# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

THE UNIVERSITY OF CALGARY

# Physical Map Covering 2.5 Mb Distal to DXS6849 (ZNF21) in Human Xp11.3-p11.23

by

Karen Leigh Stoddart

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MEDICAL SCIENCE

CALGARY, ALBERTA

SEPTEMBER, 1997

## ABSTRACT

A 2.5 Mb contig was constructed of YACs and PACs, extending from DXS6849 to a new marker EC7034R, one megabase distal to UBE1 within the p11.3 region of the human X chromosome. This contig, which has on average three-fold cloned coverage, was assembled using 37 markers, including 15 new STSs developed from YAC and PAC end-fragments, for an average inter-marker distance of 70 kb. Marker order was predicted from STS content, cell hybrid data and SEGMAP analysis. One polymorphic TC dinucleotide repeat sequence from an end-clone was identified. The contig is merged with a published physical map in the centromeric direction, and provides reagents to aid in the sequencing of the X chromosome and in the finding of genes for X-linked congenital stationary night blindness, X-linked retinitis pigmentosa, and Åland Island eye disease.

## ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Torben Bech-Hansen for his guidance throughout the course of this study. Secondly, I would like to thank the members of my supervisory committee, Drs. Fred Biddle, Robert Bell, and Paul Mains for their participation and contribution towards this effort.

I would like to thank all past and present members of Dr. Bech-Hansen's laboratory. I would especially like to thank Dr. Kym Boycott for being a wonderful friend and teacher, and Bob Winkfein for his helpful discussions and advice.

A special thanks to my parents Lorna and Ian, to my sister Allison, brother Andrew, grandmothers Kate-Ann Stephen and Marjorie Stoddart for their unconditional love and support throughout not only this work but my entire my life. I would also like to thank all of my friends, especially William Panenka, Humberto Jijon and Julie Fisher, for their encouragement, patience and support throughout this work.

This work was supported in part by grants from CGAT/MRC Canada and the RP Research Foundation to NTBH.

## DEDICATION

To Ian, Lorna, Allison, Andrew, Kate-Ann and Marjorie

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ABBREVIATIONS

| | |
|---|---|
| A | adenosine residue |
| BAC | bacterial artificial chromosome |
| bp | base pair |
| BSA | bovine serum albumin |
| C | cytidine residue |
| $^{o}C$ | degrees Celsius |
| CHO | Chinese hamster ovary |
| Ci | Curie |
| cm | centimetre |
| cM | centimorgan |
| cpm | counts per minute |
| CSNB | X-linked congenital stationary night blindness |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytidine triphosphate |
| ddATP | dideoxyadenosine triphosphate |
| ddCTP | dideoxycytidine triphosphate |
| ddGTP | dideoxyguanosine triphosphate |
| $ddH_20$ | double distilled water |
| dGTP | deoxyguanosine triphosphate |
| DNA | deoxyribonucleic acid |
| ds | double stranded |
| dsDNA | double stranded deoxyribonucleic acid |
| dTTP | deoxythymidine triphosphate |
| EDTA | ethylenediaminetetraacetic acid |

| | |
|---|---|
| EtBr | ethidium bromide |
| EST | expressed sequence tag |
| G | guanosine residue |
| g | gram |
| GDB | genome data base |
| HCl | hydrochloric acid |
| HEPES | N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid |
| HPRT | hypoxanthine phosphoribosyl transferase |
| kb | kilobase pair |
| KCl | potassium chloride |
| M | molar |
| MAOA | monoamine oxidase A |
| MAOB | monoamine oxidase B |
| µg | microgram |
| mg | milligram |
| min | minute |
| $MgCl_2$ | magnesium chloride |
| µL | microlitre |
| mL | millilitre |
| µM | micromolar |
| mM | millimolar |
| mmol | millimole |
| NaCl | sodium chloride |
| NaOH | sodium hydroxide |
| ng | nanogram |

| | |
|---|---|
| OAc | Acetate |
| O.D. | optical density |
| OLB | oligonucleotide labelling buffer |
| OTC | ornithine transcarbamylase |
| PAC | P1 artificial chromosome |
| PCR | polymerase chain reaction |
| pmol | picomole |
| RNA | ribonucleic acid |
| RNAse | ribonuclease |
| RP2 | retinitis pigmentosa 2 |
| sec | second |
| SDS | sodium dodecyl sulfate |
| STS | sequence tagged site |
| SYNI | synapsin I |
| T | thymidine residue |
| Taq | *Thermus aquaticus* |
| TEMED | N, N, N', N'-tetramethylethylenediamine |
| TIMPI | tissue inhibitor of metalloproteinase |
| TK | thymidine kinase |
| TTP | thymidine triphosphate |
| UBEI | ubiquitin-activating enzyme I |
| UHX | ubiquitin C-terminal hydrolase gene |
| UV | ultraviolet |
| V | volts |
| v/v | volume/volume |
| W | watts |

| | |
|---|---|
| w/v | weight/volume |
| YAC | yeast artificial chromosome |
| ZNF21 | zinc-finger protein 21 |

# CHAPTER 1 - INTRODUCTION

The Human Genome Project is a worldwide research effort whose primary focus is to develop detailed genetic and physical maps of the human genome and to determine the complete nucleotide sequence of human DNA (Guyer and Collins, 1995). The information generated in this project will lead to the location of the estimated 50,000 - 100,000 genes within the human genome and will provide the framework for studying how certain DNA variations among humans predispose towards various diseases (Venter *et al.*, 1996). Achieving this goal will enable early detection of disease, diagnosis, effective preventative medicine, efficient drug development, and personalized therapies. The mapping and characterization of genes whose function and/or location are unknown will also aid in the eventual identification of genes involved in additional genetic diseases. Determining the entire sequence of the human genome, although a huge scientific success, has initiated concern for ethical, legal and social implications. For this reason, other goals of the human genome project involve: the identification of issues and the development of policy options to address them; the development of policy options regarding genetic testing; the acceptance of human genetic variation; and the enhancement and expansion of public and professional awareness (Collins and Galas, 1993)

The human X-chromosome contains an estimated 160 million base pairs, and possesses an estimated 5,000 genes (Mandel *et al.*, 1992). A number of clinically important, genetically determined diseases of the eye have been mapped to the proximal region of the short arm of the X chromosome (Nelson *et al.*, 1995). The gene responsible for one form of retinitis pigmentosa, RP3, has recently been identified in this region (Meindl *et al.*, 1996; Roepman *et al.*, 1996a; Roepman *et al.*, 1996b). However, the

causative genes for other eye diseases that have been mapped to Xp remain to be identified. For example, X-linked congenital stationary night blindness (CSNB) (Bech-Hansen and Pearce, 1993; Musarella *et al.*, 1989), retinitis pigmentosa (RP2) (Thiselton *et al.*, 1996) and Åland Island eye disease (AIED) (Alitalo *et al.*, 1991; Glass *et al.*, 1993) have been mapped by linkage analysis to the region Xp21 - p11. Further linkage analysis with affected families and additional polymorphic markers has refined the critical regions of these disorders to the Xp11.3-11.22 region. Additionally, RP2 (Thiselton *et al.*, 1996) and CSNB (Bech-Hansen and Pearce, 1993; Musarella *et al.*, 1989) reportedly overlap with the Xp11.3-p11.23 region.

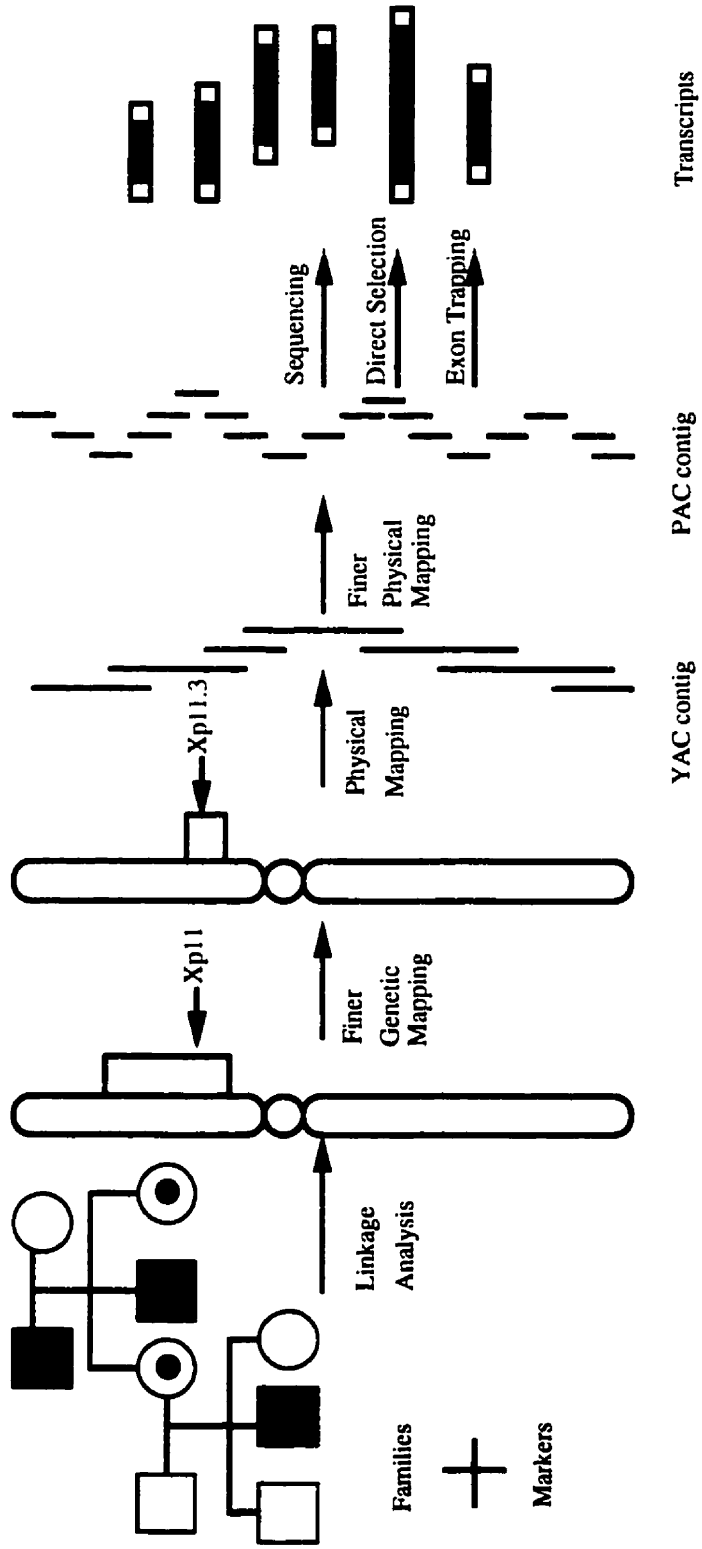
Linkage analyses represents just part of a positional cloning approach used in efforts to locate disease genes whose function is unknown. In fact, positional cloning involves mapping a disease gene to its correct location on a chromosome and assumes no functional information (Collins, 1995). Successive narrowing of a minimal interval eventually results in the identification of the specific gene by the process of excluding more and more genes. Physical mapping of the Xp11.3 region with YACs followed by the development of a PAC contig would facilitate the DNA sequencing of the region and identification of the aforementioned genes. By isolating and characterizing the genes responsible for these diseases, more conclusive diagnosis and, potentially, treatment of these disorders may be possible.

## POSITIONAL CLONING

The first successful application of the technique of positional cloning (see Figure 1; Collins, 1992; Collins, 1995) was reported by Stuart Orkin and colleagues with the cloning of the X-linked gene for chronic granulomatous disease (Orkin, 1986). Since then, approaches based on this technique have become widely applied to gene discovery.

Figure 1.

Figure depicting the approach taken for positional cloning. Positional cloning involves mapping the responsible gene to a region on a chromosome and assumes no functional information. Successive narrowing of the candidate interval eventually results in the identification of a candidate interval that can be sequenced to determine the presence of transcripts (modified from Figure 2 (Collins, 1992)).


Families
Markers
Linkage Analysis
Xp11
Finer Genetic Mapping
Xp11.3
Physical Mapping
YAC contig
Finer Physical Mapping
PAC contig
Sequencing
Direct Selection
Exon Trapping
Transcripts

Positional cloning begins with the collection of pedigrees in which the diseased gene is segregating. These families are studied with multiple polymorphic markers until evidence for linkage is identified (Collins, 1992). Once a disease has been linked to markers on a specific chromosome, additional fine mapping can be applied to narrow the minimal region. The information that one is able to extract from linkage analysis is limited by the number of informative meiotic events (Collins, 1992). Once a minimal genetic region has been defined, the interval can be surveyed for any candidate genes. If no candidate genes are identifiable, transcripts must be identified from the region. Identification of patients who have visible cytogenetic rearrangements can greatly assist the low and high resolution mapping of the responsible gene (Collins, 1992). Many of the positional cloning successes have relied on such rearrangements, and many others have been aided by the identification of patients carrying deletions of tens or hundreds of kilobases (kb) (Collins, 1992).

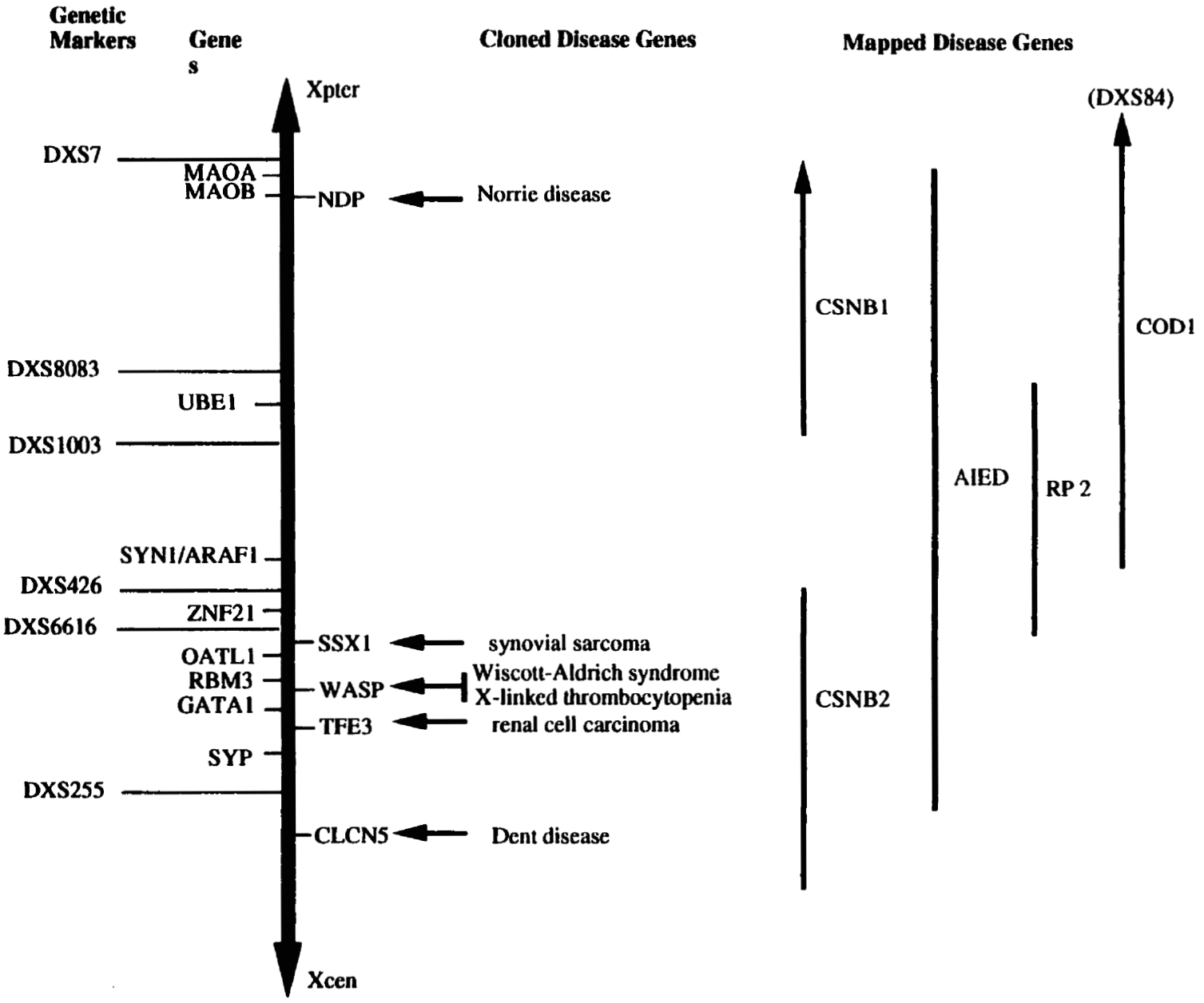
## HUMAN DISEASES MAPPED TO Xp11

Positional cloning has successfully been used to map many diseases to the Xp11 region of the human X chromosome (see Figure 2). Three retinal disorders are of particular interest, namely X-linked congenital stationary night blindness, X-linked retinitis pigmentosa and Åland Island Eye disease.

Retinitis pigmentosa is a group of retinal degenerations characterized by progressive visual field loss, night blindness and pigmentary retinopathy (Aldred *et al.*, 1994). X-linked retinitis pigmentosa (XLRP) is one of the most severe forms of RP, with onset in males in the first decade, progressing to blindness by the third or fourth decade (Berson *et al.*, 1980). There is convincing evidence that many different loci exist on the X-chromosome for XLRP (McGuire *et al.*, 1995; Musarella *et al.*, 1988; Thiselton *et al.*, 1996; Wirth *et al.*, 1988). The RP3 locus has been shown to map to a 1 Mb region

Figure 2.

Figure depicting a gene map of Xp11.3-p11.23. Genetic markers and known genes in Xp11.3-p11.23 are shown to the left of the chromosome. Genes implicated in various genetic disease and candidate intervals for genes involved in ocular diseases are shown to the right.


Genetic Markers
Genes
Cloned Disease Genes
Mapped Disease Genes
Xpter
(DXS84)
DXS7
MAOA
MAOB
NDP
Norrie disease
CSNB1
COD1
DXS8083
UBE1
DXS1003
AIED
RP 2
SYN1/ARAF1
DXS426
ZNF21
DXS6616
SSX1
synovial sarcoma
OATL1
RBM3
WASP
Wiscott-Aldrich syndrome
X-linked thrombocytopenia
GATA1
TFE3
renal cell carcinoma
CSNB2
SYP
DXS255
CLCN5
Dent disease
Xcen

between the markers DXS1110 and OTC, in the Xp21.1 region (Nussbaum *et al.*, 1985; Ott *et al.*, 1990). Recently the identification of a 6.4 kbp microdeletion in an affected patient resulted in the identification of the RP3 gene (Meindl *et al.*, 1996; Roepman *et al.*, 1996a; Roepman *et al.*, 1996b). Another XLRP locus, RP2, has been localized to a 5 cM region between the markers DXS8083 and DXS6616 (Thiselton *et al.*, 1996) within Xp11.3 -p11.23.

Congenital stationary night blindness is a disorder that presents with autosomal or X-linked recessive patterns of inheritance. Clinical features of the X-linked form in affected males usually includes impaired visual acuity, impairment of night vision and, in some cases, myopia and nystagmus. Linkage studies initially mapped CSNB to the Xp11 region near DXS7 and DXS255 (Figure 2) (Bech-Hansen *et al.*, 1990; Gal *et al.*, 1989; Musarella *et al.*, 1989). Further studies have led to the suggestion that X-linked CSNB is heterogeneous with potentially three distinct locations for genes determining this phenotype (Nelson *et al.*, 1995). The laboratory of Dr. Bech-Hansen has mapped one CSNB locus to the proximal side of TIMP1 in Xp11.23 (Bech-Hansen and Pearce, 1993). Two other loci map within the region of interest to this research project. One form has been mapped to the region between markers OTC and DXS1003 (Berger *et al.*, 1995) and another has been mapped proximal to MAOA (Bech-Hansen *et al.*, 1992; Bergen *et al.*, 1994).

Whether CSNB locus is allelic to other eye diseases in this region, such as RP2, RP3 and ÅIED, will only become clear once candidate genes have been isolated and tested for mutations.

Finally AIED, thought by some to be the same condition as X-linked CSNB (Glass *et al.*, 1993; Weleber *et al.*, 1989), is characterized in males by reduced visual acuity, axial myopia and hypoplasia of the fovea (McKusick, 1986). Linkage analysis

has refined AIED to the region between DXS7 (Schwartz and Rosenberg, 1991) and DXS255 (Glass *et al.*, 1993).

## PHYSICAL MAPPING

Retinitis pigmentosa, CSNB, and AIED have all been placed within the Xp11 region based on linkage analysis. The positioning of disease loci on chromosomes is done by genetic linkage analysis in context of a genetic map, followed by the development of a physical map. A physical map of a given genetic region provides information on distances between genes and genetic markers, in terms of DNA, which in turn provides information on a gene's relative position. Obtaining this infromation with regards to the Xp11 region will lead to the eventual identification of the genes responsible for the retinal disorders described above.

As genetic markers, sequence tagged sites (STS) have expedited the process of developing physical maps. STSs are defined as DNA sequences that are operationally unique (meaning the STS can be specifically detected using PCR in the presence of all other genomic sequences) (Olson *et al.*, 1989). A subset of STSs, namely expressed sequence tags (ESTs), are partial cDNA sequences that provide the additional feature of pointing directly to an expressed gene. Olson et al. (1989) STSs could be used as the common landmarks for physical mapping. Overlapping clones would be described by their common STS content to develop a contig of a region. A contig therefore represents the structure of contiguous regions of the genome by specifying the overlapping relationships among a set of clones (Olson *et al.*, 1989). Development of a physical contig provides the framework necessary for the eventual isolation of genes responsible for genetic disease.

