PHYSICAL AND GENETIC MAPPING IN THE PROXIMAL SHORT ARM OF THE X CHROMOSOME

SMARAGDA KAMAKARI

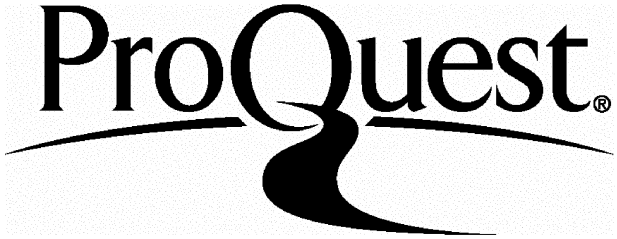
A thesis submitted for the degree of Doctor of Philosophy

UNIVERSITY OF LONDON

INSTITUTE OF OPHTHALMOLOGY

DEPARTMENT OF HUMAN MOLECULAR GENETICS

May 1994



ACKNOWLEDGEMENTS

I would like to thank my supervisors Susan Lindsay and Shomi Bhattacharya for their constant guidance and support throughout these last years. I would also like to thank Shomi not only for his help but also for his friendship and kindness. I also thank Susan for her friendship but also for her enthusiasm and encouragement which gave me confidence when I most needed it.

I thank Dr Alison Hardcastle for very useful suggestions during the editing of this thesis.

I thank all my colleagues at both Newcastle and London and especially, Dr Chris Inglehearn (for his integrity), Mr Neil Ebenezer (for excellent technical support), Mr Andy Curtis (for being a pleasure to work with), Miss Dawn Thiselton (for sharing xlRP agonies), Dr Tim Stout (for explaining ophthalmology brilliantly), Dr Jude Fitzgibbon (for encouragement during the writing up period), Dr Tim Jordan (for installing my computer), Drs Cheryl Gregory and Kevin Evans (for elucidating computer mysteries) and Mr Jeff Keen (for being a nice person at the end of the day).

I thank Amalia Sertedaki for sharing home sickness, secrets and wine in the (absurdly) cold nights at Newcastle.

I thank Simon Carter for being a good friend and source of sanity during some rather difficult times.

I thank all my friends in Greece and especially Eva, Lena and Milton for their constant concern and affection and including Sotiris for offering me an idea of how I would like my "Groundhog" day to be.

I thank Mani Nayudu for reminding me the merit of a genuine friendship and for sharing a flat with me without significant incidents of bloodshed. I also thank Alwyn for contribution to a peaceful homely environment.

Finally, I am deeply indebted to my family without the support and encouragement of whom my studies would not have been complete.

ABSTRACT

The aim of this thesis was to further develop the techniques of Fluorescence In Situ Hybridisation (FISH) for the localisation of markers on the X chromosome and to contribute to the physical mapping of the proximal short arm with particular emphasis on the region where the gene responsible for one form of Retinitis Pigmentosa (RP2) has been mapped by genetic linkage studies.

FISH technique was initially used to sublocalise markers to X chromosome regions. Two-colour FISH was subsequently used to order selected X specific markers because of their position in proximal Xp relative to reference markers DXS7 and DXS426. These markers are known to be tightly linked to the RP2 locus. On the basis of the in situ results one of the newly isolated markers was eventually mapped genetically as it appeared to map in the region of interest. Therefore, a microsatellite was isolated from this cosmid and sequenced. The unique sequences flanking the repeat were used to design Polymerase Chain Reaction (PCR) amplification primers. After establishing allele number, length and frequency of this marker, its genetic position was defined by following its inheritance in twelve X linked Retinitis Pigmentosa families. These families had been previously analysed and linkage data for a number of other proximal short arm markers were available. The new marker eventually mapped distal to the region of interest and, therefore, was not included in further physical mapping studies.

An alternative approach to generating markers was to construct a Yeast Artificial Chromosome (YAC) contig of the region using the DXS426 locus as a starting point for chromosome walking. A genomic YAC library was, therefore, screened by PCR and two YACs containing the DXS426 locus were isolated. The orientation of the two YACs was defined by end-cloning. Primers of the proximal end clone were used in order to isolate an overlapping YAC. The end clones of the overlapping YAC were isolated, sequenced and PCR amplification primers were designed. The above YAC clones covered a physical distance of approximately 1 Megabase and physically linked the DXS426 locus with the OATL1 locus in the proximal short arm of X chromosome.

This work has contributed to the high resolution physical and genetic linkage map of the proximal short arm of the X chromosome and has provided a resource for the identification of candidate genes involved in inherited diseases localised to this part of the chromosome.

PUBLICATIONS

Curtis, A. R. J., Roustan, P., Kamakari, S., Thiselton, D., Lindsay, S., and Bhattacharya, S. S. (1992). Two dinucleotide repeat polymorphisms at the DXS571 locus. *Human Molecular Genetics* **1**: 776.

Roustan, P., Curtis, A. R. J., Kamakari, S., Thiselton, D., Lindsay, S., and Bhattacharya, S. S. (1992). Dinucleotide repeat polymorphism at the DXS559 locus. *Human Molecular Genetics* **1**: 778.

Porteous, M. E. M., Curtis, A., Lindsay, S., Williams, O., Goudie, D., Kamakari, S., and Bhattacharya, S. S. (1992). The gene for Aarskog syndrome is located between DXS255 and DXS566 (Xp11.2-Xq13). *Genomics* **14**: 455-461.

Curtis, A. R. J., Headland, S., Lindsay, S., Thomas, N. S. T., Boye, E., Kamakari, S., Roustan, P., Anvret, M., Wahlstrom, J., McCarthy, G., Clarke, A. J., and Bhattacharya, S. S. (1993). X chromosome linkage studies in familial Rett syndrome. *Hum. Genet.* **90**: 551-555.

Roustan, P., Curtis, A. R. J., Kamakari, S., Thiselton, D., Lindsay, S., and Bhattacharya, S. S. (1993). Dinucleotide repeat polymorphism at the DXS573 locus. *Human Molecular Genetics* **2**: 92.

Thiselton, D., Kamakari, S., Curtis, A. R. J., Coleman, M. P., Davies, K. E., Linsday, S., and Bhattacharya, S. S. (1993). Dinucleotide repeat polymorphism at the DXS556 locus, *Human Molecular Genetics* **2**: 613.

Lindsay, S., Curtis, A. R. J., Roustan, P., Kamakari, S., Thiselton, D., Stephenson, A., and Bhattacharya, S. S. (1993). Isolation and characterisation of three microsatellite markers in the proximal long arm of the human X chromosome. Genomics **17**: 208-210.

Kamakari, S., Thiselton, D., Hardcastle, A. J., Hampson, M., Nayudu, M., Jones, S., and Bhattacharya, S. S. (1994). Physical mapping of Xp11: an interval containing the RP2 gene. 5th X Chromosome Workshop, Heidelberg, Germany, April 24-27 (In press).

DECLARATION

Unless otherwise stated, I declare that the experimental data contained within this thesis submitted for the degree of Doctor of Philosophy is my own original work, and has not been submitted for a degree at any other university.

Smaragda Kamakari MSc

CONTENTS

CHAPTER 3: FLUORESCENCE IN SITU HYBRIDISATION.

TABLES

LIST OF FIGURES

Figure 34: Autoradiographs of DNA sequencing gels showing the poly (CA) sequence within the PCR products obtained as described in fig. 33a.

Figure 35a-b: Sequence flanking the poly CA repeat near the proximal end of YAC 33CA11.

Figure 36: Ethidium bromide-stained agarose gel of PCR products from an X-chromosome somatic cell hybrid panel and YACs using primers flanking the CA repeat near the end of YAC 33CA11.

Figure 37: Autoradiograph showing PCR products of identical size for 10 unrelated individuals.

Figure 38: Sequence comparison of right end of 33CA11 YAC versus GenBank showing 83% homology with an Alu repeat.

Figure 39: An example of PCR screening of secondary ICI YAC pools using specific primers for the right end of YAC 33CA11.

Figure 40: a. PFGE of DNA plug samples from YACs 20CC8, 36BG12 and 19DH7.
b. Autoradiograph showing the size of the YACs after Southern blotting gel (a) and hybridisation with total human DNA.

Figure 41: Vectorette system for the isolation of YAC termini from YAC 20CC8.

Figure 42: Sequence of 20CC8 ends clones.

Figure 43: Ethidium-bromide stained agarose gel of PCR products from YACs36BG12 and 19DH7 using primers for the DXS1003 (276) locus.

Figure 44: Ethidium bromide-stained agarose gel of PCR products from an X-chromosome somatic hybrid panel using specific primers for the right end clone of YAC 33CA11.

Figure 45: Ethidium bromide-stained agarose gels of PCR products from YACs using primers from the left and right ends of 33CA11 YAC.

Figure 46: Ethidium bromide-stained agarose gels of PCR products from YACs using primers from the left and right ends of 20CC8 YAC.

Figure 47: A YAC contig around DXS426 locus.

CHAPTER 1: INTRODUCTION

1 Human gene mapping

1.1 Historical highlights

Human genetics is undergoing a dramatic metamorphosis. A subject that only a few decades ago was characterized principally by the cataloging of genetic diseases and their symptoms, the determination of their patterns of inheritance, and the measurement of their frequencies of occurence in various populations, human genetics has been revolutionized by molecular biology. The analytic precision and power of human genetics are increasing at a breathtaking rate.

The birth of molecular biology is often considered to be 1953, when Watson and Crick published their model for the double-helical structure of DNA. Since then, molecular genetics has built upon the foundation of transmission genetics established by Mendel in 1865 and his successors in the first half of the twentieth century. One of Mendel's greatest insights was the realization that there are *paired units of inheritance.* Mendel's work was forgotten for almost half a century and rediscovered simultaneously by de Vries, Correns, and Tschermak in 1900. It was almost immediately applied to human genetics by Garrod in 1902 in a paper on the "The incidence of alkaptonuria: a study in chemical individuality". In the first two decades of the twentieth century there were many fundamental advances in genetics including the discovery of concepts of linkage and recombination and the demonstration that the different blood types in humans are caused by multiple alleles at the same locus. In 1911, E. B. Wilson initiated human gene mapping by deducing that the gene for color blindness is located on the X chromosome. Two events in the 1940s foreshadowed the emergence of molecular genetics. First, there was the *one gene-one enzyme hypothesis* of Beadle and Tatum (1941) and second, the proof by Avery, McLeod and McCarty in 1944 that DNA carries inheritable information. In the following years, another series of events with special significance for human genetics involved the condition sickle cell anaemia (Neel, 1949 ; Pauling et al., 1949) and led to the recognition in 1957 that this genetic

disease was caused by an aminoacid substitution in beta-globin (Ingram, 1957). The transformation of human genetics into a discipline dominated by molecular analysis and molecular explanations of disease has accelerated enormously since the development of recombinant DNA technology in the mid-1970s. The first human gene to be cloned and completely sequenced was the gene for beta-globin (reviewed by Maniatis et al., 1980). Now, there are hundreds of cloned human genes. A panoply of rapid and sensitive techniques for the manipulation of DNA, RNA and protein stimulated an explosion of activity that touches virtually every aspect of human genetics, ranging from the research laboratory to the clinic.
The most direct impact of the rapidly increasing knowledge of the human genome and its variations will be on the diagnosis and treatment of genetic disease. In principle, all genetic diseases can be analysed in molecular terms, and a great many of them should eventually be treatable and/or curable because of precise molecular descriptions of the underlying defects in DNA and metabolic consequences. Although genetic disease is currently treated by non-gene-level methods, gene therapy i.e. the treatment of disease by replacement or correction of defective genes is a major goal in the application of genetic engineering technology to human disease.

1.2 The Human Genome Project

The largest and the most expensive biology project ever undertaken -the Human Genome Project- is underway. This is a truly international effort for the development of a detailed genetic and physical map of the human genome, to locate and identify virtually all human genes, and to determine the DNA sequence of most or all of the human genome. The seeds of the Human Genome Project were sown in 1977, when simple and efficient methods for sequencing were described. As soon as public discussion of the possibility of completely analysing the human genome began in 1984-1986, questions about the feasibility of the project dealt with problems of cost, optimal use of trained personnel, whether the full DNA sequence

should be the primary goal, and uses of the information to be obtained. Social, ethical and legal questions related to potential misuse of the genetic information were raised, the most profound of which concerned the possible application of genetic data for altering the basis of human disease, human talents, and social behaviour. Although debate continues, it is now recognised that biology and medicine have already benefited and research on physiology and development of organisms, the molecular basis of evolution, and other fundamental problems in biology could all be facilitated in the long run by the genome project.

1.3 Genome mapping strategies

1.3.1 General considerations

The ultimate goal of human gene mapping is to know the chromosomal location of every human gene and the order in which the genes are arranged. If we are going to search for genes that are involved in human disease we ought to have some idea of the size of the human genome and hence the magnitude of the task. The human haploid nuclear genome contains over three billion base pairs of DNA, distributed among the 23 pairs of chromosomes. The total number of human genes is still unknown. Estimates range between 30000 to more than 100000 (McKusick, 1992). There are two general approaches to mapping the human genome. The first, which owes its origin to the classical studies of T. H. Morgan and his colleagues on Drosophila, is genetic mapping, that is the use of linkage to assign the order of genes on particular chromosomes. The second, physical mapping, is more direct in that it entails the structural analysis of chromosomes with the objective of building up a map of their constituent genes. At its simplest level of resolution this involves chromosome banding, whereas the ultimate physical map would be a complete nucleotide sequence of the human genome. In practice gene hunting involves the combined use of both genetic and physical approaches.

1.3.2 Genetic and physical mapping strategies

Genetic linkage maps are mainly made by studying families and measuring the frequency with which a particular marker and the disease or other trait are inherited together, or linked. As the distance between the marker and the disease locus decreases there will be less chance of crossing-over and the number of recombinants will become fewer, i.e. the recombination fraction falls. It follows that this fraction (called θ) will vary from 0, which indicates tight linkage, to 0.5 indicating independent assortment. In practice it is usually necessary to work out the most likely value of the recombination fraction by making assumptions about the phase of linkage and combining data from more than one family. The probability of linkage over an assumed range of recombination fractions is estimated and divided by the probability of obtaining the same data as a result of independent assortment. Odds in favour of linkage are expressed on a logarithmic scale to produce a lod ($\log_{10}$ of the odds) score. By convention, a lod score of 3 (odds=1000:1) is considered to be "proof" of linkage, and a lod score of 2 (odds=100:1) is "strong evidence". A score of -2 is taken to rule out linkage at the θ value that leads to that score. It is apparent that distance on genetic maps is a measure of recombination frequency. Recombination fraction is converted into genetic distances by the use of a mapping function e.g. Kosambi mapping function. The correlation between recombination fraction and the genetic distance (cM) is linear between 0 and 0.1. Therefore, a θ value of 0.1 usually refers to 10cM. It has been calculated that the haploid genome is of approximately 30 Morgans.

In a seminal paper in 1980, Botstein and co-workers pointed out that inter-individual variations in nucleotide sequence, some of which could be detected as restriction fragment length polymorphisms (RFLPs), were potentially a treasure house of marker loci for linkage analysis. They predicted that only 150 different markers would be needed to link all human genes to chromosomal regions containing such markers. Soon, however, it became apparent that hundreds of DNA probes for highly polymorphic sequences, scattered widely over the genome, were required for

a complete human linkage map (White et al., 1985). Since then, great progress has been made toward such a linkage map. In late 1987, White et al. (HGM9, 1987) reported a 475-locus map separated on average by 10cM covering 17 of 23 chromosomes with an average resolution of 10cM and Donnis-Keller et al. reported the construction of their own 403-RFLP map of the human genome with linkage groups on all chromosomes and a resolution of 9cM. By 1991, 3000 human polymorphic markers had been isolated (Williamson et al., 1991), but 90% of them had a heterozygosity of less than 50% and they were unevenly spaced throughout the genome. The most informative of them, VNTRs (variable number of tandem repeats) or minisatellites (Nakamura et al., 1987), tended to cluster near the ends of the chromosomes (e.g. Royle et al., 1988). Although maps based on these markers had contributed greatly to the primary mapping of a number of diseases, they were insufficient for many applications such as mapping rare monofactorial diseases or refining linkage intervals to distances suited for gene identification. The efforts for the construction of a higher density linkage map with more informative markers were greatly facilitated by the recognition that polymorphisms originating from short interspersed tandem repeats or microsatellites, abundant and ubiquitous throughout the genome, could be exploited for genetic mapping if they were assayed by Polymerase Chain Reaction (PCR) (Weber and May, 1989 ; Litt and Luty, 1989). A genetic linkage map of the human genome, reported by the NIH/CEPH Collaborative Mapping Group in 1992 consisted of 1416 loci, 339 of which were microsatellites. Soon after another linkage map constructed by J. Weissenbach and colleagues (1992) was based entirely on the segregation analysis of 814 newly characterised highly polymorphic microsatellites with an average resolution of 5cM. The indispensable link between medicine and the genome has been the genetic map, with its ability to place inherited disease states with a resolving power determined by restriction fragment length polymorphisms (RFLPs) and microsatellite analysis at about 5-10 cM (or about 5-10Mb) between current markers. Once the chromosomal location of a disease-producing gene has been determined, fine

structure genetic mapping narrows the region to be searched. Subsequent studies often include examination of known genes mapping to the area and isolation of new genes by positional cloning strategies (Collins, 1992). The success of this approach has been stunning, and the pace is accelerating rapidly, due in large to the availability of sets of mapped genetic markers and improvements in physical mapping methods.

A physical map must include an ordered array of markers that faithfully represent their order and relative distances in the corresponding chromosomal DNA, and the map must be flexible enough to allow further refinement at the nucleotide level. The traditional methods of physical mapping to complement genetic localisations have been the cytogenetic techniques. The techniques include the use of rodent-human hybrid cells and in situ hybridisation, and both have been sharply improved in recent years. Rodent-human hybrid cell lines containing individual human chromosomes in whole or in part can localize a marker of interest to the level of a cytogenetic band on a chromosome. Hybrid cell lines can be used in various ways, including in situ hybridisation analysis and the development of chromosome or chromosomal region-specific libraries. In situ hybridisation to prometaphase chromosomes can be used to localise DNA segments to mapping intervals of specified regions of a chromosome or to correlate the actual distance between markers and distance observed on chromosomal spreads (Lawrence et al., 1988). In one application chromosomal in situ suppression hybridisation (CISS hybridisation) (Landegent et al., 1987 ; Lichter et al., 1988 ; Pinkel et al., 1988 ; Cremer et al., 1988) combined with image digitation and computer analyses resulted in a map of human chromosome 11 at 1-3cM resolution (Lichter et al., 1990). Simultaneous in situ hybridisation with cosmid clones detected with different coloured fluorescent dyes can be used to map loci located as close as 50kb in interphase nuclei (Trask et al., 1989, 1991). This method is a potentially powerful way of ordering a collection of cosmid clones on a chromosome of choice. Cytogenetic methods remain strictly

limited to locating the positions of DNA markers on the chromosomes or to mapping relative positions with respect to each other. Increased resolution in mapping has been provided by advances both in fractionation and in cloning of DNA. The important advance in fractionation of DNA was the development of pulsed-field electrophoresis (PFGE; Schwartz and Cantor, 1984), which allows modified electrophoretic apparatus to separate DNA molecules as large as 10Mb. Combined with the use of restriction enzymes which cut DNA rarely, PFGE has allowed construction of physical maps. Markers are physically linked when they hybridise to a common restriction fragment separated on a gel. Construction of such physical maps in the megabase range requires an adequate supply of markers and favourably placed restriction sites in the region of interest. This method has the intrinsinc limitation that it does not provide the DNA in a form that facilitates further study. Higher resolution or further analysis is largely dependent on cloned DNA fragments which must be arranged by overlaps to achieve long-range contiguity. The collections of clones can be of various types i.e. plasmids, phage, cosmids, but it is intuitively obvious that larger clones make physical mapping easier. The development of yeast artificial chromosomes (YACs ; Burke et al., 1987) as a cloning vector has promoted rapid progress in mapping in recent years. YACs provide isolated fragments of human DNA of up to a megabase or more in length, so that a few hundred would be sufficient to span an entire chromosome. In current practice, satisfying the requirements for the construction of a physical map has come to mean that the map would be made using overlapping YAC clones that provide long-range continous coverage, and which are aligned by markers at appropriate distances.

The markers which effectively format the map and make it useful can include, for example, genes, probes that detect polymorphisms, and sequence-tagged sites (STSs) (Olson et al., 1990) using primer pairs defining unique polymerase chain reaction products which can easily be resynthesised in any laboratory with an interest in genome analysis.

1.3.3 The first-generation physical map of the human genome

The first-generation physical map of the human genome (Cohen et al., 1993) was based on the extensive analysis of the CEPH yeast artificial chromosome library which contains 33000 clones with an average insert size of 0.9Mb covering the equivalent of 10 haploid genomes. Four major physical mapping techniques were used to generate structural and positional data from this large number of YAC clones; 1. fingerprints for all 33000 YAC clones were produced by detection of medium-repeat sequence containing fragments generated by three enzymes 2. 2100 polymorphic genetically mapped STSs were used to screen partially or totally the same 33000 YACs. This resulted in STS content mapping of 6580 clones (20% of the total library). 3. Homology relations between YAC clones were established as follows: a. preparation of high-density membranes containing pools of PCR products which were specific for each clone in a subset of 25000 YACs b. hybridisation of these membranes using inter-Alu PCR probes derived by amplification of individual YACs in a second subset of 5332 YAC clones. 4. Five hundred YACs containing genetically mapped polymorphic STSs (one every 7.4 cM) were positioned on metaphase chromosomes by the fluorescent in situ hybridisation (FISH) technique. The combination of the four techniques allowed the intergration of genetic, physical and cytogenetic maps. By combining the data it should be possible to reconstruct a continuum of overlapping clones spanning all chromosomes. However, the situation is far from optimal. The main disadvantage of YAC cloning is the presence of a large proportion (40-50%) of chimaeric clones, containing artefactually linked segments from non-contiguous regions of the genome. Another source of artefacts is the presence in the human genome of homologous regions which can also result in erroneous contig assembly. Finally, the investigators undertook an alternative multilevel mapping approach in which physical map construction was directly based

on inte gration with the genetic map. Therefore, continua of overlapping clones were assembled for short, genetically defined adjacent intervals by screening the CEPH library with polymorphic STSs corresponding to the 1993 meiotic recombination map of Genethon. The result of this approach was an 87% physical coverage of the total genetic length of the human genome. The efficiency of the STS screening was not uniform for all regions of the genome. Poor coverage, for example of chromosomes 19, 17 and 1p was due mainly to an inability to obtain at least one positive YAC per STS in these regions. Despite the limitations of the present map, access to the structural and positional information for individual YACs will certainly help the current mapping effort.

2 X chromosome

2.1 General considerations

The X chromosome is the most extensively studied of all human chromosomes due to a wide interest in X-linked diseases and in the phenomenon of X chromosome inactivation. Both of these features are related to the different dosage of X in males and females.

Because males have only a single X chromosome, recessive diseases tend to be revealed; this accounts for the large numbers of X-linked diseases (McKusick, 1992) and for their characteristic inheritance pattern. The need to understand these diseases and to provide diagnostic tools for genetic counselling in affected families was a major impetus for mapping studies. As a result of this driving force, the X chromosome was the first to have a genetic map based on restriction fragment length polymorphisms (RFLPs) and systematic approaches to physical coverage were undertaken that have been expanded as part of the Human Genome Initiative. Genes for the two X-linked diseases (chronic granulomatous disease and Duchenne muscular dystrophy ; Royer-Pocora et al., 1986 ; Monaco et al, 1986) were the first to be isolated by a mapping approach in the absence of functional information about

the gene, by what is now called positional cloning (Collins, 1992). The isolation of the Duchenne muscular dystrophy gene (DMD) was an enormous undertaking; it is by far the largest gene known in any organism (2.4Mb), contains many introns of 100 to 200kb, and may take 24 hours to be transcribed from end to end (Mandel et al., 1992). More recently, the study of two other diseases (the fragile X mental retardation syndrome and spino-bulbar muscular atrophy) uncovered a new and unexpected mutation mechanism, the expansion of trinucleotide repeats (Fu et al., 1991 ; La Spada et al., 1991) detailed later in the introduction (p.23)

2.2 X chromosome physical and genetic map

A unified genetic, physical and functional map of the human X chromosome is being built through a concerted, international effort. About 60% of the 160Mb of the X chromosome DNA have been cloned in overlapping, ordered contigs derived from yeast artificial chromosomes. There are 85 cloned, expressed and localized genes, (44 in p arm and 41 in q arm) and 93 breakpoints (54 in p arm, 38 in q arm and 1 in the centromere) (Report of the Fourth International Workshop on Human X Chromosome Mapping ; XCW4, 1993). Contig assembly is based on a variety of methods using STS content and hybridisation probes. Mapping strategies include all-walking methods with YAC-end inserts, the use of random STSs to recover YACs for defined chromosomal regions and the use of Alu-PCR products from YACs, somatic cell hybrids or radiation hybrids to screen reference libraries and detect overlaps. The longest set of unambigously ordered and oriented contigs have been assembled across the Xq24-qter region. More than 95% of this 50Mb region has been covered in contigs. It is currently the best characterised region of the X chromosome with large stretches further studied at 50-300kb resolution for their overall base content of GC, for the presence of rare-cutter restriction sites and CpG islands, and for the distribution of Alu and L1 repetive elements. Other well characterised regions are the proximal long arm from the centromere to Xq13 with more than 50% coverage in contigs (Vulpe et al., 1993) and band Xq22 which is

almost entirely covered by a 7.5Mb contig (Vetrie et al., 1993). The contig results from many groups are now being intergrated by the use of STSs and linkage probes that are in common. Physical mapping is now reaching the phase of "closure". It is a truism, however, that the last 10% of the map can require 90% of the effort, with contigs presenting a variety of problems including: regions that are unstable, recombined, or unrecovered in cloned DNA ; sequence elements that are repeated along the X ; clones that contain large internal deletions or bring together cloned fragments from disparate regions of the X. A major aid to closure, and of course a major route to the identification of disease genes, is the corresponding genetic map. There are four framework maps for the X chromosome consisting of RFLPs and microsatellite markers (about 70 microsatellite markers have been characterised for the X chromosome). The maps were constructed by different groups in the last two years and they are in good overall agreement (Fain and Barker, XCW3 and Chromosome Coordinating Meeting, 1992 ; Weissenbach et al., 1992 ; Murray et al, XCW, 1993 ; Mulley et al., XCW, 1993). The genetic distance between marker DXS143 at Xp22.31 and markers DXS52/DXS15 at Xq28 is 210, 205, 173 and 180cM in the maps of Murray et al., Mulley et al., Fain and Barker, and Weissenbach et al., respectively.

A striking feature of the genetic map is the tenfold difference in recombination between males and females for the 2.6Mb pseudoautosomal region (PAR) shared between the tips of the short arms of the X and Y chromosomes. Comparison of genetic and physical distances between markers in this region has demonstrated that the overall recombination frequencies are approximately tenfold lower in female, compared to male meiosis, except for the most telomeric 80kb. That interval has showed comparable recombination rates in males and females, and constitutes a recombination hot spot, with a 30-fold higher rate than expected from its physical size (Rappold, XCW4, 1993, abs. 6). The pseudoautosomal region is particularly rich in highly polymorphic minisatellite markers and it remains to be seen whether these tandem repeat sequences are involved in the high male recombination rate.

Increased recombination is also found within the dystrophin gene (DMD) (Abbs et al., 1990) and in regions flanking the fragile X site in Xq27.3 (FRAXA), although the fragile locus itself is not a hot spot for recombination (Richards et al., 1991). The pericentromeric region shows less recombination than expected from its size (e.g. Mahtani et al., 1991) perhaps reflecting a centromere effect on recombination frequency.

A microsatellite marker at Xq28 shown recently to recombine between X and Y in about 2% of male meioses defines a new pseudoautosomal region (Freije et al., 1992).

2.2.1 Xp21: The region containing the DMD gene

Duchenne mascular dystrophy (DMD) and the less severe Becker dystrophy (BMD) are X-linked myopathies arising from mutations in a large genomic region encoding a 14 kb mRNA which has been completely sequenced (Monaco et al., 1986 ; Burghes et al., 1987 ; Koenig et al., 1987, 1988). The first indication that the gene responsible for DMD and BMD is in band Xp21 came in the late 1970s and early 1980s when rare females with the Duchenne or Becker phenotype were described, each of whom had a de novo X autosome translocation with a breakpoint in band Xp21 (reviewed in Worton and Thompson, 1988). Linkage studies had also implicated Xp21 as the locus for the DMD gene (e. g. Murray et al., 1982). The third line of evidence came from a small set of patients with complex phenotypes including DMD, and particularly from study of BB, a male patient with a cytogenetically visible deletion in bands Xp21 and a pathology including DMD, chronic granulomatous disease, McLeod syndrome and pigmentary retinopathy (Francke et al., 1985). Two strategies designed to take advantage of the unique patients with structural rearrangements at the site of the gene have been successful in isolating genomic clones from within the gene. The approach of Kunkel and his colleagues (1985) depended on the isolation of multiple clones from within the region known to be deleted in BB. DNA from an XXXXY male was cleaved with the

restriction endonuclease MboI. Sheared DNA from patient BB was added in a large excess in a competitive hybridisation reaction (the phenol enhanced reassociation technique-pERT) to compete selectively with all sequences except those from the deleted segment. Among the reassociated molecules, only the perfectly aligned MboI-digested molecules were clonable, and these were enriched for sequences missing from the patient BB. One of the eight pERT clones derived from the BB deletion (called pERT87) detected DNA deletions in a proportion of male patients with DMD and was tightly linked to the disease in family studies. The pERT87 sequence became the start point for a bidirectional chromosome walk along a normal X chromosome which led to the isolation of 220kb of contigous DNA in this region, later designated as the DXS164 locus. The second approach leading to the DMD gene was undertaken by Worton and his group (1984). The approach was dependent on the identification of an X ; 21 translocation female, with the breakpoint in chromosome 21 in the ribosomal DNA cluster on the short arm (Verellen-Dumoulin et al., 1984 ; Worton et al., 1984). Ribosomal DNA probes were used to identify and clone from the patient a segment of DNA that spanned the translocation junction (Ray et al., 1985). This cloned segment was obtained from a mouse-somatic cell hybrid containing the der(X). The junction clone designated XJ1 contained 620bp of rDNA at one end, and about 11kb of X-chromosome sequences at the other. Chromosome walking from XJ1 along a normal X chromosome yielded about 120kb of the human X chromosome DNA (the DXS206 locus) derived from both sides of the junction site. Linkage analysis showed that the DXS206 locus was within or very close to the DMD gene and deletions in patients confined to this locus confirmed that it came from the DMD gene. The strategy to locate exons of the gene was to test individual genomic subclones from DXS164 and DXS206 for hybridisation with expressed sequences from muscle and/or to test for sequences conserved across species. Monaco et al. (1986) reported the first exons in the DXS164 locus by identifying a subclone that hybridised with monkey, bovine, mouse, hamster and chicken DNA in a "zoo blot". This clone revealed a message of

14kb on Northern analysis of muscle, which subsequently was shown to code for the protein that is defective in DMD. The genomic locus in Xp21.2 consists of a minimum of 65 exons (Koenig et al., 1987) spread over 2300kb, as measured by pulsed-field gel electrophoresis (van Ommen et al., 1986 ; Kenwrick et al., 1987 ; Burmeister et al., 1988 ; Den Dunnen et al., 1989). The study of a brain-specific promoter (Feener et al., 1989) located proximal to the muscle-specific promoter extends the genomic locus 100kb to approximately 2400kb (Boyce et al., 1991). Perhaps due to the large size of this locus, the majority of mutations giving rise to DMD and BMD are intragenic deletions (~65%) or duplications (~5%) (Forrest et al., 1987 ; Koenig et al., 1987 ; Den Dunnen et al., 1987 ; Hu et al., 1990). The severity of the phenotype in most cases depends on the effect of deletions and duplications on the open reading frame of the spliced transcript (Monaco et al., 1988 ; Malhotra et al., 1988 ; Hu et al., 1990), and exceptions may be due to alternatively spliced transcripts (Chelly et al., 1990). The product of the DMD gene, dystrophin, is a large 427 kDa protein with aminoacid sequence homology to the spectrin family of membrane cytoskeletal proteins. Dystrophin is found in smooth muscle throughout the body and in brain but is most abundant in skeletal and cardiac muscle, where it has been localised to the inner face of the plasma membrane (Zubrzycka-Gaarn et al., 1988). Two YAC contigs containing the complete DMD gene have been constructed in two parallel studies: a 3.2 Mb contig of 36 overlapping YACs (Monaco et al., 1992) and a 2.6 Mb contig of 34 overlapping YACs (Coffey et al., 1992). The former contig was constructed using a combined analysis of the exon content and the long-range SfiI restriction map of the YACs (compared to the known long-range restriction map of the DMD gene) and the latter using an STS-based approach (to identify and overlap the YAC clones ; see also chapter 5, section 1.2.4.1). A minimum of 7 and 12 overlapping YACs represent the complete 2.4 Mb DMD gene in the 3.2 Mb and 2.6 Mb contigs, respectively. These overlapping YAC clones were recombined in yeast to reconstruct the entire genomic locus in a single YAC (with the exception of a 100kb fragment containing exon 60 ; Den-Dunnen et

al., 1992). A single recombined YAC clone containing the complete DMD gene with exons, introns, and regulatory elements would be useful for future gene expression and possible gene therapy studies.

Proximal of this contig, in Xp21.1, a large 1.7 Mb YAC contig contains the genes for McLeod syndrome (XK), chronic granulomatous disease (CYBB), retinitis pigmentosa 3 (RP3) and ornithine transcarbamylase (OTC) (XCW4, 1993, abs. 28 and 47). Cosmid contigs underlying several of the YACs are being used to search for candidate genes for XK and RP3 (see section 3.3.2). The characterisation of a 5.5 kb transcript as a candidate gene for McLeod syndrome has been reported (XCW4, 1993, abs. 47).

2.2.2 The proximal short arm: Xp11.3-Xpcen

The region between Xp11.3 and Xp11.22 consists of two main YAC contigs as reported in 1993 during the Fourth International Workshop on Human X chromosome : a 1.3 Mb contig containing DXS7, monoamine oxidase A and B genes (MAOA and MAOB) and the gene for Norrie's disease, and a second 1.7 Mb YAC contig containing the genes for murine sarcoma viral oncogene homolog, ARAF1, neuron-specific phosphoprotein synapsin I, SYN1, tissue inhibitor of metalloproteinases, TIMP, serum glycoprotein properdin P factor, PFC, the polymorphic loci DXS1003 and DXS426, the zinc finger genes ZNF21 and ZNF81, the ETS-related gene ELK1 and the ornithine aminotransferase-like sequence cluster (OATL1).

The MAOA and MAOB genes were mapped to the short arm of the X chromosome in band Xp11.23 (Lan et al., 1989) and shown to be linked to the Norrie disease gene (Sims et al., 1989b ; see section 3.3.3) causing an X-linked recessive neurologic disorder characterised by congenital blindness, sensory neural deafness and mental retardation A Norrie kindred from Finland with an atypical phenotype was shown to have a submicroscopic chromosomal deletion encompassing the DXS7 locus (de la Chapelle et al., 1985) and the MAOA and MAOB genes (Sims et

al., 1989a) and to have a functional MAO enzyme defficiency state. (Sims et al., 1989a ; Murphy et al., 1990). The MAO genes were considered candidate genes for this disorder based on the role of biogenic amines in neuronal development (Lauder, 1985) and retinal transmission (Hadjiconstantinou et al., 1983 ; Osborne, 1981), as well as the developmentally regulated expression of MAOA and MAOB genes in the brain (Levitt et al., 1982 ; Thorpe et al., 1987) and the retina. MAOA and MAOB were excluded as candidates by DNA, enzyme and metabolic studies (Sims et al., 1989b). DXS7 locus and MAO genes are separated by 140kb (Chen et al., 1992a). MAO genes consist of 15 exons and spanning about 90kb each and they are found in a tail-to-tail configuration with the 3' coding sequences separated by about 40kb (Sims et al, 1992). Both MAOA and MAOB genes contain $(GT)_n$ polymorphisms in intron 2 (Black et al., 1991 ; Konradi et al., 1992). A recombination between the $(GT)_n$ polymorphism in MAOB gene and the NDP locus in combination with a submicroscopic deletion in a NDP patient defined the obligate region containing this gene to a chromosomal segment of 150kb and led to its identification (Sims et al., 1992 ; see section 3.3.3). The orientation of DXS7, MAO and NDP genes in this region has been shown to be: tel-DXS7-5' MAOA 3'-3' MAOB 5'-3' NDP 5'-cen (Sims et al., 1992 ; Chen et al., 1992b).

Genes for SYN1, ARAF1, TIMP and PFC have been cloned (Kilimann and DeGennaro, 1985 ; Goundis and Reid, 1988 ; Nolan et al., 1991) and mapped to the proximal short arm of the human X chromosome (Yang-Feng et al., 1989 ; Huebner et al., 1986 ; Goundis et al., 1989). Synapsin I is a phosphoprotein composed of two alternatively sliced forms, synapsin Ia and Ib, which are concentrated at the membranes of the synaptic vessicles and are thought to be involved in the regulation of neurotransmitter release into the synaptic cleft (Bahler et al., 1990). A-raf-1 is a cellular oncogene first identified by virtue of its homology to the v-raf oncogene, an acutely transforming retroviral gene (Huleihel et al., 1986 ; Huebner et al., 1986). Its homology to c-raf-1 suggests that it is a serine/threonin kinase (Huebner et al., 1986). TIMP encodes a glycoprotein inhibitor of collagenase,

stromolysin, and gelatinase, acting via the irreversible formation of inactive complexes, and is therefore extremely important in the regulation of connective tissue breakdown. It also appears to act as an erythroid hematopoietic growth factor (Gasson et al., 1985). Properdin is an important positive regulator of the alternative pathway of the complement system via interactions with the C3 and C5 convertase enzyme complexes. Properdin-deficient individuals display severe, fulminant, pyogenic bacterial infections with a high mortality rate (Holme et al., 1989). A deficiency locus has been mapped to Xp21.1-Xcen. The four genes are entirely contained within a 70-kb contigous stretch. TIMP lies within an intron of the SYN1 (downstream of exons 2-5) gene. The direction of transcription for TIMP gene has not been determined. ARAF1 lies in the proximity of the 3' end of SYN1 and is transcribed in opposite direction and contains a highly polymorphic dinucleotide repeat located 1kb downstream of the synapsin gene. PFC lies within 5kb of the 5' end of the SYN1 gene. The orientation of the four loci as suggested by Derry and Barnard is: Xpter-5' ARAF1 $(CA)_n$ 3' - 3' SYN1- TIMP-SYN1-5'- 3' PFC 5' Xpcen. Coleman et al. (1991) described a microsatellite for PFC and demonstrated by YAC analysis that PFC and the polymorphic locus DXS426 (Luty et al., 1989) lie within 390kb. In the same study a single obligate recombinant between (DXS426, PFC) and (TIMP, OATL1) in a CEPH family indicated that (PFC, DXS426) lay distal to (TIMP, OATL1). In addition, Kirchgessner et al. (1991) reported recombination in CEPH families between the ARAF1/SYN1 CA and a polymorphism identified by a probe defining the TIMP locus which suggested that TIMP was proximal to SYN1, with an estimated genetic distance of 4.2cM. The combined linkage data suggested the order SYN1/ARAF1 CA - (PFC, DXS426)-TIMP which was difficult to reconcile with the data that TIMP lies within one of the introns of SYN1 and within 70kb of the SYN1/ARAF1 CA. Additionally, given the physical proximity of these loci, the estimated genetic distance of 4.2cM was surprising unless the TIMP-CA interval represented a recombination hot spot or one or both loci were duplicated and lie at different chromosomal loci (Derry and Barnard, 1992). Therefore, the relative order

of DXS426, PFC, TIMP, SYN1 and ARAF1 was uncertain until recently. Based on further physical data, the order of these loci in the 1993 X-chromosome workshop was suggested as: Xpter-ARAF1-3'SYN1-TIMP-5'SYN1-(DXS426, PFC)-OATL1-Xpcen.

Ornithine-δ-aminotransferase (OAT) is a mitochodrial matrix enzyme that catalyses the interconversion of ornithine to D1 pyrroline-5-carboxylate (Valle and Simell, 1989). Inherited deficiency of OAT causes gyrate atrophy of the choroid and the retina, a slowly progressive retinopathy. The OAT cDNA and the structural gene have been cloned and the gene has been mapped to chromosome 10q26 (Mitchell et al., 1988). Additionally, several OAT-related sequences have been mapped to the short arm of the X chromosome (XOAT), each of which is contained in a distinct EcoR1 fragment (Geraghty et al., 1993). Sequence analysis of three XOAT shows that they are processed pseudogenes, all of which have undergone deletions and/or insertions (Looney et al., 1987). Using hybrid cell panels, it was demonstrated that there are two closely spaced clusters of XOAT (Lafreniere et al., 1991). The distal locus (OATL1 at Xp11.23) contains eight XOAT EcoR1 fragments; the proximal locus (OATL2, between Xp11.22 and Xp11.21) contains four. Several retinal degenerative diseases have been mapped in the vicinity of OATL1 (see section 3.3). Additionally, YACs containing either OATL1 or OATL2 clusters cross two independent synovial sarcoma translocation breakpoints (Leeuw et al., 1992 ; Knight et al., 1992), indicating that perhaps two synovial sarcoma loci may have duplicated with the OAT sequence cluster. Based on the speculation that the OAT-related sequences in this region may contribute to transcription units and second, that mutations in these sequences may contribute to the pathogenesis of one or more of these disorders, a YAC containing the most distal part of OATL1 has been recently used as a probe to screen a retinal cDNA library (Geraghty et al., 1993). The investigators, however, concluded that the XOAT sequences in this region are processed pseudogenes. Furthermore, they failed to establish any relationship

between the four isolated and characterised cDNAs and the retinal degeneration disorders or the synovial sarcoma in this region.

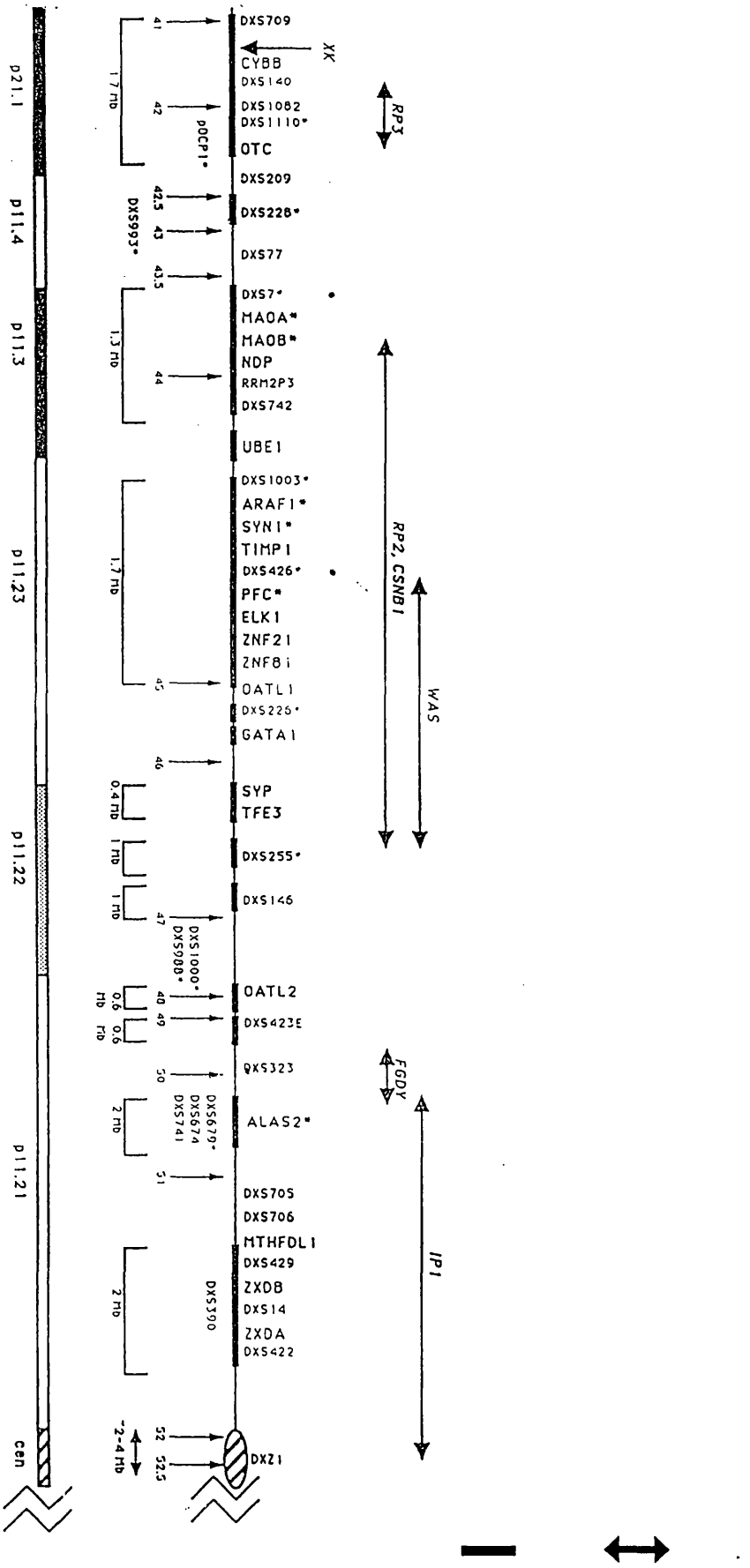
Other genes and loci in Xp11.23 and Xp11.22 including the genes for ubiquitin, synaptophysin (SYP), transcription factor binding to IGHM enhancer 3 (TFE3) and the locus DXS255 (Fraser et al., 1987) have been isolated in YACs but have not yet been connected to the two large contigs. The most recently suggested overall consensus order of these loci is shown in fig. 1. The estimated genetic length between DXS7 and DXS426 is 11.6cM (Coleman et al., 1991), between MAOA and PFC is 13.7cM (Murray et al., XCW1993) and between DXS7 and DXS255 is 15cM (Fain and Barker, XCW3, 1992 ; Mulley et al., XCW4, 1993 ; Murray et al., XCW, 1993).

In the interval between marker DXS255 (Xp11.22) and the centromeric repetitive probe DXZ1, two main 2Mb contigs have been constructed, the proximal of which contains the delta-aminolevulinate synthase gene (ALAS2) gene and the distal is centered around the locus DXS14 (fig. 1)

2.3 Mapping X-linked diseases

At least 16 X-linked disease genes have been cloned on the basis of prior knowledge of the defective protein. However, for most of the remaining diseases, the biochemical defect was unknown or very uncertain and it was thus necessary to use mapping strategies to identify the corresponding genes. Genes for seven X-linked diseases have been isolated by positional cloning (Table 1), including those for Duchenne muscular dystrophy and the fragile X mental retardation syndrome, the two most common severe X-linked diseases. In several other cases, mapping studies were useful in selecting an independently cloned "candidate" gene, which could then be validated by finding gene-specific mutations in patients. The most

Figure 1: The proximal short arm of the X chromosome consensus map (not to scale)
(Report of the fourth international workshop on human X chromosome mapping, 1993)



- Order of markers above the line is known
- Microsatellite polymorphosms are denoted with an asterisk
- Markers below the line are unordered, but are known to map in the interval

Table 1: Positional cloning of X-linked diseases. T, D, and Y indicate whether translocations (T), large deletions (D) or YACs (Y) were crucial for localisation and cloning of the gene.

Year cloned	Disease	Method	References
1986	Chronic granulomatous disease	D	Royer-Pocora et al.
1986	Duchenne muscular dystrophy	D, T	Monaco et al.
1990	Choroideremia	T, D	Cremers et al.
1991	Fragile X syndrome	Fragile site, Y	Kremer et al. ; Oberle et al. ; Verkerk et al.
1991	Kallman syndrome	D, Y	Franco et al ; Legouis et al.
1992	Lowe syndrome	T, Y	Attree et al.
1992	Norrie disease	D, Y	Berger et al. ; Chen et al.

recently identified genes using these strategies include the disease genes for Bruton's agammaglobulinemia, in Xq22 (coding for a tyrosine kinase called ATK, Vetrie et al., 1993 ; Tsukada et al., 1992), adrenoleucodystrophy, in Xq28 (Mosser et al., 1993), glycerol kinase deficiency, in Xp21.3 (Sargent et al., 1993 ; Walker et al., 1993) and severe combined immunodeficiency in Xq12-q13.1 (identified as the interleukin-2 receptor gamma chain ; Noguchi et al., 1993). Linkage studies in affected families are being actively persued in many laboratories for more than 50

diseases (XCW4, 1993). Increasingly, microsatellite markers are used for more efficient and precise mapping, and provide diagnostic assays for genetic counselling.

For about ten diseases, rare affected females have been found with balanced X-autosome translocations. In these patients the normal X chromosome is generally inactive, and the translocated X active to maintain the dosage of the autosomal genes. Such translocations have provided precise localisation for the relevant disease genes and have been instrumental in the cloning of several genes (Table 1). This approach is now being used for anhidrotic ectodermal dysplasia (EDA ; developmental disorder mapping at Xq12-q13) and has been recently used for the isolation and characterisation of the Menkes disease gene (MNK, Xq13 ; Vulpe et al., 1993 ; Chelly et al., 1993 ; Mercer et al., 1993).

Males have only one X chromosome, and it was thus surprising to find rare male patients who survive with deletions (often detected cytogenetically), encompassing several megabases of DNA. In most cases the lack of function for genes in the deleted region results in a contigous gene syndrome, allowing very accurate mapping of the relevant genes (Ballabio, 1991). This was first observed in the case of the BB deletion encompassing part of DMD and genes for chronic granulomatous disease (CYBB), McLeod syndrome (XK), and retinitis pigmentosa-3 (RP3). (Franche et al., 1985). The BB deletion was instrumental for the cloning in 1986 of DMD and CYBB, and now is being used to clone XK and RP3 (Ho et al., 1992 ; Musarella et al., 1991). Large deletions have been found in the Xp22.3, Xp21.2, Xq21.2, Xq25, and Xq27 regions. Very likely such regions have relatively low gene densities, or the deletions would be lethal (Mandel al., 1992). In the Xp21.2 region this can be accounted for in part by the huge size of the DMD gene, but more generally it illustrates the great heterogeneity in gene density throughout the genome, with gene-poor regions (in general AT-rich, Giemsa dark bands) alternating with gene-rich ones, such as the distal part of Xq28 where many genes

are clustered ; e.g. nine genes have been localised in 300kb of the color vision-G6PD region. (Davies et al., 1992 ; Maestrini et al, 1992 ; XCW4, 1993).

2.4 Mutations in X-linked disease genes

For X-linked diseases that severely decrease reproductive fitness in affected males, the number of mutations in each generation decreases by one-third (since males have one-third of the X chromosomes in the population) and the particular mutation becomes extinct after a few generations. This does not apply to mutations with mild or no effect on reproductive fitness such as those responsible for glucose-6-phosphate dehydrogenase (G6PD) deficiency, color blindness, or some cases of mild hemophilia A and B. It is already clear that there is a striking difference in deletion frequency in various diseases. In X-linked icthyosis, 80% to 90% of the mutations are large deletions encompassing the entire gene (Yen et al., 1990 ; Ballabio et al., 1990). Duchenne muscular dystrophy is another disease with a high frequency of partial deletions (60 to 70%) and a significant level of partial duplications (6 to 7%) (Koenig et al., 1989). For most other diseases analysed thoroughly (notably hemophilia A and B, as well ornithine transcarbamylase and hypoxanthine phosphoribosyl transferase deficiencies) the frequency of deletions or other rearrangements detectable by Southern blot is in the order of 5 to 15%. Databases of point mutations have been established for hemophilia A and B (Gianelli et al., 1992 ; Tuddenham et al., 1991). Because of the smaller size of the coagulation factor IX gene (which has been totally sequenced), hemophilia B has been more thoroughly analysed, with about 400 mutations reported (Gianelli et al., 1992). It is interesting that two hemophilia A mutations are due to insertion of a LINE1 interspersed repeat element (Dombroski et al., 1991). A case of hemophilia B resulting from de novo insertion of an Alu repeat has also been reported (Mandel et al., 1992). These observations indicate that transposable elements can cause disease in humans, although it is rare a occurence. For G6PD, only missense point

mutations have been found, with the exception of one single codon deletion, and it is likely that mutations which totally abolish G6PD activity are early lethals (Vulliamy *et al*,1992).

Expansion of a trinucleotide repeat sequence is a newly discovered disease-causing mechanism (Caskey et al., 1992) originally described as the sole type of mutation in two X-linked diseases. The fragile X mental retardation is caused by an astonishingly unstable expansion of a CGG repeat in a 5' exon of the gene FMR1 (Kremer et al., 1991). Large "full" mutations lead to methylation of CG dinucleotides in the region, which correlates with loss of expression of the FMR1 transcript (Eichler et al., XCW4, 1993, abs. 55). In spinobulbar atrophy, the mutation is a more moderate expansion of a CAG repeat in the NH_2-terminal coding region of the androgen receptor gene (AR) (La Spada et al., 1991) while other heterogeneous mutations in AR result in the completely different phenotype of testicular feminisation. The identification of mutations in the FMR1 gene has had a profound impact on both the understanding of the peculiar inheritance of the fragile X syndrome, and its diagnosis (Kremer et al., 1991). Trinucleotide expansion has been found more recently in myotonic dystrophy and Huntington's disease, both autosomal dominant diseases (Caskey et al., 1992 ; Richards and Sutherland, 1992 ; Huntington's Disease Collaborative Research Group, 1993).

2.5 Mapping of genes controlling or escaping X inactivation

The stable inactivation of one of the two X chromosomes in females is a fascinating phenomenon as it can spread over 100Mb of DNA, turning off several thousands of genes. It is thought to occur in three steps: initiation early in embryogenesis at a site called the X inactivation centre, propagation along the length of the chromosome, and stabilisation of the inactive state of individual loci (Lyon, 1988 ; Ballabio and Willard, 1992).

The study of structurally abnormal X chromosomes in humans and mouse gave strong evidence for the existence of a cis acting locus, the X inactivation centre

(XIC), necessary for initiation of inactivation (Lyon, 1988 ; Ballabio and Willard, 1992). This locus was mapped in both species to a region between the ectodermal dysplasia locus (EDA, or *Tabby* in mouse) and the phosphoglycerate kinase gene (PGK) (Brown et al., 1991), a region of 2.5Mb in humans now largely covered by a YAC contig (XCW,1993). A gene cloned by serendipity was found to map to this region that is transcribed only from the inactive X chromosomes (Brown et al., 1991 ; Borsani et al., 1991). The portion characterised does not appear to code for a protein. The localisation and expression of this gene, named XIST (for X-inactive-specific transcripts), suggests that is involved in initiation and spreading of inactivation. Studies of two deletions in females showing nonrandom inactivation have suggested that a region near the FRAXA-IDS loci (in Xq27.3-proximal q28) may be involved in inactivation of distal loci (Clarke et al., 1992). A second surprise was the discovery that several genes along the X appear to escape inactivation (Ballabio and Willard, 1992). Although this feature was expected for genes in the pseudoautosomal region, it was also found for at least three genes in three other regions (ZFX, UBE1, and RPS4X in Xp22.1, Xp11.23, and Xq13.1, respectively). These genes are subject to inactivation also in mouse.

Stabilisation of inactivation is very well correlated with DNA methylation (at the cytosine of a CpG dinucleotide), which occurs at CpG-rich regions (CpG islands) near the 5' end of many genes (Migeon, 1990). Analysis of such differential DNA methylation near some polymorphic X loci is also used for the study of biased inactivation in female carriers of some X-linked diseases (Boyd and Fraser, 1990 ; Hendriks et al., 1992).

2.6 Of mice, humans and kangaroos

It was postulated by Ohno (1967) that genes which are X-linked in one mammalian species should be X-linked in all others. The rationale for this was that these genes function in a single dose, while autosomal genes function in double dose. Therefore exchange of a large chromosome segment from autosomes to X would result in

impaired dosage and would be expected to be lethal. The mouse X chromosome is being actively mapped, and because of the conservation of protein-coding DNA sequences between mouse and human, the same genes or conserved markers can be mapped in parallel in the two species (Mandel et al., 1992). Indeed, almost all genes tested that are X-linked in humans have their counterpart on the mouse X chromosome and five large regions can be defined that appear to have conserved gene content and order between the human and mouse X chromosomes (Brown et al., 1992). Thus, relatively few intrachromosomal rearrangements have occured on X since divergence of the two species. The correspondence established between the maps of the two chromosomes has allowed the validation of mouse mutants as homologous models for human diseases with similar phenotypes. This is based on their equivalent map positions, and may be useful for positional cloning of such loci (Tumer et al., 1992). For example, the HYP, Ta, and Mo mouse mutants correspond to the genes for hypophosphatemic rickets (HYP), ectodermal dysplasia (EDA), and Menkes disease (MNK), respectively in humans.

An interesting exception to the conservation of genes is the apparent absence in mouse of sequences homologous to the steroid sulphatase, Kallmann syndrome (KAL), MIC2, and GS1 genes, four genes that are closely located in Xp22.3 in humans (Ballabio and Willard, 1992). Deletion of this region in males results in mild phenotype, and it may have been similarly deleted in a mouse ancestor. Another surprise was the finding that all genes tested that are located on the short arm of the human X are autosomal in marsupials and monotremes. This is at variance with Ohno's hypothesis, and suggests that the short arm was of autosomal origin and was added to the X chromosome in eutherian mammals (Watson et al., 1991).

3. Diseases in the proximal short arm of the X chromosome

3.1 Wiscott-Aldrich syndrome

The Wiscott-Aldrich syndrome (WAS) is an X-linked recessive immunodeficiency disease in which several glycoproteins of the formed elements of the blood are

abnormal (Remold-O'Donnel and Rosen, 1990). Clinically the syndrome is characterised by eczema, reccurent opportunistic and pyogenic infections, and severe thrombocytopenia with small, deformed platelets. Death from infection, hemorrhage, or malignancy may occur in the first decade of life (Perry et al., 1980). The underlying intrinsic molecular defect of WAS has yet to be characterised. The direct involvement as the primary genetic defect of the T lymphocyte membrane glycoprotein sialophorin, shown to be structurally unstable in patients with WAS (Remold-O'Donnel and Rosen, 1990), has been precluded due to the localisation of the sialophorin gene to chromosome 16. Initially localised by linkage analysis to the pericentromeric region of the X chromosome (Peacocke and Siminovitch, 1987), the WAS disease was then assigned to the proximal short arm between DXS7 and DXS14 (Kwan et al., 1988). This localisation was subsequently refined further to the interval between TIMP and DXS255, at 1.2 cM distal to DXS255 (Kwan et al., 1991).

3.2 Incontinentia pigmenti 1(IP1)

A puzzling case is represented by this rare disorder which is characterised by developmental neuroectodermal abnormalities. It affects only females and is assumed to be an early lethal in affected male fetuses. Several females with different de novo X-autosome translocations have been described, with breakpoints in Xp11.1 or Xp.21. Some of these breakpoints appear to be separated by at least 1250kb (Gorski et al., 1991), suggesting that the corresponding gene, IP1, may be very large. Such a putative large gene should be a target for other types of mutations. However, a familial form of the disease has been localised to a differet region by genetic mapping, Xq28 (defining locus IP2 ; Sefiani et al., 1991). In addition, three Finnish families showed recombination with the loci in both Xq28 and Xp11 (XCW4, 1993), suggesting the possibility of yet a third IP locus.

3.3. Eye-diseases in proximal Xp.

Vision is a highly complex biochemical and physiological process requiring a large number of gene products. A wide range of phenotypes are observed in the population which are due to mutations in the genes involved in the development, function and maintenance of the eye. Recently there has been substantial progress in categorising the vast range of human retinal degeneration phenotypes. To date, the genes for approximately a dozen such diseases have been chromosomally assigned. This is the first step in the positional cloning approach for their subsequent identification and cloning. The information about the map position has led to the identification of genes responsible for two forms of adRP and choroideraemia. Several important hereditary eye disorders have been localised to Xp11.4-Xp11.2 by linkage analysis and deletion studies, including a locus for Norrie disease (NDP ; de la Chapelle et al., 1985), congenital stationary night blindness (CSNB1 ; Gal et al., 1989), Aland Island eye disease (Glass et al., 1993) and X-linked retinitis pigmentosa form 2 (RP2 ; Bhattacharya et al., 1984 ; Ott et al., 1990). Recent evidence also suggests that two genes in this chromosomal region may be involved in X-linked progressive cone dystrophy (Meire et al., 1994) and Leber hereditary optic neuroretinopathy (LHON ; Vilkki et al., 1991). The gene for Norrie disease which is characterised by retinal dysplasia and consequent congenital blindness, has recently been isolated by deletion cloning (Chen et al., 1992).

3.3.1 X-linked retinitis pigmentosa form 2

One of the most common causes of severe visual handicap in middle life is a group of retinal degenerations collectively known as retinitis pigmentosa. The most prominent symptoms are night blindness and a narrowing of the field of vision ("tunnel vision") due to the loss of photoreceptor cell function, usually leading to severe visual loss. The classic ophthalmoscopic features are bone corpuscle like

pigmentation starting in the equatorial region of the retina, narrowing of the retinal vessels, yellowish pallor of the optic disc and disappearance of the retinal pigment epithelium. Electroretinography shows very reduced or absent potentials, both photopically and scotopically. Retinitis pigmentosa is genetically heterogeneous with most cases being due to autosomal dominant, autosomal recessive or X-linked forms. The most common form of RP is probably the autosomal recessive type and in these cases the disease often starts in the first decade of life and is rapidly progressive. Dominantly inherited pigmentary retinopathy often starts later in life, and may show slow progression, so that good vision is maintained until quite late in life. X-linked recessive inheritance seems to be more severe, with onset in males in the first decade of life, progressing to severe visual handicap by the fourth decade (Jay, 1982 ; Boughman et al., 1980 ; Heckenlively, 1983). RP is thought to be primarily a disorder of the rod photoreceptor system, which is responsible for contrast sensitivity and vision at low luminance levels. Although rod function is most obviously affected, the primary abnormality could lie in any of the interacting cellular elements of the neural retina. Several genes coding for key components of the visual transduction have been cloned in recent years, but the genetic possibilities are so large that increasing emphasis has been put on positional cloning approaches to the identification of the responsible genes.

The X-linked form of retinitis pigmentosa (xlRP) is one of the most severe forms of this group of inherited retinopathies and affects 14-25% of families in the United Kingdom (Jay, 1982 ; Bundey and Crews, 1984). Affected males present with typical RP, whereas the carrier state has shown at least three retinal phenotypes, normal fundus, metallic sheen (tapetal-like reflex), or peripheral pigmentary deposits (Bird,1975 ; Fishman et al., 1986). Initially, close genetic linkage between xlRP and a restriction fragment length polymorphism identified by the probe L1.28 (DXS7, localised to band Xp11.3) was demonstrated in five British families (Bhattacharya et al., 1984). Further linkage data from this group (Clayton et al., 1986) and other

researchers (Mukai et al., 1985 ; Friedrich et al., 1985) supported the early findings. However, Francke et al. (1985) described a male patient (BB) with a deletion in Xp21 and a pathology including Duchenne mascular dystrophy, chronic granulomatous disease, McLeod syndrome and pigmentary retinopathy. This implied a locus for xlRP close to, or in, Xp21. Genetic linkage evidence for this second locus was presented by Denton et al., (1988) who found no crossovers between the gene for ornithine transcarbamylase (OTC) and xlRP in his families. Wright et al. (1987) presented more linkage data supporting localisation of xlRP between L1.28 and 58.1 (DXS14, localised to band Xp11.21). Therefore, it appears that, based on linkage data, there are at least two genes on the short arm of the X chromosome causing RP; one located proximal to L1.28 (RP2) and another in Xp21 (RP3). There may also be a third xlRP locus (RP6) lying distal to DMD between DXS28 and DXS164 (Xp21.3-p21.2 ; Musarella et al., 1990). As yet, however, there has only been one family clearly identified as harbouring RP6.

The proportion of the two genetic types, RP2 and RP3, appears to vary in different populations. A heterogeneity analysis of linkage data on 62 families from nine laboratories showed convincing evidence for two loci with odds of 6.4×10^9 in favour of this hypothesis (Ott et al., 1990). In this study, in 60-75% of families the xlRP gene was found to be located at 1 cM distal to OTC (RP3) and 5cM proximal to DXS84. This result was consistent with linkage data from large RP3 families which show a locus 0-6cM distal to OTC (Denton et al., 1988 ; Chen et al., 1989 ; Musarella et al., 1988, 1989, 1990). In a recent study, extended heterogeneity analysis of linkage data on 40 families with xlRP also placed the RP3 locus 0.4cM distal to OTC (Teague et al., 1994). Analysis of two patients with deletions that include the putative RP3 locus also supports a location just distal to OTC in Xp21.1 (Bertelson et al., 1988 ; de Saint-Basile et al., 1988 ; Wright, 1990).

Uncertainty regarding the precise location for the RP2 exists. There are differences between the locations reported by different groups. The locus has been placed 6cM proximal to DXS7 (Friedrich et al., 1985 ; Wright et al., 1987). Close linkage to

marker DXS255 has been reported (Meitinger et al., 1989 ; Bergen et al., 1994). Friedrich et al. (1992) reported a large Danish RP2 family in which the disease gene mapped distal to DXS255 (and proximal to OTC). Coleman et al. (1990) reported close linkage of xlRP to polymorphic marker DXS426 which maps between DXS7 and DXS255. More interestingly, a location of RP2 between DXS7 and DXS426 was suggested based on two informative meioses in two different xlRP families. In one (RP15), xlRP segregated with DXS426 but not with DXS7, indicating that the disease gene was proximal to DXS7. In another family (RP47), xlRP segregated with OTC and DXS7 but not with DXS426 and DXS255, indicating that the disease gene was distal to DXS426. A major part of the study undertaken in this thesis was based on this data. A second crossover distal to OTC in the same individual of family RP47, however, revealed that the disease in this family was RP3 and not RP2 as was initially considered. Based on this data and linkage data from other studies, the location of the gene for RP2 was broadened between DXS7/MAOA and DXS255. In 25-40% of the 62 xlRP families studied during the multilocus linkage analysis (Ott et al., 1990) the disease locus was found to be located halfway between DXS14 and DXZ1 (cen), a genetic distance of 117cM from the telomeric Xg blood group marker. This position differs from that found by the other investigators. However, the distal end of the support interval for this map location extends to the map location 104cM from the telomeric Xg marker, a distance which is halfway between DXS7 and OATL1. Friedrich et al. (1992) observed an apparent expansion of the genetic map between the centromere and proximal markers (DXS14, DXS146, DXS255) in their large Danish RP2-linked family compared with a previously reported RP2 kindred (Friedrich et al., 1985) and other published linkage data. This suggested the possibility of variability in recombination associated with the observed differences in the amount of centromeric heterochromatin segregating in these kindreds. According to the most recent study of heterogeneity analysis in the set of 40 kindreds with 20 polymorphic markers by Teague and co-workers (1994), the RP2 locus is located 6.5cM proximal to DXS7 in the DXS7/MAOA to

DXS426/TIMP interval. This location for RP2 is the one consistently favoured in previous analyses.

The proportion of the two genetic types, RP2 and RP3, appears to vary in different populations. The RP2 type of mutation appears to be more common in European populations than in North America. Musarella et al. (1990) for example, failed to identify any RP2 families in 20 North American families. European studies have found that 25-30% of families are of RP2 type and the remaining families are RP3 (Ott et al., 1990 ; Bergen et al., 1994 ; Teague et al., 1994). Chen et al. (1989) also found 22% of RP2 type families in an Australian sample of 9 xlRP families.

The phenotype of RP3 versus RP2 families is still unclear. The study of Kaplan et al. (1992) suggests that early onset of myopia is the presenting symptom in RP2-linked families while later onset of night blindness is the presenting symptom in RP3-linked families. In a large RP2 kindred reported by Wright et al. (1991) affected males had myopia, but night blindness was usually the presenting symptom, sometimes as early as 3 years of age. So far no reliable clinical means of distinguishing the two X-linked loci has been found.

3.3.2. Physical mapping at the potential X-linked RP3 locus

Cloned DNA fragments absent from the DNA of BB patient have been used to isolate the genes involved in CGD (Royer Pocora et al., 1986) and DMD. The relative order of these and other loci in this region, including RP3, as determined by molecular analysis and PFGE mapping is: Xpter-DMD-XK-CGD-RP3-OTC-Xcen (van Ommen., 1986 ; Bertelson et al., 1988). Because the RP3 locus appears to map at the centromeric end of the BB deletion, a DNA fragment spanning the BB deletion junction was cloned and then used for chromosome walking along a normal X chromosome (Musarella et al., 1991). A single-copy probe, positioned near the the centromeric junction but deleted in BB, was used along with a CGD cDNA probe to establish a refined long-range physical map. Both probes recognized a common Sfi I fragment of 205kb. Since the CGD gene covers 30-60kb, the RP3 locus has

been restricted to approximately 150-170kb. A CpG island, potentially marking a new gene, was identified within the Sfi I fragment at a position approximateley 35kb from the deletion endpoint in BB (Musarella et al., 1991). More recent studies include the isolation of two YAC clones containing the 205kb Sfi I fragment with the potential RP3 locus (Fujita et al., XCW, 1993, abs. 28). More than 80% of this fragment has been covered by cosmids and selected clones are being used to search for expressed sequences and polymorphic markers.

3.3.3 Norrie disease

Norrie disease (NDP) is a an X-linked recessive neurologic disorder characterised by retinal dysplasia and consequent congenital blindness (Warburg, 1966). Neuropathologic examinations have revealed poorly differentiated, malformed retinas (Warburg, 1963 ; 1966) suggesting a specific early developmental failure within sensory cells, ganglion cells, and the inner neuroblastic layer. About two-thirds of Norrie disease patients manifest mental retardation, often with dementia or psychotic features appearing in late infancy or early childhood. Approximately one-third of all Norrie disease patients develop progressive sensorineural hearing loss, which is usually not evident until the second decade of life (Warburg, 1979). Linkage studies using the DNA probe L1.28 (DXS7 ; Xp11.3-Xp11.4) had assigned the NDP gene to the proximal short arm of the X chromosome (Gal., 1985 ; Bleeker-Wagemakers et al., 1985 ; de la Chapelle et al., 1985 ; Kivlin et al., 1987). In addition, small submicroscopic deletions of the region had been detected in several NDP families by the failure of L1.28 to detect complementary DNA sequences in affected males (Gal et al., 1986 ; de la Chapelle et al., 1985, Donnai et al., 1988 ; Diergaarde et al., 1989 ; Zhu et al., 1989). These patients displayed complex and atypical phenotypes, including microcephaly, atonic seizures, myoclonus, delayed sexual development, somatic growth failure and increased susceptibility to infections. Subsequently, both monoamine oxidase A and B genes (MAOA and

MAOB) were shown to be deleted in these affected individuals resulting in a deficiency for monoamine oxidase enzyme activity (Sims et al., 1989a ; Lan et al., 1989 ; Collins et al., 1992). MAOA and MAOB have been excluded as NDP candidate genes by DNA and enzyme studies (Sims et al., 1989b). Additional linkage data had favoured the order Xpter-DXS7-NDP-DXS426-Xcen (Katayama et al., 1988 ; Ngo et al., 1988 ; Lindsay et al., 1992). A single recombination event between a microsatellite marker within MAOB and NDP had confined the disease locus to an interval of about 160kb (Sims et al., 1992). Finally, a candidate gene for Norrie's disease has been isolated and characterised by two groups using a positional cloning strategy (Chen et al., 1992 ; Berger et al., 1992). The approach undertaken by Chen et al. was to screen cDNA libraries from human fetal and adult retina using as a probe a yeast artificial chromosome subfragment of 160kb in which the disease locus had been confined. A similar approach undertaken by Berger et al. involved screening of cDNA libraries from adult human retina and fetal brain using a cosmid probe derived from a cosmid contig extending 250kb beyond the MAOB gene. Both groups identified an evolutionarily highly conserved cDNA which during northern analysis reveals a 1.9kb transcript in retina, brain and less abundantly in choroid. The cDNA detects genomic sequences which span a maximum of 50kb, and which are partly deleted in several patients with classic, i.e. non-syndromic Norrie disease. This cDNA encodes a predicted protein of 133 aminoacids which shows no regions of extended homology to sequences in protein databases. Further study on the number and spacing of 11 cysteine residues in the ND protein revealed homology with a C-terminal domain common to a group of proteins including mucins (a family of extracellular proteins) (Meindl et al., 1992). Three newly-characterised missense mutations, replacing evolutionarily conserved cysteines or creating new cysteine codons in non-deleted patients, emphasized the functional importance of these sites (Meindl et al., 1992). These findings and the clinical features of this disorder suggest a possible role for the Norrie gene in neuroectodermal cell-cell interaction.

3.3.4 X-linked congenital stationary night blindness (CSNB1)

Congenital stationary night blindness is a group of disorders that include impairment of night vision, variable myopia, reduced visual acuity, and congenital nystagmous (Pearce et al., 1990). There are two types of CSNB depending on whether the rod deficit is complete or incomplete, where the terms complete (CSNB1) or incomplete (CSNB2) are used respectively. Cone function is always affected in CSNB2, but is infrequently affected in CSNB1. On clinical grounds the two types of CSNB appear unlikely to be allelic since no single family manifest both forms. This is consistent with the hypothesis that the fundamental defect lies in the inner photoreceptor segment in CSNB1, while in CSNB2 it is in the bipolar cell layer of the retina. This assertion that CSNB1 and CSNB2 are aetiologically distinct has been disputed on electrophysiological criteria (Pearce and Bech-Hansen, 1990). Assignment of the gene responsible for CSNB2 is awaited and may facilitate clarification of possible allelism to CSNB1 which has been mapped to Xp11 (Musarella et al., 1989 ; Gal et al., 1989 ; Bech-Hansen et al., 1990) by linkage studies. Assignment of this locus to the interval DXS7/MAO and TIMP/DXS426 was suggested by Bech Hansen et al. (1991) and to the OTC-TIMP interval by Li et al. (1991). The subsequent observation of recombination between CSNB1 and DXS7 oriented CSNB1 proximal to DXS7 (Bech-Hansen et al., 1992). CSNB2 has been suggested to be clinically indistinguishable from Aland Island eye disease (Musarella et al, 1989 ; Weleber et al., 1989).

3.3.5 Aland Island eye disease (AIED)

This non-progressive condition is characterised by reduced visual acuity, reduced dark adaptation, colour vision abnormalities, infantile nystagmus, and high axial myopia. Initially, AIED was thought to be a variant of ocular albinism, on the basis of the fundal depigmentation in the original family described (Elenius et al., 1968). This has been subsequently disproved (Van Dorp et al., 1985). In AIED female carriers

are normal on clinical examination, and this is one feature which distinguishes this condition from X linked progressive cone dystrophy (Jacobson et al., 1989). Initially, linkage analysis of an AIED family enabled a localisation for the gene to be made in the pericentromeric region of the X chromosome, between DMD and PGK at Xq13 (Alitalo et al., 1991). This localisation was refined to the interval DXS7 (Xp11.3) -DXYS1 (Xq21.3) by the report of a second family with tight linkage to TIMP and DXS255 (Schwartz et al., 1991). Linkage analysis of a third family confirmed tight linkage to Xp11 loci (PFC and DXS228) and suggested a location between DMD and DXS255 (Glass et al., 1993). Assuming homogeneity of the three pedigrees, the combined information supported a localisation of the gene responsible for AIED between DXS7 and DXS255 (Glass et al., 1993). These findings were controversial with a previous putative assignment of the AIED gene to Xp21.3-21.2 made by Weleber et al. (1989). This was based on a male patient with a contigous gene deletion having a complex phenotype of Duchenne muscular dystrophy, glycerol kinase deficiency, congenital adrenal hypoplasia, AIED and mental retardation. The overall data suggested either that AIED is genetically heterogeneous or that AIED and the gene defect in the patient of Weleber et al. were distinct (Glass et al., 1993).

About 15 years ago, Krill (1977) suggested that AIED and CSNB2 might be a single entity. A recent study of a five generation family with an X linked cone and rod disease undertaken by Glass et al. (1992) supports this assertion and suggests that if the conditions are indeed allelic, CSNB2 is a less severe form of the disease than AIED. Since the genes responsible for AIED and CSNB1 have been assigned to the interval spanning DXS7-DXS426-DXS255, the question of allelism between AIED/CSNB2 and CSNB1 is also addressed and remains open.

3.3.6 X-linked cone dystrophy

Inheritance of progressive cone dystrophy is commonly autosomal dominant, although an X linked pattern of inheritance has been reported. X linked cone

dystrophy (COD_1) is characterised by the absence of nystagmus, progressive deterioration of visual acuity, myopic refraction, full peripheral fields with central scotomas, and colour vision impairment. The electroretinogram (ERG) reveals decreased cone mediated responses and normal rod mediated responses. However, in elderly affected males rod dysfunction is also observed. Evidence has been provided for a progressive cone dystrophy locus in Xp11 or Xp21-p11.1 by three groups (Bartley et al., 1989 ; Hong et al., 1991 ; Meire et al., 1994). This localisation clearly separates the COD_1 gene localisation from the gene localisation of X linked congenital cone dysfunctions, including blue cone monochromatism (Nathans et al., 1989) and the progressive cone dystrophy reported by Reichel (1989) which are localised to the distal part of the Xq arm.

3.3.7 Leber hereditary optic neuroretinopathy (LHON)

Leber hereditary optic neuroretinopathy (LHON) is a maternally inherited disease (Nikoskelainen et al., 1987) characterised by acute or subacute visual loss due to severe bilateral optic atrophy. The age at onset of visual impairment is usually between puberty and 30 years. The hypothesis of mitochodrial transmission of the disease was supported by the discovery of a mtDNA mutation associated with LHON (Wallace et al., 1988 ; Singh et al., 1989). This mutation was found in 11 of 20 Finnish families with LHON (Vilkki et al., 1989). The clinical expression of LHON varies greatly within and between different families. The variation in the clinical expression of the disease among family members has remained unexplained, but pedigree data have suggested an involvement of an X-chromosomal factor. These data had shown that approximately 50% of males and approximately 20% of females in LHON maternal lineages had optic atrophy (Seedorf, 1985 ; Nikoskelainen et al., 1987). This suggested that a mutation in mtDNA may lead to optic atrophy when matched with a specific allele of an X-linked susceptibility gene and to different phenotypes when combined with the other allele(s). Finally, linkage studies using 15 polymorphic markers on the X chromosome in six families with

LHON showed evidence that the liability to develop optic atrophy in these families is linked to the DXS7 locus at proximal Xp (Villki et al., 1991).

4 The visual pathway

4.1 General considerations

Vision is dependent on the interaction of light with the neural retina, embryologically an outgrowth of the brain, which transmits visual information to the brain via the optic nerve. The retina contains a dense array of light-sensitive photereceptors containing visual pigment molecules that initiate the neural response to light (visual transduction). There are two main morphological types of photoreceptor, rods and cones, which contain different photopigment molecules and vary in their anatomical distributions. The cornea and lens focus light onto the retina, where rods and cones subserve different but overlapping functions. Further processing of the photoreceptor signal occurs in the inner retinal layers and in specialised regions of the brain.

4.2 The human retina

The human eyeball or globe is a hollow spheroid about 25mm in diameter. It is composed of three layers and three internal zones (fig. 2). The three layers are an outer fibrous layer (the sclera and cornea), a middle vasculomascular layer (the uvea or uveal tract comprising the choroid, ciliary body and iris) and an inner neural layer (the retina). The three internal zones are those occupied by the aqueous, the lens and the vitreous.

