA	GCTCT	GTGGT	GACAG	GGGCC	AGGCA	CTGCC	1CMA+1CMB
TGAGT	CCTTA	CCTCG	CACAT	ACAGG	TTCGC	AGGGG	CTGAG	TAACG	1CMA+1CMB
TGTGC	CTGCC	GAGTT	GGT****	ACACA	CTCGT	ACTT			1CMA+1CMB
	GG	CTCAA	CCAG**	GTGTGT	GAGCA	TGAAC	GGAAC	CAGCC	1CMA+1CMB
TAAGG	AGTGA	CGAGA	GATAG	ACGTT	CCGTG	GACAC	CCATG	TGCTT	1CMA+1CMB
CGCTC	TCCAG	TC*****	TGCCG	GGTAC	AGCTC	ACACC	GGACC	TTGTG	1CMA+1CMB
CCACT	CTCCA	GAGCT	GTCGT	AGACA	CCGAC	TACAA	CTGGT	CTACA	1CMA+1CMB
GCTTC	GAGTA	GGTCC	GGTGT	GGGAC	CGTGG	GGACG	ACGAA	GGTCC	1CMA+1CMB
TTCGA	GAGGG	ACTTC	TCGGG	TCACC	GGAGG	ACAGA	TCT		1CMA+1CMB
						AGA	TCTTG	GTTGC	F3A(2ai)
						XbaI			
CAACA	CGAAA	CCTCG	TGTTA	GAGTA	CGAGG	GGTGT	CAGGT	CGTTC	F3A(2ai)
TGGGT	GGAAA	CGAGA	CGGTC	CCTGA	CATGG	GGTTG	GGGGT	CGAGG	F3A(2ai)
TAAGT	CAAAG	ACGGT	GTGAG	AACGG	TCATG	ACCTC	GGATA	AAGGG	F3A(2ai)
TTAGT	TGTAG	GGACA	ATCAC	AAGGG	CCGGG	GCGTC	GGGGT	GATTC	F3A(2ai)
GGTCA	GTCTC	GGTCC	ACACC	CTCGT	CGGAC	GTC			F3A(2ai)
					CTG	CAGAC	TGGCG	GGCCT	IIIEa(PPiii)
					PstI				
GGGAT	TCACA	AGCAC	GCCAT	GCCCT	CATCA	GGCTG	GAGCT	TGAGA	IIIEa(PPiii)
ATTTG	CCGCC	CACCA	GGAAG	GAATG	GGTTT	GCCCT	GGGGG	CCTGG	IIIEa(PPiii)
		TGGT	CCTTC	CTTAC	CCAAA	CGGGA	CCCCC	GGACC	F5A(PPiii)
TGCCC	AGCAG	CCTCT	GCTCC	CAGCA	CCATC	TTCCC	CTCCT	ACTCC	IIIEa(PPiii)
ACGGG	TCGTC	GGAGA	CGAGG	GTCGT	GGTAG	AAGGG	GAGGA	TGAGG	F5A(PPiii)
ACCCC	CTCCC	CGCCA	GCCCC	ACCCA	TTCCA	GCTCC	ACTCC	TCAG	IIIEa(PPiii)
TGGGG	GAGGx	xxxGT	CGGGG	TGGGT	AAGGT	CGAGG	TGAGG	AGTCG	F5A(PPiii)
							AGG	AGTCG	IIIEa(PPiii)
GGCCA	GGACG	ATTTC	ACGAA	GCTGG	GTGGT	GCCGA	GGACG	TATTA	F5A(PPiii)
						GCCGA	GGACG	TATTA	IIEa(PPii)
GGCCA	GGACG	ATTTC	ACGAA	GCTGG	GTGGT	GCCGA	GGACG	TATTA	IIIEa(PPiii)
AGTTC	TCCGA	TTACC	TCCAA	CCTCG	AG				F5A(PPiii)
AGTTC	TCCGA	TTACC	TCCAA	CCTCG	AGAGT	GAGCC	GGTCA	CGGAC	IIEa(PPii)
AGTTC	TCCGA	TTACC	TCCAA	CCTCG	AGAGT	GAGCC	GGTCT	CGGAC	IIIEa(PPiii)
				SstI					
TCCCC	GACGA	GACCG	GCGTC	TGCCT	GTCTG	TCCTC	GGTCC	TAGCC	IIEa(PPii)
TCCCC	GACGA	GACCG	GCGTC	TGCCT	GTCTG	TCCTC	GGTCC	TAGCC	IIIEa(PPiii)