One of the greatest advantages of STS markers is that they can be easily, and quickly, detected by the polymerase chain reaction (PCR). PCR is a method that

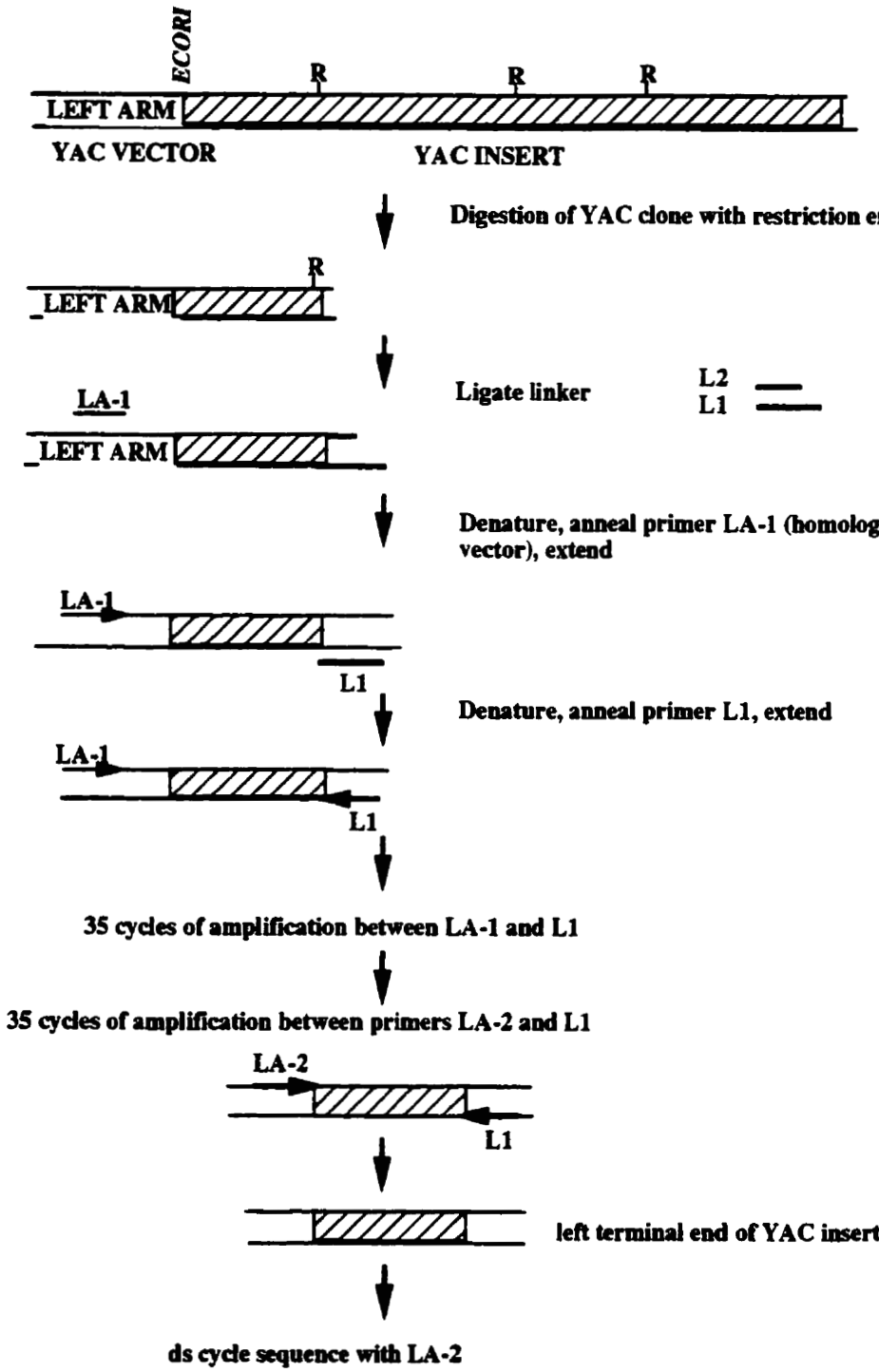
exponentially amplifies DNA sequences between two oligonucleotide primers (Sambrook *et al.*, 1989). By synthesizing PCR primers chosen to be complementary to opposite strands and opposite ends of a tract of DNA sequence, STS markers from any unique DNA sequence can be developed. Existing microsatellite markers, such as short tandem repeats (STRs) or variable number tandem repeats (VNTRs), can act as ready made STSs (Olson *et al.*, 1989).

Terminal STS markers developed from the DNA sequence of an end-fragment of a given clone can be used to screen libraries to obtain other overlapping clones, thus allowing one to 'walk' into unmapped, adjacent regions. Terminal STS markers can be developed by ligation-mediated PCR (end-cloning) (Figure 3) (Kere *et al.*, 1992). End-cloning involves isolating the end portion of the insert in a clone using an appropriate restriction enzyme, ligating a common linker to the cut end of the resulting fragment, followed by amplification of the fragment using a vector-specific primer and a linker-specific primer (Kere *et al.*, 1992). Fragments are then sequenced and analyzed to determined if the sequence has homology to any common repetitive elements in the genome. Resulting unique sequence can then be used to design primer sequences for a marker whose position along the clone of origin is precisely defined, namely the end of the insert, facilitating the orientation of these clones and thus the formatting of the final physical map. Multiple STSs generated by end-cloning allow for the development of a higher density STS map leading to the set goal of an STS map with inter-marker distances of 1 per 100kb (Nelson *et al.*, 1995).

Yeast artificial chromosome (YAC) contigs now span more than 95% of the X chromosome and the remaining gaps are delimited primarily in Xp11.4-p11.3 and Xq13 (Nagaraja *et al.*, 1997). More than 2091 STSs, including over 97 ESTs, have been placed on YACs across the 160 Mb of the X chromosome (Nagaraja *et al.*, 1997). This results in an average inter-STS distance of 75 kb although many regions, such as Xp11.3, are still

Figure 3.

Terminal STS markers are made by ligation mediated PCR (End-cloning) (Kere *et al.*. 1992). The procedure of end-cloning involves isolating the end portion of a clone using an appropriate restriction enzyme, ligating a common linker onto the cut end of the fragment and amplifying the end-fragment using a vector specific primer and a linker specific primer. Figure modified (personal communication; Dr. Kym Boycott, Ph.D. Thesis, 1997).


ECORI
R
R
R
LEFT ARM
YAC VECTOR
YAC INSERT
Digestion of YAC clone with restriction enzyme (R)
R
LEFT ARM
Ligate linker
L2
L1
LA-1
LEFT ARM
Denature, anneal primer LA-1 (homologous to YAC vector), extend
LA-1
L1
Denature, anneal primer L1, extend
LA-1
L1
35 cycles of amplification between LA-1 and L1
35 cycles of amplification between primers LA-2 and L1
LA-2
L1
left terminal end of YAC insert
ds cycle sequence with LA-2

under-represented. Markers DXS7 and ZNF21 span a 7 cM region still lacking physical coverage in Xp11.3 (Nelson *et al.*, 1995).

**LARGE INSERT VECTORS:**

The 7cM region between markers DXS7 and DXS1003 within the Xp11.3 region of the human X chromosome has yet to be represented by a physical map. The mapping and analysis of complex genomes requires techniques for the efficient cloning and manipulation of large DNA fragments. Until recently, cosmids and YACs were the primary tools for physical mapping.

Cosmids are plasmids containing a cohesive-end site (cos site) at each end of the linear molecule (Collins and Hohn, 1978). They have proved invaluable as a physical mapping tool over the past twenty years. One of the principal advantages of the cosmid system is the high efficiency of *in vitro* packaging (Ioannou *et al.*, 1994). The simplicity of DNA isolation (alkaline extraction) is another advantage of this system (Kim *et al.*, 1992). However, cosmids have an insert size restriction of 35-45 kb due to packaging in phage heads of *Escherichia coli* (Feiss *et al.*, 1977; Hohn and Collins, 1980). Also, as many as 35% of human cosmids are inherently unstable and a small fraction of these cosmids are highly unstable and undergo rearrangements and deletions shortly after bacterial transformation (Ioannou *et al.*, 1994). These inherent problems complicate the use of cosmids for the physical mapping of complex genomes.

In spite of their inherent limitations, for the past 10 years YACs have been the standard tool for the construction of chromosome and whole genome contig maps (Foote *et al.*, 1992; Spurr, 1995). They are capable of accepting inserts as large as 1-2 Mb of DNA and have enabled the isolation of DNA segments that were unclonable in conventional bacterial systems (Foote *et al.*, 1992). Although YAC inserts have been found to contain deletions or rearrangements and the frequency of chimeric clones

observed has been as high as 59% in some libraries (Foote *et al.*, 1992), YAC clones were used to construct a first-generation physical map of the whole of the human genome (Schuler *et al.*, 1996). YAC contigs now provide the framework for the future development of contigs using more stable clones.

Despite the important role that cosmids and YACs have played, both of these systems give rise to many unstable clones which undergo *de novo* rearrangements or lose large segments of insert DNA during propagation (Boycott *et al.*, 1996; Kim *et al.*, 1992). Since many mapped large-insert clones are eventually used for detailed gene studies and DNA sequencing, a high fidelity cloning system is imperative. Therefore, due to the limitations in the YAC and cosmid vector systems, alternative large insert vectors such as Fosmids, P1, PACs and BACs have been developed.

Fosmids are low copy number cosmid vectors based on the *E. coli* F factor replicon (*E. coli* fertility plasmid). Replication of the F factor in *E. coli* is strictly controlled (Shizuya *et al.*, 1992). The F plasmid is maintained in low copy number (one or two copies per cell), thus reducing the potential for recombination between DNA fragments carried by the plasmid (Shizuya *et al.*, 1992). Although still restricted in size, fosmids have been shown to stably propagate cosmid sized human DNA inserts (Kim *et al.*, 1992).

Less restricted in size, and still permitting stable propagation, the bacteriophage P1 cloning system permits *in vitro* packaging of P1 vectors containing foreign DNA inserts that are as large as 100 kb (Sternberg, 1990). The P1-based vector has been used in conjunction with a bacteriophage T4 *in vitro* packaging system enabling recovery of recombinants with inserts up to 122 kb (Rao *et al.*, 1992). The P1 system selects for recombinants over non-recombinants and has two P1 derived replication mechanisms (Sternberg, 1990). The single-copy replicon is used for stable clone propagation, while

the multi-copy replicon under control of the inducible *lac* operator, prepares clone DNA (Pierce *et al.*, 1992; Sternberg, 1990).

The bacterial artificial chromosome (BAC) system enables the cloning of large DNA fragments from a variety of complex genomic sources into bacteria. Within the bacteria the DNA is stable, easy to manipulate, and represents a single foreign DNA source (Schmitt *et al.*, 1996). BACs, like Fosmids, are based on *Escherichia coli* and their single copy plasmid F factor (Shizuya *et al.*, 1992). Standard electroporation into *E.coli* gives high-efficiency transformations of about $10^6$ transformants per microgram of DNA (Shizuya *et al.*, 1992). Size distribution of the inserts range from 10 kb to 300 kb with an average size of 100 kb. BACs with inserts exist as supercoiled circular plasmids in *E.coli* permitting easy isolation and manipulation of the large DNA in solution with minimal breaking. The BAC system results in a significantly lower incidence of chimeric clones compared to YACs, since the parA and parB genes involved in the exclusion of extraneous F factors eliminate the possibility of two BACs in a single cell (Shizuya *et al.*, 1992). Also, the co-cloning frequency of BACs has been shown to be significantly lower than that of YACs (Shizuya *et al.*, 1992). Since their introduction, BACs have been used extensively in efforts to map regions and complement existing physical maps (Schmitt *et al.*, 1996).

Lastly, P1-derived artificial chromosome vectors (PACs) are P1 derived vectors (Ioannou *et al.*, 1994). PACs have an average insert size of 130-150 kb (Ioannou *et al.*, 1994) and have a very low reported incidence of chimaerism or insert instability (Nothwang *et al.*, 1997; Osoegawa *et al.*, 1996; Zhao *et al.*, 1997). Since their introduction in 1994, PACs have been used extensively in efforts to map regions and complement existing physical maps. For example, cosmids and PACs were used to develop an integrated map of the Down Syndrome critical region (Osoegawa *et al.*, 1996), and YACs and PACs were used for developing a physical map surrounding the

SCA2 region (Nechiporuk *et al.*, 1996). In addition to their use as a physical mapping tool, PACs offer a system which enables stable propagation of large inserts and the ability to create libraries with low levels of cloning artifacts (Ioannou *et al.*, 1994). These qualities are important for detailed gene structure analysis and DNA sequence analysis.

The considerable burden of characterizing large insert clones, particularly when extended genomic regions are being analyzed, will be greatly reduced by the use of non-problematic clones. The ease of separating the recombinant DNA from host DNA by alkaline extraction procedures may be advantageous for applications such as cDNA isolation by direct selection, exon trapping, and functional studies. YAC contigs provide an STS-based-framework map that allows for the collection of clones from PAC and BAC libraries. The PAC and BAC systems should thus be useful for detailed physical mapping of large genomes, for functional genome characterization as well as for the total genome DNA sequencing effort.

## RADIATION HYBRID MAPPING

Radiation hybrid mapping is a complimentary approach to physical mapping that determines marker order and distance by a statistical analysis of marker distribution in a series of chromosomal fragments generated from one or more chromosomes by X-irradiation (Cox *et al.*, 1990). This approach may be used for the entire genome (Walter *et al.*, 1994) or for a single chromosome (Raeymaekers *et al.*, 1995; Sapru *et al.*, 1994).

By using a panel of radiation hybrids, it has been possible to generate a map of a given region of the chromosome of interest. For example, Raeymakers et al. (1995) describe a radiation hybrid map with 60 loci covering the short arm of chromosome 12 (Raeymaekers *et al.*, 1995). Screening a hybrid panel with primers for a given marker can give insight as to the marker's general location. Based on the information obtained, inferences can be made about the order of markers on a chromosome, since the closer

together two markers are on the chromosome, the less likely it is that a break occurs between them (Shaefer et al, 1993).

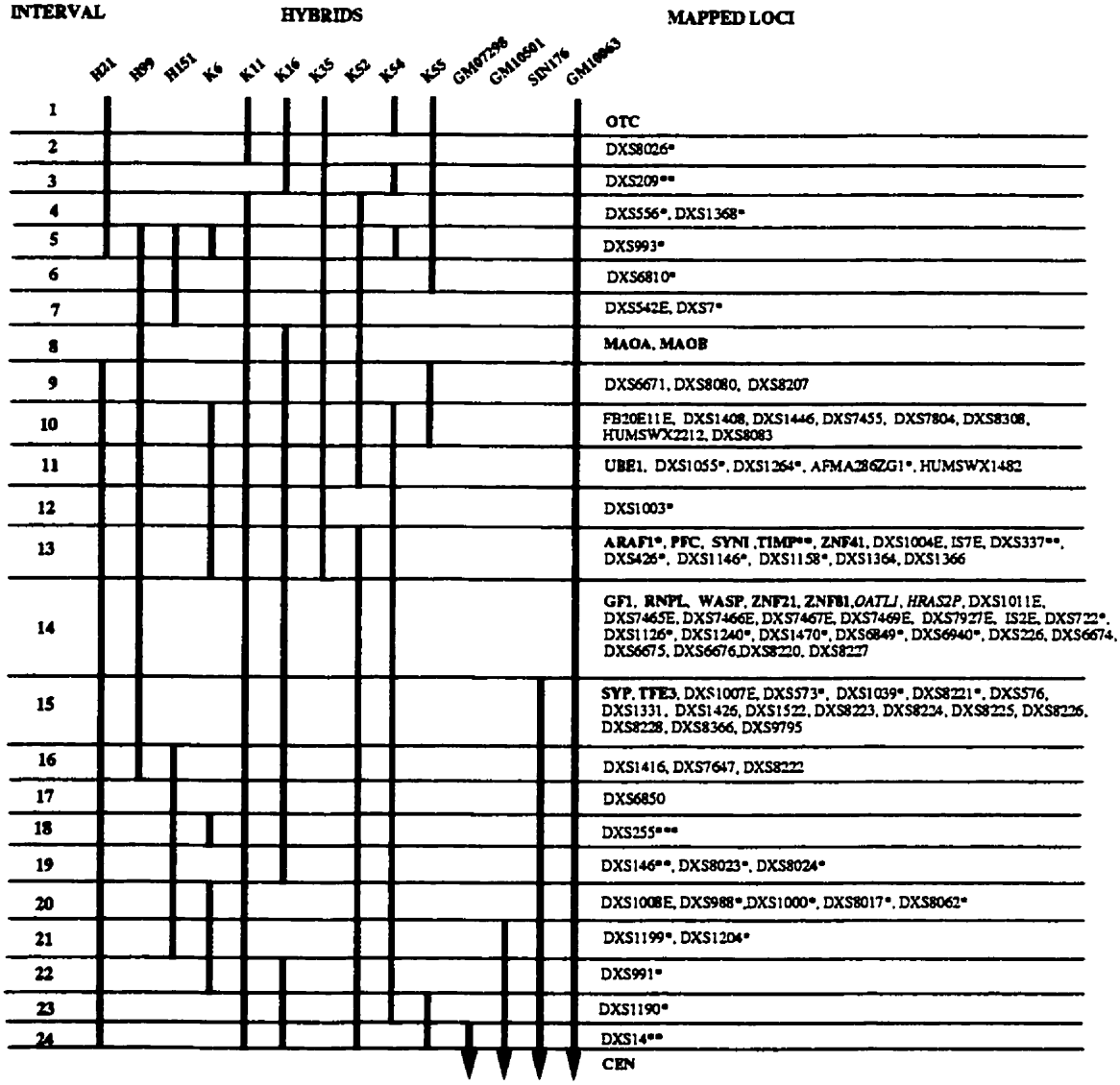
Goss and Harris (1975) developed the original method of irradiation and fusion gene transfer (IFTG) for making radiation hybrid maps (Goss and Harris, 1975). This method could not initially be widely exploited since IFTG was based on using diploid cells and there were very few markers available. However, Cox and co-workers (1992) modified the original approach by using a somatic cell hybrid containing a single human chromosome as the donor cell (Cox, 1992).

Generation of a somatic cell radiation hybrid involves the irradiation of a hamster somatic cell hybrid containing the single human chromosome of interest (Cox, 1992). Following irradiation, the human and hamster chromosomal fragments rapidly join with one another resulting in complex chromosomal rearrangements. DNA fragments from the irradiated hybrid cell are rescued by fusion to a non-irradiated hamster cell under selective conditions, such that only somatic cell hybrids formed by fusion of an irradiated cell with a non-irradiated hamster cell can grow to form viable colonies (Cox, 1992). Although this scheme does not select for the retention of human chromosomal material in the resulting viable radiation hybrids, each hybrid retains between 20-60% of the human chromosome of interest (Cox, 1992). The frequency of retention of the various human chromosomal fragments is relatively constant, irrespective of their chromosomal location, although some variation in retention frequency of individual markers has been observed (Cox, 1992).

Boycott and co-workers described the characterization of a panel of radiation hybrids for the proximal short arm of the X-chromosome (Boycott *et al.*, 1997). This panel of eleven radiation hybrids plus four conventional hybrids divide the p11 region of the X chromosome into 16 intervals wich had an average resolution of approximately 0.8 Mb (Figure 4).

Figure 4.

Detailed characterization of a panel of radiation hybrids for the small arm of the X-chromosome as described by Boycott and coworkers (Boycott *et al.*, 1997). Breakpoints in eleven hybrids and four conventional hybrids divides the p11 region of the X chromosome into 24 intervals giving an average resolution of approximately 0.8 Mb. The types of DNA markers within an interval are grouped based on their retention in the different hybrids together and are distinguished from each other: boldface indicates genes; italics indicates pseudogenes; an 'E' after the marker name indicates EST; ***, **, * indicates VNTR, RFLP, STR, respectively; and remaining markers are STSs. This database will prove invaluable as a resource and for avoiding duplication of effort between labs.


INTERVAL
HYBRIDS
MAPPED LOCI
H21
H99
H151
K6
K11
K16
K35
K52
K54
K55
GM07298
GM10591
SIN176
GM10863
1
OTC
2
DXS8026*
3
DXS209**
4
DXS556*, DXS1368*
5
DXS993*
6
DXS6810*
7
DXS542E, DXS7*
8
MAOA, MAOB
9
DXS6671, DXS8080, DXS8207
10
FB20E11E, DXS1408, DXS1446, DXS7455, DXS7804, DXS8308, HUMSWX2212, DXS8083
11
UBE1, DXS1055*, DXS1264*, AFMA286ZG1*, HUMSWX1482
12
DXS1003*
13
ARAF1*, PFC, SYN1, TIMP**, ZNF41, DXS1004E, IS7E, DXS337**, DXS426*, DXS1146*, DXS1158*, DXS1364, DXS1366
14
GF1, RNPL, WASP, ZNF21, ZNFB1, OATL1, HRAS2P, DXS1011E, DXS7465E, DXS7466E, DXS7467E, DXS7469E, DXS7927E, IS2E, DXS722*, DXS1126*, DXS1240*, DXS1470*, DXS6849*, DXS6940*, DXS226, DXS6674, DXS6675, DXS6676,DXS8220, DXS8227
15
SYP, TFE3, DXS1007E, DXS573*, DXS1039*, DXS8221*, DXS576, DXS1331, DXS1426, DXS1522, DXS8223, DXS8224, DXS8225, DXS8226, DXS8228, DXS8366, DXS9795
16
DXS1416, DXS7647, DXS8222
17
DXS6850
18
DXS255***
19
DXS146**, DXS8023*, DXS8024*
20
DXS1008E, DXS988*,DXS1000*, DXS8017*, DXS8062*
21
DXS1199*, DXS1204*
22
DXS991*
23
DXS1190*
24
DXS14**
CEN

Using somatic cell hybrids carrying single chromosomes, 100-200 clones are required to create a robust map of a single chromosome (Barrett, 1992). Therefore, radiation hybrid mapping of the entire genome would be impractical as a panel of 4,000 clones would be required for complete coverage. A solution to this dilemma involves reverting to the original protocol of Goss and Harris (Goss and Harris, 1975). Instead of using a human-rodent hybrid as a chromosome donor, a diploid human fibroblast is used. A single set of whole genome hybrids produced in this way can be used for mapping the whole human genome (Walter *et al.*, 1994). Gyapay (1996) and co-workers described a panel of 168 whole-genome radiation hybrids, constructed by irradiating donor human fibroblasts with 3,000 rads, which was suitable for high throughput mapping of random markers (Gyapay *et al.*, 1996). This panel constitutes an excellent mapping tool that can be used to localize any STS or EST and should prove a useful resource for positional cloning experiments and physical contig development.

Radiation hybrid mapping is therefore a useful mapping tool which can be used to complement other methods. A consortium of European and US Genome Centres have initiated a radiation hybrid mapping project that aims to build high resolution maps of the human genome. A radiation hybrid database has been developed, which at present, contains over 28,516 entries and can be accessed on the internet (Rodriguez-Tome and Lijnzaad, 1997)

## TRANSCRIPT MAP

Schuler and Boguski (1996) emphasized the importance of the production of a transcript map (also referred to as an 'expression map') of the genome (Schuler *et al.*, 1996). Such a map would provide data on gene density, help estimate the total number of genes and help find exons by alignment with genomic sequences (Schuler *et al.*, 1996).

Also, the rapid identification of coding sequences within large genomic regions would considerably accelerate the isolation of clinically significant genes (Lovett *et al.*, 1991).

Many strategies for the isolation of transcripts have been used including identification of CpG islands, exon trapping, direct selection and zoo blots to detect cross species conservation as described for the isolation of cDNAs for Duchenne muscular dystrophy (Monaco, 1986). CpG islands often act as sign posts for gene locations (Lindsay, 1987), and exon trapping attempts to recover transcribed sequences from cloned mammalian genomic DNA through the functional identification of cis-acting sequences required for RNA splicing (Duyk *et al.*, 1990). Direct selection, used for the identification of the genes responsible for Wilson disease (Bull *et al.*, 1993) and Wiskott-Aldrich syndrome (Derry *et al.*, 1994), is based on the use of a large genomic region to 'fish out' complementary cDNAs.

Primarily through the mapping and sequencing efforts of the Human Genome Project, partial cDNA sequences have been accumulating in public databases at an exponential rate (Strachan *et al.*, 1997). Screening of this Expressed Sequence Database (dbEST) using sequences of mitochondrial housekeeping proteins of yeast has already lead to the discovery of the gene encoding the human mitochondrial RNA polymerase (Tiranti *et al.*, 1997). Major genome centers at Washington University, Stanford, MIT, Oxford, Cambridge, and Paris have been working to map these transcripts to 0.5 Mb intervals using the whole genome radiation hybrid and yeast artificial chromosome (YAC) panel approaches. As many ESTs are derived from yet uncharacterized genes, a high density transcript map will facilitate the identification of candidates for diseased genes and increase the rate of their discovery (Wolfsberg and Landsman, 1997).

## REPETITIVE SEQUENCE

A large proportion of the human genome is comprised of repetitive DNA sequences. Two classes of repetitive sequences include the simple, dinucleotide, repeats and the more complex, Alu or SINE, repeats. Dinucleotide repeats are often polymorphic and can be used to create genetic markers. Complex repeats can be used as sign-posts for 'gene rich' or 'gene poor' regions.

### *Dinucleotide repeats*

By 1984, RFLP linkage maps had been published for the X chromosome with 21 markers spanning 185 cM (Drayna and White, 1985). In 1987, the first human genetic linkage map covering the whole genome was reported consisting of 403 polymorphic markers spaced over the 24 chromosomes with an average spacing of 10-20 cM, though there were several gaps of 50 cM in length (Donis-Keller *et al.*, 1987). The majority of these markers detected polymorphic variation at the DNA level detectable by Southern-blot analysis with radiolabelled probes.