The retina is a delicate, thin (0.4mm) layer which lines the interior of the posterior two-thirds of the globe. It consists of two main parts, the transparent neural portion internally and the pigment epithelium externally. It is common usage to refer to the neural part as "the retina". The pigment epithelium is a single layer of hexagonal cells which contain light-absorbing pigment granules. Delicate villi projecting from the internal surface of these cells surround the external segments of the apposed

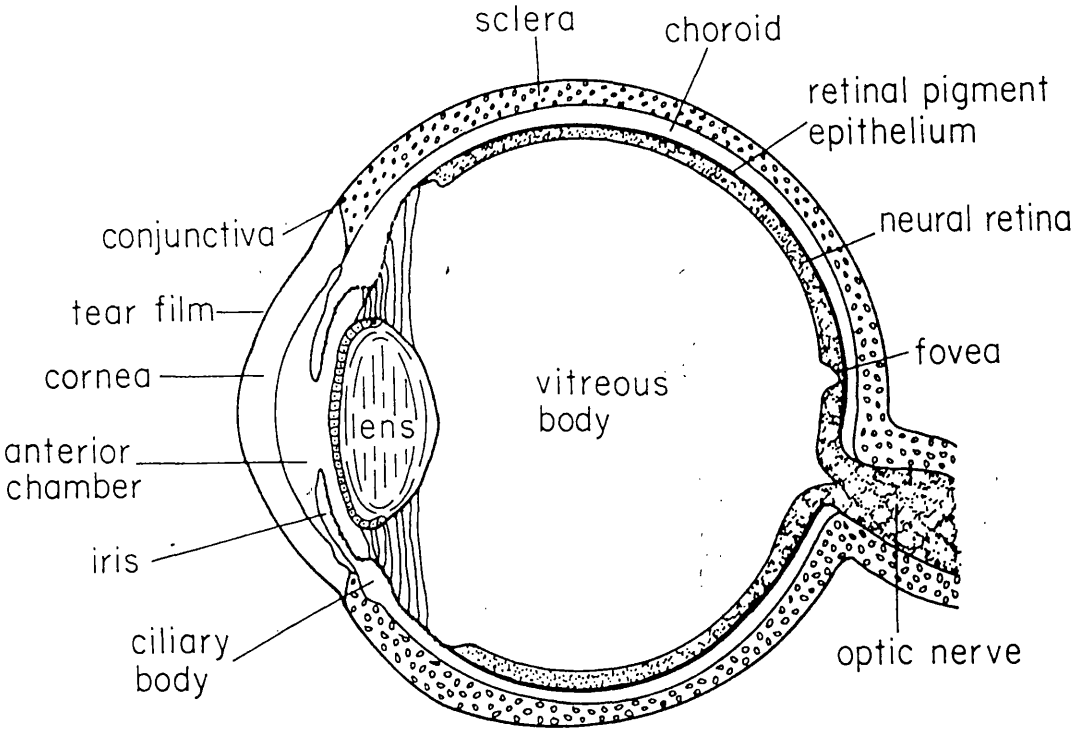


Figure 2: Transverse section of the human eye.

photoreceptors. The retinal pigment epithelial cells serve as a barrier between the choroidal circulation and the outer retinal layers, supplying nutrients, producing interphotoreceptor space glycosaminoglycans, participating in the retinol/retinal cycle required for visual transduction and phagocytosing rod and, to a lesser extent, cone outer segments.

The neural retina is attached to the outer layer of the eye at its anterior margin (the ora serrata) and at the optic disc (a pale, almost circular area, 1.5mm to 2.0 mm in diameter, 3mm nasal to the posterior pole of the globe, where the retinal nerve fibres leave and the retinal vessels enter and leave the interior of the eye). It is a most complex tissue. It contains photoreceptors (the rods and cones), five types of neurons (receptor nuclei, horizontal cells, bipolar cells, amacrine cells, and ganglion cells), a specialised type of glial cells (Muller's cells), and very many cell processes. In each eye there are about 120 million rods, 6 million cones and 1 million nerve fibres. One retinal neurone may have more than 200 connections to 20 or 30 other neurons. In histological sections it is obviously layered, there being one layer of rods and cones (their outer segments), three dark cellular (nuclear) layers, two structureless synaptic (plexiform) layers, and one nerve fibre layer (fig. 3). The outer nuclear layer contains the nuclei of rods and cones, the inner nuclear layer contains the cell bodies of bipolar, horizontal, amacrine, and Muller's cells, and the ganglion cell layer contains the ganglion cell bodies. In each plexiform layer the processes of three cell types synapse; receptor, bipolar, and horizontal cells in the outer plexiform layer, and bipolar, amacrine and ganglion cells in the inner plexiform layer. Bipolar cell dendrites are activated directly by receptors with which they synapse or indirectly through horizontal cells by more distant receptors. Ganglion cell dendrites are activated directly by bipolar cells or indirectly through amacrine cells. The axons of the ganglion cells form the innermost layer of the retina, the nerve fibre layer.

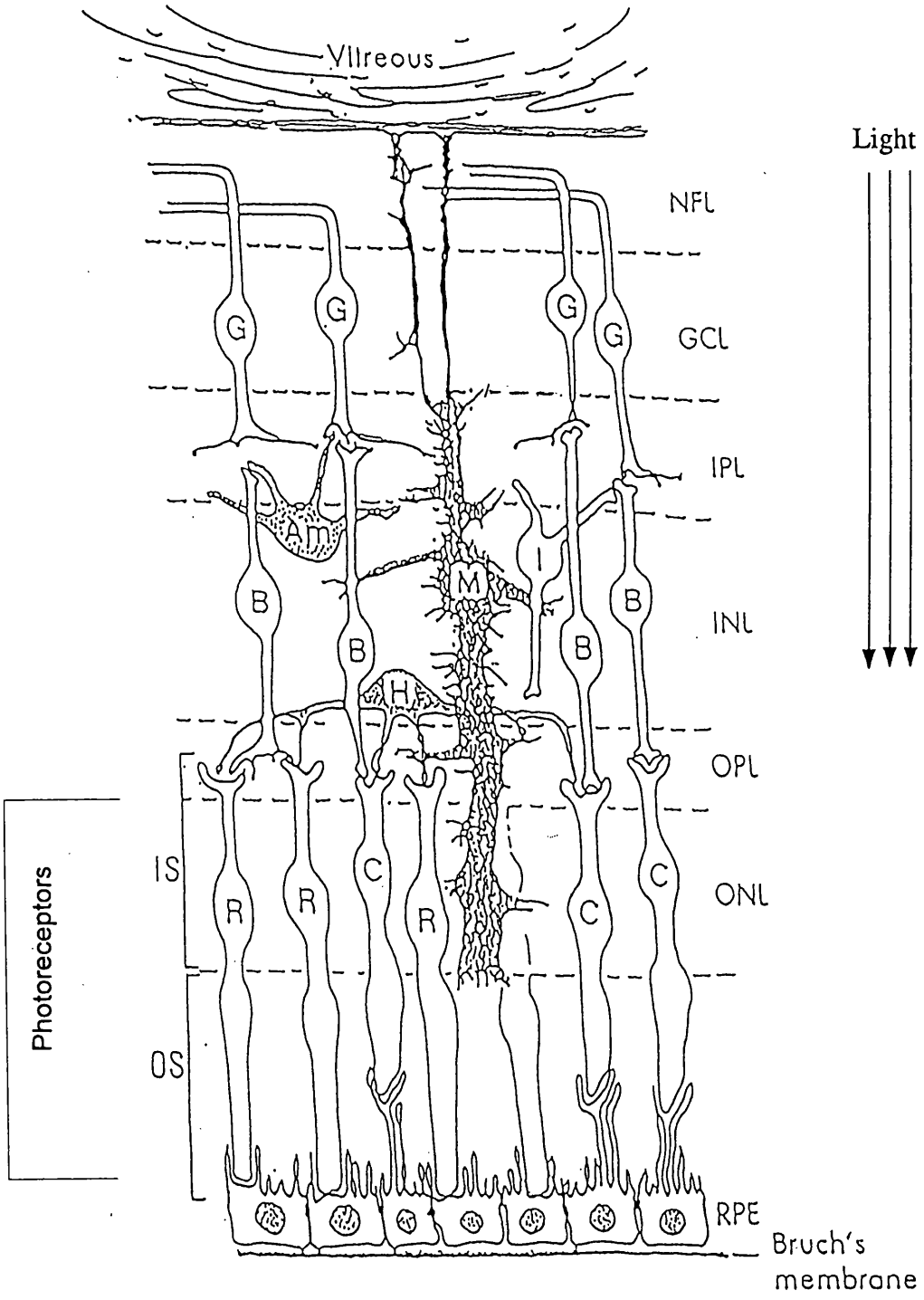


Figure 3: Structure of the human retina.

Schematic representation of a cross section of the retina (modified from Ryan et al., 1989). Symbols used: RPE, retinal pigment epithelium ; ONL, outer nuclear layer ; R, rods ; C, cones ; OS, outer segment ; IS, inner segment ; OPL, outer plexiform layer ; INL, inner nuclear layer ; B, bipolar cells ; H, horizontal cells ; Am, amacrine cell ; M, Muller cell ; IPL, inner plexiform layer ; GCL, ganglion cell layer ; NFL, nerve fibre layer.

4.3 Photoreceptors

The rod and cone photoreceptors (fig. 4) are highly polarised cells consisting of the rod- or cone-shaped, light-sensitive outer segment, a robustly metabolic inner segment which is connected to the outer segment by a connecting cilium, a cell body with its nucleus, and a highly variable region (sometimes referred to as the fiber) which connects the cell body with the synaptic terminal. The light-sensitive outer segment contains an elaborate system of stacked membranous discs whose main intrinsinc protein is photopigment. The discs develop as evaginations of the plasma membrane at the base of the outer segment. In rods, discs eventually become isolated from the plasma membrane, whereas cone discs remain continuous with one another and remain open to the extracellular space. The rod discs and, to a lesser extent, cone discs are continually being shed and phagocytosed by the retinal epithelial cells in which they are embedded. Rods contain the light-sensitive chromophore 11-cis-retinal, bound to the visual pigment apoprotein rhodopsin. They are responsible for contrast sensitivity and vision at low luminance levels. Rods are essential for "peripheral" vision, in contrast to cones, which are required for "central" vision. Cones are located at higher density packed in the macula, a small area of the central retina on which incident light is focused. Cones are responsible for vision at high luminance levels, colour sense and fine discrimination. There are three types of cones, each containing 11-cis-retinal bound to opsins absorption maxima in the red (555nm), green (530nm) or blue wave lengths (426nm).

4.4 Phototransduction

Retinal rod cells are exquisitely sensitive detectors. Psychophysical studies carried out half a century ago by Selig Hecht and co-workers revealed that a rod cell can be excited by a single photon. Biochemists and physiologists in the ensuing decades have been challenged to discover how a rod cell achieves this ultimate sensitivity. The answer is now known as a result of studies carried out by numerous investigators around the world. The flow of information in visual excitation is

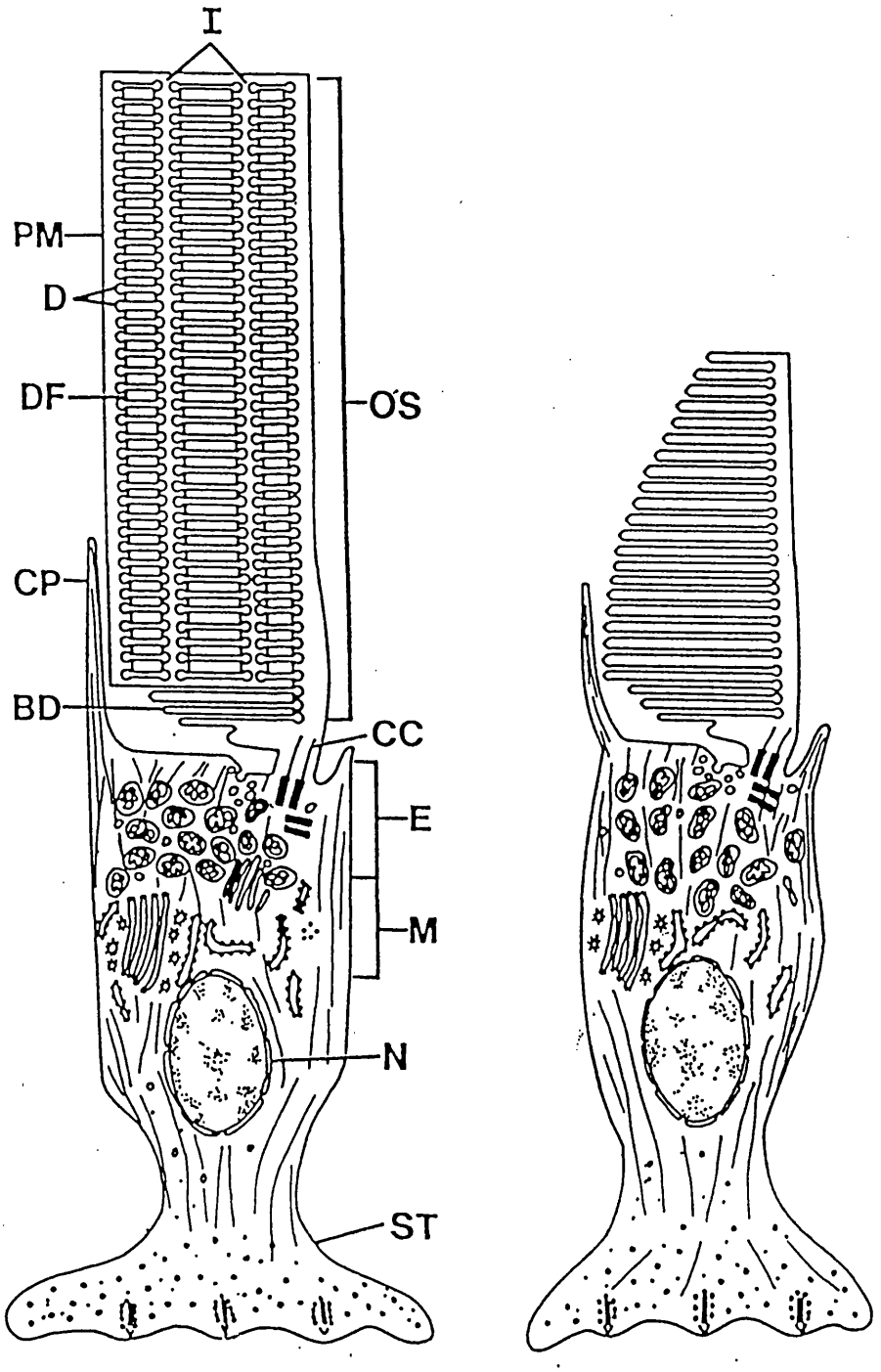


Figure 4: Diagram of rod (left) and cone (right) cone photoreceptors (Bok, 1985). Symbols used: OS, outer segment ; D, discs ; I, incisures ; PM, plasma membrane ; BD basal discs ; DF, small filaments ; CC, connecting cilium ; E, ellipsoid (which is rich in mitochondria ; M, myoid (which contains organelles for the synthesis and transport of macromolecules) ; CP, calycal processes ; N, nucleus ; ST, synaptic terminal (which protrudes into the outer plexiform layer of the neural retina).

summarised in fig. 5. In essence, light triggers a nerve signal by activating a cascade that closes cation-specific channels in the plasma membrane of the rod outer segment. Rhodopsin (R), a 40kDa protein with seven transmembrane helices (Dratz and Hargrave, 1983), contains an 11-cis retinal chromophore which is covalently linked to a lysine side chain. The photoisomerisation of the 11-cis-retinal chromophore to the all-trans form generates photoexcited rhodopsin (R*). The next step is the activation of transducin by R*. Transducin is a multisubunit peripheral membrane protein consisting of three chains: α, β and γ. The α subunit contains a binding site for GTP or GDP and a catalytic site for the hydrolysis of bound GTP. The β and γ chains form a $T_{\beta\gamma}$ subunit. T_a is associated with $T_{\beta\gamma}$ when GDP is bound (the inactive dark state), whereas they are separate when GTP is bound (the light-activated state). The GTPase activity of transducin is essential for bringing the system back to the dark state.

In the dark, nearly all of the transducin is in the inactive T-GDP state. Following illumination, T-GDP encounters R* in the plane of the disc membrane. R* induces the release of GDP from transducin and allows GTP to enter. R*-T-GTP then dissociates into T_a -GTP, $T_{\beta\gamma}$ and R*. T_a-GTP carries the excitation signal to a potent phosphodiesterase (PDE), and R* is free again to catalyse another round of GTP-GDP exchange. The activation of hundreds of molecules of transducin by a single R* is the primary stage of amplification in vision (Fung and Stryer, 1981). Activation is rapid; a molecule of R* triggers the formation of a molecule of T_a -GTP in 1ms (Vuong et al., 1984).

The phosphodiesterase consists of three kinds of polypeptide chains: α, β and γ (Baehr et al., 1979). The subunit structure of this peripheral membrane protein is $\alpha\beta\gamma_2$ (Fung et al., 1990). The α and β catalytic subunits of PDE are kept inhibited in the dark by the γ subunits, which bind very tightly (Hurley and Stryer, 1982). T_a -GTP activates PDE by carrying away its inhibitory subunits (Wensel and Stryer, 1990). The removal of γ subunits stimulates PDE more than 1500-fold. Activated PDE rapidly hydrolyses cGMP.

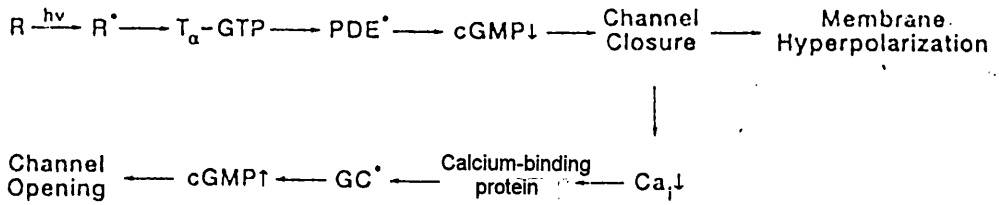


Figure 5: Flow of information in visual excitation and recovery (modified from Stryer, 1991).

Symbols used: R*, photoexcited rhodopsin ; T_α-GTP, the activated α subunit of transducin ; PDE*, activated phosphodiesterase ; GC*, activated guanylate cyclase.

In the dark, Na^+ and Ca^{2+} enter the rod outer segment through cation-specific channels, which are kept open by cGMP. The channel, an oligomer of 80 kDa subunits, has been functionally reconstituted into liposomes and planar bilayer membranes (Kaupp et al., 1989). Under physiological conditions, the channel opens and closes in times of milliseconds in response of changes in the cGMP level (Karpen et al., 1988). The light-induced decrease in the level of cGMP closes the channels, which hyperpolarizes the plasma membrane and generates the neural signal.

Channel reopening depends on the restoration of the cGMP level, which requires activation of guanylate cyclase and inhibition of the phosphodieasterase. Deactivation of PDE* requires hydrolysis of T_a -GTP to T_a -GDP. Thus, the GTPase activity of transducin determines the duration of the excitation phase. Deactivation of transducin and hence of PDE is necessary but not sufficient for the restoration of the dark state. Photoexcited rhodopsin too must be quenched. The isomerisation of all-trans-retinal to 11-cis form takes many minutes. A much faster shut-off is achieved by the phosphorylation of multiple serines and threonines in the carboxy-terminal region of R* by rhodopsin kinase, a 68 kDa cytosolic protein (Wilden et al., 1986 ; Miller et al., 1986). Arrestin, a 48 kDa cytosolic protein, then caps multiply phosphorylated R* to prevent it from interacting with transducin. Rhodopsin is regenerated many minutes later by insertion of 11-cis-retinal, release of arrestin of the COOH-terminal phosphates by protein Phosphatase 2A (Fowles et al., 1989). Restoration of the dark state also requires the stimulation of guanylate cyclase. The clue about the signal for activation of cGMP synthesis came from electrophysiological studies which revealed a negative feedback loop between cGMP and the cytosolic calcium level (Ca_i) (Yau and Nakatani, 1985). In the dark, the entry of Ca^{2+} through the cGMP-gated channel is matched by the influx of three Na+ and the efflux of one K+ (Cervetto et al., 1989). Following illumination, the influx of Ca^{2+} through the cGMP-gated channel ceases, but its export by the exchanger continues until Ca_i drops markedly, from about 500 to 50nM. A lowering

of Ca_i in this range stimulates guanylate cyclase about 5-fold in bovine rod outer segment (Koch and Stryer, 1988). Stimulation of guanylate cyclase at submicromolar Ca_i concentrations was shown to be mediated by a 23 kDa calcium-binding protein, named recoverin (Dizhoor et al., 1991). Recent observations from the same investigators, however, have suggested that recoverin is not the soluble activator of photoreceptor guanylate cyclase and its role, if any, remains to be determined (Hurley et al., 1993). Another Ca^{2+}-binding protein, calmodulin, which is present in the rod outer segment (Nagao, Yamazaki and Bitenski, 1987) has also been implied to play a role in photorecovery. Biochemical studies indicate that Ca^{2+}-calmodulin binds to a 220 kDa cGMP-gated channel associated protein (Molday, 1990). It is postulated that when cytosolic calcium levels are low calmodulin will dissociate from the channel complex, increasing the affinity of the 80 kDa subunit for cGMP (Hsu and Molday, 1993). This causes the channel to open, restoring Ca^{2+} levels to their dark level.

5. Retinal degenerations with defined mutations

5.1 Autosomal dominant retinitis pigmentosa

Close linkage of the disease locus in a large Irish family segregating adRP to a restriction fragment length polymorphism from chromosome 3q, named C17 (D3S47), was reported for the first time in 1989 (McWilliam et al., 1989). Since the gene encoding rhodopsin also mapped to 3q (Nathans et al., 1986) it became a prime candidate for the cause of autosomal dominant retinitis pigmentosa. A mutation within this gene, changing a highly conserved proline to a histidine residue at codon twenty three of exon 1, was subsequently identified in a pedigree segregating adRP (Dryja et al., 1990). A further 46 rhodopsin mutations have been reported since, accounting for 25-30% of adRP. The bulk of these are single amino acid substitutions, although a number of small deletions and at least two frame-shift mutations radically altering the sequence of the carboxy-terminus of the protein

have been described (reviewed in Humphries et al., 1993). As rhodopsin-RP was becoming well characterised attention, turned to the adRP families without rhodopsin mutation. So far, five additional loci have been identified by linkage studies in large pedigrees. Farrar and co-workers (1991a) found linkage in one of their pedigrees to the proximal short arm of chromosome 6, a region to which the gene causing the mouse retinopathy "retinal degeneration slow" had been mapped (rds ; Travis et al., 1991). Peripherin became a candidate gene in these families and, upon direct genomic sequencing, mutations were discovered which correlate with the disease (Farrar et al., 1991b ; Kajiwara et al., 1991 ; Farrar et al., 1992). Thus, for the first time, a direct correlation has been established between an animal disease model and an inherited retinopathy of man. Additional loci for adRP have also been reported on chromosomes 8 in the pericentromeric region (Blanton et al., 1991), 7p (Inglehearn et al., 1993), 7q (a region to which blue cone opsin maps ; Jordan et al., 1993) and 19qter (Al-Maghtheh et al., 1994). Thus, extensive genetic heterogeneity exists in the autosomal dominant form of RP. Since no known candidate genes reside in these regions of the genome, the cause of disease in these families remains unknown. It remains to be seen whether the six loci mentioned above account for all cases of adRP, meanwhile newly identified candidate genes in the 8q-cen, 7p, 7q or 19qter regions will be screened for mutations.

5.2 Autosomal recessive retinitis pigmentosa

The least is known about the most frequent form of RP, that is autosomal recessive RP (arRP). In 1992, however, a single arRP patient has been shown to be homozygous for a nonsense mutation at codon 249 within exon 4 of the rhodopsin gene (Rosenfeld et al., 1992) This *null* mutation should result in a functionally inactive rhodopsin which is missing the sixth and seventh transmembrane domains including the 11-cis-retinal attachment site. In 1993, four mutations in the human gene encoding the β-subunit of rod cGMP phosphodiesterase have been found in

patients with autosomal recessive RP (McLaughlin et al., 1993). The four mutations (two nonsense and two missense mutations) result in absence or alteration of the putative catalytic domain of the PDE β protein. The PDE-β gene, like the rhodopsin gene, codes for a protein involved in the phototransduction cascade and like the human RDS gene, its murine homologue is defective in a strain of mice, called *rd*, with retinal degeneration (Bowes et al., 1990 ; Pittler et al., 1991).These findings suggest that other members of this cascade may be defective in other forms of retinal degenerations.

5.3 Choroideraemia

Choroideraemia is a rare condition. Its clinical features are very similar to retinitis pigmentosa, but the appearance of the retina is different, as seen with the ophthalmoscope (Heckenlively, 1988). The vascular choroid that lies behind the outer retina and pigment epithelial cell layer degenerates earlier and to a greater extent than in retinitis pigmentosa.

The first step in the positional cloning of this gene was its localisation to band Xq21 by genetic linkage (Nussbaum et al., 1985). Then began a series of physical mapping experiments, initially with DNA from choroideraemia patients. Probes developed from Xq21 and mapped against the different deletions narrowed down the area of search and identified the closest markers (Nussbaum et al., 1987). One of these clones, defined by the locus DXS165, was deleted in several patients with classical choroideraemia (Cremers et al., 1990). Chromosome walking and jumping techniques generated new DNA markers near the locus and a 45kb genomic fragment which overlapped with most of the deletions was cloned (Cremers et al., 1990a). Using two conserved single-copy clones from this fragment, Cremers et al. (1990b) identified a candidate cDNA that was expressed in retina, choroid and retinal pigment epithelium, but also in nonretinal tissues, including lymphoblastoid cell lines. The cDNAs encompassed an open reading frame of 948 base pairs that was partially deleted or disrupted in 9 male choroideraemia patients with deletions,

and in a female patient with an X ; 13 translocation. The sequence of this cDNA predicted a protein of 316 aminoacids with partial homology to a protein involved in GTP metabolism, p25A-GDI (Fodor et al., 1991). This protein was identified and purified on the basis of its ability to inhibit the exchange of GTP for GDP bound to Rab3A, a low molecular weight guanine nucleotide-binding protein present in synaptic vesicles. Using similar mapping techniques, Merry et al. (1992) isolated a cDNA sequence which predicted a 1095 base pairs open reading frame extending in the 5' end of the sequence reported by Cremers et al. The predicted protein was thus augmented by 49 aminoacids. In addition, it was found by expression studies that patients with isolated choroideraemia had absent or reduced levels of message for this gene in the absence of any detectable structural rearrangements. Seabra et al. (1992) purified component A of Rab geranylgeranyl transferase (GG transferase), a single 95-kD polypeptide. The holoenzyme, which consists of components A and B, attaches 20-carbon isoprenoid geranylgeranyl to cysteine residues in Rab proteins, a family of GTP-binding proteins that regulate vesicular traffic. Six peptides from rat component A showed striking similarity to the products defective in choroideraemia. Seabra et al (1992) suggested that component A binds conserved sequences in Rab and that component B transfers geranylgeranyl group and a defect in this reaction may cause choroideraemia. Subsequently, he established this to be the case by demonstrating that lymphoblasts from choroideraemia subjects have a marked deficiency of component A, but not B, of Rab GG transferase (Seabra et al., 1993). The deficiency was more pronounced when the substrate was Rab3A, a synaptic vesicle protein, than it was when the substrate was Rab1A, a protein of the endoplasmic reticulum. Their studies suggested the existence of multiple component A proteins, one of which is missing in choroideraemia. The multiplicity and functional redundancy of component A genes creates a situation in which defects in one of them might cause a degenerative disease of the organ in which that particular form of component A is most essential. Finally, the investigators suggested that it will be of interest to

isolate the genes encoding the other component A proteins and to determine whether mutations in these genes underlie degenerative diseases of the nervous system or other organs.

5. Aim of the study

The aim of this study was initially to further develop the techniques of Fluorescence In Situ Hybridisation (FISH) for the localisation of markers on the X chromosome. Subsequently the development of two colour FISH was undertaken so that a finer localisation of markers within a defined region could be achieved. As described in the introduction, a number of inherited eye diseases have been localised in the proximal short arm of the X chromosome, including retinitis pigmentosa form 2 (RP2). Development of new markers should allow the construction of a high resolution genetic map of the region including the fine mapping of RP2. The ultimate aim was to initiate the construction of a YAC-based contig using a marker closely linked to the RP2 locus as the starting point. This should contribute to the development of a high resolution physical map of the proximal short arm of the X chromosome. A high resolution physical and genetic linkage map of the region should aid in the finer localisation of the xlRP2 gene, and the identification of candidate genes for this and other inherited diseases localised to this part of the chromosome.

CHAPTER 2: MATERIALS AND METHODS.

1. Fluorescence in situ hybridisation (FISH).

1.1 Preparation of metaphase spreads.

Chromosome metaphase spreads were prepared from short term cultures (48 hours) of phytohemagglutinin (PHA) stimulated whole blood in RPMI 1640 culture medium, supplemented with 15% Fetal Calf Serum. Addition of thymidine after 48 hours blocked cell division at S phase of cell cycle. Removal of the block after 19 hours allowed cell division to progress in a synchronised fashion. Colcemid was added 4 hours later, for 15 minutes, prior to harvesting, to inhibit spindle formation thus arresting the cells at metaphase. The cells were then treated with a prewarmed solution of 0.075M KCl and fixed with a 3:1 methanol:acetic acid solution.
Cell suspensions for metaphase spreads used in this project were donated by the Cytogenetics laboratory of Department of Human Genetics / University of Newcastle upon Tyne and derived from cells with a 46,XX normal female karyotype.
To prepare chromosome spreads from these suspensions, samples were centrifuged for 10 minutes at 1500 rpm, the pellet was resuspended in 0.5-1ml of freshly prepared fixative and 2-3 drops were placed onto each microscope slide. Slides were allowed to dry and then checked under a phase contrast microscope to be sure the chromosomes were properly spread and the mitotic index was adequate. Slides used for fluorescence in situ hybridisation were never more than a day old.

1.2 DNA probe labelling.

For each hybridisation, the probe DNA was labelled with either biotin or digoxigenin by the method of nick translation (Rigby et al, 1977) in a 20μl reaction volume as follows: 1μg DNA was combined with 2μl 10xsalts (0.5M Tris-HCl pH 7.8, 0.05M $MgCl_2$, 0.01M β-mercaptoethanol, 50μg/ml BSA), 2.5μl dNTPs mixture (dATP,

dCTP, dGTP, 10mM each, in 1:1:1 proportion ; Bethesda Research Laboratories : BRL), 2.5μl (2.5nmol) of either biotin-16-dUTP (Boehringer Mannheim) or digoxigenin-11-dUTP (Boehringer Mannheim) and 6.5μl H_2O, 5μl of DNAase(40pg/μl)/ DNA Polymerase I (0.4 units/μl) mixture (BRL) were added and the reaction was incubated at 15^{0}C for 90 minutes.

1.3 Column purification and precipitation of the labelled probe.

To the nick-translated probe 5μl stop buffer (300mM Na_2EDTA, pH 8.0), 1μl 5% SLS (sodium lauryl sulphate) and 25μl TNE (0.2M NaCl ; 10mM Tris-HCl, pH 8.0 ; 1mM EDTA) were added and the probe was loaded onto a Sephadex G-50 column (Pharmacia) equilibrated in TNE buffer (prewashed three times, with 400μl TNE buffer each time). 400μl TNE buffer were passed through the column and this first effluent was discarded. A further 400μl TNE buffer were added and to this second effluent containing the purified probe, 100μg sonicated salmon sperm DNA (Sigma) and 1/10 volume of sodium acetate were added and the DNA was precipitated with 2.5 volumes absolute ethanol at -20^{0}C for 1 hour . After 5 minutes centrifugation at room temperature (RT), the supernatant was removed and the pellet air dried and resuspended in 25μl TE buffer (10mM Tris- HCl pH 7.5 ; 1mM EDTA), for all cosmid probes or in 50μl TE for centromeric repeat probe.

The above procedure was slightly modified in the course of the in situ experiments by replacing the 5μl of the above stop buffer with 2μl of 0.5M EDTA, using TE buffer instead of TNE to prewash the column and 50μg of salmon sperm DNA in the second wash, reducing the time of DNA precipitation at -20^{0}C to 30 minutes followed by centrifugation of the ethanol precipitate for 30 minutes at 4^{0}C.

1.4 Biotin and digoxigenin incorporation test.

Biotin incorporation was assessed by immunological detection via a streptavidin-alkaline complex which, with NBT and BCIP as substrates, results in the production of a blue colour precipitate. Similarly anti-digoxigenin-alkaline phosphatase was

used to monitor digoxigenin incorporation. The buffers A, B, C, D and NBT (nitro-blue tetrazolium) and BCIP (5-bromo-4 chloro-3-indolyl phosphate) solutions required for the method are described in Table 2. The method was carried out as follows: **1**. Six dilutions (a-f) of the DNA probe and of the appropriate controls were prepared starting with a 1:1 dilution of the probe (1μl DNA and 1μl 20xSSC ; 20ng DNA/μl) followed by further five sequential 1 in 10 dilutions (using 10xSSC) until a final probe concentration of 0.2pg/μl was obtained (dilution f). 1μl of each of the above 6 probe dilutions (a-f ; 20ng, 2ng, 0.2ng, 20pg, 2pg and 0.2pg, respectively) was loaded on a gridded nitrocellulose filter (Schleicher and Schnell BA 85, Cellulosenitrate ; E ; 0.45μm, 88x88mm) and after drying the DNA, the filter was placed between two 3mm Whatman papers and baked at 80^0C for two hours. The number of dilutions in the course of the in situ experiments were reduced from six to three (a, b and c). **2**. The filter was then soaked in 20ml of buffer A, in a square petri dish (Sterilin, 105x105mm) for 5 minutes followed by soaking in 20ml of buffer B for 30 minutes. The filter was again incubated in 20ml of buffer B for 30 minutes containing streptavidin-alkaline phosphatase complex (diluted 1:1000 ; Boehringer Mannheim) for detection of biotin or anti-digoxigenin-alkaline phosphatase (diluted 1:5000 ; Boehringer Mannheim) for detection of digoxigenin. **3**. The filter was then washed 3 times in buffer A for biotin detection (5 minutes each) and twice (15 minutes each) for digoxigenin detection, equilibrated in buffer C (5 and 2 minutes for biotin and digoxigenin, respectively) and incubated in 20ml of buffer C containing 88μl of 75mg/ml NBT (in 70% dimethylformamide ; Boehringer Mannheim) and 68μl of 50mg/ml BCIP (in dimethylformamide ; Boehringer Mannheim) in the dark for at least 3 hours. The reaction was stopped by washing the filter in buffer D (stop buffer) for 5 minutes. The filter was wrapped in clingfilm, photocopied if desired and stored in refrigerator in the dark.

Table 2: Buffers required for biotin and digoxigenin detection.

	Biotin detection	Digoxigenin detection
Buffer A	0.1 M Tris-HCl pH 7.5 ; 1xSSC, 100ml	0.1M Tris-HCl pH 7.5 ; 0.15M NaCl, 100ml
Buffer B	0.1M Tris-HCl pH 7.5 ; 1xSSC ; 3% BSA fraction V, 50ml	0.25g of Boehringer Mannheim Blocking Reagent in 50ml of buffer A, dissolved at 50-70^{0}C for 1 hour.
Buffer C	0.1M Tris-HCl pH 9.5 ; 1xSSC ; 0.05M MgCl2, 50ml	0.1 Tris-HCl pH 9.5 ; 0.1 NaCl ; 0.05M MgCl2, 50ml
Buffer D (stop buffer)	0.01 Tris-HCl pH 7.5 ; 1mM EDTA pH 8.0, 20ml	Same as for biotin

1.5 In situ hybridisation.

1.5.1 *Metaphase spread slide treatment.*

Slides were dipped in 100μg/ml RNAase (Sigma) in 2xSSC at 37^{0}C for 1 hour, dehydrated by passage through an ethanol series (70%, 90%, 100%) allowing 5 minutes immersion in each. They were then air dried and denatured in prewarmed 2xSSC, 70% formamide (Analar) solution at 74^{0}C for two minutes and immediately passed through a second ethanol series and air dried. The slides were now ready for probe application.

1.5.2 *Hybridisation cocktail for repetitive probe DXZ1.*

X-specific centromeric probe DXZ1 was the first probe to be selected for the FISH technique and is, therefore, extensively described in the materials and methods section of the single in situ results chapter (sections 2.1 and 2.2 of chapter 3)

1.5.3 *Hybridisation cocktail for cosmid probes.*

Approximately 600ng of labelled probe (15μl of the 25μl labelled probe) were combined with 600μg of Total Human DNA (1:1000 competition) in the presence of 5xSSC in a 100μl reaction volume, denatured at 100^0C for 10 minutes, placed in ice for 1 minute and preannealed at 60^0C for 10 minutes. To the pre-annealed probe DNA 100μl of 20% dextran sulphate in 100% formamide and 10μl of 10mg/ml carrier E. coli t-RNA were added to make a 210μl hybridisation cocktail containing 3ng/μl probe DNA, 2.5xSSC, 500μg/ml E. coli t-RNA, 10% dextran sulphate and 50% formamide. 40μl of this hybridisation cocktail were applied on clean coverslip and a metaphase spread slide (treated as descibed in 1.5.1) was inverted on top of the coverslip ensuring that no air bubbles were trapped and sealed with rubber solution. When the seal was completely dry, the slide was transferred to a high humidity chamber, saturated with 20xSSC and incubated at 37^0C for 18 hours. If centromeric probe DXZ1 was co-hybridised, its hybridisation cocktail was prepared separately as described in section 2.1.1 of chapter 3 and 20μl of DXZ1 (20ng of DNA) were combined with 20μl (60ng of DNA) of the cosmid hybridisation mixture and applied on the slide.

The preparation of the above hybridisation cocktail was simplified as described below in the course of the in situ experiments although the final concentration of its components, with the exception of the probe concentration (2 instead of 3ng/μl), was not altered: 80ng (2μl of the 25μl labelled probe suspension) was combined with 80μg of Total Human DNA, 20xSSC, 20μl 20% dextran sulphate in 100% formamide, 2μl 10mg/ml E. coli t-RNA and 4μl H2O in a 40μl hybridisation cocktail,

denatured at 75^{0}C for 10 minutes and incubated (pre-annealing) at 37^{0}C for 45 minutes (modified from Trask et al, 1991). The entire volume was applied on the slide. If two or three probes were co-hybridised the above hybridisation cocktail was adjusted accordingly as follows: if two or three cosmids were co-hybridised, 40ng of each cosmid DNA was used instead of 80ng (total amount of DNA in the hybridisation cocktail 80 and 120ng, respectively) and if centromeric probe DXZ1 was co-hybridised with one or two cosmids, 20ng of DXZ1 DNA was also added to the mixture without the need for separate hybridisation mixture for it as previously described.

1.6 Post hybridisation washes.

Slides hybridised with biotinylated probes were washed for 3 minutes each four times, first in prewarmed 50% formamide, 2xSSC at 45^{0}C followed by only 2xSSC washes at 45^{0}C and finally immersed briefly in 4xSSC, 0.05% Triton X-100 at RT. Post hybridisation washes for digoxigenated alone or biotinylated and digoxigenated probes in combination were similar as above except that the time for washes was four instead of three minutes each and Triton X-100 was replaced by Tween 20.

1.7 Signal detection.

1.7.1 ***FITC signal detection from biotinylated probes.***

The signal detection procedure for single hybridisations was carried out as follows:

1. Slides were incubated for five minutes at RT with 40μl per slide of blocking buffer consisting of 4xSSC, 0.05% Triton X-100 and 5% low fat milk powder under coverslips.
2. This was followed by a 30 minutes incubation at 37^{0}C in 40 μl (per slide) of the first antibody, FITC (fluorescein isothiocyanate)-conjugated avidin DCS (Vector Laboratories), at final concentration 4μg/ml in blocking buffer.

3. Slides were washed three times, for two minutes each in 4xSSC, 0.05% Triton X-100 solution at 45^{0}C.

4. 40μl (per slide) of the second antibody, biotinylated goat anti-avidin (Vector Laboratories) at final concentration 5μg/ml in blocking buffer were applied and incubated at 37^{0}C for 30 minutes.

5. Slides were washed as described in step 3.

[Note: The whole procedure was repeated twice finishing with a final incubation in FITC-avidin DCS].

6. After the final wash, the slides were mounted in anti-fading glycerol (Citifluor) solution containing, initially, 3μg/ml propidium iodide (Sigma), progressively decreased to 0.5μg/ml (40μl per slide). Slides were dry blotted, sealed with rubber solution and observed under the fluorescence microscope.

1.7.2 *FITC signal detection from digoxigenated probes.*

After attempting different combination of antibodies for digoxigenated probes in single hybridisations (described in materials and methods of single in situ hybridisation results chapter) the following procedure was carried out:

1. Slides were incubated with 40μl per slide of blocking buffer consisting of 5% low fat milk powder in 4xSSC at RT for 20 minutes after which they were washed twice, for five minutes each in 4xSSC, 0.05% Tween 20 and once, for five minutes in a solution containing 0.1M Tris-HCl, 0.15M NaCl and 0.05% Tween 20 (TNT) at RT.

2. Slides were incubated with the first antibody, mouse anti-digoxigenin (Boehringer Mannheim) at final concentration 0.2μg/ml in TNB solution (0.5% Boehringer Blocking Reagent in 0.1M Tris-HCl, 0.15M NaCl) at 37^{0}C for 30 minutes.

3. Slides were washed three times for five minutes in TNT solution at RT.

4. After excess fluid was drained, slides were incubated with the second antibody, rabbit, anti-mouse-FITC (diluted 1:500 in TNB solution ; Sigma) at 37^{0}C for 30 minutes and washed as described in step 3.

5. Slides were finally incubated with the third antibody, goat, anti-rabbit-FITC (1:1000 or 1:1500 dilution in TNB solution ; Sigma) at 37^{0}C for 30 minutes and washed as described in step 3.

6 After the final wash, slides were counterstained with propidium iodide as described in step 6 of section 1.7.1.

The above detection method consists of two layers of detectable antibodies and it was applied for both centromeric repeat and cosmid probes.

1.7.3 *Simultaneous AMCA and FITC signal detection from biotinylated and digoxigenated probes.*

Slides were incubated with combined solutions of AMCA-conjugates (for biotin) and FITC-conjugates (for digoxigenin). The first incubation with blocking buffer and washes between incubations were as described in section 1.7.2. The sequential incubations were as shown in table 3.

Table 3: Antibodies used in each step of two colour FISH detection.

Incubation stage		Stock antibody concentration	Antibody solution and final concentration (where applicable)	Incubation temperature	Incubation time
1		0.5mg/ml	AMCA-avidin D, 30μg/ml	RT	20 minutes
2	a	0.1mg/ml	Mouse anti-digoxigenin, 0.2μg/ml	37^{0}C	30 minutes
	b	0.5mg/ml	Biotinylated goat anti-avidin D, 5μg/ml		
3	a		Rabbit anti-mouse FITC	37^{0}C	30 minutes
	b	0.5mg/ml	AMCA-avidin D, 30μg/ml		
4	a		Goat anti-rabbit FITC	37^{0}C	30 minutes
	b	0.5mg/ml	Biotinylated goat anti-avidin D, 5μg/ml		
5		0.5mg/ml	AMCA-avidin D, 30μg/ml	37^{0}C	30 minutes

1.8 Signal visualisation.

Signals on metaphase chromosomes were viewed using a Leitz Diaplan microscope. Photographs were taken using a Wild MPS 46/52 camera (Leitz) and 400 ASA Kodak Gold film. FITC was visualised using filter I3 and AMCA using filter A (Leitz). In double hybridisations filter block interchange and double exposure of the same metaphase on one photographic frame permitted simultaneous visualisation of both FITC and AMCA signals.

1.9 Preparation of solutions used in FISH.

All solutions were made to volume with Sterile Distilled Water (SDW) unless otherwise stated.

Biotin incorporation test.

Buffer A : 0.1M Tris-HCl pH 7.5 ; 1xSSC, 100ml.

1M Tris-HCl pH 7.5	10ml
20xSSC	5ml

Buffer B: 0.1M Tris-HCl pH 7.5 ; 1xSSC ; 3% BSA (bovine serum albumin) fraction V, 50ml

1M Tris-HCl pH 7.5	5ml
20xSSC	2.5ml
BSA	1.5g

Buffer C: 0.1M Tris-HCl pH 7.5 ; 1xSSC ; 50mM $MgCl_2$, 50ml

1M Tris-HCl pH 7.5	5ml
20xSSC	2.5ml
1M $MgCl_2$	2.5ml

Stop Buffer: 0.01M Tris-HCl pH 7.5 ; 1mM EDTA pH 8.0, 20ml

1M Tris-HCl pH 7.5 200µl

0.5M EDTA 40µl

Digoxigenin incorporation test.

Buffer A: 0.1M Tris-HCl pH 7.5 ; 0.15M NaCl, 100ml

1M Tris-HCl pH 7.5 10ml

5M NaCl 3ml

Buffer B: 0.25g of Boehringer Mannheim Blocking Reagent in 50ml of buffer A, prepared in advance and dissolved at 50-70^{0}C.

Buffer C: 0.1M Tris-HCl pH 9.5 ; 0.1 NaCl ; 0.05M $MgCl_2$, 50ml

1M Tris-HCl pH 9.5 5ml

5M NaCl 1ml

1M $MgCl_2$ 2.5ml

Stop Buffer as for biotin.

Metaphase spreads slide treatment.

RNAase: 100µg/ml in 2xSSC, 100ml

10mg/ml RNAase 1ml

20xSSC 10ml

Chromosome denaturing solution: 2xSSC ; 70% formamide, 500ml

20xSSC	50ml
100% formamide	350ml

Post hybridisation washes for biotinylated probes.

Solution 1: 50% formamide ; 2xSSC, 1 liter

100% formamide	500ml
20xSSC	100ml

Solution 2: 2xSSC, 1 liter

20xSSC	100ml

Solution 3: 4xSSC ; 0.05% Triton X-100, 2 liters

20xSSC	400ml
Triton X-100	1ml

All three solutions were prewarmed at 45^{0}C. Triton X-100 dissolved after solution 3 was warmed. The same solutions were used for digoxigenated probes except for solution 3 in which Triton X-100 was replaced with Tween 20.

Detection solutions for biotinylated probes.

1. Blocking buffer: 4xSSC ; 0.05% Triton X-100 ; 5% low fat milk powder, 10ml

0.5g low fat milk powder were dissolved in 10ml of washing solution 3 (see above).

2. Antibody solutions.

a. Primary antibody: FITC-avidin DCS, 4μg/ml

2μl of the 2mg/ml stock antibody solution in 1ml blocking buffer.

b. Secondary antibody: Biotinylated goat anti-avidin, 5μg/ml

1μl of the 0.5mg/ml stock antibody solution in 100μl blocking buffer.

3. Propidium iodide (Sigma) in glycerol,Citifluor:

Final concentration: 0.5μg/ml

Stock solution: 3mg/ml in SDW.

Detection solutions for digoxigenated probes.

1. Blocking buffer: 4xSSC ; 5% low fat milk powder, 10ml

0.5g low fat milk powder was dissolved in 10ml of 4xSSC solution.

2. TN (Tris-NaCl): 0.2M Tris-HCl ; 0.15M NaCl, 2.5 liters

2M Tris-HCl	125ml
5M NaCl	75ml

3. TNB

0.125g of Boehringer Mannheim Blocking Reagent in 25ml of TN solution.

4. TNT

1250μl of Tween 20 added in 2.5 liters of TN solution.

5. Antibody solutions:

a. Mouse anti-digoxigenin, 0.2μg/ml

1μl of 0.1mg/ml stock concentration in 500μl of TNB solution.

b. Rabbit, anti-mouse IgG-FITC, 1:500 dilution.

1μl of stock solution in 500μl of TNB solution.

c. Goat, anti-rabbit IgG-FITC, 1:1000 or 1:1500 dilution.

1μl of stock solution in 500μl of TNB solution

Simultaneous detection of biotinylated and digoxigenated probes.

Solutions 1, 2, 3 and 4 as for detection of digoxigenated probes.

5. Antibody solutions were as shown in table 3.

.

2. Clone DNA preparation.

2.1 Alkaline lysis DNA miniprep of plasmids or cosmids (Birnboim and Doly, 1979)
The method described below is the most commonly used. It is appropriate for preparing of DNA from 5ml cultures of plasmid or cosmid containing bacteria. Colonies from a bacterial stab or a glycerol stock were streaked out on agar plates supplemented with the appropriate antibiotic and incubated overnight at 37^{0}C. A single colony was picked from the plate and used to inoculate 5ml of sterile LB broth (1% tryptone, 0.5% yeast extract and1% NaCl per liter of H_2O ; pH 7) containing antibiotic. Following incubation at 37^{0}C overnight with agitation, 1.5ml of this culture was transferred to an Eppendorf tube and spun in a microfuge for 2 minutes at 13000rpm*1. The supernatant was removed and the pellet resuspended in 100µl of 50mM glucose, 10mM EDTA and 25mM Tris-HCl solution (Solution I). 400µl of 0.2M NaOH and 1% SDS solution (Solution II ; freshly prepared) was added and mixed with the resuspended pellet (without vortexing) and then placed on ice for 5 minutes. 300µl of 3M potassium acetate and 2M acetic acid solution (Solution III) was added and mixed thoroughly. The sample was placed on ice for at least 5 minutes followed by centrifugation for 2 minutes. The supernatant was removed (approximately 750µl) and the DNA precipitated with 0.6 volumes of isopropanol (approximately 450µl). The pelleted DNA was washed with 70% ethanol, air dried and resuspended in 20µl TE (10mM Tris-HCl and 1mM EDTA ; pH 7.5-8). 2µl of the preparation were electrophoresed along with size markers on a 0.8% agarose gel to check the quality of DNA preparation.

*1 All subsequent centrifugations of Eppendorf tubes described in this section were carried out in a microfuge at 13000rpm.

2.2 Alkaline lysis DNA midiprep of plasmids or cosmids.

The method below describes the isolation and purification of larger amounts of DNA.

50ml of LB broth supplemented with the appropriate antibiotic was inoculated with a single colony obtained as described above and incubated in a sterile 200ml flask at 37^{0}C overnight with agitation. After preparing glycerol stocks from the culture (appropriate volume of culture mixed with 0.25 volumes of 50% glycerol under sterile conditions and stored at -20^{0}C or -70^{0}C), the culture was transferred into 50ml conical tubes (Greiner) and centrifuged at 3000g for 5minutes. The pelleted cells were resuspended in 4ml of Solution I (see section above) and left at RT for 5 minutes. 8ml of Solution II (see section above) was added and mixed with the pelleted cells which were then placed on ice for 5 minutes. 4ml of Solution III (see section above) was then added and thoroughly mixed. The sample was placed on ice for 15 minutes. After the addition of 1ml of H_2O, the sample was centrifuged at 3000g for 10 minutes. The supernatant was strained into fresh 50ml Falcon tubes and the DNA precipitated with 0.6 volumes isopropanol at 3000g for 5 minutes. The pelleted DNA was washed with 70% ethanol, resuspended in 0.4ml TE buffer and transfered to an Eppendorf tube. 15μl of 3M sodium acetate and 20μl of 10mg/ml RNAase were added to the DNA suspension which was then incubated at 37^{0}C for 30 minutes. The DNA was then extracted with one volume of phenol, followed by one volume of phenol/chloroform (1:1) and one volume of chloroform. 35μl of 3M sodium acetate was added to the final aqueous phase, and the DNA precipitated with 2.5 volumes of absolute ethanol. The pelleted DNA was washed with 70% ethanol, vacuum dried for 10 minutes and resuspended in 200μl TE buffer. 1μl of this DNA suspension was electrophoresed along with size markers on a 0.8% agarose gel.

2.3 Digestion of DNA with Restriction Endonucleases.

In principle, 1unit of restriction enzyme digests 1μg of λDNA in 60 minutes. Crude DNA preparations often require more enzyme and/or more time for complete digestion. The volume of restriction enzyme added should be less than 1/10 the volume of the reaction mixture, as glycerol in the storage buffer may interfere with the reaction.

In a typical 20μl reaction, xμl DNA (0.1-4μg in H_2O or TE buffer) was combined with 2μl of 10x appropriate restriction buffer and 1 to 5u/μg DNA of selected restriction endonuclease. Spermidine and BSA to final concentrations 1mM and 0.1mg/ml, respectively, were occasionally added. The reaction was incubated for 1 hour at the recommended temperature (usually 37^0C). 2μl of 10x loading buffer (25% Ficoll Orange ; 1/10 volume of the reaction) was added to the reaction mixture and the sample was analysed by gel electrophoresis.

2.4 Conventional agarose gel electrophoresis.

Gel electrophoresis is generally the fractionation method of choice for nucleic acids.The protocol described here is used to separate DNA fragments between 0.5 and 25 kb.

The gel was prepared, using electrophoresis buffer (1x TAE) and electrophoresis-grade agarose (Sigma ; see table 4) by melting in a microwave oven, mixing, adding ethidium bromide stain, cooling to 55^0C and pouring into a sealed gel casting platform.

Table 4: Appropriate agarose concentrations for separating DNA fragments of various sizes.

Agarose%	Effective range of resolution of linear DNA fragments (kb)
0.5	30 to 1
0.7	12 to 0.8
1.0	10 to 0.5
1.2	7 to 0.4
1.5	3 to 0.2

DNA samples were prepared with an appropriate amount of 10x loading buffer and loaded into wells. Appropriate DNA molecular weight markers were also loaded . The samples were electrophoresed in 1xTAE buffer. The gel was photographed on a UV transilluminator with a Polaroid MP4 camera using a Kodak, plus-X, Estar film.

2.5 Restriction analysis of DNA fragments by blotting and hybridisation.

Frequently it is desirable to identify an individual fragment in a complex mixture that has been resolved by gel electrophoresis. This is accomplished by a technique first described by Southern (1975), in which the fragments are transferred from the gel to a nitrocellulose membrane and the fragment of interest is identified by hybridisation with a complementary labelled nucleic acid probe (usually radioactive).

2.5.1 ***Bidirectional Southern blotting.***

The gel to be blotted was photographed with a cut corner (to aid later orientation of the samples) and then rinsed with distilled H_2O, denatured in 0.5M NaOH plus 1.5M NaCl solution for 30 minutes, rinsed with distilled H_2O, placed in neutralizing solution (0.5M Trisma Base, 3M NaCl, 0.3M trisodium citrate) for 30 minutes and rinsed with distilled H_2O. The gel was then placed in 20x SSC for 1 hour. For the blotting procedure, two pieces of Hybond N (Amersham) filter were placed on either

side of the gel. Paper towels and 3MM Whatman paper were stacked appropriately. The transfer was allowed to proceed for 2-16 hours. The filters were washed briefly in 2xSSC and air dried. DNA was then fixed by UV crosslinking (for 5 minutes).

2.6 Sequencing

2.6.1 *Sequencing reactions.*

Sequencing reactions are described individually for HX43 subclone (double stranded DNA) and the end clones of the YACs (single stranded DNA) in sections 3.3 and 4.7, respectively.

2.6.2 *Sequencing gel (polyacrylamide denaturing gel) preparation.*

The products of DNA sequencing reactions were analysed on 6% polyacrylamide denaturing gels which were prepared as follows:

The gel plates (Biorad) were washed extensively with soap and water, rinsed with ethanol and air dried. A film of Sigmacote (silanizing solution ; Sigma) was applied to the inner side of the back plate (the one bearing the reservoir). The gel plates were assembled according to manufacturer's instructions (Biorad) using 0.4mm spacers. For 100ml of 6% polyacrylamide gel the following were mixed thoroughly: 66ml of 8.3M urea solution, 10ml 10xTBE (108g Tris-base, 55g boric acid, 40ml 0.5M EDTA, pH 8.0 per liter of distilled H_2O), 24ml of 25% acrylamide solution (company name). For 21x50cm sequencing gels 100ml of this solution was required. 25ml of the above 100ml solution was mixed with 150µl of 25% ammonium persulphate and 150µl TEMED and immediately poured in the sealing tray to form a sealing gel at the base of the plates. The remaining 75ml of the acrylamide solution was mixed with 300µl of 25% ammonium persulphate and 34µl TEMED and slowly poured into the plates along one side at a 45^0 angle. The angle of the plates and rate of flow were adjusted so that the acrylamide ran in smoothly without bubbles. The comb was inserted carefully in the top end of the plates and the gel was allowed to polymerise for approximately 45 minutes. After the gel had

polymerised, it was removed from the sealing tray and placed in the electrophoresis apparatus. The lower and upper reservoirs were filled with 500ml of 1xTBE buffer. The comb was removed and the gel was prerun (60W constant) until the temperature of the gel reached 50^{0}C (30 to 45 minutes). After the gel had pre-electrophoresed, the samples were denatured at 90^{0}C for 5 minutes. Excess urea was removed from the wells and an aliquot (2-3μl) of denatured sequencing reactions were loaded into adjacent wells, in sets of four (A, C G, T) for each template. The gel was electrophoresed at constant power. The electrophoresis time was adjusted knowing that for 6% denaturing polyacrylamide gels, oligonucleotides of 26 bases co-migrate with the dye marker Bromophenol Blue and oligonucleotides of 106 bases co-migrate wih Xylene Cyanol (both markers contained in the "stop solution" for sequencing reactions). After electrophoresis, the two plates were carefully separated. The gel was adherent to front plate, which was then placed in a tray and gently covered with a fixative (10% methanol, 10% glacial acetic acid) for 5-10 minutes. Using a piece of 3MM Whatman paper the gel was gently peeled away from the plate, covered with plastic wrap, placed between two more pieces of Whatman paper and vacuum dried at 80^{0}C for 1 hour. The gel was then exposed to Kodak X-OMAT film at RT without an intensifying screen or -70^{0}C with an intensifying screen.

2.7 Preparation of solutions:

LB medium (Luria-Bertani)

	per liter
Tryptone	1% w/v
Yeast extract	0.5% w/v
NaCl	1% w/v

pH was adjusted to 7.0.

Solution I

50mM Glucose

10mM EDTA

25mM Tris-HCl

Prepared from 0.5M EDTA and 2M Tris-HCl stocks.

Solution II (prepared immediately before use):

0.2M NaOH

1% (w/v) SDS

Prepared from 5M NaOH and 10% SDS stocks.

Solution III:

Final concentration	per 100ml
3M Potassium acetate	29.44g
2M Acetic acid	12ml

TE buffer

Final concentration	per 500ml	stock concentration
10mM Tris-HCl pH 7.5	25ml	2M stock
1mM EDTA	100µl	0.5M stock

pH 7.5-8.0

3M Na acetate

102.1g resuspended in 250ml SDW. pH adjusted to 4.8 with glacial acetic acid.

TAE electrophoresis buffer

1x working solution:	50X stock solution pH ~8.5: per liter
40mM Tris acetate	242g Tris base
2mM EDTA	57.1ml glacial acetic acid
	37.2g $Na_2EDTA.2H2O$

10x loading buffer: 25% Ficoll Orange

10x stock	per 100ml
250mg	Orange G
25g	Ficoll
9.3g	EDTA

Denaturer

Final concentration	per liter
0.5M NaOH	20g
1.5M NaCl	87.66g

Neutralizer

Final concentration	per liter
0.5M Trisma base	60.5g
3M NaCl	175g
0.3M Na_3Citrate	88.25g

20x SSC (standard saline citrate)

Final concentration	per liter
3M NaCl	175g
0.3M Na_3citrate	88g

pH adjusted to 7.0 with concentrated HCl.

TBE electrophoresis buffer

1x working solution	10x working solution: per liter
89mM Tris base	108g
89mM boric acid	55g
2mM EDTA	40ml 0.5M EDTA

25% ammonium persulphate

0.25g ammonium persulphate resuspended in 1ml SDW.

3. Isolation and characterisation of a microsatellite from cosmid HX43.

3.1 Construction of a mini library from cosmid HX43

3.1.1 *Vector and cosmid restriction digests and DNA purification.*

5 μg each of vector (pT7T3) and cosmid DNA were digested with 20 units of BamHI and SauIIIa, respectively, in a 50 μl reaction at 37^{0}C for 2 hours. Completion of digestion was monitored by gel electrophoresis.

The digested DNAs were extracted with 50 μl of phenol, phenol / chloroform (1:1), and finally chloroform. 1/10 volume of 3M sodium acetate was added to the final aqueous phase, and the DNA precipitated with 2 volumes of absolute ethanol. Following centrifugation for 10 minutes at 4^{0}C, the pelleted DNA was washed with 70% ethanol, air dried and resuspended in 10 μl TE. An aliquot (1μl) of each purified DNA was electrophoresed on a 0.8% agarose gel. Concentrations were estimated at 250ng/μl for the SauIIIa digested cosmid HX43 and 100ng/μl for the BamHI digested vector pT7T3.

3.1.2 *Dephosphorylation and ligation.*

Cosmid HX43/SauIIIa DNA (1 μg) was dephosphorylated in a 10μl reaction. The temperature of the reaction was adjusted to 30^{0}C before the addition of 1unit of HK™ phosphatase (Epicentre Technologies). Incubation continued for 1 hour. The enzyme was heat (65^{0}C) inactivated.

An aliquot (50ng) of the dephosphorylated insert (HX43/SauIIIa) was ligated to 150ng (an equivalent number of molecules) of digested vector (pT7T3/BamHI) in a 10μl ligation reaction. A sample (1μl) of the reaction mixture was removed for use as a control before adding 1μl (0.1unit/μl) of T4 DNA ligase (Biolabs) to the remaining 9μl of the reaction. The ligation reaction was incubated at 12^{0}C overnight. Ligation control reactions were also carried out containing the following DNA samples as the only DNA present in the reaction: 1. Vector pT7T3/BamHI (100ng ; positive control)

2. Dephosphorylated insert (100ng of HX43/SauIIIa DNA ; negative control) 3. Phosphorylated insert (100ng ; positive control) 4. λHindIII (100ng ; positive control).

3.2 Isolation of a subclone from the HX43 mini library containing a CA repeat.

3.2.1 *Transformation of competent cells with ligated products.*

Competent cells were transformed with products of the mini library construction as follows: 2μl (40ng total DNA) of the above ligation mixture was added to 250μl of previously thawed competent E. coli XL1-Blue cells (ref to protocol). No DNA (2μl of H_2O) as well as 20ng of vector pT7T3/BamH1 ligated to itself and 1ng of wild type vector pT7T3 were used as controls. The transformation mixtures were placed on ice for 30 minutes, heat shocked at 42^{0}C for 2 minutes and placed on ice for 1 minute. 800μl of prewarmed LB broth was added to each mixture and placed at 37^{0}C for 1 hour. After centrifugation at low speed (4000rpm ; microcentrifuge) for 1 minute, the supernatant was removed and the cells resuspended in 100μl LB broth. These cells were plated onto sterile Hybond N filters on top of agar plates supplemented with 50μg/ml ampicillin, 15.7μg/ml IPTG (isopropylthio-β-D-galactoside) and 60μg/ml Xgal (5-bromo-4-chloro-3-indolyl-β-galactoside). The plates were incubated at 37^{0}C overnight.

3.2.2 *Preparation of replica filters.*

Replica filters, one for the experimental plate and one for the control (vector ligated to itself), were prepared as follows: the master filter was peeled off each transformation plate and laid (colony side up) on a sterile 3MM Whatman paper. A fresh sterile filter was laid on the master filter. The two filters were pressed firmly together and oriented asymmetrically by making a series of holes with a needle. The master filter was placed back onto the original plate and the replica filter on a fresh LB agar plate supplemented with 50μg/ml ampicillin. The two plates were incubated at 37^{0}C for 2 hours. The original plates were stored at 4^{0}C. Replica filters were

prepared for hybridisation as follows: They were laid on 3MM Whatman papers soaked in a) 10% SDS (Sodium Dodecyl Sulphate) for 3 minutes, b) denaturer (0.5M NaOH, 1.5M NaCl) for 5 minutes, c) neutralizer (0.5M Trisma Base, 3M NaCl, 0.3M Tri Na Citrate) for 5 minutes and d) 2xSSC (0.3M NaCl, 0.03M trisodium citrate) for 5 minutes. They were air dried and the DNA was fixed with UV irradiation for 5 minutes.

3.2.3 *Hybridisation of replica filters with oligonucleotide probe $(GT)_{11+1}$.*

Oligonucleotide $(GT)_{11+1}$ (20pmoles) was end labelled in a 20μl labelling reaction containing polynucleotide kinase buffer (one-phor-all buffer, Pharmacia), 50μCi of γ-^{32}P-dATP and 1μl (9.3 units) of polynucleotide kinase (Pharmacia) at 37^{0}C for 30 minutes.

The replica filters were prehybridised in 20 ml hybridisation mixture containing 0.125M NaPi (pH 7.2 ; 68.4ml Na_2HPO_4 and 31.6ml NaH_2PO_4 per liter of distilled H_2O), 4M NaCl, 1mM EDTA, 7% SDS and 10% PEG (Polyethylene Glycol) at 65^{0}C for 5 minutes and hybridised with the end labelled probe at 65^{0}C overnight.

3.2.4 *Posthybridisation washes*

Following hybridisation, the filters were washed twice (5 minutes each) in 4xSSC, 0.1% SDS at RT and twice (15 minutes each) in the same solution at 65^{0}C.

3.2.5 *Autoradiography*

The filters were exposed to X-OMAT AR Kodak autoradiographic film using an intensifying screen overnight at -80^{0}C.

3.2.6 *Isolation of the $(CA)_n$ positive clones*

Following autoradiography, the orientation holes of the replica filter were marked on the autoradiograph which was then matched to the master plate (bearing the same orientation holes) thus identifying the $(CA)_n$ positive colonies (corresponding to

positive signals). These colonies were isolated and DNA was prepared using the alkaline lysis DNA mini- or midiprep method. An aliquot of each preparation (2μl) was electrophoresed (0.8% agarose gel, at 60V for 1 hour), the gel blotted and the filter hybridised with the $[(GT)_{11}G]$ probe to confirm that the clone contained a $(CA)_n$ repeat.

3.3 Sequencing the HX43 subclone.

DNA (5 μg) was denatured in 20 μl of solution containing 0.2M NaOH and 0.2 mM EDTA at 37^{0}C for 30 minutes. Following incubation, 1/10 volume of 3M sodium acetate was added and the DNA was precipitated with 2.5 volumes of absolute ethanol. The DNA was pelleted by centrifugation at 4^{0}C for 10 minutes, washed with 70% ethanol and resuspended in 7μl H_2O. DNA sequencing was carried out using a United States Biochemicals (USB) sequencing kit as follows: The resuspended DNA was annealed to 3ng (1μl) Universal primer (-40) in a 10μl reaction containing 2μl Sequenase reaction buffer at 65^{0}C for 2 minutes followed by a 30 minute incubation period so that the temperature fell below 30^{0}C. To the annealed template-primer mixture 1μl 0.1M DTT, 2μl diluted labelling mix (1.5μM dGTP, 1.5μM dCTP, 1.5μM dTTP), 2μl diluted enzyme Sequenase version 2.0 (1:8 in ice-cold dilution buffer) and 0.5μl α-^{35}S-dATP (10μCi/μl) were added and the labelling reaction was incubated for 2-3 minutes at RT. Following labelling, 3.5μl were removed and transferred to each of 4 tubes containing 2.5μl of each of the 4 chain-terminating nucleotide mixtures. Incubation continued for 5 minutes at 37^{0}C and the reactions were terminated by adding 4μl of stop solution to each. The samples were denatured at 75-80^{0}C for 2 minutes before loading on the sequencing gel and 2-3μl were loaded in each lane.

3.4 PCR amplification of subclone pHX43-10.

Following sequencing of subclone pHX43-10, the unique sequences flanking the CA repeat were used to design PCR amplification primers. Primers were: pHX43/ II: 5'

CTC ATA CAC ACA GCT GAC TG 3' (GT strand) and pHX43/ III: 5' CTT CTT TCT CAT TCA CAG AC 3' (CA strand) (EMBL accession number X67604).

PCR amplifications were carried out both non-radioactively and radioactively. Non-radiactive PCR amplifications were carried out in 25μl reaction volumes containing 300pg of cosmid or plasmid DNA, 0.2μM of each primer, 200μM of each dNTP and 1unit of Taq polymerase (Northumbria Biologicals Ltd). Radioactive PCR amplifications were carried out with a ^{32}P end labelled primer. End labelling of 50pmoles of the primer (CA strand) was carried out in a 20μl reaction with 3μCi of γ-^{32}P-dATP and 1μl (9.3 units) polynucleotide kinase at 37^{0}C for 30 minutes. PCR reaction mix (15μl) consisted of 150ng of genomic DNA template, 1pmole of each primer (end labelled and unlabelled), 200μM of each dNTP and 0.5unit of Taq polymerase. Following an initial denaturation at 95^{0}C for 1 minute, samples (non-radioactive and radioactive) were processed through 30 cycles of denaturation at 92^{0}C for 1 minute, annealing at 50^{0}C for 1 minute and extension at 72^{0}C for 2 minutes and finally, an elongation step at 72^{0}C for 10 minutes. Non-radioactive PCR products were analysed on 2% agarose gels and radioactive products on 6% polyacrylamide sequencing gels.

4. Construction of a Yeast Artificial Chromosome (YAC) contig.

4.1. YAC library screening by PCR.

The ICI YAC genomic library (Anand et al, 1990) was screened with Sequence Tagged Sites (STSs) by PCR. The primary library consists of 35,000 clones with an average insert size of ~350kb representing a >3.5 times coverage of the human genome. The screening was carried out in three stages:

4.1.1 *Screening the YAC library primary pools.*

The ICI library (obtained from the U.K. Human Genome Mapping Project Resource Center) is represented in 40 primary pools of 864 YAC clones each (34,560 clones in total) in agarose plugs, stored in 5mM EDTA. Each plug was rinsed three times, for five minutes each in distilled sterile H_2O, melted at 65^0C for 20 minutes and diluted 1:1 with distilled sterile H2O. The diluted plugs were used for the experiments described and were stored at either 4^0C or - 20^0C. For specific PCR amplification using the primers of interest, each diluted pool was melted at 65^0C, and an aliquot (3μl) per 25μl volume reaction was used for analysis. Before starting the required PCR cycle, additional denaturation of pool DNA was carried out at 95^0C for 5-10 minutes. Following this cycle, the samples were cooled to 90^0C and 1 unit of Taq DNA polymerase was added. PCR conditions for each STS are described in relevant Results chapters (Chapter 5, section 2.1).

4.1.2 *Screening the YAC library secondary pools.*

Each primary pool (864 clones) of the ICI library is represented in 9 secondary pools containing one plug of 96 YAC clones each. The preparation and screening by PCR was carried out as described for the primary pools. Screening of the 40 primary pools resulted in one or more positive pools, for the specific sequence tested. For each positive primary pool, 9 secondary pools, labelled A-I, were screened using the primers from the same sequence of interest.

4.1.3 *Screening the YAC library tertiary pools.*

Each secondary pool of the ICI library is represented in 20 tertiary pools containing 20μl of cells each corresponding to a collection of individual clones as follows: pools 1-8 contain 12 individual clones each corresponding to one row (Table 5) and pools 9-20 contain 8 individual clones each corresponding to one column (Table 5). An aliquot (5μl) of each of the 20 pools corresponding to a positive secondary pool was used for PCR amplification (in a 25μl reaction) using the same primers of interest as for the primary and secondary screening. Screening the 20 tertiary pools resulted in two positive pools (one among pools 1-8 and one among 9-12) the combination of which identified the individual clone (see Table 5).

Table 5: ICI library tertiary pools.

	Pool 9	Pool 10	Pool 11	Pool 12	Pool 13	Pool 14	Pool 15	Pool 16	Pool 17	Pool 18	Pool 19	Pool 20
Pool 1	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
Pool 2	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
Pool 3	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Pool 4	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Pool 5	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
Pool 6	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Pool 7	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
Pool 8	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

4.2 YAC clone DNA preparation in plugs.

YAC clone DNA was prepared according to Vollrath and Davies (1987), with slight modification. Yeast clones were streaked out onto YPD^- plates and incubated at 30^0C overnight. A single large colony was picked into 10ml YPD^- media and grown at 30^0C for 24 hours; the culture was subsequently used to seed a 100ml or 200ml culture in YPD^- media, grown with agitation at 30^0C for 40 hours after which the $O.D._{600}$ was measured and the number of cells were calculated ($O.D._{600}$ of 0.3 corresponds to $3.3x10^6$ cells/ml).

Cells were pelleted by centrifugation for 10 minutes at 4K, and resuspended in 50ml SCE solution (1M sorbitol, 0.1M NaCl, 10mM EDTA). The cells were repelleted and resuspended in SCEM (SCE with 30mM β-mercaptoethanol, made fresh each time) to a final concentration of $2.5 x10^9$ cells per ml. Lyticase (Sigma) was added to a final concentration of 120μg/ml and incubation was performed at 30^0C for 1 hour. Low melting point (LMP) agarose (1%) was prepared in 1M sorbitol and kept molten at 55^0C. Yeast cells (0.5ml) were mixed with an equal volume of LMP agarose and 200μl aliquots were dispensed into plastic moulds and left to set at 4^0C.

The plugs were placed in proteinase K/sarcosyl solution (1% sodium lauryl sarcosine, 0.5M EDTA pH 8.0, 2mg/ml proteinase K ; 1ml per plug) and incubated at 50^0C for 2 days. Following extensive rinsing with TE, proteinase K was inactivated by treatment with 0.04mg/ml PMSF (phenyl-methyl-sulphonyl-fluoride) in TE with incubation at 50^0C for 30 minutes. The plugs were then rinsed three times in TE and stored at 4^0C.

4.3 Pulsed field gel electrophoresis (PFGE)

The practical usefulness of conventional agarose gel electrophoresis is limited to the separation of DNA up to 50kb in size. PFGE, a type of electrophoresis first described by Schwartz and Cantor in1984, is capable of separating chromosome size DNAs from the yeast genome . By alternating the electric field between spatially distinct pairs of electrodes, DNAs of the order of 10Mb are able to re-

orientate and move differentially through the pores in an agarose gel. The time each field for a particular electrode pair is on is called the pulse time. The shift between alternate fields is essentially instantaneous.

The experiments described here were conducted on the CHEF-DR II apparatus (Biorad) based on the CHEF (clamped homogeneous electric fields) technique (Carle and Olson, 1984).

4.3.1 *Casting and loading the gel.*

100ml of 1% agarose (Molecular Biology Certified ; Biorad) in 0.5x TBE buffer (Tris-Borate buffer ; 45mM Tris base, 45mM boric acid, 1mM EDTA) was poured into the casting stand (standard casting stand, Biorad) and allowed to set for 1 hour at RT. Sample plugs were loaded while the gel remained in the casting stand. Care was taken so that samples were less than 90% of the height of the well. They were then sealed in place by filling each sample well with 1% Low Melting Agarose (Biorad). The agarose was allowed to solidify for 15 minutes.

4.3.2 *Gel electrophoresis*

The electrophoresis chamber was rinsed briefly with two liters of pre-cooled (4^0C, for 1 hour) distilled H_2O. Two liters of precooled 0.5x TBE electrophoresis buffer were poured into the chamber and allowed to circulate briefly through the recirculating pump. The gel was placed safely in the center of the chamber (according to manufacturer's instructions). The buffer flow through the pump was adjusted to 1L/hour so that the gel was not disturbed during electrophoresis. The following electrophoresis parameters and voltage were selected to separate DNA in the range 0.1-2.0Mb (accurate operating instructions can be found in the Biorad CHEF- DR II manual): a. Pulsewave 760: Initial A Time: 60 seconds, Final A time: 90 seconds (Note: this pulse mode is called ramping), Run Time: 22 hours. b. Model 200/2.0 power supply: Voltage: 180V, Run Time: 22 hours. The temperature during electrophoresis was 14^0C.

4.3.3 *Gel staining.*

The gel was placed into 0.5μg/ml ethidium bromide solution (freshly prepared) and stained for 30 minutes. It was then destained in distilled H_2O for 1 hour, visualised on a UV transilluminator and photographed.

4.4 Alkali blotting of Pulsed Field Gel.

The gel to be blotted was depurinated in 0.25M HCl solution until the dyes xylene cyanol and bromophenol blue (previously electrophoresed on a conventional agarose gel for use as depurination indicators) had changed colour and for an additional 10 minutes. The gel was then rinsed in distilled water and placed in blotting buffer, 0.4N NaOH, for 15 minutes. For the blotting procedure, a sheet of Hybond-N+ (Amersham) membrane (cut to size) was placed on top of the gel (on the opposite side of the wells). Paper towels and 3MM Whatman paper were stacked appropriately. The transfer was allowed to proceed for 24 hours. The membrane was washed in 0.5M Tris-HCl for five minutes, rinsed briefly (1 minute) in 2xSSC and air dried. There was no need to fix DNA because of the alkali blotting procedure.

4.5 Hybridisation with Total Human DNA.

4.5.1 *Labelling Total Human DNA .*

50ng of Total Human DNA (Sigma) was labelled using random nucleotide prime synthesis (USB kit) in a 30μl reaction containing 3μl reaction mixture (random hexanucleotide mixture), 3μl of dNTPs mixture (dATP, dGTP, dTTP in 1:1:1 mixture ; 0.5mM each nucleotide), 5μl of α-^{32}P-dCTP (10μCi/μl) and 2μl (4units/μl) of Klenow enzyme at 37^{0}C for two and a half hours, after which % incorporation was measured by the following simple procedure: 0.5μl of the above reaction was placed on a 2.5cm Whatman round glass microfibre filter and a Geiger monitor was appropriately adjusted to read 100 counts. The filter was then washed under vacuum with 15ml 5% TCA (Trichloroacetic acid) and

monitored again. The new counts gave a rough measure of percentage of the incorporated label. If incorporation was efficient, the probe DNA was purified and precipitated as described in section 1 (in situ section) and resuspended in 100µl H_2O. It was then denatured at 100^0C for 5 minutes and briefly centrifuged for 30 seconds. The probe was then ready for hybridisation.

4.5.2 *Hybridisation.*

The filter was prehybridised in NaPi hybridisation buffer (section 3.2.3 of this chapter ; in a petri dish) at 65^0C for 20 minutes, after which the probe was added and hybridisation was performed at 65^0C overnight.

4.5.3 *Posthybridisation washes and autoradiography.*

After hybridisation the filter was washed in 2xSSC / 0.1% SDS, twice for 5 minutes each at RT ; in 2xSSC / 1% SDS, twice for 30 minutes each at 65^0C ; and in 0.2xSSC / 0.1% SDS, once for 30 minutes at 65^0C.
The filter was exposed to X-OMAT AR Kodak autoradiographic film ovenight at RT.

4.6 Isolation of YAC ends using the Vectorette system.

The development of Vectorette PCR (Riley et al, 1990) has enabled the recovery of short fragments of DNA adjacent to a known sequence by PCR. The use of known vector sequence to obtain the uknown insert sequence next to it for chromosome walking (Anand et al, 1991) is one application of this method. The method described below applies to YAC clones and was generally carried out according to the Chemical Genetics protocol (Cambridge Research Biochemicals) with modifications (especially concerning PCR reaction mixtures, conditions and sequencing) to satisfy individual experiment requirements.

4.6.1 *Construction of Vectorette library*

A segment (approx. 1/4) of a 200μl YAC DNA plug was equilibrated in 1ml of cold TE overnight. TE was replaced with 1ml of the relevant 1x restriction buffer for 1 hour on ice. The buffer was then replaced with 100μl of fresh 1x restriction buffer and 20 units of the chosen restriction enzyme and incubation was performed overnight at 37^0C.

The restriction buffer/enzyme solution was replaced with 1ml of 1x ligase buffer (Gibco BRL or Pharmacia) and the plug equilibrated for 1 hour on ice. The buffer was then replaced with 100μl of fresh ligase buffer and 5μl (3 pmol) of the appropriate blunt-ended Vectorette units (CRB) were added. The plug was melted at 65^0C for 10 minutes, followed by cooling to 37^0C. ATP to final concentration of 1mM and 9 units of T4 DNA ligase (Gibco BRL or Pharmacia) were added and the ligation reaction was incubated at 37^0C for 2 hours. 100μl of distilled sterile H_2O was added and the Vectorette library was stored in aliquots at -20^0C.

