NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

LATERAL GENETICS: AN APPROACH TO THE CLONING OF POTENTIAL HUMAN CHORIO-RETINAL X-LINKED GENES

PAUL W. WONG

Thesis submitted to the Department of Biochemistry
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

University of Ottawa
Ottawa, Ontario, Canada

ISBN 0-315-75081-2

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

At present the cloning of disease genes relies on the application of forward and reverse genetics which requires prior knowledge of either the protein or the location of the gene locus underlying the disorder. In the present study a third alternative, a lateral approach, which focuses in at the level of expressed sequences (mRNA) is examined. This approach is used in the present study to explore its potential in the isolation of X-linked chorio-retinopathy candidate genes.

The strategy developed for this approach uses a probe derived by the PCR amplification of a chorio-retinal cDNA library to screen a hamster/human X hybrid genomic library made in λ phage vectors. The immediate achievement of this cross screen is the direct association of tissue expression with chromosome specificity. The initial application and evaluation of this approach is presented in the following thesis.

DEDICATION

The road from start to finish of this present journey has been a long, dark and rocky one full of unexpected obstacles and distractions. I am eternally grateful for the love and support of my family and friends during this period of insanity who have so valiantly kept me on the straight and narrow and who have so courageously fought to keep the trolls at bay. A smile and a friendly voice is often food for the soul in a cold world. No man is an island.... but then of course who would want to be a piece of dirt in the middle of the ocean?

Simply put, this thesis is dedicated to my family and friends who so carefully surround me and who I cherish.

ACKNOWLEDGEMENTS

For the duration of the present project I have been truly blessed by the generosity of others. I would like to thank Dr. M. Brennan (Bethesda, Maryland) for so kindly providing an aliquot of human genomic DNA cross linked to cellulose and Dr. J. Nathans (Baltimore, Maryland) for making the HS7 plasmid available. I would like to thank Dr. J.L. Hamerton and Dr. C. Gregory (Winnipeg, Manitoba) for their overwhelming generosity in providing the 82082a and 1103 cell lines, McCoy's 5A medium, and DNA for the human X-panel. I would also like to thank the RP foundation, Medical Research Council of Canada and OGS for their support.

I am extremely grateful to the many individuals who have so generously provided choriod and retina samples and I would like to thank Dr. I. MacDonald for his persistence in making these human choroid/retinas available, his patience and his support.

I am deeply indebted to the cast of crazys both past and present who haunt the Tenniswood fold for making this particular basketcase feel so much at home. You are all rather a tolerant lot and I thank you all for putting up with the loud music, the deranged laughter, the hand, the sarcasm, and the bubble gum. I thank you all for sharing your thoughts and expertise and rest assured I have learned something special from each and everyone of you.

In particular, I would like to thank Ken Lawless, Mike Montpetit, and Jocelyne Leger for taking the time and having the patience to break in a naive and ignorant non-molecular biologist into this rather overwhelming field.

I am truly grateful to Tim Gleeson for his kind help in the S1 nuclease protection experiments and I would like to especially thank Richard Pilon for his invaluable help and aid in the later stages of the project. I am indebted to you both. I am eternally grateful to Mike Montpetit, Jocelyne Leger, Michele Rouleau, Sean Guenette, Valerie Weaver and France Croze for their support and numerous brain storming/therapy sessions. I am especially thankful to H.S. Wang for her insightful queries and the soul searching sessions. Thanks guys, paid professionals could not have done a finer job! I also wish to thank Jean Pineault, Jonathon Lakins, Marilyn Mooibroek, Lars Daehlin, and Dr. Ruth for lending the more than occasional ear (you are all much too kind....or deaf?). Most of all I wish to thank all the above for their friendship and smiles.

Last but not least I would like to sincerely thank my supervisor, Dr. Martin Tenniswood, for his uncanny ability to give direction without pointing, his wisdom in not accepting the towel at times when I was ready to throw it in, for being assertive but not pushy, the bad jokes and the continual re-telling of them, for understanding the concept of positive procrastination and thought experiments, his overwhelming generosity, enthusiastic support, unwavering patience, trust and f[illegible]h. (Can I have a raise?) Live long and prosper!

TABLE OF CONTENTS

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

CHAPTER EIGHT

CHAPTER NINE

CHAPTER TEN

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.000)INTRODUCTION

The human choroid and retina are affected by many genetically inherited disorders, at least 30 of which are known to be located on the human X-chromosome. Our initial interest in retinal expressed sequences arose from family studies examining two distinct X-linked retinal degenerative disorders, choroideremia and Norrie's disease (MacDonald *et al.*, 1987a, b; Polomeno *et al.*, 1987; Wong *et al.*, 1988).

In general the symptoms and manifestations of chorio-retinal disorders are highly heterogeneous, however, a common theme is an impairment of the visual process. A better understanding of individual genetic disorders that affect the retina will further the understanding of the biology and biochemistry of the retina. The present study was designed to further our understanding of the choroid and retina by identifying X-specific expressed sequences in the tissue, a first step in the application of a candidate gene approach to the isolation of genes underlying known X-linked chorio-retinal disorders.

1.100) GENES AND ALLELIC VARIATION

One of the cornerstones of genetic dogma is that an active structural gene codes for an RNA transcript, which in turn gives rise to a functional protein. Each gene will have a specific location in the genome on a specific chromosome, a gene locus. Variations within an individual gene may arise through mechanisms such as

point mutations and recombination. If these changes occur in the germ line, they can be passed on to future generations. Thus different variants of a specific gene may occupy the same gene locus on independent chromosomes of the same type. Individual genes which occupy the same gene location are said to be "allelic". Individuals who have two alternate alleles at a specific autosomal gene locus (one on each of the chromosomes of an autosome pair) are said to be heterozygous at that gene locus; in cases where both alleles are identical, the individual is said to be homozygous. With respect to X specific gene loci, normal females have the potential to be genetically homozygous or heterozygous while normal males are almost always hemizygous (since they have only one copy of the gene).

Changes in the genetic code which are reflected in the structural gene product may affect the quality or the quantity of the finished gene product which in turn may result in the formation of a normal variant or abnormal variant. If the abnormal variant differs significantly from normal function then the inheritance of such a gene may have pathological and/or clinical consequences.

1.200) CLINICAL PERSPECTIVES

In most cases the understanding of a given genetic disorder is limited. In the very simplest cases, genetic disorders are inherited directly as either a mendelian dominant or recessive trait. A recessive condition arises when a heterozygous condition (with the presence of the disease allele) at the disease locus does not give

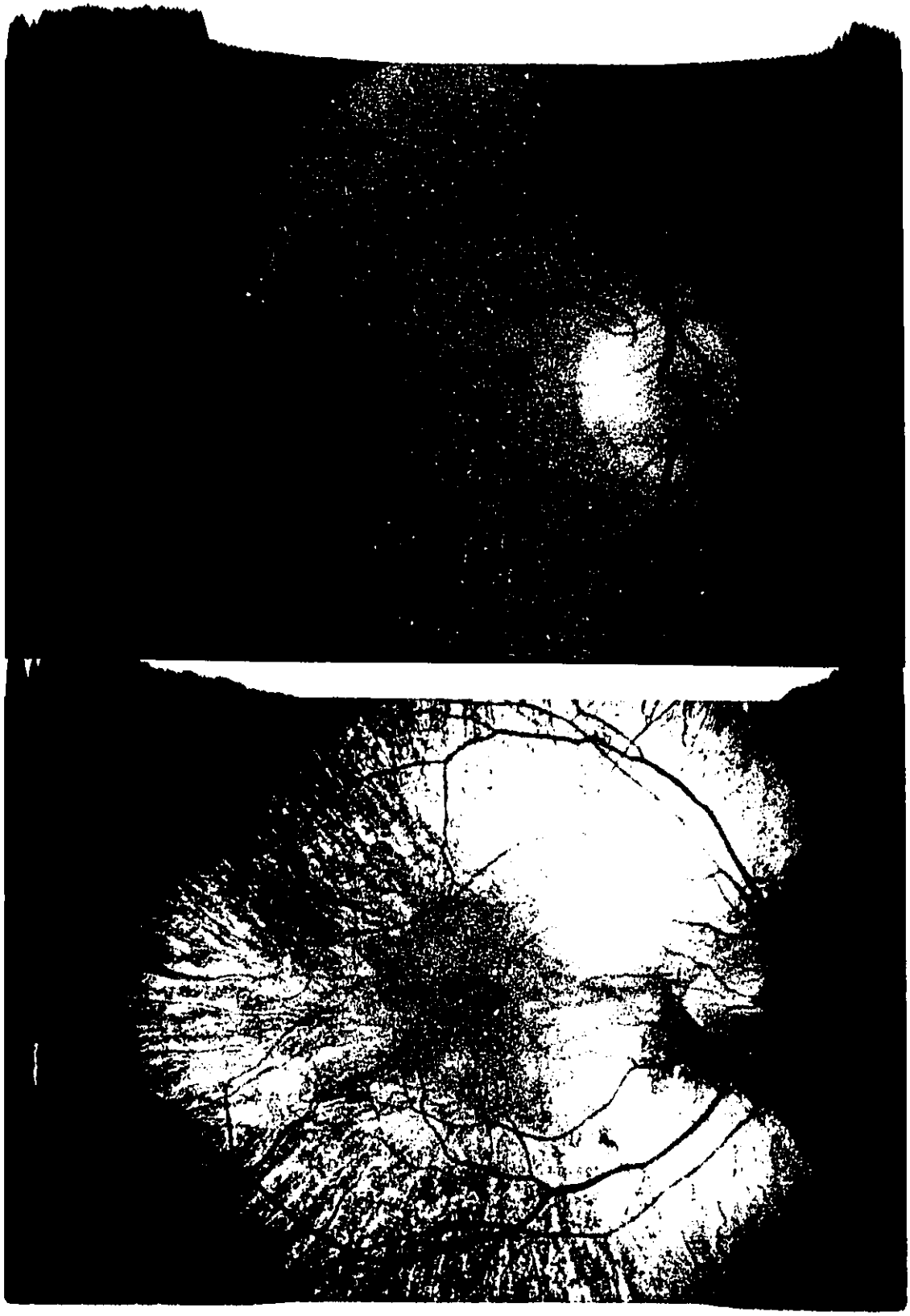
rise to the disorder. A dominant trait leads to the disease condition regardless of heterozygosity or homozygosity of the disease gene, while recessive traits will arise only if the disease allele is homozygous or hemizygous. Often the segregation of X-linked and autosomal traits follows a distinct pattern of inheritance and careful family analysis of specific disorders is sufficient to conclude or refute localization to the X-chromosome.

Hereditary chorio-retinal dystrophies encompass a clinically diverse group of genetically distinct disorders that play a primary role in the etiology of blindness in man. The manifestation and progression to blindness in disorders such as retinitis pigmentosa (RP), choroideremia, and Norrie's disease are wide and varied. The basic scenario is that a normal retina (Fig. 1A) degenerates, as in the case of choroideremia (Fig. 1B), to the state that it is no longer functional or present. Similarly there appears to be a great heterogeneity even within a given disorder. RP for example defines a disorder that is characterized by progressive degeneration of the retina, leading to loss of visual field, night blindness, and ultimately complete loss of sight (Inglehearn *et al.*, 1990). There are however X-linked, autosomal dominant, and autosomal recessive forms, with phenotypic heterogeneity in all categories. At present there are believed to be at least 4-5 distinct RP loci in the human genome. The familial mode of inheritance and the time of disease onset are both criteria for the clinical evaluation of the subtype of RP.

FIG. 1 NORMAL AND DISEASED HUMAN RETINAS.

Panel A: Photograph of the fundus of the right eye of a normal individual with an intact retina and retinal pigment epithelium (orange). The central spot is a fixation device artifact.

Panel B: Photograph of the fundus of the right eye of a male with Choroideremia showing widespread choriod/retinal atrophy with sparing of the macula.



The clinical profile of any specific disorder is important in the diagnosis and management of the disorder, however it does not necessarily provide a solid foundation for elucidating the underlying mechanism. In most disease states the observable features of the disease at the clinical level are the end products of the underlying lesion and not necessarily the lesion itself. Clinical evaluation does, however, provide evidence that the disease locus exists.

1.300) THE BIOCHEMICAL APPROACH

In the past twenty years extensive efforts have been made to study genes underlying genetic disorders at the DNA level. The primary lesion underlying most genetic disorders is not known. There are two approaches that are used to clone and characterize a disease locus at the DNA level: forward and reverse genetics.

1.310) Forward Genetics

Until recently most methods for cloning specific genes have followed the strategy of forward genetics, which relies on the prior characterization of the protein. By knowing the structure of the protein it is possible to devise strategies to isolate the corresponding gene. With respect to known human genetic disorders the approach of forward genetics has found a direct application in the characterization of inborn errors of metabolism. In these genetic disorders the aberrant biochemical metabolic defect has been well characterized and in a majority of cases the protein responsible has been purified.

Gyrate atrophy, for example, is a rare autosomally inherited chorio-retinal degeneration caused by deficiency of ornithine amino transferase which leads to blindness (Simell and Takki, 1973). Ornithine amino transferase is constitutively expressed in most tissues, but is virtually absent in all tissues of patients with gyrate atrophy. This particular transferase has been purified and biochemically characterized (Sanada *et al.*, 1976; Lyons *et al.*, 1976; Smith *et al.*, 1977); it is important in the intracellular production of glutamate and proline from ornithine (Volpe *et al.*, 1969), and in shuttling excess dietary amino acid carbons to the gluconeogenic pathway via the tricarboxylic acid cycle (Boernke *et al.*, 1981; Shih, 1981). The ornithine amino transferase gene was subsequently isolated using anti-human ornithine amino transferase antibodies raised against purified human liver ornithine amino transferase to screen a liver cDNA λgt11 library (Innana *et al.*, 1986; Mitchell *et al.*, 1988).

Unfortunately for most human genetic disorders the gene product or products responsible for the individual disorder have not yet been identified and in these cases the application of forward genetics is futile.

1.320) Reverse Genetics

What is observed and recognized in most genetic disorders are the manifestations of the disorder and not the primary lesion which underlies it. The isolation of such disease genes, without prior insight into the structural, physical, or functional aspects of the primary gene or gene product, has relied heavily on the

localization of the putative disease locus to the human genetic map by standard mapping procedures and subsequent application of reverse genetics. Once a disease locus has been mapped the basic strategy has been to focus on the isolation and characterization of genomic sequences in the vicinity. Thus the ability to limit the region of localization by the identification of patients with chromosomal anomalies such as deletions or translocations, fine physical mapping, or the identification of closely linked flanking markers expedites the search. Each step of the process (the study of the disorder at the clinical level, the genetic mapping, and the eventual molecular isolation and characterization of a putative disease gene) however is intensely time consuming. This strategy focuses on the identification of a specific locus without prior knowledge of what it is, except that an allele(s) at this locus gives rise to a disease state.

1.321) Genetic Mapping

The initial step in the reverse genetics approach relies on the ability to map a specific disease gene locus to a specific region of the genome. The localization of a gene to a small region of a single chromosome greatly limits the amount of genetic material that must be screened to isolate the gene.

Although there are several ways to map a given marker such as exclusion mapping, deletion mapping, irradiation hybrid mapping, somatic cell panel mapping, and chromosomal *in situ* hybridization, the task of systematically mapping a disease

locus to a specific region of the genome is largely reserved for linkage mapping. Genetic markers with known locations (marker loci) are tested for linkage with the disease gene of interest. These markers may correspond to known genes or to polymorphisms that do not necessarily correspond to a gene. In the past, the available markers used for linkage studies were limited to blood protein polymorphisms, enzyme polymorphisms, chromatin polymorphisms and the number of markers was not large enough to provide a complete representation of the genome. The development of polymorphic DNA markers has revitalized this procedure.

Linkage is the occurrence of two loci sufficiently close together on the same chromosome so that their alleles do not assort independently. The distance between these two loci can be measured indirectly by the amount of recombination that occurs between them. It is generally assumed that the more distant two loci are, the more probable that recombination will occur between them. Thus alleles at loci which are far apart on the same chromosome, or on different chromosomes tend to assort independently.

The rationale in applying classical linkage analysis to gene mapping is as follows: if a gene, A, is localized to a specific region, W, of chromosome K and it is shown that gene B is closely linked to gene A, then logic dictates that gene B must also be situated near area W of chromosome K.

Linkage data are obtained from family studies and focus on the segregation of alleles at specific loci or markers in relation to other loci or markers. Non-allelic genes which consistently segregate together are probably linked. Matings in which clear segregation of the alleles at two loci is evident provide informative linkage data whereas matings in which clear segregation cannot be observed are not informative. Linkage between two loci is determined mathematically by calculating a LOD score (log of odds) which compares the likelihood of obtaining the observed data under the assumption of various values of recombination between two loci with those which would be obtained if there were independent assortment of alleles at those loci.

Mapping studies have served to localize a number of ophthalmic disorders including X-linked RP to Xp11 (Bhattacharya *et al.*, 1984), Norrie's disease to Xp11 (Bleeker-Wagemaker *et al.*, 1985), choroideremia to Xq21 (Nussbaum *et al.*, 1985), Lowe's syndrome to Xq24-26 (Silver *et al.*, 1986), blue monochromy to Xq28 (Lewis *et al.*, 1987), retinoschisis to Xp22 (Altalo *et al.*, 1989), stationary night blindness to Xp11.3 (Mussarella *et al.*, 1989), autosomal retinitis pigmentosa to 3q (McWilliam *et al.*, 1989), and Nance Horan syndrome to Xp22.2-22.3 (Lewis *et al.*, 1990).

1.322) Isolation Strategies

The ideal mapping situation prior to attempts at molecular cloning is to have two separate markers which are closely linked to the disease locus of interest and

which flank the disease locus. The usual strategy in this scenario has been to begin chromosomal walking or jumping from one marker towards the other. Chromosomal walking involves the isolation of a small DNA fragment from one end of a genomic recombinant clone(a) and its use to rescreen a genomic library to isolate a second clone(b) containing additional DNA sequences adjacent but not in clone(a)(Cantor, 1984).

Chromosomal jumping involves the use of a jumping library (Collins *et al.*, 1984). Inserts of clones in these libraries consists of a genomic fragment (A), a selectable marker and a second genomic fragment (B) that is not adjacent to A but is located a sizable distance upstream or downstream from A. Isolation of clones from these libraries therefore enables the isolation of two sequences that are located in the genome a significant distance from each other, a jump. In either case, walking or jumping, the DNA spanning the 2 flanking markers is systematically cloned to isolate the gene in question.

The use of linkage data and the genetic map as sign posts to begin the isolation of a gene has several inherent risks. The first and foremost is the fact that the linkage map and the physical map do not always coincide. Because the linkage map is assigned on the basis of recombination rates, which are known to vary from genomic region to genomic region, it is possible that loci which are close on the genetic map may in fact be far on the physical map. Conversely it is possible that

what appears to be genetically distant be in reality physically close. Regardless loci which are genetically linked are physically syntenic, that is located on the same chromosome.

The identification of chromosomal abnormalities in the region of interest has often aided the search of a specific gene. If the individual with the chromosomal abnormality also suffers from the disorder under study this provides an independent confirmation of the localization of the disease gene locus. In addition the deletion or translocation breakpoints will provide a physical boundary for the gene locus. The identification of such translocations and micro-deletions have already played a major role in the isolation of several disease loci by reverse genetics: these include the isolation of the retinoblastoma gene (Friend *et al.*, 1986), Duchenne muscular dystrophy gene (Monaco *et al.*, 1986), and chronic granulomatous disease (Royer-Pokora *et al.*, 1986). Strangely enough, only the isolation of the cystic fibrosis gene has been accomplished purely on the basis of reverse genetics without the use of defined chromosomal aberrations (Rommens *et al.*, 1989; Riordan *et al.*, 1989; Kerem *et al.*, 1989).

Chromosomal abnormalities have been identified in a number of eye disorders including Lowe's syndrome (Hodgson *et al.*, 1986), Norrie's disease (Bleeker-Wagemaker *et al.*, 1985; Donnai *et al.*, 1988; De La Chapelle *et al.*, 1985; Ohba *et al.*, 1986; and Zhu *et al.*, 1989), X-linked RP (Franke *et al.*, 1985; de Saint-Basile *et al.*,

1989), choroideremia (Schwartz *et al.*, 1986; Hodgson *et al.*, 1987; Sui *et al.*, 1988), and ocular albinism (Schnur *et al.*, 1989).

One of the best illustrations of reverse genetics in ophthalmic research is its application in the identification of the choroideremia gene locus (TCD). Early mapping studies established a close linkage between TCD and DXYS1, an anonymous DNA marker known to map to Xq13-21 (Nussbaum *et al.*, 1985; Lewis *et al.*, 1985). Additional studies centered on testing other existing markers known to be in the region Xq21-22 for linkage with TCD to confirm and refine this location (Gal *et al.*, 1986; Schwartz *et al.*, 1986; Lesko *et al.*, 1987; MacDonald *et al.*, 1987a, b; Sankila *et al.*, 1987). The TCD map position has been confirmed independently from linkage studies by the identification of individuals with complex deletions and translocations in Xq21 who also have choroideremia (Rosenberg *et al.*, 1986; Schwartz *et al.*, 1986; Hodgson *et al.*, 1987; Nussbaum *et al.*, 1987; Cremers *et al.*, 1987; Cremers *et al.*, 1988; Cremers *et al.*, 1989a; Schwartz *et al.*, 1988). The analysis of these deletion and translocation patient DNAs allowed the division of Xp21 into seven distinct physical regions (Cremers *et al.*, 1989).

The fine mapping and generation of anonymous Xq probes revealed that markers DXS95, DXS165 and DXS233 coincided to the region of Xq21 where the putative TCD gene might be (Cremers *et al.*, 1989b; Merry *et al.*, 1989). Analysis of choroideremia patient DNA showed that the marker DXS165 was deleted in several

cases. New markers were generated near the DXS165 locus (Cremers *et al.*, 1987; Cremers *et al.*, 1989a; van de Pol *et al.*, 1990; Merry *et al.*, 1990), one of which was found to overlap a majority of the TCD deletions. Portions of this clone were used to screen a human chorio-retinal cDNA library, and a number of cDNA clones were isolated, which upon patient analysis, were also found to overlap the defined TCD deletion region. In addition the DNA analysis of a female with symptoms of choroideremia and a t(X;13) translocation revealed that the X breakpoint disrupted the cDNA sequence in such a way that part of the cDNA sequence was deleted (Cremers *et al.*, 1990). The combined data strongly suggests that the cDNA clone isolated by Cremers is part of the choroideremia gene. The most surprising finding of this research, perhaps, is that the putative choroideremia gene is not a retinal specific gene but is expressed in a variety of tissues.

1.330) Forward genetics updated

A twist in the forward strategy is the application of the candidate gene approach. Genes encoding proteins which are known to play an important biological role in a given tissue are cloned, mapped, and examined as potential candidate genes for known disease loci which co-localize in the same map region. However it is often difficult to isolate and purify low abundant proteins or RNAs and it is often difficult to obtain large volumes of human tissue for extraction. An alternative strategy is to purify an important protein from another species and to isolated its corresponding gene. If the function of the relevant gene product is important the sequences will be

highly conserved between species, and it should be possible to isolate the human gene by using the alternate-species gene as the probe.

In the past 10 years this approach has been used successfully to define the underlying lesions of several known human retinal disorders, the most impressive of which has been the human color vision disorders. Using a bovine rhodopsin gene as probe (Nathans and Hogness, 1983), genes for the four human visual pigments were isolated and characterized (Nathans *et al.*, 1986a). The rod pigment (rhodopsin) was localized to 3q, the blue pigment gene to 7q, and the red and green pigment genes to Xq. The region of localization of the red and green pigment genes were in the identical area that encompassed the gene loci known to be responsible for X-linked inheritance of human color blindness (red and green). Analysis of the red and green pigment loci in individuals known to be color blind has confirmed that these genes underlie the disorder (Nathans *et al.*, 1986b): dichromats with no green sensitivity were missing the green pigment genes whereas dichromats with red sensitivity had an alteration in the red cone gene. Subsequently blue monochromy was shown to be due to the absence of both the red and green pigment genes (Nathans *et al.*, 1989).

1.340) Isolation based on markers of gene sequences

A number of strategies have been developed which attempt to exploit the cloning of expressed sequences. A major branch of research has focused on exploiting unique characteristics which set aside DNA sequences which code for a gene from

those which do not, they work on exploiting landmarks within the structural genes themselves.

Genomic clones can be assayed for the presence of rare cutting restriction sites which identify genomic regions having a high concentration of the dinucleotide CpG, a Hpa II tiny fragment (HTF) island. The presence of HTF islands has been found to be associated with the presence of gene sequences (Bird, 1986; Lindsay and Bird, 1987). This association is stronger for ubiquitously expressed genes than for tissue specific genes. The strategy is to identify genomic clones which digest with CpG restriction enzymes (an indication of a HTF island) and then to examine if these clones harbor gene sequences. This screening technique is not specific for any particular chromosome. However the genomic region can be limited by using genomic libraries made from DNA isolated from specific somatic cell hybrids, which contain only a subset of human chromosomes (Weiss and Green, 1967). Somatic cell hybrids with a single human X-chromosome have already been used to explore potential HTF/gene sites on that chromosome (Lindsay *et a*l., 1988)

Another genetic assay to identify transcribed regions utilizes the functional identification of sequences required for RNA splicing as indicators for exons (Duyk *et al.*, 1989, Duyk *et al.*, 1990; Auch and Reth, 1990). The basis of the strategy stems from the observation that during the RNA phase of the retroviral life cycle genomic sequences of non viral origin are correctly spliced and can be recovered in the

spliced form. The assay involves the use of a SV40/retroviral exon trap vector. The primary feature of this vector is that it possesses a single splice donor site next to a genetically marked intervening sequence followed by a multicloning site. During the RNA phase of the retroviral cycle, constructs which form an exon-intron-exon motif undergo a splicing event which removes the genetically marked intervening sequence (which is part of the vector). Following recovery of the DNA form of the virus, the loss of the intervening sequence is assayed by a genetic screen in bacterial cells. Clones which do not have the genetic marker become potential candidates as clones which harbor gene sequences.

1.350) **Problems in the existing methods**.

The major problem in the forward approach is that a knowledge of the primary protein structure is required prior to cloning. In most inherited genetic disorders the exact molecular lesion underlying the disorder is not known. The forward approach therefore is useful in identifying only a small number of genetic disorders. However the approach is potentially useful since the random cloning of genes that are expressed at the protein level in any specific tissue may in fact lead to the isolation of potential candidate disease genes.

The major drawback in the application of reverse genetics is that it requires a considerable commitment of time and manpower to progress from the initial

localization of the gene to the isolation of that gene. In the case of both Duchenne muscular dystrophy and cystic fibrosis it took several years of effort by large laboratories. Undoubtedly a large part of the delay is due to the extremely time consuming analysis (walking and jumping) inherent to this approach.

As a general approach, the major drawback in the isolation of HTF islands and the exon trapping technique is that they also require an immense amount of time and work. An additional problem in the HTF island method is that not all genes have associated HTF islands, thus a negative finding (no island) does not necessarily mean that there is no gene sequence. In the case of the exon trapping technique the background of false consensus like sequences and genes that may not have introns are difficult to estimate; likewise a positive finding (a splice site) does not absolutely correlate with the presence of a gene sequence and a negative finding (no splice site) is not necessarily confirmation that there is no gene sequence.

1.400) ALTERNATIVE APPROACHES

1.410) A need for more candidate gene markers

Recently the gene for type 1 autosomal dominant retinitis pigmentosa (ADRP) was localized and isolated. This feat was accomplished by a combination of the forward and reverse approaches. Standard linkage studies, the initial step in the reverse approach, in a single large Irish kindred established linkage between type 1 ADRP with D3S47, an anonymous marker known to be localized to 3q (McWilliam

et al., 1989). At the same time forward genetic approaches mapped a number of retinal genes, particularly those for retinol-binding proteins 1 and 2 and rhodopsin to this region of the genome (Colantuoni *et al.*, 1986; Demmer *et al.*, 1987; Nilsson *et al.*, 1988; Nathans *et al.*, 1986a) and these loci immediately became potential candidate genes underlying type 1 ADRP. Within a year of the localization of Type 1 ADRP it was determined that in some cases a point mutation, a base transversion of C to A resulting in the change of a proline to a histidine residue, in the rhodopsin gene leads to type 1 ADRP (Dryja *et al.*, 1990).

The successful isolation of the type 1 ADRP gene illustrates both the need for careful mapping studies and for genetic markers which correspond to expressed sequences as opposed to anonymous markers.

1.420) A potential new source of markers

The source for genetic markers corresponding to expressed sequences has so far relied on forward genetics. The study of disorders at the biochemical level however is often impeded by the lack of suitable amounts of human tissue for study. In the case of disease tissue it is difficult to obtain retinal tissue in a state that is informative from individuals with any of the retinal degenerative disorders such as RP or choroideremia. The reason for this is that the disorder often manifests at an early age and the individual often succumbs to complete retinal degeneration and

blindness long before death. Obtaining postmortem tissue for study of the retina in these cases would therefore be a futile exercise.

Normal retinal tissue, however, is available for study, although still not in any great quantities. On average an adult retina weighs approximately 160 mg. Thus it is virtually impossible to obtain sufficient amounts of retinal proteins from human tissue for meaningful study. However it is possible to exploit the intermediary between expressed gene and final gene product, the mRNA population of the tissue.

The human genome spans approximately 3.3×10^9 nucleotide base pairs. Within this mass of DNA there are approximately 100,000 functional genes distributed amongst the majority of DNA, which does not encode for a given gene product. It is estimated that a maximum of 37,000 of these genes are expressed, represented in the mRNA pool, in a given tissue at a given time (Williams, 1981). Although the sequences that are expressed in a given tissue may not be tissue specific, the expression profile is characteristic for each tissue.

When considering a genetic disorder it is important to attempt to understand the underlying mechanism for the disease condition. In the case of X-linked recessive disorders such as Norrie's disease, choroideremia, retinoschisis, and RP the disease state arises not because of the presence of an abnormal gene product but because of the absence of the normal gene product, since males who are hemizygous for the

disease gene are affected by the disorder, while females with a single disease gene are genetically heterozygote and functionally normal. It is expected therefore that a mRNA representative of each of the X-linked recessive disease gene loci will be present in normal retinal tissue. A compounding problem in many of the X-linked eye disorders, including choroideremia, retinitis pigmentosa and Norrie's disease, is that these diseases affect both the retina and the choroid of the eye. For this reason both tissues have been used for analysis in the present study.

1.430) Analysis of a cDNA library

A cDNA library is a cloned bank of sequences which reflects the mRNA population from the tissue from which it was extracted. The identification of candidate disease clones for any genetic disorder in question could be achieved by the sequential chromosome mapping of each unique clone from a chorio-retinal cDNA library to generate a subset of clones which localize to the chromosomal region of interest, the region to where the disease gene maps, in a tissue which is known to be affected by the disorder. Because different messages exist at different abundances in the mRNA population it is possible for some clones in a cDNA library to become under represented by chance, thus at least 1.69×10^5 independent cDNA clones need to be mapped at random to ensure a 99% probability of characterizing a given low abundant sequence (Clarke and Carbon, 1976). Furthermore the vast majority of the clones mapped in this fashion would localize to chromosomes other than the X chromosome.

1.440) Differential cross screening

cDNAs present in one type of cell (+cell) but not in another (-cell) can be identified by differential hybridization. Typically duplicate filters carrying cDNA clones derived from the (+cell) are screened with radioactive cDNA from the (+cell) and (-cell) respectively. Clones which hybridize with the (+) but not the (-) probe are (+cell) specific. An alternative modification is to screen the (+cell) cDNA library with a radioactive (+cell) cDNA probe that has been subtracted with non-radioactive (-cell) cDNA.

With respect to ophthalmic disorders the traditional differential cross screening methodology has been taken in order to isolate abundant retinal specific sequences on the assumption that abundant retinal specific sequences are likely to encode proteins of major importance to the tissue (McInnes *et al.*, 1988). Disruption of these genes will therefore underlie tissue specific diseases. The differential screen involved the screening of a human retinal cDNA library with cDNA probes derived from bovine retina mRNA and other human tissues. The frequency of non-rhodopsin, abundant, retinal specific cDNA clones was found to be approximately 1 %. From this analysis the cloning of ROM1, which is a gene encoding a membrane protein located in the outer segment of the rod photoreceptor, and HOX10, which encodes a developmentally regulated homeobox gene in the human retina, have already provided two additional genes which are highly applicable to the candidate gene approach (Bascom *et al.*, 1989; Bascom *et al.*, 1990; De Chen *et al.*, 1989; De

Chen *et al.*, 1990). The application of this strategy, however, does not isolate genes from a predefined region of the genome (for example the X chromosome) nor will it identify genes which underlie retinopathies which do not display tissue specific expression.

1.450) hnRNA

In another strategy, cDNA libraries are made from mRNA of hybrid cell lines. The human chromosomes in such a hybrid continues to express some of its genes. A cDNA library made from such hybrids will therefore contain clones derived from both human and rodent mRNA. A subset of human clones can be differentiated on the basis of human repetitive sequences (Neve *et al.*, 1986). This approach is applicable to any chromosome or sub-chromosomal region as defined by the specific somatic cell hybrid in hand and attempts to specifically isolate expressed sequences from a predefined subregion of the genome. The type of clones that will be isolated with this strategy is limited since the detection system requires that the specific sequence be expressed in the hybrid cell line and that the expressed sequence itself contain a human specific repetitive element.

A modification in the above strategy has been to focus on the use of hybrid cell line heterogeneous nuclear RNA (hnRNA), RNA isolated from the nucleus consisting of unprocessed primary transcripts, as template for cDNA library synthesis instead of mature mRNA. cDNA synthesis from hnRNA is primed with

oligonucleotides derived from intronic 5' consensus splice donor sites (Liu *et al.*, 1989). The cDNAs generated are expected to be copies of the upstream exon and the next intron. As intron sequences are more likely to contain highly repeated sequences, the resulting cDNA library is screened with labeled human genomic DNA to identify clones containing human cDNAs among the rodent cDNAs.

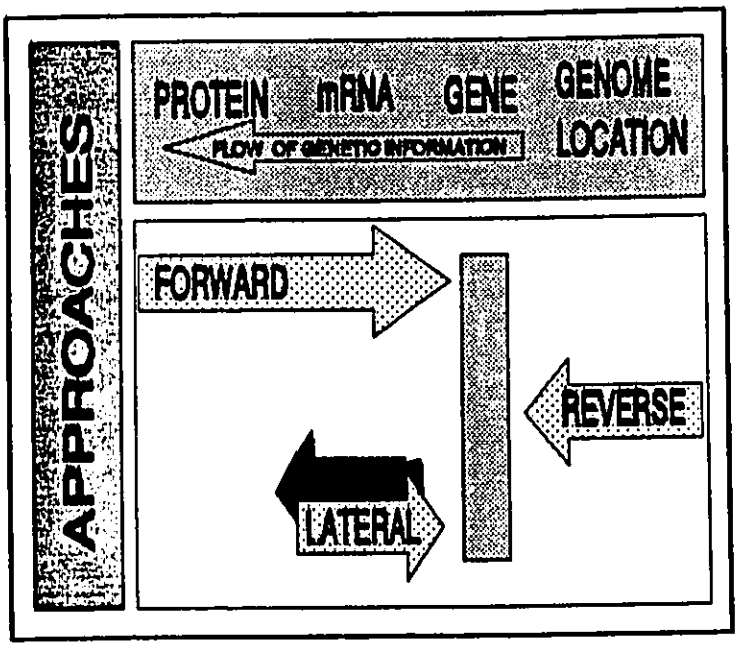
Another modification of the hybrid cDNA approach makes use of oligonucleotide primers derived from Alu sequences (Corbo, 1990). The primers used in this method are human specific (Nelson *et al.*, 1989). Thus, the cDNA library generated in this case will be of human origin.

The major drawback in use of hybrid cell lines in the above strategies is that only a subset of human genes within the genomic region under scrutiny will be expressed. Clones that are isolated, therefore, will not necessarily represent a total set of genes and in such a system it is difficult to anticipate what will be expressed.

1.460) Lateral Genetics

If the forward approach is the isolation of a gene on the basis of protein information and the reverse approach is the isolation of a gene based on its genomic location (DNA), then the isolation of a gene at the level of mRNA can perhaps be viewed as a lateral approach which bridges the forward and reverse strategies in its ability to shed light at both the protein and DNA level of organization (Fig. 2).

FIG. 2 APPROACHES TO THE ISOLATION OF DISEASE GENES.


APPROACHES
PROTEIN
mRNA
GENE
GENOME LOCATION
FLOW OF GENETIC INFORMATION
FORWARD
REVERSE
LATERAL

The ability to bring together a probe which is representative of the expressed sequences within a given tissue together with genomic libraries representing selected human chromosomes is intriguing. The successful marriage of these two will allow the direct association of tissue gene expression with chromosome specificity.

The use of a probe reflecting the mRNA population, whether it be radiolabeled mRNA or cDNA derived from a given tissue has been used sparingly in the screening of genomic libraries (Paulsen *et al.*, 1986; Scherer *et al.*, 1981; Tashima *et al.*, 1981; Hochgeschwender *et al.*, 1989). The primary reason for this has been the finding that this cross screening tends to lead to the selection of false positives (Williams, 1981). This problem however can be overcome by removal of the repetitive fraction of the probe by prehybridization of the probe to repetitive sequences bound to a membrane (Tokarskaya *et al.*, 1980), competition of the probe with an excess of repetitive sequences during hybridization (Tashima *et al.*, 1981), pre-association of the probe with a source of repetitive DNA (Paulsen *et al.*, 1986; Sealy *et al.*, 1985), and removal of repetitive sequences using genomic DNA linked to DNA-cellulose (Brison *et al.*, 1982; Hochgeschwender *et al.*, 1989). Other investigators however have been able to isolate specific genomic clones harboring expressed sequences without removal of the repetitive fraction of the mixed cDNA probe (Scherer *et al.*, 1981).