In the past five years, there have been rapid changes in the construction of linkage maps. This has occurred due to the discovery of polymorphic simple repeat sequences, termed microsatellites, which are based on short, mainly dinucleotide, repeats such as $(CA)_n$ or $(TC)_n$ (Litt and Luty, 1989; Weber and May, 1989). In the human genome, there are 50,000 to 100,000 interspersed microsatellite blocks, with the range of repeat units being roughly 2-30. As these repeats are often polymorphic, they can be used as markers. If spaced evenly throughout the genome, this would place them every 30-60 kb (Weber and May, 1989). Although their function, if any, is unknown, it has been speculated that these repeats participate in gene regulation (Hamada *et al.*, 1984) or serve as hot spots for recombination. These DNA markers have become the cornerstone of all

the subsequent high-density genetic maps produced in the past few years (Spurr, 1995). International collaborative efforts have been made to develop whole genome genetic linkage maps. These detailed genetic maps have been used as frameworks for the integration of genetic markers into physical contig maps.

*Complex repetitive elements*

Though repetitive elements are known to be dispersed across the entire genome, some elements have only been observed on specific chromosomes (Willard, 1985). Interspersed repeated sequences are classified as SINES (short interspersed repeated sequences) or LINES (long interspersed repeated sequences) (Singer and Skowronski, 1985). The most common human SINE is the Alu DNA sequence family. It has a consensus sequence of about 300 bp and is reiterated 300,000 to 900,000 times in the human genome (Hwu *et al.*, 1986). Another common SINE is the alpha satellite family. These human DNA repeats consist of adjacent, but diverged, copies of a basic 171 bp or 169 bp unit (Manuelidis, 1978a; Manuelidis, 1978b). Each chromosome is characterized by specific and different subsets of alpha satellite DNA, defined both by restriction enzyme periodicity and by primary DNA sequence (Willard, 1985). Alpha repeats are more commonly found clustered around the centromeric region of all chromosomes (Willard, 1985). The major human LINE sequence family is L1 and has a consensus sequence of about 6.4 kb. The genome is organized such that regions high in GC content correspond with an increased concentration of Alu sequences and expressed genes whereas regions high in AT content are high in L1 sequences and have few genes (Korenberg and Rykowski, 1988).

## PROGRESS OF THE HUMAN GENOME PROJECT

The Human Genome Project is expected to produce a sequence of DNA representing the functional blueprint and evolutionary history of the human species by 2005 (Marshall, 1995). The past five years have focused on genetic and physical mapping (Venter *et al.*, 1996). Now that much of the human genome has been physically mapped, the focus is changing to sequencing (Marshall, 1995) and structural genomics devoted to determining the exact chromosomal positions and structural organizations of our 50,000 to 100,000 genes (Strachan *et al.*, 1997).

The PAC and BAC systems can be applied to convert long-range YAC contigs into higher resolution, less chimaeric contigs through the use of probes derived from the YACs. These contigs, in turn, provide more information on the orientation and order of markers and the templates required for the eventual sequencing of the genome.

Currently, high resolution physical maps have nearly been completed for chromosome 7 (Bouffard *et al.*, 1997) and the X chromosome (Nagaraja *et al.*, 1997). Three additional human chromosomes (16, 19, and 22) have also been mapped, in less detail (Ashworth *et al.*, 1995; Collins *et al.*, 1995; Doggett *et al.*, 1995). Sequencing of these, and other chromosomes, is beginning. Roughly 1.29% of the human genome has been sequenced to date. Over 5500 genes have been mapped to specific cytogenetic locations within the human genome with 388 of these genes being located on the X chromosome. Informal estimates place the world's large scale sequencing capacity at around 20 Mb per year. However, the rate of sequencing is increasing. Summary updates on the Genome Project can be reviewed at the following website: (http://www.ornl.gov/TechResources/Human_Genome/project/timeline.html; August 22, 1997)

In parallel to the Human Genome Project, total gene maps are now available for those organisms whose complete genomic sequence has been determined, including 141 viruses, 51 organelles, two eubacteria, one archeon, and one eukaryote *(Saccharomyces cervisiae)* (Schuler *et al.*, 1996). These model sequences are of interest as putative function of a newly isolated human disease gene may be revealed by its sequence similarity to a well-studied gene in another organism.

## SUMMARY

Physical YAC contigs now span much of the human X chromosome (Nagaraja *et al.*, 1997) and efforts have already begun to collect clones suitable for developing the total sequence of this chromosome (Nagaraja *et al.*, 1997). The region between DXS7 and ZNF21 in the Xp11.3 region of the human X chromosome spans approximately 7 cM and is one of the few remaining gaps on the X-chromosome without physical coverage (Nelson *et al.*, 1995). This region is also of interest because, as previously mentioned, it overlaps with the critical regions for RP2, AIED and CSNB.

This thesis research project extended the existing physical contig of the human Xp11.3 region between a new marker EC7034R and ZNF21 using yeast artificial chromosomes (YACs). Then, using existing markers, an initial set of PACs were collected to begin the construction of a DNA sequence-ready template of the region. Eventually, the sequencing of this region will lead to the identification of all of the transcripts in the Xp11.3 region and potentially to the identification of each of the genes responsible for the genetic disorders which map to this region.

# CHAPTER 2 - MATERIALS AND METHODS

## General Materials and Methods

### *Hybrid Cell Lines*

All somatic cell hybrids used in these studies were part of a hybrid mapping panel characterized in the laboratory of Dr. N. Torben Bech-Hansen (Boycott *et al.*, 1997). The hybrids H21, H99 and H151 were generated by subjecting the hybrid line C12D ($HPRT^-/TK^+$; carrying a human X chromosome in a hamster background) to 20,000 rads of radiation followed by selective fusion rescue with the hamster cell line A23 (J. Knight and P. Goodfellow, personal communication). Hybrids K6, K11, K16, K35, K52, K54 and K55 were produced by subjecting the hybrid line AG9.1 ($HPRT^-/TK^+$; carrying a human X chromosome in a hamster background) to 7,000 rads of radiation followed by fusion rescue with hamster cell line A23 (Boycott *et al.*, 1997). These hybrids were developed by Dr. H. Zoghbi and K. Ellison.

Four conventional somatic cell hybrids with various human X chromosomes with short arm deletions were also used: GM97298, GM10063, GM10501 and SIN176, a human hamster hybrid with del (Xp) (pter-p22.11::p11.23-qter) (Ingle *et al.*, 1985). GM06318B, GM07298, GM10063 were obtained from the NIGMS Human Genetic Cell Repository, and hybrid SIN176 was provided by Dr. A. Zinn (University of Texas Southwestern Medical School, Dallas, Texas)

### *Agarose gel electrophoresis*

Analysis of nonradioactive STS PCR products was carried out in 1.5% SEPARIDE (GIBCO/BRL) / 1% agarose (GIBCO/BRL), with 1X TBE (89 mM Tris, 1 mM boric acid, 2 mM EDTA) running buffer or in 1.5% Nusieve GTG agarose (FMC BioProducts) / 1% agarose (GIBCO/BRL) with 1X TAE (40 mM Tris, 20 mM acetic

acid, 10mM EDTA) running buffer. Ethidium bromide was added to the molten gel prior to casting for a final concentration of 0.5 μg/mL. Before loading, PCR reactions were mixed with loading dye (final 1X, 5% glycerol and 0.035% orange G). PCR products were electrophoresed for 1 hr at 80 V on a 10 cm gel tray. pBluescript cut with *HaeIII* was routinely used as the DNA marker standard. DNA was then visualized with an ultraviolet transilluminator set on 302 nm and photographed with Polaroid Type 52 film using a red filter.

Electrophoresis of PAC DNA digests was carried out in 0.8% agarose, with 1X TBE running buffer. Ethidium bromide, to make a final concentration of 0.5 μg/ml, was added to the molten gel prior to casting. Digests were electrophoresed at 40 V for 16 hrs on a 25 cm gel tray.

*DNA transfer (Southern transfer)*

DNA was transfered onto Hybond - $N^+$ nylon membranes (Amersham) using a vacuum transfer system (Tyler Research Instruments, Edmonton, Alberta). To facilitate this transfer, high molecular weight DNA (from PACs or YACs) contained in the gel was first depurinated in 0.25 N HCl with gentle shaking for 15 min. DNA kept denatured during the transfer by keeping the gel immersed in 0.4 N NaOH. Sixty cm of water vacuum was applied to the system for 90 mins. Following the transfer, filters were neutralized in 2X SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0), blotted on 3MM Whatman paper and heat-sealed in a Seal-a-Meal bag (Dazey).

*Radiolabeling of DNA*

DNA probes were labelled using the random hexanucleotide primer method (Feinberg and Bogelstein, 1984). Unique or human genomic probes were labelled using 150 ng DNA in 50 μL reactions. DNA combined with sterile distilled water to a volume

of 30 µL was denatured by boiling for 5 mins and then quenched on ice before adding 13 µL OLB (195 mM Tris pH 8.0, 780 mM HEPES, 40 mM beta-mercaptoethanol, 1.87 mg/ml BSA, 19.5 mM $MgCl_2$, 109 µM of each of dATP, dGTP, dTTP, and 716 µg/ml hexanucleotides) 2 units of the Klenow fragment of DNA polymerase I and 50 µCi [$^{32}$P] dCTP at a specific activity of 3000 Ci/mmol. The reaction was vortexed, spun down and allowed to sit for 1.5-2.0 hrs at room temperature.

Labelled DNA was separated from unincorporated nucleotides using MicroSpin G-50 Columns (Pharmacia Biotech). The eluant was then boiled for 5 mins to denature the $^{32}$P-labelled DNA prior to hybridization.

*Hybridization conditions/stripping*

Filters to be probed were incubated at 65°C for 1.5-2 hrs in bags containing prehybridization buffer (0.5 M sodium phosphate: 1 mM EDTA, 1%w/v BSA, 5% W/V SDS, 30% formamide). 267 µg/µL salmon sperm DNA was included in the prehybridization solution and 100 µg/ml in the hybridization solution to block non-specific binding of the probe to the filter. Following the incubation period, the prehybridization buffer was removed from the bags. Hybridization buffer with 50 mg/ml dextran sulfate and the labelled probe were then added to the bag, which was then sealed and incubated at 55°C for 16-20 hrs.

The following day, filters were removed from their bags and washed twice in 2X SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0) / 0.1% SDS for 15 min per wash. Filters were then dried briefly on 3 MM Whatman paper, wrapped in plastic wrap (Sealwrap) and exposed to Kodak XAR 5 film with Dupont Quanta III intensifying screens at -70°C. Films were developed in a RGII FUJI X-Ray Film Processor.

Before filters could be re-probed, probe was removed from filters by washing the filters twice for 25 min in boiling water with 0.03 X SSC and 0.5% SDS. Filters were

tested with a radiation detector (Berthold) and/or exposed to film to ensure the majority of traces of the probe had been removed prior to being dried and probed again as described above.

*PCR*

Non-radioactive PCRs were used for screening the hybrid panel and for assessing STS content of YACs. Reaction mixes of 5μL were set up using the TNK buffer system (Blanchard *et al.*, 1993) (10 mM Tris pH 8.3, 1.5 mM $MgCl_2$, 5 mM $NH_4Cl$, 100 mM KCl) 100 ng DNA (human genomic hybrid or YAC DNA), 0.4 μM of each primer, dNTP mix (100 μM of each dATP, dCTP, dGTP and dTTP) and 0.5 units of *Taq* DNA polymerase (GIBCO/BRL).

Radioactive PCR was carried out in 20 μL reactions using 50 ng DNA (human genomic, hybrid or YAC), 300 μM MgCl buffer (1 M KCl, 1 M Tris pH 8.3, 10 mg/ml BSA, Tween-20, NP-40 and 1.5 mM $MgCl_2$), dNTP mix (20 μM of each dATP, dCTP, dGTP and dTTP), 0.25 μM of each primer, 3,000 Ci/mmol (Amersham) and 0.5 units *Taq* DNA polymerase (GIBCO/BRL).

All PCR reactions were overlaid with paraffin oil and subjected to 7 min of denaturation at 95°C in the thermocycler (Perkin Elmer 480, or MJ Research DNA Engine) prior to cycling. Nonradioactive STS PCR reactions involved 35 cycles of 30 secs denaturation at 94°C, 45 secs annealing at 55°C and 45 secs elongation at 72°C. Radioactive PCR reactions underwent 27 cycles of 1 min denaturation at 94°C, 2 min annealing at 55°C and 1 min elongation at 72°C. Once all cycles were completed, a final elongation step at 72°C for an additional 7 min was executed.

*Preparation of DNA size markers*

Bluescript size markers were prepared by digesting 100 µg pBluescript DNA with 200 units of *HindIII* or *HindIII / EcoRI* in the specified react buffer system. Samples were then incubated at 37°C for 1 hr. In order to be ready for loading, samples were then combined with low TE and loading dye and stored at -20°C.

Marker for determining YAC size was prepared from YAC DNA YP149 as described for making agarose plugs in the 'Analysis of YACs' section.

Radiolabelled size standard marker was prepared as described in the Sequenase Version 2.0 DNA sequencing Kit. One µg M13 DNA was sequenced using [$^{35}$S] dATP as described.

*Isolation of DNA from agarose for sequencing*

To obtain a clean template for DNA sequencing, end-clone DNA was agarose gel purified. Initially, DNA was isolated by the GELase procedure. It was found that isolation of DNA was more successful using DEAE paper, so later experiments used this method.

DNA isolated using the GELase procedure was initially run in a 1% agarose (GIBCO/BRL), 1.5% SEPARIDE (GIBCO/BRL) gel. After having run for 1hr at 50 V, bands were run into 1% low melting point agarose (GIBCO/BRL), cut out and weighed. Gel slices were melted at 70°C with 1 µL of 50X GELase buffer (1X buffer consists of 40 mM Bis-Tris-pH 6.0 and 40 mM NaCl) per 50 mg of gel then equilibrated at 45°C for 10 min. After melting, 0.5 units of GELase enzyme (Cedarlane Laboratories) was added per 200 mg of gel and digested for 2 hrs. After one hour of digestion, reactions were boosted with half the initial GELase volume. One hour later, the DNA samples were ethanol precipitated.

Ethanol precipitations involved adding one volume 5 M ammonium acetate and 2 volumes 95% ethanol and centrifuging at 10,000 rpm for 30 min. Supernatants were then removed and the pellets were washed with 70% ethanol. Samples were centrifuged briefly at 10,000 rpm, the ethanol was removed and the pellets were allowed to dry before resuspending them in 40 µL double distilled water.

For DNA isolated using a DEAE membrane (Schleicher & Schuell NA45), samples were run in a 1% agarose, 1.5% Sepharide gel at 50 V for 2 hrs. Bands were then visualized under ultraviolet light (short wave UV-254 nm) and agarose behind the bands was cut away to reduce the chance of multiple bands being isolated. The gel was then allowed to run for an additional 30 min at 50 V. After electrophoretic separation in the gel, a strip of DEAE paper was placed in an incision just ahead of the band of interest. Electrophoresis was continued until binding of DNA to the membrane was completed as judged by visualization of ethidium bromide fluorescence using a hand held short wave UV-254 nm light. The DEAE strips were rinsed with buffer to free them of residual agarose. The bound DNA was eluted into 500 µL DEAE buffer (1.0M NaCl, 0.1 mM EDTA, 20 mM Tris pH 8) by incubating strips in buffer at 65°C for 45 min. The strips were removed from the buffer and the buffer was subjected to extraction with phenol followed by chloroform. DNA in the buffer was precipitated by adding 2.5 volumes of isopropanol and centrifuging at 10, 000 rpm for 30 min. Pellets were washed with 70% ethanol and dried prior to resuspending in 40 µL water.

*dsDNA cycle sequencing*

YAC and PAC end-clones were sequenced using the GIBCO/BRL dsDNA Cycle Sequencing System (GIBCO/BRL). Prior to the sequencing reactions, primers were end-labelled with [$^{33}$P] ATP or [$^{32}$P] ATP in a reaction consisting of 0.25 µM primer (left or

right internal vector primer ), 5X Kinase buffer (300 mM Tris-HCl (pH 7.8), 50 mM MgCl2, 1M KCl), T4 kinase (0.2 units) and [$^{33}P$ or $^{32}P$] ATP.

Each sequencing reaction contained approximately 10 ng DNA, labeled primer, Taq sequencing buffer (300 mM Tris-HCl (pH 9.0), 50 mM MgCl2, 300 mM KCl, 0.5% (w/v) W-1), Taq DNA polymerase (0.5 units) and one of four termination mixes: A (2 mM ddATP and 100 μM each dATP, dCTP, 7-deaza dGTP, and dTTP), C (1 mM ddCTP and 100 μM each dATP, dCTP, 7-deaza-dGTP, and dTTP), G (0.2 mM ddGTP and 100 μM each dATP, dCTP, 7-deaza-dGTP, and dTTP), or T (2 mM ddTTP and 100 μM each dATP, dCTP, 7-deaza-dGTP, and dTTP). Reactions were covered with silicone oil and subjected to 20 cycles of 30 secs at 95°C, 30 secs at 55°C and 60 secs at 72°C followed by 10 cycles of 30 secs at 95°C and 60 secs at 72°C. Reactions were then terminated with Stop solution (95% (v/v) formamide, 10mM EDTA (pH 8.0), 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol) and run on an acrylamide gel as described.

*Electrophoresis of sequencing reactions and radioactive PCRs*

Radioactive PCR reactions and sequencing reactions were run on 31.0 cm X 38.5 cm denaturing polyacrylamide gels using a model S2 apparatus (GIBCO/BRL). A 75 ml acrylamide solution [8 M urea, 1X TBE, 6% acrylamide (from a 40% acrylamide:bis acrylamide (38:2) stock solution)] was made and left mixing over low heat until the urea had completely dissolved. The mixture was then filtered through a 0.22 micron Millipore filter attached to a 60 ml syringe. Prior to casting the gel, a fresh 10% solution of ammonium persulfate dissolved in water was added to the mixture to make a final concentration of 0.1% APS along with 0.032% v/v TEMED. The mixture was mixed by inverting and cast between two glass plates with 0.4 mm spacers using the Otter-Owl Scientific gel casting apparatus. Two 25-well sharktooth combs for sequencing or one well-forming comb for analysis of PCR products were inserted at the top of the gel before

leaving the gel to polymerize for 30 min. The gel was then loaded in 1X TBE running buffer and left to run at 70 watts for the duration of the run. Sequencing gels were run for 1.5 hrs, loaded again and allowed to run for another 1.5 hrs. PCR reaction gels were run for 1.5-6 hrs depending on the size of product. Once the runs were completed, gels were transfered to Whatman 3MM filter paper, covered with plastic wrap and dried on a Bio-Rad model 583 Gel dryer. Dried gels were exposed to Kodak XAR-5 film overnight at room temperature.

*DNA quantitation*

To determine the concentration and purity of DNA, spectrophotometric readings were taken at 260 nm and 280 nm with a Beckman DU-65 spectrophotometer. A ratio between the optical density readings at 260 nm and 280 nm (OD260/OD280) of 1.8 was taken to indicate a sufficiently clean DNA preparation. For calculating DNA concentrations, it was assumed that 1 O.D. at 260 nm corresponded to 50 µg/mL of double-stranded DNA and 34 µg/ml of single-stranded DNA.

*Recovery of DNA from agarose slices for probe preparation*

PCR products were pooled and then separated on a 1% agarose, 1.5% SEPARIDE gel for 30 min at 70 V, then run into 1% low melt agarose. DNA bands were cut out and heated at 65°C to melt. Heated samples were then placed on dry ice/EtOH to cool the melted agarose. Tubes were centrifuged for six min and the supernatant was harvested immediately to prevent agarose from reabsorbing the liquid. DNA in the supernatant could then be used for labelling by random priming. This method was provided complements of Dr. Doug Demetrick (University of Calgary).

*Restriction enzyme digest*

Genomic DNA, normally 5 μg, was digested with 40 units of enzyme (NotI, EcoRI or BamHI). The manufacturers buffer system was used and sterile distilled water was used to bring reactions to the proper volume. Digests of YAC and PAC DNA were allowed to proceed for two hours and genomic DNA for four hours with an addition of 20 units of enzyme after two hours of incubation. Incubations were run at 37°C. Orange G loading dye (1X final concentration) was added to each reaction prior to electrophoresis.

**Isolation of Large Insert DNA Clones**

In a collaboration between Dr. Bech-Hansen's laboratory and The Center for Genetics in Medicine (Washington University, St. Louis, MI) a series of seven YAC libraries (A, B, C, E, F, I, M) were screened to identify YACs for this project. Libraries A, B and C were derived from a 46 XY lymphoblast cell line, E from 49 XXXXX cell line, and F from a Xpter-Xq27.3 hamster-human somatic hybrid cell line (Nagaraja *et al.*, 1994). Libraries I and M are the ICI YAC library (Nagaraja *et al.*, 1994) and the Mega CEPH Library (Cohen *et al.*, 1993), respectively.

In this collaboration with Drs. David Schlessinger and R. Nagaraja (St. Louis), screening was carried out by PCR, with robotic assistance, using a uniform temperature regimen and the TNK buffer system. Isolated YAC clones were sent as stabs in YAC trp-, ura- selective agar (pH 7.0) (0.7% w/v yeast nitrogen base without amino acids, w/v 2% glucose, w/v 0.005% adenine sulphate, w/v 0.005% tyrosine, w/v 1.4% vitamin assay casamino acids and 1.5% w/v agar).

*Analysis of YACs*

Upon receipt in Calgary, YACs were streaked on YAC selective agar plates ($trp^-$, $ura^-$) and incubated at 30°C for 48 hrs. Single clonal isolates were inoculated into 5 ml

of YAC media (YAC trp-, ura- selective agar without agar) and grown aerobically at 30°C for 48 hrs. Fifty percent glycerol stocks were prepared with 500 µL of each culture and stored at -70°C. The remaining culture was used to prepare four intact yeast chromosomal DNA plugs.

To prepare such plugs, YAC cultures were pelleted for 15 min at 2700 rpm, washed in 100 mM EDTA pH 8.0 and then re-pelleted. Pellets were then resuspended in CPES buffer pH 5.6 (20 mM citrate-PO4 buffer, 50 mM EDTA pH 8.0, 0.9 M Sorbitol) supplemented with 5 mM DTT. A solution of 8 mg/mL yeast lysing enzyme (10 mg/ml from *Trichoderma harzianum* in 10 mM sodium phosphate pH7.5 and 50% glycerol) in CPE buffer pH 5.6 (20 mM Citrate-PO4 buffer, 50 mM EDTA pH 8.0) and 4.0% molten LMP agarose in CPE buffer was combined with the sample and each sample quickly dispensed into four chilled blocks. The agarose plugs were chilled for 30 min and then dispensed into CPE buffer (four blocks per sample) and incubated for one hour. CPE buffer was then removed from the plugs and solution A (0.5 M EDTA, 1% sodium sarkosine, 0.5 mg/mL proteinase K) was added and the plugs incubated at 50°C for 48 hrs. Solution A was then be removed and four plugs per sample were stored in 0.5 M EDTA at 4°C. Blocks were then ready to be loaded directly into gels for pulsed-field gel electrophoresis (PFGE) for analysis of the YAC DNA.