4.6.2 *Vectorette PCR.*

PCR amplification of the Vectorette library was performed using the Universal Vectorette Primer (CRB) and the pYAC4 vector left or right arm primers (CRB). The primers were phoshorylated or unphosphorylated depending on which side of the PCR YAC end product was chosen to sequence from as discussed below in the sequencing section. For PCR amplification, the following were combined in a sterile microtube: 1μl of Vectorette library, 10μl of 10x PCR buffer (N.B.L.), 50μM of each dNTPs, 1μM of Universal Vectorette Primer and 1μM of specific primer. The volume was made up to 99μl with distilled sterile H_2O and the sample was overlaid with mineral oil. The sample was heated to 94^0C for 3 minutes and 2.5 units of Taq polymerase added. It is important that the temperature does not fall below 94^0C as it was found to be vital to the specificity of the reaction. PCR conditions were as follows: denaturation at 94^0C for 1

minute, annealing at 60^{0}C for 1 minute and elongation at 72^{0}C for 2.5 minutes for 40 cycles. The Vectorette PCR products were analysed on a 1% agarose gel stained with ethidium bromide.

The above PCR reaction mixture was frequently modified as discussed in relevant results chapter.

4.7 Direct sequencing of Vectorette PCR products.

4.7.1 *Preparation of single stranded DNA*

In the method described here, one of the two PCR primers has a 5' phosphate group. The strand containing this primer is degraded by lambda exonuclease while the other strand containing a primer without a 5' phosphate remains intact. Therefore, the sequencing primer must derive from the strand containing the phosphorylated primer. The phosphate group was chemically added to the primers used during oligonucleotide synthesis (CRB). The sequencing primers used were internal to the PCR primers and also came from CRB. The methods used are described below:

a. Lambda exonuclease treatment of a partially purified PCR product:

After checking an aliquot (15μl) of the PCR reaction on an agarose gel, the remaining 85μl was extracted with an equal volume of phenol and centrifuged for 5 minutes. To the top aqueous layer two volumes of pre-cooled absolute ethanol were added and the sample incubated at -70^{0}C for 10 minutes and centrifuged for 10 minutes. The supernatant was removed and the pelleted DNA air dried and resuspended in 20 μl of sterile distilled H_2O. To the purified PCR product 2.5μl 10x λ-exonuclease buffer (67mM glycine-NaOH pH 9.4, 2.5mM $MgCl_2$), 1.5 μl sterile distilled water (SDW) and 1μl λ-exonuclease (4units/μl ; Gibco BRL) were added and incubation was performed at 37^{0}C for 30 minutes. After incubation, 25μl of SDW was added to the sample and two phenol

extractions were performed and the DNA precipitated with ethanol. The pelleted DNA was air dried and resuspended in 10μl of sterile distilled water.

b. Lambda exonuclease treatment of column purified PCR product:

In later experiments column purification (Microspin Columns ; Pharmacia) of the PCR product was performed as a more efficient method of purification. The amount of λ exonuclease was increased on the assumption that more single-stranded DNA for sequencing would be yielded.

85μl of the PCR product divided into 4 samples was extracted once with an equal volume of phenol/chloroform/isoamyl alcohol and centrifuged for 1 minute. The aqueous layer of each sample was transferred to the top of the compacted column resin (prepared according to manufacturers' instructions ; Pharmacia) which was then centrifuged for 1 minute at 3000rpm. Each effluent contained the purified PCR product (~22μl) to which 2.5μl of 10x Lambda Exonuclease buffer (Pharmacia) and 1μl of Lambda exonuclease (8-10units/μl ; Pharmacia) were added and incubated at 37^{0}C for 20 minutes. After incubation, the samples were phenol extracted once and passed through the column as before. The four effluents containing the single-stranded DNA were pooled and ethanol precipitated. The pelleted DNA was resuspended in 10μl SDW, and used for the sequencing reaction.

4.7.2 *DNA sequencing.*

Sequencing was carried out using γ-^{32}P end labelled sequencing primer or, alternatively, α-^{32}P incorporation labelling utilising the Pharmacia sequencing kit in both cases. The protocols followed are described below:

a. Using ^{32}P end labelled primer.

1pmol of sequencing primer was end labelled in a 10μl labelling reaction containing 1μl 10x kinase buffer, 1μl [γ-^{32}P]-dATP (10μCi/μl), 6μl SDW and 1μl T4 polynucleotide kinase at 37^{0}C for 30 minutes. 5μl of the labelled primer was

annealed to 5μl single-stranded PCR product (prepared as described in 4.7.1a) in a 15μl volume reaction containing 3μl annealing buffer and 2μl SDW at 70^{0}C for 5 minutes followed by cooling to 45^{0}C for 10 minutes. During this time, four tubes with 4μl of "termination mix" (A, C, G and T) were prepared and kept on ice until use. 2μl of T7 DNA polymerase (diluted 1:4 in cold enzyme dilution buffer) was added to the annealed template and 4μl of the mixture transferred into each tube containing the "termination mix". The microtubes were incubated for 2-5 minutes at 37^{0}C. After incubation the reactions were terminated with 4μl "stop solution" for each tube. The samples were analysed on a 6% acrylamide denaturing gel.

b. Using ^{32}P incorporation labelling:

10μl of single-stranded PCR product (prepared as described in 4.7.2b) was annealed to 1.5pmol of sequencing primer in a 14μl reaction containing 2μl annealing buffer at 60^{0}C for 10 minutes followed by cooling to RT for 10 minutes. 2μl of T7 DNA polymerase (diluted 1:4 in cold enzyme dilution buffer), 3μl of labelling mix-dCTP diluted 1:4 in 330mM NaCl and 1μl of [α-^{32}P]-dCTP were added to the annealed template. The reaction was incubated at RT for 5 minutes. During this time, four tubes with 2.5μl of "termination mix" (A, C, G and T) were prepared and prewarmed at 37^{0}C. Following labelling, 4.5μl aliquots were transferred into each of the four sequencing mixes and the samples incubated at 37^{0}C for 4 minutes. The reactions were terminated with 5μl of "stop solution".

4.8 Preparation of solutions.

(Section 4.2)

YPD$^-$ medium

Final concentration

0.7% (w/v)	Yeast nitrogen base
2% (w/v)	Glucose
0.0055% (w/v)	Adenine hemisulphate
0.0055% (w/v)	Tyrosine
1.4% (w/v)	Casamino acids

YPD$^-$ agar: 1.5% bacteriological agar was added in 100ml YPD$^-$ medium.

SCE

Final concentration

1M	D-sorbitol
0.1M	NaCl
10mM	EDTA

SCEM

This is SCE solution containing 30mM β-mercaptoethanol.

24μl of 12.5M β-mercaptoethanol were added to 10ml SCE.

1M Sorbitol

36.44g D-sorbitol (MW 182.2) were resuspended in 200ml SDW. The solution was autoclaved.

1% Low Melting Point agarose:

0.1g LMP agarose melted in 10ml 1M sorbitol.

Proteinase K solution:

Final concentration

1% w/v Sodium lauryl sarcosine

0.5M EDTA

2mg/ml Proteinase K

PMSF

Final concentration: 0.04 mg/ml in TE.

17.4mg PMSF were resuspended in 10ml isopropanol. This was the 10mM stock solution (1.74mg/ml) kept in aliquots at -20^{0}C. 1.14ml of stock solution were added to 50ml TE to give the above final concentration.

1.2 Chromosome analysis by FISH.

1.2.1 *General considerations.*

In situ hybridisation provides the most direct way to study the localisation of DNA sequences. The method involves hybridisation of labelled nucleic acid probe with target nucleic acids in cells, tissue sections or chromosomes. It is based on the pioneering studies of Gall and Pardue in 1969 who used 18S and 28S tritium-labeled ribosomal RNA to detect these genes in oocytes of the toad xenopus laevis, and Jones who in 1970 localised mouse satellite DNA to the pericentromeric heterochromatic regions of mouse chromosomes. Since then, routine isotopic in situ hybridisation protocols were established in many laboratories during the 1970s, but between 1969 and 1981 the technique was limited to detection of repetitive sequences in eucaryotic genomes (Saunders et al, 1972, Jones et al, 1973, Gosden et al, 1975, Steffensen et al, 1974, Evans et al, 1974). In 1981 Gerhard et al, Malcolm et al, and Harper et al were the first to demonstrate the possibility of localising single-copy sequences cloned from individual genes by isotopic in situ hybridisation. Detailed protocols for isotopic chromosomal localisation of DNA probes have been described elsewhere (Pardue 1985, Buckle and Craig, 1986). In the late 1970s and early 1980s non-isotopic in situ hybridisation techniques were introduced and developed (Rudkin and Stollar, 1977, Bauman et al, 1980). Since then, this approach has become increasingly popular and has found applications in a number of fields such as cytogenetics, for example for the analysis of numerical and structural chromosome aberrations in metaphase and also directly in the interphase nucleus (interphase cytogenetics ; Cremer et al, 1986), the latter are often associated with tumor diagnosis (Cremer et al, 1988, Hopman et al, 1988). In situ methods are also used in histopathology for the detection of virus in infected human tissue (e.g. Wells et al., 1987), in basic biological studies of chromosome structure and organisation (e.g. Baumgartner et al, 1991) and molecular genetics for

gene mapping. The expanding use of non-isotopic in situ hybridisation techniques and in particular of fluorescent in situ hybridisation (FISH) is a result of several factors, notably: **1)** the advantages over isotopic methods including high specificity, improved signal resolution, rapid results, probe stability and, in the case of gene mapping, direct mapping without extensive statistical analysis **2)** the sensitivity of the technique approaches that of isotopic methods (2-5kb). Single-copy sequences of about 1kb have also been successfully mapped although the discrimination of weak specific signals against background hybridisation may require statistical analysis and/or very sensitive camera systems **3)** the ability to combine several techniques for simultaneous multicolour detection of different sequences within the same preparation (e.g. Nederlof et al, 1991) **4)** the steadily improving quality of fluorescence microscopes over the last decade **5)** the ability to specifically highlight the entire genome of a particular species, entire chromosome, chromosomal subregions, or single-copy sequences, depending on the complexity of the probes used **6)** the development of in situ suppression hybridisation protocols (Landegent et al, 1987) in which the hybridisation signal from ubiquitous repeated sequences, such as the Alu and Kpn elements, is suppressed by prehybridisation of probes with unlabeled genomic DNA. This is crucial for localisation of unique sequences contained in large insert probes, such as cosmids and yeast artificial chromosomes (YACs).

The principles of the technique are as follows: DNA or RNA sequences are first labeled with reporter molecules. The probe and target chromosomes or nuclei are denatured. Complementary sequences in the probe and target are then allowed to reanneal. After washing and incubation in fluorescently labelled affinity reagents, a discrete fluorescent signal is visible at the site of probe hybridisation.

1.2.2 Methodological considerations.

1.2.2.1 Chromatin preparation and denaturation.

Cells are hypotonically swollen and fixed on slides by conventional procedures. Slides are initially treated with RNAase to remove endogenous RNA. They are then incubated briefly at 70^0C in a solution containing 70% formamide to dissociate the DNA into single strands, and then fixed in cold ethanol to reduce strand reannealing before the addition of probe.

1.2.2.2 Probe labelling.

A number of sensitive non-isotopic labelling and detection procedures are available. Probes may be directly conjugated with fluorescent molecules (Bauman et al, 1980), however, the most widespread approach is to label probes with reporter molecules which, after hybridisation, bind fluorescent affinity reagents. Typical reporter molecules include biotin, digoxigenin, dinitrophenyl (DNP), aminoacetylfluorene (AAF), mercury and sulfonate (reviewed in Bauman et al, 1990). Two labelling and detection techniques are now routinely used in many laboratories because of the high sensitivity and commercial availability of the materials required : probe labelling with either biotin or digoxigenin as reporter molecules and indirect detection of hybridised probe via avidin or antibodies conjugated with fluorescent dyes. Biotin and digoxigenin (and also DNP) are incorporated as labelled nucleotides using nick translation. The other reporter molecules are attached to DNA through chemical reactions. The size of the probe is critical for in situ hybridisation and should be between 100 to 500 nucleotides, a size range that maximises specific hybridisation and decreases background hybridisation (Lawrence et al, 1985). This size range can be obtained by adjusting the DNAase concentration in a nick translation reaction (see discussion) or by sonication in case of chemical modification.

1.2.2.3 Hybridisation.

Labelled probes are mixed in a hybridisation buffer containing formamide, salt, dextran sulphate and blocking DNA (e.g. salmon sperm DNA) or E. coli t-RNA. For repetitive probes the mixture is denatured and applied to slides after chromosome

denaturation or is directly applied to slides and denatured simultaneously with the chromosomes. For unique sequence contained in large-insert probes, the probe is preannealed with excess unlabeled genomic or Cot1 DNA to reduce repetitive sequence binding to the target, denatured and applied to slides. Slides are incubated overnight (16-18h) at 37^{0}C, although DNA reassociation is presumed to take place more quickly. The stability of DNA-DNA (or RNA-DNA) hybrids can be assesed from the melting temperature (T_m). T_m is the temperature at which 50% of the nucleic acid duplexes dissociate. The DNA-DNA reassociation temperature (T_r) occurs 25^{0}C below T_m. Lower salt concentration and higher temperature favour accurate base pairing. However, the presence of formamide in the hybridisation mixture lowers the temperature required for accurate base pairing and so helps preserve the morphology of chromosomes. Dextran sulphate has been shown to increase the rate of reassociation of DNA (Wahl et al, 1979). Blocking DNA or E. coli t-RNA is included in the hybridisation mixture to reduce non-specific binding of the probe to chromatin and glass.

1.2.2.4 Stringency washes.

In the case of DNA-DNA hybrids non-specifically bound DNA probe is removed by washes in salt solutions of varying ionic strengths, or by incubation in low salt (with or without formamide) at high temperatures. Incubations at high temperatures for prolonged periods destroys the architecture of the chromosomes and should be avoided. The standard post-hybridisation washes are performed with 50% formamide in 2xSSC followed by 2xSSC at a temperature 5-7^{0}C higher than the hybridisation temperature. The number and duration of washes may vary according to the probe size.

1.2.2.5 Fluorescent detection.

Biotin, a ubiquitous small molecular weight vitamin, was originally employed in non-isotopic in situ hybridisation in order to utilize its high binding affinity to the naturally occuring glycoprotein (found in egg white) avidin (Langer et al, 1981). Thus, biotinylated probes are detected with fluorochrome-conjugated avidin. Anti-biotin antibodies can also be used. Digoxigenin is the aglycone derivative of digoxin (a derivative of the foxglove Digitalis purpurea, initially used for the treatment of heart disease) and antibodies to digoxin have been shown to cross-react with it with high affinity (Smith et al, 1970 ; Monji et al, 1980 ; Valdes et al, 1984). Thus, the well-established antisera used in the assessment and treatment of digoxin toxicity were applied to the detection of digoxigenin in non-isotopic in situ hybridisation. Digoxigenated probes are detected with specific antibodies followed by fluorescently labeled anti-immunoglobulins. Similar approaches are used for the detection of DNP, AAF and sulfonate. Different fluorescent labels are used in FISH, distinguished by their absorbtion and emission spectra, that is, the range of light wavelengths required for excitation and those re-emitted by that substance. The relationship between the wavelength of the "exciting" light and that of the "emitted" light is governed by Stokes' law. This states that light re-emitted by a substance is always of a longer wavelength and lower intensity than the excitation light. The most commonly used fluorescent labels are fluorescein isothiocyanate (FITC ; Riggs et al,1958) which is excited at 495nm and emits in the green-yellow region at 515nm, rhodamin and Texas Red which are excited at 550nm and 595nm and emit in the red region at 575nm and 615nm, respectively. Another fluorochrome, used in this work, is 7-amino-4-methyl coumarin-3-acetic acid (AMCA ; Khalfan et al, 1986) which is excited by UV light (350nm) and emits in the blue region at 450nm.

1.2.2.6 Chromosome counterstaining and banding.

The most commonly used fluorescent dye for chromosome staining is propidium iodide which excites at 340nm and emits in the red region at 600-610nm. Banding methods include using actinomycin/diamidinophenylindole (DAPI ; emission in the

blue region) staining for G-bands, quinacrine (emission in the yellow region) staining for Q-bands, DAPI alone (without additional stains such as actinomycin or distamycin) staining on heat denatured chromosomes for a Q-banding-like pattern, growth in bromodeoxyuridine, Hoechst 33258 staining and UV irradiation for R-bands, or FISH of cloned Alu sequences for an R-banding-like pattern.

1.2.2.7 Microscopy.

Signals from painted whole chromosomes, stained sub-chromosomal regions, or localised single probes are generally visible when an epifluorescence microscope is used. It is very important to have appropriate objectives and filter sets for the visualisation of the essential fluorochromes. Optical filters perform two tasks: a) they select the optimum wavelength of light for excitation of the fluorochrome (excitation filters) b) they suppress the excess exciting light and select out the emission wavelength of the fluorochrome (barrier filters). For some applications a filter set may be used which is very selective for a given fluorochrome, while in other applications filter sets may be used which allow the simultaneous detection of two fluorochromes, such as FITC for signal detection and propidium iodide for chromosome staining. Double band-pass filter sets have become available which allow the simultaneous recording of fluorochromes, such as FITC and rhodamine. Such filter sets may become very important in two colour mapping experiments. Alternatively, a change of filter sets can be used in two colour mapping experiments although these can cause more or less pronounced image shifts detectable after double exposure of photographic films (registration problem). Images of hybridised chromosomes or cells can be recorded with high resolution on high speed colour film or into a computer memory via a static CCD (charge coupled device) camera or a confocal laser scanning microscope. CCD cameras and confocal laser scanning microscopes allow the generation of digitized images. The former system uses the most sensitive and expensive camera system and enables the recording of signals which are not visible to the observer's eye.The second system can greatly reduce

the out-of-focus fluorescence allowing generation of high quality optical sections through the labelled specimen. It is especially suitable, therefore, for three-dimensional work.

2. Materials and methods.

2.1 In situ hybridisation of DXZ1 clone: biotin labelling and FITC detection.

2.1.1 *Probe preparation and labelling*

DXZ1, a 2kb BamH1 X-specific centromeric probe defining the higher repeat unit of the X chromosome alpha satellite (Mahtani and Willard, 1988) was used as a control for the technique. It was obtained from the American Type Culture Collection. DNA was prepared using the alkaline lysis midiprep method and resuspended in 100μl TE. An aliquot (5.5μg) of DNA was digested with BamH1 (5units ; NBL) in a 10μl reaction containing the manufacturer's recommended buffer supplemented with 1mM spermidine and 0.1mg/ml BSA at 37^0C for 1 hour. The sample was electrophoresed on a 0.8% agarose gel to check DNA quality and insert size.

The probe was labelled with biotin-16 dUTP using 1) random oligonucleotide primed synthesis 2) nick translation. Random oligonucleotide primed synthesis was carried out in a 20μl labelling reaction as follows: 1μg of DNA was denatured by boiling for 5 minutes, 3μl of dNTP mixture (dATP, dCTP, dGTP, 0.5mM each ; Bethesda Research Laboratories), 2μl of hexanucleotide mixture (Boehringer Mannheim ; random primed DNA labelling kit), 2.5μl (2.5nmol) of biotin-16-dUTP (Boehringer Mannheim), 10.5μl of H_2O and 1μl (2units/μl) Klenow enzyme were added. The reaction was incubated for 45 minutes at 37^0C. Labelling by nick translation has been described in chapter 2, section 1.2. Probe purification and precipitation and biotin incorporation test have been described in chapter 2, sections 1.3 and 1.4.

2.1.2 *Preparation of hybridisation cocktaiis*

A total volume of 100μl hybridisation cocktails containing 1ng/μl or 2.5ng/μl DXZ1 probe DNA, 2xSSC, 500μg/ml E. coli t-RNA, 10% dextran sulphate and 50% formamide were prepared. A "no DNA" control hybridisation cocktail was also prepared. 40μl of these hybridisation cocktails were applied per slide (probe DNA concentration 40 and 100ng/slide). Probe and chromosome denaturation were performed simultaneously by transferring the slide to a plastic tray floating in a waterbath at 75^{0}C for 10 minutes. Hybridisation was then carried out at 37^{0}C for 18 hours.

Post hybridisation washes and signal detection have been described (chapter 2, sections 1.6 and 1.7.1, respectively).

2.1.3 *Signai visualisation*

Signals on metaphase chromosomes were viewed and photographed using a Zeiss Photomic 3 microscope (with epifluorescence attachment) and a 1000 ASA Kodak Gold film. This also applies to signals from cosmids HX5, HX10, HX60, HX79, HX82 and HX97 in single in situ hybridisations (detection with FITC-conjugates). For the remaining cosmids signal visualisation was carried out as described in section 1.8 of chapter 2.

2.2 In situ hybridisation of DXZ1 clone: digoxigenin labelling and FITC detection

Conditions needed to be altered in order to use digoxigenin successfully. Alterations were carried out in two main areas of the protocol:

a) Nick translation: Modified protocols for labelling the probe by nick translation were used. A comparison between the previously used nick translation reaction (standard reaction, section 1.2 of chapter 2 ; modified from a protocol provided by Dr John Gosden) and the modified one (modified from a protocol supplied by Prof. v. d.

Ploeg, Dr Ton Raap and Dr Hans Tanke, Department of Cytochemistry and Cytometry, Leiden, Netherlands through Boehringer Mannheim) is shown in table 6:

Table 6: Comparison of the modified nick translation with the standard reaction.

Components, stock concentration.	Modified reactions: final concentration	Standard reaction: final concentration
DNA ,1μg/μl	1μg	1μg
10x buffer, 0.5M Tris-HCl pH 7.8 ; 50mM $MgCl_2$; 0.5mg/ml BSA	1x	1x of similar buffer (section B.1.2)
dATP, dGTP, dCTP, 0.5mM each	0.05mM each	0.4mM each
Digoxigenin-11-dUTP, 1mM	0.04mM	0.125mM
dTTP, 0.5mM	0.01mM	a. 0 or b. 0.025mM
DTT, 100mM	10mM	-
DNAase I , 1μg/ml (1000pg/μl)	A: 5000pg or B: 500pg	200pg
DNA Polymerase I, 10units/ μl	10units	2units
Total volume	50μl	20μl

b) Detection of digoxigenated DXZ1

Three different sets of antibodies were initially used as follows: 1) sheep, anti-digoxigenin, Fab fragments (2.5μg/ml ; primary antibody, Boehringher Mannheim) and donkey, anti-sheep IgG-FITC (7μg/ml ; secondary antibody, Sigma) (Lichter et al., 1990) and 2) sheep, anti-digoxigenin, whole molecule (0.4μg/ml ; Boehringer Mannheim) and rabbit, anti-sheep IgG (20μg/ml, Vector ; Trask et al., 1991) and 3) sheep, anti-digoxigenin, whole molecule (0.4, 0.8 and 4μg/ml) and donkey, anti-sheep IgG-FITC (5, 7, 10, 20 and 100μg/ml). For this set, probe concentrations of 80 and 120ng/slide were also used (in addition to the normally used concentration of 40ng/slide)

None of the above antibody sets or different antibody and probe concentrations (in the case of the third set) led to signal detection. Therefore, an additional set of antibodies (mouse, anti-digoxigenin, rabbit, anti-mouse-FITC and goat, anti-rabbit-FITC ; based on data supplied by Prof. v. d. Ploeg, Dr Ton Raap and Dr. Hans Tanke, Department of Cytochemistry and Cytometry, Leiden, Netherlands through Boehringer Mannheim) and an overall altered detection protocol were used (chapter 2, 1.7.2).

2.3 In situ hybridisation of cosmid HX79

2.3.1 *Probe preparation*:

Biotin labelling by nick translation, purification and precipitation of cosmid HX79 DNA have been described in chapter 2, sections 1.2 and 1.3. DNA was resuspended in 50μl TE (instead of 25μl). A biotin incorporation test was performed (chapter 2, 1.4).

2.3.2 *Preparation of hybridisation cocktails*

The hybridisation cocktail for cosmids described in chapter 2, 1.5.3 was based on previous attempts for successful FISH with cosmid HX79. For this cosmid, 250ng of labelled probe were initially combined with 35μg total human DNA (1:140 ratio) in the presence of 5xSSC in a 40μl reaction volume, denatured at 100^0C for 5 minutes, placed in ice for 1 minute and preannealed at 60^0C for 10 minutes. To the pre-annealed probe DNA 50μl of 20% dextran sulphate in 100% formamide and 5μl of 10mg/ml carrier E. coli t-RNA were added to make a 100μl hybridisation cocktail containing 2.5ng/μl probe DNA, 2xSSC, 500μg/ml E. coli t-RNA, 10% dextran sulphate and 50% formamide. On the basis of the in situ results, this hybridisation cocktail was modified so that competition was increased to 1:1400. Final concentrations of the other components were only slightly modified (2ng/μl probe DNA and 2.5xSSC). Further improvement led to the preparation of the hybridisation cocktails described in chapter 2, 1.5.3.

40μl of hybridisation cocktail were applied per slide (or 20μl when co-hybridised with centromeric probe DXZ1). Post hybridisation washes have been described in chapter 2, 1.6. Application of two, three and four layers of antibodies was attempted for the enhancement of signal.

2.4 In situ hybridisation of cosmid clones:

Cosmid clones in this study were derived from an X-specific somatic cell hybrid library (Lindsay and Bird, 1987). The insert DNA was partially digested with the enzyme MboI and cloned into the BamH1 site of cosmid vector pJB8 with an average insert size of 40kb. They had been previously tested for X specificity (Lindsay and Bird, 1987) and were selected for sublocalisation to X chromosome

regions using FISH because each of the cosmids contained a CA repeat. Before in situ hybridisation, an aliquot of each cosmid DNA was digested with HindIII and electrophoresed along with an undigested aliquot and an appropriate size standard marker (λDNA/HindIII) on a 0.8% agarose gel to check the quality of preparation, estimate concentration and confirm that the cosmid had the expected HindIII digest pattern (data not shown).

Cosmids HX5, HX10, HX60, HX82, HX97 and HX79 were labelled with biotin-16-dUTP and cosmids HX14, HX15, HX46, HX65 and HX100 were labelled with digoxigenin-11-dUTP. All cosmids were detected with FITC-conjugates. The methods used have been described in chapter 2, section 1.

2.5 Two colour in situ hybridisation.

2.5.1 Detection of DXZ1 and cosmid clones

Biotinylated DXZ1 probe was detected using AMCA-avidin D. Concentrations of 1, 5, 6 and 7μg/ml in either 4xSSC, 0.05% Triton-100 and 5% low fat milk powder or 4xSSC, 0.05% Tween 20 buffers were used.

Biotinylated cosmids were detected with 6, 7, 20 and 30μg/ml of AMCA avidin D dilluted in the above buffers.

2.5.2 Mapping of cosmids relative to reference markers HX74 and L1.28

Cosmid clones HX4, HX20, HX39, HX43 and HX91 were derived from the same library as described above and they were mapped by FISH (by Dawn Thiselton) to the proximal short arm of the X chromosome. An additional cosmid from the same library, HX99, had been previously mapped by somatic cell hybrids to a similar position in proximal Xp (Susan Lindsay, unpublished observation). The above cosmids and cosmids HX65 and HX97 (see results) were selected to be mapped relative to HX74 (DXS426 ; Luty et al, 1990) and L1.28 (DXS7 ; Bhattacharya et al,

1984) cosmids by the two colour FISH technique. HX74, a cosmid derived from the same library, had been shown to contain the DXS426 locus (by Dawn Thiselton) by PCR amplification using DXS426 specific primers 5' CCT TCA TCT CTA CCA AGA TA 3' and 5' CTG CAC TCC AGC CTG AAT AA 3' (Luty et al, 1989, Coleman et al, 1990). The LI28 cosmid was a gift from A. P. Monaco (Institute of Molecular Medicine, Oxford).

Cosmid HX20 was initially cohybridised with the reference markers in two triple hybridisation experiments. In the first one, HX20 was biotinylated and detected with AMCA-conjugates and the reference markers were digoxigenated and detected with FITC-conjugates. In the second experiment, HX20 was digoxigenated and detected with FITC-conjugates and the reference markers were biotinylated and detected with AMCA-conjugates (data not shown). Based on these results the first combination of labelling and detection was selected to localise the remaining cosmids. Additionally, each cosmid in question was hybridised separately with each reference marker. The following experiments were carried out:

1) Biotinylated HX4, HX39, HX43, HX65, HX91 and HX99 were cohybridised with digoxigenated reference marker HX74.

2) Cosmids HX4, HX43 and HX91 were cohybridised with reference marker L1.28 and also with L1.28 and DXZ1.

3) HX97 was cohybridised with both reference markers.

Two metaphase spreads slides for each combination of cosmids and a "no DNA" control for each experiment were routinely prepared.

3 Results.

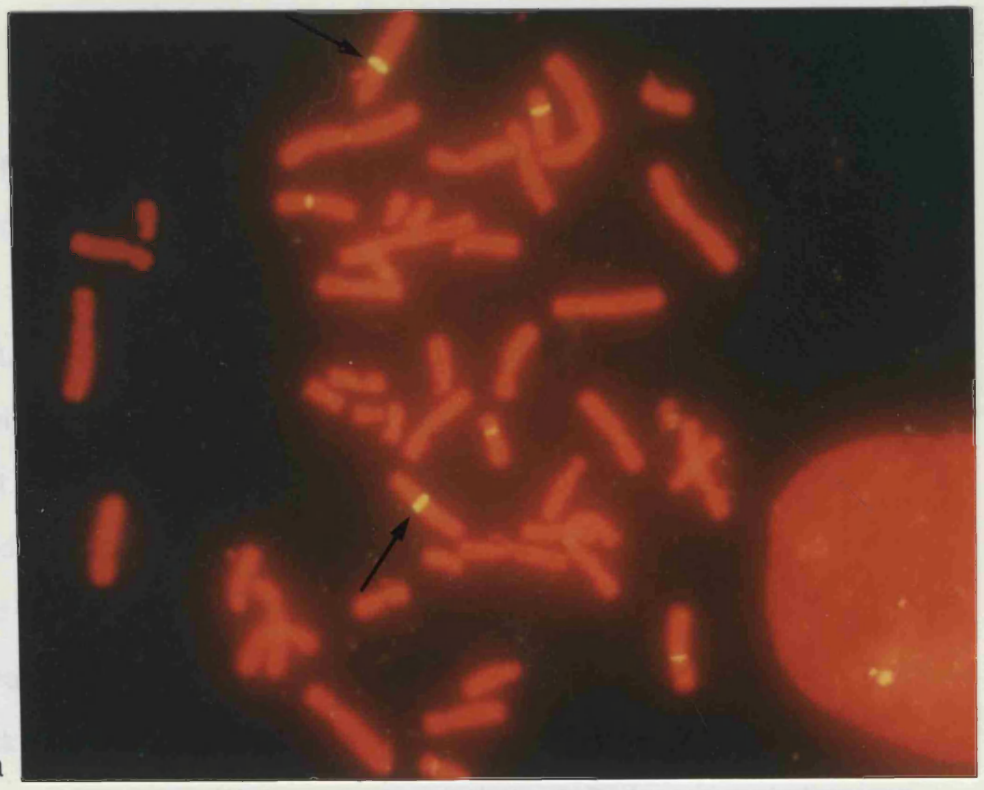
3.1 <u>The controls</u>.

DXZ1 was chosen as a control for the effectiveness of the FISH technique because as an X-specific alpha centromeric probe it would be easily detectable and also identify the chromosome of interest. It was labelled with biotin-16-dUTP using 1) random oligonucleotide primed synthesis and 2) nick translation. A colorimetric test

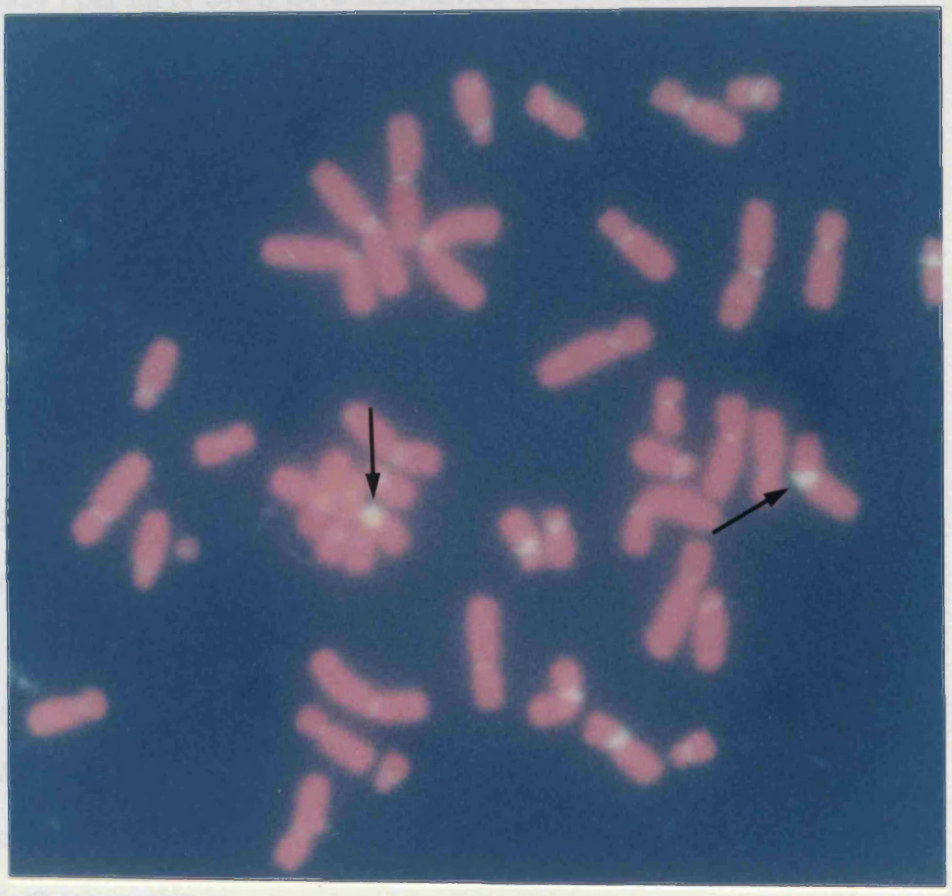
for the quality of biotinylation of the probe was performed. The colour products of the test were evaluated according to their intensities compared to the intensities of control DNAs (X-alphoid probe and salmon sperm DNA, gifts from Dr John Gosden, MRC unit Human Genetics Unit, Edinburgh) at the corresponding concentrations. The test indicated efficient labelling for both oligolabelled and nick translated probes.

The range of probe DNA concentration recommended for in situ hybridisation was 0.2ng/μl to 2ng/μl of hybridisation mixture (Dr John Gosden, personal communication) i.e. 8-80ng per slide if 40μl of hybridisation mixture were used per slide. Eventually, two probe concentrations were selected, 1 and 2.5ng/μl. Signal detection was carried out using FITC-conjugates with one amplification step. Chromosomes were counterstained with propidium iodide. Despite the fact that biotin incorporation test had indicated successful biotin incorporation, signal from the oligolabelled DXZ1 was not detected. In contrast, signal from the nick translated probe was detected with a high hybridisation efficiency. All metaphases examined showed signals on both chromatids of both homologous chromosomes with virtually no background spots. Weaker signals also marked the centromeres of a number of other chromosomes (fig. 6). Similar hybridisation efficiency for DXZ1 was observed at concentrations 20 and 10ng per slide, however, the intensity of the signal was reduced. At a lower concentration of 5ng per slide, hybridisation efficiency was reduced considerably ; in most metaphases signal from only one X chromosome was detected while in a small number of metaphases no signal was detected (data not shown). Similar results were obtained when DXZ1 probe was labelled with digoxigenin and detected with FITC-conjugates (fig. 6a) or labelled with biotin and detected with AMCA-conjugates (blue signal due to AMCA ; fig. 6b).

The standard nick translation protocol used for biotin labelling was also used for digoxigenin-11-dUTP labelling of DXZ1 probe. Probe concentration of 40ng per slide was used during hybridisation. The antibodies, sheep, anti-digoxigenin and donkey, anti-sheep-FITC were initially used for signal detection. Since signal was



a



b

Figure 6: FISH result with the X-specific alpha centromeric probe DXZ1 on metaphase chromosomes from a normal female. Chromosomes are counterstained with propidium iodide and fluoresce red. a. The probe is labelled with digoxigenin. Regions of hybridisation are detected with FITC and appear yellow (arrows) b. The probe is labelled with biotin. Regions of hybridisation are detected with AMCA and appear blue (arrows).
Weaker signals mark the centromeres of a number of other chromosomes.

not detected, probe and antibody concentrations were altered. Different antibodies were also used (section 2.2b of this chapter). Since these modifications did not lead to successful signal detection, a combination of new antibodies (1. mouse, anti-digoxigenin, 2. rabbit, anti-mouse-FITC, 3. goat, anti-rabbit-FITC) and a modified nick translation protocol (shown in table 5) were introduced. Specific signal from DXZ1 was not detected, unspecific signals, however, along the chromosomes (data not shown) indicated that the new antibodies were probably functional, but the size of the labelled probe was too small. Therefore, the amount of DNAase I in the nick translation reaction was reduced (from 5000 to 500pg per μg of DNA). Two additional nick translation reactions with and without addition of TTP (and using 200pg of DNAase I per μg of DNA) were also carried out. Probe concentration was 40ng per slide and detection was carried out using the new antibodies. Signals were detected in all three cases with a high hybridisation efficiency (fig. 6a).

HX79 was the first cosmid to be hybridised in situ. As a cosmid it contained, in addition to the unique sequence, interspersed repetitive sequences, such as SINEs (small interspersed repetitive elements, e.g. Alu-elements) and LINEs (large interspersed repetitive elements, e.g. L1-elements) which had to be suppressed. Therefore, 140fold excess of competitor total human DNA was used. Signal was detected using FITC-conjugates, with two amplification steps. High background did not allow definite localisation of HX79, although there was indication that it mapped to Xq22-q23 (fig. 7a). Increasing the amount of total human DNA to 1400 fold excess during competition eliminated the background and led to unequivocal localisation of HX79 to Xq22-q23 (fig. 7b). The use of 1:1000 fold excess total human DNA also proved sufficient and was used in subsequent experiments for the localisation of other cosmids.

Signal for cosmid HX79 was also detected using AMCA-conjugates, with two amplification steps. Concentrations of AMCA-avidin at 1, 5, 6 and 7μg/ml of blocking buffer, previously used to successfully detect signal from DXZ1 probe (fig. 6b), were insufficient for the detection of signal from HX79. Signal of satisfactory intensity was

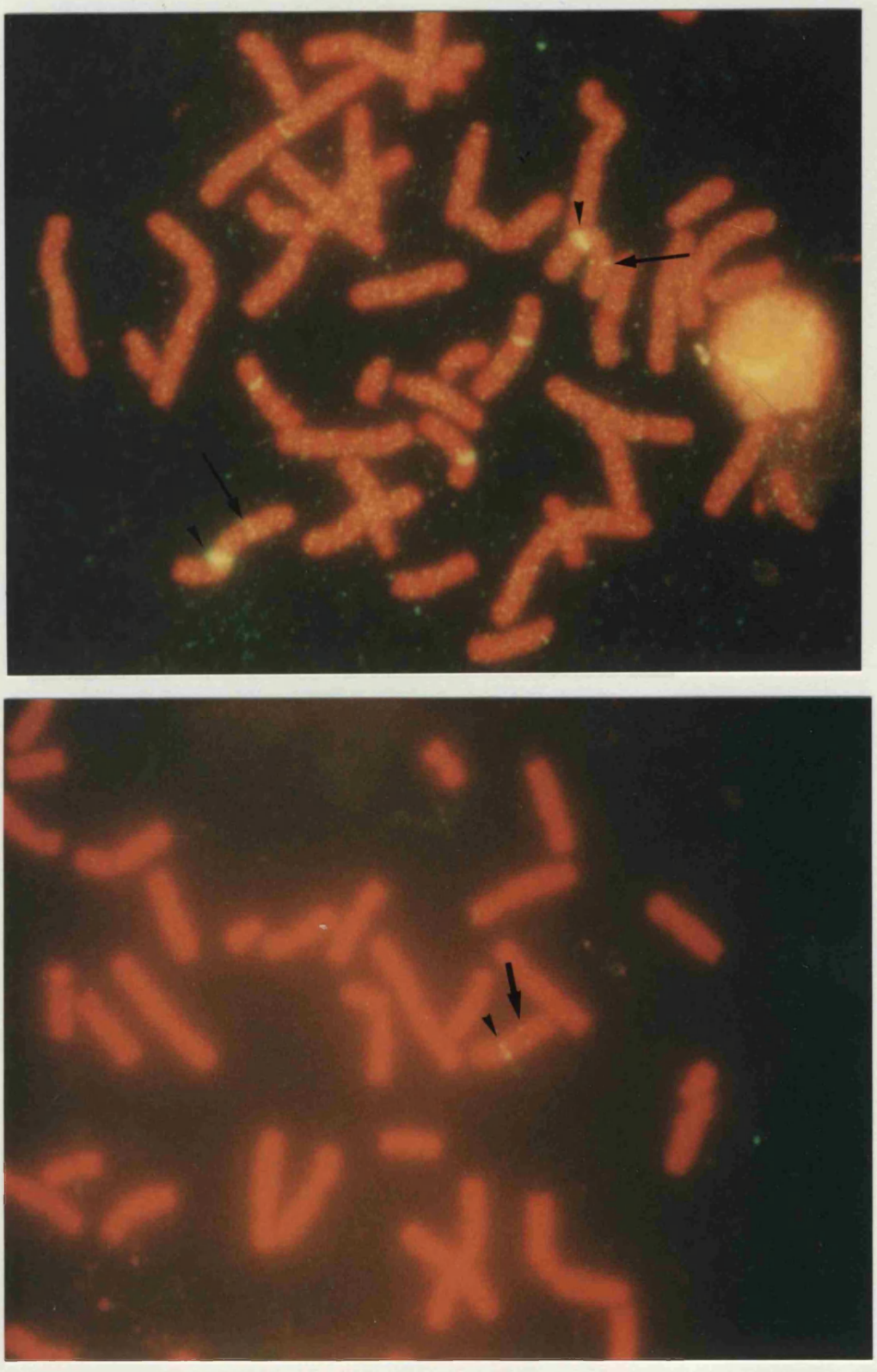


Figure 7: FISH result with the X-specific cosmid HX79 using different competition conditions. HX79 was labelled with biotin and detected with FITC following two amplification steps. a: High background observed following preannealing of cosmid DNA with 140fold excess of competitor total human DNA. Signals from HX79 are not clearly distinguishable (arrow) b: High background was eliminated when total human DNA was increased to 1400fold excess. Signal from cosmid HX79 is unequivocal (arrow). Signals from DXZ1 are shown with an arrowhead.

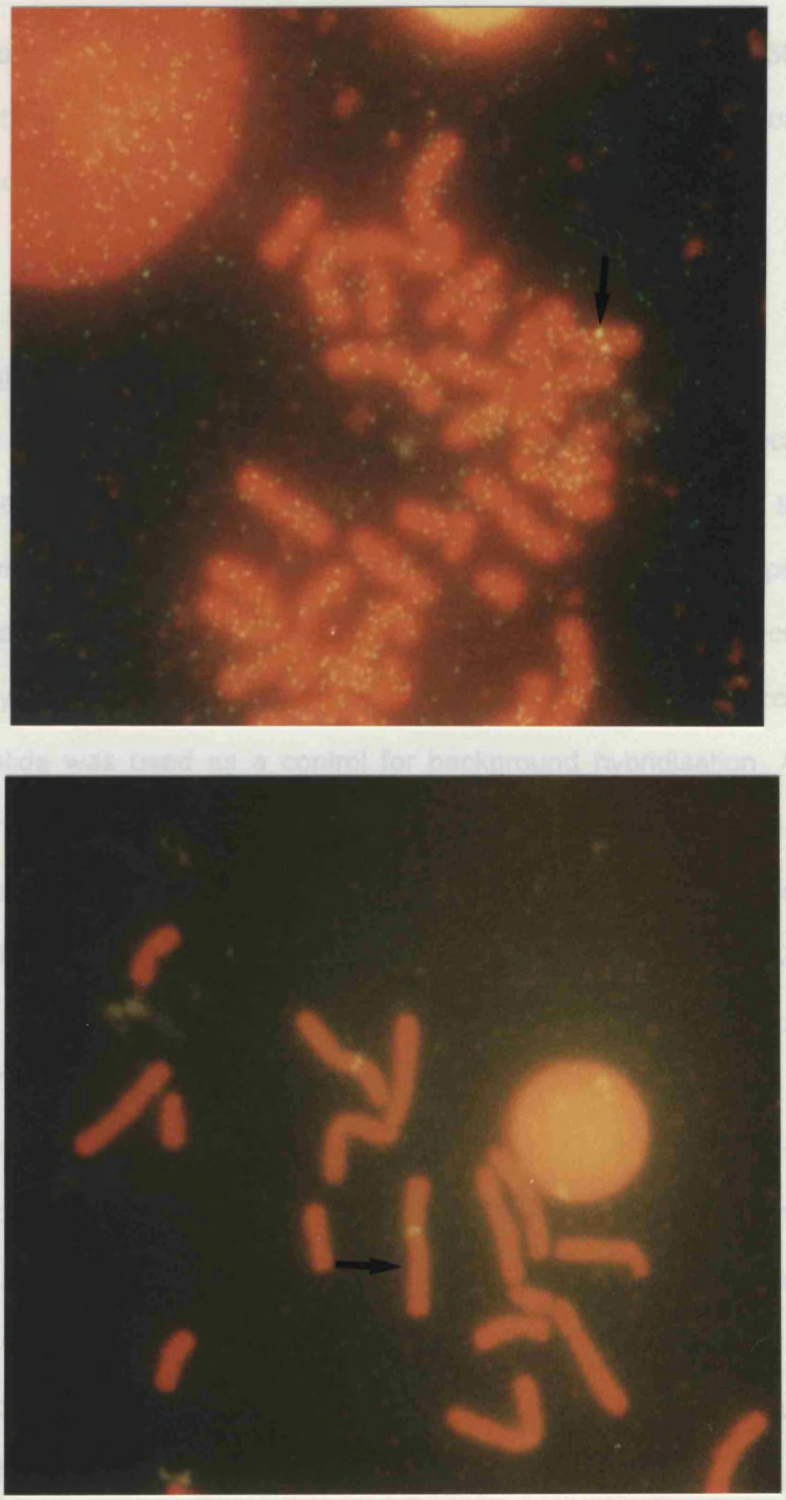


Figure 8: Detection of cosmid HX79 following application of four (a) and two (b) layers of FITC-conjugated antibodies. a: Detection with four layers of antibodies enhanced the intensity of the signal from cosmid HX79 (arrow), but also unspecific signals from background hybridisation. b: Detection with two layers of antibodies was efficient for repetitive probe DXZ1, but not for the cosmid.The signal from HX79 is hardly visible (arrow). Competition was as in fig. 7b.

obtained, however, when the concentration was increased to 20 and 30μg/ml (fig. $9j_2$). Alternative dilution buffers for AMCA did not improve the intensity of the signal (data not shown).

3.2 Localisation of markers on the X chromosome.

11 cosmids from an X-specific somatic cell hybrid library were sublocalised to X chromosome regions by FISH (fig. 9a-k). Two metaphase spread slides were routinely hybridised with each cosmid and at least 20 metaphase spreads were analysed per cosmid and a number of them were recorded using conventional photography. The X-centromeric probe DXZ1 was used as a positive control and a "no DNA" slide was used as a control for background hybridisation. All cosmids yielded highly specific signals on metaphase chromosomes. More than 90% of all metaphases showed signals on both chromatids of both chromosome homologues. Background spots were virtually absent. Four cosmids, HX5, HX65, HX82 and HX97 were found to map to the proximal half of the short arm of the X chromosome. The remaining 7 were found to map to the long arm of the X chromosome, with HX10, HX60 and HX100 mapping to the proximal half and HX14, HX15 and HX46 to the distal half. Localisation was estimated relative to the distance of the signal from the centromere. This gave a broad map position. Although chromosome banding was not performed, a tentative map position for each cosmid is given in table 7. However for some of the cosmids, this map position has been confirmed using other methods. This information is also included in the same table.

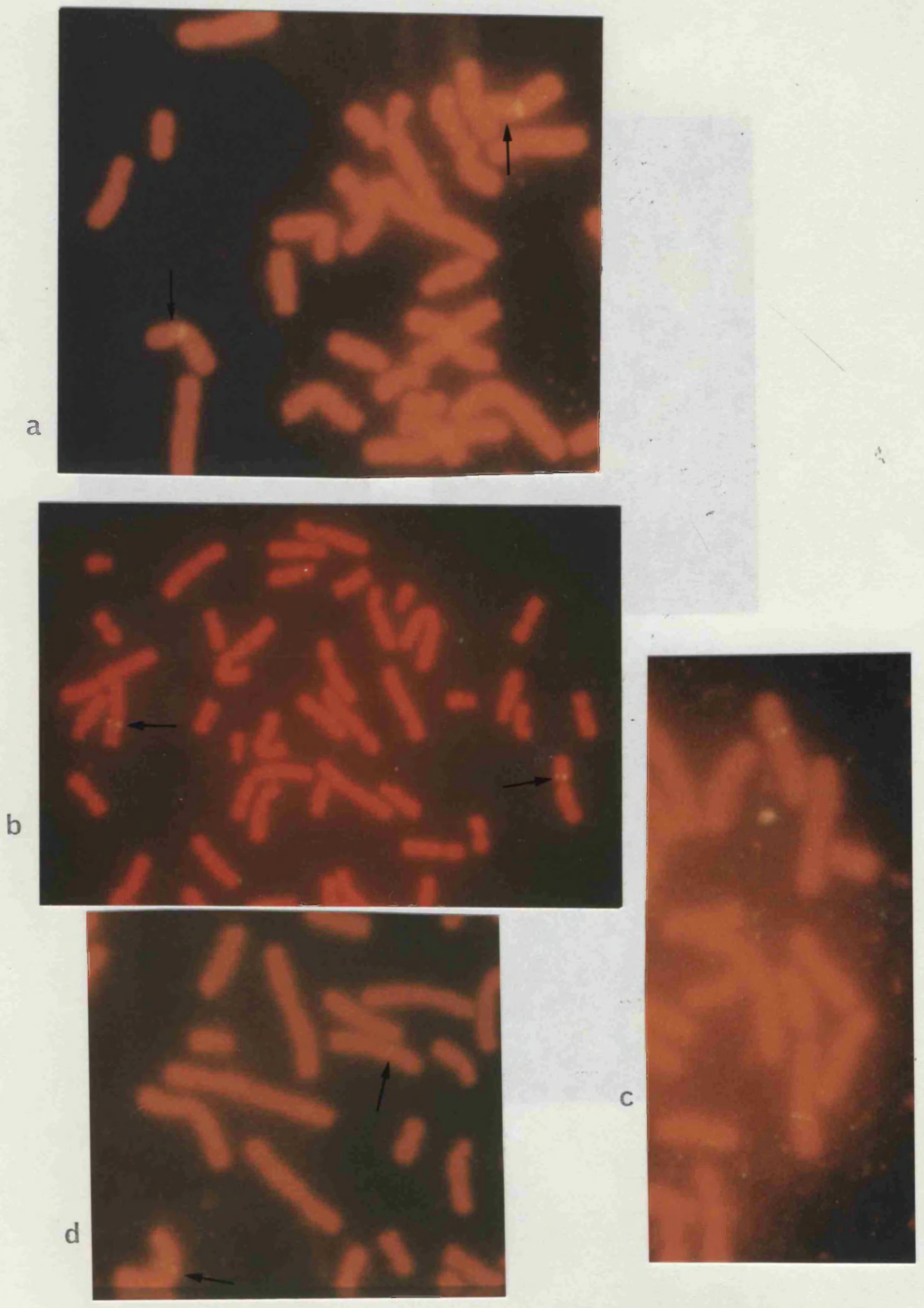


Figure 9: Sublocalisation of 11 X-specific cosmids to X chromosome regions by FISH. Cosmids in **a, c, d, e, g** and **j** were labelled with **biotin** and cosmids in **b, f, h, i and k** were labelled with **digoxigenin.** Signals from all cosmids were detected with FITC-conjugates except HX79 in j_2 which was also detected with AMCA-conjugates. **a:** Cosmid HX97 **b:** Cosmid HX65 **c:** Cosmid HX82 **d:** Cosmid N15. All four cosmids in a, b, c and d map in the proximal half of the short arm of the X chromosome (black arrows) with N15 very close to the centromere.

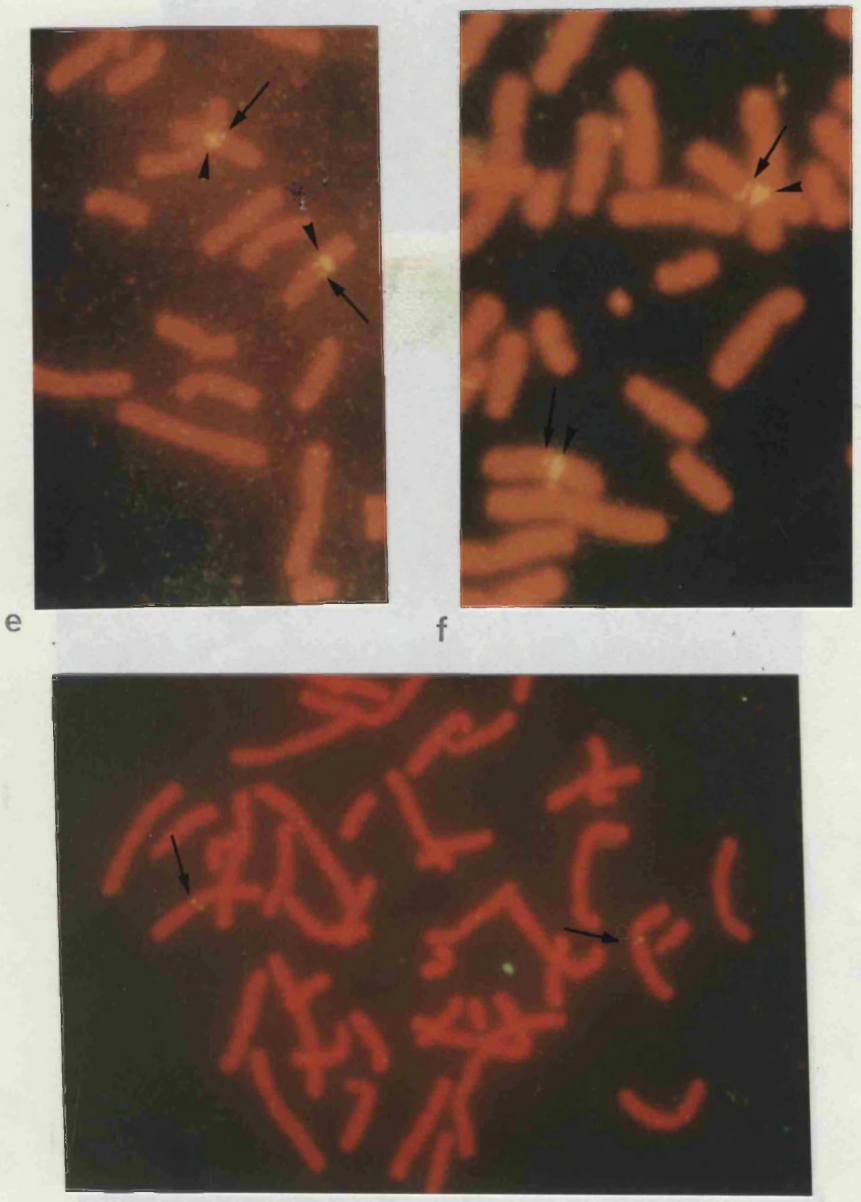


Figure 9 (continued): e: Cosmid HX10 f: Cosmid HX60 g: Cosmid HX100. All three cosmids in e, f and g map in the proximal long arm of the X chromosome (arrows). Centromeric probe DXZ1 in e and f is also shown (arrowhead).

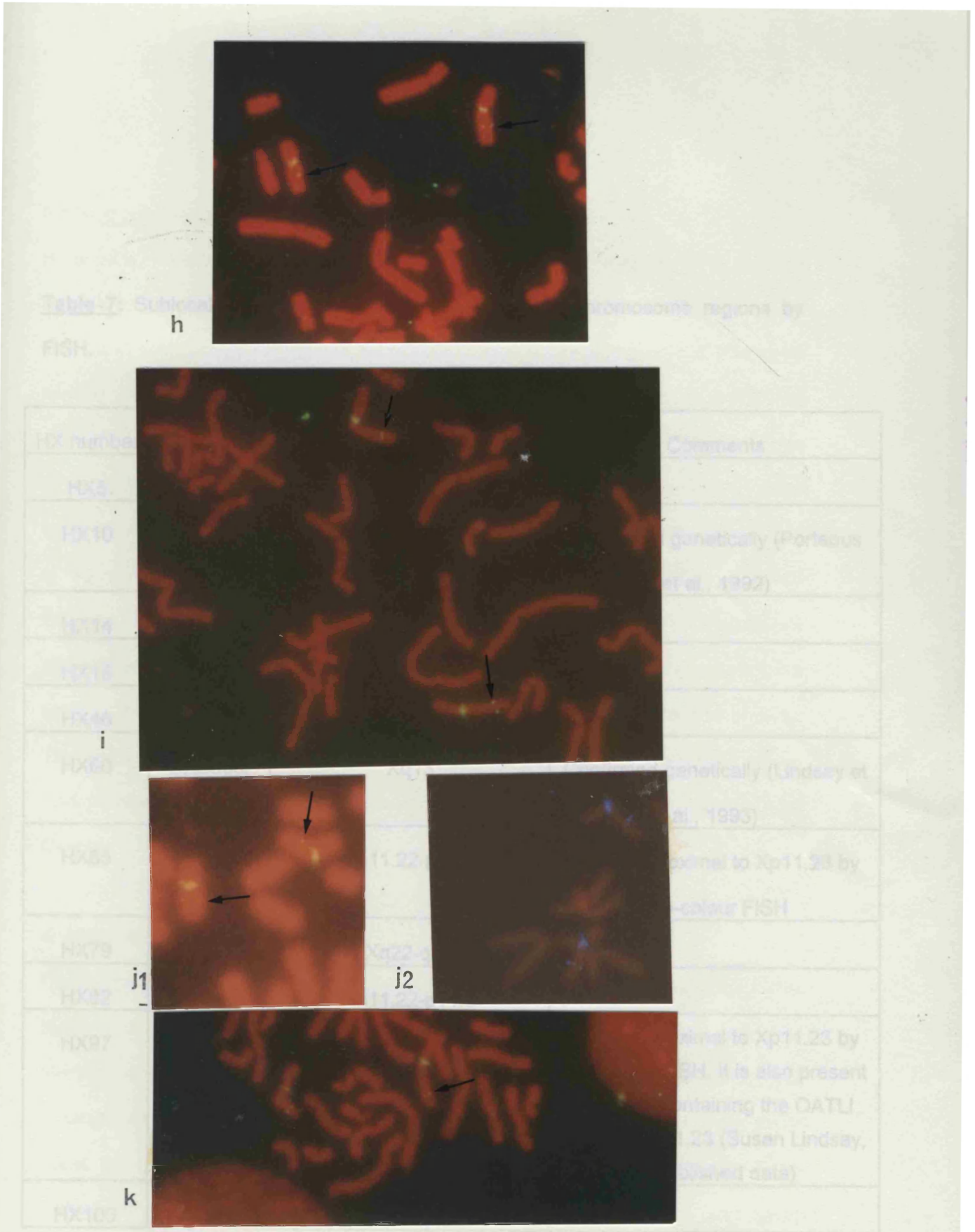


Figure 9 (continued): h: Cosmid HX14 i: Cosmid HX46 j1: Cosmid HX79 (FITC detection) j2: Cosmid HX79 (AMCA detection) k: Cosmid HX15. All four cosmids in h, i, j and k map in the distal half of the long arm of the X chromosome.

Table 7: Sublocalisation of 11 X-specific cosmids to X chromosome regions by FISH.

HX number	D number	Suggested location	Comments
HX5	DXS976	Xp11.1-p11.22	
HX10	DXS554	Xq12	Confirmed genetically (Porteous et al., 1992)
HX14		Xq22-q23	
HX15		Xq25-q26	
HX46		Xq25-q26	
HX60	DXS566/ DXS441	Xq13	Confirmed genetically (Lindsay et al., 1993)
HX65		Xp11.22-p11.3	Mapped proximal to Xp11.23 by two-colour FISH
HX79	DXS571	Xq22-q23	
HX82	DXS573	Xp11.22-p11.3	
HX97	DXS722	Xp11.22-p11.3	Mapped proximal to Xp11.23 by two colour FISH. It is also present in a YAC containing the OATLI locus in Xp11.23 (Susan Lindsay, unpublished data)
HX100		Xq12-q13	

3.3 Two colour fluorescence in situ

8 cosmids, selected because of their map position on proximal Xp were ordered relative to the reference markers L1.28 and HX74 by two colour FISH. L1.28 was the distal (containing sequences from DXS7) and HX74 the proximal reference marker (containing sequences from DXS426 ; HGM11, 1991). The map position of cosmids L1.28 and HX74 in proximal Xp was confirmed (fig. 10a-b) before their use as reference markers in the FISH experiments. Cosmids HX4, HX20 and HX91 were found to map distal to DXS7 (fig. 11a-c) whereas cosmids HX39, HX65, HX97 and HX99 were found to map proximal to HX74 (fig. 11d-g). Signal from cosmid HX43 overlapped with signal from the distal reference marker L1.28 in 10 metaphase chromosomes (fig. 12a-c). However, in a further 14 metaphase chromosomes the signal from HX43 appeared to be distal to that from DXS7 (data not shown). Therefore, it was not possible to resolve its precise position by this method alone. The methods used to resolve the position of HX43 are the subject of the next chapter.

4. Discussion.

Unique sequences, chromosomal subregions, or entire genomes can be specifically highlighted in metaphase or interphase cells by FISH. The technique can be used to identify chromosomes, detect chromosomal abnormalities, determine the chromosomal location of specific sequences or directly order genomic clones. FISH plays an increasingly important role in a variety of research areas, including cytogenetics, prenatal diagnosis, tumor biology, gene amplification and gene mapping. In this study the technique was used for the sublocalisation of 11 X-specific cosmid clones and for the ordering of 8 cosmids, mapping to the proximal short arm of the X chromosome, relative to reference markers DXS7 and DXS426.

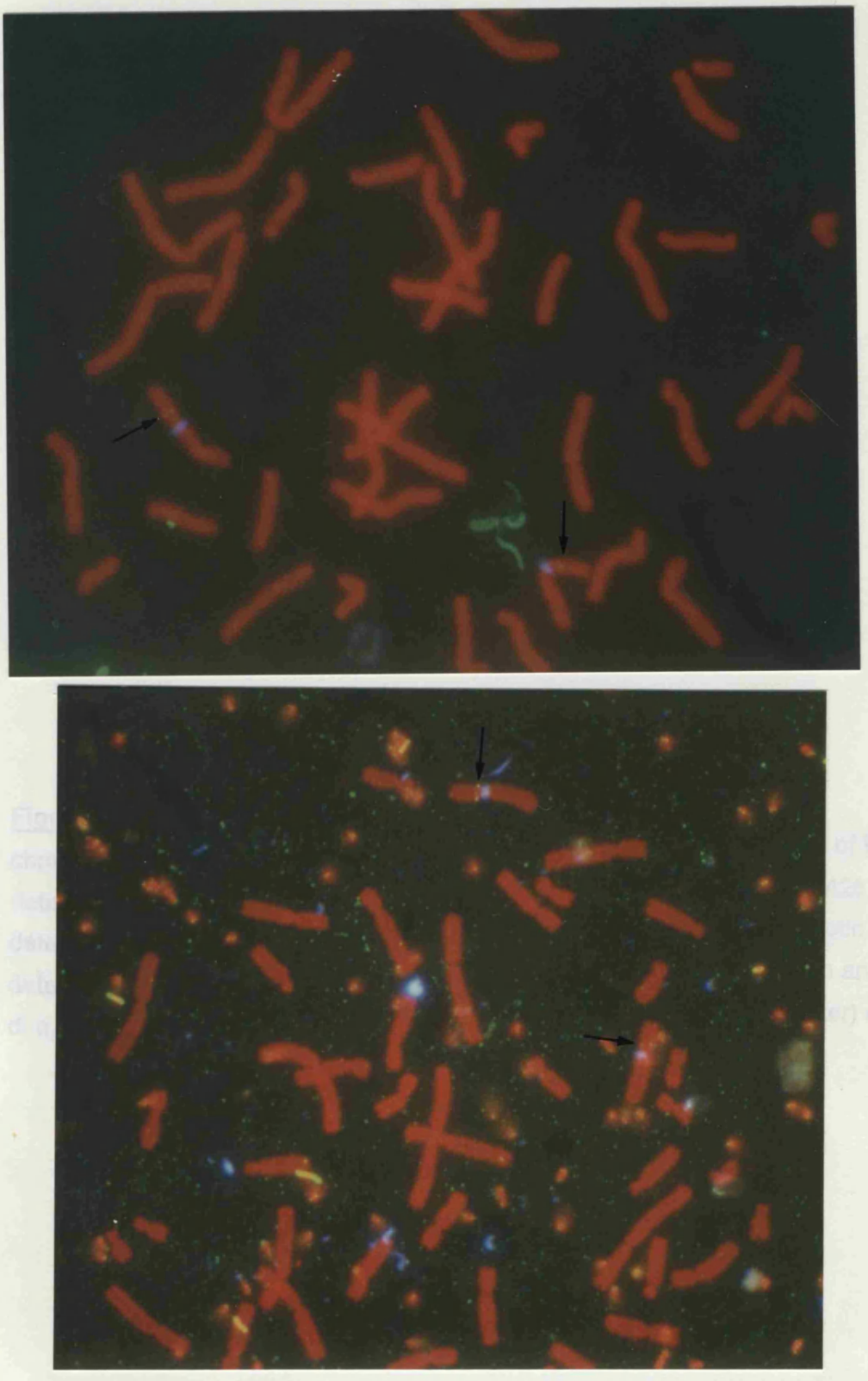


Figure 10: Verification of the map position of two reference markers by FISH prior to their use in two colour FISH experiments The markers were labelled with digoxigenin and detected with FITC. a: Signals from the distal reference marker, cosmid L1.28 (containing sequences from the DXS7 locus) (arrow). b: Signals from the proximal reference marker HX74 (containing sequences from the DXS426 locus) (arrow).
The centromeric probe DXZ1 marking the X was labelled with biotin and detected with AMCA (blue).

Figure 11: Order of X-specific cosmids mapping in the proximal short arm of the X chromosome relative to reference markers L1.28 (DXS7) and HX74 (DXS426) as determined by two colour FISH. Cosmids in question were labelled with biotin and detected with AMCA and reference markers were labelled with digoxigenin and detected with FITC. Cosmids in a, b and c map distal to L1.28 (distal marker) and in d, e, f and g proximal to HX74 (proximal marker).

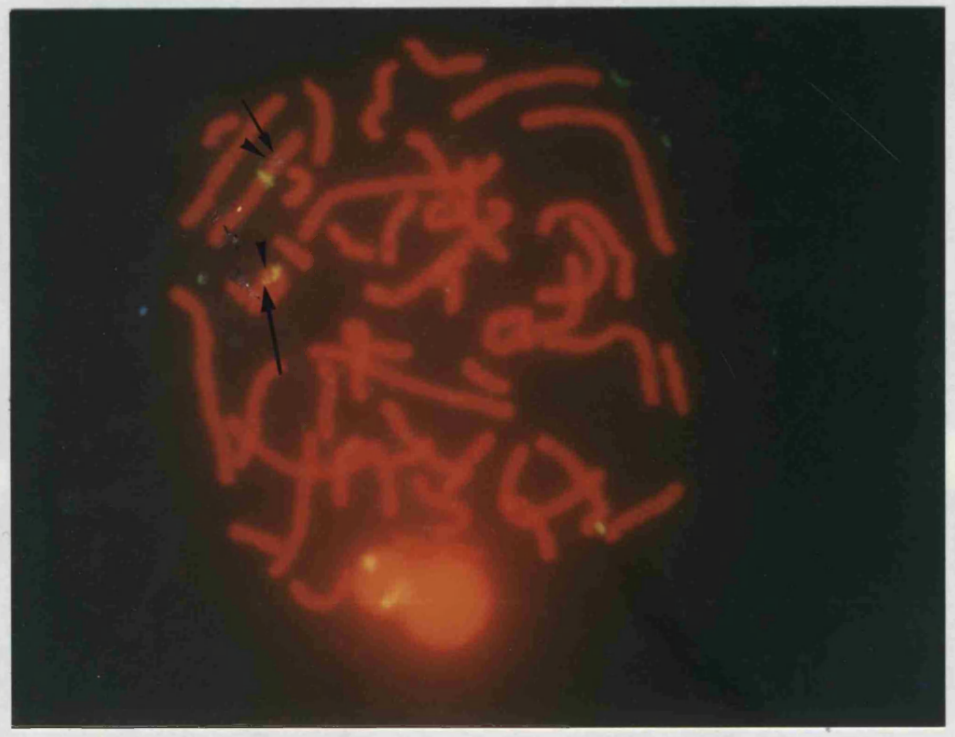


Figure 11a: Cosmid HX4 (blue signal - see arrow) maps distal to L1.28 (yellow signal-see arrowhead). Simultaneous visualisation of FITC and AMCA was achieved by filter block interchange and double exposure of the same metaphase

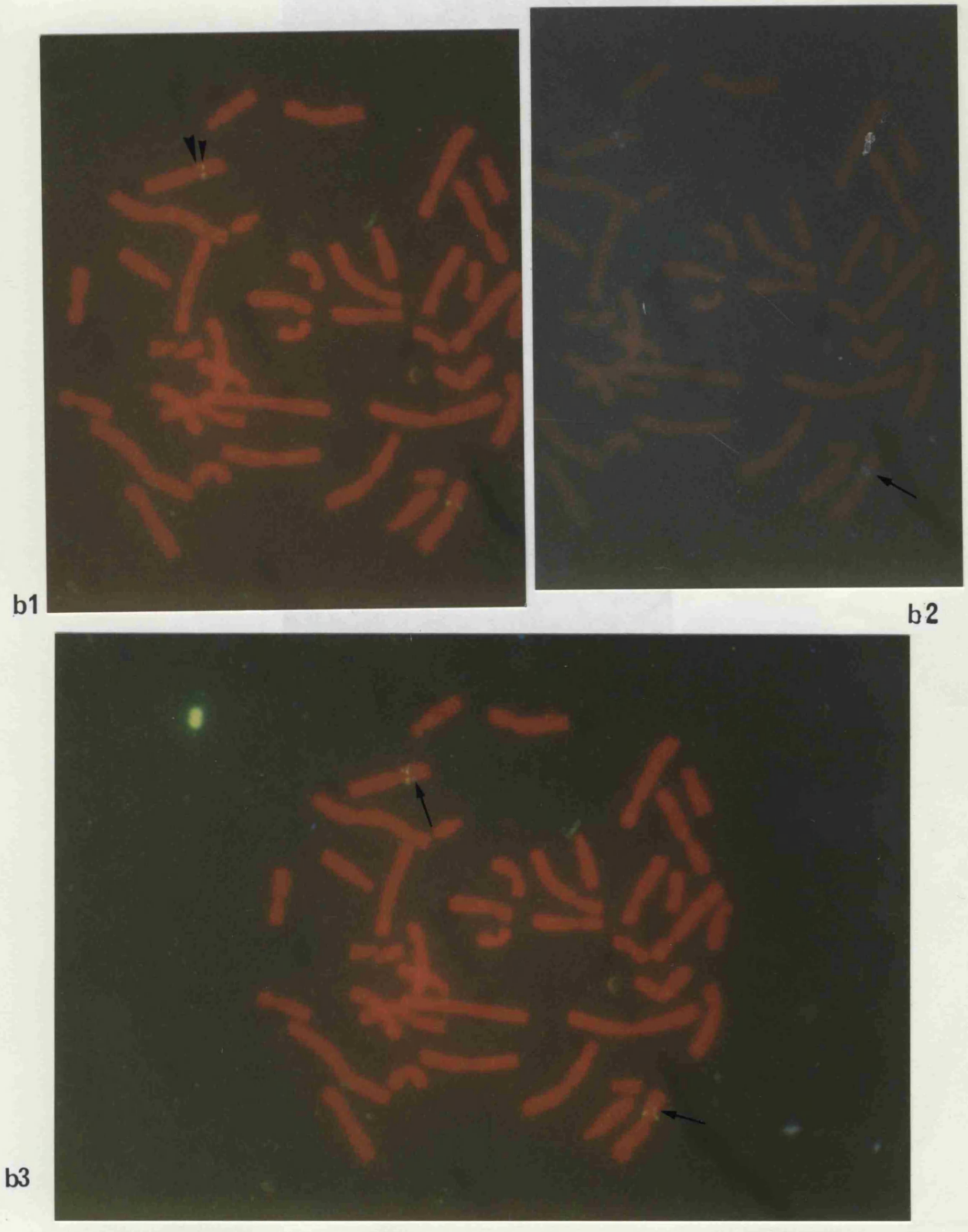


Figure 11b: Co-hybridisation of cosmid HX20 with both reference markers b1: Signals from reference markers L1.28 and HX74 (arrowheads) visualised using the FITC filter b2: Signal from cosmid HX20 in the same metaphase as b1 visualised using the AMCA filter b3: Double exposure of the same metaphase. Cosmid HX20 maps distal to both reference markers (arrows).

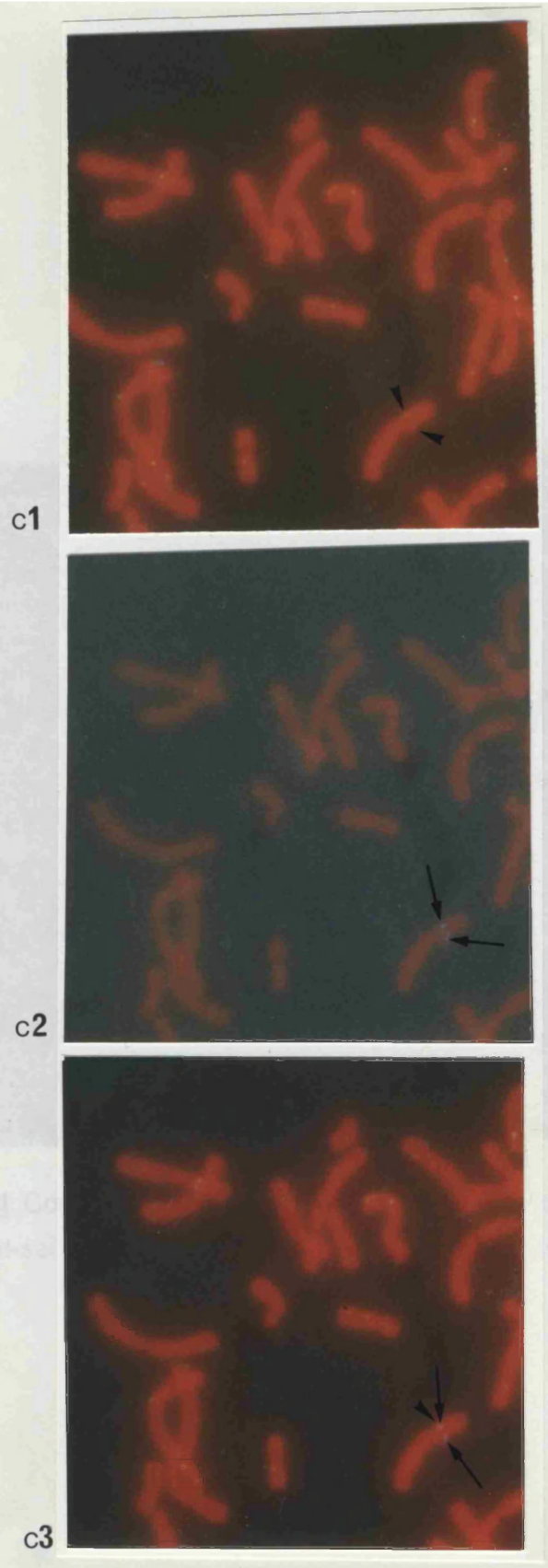


Figure 11c: Cosmid HX91, c1: Reference marker L1.28 (arrowhead) visualised with FITC filter c2: Cosmid HX91 visualised with AMCA filter (arrow) c3: Double exposure of the same metaphase. Cosmid HX91(blue signals-see arrowheads) maps distal to L1.28 (yellow signal-see arrow).

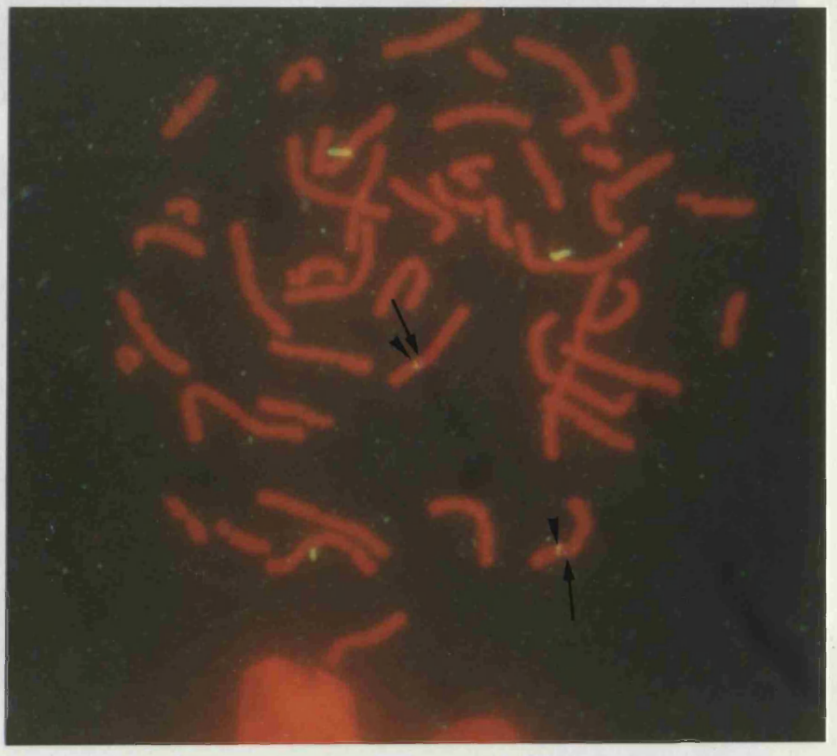
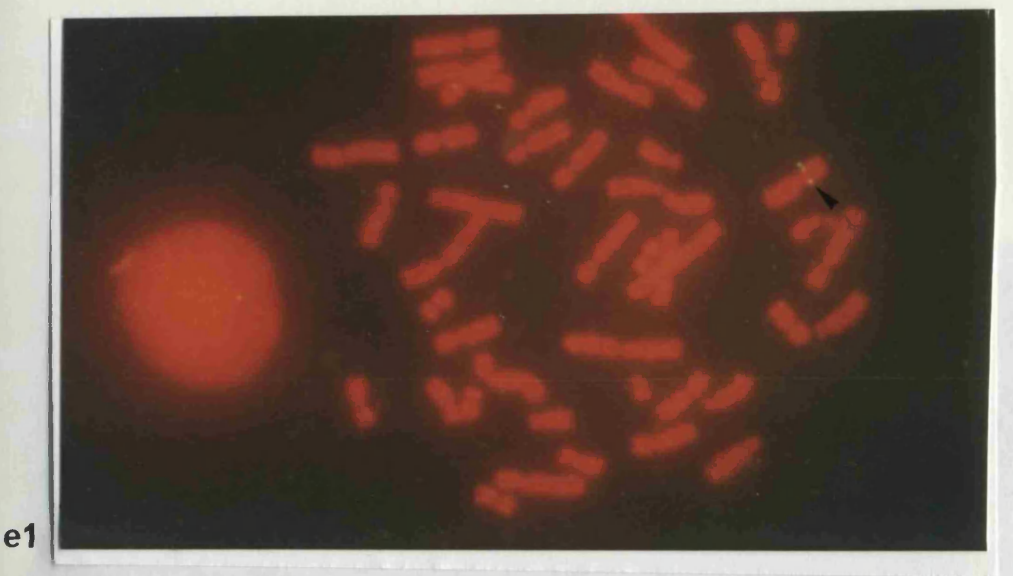
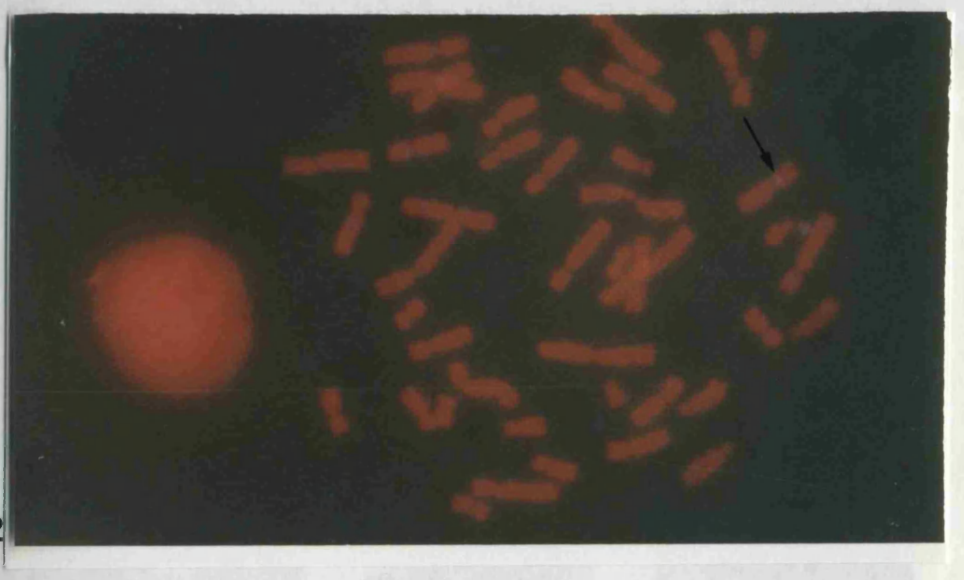


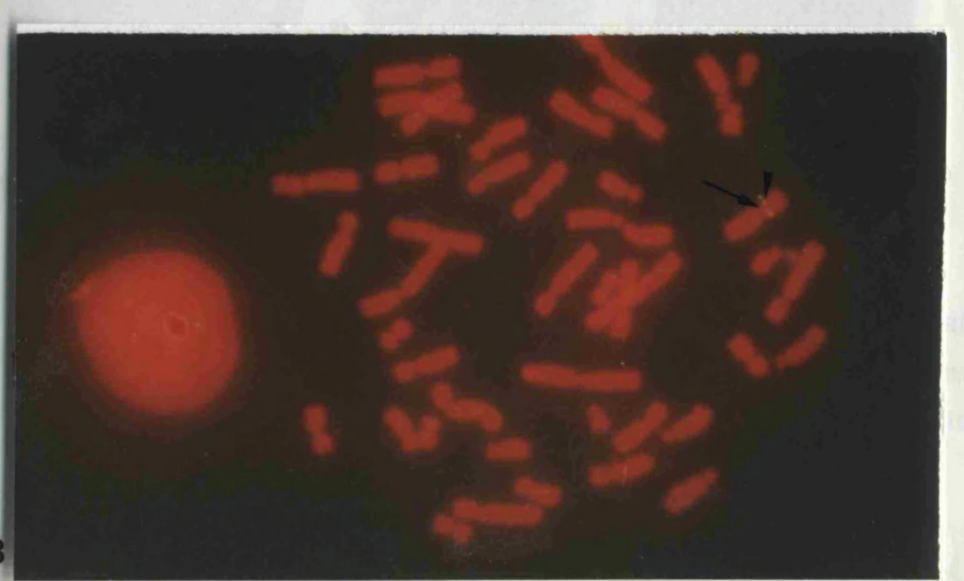
Figure 11d: Cosmid HX39. Double exposure showing localisation of the cosmid (blue signal-see arrow) proximal to HX74 (yellow signal-see arrowhead).



e1



e2



e3

Figure 11e: Cosmid HX65, e1: Reference marker HX74 (yellow signal-see arrowhead) visualised with FITC filter e2: Cosmid HX65 visualised with AMCA filter (blue signal-see arrow) e3: Double exposure of the same metaphase showing localisation of the cosmid proximal to HX74.