The major problem in studying diseases of the human eye is not necessarily the technical aspects but rather the low amounts of specific human tissue that are available for scientific research. In general, approximately 200-250 μg of total RNA can be extracted from a pair of human choroid/retinas. In this total extracted RNA, 2-5% will represent poly(A)+ RNA (4-10 μg), a portion of the total RNA that is enriched in mRNA sequences. The optimized quantity of poly(A)+ for cDNA synthesis is approximately 10 μg. In a good year it is feasible to collect 2-3 pairs of choroid/retinas from organ donors: this collection however occurs sporadically. It is therefore difficult to anticipate the total amount of human chorio-retinal RNA that would be available for study at any given time.

In order to overcome this problem a simple alternative strategy has been devised and will be discussed in this thesis. Briefly, instead of relying on an unpredictable supply of choroid/retinas for mRNA as the probe source, chorio-retinal poly(A)+ RNA was used to generate a human chorio-retinal cDNA library. A method was devised which consisted of the amplification of the inserted human chorio-retinal cDNA sequences by the polymerase chain reaction (PCR)(Saiki *et al.*, 1985, Mullis and Faloona, 1987, Saiki *et al.*, 1988). The specifications of this procedure enabled the screening of genomic libraries constructed in similar based vectors as that used for the cDNA library. The cDNA library together with the amplification procedure represents a continual probe source.

In order to isolate potential X-specific chorio-retinal expressed sequences a genomic library was constructed from a human/hamster hybrid cell line, whose only human component is a single X-chromosome. This library was screened with the PCR amplified human chorio-retinal cDNA library probe. Theoretically human clones identified in this way should be derived from the human X-chromosome and should harbor chorio-retinal expressed sequence. Clones which are verified with respect to these qualifications can then be used in a candidate gene approach to see if they are relevant to any of the known X-linked chorio-retinopathies. The development and critical examination of this lateral screening strategy will be the subject of this thesis. The proposed study involves the completion of four distinct steps:

1. The synthesis of a representative chorio-retinal cDNA library.
2. The synthesis of a random human X-specific genomic library.
3. The development of the library cross screening procedure.
4. Initial verification of the results of the proposed cross screen.

CHAPTER TWO

MATERIALS AND METHODS

2.100) RNA

2.110) Eye samples:

The human retina lies between the vitreous and the choroid layer of the eye (Fig. 3). As stated earlier both the choroid and the retina are affected in a number of X-linked eye disorders, as such both tissues have been obtained for the present study.

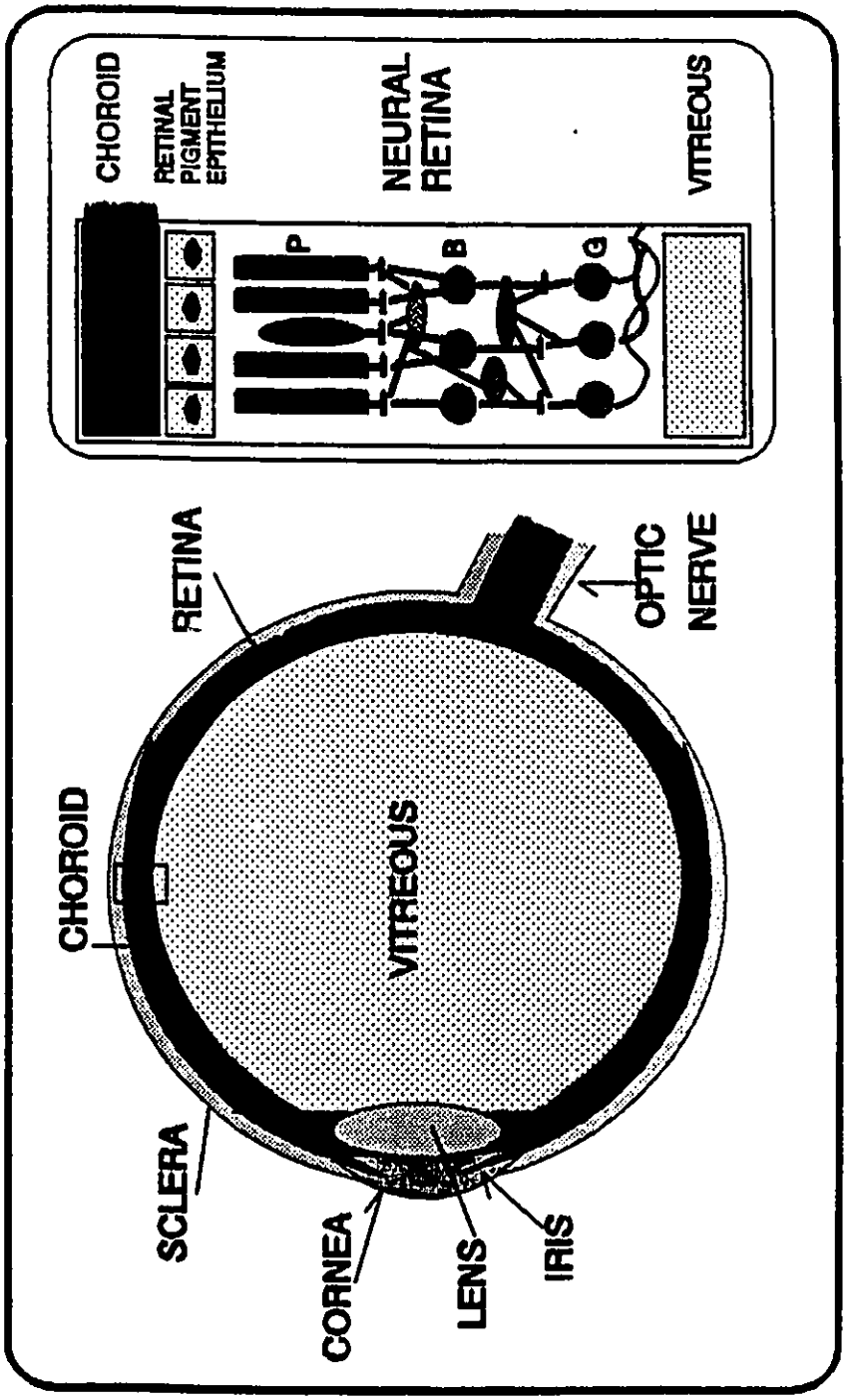
Human choroid/retina samples were provided by Dr. I. MacDonald (Department of Ophthalmology, Ottawa General Hospital). Human eyes were enucleated by the ophthalmic surgeon within 6 h, postmortem. A stab incision was made 2 mm posterior to the limbus and the corneal button was removed with a 2 mm cuff of sclera. A complete cyclo-dialysis was performed along with 3 axial cuts of the scleral coat toward the optic nerve. The anterior segment, iris, lens and cilary body were removed and in all cases an attempt was made to expulse the vitreous. The remaining choroid and retina, from pars plana posteriorly was then amputated at the optic nerve and snap frozen in liquid nitrogen and stored at -70°C until RNA extraction. In cases where the vitreous would not cleanly separate from the retinal layer, the entire choroid/retina/vitreous was snap frozen in liquid nitrogen and stored as described above.

2.120) Isolation of RNA

2.121) Isolation of Total RNA: Total RNA was isolated from flash frozen human choriod/retina tissue using the LiCl/urea procedure (Auffray and Rougeon, 1979)

FIG. 3 GENERAL ANATOMY OF THE HUMAN EYE.

The retina lies between the vitreous and the choroid. Located between the choroid and the retina is the retinal pigment epithelium. The neural retina consists of over 50 distinct cell types. Shown in the inset are the neural cells involved in the direct neural pathway to the brain. This neural relay includes the photoreceptor cells (P; both cone and rod photoreceptor cells), the bipolar cells (B), and the ganglion cells(G).


CHOROID
RETINAL PIGMENT EPITHELIUM
NEURAL RETINA
VITREOUS
P
B
G
RETINA
OPTIC NERVE
CHOROID
VITREOUS
SCLERA
CORNEA
LENS
IRIS

with minor modifications (Tenniswood and Simpson, 1982). Frozen tissue was weighed and ground to a fine powder in liquid nitrogen, homogenized on ice in a fixed initial volume of extraction buffer (20 ml/gm of tissue: 3 M LiCl, 6 M urea in 50 mM NaOAc, pH 5.5) and sonicated (3X 15 sec bursts set at 30 cycles, with 1 minute cooling intervals between bursts). SDS was added to 0.1% and the sample was incubated at 4°C overnight.

The sample was centrifuged at 16,000 rpm for 30 min and the pellet was recovered, drained, and resuspended in wash buffer (4 M LiCl, 8 M urea in 50 mM NaOAc, pH 5.5) with 0.1% SDS up to the initial volume. The sample was re-centrifuged as described above and the pellet was recovered and dissolved in half the initial volume of 50 mM NaOAc pH 5.5 and extracted with an equal volume of phenol (saturated in 50 mM NaOAc pH 5.5) and an equal volume of chloroform. The organic and aqueous phases were separated by centrifugation (1000 rpm, 5 min, room temperature) and the aqueous phase was retained. The organic phase was re-extracted with 50 mM NaOAc and the aqueous phases were pooled. The combined aqueous phases were re-extracted in an equal volume of phenol/chloroform and once with chloroform. The RNA was precipitated from the aqueous phase by adding 1/10th volume 3 M NaOAc pH 5.5 and 2.5 volumes of cold ethanol. The RNA was precipitated at -20°C overnight and pelleted at 10,000 rpm for 1 h at 0°C. The pellet was recovered, dried and redissolved in water. The final product was stored at -70°C.

2.122) Isolation of Poly(A)+ RNA: Poly(A)+ RNA was prepared using oligo-(dT)cellulose chromatography (Aviv and Leder, 1972). Total RNA was boiled for 2 min and chilled on ice for 7 min. An equal volume of 2X binding buffer (1 M NaCl, 40 mM Tris pH 7.6, 2 mM EDTA) was added and the RNA was added to oligo(dT) cellulose that had been equilibrated with 2X binding buffer, and incubated at 37°C on a rotary mixer for 4 h. The slurry was packed into a column. The poly(A)- fraction was first eluted by passing a minimum of 10 bed volumes of binding buffer through the column. Poly (A)- RNA, was precipitated by addition of cold ethanol and an overnight incubation at -20°C. The poly(A)- RNA was pelleted by centrifugation (13,000 rpm, 60 min, at 0°C) and redissolved in sterile double distilled water.

Poly (A)+ RNA was eluted from the column by passing 3 bed volumes of sterile water through the column and collecting the eluate. Poly (A)+ RNA was precipitated overnight at -20°C after addition of 1/10th volume 3 M NaOAc and 2.5 volumes ethanol. The poly (A)+ RNA was pelleted at 10,000 rpm for 90 min at 0°C. The pellet was air dried and redissolved in sterile double distilled water and stored at -70°C.

2.123) Quantification of RNA: The concentration of RNA was estimated by measuring optical density (O.D.) at 260 nm wavelength. The ratio of $O.D._{260}/O.D._{280}$ was used to assess purity. Ratios of 1.5-2.0 were considered to indicate negligible

protein or phenol contamination. Typical yields were approximately 80 μg of total RNA per gram wet weight of choroid/retina/vitreous tissue.

2.130) Characterization of RNA

The functional quality of the isolated RNA samples were examined by *in vitro* translation in an mRNA dependent rabbit reticulocyte lysate system (Pelham and Jackson, 1976) using [^{35}S]-methionine as the radioactive label. The reticulocyte lysate system was supplied by Promega Biotec (Madison, Wisconson). *In vitro* translation was carried out by adding 2 μg of RNA to 35 μl of nuclease treated lysate, 20 μM amino acid mixture (minus methionine), 50 μCi of ^{35}S-methionine in a total volume of 50 μl. The cocktail was allowed to incubate at 30°C for 1 h. Aliquots of 2μl were removed, spotted onto GF/C filters, and the amount of incorporated radiolabel was determined by trichloroacetic acid precipitation. The remaining *in vitro* translation product was stored at -20°C.

The translation products were analyzed by one and two dimensional polyacrylamide gel electrophoresis, PAGE (O'Farrell, 1975; Newbold *et al.*, 1982). One dimensional PAGE was carried out in 10 or 15 % SDS slab gels under denaturing conditions. Bio-Rad low molecular weight SDS-PAGE protein markers were used as molecular weight standards. After electrophoresis the gels were stained with Coomassie blue, destained, and fixed with 7% acetic acid. After fixing, the gels were prepared for fluorography (Bonner and Laskey, 1974), dried down under

vacuum and then exposed to Cronex X-ray film in a Kodak X-Omatic cassette with intensifying screens.

In the case of two dimensional PAGE, the labeled translation products were first electrophoresed in isofocusing tube gels using LKB pH 3-10 ampholines, followed by one dimensional PAGE for the second dimension (Newbold *et al.*, 1982). Isofocusing gels using unlabeled reticulocyte lysate were run concurrently, cut into 1 cm long discs, placed in 1 ml of distilled water and the pH measured (these served as a measure of pI for the sample gels). After second dimension electrophoresis the gel was processed as described for one dimensional PAGE.

2.200) PREPARATION OF CHORIO-RETINAL cDNA LIBRARY

2.210) cDNA Synthesis

Poly(A)$^+$ RNA was used as the template for cDNA synthesis. cDNA was synthesized using a modification (Rutledge *et al.*, 1988) of the method described by Gubler and Hoffman, 1983 (Fig. 4).

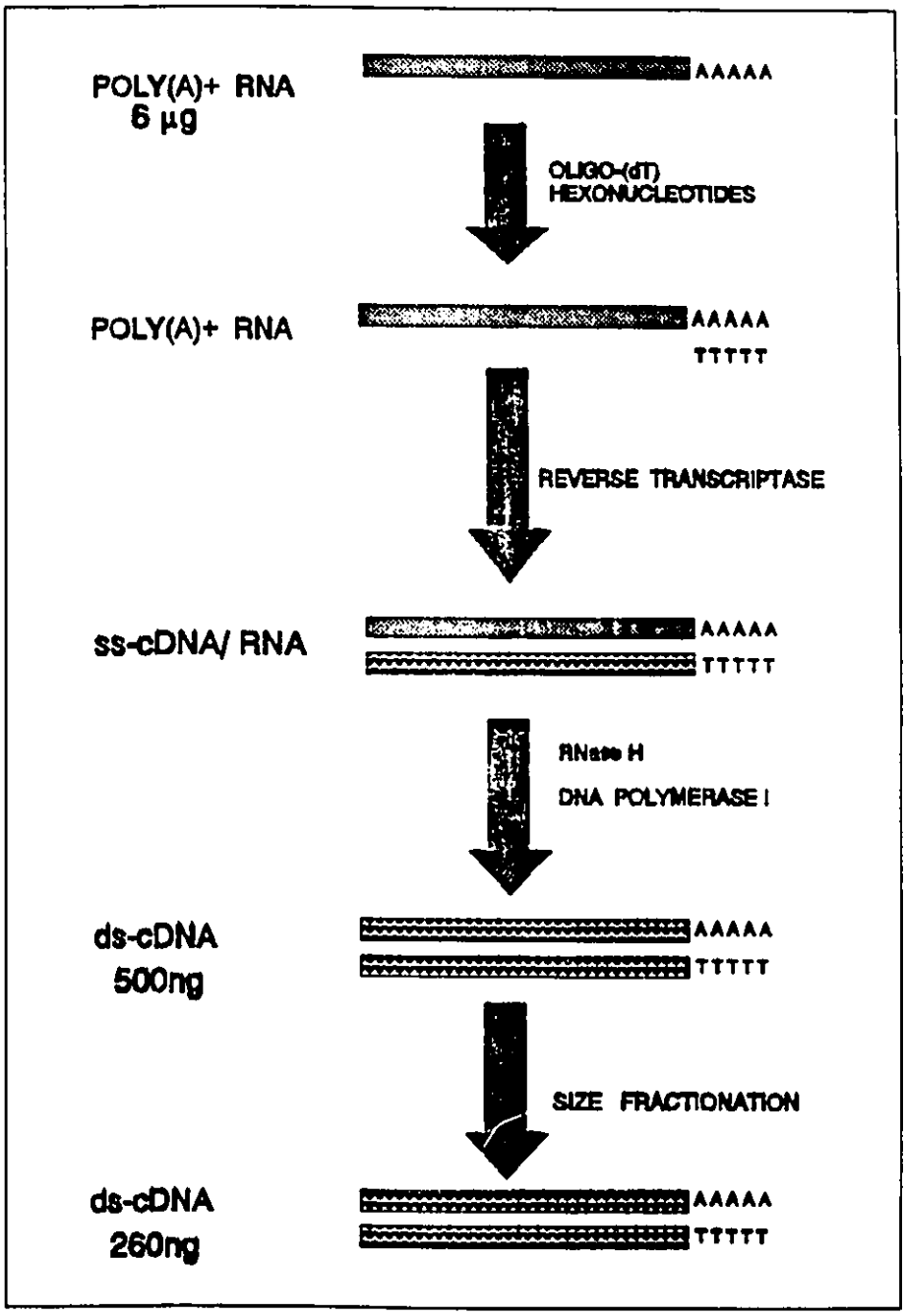
FIG. 4 SYNTHESIS OF ds-cDNA

For first strand synthesis, 6 μg of poly(A)+ RNA was incubated in a final volume of 30 μl with 100 mM Tris-HCl pH 8.3, 130 mM KCl, 10 mM $MgCl_2$, 2.5 mM DTT, 4 μg oligo$(dT)_{10\text{-}18}$, 1 mM dATP, 1 mM TTP, 1 mM dGTP, 1 mM dCTP, and 50 μCi of $\alpha[^{32}P]$-dATP. After heating to 60°C for 2 min,

24 units of AMV reverse transcriptase was added and the mix was incubated at 42°C for 30 min.

For second strand synthesis the reaction volume was brought up to 100 μl with 50 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$ 20 mM DTT, and 50 μg/ml BSA. RNase H (2 units) and DNA polymerase I (25 units) were added and the reaction was incubated at 15°C for 2 h.

The sample was then extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and passed through a 1 ml G-50 spun column to remove unincorporated nucleotides (Maniatis, Fritsch and Sambrook, 1982). Subsequently the ds-cDNA was size fractionated on a Sepharose 4B column.


POLY(A)+ RNA
6 µg
AAAAA
OLIGO-(dT)
HEXONUCLEOTIDES
POLY(A)+ RNA
AAAAA
TTTTT
REVERSE TRANSCRIPTASE
ss-cDNA/ RNA
AAAAA
TTTTT
RNase H
DNA POLYMERASE I
ds-cDNA
500ng
AAAAA
TTTTT
SIZE FRACTIONATION
ds-cDNA
260ng
AAAAA
TTTTT

A 1 μl aliquot of the eluate, in 5 ml of scintillation fluid was counted in a Beckman L1800 scintillation counter. The total incorporation of radioactivity was used to estimate the final yield of ds-cDNA using the equation:

EQUATION 1

$$\text{ds-cDNA (ng) yield} = \frac{\text{nmole cold dNTP added} \times 326 \times \text{total counts} \times 4}{\mu\text{Ci dNTP used} \times 2.2 \times 10^6}$$

2.220) Size fractionation of cDNA

Prior to cloning, the cDNA was size fractionated in a 1 ml Sepharose 4B column to size select for the larger cDNA fragments. One drop (30 μl) fractions were collected and radioactivity measured by the Cerenkov method in a Nuclear Chicago Isocap 300 scintillation counter. From the profile of incorporated counts vs fraction, the fractions containing the larger cDNA species were pooled. The radioactivity associated with this pool was determined by counting a 1 μl aliquot in a Beckman L1800 scintillation counter and the total yield of ds-cDNA was estimated as above. The cDNA pool was precipitated at -20°C overnight in the presence of 0.3 M NaCl and ethanol and redissolved in sterile double distilled water.

One μl of the size fractionated cDNA pool was sized by electrophoresis on a 2% slab gel in a Tris-acetate buffer system. A Hind III digest of λ DNA (H3L) was used as a standard molecular weight marker. The gel was stained with ethidium bromide (1 μg/ml) and photographed. The gel was dried down under vacuum on a

Savant gel drier and autoradiographed using Cronex X-ray film in a Kodak X-Omatic X-ray cassette with intensifying screens.

2.230) Cloning into λgt10

The size selected cDNA was inserted into the Eco RI site of the phage vector λgt10 (Hyunh *et al.*, 1985) using a commercially available cloning kit (Amersham, Oakville, Ontario). A summary of the procedure is presented in Fig. 5.

The bacterial host used for the human retinal cDNA library was NM514. The bacteria L87 was used as a control host in the assaying of cloning efficiency. In both cases frozen liquid cell stocks of each cell line were established and stored at -70°C. Cells to be used for phage infection were grown in 50 ml of L-broth supplemented with 0.4% maltose at 37°C on a shaker until the $O.D._{600}$ reached 0.5. The bacteria were cooled on ice, pelleted at 4°C and resuspended in 15 ml of cold 10mM $MgSO_4$ or SM buffer. In all cases prepared plating cells were stored at 4°C and used within 3 days.

In the plating of phage a fixed number of phage particles was added to 200 μl of plating cells and incubated at 37°C for 30 min. The phage and cells were then added to melted top agar (3 ml for small plates and up to 6 ml for large plates) mixed and poured onto a L-agar plate. The layer of top agar was allowed to solidify and the plate was incubated at 37°C overnight.

FIG. 5 cDNA CLONING INTO λGT10

cDNA cloning into λgt10 was carried out using a λgt10 cloning kit (Amersham, Oakville, Ontario). In all cases enzymes required as well as their respective buffers were supplied. Following is a brief summary of the cloning protocol.

1) Preparation of the ds-cDNA for ligation into the vector
The ds-cDNA to be inserted (up to 1 μg) was first methylated with Eco RI methylase (20 U) at 37°C for 1 h to protect internal Eco RI sites. The reaction was stopped by heating the reaction tube at 70°C for 10 min.

The ds-cDNA was ligated to Eco RI linkers (1 μg) by the addition of 5 U T4 Ligase. The ligation was carried out at 15°C for 20 h. The reaction was stopped by heating the reaction tube at 70°C for 10 min.

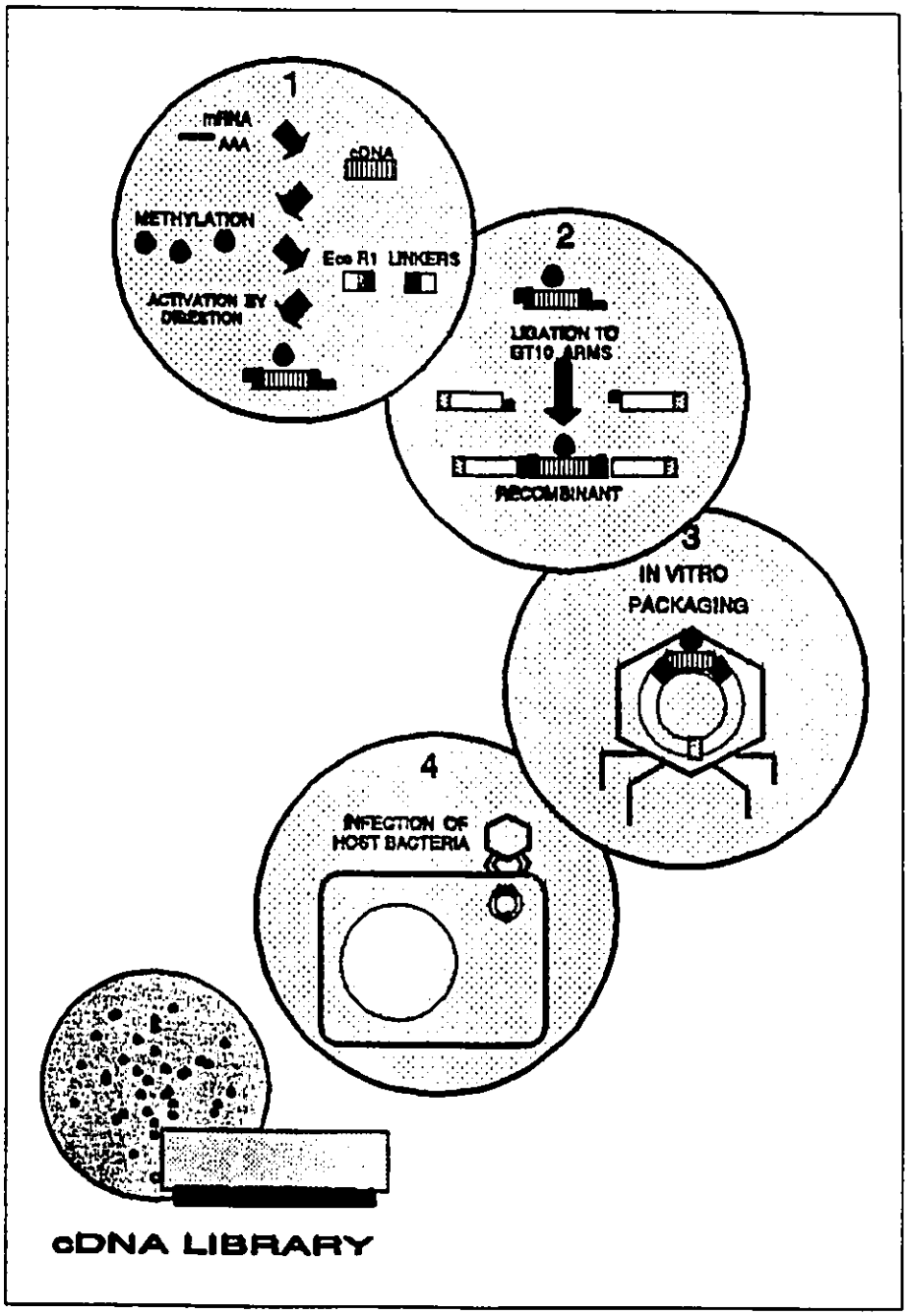
The ds-cDNA was digested with Eco RI (100 U) at 37°C for 5 h in order to activate the linkers. The reaction was stopped by heating the reaction tube at 70°C for 10 min.

The reaction mix was loaded onto a column in order to remove excess linkers. 200 μl fractions were eluted and collected by the addition of 200 μl of STE. In total 10 fractions were collected. The radioactivity was measured by Cerenkov counting. The two major peaks within the first 6 fractions were pooled and the DNA was precipitated.

2) Ligation of ds-cDNA to λgt10 arms
50 ng of ds-cDNA was ligated to 1 μg of λgt10 arms with 2.5 U of T4 ligase in a final reaction volume of 10 μl. The reaction was carried out at 15°C for 16-20 h.

3) Packaging of recombinants into viable phage
The DNA was precipitated and resuspended in 2.5 μl of TE buffer. The packaging extracts provided by Amersham were added directly to the DNA and the mix was incubated at 20°C for 2 h.

4) Storage and infection of bacteria
Packaged libraries are stored in 0.5 ml of SM buffer with the addition of 10 μl of chloroform. Libraries generated were plated out as outlined in materials and methods using bacterial cell lines NM514 and L87.


1
mRNA
AAA
cDNA
METHYLATION
Eco R1 LINKERS
ACTIVATION BY DIGESTION
2
LIGATION TO GT10 ARMS
RECOMBINANT
3
IN VITRO PACKAGING
4
INFECTION OF HOST BACTERIA
cDNA LIBRARY

2.300) DNA EXTRACTION

2.310) Blood

Peripheral blood from individuals were collected in vials containing EDTA, and DNA was isolated according to the method of Dr. D. Hoar (Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary). Ten ml of whole blood was incubated at 37°C for 5 min with 50 ml of NH_4Cl/Tris solution (17 mM Tris-HCl, 140 mM NH_4Cl, pH 7.65; prewarmed to 37°C) and centrifuged at 2000 rpm for 10 min at room temperature. The supernatant was decanted and the pellet was washed twice with saline (0.85% NaCl). The remaining cell pellet was resuspended in 3 ml of high TE (100 mM Tris-HCl (pH 7.5), 10 mM EDTA). Immediate cell lysis was induced by quick injection of 2 ml of lysis buffer (100 mM Tris-HCl, 40 mM EDTA, 0.2% SDS, pH 8.0) with a 10 ml syringe through a 26 gauge needle. Ten ml of high TE were added and the solution was extracted twice with an equal volume of phenol/chloroform (saturated in high TE) and once with an equal volume of chloroform/isoamyl alcohol (24:1). The organic and the aqueous phases were separated by centrifugation at 4000 rpm for 10 min at room temperature.

The DNA was precipitated by adding one tenth volume of 3 M NaOAc and 2.5 volumes of ice cold ethanol. In most circumstances the DNA precipitated immediately and was spooled onto a sterile siliconized glass hook, transferred to 10 ml of TE buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA) and dissolved at 4°C overnight on a rotary mixer. When the DNA was in solution, DNAse-free RNAse A

was added to a final concentration of 100 μg/μl and the sample was incubated at 37°C for 1 h. SDS was added to a final concentration of 1% and proteinase K was added to a final concentration of 100 μg/μl, the tube was inverted and incubated at 37°C overnight.

The DNA solution was extracted twice with an equal volume of phenol/ chloroform and once with chloroform/isoamyl alcohol (24:1). The DNA was precipitated and redissolved in 1 ml of TE buffer. In cases where the DNA did not redissolve after one day additional TE buffer was added.

2.320) Human placenta

Human placenta was obtained fresh, flash frozen in liquid nitrogen and kept at -70°C until DNA extraction (Gross-Bellard *et al.*, 1973). Frozen placental tissue was ground to a fine powder in liquid nitrogen. Seven grams of ground tissue was added to 45ml of solution 1 (100 mM EDTA, 200 mM Tris-HCl pH 8.0, 1% SDS) and the tissue was homogenized. Proteinase K was added to 30 μg/ml and the tube was incubated at 60°C overnight.

The tube was allowed to cool slowly to 25°C. 1/250 th volume of diethyl pyrocarbonate was added and the solution was incubated at 25°C for 30 min. NaCl was added to a final concentration of 200 mM and the solution was extracted twice with an equal volume of phenol/chloroform and once with chloroform/isoamyl

alcohol (24:1). DNA was precipitated by addition of 2.5 volumes of ice cold ethanol and was collected by spooling. The DNA was dissolved in 10 ml of TE overnight at 4°C on a rotary mixer.

DNAse-free RNAse A was added to 50 μg/ml and incubated at 37°C for 1 h. Proteinase K was added to 250 μg/ml and incubated at 60°C for 3 h. The solution was then extracted once each with phenol/chloroform and chloroform/isoamyl alcohol (24:1). The DNA was precipitated by the addition of cold ethanol and retrieved by spooling. The DNA was re-dissolved in TE overnight at 4°C on a rotary mixer at 4°C. The DNA solution was incubated at 60°C for 1 h to ensure that all the DNA was in solution.

2.330) Cell cultures

Two cell lines, 82082a and 1103, were obtained from Dr. J.L. Hamerton (University of Manitoba, Winnipeg, Manitoba). 82082a is a hamster/human somatic cell hybrid cell line with a single human X chromosome as its only human component (Riddell *et al.*, 1986). 1103 is the hamster parent of 82082a and is HPRT$^-$.

1103 cells were grown in McCoy's 5A medium (HSC modified, Gibco) supplemented with 2.2% $NaHCO_3$, 100 U/ml Penicillin G, 100 μg/ml streptomycin and 10% bovine serum albumin. 82082a cells were grown in the same medium as defined above except that it was further supplemented with 1 μM thymidine, 10 μM hypoxanthine, 3.2 μM aminopterin, and 10 pM glycine (HAT selective medium) in

order to maintain a selection for the human X-chromosome by requiring the expression of the human HPRT gene. In both cases cell lines were maintained at 37°C in 4% CO_2.

DNA was extracted from the cell lines using modifications of standard protocols (Maniatis *et al.*, 1982). Cells were disaggregated from the culture plates using 0.05% trypsin and 0.5 mM EDTA in a buffered saline solution and harvested. Cells were spun down and washed three times with buffered saline solution and resuspended in TE and homogenized with 10 strokes in a glass/teflon homogenizer. Proteinase K was added to a final concentration of 20 μg/ml and the solution was incubated at 60°C overnight. The remaining steps in the extraction protocol were essentially as described for the isolation of DNA from human placental tissue.

2.340) X-cell panel DNA

DNA from various established cell lines (676x175; 749; 617x347) were a gift from Dr. J.L. Hamerton. Taken collectively the samples generate a human chromosome X panel for mapping of DNA markers.

2.400) MEASUREMENT OF DNA CONCENTRATION

The concentration of DNA was estimated by measuring O.D. at 260 nm wavelength. The ratio of O.D.260/O.D.280 was used to assess purity. Ratios of 1.5-2.0 were considered to indicate negligible protein or phenol contamination.

2.500) PREPARATION OF THE 82082a GENOMIC LIBRARY

2.510) Digestion of genomic DNA

High molecular weight (HMW) DNA (20-40 μg), extracted from the 82082a cell line as described above, was digested with Mbo I in a series of enzyme dilutions (Fig. 6). The tubes were incubated at 37°C for 1 h and placed on ice. The reaction was stopped by addition of EDTA to 20 mM. 10 μl of T1, 8 and 9 were set aside. The remainders of T1-8 were pooled.

2.520) Isolation and purification of DNA fragments

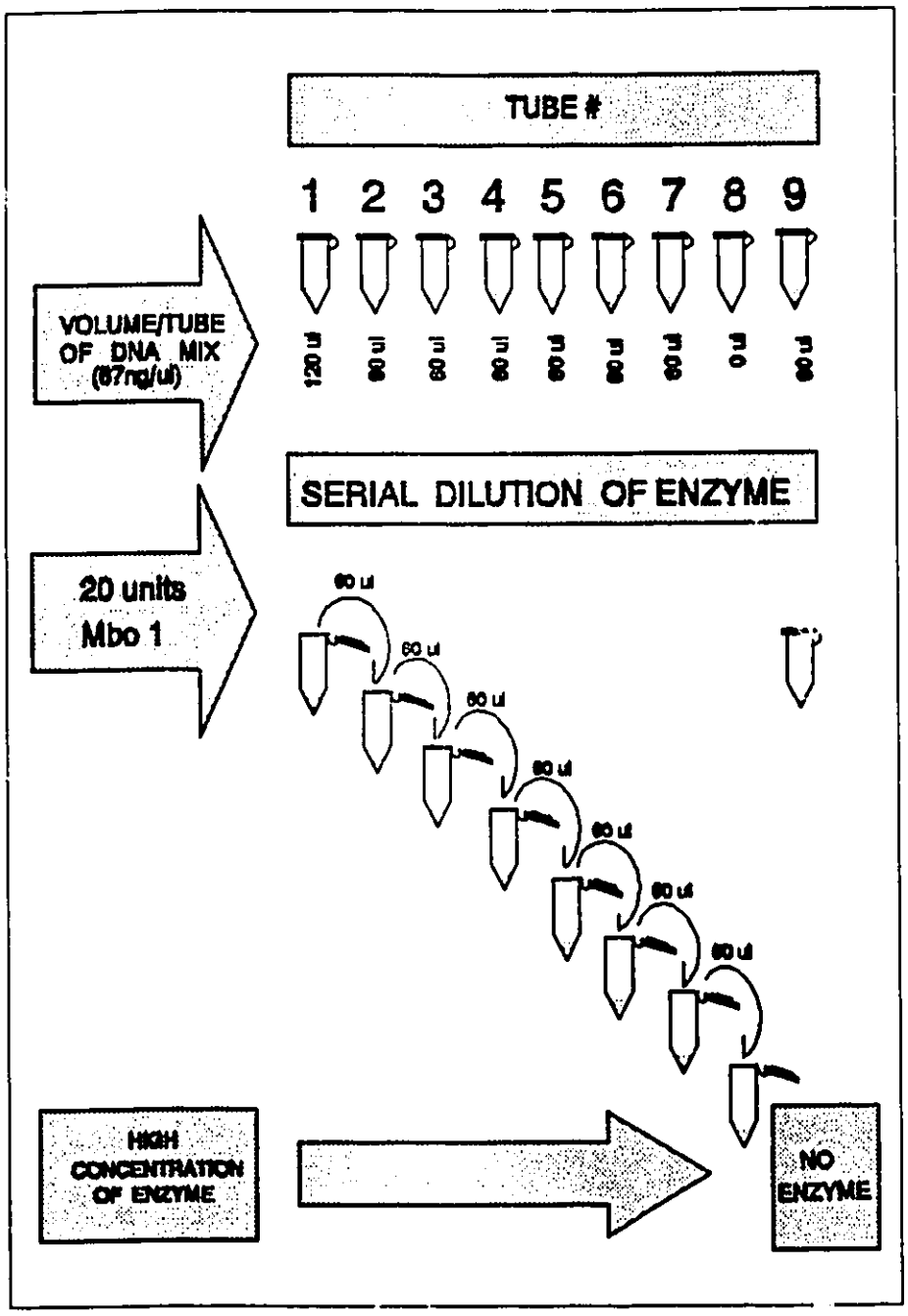
The aliquots of T1, 8, 9, the pooled samples and molecular size markers were electrophoresed at 4°C at 40V for 16-18 h in 0.5% low melting agarose. The lanes consisting of T1, 8, 9 and molecular size markers were stained with ethidium bromide (1 μg/ml) and visualized by 306 nm UV transillumination. After viewing these control lanes to ensure that adequate digestion occurred, the gel area corresponding to 17-23 Kb fragments was excised from the unstained pooled DNA lanes and extracted using a modification of Maniatis *et al.*, (1982). The excised gel pieces were dissolved in 5 volumes of 20 mM Tris pH 8, 1 mM EDTA (preheated to 65°C) and incubated at 65°C for 1-2 h. The sample was then left at room temperature for 30 min to cool and extracted with an equal volume of phenol saturated in high TE. The aqueous phase was isolated by centrifugation at 20°C and re-extracted once with phenol/chloroform and once with chloroform/isoamyl alcohol (24:1). The DNA was precipitated, air dried briefly and resuspended in up to 200 μl of TE buffer.