CCACC	CATCG	TTTTC	CTGGG	AGTCT	TTCCC	ACCTC	CTTAG	AACCC	IIEa(PPii)
CCACC	CATCG	TTTTC	CTGGG	AGTCT	TTCCC	ACCTC	CTTAG	AACCC	IIIEa(PPiii)
GAACT	CCGGT	CAAGA	CTGTC	CAGGA	CCGCC	CCCTC	CGGTT	CCGTC	IIEa(PPii)
GAACT	CCGGT	CAAGA	CTGTC	CAGGA	CCGCC	CCCTC	CGGTT	CCGTC	IIIEa(PPiii)
TCGCC	GTTCC	ACCCC	CACCT	AGG					IIEa(PPii)
TCGCC	GTTCC	A	BamHI						IIIEa(PPiii)
			GGA	TCCAC	TGGAC	TTGGC	CG		IIIPa(PPiii)
TCCCC	TGGCC	CCACC	AATAA	ACACT	TGAAA	AAACT	GCTTG	TGAAA	IIIPa(PPiii)
GCACA	CTGGG	CAAGC	TGTCC	GTGAC	AGACA	CACCC	CCGGG	CACAG	IIIPa(PPiii)
GCAGC	ACACC	CCTCC	CTGGA	ACAAT	CCAAC	AGACA	ATTAC	ACCCA	IIIPa(PPiii)
CACAG	GACGG	TCCAA	CAGTT	AAGAA	CCCCC	ACAGG	ACACA	GTCGA	IIIPa(PPiii)
CAAAA	CCACC	TCCTG	CACAG	AACAG	CCAGC	CTAAC	CTGAC	AGCCT	IIIPa(PPiii)
TT	GGTGG	AGGAC	GTGTC	TTGTC	GGTCG	GATTG	GACTG	TCGGA	IIIPa(PPiii)
CCACC	TTCCC	AGGG							
GGTGG	AAGGG	TCCCG	GGTCC	GACTA	CCTGT	TAGTG	TGGGT	GTGTC	IIIPa(PPiii)
CTGTA	CCTGT	CTTCG	ACGGC	CCGGG	GGGGA	CGGAC	CTGAC	CTGTC	IIIPa(PPiii)
CTCGG	TCTCC	TTCTC	AAGGT	TCAGT	ACGAG	AGTGC	CAACC	CCGGT	IIIPa(PPiii)
TTCCC	CATGG	TGGGA	CCCAG	ACTGG	ATATC	GGTGG	AGTGA	TGGAC	IIIPa(PPiii)
TACAG	GGAGA	AGGTG	GTAGA	TTGGG	TCCCG	TCGTC	CTCCC	TGACC	IIIPa(PPiii)
CTTGG	GTCCC	CCGAC	CGTCG	AC					IIIPa(PPiii)
			PvuII						

Sequence of cloned normal chromosome 1 DNA segments.