PFGE was used to determine size and integrity of the human DNA inserts in YAC clones. Individual agarose plugs were loaded into one of 20 wells of a 1% agarose gel (0.5 X TBE ) in a 14 cm X 20 cm gel tray. Once loaded into wells, plugs were covered with molten 1% low melting point agarose in 0.5 X TBE. Gels were run on a CHEF-DR III System in 0.5X TBE running buffer at 14°C for 28 hrs with initial and final switch times of 3.0 sec and 30 sec respectively, a field strength of 6.0 volts/cm and an included angle of 120°. Size estimation of the YACs was based on co-migration of the *Saccharomyces cerevisiae* marker chromosomes (YP149). Following electrophoresis,

gels were stained with EtBr (0.5 μg/mL) for 30 min, visualized with an ultraviolet transilluminator set on 302 nm and photographed with Polaroid Type 52 film using a red filter.

Yeast chromosomal DNA was then transferred from the gel onto a Hybond-$N^+$ membrane using Southern transfer and was detected by hybridizing the filters with human $^{32}P$ labelled genomic DNA as described in the 'General Materials and Methods' section .

One colony, usually the clone with the largest DNA insert, was inoculated from the glycerol stock into YAC selective media and grown aerobically at 30°C for 48 hrs. Cells were pelleted at 2700 rpm for 15 min, washed with 100 mM EDTA pH 8.0 and pelleted again. Cells were then lysed by incubating at 37°C for one hour with 20 μg/ml yeast lysing enzymes in 1 ml CPES buffer. Cells were then pelleted gently and the supernatant removed. DNA extraction buffer (10 mM Tris-HCl pH 8.0, 0.1 M EDTA pH 8.0, 0.5% SDS, 20 μg/ml RNAse A) was added to the cells and allowed to incubate for 1 hour at 37°C. Subsequently, 100 μg Proteinase K was added followed by an additional incubation for 16 hrs at 50°C to remove protein. Samples were purified by subjecting them to three extractions with phenol/$CHCl_3$:isoamyl alcohol (25:24:1) and one final $CHCl_3$/isoamyl alcohol (24:1) extraction and DNA was subsequently precipitated by adding 10% volume 4 M NH4OAc, 2 volumes 95% cold ethanol and placing at -70°C for 15 min. Samples were spun for 40 min at 4°C at 10,000 rpm. The supernatants were removed and the pellets washed with 70% ethanol. Pellets were resuspended in 150 μL low TE (10 mM Tris.Cl pH 8.0, 1 mM EDTA pH 8.0). Yeast DNA preparation typically yielded 75-100 μg DNA per 5 ml culture.

*DNA isolation from P1 artificial chromosome (PAC) clones*

PAC clones were isolated as described by Ioannou and co-workers (Ioannou *et al.*, 1994). Single bacterial colonies were inoculated into 2 ml LB media supplemented with 25 μg kanomycin and grown for 16 hrs at 37°C. Half a millilitre of the resulting culture was combined with 300 μL sterile glycerol and stored at -70°C for future use. Remaining cells were transfered to 1.5 ml microfuge tubes and centrifuged for 10 min at 3000 rpm. Pellets were resuspended in 300 μL P1 solution (15 mM Tris pH 8, 10 mM EDTA, 100 μg/ml RNaseA), and 300 μL P2 solution (0.2 N NaOH, 1%SDS) was added and the samples left at room temperature for 5 min. Slowly 300 μL P3 solution (3 M KOAc pH 5.5) was added and the samples were placed on ice for 5 min. Samples were then centrifuged at 10,000 rpm for 10 min at 4°C. All steps to follow were performed at 4°C to enhance the DNA isolation. Supernatant was then transfered into tubes containing ice-cold isopropanol, samples were mixed by inversion and left on ice for 5 min. Samples were centrifuged for 15 min at 4°C, the supernatant was discarded and pellets washed with 70% EtOH and dried before resuspending them in 40 μL water.

PAC DNA was subjected to enzyme digestion as described in 'General Materials and Methods'. For sizing, PACs digested with *NotI* were run out on a pulsed-field gel. Otherwise, PACs digested with *EcoRI* or *BamHI* were subjected to standard agarose gel electrophoresis as described in 'General Materials and Methods'.

PACs digested with *NotI* were run with a low range PFG marker (New England Biolabs) for 16 hrs with an initial and final switch time of 6.75s and 13.58s respectively, field strength of 6 V/cm and an included angle of 120. Gels were stained with EtBr, photographed and transfered as described in 'General Materials and Methods'.

*End clone STSs*

The ends of YACs or PACs were isolated by a ligation-mediated PCR method (Kere *et al.*, 1992). Such end-fragments were DNA sequenced to develop new STS markers to facilitate 'chromosome walking' (Figure 3). One hundred ng of PAC or YAC DNA was digested for 2 hrs at 37°C with 5 units of enzyme (*AluI, PvuI, EcoRV, RsaI* or *ScaI*) and the recommended reaction buffer. Once the YACs or PACs had been cut, linker (2 pmol L1 and 2 pmol L2) was ligated onto the ends of fragments. The ligation reaction, which consisted of the digestion products, linker, 20 units T4 ligase and 5X ligation buffer (as specified by the manufacturer - New England Biolabs), was incubated at 16°C for 16 hrs. The ligation mix was diluted with 60 µL ddwater, denatured by heating for 10 min at 95°C, quenched on ice and subject to a first round of DNA amplification using left or right external vector primers (Table 1). Two microliters of the ligated mixture was amplified by PCR in 10µL reactions using the linker primer and a YAC vector arm specific primer (5 pmol vector primer, 0.5 pmol linker L1). Temperature cycling conditions were 94°C for 30 secs, 65°C for 45 secs, and 72°C for 90 secs for 35 cycles using TNK-50 buffer system (10 mM Tris pH 8.3, 1.5 mM $MgCl_2$, 5 mM $NH_4Cl$, 50 mM KCl), dNTPs (125µmol of each dATP, dCTP, dGTP and dTTP) and 0.25 units Taq. The reaction mixture was then diluted with 150 µL water and 5 µL of the mixture was reamplified using linker primer and internal YAC vector arm-specific primer under identical conditions. The reamplification products were purified on a low-melting-point gel, excised and sequenced by double stranded Cycle Sequencing as described in 'General Materials and Methods'.

**Table 1. Primer Sequences used for End-clone Generation**

| Primer Name | Primer Sequence 5'-3' |
|---|---|
| YAC Left Vector Primer | CACCCGTTCTCGGAGCACTGTCCGACCGC |
| YAC Left Vector Primer Internal | TCTCGGTAGCCAAGTTGGTTTAAGG |
| YAC Right Vector Primer | ATATAGGCGCCAGCAACCGCACCTGTGGCG |
| YAC Right Vector Primer Internal | TCGAACGCCCGATCTCAAGATTAC |
| | |
| PAC Left Vector Primer | CATTGTAGGACTATATTGCT |
| PAC Left Vector Primer Internal | TAATACGACTCACTATAGGG |
| PAC Right Vector Primer | TAGACGGCAAAGCTAGGAGG |
| PAC Right Vector Primer Internal | CGGCAGCTGTAAATCCACTG |
| Linker Primer L1 | GCGGTGACCCGGGAGATCTGAATTC |
| Linker Primer L2 | GAATTCAGATC |

[a] primer sequences obtained from PAC vector sequence (Ioannou *et al.*, 1994), or as published for YAC end-cloning (Kere *et al.*, 1992).

DNA sequences determined from YAC and PAC end-clones were analyzed using the Blast search program (Altschul *et al.*, 1990) of the National Center for Biotechnology Information (NCBI) accessed through email (blast@ncbi.nlm.nih.gov). Sequence homology ratings over 100 indicated that the sequence being analysed potentially contained a high degree of similarity with other sequences in the database. This information was critical when designing primers since designing primers depends on unique sequence.

DNA sequences that were determined to be unique by Blast analysis were used to design primer sequence using the software program Primer Designer - (Version 1.01, Scientific and Educational Software). Twenty bp primers (forward and reverse) were

designed with a percent GC content between 40-60%, melting temperatures between 55-72 $^{o}$C and no runs longer than four identical bases. Resulting markers were ideally 100-300 bp in size.

Oligonucleotide primers were synthesized using a Beckman Oligo 1000 DNA synthesizer. Oligonucleotides were then eluted from the columns and lyophilized as instructed in the Beckman's UltraFast cleavage and deprotection kit. Lyophilized oligonucleotides were resuspended in 500 µL sterile distilled water and quantitated as described in 'General Materials and Methods'.

# CHAPTER 3 - RESULTS

Development of an extended physical contig in the Xp11.3 area was tackled in two stages. The first stage involved the positioning of existing markers, the creation of new STS markers to increase the density of markers in the region and the identification and characterization of YAC clones to develop a YAC contig. The second stage involved using existing and new markers to seed the target region with PACs. These PACs were used to orient some of the markers whose relative placement could not be determined by the analysis of YAC data.

## STS Markers

Earlier studies in the laboratory of Dr. N. Torben Bech-Hansen resulted in the characterization of a panel of radiation hybrids for the short arm of the human X-chromosome (Boycott *et al.*, 1997). This panel of eleven hybrids plus four conventional hybrids divides the p11 region of the X chromosome into 24 intervals and has an average resolution of approximately 0.8 Mb (Figure 4 in Introduction, pg.19). The genetic region between DXS8083 and ZNF21 spans five intervals, 10 to 14 inclusive, on the mapping panel. An initial set of 19 markers were sublocalized to Xp11 based on their retention in conventional and radiation reduced hybrids using PCR (Table 2). The presence of STS marker DXS7804 on hybrid DNA H21, H99, K6, K11, K16, K35, K52, K54, K55 and GM10063 illustrates the application of this panel to the mapping of STS markers (Figure 5). This retention pattern is consistent with DXS7804 being localized to interval 10.

Subsequent to the analysis of the initial collection of markers, 19 new region-specific STSs/STRs/ESTs were obtained or developed to increase DNA marker density. Three ESTs (Table 3) were obtained from published sources and from the internet, and 15 STSs, including one STR (EC8058), were designed from selected end-fragments of

Table 2. DNA markers used for initial construction of physical YAC contig in Xp11.3

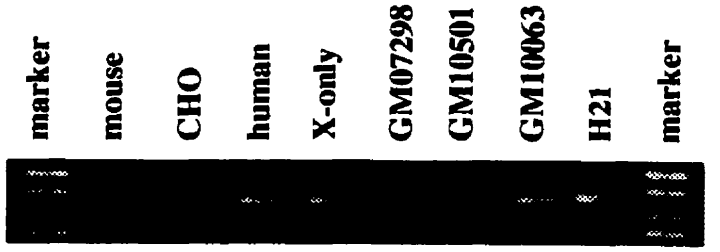
| Marker | Hybrid Interval | Description[a] | Source |
|---|---|---|---|
| DXS6849 (ZNF21) | 14 | STR | Huebner *et al.*, 1991 |
| DXS426 | 11 or **13**[c] | STR | Luty *et al.*, 1990 |
| SYN | 11 or **13** | EST | Derry and Barnard, 1992 |
| DXS1266 | 11 or **13** | STS | Coleman *et al.*, 1994 |
| DXS1146 | 11 or **13** | STR | Hong *et al.*, 1993 |
| DXS1364 (A1.72)[b] | 11 or **13** | STR | Boycott *et al.*, 1997 |
| AFMA286ZG1 | 11 or **13** | STR | Weissenbach *et al.*, 1992 |
| UBE1 | 11 or **13** | EST | Lafreniere *et al.*, 1991 |
| DXS1003 | 12 | STR | Weissenbach *et al.*, 1992 |
| DXS 1366 (A1.78)[b] | **11** or 13 | STR | Boycott *et al.*, 1997 |
| DXS1264 | **11** or 13 | STS | Coleman *et al.*, 1994 |
| DXS1055 | **11** or 13 | STR | Gyapay *et al.*, 1994 |
| sWXD2212 | 10 | STS | GDB |
| DXS8308 | 10 | STS | GDB |
| DXS1408 | 10 | STS | GDB |
| DXS7455 | 10 | STS | GDB |
| DXS7804 | 10 | STS | GDB |
| FB20E11 | 10 | EST | GDB |
| DXS1446 | 10 | STS | GDB |

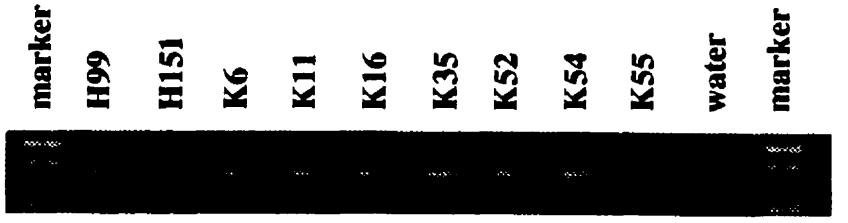
[a] STS - sequence tagged site; EST - expressed sequence tag; STR - sequence tagged repeat.

[b] laboratory designations

[c] bold face indicates interval marker was mapped to by STS analysis of YAC and PAC clones.

Figure 5. Somatic cell hybrid panel mapping of marker DXS7804. Retention of DXS7804 on a set of 10 radiation-reduced hybrids and four somatic cell hybrids was evaluated using PCR as described in the 'General Materials and Methods' section. Names of hybrids, control lanes human genomic, X-only, Chinese hamster (CHO), mouse and dd$H_2O$, are indicated along the top. PCR products, of 160 bp in size, were separated on 2% Nuseive agarose/1% agarose gels and visualized by EtBr staining as described in the 'General Materials and Methods' section.


marker
mouse
CHO
human
X-only
GM07298
GM10501
GM10063
H21
marker


marker
H99
H151
K6
K11
K16
K35
K52
K54
K55
water
marker

YACs or PACs (Tables 7 and 9). One additional STR, DXS8083, was obtained from a recent publication (Thiselton *et al.*, 1996). Due to difficulties in amplification efficiency, new primers were designed from an existing sequence in Genbank (Table 4).

Table 3. Additional ESTs obtained from public sources during study

| Name | Size (bp) | Source |
|---|---|---|
| WI18364 | 104 | GDB |
| UHX | 140 | Swanson *et al.*, 1996 |
| PCTK1 | 400* | Carrel *et al.*, 1996 |

* PCR product is 400 bp when amplifying genomic DNA.

Table 4. Re-designed primer pair for polymorphic STR DXS8083

| Marker | Primer Sequence | Size (bp) | Genbank Accession No.* |
|---|---|---|---|
| DXS8083 | CTTCTGCACAGCAAAGGAAA<br>ACTCCAGGAGGCCGTATGTC | 232-248 | Z54041 |

* Published DNA sequence of this locus.

Whereas most of the markers studied were uniquely localized in Xp11.3, analysis of the retention pattern of marker DXS8083 on the radiation hybrid panel suggested that these primer sequences were duplicated in the genome. Amplification of the hybrid mapping panel with DXS8083, resulted in two bands being observed in the X-only hybrid and genomic DNA. Furthermore, the larger band of the doublet was also seen in conventional hybrids that covered Xq. The smaller band of the doublet was only seen in all radiation hybrids and conventional hybrids which mapped to interval 10 of the mapping panel.

**Development of a YAC contig in Xp11.3**

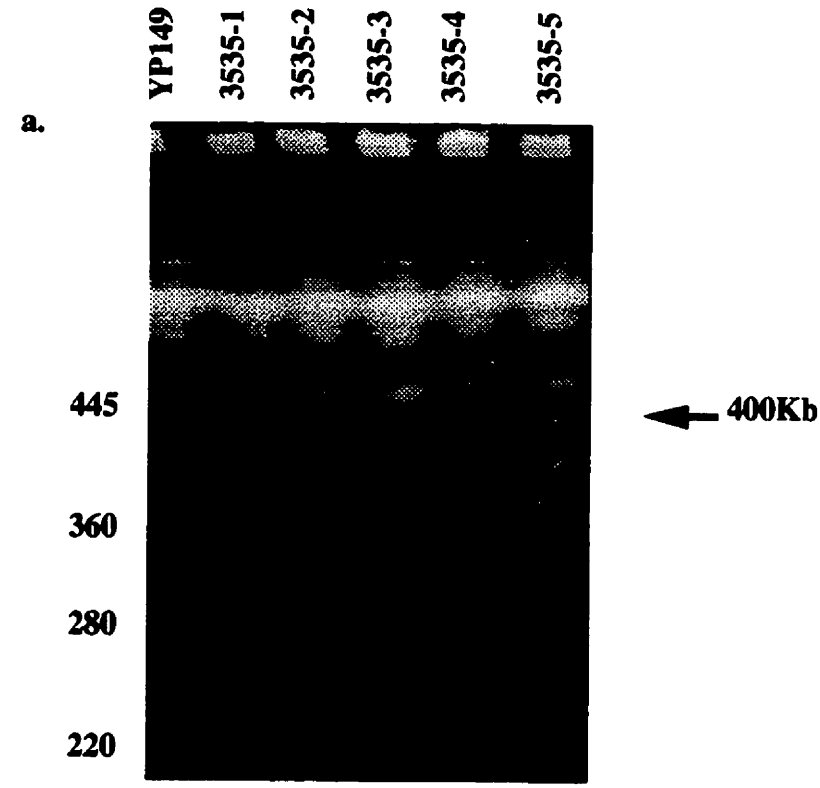
*Characterization of clones*

The region between the markers MAOB and ZNF21 is one of the few regions on the X chromosome still without continuous YAC coverage. This region of approximately 7 cM was chosen to study due to the absence of physical coverage and, as described in the Introduction, the potential presence of genes responsible for several retinal disorders.

A total of seventy-five YAC clones were characterized in this study to develop a YAC contig. The resulting contig was composed of 26 YAC clones which covered over 2 Mb between markers EC7034 and ZNF21. The size and stability of each YAC clone was determined by pulsed field gel electrophoresis. Five clonal isolates for each YAC were run on a pulsed field gel along with the size standard YP149. To visualize the YAC DNA in a yeast background, gels were transferred to a Hybond-$N^+$ membrane and probed with $^{32}$P-labeled human genomic DNA. YAC stability was determined by assessing the sizes of the five clonal isolates. Any size variation observed in these YACs was attributed to clone instability. Figure 6a shows a pulse field gel of YAC yWXD3535, and Figure 6b shows the results after being probed with $^{32}$P-dCTP labeled genomic DNA. By comparing the position of the YAC with the bands in the YP149 standard, the size of each clonal isolate of YAC yWXD3535 was determined to be approximately 400 kb and therefore this YAC was considered to be stable.

In contrast, clonal isolates of yWXD3234 were not found to be the same size. Figures 7a and 7b show a pulsed-field gel of clonal isolates of YAC yWXD3234 and the subsequent filter probed with $^{32}$P-labeled human genomic DNA respectively. Lanes 1 and 2 show clones with an estimated size of 260 kb. Lanes 3 and 4 show clones of 110 kb. Lane 5 shows two bands and, presumably, two derivatives from one clone. As this YAC gave rise to products of different sizes, the YAC yWXD3234 was classified as unstable.

Figure 6. Pulsed-field gel stained with EtBr showing YAC yWXD3535 clonal isolates (a) and the subsequent probing of the filter with $^{32}$P-labeled human genomic DNA (b). Each clonal isolate and the YP149 marker is shown across the top of the gel. By comparison with its position relative to the DNA size standard YP149, yWXD3535 was estimated to be 400 kb.


a.
YP149
3535-1
3535-2
3535-3
3535-4
3535-5
445
360
280
220
400Kb

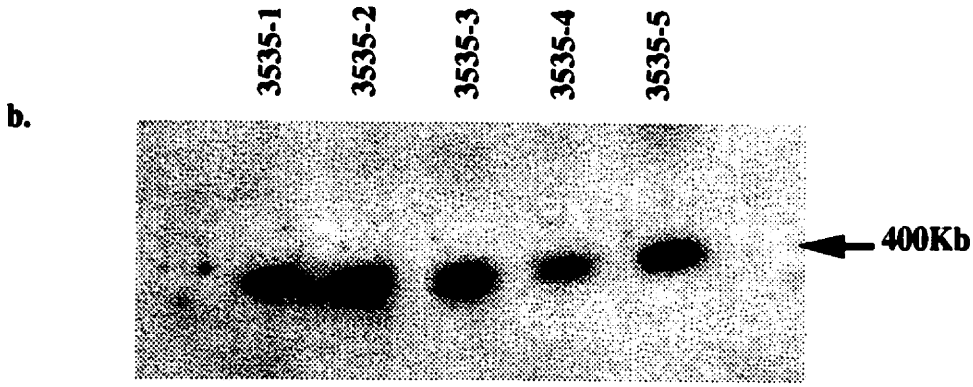
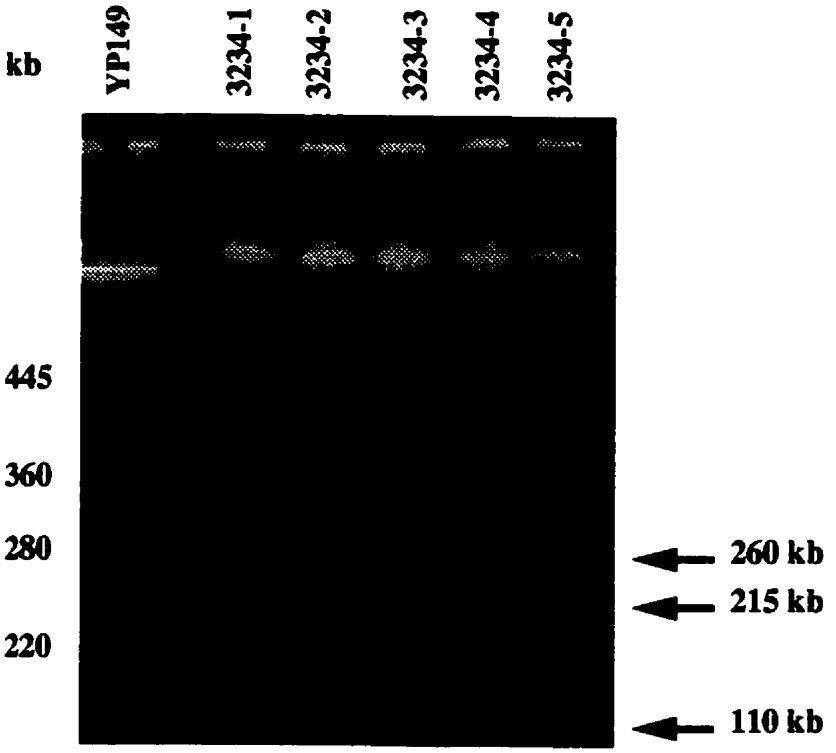
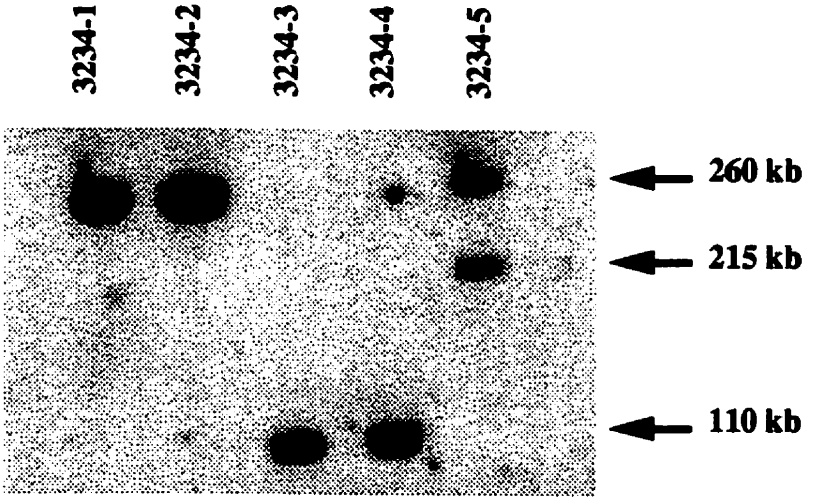

b.
3535-1
3535-2
3535-3
3535-4
3535-5
400Kb

Figure 7. Pulsed-field gel stained with EtBr showing (a) YAC yWXD3234 clonal isolates and (b) the subsequent probing of filter with $^{32}$P- labeled human genomic DNA. Each clonal isolate and the YP149 marker is shown across the top of the gel. By comparison with its position relative to the DNA size standard YP149, yWXD3234 was estimated to have two different sizes of 110 kb or 260 kb. Lane five presumably contained two different different sizes of the same clone. As this YAC gave rise to clonal isolates with varying sizes, it is considered to be unstable.

a.


kb
YP149
3234-1
3234-2
3234-3
3234-4
3234-5
445
360
280
220
260 kb
215 kb
110 kb

b.