FIG. 6 GENERATION OF GENOMIC FRAGMENTS FOR CLONING

High molecular weight DNA (40 µg), extracted from the 82082a cell line as described above, was digested with Mbo I in a series of enzyme dilutions. The DNA (40 µg) was diluted to 0.067 µg/µl in 1X Mbo I buffer and aliquoted into 9 microfuge tubes on ice, 1/5 th of the volume into tube 1 and 1/10th into each of tubes 2-9, excluding tube 8.

Mbo I (2.5 U/µg of DNA) was added to tube 1 and mixed gently. Half the volume of tube 1 was transferred to tube 2 and mixed gently. Subsequent two fold dilutions were carried out to tube 8. The tubes were incubated at 37°C for 1 h and placed on ice. EDTA was added to 20 mM in order to stop the reaction.

For electrophoresis 10 µl of tube 1, 8 and 9 were used as controls to monitor the digestion profile. The remainders of tube 1-8 were pooled and run in the minimum number of wells that would hold its volume.


TUBE #
1 2 3 4 5 6 7 8 9
VOLUME/TUBE OF DNA MIX (67ng/ul)
SERIAL DILUTION OF ENZYME
20 units Mbo 1
60 ul
HIGH CONCENTRATION OF ENZYME
NO ENZYME

2.530) Construction of Genomic library

Genomic fragments were cloned into the phage vector EMBL3 (Promega Biotec, Madison, Wisconsin) using the protocol supplied by the manufacturer for ligation and packaging of recombinant phage forming units. In brief 0.35 μg of genomic insert was ligated to 2 μg of vector arms with 2.5 U of T4 DNA ligase at 21°C for 6 h before packaging (Fig. 7).

The bacterial host strain used for this library was the *E. coli* strain MB406. Frozen liquid cell stocks of the cell line was established and stored at -70°C. Cells to be used for phage infection were grown in 50 ml of L-broth supplemented with 0.4% maltose at 37°C on a shaker until the $O.D._{600}$ reached 0.5. The cultures were then cooled on ice and pelleted at 4°C. The cell pellet was resuspended in 15 ml of cold 10 mM $MgS0_4$ or SM buffer. In all cases plating cells were stored at 4°C and used within 3 days.

The phage were plated as follows: a fixed number of phage particles was added to 200 μl of plating cells and incubated at 37°C for 30 min. The phage and cells were added to melted top agar (3 ml for small plates and up to 6 ml for large plates), mixed and poured onto an L-agar plate. The layer of top agar was allowed to solidify and the plate was incubated at 37°C overnight.

FIG. 7 GENOMIC CLONING IN EMBL3

1) Generation of cloning fragments

Genomic fragments for cloning were generated as already described. In brief, cloning fragments were generated by a series of Mbo I digestions of HMW genomic DNA. The fragments were sized by electrophoresis and the specific size of fragments to be cloned were isolated directly from the preparative gel. For cloning 0.35 µg of genomic cloning DNA fragments was precipitated and resuspended in 2 µl of sterile water, heated at 52°C for 10 min and then chilled on ice for 7 min prior to ligation.

2) Ligation of genomic fragments to EMBL3

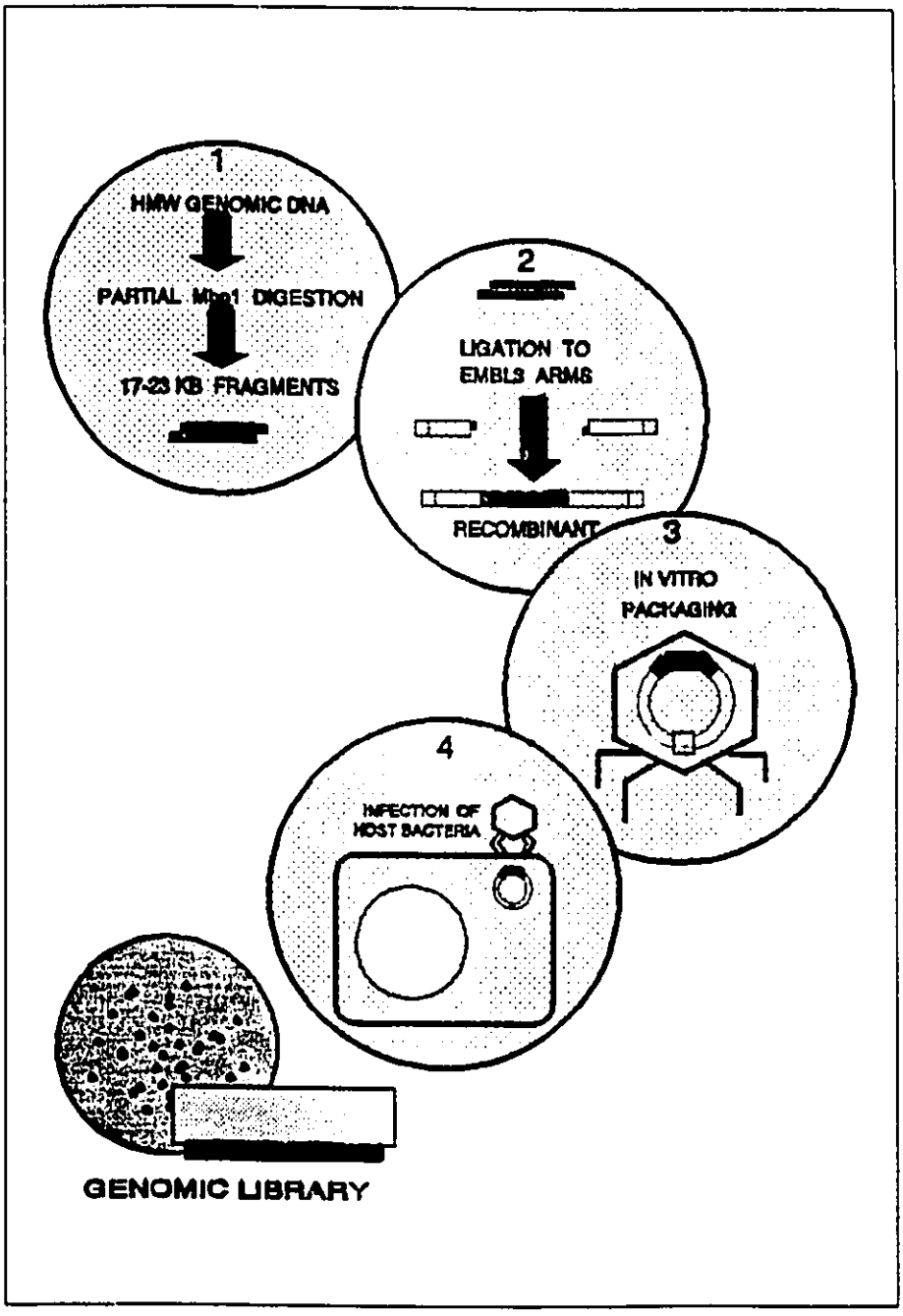
Ligation of the genomic DNA fragments to 2 µg of vector arms was accomplished in a final reaction volume of 10 µl with 2.5 U of T4 DNA ligase at 21°C for 6 h. EMBL3 vector arms were obtained from Promega Biotec (Madison, Wisconsin).

3) Packaging of recombinants into viable phage

The packaging system used was obtained from Promega Biotec (Madison, Wisconsin). An entire tube of packaging extract was added to the ligation tube, mixed and incubated at 21°C for 2 h.

4) Storage and infection of bacteria

Packaged libraries were stored in 450 µl of SM buffer with the addition of 20 µl of chloroform. Libraries were plated out as defined in materials and methods using the bacterial cell line Mb406.


1
HMW GENOMIC DNA
PARTIAL Mbo1 DIGESTION
17-23 KB FRAGMENTS
2
LIGATION TO
EMBL3 ARMS
RECOMBINANT
3
IN VITRO
PACKAGING
4
INFECTION OF
HOST BACTERIA
GENOMIC LIBRARY

2.600) SUBCLONING OF INDIVIDUAL λ CLONES

Specific EMBL3 genomic clones and λgt10 cDNA clones were subcloned into the vector PTZ19R. EMBL3 clones were digested with either Bam HI, Eco RI, or Sal I/Eco RI while inserts from λgt10 clones were released by an Eco RI digestion. Fragments generated from 500 ng of intact DNA were ligated to similarly linearized dephosphorylated vector (500 ng) with 1 U of T4 ligase at 21°C for 3h. The resulting recombinant constructs were used to transform competent E. coli DH5α bacteria by heat shock. Transformation was performed by adding 200 μl of competent cells to each ligation reaction and incubation on ice for 30 min and 42°C for 3 min. 400 μl of 2YT medium was added to each tube and the cells were further incubated at 37°C for one h and then on ice for 5 min. The transformed bacteria were plated out on LB-Amp plates containing isopropylthio-B-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-B-D-galactoside (X-gal) and incubated at 37°C overnight. In the selection system imposed by the plasmid, colonies with recombinant plasmids are white while those with intact plasmid (unaltered) are blue.

Competent cells were derived from a growing culture of E. coli DH5α cells ($O.D._{600}$ of 0.5) in 100 ml of 2YT broth. The cells were chilled on ice and pelleted by centrifugation. The cell pellet was resuspended in 40 ml of TfbI (30 mM KOAc, 100 mM $RbCl_2$, 10 mM $MnCl_2$, 15% glycerol, pH5.8) and incubated on ice for 5 min. The cells were pelleted and resuspended in 4 ml of TfbII (10 mM MOPS, 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15% glycerol, pH 5.8) and incubated on ice for 15 min. Cells

were used directly for transformation or stored in 200 μl aliquots at -70°C and thawed just prior to use.

2.700) ISOLATION OF INTACT CLONES

2.710) Preparation of plasmids

Plasmids were grown in selective medium as defined by the specific suppliers. Plasmid DNA was extracted from 250 or 20 ml cultures using standard protocols (Birnboim and Doly, 1979).

2.720) λgt10 Clones

Recombinant λgt10 DNA was isolated from 10^5 independent clones from the unamplified human chorioretinal cDNA library using a plate lysate technique (Maniatis *et al.*, 1982). The extraction of phage DNA from the lysate was performed using a modification of the protocol of Y. Myal (University of Manitoba, Department of Physiology, Winnipeg, Manitoba, unpublished). MB406 bacteria were infected with phage in order to obtain high density plating on L-agar plates, incubated at 37°C overnight, and the phage were recovered by elution into SM buffer (20 ml per plate). The phage lysate was incubated at 37°C with DNAse (10 μg/ml) and RNAse A (100 μg/ml) for 1 h. Cellular debris was pelleted by centrifugation at 3000 rpm for 30 min at 4°C. The supernatant was isolated, NaCl was added to 1 M, PEG 8000 was added to 1% and allowed to dissolve completely and incubated on ice for 1 h. The sample was subsequently mixed on a rotary mixer at 4°C overnight.

Phage were pelleted by centrifugation at 11,000g for 10 min at 4°C and the pellet was redissolved in SM (500 μl per original volume). SDS was added to 100 mM and proteinase K was added to 300 μg/ml and incubated at 65°C for 1 h. The sample was then extracted, once with an equal volume of phenol (saturated with TE), phenol/chloroform and once with chloroform/isoamyl alcohol (24:1) During each extraction the sample was gently mixed for 5 min before micro-centrifugation and subsequent recovery of the aqueous phase. The DNA was precipitated with an equal volume of isopropanol, recovered by spooling, washed with 70% ethanol, and redissolved in TE.

2.730) EMBL3 Clones

Recombinant genomic EMBL3 DNA was isolated using a column technique (Helms *et al.*, 1987). High titre stocks (10^9 pfu/ml) of each respective clone were made by successive amplification of phage titres. Five μl of the phage stock was dotted onto a lawn of MB406 cells on an L-agar plate and incubated at 37°C for 8 h. The area of lysis was transferred into 200 μl of SM buffer and the phage was eluted at 4°C overnight. 300 μl of MB406 plating cells were added to the phage eluate and incubated at 37°C for 1 h. The total contents were transferred to 5 ml of prewarmed L-broth supplemented with 100 mM $MgSO_4$ and incubated at 37°C for 6 h with shaking. 100 μl of chloroform was added to the culture to complete cell lysis. An equal volume of water was added to the phage and the entire solution was centrifuged to pellet cell debris. The subsequent supernatant was isolated.

The supernatant was passed through a column made in a 10 ml syringe with a sterile glass wool plug and packed with pretreated DE52 preswollen beads (DE52 was acidified to a pH of below 2 with 0.1 N HCl and then brought back up to a pH above 7.0 by successive washes with 0.1 M Tris-HCl pH 8). 5 ml of chase buffer (10 mM Tris-HCl pH 8, 10 mM MgOAc, 60 mM NaOAc) was applied to the column followed by 1 ml of elution buffer (10 mM Tris-HCl pH 8, 50 mM MgOAc). Subsequently 3 additions of 0.6 ml of elution buffer were made (in these cases the eluate from the column was isolated into separate tubes for DNA extraction).

Each tube was extracted once each with an equal volume of phenol, phenol/chloroform and chloroform/isoamyl alcohol (24:1). In each extraction the tubes were mixed by inversion for 5 min and the phases were separated by micro-centrifugation for 5 min at room temperature. Subsequently 5 μl of phenolized dextran, 1/10 th volume of 3 M NaOAc and 2.5 volumes of ice cold ethanol were added to each tube. The tubes were incubated at -70°C for at least 30 min to complete DNA precipitation. The DNA was pelleted by micro-centrifugation at 4°C for 30 min. The supernatant was decanted and the pellet was allowed to air dry before redissolving in 30 μl of TE.

2.800) SYNTHESIS OF DNA PROBES

2.810) Isolation of cloned DNA fragments

2.811) Isolation on glass beads: Clones were digested with specific restriction

enzymes to release the insert from the vector under the conditions suggested by the enzyme supplier. The digested samples were electrophoresed on a 0.7% agarose gel using a TAE buffer system to size fractionate the DNA fragments (Maniatis *et al.*, 1982). The gel was stained with ethidium bromide (1 μg/ml) and the electrophoretic pattern was visualized on a 306 nm UV transilluminator. The fragment corresponding to the cloned insert was cut out of the gel and retrieved by binding the DNA to glass beads (Gene Clene™, Bio101, La Jolla, CA) in the presence of NaI, followed by the elution into TE buffer, pH 8 (Vogelstein and Gillespie, 1979).

2.812) Paper isolation: Insert DNA was isolated using NA-45 DEAE Membrane (Schleicher and Schuell Inc, Keene, N.H.) as suggested by the manufacturer. Clones were digested with specific restriction enzymes and electrophoresed as above except that ethidium bromide was added to the agarose prior to casting the gel. The electrophoretic pattern was visualized by UV transillumination. The fragment corresponding to the cloned insert was visualized and a piece of treated DEAE membrane was inserted into the gel just below it. The insert fragment was electrophoresed onto the membrane. The DNA was subsequently eluted off the membrane in a high salt buffer (1 M NaCl, 0.1 mM EDTA, 20mM Tris pH 8) at 65°C (Dretzen *et al.*, 1981).

2.813) Low Melting Agarose: Clones were digested and size fractionated as described above except that low melting agarose was used. The gel was stained with ethidium

bromide (10 μg/ml) and the electrophoretic pattern was visualized by UV transillumination. The fragment corresponding to the cloned insert was excised and weighed. Three volumes of sterile water was added to the sample and boiled for 10 min. Before cooling the sample was divided into 50 μl aliquots and stored at 4°C.

2.814) Polymerase Chain Reaction (PCR): Recombinant λgt10 insert DNA from intact clones were PCR amplified (Saiki *et al*, 1985) using the GeneAmp DNA amplification Reagent Kit (Perkin-Elmer Cetus, Ottawa, Ont.) and λgt10 Eco RI forward and reverse sequencing primers (New England Biolabs, Waverly, MA). In total, 25-30 cycles as described in Fig. 8 were performed.

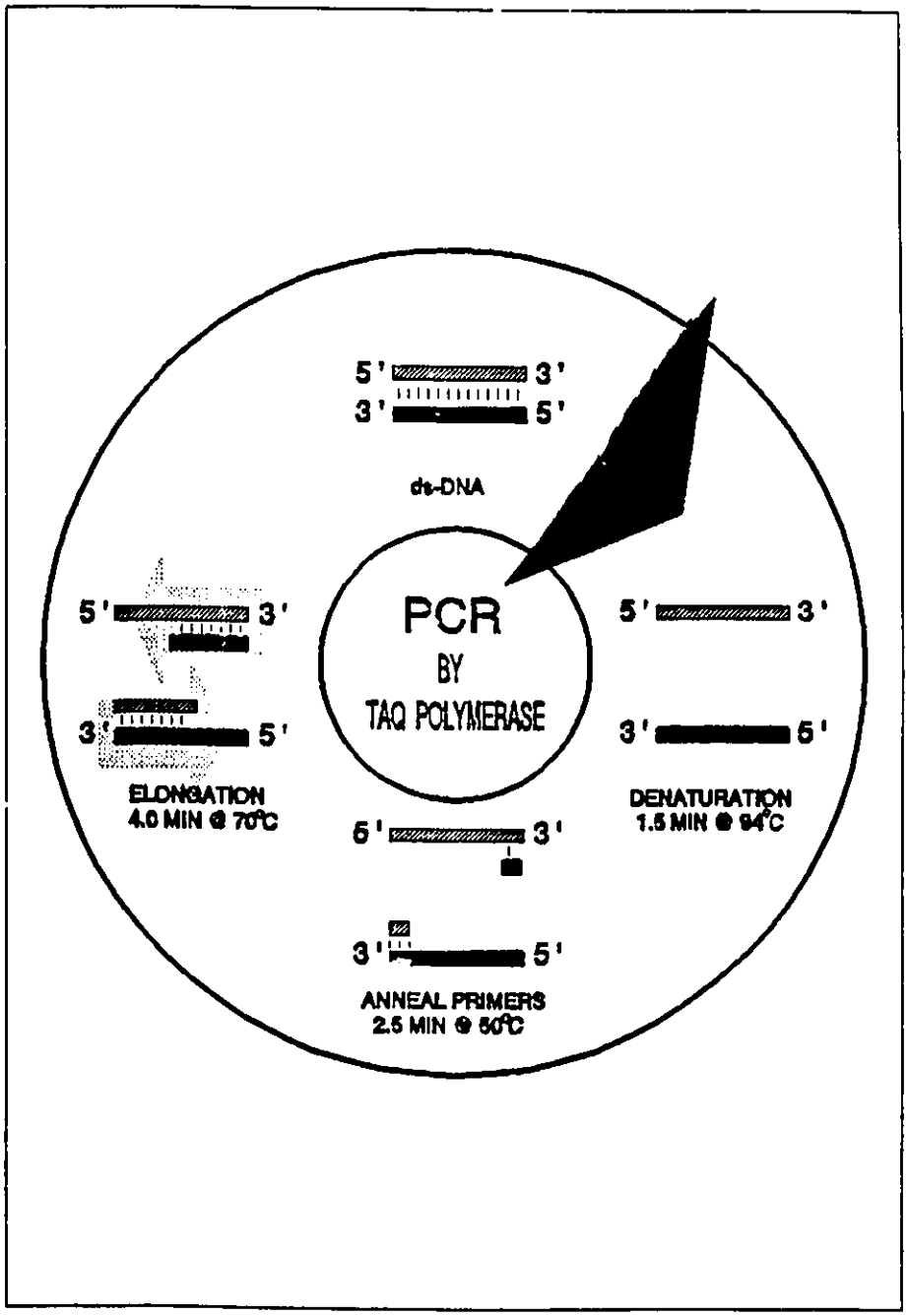
2.820) <u>Radioactive labeling of probes</u>

2.821) End labeling of single stranded oligonucleotides (oligo): Oligos were end labeled according to Sambrook *et al.*, (1989). 100 ng of oligo was mixed with 1x polynucleotide kinase buffer (50 mM Tris/HCl pH 7.5, 10 mM $MgCl_2$, 5 mM dithiothreitol, 0.1 mM spermidine-HCl, 0.1 mM EDTA pH 8.0), 5 μCi γ^{32}P ATP, and 20 units of polynucleotide kinase to a final volume of 50 μl and incubated for 30 minutes at 37°C. The reaction was stopped by the addition of 1 μl of 500 mM EDTA and unincorporated nucleotides were removed by passing the reaction down a Sephadex G50 spun column. Oligomers were custom synthesized by the University of Ottawa Biotechnology Research Institute (Ottawa, Ontario). The specific activity of the final probe ranged from 10^7-10^8 cpm/μg of DNA.

FIG. 8 THE POLYMERASE CHAIN REACTION (PCR).

The polymerase chain reaction consists of the continuous cycling of three steps, denaturation of the double stranded DNA template, annealing of specific DNA primers which flank the segment of DNA to be amplified, and the 5' elongation from these primers to regenerate the respective second strands.

Amplification of the chorio-retinal cDNA library consisted of 25-30 cycles as defined in the figure except that the denaturation period of the initial cycle was extended to 10 min and the elongation of the final period was extended to 13 min. The final PCR product was extracted with an equal volume of chloroform and precipitated. The yield of amplified DNA could be estimated by gel analysis, spectrophotometric analysis, or addition of $\alpha[^{32}P]$dATP and $\alpha[^{32}P]$dCTP during amplification.


5' 3'
3' 5'
ds-DNA
PCR
BY
TAQ POLYMERASE
5' 3'
3' 5'
ELONGATION
4.0 MIN @ 70°C
5' 3'
3' 5'
DENATURATION
1.5 MIN @ 94°C
5' 3'
3' 5'
ANNEAL PRIMERS
2.5 MIN @ 50°C

2.822) Nick translation: Plasmids and genomic DNA were nick translated following the standard procedure of Rigby *et al.*, (1977) using a commercially available kit (Amersham, Oakville, Ont.). The specific activity of the final probe ranged from 10^7-10^8 cpm/μg DNA.

2.823) Multiprime: In initial experiments isolated DNA was oligo-labeled as described by Feinberg and Vogelstein (1983) using a commercially available kit (Amersham, Oakville, Ont.). In later experiments oligo labeling was performed using the NEN Random Primer extension labeling system (NEN, Dupont, Lachine, Quebec). The specific activity of the final probe ranged from 10^8-10^9 cpm/μg DNA.

2.824) Depletion of repetitive sequences from probes: Two methods for depletion of repetitive sequences were used:

i) Preassociation: Radioactive probe was mixed with an excess of sheared human genomic DNA in the ratio of 1:10,000 probe to genomic DNA (unless otherwise noted) in 5XSSC. The DNA mix was boiled for 10 min and incubated at 68°C for 15 minutes (Sealey *et al.*, 1985). The preassociated probe was used directly for hybridizations.

ii) Subtraction: Denatured multiprimed labeled probe was mixed with a maximum of 100 μg of human genomic DNA-bound cellulose in a ratio of 1 to 500 (probe:DNA-cellulose) in hybridization solution (50% formamide, 0.75 M NaCl, 50 mM sodium phosphate buffer (pH 6.8), 5 mM EDTA, 1 x Denhardt's, 100 μg/ml

denatured calf thymus DNA). The mix was incubated at 80°C for 2 min and then transferred to 37°C (with mixing) for 12 h. The DNA-cellulose was pelleted and the supernatant was isolated. The probe in the supernatant was remelted at 80°C for 10 min, cooled to 37°C and remixed with the DNA-cellulose pellet. The tube was again incubated at 37°C for 12 h with mixing. The cycle of melting and reannealing of probe to DNA-cellulose was carried out for 3 days. The DNA-cellulose was prepared according to Hochgeschwender *et al.*, (1989) and the specific batch of human genomic DNA-bound cellulose used in the present study was a kind gift from Dr. M. Brennan (Bethesda, Maryland).

2.900) ANALYSIS OF NUCLEIC ACIDS

2.910) Clone Analysis

2.911) Plaque lift analysis: Plates of the respective phage library were chilled at 4°C for 1 h before lifts were taken. Lifts were taken using Hybond N filters (Amersham, Oakville, Ont) and processed according to Allday and Jones (1987). Prehybridization, hybridization, filter wash conditions and autoradiography were as defined in section 2.932.

2.912) Phage dot analysis: Five μl of isolated phage stocks were dotted onto a pre-poured lawn of competent bacterial host cells in a grid pattern. The plate was allowed to absorb the liquid phage and was then incubated at 37°C overnight. For further analysis a plaque lift was taken and processed as described above (sec.2.932).

Prehybridization, hybridization, filter wash conditions and autoradiography were as defined in section 2.932.

2.913) Grunstein-Hogness Hybridization: Transformed bacterial colonies were picked onto a Hybond membrane placed on top of an L-plate with ampicillin. The colonies were allowed to grow at 37°C overnight. The membranes were then submitted to four 5 min treatments as follows. Membranes were placed onto 10% SDS (colony side up) and in succession transferred to solution 1 (0.5 N NaOH, 1.5 M NaCl), solution 2 (1.5 M NaCl, 0.5 M Tris-HCl pH 8.0), and solution 3 (2xSSC). The lifts were allowed to dry before heat baking at 80°C (under vacuum) for 2 h. Prehybridization, hybridization, filter wash conditions and autoradiography were as defined in section 2.932.

2.920) Electrophoresis

2.921) RNA: RNA was electrophoresed in formaldehyde gels (Sambrook *et al.*, 1989). RNA samples were prepared for electrophoresis by diluting in a solution of final concentration of 1xMOPS buffer (20 mM morpholinopropanesulfonic acid, 5 mM NaOAc, 1 mM EDTA, pH 7.0), 17.5% formaldehyde, 50% deionized formamide, followed by an incubation at 55°C for 15 min and quick cooling on ice. If bands were to be visualized ethidium bromide was added to a final concentration of 25 μg/ml prior to incubation. 10 μl of loading buffer (0.25% bromophenol blue (w/v), 0.25% xylene blue (w/v), 1mM EDTA pH 8.0, 50% glycerol) was added to

each sample and loaded onto a 2.5% agarose gel (made in 18% formaldehyde and 1xMOPS buffer). Electrophoresis was carried out in a buffer consisting of 8.3% formaldehyde and 1xMOPS buffer. Gels were run at 30 volts for 16 h. Ethidium bromide stained samples were visualized by UV transillumination and photographed using an orange filter and Ilford FP4 film.

2.922) DNA: DNA samples were either run intact or predigested with restriction endonucleases. In the case of EMBL3 clones digestion was carried out essentially as recommended by the suppliers, with the exception of the digestion time which was extended to 6 h. Full digestion of genomic DNA was carried out overnight. In all cases 1/10th volume of 10X TAE loading buffer was added to the DNA samples and incubated at 65°C for 5 min prior to loading.

DNA samples were electrophoresed on 0.5-2.0% Tris-acetate/agarose gels (TAE buffer system) depending on the specific experiment (Maniatas *et al.*, 1982). In the case of low melting agarose gels were cast and run as previously described (Maniatis *et al.*, 1982). After electrophoresis the gels were stained with ethidium bromide (1 μg/ml) and photographed as described above.

2.930) <u>Southern analysis</u>

2.931) Transfer: DNA samples fractionated on agarose gels were processed and transferred directly to nylon membranes (Hybond N or HybondN+, Amersham, Oakville, Ont) by capillary blotting using the protocols provided by the supplier.

In the case of the Hybond N membrane, gels were pre-soaked in 0.25 M HCl for 15 min, denatured in 1.5 M NaCl, 0.5 M NaOH for 30 min, and neutralized in 1.5 M NaCl, 0.5 M Tris-HCl pH 7.2 for 30 min prior to capillary blotting using 20XSSC as transfer buffer. When 1.5 M NaCl, 0.25 M NaOH was used as transfer buffer, the gel was presoaked in 0.25 M HCl for 15 min and denatured in 1.5 M NaCl, 0.5 M NaOH for 30 min before setting up for capillary blotting. In both cases the transfer was carried out overnight and the final blot was baked at 80°C under vacuum for 2 h prior to prehybridization.

In the case of Hybond N+, gels were soaked in 0.25 M HCl for 20 min prior to capillary blotting using 0.4 N NaOH as transfer buffer. Blotting was carried out for 2-3 h. The resultant blot was immediately ready for prehybridization.

2.932) Hybridization: Prior to hybridization, blots were prehybridized in 6XSSC, 5xDenhardt's solution, 0.5% SDS and 3 mg of heat denatured sonicated calf thymus DNA at 65°C for 4 h. Hybridization was carried in the same buffer with 100 ng of denatured radiolabeled probe at 65°C for 24 h. In cases when the amount of radiolabeled probe was low 10% dextran sulphate was added to both the prehybridization and hybridization solutions to enhance the degree of hybridization. In cases where the probe to be used was derived from a EMBL3 genomic clone, the probe was pre-annealed to 100 μg of total human DNA at 65°C for 15 min prior to incubation with the blot. Filters were subsequently washed at 65°C (two washes in

2 x SSC for 15 min, one wash in 2 x SSC and 0.1% SDS for 30 min). A final high stringency wash in 0.1 x SSC for 10 min at 65°C was included as required. Blots were wrapped in plastic and autoradiographed using Cronex X-ray film in a Kodak X-Omatic X-ray cassette with intensifying screens.

2.940) Slot blot analysis:

The slot blot apparatus supplied by Schleicher and Schuell Inc (Keene, N.H.) and the manufacturer's instructions for DNA samples were used. The resultant slot blot filter was baked at 80°C for 2 h. Prehybridization, hybridization, filter wash conditions and autoradiography were as defined above (section 2.932).

2.950) Nucleic acid sequencing:

Sequencing was performed by the chain termination method (Sanger, Milken and Coulson, 1977) in the presence of $\alpha[^{35}S]$-dATP and T7 and SP6 primers, using Sequenase™, a modified version of the T7 DNA polymerase (Tabor and Richardson, 1987). Samples were run on 6-8% polyacrylamide, 8M urea gels at 50W, 2000 V, and 30mA for 2.5-5 h. The DNA was fixed within the gel with 5% methanol, 5% acetic acid prior to drying. Autoradiography was performed at room temperature using Cronex-4DC film and a single intensifying screen. Computerized analysis of the sequences obtained and comparison with the GenBank database were performed using Microgenie™ computer software (Queen and Korn, 1984) (Beckman, Mississauga).

2.960) S1 Nuclease Protection Assay:

The protocol used is essentially as defined by Sambrook *et al.*, (1989). A minimum of 1 μg of total RNA in 30 μl of hybridization buffer (40 mM PIPES pH 6.4, 1 mM EDTA, 0.4 M NaCl, 80% deionized formamide) was mixed with approximately $1x10^5$ to $1x10^6$ cpm of labeled oligomer probe. The mix was heated at 85°C for 10 min and then allowed to hybridize at 30°C overnight. The reaction was chilled on ice and 300 μl of pre-chilled S1 buffer (0.21 M NaCl, 5 mM NaOAc pH 4.5, 4.5 mM $ZnSO_4$, 20 μg/ml single stranded DNA, 200 U/ml S1 nuclease) was added. The S1 nuclease digestion was performed by incubating the reaction at 37°C for 2 h. The reaction was chilled on ice and 80 μl of stop buffer (4 M NH_4OAc, 50 mM EDTA, 50 μg/ml yeast tRNA) was added. The reaction products were precipitated and the pellets were redissolved in sequencing loading buffer and analyzed on a 6% polyacrylamide sequencing gel.

CHAPTER THREE

ISOLATION OF CHORIO-RETINAL RNA AND THE SYNTHESIS OF cDNA LIBRARIES

3.000 INTRODUCTION

At the inception of the present project in 1986 no known human chorio-retinal cDNA libraries were available; this necessitated the need to construct one for the present study. Later in that year however Nathans and co-workers published their work on the isolation of the human opsin genes in which they had made and screened such libraries. In the last three years other groups have made retinal cDNA libraries and such libraries are now commercially available. The chorio-retinal library generated in the present study represents one of the first libraries for the human eye.

3.100) RNA

3.110) RNA Yield

The isolation of intact RNA from human choroid/retina tissue depends greatly on the ability to obtain fresh tissue. On samples obtained approximately 6 h post mortem the yield of total RNA is approximately 200 $\pm$ 70 μg per pair of human choroid/retinas (Table 1). The yield of total RNA on retina samples obtained 12 h or more post mortem and samples which were initially snap frozen without predissection of vitreous material gave lower yields (not shown). Although the LiCl/urea method for RNA isolation has been used successfully in the isolation of RNA from many tissues the present application is the first with human choroid/retina tissue. As such, it is a necessity to verify the quality of the RNA.

TABLE 1 SUMMARY OF TOTAL RNA YIELDS EXTRACTED FROM HUMAN CHORIOD/RETINA SAMPLES.

Tissue samples used for extraction were flash frozen in liquid nitrogen within 6 hours post mortem and stored at -70°C until needed. Total RNA was extracted using a LiCl/urea method (Auffray and Rougeon, 1979; Tenniswood and Simpson, 1982). RNA yields from 11 independent extractions have ranged from 136 μg of total RNA per human choriod/retina sample to 294 μg RNA per sample. The mean yield of total RNA extracted was determined to be 207 μg of RNA with a standard deviation (sd) of 66 μg per individual choriod/retina sample. One sample consists of a pair of choriod/retinas isolated from the same donor.

EXTRACTION (N)	# RETINA SAMPLES	TOTAL RNA YIELD(μg)	RNA YIELD PER RETINA SAMPLE(μg)
1	1	136	136
2	2	528	264
3	2	560	280
4	2	590	295
5	1	152	152
6	1	138	138
7	1	246	246
8	1	284	284
9	1	146	146
10	1	160	160
11	1	174	174
		SAMPLE SIZE (N) MEAN YIELD sd	11 207 66

3.120) Integrity of the Total RNA extracted

Total RNA consists of rRNA, tRNA, hnRNA as well as mRNA species. The two most prominent RNAs are the 28 and 18S ribosomal subunits which are clearly visible upon denaturing gel analysis (Fig. 9). General degradation will affect the RNA pool as a whole and any signs of degradation of total RNA can be reflected visually in the integrity of the 28S and 18S bands. The tightness of the 28S and 18S bands of the chorio-retinal total RNA extracted suggest that there is no evident degradation thus demonstrating that the LiCl/urea method is satisfactory for isolation of intact RNA from human choroid/retina tissue.

3.130) Analysis of translation products

The functional integrity of the total chorio-retinal RNA was tested by translation in an *in vitro* translation system and analysis of the translation products. Discrete protein bands are present in both the PAGE (Fig. 10) and 2 dimensional gel analysis (Fig. 11) of chorio-retinal translation products. In the SDS-PAGE analysis proteins are separated solely on the basis of molecular weight. In the 2 dimensional analysis proteins are first separated on the basis of isoelectric focusing points and then according to molecular weight. The control reaction for the *in vitro* translation assay is to run the system without the addition of RNA. Analysis of this control reaction reveals the translation products inherent to the reticulocyte system which represents the background of the system (Fig. 10, lane 1). Analysis of the *in vitro* translation products of chorio-retinal RNA by both methodologies revealed

FIG. 9 ANALYSIS OF TOTAL HUMAN CHORIO-RETINAL RNA BY FORMALDEHYDE GEL ELECTROPHORESIS.

Total RNA (10 μg) isolated from 8 independent samples was visualized in a 1% formaldehyde gel after ethidium bromide staining. Electrophoresis was carried out at 30V for 16 h as defined in material and methods.

The two most prominent RNA bands are the 28S and 18S ribosomal bands. In all cases they present as tight discrete bands without any signs of degradation, thus inferring that the method of RNA isolation permits the structural preservation of the chorio-retinal RNA.

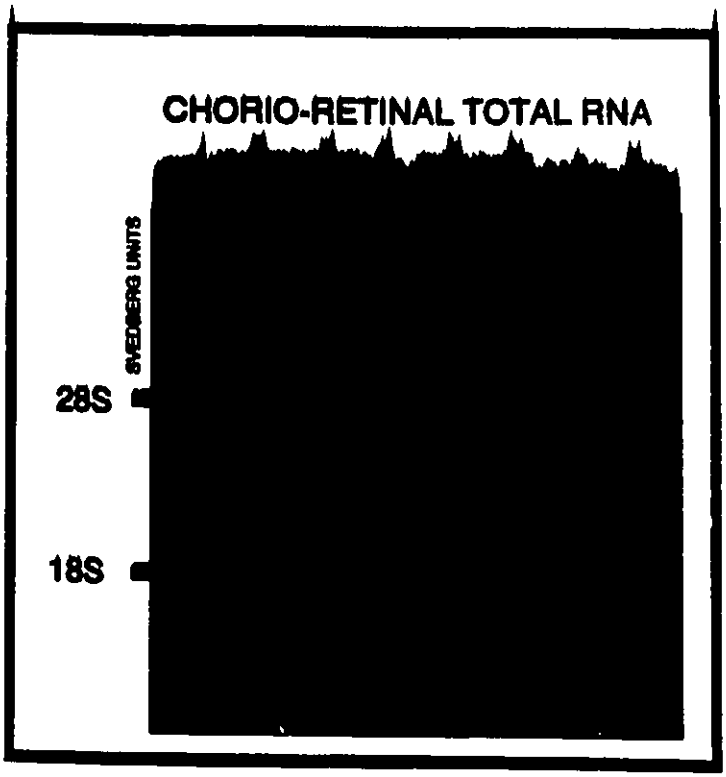

CHORIO-RETINAL TOTAL RNA
SVEDBERG UNITS
28S
18S

FIG. 10 PAGE ANALYSIS OF CHORIO-RETINAL RNA in vitro TRANSLATION PRODUCTS.