	CAGCT GCTCC	TCTGT	TGGAG	TGGAC	TCGCC	AGCCC	CAATG	16i(A1)
	PvuII			CTG	AGCGG	TCGGG	GTTAC	15ii(A1)
CCTGC	CAGAC	ACTCT	GCCTG	GGGCA	GGGTG	GGGCC	AAGGC	TGTGC	16i(A1)
GGACG	GTCTG	TGAGA	CGGAC	AAAGT	CCCAC	CCCGG	TTCCG	ACACG	15ii(A1)
CGCTT	ACAAC	TCCAC	TCCTC	TAGGC	CATTT	GGGAA	GTTTT	CAGGC	16i(A1)
GCGAA	TGTTG	AGGTG	AGGAG	ATCCG	GTAAA	CCCTT	CAAAA	GTCCG	15ii(A1)
TCTTG	TTCCT	GAAGA	CAGAC	AAGCT	GACCT	GAGGC	CTGAG	CCTTG	16i(A1)
AGAAC	AAGGA	CTTCT	GTCTG	TTCGA	CTGGA	CTCCG	GACTC	GGAAC	15ii(A1)
AGGAG	TAAGA	CACCT	GACAG	TAAGC	TAGGG	GACCT	CCCGT	CCCAT	16i(A1)
TCCTC	ATTCT	GTGGA	CTGTC	ATTCG	ATCCC	CTGGA	GGGCA	GGGTA	15ii(A1)
GCCCC	CAACT	CTCAT	GAAGC	ATCTT	TAATA	GTCTC	TGAAG	CACAA	16i(A1)
CGGGG	GTTGA	GAGTA	CTTCG	TAGAA	ATTAT	CAGAG	ACTTC	GTGTT	15ii(A1)
TTTGC	CCACC	ACAAA	TAGTT	CCTCT GCAG					16i(A1)
AAACG	GGTGG	TGTTT		PstI					15ii(A1)
				CT GCAGC	TCAAC	TGGTC	ACCGA		15iii(A1)
ATGGG	CTGGC	AGCTG	TGCCT	CTCCA	GAACC	ACCCT	CTCCC	ATGAC	15iii(A1)
AGACA	GCGAC	TCCTT	GATAA	GTGTA	TTTTG	GTTGA	CAGCC	TTGAC	15iii(A1)
AGCCT	AATTC	TCTCT	GCTCT	GAGTA	AGCAA	GTCTG	ACCAC	ATGGC	15iii(A1)
CCCAA	AGAGT	CGGAC	AGGCC	ACCGC	GATCA	CCCAC	CACGT	GCACT	15iii(A1)
GGGTG	GAGAG	TCAGA	CGACG	CCCCT	CAGGA	AGCCG	CGCTA	AGCAC	15iii(A1)
TTGGC	GTGAT	GGAAT	CATCA	CTGCA	AGGAG	CCATG	TCCGC	TTTCC	15iii(A1)
					CCTC	GGTAC	AGGCG	AAAGG	18vi(A1)
TTTAC	AATCC	CCTGG	GTGGA	GACAG	TTGTA	CTTTT	TTTTT	TT	15iii(A1)
AAATG	TTAGG	GGACC	CACCT	CTGTC	AACAT	GAAAA	AAAAA	AACTC	18vi(A1)
TACCT	CAGAA	CGAGA	AAGTG	GTTCG	ACCTC	ACGTC	ACCGC	GCTGG	18vi(A1)
AACCG	AGTGA	CGTTG	GAGAC	CAAGT	TCGCT	AAAAG	GATGG	AGTCG	18vi(A1)
GAGGA	CTCAT	CGACC	CTAAT	GTCCG	TGCGC	GGTGG	TGCGG	GTCGA	18vi(A1)
TTAAA	AACAT	CAAAA	TCATC	TCTGC	CCCAA	AGTGG	TACAA	CCGGT	18vi(A1)
CCTAC	CAGAG	CTAGA	GGACT	GGAGC	ACTAG	ACGGA	CGGAG	TCGGA	18vi(A1)
GGGTT	TCACG	ACCCT	AATGT	CCGCA	CCGGT	GGTGT	GGGCC	AGTTG	18vi(A1)
TCAAA	TTGAA	AATTA	CTGTA	GGTAG	ACATG	TTCTT	CCGAT	CCACT	18vi(A1)
CCCCT	GTCAT	ACTCC	CATCT	TCCCG	AGAGT	CCCGG	GGAAG	ACCCG	18vi(A1)
TGTAC	ATATG	GTCTG	GGTCA	CACAA	GAGTC	CTTTC	CCGTT	CCACC	18vi(A1)

GGTCG AAATC AGTGG GGATC CGACA CTCGT ATCGT CTCTA ACCTT 18vi(A1)

TGGGG TTTTG TCAGG GGTCC AAGGA GAGAC GACCC TTGAA CCTGA 18vi(A1)

AGGTC TTTCG TGTGA CAGGT GAGCG AGGGT CGTTT CTTGT AGACC 18vi(A1)

TTGTG TGGTC CACGA GTGGG CTCCG TTGTA TTGGT GTCCC CGAGT 18vi(A1)

GTCGA C 18vi(A1)

CAGCT GAACA CAAGG GTTCC TCCAG GAAGG CTCCA GGAAG ACCTG 19iii(A1)

PvuII

GCCAG AGGAT AGGCT CACCT GGCCT GGTCC GTTCC TGAGG CCTTA 19iii(A1)

CTACC CCAAA GCCCA GAGCC ATGGT GGTTC TGGAG CCCAC CATGC 19iii(A1)

TGCGG CAGGC AGGAG GATGT ACTGA AGGAA AGCCA GGCCG GGTAT 19iii(A1)
CCATA 15i(A1)

GCATG CATTT ACACA CAGAT AAGTG TGCCC GTGTG TGCAC ACACA 19iii(A1)
CGTAC GTAAA TGTGT GTCTA TTCAC ACGGG CACAC ACGTG TGTGT 15i(A1)

CATGC ACATA CACCA AGAAG TCGGA TTGGA TGGCT CCAAA TTCCC 19iii(A1)
GTACG TGTAT GTGGT TCTTC AGCCT AACCT ACCGA GGTTT AAGGG 15i(A1)

AT 19iii(A1)
TACCC GAGGC CCGGG TACGC CGTTC TGGAT GTTTT CTATG TTTAC 15i(A1)

GAGGG TTCTC ACCGA CGAAG GGGTA CCCGG ACGTT CCTTC CTTCC 15i(A1)

GCTTC ATCAC GCCTG CGGTC CTTGA ACGGG TCACC CTCGA CAACG 15i(A1)

ACGAC GACGT C 15i(A1)