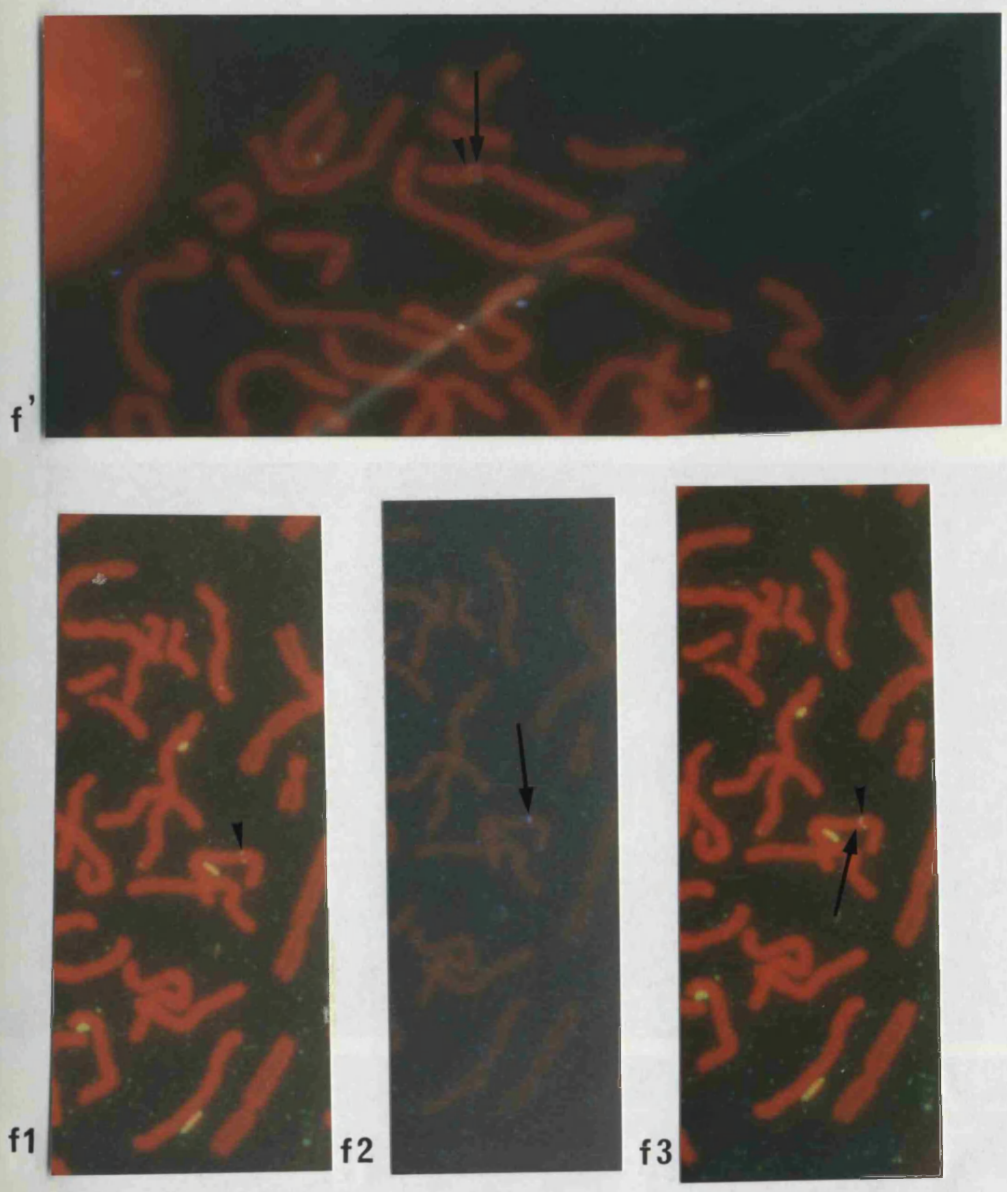


Figure 11f: Cosmid HX97, f': Cosmid HX97 (arrow) maps proximal to reference marker L1.28 (arrowhead ; L1.28 is distal to HX74) f1: Marker HX74 visualised with FITC filter (arrowhead) f2: Cosmid HX97 visualised with AMCA filter (arrow) f3: Double exposure showing localisation of the cosmid proximal to HX74.

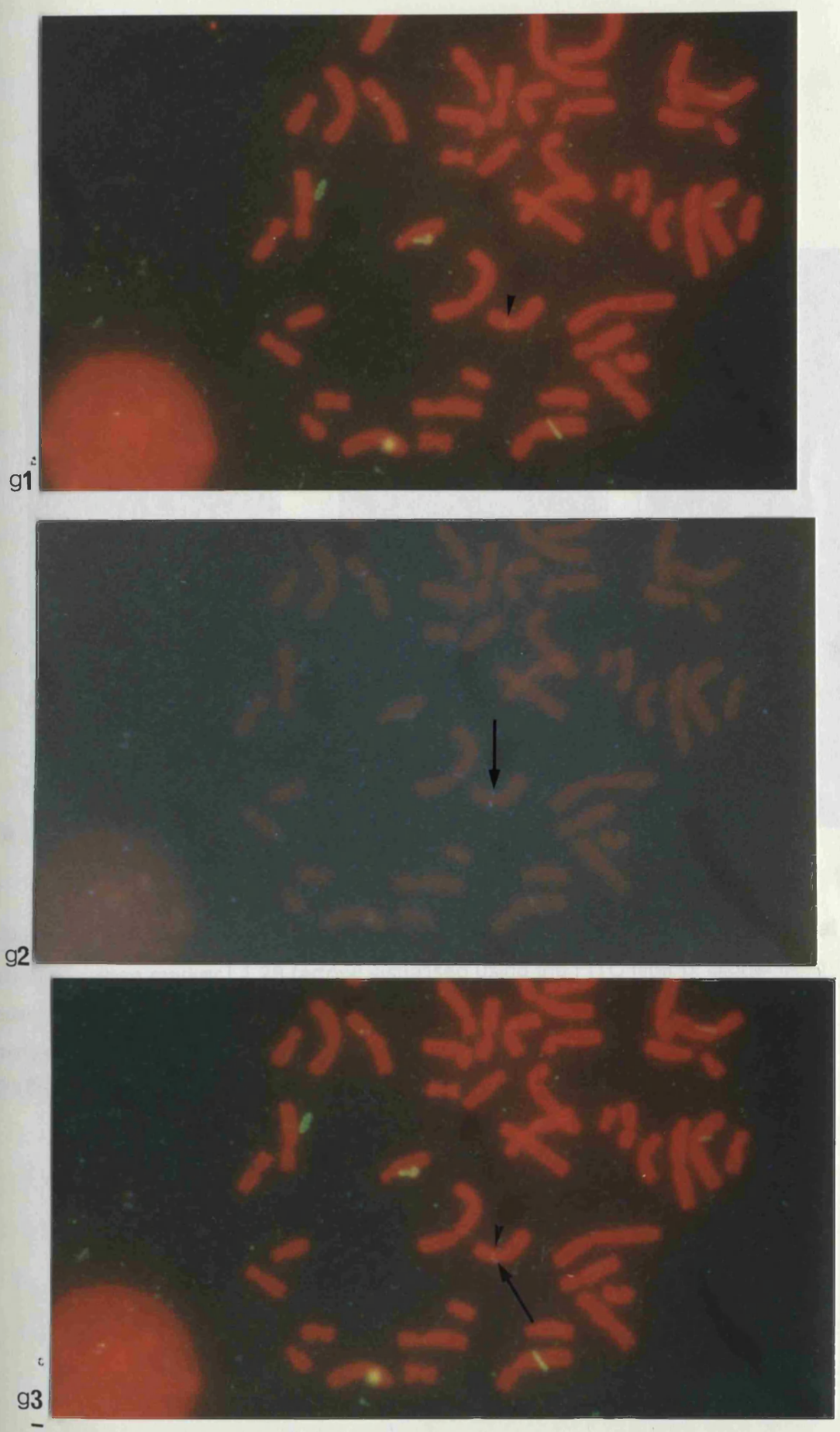


Figure 11q1.2.3: Cosmid HX99 maps proximal to HX74 (g3: double exposure)

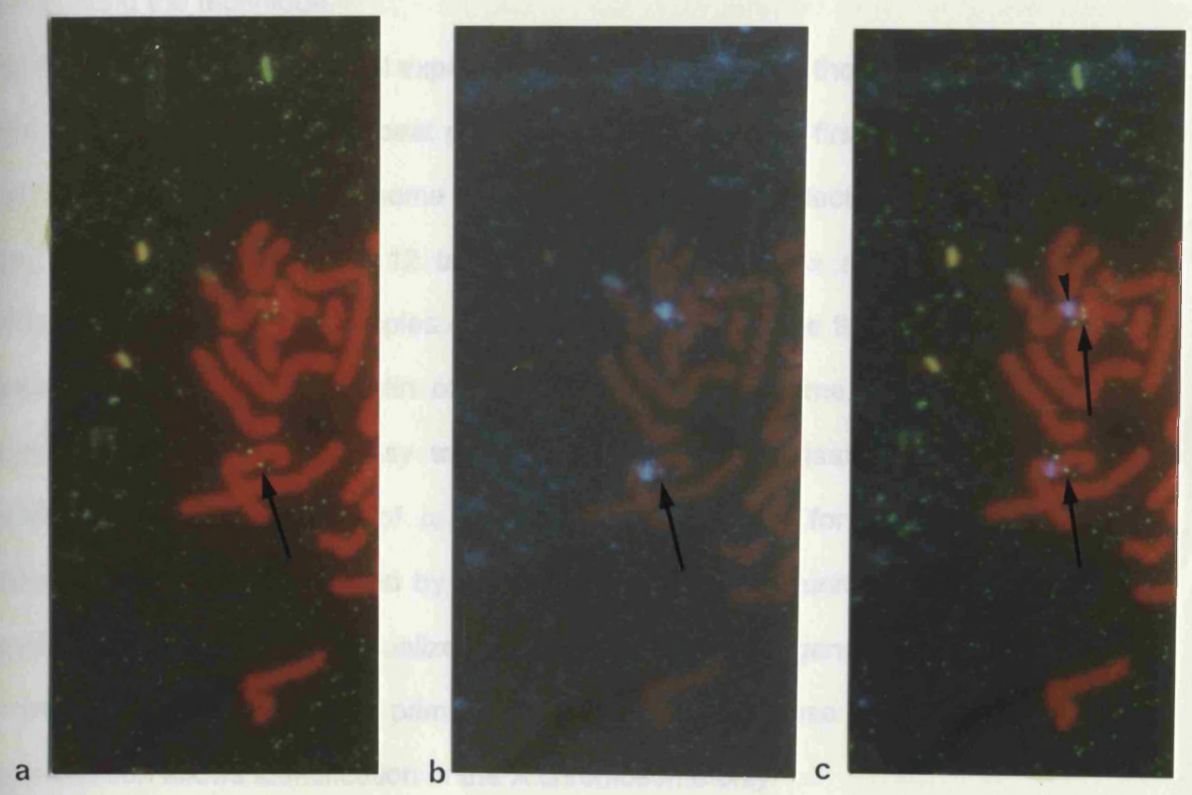


Figure 12: Co-hybridisation of cosmid HX43 with reference marker L1.28. a: Signal from reference marker L1.28 (arrow) visualised using the FITC filter b: Signal from cosmid HX43 visualised using the AMCA filter c: Double exposure of the same metaphase. Signal from cosmid HX43 overlaps with signal from cosmid L1.28. DXZ1 is labelled with biotin and detected with AMCA (arrowhead).

These markers are known to be tightly linked to the X-linked Retinitis Pigmentosa form 2.

4.1 Testing the technique.

As there was no prior practical experience of FISH mapping in the laboratory, DXZ1, the X-specific centromeric repeat probe was selected as the first probe for testing the technique. The X chromosome higher-order repeat unit detected by DXZ1 is 2kb long, and is comprised of 12 tandemly repeated 171bp α satellite monomers (Willard et al, 1986). 5000 copies of this 2kb repeat constitute the α satellite DNA, present in the heterochromatin of the human X chromosome. Since the target region is large, DXZ1 is easy to detect by in situ hybridisation. Specificity of centromeric tandem arrays of α satellite DNA monomers for individual human chromosomes is characterized by the length of the repeat unit, by the particular restriction enzyme used to visualize the repeat in digests of genomic DNA (BamH1 in case of DXZ1), and by its primary DNA sequence. The use of DXZ1 in in situ hybridisation allows identification of the X chromosome only.

FISH of DXZ1 labelled by nick translation resulted in signal detection with a high hybridisation efficiency. In contrast, random oligonucleotide labelling gave negative results even when the biotin incorporation test had indicated efficient labelling (data not shown). This was probably due to the size of the probe molecules after oligolabelling which is an important requirement for a good in situ hybridisation probe. This should be between 100 and 500 nucleotides in length. Primer extension of genomic fragments often results in probe molecules of a higher size-range. In a nick translation reaction the required range can be easily obtained by adjusting the DNAase concentration thus making it preferable technique for labelling. The amount of DNAase for nick translation is critical. In this study when 5ng of DNAase was used per μg of DXZ1 DNA, unspecific signals were observed on the chromosomes. This suggested that the probe molecules were of a lower size-range than required

for in situ hybridisation. Reducing the amount of DNAase to 0.5ng per μg of DNA resulted in specific signal detection (fig. 6a).

Nick translation protocols use dTTP as an additional nucleotide in the nucleotide mixture during labelling with digoxigenin. In this study addition of dTTP in the nick translation reaction did not prove essential since its presence or absence gave identical in situ hybridisation results. The simplest nick translation reaction (using 0.2ng of DNAse I and 2units of Polymerase I per μg of DNA) for both biotin and digoxigenin labelling proved efficient throughout the use of the FISH technique.

The combination of the antibodies, sheep anti-digoxigenin and donkey anti-sheep-FITC proved unsuitable for the detection of signal from digoxigenated DXZ1 probe. Although the reason for this remains unclear it may be that the quality of the antibody preparations was not appropriate or that signal amplification was required. The replacement with a different set of antibodies, (anti-digoxigenin in mouse, anti-mouse-FITC in rabbit and anti-rabbit-FITC in goat) provided an amplification step and resulted in successful signal detection (fig. 6a).

4.2 Localisation of markers on the X chromosome.

The FISH technique was applied to sublocalise 11 cosmids from an X-specific somatic cell hybrid library to X chromosome regions. Competition was performed with total human DNA to remove repetitive sequences. This method which has been called chromosomal in situ suppresion (CISS) hybridisation (Lichter et al, 1988) eliminated the need to identify and isolate single-copy subclones from each large genomic DNA fragment (Landegent et al, 1985), thus, substantially decreasing the effort required to map a genomic clone. If interspersed repetitive elements (IRS) are not removed from the probe, hybridisation signals due to the ubiquitous presence of IRS throughout the genome, are distributed over the whole chromosome complement. Insufficient competition (1:140) during the initial application of the

technique resulted in such a distribution of hybridisation signals. Increasing the amount of competitor DNA (competition 1:1400 or 1:1000) eliminated the problem.

A range between 40 and 120ng of cosmid probe per slide has been successfully used in this study which indicates that large amounts of DNA are not required for signal production. The use of 1μl (instead of the initial 15μl ; see section 1.5.3 of chapter 2) of labelled cosmid probe in the hybridisation mixture proved useful since 25μl of labelled probe yielded sufficient probe to hybridise 25 slides.

Application of two layers of FITC-conjugated antibodies resulted in an almost invisible signal (fig. 8b) from the biotinylated cosmid probe whereas application of four layers intensified the specific signal but also unspecific signals from background hybridisation (fig. 8a). Application of three antibody layers resulted in the detection of a discrete signal with virtually no background spots. The same result was obtained from digoxigenated probes after the application of only two antibody layers which indicates a higher sensitivity of the digoxigenin detection system.

The detection system for digoxigenin was direct and comprised of three steps. The detection system for biotin was indirect and comprised of five steps including alternating layers of antibodies conjugated with a reporter group (biotinylated, anti-avidin) and antibodies conjugated with a fluorochrome (this detection procedure is called sandwich amplification). Therefore, digoxigenin-labelled probes had the advantage of a considerably less-time consuming detection. However, the ability of amplifying the signal intensity from biotinylated probes by alternating layers of antibodies is advantageous for the detection of small or weakly hybridizing probes.

Signals from the cosmids were visualised using conventional fluorescence microscopy and a filter set which allowed the simultaneous visualisation of FITC for signal detection and propidium iodide for chromosome staining. The map position of each cosmid could be roughly assigned by visual inspection. Cytological banding was not performed since fine mapping was desired only for cosmids mapping in the proximal short arm of the X chromosome. This was achieved by two colour FISH.

For the remaining cosmids, localisation was estimated relative to the distance of the signal from the centromere which gave a broad map position. The localisation of two of these cosmids (HX10 and HX60) was confirmed genetically by multipoint linkage analysis using other polymorphic markers from the region (Porteous et al., 1992 ; Lindsay et al., 1993).

In studies which use digital imaging microscopy to collect and analyse the hybridisation data, the map position of the probe can be expressed as the fractional length (FL) of the total chromosome relative to an arbitrarily chosen fixed reference point e.g. pter which is then designated FLpter (Lichter et al., 1990).

4.3 Two colour fluorescence in situ hybridisation.

Two colour FISH was used to order 8 cosmids, selected because of their map position in proximal Xp, relative to the markers HX74 (containing primer sequences for DXS426) and L128 (containing sequences for DXS7). Linkage analysis in families listed in the Centre d' Etude du Polymorphisme Humain (CEPH) had indicated that the genetic distance between DXS7 and DXS426 is 11.6cM (Coleman et al., 1991). Multiply informative meioses in two RP2 families (see introduction) had indicated that the most likely location of the RP2 gene was between these two markers. It was, therefore, considered necessary to order cosmids mapping in this region relative to these markers in order to aid the finer localisation of the RP2 gene. Lack of highly informative markers between DXS7 and DXS426 was an additional reason to perform this task. Cosmids in question were labelled with biotin and detected as a blue signal (due to AMCA fluorochrome) and reference markers were labelled with digoxigenin and detected in yellow (due to FITC fluorochrome). The same preparation of labelled reference markers was used for all experiments thus maintaining a stable parameter. AMCA gave a signal of weak intensity compared to the signal produced by FITC and a 7-fold concentration (30μg/ml final concentration of AMCA-avidin D compared to 4μg/ml of FITC-avidin DCS) was needed to produce a signal of satisfactory intensity. It was also more

prone to fading than FITC which was another disadvantage since the time interval between visualising and recording the signal had to be istantaneous.

FITC and AMCA were visualised using two filter sets. Image shift was not observed after double exposure during photography. The image shift can occur when filters are changed and it is detectable after double exposure of photographic films or even when digital images are electronically overlayed. The use of double band-pass filters offers a solution to this problem.

Cosmid order was derived from the order of blue and yellow fluorescent dots on metaphase chromosomes. The position of the blue site above or below the yellow site determined probe order relative to the reference marker. Cosmid HX20 was cohybridised with both reference markers in a three probe labelling and detection experiment (HX20 detected in blue and reference markers in yellow) and a pattern blue-yellow-yellow-centromere was observed mapping HX20 distal to both reference markers (fig. 11b). Simultaneous visualisation of the reference markers on the chromosome allowed a rough estimation of a physical distance of approximately 6Mb between them. This estimation was based on **1)** the knowledge that metaphase mapping has a resolution of 1-2Mb (Trask et al., 1991) **2)** the observation that the two markers were clearly resolved potentially allowing a third marker mapping between them to be resolved and **3)** the consideration that the chromosomes used were not optimally elongated (in this case the distance between the two markers would appear smaller than it actually is).

Cosmid HX97 was mapped relative to both reference markers (separately). The remaining cosmids were first mapped relative to proximal marker HX74 (DXS426). If they were found to map proximal to this marker, they were not further analysed. If they were found to map distal to this marker, further analysis relative to the distal marker L1.28 (DXS7) was undertaken. The map position of cosmid HX43 was not resolved. Signal from this cosmid overlapped with signal from distal reference marker L1.28 in 10 metaphase chromosomes and it appeared distal to it in a further 14 metaphase chromosomes. This indicated that the two markers were probably 1-

2Mb apart. The precise location of HX43 could be resolved by two colour interphase mapping which has a resolution of 50-100kb, although a closely spaced marker of known orientation would be required. It was eventually resolved genetically as discussed in the following chapter.

The cosmids which were sublocalised by single FISH can be used in linkage studies for disease genes mapping to the relevant regions and can contribute to the physical and genetic mapping of the human X chromosome. Cosmids which were more finely mapped to proximal Xp are currently being characterised by linkage in a panel of X-linked RP families allowing a direct comparison of the physical and genetic map of the proximal short arm of the X chromosome.

CHAPTER 4: ISOLATION, CHARACTERISATION AND GENETIC MAPPING OF A MICROSATELLITE FROM COSMID HX43

1. Introduction

1.1 Characterisation of a microsatellite in this study

HX43 was one of the 8 microsatellites containing cosmids which were selected to be ordered relative to reference markers DXS7 and DXS426 known to be flanking the RP2 locus in the proximal short arm of the X chromosome (Coleman et al., 1990). This task was performed using two colour FISH (see chapter 3). Signal from cosmid HX43 in 10 metaphases overlapped with signal from DXS7 while in a further 14 metaphase chromosomes the signal from HX43 appeared distal. Therefore, its precise location could not be resolved in this way. It was critical, however, to resolve the precise location since it appeared to be a candidate marker mapping between DXS7 and DX426. HX43 could, therefore, be useful as a polymorphic marker and had the potential to refine the genetic localisation of the RP2 gene. It could also be used as a starting point for the construction of a physical map of the region. To resolve the location of this marker, isolation, characterisation and genetic mapping of a microsatellite was carried out. Therefore, a mini library from cosmid HX43 was generated by ligation of sequencing vector (pT7T3) with cosmid DNA restricted with a frequent-cutting enzyme (SauIIIA). The library was screened for a subclone containing the microsatellite by hybridisation with a simple oligonucleotide probe (poly dG-dT) and the positive clone sequenced. Using the polymerase chain reaction to amplify DNA within and immediately flanking the repeat block from 46 (92 X chromosomes) unrelated caucasian females, the polymorphism and heterozygosity (allele number, length and frequency) of HX43 microsatellite were established. The position of HX43 marker was then defined by following its inheritance in twelve XLRP families (comprising 91 meiosis) for which linkage data was available for the markers OTC (ornithine transcarbamylase ; Xp21 ; Petty et al., 1991), DXS228 (Coleman et al., 1991), DXS7 (Moore et al., 1992) and DXS426

(Luty et al., 1990). The map position and order of these loci were as follows: Xpter-OTC-DXS228-DXS7-DXS426-cen (Report of the Fourth International Workshop on Human X Chromosome Mapping, 1993).

1.2 Microsatellites for linkage analysis of genetic traits

1.2.1 *General considerations*

Two major advances in molecular biology have made possible the rapid development of highly informative markers for genetic mapping and linkage analysis. First, in 1985, it was discovered that short segments of DNA could be amplified in vitro using DNA polymerase and temperature cycling, by the polymerase chain reaction (PCR ; Saiki et al, 1985 ; Mullis, 1990). Second, in 1989, four laboratories (Weber and May, 1989 ; Tautz, 1989 ; Litt and Luty,1989 ; Smeets et al., 1989) used PCR to demonstrate a high level of polymorphism or allelic variation in the repeat number for simple sequence tandem repeats known as microsatellites.

Microsatellites consist of around 10-50 copies of a 1 to 6 bp motif that can occur in perfect tandem repetition, as imperfect repeats or together with another repeat type (Weber, 1990). Microsatellites occur frequently and randomly throughout most eucaryotic genomes (Hearne et al., 1992) and are named by analogy with the larger minisatellite arrays (Jeffreys et al., 1985) or variable number of tandem repeats (VNTRs ; Nakamura et al., 1987). These arrays are highly polymorphic but are less common than microsatellites and have larger sequence motifs (29-64bp) extending over more than 1kbp, making them less amenable to PCR analysis (Horn et al., 1989).

1.2.2 *Microsatellite types and frequencies*

Beckman and Weber (1992) calculate that a total of 745kbp of human genomic DNA sequence contains an average of one microsatellite every 6kbp. 76% of repeat

types are A, AC, AAAN, AAN or AG in decreasing order of abundance. On average, AC repeats occur every 30kbp (Litt and Luty, 1989 ; Hamada et al., 1982 ; Stallings et al., 1991 ; Beckman and Weber, 1992) distributed equally in 5'- and 3'-untranslated regions and introns, and these are the most common type of microsatellite. Around 80% of human A, AAAN and AAN repeats and 50% of AT microsatellites occur 3' to Alu elements (Beckman and Weber, 1992 ; Economou et al., 1990). Tri- and tetranucleotide repeats occur every 300-500kbp on the human X chromosome (Edwards et al., 1991).

The majority of published human microsatellite sequences have been developed from randomly cloned fragments of total human genomic DNA (Weissenbach et al., 1992). These are detailed in the Genome Database (GDB) sponsored by the Howard Hughes Medical Institute and the John Hopkins University (Baltimore, USA).

1.2.3 *Microsatellite characterisation*

New microsatellites can be cloned directly from total genomic DNA or DNA enriched for specific chromosomes. DNA libraries are made using frequent-cutting enzymes (Saulll a or Alul) and cloned in sequencing vectors. Recombinant clones are then screened for microsatellites with simple repetitive oligonucleotide probes and DNA templates from positive clones are sequenced. For libraries of larger insert size a sequencing primer may be annealed directly to the repeat to determine one of the flanking sequences (Yuille et al. 1991). PCR amplification primers on either side of the repeat are chosen using either a computer programme (Lowe et al. 1990) or simply manual inspection. During primer design, complementary sequences around the 3' bases (to prevent self-annealing) and homologies with repeated sequencies (Alu, Kpn) are avoided and roughly equal proportions of each base are selected.

1.2.4 *Microsatelllte typing*

Following PCR amplification, the allelic system of a microsatellite can be visualised by ethidium bromide staining after nondenaturing gel electrophoresis (Hearne et al.

1992). Greater detection sensitivity is obtained in native gels with silver staining (Todd et al. 1991). These have the advantage of being nonradioactive methods. Both ^{32}P and ^{35}S may be used to label microsatellite products; the former is widely used to label one end of a primer, and either isotope may be incorporated into the product during amplification (Litt, 1991). Products are resolved on DNA sequencing (denaturing) gels.

Gel resolution may be improved by the fluorescent labelling of PCR primers (Edwards et al. 1991 ; Carrano et al. 1989). Fluorescence is quantitative for much of the amplification reaction and genuine alleles can be distinguished by their relative intensities. Four fluorescent dyes are currently available to label PCR primers for high sensitivity laser detection of four different microsatellites. Computer analysis of the output signal from a scanning laser enables the different primer products to be distinguished by their wavelengths even when their sizes overlap. Allele sizes may be determined by the inclusion of fluorescent size standards in each sample. Simultaneous amplification of different loci (multiplex PCR) has been greatly enhanced by use of this fluorescent-labelling technology.

1.2.5 *Genetic mapping*

The international effort to quickly and systematically develop a high-resolution genetic map of the human genome have been greatly facilitated by the development of microsatellite markers. A genetic linkage map of the human genome was constructed by the NIH/CEPH Collaborative Mapping Group (1992). The map covering at least 92% of the autosomal length of the genome and 95% of the X chromosome consists of 1416 loci, 339 of which are microsatellites. Soon after another linkage map constructed by J. Weissenbach and colleagues (1992) was based entirely on the segregation analysis of 814 newly characterised polymorphic loci containing $(CA)_n$ repeats. Most of these markers (~760) contained a minimum of 12 CA doublets, they were tested for polymorphism on a sample of four individuals (eight alleles) and their highest allele frequency was <0.5 among these eight alleles.

Consequently, they were highly polymorphic markers with a mean level of heterozygosity close to 0.75. This map covered a linkage distance spanning 90% of the genome as estimated from the NIH/CEPH map with an average resolution of 5 cM. Despite several gaps of 20cM or more in these maps, they provide a general tool for primary linkage mapping of most single gene disorders even when only a limited number of affected families are available. This is due entirely to the high informativeness of the microsatellite markers. Continued generation of new markers should help to fill the remaining gaps in the present effort to construct the index map of the human genome with markers of heterozygosities above 0.7 and with no gaps greater than 15cM. These markers are also "ready-made" sequence tagged sites (STSs), providing useful reagents for physical mapping.

1.2.6 *Pathogenic simple repeat sequences*

Recent evidence indicates that trinucleotide repeat length variation causes a number of human diseases. Fragile X syndrome is associated with expansion of a $(CGG)_n$ repeat of usually more than 52 copies. This repeat is to lie in the 5' untranslated region of the FMR1 gene (Fu et al., 1991). Similarly, repeat number expansion of a CTG motif in the 3' untranslated region of the myotonin protein kinase gene leads to myotonic dystrophy (Aslanidis et al., 1992 ; Brook et al., 1992 ; Mahadevan et al., 1992). In the case of Huntington's disease, the unstable $(CAG)_n$ repeat (42 to 66 copies in HD patients) may be within the coding sequence of the Huntingtin (IT15) gene (The Huntington's Disease Collaborative Research Group, 1993), a feature shared with spino-bulbar muscular atrophy, an X-linked recessive disorder caused by expansion of a $(CAG)_n$ repeat in the coding sequence of the androgen receptor gene (LaSpada et al., 1991). These diseases show "genetic anticipation"- their severity increasing in successive generations- and this is now known to be due to increase in the repeat length. It will be interesting to see how often simple repeat expansion is involved in human disease, and what mechanisms cause this expansion beyond the normal range of stable inheritance.

2. Results

2.1 Construction of a mini library from cosmid HX43

Vector and cosmid restriction digests and DNA purification

Figure 13 shows aliquots of cosmid HX43 digested with SauIIIa, before and after purification, and an aliquot of vector pT7T3 digested with BamH1, after purification, electrophoresed on a 0.8% agarose gel alongside λHindIII marker. Vector pT7T3 had the expected size of 2.9kb and cosmid HX43 was digested to an average fragment size of 1kb. Concentrations were roughly estimated at 250ng/μl for HX43/SauIIIa and 100ng/μl for pT7T3/BamH1.

2.2 Ligation of dephosphorylated cosmid

The insert was ligated to vector pT7T3 as shown in figure 14. The first sample (lane 1) is an aliquot (1μl) removed from a 10μl ligation reaction before the addition of T4 DNA ligase (stored at -20^{0}C until electrophoresis) and the second sample (lane 2) is an aliquot (2μl) of the same reaction after ligation. A "shift up" of the ligated sample compared to the sample before ligation indicated successful ligation. Appropriate control ligation reactions were also carried out (data not shown). Samples were electrophoresed on a 0.8% agarose gel.

2.3 Transformation of competent cells with ligated products

After transformation (see chapter 2, section 3.2.1), colonies of bacterial cells containing a recombinant cosmid were white in colour because of the absence of a functional β-galactosidase activity. Colonies of bacterial cells which did not contain a recombinant molecule or they had been transformed with either of the two control samples (self ligated vector or wild type vector DNA) were blue in colour due to conversion of the colourless chromogenic substrate Xgal into a blue product by the E. coli enzyme β-galactosidase. No colonies were observed in the No DNA control plate as expected.

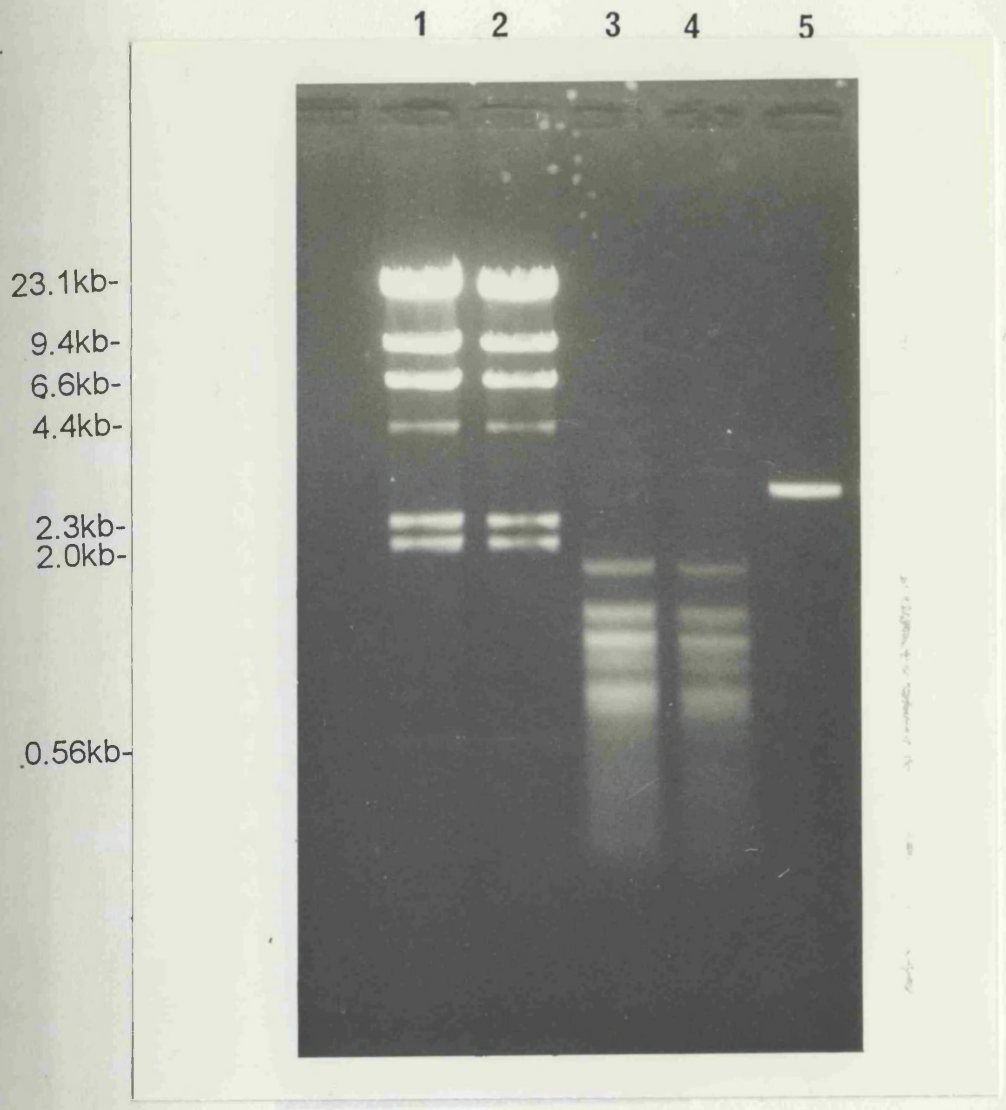


Figure 13: Electrophoresis on a 0.8% agarose gel of cosmid HX43 and vector pT7T3 DNAs digested with SauIIIa and BamH1, respectively. Lane 1: λHindIII marker (1μg) Lane 2: λHindIII marker (0.5μg) Lane 3: Cosmid HX43 (0.5μg) digested with SauIIIa before purification Lane 4: Cosmid HX43 digested with SauIIIa after purification Lane 5: pT7T3 digested with BamH1 (after purification)

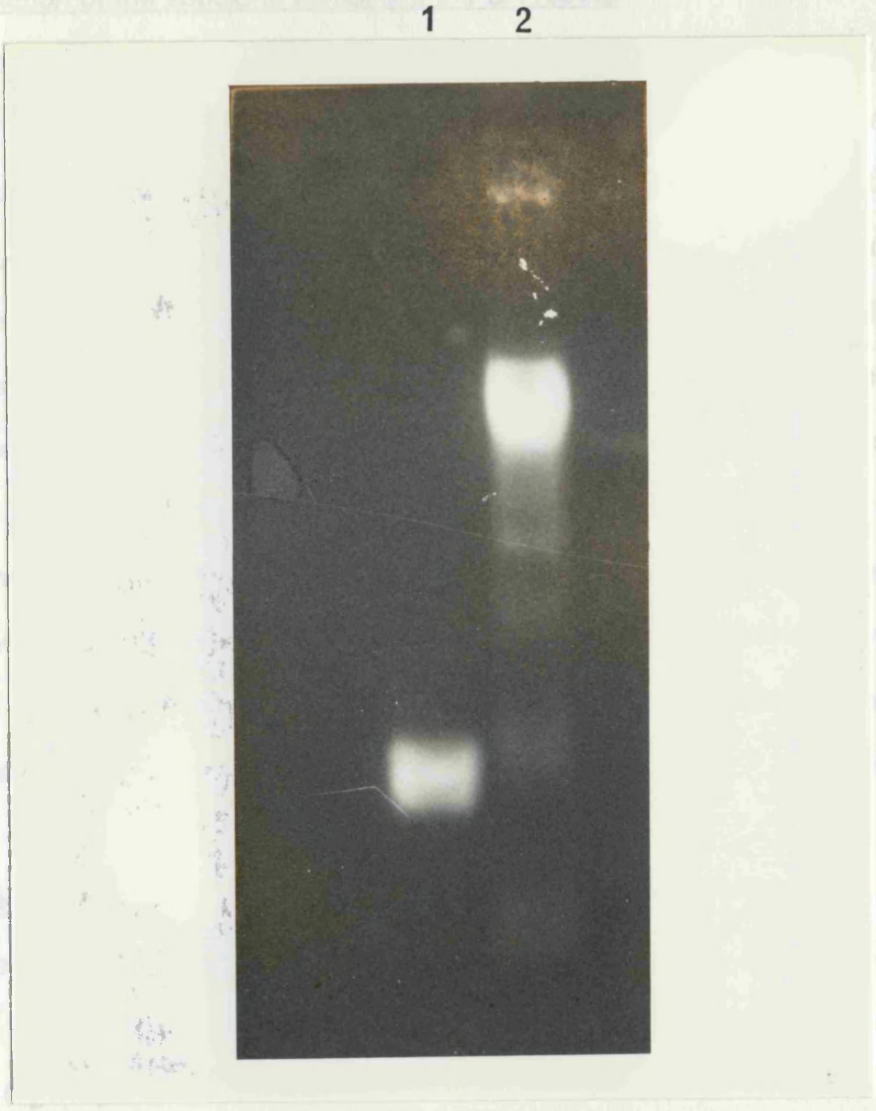


Figure 14: Electrophoresis on a 0.8% agarose gel of the ligated insert (HX43/SauIIIa) to vector (pT7T3/BamH1) Lane 1: Unligated vector and insert DNA where the insert DNA is not visible due to low quantity of loading. Lane 2: After ligation, a clear demonstration of the increase in size due to ligated product.

2.4 Isolation of the subclone containing the CA repeat

Replica filters, one for the experimental plate and one for a control plate (vector ligated to itself ; used as a negative control for the hybridisation), were prepared as described (chapter 2, sections 3.2.2) and screened by hybridisation with the probe $(GT)_{11+1}$. A filter with cosmids (derived from Southern blotting) known to contain CA repeats was included in the hybridisation experiment as a positive control. Positive signals were observed after autoradiography in the experimental plate and the positive control filter (fig. 15). No signals were observed in the negative control. Single colonies corresponding to positive signals were isolated and DNA was prepared using the alkaline lysis midiprep method. An aliquot of each preparation (2μl) was electrophoresed, the gel blotted and the filter hybridised with the $(GT)_{11+1}$ probe. Following autoradiography (exposure for two hours, at -70^{0}C), two positive signals were observed corresponding to DNAs from subclones 7 and 10 (fig. 16).

2.5 Sequencing of the pHX43-10 subclone

Before sequencing, the insert size of pHX43-10 subclone was determined. Saulla digest releases the insert from the vector and this was found to be approximately 1.4kb (fig. 17).

Subclones pHX43-7 and pHX43-10 and vector pT7T3 were, initially, sequenced using the Universal sequencing primer (-40, USB ; residues 120-136 in pT7T3 vector sequence retrieved from Genetics Computing Group via HGMP). Comparison of the sequence data revealed that sequence of the subclones were different after the BamH1 cloning site (residues 217-222 in pT7T3) compared to the vector sequence confirming presence of an insert in the subclones. Sequences of the two subclones were identical and, therefore, only one of the subclones (pHX43-10) was further analysed. At this stage, sequence data (183bp) did not reveal a stretch of poly CA repeat (fig. 18a). Therefore, a primer within this sequence was designed (primer I, residues 73-93, fig. 20) and used to extend the sequence of

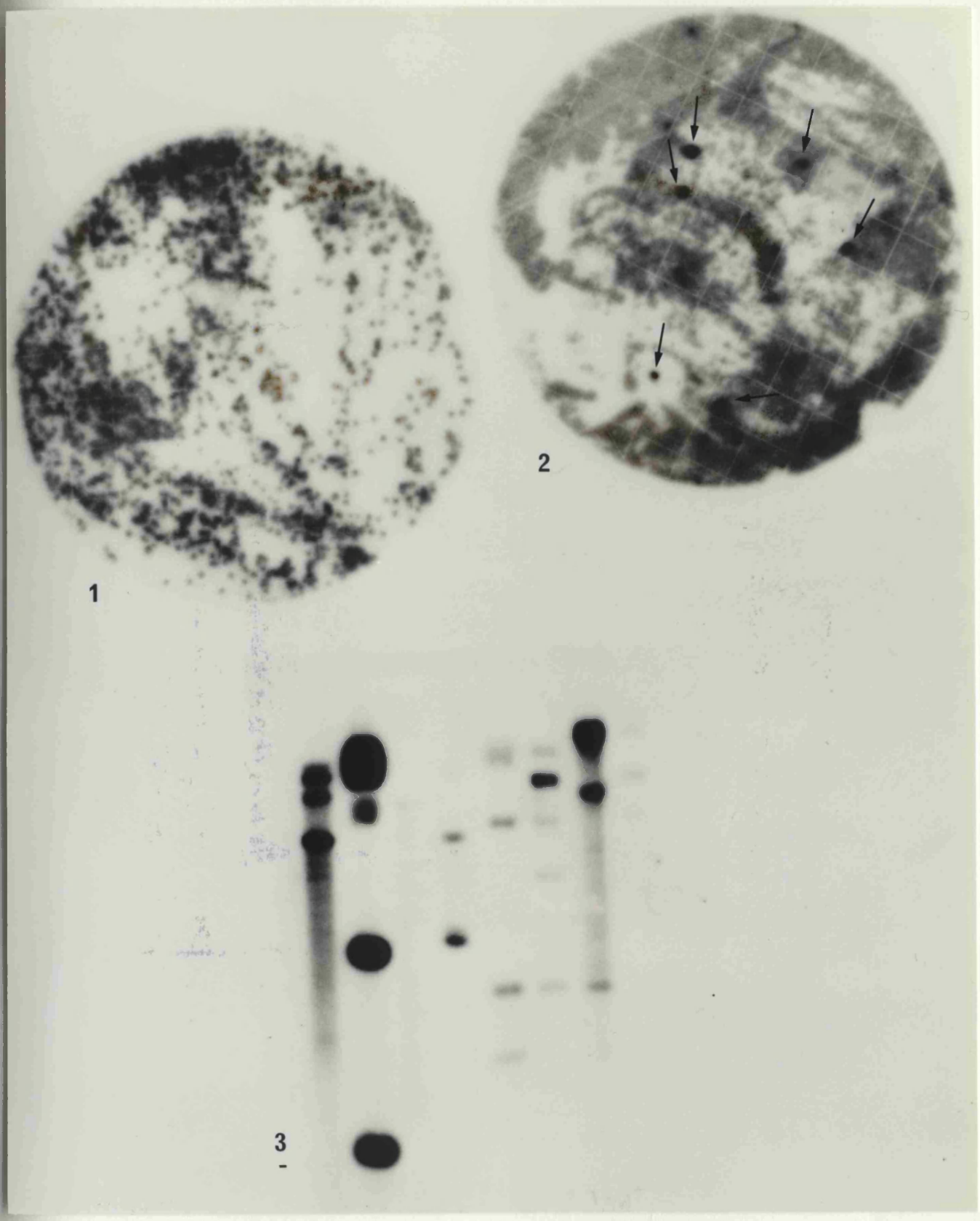


Figure 15: Positive signals (arrows in 2) after hybridisation (with probe [GT]n) and autoradiography of bacterial colonies transformed with the cosmid HX43 mini library subclones. Number 1 is the negative control (colonies transformed with vector ligated to itself) and number 3 is the positive control (filter with cosmids known to contain CA repeats).

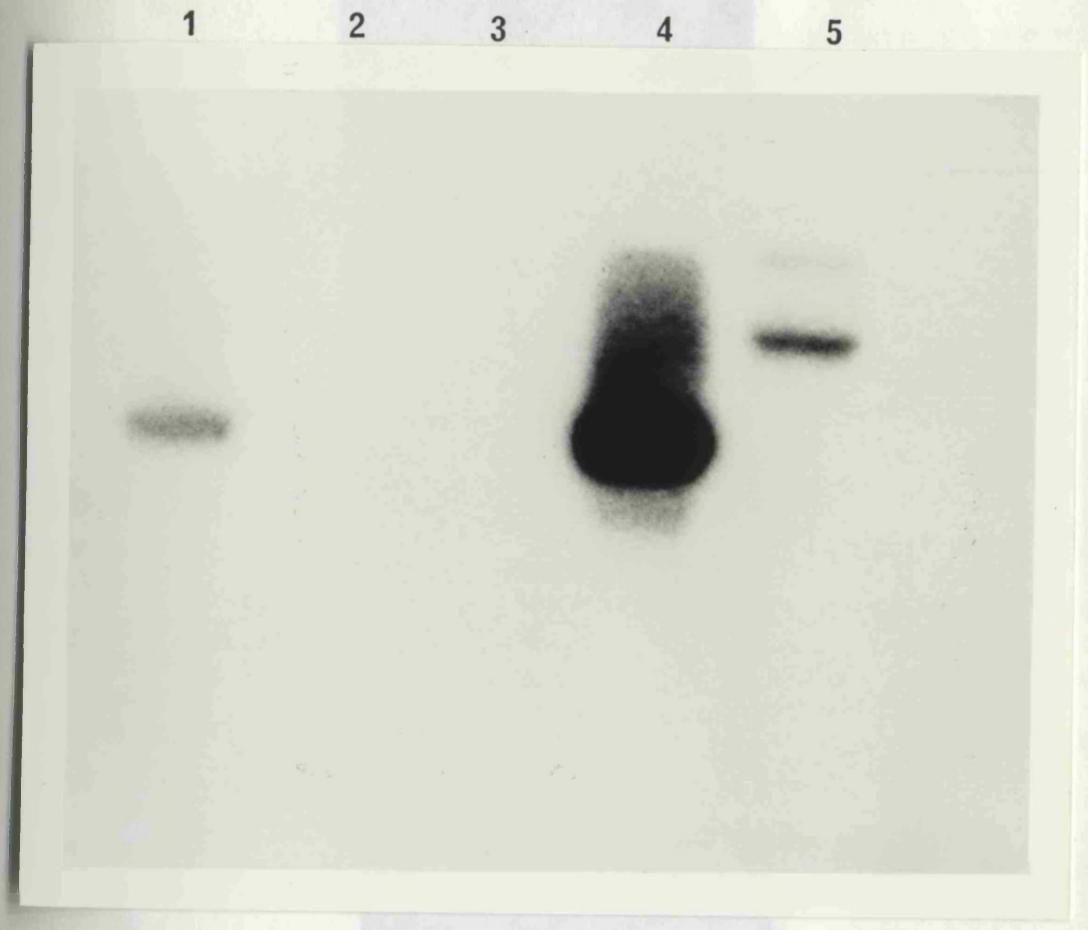


Figure 16: Southern blot hybridised with a poly (GT)n probe indicating the presence of a poly CA repeat (lanes 1 and 4) in subclones of cosmid HX43. Lane 5 is a positive control.

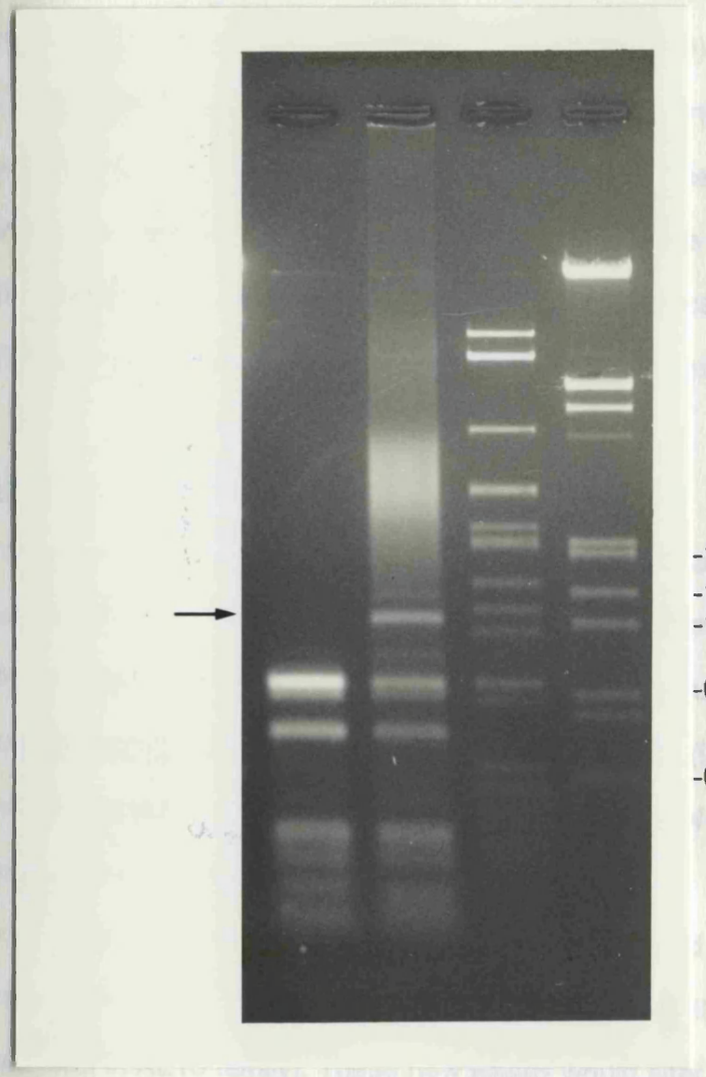


Figure 17: Gel electrophoresis of subclone pHX43-10 and vector pT7T3 digested with SauIIIa. Lane 1: Vector pT7T3 digested with SauIIIa Lane 2: Subclone pHX43-10 digested with SauIIIa Lane 3: λAvaII marker Lane 4: λHindIII/EcoR1 marker. An additional fragment of size ~1.4kb is observed in lane 3 which is the insert of subclone pHX43-10.

pHX43-10. This sequencing step provided data (fig. 18b-c) between residues 101-445 (fig. 20) containing a $(GT)_{18}$ repeat. Since the sequence after $(GT)_{18}$ repeat was unclear, a primer before the repeat was designed (primer II, residues 334-353, fig. 20) and the subclone sequenced further (residues 404-569). To confirm and clarify all the ambiguities, sequence of the reverse strand was generated (using primer III, residues 456-437) between residues 404-235 containing the $(AC)_{18}$ repeat (fig.19). Overall, the various sequencing steps provided a sequence of 569bp of subclone pHX43-10 as shown in fig. 20. The size of the PCR amplification product containing the $(GT)_{18}$ using primers II and III from the HX43 cosmid was 126bp.

2.6 Polymorphism of the HX43 $(GT)_{18}$ microsatellite

Using the polymerase chain reaction to amplify DNA within and flanking the HX43 microsatellite in 46 unrelated caucasian females (92 X chromosomes), 6 different allelic fragments (fig. 21) were detected. The sizes and frequencies of the 6 allleles are shown in table 8. Using these allele frequencies, a PIC (Polymorphic Information Content ; Botstein et al., 1980) of 0.52 and heterozygosity of 0.59 were calculated. The observed heterozygosity was 0.54.

Two more alleles, A7 (122bp) and A8 (120bp) were detected later during the genetic analysis of HX43 in the XLRP families (allele A8, individual III18 and allele A7, individual III59 in XL15 family). These new alleles would alter the PIC value slightly if included in its calculation.

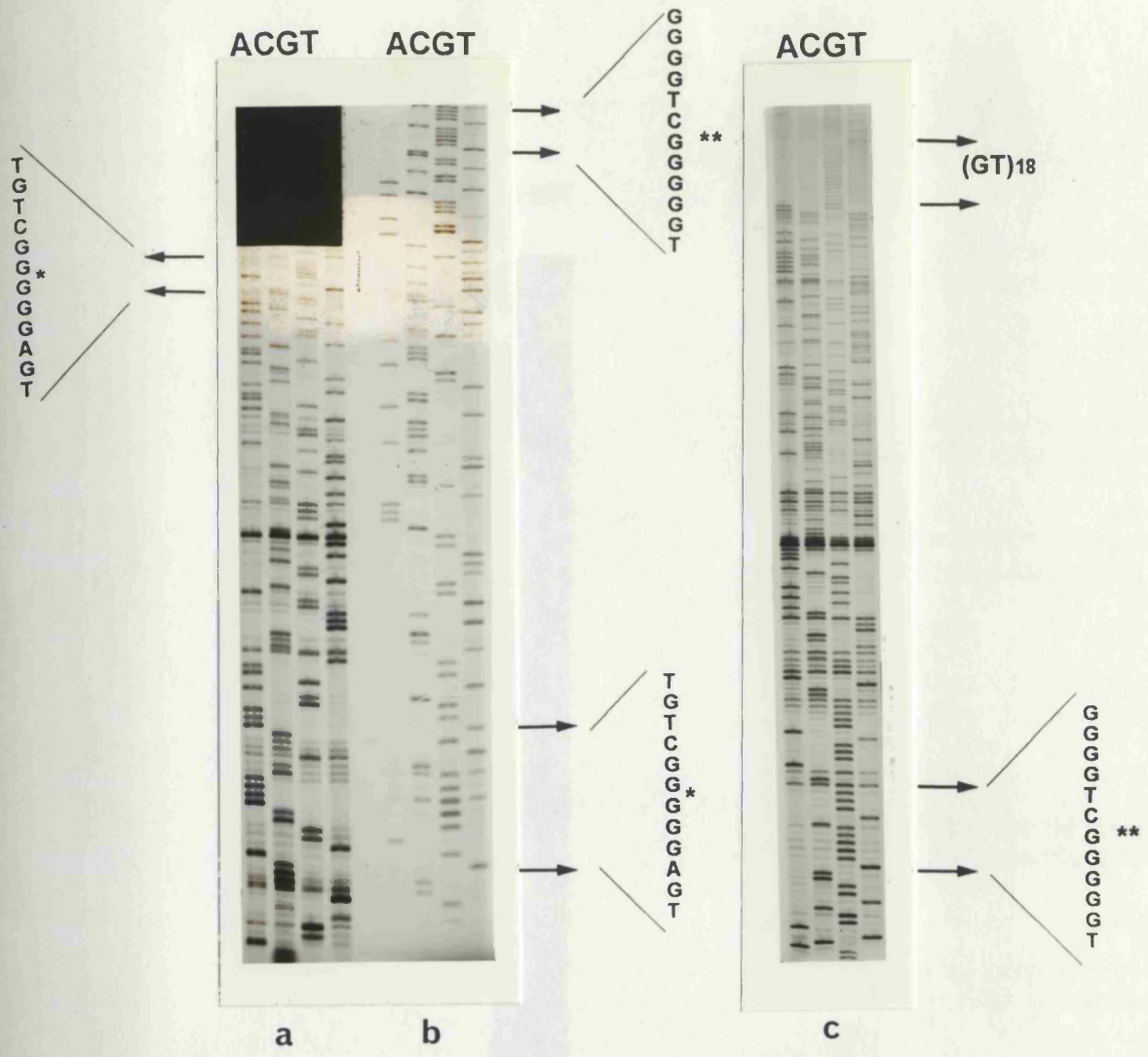


Figure 18: Autoradiographs of DNA sequencing gels showing the sequence of pHX43-10 subclone.

Sequence of pHX43-10 subclone using the Universal sequencing primer is shown in a. A primer was designed within this sequence for further sequencing. The obtained sequence after shorter and longer periods of electrophoresis is shown in b-c, respectively. A $(GT)_{18}$ repeat revealed at the end of the readable sequence after longer period of electrophoresis is shown in c.

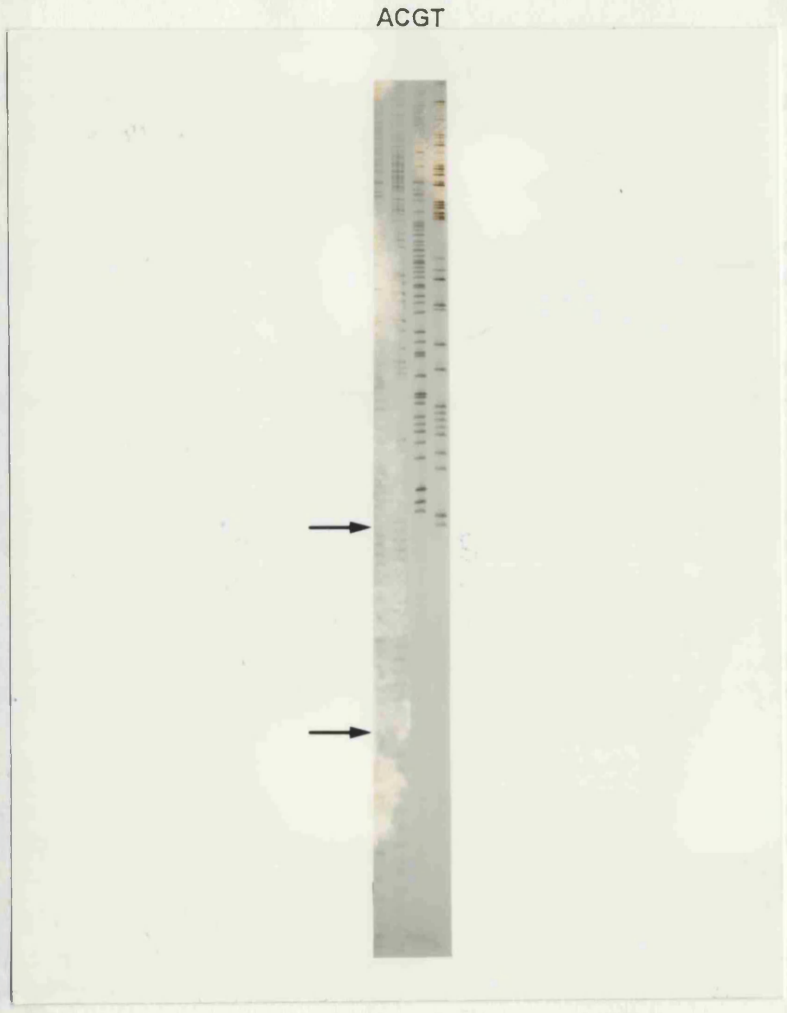


Figure 19: Autoradiograph of a DNA sequencing gel showing part of the pHX43-10 sequence containing an $(AC)_{18}$ repeat (between the arrows ; complementary strand to that shown in fig. 18c).

Sequence Range: 1 to 569

```
        10          20          30          40          50          60
GTCTNCCCAT GGCCAAAACC GCCCAAAGGA GAATNCNCTT TGGACTGGCT CTNCTGGGCT
CAGANGGGTA CCGGTTTTGG CGGGTTTCCT CTTANGNGAA ACCTGACCGA GANGACCCGA

        70          80          90          100         110         120
                                                        *
                          Primer I
CCTGCTTGAT GGCTCAGAAT ACTCCAGAAC TAGGATGACA GGCCTGAGGG GGCTGTGGCT
GGACGAACTA CCGAGTCTTA TGAGGTCTTG ATCCTACTGT CCGGACTCCC CCGACACCGA

       130         140         150         160         170         180
GGTCCTCTGG TTTGGCAAAT GCCTTCACCC TACCTGGACC CAGCTCTGCC TCTTCTGTCT
CCAGGAGACC AAACCGTTTA CGGAAGTGGG ATGGACCTGG GTCGAGACGG AGAAGACAGA

       190         200         210         220         230         240
                    *
AGGATGGGAC TAGGCTGGCC TGGGGGCTGG GGCGAGGGTA GGGGGCCTAG GGCGTCCTTG
TCCTACCCTG ATCCGACCGG ACCCCCGACC CCGCTCCCAT CCCCCGGATC CCGCAGGAAC

       250         260         270         280         290         300
                                                                    *
CAGAGAGGCA GAAAGCCCCT CTGCTTGCTC GTCTTCACCC TCACGGCAGG CCTGGCTGCA
GTCTCTCCGT CTTTCGGGGA GACGAACGAG CAGAAGTGGG AGTGCCGTCC GGACCGACGT

       310         320         330         340         350         360
                                                  Primer II
GACTTGTTTG CTTTCAGTCC TGTGAGCTTT TTCCTCATAC ACACAGCTGA CTGATTTTCT
CTGAACAAAC GAAAGTCAGG ACACTCGAAA AAGGAGTATG TGTGTCGACT GACTAAAAGA

       370         380         390         400         410         420
                                            *
TCTCAGAGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTTTGTGTG GTGTGTGGTG
AGAGTCTCAC ACACACACAC ACACACACAC ACACACACAC ACAAACACAC CACACACCAC

       430         440         450         460         470         480
TGGGTGGTGG GGGNTTGTCT GTGAATGAGA AAGAAGTAAA TAAAATACCT CAGGCGGGTG
ACCCACCACC CCCNAACAGA CACTTACTCT TTCTTCATTT ATTTTATGGA GTCCGCCCAC
                     Primer III

       490         500         510         520         530         540
                    *
GGTCTCTGTC CTGTGTGCCC AGGCTGCCAG CTGCACCTGC CCAAGGTCTC TGGGCAAAGC
CCAGAGACAG GACACACGGG TCCGACGGTC GACGTGGACG GGTTCCAGAG ACCCGTTTCG

       550         560
TGGTCCACGG AGGAGCGGGG CTGTCTGGA
ACCAGGTGCC TCCTCGCCCC GACAGACCT
```

Figure 20: Sequence of pHX43-10 subclone.

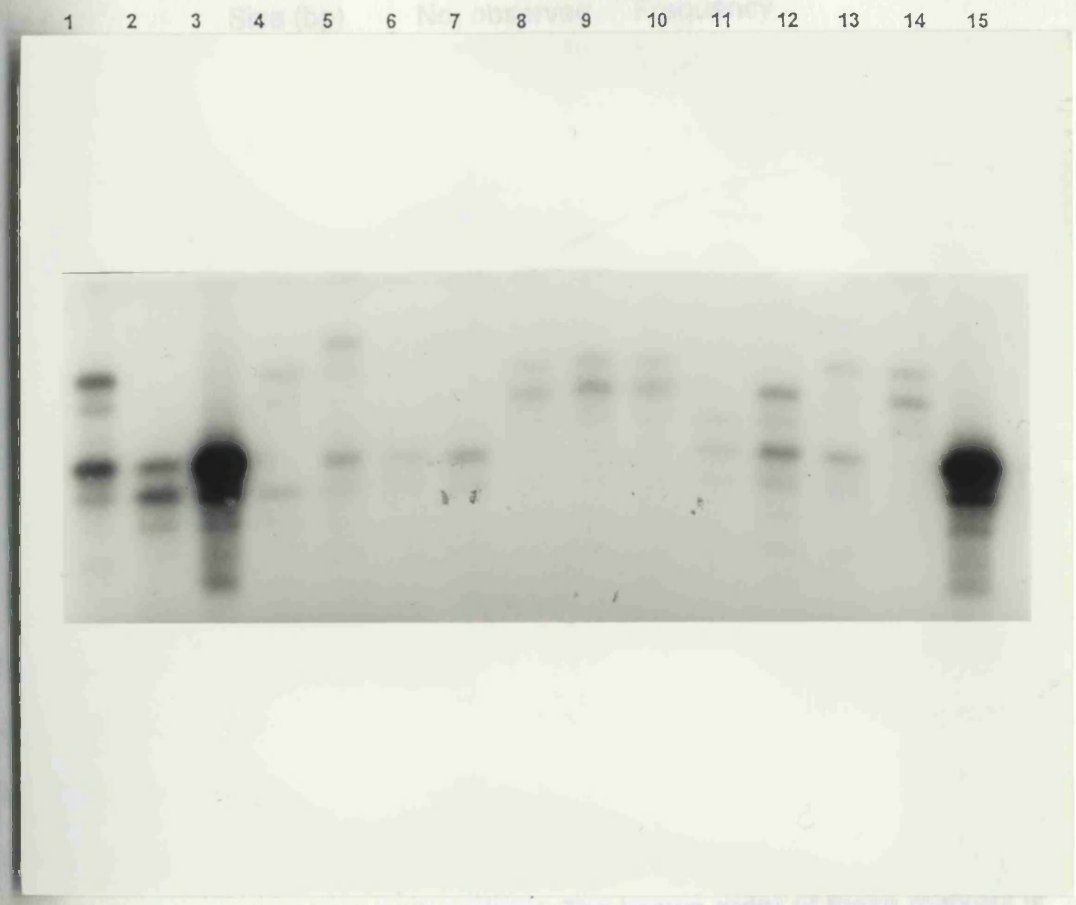


Figure 21: Autoradiograph of a DNA sequencing gel showing 6 alleles of marker HX43 detected after PCR amplification in 46 unrelated females.
Lane 1 and 13: A2/A5. Lane 2: A5/A6. Lane 3: subclone pHX43-10 (126bp). Lane 4: A2/A6. Lane 5: A1/A5. Lanes 6 and 7: A5/A5. Lanes 8, 9, 10 and 14: A2/A3. Lane 11: A4/A5. Lane 12: A3/A5. Lane 15: subclone pHX43-10.

Table 8: Sizes of HX43 allelic fragments and their frequencies in 46 unrelated caucasian females (92 X chromosomes)

Allele	Size (bp)	No. observed	Frequency
A1	134	1	0.01
A2	132	29	0.32
A3	130	5	0.05
A4	128	1	0.01
A5	126	50	0.55
A6	124	6	0.06

2.7 Localisation of HX43 marker by genetic analysis

The inheritance of marker HX43 was followed in 12 X-linked Retinitis Pigmentosa (XLRP) families for which prior linkage data for a number of markers were available (fig. 22a-l). The loci OTC, DXS228, DXS7 and DXS426 were used as reference markers to define the position of HX43 genetically (typing for reference markers was performed by other members of the group). The known order of these markers is telomere-OTC-DXS228-DXS7-DXS426-centromere (Report of the Fourth International Workshop on Human X Chromosome Mapping, 1993). The data were analysed pairwise using the ILINK programme to estimate recombination fractions (θ) and corresponding LOD scores (Lathrop and Lalouel, 1984). The results of two-point analyses between HX43 and OTC, DXS228 and DXS7 are shown in table 9. It is clear that HX43 is closely linked to OTC, DXS228, DXS7 (and DXS426), which is consistent with the physical localisation observed by two colour FISH (chapter 3).

Table 9 Lod-score table for linkage of HX43 to OTC, DXS228 and DXS7.

	θ_{max}	Lod_{max}
HX43/OTC	0.074	5.813
HX43/DXS228	0.055	5.737
HX43/DXS7	0.081	7.418
OTC/DXS7	0.120	5.710
DXS228/DXS7	0.0	10.386

Several muliple informative crossovers allowed localisation of HX43 with respect to the other markers. In four informative meioses in families XLRP11 (IV6 ; fig. 22b), XLRP15 (III25 ; fig. 22c and 23b), XLRP45 (III1 ; fig. 22e) and XLRP47 (V1 ; fig. 22f) HX43 co-segregates with marker DXS228 and crosses over with OTC, strongly suggesting that HX43 lies proximal to OTC. Similarly in further eight informative meioses in families XLRP11 (IV6 ; fig. 22b), XLRP15 (III3, III18, III21 and III25 ; fig. 22c and 23a), XLRP73 (IV4 and IV7 ; fig. 22i) and XLRP74 (IV4 and V8 ; fig. 22j), HX43 co-segregates with marker OTC and crosses over with DXS228, indicating that HX43 lies distal to DXS228 (and consequently distal to marker DXS7). Therefore, it is concluded that the order of loci must be tel-OTC-HX43-DXS228-DXS7-cen.

3. Discussion

The polymorphic nature of dinucleotide repeats and their exploitation as genetic markers have been well documented (Weber, 1990 ; Litt and Luty, 1989 ; Weber and May, 1989 ; Tautz, 1989 ; Smeets et al., 1989). In most cases, small-insert plasmid libraries were constructed and clone containing CA repeats were identified by nucleic acid hybridisation. Such an approach was used in this study for the isolation of a microsatellte from the X-specific cosmid HX43. In a similar method,

XLRP5

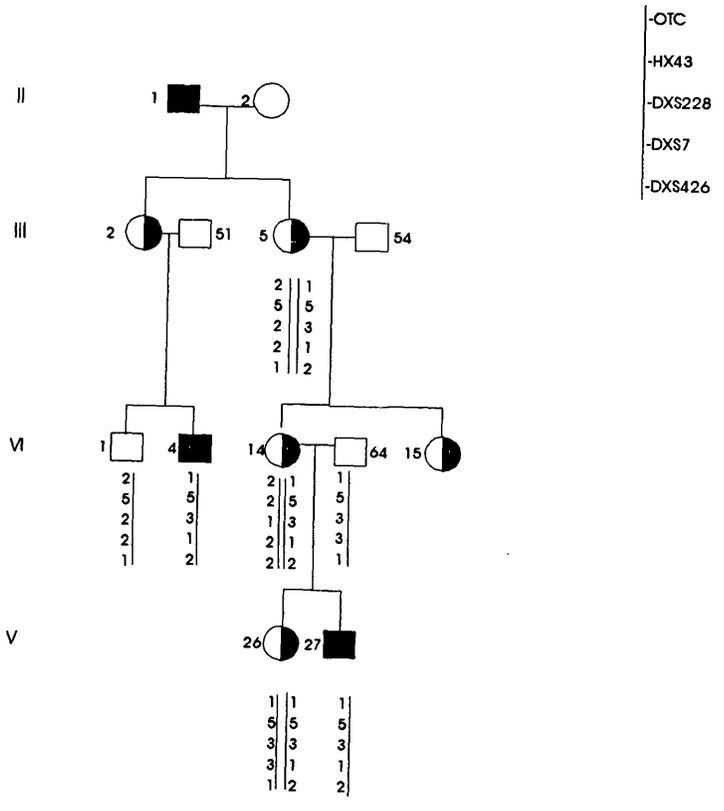


Figure 22a: No crossovers

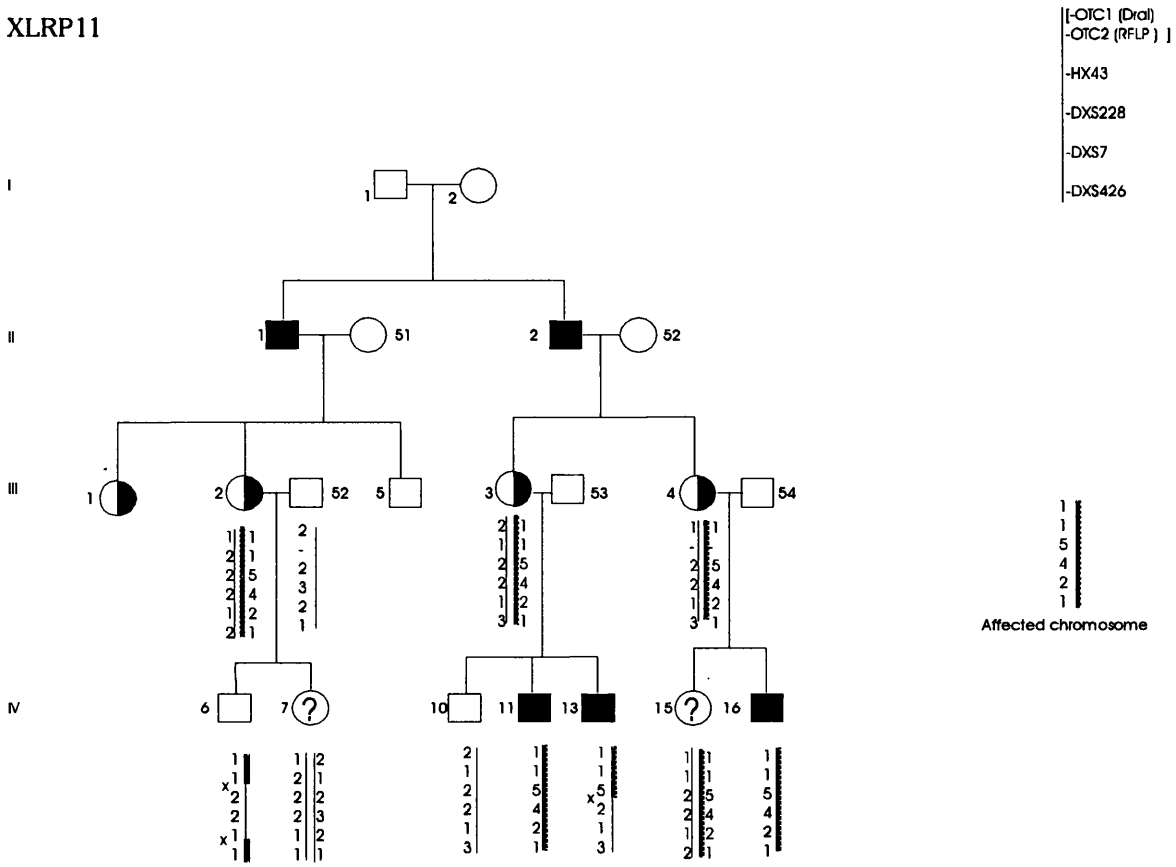


Figure 22b: Crossover between DXS7 and DXS426, individual IV6.

Crossover between OTC and HX43 (individual IV6) places HX43 proximal to OTC.

Crossover between HX43 and DXS228 (individual IV13) places HX43 distal to DXS228.

XLRP15

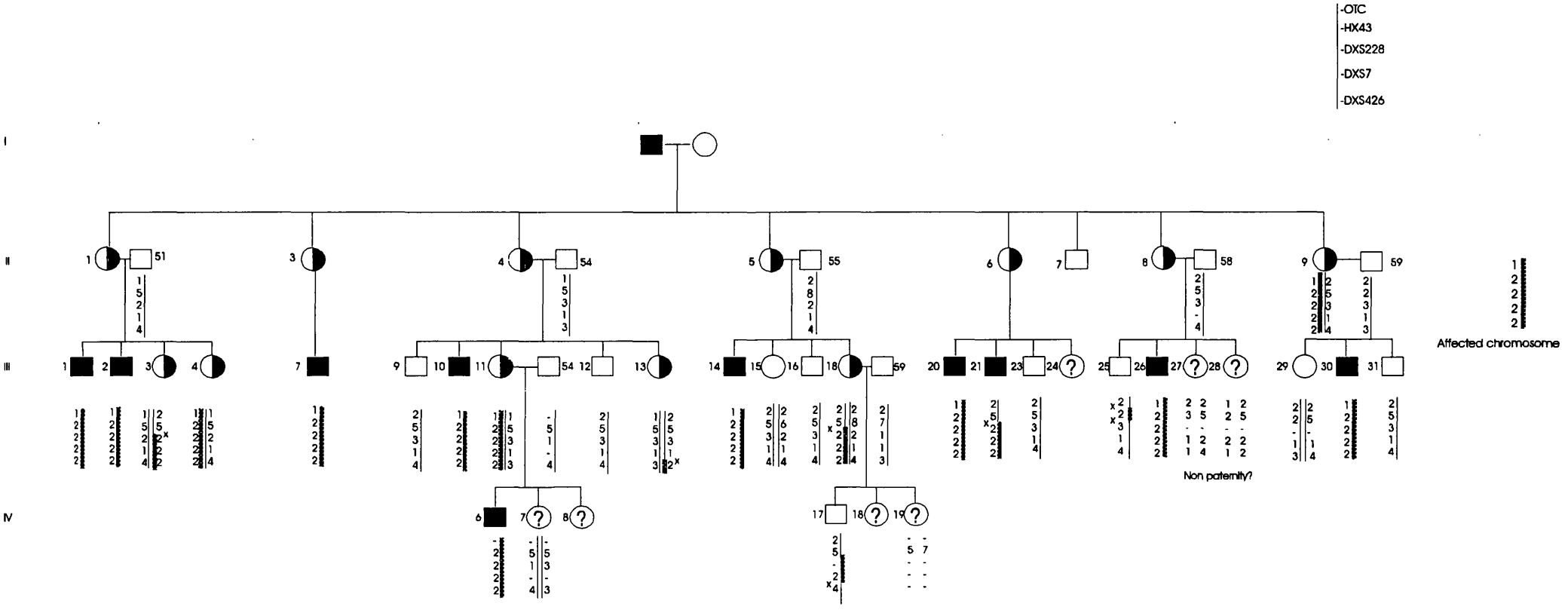


Figure 22c: Crossover between HX43 and DXS228 (individual III3) places HX43 distal to DXS228.
Crossover between DXS7 and DXS426 in individual III13.
Crossover between HX43 and DXS228 (individual III18) places HX43 distal to DXS228.
Crossover between DXS7 and DXS426 in individual IV17.
Crossover between HX43 and DXS228 (individual III21) places HX43 distal to DXS228.
Crossovers between a. OTC and HX43 b. HX43 and DXS228 (individual III25) placing HX43 proximal to OTC and distal to DXS228.