Chorio-retinal RNA was translated using a mRNA dependent rabbit reticulocyte lysate system. Shown here are translation products from two independent translation reactions which were analyzed on 10% SDS slab gels under denaturing conditions. The lysate control represents the inherent translation products of the reticulocyte lysate system used and is generated by running an translation reaction without the addition of any external RNA. In each case approximately 10^5 cpm of translation product was loaded per lane.

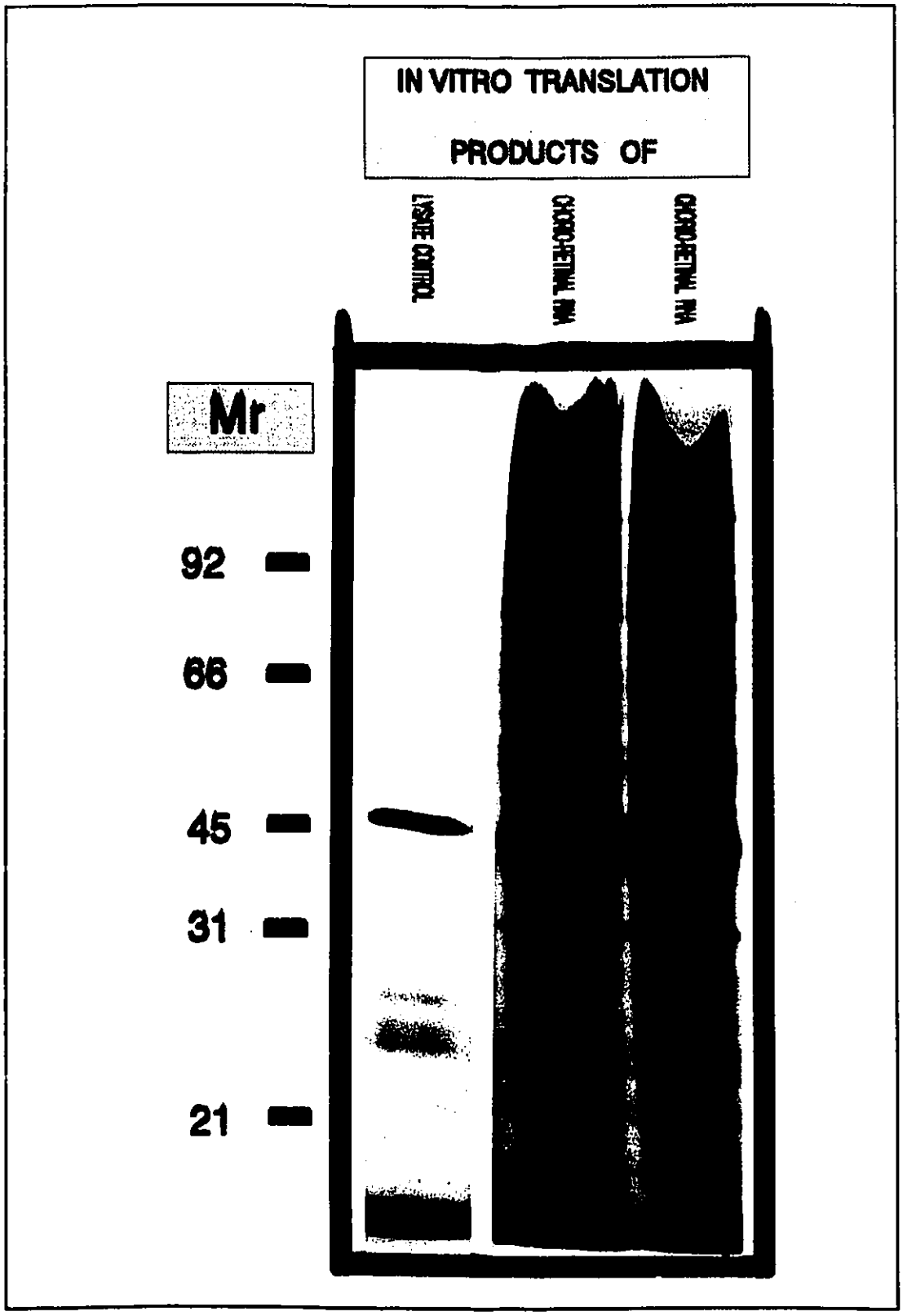
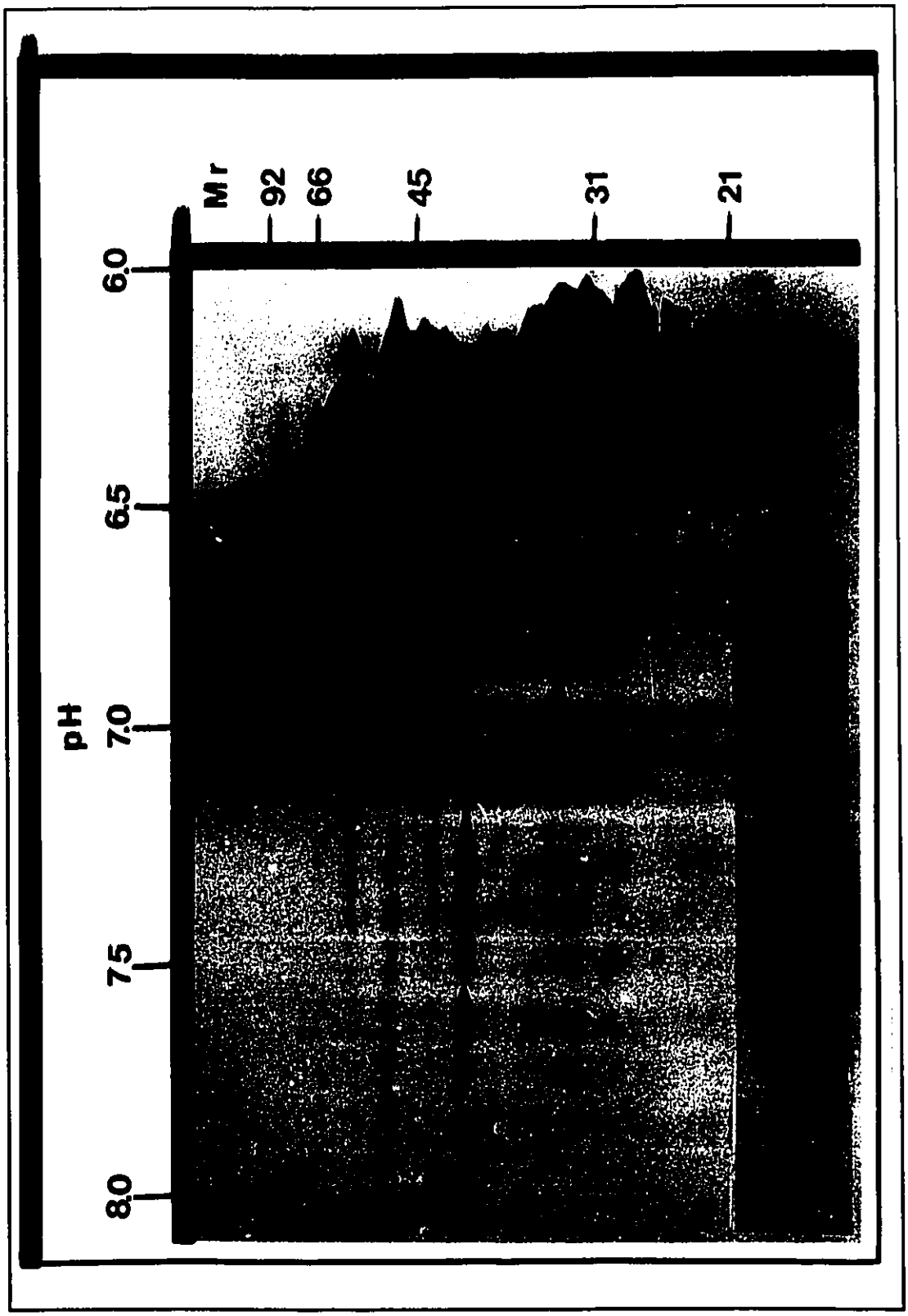

IN VITRO TRANSLATION
PRODUCTS OF
LYSATE CONTROL
CHOROIDAL RNA
CHOROIDAL RNA
Mr
92
66
45
31
21

FIG. 11 TWO DIMENSIONAL PAGE ANALYSIS OF CHORIO-RETINAL in vitro TRANSLATION PRODUCTS.

A representative two dimensional PAGE analysis of chorio-retinal RNA translation products is shown here. The first dimension was performed in an isofocusing tube gel using LKB pH 3-10 ampholines followed by a second electrophoresis on a 15% SDS slab gel under denaturing conditions. In total 3×10^5 cpm of translation product was loaded onto the isofocusing tube gel. The background bands inherent to the reticulocyte lysate system have been boxed in; the background indicated was determined by two dimension PAGE analysis of reticulocyte lysate translation products without addition of external RNA (gels not shown).


Mr
92
66
45
31
21
pH
6.0
6.5
7.0
7.5
8.0

discrete protein bands indicating that the chorio-retinal RNA isolated was functionally intact.

3.140) Isolation of chorio-retinal poly(A)+ RNA

Because of the relatively low amounts of total RNA extracted per retina pair, total extracted RNA from several samples were pooled prior to the isolation of poly (A)+ RNA. The yield of poly(A)+ enriched RNA from total RNA using the oligo(dT)-cellulose methodology was approximately 2%-5%. The electrophoretic profile of both the poly(A)+ and the poly(A)- fractions are presented in Fig. 12. As stated earlier the two most abundant RNAs are the 28S and 18S ribosomal subunits. These two RNAs segregate into the poly(A)- fraction. The level at which they remain in the poly(A)+ fraction is therefore a measure of the efficiency of the separation procedure. As is evident in lane 2 of Fig. 12 (poly(A)+ sample) both the bands that correspond to the 28 and 18S bands that are very prominent in the poly(A)- lane are not detectable over the intensity of the other species of RNA in that fraction.

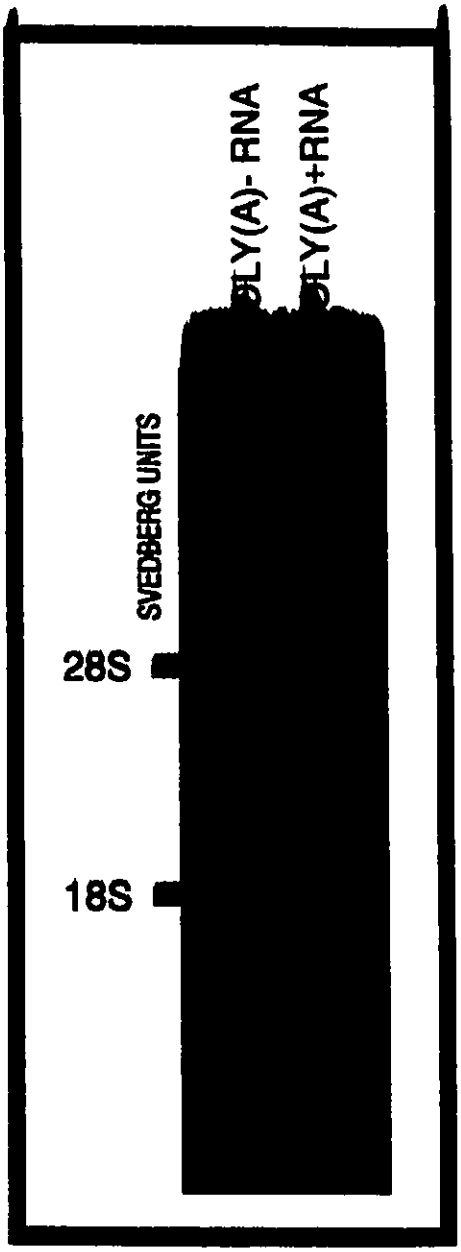
3.200) cDNA: INTRODUCTION

A cDNA library represents a clone bank of the expressed messages within a given tissue. The generation of such a library consists of the isolation of high quality total RNA and poly(A)+ RNA, synthesis of a cDNA copy of each RNA species and the subsequent incorporation of this cDNA into a vector.

FIG. 12 ANALYSIS OF THE POLY(A)- AND POLY(A)+ FRACTIONS ISOLATED FROM TOTAL HUMAN CHORIO-RETINAL RNA.

Total chorio-retinal RNA was fractionated into a poly(A)- and poly(A)+ fraction by oligo-(dT) cellulose chromatography. Shown here is the electrophoretic analysis of 10 μg of the poly(A)- fraction and 5 μg of the poly(A)+ fraction. Electrophoresis was carried out in a 1% agarose gel using a formaldehyde buffer system for 16 h at 30 volts.

In general the poly(A)+ fraction represents a RNA fraction enriched in mRNA species, while the poly(A)- fraction represents a mRNA depleted fraction. What is clear from the analysis is that the 28S and 18S ribosomal sequences clearly segregate into the poly(A)- fraction and appear as discrete bands, thus indicating that the structural integrity of the RNA has been preserved. In addition visualization of the 18 and 28S bands cannot be observed above the RNA smear in the poly(A)+ fraction suggesting that the method of fractionation is efficient.


POLY(A)- RNA
POLY(A)+RNA
SVEDBERG UNITS
28S
18S

3.210) Synthesis of ds-cDNA and size fractionation

6μg of poly(A)+ RNA was used as template in the synthesis reaction. Based on the amount of incorporated radioactivity in the resulting sample it was calculated (using equation 1) that approximately 500 ng of ds-cDNA was synthesized.

EQUATION 1

$$\text{ds-cDNA yield (ng)} = \frac{\text{nmole cold dNTP added} \times 326 \times \text{total counts} \times 4}{\mu\text{Ci dNTP used} \times 2.2 \times 10^6}$$

$$= \frac{30\text{ nm} \times 326 \times 1.3 \times 10^6 \times 4}{50\mu\text{Ci} \times 2.2 \times 10^6}$$

$$= 470 \text{ ng ds-cDNA}$$

The synthesized ds-cDNA was size fractionated prior to cloning in order to assure that only longer sequences would be cloned. The elution profile demonstrates that the first peak of ds-cDNA was eluted from the column between fractions 7-15 (Fig. 13). These fractions were pooled, precipitated and used for the cloning experiments. In total 260 ng of size fractionated ds-cDNA was recovered for further manipulations.

In order to determine the average size of the pooled fractions isolated above, 1 μl of the pooled eluate was electrophoresed (Fig. 14). From this analysis it was determined that the size of the fractionated ds-cDNA isolated ranged from 200 bp

FIG. 13 THE ELUTION PROFILE OF SIZE FRACTIONATED ds-cDNA

Prior to cloning the ds-cDNA was size fractionated on a Sepharose 4B column. One drop fractions were collected and radioactivity was measured by the Cerenkov method in a nuclear Chicago Isocap 300 scintillation counter. The data is presented as the percentage of recovered radioactivity versus fraction number. The bulk of the radioactivity elutes in two peaks. One which spans from fraction 7 to 15 and a second one which spans fractions 19-22. The early peak harbors the larger ds-cDNA species and as indicated by the region highlighted in grey. Fractions 7-15 were pooled and used for the cloning procedure.

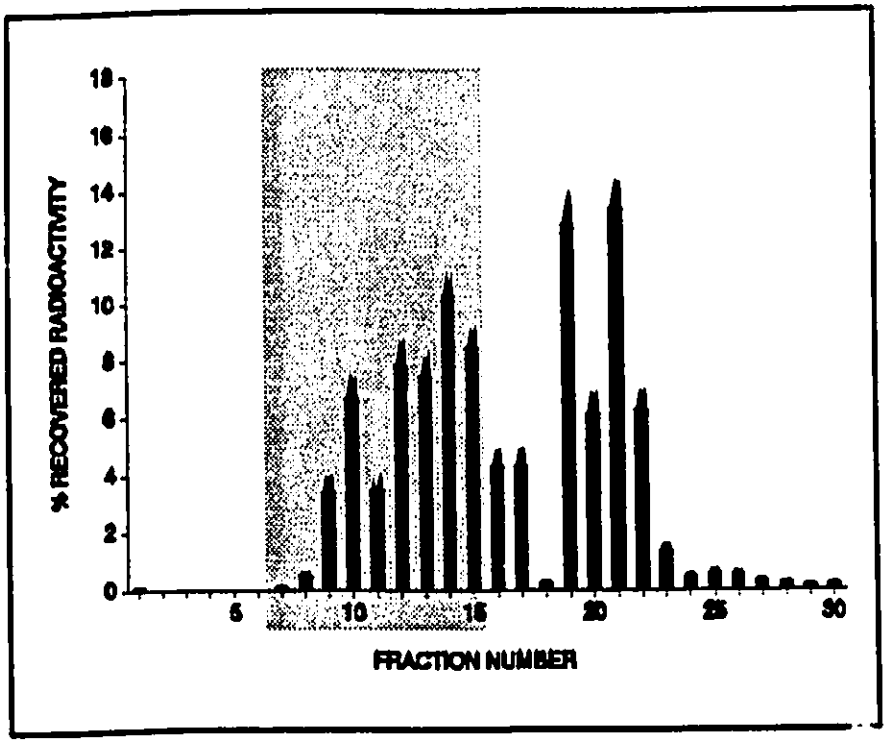
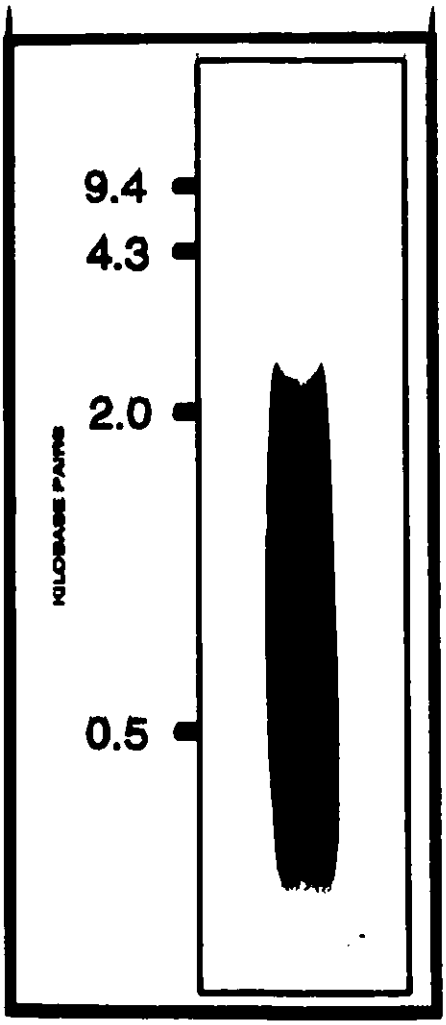

% RECOVERED RADIOACTIVITY
0
2
4
6
8
10
12
14
16
18
5
10
15
20
25
30
FRACTION NUMBER

FIG. 14 DETERMINATION OF THE AVERAGE LENGTH OF ds-cDNA IN THE SIZE FRACTIONATED ds-cDNA POOL.

In order to determine the average length of the ds-cDNA to be used for cloning 1 μl of the pooled fraction gathered after fractionation on a Sepharose 4B column was sized by electrophoresis on a 2 % agarose gel using a TAE system. The gel was dried down under vacuum on a Savant gel drier and autoradiographed. Analysis of the autoradiograph shows that the length of ds-cDNA synthesized ranged from 200 bp up to 2000 bp. The average length is approximately 500 bp. Hind III digested λDNA was used to generate a molecular size marker during electrophoresis and was visualized after ethidium bromide staining and UV transillumination prior to the drying down of the gel.


9.4
4.3
2.0
0.5
KILOBASE PAIRS

to 2000 bp with an average length, based on the maximum intensity on the autoradiograph, of about 500 bp.

3.220) Cloning into the vector

The ds-cDNA at this point was blunt ended. However, since the cloning of blunted ended sequences is not efficient, Eco RI linkers were added to the ends of ds-cDNA and the modified ds-cDNA was shuttled into the Eco RI λgt10 cloning site (Fig. 5). In total, starting from 6 μg of poly(A)+ RNA, 180 ng of ds-cDNA with activated Eco RI linkers was recovered.

Approximately 50 ng of ds-cDNA was used in single ligations to insert into the Eco RI site of λgt10 arms (1 μg) which were subsequently packaged *in vitro* into viable phage. In total three such libraries and an additional arms control library (arms library) were made. The arms library was constructed exactly as the retinal cDNA libraries except that addition of chorio-retinal cDNA insert was omitted.

In the λgt10 cloning system the Eco RI insertion site occurs within the phage repressor gene (cI-). In the normal bacteriophage life cycle it is the expression of this gene that determines that the phage takes the lysogenic pathway of integration. When a sequence is inserted into the Eco RI site it disrupts the cI- gene and inactivates it. In the absence of cI- expression the lytic pathway of phage integration is favoured. In a permissive cell host (L87) cI+ phage will undergo lysogeny giving rise to turbid

plaques while cI- phage will favour the lytic pathway and give rise to clear plaques. The L87 titre therefore represents the total end products of the cloning procedures, the majority of which are religated parental λgt10. However selective cell strains (NM514) with the Hfl+ mutation more effectively differentiate between cI- and cI+ phage. When a Hfl+ (high frequency lysogeny mutation) bacteria is infected with cI+ phage, the majority of cI+ phage are repressed to the extent that plaque formation is repressed, cI- phage however undergo the lytic pathway of infection with normal efficiency (Huynh *et al.*, 1985). Plating the libraries on a Hfl+ cell line decreases the background of non recombinant plaques.

3.230) Initial characterization of the library

For the initial analysis, Each of the libraries were titred on both the L87 (permissive) and NM514 (selective) bacterial host cell lines (Table 2). The titres of the arms library gives an indication that the biological selection system of NM514 is active but not 100% efficient. The NM514 titre (2.0×10^3 pfu/ml) is however much lower than the L87 titre (5.1×10^5 pfu/ml); the ratio of the arms library L87 to NM514 titres (arms selective ratio) represents the proportion of cI+ phage clones that escape the NM514 selection process. The arms selective ratio is a constant for all the libraries made using the same batch of λgt10 arms and packaging extracts. As such the ratio can be used to estimate the non recombinant phage (cI+) background in each of the retinal cDNA libraries constructed. Since the L87 titres are

TABLE 2 TITRATION OF CONSTRUCTED HUMAN CHORIO-RETINAL cDNA LIBRARIES.

In total 3 cDNA libraries derived from human chorio-retinal poly(A)+ RNA were constructed in λgt10. The bacterial cell lines NM514 and L87 were used as the hosts for infection in the determination of phage titres. The average titres of these libraries is $2.0x10^6$ pfu/μg of ds-cDNA insert or $1.0x10^5$ pfu/μg of vector. The combined libraries consists of approximately $3.1x10^5$ independent clones.

The controls for the cloning experiments included:

1. The packaging of intact gt10 phage DNA (intact phage) in order to test the packaging efficiency of the system.
2. The ligation and packaging of the vector arms alone (gt10 arms) to determine the inherent background of the system.

DNA SOURCE	INTACT PHAGE	gt10 ARMS	CHORIO-RETINAL cDNA	CHORIO-RETINAL cDNA	CHORIO-RETINAL cDNA
INSERT DNA	500 ng	----	50 ng	50 ng	50 ng
gt10 ARMS	----	1 μg	1 μg	1 μg	1 μg
L87 TITRE pfu/μg VECTOR	$5.0x10^6$	$5.1x10^5$	$2.2x10^5$	$1.6x10^5$	$1.4x10^5$
NM514 TITRE pfu/μg VECTOR	$1.2x10^3$	$2.0x10^3$	$1.1x10^5$	$1.1x10^5$	$0.9x10^5$
L87/NM514 RATIO	431	255	2	1.5	1.6
BACKGROUND pfu/μg VECTOR	----	----	862	627	549
ADJUSTED TITRE pfu/μg VECTOR	----	----	$1.1x10^5$	$1.1x10^5$	$0.9x10^5$
%RECOMBINANT	----	----	98.2	99.1	98.9
pfu/μg cDNA	----	----	$2.2x10^6$	$2.2x10^6$	$1.8x10^6$

predominantly religated cI+ parental λgt10 phage, the NM514 cI+ background can be estimated by using equation 2. The number of true cI+ recombinants in each library plated on NM514 can therefore be estimated using equation 3 and the final cloning efficiency can be calculated using equation 4. The percentage of recombinants of each library when plated on NM514 can be calculated using equation 5. Results are summarized in Table 2.

EQUATION 2

$$\text{NM514 cI+ background} = \frac{\text{L87 titre}}{\text{Arms selective ratio}}$$

EQUATION 3

$$\text{\# true recombinants plated on NM514 (\#TR)} = \text{NM514 titre} - \text{NM514 cI+ background}$$

EQUATION 4

$$\text{Cloning efficiency per } \mu\text{g cDNA} = \frac{\text{\#TR} \times 1000}{\text{ng cDNA used}}$$

EQUATION 5

$$\text{\% recombinant clones (cI-)} = \frac{\text{\#TR}}{\text{NM514 titre}} \times 100$$

3.240) Physical verification of cloning

In order to determine the physical background of non-recombinants, 11 clones from the chorio-retinal cDNA library were taken at random and submitted to PCR amplification. Of these 11 clones 10 were found to have inserts ranging from 200 bp up to 1 kbp. In order to acquire a better evaluation of the insert size profile of the library, 1 ng of DNA isolated from a plate lysate of the entire library was PCR amplified and electrophoretically analyzed (Fig. 20, chapter five). On the basis of the PCR profile the majority of inserts range from 200bp to 1000bp.

3.250) Theoretical considerations

The typical cell has an estimated 37,000 independent mRNA species (Williams, 1981). These individual species can be categorized according to their relative abundances within the cell. mRNA sequences which exist at 12,000 copies per cell are considered to be abundant sequences, those which exist at 300 copies per cell are categorized as middle abundant sequences, while sequences less than 15 copies per cell are viewed as rare messages (Alberts *et al.*, 1983). The exact number and type of unique expressed sequences and their relative abundance will vary from cell type to cell type.

In the rodent retina the sequence complexity of expressed sequences has been estimated to be approximately 10,000 unique messages (Susan Lindsay, Newcastle, England, personal communication). If this is the case then, one might expect to be

able to find any expressed sequence within a minimum of 10,000-37,000 clones from a retinal cDNA library if all unique sequences have an equal chance of being cloned. However, because of the differences in relative abundances of different mRNA species, abundant sequences will have a higher probability and a higher frequency of being cloned than rare sequences. In order to ensure the representation of any sequence irrespective of its abundance a minimum number of clones that must be present can be estimated using equation 6 (Clarke and Carbon, 1978).

EQUATION 6

$$N = \frac{\ln(1-P)}{\ln [1- (1/n)]}$$

P= the probability of obtaining a given sequence
n= the minimum number of clones to be representative of the mRNA population

N= the number of clones needed to be representative of the mRNA population

To ensure that a given low abundant sequence is present in the population of clones (within a 99% probability) it is necessary to have between 4.6×10^4 - 1.7×10^5 clones. As shown in Table 2 there are a total of 3.1×10^5 clones in the three human chorio-retinal cDNA libraries constructed, roughly twice the upper limit of clones necessary to ensure full representation of expressed sequences.

3.260) The cDNA libraries are representative

It is difficult to determine whether the libraries are truly representative of the mRNA population. One of the few groups of messages that have been well studied at the molecular level in the retina are the opsins. The most abundant of the opsins is rhodopsin which has been estimated to represent 0.5-1 % of the mRNA population of the human retina (Nathans *et al.*, 1986a; McInnes *et al.*, 1988). The red and green opsins are slightly less abundant. These two genes are highly homologous to each other and are expressed at a level which is approximately 10-20 times lower than that observed for rhodopsin (Nathans *et al.*, 1986a). In order to determine whether the present retinal cDNA library was representative, the distribution of the red and green opsins was examined. A clone, HS7, corresponding to the cDNA sequence of the red pigment gene was used to screen a small portion of the unamplified chorio-retinal library. The relative abundance of the red opsin sequence in the library was determined to be roughly 0.1 %, a number which agrees with previous estimations (Nathans *et al.*, 1986a).

CHAPTER FOUR

SYNTHESIS OF A HUMAN X-SPECIFIC GENOMIC LIBRARY

4.000) SOURCE OF DNA FOR GENOMIC LIBRARY

The human genome spans approximately 3.3 x 10^9 nucleotide base pairs and is organized into 23 discrete chromosomal units. The human X-chromosome represents 5.84% of the genetic material (Maynard-Smith *et al.*, 1961). Since we were attempting to isolate X-linked retinal sequences we decided to limit the genetic region to be manipulated to only that of the X-chromosome. This was achieved by the use of the human hamster hybrid somatic cell line 82082a as the DNA source for the genomic libraries made. 82082a consists of a single human X-chromosome in a hamster genetic background (Riddell *et al.*, 1986). The added advantage of this particular line is that it also lacks intact hamster X-chromosomes (H.S.Wang, Ottawa, Ont, personal communication). Although human chromosomes are lost at random in such somatic cell hybrids the human X chromosome can be selected for by growing the cells in HAT medium. In such a medium expression of the human hypoxanthine phosphoribosyltransferase, a gene locus present on the human X-chromosome is required, ensuring a continued selection for the human X-chromosome.

4.100) BACTERIOPHAGE LAMBDA BIOLOGY

The vector chosen for the present study was EMBL3 which is a lambda phage based replacement vector. In this vector enzymatic cleavage of the vector DNA generates three distinct regions, a left and right arm which represent the genetic information required for propagation of the phage and a middle fragment (stuffer

region), which can be replaced by non-lambda DNA. In general, recombinant phage molecules which are smaller than 78% or larger than 105% of the native phage genome size are not packaged into viable phage efficiently. Since the normal phage genome size is approximately 50 kb this places a restriction of 9-23 kb as the size of foreign DNA that can be inserted into the vector.

An added advantage to using the EMBL3 vector is its inherent high cloning efficiency without the requirements for a recombinant permissive host, which have been implicated in selecting against the cloning of certain types of sequences (Wyman *et al.*, 1986; Frischauf *et al.*, 1987). In the present study the bacterial cell line MB406 which is a recombinant non-permissive host was used in the propagation of the EMBL3 library.

4.200) INSERT ISOLATION

4.210) Theory

Generation of a representative genomic library relies on the ability to clone and maintain a total genomic set of random genomic DNA fragments. Generation of such cloning fragments is maximized by mechanical shearing of DNA. However, since the efficiency of cloning such fragments is not very high, most methods rely on the generation of quasi-random fragments by partial digestion of DNA with restriction enzymes that cleave the target DNA frequently. Under normal circumstances a digestion profile of HMW DNA is performed in order to determine

the conditions at which the maximum number of fragments are present at the fragment size of choice for cloning (9-23 kb). When this parameter is known a large scale digestion under those conditions is performed. The resulting fragment pool is size fractionated by either density gradient centrifugation or preparative gel electrophoresis to select those fragments whose size is suitable for cloning. Although this method is widely used to prepare genomic libraries it is questionable whether it gives the best representation of the genome. The basic assumption in the rationale of the restriction generated fragment profile is that the number of restriction sites for a frequent cutting 4-mer enzyme such as Mbo I is equally and randomly distributed in the genome, therefore giving an expected restriction site once every 256 base pairs.

The recognition site for Mbo I is GATC, and the frequency at a statistical level will rely on the GC content of a specific region of the genome. The GC content of genomic DNA is known to vary from region to region from as much as 73 to 22% (Normore *et al.*, 1976, Bernardi *et al.*, 1985; Ikemura *et al.*, 1990). The repetitive DNA in higher eukaryotes further biases restriction site representation (Tartoff, 1975). In addition endonuclease activity also varies greatly with the DNA substrate and activity for each site can be modified by neighbouring sequences (Brooks 1987). It is likely that certain regions of the genome will have more Mbo I sites than others and that the susceptibility of any specific site to endonuclease digestion will vary. If this is the case then each genomic region (and in fact each site)

will have its own set of individual optimal digestion conditions. Thus, the use of a single optimized digestion condition to generate genomic fragments for insertion is unlikely to produce a representative collection of genomic fragments.

In order to account for these variables, a modification in the approach of genomic insert isolation was made in the present study. Instead of a single set of optimized digestion conditions an entire digestion profile was performed on small amounts of HMW genomic DNA from the hamster/ human X hybrid cell line 82082a (Riddell *et al.*, 1986) and the entire profile was pooled. Fragments of 17-23 kb were then isolated from this mix by manual insert isolation following low melting agarose electrophoresis. In addition the insert size used was a minimum of 17 kb, to ensure that DNA fragments that are cloned represent single fragments, and not the cloning of two independent DNA fragments that have been ligated together and cloned.

It is difficult to show unequivocally whether the precautions taken in the making of the 82082a library are warranted. However the recent cloning of the 5' end of the human prolactin inducible protein gene (PIP) has suggested that the genomic cloning methodology defined in this thesis may enhance the cloning of sequences previously defined as unclonable sequences (Myal *et al.*, 1991). In a number of screening experiments with a partial PIP clone the 5' end of the gene was not found to be represented in a total of five independent human genomic libraries constructed using standard methodologies and using a recombination permissive host,

this included a commercially available EMBL3 genomic library made with genomic fragments generated by a partial Mbo I digestion. The screening of an amplified EMBL3 human genomic library made using the modified methodology (defined in the present thesis) identified a single phage clone encompassing the entire PIP gene.

The modifications used in the present methodology included the use of a recombination non-permissive host and the manner in which the inserts are isolated. A fundamental question therefore is whether the presence of the 5' end of the PIP gene in the library is due to the use of the recombination non-permissive host or due to the method of insert isolation employed. If the maintenance of the 5' end of PIP is due to the use of a recombination non-permissive host then placing the clone with the 5' end into a bacterial host that allows recombination should result in the eventual loss of the 5' end fragment. Studies in which the 5' end PIP clone has been propagated in a recombination permissive host for at least two generations has not resulted in the loss of the 5' fragment (Wong *et al.*, 1991). It would appear therefore that the insert isolation technique employed is the major factor in the successful cloning of sequences that might not otherwise be cloned.

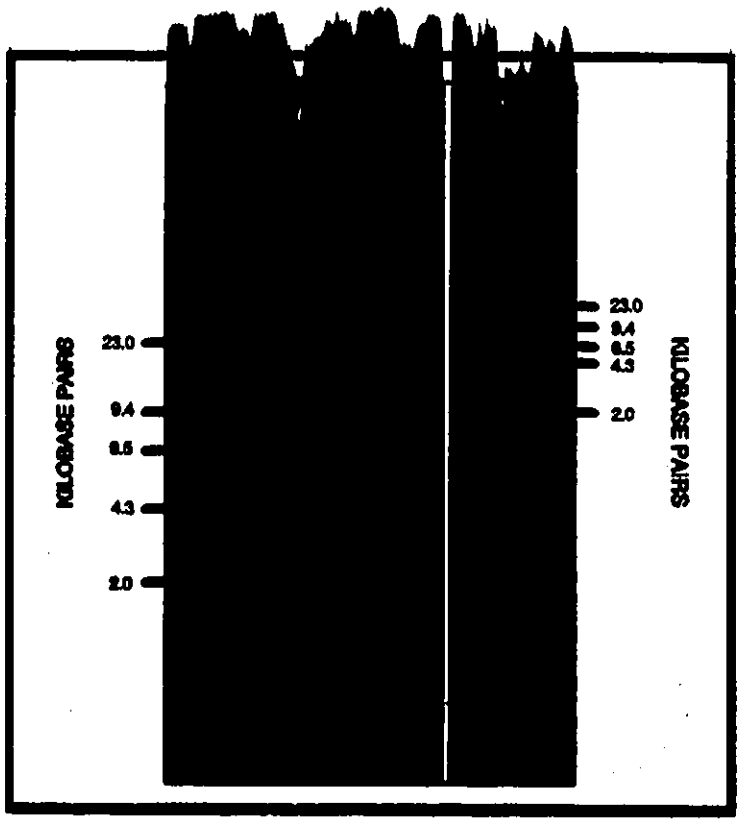
4.220) Isolation of DNA fragments from 82082a

In general 20-40 μg of HMW genomic DNA was digested by Mbo I as described in methods and analyzed by agarose gel electrophoresis (Fig. 15 A). The gel area corresponding to 17-23 kb fragments was excised from the unstained pooled

FIG. 15 ANALYSIS OF THE GENOMIC FRAGMENTS TO BE USED FOR CLONING.

Panel A: The DNA fragments generated by Mbo I digestion (FIG. 6) were electrophoresed in a 0.5% LMA gel using TAE as buffer. Electrophoresis was carried out at 4°C at 40 V for 16 to 18 h. Analysis of DNA from tubes 1, 8, and 9 of the digestion profile give an indication of the efficiency of the digestion over all. Tube 9 represents undigested DNA. Tube 1 should show a high degree of digestion as this is the tube with the highest concentration of restriction enzyme while tube 8, with the lowest concentration of enzyme, should show much less evidence of restriction. Shown are lanes 1-5 which consisted of control lanes which have been ethidium bromide stained.
Lane 1: 1 µg of intact lambda DNA
Lane 2: 10 µl of intact genomic DNA (tube 9)
Lane 3: 10 µl of DNA mix from tube 1
Lane 4: 10 µl of DNA mix from tube 8
Lane 5: 1 µg of Hind III digested lambda DNA
The remaining lanes (not shown) consisted of the pooled sample, which was loaded into the minimal number of wells that would hold its volume. These lanes which were to be used for the generation of cloning fragments were not at any time stained with ethidium bromide.