3234-1
3234-2
3234-3
3234-4
3234-5
260 kb
215 kb
110 kb

In the cases where size variation was observed, the largest clonal isolate was chosen for further characterization, based on the assumption that size variations were likely due to deletions. The sizes of the YACs analyzed ranged from 120 kb to 600 kb (Table 5) and six YACs were observed to have heterogeneous sizes.

To ensure that the chosen YACs were representative of the region, STS PCR was performed on at least one clonal isolate of every YAC clone. Figure 8 shows that marker DXS7804 lies on YAC lones yWXD3535, yWXD4571, yWXD8086 and yWXD8058. Most clones were assessed for STS content by non-radioactive PCR. However, due to difficulties encountered in scoring the PCR products visualized on an EtBr gel, radioactive PCR was performed for analysis of marker retention for several polymorphic markers (Table 6). STS content data obtained in this study was not consistent with the results I had expected. Many of the clonal isolates did not contain all of the STS markers they had been reported as containing. Since these STS content discrepancies might have been due to deletions in the clonal isolates chosen for analysis, STS PCR was also performed on bulk DNA preparations of all 25 YACs chosen for the study. In two cases (data not shown), PCR products were positive for the bulk preparations and not for the chosen clonal isolate. PCR on the bulk preparation of YAC yWXD1149 resulted in products with markers DXS1366 and EC2691R whereas the chosen clonal isolate was negative for both these markers. Similarly, the bulk preparation of yWXD3376 was positive for UBE1. In these cases the bulks were scored as positive since the chosen clonal isolates likely contained a deletion not present in all clones of the bulk culture. YACs which were smaller than 100 kb or did not contain more than one STS were not included in the final study.

Table 5. YAC clones used in the final physical map of this study

| YAC name[a] | Size (kb) | Library location | Chimeric |
|---|---|---|---|
| yWXD1149 | 570 | E122H7 | |
| yWXD1909 | 300 | I37GA5 | |
| yWXD1910 | 285 | I8DF4 | left |
| yWXD2689 | 360 | I8GF7 | left |
| yWXD2690 | 580 | I18CG12 | |
| yWXD2691 | 345 | I34HA4 | |
| yWXD3283 | 100 | F3H9 | |
| yWXD3234 | 280 | F3D8 | |
| yWXD3376 | 120 | F5B1 | |
| yWXD3535 | 390 | F6G5 | |
| yWXD3882 | 150 | F10E10 | |
| yWXD4571 | 280 | F18A7 | |
| yWXD7034 | 320 | E192D8 | |
| yWXD8049 | 600 | I192C4 | |
| yWXD8057 | 280 | I47E8 | |
| yWXD8058 | 300 | I107E9 | |
| yWXD8084 | 300 | I253F12 | right |
| yWXD8085 | 330 | I257E10 | left |
| yWXD8086 | 500 | I184D8 | |
| yWXD8180 | 100 | E173C4 | right |
| A0120 | 590 | ICRFy900A0120 | |
| A1220 | 730[a] | ICRFy900A1220 | |
| C1022 | 400[b] | ICRFy900C1022 | |
| C1228 | 330 | ICRFy900C1228 | |
| D1551 | 530[a] | ICRFy900D1551 | |
| F0892 | 480 | ICRFy900F0892 | |
| F05131 | not sized | ICRFy900F05131 | |

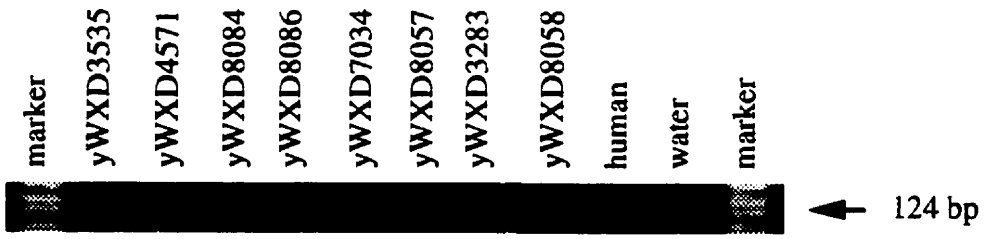
[a] published size (Coleman *et al.*, 1990); not confirmed

[b] published size (Hagemann *et al.*, 1994; Kwan *et al.*, 1995); not confirmed

Table 6. Polymorphic markers amplified by radioactive PCR

| Marker | Primer Sequence 5'-3' | Size (bp) | Reference |
|---|---|---|---|
| DXS1055 | TTAAACAATGCACAACTGGG | 81-95 | Gyapay *et al.*, 1994 |
| | ATGGGATACACTGTTCTGGG | | |
| DXS1003 | CCATTCCTCACTGGCAAG | 169-195 | Weissenbach *et al.*, |
| | TTCACCCATAGAAGCCGT | | 1992 |
| EC8058R | GGAAATCACATGGTAAGAGA | 315-335 | this study |
| | TACCAACAAGACATACCTTT | | |
| DXS8083 | CTTCTGCACAGCAAAGGAAA | 232-248 | Thiselton *et al.*, |
| | ACTCCAGGAGGCCGTATGTC | | 1996 |

Figure 8. STS content analysis of a panel of YACs using marker DXS7804. Retention of the marker was determined by non-radioactive PCR analysis as described in 'General Materials and Methods' section. The name of each YAC clone is described on top along with human genomic DNA and dd$H_2O$ control lanes. DXS7804 is present on YAC clones yWXD3535, yWXD4571, yWXD8086 and yWXD8058. Products were separated on a 1.5% SEPARIDE/ 1% agarose gel and visualized by EtBr staining. The DNA size standard is pBst cut with HaeIII.


marker
yWXD3535
yWXD4571
yWXD8084
yWXD8086
yWXD7034
yWXD8057
yWXD3283
yWXD8058
human
water
marker
124 bp

*End-cloning*

After determining the clonal stability and STS content of an initial set of 10 YAC clones, efforts were made to isolate end-fragments from these clones by ligation-mediated PCR (end-cloning). In addition to providing increased STS density and beginning the 'walk' out of the seed contig, the isolation of YAC ends by ligation-mediated PCR and the subsequent development of new STS markers from such YAC ends, provided another method for evaluating the integrity of clones.

Figure 9 shows the fragments isolated from the left and right ends of YAC yWXD3234 using restriction enzymes *RsaI* and *AluI*. Fragments obtained that were over 200 bp were sequenced using the cycle sequencing system, as described in the 'General Materials and Methods'. In the case of the YAC yWXD3234 depicted in Figure 9, Left 3234-*RsaI* and Right 3234-*AluI* were chosen for sequencing. DNA sequence obtained from such end-clone fragments was then examined for homology with sequences in the GenBank database using BLAST analysis (Altschul *et al.*, 1990). Primers were designed from any sequence that did not show homology with known repetitive sequences (e.g. Alu and L1 repeats). Two end-clones, EC1909R and EC1228R, were discarded after BLAST analysis showed these sequences had homology with Alu or L1 repeats. Figure 10 shows the sequence used to design the forward primer corresponding to the right-end fragment of yWXD3234. All new STS markers were used to amplify the hybrid panel described in the Introduction to determine if they were region specific or if they had been derived from a chimeric parent clone. To ensure that the end fragment, and therefore the new STS, was not from a chimeric YAC end, EC3234R (right end of yWXD3234) was tested on the mapping panel of radiation and conventional hybrids (Figure 11). EC3234R was mapped to interval 11 or 13 (both of these intervals have identical mapping patterns). Once EC3234R was assessed to be region-specific, it was used to increase further the

Figure 9. Fragments isolated by ligation mediated PCR from the left and right ends of YAC yWXD3234 using the restriction enzymes *RsaI* and *AluI*. Bands in EtBr stained gels were sized by comparison with the size standard pBst cut with *HaeIII* and the largest, clearest bands from each end were then sequenced.

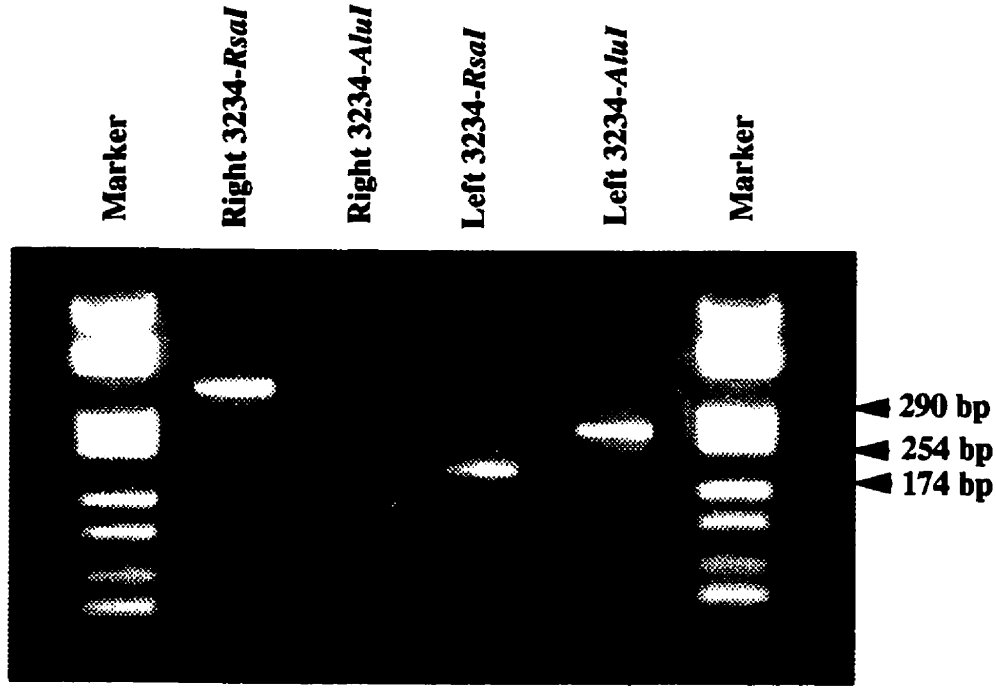

Marker
Right 3234-RsaI
Right 3234-AluI
Left 3234-RsaI
Left 3234-AluI
Marker
290 bp
254 bp
174 bp

Figure 10. DNA sequence from part of the right end-clone isolated from YAC yWXD3234 resolved on a 6% acrylamide gel. This sequence was used to design the forward primer for the marker EC3234R described in Table 5.

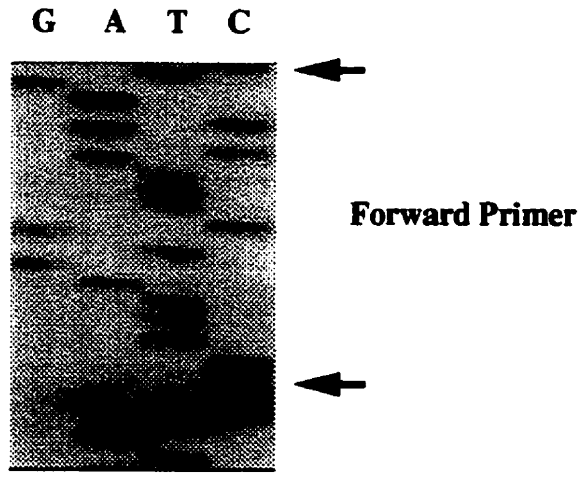
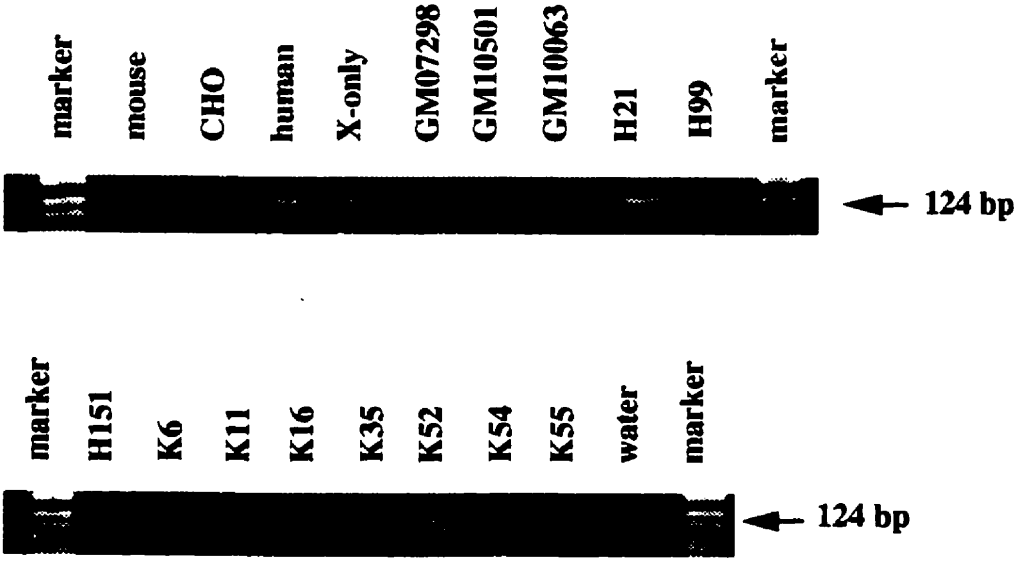

G A T C
Forward Primer

Figure 11. Somatic cell hybrid panel mapping of marker EC3234R. Retention of EC3234R on a set of 10 radiation-reduced hybrids and four somatic cell hybrids was evaluated using PCR as described in the 'General Materials and Methods' section. Names of hybrids are indicated along the top including control lanes with human genomic, human X-only, Chinese hamster (CHO), mouse and dd$H_2O$. PCR products, of approximately X bp in size, were separated on 2% Nuseive agarose/1% agarose gels and visualized by EtBr staining as described in the 'General Materials and Methods' section. Retention of EC3234R on hybrids H21, H99, K6, K11, K16, K35, K52, K54 and GM10063 is consistent with EC3234R mapping to interval 11 or 13 of the hybrid mapping panel. The specific location of EC3234R was resolved by physical mapping data (see Figure 10).


marker
mouse
CHO
human
X-only
GM07298
GM10501
GM10063
H21
H99
marker
124 bp
marker
H151
K6
K11
K16
K35
K52
K54
K55
water
marker
124 bp

STS content data of the 25 YAC clones, and determine overlap. Figure 12 shows that yWXD3234 overlaps with yWXD1910, yWXDy900F0892 and yWXD2691. Furthermore, this data indicated that EC3234R belongs within interval 13, and not 11, of the mapping panel. Region-specific end-clones that were found, by STS content analysis, to lie in a physically unmapped region were then used to screen YAC libraries to obtain further clones. This process allowed unmapped regions to be 'walked' into.

In total, twenty-one YAC end-clone fragments were isolated and sequenced. Of these, 12 were found to be region-specific (Table 7), five consisted of chimeric ends and four ends had homology to Alu or L1 repetitive sequences. Of the five chimeric ends, the four ends derived from the I library, and the one chimeric end from the E library were derived from another human chromosome.

Analysis of the retention pattern of marker EC3535R on the hybrid panel suggested, as observed earlier for DXS8083, that these primer sequences were duplicated in the genome. Amplification of the hybrid panel with EC3535R, resulted in two bands being observed in the X-only hybrid and genomic DNA. Furthermore, the smaller band of the doublet was also seen in conventional hybrids that covered Xq. The larger band of the doublet was only seen in all radiation hybrids and conventional hybrids that mapped to interval 10 of the mapping panel.

*New Polymorphic Marker*

Sequencing of the right end-clone of yWXD8058 revealed a TC dinucleotide repeat. Figure 13 shows a portion of the repeat and sequence on one side. Since dinucleotide repeats have often been found to be polymorphic, primers were designed to develop a marker that included the TC repeat. Unfortunately, sequence could not be obtained from both sides of the repeat by sequencing from the right internal vector primer. Therefore, the end-clone was sequenced from both the linker specific primer and

Figure 12. STS content analysis of a panel of YACs using marker EC3234R. Retention of the marker was determined by non-radioactive PCR analysis as described in 'General Materials and Methods' section. The name of each YAC clone is described on top along with human genomic DNA and $ddH_2O$ control lanes. EC3234R is present on YAC clones yWXD1910, ICRFy900F0892, yWXD2691, and its parent clone sWXD3234. Products were separated on a 1.5% SEPARIDE/ 1% agarose gel and visualized by EtBr staining. The DNA size standard is pBst cut with HaeIII.

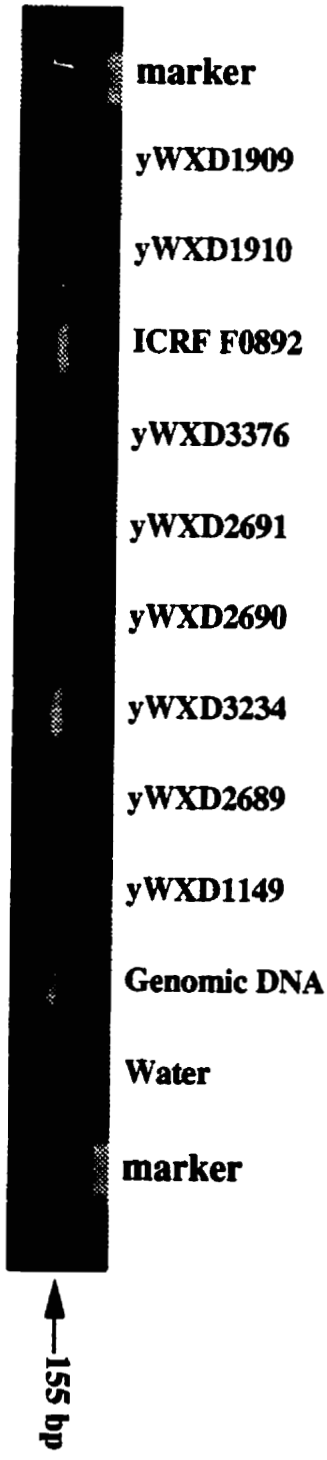

marker
yWXD1909
yWXD1910
ICRF F0892
yWXD3376
yWXD2691
yWXD2690
yWXD3234
yWXD2689
yWXD1149
Genomic DNA
Water
marker
155 bp

Figure 13 Sequence of the right end-clone of yWXD8058 showing part of the TC dinucleotide repeat. The dideoxy G, A, T, C lanes are indicated and the termination products have been resolved on a 6% acrylamide gel.

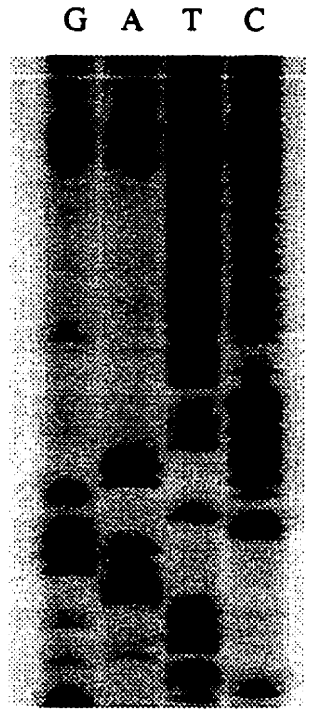

G A T C

Table 7. STS Markers Designed From YAC End-Clones

| Marker Name | Hybrid Interval[b] | Lab name | Primer Sequence 5'-3' | Size (bp) |
|---|---|---|---|---|
| sWXD3383[a] | 11 or **13** | EC1909L | TCTCAACAAATGTCAGCAAC | 134 |
| | | | CCCTCCTCTGAAATCTGCTA | |
| DXS9937 | 11 or **13** | EC3376R | GGGACTGAATCATACCTAAC | 127 |
| | | | CCACTTGTGTATTTGAGGGT | |
| DXS9936 | 11 or **13** | EC3234R | TTTCTGCCAGCAGTTACTTT | 155 |
| | | | CCCTTTAGTGCTTTACACAGT | |
| sWXD3384[a] | **11** or 13 | EC2690R | AGGCAACCCTGGGCAGTGTG | 83 |
| | | | GGATGACCCAAAGGTTCTGT | |
| N/A[a] | **11** or 13 | EC2691R | TAATGTTCCATTGTTTGCCC | 121 |
| | | | TTACTGCAAAGAAATACAAG | |
| DXS9935 | 10 | EC1149L | TCATGCCAATGAATAACCAG | 143 |
| | | | GCTAAATCAATCCCAATAAA | |
| DXS9934 | 10 | EC3535R | GGATGATATGGTTGGATTTG | 96 |
| | | | ATGATCCAATCACCTCCCAC | |
| DXS9933 | 10 | EC3535L | CTCAGGAGCAGGCGGGAAAC | 90 |
| | | | TTTCTAGCCAGGTAATTTGC | |
| DXS9932 | 10 | EC8057R | CATAGTGACTGGTCCTTTATGT | 91 |
| | | | GTATGAAGATTTGGAGTTG | |
| DXS9931 | 10 | EC8058R[c] | GGAAATCACATGGTAAGAGA | 315- |
| | | | TACCAACAAGACATACCTTT | 335 |
| DXS9930 | 10 | EC7034RI | GTTGATACCACAGAACTACA | 130 |
| | | | AACTCAATTAGATGTGTGAT | |

[a] DXS number has not yet been released.

[b] bold face indicates interval marker was mapped to by STS analysis of YAC and PAC clones.

[c] primer pair flanks polymorphic $(TC)_n$ repeat.

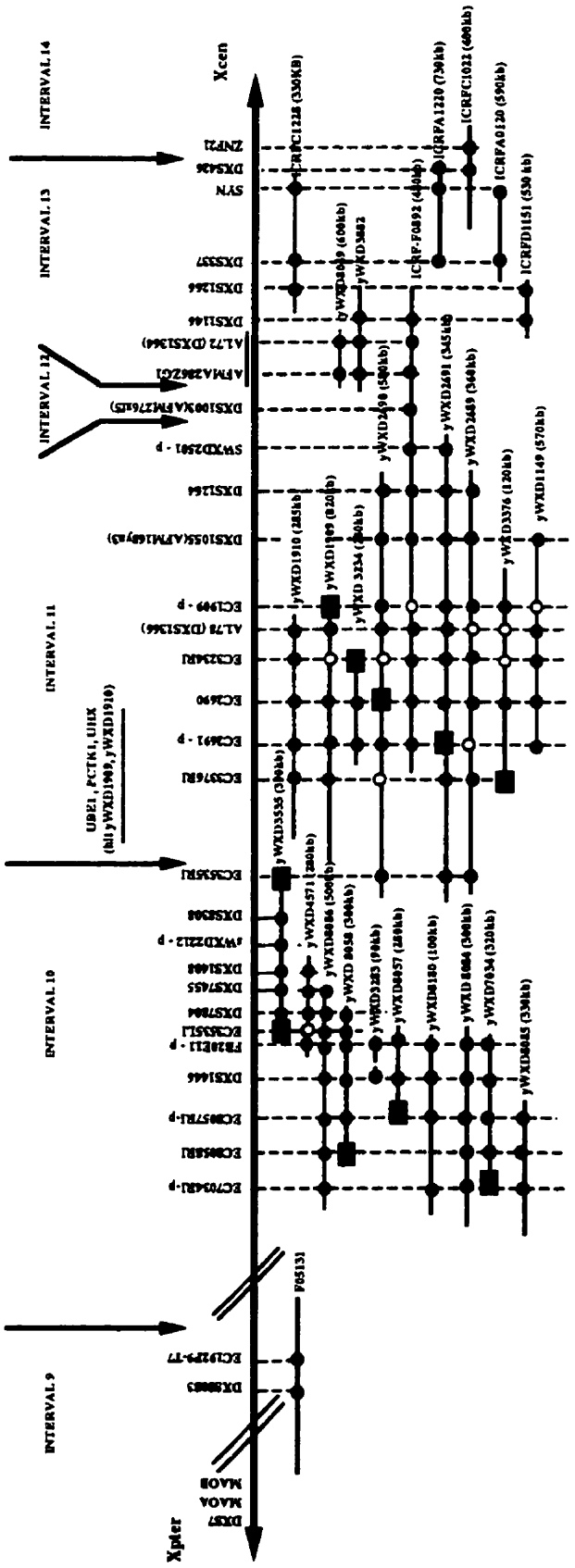
vector specific primer to get sequence flanking the TC repeat . Amplification of DNA isolated from 50 unrelated Caucasian females showed the resulting marker to be polymorphic with allele frequencies of 0.4 (313 bp product), 0.16 (315 bp product) and 0.44 (317 bp product) (data not shown). This polymorphic marker adds density to the genetic map in this region and could potentially help refine the location of RP2.