XLRP38

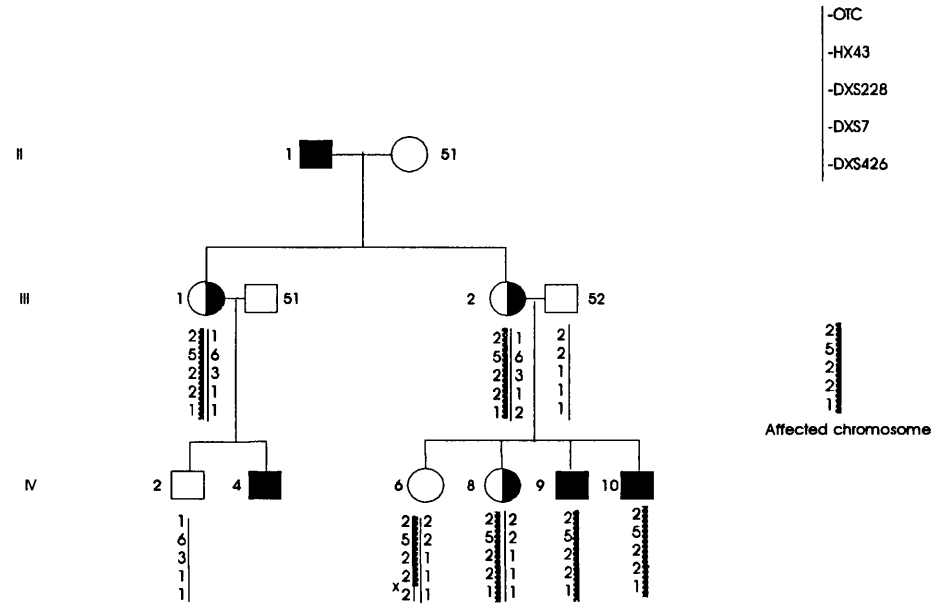


Figure 22d: Crossover between DXS7 and DXS426 in individual IV6.

XLRP45

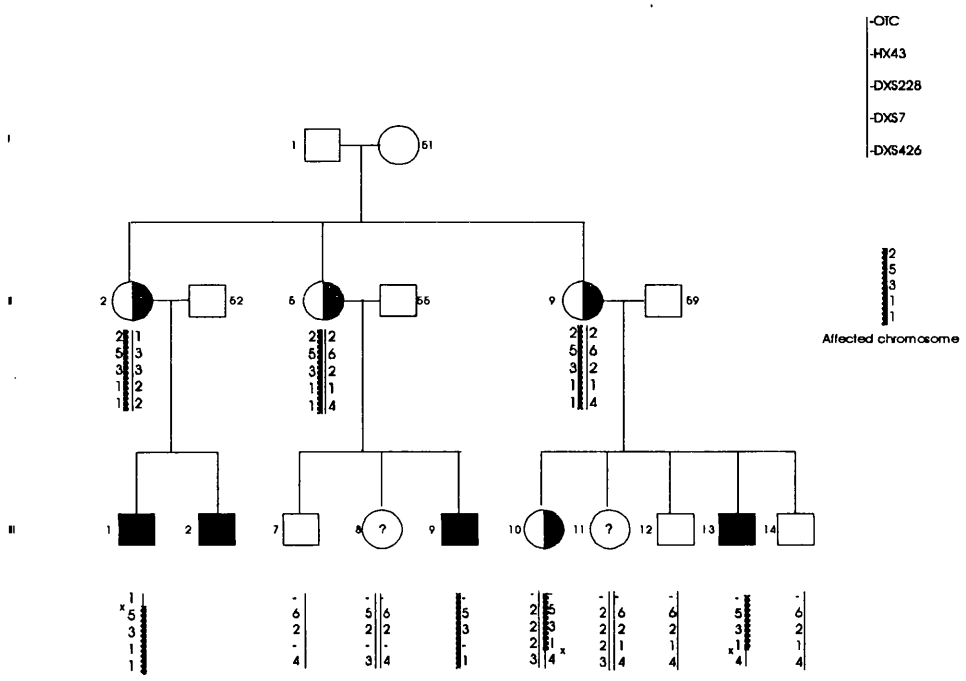


Figure 22e: Crossover between OTC and HX43 in individual III1.

Crossover between DXS426 and disease in individuals III10 and III13.

XLRP47

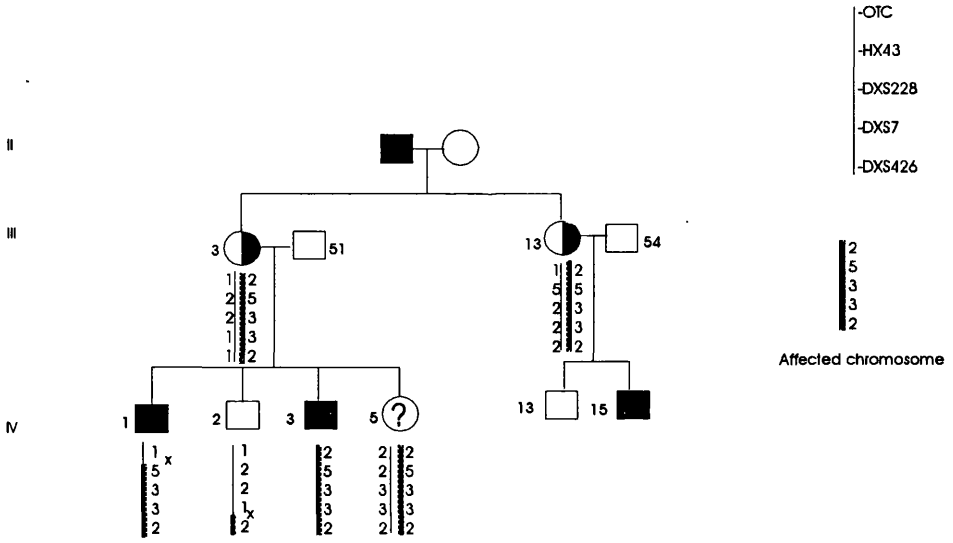


Figure 22f: Crossover between OTC and HX43 (individual IV) places HX43 proximal to OTC.

Crossover between DXS7 and DXS426 in individual IV2.

XLRP71

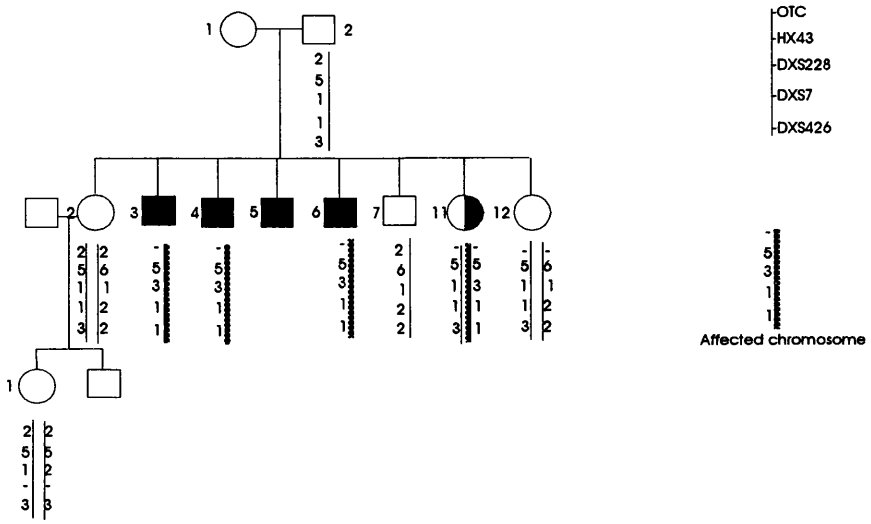


Figure 22g: No crossovers

XLRP72

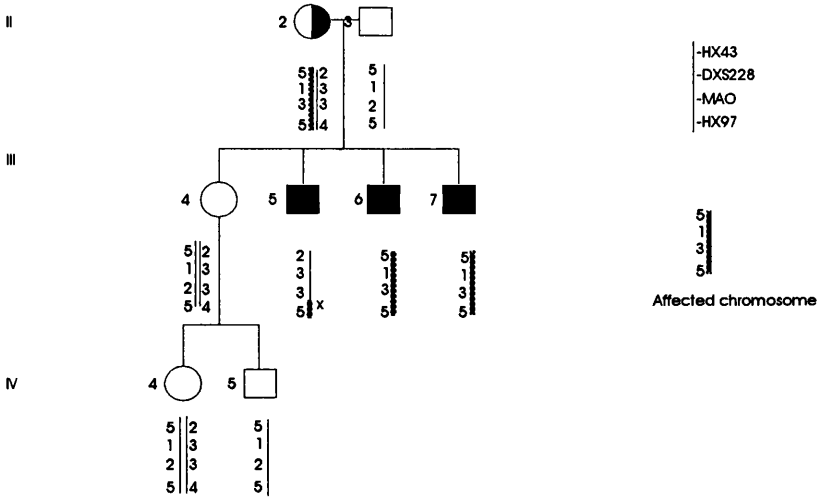


Figure 22h: DXS7 and DXS426 uninformative
MAO and HX97 were used instead
(Crossover between MAO and HX97 in individual II5)

XLRP73

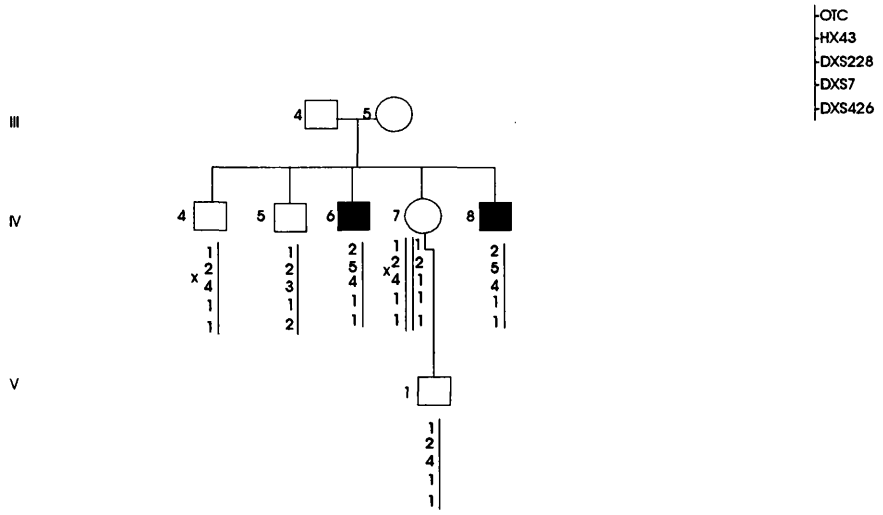


Figure 22i: Crossovers between HX43 and DXS228 in individuals IV4 and IV7

XLRP74

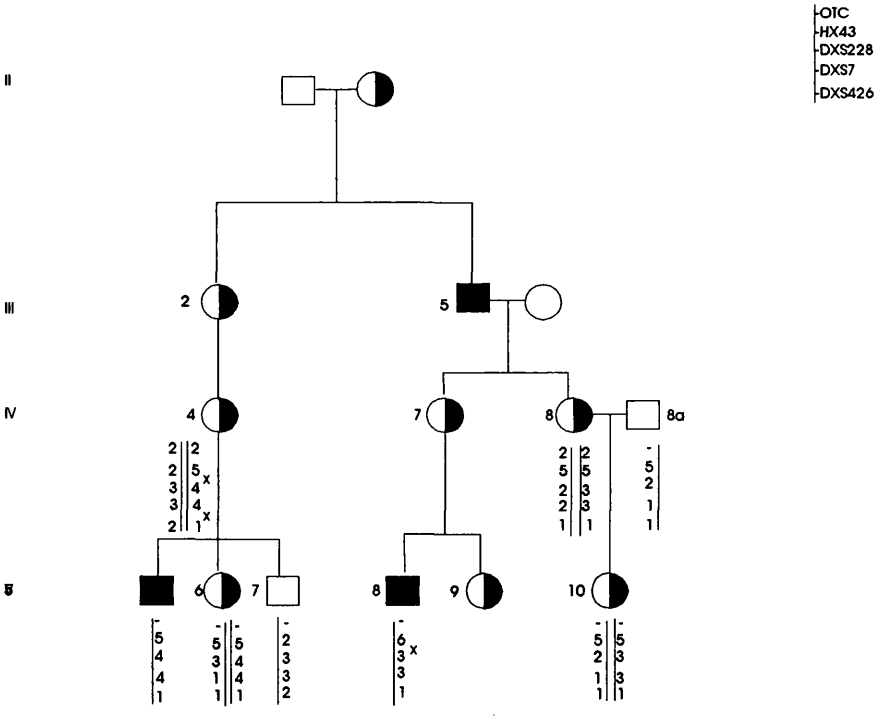


Figure 22j:

Crossover between HX43 and DXS228 and DXS7 and DXS426 in individual IV4 places HX43 distal to DXS228.

Crossover between HX43 and DXS228 (individual V8) places HX43 distal to DXS228.

XLRP75

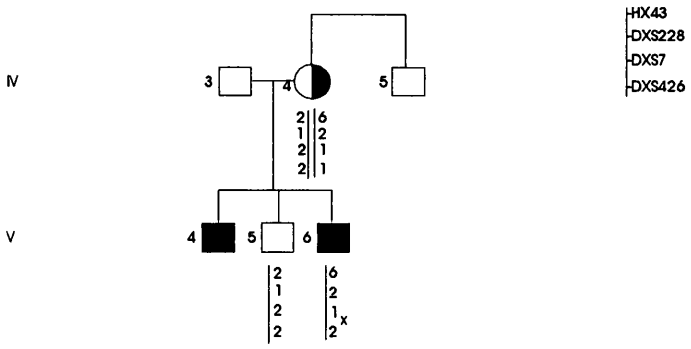


Figure 22k

XLRP77

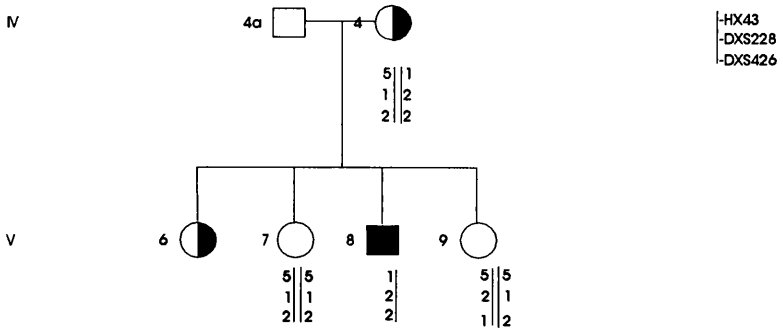


Figure 22l: Non-paternity either for V7 or V9.

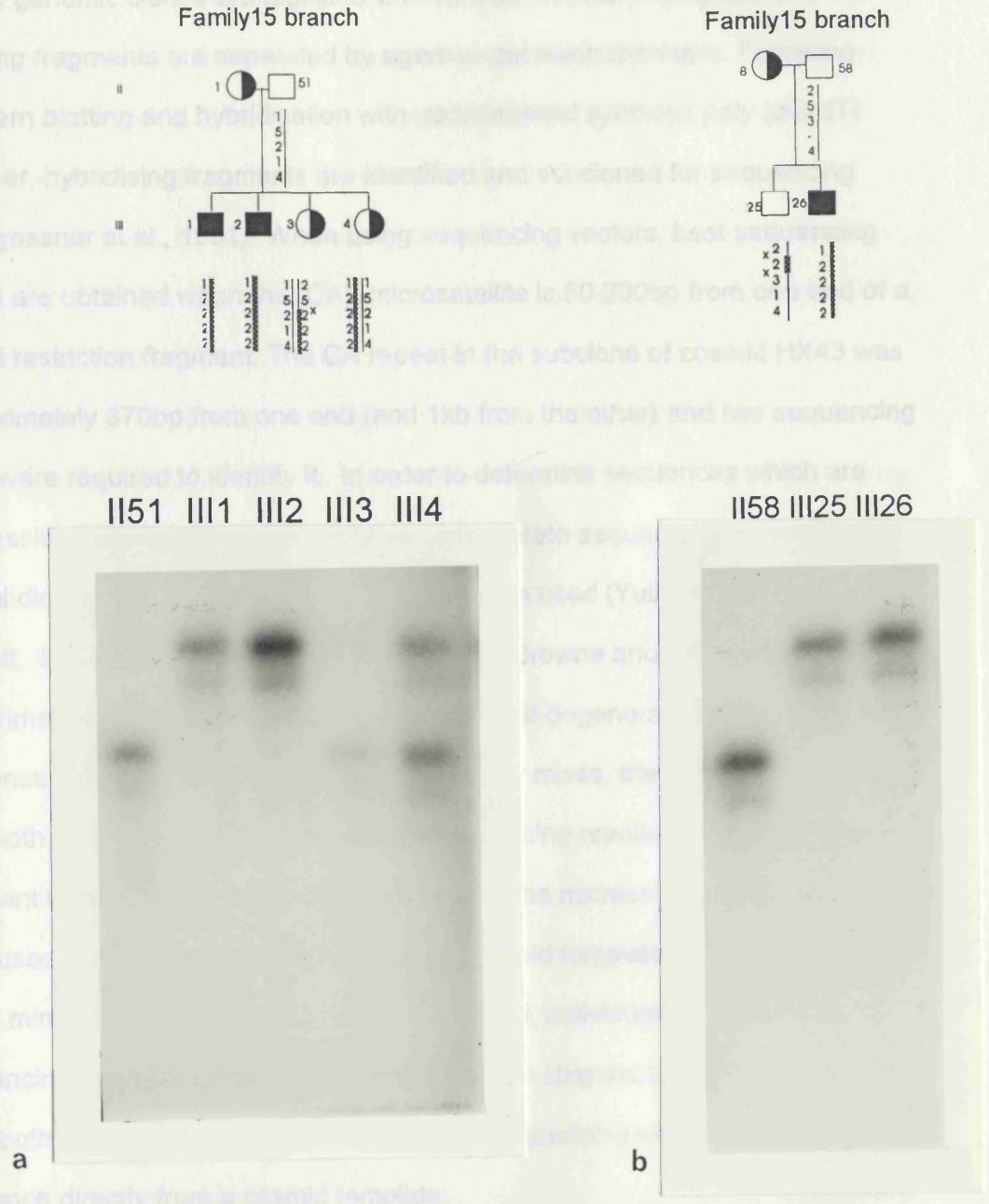


Figure 23a-b:

a: Crossover between DXS228 and HX43 (individual III3) in family F15.
Autoradiograph of HX43 segregation and family branch showing the reference markers segregation.

b: Crossovers between HX43 and OTC and HX43 and DXS228 (individual III25) in family F15.
Autoradiograph of HX43 segregation and family branch showing the reference markers segregation.

human genomic clones are digested with various restriction enzymes and the resulting fragments are separated by agarose-gel electrophoresis. Following Southern blotting and hybridisation with radiolabelled synthetic poly (dG-dT) oligomer, hybridising fragments are identified and subcloned for sequencing (Kirchgessner et al., 1991). When using sequencing vectors, best sequencing results are obtained when the $(CA)_n$ microsatellite is 50-200bp from one end of a cloned restriction fragment. The CA repeat in the subclone of cosmid HX43 was approximately 370bp from one end (and 1kb from the other) and two sequencing steps were required to identify it. In order to determine sequences which are inaccessible with vector-derived primers, degenerate sequencing primers which anneal directly to $(CA)_n$ microsatellites have been used (Yuille et al., 1991 ; Browne and Litt, 1992 ; LeBlanc-Straceski et al., 1994). Browne and Litt used two sets of four primers each [(CA) or (GT)] with a twelve-fold degeneracy at the 3' end to sequence plasmid templates. Using these primer mixes, they obtained sequence from both strands by running two sets of sequencing reactions. Their method was sufficient to obtain DNA sequence from 88% of the microsatellites attempted. Yuille et al. used a similar approach to sequence plasmid templates. They used primers with a minimum level of degeneracy at the 3' end, individually in six sets of sequencing reactions to sequence one of the two strands. LeBlanc-Straceski et al. used both methods in combination with cycle sequencing which allowed them to sequence directly from a cosmid template.

PCR-based procedures for the isolation of microsatellites have also been reported. The basic principle involves restriction digestion of the DNA (from a cosmid, a λ clone or a YAC) followed by ligation of the fragments to appropriate linkers (Edwards et al., 1992 ; Pandolfo, 1992). PCR is performed on the ligation products, using primers for the microsatellite and the linker to amplify the intervening sequence, which is then sequenced either directly or after cloning. The linkers described in both papers were based on the principle outlined by Riley et al. (1990 ; marketed as the "Vectorette" system). This method was used for the isolation of a

microsatellite in this study and is described in the following chapter. A similar method (Koref et al., 1993) uses ligation of restricted DNA to dephosphorylated vectors (e.g. pUC8) and PCR on the ligated products, using a primer for the microsatellite and a primer specific to the vector.

The methods which involve subcloning of short inserts and hybridisation screening of colony lifts for CA-containing clones are more-labor intensive than the methods using degenerate primers or PCR. This approach was used, however, by Weissenbach et al. (1992) in their genome-wide effort to generate a genetic map based on microsatellite polymorphisms. After sequencing the CA-positive clones from a library of AluI-digested human genomic DNA, they selected repeats with greater than 12 CAs for producing STSs. In all, 12,014 clones were sequenced, producing 814 polymorphic markers, twenty-five of which mapped on the X chromosome. Heterozygosities of the markers for the whole genome ranged from 0.32 to 0.96 with 97% showing heterozygosities above 0.5 and 74.4% above 0.7. HX43 marker contained 18 CAs and its heterozygosity was 0.59. Its inheritance was followed through 12 XLRP families since it was critical to resolve its localisation relative to marker DXS7 in the proximal Xp. Genetic studies in two of the X-linked Retinitis Pigmentosa families (Coleman et al., 1990), RP15 and RP47, had indicated that the RP2 gene was located between DXS7 and DXS426, provided that the same gene was indeed responsible for both families. It was important, therefore, to search for markers mapping between DXS7 and DXS426 (especially because of lack of informative markers in this region) and ideally between the two critical crossovers in the two families thus aiding the finer localisation of the RP2 gene. As part of this thesis FISH studies undertaken to localise a group of markers (including HX43) mapping to proximal Xp suggested that HX43 was the best candidate to lie between DXS7 and DXS426. Its precise position was, therefore, resolved genetically as described in this chapter. The order of loci in relation to other markers was determined as: tel-OTC-HX43-DXS228-DXS7-DXS426-cen. Since HX43 clearly maps distal to the region of interest (i.e. between DXS7 and DXS426 ; fig.24) it was

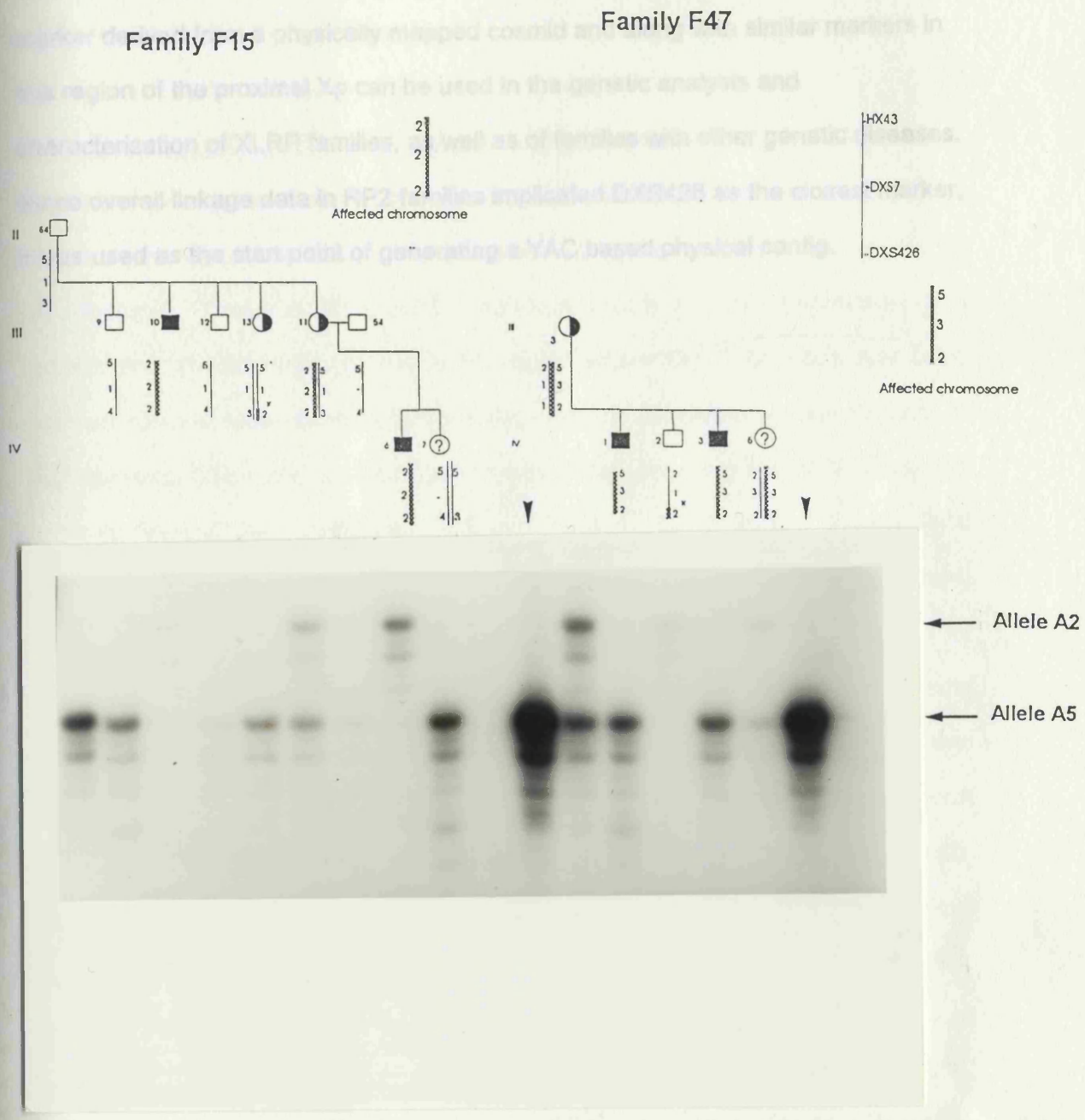


Figure 24: HX43 segregates with DXS7 in the critical crossovers in families F15 (individual III13) and F47 (individual IV2).
Autoradiograph of HX43 segregation and family branches showing the DXS7 and DXS426 segregation (PCR products from subclone pHX43-10 are shown with arrowheads).

not included in further physical mapping studies. It is, however, a polymorphic marker derived from a physically mapped cosmid and along with similar markers in this region of the proximal Xp can be used in the genetic analysis and characterisation of XLRP families, as well as of families with other genetic diseases. Since overall linkage data in RP2 families implicated DXS426 as the closest marker, it was used as the start point of generating a YAC based physical contig.

CHAPTER 5: CONSTRUCTION OF A YAC CONTIG

1. Introduction

1.1 Construction of a YAC contig around the DXS426 locus

Physical mapping and ordering of polymorphic markers by FISH was used as one approach for the generation of markers in the proximal short arm of the X chromosome. These markers would provide a source for the construction of a genetic and physical map of the Xp11 region where the RP2 locus has been mapped. On the basis of the FISH results, HX43, a candidate marker thought to map between DXS7 and DXS426 (which are closely linked to the RP2 locus and probably flank it), was genetically mapped distal to the region of interest and therefore not included in further physical studies. While this work was in progress, the location of the gene for RP2 was broadened, based on re-evaluation of genotyping data in critical families from our own laboratory. The new flanking markers were monoamine oxidase B (MAOB, proximal to DXS7 at Xp11.3) and DXS255, mapping proximal to DXS426 at Xp11.22 (unpublished results and Report of the Fourth International Workshop on Human X chromosome Mapping, 1993). The construction of a physical map of the region based on Yeast Artificial Chromosomes was a different approach to developing further markers and would provide a source for the isolation of candidate genes for RP2. A YAC contig was constructed using PCR-based end cloning and Sequence-Tagged Site (STS) mapping techniques. The DXS426 locus was used as a starting point for chromosome walking.

1.2 The analysis of YAC clones

1.2.1 General considerations

The analysis of complex genomes involves a multitude of techniques with varying degrees of resolution. Until recently a large gap existed between conventional

genetic analysis with resolution in centiMorgan (cM), which in physical terms approximates a megabase (Mb), and the size of DNA fragments that could be cloned in cosmids (~50kb). A major advance towards bridging this gap was made by Schwartz and Cantor (1984) when they described the technique of pulsed-field gel electrophoresis (PFGE), capable of resolving DNA fragments between 50-10000kb. This is approximately the size of an average human metaphase chromosome band. PFGE can be used to analyse different aspects of genome structure i.e. creating long-range maps, ascertaining the position of genes and localising genes of particular interest. However, a problem still remained due to the inability to move from long-range maps to the detailed analysis of any long stretch of DNA. This was resolved following the description of an elegant vector system (Burke et al., 1987) capable of cloning several hundred kilobase fragments of linear DNA as yeast artificial chromosomes (YACs).

1.2.2 Vector system

The vector (pYAC4) incorporates all necessary functions into a single plasmid to allow replication in *Escherichia coli*. This plasmid, called a "yeast artificial chromosome" vector supplies a cloning site within a gene whose interruption is phenotypically visible (ochre-suppressing allele of a tyrosine transfer RNA gene, SUP4), an autonomous-replication sequence (ARS1), a centromere (CEN4), selectable markers on both sides of the centromere (TRP1 and URA3), and two sequences that seed telomere formation in vivo (TEL). ARS1 and CEN4 are sequences that are naturally adjacent to TRP1 on yeast's chromosome IV. The TEL sequences are derived from the termini of the *Tetrahymena* macronuclear ribosomal DNA (rDNA) molecules. Double digestion of vector pYAC4 with BamH1 and EcoRI yields three fragments, which can be regarded as a left chromosome arm (6kb including the centromere), a right chromosome arm (3.4kb), and a replacement region which contains the his3 gene. This separates the two TEL sequences in the circular plasmid. The two arms are treated with alkaline phosphatase to prevent

religation, and then ligated onto large insert molecules derived from the source DNA by partial or complete digestion with an enzyme that leaves compatible ends. The ligation products are then transformed into yeast (Saccharomyces cerevisiae) spheroplasts by standard methods, which involve embedding the transformed spheroplasts in agar on a selective medium. Transformants are selected for complementation of the host ura3 and trp1 markers by the URA3 (right arm) and TRP1 (left arm) genes on the vector, which ensures that the artificial chromosomes have derived both their arms from the vector. Insertion of exogenous DNA at the EcoR1 cloning site results in loss of expression of the ochre suppressor SUP4. SUP4 is particularly advantageous interruptible marker since *ade2-ochre* host cells that are expressing the suppressor form white colonies and those in which the suppressor has been inactivated (recombinant clones) form red colonies (due to the accumulation of a red adenine pre-metabolite).

Three significantly different vector systems compared to the original series of pYAC vectors have also been described. The first system uses single yeast telomere-containing vectors for the cloning of human telomeres by complementation (Brown et al., 1989 ; Cross et al., 1989). A second is a two-vector system (Shero et al., 1991) incorporating the advantages of multicloning sites withT7 bacteriophage promoters. Finally, a significant addition to the YAC vectors has been a system which allows the amplification of the artificial chromosome in yeast (Smith et al., 1990).

1.2.3 YAC DNA isolation

Two approaches for the isolation of YAC DNA from yeast cells are used: preparation of DNA in solution or as intact yeast chromosomes in agarose plugs. Preparation in solution yields DNA of a size range (50-200kb) sufficient for PCR analysis and restriction mapping by standard techniques, neither of which require high molecular weight DNA. High molecular weight DNA is impossible to prepare by conventional solution methods because of the extreme sensitivity to shear damage.

Protection from shear forces during preparation is, therefore, achieved by encapsulating cells prior to lysis in agarose beads or blocks (plugs). An enzyme (e.g. lyticase) is initially used to degrade the polysaccharide cell wall material of the host organism Saccharomyces cerevisiae and make spheroplasts (cell wall-less cells). The spheroplasts are then embedded in low-gelling agarose and lysed in a EDTA/sarcosine/proteinase K solution at 50^0C for 2 days. Lysis occurs sooner than the 2-day incubation period which is, however, necessary to remove bound material from the DNA molecules completely. DNA molecules with bound protein may alter electrophoretic mobility or be resistant to some restriction enzyme digestion. The procedure is simple and the samples can be used for conventional electrophoresis and cloning experiments.

1.2.4 Identification of YAC clones

YAC libraries can be screened by PCR or hybridisation. Generally, YAC libraries are stored as individual clones in microtitre plates (Brownstein et al., 1989 ; Abidi et al., 1990 ; Anand et al., 1990). Clones are gridded directly from these plates onto filters for hybridisation either manually or by the use of a robotic device. Using this system, for example, the 60000 clones of the Washington University YAC library (Brownstein et al., 1989) are represented on 38 filters and primary hybridisation screens can be performed easily (Bentley et al., 1992 ; Nizetic et al., 1991).
For screening by PCR (Green and Olson, 1990a), aliquots of DNA from pools of clones representing the library (DNA pooled from master plates on which the library is stored) are used as templates in PCR with a pair of primers that amplify the appropriate specific target sequence. The above mentioned Washington University library is arranged in 38 pools, each pool consisting of 1536 YAC clones. The ICRF library (Larin et al., 1991) is arranged in 39 pools of 384 clones each and the ICI library (Anand et al., 1990) in 40 pools of 864 clones each. The presence of the target sequence in one of the clones of a pool directs amplification of the appropriate PCR product. In a second screen, aliquots of DNA from pools of a

number of clones which subdivide the primary pool (for example, 9 pools of 96 clones each subdivide the primary pools of the ICI library) are analysed. In the final step, PCR amplification of DNA from pools which subdivide the secondary pool leads to the identification of the individual positive clone. For example, PCR on pools of 8 rows and 9 columns from a single secondary pool of the ICI library identifies the co-ordinates of the positive clone.

The choice of screening method used depends upon each application. Hybridisation screening is preferable when using a probe which detects a multigene family, or when using complex probes, since the large number of clones identified would require an excessive number of PCRs to identify every positive clone. Conversely, screening exclusively by PCR is non-radioactive, more efficient when searching for a single clone, and is particularly suited to a sequence-tagged site (STS) strategy (Olson et al., 1989) for genome analysis and a YAC contig assembly.

1.2.5 Assembly of contigs

A major application of YACs is to determine the order and physical distance between landmarks that have been been previously assigned to specific regions of the genome. In the simpest cases, a single YAC is of sufficient length to link two or more landmarks directly. To span larger distances, however, it is necessary to assemble a contig of overlapping YACs. Contig assembly can be divided in two steps: isolation of the necessary YACs, and detection of overlaps. The strategy used to assemble contigs depends on the density of probes available, the lengths of the YAC inserts in the clones, and the compexities of the libraries being screened.

1.2.5.1 STS / Probe Content mapping

The initial YAC contigs to be assembled were in regions where the available probes or sequence data was of sufficient density both to isolate all the necessary YACs and to detect the majority of overlaps. Green and Olson (1990b) used 16 STSs to isolate and overlap 30 YACs spanning 1.5 Mb region containing the cystic fibrosis

transmembrane regulator (CFTR) gene on chromosome 7. STSs were derived from a combination of known probes, clones obtained by chromosome jumping, and sequences of individual exons of the CFTR gene.

A 2.6 Mb contig spanning the dystrophin gene was assembled using a similar STS-based approach (Coffey et al., 1992), in which 27 STSs were used to identify 34 YAC clones and detect all except one overlap (intron 1). The overlap in intron 1 was detected using Alu PCR fingerprinting (Nelson et al., 1989) of clones already isolated on either side of the intron, and confirmed by rescuing the end of the appropriate clone by Vectorette PCR. In a parallel study of the dystrophin gene (Monaco et al., 1992), YACs were isolated by hybridisation screening from a library with larger inserts and ordered by hybridisation with probes representing the full-length cDNA, without the need to analyse the ends of any YACs.

In chromosomal regions with insufficient probes to assemble contigs directly, two strategies can be adopted: (1) generate more (region-specific) probes or (2) perform chromosome walking. Several methods have been used to generate novel region-specific probes. Libraries of flow-sorted chromosomes in bacteriophage lambda (Fuscoe et al., 1989) or cosmids (Nizetic et al., 1991) have been constructed and used to generate probes or STSs. For small regions, microdissection of G-banded metaphase chromosomes spreads is a precise and direct way to isolate region-specific fragments (Ludecke et al., 1989).

Somatic cell hybrids have also been used for isolation of probes. The traditional approach is to make a genomic library from the DNA of a somatic cell hybrid and identify the human-specific clones (Gusella et al., 1980). A more efficient method is to use repeat specific primers to amplify DNA between repeats that are sufficiently close together (Nelson et al., 1989 ; Ledbetter et al., 1990). Cole et al. (1991) were able to obtain Alu PCR products at an estimated density of one per 75-100kb from a hybrid containing a segment of Xq26 and generate sufficient STSs to isolate and identify YACs spanning the entire region.

1.2.5.2 Chromosome walking: Isolation of YAC ends

Chromosome walking relies on isolation of a DNA fragment at or near an end of a cloned insert for use as a probe to screen the library and identify more clones. This strategy has been applied extensively to construct contigs in cosmids and lambda clones. To recover the end of a YAC is more difficult as the YAC is maintained as a stable single-copy chromosome in the host cell.

A number of methods have been developed to isolate YAC ends:

1. Subcloning the yeast DNA into bacteriophage lambda and screening with vector sequences. However, this is a labour-intensive approach (Green and Olson, 1990b).

2. Plasmid rescue: the pBR322-derived sequence of the ampicillin-resistance gene and origin of bacterial replication in the left arm of pYAC4 can be used to rescue an insert fragment attached to it following digestion and circularisation of the total YAC DNA and transformation of the products into E. coli. The right arm of YAC4 does not contain the bacterial sequences necessary for plasmid rescue, but they can be inserted before plasmid rescue.

3. By inverse PCR (Ochman et al., 1988 ; Triglia et al., 1988 ; Silver and Keerikatte, 1989), YAC clone DNA is digested with an appropriate enzyme and re-ligated. Among the circularized molecules are those containing vector sequences joined to the end of the YAC insert. Digestion with a second restriction enzyme which cuts exclusively in the vector sequence yields a linear molecule with the end of the insert flanked by two segments of the vector arm. Two vector primers can be used to amplify the required insert sequence (Silverman et al., 1991).

4. In the Vectorette PCR method (Riley et al., 1990), the YAC clone DNA is digested and ligated to the Vectorette, a pair of annealed oligonucleotide universal adaptors which have an internal region of non-complementarity, or "bubble" (fig. 25). Terminal sequences are amplified using a vector arm specific primer and a primer (Vectorette primer) which has the same sequence as the bottom strand of the mismatched region of the Vectorette. The vector arm specific primer (directed towards the sequence of interest) produces a complementary sequence using the bottom strand

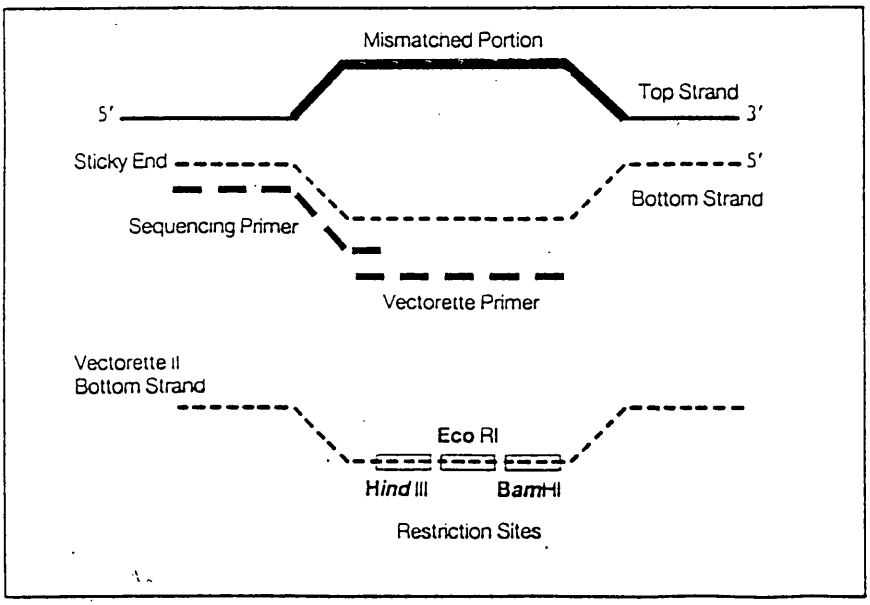


Figure 25: Schematic of a Vectorette unit or "bubble".

of the Vectorette as template in the first cycle of PCR. This provides the only available target for the Vectorette primer. PCR continues normally and results in synthesis of a product which contains vector, insert end and a part of the Vectorette sequence. The same Vectorette library can be used with different specific primers in combination with the bubble primer to amplify any sequence which is adjacent to a known sequence. The Vectorette method was first used to walk between D7S8 and D7S23 thus establishing a 1.5-Mb contig spanning the CFTR gene region (Anand et al., 1991).

5. An alternative approach is to use Alu- and vector-specific primers in combination to amplify between either end of the vector and the nearest human repeat sequence in the insert (Nelson et al., 1991 ; Ragoussis et al., 1991). The method has the advantage of being applicable directly to YAC colonies, and is thus the only method which avoids preparation of YAC clone DNA.

1.2.5.3 Random assembly / Fingerprinting

The methods described above rely on extensive characterization and use of single-copy probes and/or STSs to determine overlaps between YAC clones. For efficient handling of large numbers of clones, or for analysing YACs in regions where there is no prior information, a pattern of fragments (fingerprint) containing repeat sequences present in each YAC may be compared. This can be achieved by Southern analysis of individual YACs using repetitive sequence probes (Alu fingerprinting ; Wada et al., 1990). An alternative is to use Alu PCR (Nelson et al., 1989) which can be applied directly to YAC colonies, and generates an average of 10 bands between 0.1 and 2.0 kb per YAC. This method has been used, for example, to confirm the overlap between two YAC clones in the class II region of the human MHC (Ragoussis et al., 1991).

1.2.5.4 Chimaeric YACs

Detailed analysis of YACs has resulted in the identification of a proportion of clones in most libraries which are chimaeric i. e. they contain ligated fragments derived from different regions of the genome. Chimaeric clones will affect the overlapping and mapping of YAC clones to varying degrees depending on the strategy used. Thus STS or probe content mapping is relatively unaffected by chimaeric clones (Green and Olson, 1990b ; Coffey et al., 1992 ; Monaco et al., 1992), whereas chromosome walking is critically dependent on detecting and circumventing chimaeric ends of YACs. All YAC ends used for walking should be analysed (e. g. Silverman et al., 1991). Chimaeric clones may be detected by in situ hybridisation of the YAC to metaphase chromosome spreads, in which the resulting signal is observed at two distinct genomic sites. Alternatively, the ends of the YAC may be recovered and used for analysis of a monochromosome somatic cell hybrid panel (Anand et al., 1991).

Two mechanisms have been proposed to explain the occurence of chimaeras. One is that two fragments of insert DNA are co-ligated prior to transformation. The alternative is that co-transformation of more than one ligated DNA recombinant into the same cell is followed by recombination between homologous sequences. The second model was supported by the results of detailed analysis of the chimaeric clone yCF1 (Green and Olson, 1990), in which the sequence at the chimaeric breakpoint was found to consist of an Alu repeat, and no EcoR1 site (the enzyme used to prepare the insert DNA prior to ligation) was found (Green et al., 1991).

1.2.5.5 Construction of cosmid and bacteriophage contigs from YACs

The isolation of cosmids in a contiguous array representing the genomic insert of a YAC is useful for detailed analysis and identification of new landmarks. Cosmid or phage clones representing a YAC insert can be used as a basis for the isolation of highly informative polymorphic sequences such as CA repeats (Yulle et al., 1991) or novel coding regions. To obtain complete representation of a YAC insert in a

bacterial cloning system two approaches can be adopted, 1. cosmid or bacteriophage libraries can be constructed from complete yeast DNA of a YAC clone. The libraries are then screened with human DNA to identify clones representing the YAC insert. The clones are then overlapped by fingerprinting (Bellane-Chantelot et al., 1991) or by hybridising a panel of probes isolated from the ends of some clones to a filter containing the clones (Pieretti et al., 1991). 2. In the second approach, comprehensive resources of cosmids in microtitre plates, constructed for example from flow-sorted chromosomal DNA (Nizetic et al., 1991), are used as a source of bacterial clones representing a YAC insert. The YAC insert is gel-purified, labelled and used as a probe after suppression of the repetitive sequences to identify the corresponding cosmids. An important consequence of this approach is that it provides useful overlapping information when related YACs are used in separate hybridisations of the same cosmid resource (Coulson et al., 1988).

2. Materials and methods

2.1 PCR conditions for different STSs

2.1.1 YAC library screening by PCR (see section 4.1 of chapter B for details)

The ICI YAC genomic library was screened with three STSs by PCR. Preparation of DNA pools for PCR amplification is described in section 4.1.1 of chapter 2. PCR amplifications were carried out in a 25μl reaction volume containing an aliquot (3μl) of each pool, 0.2μM of each primer, 200μM of each dNTP and 1 unit of Taq polymerase. A hot start step was performed before the addition of the enzyme (section 4.1.1, chapter 2). Samples were processed through 30 cycles of denaturation (94^0C for 1 minute), annealing for 1 minute (for DXS426 at 52^0C, for DXS1003 and for 33CA11/ Right end clone at 60^0C) and extension (72^0C for 1 minute). For each STS the following PCR primers were used:

1) DXS426 locus

5' CCT TCA TCT CTA CCA AGA TA 3' and 5' CTG CAC TCC AGC CTG AAT AA 3' (Luty et al., 1989 ; Coleman et al., 1990). Expected size of PCR product: 290bp.

2) DXS1003 (AFM276xf5) locus (Weissenbach et al., 1992)

Mappairs primers (Research Genetics) were used. Expected size of PCR product: 185bp.

3) STS from 33CA11/ Right end clone

5' GAA TTC CAG CTG AGA AAT GC 3' and 5' AGG CTC GGT CTC AAA TGC CT 3'. Expected size of PCR product: 134bp.

2.1.2 STS analysis of YACs

The newly isolated YAC clones were tested for the presence of STSs by PCR. An aliquot of each plug (1/8 of a 200μl plug) was rinsed three times in distilled sterile H_2O, melted at 65^0C and diluted 1:1 with distilled H_2O. An aliquot (3μl) of the diluted sample was used for PCR analysis in a 25μl reaction volume containing 0.1 μM of each primer, 200μM of each dNTP and 1 unit of Taq polymerase. For colony PCR (Huxley et al., 1990), a small part of colony picked directly from the agar plate or stab was diluted in 10-20μl of TE. An aliquot (5μl) of the diluted sample was used for PCR analysis. PCR samples were denatured for 5 minutes before the addition of the enzyme. Samples were processed through 35 cycles of denaturation (94^0C for 1 minute), annealing for 1 minute (for properdin P factor, synapsin I, UBE 1 and 20CC8/Right end clone at 60^0C, for 33CA11/Left end clone at 45^0C and for 20CC8/Left end clone at 65^0C or 68^0C) and extension at 72^0C for 1 minute (2 minutes for UBE1). For each STS the following primers were used:

1) DXS426

PCR conditions were described in the previous section.

2) Properdin P factor (PFC ; Coleman et al., 1991)

5' CCT GAG GAT AGT GTC AGC GAT 3' and 5' CTT TCA GGG CTA CTG GTC ACT 3'. Expected size of PCR product: 224bp.

3) Synapsin I (Kirchgessner et al., 1991)

5' GAA TAT GAA AGC TCA GAG G 3' and 5' CCA GAG CCC TGT GGG CAT C 3'.

Expected size of PCR product: 202bp.

4) UBE 1 (ubiquitin gene)

5' GAG CGG GGA CTT TGT CTC CT 3' and 5 ' CTT TGA CCT GAC TGA CGA T 3'

Expected size of PCR product: 700bp. Human DNAs were used as positive controls.

5) STS from 33CA11/Left end clone

5' ACT ACG GAA TTC CCA CTT 3' and 5' AAC TTA CTT GGT CTC TTG 3'.

Expected size of PCR product: 111bp.

6) STS from 33CA11/Right end clone

Described in the previous section.

7) STS from 20CC8/Left end clone

5' AGG CCA GGT GCA GTG ACT CAT G 3' and 5' TTG CCA TGT TGC CCA GGC TGG 3'. Expected size of PCR product: 107bp

8) STS from 20CC8/Right end clone

5' CTG ATG CCA CTG AGT TAT AGG 3' and 5' CTT AGG TAA ATA TTC TGG GGC C 3'. Expected size of PCR product size: 101bp.

2.2 Isolation of YAC ends using the Vectorette system

2.2.1 YAC 33CA11

Two libraries were constructed (as described in section 4.6.1 of chapter 2) using the enzymes RsaI and HaeIII. PCR amplification of the RsaI library was carried out for the left end of the clone and of both libraries for the right end of the clone. Protocols were as described in section 4.6 of chapter 2. During PCR amplification, the Universal Vectorette primer was unphosphorylated and the vector arm primers phosphorylated. PCR products were analysed on 1% agarose gel. They were then partially purified by phenol extraction and converted into single stranded DNA using

the Lambda exonuclease method (4.7.1 a, chapter 2). The products were sequenced using the ^{32}P end labelled left and right vector arms sequencing primers (4.7.2 a).

2.2.2 YAC 20CC8

Two libraries were constructed (section 4.6.1, chapter 2) using the enzymes RsaI and EcoRV. Protocols for PCR amplification of the two libraries were modified. PCR amplifications were carried out for 40 cycles of denaturation at 94^0C for 1 minute, annealing at 62^0 (or at 65^0C) for 1 minute and extension at 72^0C for 3 minutes. PCR products were analysed on a 2.5% agarose gel. Components in the PCR reactions, size of PCR products and sequencing method were as follows:

1) Right arm of the YAC

PCR components: 5μl of Vectorette library,10μl of 10x PCR buffer, 100μM of dNTPs, 1μM of phosphorylated pYAC4 right arm primer, 1μM unphosphorylated Universal Vectorette primer. PCR products of 400bp and ~1kb were obtained for RsaI and EcoRV libraries, respectively.

2) Left arm of the YAC

PCR components: 10μl of Vectorette library, 10μl of 10x PCR buffer, 100μM of dNTPs, 1μM (Rsa I library) or 0.5μM (EcoRV library) of phosphorylated pYAC4 left arm primer, 1μM (RsaI library) or 0.5μM (EcoRV library) unphosphorylated Universal Vectorette primer. PCR products of 500bp and 400bp were obtained for RsaI and EcoRV libraries, respectively.

The RsaI library PCR products were column purified, converted into single stranded DNA using the Lambda exonuclease method (4.7.1 b, chapter 2) and sequenced using YAC arm sequencing primers for ^{32}P incorporation labelling (4.7.2 b, chapter 2). Sufficient sequence data were obtained, therefore EcoRV PCR products were not further analysed.

2.3 Isolation of a $(CA)_n$ repeat from 33CA11 YAC

2.3.1 Isolation of cosmid clones at the end of 33CA11 YAC

Following the isolation and sequencing of the end clones of YAC 33CA11, PCR amplification primers were designed and used to amplify 33CA11 YAC DNA and genomic DNA. The PCR products of the right end clone were used to screen the ICRF (Hans Lehrach et al., 1990) reference X-specific cosmid library no. 104 (L4/FCS X). This library was constructed by Dean Nizetic (Nizetic et al., 1991) from digests of DNA from "Flow-sorted" Human X-chromosome and ligated into Lawrist4 vector. The screening was performed (by Dr Fiona Francis) in ICRF laboratories and resulted in six positive clones (ICRFc104 C096, ICRFc104 D0480, ICRFc104 F1080, ICRFc104 D01210, ICRFc104 D08214 and ICRFc104 D10158). Bacterial stabs of these cosmid clones were streaked out on LB agar plates supplemented with 50μg/ml kanamycin and incubated at 37^{0}C overnight. A single colony was picked from each plate and used to inoculate 10ml of sterile LB broth containing 50μg/ml kanamycin. Following incubation at 37^{0}C with agitation, DNA was isolated and purified using the Magic Miniprep^{TM} DNA Purification System (Promega) according to manufacturer's instructions (except that ~9ml of culture was used instead of the recommended 1-3ml). DNA was eluted from the purification Minicolumn in 50μl TE and an aliquot (5μl) of each preparation was electrophoresed on a 0.8% agarose gel to check sample quality and estimate DNA concentration. An aliquot (3μl of 1:1000 dilution) of each cosmid DNA was PCR amplified using the primers of 33CA11/Right end clone to identify true positive clones.

2.3.2 A search for CA repeats in a pool of cosmids

An aliquot (10-20μl) of each positive cosmid clone was digested with 20units HindIII in a 25μl reaction containing 2.5μl of 10x appropriate restriction buffer at 37^{0}C for 2 hours and the entire reaction was electrophoresed on a 0.8% agarose gel alongside λHindIII and a 100bp ladder (Gibco BRL) as size markers. The gel was blotted

(using Hybond N+ membrane from Amersham, according to manufacturers' instructions) and the filter hybridised with an oligonucleotide $(GT)_{11+1}$ probe to identify $(CA)_n$ repeats in the cosmids.

2.3 3 Isolation of a $(CA)_n$ repeat from a cosmid using the Vectorette system

2.3.3.1 Construction of Vectorette library

A Vectorette library was constructed from cosmid ICRFc104 D0*1210* using the enzyme AluI as follows: An aliquot (15µl) of cosmid DNA was digested with 20units of AluI (Pharmacia) in a 50µl reaction volume containing 1x restriction buffer at 37^{0}C for 11/2 hours. Following digestion, 5µl (3pmol) of the appropriate Vectorette units (CRB), 1µl of a 100mM ATP and 5units of T4 DNA ligase (Pharmacia) were added to the same microtube. The ligation reaction was incubated at 37^{0}C for two hours after which 100µl of SDW was added and the Vectorette library was stored in small aliquots at -20^{0}C

2.3.3.2 Vectorette PCR

PCR amplification of the Vectorette library was performed using the phosporylated Universal Vectorette Primer (CRB) in combination with either primers C_4 $(CA)_{11}$ or G_4 $(GT)_{11}$ (both unphosporylated). For PCR amplification 2µl and 5µl of Vectorette library, 10µl of 10x PCR buffer, 100µM of each dNTP and 1µM (or 0.5µM) of each primer were mixed and the volume made up to 99µl with distilled sterile H_2O. The sample was denatured at 95^{0}C for 5 minutes and 2.5units of Taq polymerase added at 90^{0}C. PCR conditions were as follows: denaturation at 94^{0}C for 1 minute, annealing at 65^{0}c for 1 minute and extension at 72^{0}C for 40 cycles.

2.3.3.3 Direct sequencing of the Vectorette PCR products

Single stranded DNA was prepared using the Lambda exonuclease method (section 4.7.1b, chapter 2) and sequenced using the Vectorette sequencing primer (section 4.7.2b, chapter 2).

2.3.3.4 PCR amplification

The primers used for PCR amplification of cosmid or genomic DNA were: 5' TTA GGG TGG GCT TTA ATC TAA 3' (CA strand) and 5' GTG TTT CTT GGC TTA CAG ATG 3' (GT strand). PCR ampifications were carried out non-radioactively and radioactively as described in section 3.4 of chapter B. PCR condtitions for both non-radioactive and radioactive reactions were as follows: An initial denaturation at 95^0C for 1 minute followed by 30 cycles of denaturation at 94^0C for 1 minute, annealing at 56^0C (or 58^0C) for 1 minute and extension at 72^0C for 1 minute. Non-radioactive PCR products were analysed on 2% agarose gels and radioactive products on 6% polyacrylamide sequencing gels.

2.3.4 Verification of the CA repeat

Following PCR amplification of cosmid (from which the CA repeat derived) and genomic DNA, PCR products were electrophoresed on a 2% agarose gel alongside pBR322 digested with AvaII and EcoRI as size marker. The gel was blotted (using Hybond N+ membrane from Amersham, according to manufacturers' instructions) and the filter hybridised with an oligonucleotide $(GT)_{11+1}$ probe to verify the presence of the CA repeat.

3 Results

3.1 Isolation of YACs containing the DXS426 locus

The ICI YAC genomic library was screened by PCR with DXS426. A YAC containing DXS426 locus (DXS426/4, a gift from Dr Mick Coleman ; positive control) and a no

DNA control were included in the screening experiments. Two primary positive pools, 4 and 33, were identified. Secondary pools 4H and 33C were positive. The tertiary pools 7 and 10 for secondary pool 4H, corresponding to clone G2 (see table in section 4.1.3, chapter 2), and pools 1 and 19 for secondary pool 33C, corresponding to clone A11, were positive (data not shown). The combined results from the three step screening of the library led to the identification of two positive for DXS426 YAC clones, 4HG2 and 33CA11.

DNA from these clones was prepared in plugs (section 4.3, chapter 2) and electrophoresed by pulsed field gel electrophoresis (section 4.3, chapter 2) to check quality of the preparation (fig. 26a). Saccharomyces cerevisiae (Strain YNN295 ; Biorad) was used as size marker. The two YACs were not instantly visible but there was indication that they coincided with yeast chromosomal bands. Alkali blotting of the pulsed field gel and hybridisation with total human DNA (sections 4.4 and 4.5, chapter 2) confirmed this, sizing clone 4HG2 at 245kb and clone 33CA11 at 290kb (fig. 26b).

3.2 STS analysis in YACs 33CA11 and 4HG2

YACs 4HG2 and 33CA11 were analysed for the presence of DXS426, PFC and synapsin loci by PCR. Two YACs containing DXS426 locus (DXS426/1 and DXS426/6, gifts from Dr Mick Coleman) and genomic DNA were used as positive controls. A YAC from a different region was used as a negative control. Both clones contained the DXS426 locus (fig. 27) thus confirming the library screening results. Clone 4HG2 contained PFC whereas 33CA11 did not (fig. 28). Synapsin was absent from both YACs (data not shown).

3.3 Isolation of end clones from YAC 33CA11

Amplification of an RsaI Vectorette library gave single PCR bands of ~0.35kb for the left end and of ~0.3kb for the right end (fig. 29). The pYAC4 left and right arm specific primers used for the ampilification of the library are approximately 250

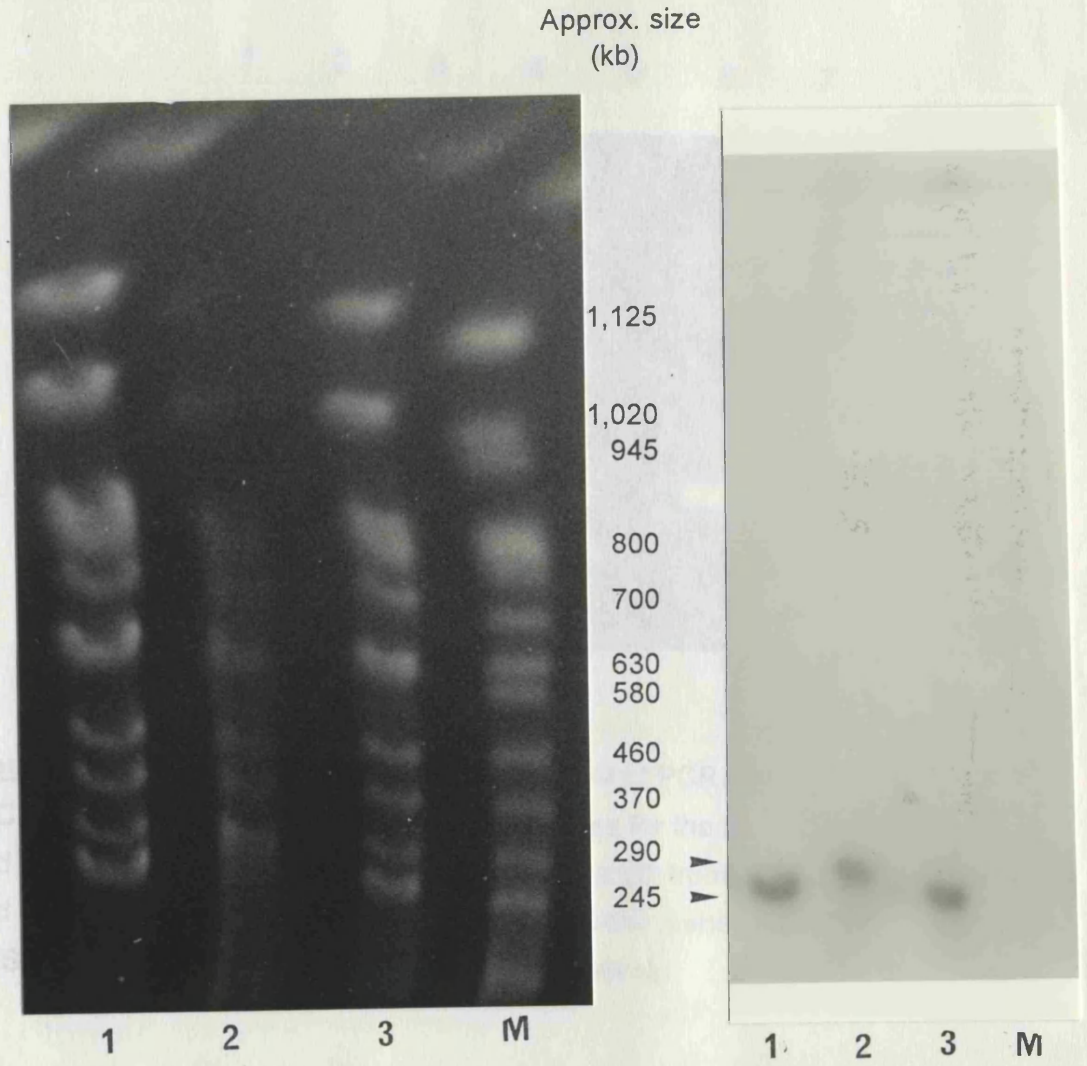


Figure 26: a. PFGE electrophoresis of 33CA11(lane 2) and 4HG2 (lane 3) YACs. Lane 1 is a YAC clone of known size (245kb). M denotes the Saccharomyces cerevisiae (strain YNN295) size marker.

b. Autoradiograph showing hybridisation of total human DNA to a Southern blot of gel (a). 33CA11 was sized at 290kb and 4HG2 at 245 kb.

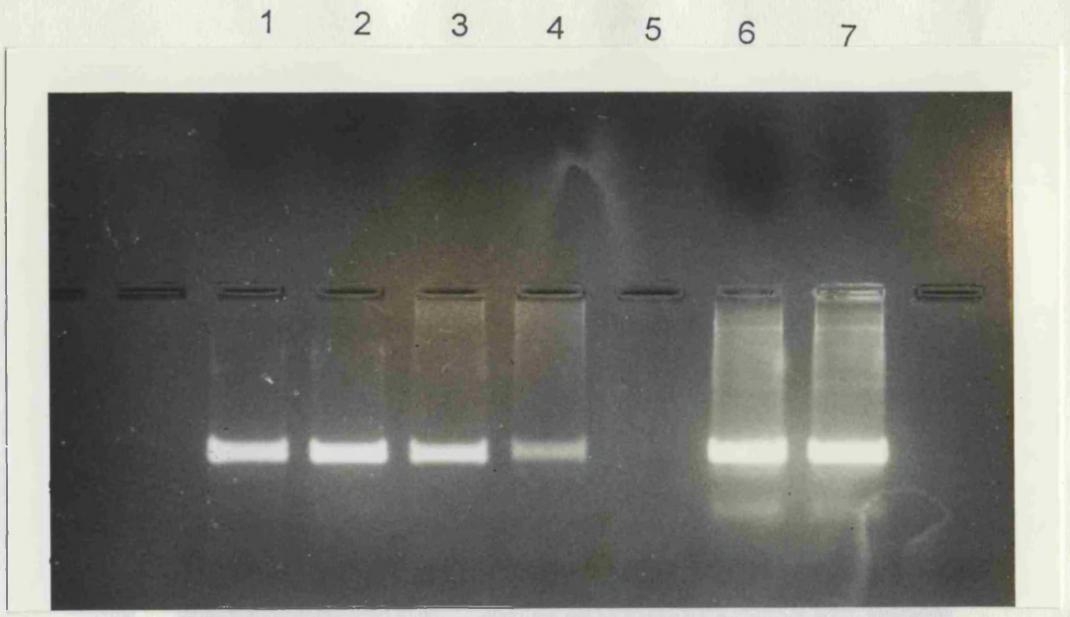


Figure 27: Ethidium bromide-stained agarose gel of PCR products from YACs 33CA11 (lane 6) and 4HG2 (lane 7) using primers for the DXS426 locus. Lanes 1 and 2 are positive controls (two DXS426 YACs, a gift from Dr M. Coleman). Lanes 3 and 4: Genomic DNA also used as positive controls. Lane 5: A YAC containing the DXS255 locus which was used as a negative control.

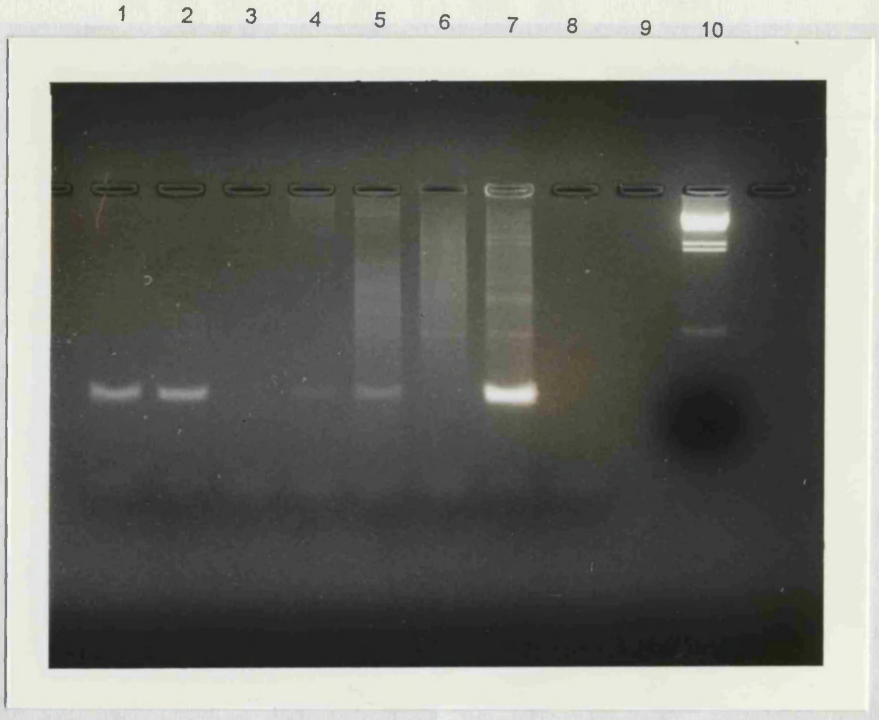


Figure 28: Ethidium bromide-stained agarose gel of PCR products from 33CA11 and 4HG2 YACs using primers for the properdin P factor (PFC). The PFC locus is present in the distal 4HG2 YAC (lane 7) and absent from the proximal 33CA11 YAC (lane 6). YACs DXS426/1 and DXS426/4 (lanes 1 and 2 ; gifts from Dr M. Coleman) were known to contain the PFC locus. Lane 3 is a "no DNA" control. Lanes 4 and 5: Human genomic DNA also used as positive control. The size marker M is λ DNA digested with HindIII.

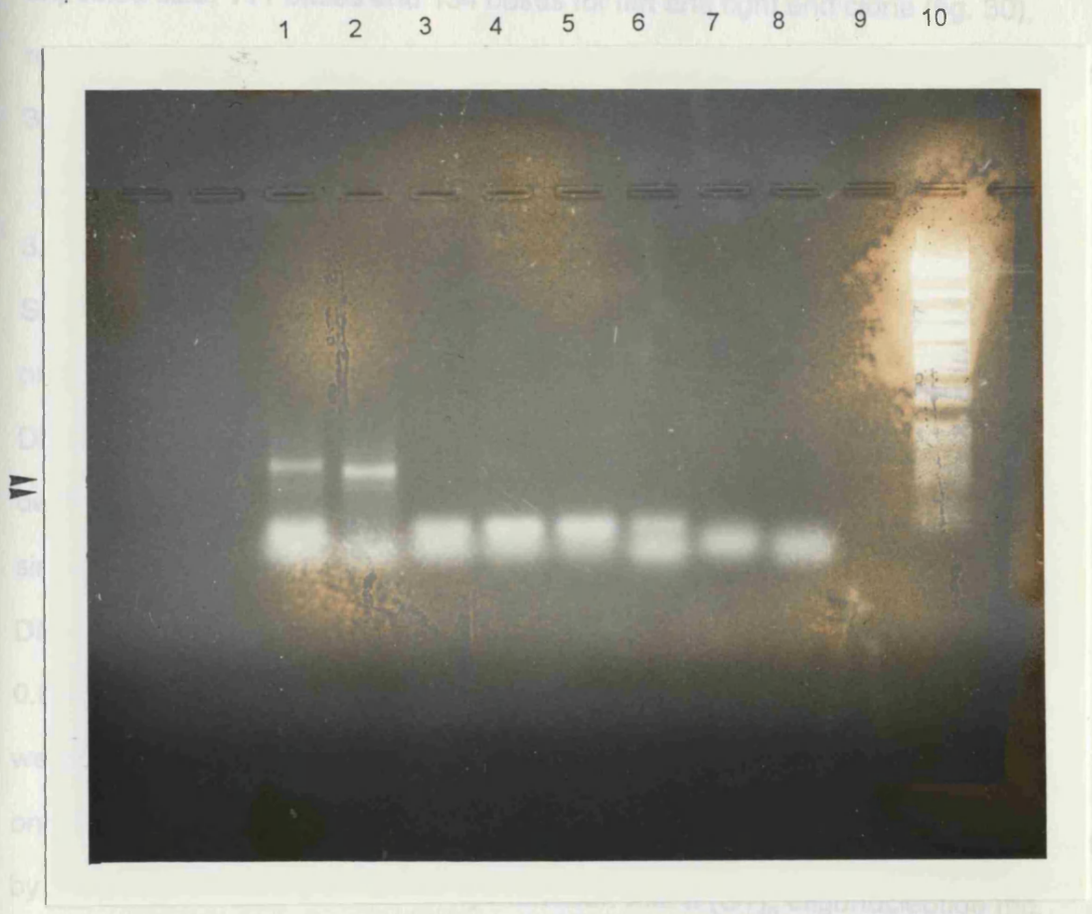


Figure 29: Vectorette system for the isolation of YAC termini from YAC 33CA11. PCR amplification of an RsaI Vectorette library of YAC 33CA11 with Vectorette Universal PCR primer and primers specific for the left arm (lane 1 ; ~ 0.35kb) or the right arm (lane 2 ; ~ 0.3kb) of pYAC4. Lanes 3-8 are negative controls. Lane 10: Size marker: λ DNA digested with AvaII.

bases and 140 bases, respectively, away from the EcoRI cloning site. Left and right sequencing primers are 7bp and 19bp away from the EcoRI cloning site. Therefore, end clones of considerably smaller size, compared to that of the PCR products, were expected. Following sequencing of the PCR products, sequences of the expected size, 111 bases and 134 bases for left and right end clone (fig. 30), respectively, were obtained. Primers were designed and used for PCR analysis of 33CA11. Positive results of the expected size confirmed their origin (fig. 45).

3.4 Isolation of a CA repeat from the right end of YAC 33CA11

Six cosmids were isolated from the ICRF X-specific cosmid library using the PCR product for the right end clone of YAC 33CA11 as a probe. PCR amplification of DNA from these cosmids using primers from the right end clone of the YAC demonstrated that cosmids D0480, F1080 and D01210 were true positive clones since they contained this STS (fig. 31).

DNA from the positive cosmid clones digested with HindIII was electrophoresed on a 0.8% agarose gel. The different HindIII pattern of the cosmids indicated that they were not identical (fig. 32a). They had, however, three common HindIII fragments, one of which was approximately 5kb in size and contained a $(CA)_n$ repeat identified by Southern analysis followed on hybridisation with a $(GT)_n$ oligonucleotide (fig. 32b).

An AluI Vectorette library constructed from cosmid D01210 was PCR amplified using the Vectorette Universal primer in combination with either a $(GT)_n$ or a $(CA)_n$ primer. Two PCR products were obtained at ~200 base pairs and 450 base pairs with $(GT)_n$ and $(CA)_n$ primers, respectively (fig. 33). The PCR products were purified, prepared as single stranded DNA and sequenced using the Vectorette sequencing primer. A stretch of 11 CAs was revealed at the end of the 200bp PCR product (fig. 34a) while a stretch of 11 GTs was revealed at the end of the 500bp PCR product (fig. 34b). A total of 83 bases from the CA strand and 213 bases from the GT strand are shown in fig. 35a-b.

Figure 30a: Nucleic Acid Sequence from YAC 33CA11 Left End Clone

```
        10          20          30          40          50          60
Forward primer---->
ACTACGGAAT TCCCACTTNN NTCTGTCAAA CGTATACCTC TCTCGATGAC CCTGTAGTAG
TGATGCCTTA AGGGTGAANN NAGACAGTTT GCATATGGAG AGAGCTACTG GGACATCATC

        70          80          90          100         110

CAACCAGAAC CCCTCACTTT GGCTCAGTCT TCCCAAGAGA CCAAGTAAGT TAAGGAG
GTTGGTCTTG GGGAGTGAAA CCGAGTCAGA AGGGTTCTCT GGTTCATTCA ATTCCTC
                                  <-----Reverse primer<-----

AGGACAGCGA TTCTCGTACG AAC
TCCTGTCGCT AAGAGCATGC TTG
----Vectorette unit------>
```

Figure 30b: Nucleic Acid Sequence from YAC 33CA11 Right End Clone

```
        10          20          30          40          50          60
Forward primer------>
GAATTCCAGC TGAGAAATGC TAATAAAAGC AATGATCACG AGGTCCAGGT GCAGTGACTC
CTTAAGGTCG ACTCTTTACG ATTATTTTCG TTACTAGTGC TCCAGGTCCA CGTCACTGAG

        70          80          90          100         110         120

ATGCCTGTAA TTCCAGCACT TTGGGAGGCC AAGGTGGGTA GATCGCTTCA GCCCAGGCAT
TACGGACATT AAGGTCGTGA AACCCTCCGG TTCCACCCAT CTAGCGAAGT CGGGTCCGTA
                                                          <-----
       130

TTGAGACCGA GCCTGG
AACTCTGGCT CGGACC
-Reverse primer
```

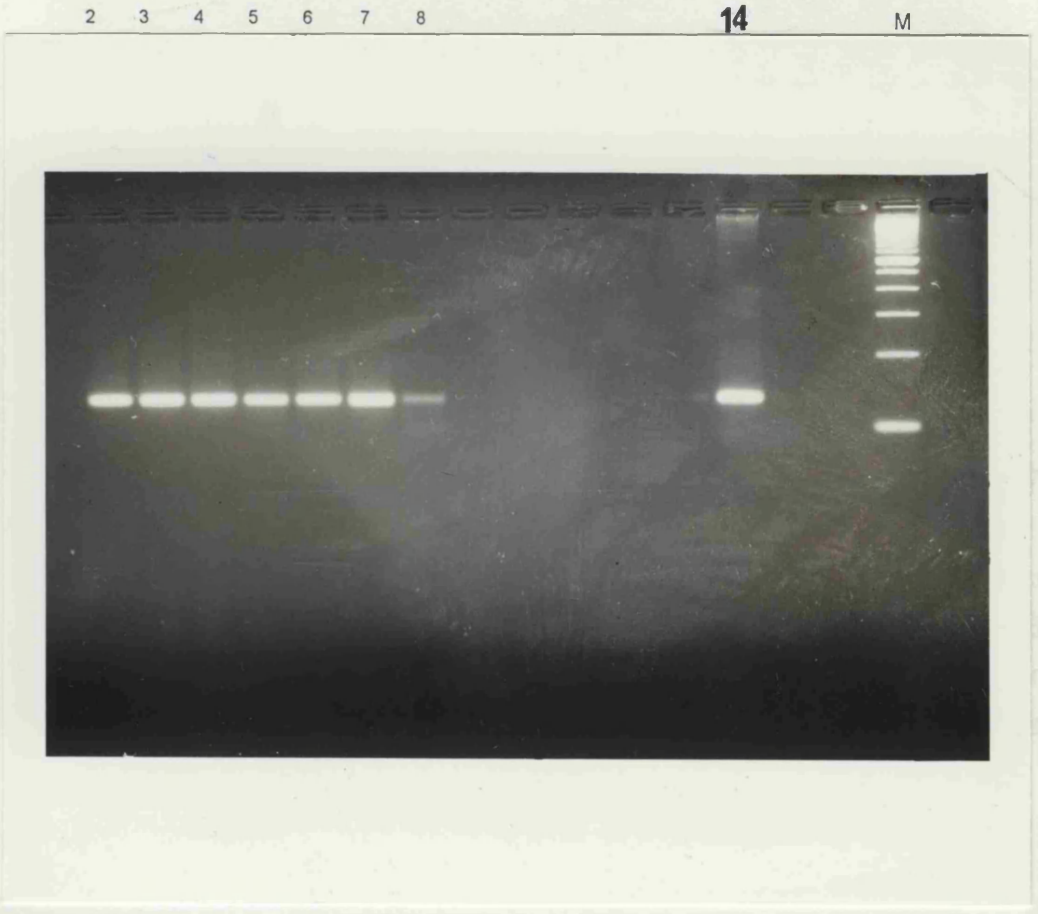


Figure 31: Ethidium bromide-stained agarose gel of PCR products from ICRF cosmids using specific primers for the right end of YAC 33CA11.
Lanes 2, 3 and 4: Cosmid ICRFc104 D0480. Lanes 5 and 6: Cosmid ICRFc104 F1080. Lanes 7 and 8: Cosmid ICRFc104 D01210.
Lane 13: Cosmid HX43 was used as a negative control. Lane 14: YAC 33CA11 was a positive control. Lane 15: A "no DNA" control. M denotes a 100 base pairs ladder size marker.

Figure 32: a. Ethidium bromide-stained agarose gel electrophoresis of DNA from cosmids D0480 (lanes 3 and 4), F1080 (lanes 5 and 6) and D01210 (lanes 7 and 8) digested with HindIII. Lane 10: DNA from cosmid HX43 digested with HindIII Lane 11: undigested DNA from cosmid HX65. Cosmids in lanes 10 and 11 were known to contain a microsatellite. Lanes 1 and 13: Size markers (λ/HindIII) Lane 14: Size marker: a 100bp ladder. b. Autoradiograph indicating the presence of a CA repeat in a ~5kb HindIII fragment (marked with arrowhead in a) following blotting of gel (a) and hybridisation with a $(GT)_n$ oligonucleotide.