Panel B: DNA (350 ng) recovered for cloning has been electrophoresed in 0.9% agarose using TAE as buffer (lane 2). The insert migrates as a tight band, indicating very little degradation if any, in the size region from which it was initially isolated (between the 23 and 9 kbp fragments of the Hind III digested λDNA standards, lane 1 and 3).


KILOBASE PAIRS
23.0
9.4
6.5
4.3
2.0
23.0
9.4
6.5
4.3
2.0
KILOBASE PAIRS

DNA lanes and the DNA was extracted. On average 5-10% of the initial starting HMW DNA were recovered as clonable fragments. A second gel was run on the recovered cloning fragments to ensure the quality and size of the DNA (Fig. 15 B).

4.300) SYNTHESIS OF 82082a GENOMIC LIBRARY

For the construction of EMBL3 genomic libraries 0.35 μg of this DNA was ethanol precipitated and cloned according to protocols from the supplier (Promega Biotec, Madison, Wisconsin). Several libraries were constructed using these protocols and on average the titres of these libraries were 2.1x 10^6 pfu/μg insert (Table 3).

4.310) Average insert size

Because of the nature of the vector system all phage constructs within the library will contain an insert. Specific clones from the library were analyzed to determine the average insert size within the library (Fig. 16). In all cases inserts were liberated from the vector by a prolonged Sal I digestion. A number of clones appear to consist of at least two different independent phage recombinants, the majority of clones, however, contain a single insert, with an average insert size of 19 kb (Table 4).

4.320) 82082a library consists of hamster and human clones

82082a consists of a single human X chromosome in a tetraploid hamster genetic background. The hamster genetic component of 82028a is estimated to be

TABLE 3 TITRATION OF CONSTRUCTED 82082a GENOMIC LIBRARIES.

In total 3 genomic libraries derived from 82082a genomic DNA were constructed in the λ phage based vector EMBL3. The bacterial cell line MB406 was used as the host for infection. The average titres of these libraries is $2.1x10^6$ pfu/μg of genomic insert or $3.7x10^5$ pfu/μg of vector. The combined libraries consists of approximately $1.1x10^6$ independent clones.

The controls for the cloning experiments included:

1. The ligation and packaging of the vector arms alone (EMBL3) to determine the inherent background of the system.
2. The ligation and packaging of a 16 kbp genomic control insert to test the ligation efficiency and packaging efficiency of the cloning system.
3. The packaging of intact phage DNA (cI857 Sam 7 DNA) in order to test the packaging efficiency of the system.

DNA SOURCE	EMBL3 ARMS	CONTROL INSERT	INTACT PHAGE	82082a GENOMIC	82082a GENOMIC	82082a GENOMIC
INSERT DNA(μg)	-----	0.25	0.5	0.35	0.35	0.35
EMBL3 ARMS(μg)	0.25	0.25	-----	2.00	2.00	2.00
BACTERIAL HOST	Mb406	Mb406	LE392	Mb406	Mb406	Mb406
pfu/ml	0	2.9×10^6	1.2×10^8	8.5×10^5	8.8×10^5	5.0×10^5
pfu/μg INSERT	-----	1.1×10^7	-----	2.4×10^6	2.5×10^6	1.4×10^6
pfu/μg VECTOR	0	1.1×10^7	2.4×10^8	4.2×10^5	4.4×10^5	2.5×10^5

FIG. 16 ELECTROPHORETIC ANALYSIS OF INSERT SIZES FROM INDEPENDENT CLONES ISOLATED FROM THE 82082a GENOMIC LIBRARY.

An aliquot of the 82082a genomic library was used to infect MB406 bacterial cells and low density plating was performed. A number of plaque isolates were taken. DNA was extracted from each of the isolates. Inserts from thirteen 82082a genomic clone isolates were released from the vector by prolonged Sal I restriction digestions. Electrophoresis of the digested products was carried out in a 0.8% agarose gel using a TAE buffer system. Bands were visualized under UV transillumination after ethidium bromide staining. Bands migrating consistently throughout the lanes at approximately 20 and 9 kbp represent the left and right EMBL3 vector arms respectively. The sizes of inserts observed ranged from 20 kbp (lanes 9, 10, 12, 13) to 8.6 kbp (lane 7, the second band). In a number of instances (lanes 3, 6, 7, 10,12, and 13) a minimum of two distinct recombinants were evident. An analysis of individual insert sizes is presented in Table 4. In all cases 1 µg of digested DNA has been loaded per lane.

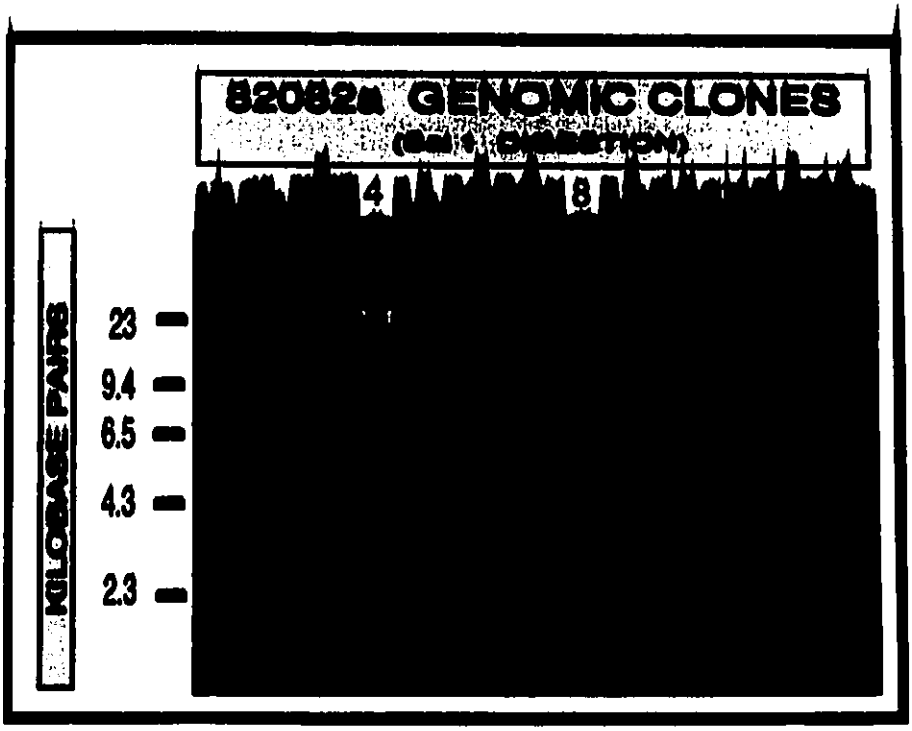

82082a GENOMIC CLONES
4
8
KILOBASE PAIRS
23
9.4
6.5
4.3
2.3

TABLE 4 MOLECULAR SIZING OF CLONE ISOLATES TAKEN FROM THE 82082a GENOMIC LIBRARY.

Genomic inserts from each isolate were liberated by a prolonged Sal I digestion. The molecular size was determined from the electrophoretic profile of the digested DNA (Fig. 16). A summary of the determined insert sizes is presented in this table. In a number of cases specific isolates appear to contain two independent insert bands. Because the EMBL3 vector is limited to accepting genomic fragments ranging in size from 9-23 kbp in length, a number of the isolates with two inserts consist of two distinct recombinants while others are more likely to consist of a single recombinant sequence with an internal Sal I site. In total 19 distinct recombinants, ranging in insert sizes from 8.6 to 20 kbp, were observed. The average insert size observed was 17.9 kbp with a standard deviation of 2.6 kbp.

LANE IN FIG. 16	BAND SIZES (Kbp)	# CLONES REPRESENTED	APPARENT CLONE SIZE (Kbp)
1	12.5 6.0	1	CLONE 1A: 18.5
2	17.4	1	CLONE 2A: 17.4
3	19.1 15.8	2	CLONE 3A: 19.1 CLONE 3B: 15.8
4	17.4	1	CLONE 4A: 17.4
5	17.4	1	CLONE 5A: 17.4
6	19.1 17.4	2	CLONE 6A: 19.1 CLONE 6B: 17.4
7	19.1 8.6	1	CLONE 7A: 19.1 CLONE 7B: 8.6
8	17.4	1	CLONE 8A: 17.4
9	20.0	1	CLONE 9A: 20.0
10	20.0 15.8	2	CLONE 10A: 20.0 CLONE 10B: 15.8
11	15.8 2.5	1	CLONE 11A: 18.3
12	20.0 19.1	2	CLONE 12A: 20.0 CLONE 12B: 19.1
13	20.0 19.1	2	CLONE 13A: 20.0 CLONE 13B: 19.1

9 times (900%) that of the human genome, while the human X chromosome represents 5.84% of the human genome. Thus the ratio of human DNA to hamster DNA is roughly 1:154. If all clones are of the same size then one should find one human genomic clone for every 154 82082a clones screened (0.66 %).

Human and hamster clones are differentiated from each other by screening the library with radiolabeled total human genomic DNA to identify human clones and total hamster genomic DNA to identify hamster specific sequences (Fig. 17). The detection system is based on the presence of species specific repetitive DNA in the clones (Gussella *et al.*, 1980; Korenberg *et al.*, 1987). In total 8000 clones from the 82082a library were screened with radioactively labeled total human genomic DNA. Fifty one clones were found to hybridize with the human probe, the remaining clones that did not hybridize with the human probe were assumed to be of hamster origin. The expected frequency of human clones in the 82028a library is 0.66%. The observed number of human hybridizing clones is 51 in 8000 screened. Application of a goodness of fit test, with 95% confidence limits (Table 5), reveals that the observed frequency of human clones is in agreement with the expected frequency. It would appear therefore that the human and hamster fragments in the cloned library are representative of the genomic DNA pool with no species preference in the cloning procedure.

FIG. 17 SCREENING OF THE 82028a GENOMIC LIBRARY FOR HUMAN SEQUENCES.

Panel A: A plaque lift of approximately 1000 genomic clones was screened with nick translated human genomic DNA (specific activity; 10^8 cpm/μg). Labeling, hybridization and wash conditions were as defined in material and methods.

Panel B: A duplicate plaque lift was subsequently screened with nick translated hamster genomic DNA (specific activity: 10^8 cpm/μg). Areas which are circled represent plaques which show hybridization with both the human genomic and hamster genomic probes. Areas which are circled and indicated by an arrow represent clones which hybridize with the human probe only.

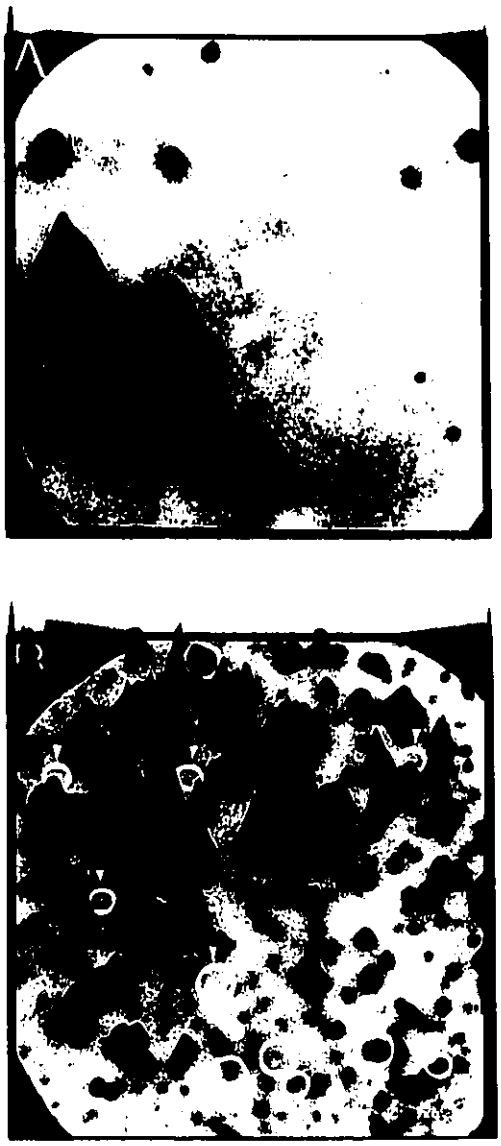


TABLE 5 SUMMARY OF A GOODNESS OF FIT TEST BETWEEN THE OBSERVED DISTRIBUTION OF HUMAN CLONES AND THE EXPECTED DISTRIBUTION OF HUMAN CLONES IN THE 82082a HYBRID GENOMIC LIBRARY.

HUMAN SPECIFICTIY	OBSERVED # CLONES	EXPECTED # CLONES (THEORETICAL)	X^2_o
HU+	51	52.8	0.06
HU-	7949	7947.2	0.00
		TOTAL X^2_o	0.06
		$X^2_{0.05,\ df:1}$	3.84

Although the use of total human genomic DNA can select for human specific clones it can also in some instances cross hybridize with hamster sequences as well (and vice versa). In order to examine the extent of this cross hybridization the 51 human positive clones were subsequently screened with a hamster genomic probe, three groups of clones were found. Twenty six of the clones showed strong hybridization with the hamster probe, 8 showed a much lower degree of hybridization and 17 did not hybridize. The eight clones that showed weak hybridization with the hamster probe all showed strong hybridization with the human probe and were assumed to be human clones. Thus, in total 25 clones could be confidently categorized as human specific although all clones presented with intense hybridization with the total human genomic probe.

CHAPTER FIVE

PCR AMPLIFICATION

5.000) INTRODUCTION

To perform an efficient cross screening between the chorio-retinal cDNA library and the EMBL3 library it is essential to isolate the inserted cDNA free from any contamination with λ DNA. While the isolation of insert DNA from single clones is a simple procedure, the isolation of DNA insert from a heterogeneous population of clones is a more complex endeavour. In the present study the isolation of insert DNA from an entire chorio-retinal λ cDNA library to screen a λ phage genomic EMBL3 library carries with it specific risks and difficulties.

There are several major difficulties; foremost is the isolation of the insert without phage DNA contamination. The initial attempts to screen the genomic library with conventionally isolated inserts (glass beads, electroelution, DEAE paper) from the chorio-retinal cDNA library were not successful. The primary reason for this is a λ sequence contamination of the insert preparation. This arises presumably from random shearing of vector sequences during manipulations. In the case of a given cDNA gt10 clone the vector sequence represents approximately 90% of the total sequence. The screening of the genomic library with these particular pools of cDNA inserts resulted in the hybridization to every genomic clone present in the screen (Fig. 18), a phenomenon which we named the Christmas Tree effect.

The successful isolation of insert DNA from a heterogeneous pool of cDNA clones was made possible by the use of the polymerase chain reaction (PCR) (Wong

FIG. 18 THE CHRISTMAS TREE EFFECT.

Duplicate plaque lifts of clones from the 82028a genomic library were screened with manually isolated insert generated from the chorio-retinal cDNA library and with labeled λDNA. All plaques were found to hybridize to the λ probe as well as to the cDNA library derived probe. The conclusion from these screens was that manual isolation of insert from an entire cDNA library carried with it a high enough background of lambda vector sequence contamination to obscure any specific hybridization based on the cDNA insert alone.

Panel A: A plaque lift of the genomic library was screened with a multiprime labeled insert (specific activity; 10^8 cpm/μg) isolated from the chorio-retinal cDNA library. cDNA library insert was isolated using NA-45 DEAE membrane during electrophoresis.

Panel B: A duplicate plaque lift was screened with nick translated lambda DNA (specific activity; 10^8 cpm/μg).


A

et al., 1989). This technique allows the amplification of specific regions of DNA between two specific oligonucleotide primers which hybridize to opposing strands (Saiki *et al.*, 1985; Mullis and Faloona, 1987; Saiki *et al.*, 1988). The technique has had an instant application in the analysis of single DNA fragments with respect to gene structure, nucleotide sequencing and restriction fragment length polymorphism studies, in each case focusing on a single discrete DNA sequence (Saiki *et al.*, 1985; Scharf *et al.*, 1986; Wong *et al.*, 1987; Chelley *et al.*, 1988; Stoflet *et al.*, 1988; Sarkar and Sommer, 1988; Bugawan *et al.*, 1988). In the described screening procedure, PCR is used in a different manner and that is to simultaneously amplify a heterogeneous population of sequences cloned into a common cloning vector (in this case λgt10).

5.100) DNA SOURCE FOR AMPLIFICATION

DNA was isolated from the unamplified cDNA library using a plate lysate technique. In total $5.5x10^4$ clones (an entire library) were plated out at high density and used to generate a lysate in a total of 40 mls of SM. The titre of the lysate after amplification was 9×10^{11} pfu/ml.

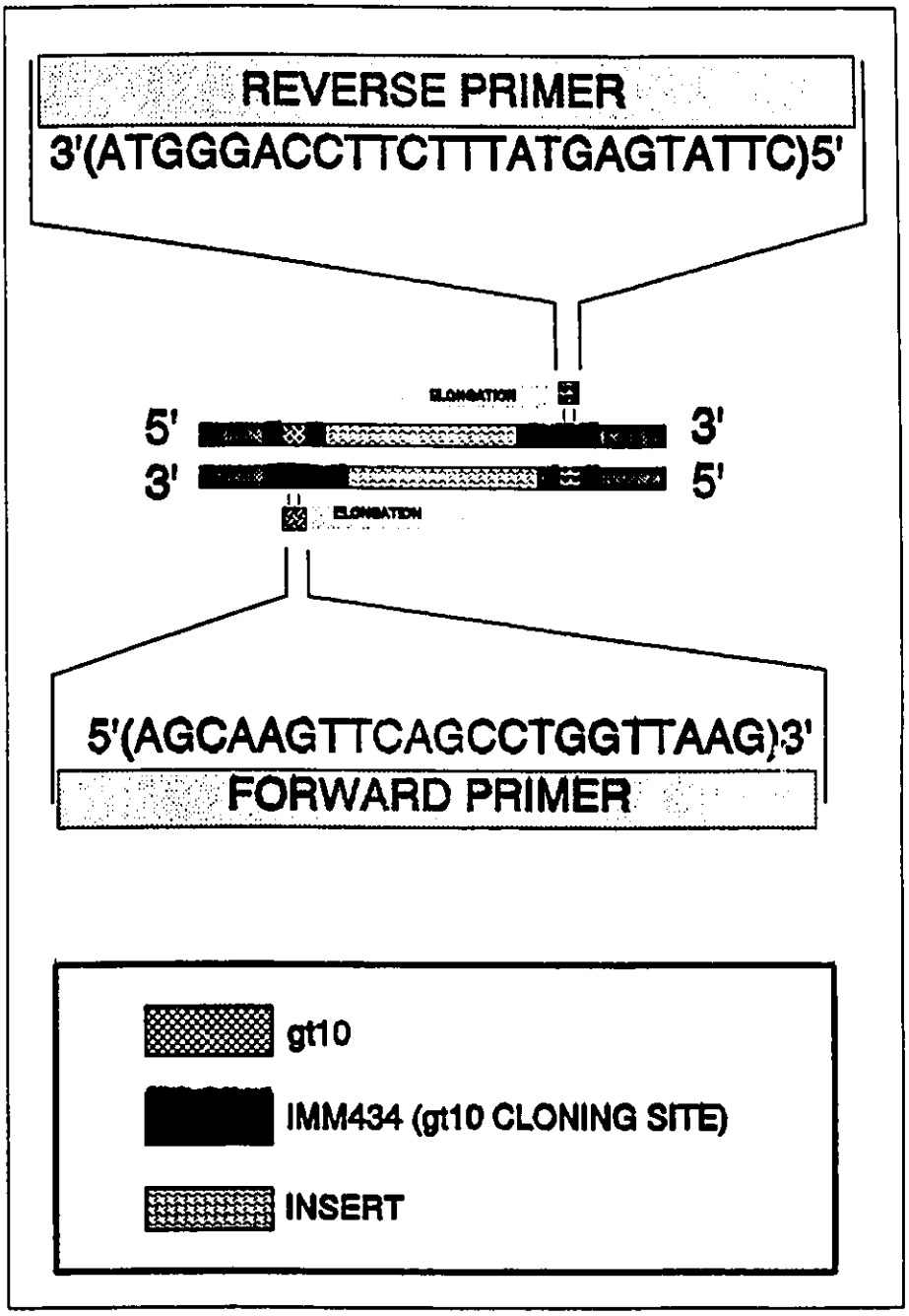
For DNA extraction 30 mls of this lysate was used. The yield from these extractions was approximately 500 μg of DNA. The DNA was stored in a final concentration of 100 ng/μl.

5.200) PRIMERS

The primers used for the reaction were the forward and reverse sequencing primers for the λgt10 vector (Fig. 19). In this way the primers selected were specific for the vector and not the human cDNA insert. Amplification of insert therefore is independent of the particular insert, although the size of the insert may affect the efficiency of amplification. This variable can be limited by extending the elongation time to 5 min, thus allowing the complete amplification of clones with large inserts.

The cross screening strategy described is made possible by the nature of the vectors used to generate the cDNA and genomic libraries respectively. Although both λgt10 and EMBL3 are λ based vectors and therefore have regions that will cross hybridize, the cloning site in each vector is different. In the case of λgt10 the cloning site is an Eco RI site in the IMM434 region of the λgt10 genome, a region which is totally absent from the EMBL3 DNA. Thus the use of λgt10 sequencing primers slightly upstream and downstream of the insert site in the PCR reaction allows the amplification of insert DNA and small portions of the λgt10 IMM434 region that will not hybridize with EMBL3 sequences. The polymerase chain reaction can therefore be used to specifically amplify and radioactively label the human retina cDNA inserts of a λgt10 phage cDNA library. The labeled DNA, which is essentially free of labeled λ arms, can be used to screen the EMBL3 genomic library made with DNA isolated from a human/hamster somatic cell hybrid cell line whose only human component is a single X chromosome. Human clones which display positive

FIG. 19 PRIMER SEQUENCES USED FOR λgt10 LIBRARY PCR AMPLIFICATION.


REVERSE PRIMER
3'(ATGGACCTTCTTTATGAGTATTC)5'
ELONGATION
5'
3'
3'
5'
ELONGATION
5'(AGCAAGTTCAGCCTGGTTAAG)3'
FORWARD PRIMER
gt10
IMM434 (gt10 CLONING SITE)
INSERT

hybridization contain, by definition, sequences from the human X chromosome which are expressed in the human retina (which may consist of both unique as well as repetitive expressed sequences).

5.300) THEORETICAL CONSIDERATIONS

In theory the polymerase chain reaction allows the doubling of a given amount of DNA template in each consecutive cycle. Thus if a single DNA sequence were amplified through 25 PCR cycles a total of 3.3×10^7 copies of this sequence would be synthesized. However because of the physical limitations of the reaction, the amplification factor is usually of the order of 10^5 (Perkin Elmer Cetus, Ottawa, Ontario). In order to optimize the reaction both the primer concentrations and the λgt10 library DNA concentrations must be taken into account. In one nanogram of DNA isolated from the λgt10 cDNA library there are approximately 10^7 individual DNA molecules (given that the average insert size is approximately 500 bp, that the size of the vector is 43.34 kbp, and that the average weight per molecule of recombinant is 2.4×10^{-20}g). If each of the inserts within these molecules is amplified 10^5 times then at least 10^{12} molecules of each primer are required for the reaction. The stock concentration of each primer used is 20 μM, or approximately 10^{13} primer molecules per μl. In the amplification procedure performed a total of 10^{14} molecules of each primer were added to the reaction mix, representing a 100 times excess. Since nucleotides for DNA synthesis are added in excess and the level of Taq polymerase was boosted after the first 10 cycles of amplification, these parameters

should not be limiting to the amplification process. Therefore the final amplification product should be representative of the cDNA library amplified.

5.400) PCR PRODUCT YIELDS

From the amount of incorporated radioactivity in the PCR product it was possible to estimate the total yield of insert DNA. On average 2-3 μg of amplified human λgt10 insert DNA was obtained per nanogram of intact λgt10 recombinant DNA. A sample of the PCR product, shown in Fig. 20 demonstrates that the PCR amplified DNA represents a heterogeneous pool of DNA fragments ranging from 200 to 1500 bp in size. Given that the λgt10 genome is 43.34 Kb in size, that the average insert size is 500 base pairs, and that the input DNA was 1 ng, the production of 2-3 μg of amplified DNA represents an amplification factor of approximately 10^5 times which is the maximum level of efficiency expected for the amplification system.

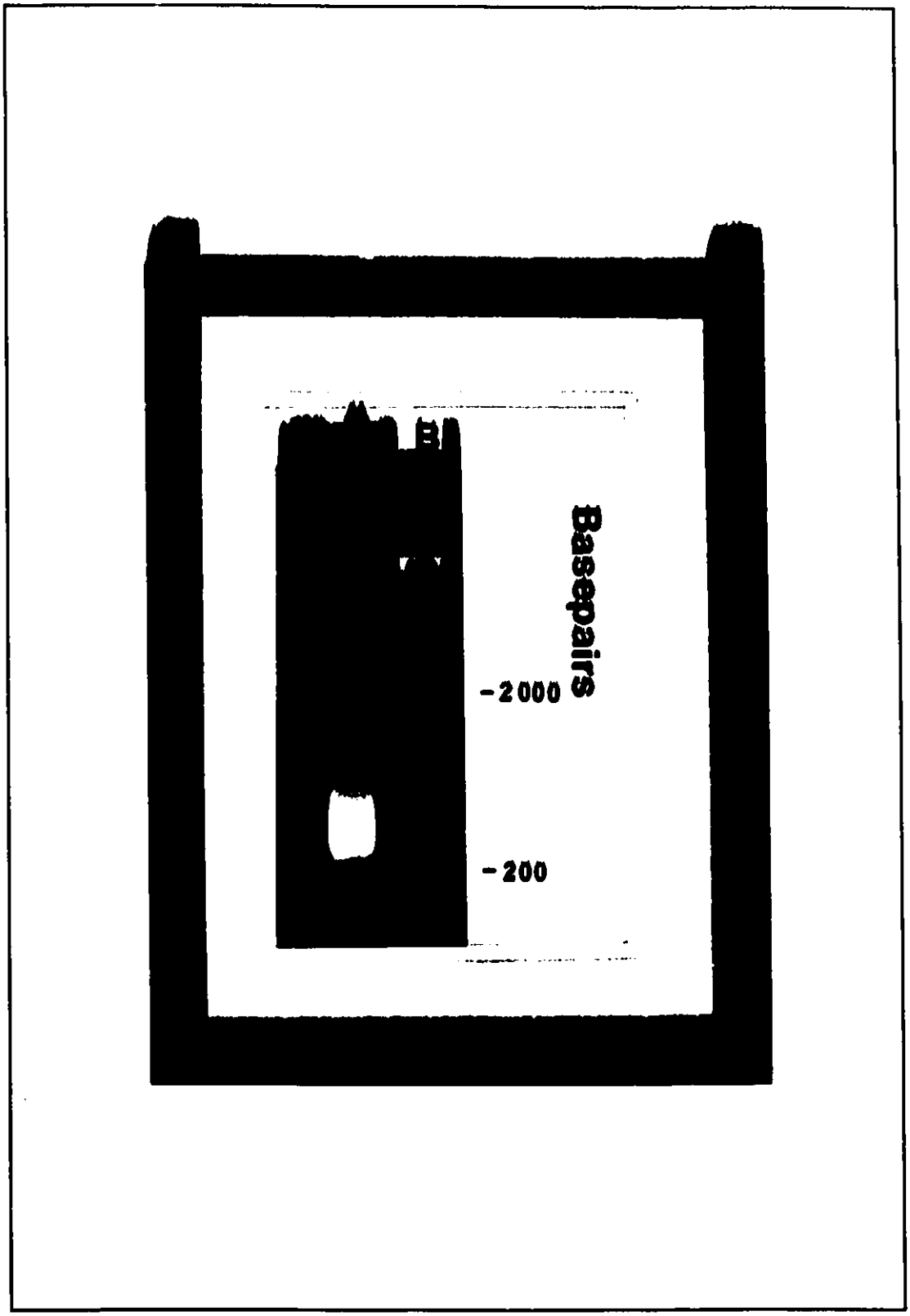
5.500) EXPECTED LEVELS OF DETECTION

The detection of a specific sequence by *in situ* screening will depend on the frequency of the sequence in the clone bank being screened and the concentration of the specific probe used. In the present case the library to be screened is a hybrid genomic library. Given that the library represents a random distribution of the respective genome then the expected frequency of any specific genomic clone is the same. The successful identification of that clone will therefore depend on the concentration of its corresponding probe. In the present case this is not easy to

FIG. 20 ELECTROPHORESIS OF THE PRODUCTS OF PCR AMPLIFICATION OF cDNA INSERTS FROM A HUMAN CHORIO-RETINAL cDNA λGT10 LIBRARY (P).

PCR amplification of 1 ng of DNA derived from the cDNA library was performed as outlined in materials and methods. Electrophoresis of the PCR product was carried out in 1 % agarose at 90V for 2 h. The range of sizes for the products of PCR span from 200 bp to 1000 bp and is a reflection of the general cDNA insert size of the respective library.

Lane A: 500 ng of PCR product
Lane B: 1 μg Hind III/Eco RI digested λ DNA. The sizes of standard marker bands (in bp) are indicated on the right margin.


Basepairs
2000
200

anticipate since the proposed cross screen uses a heterogeneous pool of sequences that should mirror the population profile of the chorio-retinal mRNA.

It is possible however to estimate the relative abundance of an mRNA sequence required for detection when used as a general probe. Given that the cDNA library mirrors the mRNA population profile, the average radioactivity of 100 ng of labeled PCR cDNA library insert is 1-2 X 10^8 cpm. Assuming that the average insert size is 500 base pairs, the radioactivity per clone in an unamplified library is (1.0 X 10^8 cpm/ 5.5 X 10^4), or approximately 2000 cpm/original clone. In the experiments described in this thesis the hybridization volume is typically 10 mls, giving a probe concentration for each original cloned insert of 200 cpm/ml. The minimum level of radioactivity for detection of any sequence is approximately 1000 cpm/ml (M. Brennan, Bethesda, Maryland, personal communication). Thus at a theoretical level any specific sequence that is present at least 5 times in the original pool will be detectable. At a crude level, therefore, clones with an original frequency of 5 in 5.5 X 10^4 or 0.01% should be detectable using the present methodology. This frequency is consistent with sequences falling within the lower end of middle abundant messages.

This number represents a very crude estimate. Factors which will undoubtedly affect it will include the probability that a given sequence is not represented in the pool of clones sampled for the probe. This will have a more drastic effect on those

sequences which are present in lower abundance in the normal mRNA. In fact, given the sample size used for the probe (5.5×10^4) the probability of finding any sequence is only approximately 80%. Another factor that will affect the level of detection of sequences will be the length of the specific insert from which the probe is made as this will alter the amount of radioactivity incorporated into each sequence. Estimations made above have been made assuming an average length of 500 base pairs, however it is known that the probe as a whole consists of a heterogeneous pool lengths. Taking into these considerations it may be more realistic to expect the method to be capable of detecting middle abundant messages, or approximately an mRNA which is represented 10-20 times in the original pool (0.02-0.04%).

In general a mixed cDNA probe should be sufficiently sensitive to detect message representing the lower levels of middle abundant sequences (20-40 copies per cell). Thirty per cent of all individual mRNA species within a typical cell, however, are of rare abundance and would not therefore be expected to be detectable with a mixed cDNA based probe (Williams, 1981). However a basic assumption is that genes which are more highly expressed are likely to have a more important role in the homeostasis of a given tissue. Thus abnormalities in these genes will more likely give rise to a disease state (McInnes *et al.*, 1988).

CHAPTER SIX

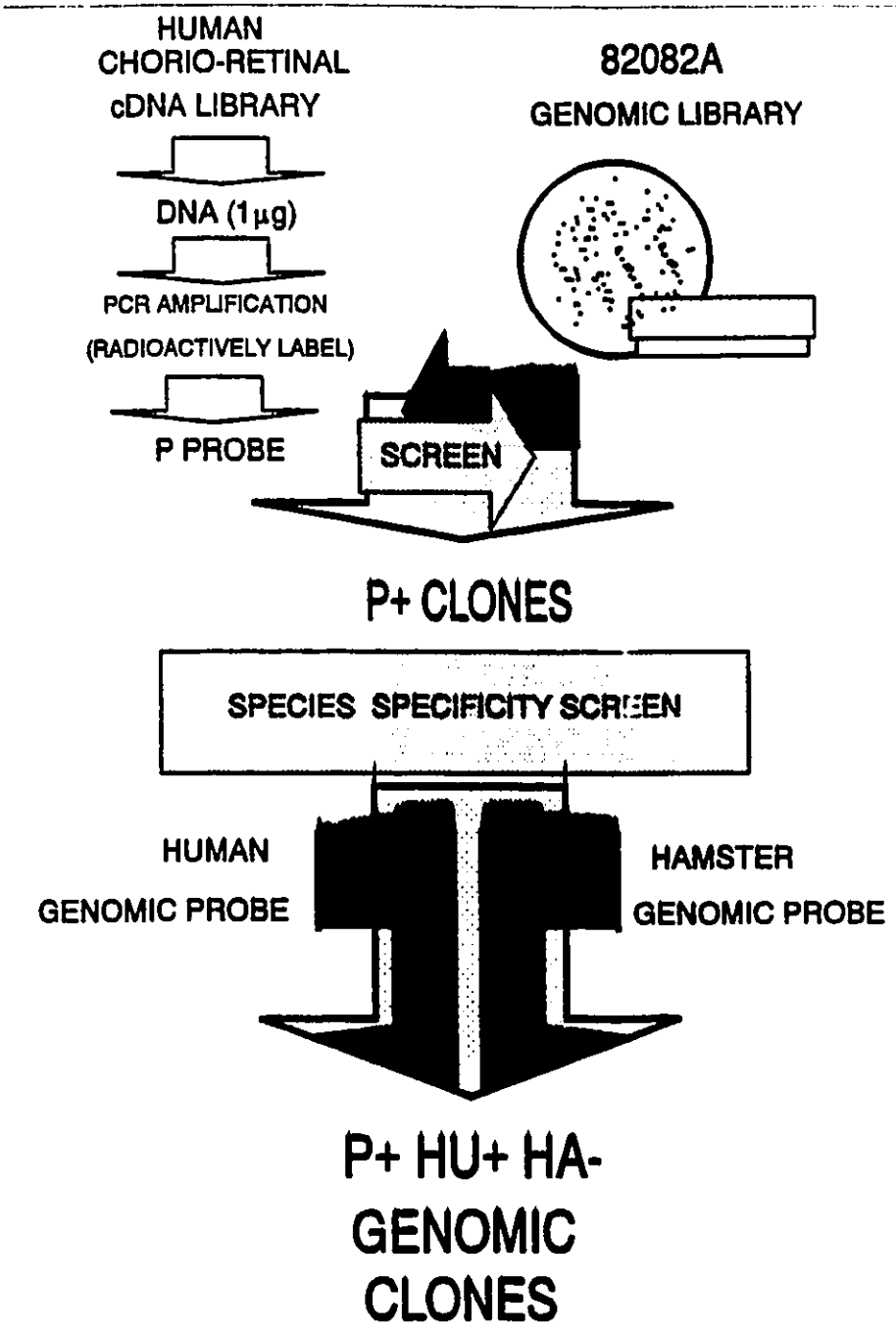
CROSS SCREENING STRATEGY

6.000) INTRODUCTION

The direct application of the proposed screening strategy is to isolate human X specific genomic clones that harbor sequences which are expressed in choroid/retina tissue. The isolation of a specific genomic sequence with its cognate cDNA is very common. The isolation of genomic sequences using a heterogeneous pool of cDNA or labeled mRNA has not been performed very often in the past. Because chorio-retinal mRNA is available in limited quantities screening with cDNA or labeled mRNA was not a feasible method in the present study. The alternative was to use one λ based library to screen a second λ based library (Wong et al., 1989).

The proposed cross screen used one library, a cDNA library (made in λgt10) to screen a genomic library (made in the λ based vector EMBL3). The probe representing the cDNA library (P probe) was generated through use of the polymerase chain reaction and was used to screen the genomic library (Fig. 21). Plaques which hybridized to the P probe (P+ clones) were further screened to determine the species specificity (human versus hamster), and to identify those which were human specific. The cDNA library used as the source of the probe was the chorio-retinal cDNA library described in chapter three and the genomic library screened was the 82082a somatic cell hybrid genomic library described in chapter four.

FIG. 21 SUMMARY OF THE PROPOSED LATERAL SCREENING STRATEGY.


HUMAN
CHORIO-RETINAL
cDNA LIBRARY
DNA (1µg)
PCR AMPLIFICATION
(RADIOACTIVELY LABEL)
P PROBE
82082A
GENOMIC LIBRARY
SCREEN
P+ CLONES
SPECIES SPECIFICITY SCREEN
HUMAN
GENOMIC PROBE
HAMSTER
GENOMIC PROBE
P+ HU+ HA-
GENOMIC
CLONES

6.100) HYBRIDIZATION OF THE P probe

The PCR product probe generated from the chorio-retinal cDNA library was used directly to screen the 82082a genomic library (Fig. 22 A). For comparison purposes a plaque lift of the 82028a genomic library has also been probed with λgt10 sequences (Fig. 22 B).