*YAC Contig*

The resulting YAC contig was assembled using 26 YACs and 37 DNA markers from the Xp11.3 region spanning the interval DXS8083 to ZNF21 (Figure 14). The contig was constructed based on the retention of DNA markers in YAC clones, as determined by non-radioactive (Figure 12) or radioactive PCR, and information on DNA marker retention in conventional and radiation hybrids (Figure 5, Figure 11). End-clones developed from YAC clones increased the STS marker density within the region, gave more information on overlaping clones, and allowed additional YACs to be acquired by screening YAC libraries. This in turn contributed to the development of the resulting physical contig.

The initial effort to assemble a contig of the Xp11.3 region is depicted in Figure 14. This contig, based entirely on data obtained from YAC analysis, has several discrepancies. Markers UBE1, PCTK1 and UHX are known to fall within interval 11 or 13 on the hybrid panel and have been published as lying distal to DXS1055 (Knight *et al.*, 1994). However, STS content analysis using YACs was not consistent with these findings. Although these markers fell on some YACs (yWXD1909 and yWXD1910), the STS data obtained did not allow markers to be placed within the contig without deletions being found in a significant number of clones. For this reason, these markers have been placed, unordered, at the top of the contig. Furthermore, the sizes of YACs are not to

Figure 14. Physical contig and DNA marker map in human Xp11.3-p11.23. The thick, dark line with bi-directional arrows indicates the chromosome, with the telomere to the left and the centromere to the right. DNA markers and genes are ordered across the top of the figure, as determined by retention in conventional and radiation-reduced hybrids, YAC content, parsimony and SEGMAP analysis. Polymorphic markers are indicated by asterisks. The hybrid breakpoints are indicated by arrows. A line under two or more markers indicates that they could not be ordered relative to each other. YAC clones are represented as lines together with their names and sizes. Deletions are indicated by open circles, end clones are indicated by boxes. UBE1, PCTK1 and UHX are in a box above the contig since they could not be placed using YAC data.


Xcen
Xpter
INTERVAL 14
INTERVAL 13
INTERVAL 12
INTERVAL 11
INTERVAL 10
INTERVAL 9
ZNF21
DXS436
SYN
DXS337
DXS1266
DXS1146
AL72 (DXS1266)
AFMA286ZG1
DXS1003(AFMB276ZD)
SWXD2501 - p
DXS1264
DXS1055(AFMA146yh5)
EC1909 - p
AL78 (DXS1266)
EC2264RI
EC2690
EC2691 - p
EC2576RI
UBE1 , PCTK1, UBK
(hl ) yWXD1910, yWXD1910)
EC2565RI
DXS8208
yWXD2212 - p
DXS1068
DXS7455
DXS7004
EC2565LI
FB20E21 - p
DXS1666
EC2057RI-p
EC2056RI
EC2703RI-p
DXS7
MAOA
MAOB
DXS8083
EC1929-T7
P05131
ICRFC1228 (350KB)
ICRFA1220 (790kb)
ICRFC1021 (400kb)
ICRFA0120 (590kb)
ICRFD1151 (530 kb)
ICRF-F0892 (400kb)
yWXD3059 (600kb)
yWXD3082
yWXD2690 (500kb)
yWXD2691 (345kb)
yWXD2689 (360kb)
yWXD3376 (120kb)
yWXD1149 (570kb)
yWXD1910 (285kb)
yWXD1965 (820kb)
yWXD 3234 (300kb)
yWXD3555 (390kb)
yWXD4571 (280kb)
yWXD0096 (500kb)
yWXD 0058 (300kb)
yWXD3223 (90kb)
yWXD4657 (280kb)
yWXD0180 (100kb)
yWXD 0086 (300kb)
yWXD7034 (320kb)
yWXD0085 (330kb)

scale since the current, most parsimonious, order of markers does not permit clones to be drawn to scale.

**Collection of PACs**

The collection of PACs in the target region was accomplished by screening the deJong pCYPAC library (Ioannou *et al.*, 1994) with 26 existing and new markers. Fourteen PAC clones resulted from the screenings and were characterized. This information helped seed the region for the future creation of a PAC contig and helped orient markers whose relative placement could not be determined by the analysis of YAC data.

The size of each PAC clone was determined by digesting the DNA with *NotI*, running the DNA on a pulse field gel and comparing the size with a midrange PFG marker. Figure 15 shows a pulsed-field gel stained with EtBr. The size of the PAC clones varies between 23 kb and 160 kb (Table 8). PAC 182B7, seen in lanes 4 and 5, has two distinct bands implying an internal *NotI* site. PAC 71L16 also has 2 distinct bands. This could also be due to an internal *NotI* site. Another possibility is that the two bands seen for PAC 71L16 and PAC182B7 are due to the presence of two separate clones resulting from a co-cloning event.

Five clonal isolates of six different PAC clones (237A9, 250J21, 235D18, 130L5, 86K24, 182B7) were digested with *BamHI*, visualized on an EtBr stained gel and their restriction patterns compared. Of the six PACs analyzed, no discrepancies were observed. These clones were thus considered to be stable.

STS content was evaluated using the technique of Southern hybridization. Figure 16 illustrates a Southern hybridization using marker DXS1264 which is shown to lie on PAC clones 71L16, 182B7, and 250J21. For many of the polymorphic markers, STS content had to be evaluated using radioactive PCR. Figure 17 shows the results of a

Figure 15. Pulse field gel stained with EtBr showing eight different PAC clones digested with *NotI*. Each clone and the mid-range PFG marker is labeled at the top of the gel. By comparison with its position relative to the DNA size standard, the size of each PAC clone could be determined.

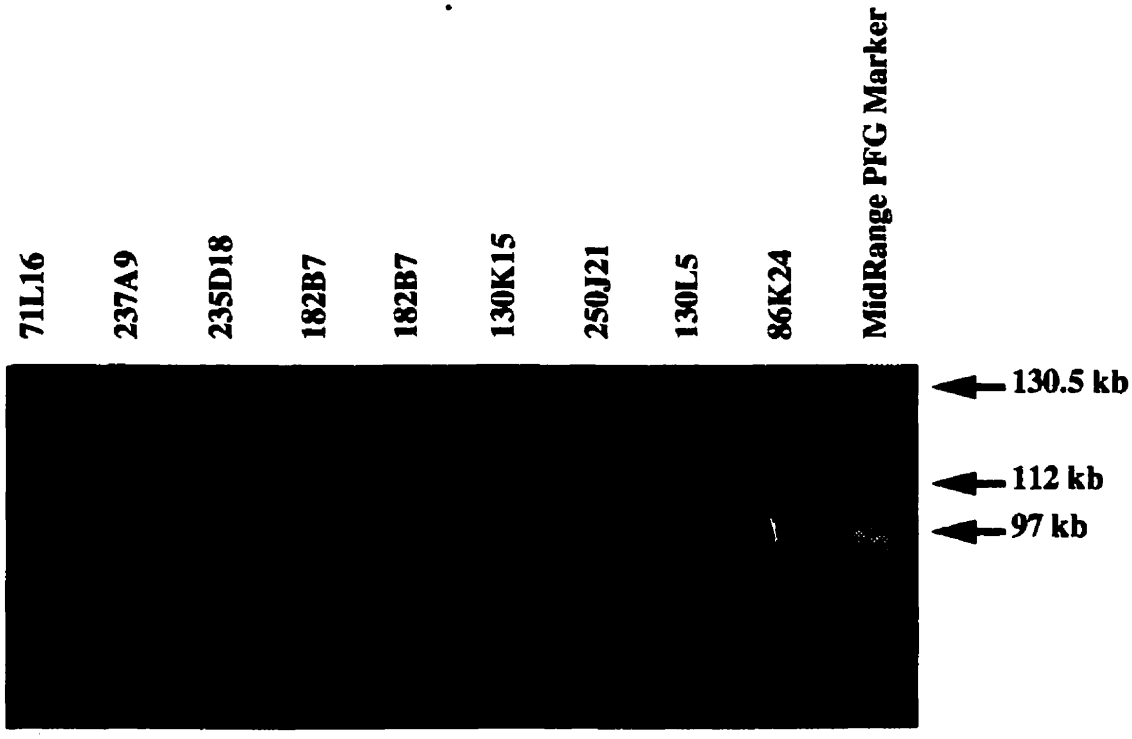

71L16
237A9
235D18
182B7
182B7
130K15
250J21
130L5
86K24
MidRange PFG Marker
130.5 kb
112 kb
97 kb

Table 8. Sizes of PAC clones

| PAC | Size (kb) |
|---|---|
| 130L5 | 112 |
| 86K24 | 110 |
| 395N6 | 120 |
| 30G7 | 110 |
| 182B7 | 160 |
| 71L16 | 120 |
| 250J21 | 110 |
| 248M5 | 120 |
| 121M13 | 23 |
| 237A9 | 115 |
| 235D18 | 115 |
| 130K15 | 110 |
| 121D15 | 97 |
| 63F12 | 130 |
| 192P9 | 145 |

Figure 16. DNA marker content analysis of PAC clones using Southersn hybridization. The PAC clone names are indicated along the top and described in Table 8. PAC clones were digested with BamHI and electrophoresed on 0.8% agarose gels, blotted and probed with the DNA marker DXS1264. The control lane is human genomic DNA. DXS1264 is contained in the PACs

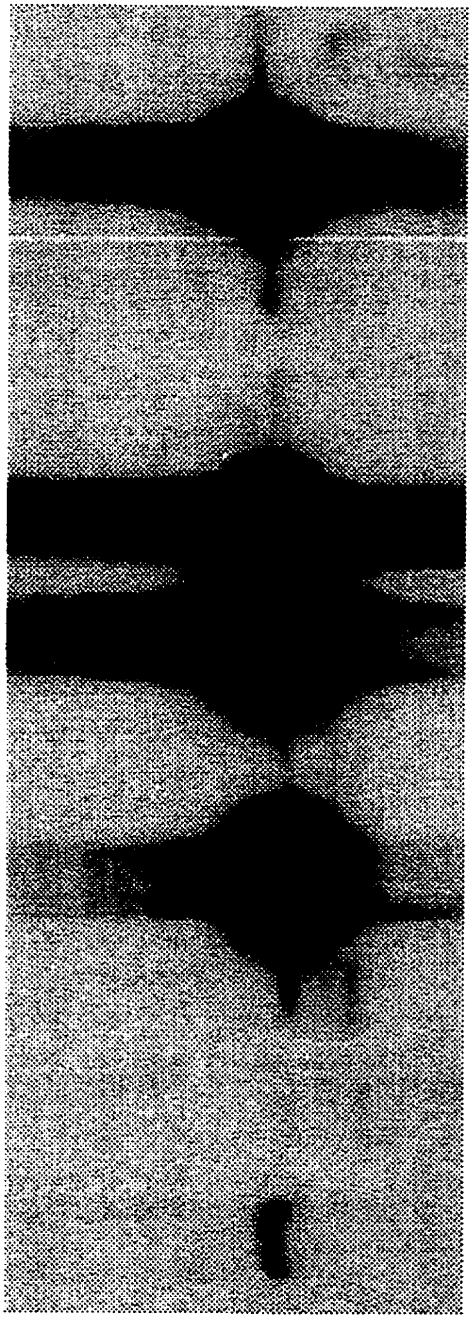
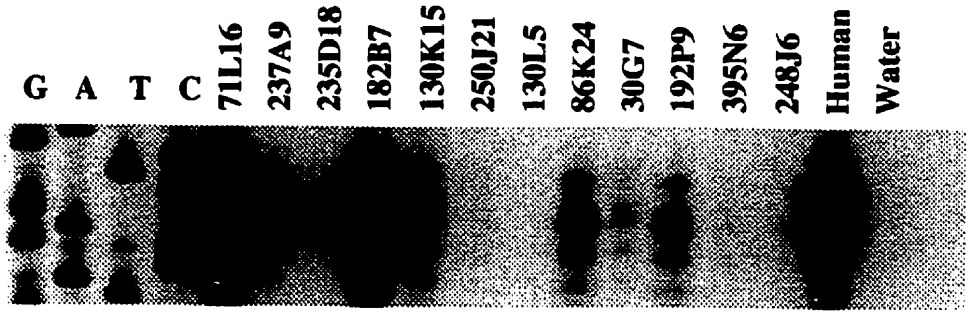

71L16
237A9
235D18
182B7
182B7
130K15
250J21
130L5
86K24
human

Figure 17. DNA marker retention analysis for polymorphic marker DXS1055 on a panel of PAC clones using radioactive PCR. The PAC names, human and water controls are indicated along the top. DXS1055, a polymorphic end-clone described in Table 7, is retained on PAC clones 71L16 and 182B7. Radioactive PCR was performed as described in 'General Materials and Methods' section. PCR products were separated on an acrylamide gel as described in the ' General Materials and Methods' section and sized by comparison with the previously described radiolabeled DNA marker from the SEQUENASE 2.0 DNA sequencing kit.


G
A
T
C
71L16
237A9
235D18
182B7
130K15
250J21
130L5
86K24
30G7
192P9
395N6
248J6
Human
Water

radioactive PCR using DXS1055. The data resulting from STS content analysis of the PACs (Figure 18) showed that a number of the clones (121M13, 237A9, 235D18, 130K15, 182B7, 71L16, 250J21 and 248M5) were clustered between markers EC3234R and EC1149R. Furthermore, 14 markers were found to fall within this region many of which were oriented with respect to one another by comparing overlaping PAC clones. For example, DXS1055, DXS1264 and A1.78 were found to lie on PAC71L16 whereas EC2691R, EC2690R, WI18364 and EC182B7-T7 were found to lie on both PAC71L16 and PAC250J21. This data showed that these two groups of markers lay apart from one another but did lie next to each other. Lastly, STS content analysis revealed that PAC71L16 was positive for markers DXS1003, EC182B7T7, WI18364, EC2690R, EC2691R, DXS1366, DXS1264 and DXS1055. However, it is also positive for marker EC7034R. This data is more consistent with a co-cloning event. There was no STS evidence to prove PAC182B7 had undergone a co-cloning event. Furthermore, STS content analysis did not provide enough data to determine wether the size of PAC182B7 was consistent with one of the bands or both bands added together.

To increase marker density, and to help elucidate marker order, external and internal primer sequences were designed from the PAC vector to allow further development of end-clone fragments. Three PAC end-clones were isolated and developed into STS markers (Table 9). All markers developed from PAC end-fragments mapped to the appropriate interval on the hybrid panel (Table 9).

Figure 18. Physical PAC contig and DNA marker map in human Xp11.3-p11.23. The thick, dark line with bi-directional arrows indicates the chromosome, with the telomere to the left and the centromere to the right. DNA markers and genes are ordered across the top of the figure, as determined by retention in conventional and radiation-reduced hybrids, PAC content and parsimony. Polymorphic markers are indicated by asterisks. The hybrid breakpoints are indicated by arrows. A line under two or more markers indicates that they could not be ordered relative to each other. PAC clones are represented as lines together with their names.

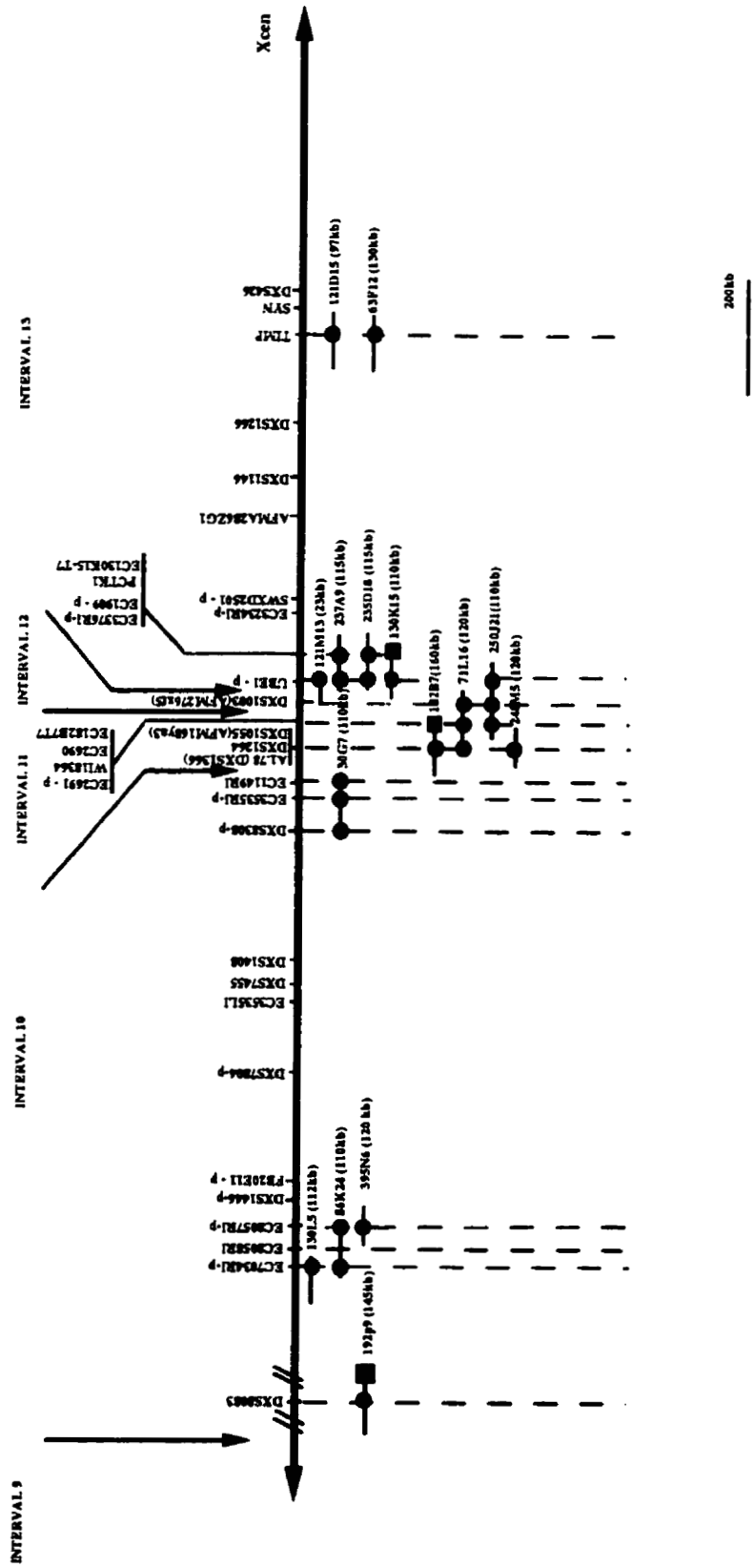

Xcen
INTERVAL 13
INTERVAL 12
INTERVAL 11
INTERVAL 10
INTERVAL 9
200kb

Table 9. STS Markers Designed From PAC End-Clones

| Marker Name | Lab Designation[a] | Primer Sequence 5'-3' | Interval[b] | Size (bp) |
|---|---|---|---|---|
| DXS9938 | EC130K15T7 | TTGAGATAACAGGTGGCACC<br>AGGACCTGCTACTCAATCTG | 11 or **13** | 145 |
| DXS9939 | EC182B7T7 | TCTCTAGTTGCATCCAGTCT<br>CTGGGTAATACAGAGGAAAG | **11** or 13 | 134 |
| DXS9940 | EC192P9T7 | CCCATTCAGTATGATAATAG<br>TCAATCAGTGTGATATATC | 10 | 149 |

[a] The naming of markers is a combination of EC (End-Clone), the name of parent clone and followed by T7 (the T7 end of a PAC).

[b] bold face indicates interval marker was mapped to by STS analysis of YAC and PAC clones.

Fourteen PACs were collected by screening a PAC library with 26 DNA markers from the Xp11.3 region spanning DXS8083 and ZNF21 (Figure 18). The contig was constructed based on the retention of DNA markers in PAC clones, as determined by radioactive PCR (Figure 17) or Southern hybridization (Figure 16). Radioactive PCR was performed for analysis of marker retention for several polymorphic markers (Table 7), because of difficulties encountered in performing Southern hybridizations with repetitive sequences.

## Integrated YAC and PAC Physical contig

### *Marker Order*

The PAC contig depicted in Figure 18 clarifies several problem areas encountered in the YAC contig. Specifically, the localization of markers UBE1, UHX and PCTK1 was elucidated. Furthermore, the 14 markers between EC3234R and EC1149R fell on a series of PACs, indicating their close proximity and orientation with respect to one-another. The integration of all PAC and YAC data resulted in a marker order of Xpter-

EC7034R-EC8058R-EC8057R-DXS1446-FB20E11-DXS7804-EC3535L-DXS7455-DXS1408-DXS8308-sWXD2212-EC3535R-EC1149R-(DXS1366, DXS1264, DXS, 1055)-(EC2691, EC2690, WI18364, EC182B7T7)-DXS1003-UBE1-(EC13K15T7, PCTK1, UHX, EC3376, EC1909)-EC3234R-sWXD2501- AFMA286ZG1-(DXS1146, DXS1364)-DXS1266-DXS337-(SYN/TIMP)-DXS426-ZNF21-Xpcen (Figure 19). PAC data was crucial for determining the order and location of the markers between EC3234R and EC1149R since YACs in this region had a large number of deletions and analyzing only YAC data led to a very different interpretation of the region. PAC data was still unable to discern the order of several groups of markers since all of these markers lie on the same sets of PACs. Lastly, a cross-over between DXS1055 and 1003 helped place DXS1055 distal to DXS1003 (Gyapay *et al.*, 1994). This was vital information as both PAC and YAC data could not orient the island of markers between EC3234R and EC1149R.