PstI

CTGC AGAAA GAAGC AGGAA CACTC TCTTC GATAG ATTTA CCAGC 2i(A1)
CTGC AGAAA GAAGC AGGAA CACTC TCTTC GATAG ATTTA CCAGC 4i(A4)

PstI

AAAAT TCGTG ATTCA TGCCT GTATA GTAAT TGCTT TTGTG ATGTT 2i(A1)
AAAAT TCTTG ATTCA TGCCT GTATA GTAAT TGCTT TTGTG ATGTT 4i(A4)

CATAT TTTTA TTTAA CCAGA TATTG GTTCT GTAGC TTA 2i(A1)
TT GGTCT ATAAC CAAGA CATCG AATAC GTACT 3i(A4)
CATAT TTTTA TTTAA CCAGA TATTG GTTCT GTAGC TTATG CATGA 4i(A4)

ACAAC CGATG AGAGG TTTCT TTAAA GTAAA ATTTA ATCAA TACGG 3i(A4)
TGTTG GCTAC TCTCG AAAGA AATTT 4i(A4)

GGAAT CCTCT TATTA CTTTG TAAAC CCCTC TTAAA ATAAT AACCC 3i(A4)

CCTAA CTTCG ATATA AACCA AATTC TTTAT GGACG GTTTA TTACG 3i(A4)

AAACT GTTAA AATTT GTATT CGAA 3i(A4)

HindIII

Sequence of cloned der(1) DNA segments from the t(X;1).

```
                                                GGAT  TCAGC  ATAAG  TCATA FF1I
                                                GGAT  TCAGC  ATAAG  TCATA FF1K

TATTT  TTAAA  ATATG  TCAGT  TCTTA  CCATT  GTGTT  AGGGA  AACCA FF1I
TATTT  TTAAA  ATATG  TCAGT  TCTTG  CCATT  GTGT   AGGGA  AACCA FF1K

CCAAG  TACAC  ATTTA  TACAT  TTTTG  CTCAG  CATTT  TTCAG  GAGAT FF1I
CCAAG  TACAC  ATTTA  TACAT  TTTTG  CTCAG  CATTT  TTCAG  GAGAT FF1K

AATCT  CAATG  ATGAT  TCAAA  GAAAA  AAATA  ATAAC  TCATA  AGAAT FF1I
AATCT  CAATG  ATGAT  TCAAA  GAAAA  AAATA  ATAAC  TCATA  AGAAT FF1K

TTTTA  TTTAC  TGACT  GCTTA  GTCAC  AAAAG  ACAGA  ACACC  ATCTC FF1I
TTTTA  TTTAC  TGACT  GCTTA  GTCAC  AAAAG  ACAGA  ACACC  ATCTC FF1K

TTGAA  GTCTC  AGAAG  GAAGG  AGAAG  AGACT  GGGCC  CATGG  CAGTC FF1I
TTGAA  GTCTC  AGAAG  GAAGG  AGAAG  AGACT  GGGCC  CATGG  CAGTC FF1K


          GG  CTCAA  CCAG**GTGTGT  GAGCA  TGAAC  GGAAC  CAGCC FF1I
          GG  CTCAA  CCAG**GTGTGT  GAGCA  TGAAC  GGAAC  CAGCC FF1K

TAAGG  AGTGA  CGAGA  GATAG  ACGTT  CCGTG  GACAC  CCATG  TGCTT FF1I
TAAGG  AGTGA  CGAGA  GATAG  ACGTT  CCGTG  GACAC  CCATG  TGCTT FF1K

CGCTC  TCCAG  TC*****TGCCG  GGTAC  AGCTC  ACACC  GGACC  TTGTG FF1I
CGCTC  TCCAG  TC*****TGCCG  GGTAC  AGCTC  ACACC  GGACC  TTGTG FF1K

CCACT  CTCCA  GAGCT  GTCGT  AGACA  CCGAC  TACAA  CTGGT  CTACA FF1I
CCACT  CTCCA  GAGCT  GTCGT  AGACA  CCGAC  TACAA  CTGGT  CTACA FF1K

GCTTC  GAGTA  GGTCC  GGTGT  GGGAC  CGTGG  GGACG  ACGAA  GGTCC FF1I
GCTTC  GAGTA  GGTCC  GGTGT  GGGAC  CGTGG  GGACG  ACGAA  GGTCC FF1K

TCCGA  GAGGG  ACTTC  TCGGG  TCACC  GGAGG  ACAGA  TCT          FF1I
TTCGA  GAGGG  ACTTC  TCGGG  TCACC  GGAGG  ACAGA  TCT          FF1K
                                          XbaI
```

Sequence of cloned DNA fragments from the der(X) of the X;1 translocation.