When plaque lifts of the 82082a genomic library were probed with the P probe only a few of the 82082a clones showed evidence of hybridization (Fig. 22 A). When the same library was screened with λgt10 DNA all plaques showed evidence of hybridization due to the presence of cross hybridization between common λ based sequences (Fig. 22 B). A comparison of the two profiles thus revealed that the P probe hybridizes through the human cDNA insert and not through the sequencing primers used in the PCR reaction.

6.200) CROSS SCREENING

6.210) Isolation of PCR+ / Human specific clones

Positive clones resulting from the initial screens were picked and used to establish individual phage stocks. Because of a varied hybridization signal from clone to clone (an expected observation if clones are being identified on the basis of different expressed sequences) each of the initial clones were screened a second time with the P probe by phage dot analysis in order to confirm their status. In total 182 confirmed P+ clones have been isolated from the screening of 19,000 genomic

FIG. 22 SCREENING OF THE 82082a GENOMIC LIBRARY WITH RADIOLABELED PCR-AMPLIFIED PRODUCT (P probe).

Panel A: Plaque lift of 1000 clones screened with P probe (100 ng, specific activity of 10^7 cpm/μg). The labeling of the P probe was performed by the addition of $\alpha[^{32}$-P]dATP and $\alpha[^{32}$-P]dCTP during PCR amplification of the chorio-retinal cDNA library.

Panel B: Plaque lift of 400 clones screened with ^{32}P-labeled λgt10 DNA (200 ng, specific activity of 10^8 cpm/μg).

Plaque lifts of the 82082a genomic library were prehybridized, probed , washed, and autoradiographed as defined in materials and methods. Arrow heads indicate clones showing weak hybridization.


A
B

clones. Clones isolated in this way represent genomic clones that harbor sequences which are expressed in choroid/retina tissue as represented in the P probe.

Because of the nature of the 82082a genomic library, being derived from both a hamster and partial human genetic background, each clone isolated was examined for both human and hamster specificity (Fig. 23 A). A summary of this analysis is presented in Table 6 A. From this second analysis four distinct subgroups of clones were found within the 182 P+ clones (Fig. 24):

1) **Human specific clones (P+ HU+ HA-):** Clones that demonstrated hybridization with radiolabeled human genomic DNA but not with radiolabeled hamster genomic DNA, and clones in which there was hybridization to both species but the hybridization with human genomic DNA was much more intense than that seen with hamster genomic DNA.

2) **Hamster specific clones (P+ HU- HA+):** Clones that demonstrated hybridization with radiolabeled hamster genomic DNA but not with radiolabeled human genomic. DNA and clones in which there was hybridization to both species but the hybridization with hamster genomic DNA was much more intense than that seen with human genomic DNA.

3) **Bi-species specific clones (P+ HU+ HA+):** Clones which showed hybridization of equal intensity to both genomic probes.

FIG. 23 ALTERNATIVE STRATEGIES IN THE ISOLATION OF RETINAL/HUMAN SPECIFIC CLONES.

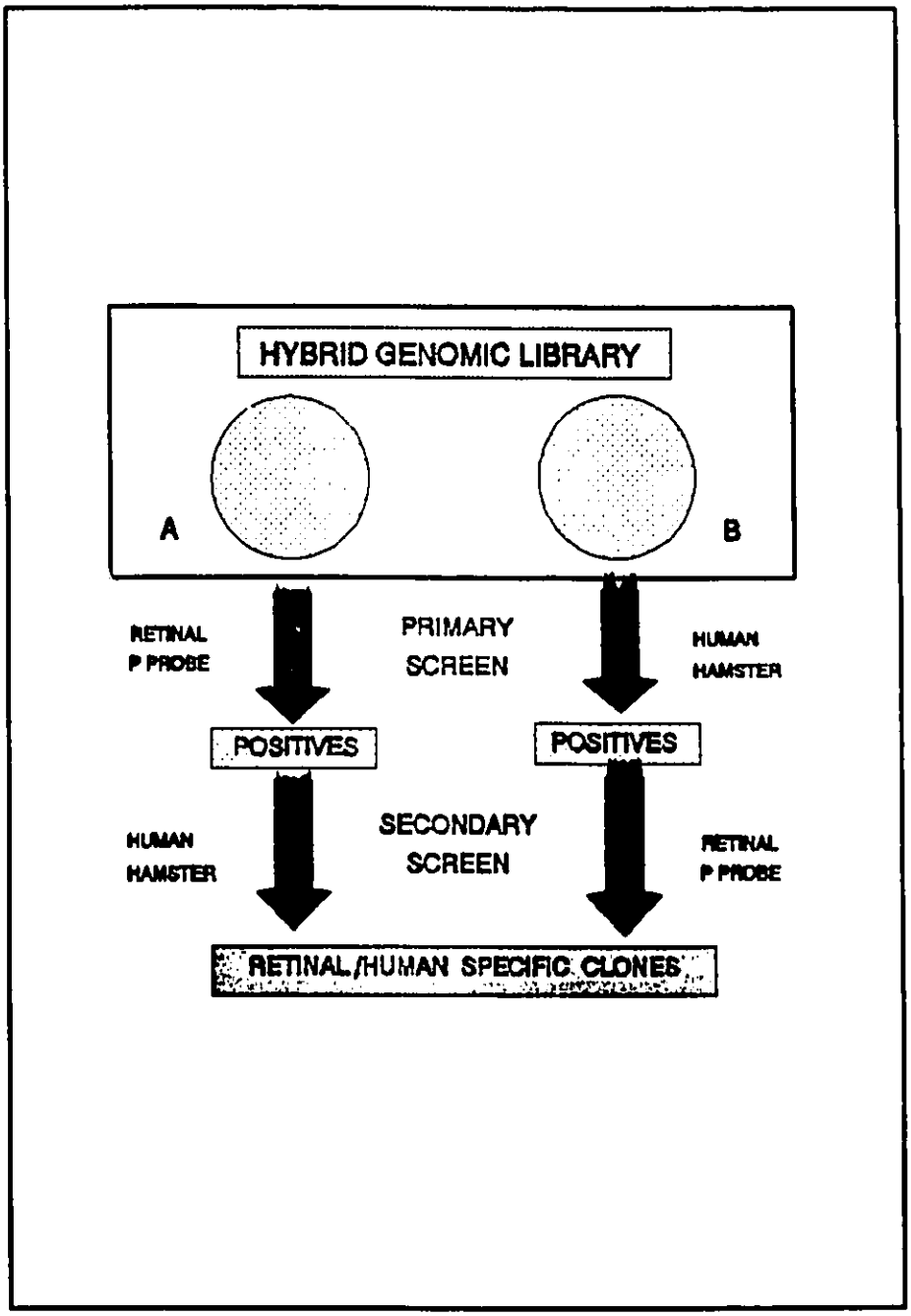

HYBRID GENOMIC LIBRARY
A
B
RETINAL
P PROBE
PRIMARY
SCREEN
HUMAN
HAMSTER
POSITIVES
POSITIVES
HUMAN
HAMSTER
SECONDARY
SCREEN
RETINAL
P PROBE
RETINAL/HUMAN SPECIFIC CLONES

TABLE 6 SUMMARY OF CROSS SCREENING RESULTS.

82082a genomic clones were screened on the basis of two criteria: with respect to species specificity (human, HU, and hamster, HA) and with respect to hybridization with the P probe. In order to determine if the order of screening for these two criteria was important, two separate aliquots of the library were screened in the alternative orders (Fig.23). Clones which gave a positive hybridization signal are indicated by (+) while clones that did not hybridize with the respective probe are indicated by a (-).

A) In total 19,000 genomic clones from the 82082a library were first screened to identify clones which hybridized with the P probe. Positive clones (P+) were subsequently screened for species specificity. In the tabulation of results P+ clones have been grouped with respect to a number of parameters including their relation with HA specificity alone, HU specificity alone, and subsequently with respect to both HA and HU specificity simultaneousiy.

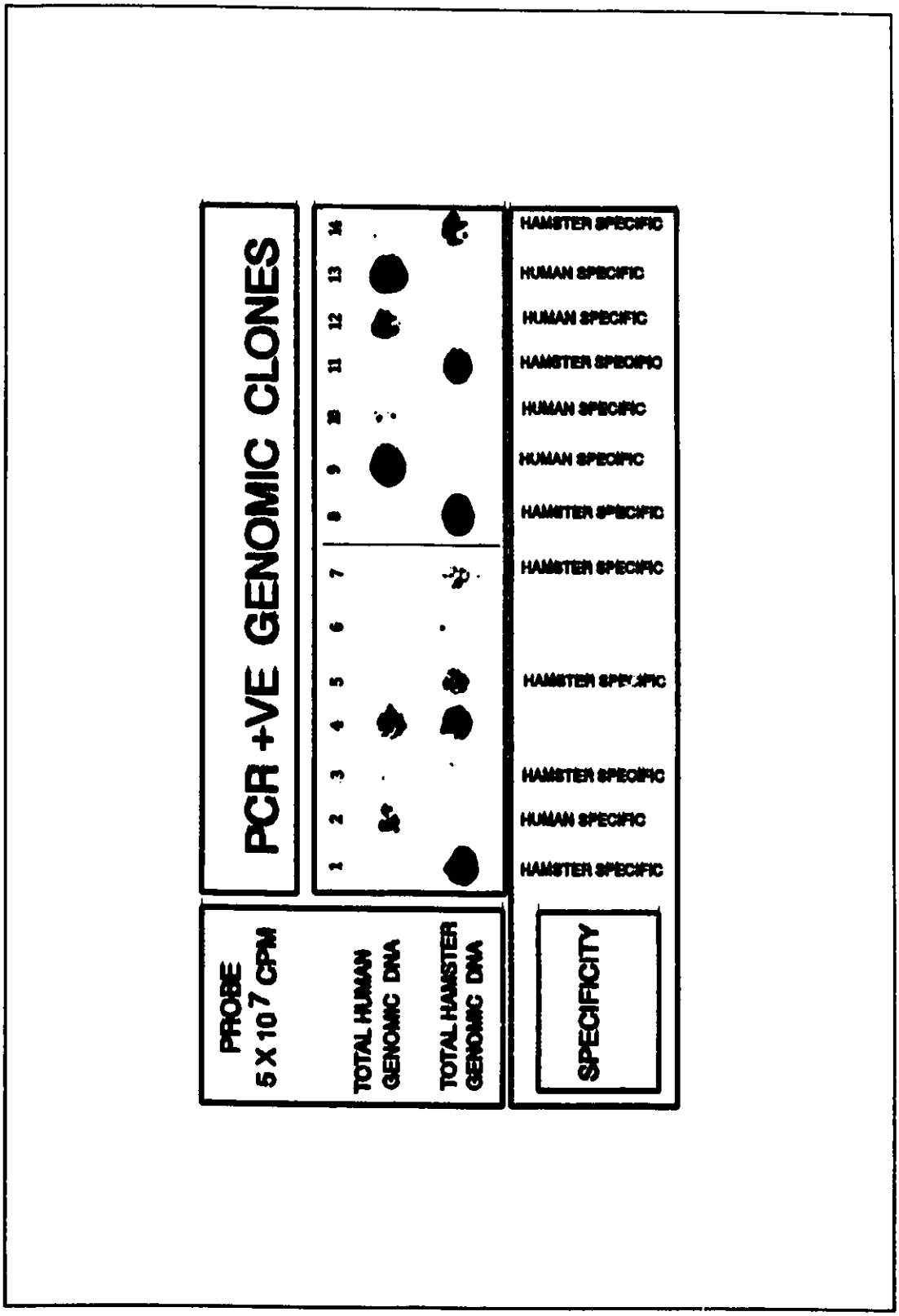
B) In total 8,000 genomic clones from the 82082a library were first screened for species specificity (HU and HA). Subsequently all HU+ clones were probed with the P probe. In the tabulation of results HU+ clones have been grouped with respect to a number of parameters including their relation with HA specificity alone, P specificity alone, and subsequently with respect to both HA and P specificity simultaneously.

A) PCR PRIMARY SCREEN (19000 82082a GENOMIC CLONES SCREENED)					B) HUMAN PRIMARY SCREEN (8000 82082a GENOMIC CLONES SCREENED)			
HYBRIDIZATION PATTERN	#CLONES	% OF PCR + CLONES	% OF TOTAL CLONES SCREENED		HYBRIDIZATION PATTERN	# CLONES	% OF HUMAN + CLONES	% OF TOTAL CLONES SCREENED
P +	182	100.0	0.96		HU+	51	100.0	0.64
P + HA- P + HA+	62 120	34.0 66.0	0.33 0.66		HU+ HA- HU+ HA+	25 26	49.0 51.0	0.31 0.33
P + HU- P + HU+	105 77	57.7 42.3	0.55 0.41		HU+ P - HU+ P +	7 44	13.7 86.3	0.09 0.55
P + HA- HU- P + HA- HU+ P + HA+ HU+ P + HA+ HU-	5 57 20 100	2.7 31.3 11.0 55.0	0.03 0.30 0.10 0.53		HU+ HA- P - HU+ HA- P + HU+ HA+ P + HU+ HA+ P -	4 21 23 3	7.8 41.2 45.1 5.9	0.05 0.26 0.29 0.04

FIG. 24 SCREENING OF P+ λ EMBL3 CLONES FOR HUMAN AND HAMSTER SPECIFICITIES.

Fourteen P+ clones were screened for species specificity by phage dot analysis. One set (A) of phage-dot lifts was probed with ^{32}P-labeled total human genomic DNA (specific activity of 10^8 cpm/μg). A duplicate set of lifts (B) was probed with ^{32}P-labeled total hamster genomic DNA (specific activity of 10^8 cpm/μg). Shown here is a representative sampling of the analysis. Genomic probes were labeled by nick translation as described in Materials and Methods. Prehybridization, hybridization, wash conditions and autoradiography were performed as defined in Materials and Methods.

From this analysis four subgroups of clones can be differentiated:
a) HU + / HA - clones: clones 2, 9, 10, 12, 13
b) HU + / HA + clones: clone 4
c) HU - / HA + clones: clones 1, 3, 5, 7, 8, 11, 14
d) HU - / HA - clones: clone 6


PCR +VE GENOMIC CLONES
PROBE
5 X 10 7 CPM
TOTAL HUMAN GENOMIC DNA
TOTAL HAMSTER GENOMIC DNA
1 2 3 4 5 6 7 8 9 10 11 12 13 14
SPECIFICITY
HAMSTER SPECIFIC
HUMAN SPECIFIC
HUMAN SPECIFIC
HAMSTER SPECIFIC
HUMAN SPECIFIC
HUMAN SPECIFIC
HAMSTER SPECIFIC
HAMSTER SPECIFIC
HAMSTER SPECIFIC
HAMSTER SPECIFIC
HUMAN SPECIFIC
HAMSTER SPECIFIC

4) **Species non-distinct clones (P+ HU- HA-):** Clones which did not show hybridization to either human or hamster genomic DNA.

6.220) **Isolation of Human/PCR+ specific clones**

In order to determine whether the nature of the primary screen affects the final identification of clones the reverse order of screening (a human genomic primary screen and a P probe secondary screen) was performed (Fig. 23 B). Human clones which were isolated in the initial characterization of the 82082a library (see previous results, chapter four) were secondarily screened with the P probe. In total 51 human positive clones which were isolated from the primary screening of 8000 (82082a) genomic clones were submitted to this secondary screen. A summary of this analysis is presented in Table 6 B. In total 4 subgroups of clones were identified:

1) **Human specific clones harboring chorio-retinal expressed sequences (HU+ HA- P+):** Clones that demonstrated hybridization with radiolabeled human genomic DNA and the P probe but did not demonstrated any hybridization with radiolabeled hamster genomic DNA.

2) **Bi-species specific clones harboring chorio-retinal expressed sequences (HU+ HA+ P+):** Clones that demonstrated hybridization with radiolabeled human genomic DNA, radiolabeled hamster genomic DNA, and the P probe. In most cases the hybridization with the human probe was stronger than with the hamster probe.

3) **Human specific clones not harboring chorio-retinal expressed sequences (HU+ HA- P-):** Clones that demonstrated hybridization to the total human genomic probe only.

4) **Bi-species specific clones not harboring chorio-retinal expressed sequences (HU+ HA+ P-):** Clones that demonstrated hybridization to both the total human and hamster genomic probes but no hybridization with the P probe.

6.300) NATURE OF HYBRIDIZATION

The hybridization of a labeled probe derived from total human genomic DNA is based on the presence of highly to middle abundant repetitive sequences. It is assumed that the hybridization of expressed sequences (mRNA) is based on the cognate cDNA that corresponds to low copy sequences in the genome and that the level of detection will depend on the relative abundance.

An important consideration that may be a compounding problem is the degree of repetitive sequences that are expressed in the RNA pool. If the amount of repetitive sequences in the derived probe is significant then there should be no difference between screening a genomic library with this probe or screening it with total genomic DNA. If this is the case, clones that show positive hybridization with the P probe should be positive with the genomic probe and conversely, clones which are negative with the P probe should be negative with the total genomic probe, resulting in a perfect correlation between positive clones isolated from a PCR screen and those obtained from a total human genomic screen.

Analysis of the data reveals that there is a significant disconcordancy between the hybridization patterns with the P probe and with the total genomic probe. Approximately 14% of the clones initially isolated using the total genomic probe were determined to be P-. Conversely 58% of clones initially isolated using the P probe were determined to be negative with the genomic probe. The conclusion from these observations is that the detection of genomic clones with the P probe differs from that of total genomic DNA. This does not exclude, however, the possibility that the two probes may share a common component as well as those in which they differ nor does it exclude the possibility that the presence of expressed repetitive sequences might obliterate the relevant signals when using the P probe.

6.400) A COMPARISON BETWEEN THE TWO CROSS SCREENS

Regardless of the screening order the proportion of P+/HU+/HA- clones should remain constant. Theoretically clones which are not potentially detectable with a specific probe will be negative regardless of whether the specificity is tested in the primary or the secondary screen. At a first glance this consistency exists with the frequency of P+/HU+/HA- clones at 0.30% (P probe primary screen) and 0.26% (Human primary screen) and it appears that there is little difference in the exact order of the cross screen. However a more careful analysis is warranted.

The genomic library screened is derived from a human/hamster hybrid cell line. Thus analysis of the two sets of positive clones for hamster orthologous

sequences may be used as an indication as to whether the two sets of clones represent the same sub-population. From an initial inspection of the data it appears that the clones isolated with the Human genomic probe as the primary screen conforms to a 1:1 ratio HA+/HA- while the clones isolated with the P probe as the primary screen adheres to a 1:2 ratio HA+/HA-. Goodness of fit tests of these ratios to the observed frequencies of HA+ and HA- clones for the alternative screening situation is presented in Table 7. The observed HA+/ HA- ratio (human primary screen) can be statistically excluded from the proposed 1:2 ratio. The observed HA+/HA- ratio (PCR primary screen) can be statistically excluded from the proposed 1:1 ratio.

A test for homogeneity (Table 8) shows clearly that the frequency of clones which hybridize with the hamster genomic probe is not independent of the nature of the primary screen. This observation however is not all together unexpected since the P probe is much less stringent than the human genomic probe in its permissive selection of hamster clones.

It is perhaps more important to determine if the HU+/P+ population resulting from the 2 screenings are similar. In a direct comparison of the clone subgroups, two of the four from each screening are directly comparable: the P+/HU+/HA+ and P+/HU+/HA- subgroups. A statistical analysis for goodness

TABLE 7 GOODNESS OF FIT TESTS BETWEEN THE OBSERVED PROPORTION OF HA+ / HA- CLONES OBSERVED IN THE SUBSET OF CLONES ISOLATED IN THE PCR CHORIO-RETINAL cDNA LIBRARY PRIMARY SCREEN VERSUS THE HUMAN SPECIFICITY PRIMARY SCREEN (AND VICE VERSA).

(A)

HAMSTER SPECIFICITY (among all P + clones)	PCR 1° SCREEN OBSERVED DISTRIBUTION	EXPECTED DISTRIBUTION BASED ON HUMAN 1° SCREEN (HA-/+) RATIOS	x^2
HA -	62	91.0	9.2
HA +	120	91.0	9.2
		TOTAL ESTIMATED X^2	18.4
		$X^2_{0.05, df:1}$	3.84

(B)

HAMSTER SPECIFICITY (among all HU+ clones)	HUMAN 1° SCREEN OBSERVED DISTRIBUTION	EXPECTED DISTRIBUTION BASED ON PCR 1° SCREEN (HA-/+) RATIOS	x^2
HA -	25	17.0	3.7
HA +	26	34.0	1.8
		TOTAL ESTIMATED X^2	5.5
		$X^2_{0.05, df:1}$	3.84

TABLE 8 ANALYSIS OF THE HA+ / HA- DISTRIBUTION WITHIN THE SUBPOPULATION OF CLONES ISOLATED FROM EACH OF THE TWO PRIMARY SCREENS.

Summary of a test for homogeneity analysis examining the distribution of hamster specificity among the P+ and HU+ genomic clones isolated as a result of the P and HU primary screens, respectively. A summary of the estimated Chi square (X^2) values are presented in the last column.

Total number of clones in the analysis	:	233
Total number of clones from the P 1° screen	:	182
Total number of clones from the HU 1° screen	:	51
Total number of HA- clones in the analysis	:	87
Total number of HA+ clones in the analysis	:	146

HA -/+ DISTRIBUTION WITHIN CLONES ISOLATED FROM THE P AND HU PRIMARY SCREENS.

CLONE STATUS	OBSERVED # OF CLONES	EXPECTED # OF CLONES	X^2_c
# OF HA- CLONES FROM THE (PCR) PRIMARY SCREEN	62	68.0	0.53
# OF HA+ CLONES FROM THE (PCR) PRIMARY SCREEN	120	114.0	0.32
# OF HA- CLONES FROM THE (HU) PRIMARY SCREEN	25	19.0	1.89
# OF HA+ CLONES FROM (HU) PRIMARY SCREEN	26	32.0	1.13
		TOTAL X^2_c $X^2_{0.05,\ df:1}$	3.87 3.84

of fit (Table 9) shows that the PCR primary screen distribution of P+ HU+ HA+ to P+ HU+ HA- clones does not fit with the distribution observed in the Human primary screen (and vice versa). A direct test for homogeneity (Table 10) with respect to these two subgroups of clones shows that hamster specificity is not independent of the nature of the primary screen. It is interesting to note that in the PCR primary screen the proportion of HA+ clones drops from 33% (when all P+ clones are considered) to 25% (when only the P+ HU+ clones are considered) whereas the distribution of HA+ clones isolated in the human primary screen is the same regardless of the grouping considered.

6.500) SUMMARY

The statistical analysis of the profile of isolated clones allows a small degree of insight into the validity of the screening procedure. What can be inferred from the analysis is that the PCR primary screen and the human genomic primary screen showed a subtle difference in the types of clones that were isolated.

TABLE 9 GOODNESS OF FIT TESTS BETWEEN THE OBSERVED PROPORTION OF HA+ / HA- CLONES OBSERVED IN THE SUBSET OF P+ / HU+ CLONES ISOLATED IN THE PCR CHORIO-RETINAL cDNA LIBRARY PRIMARY SCREEN VERSUS THE HUMAN SPECIFICITY PRIMARY SCREEN (AND VICE VERSA).

(A)

HAMSTER SPECIFICITY (among P+/HU+ clones)	PCR 1° SCREEN OBSERVED DISTRIBUTION	EXPECTED DISTRIBUTION BASED ON HUMAN 1° SCREEN (HA-/+) RATIOS	x^2
HA -	57	38.5	8.9
HA +	20	38.5	8.9
		TOTAL ESTIMATED X^2	17.8
		$X^2_{0.05,\ df:1}$	3.84

(B)

HAMSTER SPECIFICITY (among P+/HU+ clones)	HUMAN 1° SCREEN OBSERVED DISTRIBUTION	EXPECTED DISTRIBUTION BASED ON PCR 1° SCREEN (HA-/+) RATIOS	x^2
HA -	21	33.0	4.3
HA +	23	11.0	13.0
		TOTAL ESTIMATED X^2	17.3
		$X^2_{0.05,\ df:1}$	3.84

TABLE 10 ANALYSIS OF THE HA+ / HA- DISTRIBUTION WITHIN THE P+ / HU+ SUBPOPULATION OF ISOLATED CLONES.

Summary of a test for homogeneity analysis examining the distribution of hamster specificity among the P+ HU+ genomic clones isolated as a result of the P and HU primary screens. A summary of the estimated Chi square (X^2) values are presented in the last column.

Total number of clones in the analysis	:	121
Total number of clones from the P 1° screen	:	77
Total number of clones from the HU 1° screen	:	44
Total number of HA- clones in the analysis	:	78
Total number of HA+ clones in the analysis	:	43

HA -/+ DISTRIBUTION WITHIN THE P+ HU+ SUBPOPULATIONS OF ISOLATED CLONES.

CLONE STATUS	OBSERVED # OF CLONES	EXPECTED # OF CLONES	X^2_e
# OF HA- CLONES FROM THE (PCR) PRIMARY SCREEN	57	49.6	1.10
# OF HA+ CLONES FROM THE (PCR) PRIMARY SCREEN	20	27.4	1.99
# OF HA- CLONES FROM THE (HU) PRIMARY SCREEN	21	28.4	1.93
# OF HA+ CLONES FROM (HU) PRIMARY SCREEN	23	15.6	3.51
		TOTAL X^2_e $X^2_{0.05, df:1}$	8.53 3.84

CHAPTER SEVEN

AN INITIAL ANALYSIS OF THE CLONES

7.000) INTRODUCTION

The results of the lateral cross screen procedure yielded 78 P+ HU+ HA- genomic clones that are potentially X-specific sequences and harbor chorio-retinal expressed sequences. To establish the status of these clones a physical analysis of a subset from the pool of P+ HU+ HA- clones was undertaken.

7.100) ISOLATION OF INDEPENDENT CLONES

Six clones (S7, R6, R8, R20, R30, and X3) were selected at random from the pool of isolated P+ HU+ HA- specific clones for further characterization. DNA isolated from each clone was submitted to restriction analysis using several different restriction enzymes (Sal I/Eco RI; Eco RI; Bam HI; Hind III). The electrophoretic profile of each clone was found to be unique, suggesting that each clone represents an independent sequence from the human component of the 82082a genome (Fig. 25).

Southern analysis of these six profiles demonstrated that only a subset of bands within each clone hybridized with the P probe. In addition the intensity of hybridization varied from clone to clone (Fig. 26). Clone S7, which showed faint hybridization on phage dot analysis, did not show any hybridization on Southern analysis. It is not known if this lack of signal is due to limitations of detection (specificity of the probe) or whether S7 represents the inherent background in the screening strategy.

FIG. 25 RESTRICTION PROFILES OF 6 P+/HU+ GENOMIC CLONES.

Restriction digestion of clones were performed as defined in materials and methods. One μg of digested DNA was loaded per lane and electrophoresis was carried out in 1% agarose using a TAE system at 30V for 16 h. Bands were visualized after ethidium bromide staining and UV transillumination.

Lane M1 : Molecular size marker 1, Hind III digested lambda DNA.
Lane M2 : Molecular size marker 2, Eco RI digested lambda DNA.
Lane 1 : Sal I/Eco RI restriction profile.
Lane 2 : Eco RI restriction profile.
Lane 3 : Hind III restriction profile.
Lane 4 : Bam HI restriction profile.

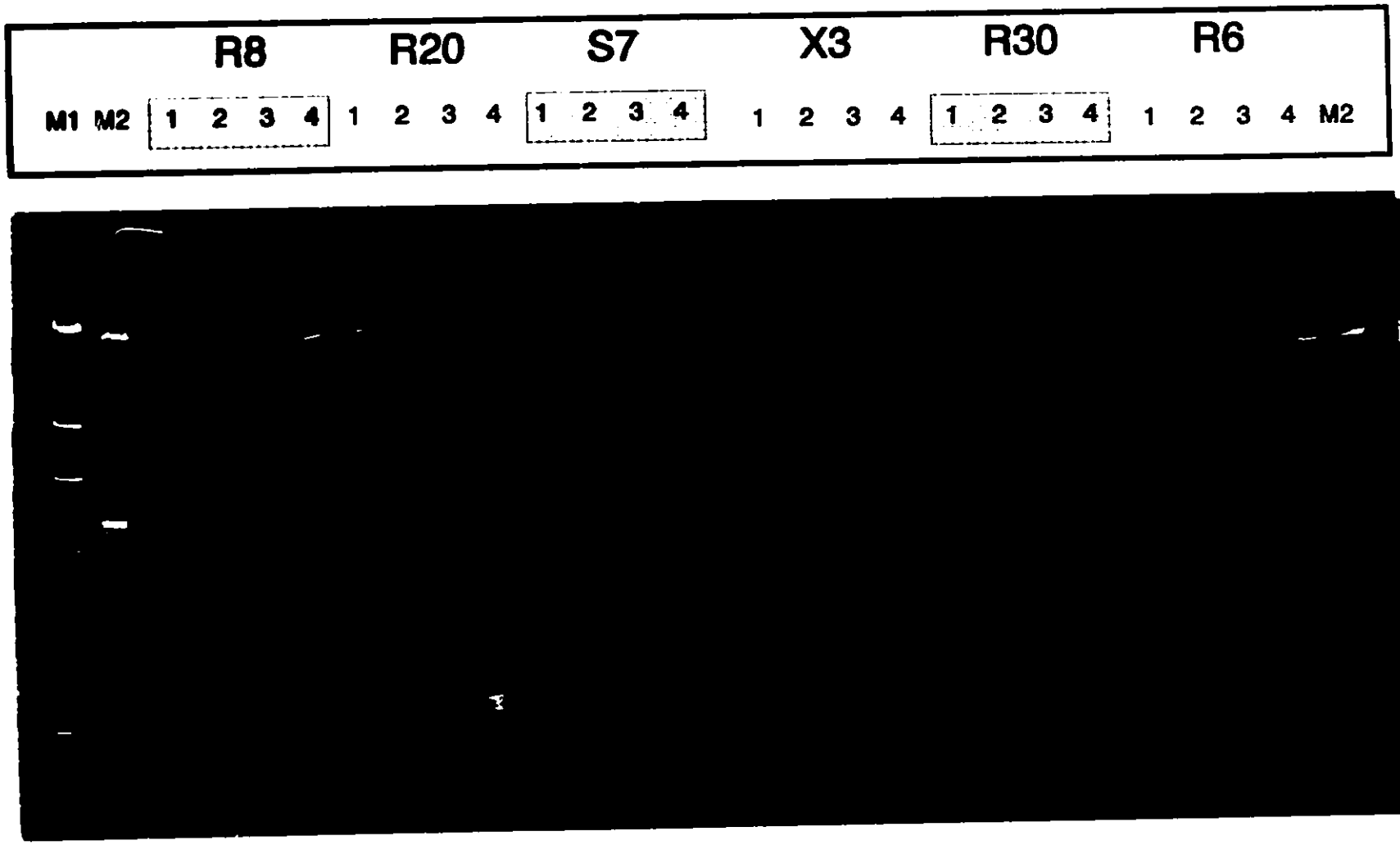
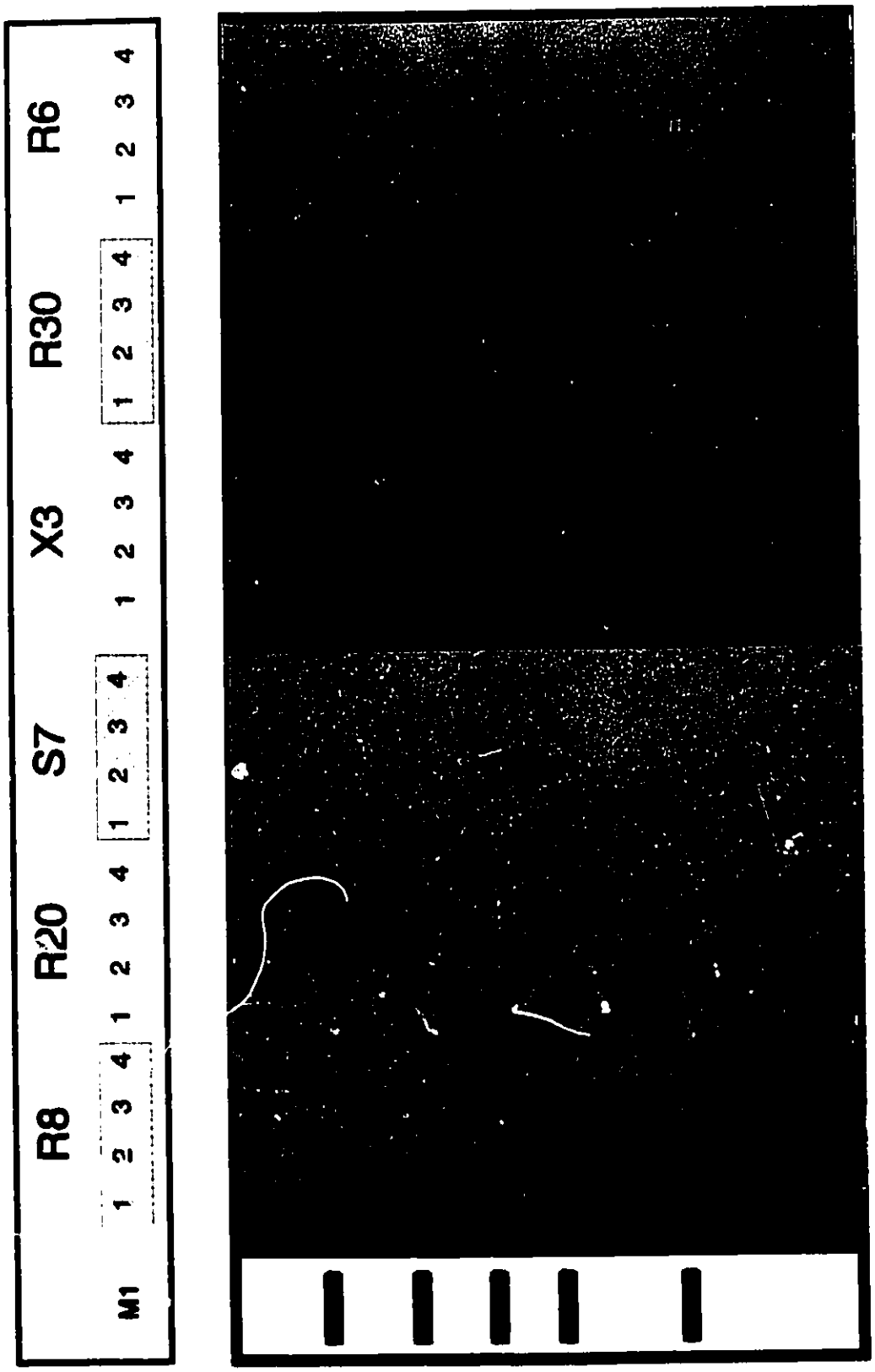

R8
R20
S7
X3
R30
R6
M1 M2 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 M2

FIG. 26 SOUTHERN ANALYSIS OF 6 P+ HU+ GENOMIC CLONES PROBED WITH P PROBE (ANALYSIS OF SAMPLES SHOWN IN FIG. 25)

Lane M1 : Molecular size marker 1, Hind III digested lambda DNA.
Lane 1 : Sal I/Eco RI restriction profile.
Lane 2 : Eco RI restriction profile.
Lane 3 : Hind III restriction profile.
Lane 4 : Bam HI restriction profile.

The labeling of the P probe was performed by the addition of $\alpha[^{32}$-P]dATP and $\alpha[^{32}$-P]dCTP during PCR amplification of the chorio-retinal cDNA library (specific activity of 10^7 cpm/μg).


R8
R20
S7
X3
R30
R6
M1
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4

7.200) NATURE OF HYBRIDIZATION: HUMAN VS PCR

To refine the nature of the differences in cross screening at a physical level a number of P+ HU+ clones were examined for P probe and human specificity using Southern analysis.

Clones R6, R8, and S7 are shown in Fig. 27. On the left hand side(A) of each panel is the digestion profile and the subsequent probing of that lane with P probe. On the right hand side (B) of each panel is a duplicate digestion profile and the subsequent probing of that lane with total human genomic DNA.

The results of this experiment show that the P probe, which reflects a profile of the expressed sequences inherent to choroid/retina tissue, hybridizes to only a select number of bands. In the R6 clone this corresponds to a 8.2 kb band. In the R8 genomic clone there is intense hybridization to a 4.5 kb fragment and very weak hybridization to a 2.5 kb fragment. No hybridization with the probe was observed in the analysis of S7.

The hybridization of human genomic DNA is much less discriminatory, hybridizing to all bands of each profile. In an extreme case P+ hybridization to a Sal I/Eco RI digestion of clone S7 failed to identify any hybridizing fragment, while the human probe appears to hybridize to all possible human restriction bands in all three clones.