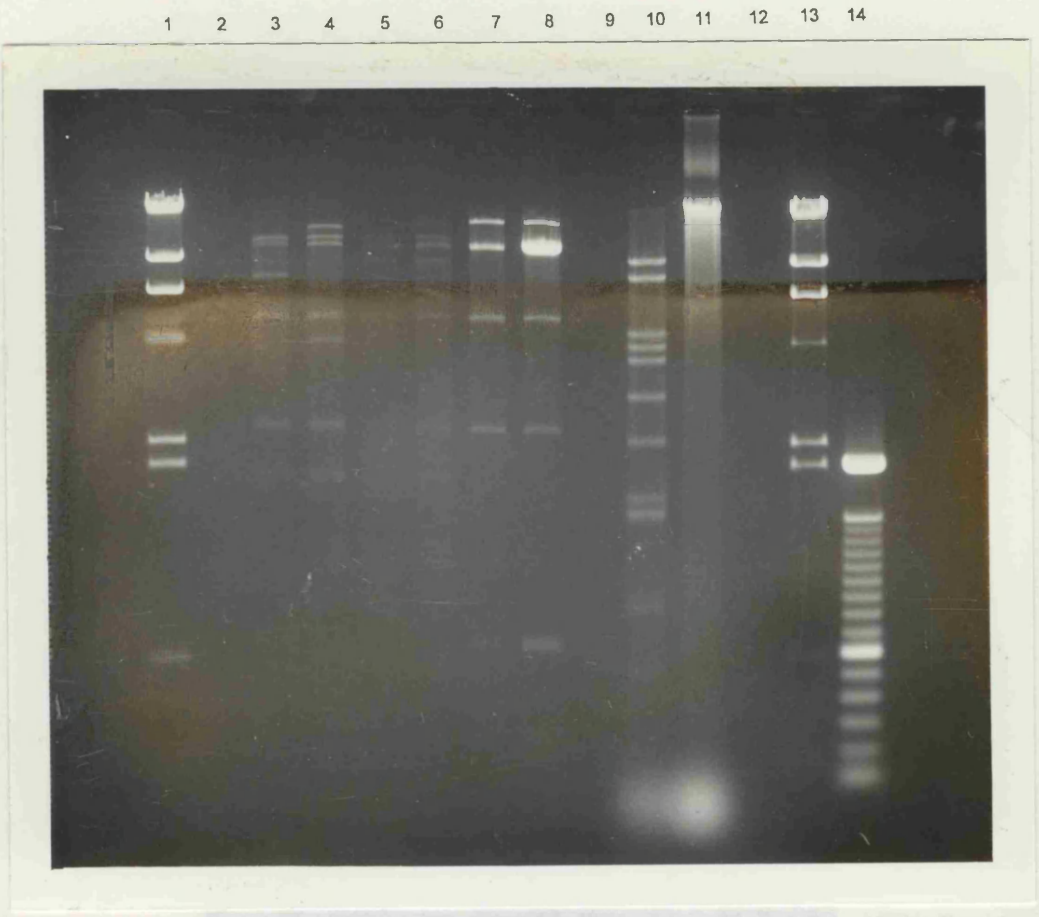


Fig. 32a

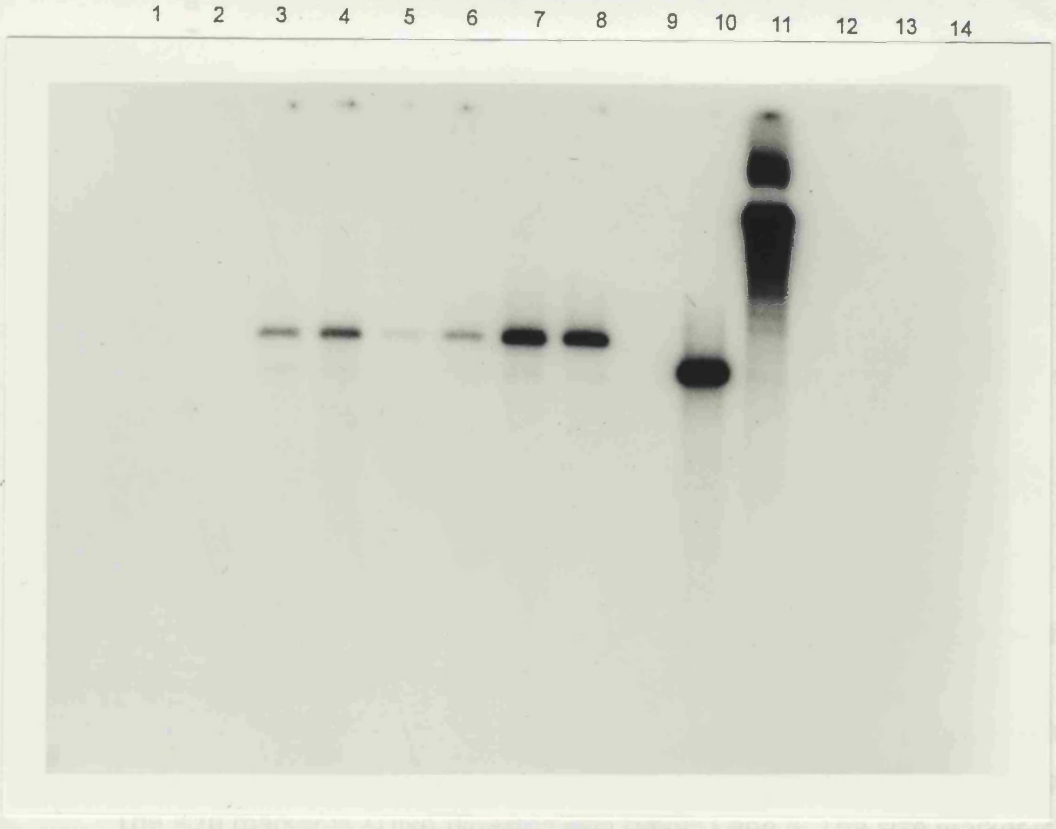


Fig. 32b

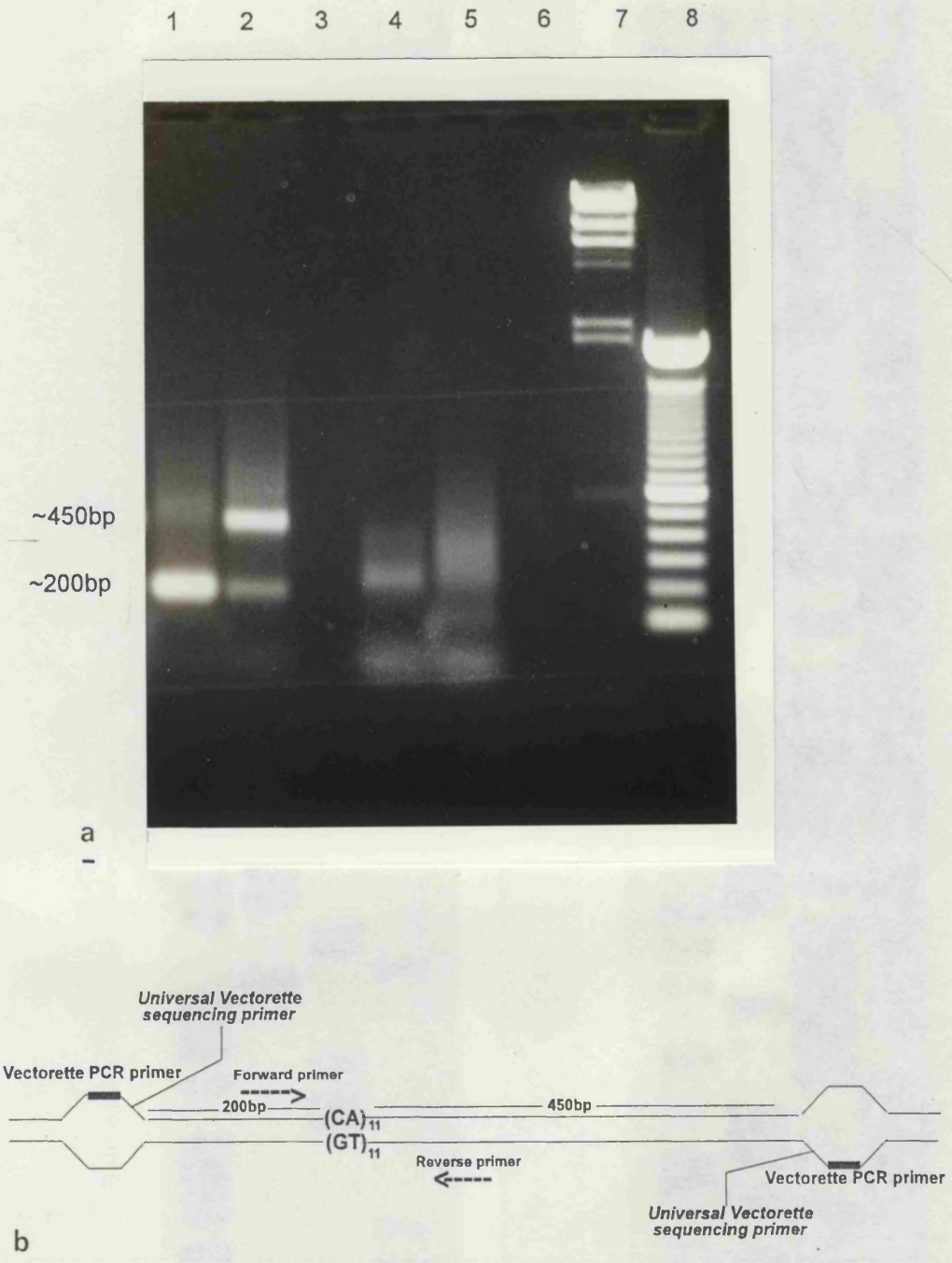


Figure 33: a. Vectorette system used for the isolation of a CA repeat from cosmid D01210: PCR products obtained from AluI Vectorette library using the phosphorylated Universal Vectorette Primer in combination with either primers G_4 $(GT)_{11}$ (Lane 1) or C_4 $(CA)_{11}$ (Lane 2). Lanes 4 and 5 are "no DNA" controls. Lane 7: The size marker is λDNA digested with HindIII Lane 8: The size marker is a 100bp ladder. b. A schematic representation of the isolation of a CA repeat from cosmid D01210 using the Vectorette system.

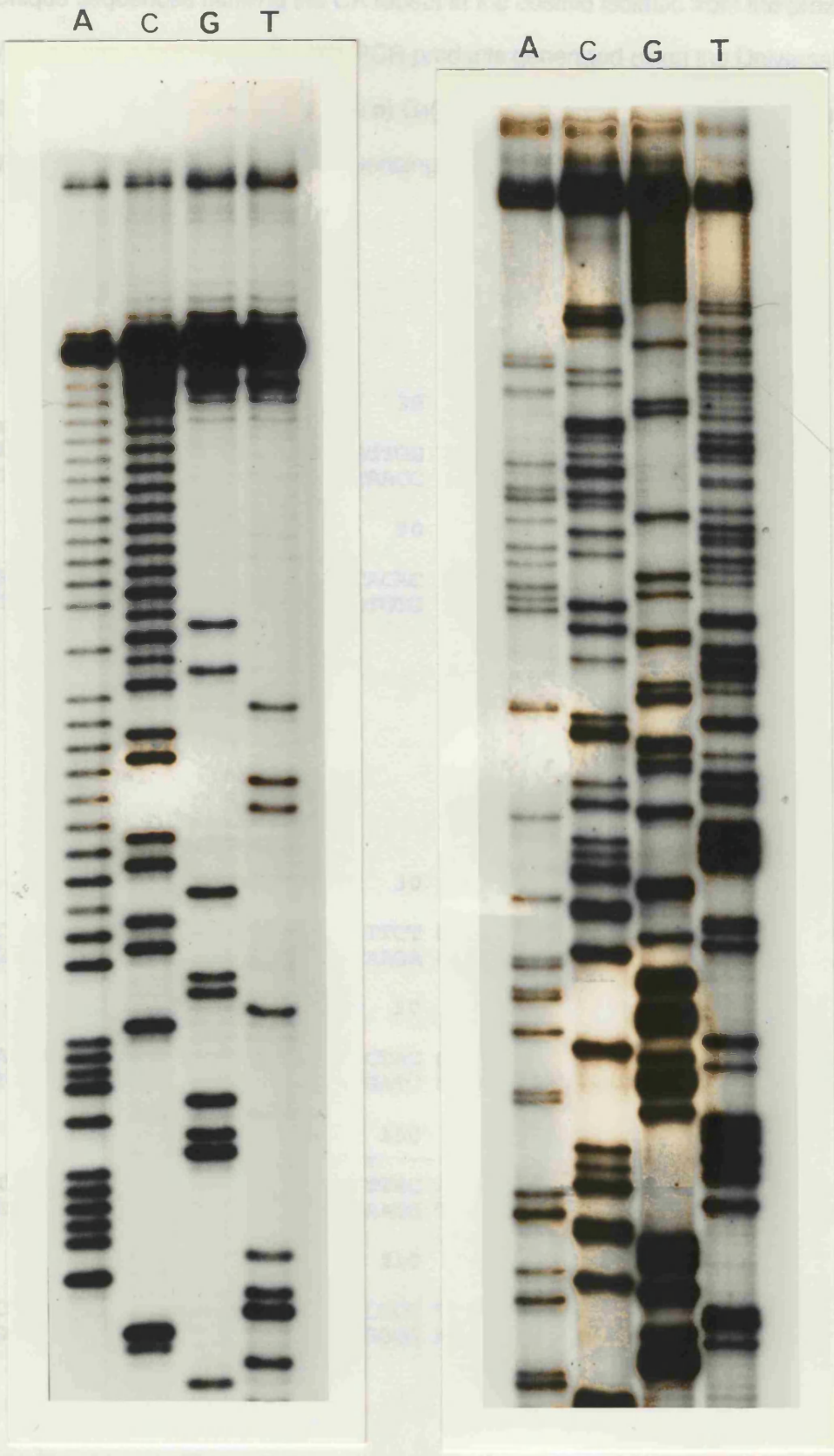


Figure 34a-b: Autoradiographs of DNA sequencing gels showing the poly (CA) sequence within the PCR products obtained as described in fig. 33a. A schematic model for the isolation of the CA repeat is shown in fig. 33b.

Figure 35: Unique sequences flanking the CA repeat in the cosmid isolated from the proximal end of YAC 33CA11. The templates are PCR products generated using the Universal Vectorette primer in combination with **a)** $G_4(GT)_{11}$ and **b)** $C_4(CA)_{11}$ primers. They were sequenced using the Vectorette sequencing primer.

a.

```
        10          20          30          40          50          60
Forward primer-------->
TTAGGGTGGG CTTTAATCTA ATATAATTGG TGTCCTTATA AAAAGGAGAA AACTGGACAC
AATCCCACCC GAAATTAGAT TATATTAACC ACAGGAATAT TTTTCCTCTT TTGACCTGTG

        70          80          90         100

AGACACATAT ACACATACGC ACGCACACAC ACACACACAC ACACA
TCTGTGTATA TGTGTATGCG TGCGTGTGTG TGTGTGTGTG TGTGT
```

b.

```
        10          20          30          40          50          60

AACAAGAGAA ATTTATTGTC TCACAGTTCT GGAGGCCAGA ACTCCAAAAC CAAGGTGTTA
TTGTTCTCTT TAAATAACAG AGTGTCAAGA CCTCCGGTCT TGAGGTTTTG GTTCCACAAT

        70          80          90         100         110         120

GCAGGGCCAT ACGCGCTTTG AAGGTGCTAG GGAAGGAACT GTTCCAGGCC TCTCTTAGCT
CGTCCCGGTA TGCGCGAAAC TTCCACGATC CCTTCCTTGA CAAGGTCCGG AGAGAATCGA

       130         140         150         160         170         180
                Reverse primer-------->
TCTGGTGGCC TCAAGTGTTT CTTGGCTTAC AGATGTATCA TTCTAGTCAC ATTCCATCTT
AGACCACCGG AGTTCACAAA GAACCGAATG TCTACATAGT AAGATCAGTG TAAGGTAGAA

       190         200         210         220         230

CTCCCTGTGT ATCTTCATAT TGTCTTCCCC TCTGTGTGTG TGTGTGTGTG TGTG
GAGGGACACA TAGAAGTATA ACAGAAGGGG AGACACACAC ACACACACAC ACAC
```

Two PCR primers were designed (shown in bold in fig. 35a-b ; expected PCR product of 183 bp) and used to amplify (non-radioactively) genomic DNA from unrelated individuals and DNA from the cosmids of origin. PCR products of the predicted size were obtained (data not shown). Southern blotting and hybridisation of the PCR products with a $(GT)_n$ oligonucleotide confirmed the presence of the CA repeat (data not shown). The X-specificity of the CA repeat was verified by PCR analysis of a an X-chromosome somatic cell hybrid panel (fig. 36). PCR analysis of YACs 33CA11, 20CC8 and a YAC containing the OATL1 locus demonstrated the presence of the CA repeat in 33CA11 and OATL1 YACs confirming its origin (fig. 36).

Finally, the primers were used to amplify radioactively DNA from 10 unrelated caucasian females (20 X chromosomes) to detect polymorphism of the repeat. PCR products of identical size from all individuals (fig. 37) indicated that the CA repeat was monomorphic.

3.5 Isolation of a YAC overlapping with DXS426 YAC

PCR amplification of the library primary pools using primers for the right end STS of 33CA11 (fig. 30 ; Forward primer: residues 29-49, underlined ; Reverse primer shown in bold) resulted in PCR products of the predicted size for all pools (data not shown). Following exclusion of library contamination, the sequence was compared against the GenBank database using the FASTA programme (via HGMP). The search demonstrated that most of the sequence (including the primers) was homologous to Alu repeat elements thus explaining the above result (fig. 38). Therefore, a new forward primer (fig. 30, shown in bold) was designed for the unique sequence of this STS (residues 1-30 ; fig. 30). Screening of the library using the new primer resulted in the identification of one positive primary pool, 20 (data not shown). Secondary pool 20C was identified (fig. 39). Screening the tertiary pools for 20C gave negative results for all pools. PCR amplification of these pools

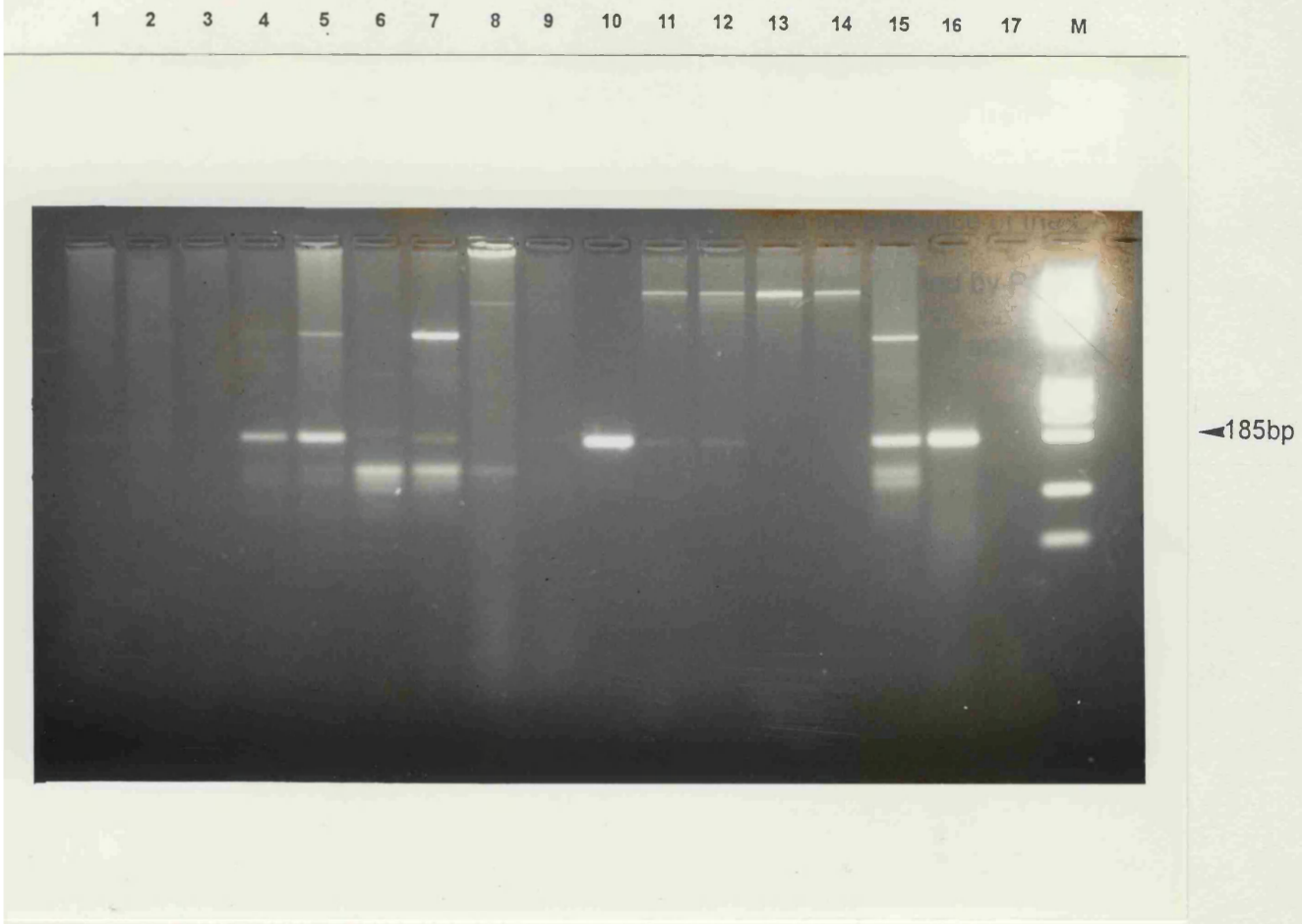


Figure 36: Ethidium bromide-stained agarose gel of PCR products from an X-chromosome somatic cell hybrid panel and YACs using primers flanking the CA repeat near the end of YAC 33CA11. Lane 1: The mouse-human X-only hybrid ThyB-X Lane2: Mouse DNA Lane 3: Hamster DNA Lanes 4 and 5: the mouse-human hybrid Kag 2.3 containing Xp21.1-Xqter as its only human component Lane 6: the hamster-human hybrid sin 176 containing an X chromosome deleted for the region Xp11.23-Xp22.11 Lane 7, 8 and 9: three hybrids containing different fragments of the long arm of the X chromosome: Xq23-Xqter, Xq25-Xqter, Xq27-Xqter, respectively Lane 10: A YAC containing the OATL1 locus Lanes 11 and 12: YAC 33CA11 Lanes 13 and 14: YAC 20CC8 Lane 15: Human genomic DNA Lane16: Cosmid of polyCA origin (D01210) . Lane 17: A "no DNA" control. M denotes a ϕ174 DNA digested with HaeIII size marker.

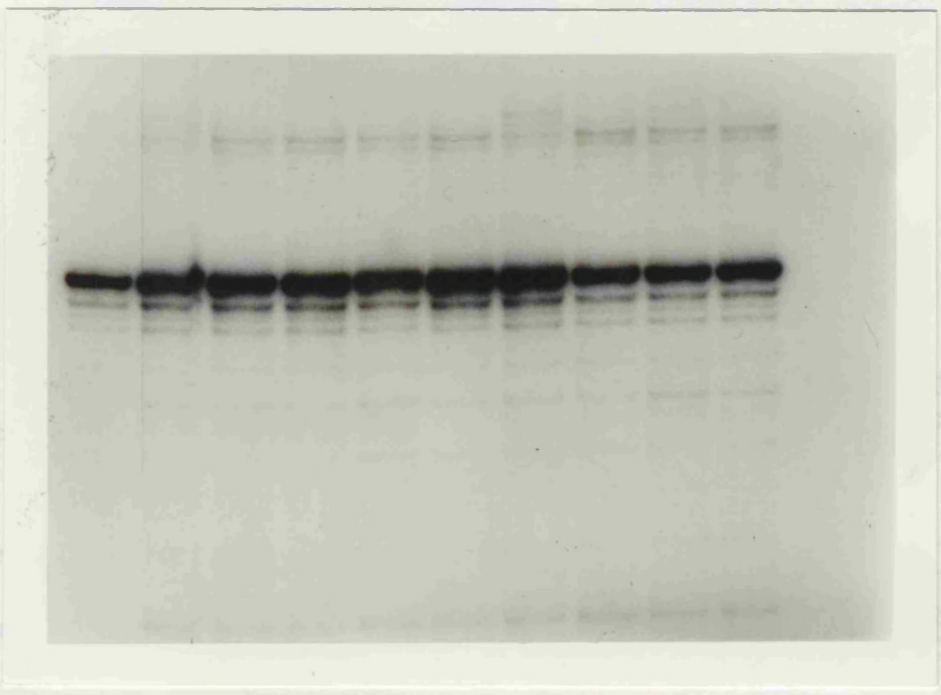


Figure 37: Autoradiograph showing PCR products of identical size for 10 unrelated individuals.

```
HUMBCRE L02935 (repeat region: alu sx) Human major bre  228   228   263
  83.0% identity in 100 nt overlap


                    10         20         30         40         50         60
Right end of 33CA11 CCAGCTGAGAAATGCTAATAAAAGCAATGATCACGAGGTCCAGGTGCAGTGACTCATGCC
                                                :: ::::: :: :: X::::: ::::::::
HUMBCR                                    AAGTTTCTCGAGG-CCGGGCGCAGTGGCTCATGCC
                                                    10         20         30

                    70         80         90        100        110        120
Right end of 33CA11 TGTAATTCCAGCACTTTGGGAGGCCAAGGTGGGTAGATCGCTTCAGCCCAGGCATTTGAG
                    :::::: :::::::::::::::::  :::  ::: :::::::: ::: :::::  :: :::
HUMBCR              TGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCGCTTGAGCTCAGGAGTTGGAG
                        40         50         60         70         80         90

                   130
Right end of 33CA11 ACCGAGCCTGG
                    ::X ::::::
HUMBCR              ACC-AGCCTGACCAACATGGTGAAACCCTGTGTCTACTAAAAATACAAAGATTAGCCAGG
                        100        110        120        130        140        150
```

Figure 38: Sequence comparison of right end of 33CA11 YAC versus GenBank showing 83% homology with an Alu repeat.

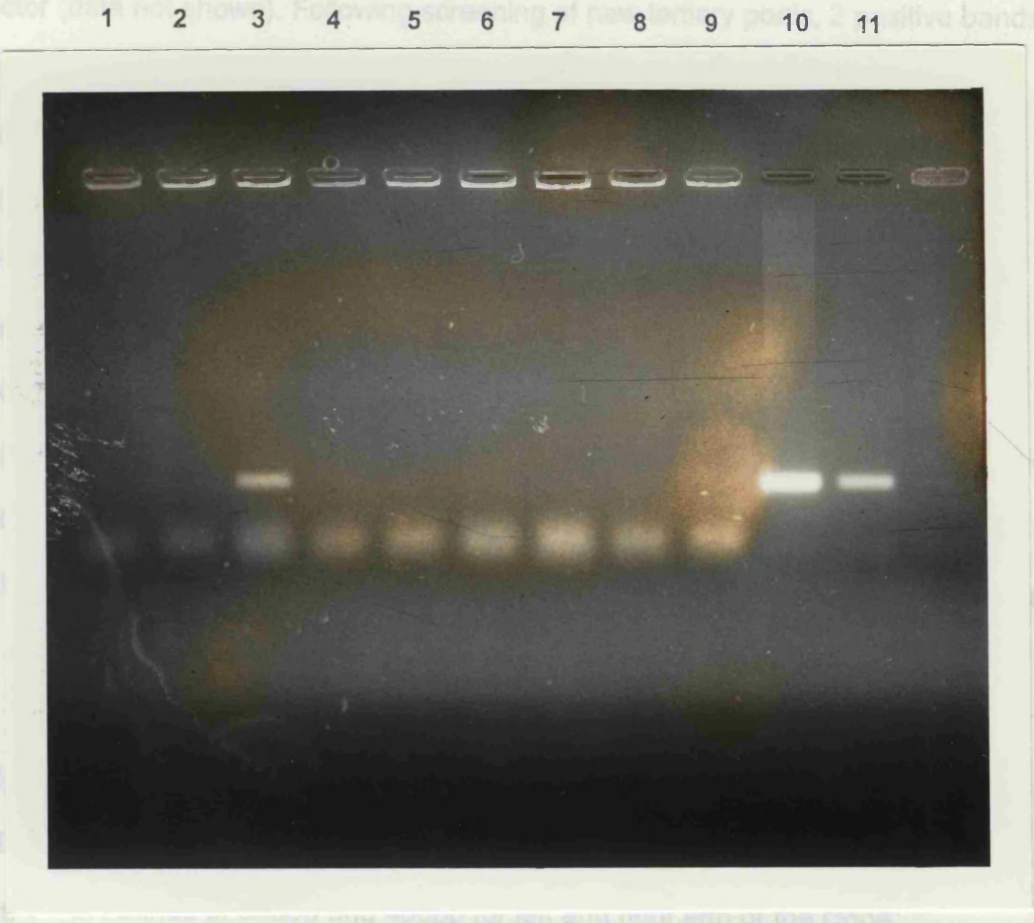


Figure 39: An example of PCR screening of secondary ICI YAC pools using specific primers for the right end of YAC 33CA11 during which pool 20C is identified. Lanes 1-9: secondary pools 20A-20I. Lane 10: YAC 33CA11. Lane 11: Primary pool 20.

using vector pYAC4 primers indicated that most of the pools did not contain the vector (data not shown). Following screening of new tertiary pools, 2 positive bands for pools 3 and 16 corresponding to clone C8 were identified (data not shown). Weaker positive bands for 3 pools made the interpretation of the results uncertain, therefore, two clones, 20CC8 and 20CC5 were selected for analysis. Colony PCR for the two clones demonstrated that 20CC8 was the true positive clone (data not shown) and was analysed further.

DNA from YAC 20CC8 was prepared in plugs and examined by PCR for the presence of the 33CA11/ right end STS. Appropriate controls were used. Clone 20CC8 contained the above STS (fig. 45b) thus confirming that it was overlapping with 33CA11. It was then sized at 580kb by pulsed field gel electrophoresis followed by alkali blotting and hybridisation with total human DNA (fig. 40).

3.6 Isolation of end clones from YAC 20CC8

Following the construction and PCR amplification of the RsaI Vectorette library, single PCR bands of 500bp and 400bp for left and right arm of the clone, respectively were obtained (fig. 41a-b). These PCR products were sequenced and a total of 122 bases for the left arm and 114 bases for the right arm product were obtained (fig. 42). Primers were designed and used for PCR analysis of 20CC8. The presence of both ends in 20CC8 confirmed their origin (fig. 46).

3.7 Isolation of YACs containing the DXS1003 locus

To resolve the orientation of the two DXS426 (33CA11 and 4HG2) YACs with respect to the centromere, YACs for the microsatellite locus DXS1003, mapping distal to the DXS426 locus (Report of the fourth international workshop on human X chromosome mapping, 1993) were isolated. The ICI YAC genomic library was, therefore, screened by PCR using primers for this locus. Human genomic DNA, pYAC4 DNA and a no DNA control were included in the screening experiments. Two primary positive pools, 19 and 36, and secondary pools 19D and 36B were

Figure 40: a. PFGE of DNA plug samples from YACs 20CC8, 36BG12 and 19DH7. Lane 1: Concatemerized λ bacteriophage DNA size standard. Lane 2: S. cerevisiae chromosomal DNA size marker. Lanes 3 and 4: YAC 20CC8. Lane 5: YAC 36BG12. Lane 6 and 7: YAC 19DH7 Lane 8: YAC 4HG2 (245kb). b. Autoradiograph showing the size of the YACs after Southern blotting gel (a) and hybridisation with total human DNA. Signal was not obtained from YAC 19DH7 (lanes 6 and 7).

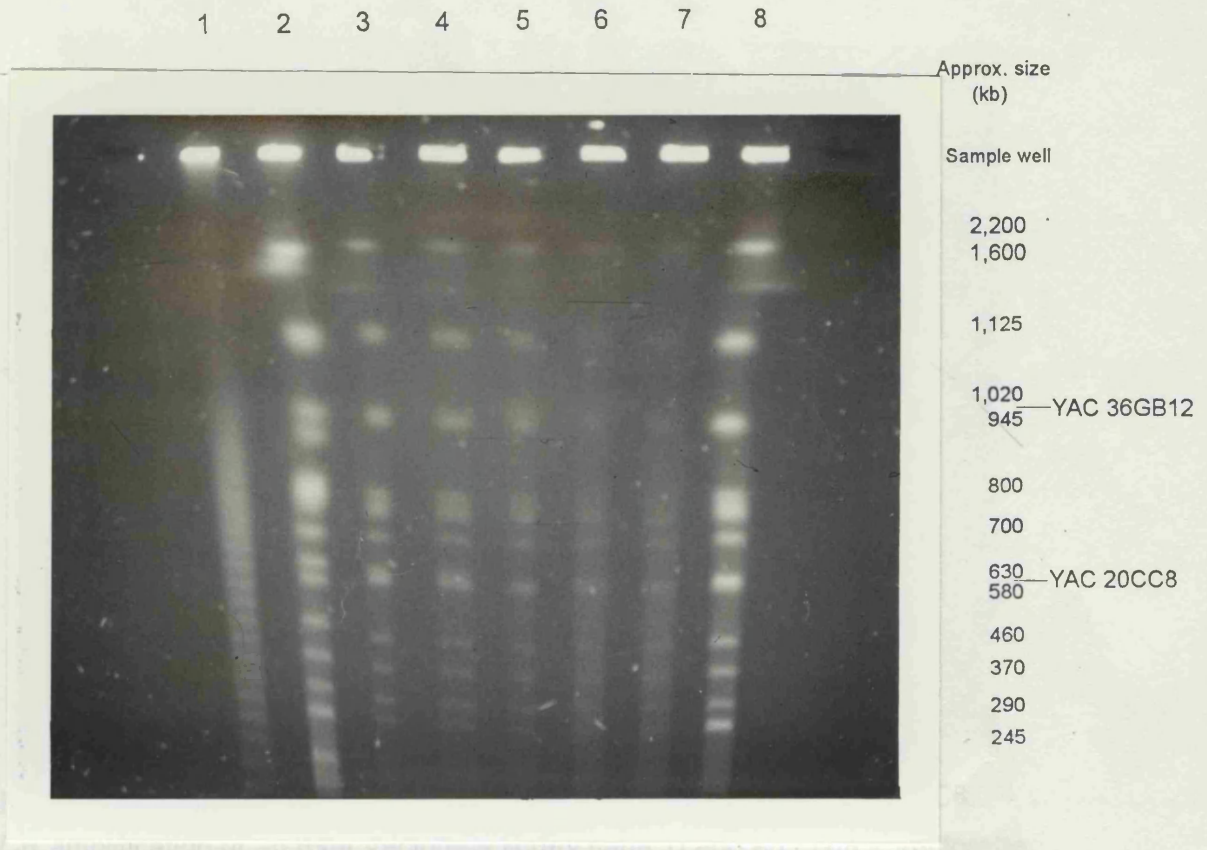


Fig. 40a

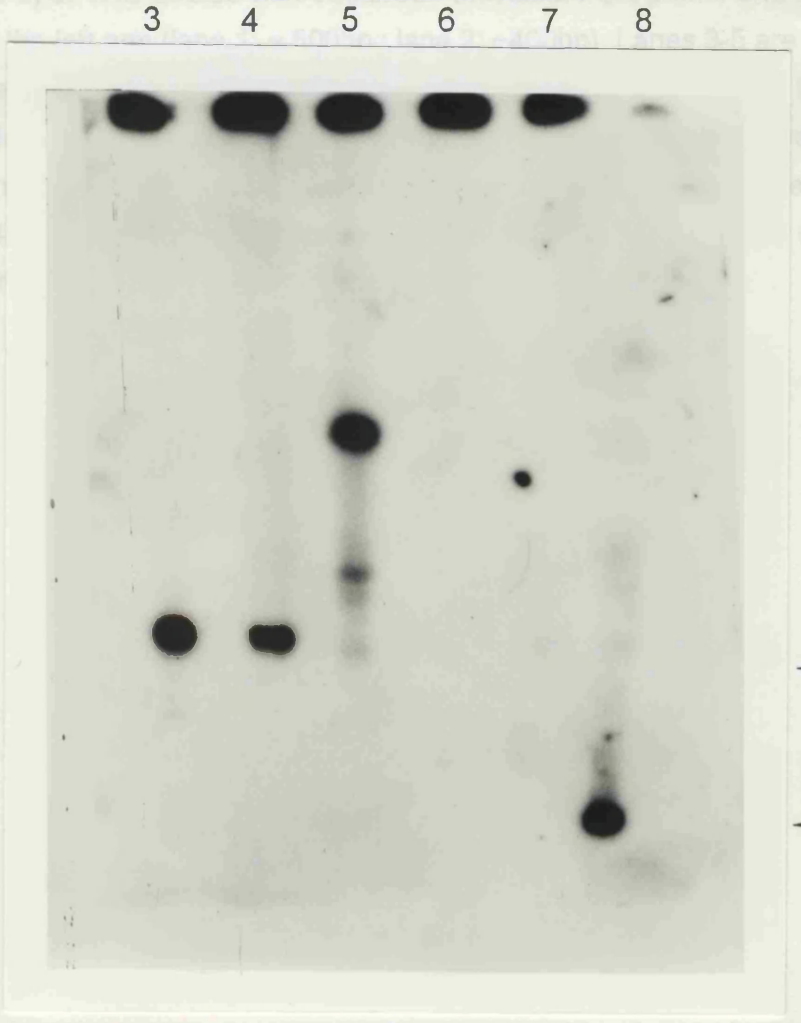


Fig. 40b

Figure 41: Vectorette system for the isolation of YAC termini from YAC 20CC8. a. PCR amplification of an RsaI Vectorette library (lane 1) and an EcoRV Vectorette library (lane 2) of YAC 20CC8 with Vectorette Universal PCR primer and a primer specific for the left arm (lane 1: ~ 500bp ; lane 2: ~400bp). Lanes 3-5 are negative controls. Lane 7: A pBR322 DNA digested with EcoRI/AvaII size marker. Lane 8: A 100bp ladder size marker. b. PCR amplification of an RsaI Vectorette library of YAC 20CC8 with Vectorette Universal PCR primer and a primer specific for the right arm (lane 1: ~ 400bp). Lanes 2 and 3 are negative controls. Lane 5: A 100bp ladder size marker. Lane 6:A pBR322 DNA digested with EcoR1/AvaII size marker.

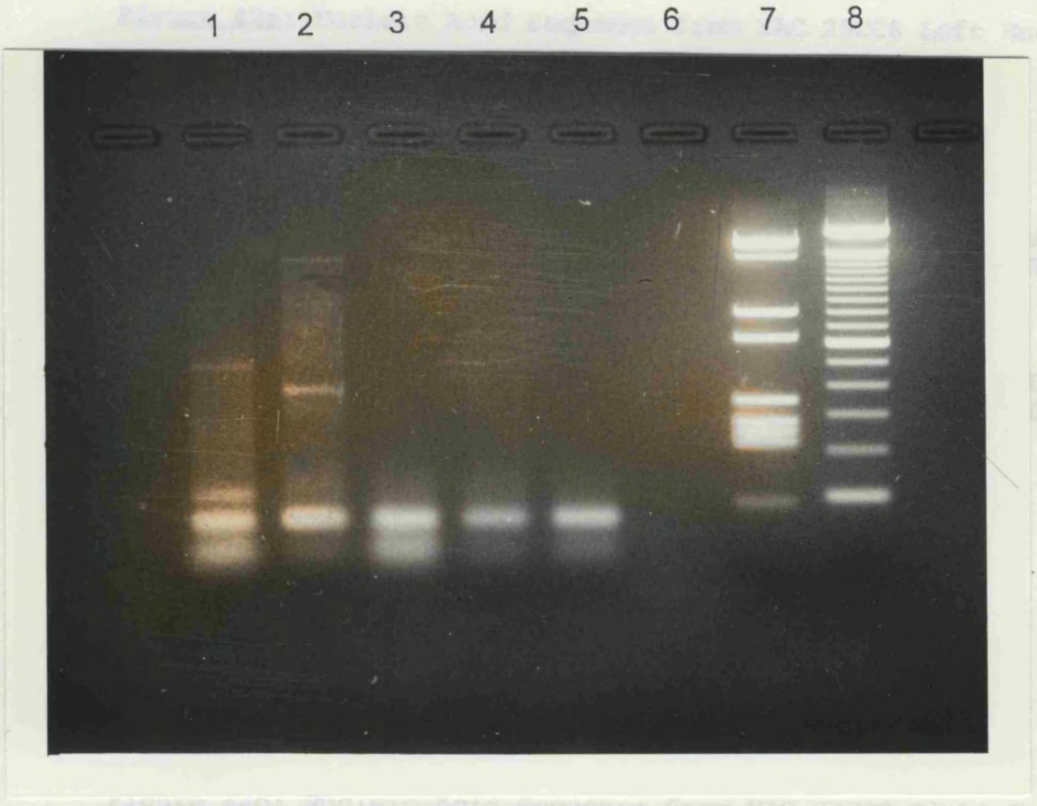


Fig. 41 a

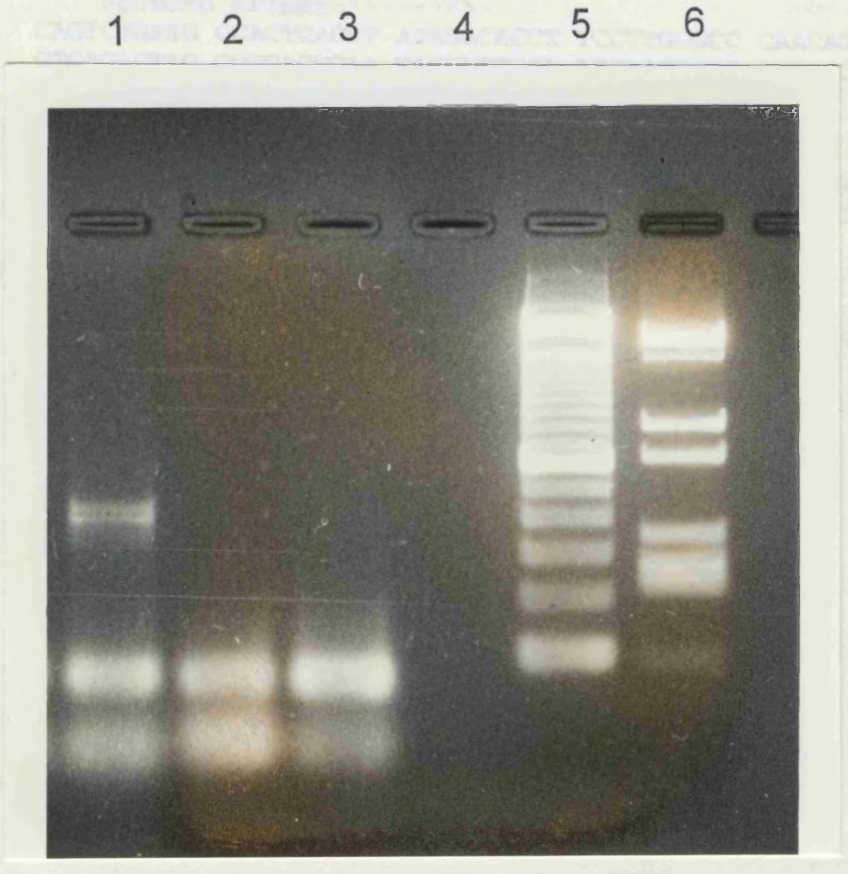


Fig. 41 b

Figure 42a: Nucleic Acid Sequence from YAC 20CC8 Left End Clone

```
       10         20         30         40         50         60
   Forward primer--------->
ACGAGGCCAG GTGCAGTGAC TCATGCCTGT AATTCCAGCA CTTTGGGAGG CCAAGGTGGG
TGCTCCGGTC CACGTCACTG AGTACGGACA TTAAGGTCGT GAAACCCTCC GGTTCCACCC

       70         80         90         100        110        120

TAGATCGCTT CAGCCCAGGC ATTTGAGACC AGCCTGGGCA ACATGGCAAA AACCCATCTG
ATCTAGCGAA GTCGGGTCCG TAAACTCTGG TCGGACCCGT TGTACCGTTT TTGGGTAGAC
                              <---------Reverse primer

TTAAGGAG AGGACAGCGA TTCTCGTACG AACGGTTACG A
AATTCCTC TCCTGTCGCT AAGAGCATGC TTGCCAATGC T
  <------------Vectorette unit------------>
```

Figure 42b: Nucleic Acid Sequence from YAC 20CC8 Right End Clone

```
       10         20         30         40         50         60
    Forward primer-------->
CAGTCTGATG CCACTGAGTT ATAGGCACCT TCCTTGGGCC CAACAGTTGG GTTTATGAAG
GTCAGACTAC GGTGACTCAA TATCCGTGGA AGGAACCCGG GTTGTCAACC CAAATACTTC

       70         80         90         100        110

ACAGGATCTG GATGTCCATT CCAGGCCCCA GAATATTTAC CTAAGATATA TGTTAAG
TGTCCTAGAC CTACAGGTAA GGTCCGGGGT CTTATAAATG GATTCTATAT ACAATTC
                          <--------Reverse primer          <--
<--

GAGAGGACAG CGATTCTCGT ACGAACGGTT ACGA
CTCTCCTGTC GCTAAGAGCA TGCTTGCCAA TGCT
----------Vectorette unit---------->
```

identified. The tertiary pools 8 and 15 for secondary pool 19D, corresponding to clone H7 (see table in section 4.1.3, chapter 2), and pools 7 and 20 for secondary pool 36B, corresponding to clone G12, were positive (data not shown). The combined results from the three step screening of the library led to the identification of two positive for the DXS1003 YAC clones, 19DH7 and 36BG12.

DNA from these clones was prepared in plugs and analysed by PCR for the presence of DXS1003. Appropriate controls were used. Both clones contained the above locus (fig. 43) thus confirming their identity. YAC 36G12 was then sized at approximately 1Mb by pulsed field gel electrophoresis followed by alkali blotting and hybridisation with total human DNA (fig. 40). YAC 19DH7 was not sized due to a lack of signal.

3.8 X-specificity of end clones

X-specificity of the end clones from YACs 33CA11 and 20CC8 was confirmed by PCR analysis of DNAs from a panel of X-specific somatic cell hybrids DNAs. The following hybrids were used: 1) the mouse-human X-only hybrid ThyB-X (Lund et al., 1983) 2) the mouse-human hybrid Kag 2.3 containing Xp21.1-Xqter as its only human component 3) the hamster-human hybrid sin 176 (Ingle et al., 1985) containing an X chromosome deleted for the region Xp11.23-Xp22.11 4) three hybrids containing different fragments of the long arm of the X chromosome: Xq23-Xqter, Xq25-Xqter, Xq27-Xqter. Mouse and hamster DNA were also included in the analysis. An example of analysis for X-specificity is shown for the right end clone of YAC 33CA11 (fig. 44).

3.9 Orientation of YACs and loci with respect to the centromere

The newly isolated YAC clones, 4HG2, 33CA11, 20CC8, 36BG12 and 19DH7, were analysed by PCR for the end-specific STSs of 33CA11 and 20CC8. A YAC containing the OATL1 locus (a gift from Dr A. Monaco) known to map proximal to DXS426 locus was included in the analysis. YACs 36BG12 and 19DH7 did not

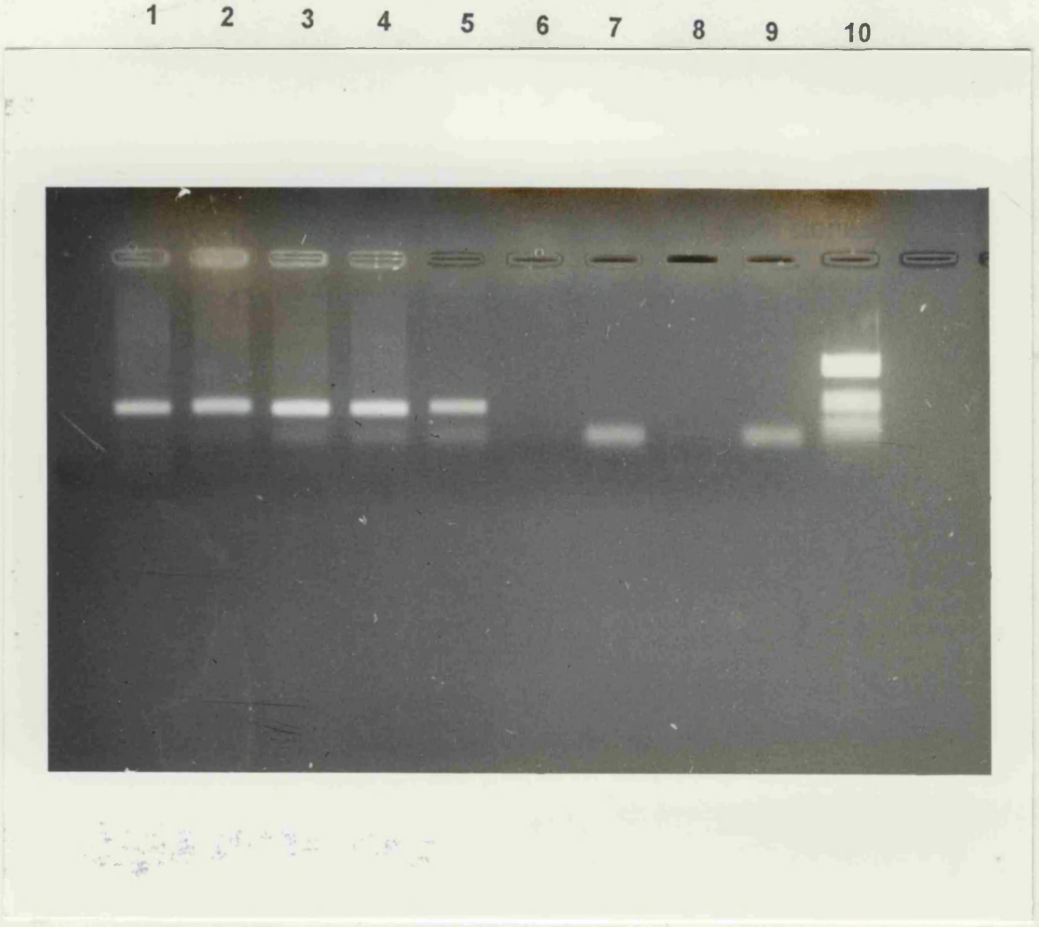


Figure 43: Ethidium-bromide stained agarose gel of PCR products from YACs36BG12 (lanes 1 and 2) and 19DH7 (lanes 3 and 4) using primers for the DXS1003 (276) locus. Lane 5: A YAC containing the DXS1003 locus (a gift from Dr M. Coleman). Lane 7: pYAC4 vector DNA Lane 8: Yeast DNA Lane 9: A "no DNA" control. Lane 10: A pBR322 DNA digested with HaeIII size marker.

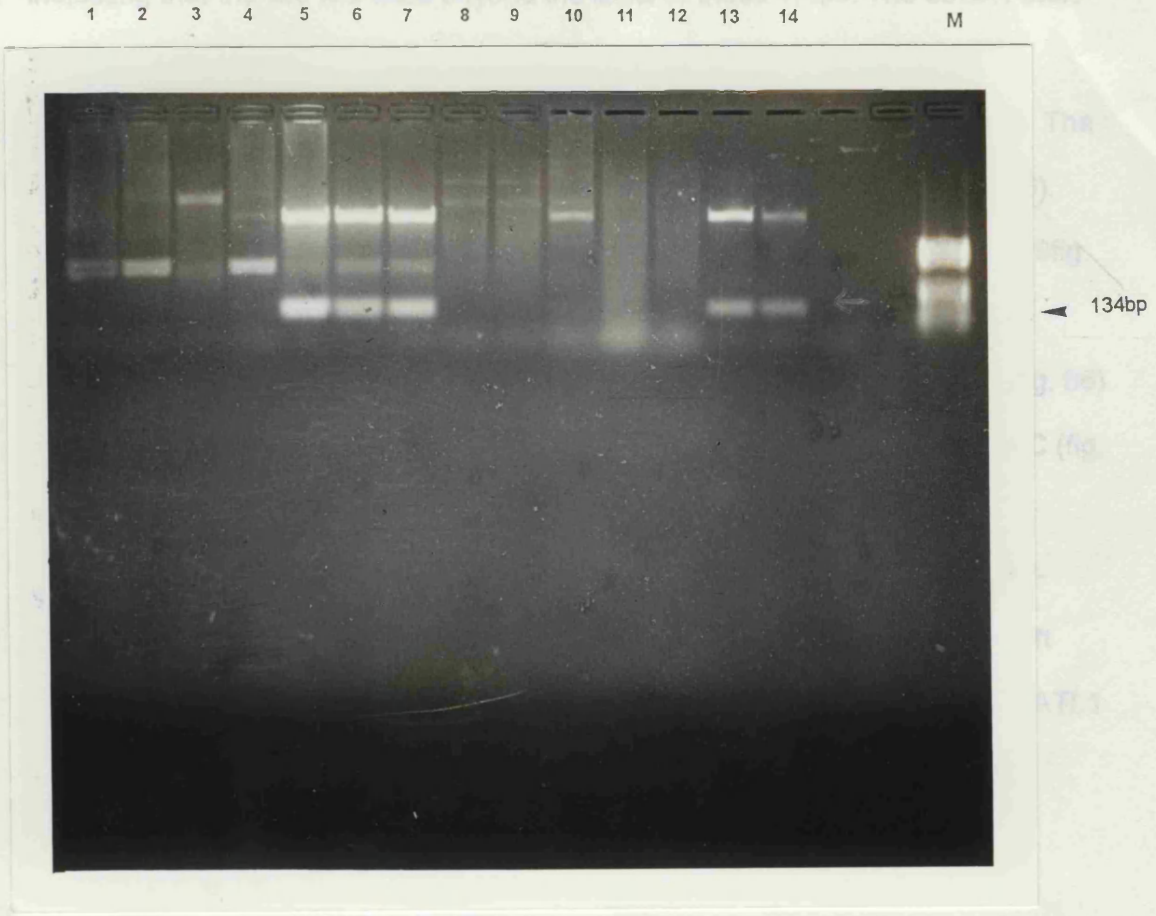


Figure 44: Ethidium bromide-stained agarose gel of PCR products from an X-chromosome somatic hybrid panel using specific primers for the right end clone of YAC 33CA11. Lane 1: The mouse-human X-only hybrid ThyB-X. Lane2: Mouse DNA. Lane 3: Hamster DNA. Lane 4: The mouse-human hybrid 12/3 containing Xpter-Xp21 as its only human component. Lanes 5, 6 and 7: The mouse-human hybrid Kag 2.3 containing Xp21.1-Xqter as its only human component. Lanes 8 and 9: The hamster-human hybrid sin 176 containing an X chromosome deleted for the region Xp11.23-Xp22.11. Lanes 10, 11 and 12: Three hybrids containing different fragments of the long arm of the X chromosome: Xq23-Xqter, Xq25-Xqter, Xq27-Xqter, respectively. Lane 13 and 14: Human genomic DNA. M denotes a pBR322 DNA digested with HaeIII size marker.

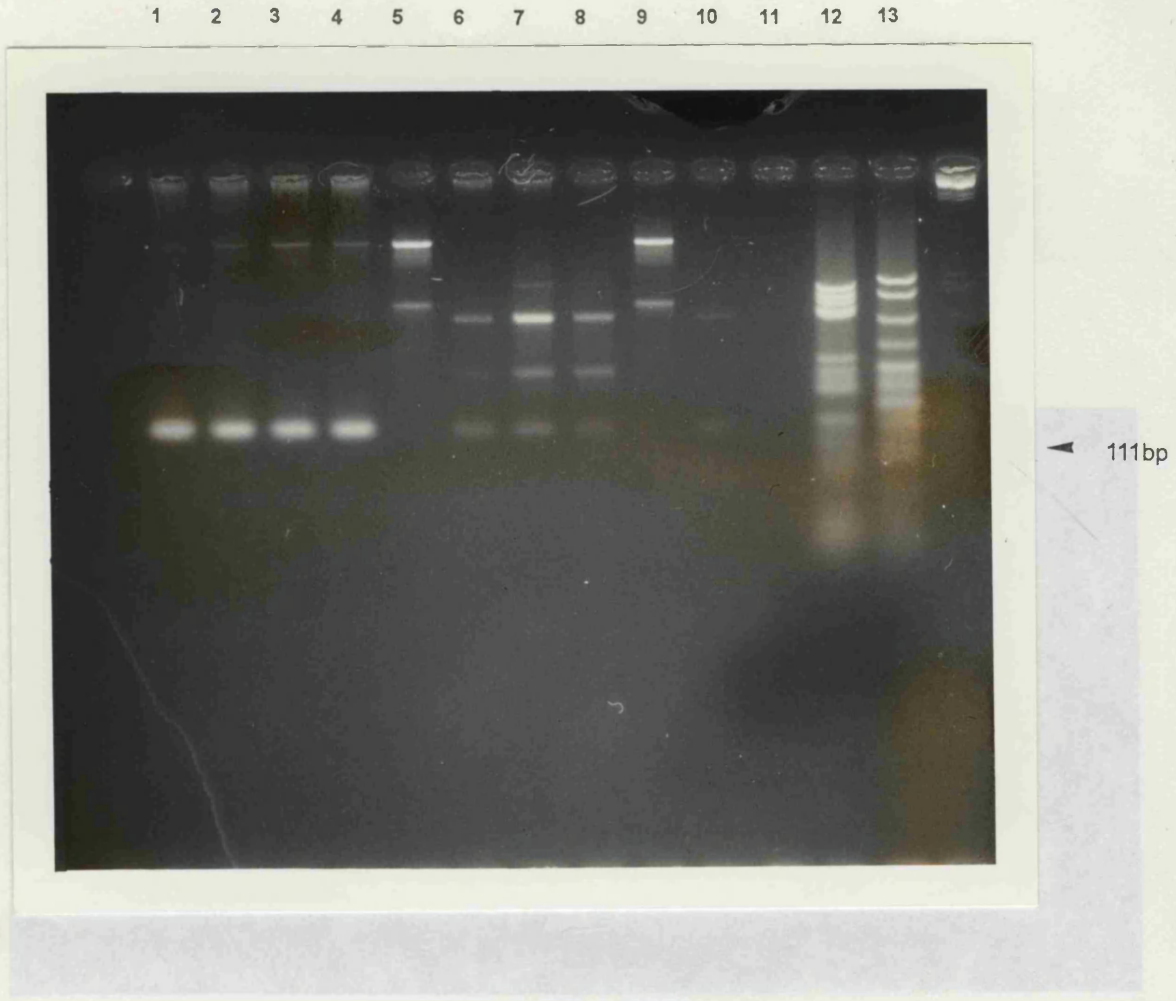
contain any of these STSs (fig. 45). The microsatellite for synapsin1 gene and the UBE1 gene, both known to map in the proximity of DXS426 and DXS1003 loci, were absent from YACs 36BG12, 19DH7, 33CA11 and 4HG2 (data not shown), indicating that the two loci were beyond the limits of these YACs. The 33CA11/left end-specific STS was present in YAC 4HG2 (fig. 45a). The 33CA11/right end-specific STS was absent from 4HG2 and present in the OATL1 YAC (fig. 45b). The 20CC8/left end-specific STS was present in 33CA11 and OATL1 YACs (fig. 46). The 20CC8/right end-specific STS was absent from both 33CA11 and OATL1 9fig. 46). Finally, as previously mentioned: 1) the CA repeat near the right end of 33CA11 was present in 33CA11 and OATL1 YACs and absent from 20CC8 (fig. 36) and 2) the PFC locus was absent from 33CA11 YAC and present in 4HG2 YAC (fig. 28).

The overall data from the STSs analysis supports the order of the YACs: Xpter-4HG2-33CA11-20CC8-Xpcen and the order of the loci: Xpter-PFC-33CA11/Left end-DXS426- 33CA11/Right end $(CA)_n$-20CC8/Left end-33CA11/Right end-(OATL1 locus-20CC8/Right end)-Xpcen (fig. 47).

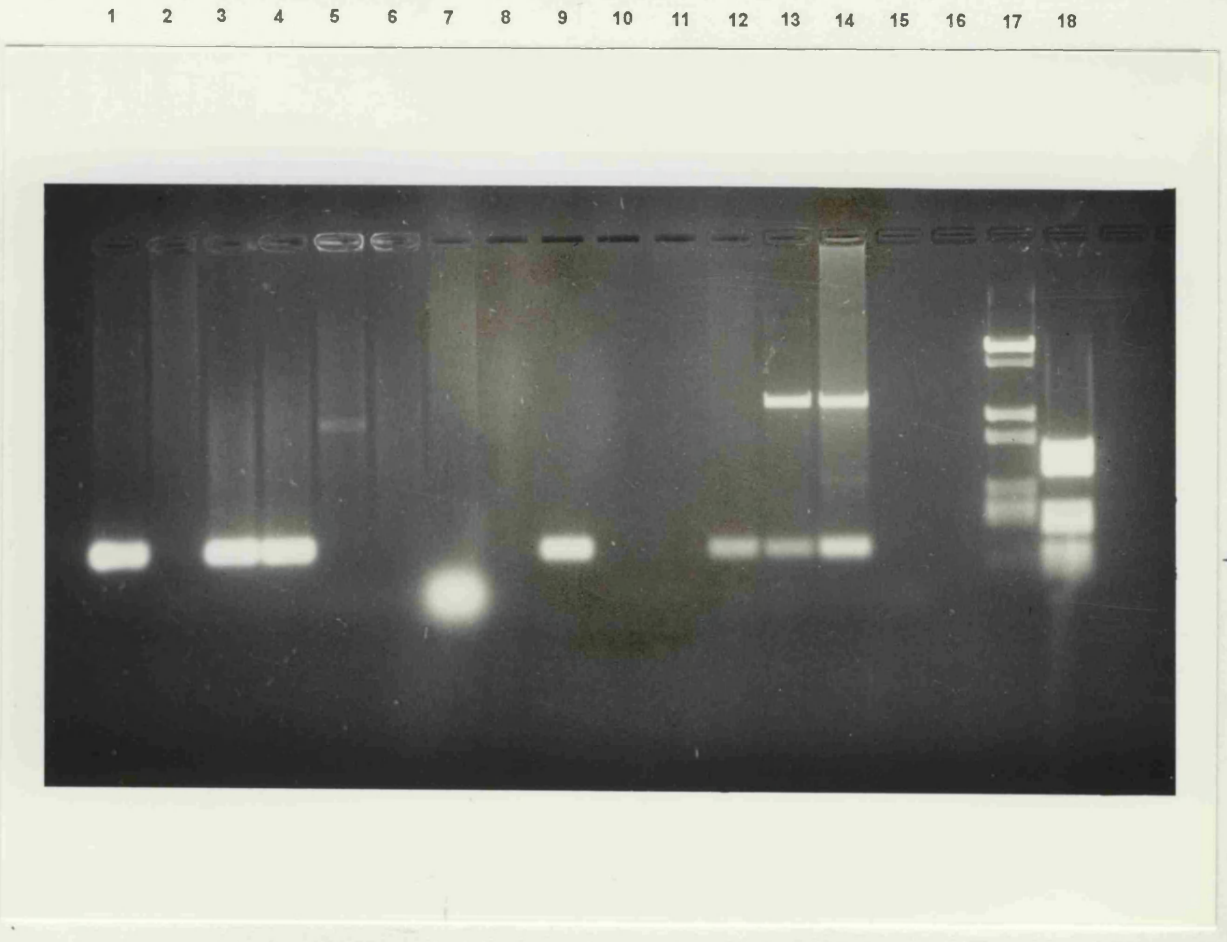
4. Discussion

A genomic YAC library was screened by PCR using the DXS426 locus primers. Two positive YACs were identified and sized by pulsed field gel electrophoresis followed by Southern blotting and hybridisation with total human DNA. The ends of one of the YACs were isolated using the Vectorette method (Riley et al., 1990) and sequenced. Simultaneous ordering of the YACs and end-specific STSs was achieved by PCR of these and other YACs in the region. The proximal end was used to isolate the overlapping YAC, the ends of which were rescued, sequenced and orientated. During this walk a microsatellite was isolated by a rapid, PCR based technique. Analysis of the DXS426 YACs resolved the orientation of properdin P factor and DXS426 loci, known to map in close proximity, and the distance between them was refined.

Figure 45: Ethidium bromide-stained agarose gels of PCR products from YACs using primers from the left (a) and right (b) ends of 33CA11 YAC.
a. Lanes 1 and 2: YAC 33CA11 Lanes 3 and 4: YAC 4HG2 Lanes 5: A YAC containing the DXS1003 locus. Lanes 6, 7 and 8: Human genomic DNA. Lane 9: A YAC containing the ARAF1 locus. Lane 10: A "no DNA" control. Lane 12: A pBR322 DNA digested with HaeIII size marker. Lane 13: A pBR322 DNA digested with MspI size marker.
b. Lane 1: YAC 33CA11 Lane 2: YAC 4HG2 Lanes 3 and 4: YAC 20CC8. Lane 5 and 6: YACs 36BG12 and 19DH7 (containing the DXS1003) locus. Lanes 7, 8 and 9: YACs OATL1-1, OATL1-2 and OATL1-6, respectively (gifts from Dr A. Monaco). Lane 10: YAC containing the locus ARAF1 (a gift from Dr M. Coleman). Lane 11: Cosmid HX97 containing the DXS722 locus (see section 3.2 of chapter 3) Lanes 12 and 13: Human genomic DNA. Lane 14: Kag 2.3, a mouse-human hybrid containing Xp21.1-Xqter as its only human component. Lane 15: A "no DNA" control. Lanes 17 and 18: Size markers: pB322 DNA digested with AvaII and EcoRI and pB322 DNA digested with HaeIII, respectively.


1 2 3 4 5 6 7 8 9 10 11 12 13
111bp

a


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
134bp

b

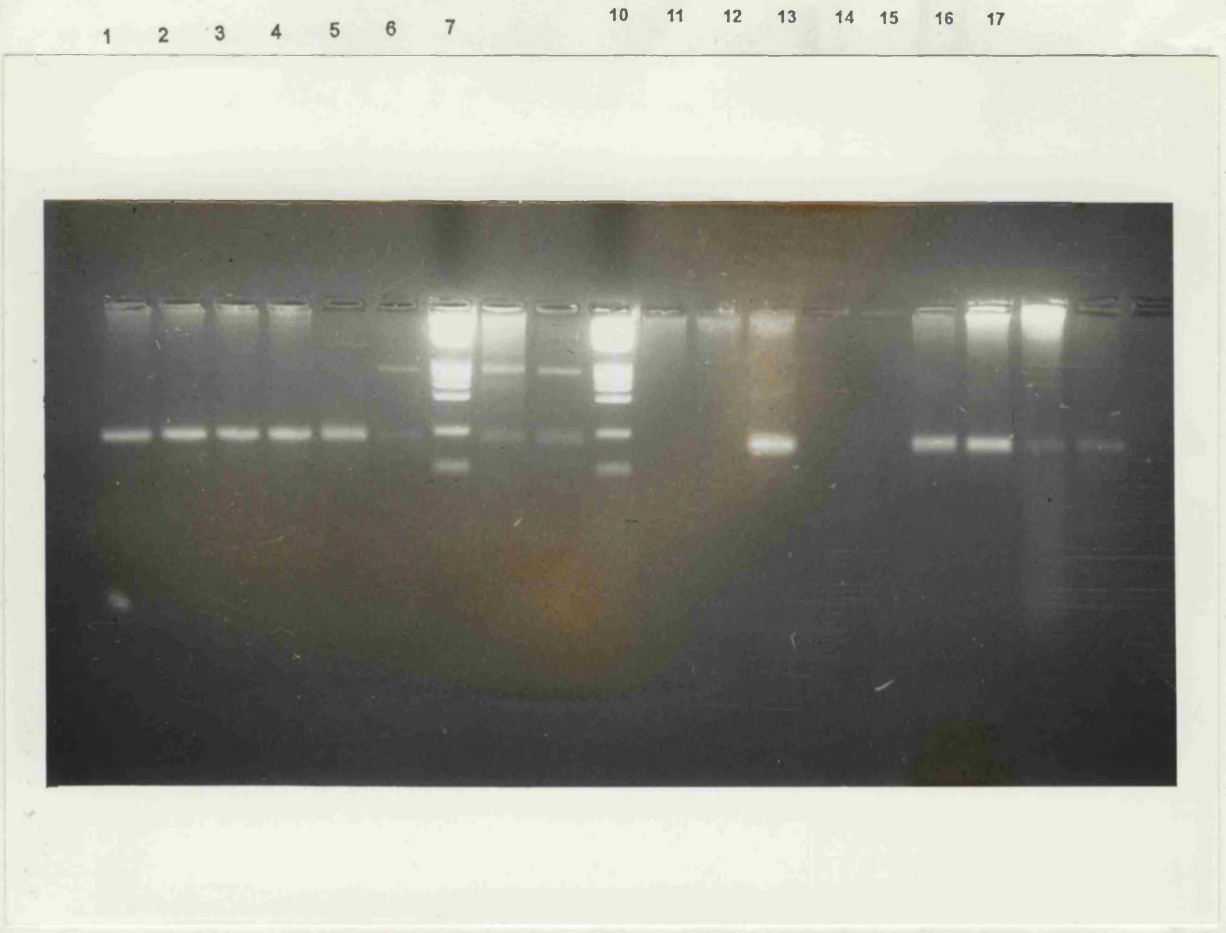


Figure 46: Ethidium bromide-stained agarose gels of PCR products from YACs using primers from the left (lanes 1-7) and right (lanes 10 -17) ends of 20CC8 YAC. Lane 1 and 2: YAC 33CA11. Lane 3 and 4: YAC 20CC8. Lane 5: A YAC containing the OATL1 locus. Lane 6 is a "no DNA" control. Lanes 7 and 10 are ϕ174 DNA digested with HaeIII size marker.

Lane 11 and 12: YAC33CA11. Lane13: YAC 20CC8. Lane 14: A YAC containing the OATL1 locus. Lane 15: A "no DNA" control. Lanes 16 and 17: Kag 2.3, a mouse-human hybrid containing Xp21.1-Xqter as its only human component.

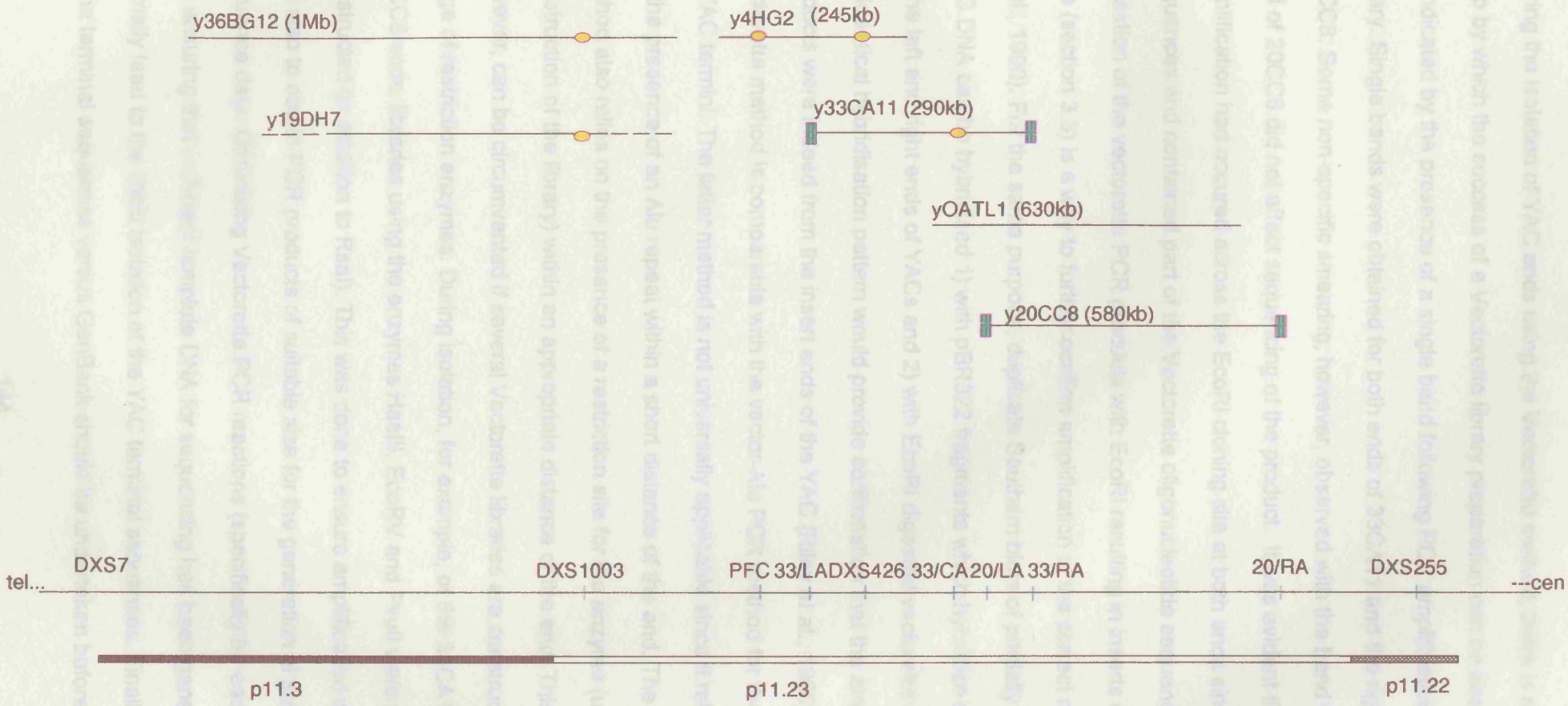


Figure 47: A YAC contig around the DXS426 locus at Xp11.23.
Green box: end clone ; Yellow dot: presence of STS.
DXS7, DXS1003, DXS426 and DXS255: microsatellite polymorphic loci. PFC: properdin P factor.
33/LA and 33/RA: left and right arm of 33CA11 YAC. 33/CA: a (CA)n repeat isolated from the proximal end of 33CA11 YAC. 20/LA and 20/RA: left and right arm of 20CC8 YAC.
yOATL1: a YAC containing part of the ornithine aminotransferase-like cluster.

During the isolation of YAC ends using the Vectorette method, there is only one step by which the success of a Vectorette library preparation can be assessed. This is indicated by the presence of a single band following PCR amplification of the library. Single bands were obtained for both ends of 33CA11 and the right end of 20CC8. Some non-specific smearing, however, observed with the band for the left end of 20CC8 did not affect sequencing of the product. It was evident that amplification had occured across the EcoRI cloning site at both ends since the sequences end contained part of the Vectorette oligonucleotide sequence. Digestion of the vectorette PCR products with EcoRI resulting in inserts of expected size (section 3.3) is a way to further confirm amplification of the correct region (Riley et al., 1990). For the same purpose, duplicate Southern blots of partially digested YAC DNA can be hybridised 1) with pBR322 fragments which hybridise specifically to the left and right ends of YACs and 2) with EcoRI digested vectorette products. An identical hybridisation pattern would provide confirmation that the amplification products were indeed from the insert ends of the YAC (Riley et al., 1990). The Vectorette method is comparable with the vector-Alu PCR method for the isolation of YAC termini. The latter method is not universally applicable since it relies entirely on the presence of an Alu repeat within a short distance of the end. The Vectorette method also relies on the presence of a restriction site for the enzyme (used during construction of the library) within an appropriate distance of the end. This problem, however, can be circumvented if several Vectorette libraries are constructed using a range of restriction enzymes. During isolation, for example, of the 33CA11 and 20CC8 ends, libraries using the enzymes HaeIII, EcoRV and PvuII were also constructed (in addition to RsaI). This was done to ensure amplification of the library and also to obtain PCR products of suitable size for the generation of sufficient sequence data. Optimising Vectorette PCR reactions (specifically for each YAC) and ensuring that sufficient template DNA for sequencing has been generated will normally lead to the rapid isolation of the YAC terminal sequences. Finally, a search of the terminal sequences versus GenBank should be undertaken before screening

the YAC library in order to prevent problems arising due to repeat sequences potentially present in the end sequence (e.g. proximal end of YAC 33CA11). Analysis of the DXS426 YACs resolved the orientation of properdin P factor and DXS426 loci, known to map in proximity, and refined the physical distance between them. Properdin P factor (PFC) is a component of the alternative pathway of the serum complement system (Pangburn et al., 1984), previously co-localised with DXS1003, DXS426 and OATL1 loci within a 1.7Mb YAC contig (order: Xpter-DXS1003-DXS426-OATL1-Xpcen ; report of the fourth international workshop on human X chromosome mapping, 1993: p. 160). It has also been shown that PFC and DXS426 lie within 390kb of each other (Coleman, 1991). A proximal localisation of PFC relative to DXS426 has been proposed (report of the fourth international workshop on human X chromosome mapping, 1993: p. 156) although the order of PFC has not been unequivocally resolved (report of the fourth international workshop on human X chromosome mapping, 1993: Abstract 42). Analysis of the two DXS426 YACs (33CA11 and 4HG2) using primers for the properdin P factor locus demonstrated the presence of properdin in the distal 4HG2 and its absence from the proximal 33CA11. This data indicates that PFC lies distal to DXS426 and that both loci lie within 245 kb (distance spanned by the 4HG2 YAC). The order of the PFC and DXS426 loci is thus resolved as: Xpter-PFC-DXS426-Xcen.

A genetic map of the region under study based on highly informative markers is now available. The local density of these markers, however, is not sufficient for the positional cloning of the RP2 and/or other genes in this region. Therefore there is a need to develop further new markers for the target region. A CA repeat present in a cosmid anchored to the proximal end of YAC 33CA11 was, therefore, developed and characterized. Methods for the isolation of microsatellites from a YAC clone have been reported. In one method, the YAC was gel-purified, digested with a frequent-cutting enzyme and ligated to synthetic amplimers. This material was amplified, denatured and annealed to a biotinylated $(CA)_{11}$ oligonucleotide. The YAC fragments containing (CA) repeats were then recovered by binding to

streptavidin coated magnetic beads and the resultant material cloned and sequenced (Hardwick et al., 1993). Another method involves the selective cloning and sequencing of $(CA)_n$-positive inter-Alu sequences (De Souza et al., 1994). In this method advantage can be taken of the inter-Alu amplification to bypass the need for purification of the human sequences. PCR-based methods for the isolation of microsatellites from DNA cloned in plasmid, cosmid, phage or YAC clones have been described (Edwards et al., 1991 ; Pandolfo, 1992 ; Koref et al., 1993). The method described here was also PCR-based according to the principles of the Vectorette technique, primarily used for the isolation of YAC ends and previously described for the isolation of $(GT)_n$ repeats from YAC clones (Pandolfo, 1992). However, the method was modified since two primers [a $C_4(CA)_{11}$ and a $G_4(GT)_{11}$ primer] instead of one were used. This was advantageous because a single amplification of the Vectorette library was required after which CA and GT strands from opposite direction were sequenced and unique sequences flanking the repeat determined. Therefore preventing the need for further steps. Additionally, the use of the phosphorylated Vectorette primer in combination with the lambda exonuclease method of single stranded DNA preparation allowed direct sequencing of the PCR products without the need for subcloning. This method was rapid compared to the subcloning of short inserts previously described in this study for the isolation of a microsatellite from cosmid HX43 (chapter 4).

Following genotyping on unrelated individuals this CA repeat was found to be monomorphic. On confirming the sequence, it was found that the repeat contained only 11 CA doublets. During construction of the linkage map of the human genome by Weissenbach et al. (1992), all clones containing 12 or less CA doublets were rejected. The investigators suggest that such markers are unlikely to prove highly polymorphic in the population. Therefore, it was not surprising that the CA repeat at the proximal end of YAC 33CA11 turned out to be monomorphic. Although this STS can not be used as a genetic marker, however, it can be used as an additional anchored physical marker in this region. Furthermore, the method of isolation can

be used to generate poly-CAs rapidly in a targeted fashion, from cosmids containing the end-specific STSs or from a cosmid library representing the YAC insert. The principle of targeted development of microsatellite markers and their use in linkage analysis will allow positioning on the genetic map thus facilitating the inte gration of the developing physical and genetic map of the region.

Orientation of YAC 20CC8 end-specific STSs was determined by their presence or absence in the distal 33CA11 YAC. The distal end of YAC 20CC8 was also present in the OATL1 YAC with the the proximal end absent thus suggesting that 20CC8 YAC extends beyond this YAC, towards the centromere. This does not imply that 20CC8 also extends beyond the OATL1 cluster since the OATL1 YAC used in this study does not contain the complete OATL1 region. This data, however, suggests that 20CC8 contains part of the OATL1 cluster and further investigation of this YAC using OAT cDNA (Mitchell et al., 1988) probes or primers is required to confirm this and assess its OAT like sequences.