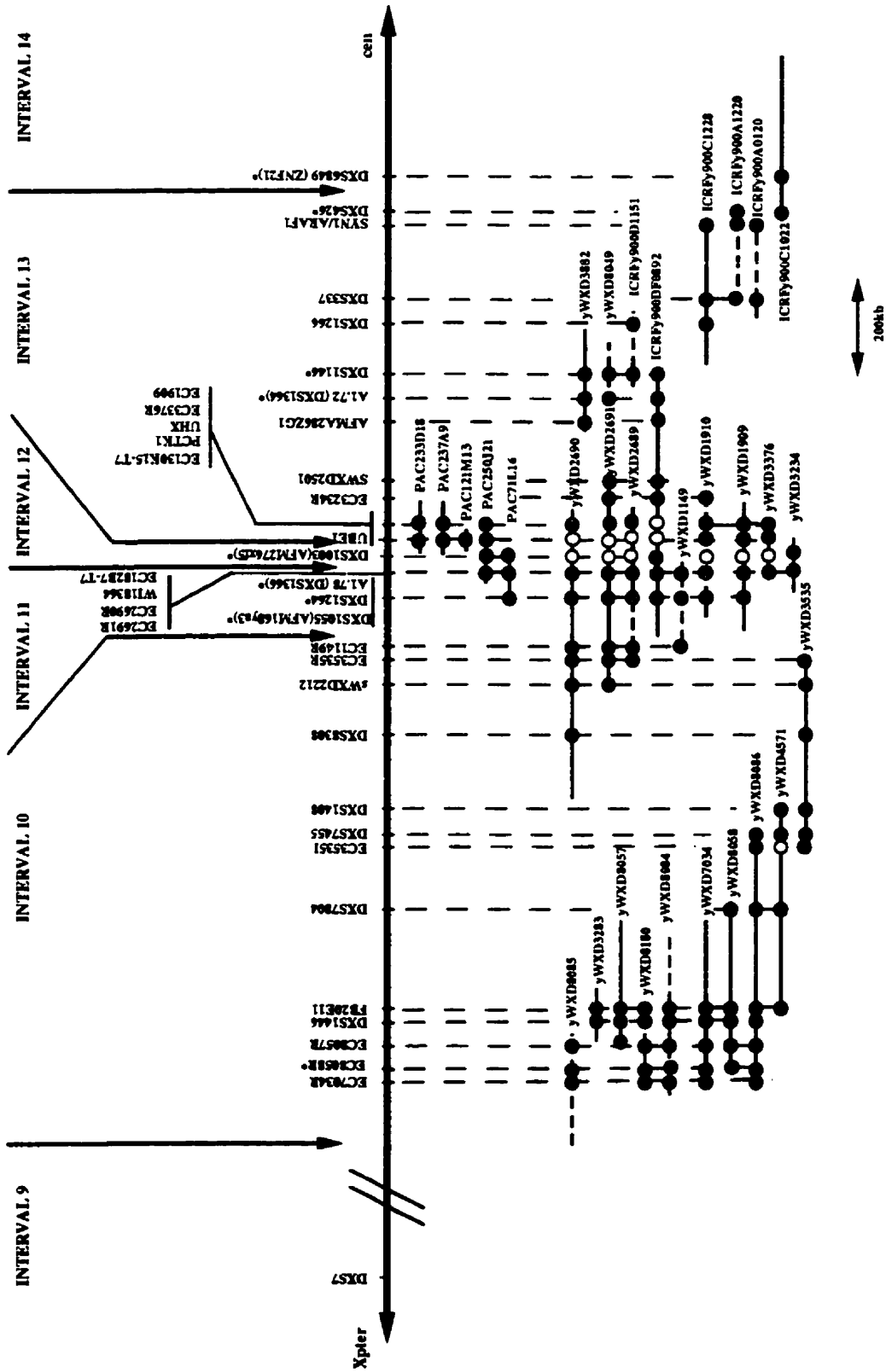
*Weak links and gaps in the Contig* :

There are four weak links in the reported contig, defined as two markers being held together by only one clone. These lie between markers DXS8308 and DXS1408, DXS1146 and DXS1266, DXS1266 and DXS337 and SYN1 and DXS426.

The first weak link between DXS8308 and DXS1408 is spanned by YAC yWXD3535. Although only yWXD3535 spans this region, markers were developed from both its ends firmly anchoring it to interval 10 and no deletions were seen in the clone. This supports the idea that this clone is stable and non-chimeric and thus a valid link.

The other set of weak regions are all found at the proximal end of the contig. All clones and markers from this region were previously published by several groups (Coleman *et al.*, 1990; Hagemann *et al.*, 1994; Knight *et al.*, 1994; Kwan *et al.*, 1995).

Figure 19. Integrated physical contig and DNA marker map in human Xp11.3-p11.23. The thick, dark line with bi-directional arrows indicates the chromosome, with the telomere to the left and the centromere to the right. DNA markers and genes are ordered across the top of the figure, as determined by retention in conventional and radiation-reduced hybrids, YAC and PAC content, parsimony and SEGMAP analysis. Polymorphic markers are indicated by asterisks. The hybrid breakpoints are indicated by arrows. A line under two or more markers indicates that they could not be ordered relative to each other. YAC and PAC clones are represented as lines together with their names. Deletions are indicated by open circles, and circles with hatched centers indicate that a YAC is not positive for all markers represented. Dashed lines associated with a YAC indicates that there is evidence of chimerism or, in the case of yWXD1149, yWXD8049, ICRFy900D1151, ICRFy900A1220, ICRFy900A0120, that the size of the YAC is too large to fit within the constraints of the map, suggesting that these clones are also chimeric.


INTERVAL 14
INTERVAL 13
INTERVAL 12
INTERVAL 11
INTERVAL 10
INTERVAL 9
cen
Xpter
DXS7
DXS6849 (ZNF21)*
DXS426*
SYN1/ARAF1
DXS337
DXS1266
DXS1146*
AL.72 (DXS1366)*
AFMA286ZC1
SWXD2501
EC23.46R
UBE1
DXS1085(AFM274xf5)*
AL.78 (DXS1246)*
DXS1246*
EC11.49R
EC35.58R
sWXD2212
DXS8368
DXS1468
EC35.51
DXS7455
DXS7864
FB20E11
DXS1444
EC80.57R
EC80.61R*
EC70.34R
EC139.R15-T7
PCTK1
UHX
EC37.6R
EC1.909
EC49.1R
EC49.9R
WI18.46
EC102.87-T7
PAC233D18
PAC237A9
PAC121M13
PAC25Q21
PAC71L16
yWXD3882
yWXD8049
ICRFy900D1151
ICRFy900DF8892
yWXD2690
yWXD2691
yWXD2689
yWXD1149
yWXD1910
yWXD1909
yWXD3376
yWXD3234
yWXD3535
yWXD4571
yWXD8086
yWXD8058
yWXD7034
yWXD8084
yWXD8057
yWXD0186
yWXD3283
yWXD8085
ICRFy900C1228
ICRFy900A1220
ICRFy900A0120
ICRFy900C1022
200kb

These clones are being reported to demonstrate the overlap of this contig with other reported contigs in the proximal region (Boycott *et al.*, 1996).

Finally, a gap remains between EC7034R and DXS8083. Although both of these markers fall within interval 10 on the hybrid panel, the distance between both markers is unknown. Exhaustive screening of the libraries at Washington University with EC192P9-T7 was unsuccessful in obtaining more clones to close this gap.

*Orientation of the Contig on the X chromosome:*

The contig was unambiguously oriented on the X chromsome using the panel of hybrids which allowed the placement of markers into an interval defined by various breakpoints (Figure 3). A linkage map of the X chromosome indicated that DXS1003 lies distal to markers DXS1126 and DXS255 (Donnelly *et al.*, 1994), two markers which are within a proximally overlapping contig (Boycott *et al.*, 1996). Furthermore, the use of YACs from the ICRFy900 series (D1151, C1228, A0120, A1220 and C1022) provided overlap with previously published contigs at the centromeric end of the contig (Coleman *et al.*, 1990; Hagemann *et al.*, 1994; Knight *et al.*, 1994; Kwan *et al.*, 1995) and extends the cloned region proximally to cover the Xp11.23 interval characterized by Boycott and co-workers (Boycott *et al.*, 1996).

# CHAPTER 4 - DISCUSSION

A contig consisting of YACs and PACs was constructed on the short arm of the human X chromosome to establish a physical map in a region of the X-chromosome lacking physical coverage and to which several retinal disorders have been mapped. The contig developed in this study contains 26 YAC clones and 15 PAC clones and was assembled from the DNA marker content of YAC and PAC clones, from DNA marker retention in radiation-reduced and conventional hybrids, and Segmap analysis. The contig spans 2.5 Mb between markers ZNF21 and EC7034R, has three fold coverage, and contains 37 DNA markers in the Xp11.3-p11.23 region with an average resolution of one marker every 75 kb (Figure 19). The use of YACs from the ICRFy900 series (D1151, C1228, A0120, A1220 and C1022) provides overlap with previously published contigs at the centromeric end of the contig and extends the cloned region proximally to connect with the Xp11.23 interval previously characterized (Boycott *et al.*, 1996).

Characterization of the YACs that were isolated for the construction of the current contig demonstrated significant rates of chimerism and instability. While 26 of 76 YACs characterized in this study were chosen to represent the Xp11.3-p.23 region, the fifty other YACs were disregarded often due to chimerism or instability as determined by STS content and end-cloning experiments. Using the 26 selected YACs, rates of 19% chimerism, and 31% internal deletions were observed and 19% of YACs had heterogeneous sizes by pulse field analysis. These are minimal percentages as not all the ends of the clones could be assessed. If the 50 YACs not used in assembling the contig are taken into consideration, both the rates of instability and chimerism increase markedly. Using the same YAC libraries, a 3.6 Mb contig of Xp22.1 was constructed with reported rates of 11% chimerism and 6-22% internal deletions (Trump *et al.*, 1996). In contrast with Trump and co-workers (1996), also using the same libraries, a 2 Mb

contig constructed spanning Xp11.23-p11.22, reported a 30% rate of chimerism and 40-80% rate of internal deletions (Boycott *et al.*, 1996). Taken together, results from these studies suggest that differences in reported rates of chimerism and instability may be due to regional variations in clone stability.

The 250 kb interval between the markers EC3234R and EC1149R (intervals 12-14) in the central portion of the contig was the most troublesome in the construction of the YAC contig due to both chimerism and instability of the YACs. It has been reported that repeated sequences in mammalian DNA could provide the substrate for recombination and be one of the contributing factors to the incidence of chimerism in YACs (Rogner *et al.*, 1994). Furthermore, it has also been observed that regions with high GC and gene content may contain sequences difficult to clone in yeast (Larionov *et al.*, 1994). The results observed in my study could have been due to a combination of high GC content and a high number of repetitive elements. This region could be gene rich as four ESTs or genes (UBE1, UHX, PCTK1, and WI18364) have already been reported. Sequences from end-clones EC1909R and EC1228R obtained in this study were disregarded due to sequence homology with Alu or Line sequences. EC1909R is found within the aforementioned problem area indicating the presence of repetitive sequence. More analysis is necessary to determine if this is a region rich in repetitive elements.

PCR products for DXS8083 and EC3535R had distinct banding patterns (data not shown) which implied the presence of a possible duplication. The resulting PCR products of both DXS8083 and EC3535R, visualized on an ethidium bromide gel, showed two bands in X-only and genomic DNA. In all of the conventional hybrids which covered Xq, only the smaller band was seen for DXS8083, and only the larger for EC3535R. The opposite band was seen in the radiation-reduced and conventional hybrids specific for Xp11.3, specifically interval 10. These observations suggest that

PCR with markers for DXS8083 and EC3535R results in the amplification of DNA within interval 10 as well as amplification of sequence on Xq. These results are consistent with the hypothesis that markers DXS8083 and EC3535R amplify a sequence which is duplicated in the genome.

The YAC libraries that were screened by our collaborators at the Human Genome Centre in St. Louis represented more than 15 X-chromosome equivalents. The final contig (Figure 19) represents three-fold YAC coverage as determined by calculating the average number of hits per STS. If the unstable YACs that were characterized in the preliminary studies were also taken into consideration, screening of the E, F, and I libraries resulted in approximately nine-fold coverage of the region (data not shown). These results are the same as the number of clones detected per probe in another study characterizing these libraries (Nagaraja *et al.*, 1994). However, both my study and that of Nagaraja and coworkers had a significant discrepancy between the number of clones obtained per probe and that expected per probe. The F library, derived from human-hamster cell lines, yielded the expected number of clones but the E and I libraries, derived from cells with five and four X chromosomes respectively, yielded numbers of YACs corresponding to only two X equivalents (Nagaraja *et al.*, 1994). Since the E and I libraries were enriched in X chromosomal DNA, one would have expected more clones to be obtained using an X-chromosome specific probe compared to an autosomal specific probe. Instead, the results obtained showed that the E and I collections contain no more X YACs than are expected for autosomes. These results suggest that most of the X chromosome DNA in the hybrid cells used for the libraries is poorly cloned into YACs, possibly because only the single active X is well cloned and the DNA from the inactive X chromosomes is perhaps at a disadvantage for cloning because of special features like relatively tight coiling (Nagaraja *et al.*, 1994).

YAC clone integrity is not a novel problem and has been observed in many reported YAC constructs. In every collection of YACs reported to date, a fraction of the clones are chimeric (Nagaraja *et al.*, 1994). For this reason, work continues in developing a yeast strain which maximizes clone stability. Recent studies have shown that rearrangements and chimeras arise from recombination during or after transformation of the yeast host in YAC library construction (Larionov *et al.*, 1994). The group of RAD52 genes and the RAD1 gene have been shown to be involved in recombination in yeast (Game and Mortimer, 1974), (Klein, 1988). To reduce the frequency of chimeric YAC clones, alternative yeast hosts with DNA repair mutants such as *rad1* and *rad52* have been used for YAC cloning (Neil *et al.*, 1990). Construction of a YAC library using a *rad52* mutant strain MHY5201 resulted in a reported rate of 8% chimerism compared to a rate of 50% in wild type host AB1380 (Haldi *et al.*, 1994). Also, *rad51/rad52* and *rad1* mutant hosts were shown to stabilize deletion prone YACs containing DNA from the human colour vision locus of Xq28 (Khono *et al.*, 1994). More recently a *rad54* mutant, another member of the RAD52 epistasis group, was investigated. Although more research must be performed, this mutant appears to represents another improved YAC cloning host (Le and Dobsen, 1997). Whether any of these recombination deficient hosts will be effective in stabilizing clones from the Xp11 region remains to be established.

Since access to recombination-deficient yeast systems is still limited and their potential not entirely elucidated, the solution to the stability problem for this study was approached by changing to an alternative, more stable vector system. Many vector systems, such as the fosmid, P1, BAC, and PAC systems, have been reported as stable alternative vector systems. In this study PACs were collected to provide a better representation of the region.

In contrast to the yeast system, the PAC system has shown no problems with clone integrity. From the analysis of 15 PACs with 26 markers, no evidence of instability

or chimerism in these clones was observed in this study. Furthermore, in the isolation of three PAC end fragments, no chimeric ends were found. From these results it can be concluded that the PAC cloning system has no apparent problems handling high GC content, high gene content and tandem repeat sequences. This is likely due to the maintenance of genomic fragments as single copy plasmids in a recombination-deficient host.

Although a 3X PAC library, and in five instances a 7X library, was screened by 26 markers in this study, relatively few PAC clones were recovered. Screening with eight markers from region between markers EC1149R and EC3234R, previously mentioned as being a problem area in YACs, resulted in a cluster of eight PACs being obtained forming solid coverage for this region. However, the region flanked by the markers EC1149R and EC8057R was particularly troublesome with only one PAC being recovered after screening with 11 different markers. The quality of the probes could be questioned but, because they all worked well in Southern hybridizations (Figure 16), and because the problem also seems to be regional, an alternative explanation likely exists. It was suggested that the problem may stem from the fact that the de Jong pCYPAC library was constructed from male stem cells, hence reducing the number of copies of X chromosomes being represented (Jack Huiguenza, personal communication). This could potentially explain the limited number of clones obtained but does not explain why there also seemed to be a regional clustering of PACs. Lastly, it is possible, as observed in the YAC system, that certain regions of the genome are also more difficult to clone in PACs. There have been several published reports on using PACs to complement YAC or cosmid contigs (Osoegawa *et al.*, 1996; Sood *et al.*, 1997). In such cases PACs aided in adding resolution or filling in gaps that could not otherwise be filled.

Relatively few PAC contigs have been reported to date. Two of the larger PAC contigs published are a 1-1.5 Mb contig on 5q and a > 1 Mb contig of 2q13 (Nothwang *et*

*al.*, 1997; Zhao *et al.*, 1997). Zhao and co-workers describe a contig made up of 78 overlapping PAC clones and Nothwang and co-workers describe a contig with 34 PAC clones. In both cases at least one of the libraries used was the same as that used in this study. There were no reported problems with chimerism or instability and seemingly no problems obtaining clones in their library screening. However, it is difficult to compare these results with those obtained in this study since there are no PAC contigs yet reported on the X chromosome and cloning problems often tend to be regional. Thus more studies will have to be performed, particularly on the X chromosome, to elucidate the possible reasons for our limited success in recovering PACs in our library screens.

STS content in PACs was evaluated using either Southern hybridization or radioactive PCR. While attempts were made to determine STS content by non-radioactive PCR, the results were variable with varying band intensities and numerous false positives. Though not fully evaluated, it is possible that at some point during the isolation of the DNA or during the PCR preparation process genomic contamination may have been occurring. Another possibility is that the PCR conditions had not been optimized. In the two previously reported PAC contigs, Nothwang and co-workers determined STS content by radioactive PCR, whereas Zhao and co-workers used non-radioactive PCR methods (Nothwang *et al.*, 1997; Zhao *et al.*, 1997). No reasons are given for the choices made by each group. Regardless of the methods used to determine STS content in PACs, this data proved vital for ordering DNA markers and estimating sizes between markers.

Fortunately, the region found to be most troublesome in YACs was extensively cloned in PACs. PACs 233D18, 237A9, 121M13 and 250J21 allowed UHX, PCTK1 and UBE1 to be placed within the contig. Furthermore, a general order for the 14 markers between EC1149R and EC3234R was determined, although several markers still cannot

be oriented with respect to one another. All PACs obtained in this study will help to seed the Xp11.3-p23 region for the future development of a complete PAC contig.

Based on the physical map constructed of integrated PAC and YAC data, the marker order in the proximal region of this contig: Xpter-(DXS1264, DXS1055)-(DXS1003,UBE1-PCTK1)-DXS1266-DXS337-SYN1-ZNF21-cen is essentially in agreement with the consensus map of the 6th X Chromosome Workshop (Nelson *et al.*, 1995). While the consensus map places DXS1264 distal to DXS1055, my study placed both markers on the same PACs, though they did not order these markers; YACs in this region were too unstable to be of any assistance in this regard. Similarly, the markers UHX and PCTK1 were not ordered. Unlike the consensus map, my study placed UBE1, UHX and PCTK1 proximal to DXS1055 and DXS1264. Lastly, the most parsimonious order based on YAC data places DXS1266 proximal to DXS337.

At the proximal end, this contig provides overlapping coverage with the 12 published markers DXS1264, UBE1, DXS1055, DXS1003, DXS1146, DXS1266, DXS337, ARAF1, SYN1, TIMP1, DXS426, ZNF21 (Coleman *et al.*, 1994; Hagemann *et al.*, 1994; Knight *et al.*, 1994; Kwan *et al.*, 1995) and adds 11 markers proximal to DXS1264. At the distal end, the contig extends one megabase past UBE1 to EC7034R placing 13 markers into this region. A gap in the physical map still exists between EC7034 distal to DXS8083 and on to MAOA/MAOB and DXS7.

The physical contig developed from my studies represents to date the most robust physical coverage of 2 Mb within the Xp11.3 region, and will aid both in DNA sequencing and gene-finding efforts within this part of the human genome.

# CHAPTER 5 - SUMMARY

Although the YAC-based physical map of the X chromosome is largely complete, several regions have been especially difficult to cover in clones useful for sequencing and gene-finding. The region between the DNA markers DXS7 and ZNF21 in Xp11.3-p11.23, spanning about 7 cM, is one of the most difficult regions to clone in YACs due to the high incidence of chimerism and instability. This region is also of note because it overlaps with the critical regions for X-linked congenital stationary night blindness, retinitis pigmentosa and Åland Island eye disease.

This thesis research project extended the existing physical contig of the human Xp11.3 region between a new marker EC7034R and ZNF21 using YACs. Then, using existing markers, an initial set of PACs were collected to begin the construction of a DNA sequence-ready template of the region.

The mapping was initiated with a set of 19 markers sublocalized within Xp11 based on their retention in the panel of conventional and radiation hybrids described by Boycott et al. (1997). The genetic region between EC7034 and ZNF21 spans six intervals (9 to 14) on this mapping panel.

Subsequently, 19 additional region-specific STSs/STRs/ESTs were used to increase DNA marker density and to provide additional YAC coverage. Fourteen STSs were designed from selected end-fragments of YACs or PACs. The remaining five STSs were derived from published sources.

The sequence of one end of the insert of the YAC yWXD8058 revealed a TC dinucleotide repeat. Amplification of DNA isolated from 50 unrelated Caucasian females showed this repeat to be polymorphic with allele frequencies of 0.4 (313 bp), 0.16 (315 bp) and 0.44 (317 bp).

The use of YACs from the ICRFy900 series (D1151, C1228, A0120, A1220 and C1022) provide overlap with previously published contigs at the centromeric end of the

contig and extends the cloned region proximally to the Xp11.23 interval previously characterized in the lab of Dr. Bech-Hansen.

The 75 YACs that were characterized in this study showed extensive chimerism and frequent deletions overall. Even the 26 selected to represent the region showed minimum rates of 19% chimerism and 31% internal deletions, with heterogeneous sizes suggestive of deletions observed in pulsed-field gel electrophoretic analyses.

Fortunately, PAC clones have shown far less clone instability or chimerism. Among 15 PACs containing more than 20 STS markers, analysis of STS content, end-cloning, restriction digest comparisons, and sizing showed no evidence of instability or chimerism. In regions where the YACs were most uncertain, PAC clones were necessary to confirm the location of several markers and permitted the ordering of some markers that could not be oriented unequivocally in YACs.

The resulting 2 Mb contig is in agreement both with DNA marker retention in radiation-reduced and conventional hybrids and with SEGMAP analysis of STS/clone content. It spans the interval between markers ZNF21 and the new marker EC7034R in 26 YAC clones and 15 PAC clones . The physical contig presented represents to date the most robust physical coverage within the Xp11.3 region, and should aid both in sequencing and gene-finding efforts.

**REFERENCES**

Aldred, M. A., Teague, P. W., Jay, M., Bundey, S., Redmond, R. M., Jay, B., Bird, A. C., Battacharya, S. S. and Wright, A. F. (1994). Retinitis pigmentosa families showing apparent X linked inheritance but unlinked to the RP2 or RP3 loci. *Journal of Medical Genetics* **31**:848-852.

Alitalo, T., Kruse, T., Forsius, H., Eriksson, A. and de la Chapelle, A. (1991). Localization of the Åland Island eye disease locus to the pericentromeric region of the X chromosome by linkage analysis. *American Journal of Human Genetics* **48**:31-38.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410.

Ashworth, L. K., Batzer, M. A., Brandiff, B., Branscomb, E., de Jong, P., Garcia, E., Garnes, J. A., Gordon, L. A., Lamerdin, J. E. and Lennon, G. (1995). An integrated metric physical map of human chromosome 19. *Nature Genetics* **11**:422-427.

Barrett, J. H. (1992). Genetic mapping based on radiation hybrid data. *Genomics* **13**:95-103.

Bech-Hansen, N. T., Field, L. L., Schramm, A. M., Reedyk, M., Craig, I. W., Fraser, N. J. and Pearce, W. G. (1990). A locus for X-linked congenital stationary night blindness is located on the proximal portion of the short arm of the X chromosome. *Human Genetics* **84**:406-408.

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

Bech-Hansen, N. T. and Pearce, W. G. (1993). Manifestations of X-linked congenital stationary night blindness in three daughters of an affected male: demonstration of homozygosity. *American Journal of Human Genetics* **52**:71-77.