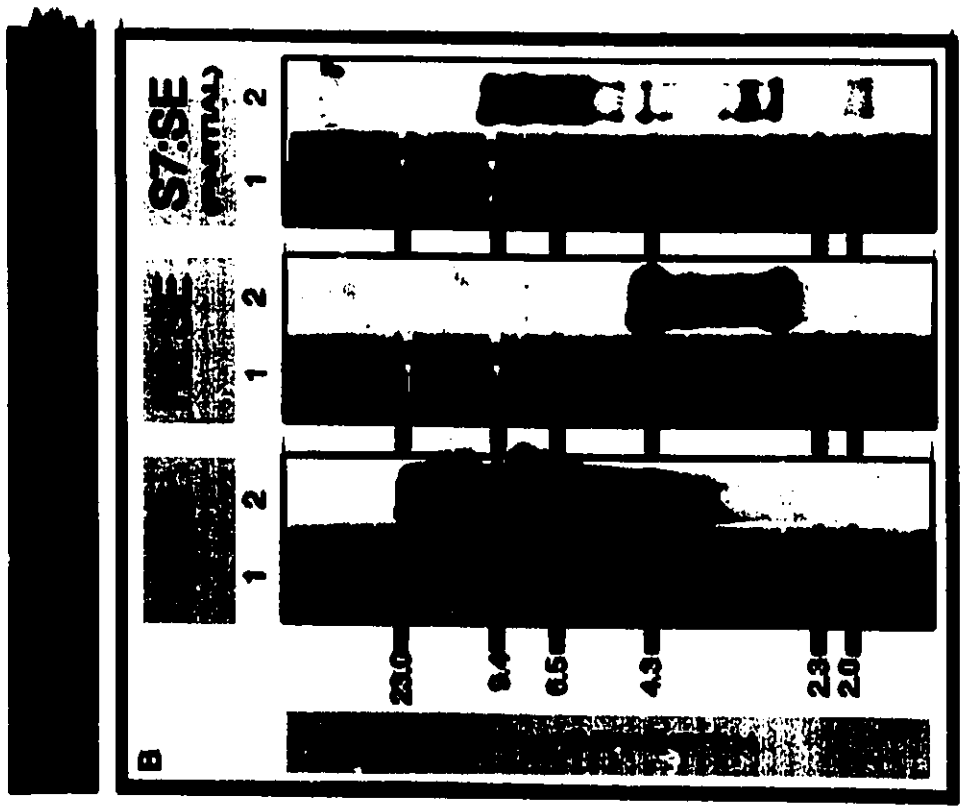
FIG. 27 SOUTHERN ANALYSIS OF SPECIFIC DIGESTION PROFILES OF THREE P+/HU+ CLONES WITH THE P PROBE AND A HUMAN GENOMIC DNA PROBE.

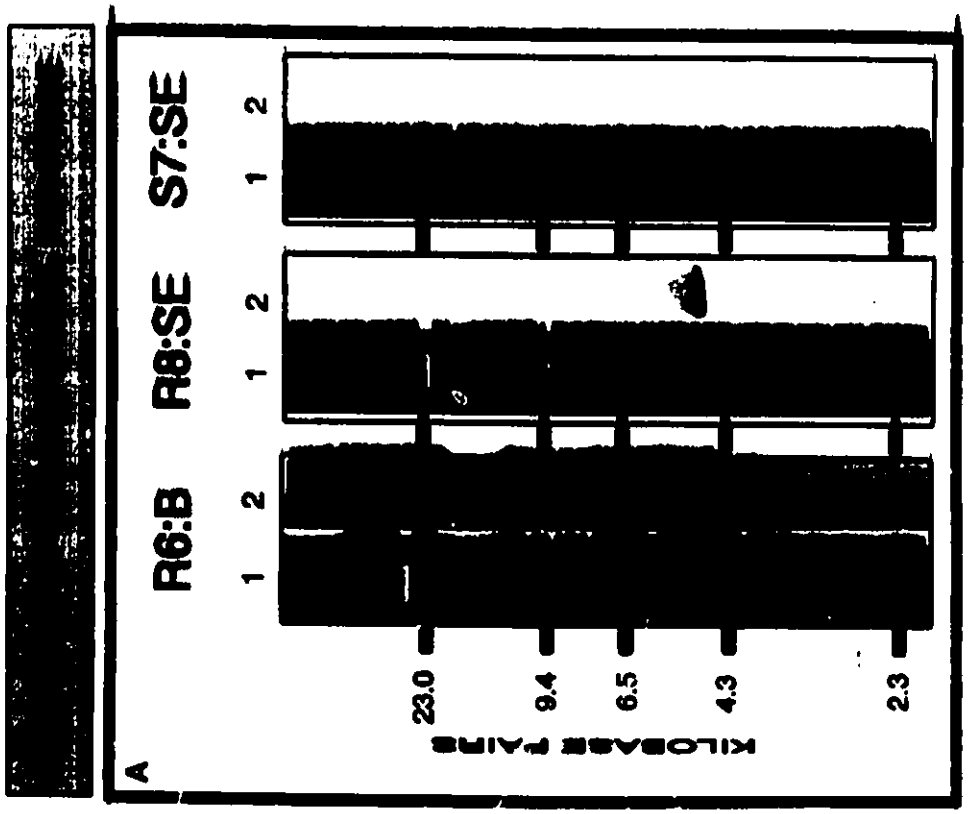
Panel A: Set of restriction profiles and the respective autoradiographic results of Southern analysis using the chorio-retinal P probe. The labeling of the P probe was performed by the addition of $\alpha[^{32}\text{-P}]$dATP and $\alpha[^{32}\text{-P}]$dCTP during PCR amplification of the chorio-retinal cDNA library (specific activity of 10^7 cpm/μg).

Panel B: A duplicate set of restriction profiles and the respective autoradiographic Southern analysis using a total human genomic DNA probe (specific activity of 10^8 cpm/μg

Lane 1: Photograph of the electrophoretic profile of 1μg of digested DNA.
Lane 2: Autoradiographic result of the respective lane after Southern analysis.

B : Bam HI restriction profile.
SE : Sal I/ Eco RI double restriction profile.


B
S7:SE
1 2
1 2
1 2
23.0
9.4
6.5
4.3
2.3
2.0


A
R6:B
R8:SE
S7:SE
1 2
1 2
1 2
KILOBASE PAIRS
23.0
9.4
6.5
4.3
2.3

There is a 1:1 correlation between the DNA fragments which hybridize with the P probe and those which hybridize with the genomic probe. However, not all DNA fragments which hybridize to the genomic probe hybridize to the P probe. This suggests that the two probes hybridize to different sequences within the restriction fragments. The hybridization of total human genomic DNA is primarily on the basis of middle abundant repetitive sequences. The nature of hybridization of the P probe is not as clearly defined, since it consists of both unique low copy and expressed repetitive sequences (which appears to constitute a small portion of the P probe). The results presented here suggest that the P probe must be hybridizing predominantly to sequences that are not abundant in the genome, a conclusion which is consistent with that made previously by Crampton *et al.*, (1981) based on work examining the practicalities of using cDNA to probe genomic libraries.

The observation that some fragments hybridize with both the human and the P probe may at first appear to confuse the issue. However, given that most genes contain exons, largely low copy sequences, and introns which contain repetitive sequences, it is not unexpected to find both repetitive and low copy sequences in restriction fragments greater that one or two kilobases.

7.300) THE EXPECTED NUMBER OF CLONES: A CALCULATION

It is difficult to anticipate the exact number of expressed retinal sequences on the X-chromosome. However it is generally accepted that there are approximately

3-4 expressed genes per 10^6 bp of genomic DNA sequence (Dr. P. O'Connell, personal communication). The X-chromosome represents roughly 5.84% of the 3.3 x 10^9 bp in the human somatic haploid genome. This means therefore that the X-chromosome consists of 1.9 x 10^8 bp of nucleotide sequences and contains roughly 580-770 active genes.

The average insert size of the 82082a genomic library is approximately 19 kbp in length. This means that 10,000 independent human clones are required to span the X-chromosome. On average therefore, if the genes are equally distributed and assuming that the gene size is not large enough to extend a given gene into two or more genomic clones, one should find an expressed gene every 13-17 human genomic clones, an expected frequency of 5.8-7.7%. In the case of the 82082a library which consists of human and hamster genomic clones, the frequency of a single human clone is approximately 0.66% or 1 in 154. Thus the expected frequency of human X-clones that harbor expressed genes in the hybrid library is approximately 0.04%-0.05%. Due to the sensitivity of the detection system it is expected that genomic clones corresponding to highly and middle abundant expressed messages will be detected: since these sequences represent 70% of the total number of different expressed species in a typical cell a more realistic expected frequency of human X-clones harboring expressed genes is 0.03%. If the sub-population of P+ HU+ HA- clones represents the group of human clones which harbor chorio-retinal expressed sequences then the overall observed frequency (0.30%) is roughly ten times that

which is theoretically expected.

7.400) FACTORS WHICH MAY CONTRIBUTE TO BACKGROUND

The screening of genomic clones using in vitro labeled RNA or cDNA often leads to the selection of false positives (Williams, 1981). It is possible that some of this background may be due to cross hybridization of the P probe to more than one gene as a result of homologies between genes within gene families or to pseudogene sequences. However, because the high wash stringencies used in the present study are likely to limit the degree of cross hybridization, it is unlikely that the apparent ten fold background is solely due to this form of cross hybridization.

A second possibility for the difference between the observed and theoretical background maybe the presence of poly(dT) tracts in mammalian DNA, which may hybridize with the poly(A) tracts present in the P probe. Since the cDNA used for the chorio-retinal library was synthesized using a oligo(dT) primer it is likely that these tracts will be present in the P probe. However at a practical level it is not clear to what degree this particular parameter would contribute to the over all background. First, the screening of the chorio-retinal cDNA library with chorio-retinal cDNA does not lead to the cross hybridization of probe to the entire cDNA library, even though most, if not all, clones will contain a stretch of poly(A). Secondly, poly(dT) tracts are present in many mammalian genomes and it is therefore likely that the calf thymus DNA used in the prehybridization step will compete out this background.

An inherent source of contamination in the preparation of poly(A)+ enriched RNA are ribosomal sequences. The degree of contamination will depend on the stringency with which the poly(A)+ enriched RNA is prepared. Since the ribosomal genes are present in multiple copies in the genome, rRNA contamination in the P probe could lead to the isolation of a number of false positive genomic clones. Ribosomal RNA, however, is copied with a very low efficiency into cDNA and its representation in the size fractionated cDNA library used in this study should not be very high (Williams and Lloyd, 1979; Lu and Werner, 1988). In addition ribosomal genomic clones should not pose a great problem because they are easily identifiable by Southern analysis. Ribosomal genes characteristically demonstrate multiple bands correlating with multiple copies of the gene within the genome. Also, mRNA used for cDNA synthesis in the present study did not have significant amounts of 28S and 18S rRNA contamination (Fig. 12, chapter three).

The final possible factor which may contribute to the background hybridization is the degree of repetitive expressed sequences present in the chorio-retinal library. Such repetitive sequences are found primarily in the 3' and 5' untranslated regions of mRNA. Not all expressed sequence are necessarily characterized with repetitive regions, and the proportion of clones harboring expressed sequences will vary from tissue to tissue. cDNA probes generated from lymphocyte or brain mRNA have a high proportion of repetitive sequences and tend to cross hybridize to greater than 90% of human clones in standard genomic libraries

(Hochgeschwender *et al.*, 1989; Paulson *et al.*, 1986; Tashima *et al.*, 1981). On the other hand, cDNA probes generated from placental mRNA have been found to cross hybridize to less than 40% of the clones in a human genomic library (Tashima *et al.*, 1981). In the present study the P probe derived from the human chorio-retinal cDNA library hybridizes to approximately 45% of the human genomic clones in the 82082a genomic library.

7.500) REPETITIVE SEQUENCES

7.510) Representation of repetitive sequences in the cDNA library

It is difficult to determine the extent of repetitive sequences in a heterogeneous pool of sequences or to anticipate the degree of background that they will present in the screening procedure. In the case of human lymphoblast amplified cDNA libraries as many as 10% of clones contain some form of repetitive sequence (Crampton *et al.*, 1981).

In order to determine the level of repetitive sequences in the unamplified chorio-retinal cDNA library which is abundant in the genome, the cDNA library from which the P probe was derived was screened with labeled human genomic DNA. In that analysis it was determined that approximately 1.4% of the clones represented in that library (43 of 3054) contained some sort of repetitive sequence that exists at reasonable frequencies in the genome. This is not to say that 1.4% of the sequence is repetitive DNA as each clone is likely to contain both repetitive and unique

sequences. The estimate established here (1.4%) for chorio-retinal cDNA clones that show hybridization to a total human genomic probe agrees well with those of others, 2% (B.Hill, Edinburgh, personal communication), and is slightly lower that those determined for amplified retinal libraries, 5% (P. Monaghan, Edinburgh, personal communication).

It has been suggested that a probe with 0.1% repetitive sequence is enough to mask the detection of specific clones in a heterogeneous pool of sequences reflecting the mRNA pool (K.Davies, Oxford, personal communication). However it would be unreasonable to assume that the repetitive sequence observed in the P probe is homogeneous. In lymphocyte mRNA the repetitive sequences belong to a number of diverse and different repetitive elements, and it has been concluded that the presence of repetitive elements within expressed sequences makes a minor contribution to the hybridization signal (Crampton *et al.*, 1981).

7.520) Contribution of repetitive sequences to hybridization

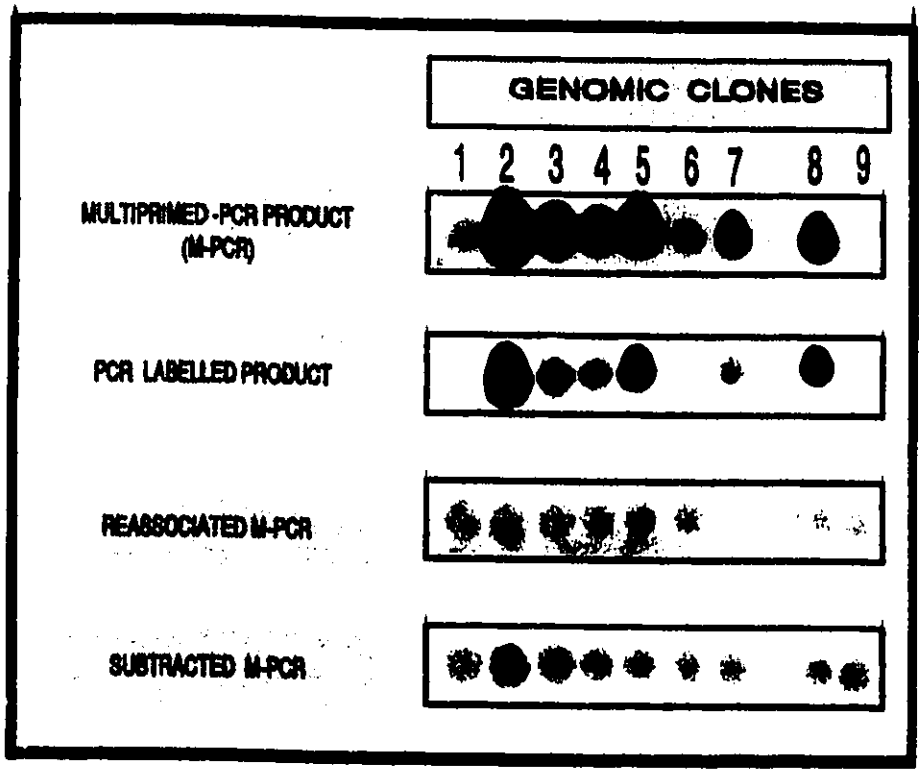
To establish the contribution of repetitive sequences within the P probe to the isolation of clones, duplicate phage dot lifts of a number of P+ HU+ HA- clones were made and probed with the P probe and a P probe from which repetitive sequences had been depleted by pre-annealing the probe with an excess amount of sheared total human genomic DNA. A 1000 fold excess of sheared genomic DNA over a cDNA probe is known to be sufficient to compete out the hybridization of repetitive sequences to genomic clones (Tashima *et al.*, 1981; Paulsen *et al.*, 1986).

In the initial study, an 11,000 fold excess of sheared human genomic DNA was used to reassociate with the P probe prior to the screening of the plaque lifts. In general the intensity of hybridization after reassociation for all clones showed a global decrease in signal (Fig. 28). However, in a total of 44 PCR+ HU+ HA- clones screened, 22 were found to show a more noticeable decrease in hybridization signal relative to the surrounding clones on the blot and to their unblocked profiles. Reassociation did not eliminate the hybridization to any of these clones. Except for the global dulling of the hybridization signals, the relative hybridization intensity to control non-human clones which were P+ was not affected. Similarly hamster clones which were P- did not show hybridization with the P or the depleted P probe (not shown).

Conceptually the reassociation of the probe to excess amounts of human genomic DNA should immobilize the repetitive fraction of the probe and therefore prevent it from hybridizing. The fact that 22 clones showed a decrease in relative hybridization signal in the first experiment is an indication that about 50% of the P+ HU+ HA- clones contain some form of repetitive sequence. This demonstration of repetitive sequences in the isolated clones is not surprising and is probably a consequence of the bias introduced into the cDNA synthesis through the use of oligo(dT) as the primer. This methodology results in the over representation of the 3'end of expressed sequences that may contain repetitive elements in the library.

FIG. 28 SCREENING OF P+ HU+ CLONES WITH P PROBE DEPLETED OF REPETITIVE SEQUENCES.

M-PCR	: Multiprimed labeled PCR chorio-retinal cDNA library probe (specific activity of 10^9 cpm/μg).
PCR-labeled product	: PCR chorio-retinal cDNA library probe labeled during PCR amplification (specific activity of 10^7 cpm/μg).
Reassociated M-PCR	: M-PCR that has been reassociated with excess amounts of sheared human genomic DNA.
Subtracted M-PCR	: M-PCR that has been hybridized to excess human genomic DNA that has been cross linked to a cellulose matrix.


GENOMIC CLONES
1 2 3 4 5 6 7 8 9
MULTIPRIMED-PCR PRODUCT (M-PCR)
PCR LABELLED PRODUCT
REASSOCIATED M-PCR
SUBTRACTED M-PCR

However, the observation that the reassociation of the P probe with a massive excess of sheared human genomic DNA did not obliterate the hybridization signal suggests either these clones have been identified on the basis of hybridization to sequences not abundant in the genome or that the depletion method was not efficient. The observed hybridization profile observed with the depleted P probe indicates that the hybridization is specific since the clones behaved differently: one subset of clones showed a marked decrease in intensity (Fig. 28: clones 2, 3, 5, 7, 8), a second did not show any relative change (Fig. 28: clones 1, 4, 6) and a third subset demonstrated a relative increase in intensity (Fig. 28: clone 9).

In order to confirm the above findings, a second method for depleting repetitive sequences from probes was used. Human genomic DNA, coupled to diazotized cellulose (DNA-cellulose), was used to subtract repetitive P probe sequences. The bound repetitive probe was separated from unbound unique probe by centrifugation. P probe subtracted in this way was used to screen duplicate phage dot lifts of the plates used for the first set of depletion studies. The results of hybridization with this probe were the same as those observed in the previous depletion study (Fig. 28), with the same clones showing a noticeable decrease in hybridization intensity.

7.530) Comparison to other studies

There are few studies in the literature which examine the cross screening of genomic libraries with a heterogenous probe pool representing the expressed sequences of a specific tissue. One of the earlier studies screened a flow sorted human X-specific genomic library for X-specific clones harboring lymphoblast expressed sequences (Paulsen *et al.*, 1986). This study demonstrated the feasibility of competition of the cDNA pool probe with sheared genomic DNA to remove repetitive sequences. However this study also reported, rather surprisingly, that an extremely low proportion of X clones harbor expressed sequences (0.12%), a number which is 20 fold less than expected for the number of potential active genes on the X chromosome (2.2%). A more recent study isolating mouse chromosome 16 brain expressed sequences using similar methodologies also reported the isolation of a rather low proportion of clones potentially harboring expressed sequences (1.5%), 4 fold less than expected (Hochgeschwender *et al.*, 1989).

In the present study the reverse observation is seen. There appears to be a ten fold higher than anticipated number of clones that potentially harbor expressed sequences (0.30% vs 0.03%). However if only the P+ HU+ HA- clones which did not show a decrease in hybridization signal after depletion of repetitive sequences are considered, then the percentage of potentially clones falls to 0.15% (bringing the numbers of clones to only a five fold excess). The present study however does have some key differences from previously published studies. In contrast to the first study

mentioned above, the genomic library screened was unamplified and used a recombinant non-permissive host in plating. In contrast to the second study, the order of cross screening is reversed. In contrast to both studies this study uses a P probe derived from a cDNA library.

7.600) SUMMARY

The results of the set of experiments presented in this current chapter suggests that there is a certain level of background of false positives due to the inherent nature of the P probe derived from the cDNA library. Although highly repetitive sequences may play a role in the intensity of hybridization of a given clone, hybridization does not appear to be exclusively due to these sequences. In addition it is possible that the biased cloning of the 3' untranslated region of the mRNAs may complicate the hybridization profile due to the presence of expressed repetitive sequences in this region. As such it is important to verify the effectiveness of the screening procedure by demonstrating that the clones isolated are human X-specific sequences and harbor chorio-retinal expressed sequences.

CHAPTER EIGHT

MAPPING OF ISOLATED CLONES

8.000) X-DOSAGE AND SPECIES SPECIFICITY ANALYSIS

The isolation of a human clone from the 82082a genomic library dictates that it will consist of a sequence from the human X-chromosome, since this is the only human component of the 82082a cell line. However it is important to determine if the clones isolated in this manner are representative of the X-chromosome as a whole, or whether the clones are derived from a biased region of the X-chromosome.

This was accomplished by mapping discrete fragments isolated from selected P+ HU+ HA- clones to a human X-specific DNA panel. The panel consists of either Eco RI or Bam HI digested DNA from a series of rodent/human X-somatic cell lines. The DNA of these hybrids was obtained from Dr. J.L. Hamerton (Winnipeg, Manitoba), in each case the cytogenetic profile of each line, as reported by Wieacker *et al.*, (1981), was confirmed on cells from the same cell passage used for DNA isolation (Dr. C. Gregory, Winnipeg, personal communication). The collected X-panel represents a gross division of the X-chromosome with the breakpoints of the three cell lines (676x175, 749, and 617x347) dissecting the X chromosome into four discrete regions (Fig. 29).

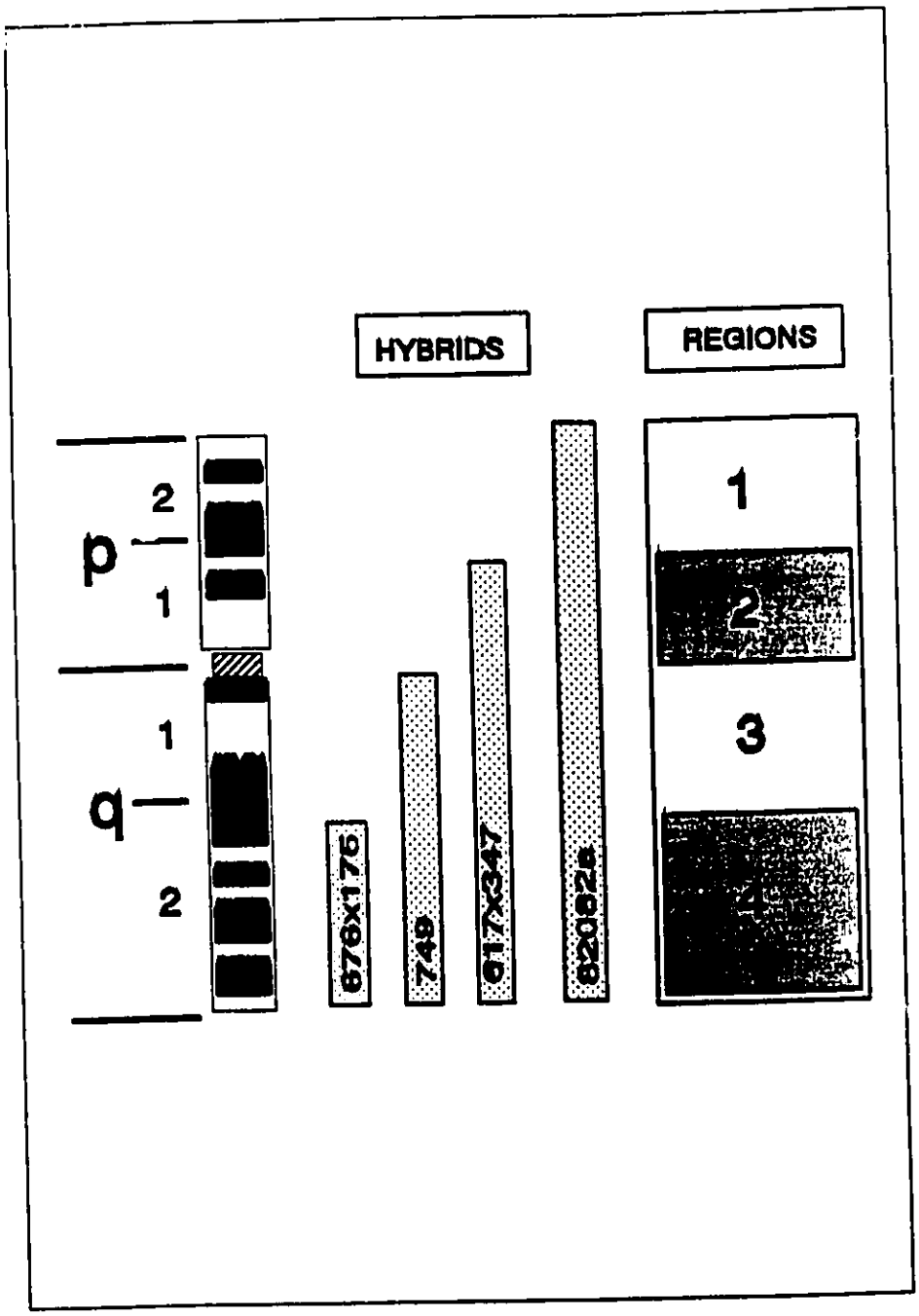
8.100) SELECTION OF GENOMIC FRAGMENTS TO BE MAPPED

The same six clones that were analyzed for independence were used as the representative group of clones to be mapped. In each case a single DNA fragment from each clone, generated by a Bam HI or a Sal I/Eco RI digestion, that

FIG. 29 SUMMARY OF THE REGIONS SPANNED BY EACH MEMBER OF THE X-SPECIFIC HYBRID PANEL USED FOR MAPPING.

The regions defined by the 4 cell lines dissects the X-chromosome into four discrete regions. The approximate region spanned by each hybrid is illustrated in the figure

a) 676x175 : Xq21.3-qter
b) 749 : Xq12-qter
c) 617x347 : Xp11.3-qter
d) 82082a : Xqter-Xpter


HYBRIDS
REGIONS
p
2
1
q
1
2
678X176
749
617X347
820828
1
2
3
4

hybridized with the P probe was used as probe. In the case of S7 where no hybridization to the P probe was evident, a series of fragments were isolate at random and tested for X-dosage by a slot blot analysis (Fig. 30). Of three fragments tested one was found to display a dosage effect and was chosen for mapping.

Insert DNA for each clone was isolated directly after electrophoresis in low melting agarose of the suitably restricted cloned DNA (Fig. 31). The isolated restriction fragments were multi-prime labeled directly in low melting agarose without prior purification and repetitive sequences were depleted from the probes by pre-annealing with an excess of sheared total human genomic DNA (Sealey *et al.*, 1985). The specific activity of probes generated ranged from 10^8-10^9 cpm/μg DNA.

8.200) SPECIFICITY OF THE X-PANEL

To confirm the specificity of the X-panel, human X-probes of known locations were mapped on the X-panel defined above. Two known and well characterized DNA probes, L1.28 and p22.33 were used to test the panel. L1.28 defines the DXS7 locus which maps to the short arm of the X-chromosome at band Xp11.3 (region 2, Fig. 29). p22.33 defines the DXS11 locus which maps to the long arm of the X chromosome between Xq24 and Xq26 (region 4, Fig. 29). In both cases labeled insert was used to screen the X-panel and in both cases the hybridization pattern was consistent with the expected pattern of hybridization for the defined regions (Fig. 32).

FIG. 30 ISOLATION OF A HUMAN X-SPECIFIC FRAGMENT FROM CLONE S7.

Screening of clone S7 with the PCR library probe did not reveal any positive bands. As such individual bands (labeled A, B, and C) from a Sal I/Eco RI digest (SE) were isolated. Each band was tested for dosage by using it as a probe against equal amounts (1 μg) of female DNA (XX) and male DNA (XY) in a slot blot analysis. In all cases probes were labeled by multiprime labeling and were reassociated with excess sheared human genomic DNA prior to use. Only one S7:A of the three bands examined showed evidence of dosage. This band was used in further studies and is regarded as a fragment which harbors a low copy X-specific human sequence.

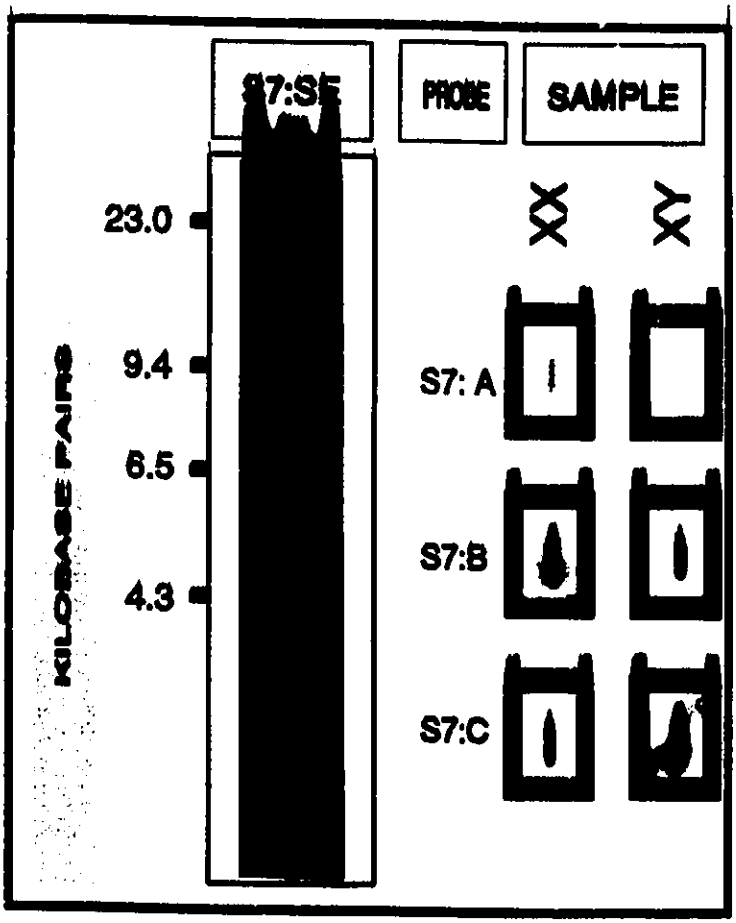

S7:SE
PROBE
SAMPLE
XX
XY
23.0
9.4
6.5
4.3
KILOBASE PAIRS
S7: A
S7:B
S7:C

FIG. 31. SUMMARY OF SPECIFIC CLONED FRAGMENTS FROM A RANDOM SUBSET OF THE HU +/ P + GENOMIC CLONES ISOLATED TO BE USED AS PROBES IN FURTHER STUDIES.

Six clones were selected at random from the group of clones already isolated. Discrete fragments were generated by a Sal I(S)/Eco RI(E) or Bam HI(B) digestion. Specific bands chosen for use as probes are indicated by an arrow. The respective insert sizes and cutting sites used to generate the fragments are indicated beneath the gel.

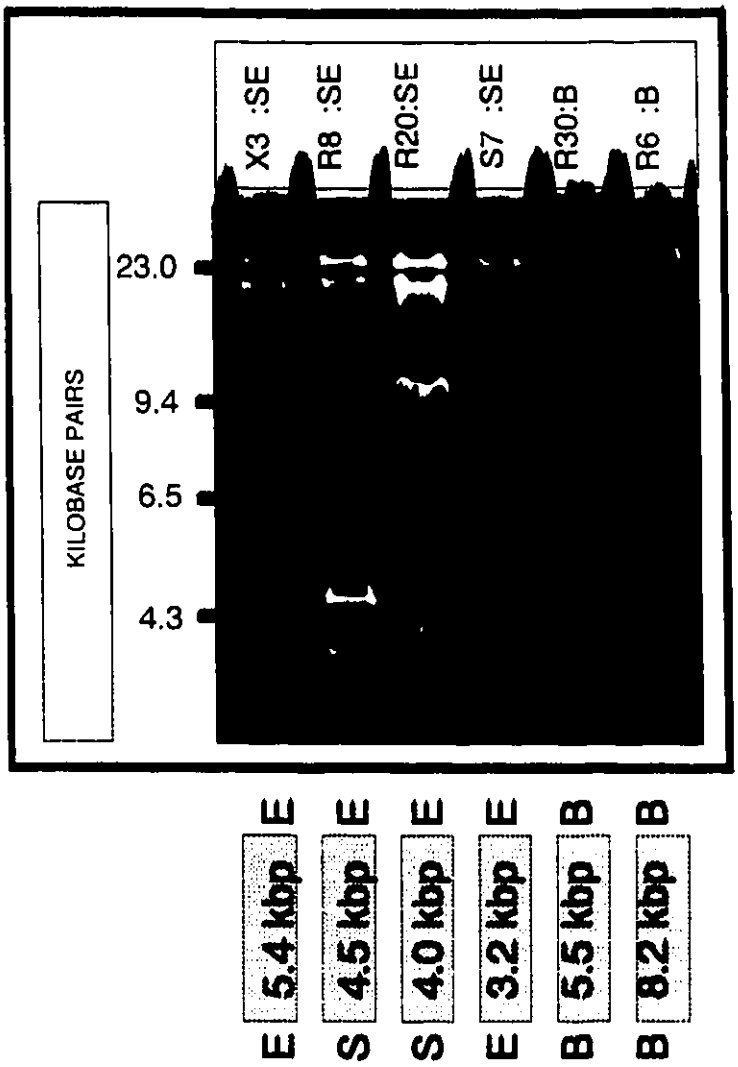

X3 :SE
R8 :SE
R20:SE
S7 :SE
R30:B
R6 :B
KILOBASE PAIRS
23.0
9.4
6.5
4.3
E 5.4 kbp E
E 4.5 kbp E
S 4.0 kbp E
E 3.2 kbp E
B 5.5 kbp B
B 8.2 kbp B

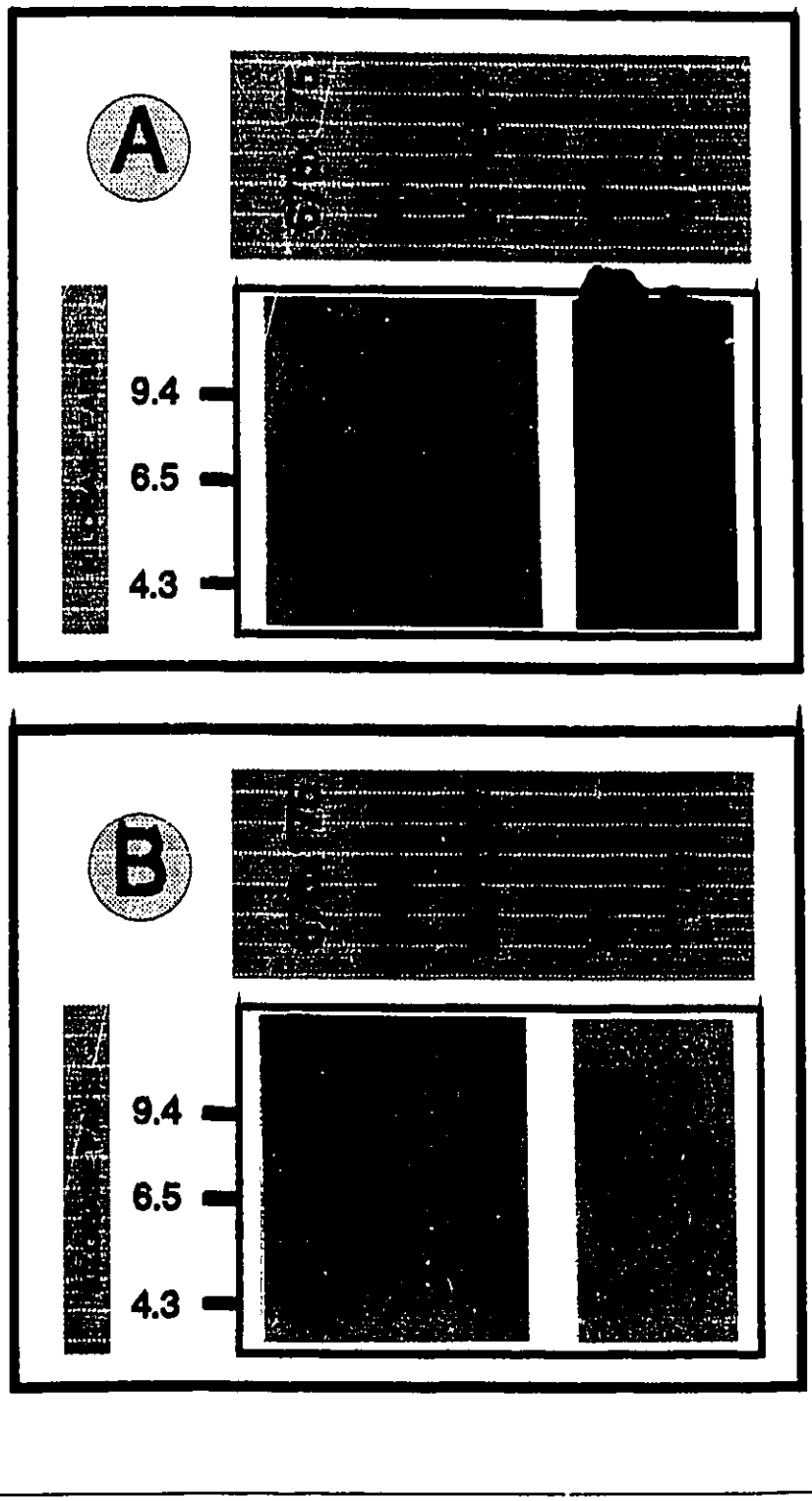
FIG. 32 HYBRIDIZATION OF KNOWN PROBES TO THE DNA X-PANEL.