The YAC clones described here physically linked the PFC locus with the distal ornithine aminotransferase-like sequence cluster (OATL1) locus in the proximal short arm of the X chromosome. PFC maps within 70kb along with the ARAF1/SYN1/TIMP gene cluster (Derry and Barnard, 1992). This suggests that this region is adjacent to 4HG2 YAC which contains PFC but not the ARAF1/SYN1 CA repeat. The CA repeat which was identified in a cosmid containing the proximal end STS of 33CA11 YAC (290kb) was not present in the 20CC8 YAC (580kb). This suggests that the two YACs overlap in a region smaller than 40kb (the size of the cosmid), therefore, the minimum region covered by the two YACs must be ~830kb. The size of the overlapping region between 33CA11 and 4HG2 YACs is not known. To determine this, further characterisation by restriction mapping or subclonig of these YACs is required. The large size of the YAC insert relative to that of other cloning systems makes detailed mapping more difficult in YACs. To overcome this difficulty, the use of partial digests to map restriction sites relative to the vector sequence at either end of the cloned insert can be done employing the approach of

Smith and Birnsteil (1976) modified by using rare-cutting enzymes and pulsed-field gel electrophoresis to fractionate the products of digestion (Grootscholten et al., 1991). An attractive method to create a set of nested deletions (fragmentation) of a YAC has been developed by Pavan et al (1990) in which linear vector DNA (chromosome fragmenting vectors, CFVs) containing a single telomere, a selectable marker and an Alu repeat (integration target sequence) is transformed into (YAC containing) yeast. Recombination of the Alu sequence with a homologous target in the YAC results in formation of a series of shortened YACs which can be distinguished from the original YAC on the basis of the selectable markers. The panel of deletion derivatives generated can be used to facilitate restriction enzyme mapping. This is accomplished by monitoring the alteration of mobility and loss of restriction fragments from sequentially shorter YACs.

The proximal end-specific STS of 20CC8 is currently being used to extend the contig towards the DXS255 locus whereas the 4HG2 YAC has been used to extend the contig and link up with the distal DXS1003 YACs.

Microsatellite markers can be isolated from these YACs using the method described here and/or other methods and positioned by linkage analysis relative to the critical crossovers in RP2 families. Flanking markers closer than the existing ones should be identified in this way thus narrowing the RP2 containing region. Finally, these YACs can be characterised further in order to construct detailed transcriptional maps, or to identify disease genes. With no biochemical clues regarding the RP2 phenotype all relevant coding sequences must be identified and isolated (Riordan et al., 1991 ; The Huntington's Consortium, 1993). Several methods have been developed to accomplish this task. Originally cross-species conserved sequences were used as indicators for expressed sequences ("zooblotting" technique ; Monaco et al., 1986). Furthermore CpG islands (Bird, 1987) were found to be associated with genes, particularly with those of "housekeeping" function (Lindsay and Bird, 1987). CpG islands are characterised by a high density of unmethylated CpG dinucleotides that are conveniently identified by their sensitivity to rare-cutting

enzymes (Brown and Bird, 1986). The cloned DNA containing such a CpG island and/or a conserved sequence is then used to screen cDNA library filters leading to the isolation of cDNA clones (Sargent et al., 1989 ; Kendal et al., 1990).

The use of whole cosmid genomic inserts or YAC clones (Elvin et al., 1990) as hybridisation probes to screen cDNA libraries provides a direct method to identify transcribed sequences.

Recently another method has been developed allowing the quick and efficient isolation of cDNAs from cloned genomic DNA, called cDNA selection (Parimoo et al., 1992 ; Lovett et al., 1992). In this method the cDNA library inserts are amplified by PCR and the PCR products are hybridised to the immobilised genomic DNA of interest. Hybridised fragments are further amplified and subcloned. This method has been successfully used for example in the identification of the X-linked agammaglobulinaemia gene (Vetrie et al., 1993).

Exon trapping does not utilise cDNAs but relies on the presence of splicing signals in genomic DNA (Duyk et al., 1990 ; Buckler et al., 1991). The genomic DNA is subcloned into a specially constructed vector which facilitates transcription and splicing of exons. These are subsequently reverse transcribed and amplified by PCR.

The identification of coding sequences employing one of these methods may provide candidate genes for RP2 and other inherited diseases, which may be genetically mapped to these regions.

CHAPTER 6: DISCUSSION

The proximal region of the short arm of the human X chromosome is known from linkage studies to carry genes responsible for a number of hereditary eye diseases, retinitis pigmentosa (RP2), congenital stationary night blindness (CSNB1), Aland Island eye disease (AIED) and Norrie disease (NDP). The gene for Norrie disease was the first eye disease gene in this region to be cloned in 1992 using a positional cloning strategy facilitated by submicroscopic deletions detected in several NDP families. Studies aimed at the fine mapping of the remaining disease genes have been hampered by the relative paucity of available DNA markers in this region of the X chromosome. Because of the long standing interest of this laboratory in retinal degenerations and particularly in RP2, this study was undertaken to generate new markers to facilitate physical and genetic mapping in the Xp11.4-p11.22 region.

At the time this study was undertaken, a probable localisation of the RP2 locus between DXS7 and DXS426 was reported in 1990 by Coleman et al. based on a single recombination event occuring between DXS426 and RP2. Subsequently, however, the family with the critical recombinant event was characterised as RP3 based on additional genotyping data. This broadened the location of the RP2 gene in a region of 15cM which is the estimated genetic distance between the flanking markers DXS7/MAO (two loci physically separated by 150kb) and DXS255.

So far, linkage studies have failed to narrow down the position of the RP2 gene. A number of inherent difficulties contribute to this: 1) Localisation relies upon the genetic mapping of a few, often small families due to the rarity of the disease 2)The chromosome segment in which the RP2 gene has been localised has a lack of polymorphic DNA markers 3) RP2 is clinically indistinguishable from RP3 and characterisation of xlRP families relies on haplotype analysis using markers between the two loci 4) Genomic rearrangements associated with the disease which would have greately facilitated the identification of the gene have not been described.

FISH mapping of randomly selected X-specific cosmid clones was one approach for the development of novel, region-specific markers on the X chromosome with an ultimate objective to select those mapping in the region of interest for finer localisation. These cosmids were potential polymorphic markers since they contained microsatellites. Other techniques which can be employed to generate new clones include chromosome dissection (Ludecke et al., 1989) which, however, has not been extensively applied and the use of irradiation reduced hybrids (Goss and Harris, 1975 ; Benham et al., 1989) containing well defined segments of the X chromosome. Provided that such hybrids are thoroughly characterised they can then be a source for the generation of novel markers from the target region by using Alu-PCR (Nelson et al., 1989 ; Ledbetter et al., 1990).

FISH plays an increasingly important role in a variety of research areas including gene mapping. In this study the technique was initially used for the sublocalisation of 11 X-specific cosmid clones to X chromosome regions. Seven cosmids were found to map to the long arm of the X chromosome. Following the localisation of one of them, namely DXS566, to Xq13, a newly isolated microsatellite aided in the finer localisation of the Aarskog syndrome locus (Porteous et al., 1992). The physical localisation by FISH of a second cosmid, namely DXS554, was confirmed and refined by genetic mapping with respect to three markers (androgen receptor, phosphoglycerate kinase, pseudogene 1 and phosphoglycerate kinase) known to lie in the proximal long arm (Lindsay et al., 1993). The proximal long arm of the X chromosome contains several interesting genetic disease loci (e.g. ectodermal dysplasia and sideroblastic anaemia), as well as the X inactivation center. As yet, there are relatively few reported microsatellites in this region. These markers should, therefore, be helpful in mapping studies in this region of the X chromosome. 8 X-specific cosmid clones mapping to the proximal short arm of the X chromosome were more finely mapped relative to reference markers DXS7 and DXS426 by two colour FISH. Since FISH did not resolve the map position of one of them, namely HX43, finer localisation of this cosmid was achieved by isolation of a microsatellite

followed by genetic mapping with respect to reference markers in the region. The order of the cosmids relative to DXS7 and DXS426 as determined by two colour FISH is: (HX4, HX20, HX91)-(HX43, DXS7)-DXS426-(HX39, HX65, HX97, HX99). The order of HX43 relative to DXS7 and other reference markers as determined genetically is: OTC-HX43-DXS228-DXS7 (chapter 4). As yet, there is only one microsatellite, DXS228, reported in the interval between OTC and DXS7 hence the HX43 microsatellite is an additional marker in this interval. Microsatellites have been isolated from cosmids HX20 and HX91 (Thiselton et al., 1993 ; Thiselton et al., 1994, submitted) and as indicated by the two colour FISH results they also map in this interval. OTC in Xp21.1 is closely linked and proximal to RP3, and DXS7 in Xp11.3 is closely linked and distal to RP2. So far there are no reliable means to distinguish clinically the two forms of this disease. Markers mapping in this interval should, therefore, help to distinguish RP2 from RP3 families and further resolve the question of heterogeneity. The physical map of this region is not well defined. There is one YAC clone reported containing marker DXS228 in band Xp11.4. Markers like HX43 should be useful to isolate new YAC clones in this region. Marker HX97 is present in a YAC containing the OATL1 locus which lies in the interval between DXS426 and DXS255 which is currently the critical region for RP2 locus. The location of marker HX65 as visualised by two colour FISH is very similar to the location of HX97, therefore, it may lie in the same interval. Since HX65 is a potentially polymorphic marker and as few microsatellites have been reported in this interval, a microsatellite could be isolated (perhaps using the rapid PCR method described in chapter 5) and characterised.

In the absence of novel genetic material, a YAC-based physical map of the region was eventually required as another approach to generate additional genetic and physical markers. DXS426 was used as a starting point for a chromosomal walk during which 5 novel STSs were generated and ordered as described in chapter 5. The orientation of the two DXS426 YACs with respect to the centromere has proved to be difficult because, while this work was in progress, the order of adjacent

markers (the ARAF1/TIMP/SYN1 cluster) relative to DXS426 was reversed, while the order of PFC which also belongs to this cluster remained unresolved (XCW4, 1993). It was not possible to resolve the order of the two DXS426 YACs based on the presence of SYN1, since neither YAC contained the ARAF1/SYN1 microsatellite. The orientation of the YACs was resolved due to the presence of an end clone of one of the YACs in a proximal YAC containing OATL1. The presence of PFC in the distal DXS426 YAC and its absence from the proximal YAC placed PFC within 245kb distal to DXS426 thus resolving their order and contributing in the refinement of the physical map of this region of the X chromosome.

The physical mapping of the X chromosome has been extensive and large regions have now been cloned into YAC contigs. Furthermore, a first generation physical map based on the CEPH "megaYAC" library (Cohen et al., 1993) has been produced with the X chromosome well represented. However, the region Xp11.4-p11.22 remains relatively poorly characterised. The YAC contig around the DXS426 locus and the two DXS1003 YACs which lie in the DXS7-DXS426 interval, described in chapter 5 will, therefore, serve as the groundwork to improve the physical map of this interval both distally and proximally. Additionally, these YACs represent an invaluable starting point for the development of new genetic markers. These markers will be positioned by linkage analysis relative to the critical crossovers in RP2 families and may narrow the RP2 containing region which at the moment is substantial. To increase the probability for this to happen, new RP2 families need to be sought. Even if such families are discovered, and they are unambiguously characterised as RP2 families which contain favourably positioned meiotic events, the re-positioning of the gene is still likely to be within a few megabases. Of equal importance as the generation of additional markers is, therefore, the identification of all transcripts using as a source the YACs described here and other YACs in the candidate region. By screening tissue-specific cDNA libraries, followed by the analysis of each positive tissue-specific gene for alterations in individuals with the disease, a candidate gene may be identified. In the

absence of alterations, the ultimate proof will reside with sequence analysis and the demonstration of mutations. Since RP2, CSNB1 and AIED map in the same interval and furthermore, since there is the speculation that they may be allelic variants, a hypothesis which is enforced by the recent description of rhodopsin mutations causing both adRP and CSNB (Dryja et al., 1993 ; Rao et al., 1994), the analysis should be extended to patients with CSNB1 and AIED. It is also possible that the cloning of the gene for RP3, which has already been localised by genetic and physical means to an approximately 200kb region (Musarella et al., 1991), will be successful in the near future. Its characterisation may implicate homologous or candidate loci for RP2, since the two forms of the disease overlap at the clinical level. It is likely that in this way retinal-specific, region-specific genes will be isolated and it is hoped that mutations responsible for ocular phenotypes may be defined in the absence of refined genetic map data.

REFERENCES

Abbs, S., Roberts, R. G., Mathew, C. R., Bentley, M., and Bobrow, M. (1990). Accurate assessment of intragenic recombination frequency within the Duchenne muscular dystrophy gene. *Genomics* **7**: 602-606.

Abidi, F., Wada, M., Little, R. D., and Schlessinger, F. D. (1990). Yeast artificial chromosomes containing human Xq24-Xq28 DNA: library construction and representation of probe sequences. *Genomics* **7**: 363-376.

Anand, R., Riley J. H., Butler, R. Smith, J. C., and Markham, A. F. (1990). A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *NAR* **18**: 1952-1956.

Anand, R., Ogilvie, D. J., Butler, R., Riley, J. H., Finniear, R. S., Powell, S. J., Smith, J. C., and Markham, A. F. (1991). A yeast artificial chromosome contig encompassing the cystic fibrosis locus. *Genomics* **9**: 124-130.

Al-Maghtheh, M., Inglehearn, C. F., Keen, J., Evans, K., Moore, A. T., Jay, M., Bird, A. C., and Bhattacharya, S. S. (1994). Identification of a sixth locus for autosomal dominant retinitis pigmentosa on chromosome 19. *Human Molec. Genet.* **3**: 351-354.

Alitalo, T., Kruse, T. A., Forsius, H., Eriksson, A. A., de la Chapelle, A. (1991). Localisation of the Aland Island eye disease locus to the pericentromeric region of the X chromosome by linkage analysis. *Am. J. Hum. Genet.* **48**: 31-38.

Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsifidia, C. et al. (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* **355**: 548-551.

Attree, O., Olivos, I. M., Okabe, I., Bailey, L. C., Nelson, D. L., Lewis, R. A., McInnes, R. R., and Nussbaum, R. L. (1992). The Lowe's oculocerebrorenal syndrome gene encodes a protein highly homologous to inositol polyphosphate-5-phosphatase. *Nature* **358**: 239-242.

Avery, O. T., McLeod, C. M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of Pneumococcal types. *J. Exp. Med.* **79**: 137-157.

Baehr, W., Delvin, M. J., and Applebury, M. L. (1979). Isolation and characterisation of cGMP phosphodiesterase from bovine rod outer segments. *J. Biol. Chem.* **254**: 11669-11677.

Bahler, M., Benfenati, F., Valtorta, F., and Greengard, P. (1990). The synapsins and the regulation of synaptic function. *Bioessays* **12** : 259-263.

Ballabio, A., Bardoni, B., Guioli, S., Basler, E., and Camerino, G. (1990). Two families of low-copy-number repeats are interspersed on Xp22.3: implications for the high frequency of deletions in this region. *Genomics* **8**: 263.

Ballabio, A., and Willard, H. (1992). Mammalian X-chromosome inactivation and the XIST gene. *Curr. Opin. Genet. Dev.* **2**: 439-447.

Bartley, J., Geis, C., Jacobson, D. (1989). X-linked progressive cone dystrophy maps between DXS7 (L1.28) and DXS206 (XJ1.1) and is linked to DXS84 (754). *Cytogenet. Cell Genet.* **51**: 959.

Bauman, J. G., Wiegant, J., Borst, P., and van Duijn, P. (1980). A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridisation of fluorochrome labelled RNA. *Exp. Cell Res.* **128**: 485-490.

Bauman, J. G., Pinkel, D., Trask, B. J., and van der Ploeg, M. (1990). I*n* Flow Cytogenetics (Gray, J. W., ed.), 275-301, Academic Press.

Baumgartner, M., Dutrillaux, B., Lemieux, N., Lilienbaum, A. Paulin, D., and Viegas-Pequignot, E. (1991). Genes occupy a fixed and symmetrical position on sister chromatids. *Cell* **64**: 761.

Beadle, G. W. and Tatum, E. L. (1941). Genetic control of biochemical reactions in Neurospora. *Proc. Natl. Acad. Sci. USA* **27**: 499-506.

Bech-Hansen, N. T., Field, L. L., Schramm, A. M., et al. (1990). A locus for X-linked congenital stationary night blindness is located on the proximal portion of the short arm of the X chromosome. *Hum. Genet.* **84:** 406-408.

Bech-Hansen, N. T., Marshal, K. J., Moore, B. J., et al. (1991). Localisation of CSNB1 to region Xp11.3-Xp11.23 between MAO and DXS426-TIMP. HGM11. *Cytogenet. Cell Genet.* **58**: 216.

Bech-Hansen, N. T., Moore, B. J., and Pearce, W. G. (1992). Mapping of locus for X-linked congenital stationary night blindness (CSNB1) proximal to DXS7. *Genomics* **12**: 409-411.

Bellane-Chantelot, C., Barillot, E., Lacroix, B., Le Paslier, D., and Cohen, D. (1991). A test case for physical mapping of human genome by repetitive sequence fingerprints: construction of a physical map of a 420kb YAC subcloned into cosmids. *Nucleic Acids Res.* **19**: 505-510.

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.

Bentley, D. R., Todd, C., Collins, J., Holland, J., Dunham, I., Hassock, S., Bankier, A., and Giannelli, F. (1992). The development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries. *Genomics* **12**: 534-541.

Bergen, A. A. B., van den Born, L. I., Schuurman, E. J. M., Pinckers, A. J. L. G., Bleeker-Wagemakers, E. M., van Ommen, G. J. B., and Sandkuijl, L. A. (1994). Multipoint linkage analysis and homogeneity tests in 15 Dutch X-linked retinitis pigmentosa families. *Eur. J. Hum. Genet.* (in press).

Berger, W., Meindl, A., van de Pol, T. J. R., Cremers, F. P. M., Ropers, H. H., Doerner, C., Monaco, A., Bergen, A. A. B., Lebo, M., Warburg, M., Zergollern, L., Lorenz, B., Gal, A., Bleeker-Wagemakers, E. M., and Meitinger T. (1992). Isolation of a candidate gene for Norrie disease by positional cloning. *Nature Genetics* **1**: 199-203.

Bertelson, C. J., Pogo, A. O., Chaudhuri, A., Marsh, W. I., Redman, C. M., Banerjee, D., Symmans, W. A., Simon, T., Frey, D., and Kunkel, L. M. 91988).

Localization of the McLeod locus (XK) within Xp21 by deltion analysis. *Am. J. Hum. Genet.* **42**: 703-711.

Bhattacharya, S. S., Wright, A. F., Clayton, J. F., Price, W. H., Phillips, C. I., McKeown, C. M. E., Jay, M., Bird, A. C., Pearson, P. L., Southern, E. M. and Evans, H. J. (1984). Close linkage betwen X-linked retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. *Nature* **309**: 253-255.

Bird, A. C. (1975). X-linked retinitis pigmentosa. *Br. J. Ophthalmol.* **59**: 177-199.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* (London) **321**: 209-213.

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

Black, G. C. M., Chen, Z.-Y., Craig, I. W., and Powell, J. F. (1991). A polymorphic dinucleotide repeat in the MAOA locus. *Nucleic Acids Res.* **19**: 689.

Blanton, S. H., Heckenlively, J. R., Cottingham, A. W., Friedman, J., Sadler, L. A., Wagner, M., Friedman, L. H., and Daiger, S. P. (1991). Linkage mapping of autosomal dominant retinitis pigmentosa (RP1) to the pericentric region of human chromosome 8. *Genomics* **11**: 857-869.

Bok, D. (1985). Retinal photoreceptor-pigment epithelium interactions. Friedenwald lecture. *Invest. Opthal. Vis. Sci.* **26**: 1659-1694.

Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., Willard, H. F., Avner, P., and Ballabio, A. (1991). Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **351**: 325-328.

Botstein, D., White, R. L., Scolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* **32**: 314-331.

Boughman, J. A., Connealy, P. M., and Nance, W. E. (1980). Population genetic studies of retinitis pigmentosa. *Am. J. Hum. Genet.* **32**: 223-235.

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. USA* **88**: 1276-1280.

Boyd, Y. and Fraser, N. J. (1990). methylatuion patterns at the hypervariable X-chromosome locus DXS255 (M27 beta): correlation with X-inactivation status. *Genomics* **7**: 182-187.

Bowes, C. et al. (1990). Retinal degeneration in the *rd* mouse is caused by a defect in the beta subunit of rod cGMP-phosphodiesterase. *Nature* **347**: 677-680.

Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H. et al. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**: 799-808.

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

Brown, W. R. A. (1989). Molecular cloning of human telomeres in yeast. *Nature* **338**: 774-776.

Browne D. L. and Litt M. (1992). Characterization of $(CA)_n$ microsatellites with degenerate sequencing primers. *Nucleic Acids Research* **20**: 141.

Brownstein, B. H., Silverman, G. A., Little, R. D., Burke, D. T., Korsmeyer, S. J., Schlessinger, D. and Olson, M. V. (1989). Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244**: 1348-1351

Buckle, V. and Craig, I. W. (1986) In situ hybridisation, *In* Human genetic diseases: a practical approach, (ed. K. E. Davies), IRL Press, Oxford.

Buckler, A. J., Chang, D. D., Graw, S. L., Brook, J. D., Haber, D. A., Sharp, P. A., and Housman, D. E. (1991). Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. USA* **88**: 4005-4009.

Bundey, S. and Crews, S. J. (1984). A study of retinitis pigmentosa in the City of Birmingham. I Prevalence. *J. Med. Genet.* **21**: 417-420.

Burghes, A. H. M., Logan, C., Hu, X., Belfall, B., Worton, R., and Ray, P. N. (1987). Isolation of a cDNA clone from the region of an X ; 21 translocation that breaks within the Duchenne/Becker muscular dystrophy gene. *Nature (London)* **328**: 434-436.

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

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

Carrano, A. V., Lamrdin, J., Ashworth, L. k., Watkins, B., Branscomb, E., Slezak, T. et al. (1989). A high-resolution, fluorescence-based, semiautomated method for DNA fingerprinting. *Genomics* **4**: 129-136

Carle, G. F. and Olson, M. V. (1984). Separation of chromosomal molecules from yeast by orthogonal-field-alteration gel electrophoresis. *Nucleic Acids Res.* **12**: 5647-5664.

Caskey, C. T., Pizzuti, A., Fu, Y. H., Fenwick, R. G., and Nelson, D. L. (1992). Triplet repeat mutations in human disease. *Science* **256**: 784-789.

Cervetto, L., Lagnado, L., Perry, R. J., Robinson, D. W., and McNaughton, P. A. (1989). Extrusion of calcium from rod outer segments is driven by both sodium and potassium gradients. *Nature* **337**: 740-743.

Chapelle, A. E. de la, Snakila, E., Lindlof, M., Aula, P., Norio, K., (1985). Norrie disease caused by a gene deletion allowing carrier detection and prenatal diagnosis. *Clin. Genet.* **28**: 317-320.

Chelly, J., Gilgenkrantz, H., Lambert, M., Hamard, G., Chafey, P., Recan, D., et al. (1990). Effect of dystrophin gene deletions on mRNA levels and processing in Duchenne and Becker muscular dystrophies. *Cell* **63**: 1239-1248.

Chelly, J., Tumer, Z., Tonneson, T., Petterson, A., Ishikawa-Brush, Y., Tommerup, N., Horn, N., Monaco, A. P. (1993). Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. *Nature Genet.* **3**: 14-19.

Chen, J.-D., Halliday, F., Keith, G., Sheffield, L., Dickinson, P., Gray, R., Constable, I., and Denton, M. (1989). Linkage heterogeneity between X-linked retinitis pigmentosa and a map of 10 RFLP loci. *Am. J. Hum. Genet.* **45**: 401-411.

Chen, Z.-Y., Sims, K. B., Coleman, M., Donnai, D., Monaco, A., Breakfield, X. O., Davies, and Craig, I. W. (1992a). Characterisation of a YAC containing part or all of the Norrie disease locus. *Human Molecular Genetics* **1**: 161-164.

Chen, Z-Y., Hendriks, R. W., Jobling, M. A., Powell, J. F., Breakfield, X. O., Sims, K. B., and Craig, I. W. (1992b). Isolation and characterization of a candidate gene for Norrie disease. *Nature Genetics* **1**: 204-208.

Chromosome Coordinating Meeting (1992). Report of the committee on the genetic constitution of the X chromosome. Cutticchia, A. J., Pearson, P. L., Klinger, H. P. (eds). *Genome Priority Reports* **1**: 588-640 (Basel, Karger, 1993).

Clarke, T. R., Wilson, P. J., Morris, C. P., Hopwood, J. J., Richards, R. I., and Sutherland, G. R. (1992). Characterization of a deletion at Xq27-q28 associated with unbalanced inactivation of the nonmutant X chromosome. *Am. J. Hum. Genet.* **51**: 316-322.

Clayton, J. F., Wright, A. F., McKeown, C. M. E., Dempster, M., Jay, B. S., Bird, A. C., and Bhattacharya, S. S. (1986). Genetic linkage between X-linked retinitis pigmentosa and DNA probe DXS7 (L1.28): further linkage data, heterogeneity testing, and risk estimation. *Hum. Genet.* **74**: 168-171.

Coffey, A. J., Roberts, R. G., Green, E. D., Cole, C. G., Butler, R., Anand R., Gianelli F. and Bentley, D. R. (1992). Construction of a 2.6-Mb Contig in Yeast artificial chromosomes spanning the human dystrophin gene using an STS-based approach. *Genomics* **12**: 474-484.

Cohen, D., Chumakov, I., and Weissenbach, J. (1993). A first-generation physical map of the human genome. *Nature* **366**: 698-701.

Cole, C. G., Goodfellow, P. N., Bobrow, M. and Bentley, D. (1991). Generation of novel sequence tagged sites (STSs) from discrete chromosomal regions using *Alu*-PCR. *Genomics* **10**: 816-826.

Coleman M., Bhattacharya S., Lindsay S., Wright A., Jay M., Litt M., Craig I., and Davies K. (1990). Localisation of the microsatellite probe DXS426 between DXS7 and DXS255 on Xp and linkage to X-linked Retinitis Pigmentosa. *Am. J. Hum. Genet.* **47**: 935-940.

Coleman M. P., Murray, J. C. Willard, H. F., Nolan, K. F., Reid, K. B. M., Blake, D. J., Lindsay, S., Bhattacharya, S.S., Wright, A., and Davies, K. E. (1991). Genetic and physical mapping around the properdin P gene. *Genomics* **11**: 991-996.

Collins, F. A., Murphy, D. I., Reiss, A., Sims, K. B., Lewis, J., and Antonarakis, S. E. (1992). Clinical, biochemical and neuropsychiatric evaluation of a patient with a contiguous gene syndrome due to a microdeletion of Xp11.3 including the Norrie disease locus and monoamine oxidase (MAOA and MAOB) genes. *Am. J. Med. Genet.* **42**: 127-134.

Collins, F. S. (1992). Positional cloning: Let's not call it reverse anymore. *Nature Genet.* **1**: 3-6.

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

Cremer, T., Landegent, J., Bruckner, A., Scholl, H. P., Schardin, M., Hager, H. D., et al. (1986). Detection of chromosome-abberations in the human interphase nucleus by visualization of specific target DNAs with radioactive and non-radioactive in situ hybridisation techniques - Diagnosis of trisomy 18 with probe L1.84. *Hum. Genetic* **74**: 346-352.

Cremer, T., Lichter, P., Borden, J., Ward, D. C., and Manuelidis, L. (1988). Detection of chromosome aberrations in metaphase and interphase tumor cells by

in situ hybridisation using chromosome-specific library probes. *Hum. Genet.* **80**: 235-246.

Cremers, F. P. M., Sankila, E.-M., Brunsmann, F., Jay, M., Jay, B., Wright, A., Pinckers, L. G., Schwartz, M., van de Pol, D. J. R., Wieringa, B, de la Chapelle, A., Pawlowitzki, I. H., and Ropers, H.-H. (1990a). Deletions in patients with classical choroideraemia vary in size from 45kb to several megabases. *Am. J. Hum. Genet.* **47**: 622-628.

Cremers, F. P. M., van de Pol, D. J. R., van Kerkhoff, L. P. M., Wieringa, B., and Ropers, H.-H. (1990b). Cloning of a gene that is rearranged in patients with choroideraemia. *Nature* **347**: 674-677

Cross, S. H., Allshire, R. C., McKay, S J., McGill, N. I., and Cooke, H. J. (1989). Cloning of human telomeres by complementation in yeast. *Nature* **338**: 771-774.

Davies K. E., Mandel, J. L., Monaco, A. P., Nussbaum, R. L., Willard, H. F. (1992). Report of the committee on the genetic constitution of the X chromosome. 11th International Workshop on Human Gene Mapping, London. *Cytogenet. Cell Genet.* **58**: 853-966.

de Saint-Basile, G., Bohler, M. C., Fischer, A., Cartron, J., Dufier, J. L., Griscelli, C., Orkin, S. H. (1988). Xp21 microdeletion in a patient with chronic granulomatous disease, retinitis pigmentosa and McLeod phenotype. Hum. Genet. 80:85-89.

de Souza, A. P., Allamand, V., Richard, I., Brenguier, L., Chumakov, I., Cohen, D., and Beckmann, J. S. (1994). Targeted development of microsatellite markers from inter-*Alu* amplification of YAC clones. *Genomics* **19**: 391-393.

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 Duchenne muscular dystrophy mutations by field inversion gel electrophoresis. *Nature (London)* **329**: 640-642.

Den Dunnen, J. T., Grootscholten, P. M., Bakker, E., Blonden, L. A. J., Ginjaar, H. B., Wapenaar, M. C., et al. (1989). Topography of the Duchenne muscular

dystrophy gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am. J. Hum. Genet.* **45**: 835-847.

Den Dunnen, J. T., Grootscholten, P. M., Dauwerse, J. G., Walker, A. P., Monaco, A. P., Butler, R., Anand, R., Coffey, A. J., Bentley, D. R., Steensma, H. Y. et al. (1992). Reconstruction of the 2.4 Mb human DMD gene by homologous YAC recombination. *Hum. Mol. Genet.* **1**: 19-28.

Denton, M. J., Chen, J. D., Serraville, S., Colley, P., Halliday, F. B., and Donald, J. (1988). Analysis of linkage relationships of X-linked retinitis pigmentosa with the following Xp loci: L1. 28, OTC, 754, XJ-1.1, pERT87, and C7. *Hum. Genet.* **78**: 60-74.

Derry, J. M. J. and Barnard, P. J. (1992). Physical linkage of the A-raf-1, properdin, synapsin I, and TIMP genes on the human and mouse X chromosomes. *Genomics* **12**: 632-638.

Diergaarde, P. J., Wieringa, B., Bleeker-Wagemakers, I. M., Sims, K. B., Breakfield, X. O., and Ropers, H. H. (1989). Physical fine mapping of a deletion spanning the Norrie gene. *Hum. Genet.* **84**: 22-26.

Dizhoor, A. M., Ray, S., Kumar, S., Niemi, G., Spencer, M., Brolley, D., Walsh, K. A., Philipov, P. P., Hurley, J. B., and Stryer, L. (1991). Recoverin: A calcium sensitive activator of retinal rod guanylate cyclase. *Science* **251**: 914-918.

Dombroski, B. A., Mathias, S. L., Nanthakamur, A. F., Scott, A. F., and Kazazian, H. H. (1991). Isolation of an active human transposable element. *Science* **254**: 1805-1808.

Donnai, D., Mountford, R. C., and Read, A. P. (1988). Norrie disease resulting from a gene deletion: clinical features and DNA studies. *J. Med. Genet.* **25**: 73-78.

Donnis-Keller, H., Green, P., Helms, C., et al. (1987). A genetic linkage map of the human genome. *Cell* **51**: 319-337.

Dratz, E. A. and Hargrave, P. A. (1983). The structure of rhodopsin and the rod outer segment disc membrane. *TIBS* April: 128-131.

Dryja, T. D., McGee, T. L., Reichel, E., Hahn, L. B., Cowley, G. S., Yandel, D. N., Sandberg, M. A., and Berson, E. L. (1990). A point mutation of thr rhodopsin gene in one form of retinitis pigmentosa. *Nature* **343**: 364-366.

Dryja, T. D., Berson, E. L., Rao, V. R., and Oprian, D. D. (1993). Heterozygous missense mutation in the rhodopsin gene as a cause of congenital stationary night blindness. *Nature Genet.* **4**: 280-283.

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

Economou, E. P., Bergen A. W., Warren, A. C. and Antonarakis, S. E. (1990). The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. *Proc. Natl Acad. Sci. USA* **87**: 2951-2954.

Edwards, A., Civitello, A., Hammond, H. A. and Caskey, T. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**: 746-756.

Elenius, V., Eriksson, A., Forsius, H. (1968). ERG in a case of X-chromosomal pigment deficiency of fundus in combination with myopia, dyschromatopsia and defective dark-adaptation. In*: The clinical value of electroretinography.* ISCRG Symposium, Ghent, 1966. Basel/New York: Karger: pp. 369-377.

Elvin, P., Slynn, G., Black, D., Graham, A., Butler, R., Riley, J., Anand, R., and Markham, A. F. (1990). Isolation of cDNA clones using Yeast Artificial Chromosome probes. *Nucleic Acids Research* **18**: 3913-3917.

Evans, H. J., Buckland, R. A. and Pardue, M. L. (1974). Location of the genes coding for 18S and 28S ribosomal RNA in the human genome. *Chromosoma (Berl.)* **48**: 405-406.

Farrar, G.J., Jordan, S. J., Kenna, P., Humphries, M. M., Kumar-Singh, R., McWilliam, P., Allamand, V., Sharp, E. M., and Humphries, P. (1991a). Autosomal retinitis pigmentosa: localisation of a disease gene (RP6) to the short arm of chromosome 6. *Genomics* **8**: 35-40.

Farrar, G. J., Kenna, P., Jordan, S. A., Kumar-Singh, R., Humphries, M. M., Sharp, E. M., Sheils, D. M., and Humphries, P. (1991b). A three base-deletion in the peripherin-RDS gene in one form of retinitis pigmentosa. *Nature* **354**: 478-480.

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

Feinberg, P. A. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**: 6-13.

Fishman, G. A., Weinberg, A. W., McMahon, T. T., (1986). X-linked retinitis pigmentosa: clinical characteristics of carriers. *Arch. Ophthalmol.* 104: 1329-1335.

Fodor, E., Lee, R. T. and O' Donnell, J. J. (1991). Analysis of choroidaeremia gene. *Nature* **351**: 614.

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

Fowles, C., Akhtar, M. and Cohen, P. (1989). Interplay of phosphorylation and dephosphorylation in vision-protein phosphatases of bovine rod outer segments. *Biochemistry* **28**: 9385-9391.

Franco, B., Guioli, S., Pragliola, A. et al. (1991). A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* **353**: 529-

Francke, U., Ochs, H. D., Martinville de, B., Giacalone, J., Lindgren, V., Disteche, 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., Brownlee, G. G., and Craig, I. W. (1987). Multi-allelic RFLP for M27-beta, an anonynous single copy genomic clone at Xp11.3-Xcen. *Nucleic Acids Res.* **15**: 9616.

Freije, D., Helms, C., Watson, M. S., and Donnis-Keller, H. (1992). Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science* **258**: 1784-1787.

Friedrich, U., Warburg, M., Wieacker, P., Wienker, T. F., Gal, A., Ropers, H.-H. (1985). X-linked retinitis pigmentosa: linkage with the centromere and a cloned DNA sequence from the proximal short arm of the X chromosome. *Hum. Genet.* **71**: 93-99.

Friedrich, U., Warburg, M., Kruse, T. A., Andreasson, S. (1992). X-linked retinitis pigmentosa: new map studies of XLRP2, and a possible human centromere effect. *Hum. Genet.* **88**: 683-687.

Fu, Y.-H., Kuhl, D. P. A., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J. M. H., Holden, J. J. A., Fenwick, R. G., Jr., Warren, S.T., Oostra, B. A., Nelson, D. L., and Caskey, C.T. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **67**: 1047-1058.

Fung, B. K., Hurley, J. B.,and Stryer, L. (1981). Flow of information in the light-triggered cyclic nucleotide cascade of vision. *Proc. Natl. Acad. Sci. USA* **78**: 152-156.

Fung, B. K., Young, J. H., Yamane, H. K., and Griswold-Prenner, I. (1990). Subunit stoichiometry of retinal rod cGMP phosphodiesterase. *Biochemistry* **29**: 2657-2664.

Fuscoe, J. C., Clark, L. M., and Van Dilla, M. A. (1986). Construction of fifteen human chromosome-specific DNA libraries from flow-sorted chromosomes. *Cytogenet. Cell Genet.* **43**: 79-86.

Gal, A., Bleeker-Wagemakers, L. M., Wienkar, T. F., Warburg, M., Ropers, H. H. (1985). Localisation of the gene for Norrie disease by linkage to the DXS7 locus. *Cytogen. Cell Genet.* **40**: 633.

Gal, A., Wieringa, B., Smeets, D. F. C. M., Bleeker-Wagemakers, L. M., Ropers, H. H. (1986). Submicroscopic interstitial deletion of the X chromosome explains a complex genetic syndrome dominated by Norrie disease. *Cytogen. Cell Genet.* **42**: 219-224.

Gal, A., Schinzelm, A., Ulrike, O., et al. (1989). Gene of X-chromosomal congenital stationary night blindness is closely linked to DXS7 on Xp. *Hum. Genet.* **81**: 315-318.

Gall, J. G. and Pardue, M. (1969). Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proc. Natl Acad. Sci. USA* **63**: 378-383.

Garrod, A. E. (1902). The incidence of alkaptonuria: a study in chemical individuality. *Lancet* **1902/II**: 1616-1620.

Gasson, J. C., Golde, D. W., Kaufman, S. E., Westbrook, C. A., Hewick, R. M., Kaufman, R. J., et al. (1985). Molecular characterization and expression of the gene encoding human erythroid-potentiating activity. *Nature* **315**: 768-771.

Geraghty, M. T., Brody, L. C., Martin, L. S., Marble, M., Kearns, W., Pearson, P., Monaco, A. P. Lehrach, H., and Valle, D. (1993). The isolation of cDNAs from OATL1 at Xp11.2 using a 480-kb YAC. *Genomics* **16**: 440-446.

Gerhard, D. S., Kawasaki, E. S., Bancroft, F. C., and Szabo, P. (1981). Localization of a unique gene by direct hybridisation in situ. P*roc. Natl. Acad. Sci. USA* **78**: 3755-3759.

Giannelli, F., Green, P. M., High, K. A., Sommer, S., Lillicrap, D. P., Ludwig, M. et al. (1992). *Nucleic Acids Research* **20**: 2027-2063.

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

Gosden, J. R., Mitchell, A. R., Buckland, R. A., Clayton, R. P., and Evans, H. S. (1975). The location of four human satellite DNAs on human chromosomes. *Exp. Cell Res.* **92**: 148-158.

Goss, S J., and Harris, H. (1977). Gene transfer by means of cell fusion. *J. Cell Sci.* **25**: 17-37.,

Goundis, D. and Reid, K. B. M. (1988). Properdin, the terminal complement components, thrombospondin and CS protein of malaria parasites contain similar sequence motifs. *Nature* **335**: 82-85.

Goundis, D., Holt, S., Boyd, Y., and Reid, K. B. M. (1989). Localization of the properdin structural locus to Xp11.23-Xp21.1. *Genomics* **5**: 397-402.

Green, E. D. and Olson, M. V. (1990a). Systematic screening of yeast artificial chromosome libraries by use of polymerase chain reaction. *Proc. Natl Acad. Sci. USA* **87**: 1213-1217.

Green, E. D. and Olson, M. V. (1990b). Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. *Science* **250**: 94-98.

Green, E.D., Riethman, H.C., Dutchik, J.E. and Olson, M.V. (1991). Detection and characterisation of chimeric yeast artificial-chromosome clones. Genomics 11: 658-669.

Grootscholten, P. M., Den Dunnen, J. T., Monaco, A. P., Anand, R., and Van Ommen, G. J. B. (1991). YAC mapping strategies applied to the DMD-gene. *Technique* **3**: 41-50.

Gusella, J. F., Keys, C., Varsanyi-Breiner, A., Kao, F. T., Jones, C., Puck, T. T. and Housman, D. (1980). Isolation and localisation of DNA segments from specific human chromosomes. *Proc. Natl Acad. Sci. USA* **77**: 2829-2833.

Hadjiconstantinou, M., Cohen, J., and Neff, N. H. (1983). Epinephrine: a potential neurotransmitter in retina. *Neurochem.* **41**: 1440-1444.

Hamada, H., Petrino, M. G. and Kakanuga, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eucaryotic genomes. *Proc. Natl Acad. Sci. USA* **79**: 6465-6469.

Hardwick, L. J., Brown, J., Dry, K. L., Lester, D. H., Aldred, M. A., Jay, M., Bird, A. C., and Wright, A. F. (1993). Isolation of microsatellite and expressed sequences from a yeast artificial chromosome from the RP3 region using biotin-streptavidin selection. *Am. J. Hum. Genet.* **53**: 1300.

Harper, M. E., Ullrich, A. and Saunders, G. F. (1981). Localization of the human insulin gene to the distal end of the short arm of chromosome 11. *Proc. Natl. Acad. Sci. USA* **78**: 4458.

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

Hecht, S., Shlaer, S., and Pirenne, M. H. (1942). *J. Gen. Physiol.* **25**: 819-840.

Heckenlively, J. R. (1988). *Retinitis Pigmentosa.* J. B. Lippencott, Philadelphia.

Hendriks, R. W., Chen, Z. Y., Hinds, H., Schuurman, R. K. B., and Craig, I. W. (1992). An X chromosome inactivation assay based on differential methylation of a CpG island coupled to a VNTR polymorphism at the 5' end of the monoamine oxidase A gene. *Hum. Mol. Genet.* **1**: 187-194.

Ho, M. F., Monaco, A. P., Blonden, L. A. J., van Ommen, G. J. B., Affara, N. A., Ferguson-Smith, M. A., and Lehrach, H. (1992). Fine mapping of the McLeod locus (XK) to a 150-380kb region in Xp21. *Am J. Hum Genet.* **50**: 317-330

Holme, E. R., Veitch, J., Johnston, A., Hauptmann, G., Uring-Lambert, B., Seywright, M., et al. (1989). Familial properdin deficiency associated with chronic discoid lupus erythematosus. *Clin. Exp. Immunol.* **76**: 76-81.

Hong, H. K., Forrel, R. E., Paul, T. O., Gorin, M. B. (1991). Localization of an X-linked progressive cone degeneration to Xp11 region. 8th International Congress of Human Genetics. *Am. J. Hum. Genet.* **138**: 703 (abstract).

Hopman, A. H. N., Ramaekers F. C. S., Raap, A. K., Beck, J. L. M., Devilee, P., van der Ploeg M., and Vooijs, G. P. (1988). In situ hybridisation as a tool to study numerical chromosome aberrations in solid tumor bladder tumors. *Histochemistry* **89**: 307-316.

Horn, G. T., Richards, B. and Klinger, K. W. (1989). Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. *Nuclic Acid Research* **17**: 2140.

Hsu, Y. T. and Molday, R. S. (1993). Modulation of the cGMP-gated channel of rod photoreceptor cells by calmodulin. *Nature* **361**: 76-79.

Hu, X., Ray, P. N., Murphy, E. G., Thompson, M. W., and Worton, R. G. (1990). Duplicational mutation at the Duchenne muscular dystrophy locus: Its frequency, distribution, origin, and phenotype genotype correlation. *Am. J. Hum. Genet.* **46**: 682-695.

Huebner, K., AR-Rushdi, A., Griffin, C. A., Isobe, M., Kozak, C., Emanuel, B. S., et al. (1986). Actively transcribed genes in the *raf* oncogene group, located on the X chromosome in mouse and human. *Proc. Natl. Acad. Sci. USA* **83**: 3934-3938.

Huleihel, M., Goldsborough, M., Cleveland, J., Gunnell, M., Bonner, T., and Rapp, U. R. (1986). Characterization of murine A-*raf*, a new oncogene related to the v-*raf* oncogene. *Mol. Cell. Biol.* **6**: 2655-2662.

Humphries, P., Farrar, G. J., and Kenna, P. (1993). Autosomal dominant retinitis pigmentosa: Molecular, genetic and clinical aspects. *Progress in Retinal Research* **12**: 231-245.

Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**: 971-983.

Hurley, J. B., and Stryer, L. (1982). Purification and characterization of the gamma-regulatory subunit of the cyclic-GMP phosphodiesterase from retinal rod outer segments. *J. Biol. Chem.* **257**: 11094-11099.

Hurley, J. B., Dizhoor, A. M., and Stryer, L. (1993). Recoverin's role: Conclusion withdrawn. *Science* **260**: 740.

Huxley, C., Green, E. D., and Dunham, I. (1990). rapid assessment of S. cerevisiae mating type by PCR. *Trends Genet.* **6**: 236.

Ingle, C., Williamson, R., De la Chapelle, Herva, R. R., Haapala, K., Bates, G., Willard, H. F., Pearson, P., and Davies, K. E. (1985). Mapping DNA sequences in a human X chromosome deletion which extends across the region of the Duchenne muscular dystrophy mutation. *Am. J. Hum. Genet.* **37**: 451-462.

Inglehearn, C. F., Carter, S. A., Keen, J., Lindsey, J., Stephenson, A. M., Bashir, R., Al-Maghtheh, M., Moore, A. T., Jay, M., Bird, A., and Bhattacharya, S. S. (1993).

A new locus for autosomal dominant retinitis pigmentosa on chromosome 7p. *Nature Genet* **4**: 51-53.

Ingram, V. M. (1957). Gene mutations in human haemoglobin.: the chemical difference between normal and sickle cell haemoglobin. *Nature* **180**: 326-328.

Jacobson, D. M., Thompson, H. S., Bartley, J. A. (1989). X-linked progressive cone dystrophy. Clinical characteristics of affected males and carriers. *Ophthalmology* **96**: 885-895.

Jay, M. (1982). On the heredity of retinitis pigmentosa. *Br. J. Ophthalmol.* **66**: 405-416.

Jeffreys, A. J., Wilson, V. and Thein, S. W. (1985). Hypervariable "minisatellite" regions in human DNA. *Nature* **314**: 67-73.

Jones, K. W., Prosser, I., Corneo, G. and Ginelli, E. (1973). The chromosomal location of human satellite DNA III. *Chromosoma (Berl.)* **42**: 445-451.

Jordan, S. A., Farrar, G. J., Kenna, P., Humphries, M. M. Sheils, D. M., Kumar-Singh, R., Sharp, E. M., Soriano, N., Ayuso, C., Benitez, J., and Humphries, P. (1993). Localization of an autosomal dominant retinitis pigmentosa gene to chromosome 7q. *Nature Genet.* **4**: 54-58.

Kajiwara, K., Hahn, L. b., Mukai, S., Travis, G. H., Berson, E. L., and Dryja, T. P. (1991). Mutations in the human RDS gene in patients with autosomal retinitis pigmentosa. *Nature* **354**: 480-483.

Kaplan, J., Pelet, A., Martin, C., Delrieu, O., Ayme, S., Bonneau, d., Briard, M. I., Hanauer, A., Larget-Piet, L., Lefrancois, P., Michel-Awad, A., Plauchu, H., Dufier, J. L., Frezal, J., Munnich, A. (1992). Phenotype-genotype correlations in X linked retinitis pigmentosa. *J. Med. Genet* **29**: 615-623.

Karpen, J. W., Zimmerman, A. L., Stryer, L., and Baylor, D. A. (1988). Gating kinetics of the cyclic-GMP-activated channel of retinal rods flash-photolysis and voltage-jump studies. *Proc. Acad. Natl. Sci. USA* **85**: 1287-1291.

Katayama, S., Wohlferd, M., and Golbus, M. S. (1988). First demonstration of recombination between the gene for Norrie disease and probe L1.28. *Am. J. Med. Genet.* **30**: 967-970.

Kaupp, B. U., Niidome, T., Tanabe, T., Terada, S., Bonigk, W., Stuhmer, W., Cook, N. J., Kangawa, K., Matsuo, H., Hirose, T., Miyata, T., and Numa, S. (1989). Primary structure and functional expression from complementary DNA of the rod photoreceptor cyclic GMP-gated channel. *Nature* **342**: 762-766.

Kendall, E., Sargent, C. A., and Campbell, R. D. (1990). Human major histocompatibility complex contains a new cluster of genes between the HLA-D and complement loci. *Nucleic Acids Research* **18**: 7251-7257.

Khalfan, H., Abuknesha, R., Rand-Weaver, R., Price, R. G., and Robinson, D. (1986). AMCA: a new fluorescent labelling agent. *Histochemical Journal* **18**: 497-499.

Killimann, M. W. and DeGennaro, L. J. (1985). Molecular cloning of cDNAs for the nerve-cell specific phosphoprotein, synapsin I. *EMBO J.* **4**: 1997-2002.

Kivlin, J. D., Sanborn, G. E., Wright, E., Cannon, L., Carey, J. (1987). Brief clinical report: further linkage data on Norrie disease. *Am J. Med. Genet.* **26**: 733-736.

Knight, J. C., Reeves, B. R., Kearny, Monaco, A. P., Lehrach, H., and Cooper, C. S. (1992). Localization of the synovial sarcoma t (X ; 18)(p11.2 ; q11.2) breakpoint by fluorescence in situ hybridisation. *Hum. Molec. Genet.* **1**: 633-637.

Koch, K. W., and Stryer, L. (1988). Highly cooperative feedback-control of retinal rod guanylate-cyclase by calcium-ions. *Nature* **334**: 64-66.

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., Muller, C. R. et al. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am. J. Hum. Genet.* **45**: 498-506.

Konradi, C., Ozelius, L., and Breakfield, X. O. (1992). Highly polymorphic $(GT)_n$ repeat sequence in intron II of the human MAOB gene. *Genomics* **12**: 176-177.

Koref, S. M., Orphanos V., and Boyle J. M. (1993). Rapid determination of sequences flanking microsatellites using dephosphorylated cloning vectors. *TIG* **9**: 43.

Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Schlessinger, D., Sutherland, G. R., and Richards, R. I. (1991). Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CCG)_n$. *Science* **252**: 1711-1714.

Krill, A. (1977). Congenital stationary night blindness. In: *Krill's hereditary retinal and choroidal diseases.* **Vol II**. Hagers-town: Harper and Row: 391-420.

Kunkel, L. M., Monaco, A. P., Midlesworth, 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. USA* **82**: 4778-82.

Kwan, S.-P., Sandkuyl, L. A., Blaese, M., Kunkel, L. M., Bruns, G., Parmley, R., Skarshaug, S., Page, D. C., Ott, J., and Rosen, F. S. (1988). Genetic mapping of the Wiscott-Aldrich syndrome with two highly-linked polymorphic DNA markers. *Genomics* **3**: 39-43.

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

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

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

Landegent, J., E., Jansen in de Wal N., van Ommen, G. J., Baas, F., de Vijlder, J. J. M., van Duijn, P., and van der Ploeg, M. (1985). Chromosomal localisation of a unique gene by non-autoradiographic in situ hybridisation. *Nature* **317**: 175-177.

Landegent, J. E., Jansen in de Wal, N., Dirks, R. W., Baas, F., van der Ploeg, M. (1987). Use of whole cosmid cloned genomic sequences for chromosomal localisation by non-radioactive in situ hybridisation. *Hum Genet* **77**: 366-370.

Langer, P. R., Waldrop, A. A., and Ward, D.C. (1981). Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. *Proc. Natl Acad. Sci.* **78**: 6633-7.

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

La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, K. H., and Fischbeck, K. H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**: 77.

Lathrop G. M. and Lalouel, J. M. (1984). Easy calculation of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* **36**: 1245-1249.

Lauder, J. M. (1985). Roles for neurotransmitters in development: possible interaction with drugs during fetal and neonatal periods. *Prog. Clin. Biol. Res.* **163**: 375-380.

Lawrence, J. B., and Singer, R. H. (1985). Quantitative analysis of in situ hybridisation for the detection of actin gene expression. *Nucleic Acids Research* **13**: 1777-1799.

Lawrence, J. B., Villanave, C. A., and Singer, R. H. (1988). Sensitive, high-resolution chromatin and chromosome mapping *in situ*: presence and orientation of two closely integrated copies of EBV in a lymphoma line. *Cell* **52**: 51-61.

LeBlanc-Straceski, J. M., Montgomery, K. T., Kissel, H., Murtaugh, L., Tsai, P., Ward, D. C., Krauter, K. S., and Kucherlapati, R. (1994). Twenty-one polymorphic markers from human chromosome 12 for intergration of genetic and physical maps. *Genomics* **19**: 341-349.

Ledbetter, S. A., Nelson, D. L., Warren, S.T. and Ledbetter, D. H. (1990). Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics* **6**: 475-481.

Leeuw de B., Berger, W., Sinke, R. J., Suijkerbuijk, T., Gilgenkrantz, S., Geraghty, M., et al. (1993). Identification of a yeast artificial chromosome (YAC) spanning the synovial sarcoma-specific t (X ; 18)(p11.2 ; q11.2) breakpoint. *Genes Chrom. Cancer* **6**: 182-189.

Legouis, R., Hardelin, J. P., Levilliers, J., Claverie, J. M., Compain, S., Wunderle, V., Millasseau, P. et al. (1991). The candidate gene for the X-linked Kallmann encodes a protein related to adhesion molecules. *Cell* **67**: 423-435.

Levitt, P., Pintar, J. E., and Breakfield, X. O. (1982). Immunocytochemical demonstration of monoamine oxidase B in brain astrocytes and serotonergic neurons. *Proc. Natl. Acad. Sci. USA* **79**: 6385-6389.

Li, Y., Bleeker-Wagemakers, E., Artlich, A., et al. (1991). The locus for X-chromosomal congenital stationary night blindness is closely linked to TIMP in two families but not in a third one. *Cytogenet. Cell Genet.* 58: 235-236.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridisation using recombinant DNA libraries. H*um. Genet.* **80**: 224-234.

Lichter, P., Tang, C-J. C., Call, K., Hermanson, G., Evans, G. A., Housman, D., and Ward, D. C. (1990). High-resolution mapping of human chromosome 11 by in situ hybridisation with cosmid clones. *Science* **247**: 64-69.

Lindsay, S. and Bird, A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327**: 336-338.

Lindsay, S., Thiselton, D. L., Bateman, J. B., Ngo, J. T., Sparkes, R. S., Coleman, M., Davies, K. E., and Bhattacharya, S. S. (1992). Localisation of the gene for Norrie disease to between DXS7 and DXS426 on Xp. *Hum Genet* **88**: 349-350.

Lindsay, S., Curtis, A. R. J., Roustan, P., Kamakari, S., Thiselton, D., Stephenson, A., and Bhattacharya, S. S. (1993). Isolation and characterisation of three microsatellite markers in the proximal long arm of the human X chromosome. Genomics **17**: 208-210.

Litt, M. and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**: 397-401.

Litt, M. (1991). in PCR: A Practical Approach (McPherson, M.C., Quirke, P. and Taylor, G., eds), 85-99, Oxford University Press.

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

Lovett, M., Kere, J., and Hinton, L. M. (1991). Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci.* **88**: 9628-9632.

Lowe, T., Sharefkin, J., Yang, S. Q., and Dieffenbach, C. W. (1990). A computer program for selection of oligonucleotide primers for polymerase chain reaction. *Nucleic Acid Research* **18**: 1757-1761.

Ludecke, H. J., Senger, G., Claussen, U. and Horsthemke, B. (1989). Cloning defined regions of the human genome by microdissection of banded chromosomes and enzymatic amplification. *Nature* **338**: 348-350.

Luty, J. A., Willard, H. F. and Litt M. (1989). Three new microsatellite VNTRs on the X chromosome. *Cytogenetics Cell Genetics* **51**: 1036

Luty, J. A., Guo, Z., Willard, H. F., Ledbetter, D. H., Ledbetter, S., and Litt, M. (1990). Five polymorphic microsatellite VNTRs on the human X chromosome. *Am. J. Hum. Genet.* **46**: 776-783.

Lund, E., Bostock, C., Robertson, M., Christie, S., Mitchen, J. L., and Dahlberg, J. E. (1983). U1 small nuclear-RNA genes are located on human chromosome-1 and are expressed in mouse-human hybrid-cells. *Molec. cell. Biol.* **3**: 2211-2220.

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

Maestrini, E., Tamanini, F., Kioschis, P., Gimbo, E., Marinelli, P., Tribioli, C., D'Urso, M., Palmieri, G., Poustka, A., and Toniolo, D. (1992). An archipelago of CpG islands in Xq28: identification and fine mapping of 20 new CpG islands of the human X chromosome. *Hum. Molecul. Genet.* **1**: 275.

Mahadevan, M., Tsilfidis, C. Sabourin, L., Shutler, G. Amemiya, C., Jansen, G. et al. (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**: 1253-1255.

Mahtani, M. M., and Willard, H. F. (1988). A primary map of the pericentromeric region of the human X chromosome. *Genomics* **2**: 294-301.

Mahtani M. M., Lafreniere, R. G., Kruse, T. A., Willard, H. F. (1991). An 18-locus map of the pericentromeric region of the human X chromosome: genetic framework for mapping X-linked disorders. *Genomics* **10**: 849-857.

Malcolm, S., Barton, P., Murphy, C., and Ferguson-Smith, M. A. (1981). Chromosomal localisation of a single-copy gene by in situ hybridisation - human beta globin genes on the short arm of chromosome 11. *Ann. Hum. Genet.* **45**: 135-141.

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**: 756-759.

Mandel, J.-L., Monaco, A. P., Nelson, D. L., Schlessinger, D., and Willard, H. (1992). Genome analysis and the human X chromosome. *Science* **258**: 103-109.

Maniatis, T., Fritsch, E. F., Lauer, J., and Lawn, R. M. (1980). The molecular genetics of human hemoglobin. *Ann. Rev. Genet.* **14**: 145-178.

McKusick, V. A. (1992). Mendelian inheritance in man. John Hopkins Univ. Press, Baltimore/London ed. 10.

McLaughlin, M. E., Sandberg, M. A., Berson, E. L., and Dryja, T. P. (1993). Recessive mutations in the gene encoding the β-subunit of rod phosphodiesterase in patients with retinitis pigmentosa. *Nature Genet.* **4**: 130-134.

McWilliam, P., Farrar, G. J., Kenna, P., Bradley, D. G., Humphries, M. M., Sharp, E. M., McConnell, D. J., Lawler, M., Sheils, D., Ryan, C., Daiger, S. P., and Humphries, P. (1989). Autosomal dominant retinitis pigmentosa (adRP): localization of an adRP gene to the long arm of chromosome 3. *Genomics* **5**: 619-622.

Meindl, A., Berger, W., Meitinger, T., van de Pol, D., Achatz, H., Dorner, C., Haasemann, M., Hellebrand, H., Gal, A., Cremers, F., and Ropers, H.-H. (1992). Norrie disease is caused by mutations in an extracellular protein resembling C-terminal globular domain of mucins. *Nature Genetics* **2**: 139-143.

Meire, F. M., Bergen, A. A. B., de Rouck, A., Leys, M., and Delleman, J. W. (1994). X linked progressive cone dystrophy. Localisation of the gene locus to Xp21-p11.1 by linkage analysis. *Br. J. Ophthalmol.* **78**: 103-108.

Meitinger, T., Fraser, N. A., Lorenz, B., Zrenner, E., Murken, J., and Craig, I. W. (1989). Linkage of X-linked retinitis pigmentosa to the hypervariable DNA marker M27β (DXS255). *Hum. Genet.* **81**: 283-286.

Mercer, J., Livingston, J., Hall, B., Paynter, J., Begy, C., Chandrasekharappa, S., Lockhart, P., grimes, A., Bhave, M., Siemieniak, D., and Glover, T. W. (1993).

Isolation of a candidate gene for Menkes disease by positional cloning. *Nature Genet.* **3**: 20-25.

Merry, D. E., Janne, P. A., Landers, J. E., Lewis, R. A., and Nussbaum, R. L. (1992). Isolation of a candidate gene for choroideraemia. *Proc. Natl. Acad. Sci. USA* **89**: 2135-2139.

Messing, J. (1983). In Methods in Enzymology 101: 20-78.

Migeon, B. R. (1990). Insight into X-chromosome inactivation from studies of species variation, DNA methylation and replication, and vice versa. *Genet. Res.* **56**: 91.

Miller, J. L., Fox, D. A., and Litman, B. J. (1986). Amplification of phosphodiesterase activation is greatly reduced by rhodopsin phosphorylation. *Biochemistry* **25**: 4983-4988.

Mitchell, G. A., Looney, J. E., Brody, L. C., Steel, G., Suchanek, M., Englehardt, J. F., Willard, H. F., and Valle, D. (1988). Human ornithine-δ-aninotransferase: cDNA cloning and analysis of the structural gene. *J. Biol. Chem.* **263**: 14228-14295.

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

Monaco, A. P., Walker, A. P., Millwood, I., Larin, Z. and Lehrach, H. (1992). A yeast artificial chromosome contig containing the complete Duchenne muscular dystrophy gene. *Genomics* **12**: 465-473.

Monji, N., Ali, H., and Castro, A. (1980). Quantification of digoxin by enzyme immunoassay: synthesis of a maleimide derivative of digoxigenin succinate for enzyme coupling. *Experimentia* **36**: 1141-3.

Molday, L. L., Cook, N. J., Kaupp, U. B., and Molday, R. S. (1990). The cGMP-gated cation channel of bovine rod photoreceptor cells is associated with a 240kDa protein exhibiting immunochemical cross-reactivity with spectrin. *J. Biol. Chem.* **226**: 21917-21922.

Moore, B. J., Kwan, S. P. and Bech-Hansen N. T. (1992). A polymorphic dinucleotide repeat at the DXS7 locus. *Nucleic Acids Res.* **20**: 929.

Mosser, J., Douar, A. M., Sarde, C. O., Kioschis, P., Feil, R., Moser, H., Poutska, A. M., Mandel, J.-L., and Aubourg, P. (1992). Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* **361**: 726-730.

Mukai, S., Dryja, T. P., Bruns, G. A. P., Aldridge, J. f., and Berson, E. L. (1985). Linkage between the X-linked retinitis pigmentosa locus and the L1.28 locus. *Am. J. Ophthalmol.* 100: 225-229.

Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. *Sci. Am.* **262**: 56-61.

Murray, J. M., Davies, K. E., Harper, P. S., Meredith, L., Mueller, C. R., 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.

Murphy, D. L., Sims, K. B., Karoum, F., de la Chapelle, A., Norio, R., Sankila, E. M., and Breakfield, X. O. (1990). Marked amine and amine metabolite changes in Norrie disease patients with an X chromosomal deletion affecting monoamine oxidase. J. *Neurochem.* **54**: 242-247.

Musarella, M. A., Burghes, A., Anson-Cartwright, L., Mahtani, M. M., Argonza, R., Tsui, L.-C., Worton, R. (1988). Localisation of the gene for X-linked recessive type of retinitis pigmentosa (XLRP) to Xp21 by linkage analysis. *Am. J. Hum. Genet.* **43**: 484-494.

Musarella, M. A., Anson-Cartwright, L., Burghes, A., Worton, R. G., Lesko, J. G., Nussbaum, R. L. (1989). Linkage analysis of a large Latin-American family with X-linked retinitis pigmentosa and metallic sheen in the heterozygote carrier. *Genomics* **4**: 601-605.

Musarella, M. A., Weleber, R. G., Murphey, W. H., et al. (1989). Assignment of the gene for complete X-linked congenital stationary night blindness (CSNB1) to Xp11.3. *Genomics* **5**: 727-737.

Musarella, M. A., Anson-Cartwright, L., Leal, S. M., Gilbert, L. D., Worton, R. G., Fishman, G. A., Ott, J. (1990). Multipoint linkage analysis and heterogeneity testing in 20 X-linked retinitis pigmentosa families. *Genomics* **8**: 286-296.

Musarella, M. A., Anson-Cartwright, L., McDowell, C., Burghes, A., Coulson, S. E., Worton, R. G., and Rommens, J. M. (1991). Physical mapping at a potential X-linked retinitis pigmentosa locus (RP3) by pulsed-field gel electrophoresis. *Genomics* **11**: 263-272.

Nagao, S., Yamazaki, A. and Bitenski, K. M. (1987). Calmodulin and calmodulin binding proteins in amphibian rod outer segments. *Biochemistry* **26**: 1659-1665.

Nakamura, Y., Leppert, M., O' Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**: 1616-1622.

Nathans, J., Thomas, D., and Hogness, D. S. (1986). Molecular genetics of human color vision: The genes encoding blue, green and red pigments. *Science* **232**: 193-202.

Nathans, J., Davenport, C. M., Maumenee, I. H., Lewis, R. A., Hejtmancik, Litt, M., et al. (1989). Molecular genetics of human blue monochromacy. *Science* **245**: 831-838.

Nederlof, P. M., Robinson, D., Abuknesha, R., Wiegant, J., Hopman, A. H. N., Tanke, H. J., and Raap, A. K. (1989). Three-colour fluorescence in situ hybridisation for the simultaneous detection of multiple nucleic acid sequences. *Cytometry* **10**: 20-27.

Neel, J. V. (1949). The inheritance of sickle cell anemia. *Science* **110**: 64-66.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirez-Solis, R., Webster, T. D., 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. USA* **86**: 6686-6690.

Nelson, D. L., Ballabio, A., Victoria, M. F., Pieretti, M., Bies, D., Gibbs, R. A., Maley, J. A., Chinault, A. C., Webster, T. D., and Caskey, C. T. (1991). Alu-primed PCR for regional assignment of 110 YAC clones from the human X chromosome: Identification of clones associated with a disease locus. *Proc. Natl Acad. Sci. USA* **88**: 6157-6161.

Ngo, J. T., Spence, A., Cortessi, S. V., Sparkes, R. S., and Bateman, B. (1988). Recombinational event between Norrie disaease and DXS7 loci. *Clin. Genet.* **34**: 43-47.

NIH/CEPH Collaborative Mapping Group. (1992). A comprehensive genetic linkage map of the human genome. *Science* **258**: 67-86.

Nikoskelainen, E. K., Savontaus, M.-L., Wanne, O. P., Katila, M. J., and Nummelin, K. U. (1987). Leber's hereditary optic neuroretinopathy, a maternally inherited disease: a genealogic study in four pedigrees. *Arch. Ophthalmol.* **105**: 665-671.

Nizetic, D., Zehetner, G., Monaco, A. P., Gellen, L., Young, B. D. and Lehrach, H. (1991). Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries. *Proc. Natl Acad. Sci. USA* **88**: 3233-3237.

Nocughi, M., Yi, H., Rosenblatt, H. H., Filipovich, A., Adelstein, S., Modi, W. S., McBride, O., and Leonard, W. (1993). Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. *Cell* **73**: 147-157.

Nolan, K. F., Schwaeble, W., Kaluz, S., Dirich, M. P., and Reid, K. B. M. (1991). Molecular cloning of the cDNA coding for properdin, a positive regulator of the halternative pathway of human complement. *Eur. J. Immunol.* **21**: 771-776.

Nussbaum, R. L., Lesko, J. G., Lewis, R. A., Ledbetter, S. A., Ledbetter, D. H. (1987). Isolation of anonymous DNA sequences from within a submicroscopic deletion in a patient with choroideremia, deafness, and mental retardation. *Proc. Natl. Acad. Sci.* **84**: 6521-6525.

Oberle, I., Rousseau, F., Heitz, D., et al. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* **252**: 1097-1102.

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

Olson, M. V., Hood, L., Cantor, C, and Botstein, D. (1990). A common language for physical mapping of the human genome. *Science* **245**: 1434-1435.

Osborne, N. N. (1981). Noradrenaline, a transmitter candidate in the retina. *J. Neurochem.* **36**: 43-47.

Ott, J., Bhattacharya, S. S., Denton, M., Donald, J., Dubay, C., felix, J., Fishman, G., Frey, D., Gal, A., Humphries, P., Jay, M., Jay, B., Litt, M., Musarella, M., Neugebauer, M., Nussbaum, R. L., Terwilliger, J. D., Weleber, R. G., Wirth, B., Wong, F., Worton, R. G., Wright, A. F. (1990). Localizing multiple X chromosome-linked retinitis pigmentosa loci using multilocus homogeneity tests. *Proc. Natl. Acad. Sci. USA* **87**: 701-704.

Pandolfo, M. (1992). A rapid method to isolate $(GT)_n$ repeats from yeast artificial chromosomes. *Nucleic Acids Research* **20**: 1154

Pardue, M. L. (1985). In situ hybridisation, *In* Nucleic acid hybridisation: a practical approach, (ed. B. D. Hanes and S.J. Higgins), IRL Press, Oxford.

Parimoo, S., Patanjali, S., shukla, H., Chaplin, D., and Weissman, S. (1991). cDNA selection: an efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc. Natl. Acad. Sci. USA* **88**: 9623-9627.

Pauling, L., Itano, H. A., Singer, S. J., and Wells, I. C. (1949). Sickle cell anemia, a molecular disease. *Science* **110**: 543-548.

Pavan, W. J., Heiter, P., and Reeves, R. H. (1990). Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. *Proc. Natl. Acad. Sci. USA* **87**: 1300-1304.

Petty, E. M., Carstens, R. and Bale, A. E. (1991). Ornithine transcarbamylase polymorphism detected by PCR introduction of DraI site. *Nucleic Acids Research* **19**: 690.

Pangburn, M. K., and Muller- Eberhard, H. J. (1984). The alternative pathway of complement. *Springer Semin. Immunopathol.* **7**: 163-192.

Peacocke, M. and Siminovitch, K. A (1987). Linkage of the Wiscott-Aldrich syndrome with polymorphic DNA sequences from the human X chromosome. *Proc. Natl. Acad. Sci. USA* **84**: 3430-3433.

Pearce, W. G. and Bech-Hansen, N. T. (1990). Gene locus for X-linked CSNB. *Genomics* 8: 743-748.

Pearce, W. G., Reedyk, M., and Coupland, S. G. (1990). Variable expressivity in X-linked congenital stationary blindness. *Can. J. Ophthalmol.* **25**: 3-10.

Perry, G. S., Spector, B. D., Schuman, L. M., Mandel, J. S., Anderson, V. E., Mchugh, R. B., et al. (1980). The Wiscott-Aldrich syndrome in the United States and Canada. *J. Pediatr.* **97**: 72-78.

Pieretti, M., Toniorenzi, R., and Ballabio, A. (1991). Rapid assembly of lambda phage contigs within YAC clones. *Nucleic Acids Res.* **19**: 2795-2796.

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. W. (1988). Fluorescence *in situ* hybridisation with human chromosome specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA* **85**: 9138-9142.

Pittler, S. J., and Baehr, W. (1991). Identification of a nonsense mutation in the rod photoreceptor cGMP phosphodiesterase beta subunit gene of the rd mouse. Proc. *Natl. Acad. Sci. U.S.A.* **88**: 8322-8326.

Ragoussis, J., Monaco, A., Mockridge, I., Kendall, E., Campbell, R. D., and Trowdale, J. (1991). Cloning of the HLA class II region in yeast artificial chromosomes. *Proc. Natl Acad. Sci. USA* **88**: 3753-3757.

Rao R. V., Cohen, G. B., and Oprian D. D. (1994). Rhodopsin mutation G90D and a molecular mechanism for congenital night blindness. *Nature* **367**: 639-641.

Ray, P. N., Belfall, B., Duff, C., Logan, C., Kean, V., et al. (1985). Cloning of the breakpoint of an X ; 21 translocation associated with Duchenne muscular dystrophy. *Nature* **318**: 672-675.

Reichel. E., Bruce, A. M., Sandberg, M. A., Berson, E. L. (1989). An electroretinographic and molecular study of X-linked cone degeneration. *Am. J. Ophthalmol.* **108**: 540-547.

Remold-O'Donnell, E. and Rosen, F. S. (1990). Sialophorin (CD43) and the Wiscott-Aldrich syndrome. *J. Exp. Med.* **159**: 1705-1723.

Report of the Fourth International Workshop on Human X Chromosome Mapping. (1993). *Cytogenetics Cell Genetics* **64, 3-4**: 148-194.

Richards, R. I., Holman, K., Kozman, H., Kremer, E., Lynch, M., Pritchard, et al. (1991). Fragile X syndrome: Genetic localisation by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. *J. Med. Genet.* **28**: 818-823.

Richards, R. I. and Sutherland, G. (1992). Heritable unstable DNA sequences. *Nat. Genet.* **1**: 7-9.

Riggs, J. L., Siewald, R. J., Burckhalter, J. H., Downes, C. M. and Metcalf, T. G. (1958). Isothiocyanate compounds as fluorescent labelling agents for immune serum. *Am. J. Pathology* **34**: 1081-1098.

Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C. and Markham, A. F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *NAR* **18**: 2887-2890.

Riordan J. R., Rommens, J. M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczac, Z., Zielenski, J., et al. (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* **245**: 1066-1073.

Rosenfeld, P. J., Cowley, G. S., McGee, T. L., Sandberg, M. A., Berson, E. L., and Dryja, T. P. (1992). A *Null* mutation in the rhodopsin gene causes rod photoreceptor dysfunction and autosomal recessive retinitis pigmentosa. (1992). *Nature Genet.* **1**: 209-213.

Royer-Pocora, B., Kunkel, L. M., Monaco, A. P., et al. (1986). Cloning the gene for the inherited disorder chronic granulomatous disease on the basis of its chromosomal location. *Cold Spring Harbor Symp. Quant. Biol.* **51**: 177-183.

Royle, N. J., Clarkson, R. E., Wong, Z. and Jeffreys, A. J. (1988). Clustering of hypervariable minisatellites in the proterminal region of human autosomes. *Genomics* **3**: 352-360.

Rudkin, G. T. and Stollar, B. D. (1977). High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence. *Nature* **265**: 472-473.

Saiki, P. K., Schorf, S., Faloona, F., Mullis, K. B., Horn, G. T., Ehrlich, H. A., Arnheim, N. (1985). Enzymatic amplification of beta globin sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

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

Sargent, C., Dunham, I. and Campbell, R. D. (1989). Identification of multiple HTF-island associated genes in the human major histocompatibility complex class III region. *EMBO J.* **8**: 2305-2312.

Sargent, C. A., Affara, N., Bentley, E., Pelmear, A., Bailey, D. M. D., Davey, P., Dow, D., Leversha, M., Aplin, H., Besley, G. T. N., and Ferguson-Smith, M. A. (1993). Cloning of the X-linked glycerol kinase deficiency gene and its identification by sequence comparison to the *Bacillus subtilis* homologue. *Hum. Molec. Genet.* **2**: 97-106.

Saunders, G. F., Hsu, T. C., Getz, M. I., Simes, E. L. and Arrighi, F. E. (1972). Location of human satellite DNA in human chromosomes. *Nature (Lond.)*, New Bilo. **236**: 244-246.

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

Seabra, M. C., Brown, M. S., Slaughter, C. A., Sudhof, T. C., and Goldstein, J. L. (1992). Purification of component A of Rab geranylgeranyl transferase: possible identity with the choroideremia gene product. *Cell* **70**: 1049-1057.

Seabra, M. C., Brown, M. S. and Goldstein, J. L. (1993). Retinal degeneration in choroideremia: deficiency of Rab geranylgeranyl transferase. *Science* **259**: 377-381.

Seedorf, T. (1985). The inheritance of Leber's disease: a genealogical follow-up study. *Acta Ophthalmol.* **63**: 135-145.

Sefiani, A., Mrad, R., Simard, L., Vincent, A., Julier, C. et al. (1991). Linkage relationships between incontinentia pigmenti (IP2) and 9 terminal X long arm markers. *Hum. Genet.* **86**: 297-

Shero, J. H., McCormick, M. K., Antonarakis, S. E., and Hieter, P. (1991). Yeast artificial chromosome vectors for efficient clone manipulation and mapping. *Genomics* **10**: 505-508.

Silver, J. and Keerikatte, V. (1989). Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an intergrated provirus. *J. Virol.* **63**: 1924-1928.

Silverman, G. A., Jockel, J. I., Domer, P. H., Mohr, R. M., Taillon-Miller, P. and Korsmeyer, S. J. (1991). Yeast artificial chromosome cloning of a two-Megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type-2. *Genomics* **9**: 219-228.

Sims, K. B., de la Chapelle, A., Norio, R., Sankila, E. -M., Hsu, Y. -P. P., Rinehart, W. B., Corey, T. J., Ozelius, L., Powell, J. F., Bruns, G., Gusella, J. F., Murphy, D. L., and Breakfield, X. O. (1989a). Monoamine oxidase defficiency in males with an X chromosome deletion. *Neuron* **2**: 1069-1076.

Sims, K. B., Ozelius, L., Corey, T. J., Rinehart, W. B., Liberfarb, R., Haines, J., Chen, W. J., Norio, R., Sankila, E. -M., de la Chapelle, A., , Murphy, D. L., Gusella, J. F., and Breakfield, X. O. (1989b). The Norrie disease gene is distinct from the monoamine oxidase genes. *Am. J. Hum. Genet.* **45**: 424-434.

Sims, K. B., Lebo, R. V., Benson, G., Shalish, C., Schuback, D., Chen, Z. Y., Bruns, G., Craig, I. W., Golbus, M. S., and Breakfield, X. O. (1992). The Norrie disease gene maps to a 150kb region on chromosome Xp11.3. *Human Molecular Genetics* **1**: 83-89.

Singh, G., Lott, M. T., and Wallace, D. C. (1989). A mitochondrial mutation as a cause of Leber's hereditary optic neuropathy. *N. Engl. J. Med.* **320**: 1300-1305.

Smeets, H. J. M., Brunner, H. G., Ropers, H-H. and Wieringa, B. (1989). Use of variable simple sequence motifs as genetic markers: application to study of myotonic dystrophy. *Hum. Genet.* **83**: 245-251.

Smith, H. O., and Birnsteil, M. L. (1976). A simple method for DNA restriction site mapping. *Nucleic Acids Res.* **3**: 2387-2398.

Smith, T. W., Butler, V., and Haber, E. (1970). Characterisation of antibodies of high affinity and specificity for the digitalis glycoside digoxin. *Biochemistry* **9**: 331-7.

Smith, D. R., Smyth, A. P., and Moir, D. T. (1990). Amplification of large PA artificial chromosomes. *Proc. Natl Acad. Sci. USA* **87**:8242-8246.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503-517.

Stallings, R. L., Ford, A. f., Nelson, D., Torney, D. C., Hildebrand, C. E., and Moyzis, R. K. (1991). Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics* **10**: 807-815.

Steffensen, D. M., Duffey, P. and Prensky, W. (1974). Location of 5S ribosomal RNA genes on human chromosome 1. *Nature (Lond.)* **252**: 741-743.

Stryer, L. (1991). Visual excitation and recovery. *The journal of biological chemistry* **266**: 10711-10714.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **17**: 6463-6471.

Teague, P. W., Aldred, M. A., Jay, M., Dempster, M., Harrison, C., Carothers, A. D., Strain, L., Brock, D. J. H., Bundey, S., Jay, B., Bird, A. C., Bhattacharya, S. S., and Wright, A. F. (1994). Heterogeneity analysis in 40 X-linked retinitis pigmentosa families. *Am J. Hum. Genet.* (in press).

Thiselton, D., Kamakari, S., Curtis, A. R. J., Coleman, M. P., Davies, K. E., Linsday, S., and Bhattacharya, S. S. (1993). Dinucleotide repeat polymorphism at the DXS556 locus, *Human Molecular Genetics* **2**: 613.

Thiselton, D. L., Lindsay, S. Kamakari, S., Hardcastle, A., Coleman, M., Davies, K. E., and Bhattacharya, S. S. (1994). Genetic and physical mapping of five novel microsatellite markers on human Xp21.1-Xp11.22 (*submitted*).

Thorpe, L. W., Westlund, K. N., Kochersperger, L. M., Abell, C. W., and Denney, R. M. (1987). Immunocytochemical localisation of monoamine oxidase A and B in human peripheral tissues and brain. *J. Histochem. Cytochem.* 35: 23-32.

Todd, J. A., et al., 1991, in PCR: A Practical Approach (McPherson, M.C., Quirke, P. and Taylor, G., eds), pp101-105, Oxford University Press.

Trask, B. J., Pinkel, D., and van den Engh, G. (1989). The proximity of DNA sequences in interphase cell nuclei is correlated to genomic distance and permits ordering of cosmids spanning 250 kilobase pairs. *Genomics* **5**: 710-717.