Bergen, A. A. B., Kestelyn, P., Leys, M. and Meire, F. (1994). Identification of a key recombinant which assigns the incomplete congenital stationary night blindness gene proximal to MAOB. *Journal of Medical Genetics* **31**:580-582.

Berger, W., van Duijnhoven, G., Pinckers, A., Smits, A., Ropers, H. H. and Cremers, F. (1995). Linkage analysis in a Dutch family with X-linked recessive congenital stationary night blindness (XL-CSNB). *Human Genetics* **95**:67-70.

Berson, E. L., Rosner, B. and Simonoff, E. (1980). Risk factors for genetic typing and detection in retinitis pigmentosa. *American Journal of Ophthalmology* **89**:763-775.

Blanchard, M. M., Taillon, M., P., Nowotny, P. and Nowotny, V. (1993). PCR buffer optimization with uniform temperature regimen to facilitate automation. *PCR Methods and Applications* **2**:234-240.

Bouffard, G. G., Idol, J. R., Braden, V. V., Iyer, L. M., Cunningham, A. F., Weintraub, L. A., Touchman, J. W., Mohr0Tidwell, R. M., Peluso, D. C., Fulton, R. S., Ueltzen, M. S., Weissenbach, J., Magness, C. L. and Green, E. D. (1997). A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Research* **7**:673-692.

Boycott, K. M., Halley, G., Schlessinger, D. and Bech-Hansen, N. T. (1996). A 2-Megabase physical contig incorporating 43 DNA markers in the human X chromosome at p11.23-p11.22 form ZNF1 to DXS255. *Genomics* **33**:488-497.

Boycott, K. M., Moore, B. J., Gratton, K. J., Stoddart, K. L., Roland, B. and Bech-Hansen, N. T. (1997). Integration of 101 DNA markers across human Xp11 using a panel of somatic cell hybrids. *Cytogenetics and Cell Genetics* **76**:223-228.

Bull, P. C., Thomas, G. R., Rommens, J. M., Forbes, J. R. and Wilson Cox, D. (1993). The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nature Genetics* **5**:327-337.

Carrel, L., Clemson, C. M., Dunn, J. M., Miller, A. P., Hunt, P. A., Lawrence, J. B. and Willard, H. F. (1996). X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme E1 and PCTAIRE-1 genes in human and mouse. *Human Molecular Genetics* **5**:391-401.

Cohen, D., Chumakov, I. and Weissenbach, J. (1993). A first-generation physical mp of the human genome. *Nature* **366**:698-700.

Coleman, M., Bhattacharya, S., Lindsay, S., Wright, A., Jay, M., Litt, M., Craig, I. and Davies, K. (1990). Localization of the microsatellite probe DXS426 between DXS255 on Xp and linkage to X-linked Retinitis Pigmentosa. *American Journal of Human Genetics* **47**:935-940.

Coleman, M., Nemeth, A., Campbell, L., Raut, C., Weissenbach, J. and Davies, K. (1994). A 1.8 Mb YAC contig in the Xp11.23: Identification of CpG islands and Physical mapping of CA repeats in a region of high gene density. *Genomics* **21**:337-343.

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

Collins, F. (1995). Positional cloning moves from perditional to traditional. *Nature Genetics* **9**:347-350.

Collins, F. and Galas, D. (1993). A new five-year plan for the U. S. Human Genome Project. *Science* **262**:43-46.

Collins, J. and Hohn, B. (1978). Cosmids: A type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proceedings of the National Acadamy of Sciences* **75**:4242-4246.

Collins, J. E., Cole, C. G., Smink, L. J., Garrett, C. L., Leversha, M. A., Soderlund, C. A., Maslen, G. L., Everett, L. A., Rice, K. M. and Coffey, A. J. (1995). A high-density YAC contig map of human chromosome 22. *Nature* **377**:367-379.

Cox, D. R. (1992). Radiation hybrid mapping. *Cytogenetics and Cell Genetics* **59**:80-81.

Cox, D. R., Burmeister, M., Price, R. E., Suwon, K. and Myers, R. M. (1990). Radiation hybrid mapping: A somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250**:245-250.

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

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.

Doggett, N. A., Goodwink, L. A., Tesmer, J. G., Meincke, L. J., Bruce, D. C., Clark, L. M., Altherr, M. R., Ford, A. A., Chi, H. C. and L., M. B. (1995). An integrated physical map of human chromosome 16. *Nature* **377**:335-365.

Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T., Bowden, D., Smith, D., Lander, E., Botstein, D., Akots, G., Rediker, K., Gravius, T., Brown, V., Rising, M., Parker, C., Powers, J., Watt, D., Kauffman, E., Bricker, A., Phipps, P., Muller-Kahle, H., Fulton, T.*et al.* (1987). A genetic linkage map of the human genome. *Cell* **51**:319-337.

Donnelly, A., Kozman, H., Gedeon, A. K., Webb, S., Lynch, M., Sutherland, G. R., Richards, R. I. and Mulley, J. C. (1994). A linkage map of microsatellite markers on the human X chromosome. *Genomics* **20**:363-370.

Drayna, D. and White, R. (1985). The genetic linkage map of the human X chromosome. *Science* **230**:753-8.

Duyk, G. M., Kim, S. W., Myers, R. M. and Cox, D. R. (1990). Exon trapping: A genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proceedings of the National Academy of Science* **87**:8995-8999.

Feiss, M., Fisher, R. A., Crayton, M. A. and Egner, C. (1977). Packaging of the bacteriophage chromosome: effect of chromosome length. *Virology* 77:281-293.

Foote, S., Vollrath, D., Hilton, A. and Page, D. C. (1992). The human Y chromosome: Overlapping DNA clones spanning the euchromatic region. *Science* **258**:60-66.

Gal, A., Schinzel, A., Orth, U., Fraser, N., Mollica, F., Craig, I., Kruse, T., Machler, M., Neugebauer, M. and Bleeker-Wagemakers, L. (1989). Gene of X-chromosomal congenital stationary night blindness is closely linked to DXS7 on Xp. *Human Genetics* **81**:315-318.

Game, J. C. and Mortimer, R. K. (1974). A genetic study of X-ray sensitive mutants in yeast. *Mutation research* **24**:281-292.

Glass, I. A., Good, P., Coleman, M. P., Fullwood, P., Giles, M. G., Lindsay, S., Nemeth, A. H., Davies, K. E., Willshaw, H. A., Fielder, A., Kilpatrick, M. and Farndon, P. A. (1993). Genetic mapping of a cone and rod dysfunction (Åland Island eye disease) to the proximal short arm of the human X chromosome. *Journal of Medical Genetics* **30**:1044-1050.

Goss, S. J. and Harris, H. (1975). New method for mapping genes in human chromosomes. *Nature* **255**:680-684.

Guyer, M. S. and Collins, F. S. (1995). REVIEW: How is the Human Genome Project doing, and what have we learned so far? *Proceedings of the National Academy of Science* **92**:10841-10848.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994). The 1993-94 Genethon human linkage map. *Nature Genetics* **7**:246-249.

Gyapay, G., Schmitt, K., Fizames, C., Lones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'Homme, J.-F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J. and Goodfellow, P. N. (1996). A radiation hybrid map of the human genome. *Human Molecular Genetics* **5**:339-346.

Hagemann, T., Surosky, R., Monaco, A., Lehrach, H., Rosen, F. and Kwan, S.-P. (1994). Physical Mapping in a YAC contig of 11 markers on the human X chromosome in Xp11.23. *Genomics* **21**:262-265.

Haldi, M., Perrot, V., Saumier, M., Desai, T., Cohen, D., Cherif, D., Ward, D. and Lander, E. (1994). Large human YACs constructed in a rad52 strain show a reduced rate of chimerism. *Genomics* **24**:478-484.

Hamada, H., Petrino, M., Kakunaga, T., Seidman, M. and Stollar, D. (1984). Characterization of genomic poly(dT-dG) . poly(dC-dA) sequences: structure, organization and conformation. *Molecular and Cellular Biology* **4**:2610-2621.

Hohn, B. and Collins, J. (1980). A small cosmid for efficient cloning of large DNA fragments. *Gene* **11**:291-298.

Hong, H.-K., Giora, R., Trucco, M. and Chakravarti, A. (1993). Dinucleotide repeat polymorphism at the DXS1146 locus. *Human Molecular Genetics* **2**:1078.

Huebner, K., Druck, T., Croce, C. M. and Thiesen, H. J. (1991). Twenty-seven nonoverlapping zinc finger cDNAs from human T cells map to nine different chromosomes with apparent clustering. *American Journal of Human Genetics* **48**:726-740.

Hwu, H., Roberts, J., Davidson, E. and Britten, R. (1986). Insertion and/or deletion of many repeated DNA sequences in human and higher ape evolution. *Proceedings of the National Academy of Sciences* **83**:3875-3879.

Ingle, C., Williamson, R., de la Chapelle, A., Herva, R., Haapala, K., Bates, G., Willard, H., Pearson, P. and Davies, K. (1985). Mapping DNA seqences in a human X chromosome deletion which extends across the region of the Duchenne muscular dystrophy mutation. *American Journal of Human Genetics* **27**:451-462.

Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A. and de Jong, P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genetics* **6**:84-89.

Kere, J., Nagaraja, R., Mumm, S., Ciccodicola, A., D'Urso, M. and Schlessinger, D. (1992). Mapping human chromosomes by walking with sequence-tagged sites from end fragments of yeast artificial chromosome inserts. *Genomics* **14**:241-248.

Khono, K., Wada, M., Schlessinger, D., D'Urso, M., Tanabe, S., Oshiro, T. and Imamoto, F. (1994). Stability of YACs containing Ribosomal or RCP/GCP locus DNA in wild-type *S.cervisiae* and RAD mutant strains. *DNA Research* **1**:191-199.

Kim, U.-J., Shizuya, H., de Jong, P. J., Birren, B. and Simon, M. (1992). Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Research* **20**:1083-1085.

Klein, H. (1988). Different types of recombination events are controlled by the RAD1 and RAD52 genes of *Saccharomyces cervisia*. *Genetics* **120**:367-377.

Knight, J. C., Grimaldi, G., Thiesen, H.-J., Bech-Hansen, N. T., Fletcher, C. D. M. and Coleman, M. P. (1994). Clustered organization of Kruppel Zinc-Finger genes at Xp11.23, flanking a translocation breakpoint at OATL1: a physical map with locus assignment for ZNF21, ZNF41, ZNF81, and ELK1. *Genomics* **21**:180-187.

Korenberg, J. and Rykowski, M. (1988). Human genome organization: Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell* **53**:391-400.

Kwan, S. P., Hagemann, T. L., Blaese, R. M. and Rosen, F. S. (1995). A high-resolution map of genes, microsatellite markers, and new dinucleotide repeats from UBE1 to the GATA locus in the region Xp11.23. *Genomics* **29**:247-251.

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

Larionov, V., Kouprina, N., Nikolaishvili, N. and Resnick, M. A. (1994). Recombination during transformation as a source of chimeric mammalian artificial chromosomes in yeast (YACs). *Nucleic Acids Research* **22**:4154-4162.

Le, Y. and Dobsen, M. J. (1997). Stabilization of yeast artificial chromosome clones in a rad54-3 recombination -deficient host strain. *Nucleic Acids Research* **25**:1248-1253.

Lindsay, S. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327**:336-338.

Litt, M. and Luty, J. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**:397-401.

Lovett, M., Kere, J. and Hinton, L. M. (1991). Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proceedings of the National Academy of Science* **88**:9628-9632.

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. *American Journal of Human Genetics* **46**:776-783.

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.

Manuelidis, L. (1978a). Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma* **66**:23-32.

Manuelidis, L. (1978b). Comples and simple sequences in human repeated DNAs. *Chromosoma* **66**:1-21.

Marshall, E. (1995). Emphasis turns from mapping to large-scale sequencing. *Science* **268**:1270-1271.

McGuire, R., Sullivan, L., Blanton, S., Church, M., Heckenlively, J. and Daiger, S. (1995). X-linked dominant cone-rod degeneration: linkage mapping of a new locus for retinitis pigmentosa (RP15) to Xp22.13-p22.11. *American Journal of Human Genetics* **57**:87-94.

McKusick. (1986). Mendelian inheritance in man - Catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes. The John Hopkins University Press, USA

Meindl, A., Dry, K., Herrmann, K., Manson, F., Ciccodicola, A., Edgar, A., Carvalho, M. R. S., Achatz, H., Hellebrand, H., Lennon, A., Migliaccio, C., Porter, K., Zrenner, E., Bird, A., Jay, M., Lorenz, B., Wittwer, B., D'Urso, M., Meitinger, T. and Wright, A. (1996). A gene (RPGR) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). *Nature Genetics* **13**:35-42.

Monaco, A. (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* **323**:646-650.

Musarella, M. A., Burghes, A., Anson-Cartwright, L., Mahtani, M. M., Argonza, R., Tsui, L. C. and Worton, R. (1988). Localization of the Gene for X-linked recessive type of retinitis pigmentosa (XLRP) to Xp21 by linkage analysis. *American Journal of Human Genetics* **43**:484-494.

Musarella, M. A., Weleber, R. G., Murphey, W. H., Young, R. S. L., Anson-Cartwright, L., Mets, M., Kraft, S. P., Polemeno, R., Litt, M. and Worton, R. G. (1989). Assignment of the gene for complete X-linked congenital stationary night blindness (CSNB1) to Xp11.3. *Genomics* **5**:727-737.

Nagaraja, R., Kere, J., MacMillan, S., Mochubeloa, M., Johnson, D., Molini, B., Halley, G., Wein, K., Trusgnich, M., Eble, B., Railey, B., Brownstein, B. and Schlessinger, D. (1994). Characterization of four human YAC libraries for clone size, chimerism, and X chromosome sequence representation. *Nucleic Acids Research* **22**:3406-3411.

Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Griffin, S., Jermak, C., Hott, C., Masisi, M., Mumm, S., Srivastava, A., Pilia, G., Featherstone, T., Mazzarella, R., Kesterson, S., McCauley, B., Railey, B., Burough, F., Nowotny, V., D'Urso, M., States, D., Brownstein, B. and Schlessinger, D. (1997). X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Research* **7**:210-222.

Nechiporuk, T., Nechiporuk, A., Guan, X., Frederick, R., Figueroa, K., Chumakov, I., Korenberg, J. R., de Jong, P. J. and Pulst, S. M. (1996). Identification of three new microsatellite markers in the spinocerebellar ataxia type 2 (SCA2) region and 1.2 Mb physical map. *Human Genetics* **97**:462-267.

Neil, D., Villasante, A., Fisher, R., Vetrie, D., Cox, B. and Tyler-Smith, C. (1990). Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic acids research* **18**:1421-1428.

Nelson, D. L., Ballabio, F., Cremers, F., Monaco, A. P. and Schlessinger, D. (1995). Report of the sixth international workshop on X chromosome mapping 1995. *Cytogenetics and Cell Genetics* **71**:307-342.

Nothwang, H. G., Strahm, B., Denich, D., Kubler, M., Schwabe, J., Gingrich, J., Jauch, A., Cox, A., Nicklin, M., Kurnit, D. and Hildebrandt, F. (1997). Molecular cloning of the interleukin-1 gene cluster: construction of an integrated YAC/PAC contig an a partial transcriptional map in the region of chromosome 2q13. *Genomics* **41**:370-378.

Nussbaum, R. L., Lewis, R. A., Lesko, J. G. and Ferrell, R. (1985). Mapping X-linked opthalmic disease: II.Linkage relationship of X-linked retinitis pigmentosa to X chromosomal short arm markers. *Human Genetics* **70**:45-50.

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

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

Osoegawa, K., Susukida, R., Okano, S., Kudoh, J., Minoshima, S., Shimizu, N., de Jong, P. J., Groet, J., Ives, J., Lehrach, H., Nizetic, D. and Soeda, E. (1996). An integrated map with cosmid/PAC contigs of a 4Mb Down Syndrome critical region. *Genomics* **32**:375-387.

Ott, J., Bhattacharya, S., Chen, J. D., Denton, M. J., Donald, J., Dubay, C., Farrar, G. J., Fishman, G. A., Frey, D., GAl, A., Humphries, P., Jay, B., Jay, M., Litt, M., Machler, M., Musarella, M., Neugebauer, M., Nussbaum, R. L., Terwilliger, J. D., Weleber, R. G., Wirth, B., Wong, F., Worton, R. G. and Wright, A. F. (1990). Localizing multiple X chromosome-linked retinitis pigmentosa loci using multilocus homogeneity tests. *Proceedings of the National Acadamy of Science* **87**:710-704.

Pierce, J. C., Sauer, B. and Sternberg, N. (1992). A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: Improved cloning efficacy. *Proceedings of the National Academy of Sciences USA* **89**:2056-2060.

Raeymaekers, P., Van Zand, K., Lin, J., Hoglund, M., Cassiman, J.-J., Van Den Berghe, H. and Marynen, P. (1995). A radiation hybrid map with 60 loci covering the entire short arm of chromosome 12. *Genomics* **29**:170-178.

Rao, V. B., Thaker, V. and Black, L. W. (1992). A phage T4 in vitro packaging system for cloning long DNA molecules. *Gene* **113**:25-33.

Rodriguez-Tome, P. and Lijnzaad, P. (1997). The radiation hybrid database. *Nucleic Acids Research* **25**:81-84.

Roepman, R., Bauer, D., Rosenberg, T., van Duijnhoven, G., van de Vosse, E., Platzer, M., Rosenthal, A., Ropers, H.-H., Cremers, F. and Berger, W. (1996a). Identification of a gene disrupted by a microdeletion in a patient with X-linked retinitis pigmentosa (XLRP). *Human Molecular Genetics* **5**:827-833.

Roepman, R., van Duijnhoven, G., Rosenberg, T., Pinkers, A., Bleeker-Wagemakers, L., Bergen, A., Post, J., Beck, A., Reinhardt, R., Ropers, H.-H., Cremers, F. and Berger, W. (1996b). Positional cloning to the gene for X-linked retinitis pigmentosa 3: homology with the guanine-nucleotide-exchange factor RCC1. *Human Molecular Genetics* **5**:1035-1041.

Rogner, U., Kioschis, P., Wilke, K., Gong, W., Pick, E., Dietrich, A., Zechner, U., Hameister, H., Pragliola, A., Herman, G., Yates, J., Lehrach, H. and Poutska, A. (1994). A YAC clone map spanning 7.5 megabases of human chromosome band Xq28. *Human Molecular Genetics* **3**:2137-2146.

Sambrook, J., Fristch, E. F. and Maniatis, T. (1989). Molecular Cloning. In: Nolan C. (ed). Vol. 2, Cold Spring Harbor Laboratory Press, pp 14.0-14.33.

Sapru, M., Gu, J., Gu, X., Smith, D., Yu, C.-E., Wells, D. and Wagner, M. (1994). A panel of radiation hybrids of human chromosome 8. *Genomics* **21**:208-216.

Schmitt, H., Kim, U.-J., Slepak, T., Blin, N., Simon, M. I. and Shizuya, H. (1996). Framework for a physical map of the human 22q13 region using bacterial artificial chromosomes (BACs). *Genomics* **33**:9-20.

Schuler, G. D., Boguski, M. S., Stewart, E. A., Stein, G., Gyapay, G., Rice, K., White, R. E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B. B., Butler, A., Castle, A. B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P. J. R., Dibling, R., Druouot, N., Dunham, I., Duprat, S., East, C.*et al.* (1996). A gene map of the human genome. *Science* **274**:540-563.

Schwartz, M. and Rosenberg, T. (1991). Åland eye disease: linkage data. *Genomics* **10**:327-332.

Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992). Cloning and stable maintenance of 300-kilbase-pair fragments of human DNA in Escherichia-coli using an F-factor-based vector. *Proceedings of the National Academy of Sciences USA* **89**:8794-8797.

Singer, M. and Skowronski, J. (1985). Making sense out of LINES: long interspersed repeat sequences in mammalian genomes. *Trends in Biological Science* **10**:119-122.

Sood, R., Blake, T., Aksentijevich, I., Wood, G., Chen, X., Gardner, D., Shelton, D., Mangelsdorf, M., Orsborn, A., Pras, E., Balow, J., Centola, M., Deng, Z., Zaks, N., Chen, X., Richards, N., Fischel-Ghodsian, N., Rotter, J., Pras, M., Shohat, M., Deaven, L., Gumucio, D., Callen, D., Richards, R.*et al.* (1997). Construction of a 1-Mb restriction-mapped cosmid contig containing the candidate region for the familial mediterranean fever locus on chromosome 16p13.3. *Genomics* **42**:83-95.

Spurr, N. K. (1995). Navigating the Genome. *Biochemical Society Transactions* **24**:285-289.

Sternberg, N. (1990). Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proceedings of the National Academy of Sciences USA* **87**:103-107.

Strachan, T., Abitbol, M., Davidson, D. and Beckmann, J. (1997). A new dimension for the human genome project: towards comprehensive expression maps. *Nature Genetics* **16**:126-132.

Swanson, D. A., Freund, C. L., Ploder, L., McInnes, R. R. and Valle, D. (1996). A ubiquitin C-terminal hydrolase gene on the proximal short arm of the X chromosome: implications for X-linked retinal disorders. *Human Molecular Genetics* **5**:533-538.

Thiselton, D. L., Hampson, R. M., Nayudu, M., Van Maldergem, L., Wolf, M. L., Saha, B. K., Bhattacharya, S. S. and Hardcastle, A. J. (1996). Mapping the RP2 locus for X-linked retinitis pigmentosa on proximal Xp: a genetically defined 5-cM critical region and exclusion of candidate genes by physical mapping. *Genome Research* **6**:1093-1102.

Tiranti, V., Savoia, A., Forti, F., D'Apolito, M.-F., Centra, M., Rocchi, M. and Zeviani, M. (1997). Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Human Molecular Genetics* **6**:615-625.

Trump, D., Pilia, G., Dixon, P. H., Wooding, C., Thakrar, R., Leigh, S. E. A., Nagaraja, R., Whyte, M. P., Schlessinger, D. and Thakker, R. V. (1996). Construction of a YAC contig and an STS map spanning 3.6 Mb in Xp22.1. *Human Genetics* **97**:60-68.

Venter, C., Smith, H. and Hood, L. (1996). A new strategy for genome sequencing. *Nature* **381**

Walter, M. A., Spillett, D. J., Thomas, P., Weissenback, J. and Goodfellow, P. N. (1994). A method for constructing radiation hybrid maps of whole genomes. *Nature Genetics* **7**:22-28.

Weber, J. L. and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**:388-396.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. and Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature* **359**

Weleber, R., Pillers, D.-A., Powell, B., Hanna, C., Magenis, E. and Buist, N. (1989). Aland island eye disease (Forsius-Eriksson Syndrome) associated with contiguous deletion syndrome at Xp21: Similarity to Incomplete congenital stationary night blindness. *Archives of Ophthalmology* **107**:1170-1179.

Willard, H. (1985). Chromosome-specific organization of human alpha satellite DNA. *American Journal of Human Genetics* **37**:524-532.

Wirth, B., Denton, M., Chen, J.-D., Neugebauer, M., Halliday, F. B., van Schooneveld, M., Donald, J., Bleeker-Wagemakers, E. M., Pearson, P. L. and Gal, A. (1988). Two Different genes for X-Linked retinitis pigmentosa. *Genomics* **2**:263-266.

Wolfsberg, T. and Landsman, D. (1997). A comparison of expressed sequence tags (ESTs) to human genomic sequences. *Nucleid Acids Research* **25**:1626-1632.

Zhao, N., Stoffel, A., Wang, P., Eisenbart, J., Espinosa, R., Larson, R. and LeBeau, M. (1997). Molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseased to 1-1.5 Mb and preparation of a PAC - based physical map. *Proceedings of the National Academy of Sciences* **94**:6948-6953.