Trask, B. J., Massa, H., Kenwrick, S., and Gitschier, J. (1991). Mapping of human chromosome Xq28 by two-color fluorescence in situ hybridisation of DNA sequences to interphase cell nuclei. *Am. J. Hum. Genet.* **48**: 1-15.

Travis, G. H., Christerson, L., Danielson, P. E., Klisak, I., Sparkes, R. S., Hahn, L. B., Dryja, T. P., and Sutcliffe, J. G. (1991). The human retinal degeneration slow (RDS) gene: Chromosome assighment and structure of the mRNA. *Genomics* **10**: 733-739.

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. *Nucleic Acids Res.* **16**: 8186.

Tsukada, S., Saffran, D. C., Rawlings, D. J., Parolini, O., Allen, R. c., Klisak, I., Sparkes, R. s., Kubagawa, h., Mohandas, T., Quan, S., Belmont, J. W., Cooper, M. D., Conley, M. E., and Witte, O. N. (1992). Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell* **72**: 279-290.

Tuddenham, E. G. D., Cooper, D. N., Gitschier, J., Hicughi, M., Hoyer, L. W., Yoshioka, A., Peake, I. R. et al. (1991). Hemophilia A: Database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene. *Nucleic Acids Research* **19**: 4821-4833.

Tumer, Z., Tommerup, N., Tonnesen, T., Kreuder, J., Craig, I. W., and Horn, N. Mapping of the Menkes locus to Xq13.3 distal to the X inactivation center by an intrachromosomal insertion of the segment Xq13.3-q21.2. (1992). *Hum. Genet.* **88**: 668

Valdes, R., Brown, B. A., and Graves, S. W. (1984). Variable cross-reactivity of digoxin metabolites in digoxin immunoassays. *Am. J. Clin. Pathol.* **82**: 210-213.

Valle, D. and Simell, O. (1989). The hyperornithinemias. *In* "The Metabolic Basis of Inherited Disease" (C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, Eds.) 6th ed., pp 599-627, McGraw-Hill, New York.

Van Dorp, D. B., Eriksson, A. W., Dellman, J. W. et al. (1985). Aland eye disease: no albino misrouting. *Clin. Genet.* **28**: 526-531.

Van Ommen, G. J. B., Verkerk, J. M. H., Hofker, M. H., Monaco, A. P., Kunkel, L. M., Ray, P., Worton, R., Wieringa, b., Bakker, E., and Pearson, P. L. (1986). A physical map of 4 million base pairs around the Duchenne muscular dystrophy gene on the human X-chromosome. *Cell* **47**: 499-504.

Verellen-Dumoulin, C. H., Freund, M., De Meyer, R., Laterre, C. H., Frederic, J., et al. (1984). Expression of an X-linked muscular dystrophy in a female due to

translocation involving Xp21 and nonrandom inactivation of the normal X chromosome. *Hum. Genet.* **67**: 115-119.

Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., et al. (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.

Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C., and Bentley, D. R. (1993). The gene involved in X-linked agammaglobulinemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**: 226-233.

Vilkki, J., Savontaus, M.-L., and Nikoskelainen, E. K. (1989). Genetic heterogeneity in Leber hereditary optic neuroretinopathy revealed by mitochondrial DNA polymorphism. *Am. J. Hum. Genet.* **47**: 95-100.

Vilkki, J., Ott, J., Savontaus, M.-L., Aula, P., and Nikoskelainen, E. K. (1991). Optic atrophy in Leber hereditary optic neuroretinopathy is probably determined by an X-chromosomal gene closely linked to DXS7. *Am. J. Hum. Genet.* **48**: 486-491.

Vulliamy, T., Mason, P., Luzzatto, L. (1992). The molecular basis of glucose-6-phosphate dehydrogenase deficiency. *Trends Genet.* **8**: 138.

Vulpe, C., Levinson, B., Whitney, S., Packman, S., Gitschier, J. (1993). Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nature Genet.* **3**: 7-13.

Vuong, T. M., Chabre, M., and Stryer, L. (1984). Millisecond activation of transducin in the cyclic-nucleotide cascade of vision. *Nature* **311**: 659-661.

Wada, M., Little, R. D., Abidi, F., Porta, G., Labella, T., Cooper, T., Della Valle, G., D' Urso, M., and Schlessinger, D. (1990). Human Xq24-Xq28: approaches to mapping with yeast artificial chromosomes. *Am. J. Hum. Genet.* **46**: 95-106.

Wahl, G. M., Stern, M., and Stark, G. R. (1979). Efficient transfer of large DNA from agarose gels to diazobenzylxymethyl-paper and rapid hybridisation by using dextran sulphate. *Proc. Natl. Acad. Sci. (Wash.)* **76**: 3683-3687.

Walker, A. P., Muscatelli, F., and Monaco, A. (1993). Isolation of the human Xp21 glycerol kinase gene by positional cloning. *Hum. Molec. Genet.* **2**: 107-114.

Wallace, D. C., Singh, G., Lott, M. T., Hodge, J. A., Schurr, T. G., Lezza, A. M. S., et al. (1988). Mitochondrial DNA mutation associated with Leber's hereditary optic neuroretinopathy. *Science* **242**: 1427-1430.

Warburg, M. (1963). Norrie's disease. *Acta Ophth.* **41**: 134-146.

Warburg, M. (1966). Norrie's disease: a congenital progressive oculo-acoustico-cerebral degeneration. *Acta Ophth. Suppl.* **89**: 1-47.

Warburg, M. (1979). Retinal malformations. *Trans. Ophthal. Soc.* U. K. **99**: 272-283.

Watson, J. D. and Crick, F. H. C. (1953). Genetical implications of the structure of deoxyribosenucleic acid. *Nature* **171**: 964-967.

Watson, J. M., Spencer, J. A., Riggs, A. D., and Marshall Graves, J. A. (1991). Sex chromosome evolution: Platypus gene mapping suggests that part of the human X chromosome was originally autosomal. *Proc. Natl. Acad. Sci. USA* **88**: 11256

Weber J. L. and May P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 338-396.

Weber, J. L. (1990). Informativeness of human $(dC\text{-}dA)_n$. $(dG\text{-}dT)_n$ polymorphisms. *Genomics* **7**: 524-530.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P. Vaysseix, and Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature* **359**: 794-801.

Weleber, R. G., Pillers, D. M., Powell, B. R., Hanna, C. E., Magenis, R. E., Buist, N. R. M. (1989). Aland Island eye disease (Forsius-Ericksson syndrome) associated with contiguous deletion syndrome at Xp21. *Arch. Ophthalmol.* **107**: 1170-1179.

Wells, M., Griffiths, S., Lewis, F. A., and Bird, C. C. (1987). Demonstration of human papillomavirus types in paraffin processed tissue from human ano-genital lesions by in situ hybridisation. *Journal of Pathology* **152**: 77-82.

Wensel, T. G., and Stryer, L. (1990). Activation mechanism of retinal rod cyclic-GMP phosphodiesterase probed by fluorescein-labelled inhibitory subunit. *Biochemistry* **29**: 2155-2161.

White, R., Lippert, M., Bishop, D. T., et al. (1985). Construction of linkage maps with DNA markers for human chromosomes. *Nature* **313**: 101-105.

Wieacker P., Davies K. E., Cooke, H. J., Pearson, P. L., Williamson, R., Bhattacharya, S., Zimmer, J., Ropers, H. H. (1984). Toward a complete linkage map of the human X chromosome: regional assignment of 16 cloned single-copy DNA sequences employing a panel of somatic cell hybrids. *Am J Hum. Genet.* **36**: 265-276.

Wilden, U., Hall, S. W., and Kunn, H. (1986). Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of the rod outer segments. *Proc. Natl. Acad. Sci. USA* **83**: 1174.

Willard, H. F., Waye, J. S., Scolnick, M. H., Schwartz, C. E., Powers, V. E. and England, S. B. (1986). Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes-specific a satellite DNA probes: Implications for development of centromere based genetic linkage maps. *Proc. Natl. Acad. Sci. USA* **83**: 5611-5615.

Williamson, R., Bowcock, K., Kidd, K., Pearson, P., Schmidtk, P., Cereva, P., Chipperfield, M. et al. (1991). report of the DNA committe and catalogues of cloned and mapped genes, markers formatted for PCR and DNA polymorphisms. Human gene mapping 11, *Cytogenet. Cell Genet.* **58**: 1190-1832.

Worton, R. G., Duff, C., Sylvester, J. E., Schmickel, R. D., and Willard, H. F. (1984). Duchenne muscular dystrophy involving translocation of the DMD gene next to ribosomal RNA genes. *Science* **224**: 1447-1449.

Worton, R. G. and Thomson, M. W. (1988). Genetics of Duchenne muscular dystrophy. *Annu. Rev. Genet.* **22**: 601-629.

Wright, A. F., Bhattacharya, S. S., Clayton, J. F., Dempster, M., Tippet, P., McKeown, C. M. E., Jay, M., Jay, B., and Bird, A. C. (1987). Linkage relationships between X-linked retinitis pigmentosa and nine short-arm markers and localisation to between DXS7 and DXS14. *Am. J. Hum. Genet.* **41**: 635-644.

Wright, A. F. (1990). Towards the identification of genes in X-linked retinitis pigmentosa. *Progr. Retina Res.* **9**: 197-227.

Yang-Feng, T. L., DeGennaro, L. J., and Franche, U. (1986). Genes for synapsin I, a neuronal phosphoprotein, map to conserved regions of human and murine X chromosomes. *Proc. Natl. Acad. Sci. USA* **83**: 8679-8683.

Yau, K. W., and Nakatani, K. (1985). Light-induced reduction of cytoplasmic free calcium in retinal rod outer segment. *Nature* **313**: 579-582.

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.

Yen, P. H., Ellison, J., Salido, E. C., Mohandas, T. M., and Shapiro, L. (1992). Isolation of a new gene from the distal short arm of the human X chromosome that escapes X-inactivation. *Hum. Molec. Genet.* **1**: 47-52.

Yuille, M.A.R., Goudie, D. R., Affara, N.A. and Ferguson-Smith, M. A. (1991). Rapid determination of sequences flanking microsatellites. *Nucleic Acids Res.* **19**: 1950.

Zhu, D., Antonarakis, S. E., Schmeckpeper, B. J., Diergaarde, P. J., Greb, A. E., and Maumenee, I. H. (1989). Microdeletion in the X chromosome and prenatal diagnosis in a family with Norrie disease. *Am. J. Med. Genet.* **33**: 485-488.

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.

PUBLICATIONS

Two dinucleotide repeat polymorphisms at the DXS571 locus

A.R.J.Curtis, P.Roustan+, S.Kamakari, D.Thiselton, S.Lindsay and S.S.Bhattacharya[1,*]

Molecular Genetics Unit, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU and [1]Department of Molecular Genetics, Institute of Opthalmology, Bath Street, London EC1V 9EL, UK

Source and Description of Clone: Cosmid HX79 was isolated from a library containing insert DNA from a mouse–human hybrid cell line which has the X chromosome as the only detectable human chromosome (1). It was then selected by hybridisation to a poly (dCA.dGT) probe (Pharmacia). A 1.7 kb HindIII fragment of this cosmid which contains two microsatellites, A) $(CA)_{15}(GA)_{10}$ and B) $(CA)_{16}$, was subcloned as pHX79. pHX79 was sequenced and the sequences flanking the repeats (EMBL accession no. X60696) were used to design PCR primers.

PCR Primers:

A) pHX79/A1 = 5'-TCTGCTGCTTTAGATGCTTTC-3' (CA Strand)
pHX79/A2 = 5'-GACTAGGTCACATCATCAAATC-3' (GT Strand)

B) pHX79/B1 = 5'-AATATTGGTGCAGGCACTGT-3' (CA Strand)
pHX79/B2 = 5'-AATCAGATGCAGTGATGGGT-3' (GT Strand)

Polymorphism: Allelic fragments were detected on DNA sequencing gels (repeat A) or on non-denaturing 7.5% polyacrylamide gels stained with ethidium bromide (repeat B). Lengths (nt) are: A1 = 133, A2 = 131, A3 = 129 for repeat A and B1 = 148, B2 = 146, B3 = 144, B4 = 136, B5 = 134, B6 = 130 for repeat B. Subclone pHX79 contains A2 and B5.

Frequencies: A) Estimated from 27 (54 chromosomes) unrelated Caucasian females: A1 = 0.06, A2 = 0.70, A3 = 0.24. Observed heterozygosity = 0.44. B) Estimated from 38 (76 chromosomes) unrelated Caucasian females: B1 = 0.01, B2 = 0.08, B3 = 0.19, B4 = 0.05, B5 = 0.62, B6 = 0.05, Observed heterozygosity = 0.46.

Chromosomal Localization: DXS571 has been assigned to Xq21.1–q21.3 by hybridisation of the cosmid HX79 to a panel of somatic cell hybrid DNAs and confirmed by fluorescence *in situ* hybridisation to metaphase chromosomes.

Mendelian Inheritance: X-linked co-dominant segregation was observed in 3 informative families.

PCR Conditions: PCR amplifications for repeat A were carried out for 24 cycles using ^{32}P end-labelled primer pHX79/A as follows: denaturation at 94°C for 1 min, annealing and extension at 56°C for 2 min with a first denaturation step at 94°C for 3 min and a final elongation step at 56°C for 10 min.

The fragment containing repeat B was amplified for 32 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 2 min and extension at 72°C for 2 min with a first denaturation step at 94°C for 3 min and a final elongation step at 72°C for 10 min.

Acknowledgements: This work was supported by grants from Fédération des Aveugles de France, Conseil Régional Midi-Pyrénées, Fondation de France, Association Française de lutte contre les Myopathies, the Sir Jules Thorn Trust, the British Retinitis Pigmentosa Society and the Wellcome Trust.

Reference: 1) Lindsay,S. and Bird,A.P. (1987) *Nature* 327, 336–338.

* To whom correspondence should be addressed

+Present address: Laboratoire d'Endocrinologie Expérimentale, Bât. L3, CHU Rangueil, 31054

Dinucleotide repeat polymorphism at the DXS559 locus

P.Roustan+, A.R.J.Curtis, S.Kamakari[1], D.Thiselton[1], S.Lindsay and S.S.Bhattacharya[1,*]
Molecular Genetics Unit, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU and [1]Department of Molecular Genetics, Institute of Opthalmology, Bath Street, London EC1V 9EL, UK

Source and Description of Clone: Cosmid HX28 was isolated from a library containing insert DNA from a mouse–human hybrid cell line which has the X chromosome as the only detectable human chromosome (1). It was then selected by hybridisation to a poly (dCA.dGT) probe (Pharmacia). pHX28 is a Sau3AI subclone of this cosmid which contains a $(CA)_{24}$ microsatellite. pHX28 was sequenced and the sequences flanking the repeat (EMBL accession no. X66054) were used to design PCR primers.

PCR Primers:
pHX28/A = 5'-ATCCTTCACCACTGCCTCCA-3' (CA strand)
pHX28/B = 5'-CTCCCTGCTCCCATCGCCAA-3' (GT strand)

Polymorphism: Allelic fragments were detected on DNA sequencing gels. Lengths (nt) are: A1 = 248, A2 = 246, A3 = 244, A4 = 242, A5 = 240, A6 = 238, A7 = 236, A8 = 230. Subclone pHX28 contains A4.

Frequencies: Estimated from 30 (60 chromosomes) unrelated Caucasian females: A1 = 0.02, A2 = 0.05, A3 = 0.02, A4 = 0.18, A5 = 0.38, A6 = 0.28, A7 = 0.05, A8 = 0.02. Observed heterozygosity = 0.63.

Chromosomal Localization: DXS559 has been assigned to Xq12 by hybridisation of the cosmid HX28 to a panel of somatic cell hybrids DNAs and confirmed by fluorescence *in situ* hybridisation.

Mendelian Inheritance: X-linked co-dominant segregation was observed in 3 informative families.

PCR Conditions: PCR amplifications were carried out for 24 cycles using ^{32}P end-labelled primer pHX28/A as follows: denaturation at 94°C for 1 min, annealing and extension at 56°C for 2 min with a first denaturation step at 94°C for 3 min and a final elongation step at 56°C for 10 min.

Acknowledgements: This work was supported by grants from Fédération des Aveugles de France, Conseil Régional Midi-Pyrénées, Fondation de France, Association Française de lutte contre les Myopathies, the Sir Jules Thorn Trust, the British Retinitis Pigmentosa Society and the Wellcome Trust.

Reference: 1) Lindsay,S. and Bird,A.P. (1987) *Nature* 327, 336–338.

* To whom correspondence should be addressed

+Present address: Laboratoire d'Endocrinologie Expérimentale, Bât. L3, CHU Rangueil, 31054 Toulouse, France

An *Rsa*I polymorphism in the human serotonin receptor gene (HTR1A): detection by DGGE and RFLP analysis

James T.Warren Jr, Michael L.Peacock and John K.Fink*
The University of Michigan Department of Neurology, 1103 East Huron, Ann Arbor, MI 48104-1687, USA

Description/Polymorphism: A genomic DNA fragment corresponding to nucleotides 192 to 557 of the serotonin (HTR1A) receptor gene (1) was amplified by PCR and analyzed by DGGE on a 50–80% gradient (100% denaturant = 7M urea, 40% formamide) at 12 mA for 20 hours. Direct sequencing of this polymorphic region from a homozygous mutant revealed an adenine at nucleotide #294, instead of the published guanine at this position. Both allele C1(GTG) and allele C2(GTA) encode for a valine at amino acid #98. Digestion of the 396 bp PCR product (GenBank accession no. M28269) (including the 30 bp GC clamp) with *Rsa*I yields 211 and 184 bp products for allele C1 and 103, 108 and 184 bp products for allele C2.

Primer Sequences:
G21 5.4
5'CGCTCCCTGCAGAACGTGGC 3'
G21 3.4GC
5'GCGGGCGGGGCGGGGGCACGGGGGCCGCCCATGCGTCGGGCGAG-ATGAGCGCACG 3'

Frequency: Evaluated in 23 unrelated African Americans. Allele C2 was not present in 78 unrelated Caucasians tested.
C1: 0.78
C2: 0.22
Observed heterozygosity = 0.34.

Chromosomal Localization: The human HTR1A receptor gene has been mapped to 5q11.2–q13.

Mendelian Inheritance: Codominant segregation was observed in 5 families.

PCR Conditions: The PCR reaction was performed using 1 μg genomic DNA, 2.2 μM of each primer, and 1.5 mM $MgCl_2$ in a final volume of 50 μl. Samples were subjected to 3 cycles of 3 min at 94°C, 45 sec at 60°C, 1.5 min at 72°C; then 35 cycles of 1.0 min at 94°C, 45 sec at 60°C and 1.5 min at 72°C.

Acknowledgements: This research is supported by NIH-NINDS:NS07222 (to J.T.W.) and grants from the Lucille P.Markey Charitable Trusts and the National Alliance for Research in Schizophrenia and Depression (to J.K.F.). We thank Dr Leonard Lerman for the use of his MELT87 and SQHTX programs.

Reference: 1) Kobilka,B.K. *et al.* (1987) *Nature* 329, 75–79

* To whom correspondence should be addressed

GENOMICS 14, 298–301 (1992)

The Gene for Aarskog Syndrome Is Located between DXS255 and DXS566 (Xp11.2–Xq13)

M. E. M. PORTEOUS,* A. CURTIS,* S. LINDSAY,* O. WILLIAMS,*
D. GOUDIE,† S. KAMAKARI,* AND S. S. BHATTACHARYA*

*Department of Human Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4AA; and †Department of Genetics, University of Cambridge, Cambridge, United Kingdom

Received December 23, 1991; revised June 19, 1992

Aarskog syndrome has been mapped to Xq13 on the basis of a patient carrying an Xq13:8p21.2 translocation. We have identified a new microsatellite marker in a clone mapping to this region (HX60;DXS566). Using primers flanking this microsatellite along with primers detecting a microsatellite at PGK1P1 and DXS255, and DXS72, we have performed a multipoint analysis in a large kindred with Aarskog syndrome. Our results suggest that the Aarskog locus lies proximal to Xq13. This is supported by the recent redefining of the breakpoint of the original translocation as between DXS14 (Xp11.21–p11.1) and DXS146 (Xp11.23–p11.22).

INTRODUCTION

Aarskog syndrome (facio-digital-genital syndrome) is characterized by a typical facial appearance, short stature, hand anomalies, and an overriding or "shawl scrotum" in males. It was first described in 1970 by Dagfin Aarskog (1970) and since then over 100 cases have been described (Porteous and Goudie, 1991).

The pedigree structure of most reported cases has suggested X-linked inheritance, and in 1984, Bawle *et al.* (1984) mapped Aarskog syndrome to Xq13 on the basis of a familial translocation (46,X,t(X;8)(q13:p21.2) occurring in a mother and son, both of whom exhibited characteristic features of Aarskog syndrome. However, male to male transmission has been documented in two large pedigrees with affected males showing typical features of the disorder (Grier *et al.*, 1983; van de Vooren *et al.*, 1983). This suggests that Aarskog syndrome may be a heterogeneous disorder with both X-linked and autosomal inheritance patterns.

Mapping studies using polymorphic markers for the Xq13 region are indicated to confirm that Aarskog syndrome is indeed genetically heterogeneous and not solely a sex-limited autosomal dominant disorder. However, mapping in this area of the genome has been hampered by the lack of highly polymorphic markers. The discovery of microsatellites (Tautz and Renz, 1984) and the use of the polymerase chain reaction (PCR) to analyze them (Litt and Luty, 1989; Weber and May, 1989) has been particularly useful in linkage studies. We have isolated and characterized a poly(CA) region from a cosmid (HX60) that maps to Xq13. Primers that flank this microsatellite and others flanking the microsatellite at the PGK1P1 locus (Browne *et al.*, 1991), along with polymorphisms at the DXS255 (Fraser *et al.*, 1987) and DXS72 loci, have been used on conventional Southern blots in linkage studies in a large Aarskog pedigree.

MATERIALS AND METHODS

Family studies. A large West of Scotland family in which the gene for Aarskog syndrome was segregating was ascertained through the proband IV-1 (Fig. 2), who was referred to the Genetics Department because of short stature and facial anomalies. Twenty-one family members were seen and examined for clinical features of Aarskog syndrome. Individuals were examined independently by two authors (M.P. and D.G.) to determine clinical phenotype; agreement was reached in all cases. Blood samples were obtained, and DNA was extracted according to standard methods (Maniatis *et al.*, 1982).

Isolation and characterization of DXS566. HX60 was isolated from a cosmid library prepared with insert DNA from a man–mouse hybrid cell line (Lindsay and Bird, 1987). The human X chromosome is the only cytologically visible human chromosome in this cell line. HX60 has now been assigned to DXS566 and mapped to Xq13 by nonradioactive *in situ* hybridization to metaphase spreads (data not shown).

Simple repeat sequences were identified within DXS566 by hybridization with a poly(CA) probe (Pharmacia) to *Hin*dIII-digested DNA, which had been subjected to electrophoresis and blotted onto Hybond N (Amersham) according to the manufacturer's instructions. The *Hin*dIII fragment containing the poly(CA) sequences was subcloned into pT7T3 (Pharmacia) and then sequenced using the standard Sequenase kit and protocol (USB). The sequence has been deposited with the EMBL Database under Accession No. X60695. Primers flanking the poly(CA) sequence are

Primer 1:5′ CATTGCCAGAATATATGAAACAT 3′ (CA strand);
Primer 2:5′ ACTTGTCAAATATGGCTCATA 3′ (GT strand).

DXS566 and PGK1P1. PCR reactions were carried out in a 25-μl volume and contained 250 ng of genomic DNA, 1 μ*M* of each primer, 200 μ*M* of each dNTP, 1 unit of *Taq* polymerase (NBL), and the manufacturer's recommended buffer. The reactions were overlaid with 50 μl mineral oil. DXS566 reactions were processed through 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 3 min, with a final elongation cycle at 72°C for 10 min. PGK1P1 reactions were processed through 39 cycles of

0888-7543/92 $5.00

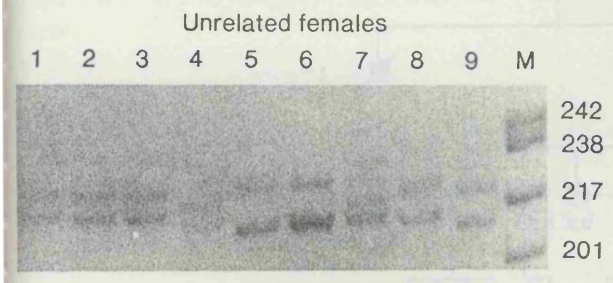


FIG. 1. DXS566 alleles in unrelated caucasian females. PCR products after amplification using the primer pair derived from clone HX60 (DXS566) have been electrophoresed for 18 h at 90 V in a 6% polyacrylamide gel that had been stained with ethidium bromide. Five different alleles can be seen. M, *Msp*I-digested PBR322 as a marker. The sizes of the marker fragments are given in basepairs (bp) on the right.

denaturation at 94°C for 1 min, annealing at 50°C for 2 min, and extension at 72°C for 2 min, with a final elongation at 72°C for 10 min.

DXS566 alleles were visualized using electrophoresis in a nondenaturing 6% polyacrylamide gel stained with ethidium bromide. Allele size was determined by comparison to the PCR product of clone DNA and standard markers and by analysis of labeled products by denaturing polyacrylamide gel electrophoresis and autoradiography. The dinucleotide repeat in the clone is $(AC)_{19}$. The PCR product in the clone is 217 bp. PGK1P1 alleles were visualized under identical conditions, the PCR product of the clone being 213 bp.

DNA analysis (DXS255 and DXS72). Aliquots of 5 μg of DNA were digested to completion with 20 units of *Pst*I (DXS255) or *Hin*dIII (DXS72) at 37°C. Samples were size fractionated on a 0.8% agarose gel and blotted onto Hybond N+ (Amersham) using 0.4 *M* NaOH. The membrane was then prehybridized in 4× SSC, 10× Denhardt's, and 1% SDS. DXS255 and DXS72 were oligolabeled as described by Feinberg and Vogelstein (1983) using a kit (Bio-Rad). Stringent washes were at 65°C in 0.5× SSC and 0.1% SDS. Autoradiographic images were obtained with Fuji RX film.

Linkage analysis. Linkage analysis was carried out by the lod score method devised by Morton (1955), using the computer program LINKMAP (Lathrop *et al.*, 1984) and LIPED (Ott, 1974). A lod score (*Z*) of 3 or more is generally accepted as evidence of linkage between a marker and a disease locus, although a lod of 2 may be acceptable in an X-linked condition. A gene frequency of 0.00001 for the Aarskog gene was assumed on the basis of the estimated prevalence of Aarskog syndrome in the West of Scotland of 8/1,000,000. The disorder was assumed to be fully penetrant.

Recent work shows DXS566 to be tightly linked to PGK1 with a lod score of 10.84 at a recombination fraction of 0 (Lindsay *et al.*, 1992). The consensus map for the loci used was reported at HGM 11 to be Xp11.3–DXS255–PGK1P1–PGK1–Xq13. DXS7 maps telomeric to DXS255 (Davies *et al.*, 1991). The distance between DXS7 and PGK1 is estimated at approximately 30 cM, while the distance between DXS1 and DXS72, which maps telomeric to PGK to Xq22, is estimated at 32 cM (Davies *et al.*, 1991). For the purpose of multipoint analysis, a recombination fraction of 0.16 was assumed between DXS255 and PGK1P1 and 0.1 between PGK1P1 and DXS566.

TABLE 1

DXS566 Alleles of 44 X Chromosomes from Unrelated Individuals

Allele (bp)	Frequency	Number of chromosomes
219	0.068	3
217	0.455	20
215	0.023	1
213	0.227	10
211	0.227	10

RESULTS

Figure 1 shows the PCR products after amplification with the DXS566-specific primers in human females. It can be seen that there are five alleles. A more extensive survey of 44 X chromosomes from unrelated Caucasians revealed a heterozygosity of 59% (data not shown). Table 1 shows the observed alleles and their frequencies. The PIC value is 0.64. X-linked codominant inheritance was observed in two non-Aarskog informative families in addition to the Aarskog family described here (data not shown).

Figure 2 shows the pedigree of a large family with members affected by Aarskog syndrome. The alleles produced by PCR at the DXS566 locus can be seen for each individual. In this family, the smallest allele (d) is segregating with the disease. There is one recombinant, individual III-6, who appears to have passed her grandmaternal allele (a) on to her affected son. The same individual was a nonrecombinant with DXS255 and PGK1P1. Four informative meioses with PGK1P1 showed no recombinants, while individual IV-2 was a recombinant with DXS255.

Table 2 shows the pairwise lod scores between the Aarskog locus and the four X chromosome markers, DXS255 (Xp11.22), PGK1P1 (Xq12–q13), DXS566 (Xq13), and DXS72 (Xq22).

Figure 3 shows the results of the multipoint analysis in graph form. A locus order DXS255–cen–PGK1P1–DXS566 was assumed with $\theta = 0.16$ between DXS255 and PGK1P1 and $\theta = 0.1$ between PGK1P1 and

TABLE 2

Pairwise Linkage Results for Aarskog Syndrome with X-chromosome Markers

Locus	Cytogenetic location	Lod score at recombination fraction (θ) of: 0	0.05	0.1	0.2	0.3	0.4
DXS255	Xp11.22	$-\infty$	2.073	2.042	1.636	1.049	0.394
PGK1P1	Xq12–q13	1.204	1.063	0.922	0.639	0.359	0.112
DXS566	Xq13	$-\infty$	2.095	2.087	1.726	1.171	0.499
DXS72	Xq22	$-\infty$	0.508	0.877	0.980	0.783	0.432

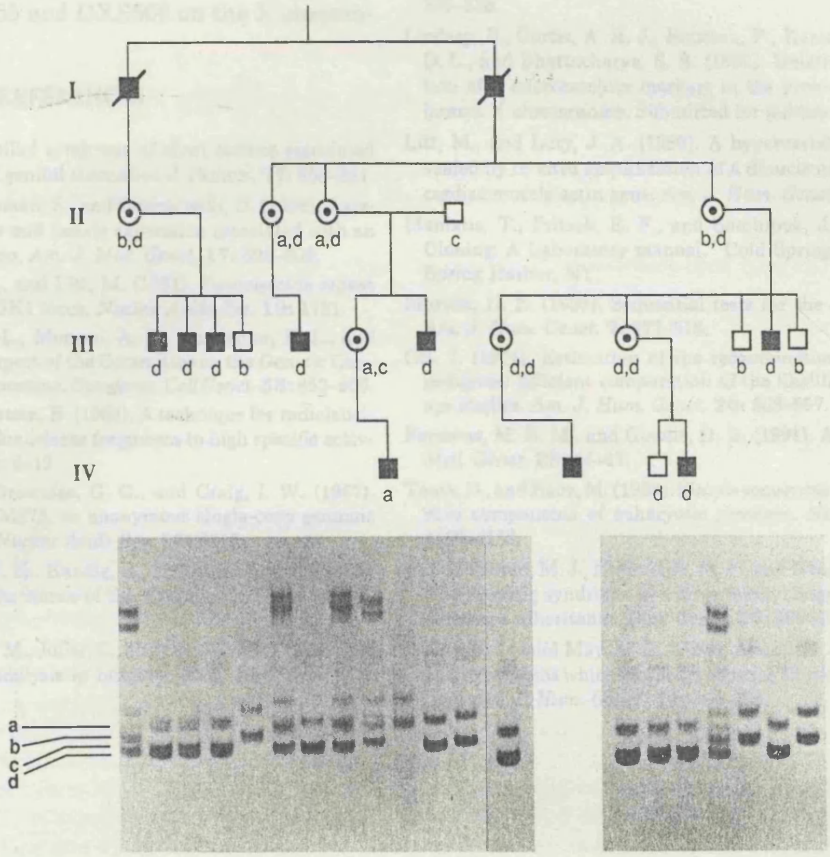


FIG. 2. Family tree. Affected individuals represented by black symbols. Lanes beneath show results with DXS566. This family has four alleles, a, b, c, d. The genotype of each family member is recorded below the respective symbol. It can be seen that Aarskog syndrome appears to be transmitted with allele d, except in the case of III-6, who is a recombinant inheriting both the Aarskog gene and allele a from her mother.

DXS566. As there were no recombinants observed in four informative meioses between Aarskog syndrome and PGK1P1, a maximal LOD was obtained at PGK1P1 (3.4). Two-point analysis was performed between Aarskog syndrome and DXS72, which maps more distally to DXS566. Individual III-6 showed recombination, again passing the grandmaternal allele on to her son, as did IV-2.

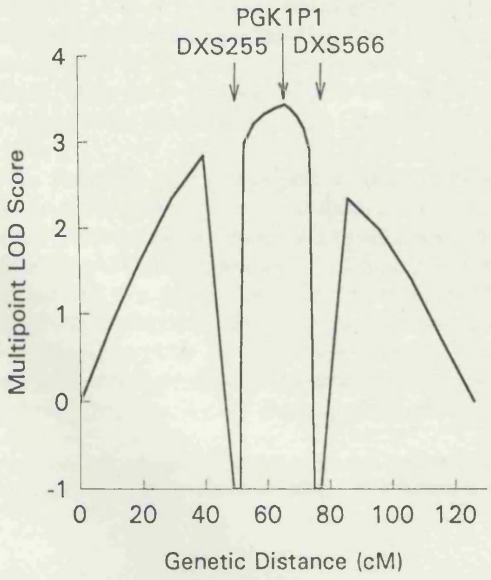


FIG. 3. Multipoint analysis showing a maximum lod score of 3.4 at the PGK1P1 locus.

DISCUSSION

Our analysis highlights the value of microsatellite markers in the mapping of genetic disease. DXS566 will be useful for studies in the proximal region of Xq. For the purposes of multipoint analysis, distances between markers had to be estimated because intermarker genetic distances were not available. However, substitution of different distances made little difference to the maximum LOD of 3.4 (data not shown), which was obtained at the PGK1P1 locus.

Since the publication of the Aarskog translocation breakpoints by Bawle *et al.* in 1984, there has been a reassessment of the case and a modification of the location of the breakpoint from Xq13 to Xp11.21–p11.22 (Thomas Glover, personal communication). Our data are consistent with the gene lying proximal to Xq13 as we have observed recombination in individual III.6 in our family with both DXS566 and DXS72. The existence of an autosomal locus for Aarskog syndrome is

indicated by male to male transmission in documented families. Our data support the existence of a second locus between DXS255 and DXS566 on the X chromosome.

REFERENCES

Aarskog, D. (1970). A familial syndrome of short stature associated with facial dysplasia and genital anomalies. *J. Pediatr.* **77:** 856–861.

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

Browne, D. L., Zonana, J., and Litt, M. (1991). Dinucleotide repeat polymorphism at the PGK1 locus. *Nucleic Acids Res.* **19:** 1721.

Davies, K. E., Mandel, J-L., Monaco, A. P., Nussbaum, R. L., and Willard, H. F. (1991). Report of the Committee on the Genetic Constitution of the X-Chromosome. *Cytogenet. Cell Genet.* **58:** 853–966.

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

Fraser, N. J., Boyd, Y., Brownlee, G. G., and Craig, I. W. (1987). Multi-allelic RFLP for M27β, an anonymous single-copy genomic clone at Xp11.3–Xcen. *Nucleic Acids Res.* **15:** 9616.

Grier, R. E., Farrington, F. H., Kendig, R., and Mamunes, P. (1983). Autosomal dominant inheritance of the Aarskog syndrome. *Am. J. Med. Genet.* **15:** 39–46.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. USA* **81:** 3443–3446.

Lindsay, S., and Bird, A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327:** 336–338.

Lindsay, S., Curtis, A. R. J., Roustan, P., Kamakari, S., Thistleton, D. L., and Bhattacharya, S. S. (1992). Isolation and characterisation of 3 microsatellite markers in the proximal long arm of the human X chromosome. Submitted for publication.

Litt, M., and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44:** 397–401.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). "Molecular Cloning: A Laboratory manual," Cold Spring Harbor Press, Cold Spring Harbor, NY.

Morton, N. E. (1955). Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* **7:** 277–318.

Ott, J. (1974). Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Am. J. Hum. Genet.* **26:** 588–597.

Porteous, M. E. M., and Goudie, D. R. (1991). Aarskog syndrome. *J. Med. Genet.* **28:** 44–47.

Tautz, D., and Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* **12:** 4127–4138.

van de Vooren, M. J., Niermeijer, M. F., and Hoogeboom, J. M. (1983). The Aarskog syndrome in a large family: Suggestive for autosomal dominant inheritance. *Clin. Genet.* **25:** 288–294.

Weber, J. L., and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44:** 388–396.

Hum Genet (1993) 90:551–555

human genetics

X chromosome linkage studies in familial Rett syndrome*

Andrew R.J. Curtis[1], Sophie Headland[2], Susan Lindsay[1], Nicholas S.T. Thomas[2], Eileen Boye[3], Smaragda Kamakari[1], Paul Roustan[1], Maria Anvret[4], Jan Wahlstrom[5], Gillian McCarthy[6], Angus J. Clarke[2], Shomi Bhattacharya[1]*

[1]Department of Human Genetics, University of Newcastle upon Tyne, 19 Claremont Place, Newcastle upon Tyne, UK
[2]University of Wales College of Medicine, Institute of Medical Genetics, Heath Park, Cardiff, UK
[3]Paediatric Research Unit, Guys Hospital, London, UK
[4]Department of Clinical Genetics, Karolinska Hospital, Stockholm, Sweden
[5]Department of Clinical Genetics, East Hospital, Gothenburg, Sweden
[6]Chailey Heritage Hospital, North Chailey, East Sussex, UK

Received: 14 February 1992 / Revised: 10 June 1992

Abstract. Four families, each with two individuals affectecd by Rett Syndrome (RS), were analysed using restriction fragment lenght polymorphisms and microsatellite markers from the X chromosome. In two of the families, X-linked dominant inheritance of the RS defect from a germinally mosaic mother could be assumed. Therefore, maternal X chromosome markers showing discordant inheritance were used to exclude regions of the X chromosome as locations of the RS gene. Much of the short arm could be excluded, including regions containing three candidate genes, OTC, synapsin 1 and synaptophysin. Although most of the long arm was inherited in common it was possible to exclude a centromeric region. Inheritance of X chromosome markers is also presented for two families with affected aunt-niece pairs, one of which has not been previously studied at the DNA level.

Introduction

Rett Syndrome (RS) is a mental retardation disorder, characteristically only affecting females, with an incidence of approximately 1 in 15,000 live-born females. Normal development occurs for 6–12 months followed by developmental regression in progressive stages (Trevathan and Naidu 1988).

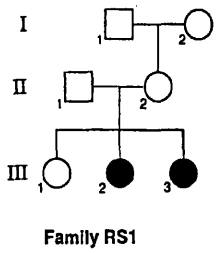
The vast majority of RS cases are isolated females and are considered to be the result of spontaneous, new mutations. However, about 12 familial cases are known world-wide (Zoghbi 1988; Buhler et al. 1990). The best evidence that RS is a genetic disorder is the complete concordance of its occurrence in monozygotic twins compared with dizygotic twins (Zoghbi 1988). That it is an X-linked disorder is suggested by the absence of affected males and supported by studies showing non-random X chromosome inactivation (Zoghbi et al. 1990a) or defects of the normal late-replication process (Riccardi 1986; Martinho et al. 1990) in some cases. In addition, two translocation cases with X chromosomes breakpoints have been described, one at Xp22.11 (Zoghbi et al. 1990b), recently revised to Xp21.3 (Ellison et al. 1991), and the other at Xp11.22 (Journel et al. 1990). Thus, the simplest theory of inheritance is of an X-linked dominant mutation, lethal in hemizygous males and with no reproduction in the affected females (Killian 1986).

Three loci on the short arm of the X chromosome have been proposed as possible candidate genes for RS: ornithine transcarbamylase (OTC), synapsin 1 (Yang-Feng et al. 1986) and synaptophysin (Ozcelik et al. 1990). Although, the results of several studies on OTC and synapsin 1 have revealed no evidence of their involvement in RS (Hanefeld et al. 1986; Clarke et al. 1990; Ferlini et al. 1990), they have not yet been definitively disproved as possibilites.

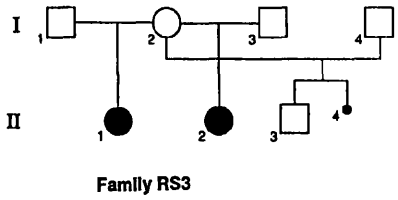
Romeo et al. (1986) have suggested a strategy for mapping the RS locus in families with more than one affected girl by using markers from the entire length of the X chromosome to determine which sections were transmitted in common from their mother. This rests on the assumption that the defect in the affected girls is inherited from their mother. Inheritance through maternal lines appears to be prominent in several of the reported familial cases (Zoghbi 1988).

The aim of the present study is to contribute to the exclusion mapping of the X chromosome in two families with affected sisters or half-sisters (RS1 and RS3). Family RS3 has been studied by several groups and results from two linkage studies have been reported recently (Archidiacono et al. 1991; Ellison et al. 1992). Data is also presented for two other families (RS5 and RS6) with affected aunt-niece pairs where the probable inheritance is, therefore, much less clear. Additional markers are reported to those already published for RS5 (Anvret et al. 1990).

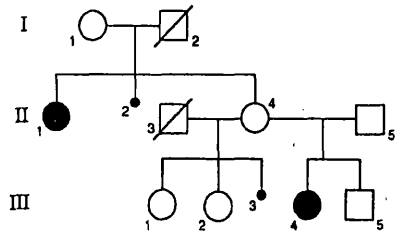
* *Present address and address for correspondence:* Department of Molecular Genetics, Institut of Ophthalmology, Bath Street, London EC1 V9EL, UK



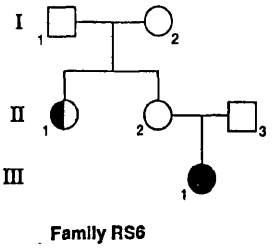
Family RS1



Family RS3



Family RS5



Family RS6

Fig. 1. Pedigress of the four families studied. In family RS6, individual II-1 is *partially shaded* to indicate her less severe form of the disease

Table 1. Markers informative in the study. Reference details for all markers not indicated are given in Davies et al. (1991)

Location	Locus	Probe	Enzyme	Reference
Xp22.32	DXS237	GMGX9	*Hin*dIII	
Xp22.3–p22.2	DXS85	L782	*Eco*RI	
Xp22.2	DXS9	RC8	*Taq*I	
Xp22.1	DXS92	pXG-16	*Hin*dIII	
XP21.2	DMD	Cf56a	*Msp*I	
Xp21.1	DXS206	XJ1.1	*Taq*I	
Xp21.1	DMD (5′ end)	(DYSI)[a]		Feener et al. (1991)
Xp21.1	DMD (5′ end)	(DYSII)[a]		Feener et al. (1991)
Xp21.1	DXS84	L754	*Pst*I	
Xp21.1	OTC	pOTC	*Msp*I	
Xp11.4–11.3	DXS7	L1.28	*Taq*I	
Xp11.3–p11.23	DXS426	[a]		
Xp11.22	DXS255	M27B	*Pst*I	
Xp11.21	DXS14	p58-1	*Msp*I	
Xq12–q13		\5BC	*Xba*I	Unpublished
Xq13.1	DXS159	cpx289	*Pst*I	
Xq13	DXS566	[a]		Porteous et al. (1992)
Xq21.1	TCD	559	*Msp*I	Cremers et al. (1990)
Xq21–q22	DXS571	[a]		Lindsay et al. (1992)
Xq24–q25	DXS424	[a]		
Xq25–q27.1	DXS425	[a]		
Xq28	DXS304	U6.2	*Taq*I	
Xq28	DXS52	F814	*Bcl*I	
Xq28	F8C	F814	*Bcl*I	

[a] Dinucleotide repeat polymorphism

Materials and methods

The pedigrees of the families studied are given in Fig. 1. Three of these families have been described previously: RS1, RS3 and RS5 (Clarke et al. 1990; Hagberg 1985; Anvret et al. 1990, respectively). RS6 has more recently been identified and will be described elsewhere.

Standard procedures (Sambrook et al. 1989) were used to extract DNA from blood lymphocytes, digest it with restriction enzymes, separate the fragments by electrophoresis, blot onto nylon membranes and hybridise with radioactively labelled marker DNA. The full list of markers used is given in Table 1.

Microsatellite polymorphisms were studied using the polymerase chain reaction (PCR). Primers were either synthesised from published sequences (DYSI, DYSII, DXS424, DXS425 and DXS426) or from our own sequence data (DXS566 and DXS571). Standard PCR reactions of 25 µl contained 250 ng DNA, 0.2 m*M* dNTPs, each primer at 0.1–1 µ*M* and 1 unit of enzyme with the appropriate manufacturer's buffer. Amplification of 34 cycles was according to published conditions (see Table 1). Alleles were separated by elec-

trophoresis in a non-denaturing polyacrylamide gel, which was stained with ethidium bromide. Fragments resulting from *Hae*III and *Msp*I digestion of pBR322 were used as molecular size markers.

Results

Concordant or discordant inheritance of markers in the pairs of sisters in RS1 and RS3 is shown in Table 2.

In RS1, access to grandparents' DNA enabled full establishment of phase in the mother for all but 1 of the 16 markers. The two girls have discordantly inherited six markers from the region of the X chromosome short arm defined by DXS92 (Xp22.1) and DXS255 (Xp11.22). This region spans the OTC gene, previously reported to be discordant in the affected girls (Clarke et al. 1990) and includes two other candidate genes, synapsin 1 and synaptophysin. Synapsin 1 is located on the same 390-kb YAC (yeast artificial chromosome) as DXS426 (Coleman et al. 1991) and synaptophysin lies distal to DXS255 at Xp11.23–p11.22 and both of these markers were inherited discordantly. The evidence from RS1 thus suggests that none of these candidates are likely to be the cause of RS in this family. DXS159 (Xq13.1) and \5BC (Xq12–q13) on proximal Xq have also been inherited discordantly by the two affected girls while the other informative markers on Xq were inherited concordantly. Thus, Xp22.2–pter (excluding DXS237) and Xq13–qter remain as likely candidate areas for the RS locus in this family. In addition, two large regions at Xp21.3 (between DXS92 and DMD) and Xp11.21–Xq12 (between DXS255 and \5BC) have not been thoroughly tested although both have discordant flanking markers.

Table 2. Summary of results in families RS1 and RS3. ND, not determined; NI, not informative; +, inherited concordantly in the affected girls; −, inherited discordantly in the affected girls

Locus	RS1	RS3
DXS237 (Xp22.3)	−	ND
DXS85 (Xp22.3–p22.2)	+	ND
DXS9 (Xp22.2)	+	+
DXS92 (Xp22.1)	−	ND
DXS206 (Xp21.1)	NI	+
DMD (5′ end–Xp21.1)	−	NI
OTC (Xp21.1)	−	ND
DXS7 (Xp11.4–11.3)	−	+
DXS426 (Xp11.3–p11.23)	−	ND
DXS255 (Xp11.22)	−	+
DXS14 (Xp11.21)	NI	NI
\5BC (Xq12–q13)	−	ND
DXS159 (Xq13.1)	−	ND
DXS566 (Xq13)	NI	−
TCD (Xq21.1)	+	ND
DXS571 (Xq21–q22)	NI	+
DXS424 (Xq24–q26)	+	+
DXS425 (Xq26–q27.1)	+	+
DXS52 (Xq28)	+	ND
F8C (Xq28)	+	ND

Table 3. Inheritance of grandparental markers in individuals III.4 (RS5) and III.1 (RS6). NI, not informative; ND, not determined; GM, grandmother; GF, grandfather; +, compatible with inheritance from grandparent given at head of column; −, not compatible with inheritance from grandparent given at head of column

Locus	RS5		RS6	
	GM	GF	GM	GF
DXS237	ND	ND	NI	−
DXS85[a]	−	NI	ND	ND
DXS43[a]	−	+	ND	ND
DXS41[a]	−	+	ND	ND
DXS67[a]	−	+	ND	ND
DMD (cf56a)	ND	ND	−	NI
DMD (DYSII)	−	−	NI	−
DXS84[a]	ND	ND	NI	−
OTC[a]	−	−	ND	ND
DXS426	NI	−	ND	ND
DXS255	ND	ND	+	−
DXS14[a]	NI	NI	ND	ND
DXS566	+	−	+	−
DXS571	NI	−	−	NI
DXS424	+	−	−	−
DXS425	NI	NI	ND	ND
DXS304	ND	ND	NI	−
F9[a]	NI	NI	ND	ND

[a] Data from Anvret et al. (1990)

In RS3, phase was not known in the mother (I-2, Fig. 1) but could be established for all markers in II-1 whose father's genotype was determined, and assumed for 8 of the 10 markers in her half-sister (II-2). Seven markers show concordant inheritance in the two half-sisters. Only DXS566 has been inherited discordantly, allowing exclusion at this locus.

A summary of the inheritance of grandparental markers in RS5 and RS6 is given in Table 3. In RS5, phase was known for all markers in III-4, inferred for all markers in the two unaffected half-sisters (III-1 and III-2) and for all except DXS425 in II-1 and II-4. Assuming that the disease has not arisen separately in II-1 and III-4, then there are two possible modes of inheritance: from I-1 (grandmother; GM) or from I-2 (grandfather; GF). If we consider inheritance from GM first, then any region of her X chromosome that is inherited discordantly by II-1 and II-4 or not inherited by III-4 can be excluded. This applies only to DMD (DYSII; Xp21.2). If inheritance from GF is considered then regions inherited from GM in III-4 can be excluded, which applies to all loci that are informative in our study, i.e. DMD, DXS426, DXS566, DXS571 and DXS424.

Access to paternal DNA enabled phase to be established in the family RS6 for II-1, II-2 and III-1 for all nine informative markers. III-1 had not inherited her grandfather's (I-1) X chromosome at seven of these nine markers. DMD (cf56a) and DXS571 were uninformative in II-2. Only Xp21.1 and Xq21–q22, of the regions examined, remain as candidates if inheritance from GF is postulated. If inheritance from GM is assumed, then

DMD, DXS571 an DXS424 are excluded as they are inherited discordantly in II-1 and II-2. In this case, DXS237 (GMGX9), DMD (DYSII), DXS84 and DXS304 were uninformative and the regions around DXS255 and DXS566 remain as candidates.

Discussion

The strategy proposed by Romeo et al. (1986), that is exclusion of discordantly inherited regions, is applicable where there is evidence of maternal inheritance. In RS1, this is provided by the mother sharing a metabolic defect with one of her affected daughters similar to that seen in OTC-carrier females (Clarke et al. 1990) and also observed in some other RS cases (Thomas et al. 1990). In RS3, the two affected half-sisters are related through their mother (Hagberg 1985; Archidiacono et al. 1991). Assuming that the gene defect is at the same locus in RS1 and RS3 it is, therefore, possible to combine evidence of exclusion. Regions excluded in our study by a comparison of alleles inherited in the two pairs of sisters are: DXS237, Xp22.32; DXS92, Xp22.1; DMD (5′ end) to DXS255, Xp21.1–Xp11.22; \5BC to DXS566, Xq12–Xq13. Double recombination events are unlikely given the density of markers tested, particularly between DMD and DXS255, but they cannot be excluded.

Archidiacono et al. (1991) showed in RS3 that the sisters had inherited alleles discordantly at DXS278 (p22.3), DXZ1 (centromere) and DXYS1X (q21.3). Ellison et al. (1992) (RS3 reported as their family 2) also showed discordance at DXS704 (p21.3), PGK1 (q13), DXS453 (q13), DXS13 (q21.31–q21.33), DXS454 and DXS458 (both in q21.1–q23) further supporting the exclusion of the proximal long arm (\5BC to DXS458) as a location of the RS gene. They also presented data from a second half-sister pair that, similarly to RS1 in this study, showed discordant inheritance of a large region of the short arm (DMD to DXS14) including the OTC, synapsin 1 and synaptophysin loci.

Evidence for a location of RS at Xp11.22, the breakpoint in one reported translocation case is not supported by the exclusion of the region defined by DXS255 in RS1. The region of the Xp21.3 translocation case breakpoint has not been tested in our study or in detail in Ellison et al. (1992) and so reamins as a possible candidate.

The present study, in conjunction with previous reports, supports the conclusion that the RS locus lies outside the region defined by DMD (p21.2) to DXS458 (q21.1–23). The loci DXS278 (p22.3), DXS237 (p22.3) and DXS92 (p22.1) can also be excluded. The region between the markers DXS9 (Xp22.2) and DXS92 (Xp22.1) remains untested. The change from concordance to discordance in both RS1 and RS3 makes continued study of this region essential. The distal portions of both the long and short arms also require further testing. To do this, more families with affected sisters or half-sisters with evidence of maternal inheritance are needed. Highly informative microsatellites will be invaluable as it is important that each of the limited number of families is fully informative at as many loci as possible.

The aunt-niece pair families represented by RS5 and RS6 remain a problem. If is of course possible that the affected individuals may have different diseases or that RS has arisen independently by sporadic mutation. If a simple model of X-linked inheritance is assumed, the absence of the syndrome in the mother (II-4 in RS5 and II-2 in RS6) has to be explained. It is possible that there is a very favourable and skewed X inactivation pattern in II-4 (RS5) and II-2 (RS6) in the relevant cell type. Alternatively, more complex models of inheritance may have to be invoked to provide an explanation (Johnson 1980; Buhler et al. 1990). The results of the study of RS5 by Anvret et al. (1990) are summarised with our own in Table 3. The regions around DMD (5′ end; p21.1) and DXS424 (p11.3–p11.23) are excluded for both testable hypotheses, i.e. inheritance from GM or from GF.

Acknowledgements. We would like to acknowledge the assistance and helpful discussions from the staff of the Molecular Genetics Unit, in particular Daisy Haggerty. We also thank Dr. Huda Zogbhi for generously providing us with her manuscript prior to publication. This research was funded by the Sir Jules Thorn Trust (A.R. J.C.) and the Wellcome Trust (S.L.; grant number 03229/Z/91/1.27/DG).

References

Anvret M, Wahlstrom J, Skogaberg P, Hagberg B (1990) Segregation analysis of the X-chromosome in a family with Rett syndrome in two generations. Am J Med Genet 37:31–35

Archidiacono N, Lerone M, Rocchi M, Anvret M, Ozcelik T, Francke U, Romeo G (1991) Rett syndrome: exclusion mapping following the hypothesis of germinal mosaicism for new X-linked mutations. Hum Genet 86:604–606

Buhler EM, Malik NJ, Alkan M (1990) Another model for the inheritance of Rett syndrome. Am J Med Genet 36:126–131

Clarke A, Gardner-Medwin D, Richardson J, McGann A, Bonham JR, Carpenter KH, Bhattacharya S, et al (1990) Abnormalities of carbohydrate metabolism and of OCT gene function in the Rett syndrome. Brain Dev 12:199–124

Coleman MP, Murray JC, Willard HF, Nolan KF, Reid KBM, Blake DJ, Lindsay S, Bhattacharya SS, Wright A, Davies KE (1991) Genetic and physical mapping around the Properdin P gene. Genomics 11:991–996

Cremers FPM, van de Pol DJR, van Kerkhoff LPM, Wieringa B, Ropers H-H (1990) Cloning of a gene that is rearranged in patients with choroideremia. Nature 347:674–677

Davies KE, Mandel JL, Monaco AP, Nussbaum RL, Willard HF (1991) Report of the committee on the genetic constitution of the X chromosome. (11th International Workshop on Human Gene Mapping) Cytogenet Cell Genet 58:853–966

Ellison KA, Chinault AC, Zoghbi HY (1991) The use of radiation hybrids and YAC clones in the efforts aimed at cloning the X chromosomal breakpoint in a patient with Rett syndrome. Poster 2065 presented at the Sixth Interantional Congress of Human Genetics, Washington, USA, October 6–11

Ellision KA, Fill CP, Terwillinger J, DeGennaro LJ, Martin-Gallardo A, Anvret M, Percy AK, Ott J, Zoghbi H (1992) Examination of X chromosome markers in Rett syndrome: Exclusion mapping with a novel variation on multilocus linkage analysis. Am J Hum Genet 50:278–287

Feener CA, Boyce FM, Kunkel LM (1991) Rapid detection of CA polymorphisms in cloned DNA: application to the 5′ region of the dystrophin gene. Am J Med Genet 48:621–627

Ferlini A, Ansaloni L, Nobile L, Forabosco A (1990) Molecular analysis of the Rett syndrome using cDNA Synapsin I as a probe. Brain Dev 12:136–139

Hagberg B (1985) Rett syndrome: Prevalence and impact on progressive severe mental retardation in girls. Acta Paediatr Scand 74:405–408

Hanefeld F, Hanefeld U, Wilichowski E, Schmidtke J (1986) Rett syndrome – search for genetic markers. Am J Med Genet 24: 377–382

Johnson WF (1980) Metabolic interference and the +/− heterozygote. A hypothetical form of simple inheritance which is neither dominant nor recessive. Am J Hum Genet 32:374–386

Journel H, Melki J, Turleau C, Munnich A, de Grouchy J (1990) Rett phenotype with X/autosome translocation: possible mapping to the short arm of chromosome X. Am J Med Genet 35: 142–147

Killian W (1986) On the genetics of Rett syndrome: analysis of family and pedigree data. Am J Med Genet 24:369–376

Martinho PS, Otto PC, Kok F, Diament A, Marques-Dias MJ, Gonzalez CH (1990) In search of a genetic basis for the Rett syndrome. Hum Genet 86:131–134

Ozcelik T, Lafreniere RG, Archer BT, Johnston PA, Willard HF, Francke U, Sudhof TC (1990) Synaptophysin: structure of the human gene and assignment to the X chromosome in man and mouse. Am J Hum Genet 47:551–561

Porteous MEM, Curtis A, Lindsay S, Williams O, Govdie D, Kamakari S, Bhattacharya SS (1992) The gene for Aarskog syndrome is located between DXS255 and DXS256. Genomics 14:298–301

Riccardi VM (1986) The Rett syndrome: genetics and the future. Am J Med Gent 24:389–402

Romeo G, Archidiacono N, Ferlini A, Rocchi M (1986) Rett syndrome: Lack of association with fragile site Xp22 and strategy for genetic mapping of X-linked new mutations. Am J Med Genet 24:355–359

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Thomas S, Oberholzer V, Wilson J, Hjelm M (1990) The urea cycle in the Rett syndrome. Brain Dev 12:93–96

Trevathan E, Naidu S (1988) The clinical recognition and differential diagnosis of Rett syndrome. J Child Neurol 3 [Suppl]:6–16

Yang-Feng TL, DeGennaro LJ, Francke U (1986) Genes for synapsin 1, a neuronal phosphoprotein, map to conserved regions of human and murine X chromosomes. Proc Natl Acad Sci USA 83:8679–8683

Zoghbi H (1988) Genetic aspects of Rett syndrome. J Child Neurol 3 [Suppl]:76–78

Zoghbi HY, Percy AK, Schultz RJ, Fill C (1990a) Patterns of X chromosome inactivation in the Rett syndrome. Brain Dev 12: 131–135

Zoghbi HY, Percy AK, Schultz RJ, Fill C (1990b) A de novo X;3 translocation in Rett syndrome. Am J Med Genet 35:148–151

Polymorphism of the protein C inhibitor (PCI) gene on chromosome 14

Katherine J.Ashbourne[1], Barbara C.Byth[1], Joost C.M.Meijers[2] and Diane W.Cox[1,3]*
[1]Research Institute, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada, [2]Department of Haematology, University Hospital Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands and [3]Departments of Paediatrics and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5G 1X8, Canada

Source/Description of Clone: A single copy, human genomic 3 kb *Hin*dIII fragment cloned into pUC19, designated as PCI Hind3. This is a portion of the gene for protein C inhibitor (1), a member of the serine protease inhibitor superfamily, which may be identical to plasminogen activator inhibitor 3 (2).

Polymorphism: A two allele polymorphism detected with *Eco*RI, with bands at 3.1 kb and 5.3 kb, and a constant band at 1.2 kb.

Frequency: Studied in DNA from 64 parents in CEPH families. Frequency of 5.3 kb fragment (A1): 0.46. Frequency of 3.1 kb fragment (A2): 0.54. Heterozygosity = 0.50.

Not Polymorphic For: *Bam*HI, *Bgl*II, *Hin*dIII, *Msp*I, *Pst*I, *Sac*I, *Taq*I, *Xba*I.

Chromosomal Localization: Localised to chromosome 14 using somatic cell hybrids (1), and to the protease inhibitor cluster at q32.1 using pulsed-field gel electrophoresis (3).

Mendelian Inheritance: Codominant segregation shown in two CEPH families (26 individuals).

Probe Availability: Write to Drs J.C.M.Meijers or D.W.Cox at the above addresses.

Other Comments: Wash stringency of 0.1% SDS, 0.1×SSC at 65°C. The probe PCI Hind3 requires pre-annealing with sheared human genomic DNA before use.

Acknowledgements: This study was supported by grant MT-5426 from the Medical Research Council of Canada to D.W.C., and a fellowship of the Royal Netherlands Academy of Arts and Sciences to J.C.M.M.

References: 1) Meijers,J.C.M. and Chung,D.W. (1991) *J. Bio. Chem.* **266**, 15028–15034. 2) Stief,T.W., Radtke,K.-P. *et al.* (1987) *Biol. Chem.* **368**, 1427–1433. 3) Billingsley,G.D., Walter,M.A. *et al.* (1992) *Am. J. Hum. Genet.* in press.

* To whom correspondence should be addressed at: The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada

Dinucleotide repeat polymorphism at the DXS573 locus

P.Roustan+, A.R.J.Curtis, S.Kamakari, D.Thiselton, S.Lindsay and S.S.Bhattacharya*
Molecular Genetics Unit, University of Newcastle upon Tyne, Newcastle NE1 7RU, UK

Source and Description of Clone: Cosmid HX82 was isolated from a library containing insert DNA from a mouse-human hybrid cell line which has the X chromosome as the only detectable human chromosome (1). It was then selected by hybridisation to a poly (dCA.dGT) probe (Pharmacia). pHX82-2 is a Sau 3A 1 subclone of this cosmid which contains a $(CA)_{22}$ microsatellite. pHX82-2 was sequenced and the sequences flanking the repeat (EMBL accession number X66056) were used to design PCR primers.

PCR Primers: pHX82-2/A=5′-GCCACCCAATCTAAA-GTTGC-3′ (CA strand)
pHX82-2/B=5′-GGTGATGATGAGTGATAAGG-3′ (GT strand)

Polymorphism: Allelic fragments were detected on DNA sequencing gels. Lengths (nt) are: A_1=145, A_2=143, A_3=141, A_4=139, A_5=137. Subclone pHX82-2 contains A_4.

Frequencies: Estimated from 50 (100 chromosomes) unrelated caucasian females: A_1=0.01, A_2=0.42, A_3=0.12, A_4=0.30, A_5=0.15. Observed heterozygosity = 0.72.

Chromosomal Localization: DXS573 has been assigned to Xp11.1–p11.22 by hybridisation of the cosmid HX82 to a panel of somatic cell hybrids DNAs and confirmed by fluorescence *in situ* hybridisation.

Mendelian Inheritance: X-linked co-dominant segregation was observed in 5 informative families.

PCR Conditions: PCR amplifications were carried out for 24 cycles using ^{32}P end labelled primer pHX82-2/A as following: denaturation at 94°C for 1 min, annealing and extension at 56°C for 2 min with a first denaturation step at 94°C for 3 min and a final elongation step at 56°C for 10 min.

Acknowledgements: This work was supported by grants from Fédération des Aveugles de France, Conseil Régional Midi-Pyrénées, Fondation de France, Association Française de lutte contre les Myopathies, the Sir Jules Thorn Trust, the British Retinitis Pigmentosa society and the Wellcome Trust.

Reference: 1) Lindsay,S. and Bird,A.P. (1987) *Nature* **327**, 336–338.

* To whom correspondence should be addressed

+Present address: Laboratoire d'Endocrinologie Expérimentale, Bât. L3, C.H.U. Rangueil, 31054 Toulouse, France

Dinucleotide repeat polymorphism at the DXS556 locus

D.L.Thiselton+, S.Kamakari+, A.R.J.Curtis, M.P.Coleman[1], K.E.Davies[1], S.Lindsay and S.Bhattacharya*,+
Molecular Genetics Unit, University of Newcastle upon Tyne, NE1 7RU and [1]Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

Source and Description of Clone: Cosmid HX20 was isolated from a library constructed using DNA from a mouse–human hybrid cell line, Thy-B1-33-12, which has the X chromosome as the only detectable human chromosome (1). It was then selected by hybridisation to a poly (dCA.dGT) probe (Pharmacia). pHX20 is a Sau3AI subclone of this cosmid which contains a $(CA)_{21}$ microsatellite. pHX20 was sequenced and the sequences flanking the repeat (EMBL accession number X67603) were used to design PCR primers.

PCR Primers:

pHX20A = 5'-AGTTTGAGGGCTTCGTTTAC-3' (CA strand)
pHX20B = 5'-TATGAAGACAGCCAACTTAGA-3' (GT strand)

Polymorphism: Allelic fragments were detected on DNA sequencing gels. Lengths (nt) are:

A1 = 192	A2 = 190	A3 = 188	A4 = 186
A5 = 184	A6 = 182	A7 = 178	A8 = 176

Subclone pHX20 contains A5.

Frequencies: Estimated from 36 (72 chromosomes) unrelated Caucasian familes:

A1 = 0.04	A2 = 0.06	A3 = 0.07	A4 = 0.13
A5 = 0.37	A6 = 0.04	A7 = 0.01	A8 = 0.29

Observed heterozygosity = 0.73.

Chromosomal Localization: DXS556 has been assigned to Xp11.23–p11.3 by hybridization of the cosmid HX20 to a panel of somatic cell hybrids DNAs and confirmed by fluorescence *in situ* hybridisation.

Mendelian Inheritance: X-linked co-dominant segregation was observed in 5 informative families.

PCR Conditions: PCR amplifications were carried out for 30 cycles using ^{32}P end labelled primer pHX20A as follows: denaturation at 94°C for 1 minute, annealing at 57°C for 1 minute and extension at 72°C for 2 minutes, with a first denaturation step at 94°C for 3 min and a final elongation step at 72°C for 10 min.

Acknowledgements: This work was supported by the British Retinitis Pigmentosa Society, the Sir Jules Thorn Trust and the Wellcome Trust (SL grant no. 033229/Z/91/Z/1.27/DG).

Reference: Lindsay,S. and Bird,A.P. (1987) *Nature* **327** 336–338.

* To whom correspondence should be addressed at +present address: Department of Molecular Genetics, Institute of Ophthalmology, London EC1V 9EL, UK

Dinucleotide repeat polymorphism at D21S49 (21q22.3)

I.N.Bespalova[1,3], S.Kiousis[3], J.S.Chamberlain[3] and M.Burmeister[1,2,3,*]
[1]Mental Health Research Institute, [2]Department of Psychiatry and [3]Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109-0720, USA

Source/Description: A $(GT)_{19}$ repeat was identified in phage SF58, which defines the locus D21S49 (1). A 1 kb EcoRI fragment containing the repeat was subcloned and partially sequenced. (Genbank accession no. L07094).

Primer Sequences:

Primer 1 — 5'-TTCCTGTAAGTCACCTGCATCTAAC-3'
Primer 2 — 5'-CGACTGTCTTTCTTAGGCTTCTGAC-3'

Frequency: 80 chromosomes from the parents of the CEPH pedigrees were analyzed. Heterozygosity = 0.70

Allele size	(bp)	Frequency
A1	162	0.05
A2	160	0.04
A3	158	0.54
A4	156	0.29
A5	154	0.03
A6	152	0.06

Reference CEPH genotypes:

13291-1	156/156
13291-2	158/160

Chromosomal Location: D21S49 is known to be on chromosome 21q22.3, between Mx and PFKL, near D21S42 and D21S19 (2).

Mendelian Inheritance: Mendelian inheritance was observed in four CEPH pedigrees.

PCR Conditions: Amplification is performed in a volume of 25 µl. Each reaction contains 20 ng genomic DNA, 250 µM each of dATP, dGTP, dTTP, and dCTP, 10 pmol of each primer, 0.5 U Taq polymerase, 30 mM Tris-HCl, pH 8.5, 50 mM KCl, 2 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.01% gelatin, 0.1% Thesit. Amplifications conditions are 40 cycles of 92°C for 30 sec, 65°C for 30 sec, and 72°C for 1 min. Amplification products can be resolved on 15% polyacrylamide gels.

Acknowledgement: Supported by a grant from the Epilepsy Foundation of America (M.B.).

References: 1) Korenberg *et al. (1987) Am. J. Hum. Genet.* **41**, 963. 2) Burmeister *et al. (1991) Genomics* **9**, 19.

* To whom correspondence should be addressed

SHORT COMMUNICATION

Isolation and Characterization of Three Microsatellite Markers in the Proximal Long Arm of the Human X Chromosome

S. LINDSAY,* A. R. J. CURTIS,* P. ROUSTAN,*,[1] S. KAMAKARI,† D. L. THISELTON,† A. STEPHENSON,* AND S. S. BHATTACHARYA†,[2]

*Molecular Genetics Unit, Division of Human Genetics, University of Newcastle upon Tyne, United Kingdom; and †Department of Molecular Genetics, Institute of Ophthalmology, Bath Street, London EC1V 9EL, United Kingdom.

Received June 19, 1992; revised March 3, 1993

Three microsatellites have been identified in cosmids from the human X chromosome. The cosmids have been assigned locus numbers DXS554, DXS559, and DXS566 and have been localized to Xq12–q13 (DXS554 and DXS559) and Xq13 (DXS566). In addition, they have been genetically mapped in relation to the androgen receptor (AR), phosphoglycerate kinase 1, pseudogene 1 (PGK1P1), and phosphoglycerate kinase (PGK1) loci in the proximal long arm. Genetically, the localization of microsatellites at DXS554 and DXS566 is indistinguishable from PGK1, whereas that at DXS559 maps between AR and PGK1, close to PGK1P1. DXS566 is identical to the independently identified DXS441 marker. These markers should be useful for physical and genetic mapping in this region.

The proximal long arm of the human X chromosome contains several interesting genetic disease loci (e.g., ectodermal dysplasia, Menkes syndrome, and sideroblastic anaemia), as well as the X inactivation center (2). As yet, there are relatively few reported microsatellites identified in this region. The microsatellites described here should, therefore, be helpful in mapping studies in this region of the X chromosome.

We have isolated three cosmids (HX10, HX28, and HX60), from a library that has been previously described (5), which map to the proximal long arm of the X chromosome. They were initially shown to map to the X chromosome by hybridization signal dosage (data not shown). HX10, HX28, and HX60 have been assigned human gene mapping (HGM) symbols DXS554, DXS559, and DXS566, respectively. Sublocalization of the cosmids was carried out by *in situ* hybridization, which showed that HX10 and HX28 map to Xq12–q13, whereas HX60 maps to Xq13 (data not shown).

The cosmids were shown to contain a microsatellite by hybridization with a poly(dCA·dGT) probe (Pharmacia). The relevant fragments were then subcloned into a pT7T3 vector (Pharmacia) and sequenced with the Sequenase kit (USB) and the sequences flanking the repeat were used to design PCR primers.

DXS554 pHX10 is a *Sau*3A subclone of cosmid HX10 which contains a $(CA)_{20}$ microsatellite (EMBL Accession No. X66053). The PCR primers used are: 5′-TTGCATAGTATTGCGGGGTG-3′ and 5′-CAACCTAAGGTGTCCATCAG-3′. Allelic fragments of the following lengths (in nucleotides) were detected: $A_1 = 111$, $A_2 = 109$, $A_3 = 107$, $A_4 = 105$. Subclone pHX10 contains A_1. The allele frequencies were estimated from 17 (34 chromosomes) unrelated caucasian females, $A_1 = 0.79$, $A_2 = 0.15$, $A_3 = 0.03$, $A_4 = 0.03$, giving an observed heterozygosity of 29% and an expected heterozygosity of 0.35.

DXS559 pHX28 is a *Sau*3A subclone of cosmid HX28 which contains a $(CA)_{24}$ microsatellite (EMBL Accession No. X66054). The PCR primers used, allele lengths, and frequencies were described by (8). The observed heterozygosity was 63% in 30 unrelated caucasian females and the expected heterozygosity is 0.74.

The microsatellite at DXS566 was reported in Porteous *et al.*, (6). Although the present work was in progress Ram *et al.* (7) reported a microsatellite at DXS441, located at Xq13. On analyzing the sequences flanking the microsatellite in HX60 we identified the primer sequences that they described to amplify DXS441. In our study the primers used were 5′-CATTGCCAGAATATATGAAAC-3′ and 5′-ACTTGTCAAATATGGCTCATA-3′ as previously described (6).

PCR amplifications were carried out for 24 cycles using one ^{32}P-labeled primer as follows: denaturation at 94°C for 1 min, annealing at 52°C (DXS554), 55°C (DXS566), or 56°C (DXS559) for 2 min, and extension at 72°C for 2 min, with a first denaturation step at 94°C for 3 min and a final elongation step at 72°C for 10 min. At all three loci Mendelian inheritance of alleles was observed in at least five informative families (data not shown).

To confirm and refine the physical localizations, the three clones were also mapped genetically with respect to three markers known to lie in the proximal long arm and whose map positions were established (2): AR (Xq11.2–q12), PGK1P1 (Xq11.2–q12), and PGK1 (Xq13.3). Their order on Xq is also known to be cen–

[1] Present address: Laboratoire d'Endocrinologie Experimentale, Bat. L3, C.H.U. Rangueil, 31054 Toulouse, France.

[2] To whom correspondence should be addressed.

GENOMICS 17, 208–210 (1993)
0888-7543/93 $5.00

AR–PGK1P1–PGK1–tel (2). The conditions and primers used for AR, PGK1P1, and PGK1 are as given in Sleddens *et al.* (9), Browne *et al.*, (1), and van Kamp *et al.* (10), respectively. The loci were analyzed in 12 families comprising a total of 110 meioses. The results were analyzed pairwise to determine the θ value at maximum LOD using the ILINK program (4) and are given in Table 1. From this table it can be seen that DXS566 (HX60) and DXS554 (HX10) could not be distinguished genetically from PGK1 or from each other, although the data are more convincing for DXS566 because of its higher level of informativeness. The localization of DXS566 is consistent with information from DXS441 (3). From the results given in Table 1, DXS559 (HX28) was genetically distinguishable from DXS566 and PGK1 and appears to lie close to PGK1P1. That the microsatellites are different in HX28 and PGK1P1 is shown by their distinctly different levels of heterozygosity (63 and 42%, respectively), the absence of the PGK1P1 primer sequence in our flanking sequence of the HX28 microsatellite and the fact that PGK1P1 primers do not amplify clone HX28 DNA (data not shown). The microsatellite at PGK1P1 is not highly polymorphic, which means that there is less information about its position relative to the other markers.

Figure 1 shows three recombinant meiotic events informative for ordering the loci. In meiosis 1 there is a recombination between the AR locus and PGK1P1 but not between PGK1P1 and DXS566. As DXS566 has been shown to be near PGK1 (Table 1), this indicates that PGK1P1 lies distal to the AR locus. In meiosis 2, there is a recombination between AR and DXS559 but not between DXS559 and DXS566 or PGK1. This suggests that DXS559 also lies distal to the AR locus, a conclusion supported by the linkage data given in Table 1. Finally, in meiosis 3 there is a recombination between DXS559 and DXS566 but not between DXS566 and PGK1. This places DXS559 proximal to DXS566 and PGK1.

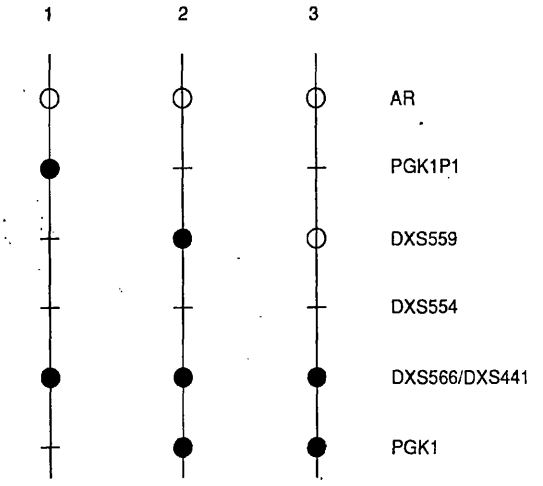


FIG. 1. Representation of recombinant meiotic events. Open and filled circles represent alleles corresponding to one or other of the maternal chromosomes, for informative loci. DXS554 was not informative in these three events.

An ILINK order of preference was carried out on the six loci. This was done in overlapping groups of four loci as there were too many possible haplotypes to permit all six loci to be analyzed at once. From this analysis (data not shown) the order of preference for the loci is AR–(PGK1P1, DXS559)–(DXS554, DXS566/DXS441, PGK1). From Table 1 and the multiply informative meioses (Fig. 1) it was clear that DXS559 is located between AR and PGK1, whereas DXS566 is very close to PGK1. The ILINK analysis could not separate PGK1P1 and DXS559 nor the group DXS554, DXS566, and PGK1.

TABLE 1

Two-Point Analyses of Locus Relationships

Locus 1	Locus 2	Theta-max (1 v 2)	LOD-max
AR	PGK1P1	0.053	2.09
AR	DXS559	0.055	5.94
AR	DXS554	0.139	0.94
AR	DXS566	0.112	5.59
AR	PGK1	0.114	2.70
PGK1P1	DXS559	0.0	1.81
PGK1P1	DXS554	0.06	3.06
PGK1P1	DXS566	0.07	5.30
PGK1P1	PGK1	0.086	2.71
DXS559	DXS554	0.0	1.74
DXS559	DXS566	0.021	10.17
DXS559	PGK1	0.042	3.43
DXS554	DXS566	0.0	4.59
DXS554	PGK1	0	4.11
DXS566	PGK1	0.0	10.84

ACKNOWLEDGMENTS

This work was supported by The Wellcome Trust (S.L.; Grant 033229/Z/91/Z/1.27/DG), the British Retinitis Pigmentosa Society, Ulverscroft Foundation, the Sir Jules Thorn Trust, Federation des Aveugles de France, Conseil Regional Midi-Pyrenees, Fondation de France, and Association Francaise de lutte contre les Myopathies.

REFERENCES

1. Browne, D. L., Zonana, J., and Litt, M. (1992). Dinucleotide repeat polymorphism at the PGK1P1 locus. *Nucleic Acids Res.* **20:** 1169.
2. Davies, K. E., Mandel, J. L., Monaco, A. P., Nussbaum, R. L., and Willard, H. E. (1991). Report of the committee on the genetic constitution of the X chromosome. *Cytogenet. Cell Genet.* **58:** 853–9966.
3. Lafreniere, R. G., Brown, C. J., Powers, V. E., Carrel, L., Davies, K. E., Barker, D. F., and Willard, H. F. (1991). Physical mapping of 60 DNA markers in the p21.1–q21.3 region of the human X chromosome. *Genomics* **11:** 352–363.

4. Lathrop, G. M., and Lalouel, J. M. (1984). Easy calculation of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* **36:** 1245–1249.
5. Lindsay, S., and Bird, A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327:** 336–338.
6. Porteous, M. E. M., Curtis, A., Lindsay, S., Williams, O., Goudie, D., Kamakari, S., and Bhattacharya, S. S. (1992). The gene for Aarskog syndrome is located between DXS255 and DXS566 (Xp11.2–q13). *Genomics* **14:** 298–301.
7. Ram, K. T., Barker, D. F., and Puck, J. M. (1992). Dinucleotide repeat polymorphism at the DXS441 locus. *Nucleic Acids Res.* **20:** 1428.
8. Roustan, P., Curtis, A. R. J., Kamakari, S., Thiselton, D. L., Lindsay, S., and Bhattacharya, S. S. (1992). Dinucleotide repeat polymorphism at the DXS559 locus. *Hum. Mol. Genet.* **1:** 778.
9. Sleddens, H. F. B. M., Oostra, B. A., Brinkmann, A. O., and Trapman, J. (1992). Trinucleotide repeat polymorphism in the androgen receptor gene (AR). *Nucleic Acids Res.* **20:** 1427.
10. van Kamp, H., Jansen, R., Willemze, R., Fibbe, W. E., and Landegent, J. E. (1991). Studies on clonality by PCR analysis of the PGK1 gene. *Nucleic Acids Res.* **19:** 2794.