The mapping panel was tested by the localization of two X-specific DNA probes of known map positions.

Panel A: Southern analysis of the Eco RI profile of the proposed X-panel probed with p22.33. p22.33 recognizes the DNA locus DXS11 and has been localized to Xq21-Xq26. In the Eco RI profile the probe hybridizes to a non-polymorphic 5.0 kbp fragment. The hybridization profile of (676x175 +: 749 +: 617x347 +) is consistent with its known map position.

Panel B: Southern analysis of the Bam HI profile of the proposed X-panel probed with L1.28. L1.28 recognizes the DNA locus DXS7 and has been localized to Xp11.3. In the Bam HI profile the probe hybridizes to a non-polymorphic 6.0 kbp fragment. The hybridization profile (676x175 -: 749 -: 617x347 +) is consistent with its known map position.

1103 : Digested genomic DNA isolated from a hamster cell line.
XY : Digested human genomic DNA isolated from a normal male


A
9.4
6.5
4.3
B
9.4
6.5
4.3

8.300) MAPPING OF SPECIFIC GENOMIC CLONES

8.310) Mapping of R8

The R8 subfragment (SF) used for localization was a 4.5 kbp Sal I/Eco RI fragment which hybridized to a single 9.4 kbp fragment on Southern analysis (Fig. 33). The R8SF hybridizes to 82082a but not with 1103, thus localizing it to the human X-chromosome. Hybridization of the fragment to 617x347 but not to 676x175 nor 749 further localizes the R8SF to Xp11.4-q11.

8.320) Mapping of X3

The X3SF used for localization was a 5.4 kbp Eco RI fragment. In the initial phage dot analysis, it was not possible to clearly conclude the species of origin of this clone since equal intensity hybridization was observed to both human and hamster sequences. However, the stronger hybridization to hamster genomic DNA coupled with the hybridization of the probe to a 5.4 kbp (as well as to a number of other hamster specific bands) hamster band observed in the Southern analysis indicates that X3 is a hamster derived clone (Fig. 34). In longer exposures of the autoradiographs a very distinct 7.4 kbp human band appears in the 82082a and the XX and XY human lanes and is noticeably absent in the 1103 lane. This fragment is present in the analysis of 617x347 but not to 749 nor 676x175 thus localizing a X3 orthologous sequence to human Xp11.4-q11.

FIG. 33 MAPPING OF CLONE R8

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines, was digested with Eco RI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 4.5 kbp Sal I/Eco RI fragment isolated from clone R8. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a human 9.4 kbp fragment (↗). The hybridization pattern observed is as follows:

676x175	-
749	-
617x347	+
82082a	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.

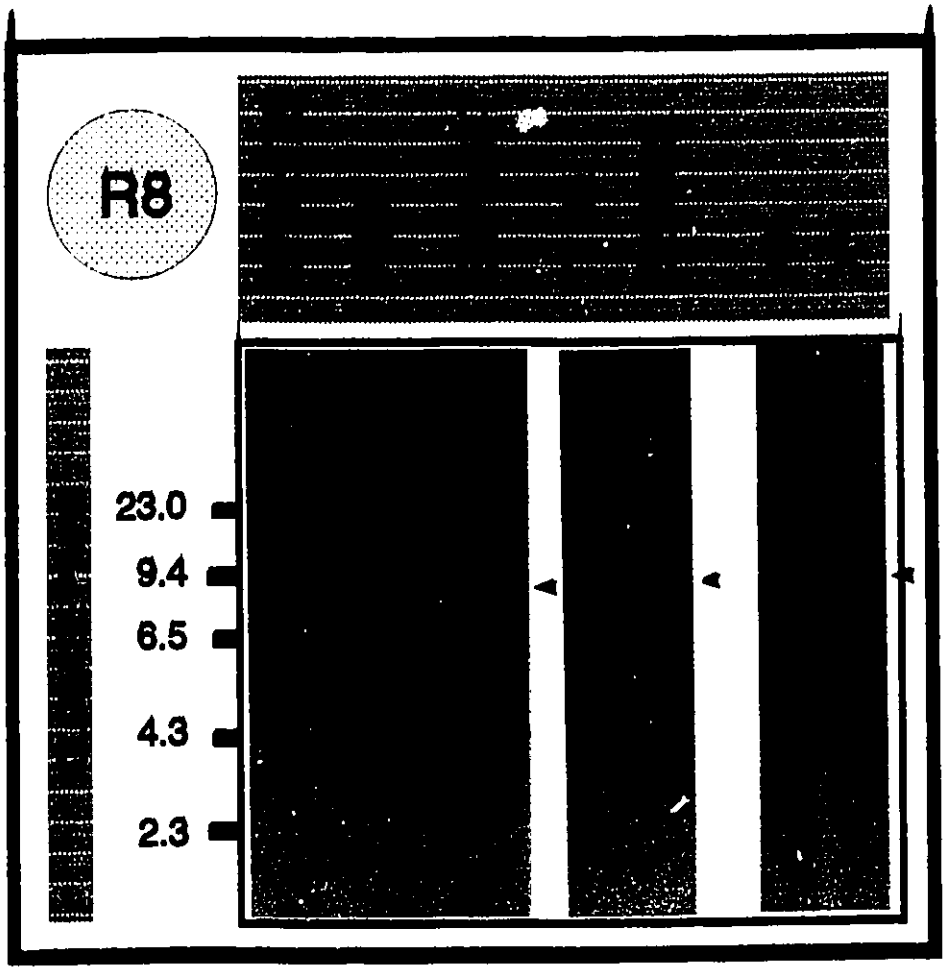

R8
23.0
9.4
6.5
4.3
2.3

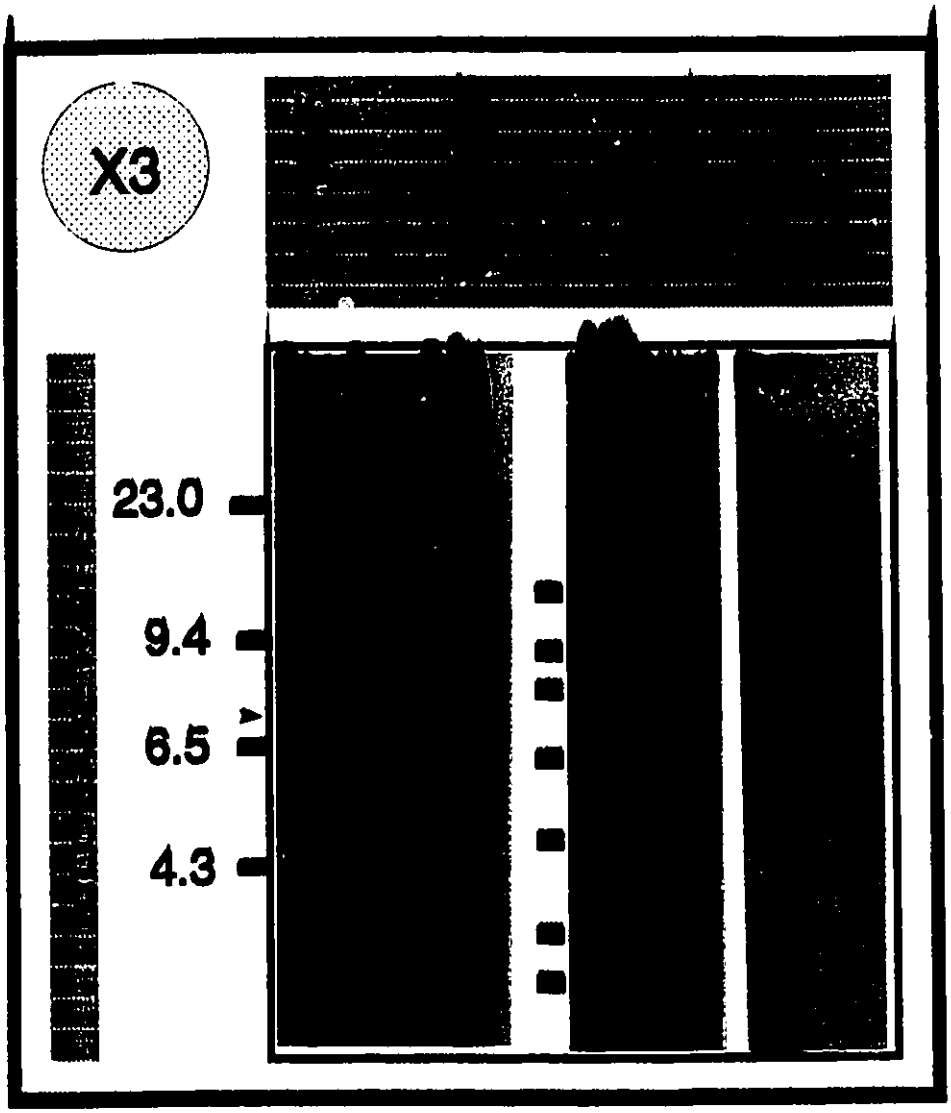
FIG. 34 MAPPING OF CLONE X3

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines was digested with Eco RI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 5.4 kbp Eco RI fragment isolated from clone X3. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a number of hamster fragments (as indicated diagramatically by boxes next to the 1103 lane) including to a hamster specific 5.4 Kbp fragment(*) and to a 8.4 kbp human specific fragment (↗). The hybridization pattern observed for this 8.4 Kbp band is as follows:

676x175	-
749	-
617x347	+
82082a	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.


X3
23.0
9.4
6.5
4.3

8.330) Mapping of R30

Clone R30 also hybridizaed with both species probes in the phage dot analysis. However, the relative hybridization signal with the human genomic probe was much stronger than with the hamster. The R30SF used for localization was a 5.5 kbp Bam HI fragment. The fragment did not hybridize to 1103, thus confirming that R30 was a human clone (Fig. 35). R30SF hybridized to a 5.5 kbp fragment in both 617x347 and 749, but did not show evidence of hybridization to 676x175 thus localizing R30 to Xq21-q11.

8.340) Mapping of S7

The S7SF used for localization was a 3.2 kbp fragment. In the Southern analysis it did not hybridize to 1103 but did identify both a 3.2 kbp fragment as well as a 2.5 kbp fragment in the human genomic lanes and 82082a (Fig. 36). In the case of the human lanes there appears to be a clear dosage effect between male DNA and female DNA, demonstrating that clone S7 is a human X-specific clone. Analysis of the X-panel with this S7SF identified the same two bands in all three of the hybrids, thus localizing S7 to Xq21-qter.

8.350) Mapping of R6

R6, based on the phage dot analysis, was clearly a human specific clone. The R6SF used for localization was a 8.2 kbp Bam HI fragment. This fragment did not hybridize with 1103 but did hybridize to a 8.0 kbp fragment in all three hybrid lanes

FIG. 35 MAPPING OF CLONE R30

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines was digested with Bam HI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 5.5 kbp Bam HI fragment isolated from clone R30. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a human 5.5 kbp fragment (↗). The hybridization pattern observed is as follows:

676x175	-
749	+
617x347	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.

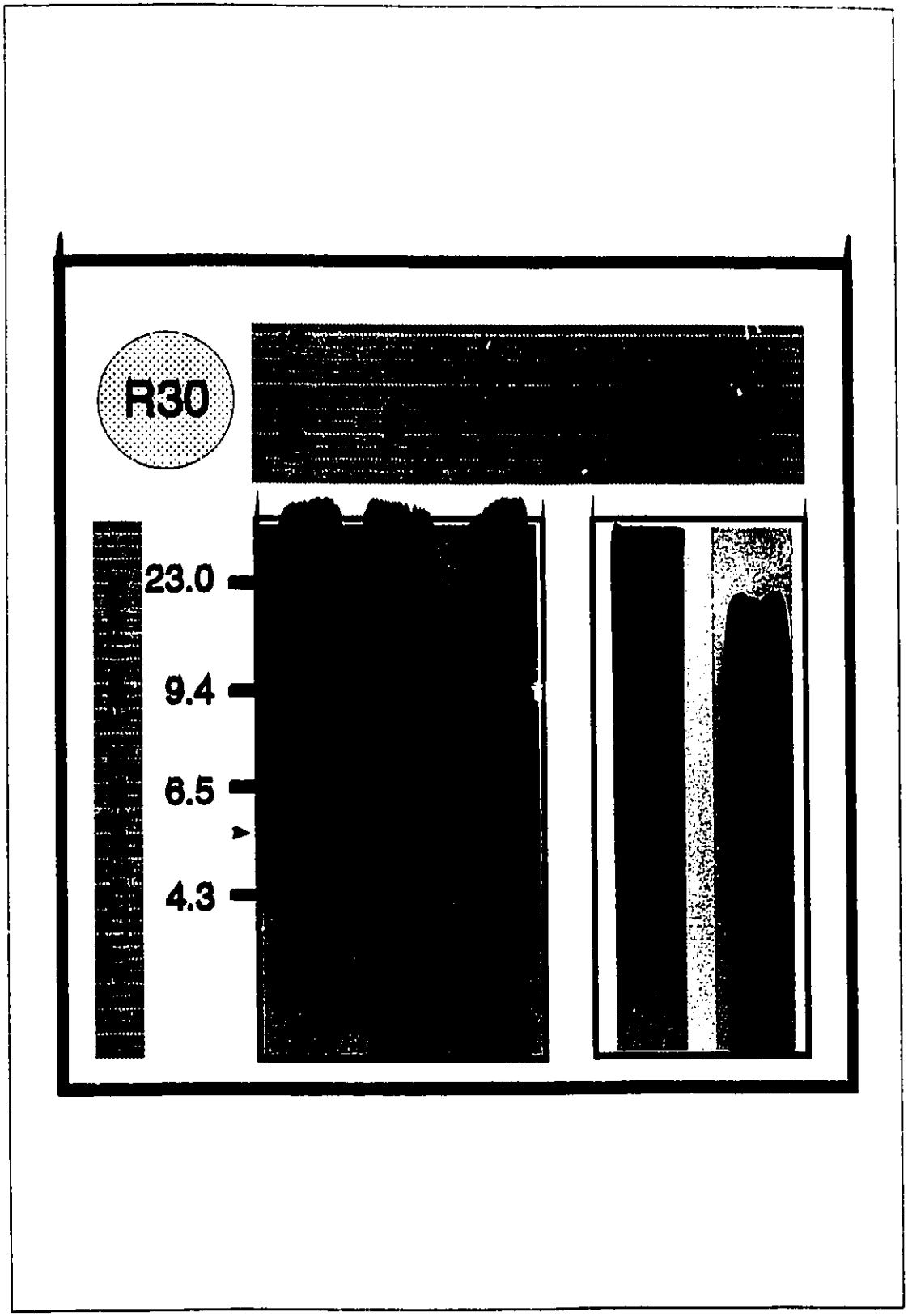

R30
23.0
9.4
6.5
4.3

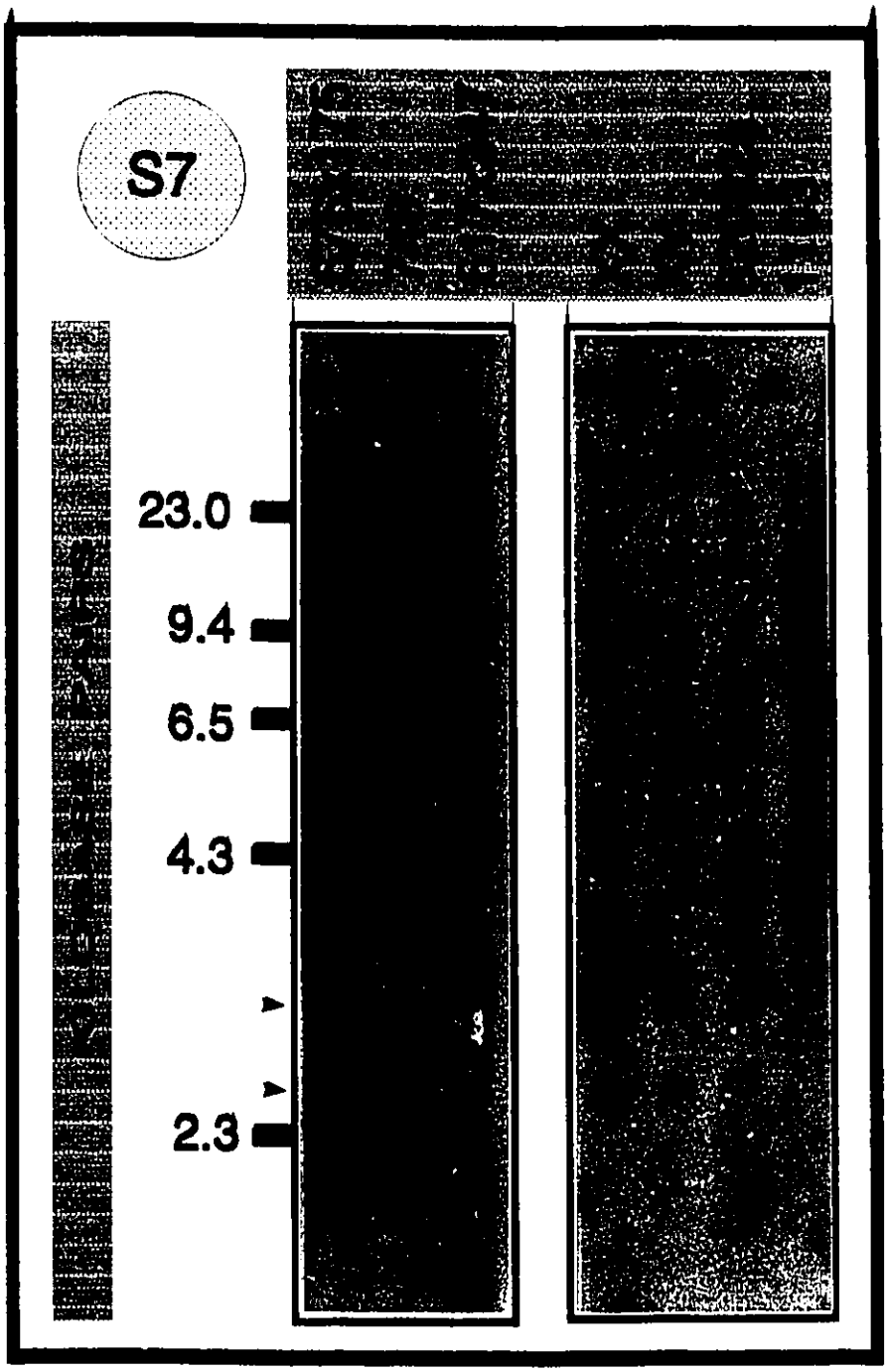
FIG. 36 MAPPING OF CLONE S7

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines was digested with Eco RI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 3.2 kbp Eco RI fragment isolated from clone S7. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a human 3.2 and 2.5 kbp fragment (↗). The hybridization pattern observed is as follows:

676x175	+
749	+
617x347	+
82082a	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.


S7
23.0
9.4
6.5
4.3
2.3

(Fig. 37). The location of R6 is therefore within Xq21-qter.

8.360) Mapping of R20

The most interesting of the clones may in fact be R20 which shows a very complex banding pattern in the analysis of human genomic DNA. The R20SF is a 4.0 kbp Sal I/ Eco RI fragment. It hybridizes very faintly to what appears to be rodent orthologous sequences presenting with a pair of bands between 9 and 6 kbp in size. A 5.0 kbp fragment hybridizes with specificity to 82082a and the human samples but not to 1103 (Fig. 38). It is assumed that this band is of human origin and that this particular band stems from the X chromosome. Analysis of the X- panel shows that this band is not present in any of the cell lines tested. Based on this data, the 5.0 kbp band identified by the R20SF can be provisionally placed in region 1(Xp11.4-pter).

What is very clear is that R20 is human and as such must be derived from the human X-chromosome. The complex banding pattern seen in the human lanes indicate that the R20SF cross hybridizes to a multiplicity of similar gene loci. Since only one of these bands can be localized to the X-chromosome, it can be concluded that the remaining bands must be derived from other chromosomal origins.

Since the R20SF probe was pre-associated with total human genomic DNA prior to its use it can be assumed that repetitive sequences greater than 100 copies

FIG. 37 MAPPING OF CLONE R6

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines was digested with Bam HI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 8.2 kbp Bam HI fragment isolated from clone R6. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a human 8 kbp fragment (↗). The hybridization pattern observed is as follows:

676x175	+
749	+
617x347	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.

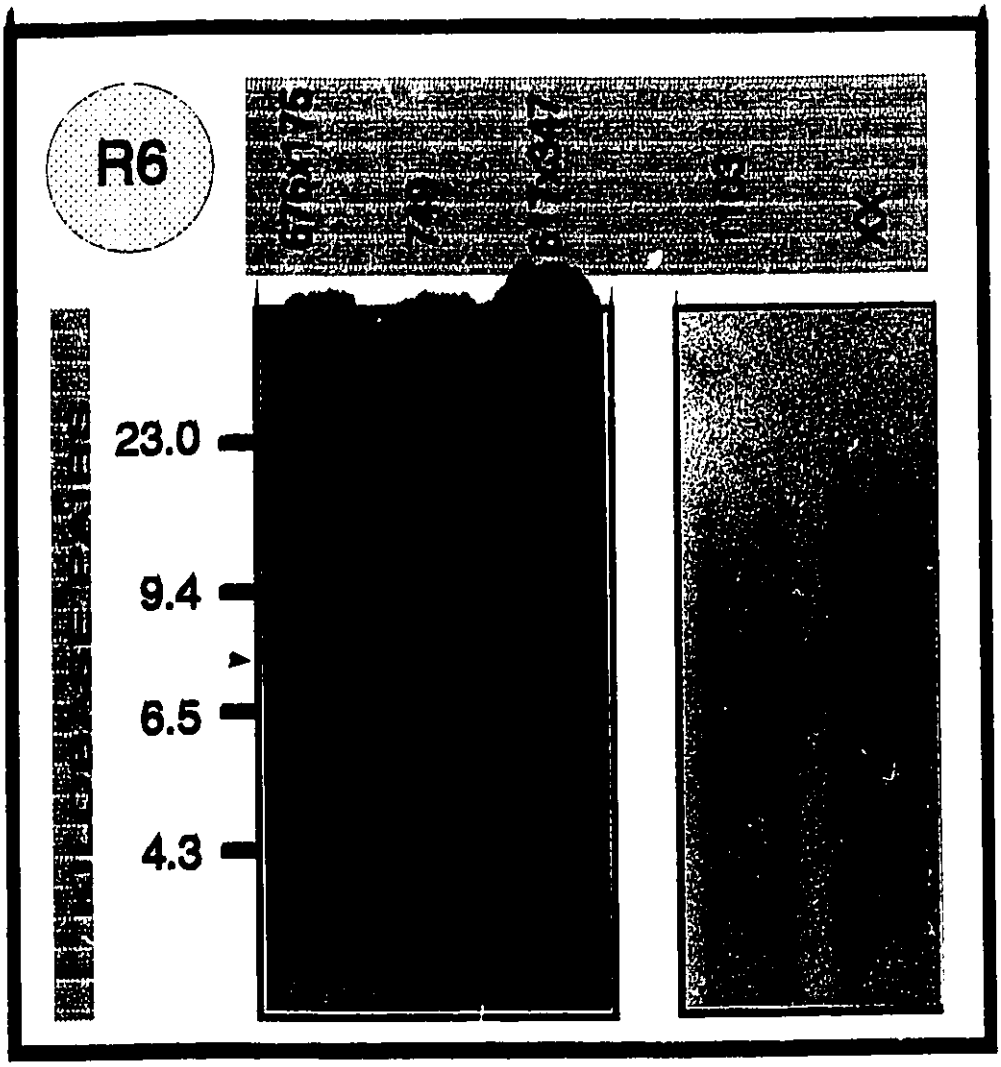

R6
23.0
9.4
6.5
4.3

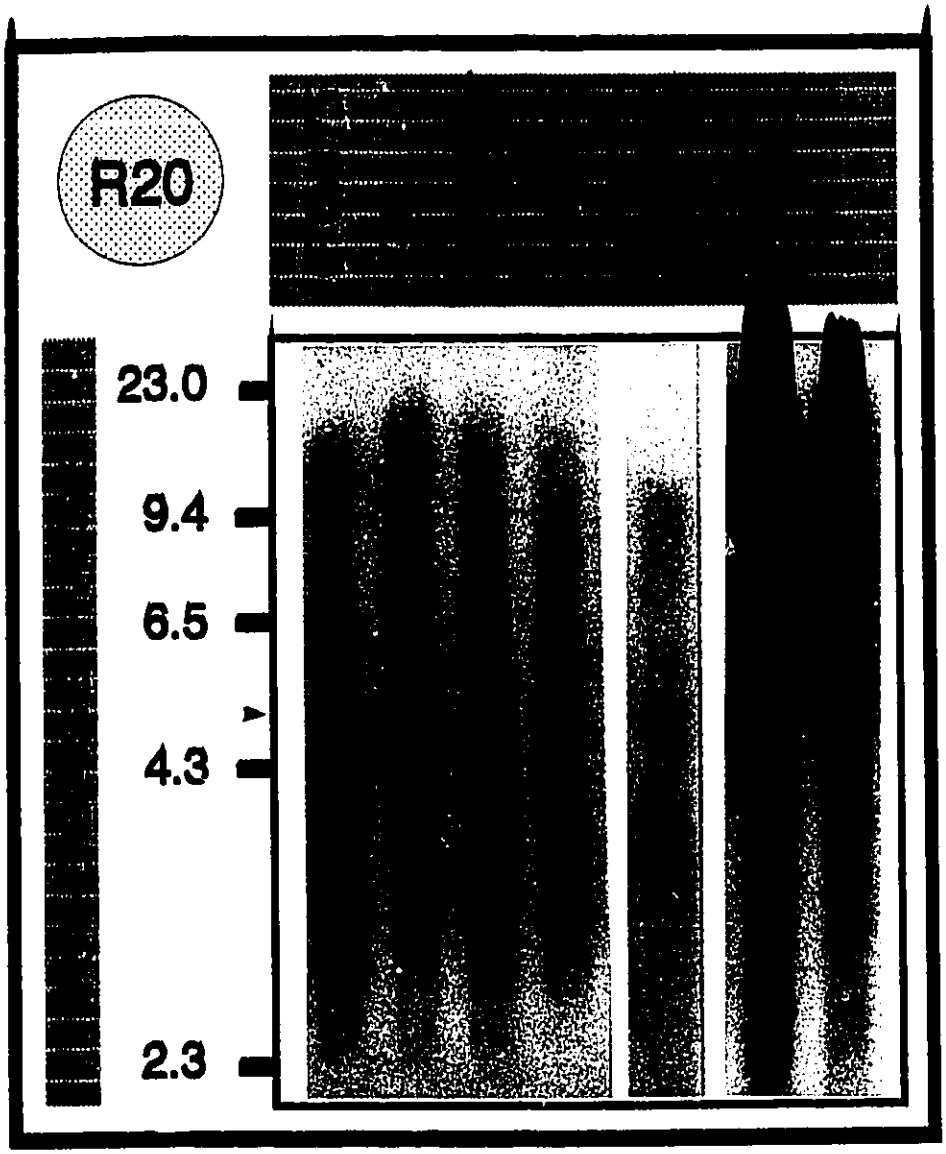
FIG. 38 MAPPING OF CLONE R20

DNA (5 μg) isolated from chromosome X-specific somatic cell hybrid cell lines was digested with Eco RI and subjected to Southern analysis to generate a human X chromosome specific DNA panel. The panel was probed with a 4.0 kbp Sal I/Eco RI fragment isolated from clone R20. The probe was labeled using the multiprime labeling method and was preassociated before use. Prehybridization, hybridization, wash and autoradiographic conditions were as outlined in Material and Methods.

The probe hybridizes to a human 5.3 kbp fragment (↗). The hybridization pattern observed is as follows:

676x175	-
749	-
617x347	-
82082a	+
1103	-

where (+) indicates hybridization and (-) indicates the absence of hybridization.


R20
23.0
9.4
6.5
4.3
2.3

per genome are not factors which affect the observed results. Perhaps a likely conclusion then is that the R20 clone corresponds to a ribosomal gene sequence, which are known to exist in multiple copies within the genome.

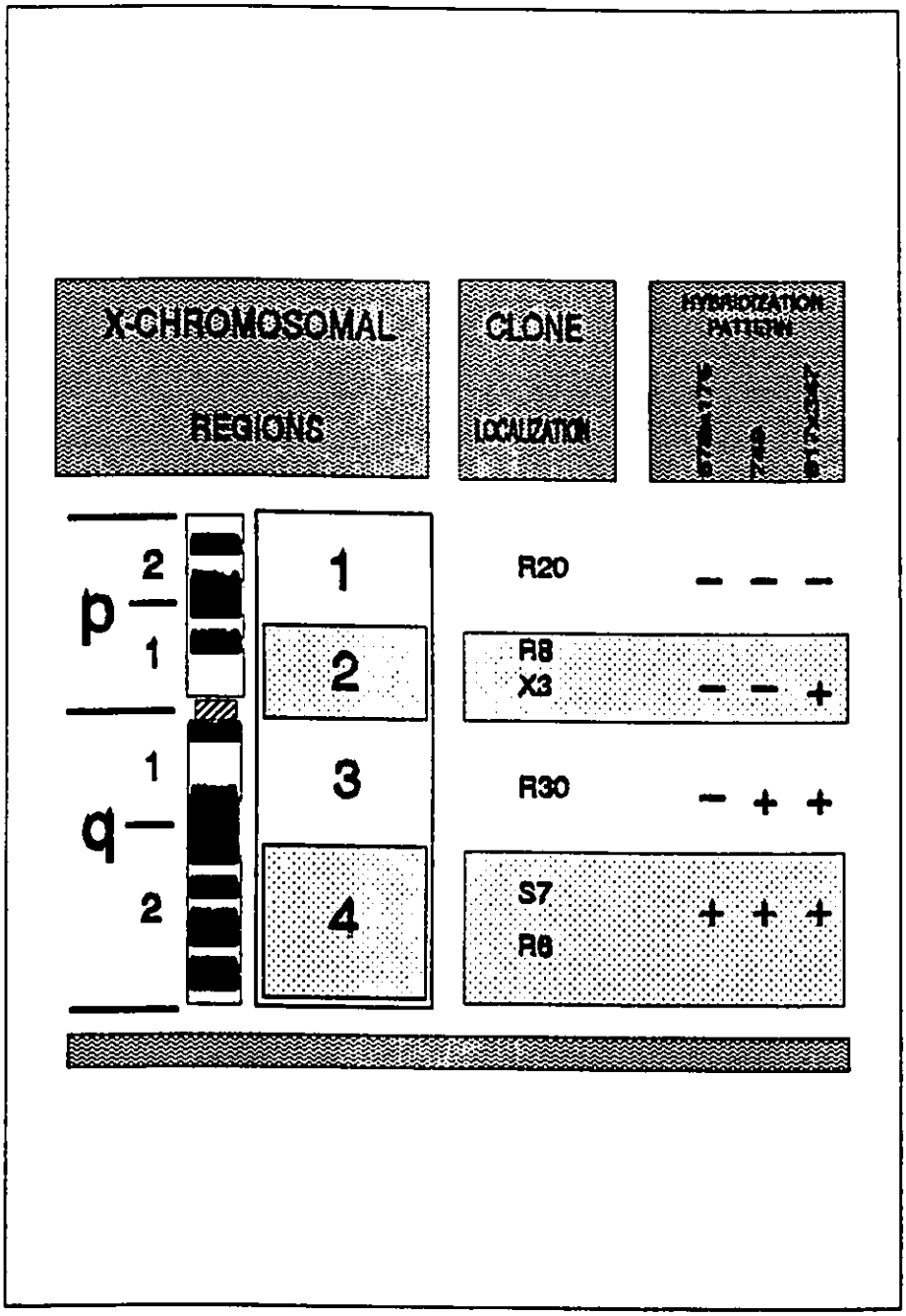
8.500) SUMMARY

In total six of the isolated clones were mapped by a somatic X specific cell hybrid panel (Fig. 39). In each case the clones demonstrated X-specificity and localized throughout the length of the X-chromosome suggesting that the genomic library used was representative of the X chromosome and that the isolation of clones was not biased to any specific region of the X-chromosome.

The mapping results also give an indication of the inherent background of false positives. One of the six clones (X3) initially thought to be of human origin was found to actually be of hamster origin. Another clone, R20, hybridizes to a complex network of bands and thus is repetitive in nature. A crude estimate of the background (false positives) based on this small survey of clones would suggest a level of approximately 1/3 or 33.3 %.

FIG. 39 REGIONAL LOCALIZATION OF P+ HU+ HA- GENOMIC CLONES TO THE X-CHROMOSOME BY A X-SPECIFIC HYBRID DNA PANEL, A SUMMARY.

(-) : No hybridization to the respective hybrid cell line DNA.
(+) : hybridization to the respective hybrid cell line DNA.


X-CHROMOSOMAL
REGIONS
CLONE
LOCALIZATION
HYBRIDIZATION
PATTERN
p
q
2
1
1
2
1
2
3
4
R20
R8
X3
R30
S7
R6

CHAPTER NINE

EXPRESSION

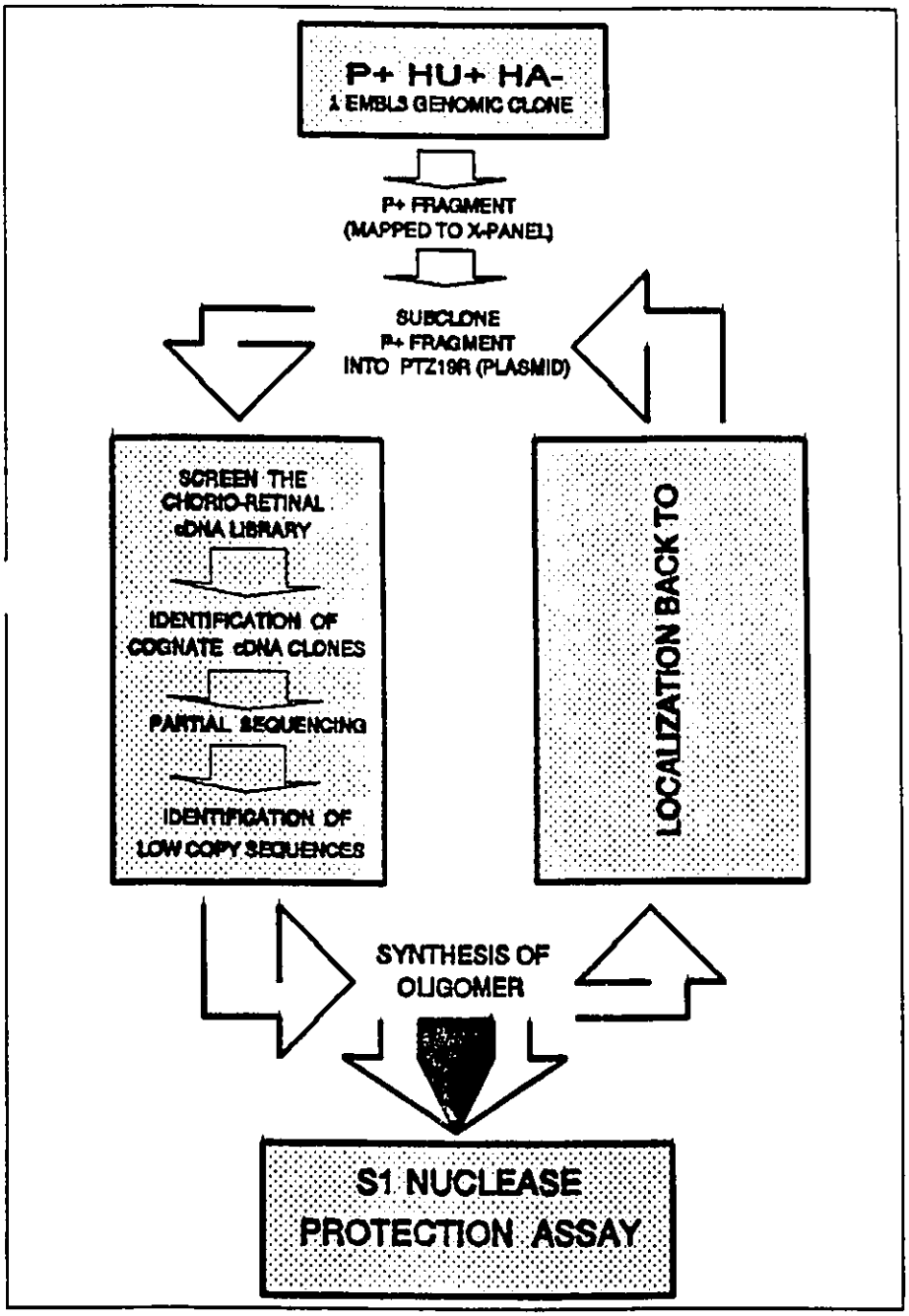
9.000) ALTERNATE STRATEGY

There are a number of well established techniques to examine whether or not a specific genomic fragment is likely to contain an expressed sequence. These include the screening of genomic clones for HTF islands and the analysis of zooblots. In the present case the genomic clones isolated are already suspected to harbor some form of chorio-retinal expressed sequence because of the nature of the initial library screenings (the use of a PCR amplified chorio-retinal cDNA library probe).

The most direct method to verify whether or not a given genomic fragment harbors an expressed sequence is to use that genomic fragment to probe Northern blots of RNA derived from the relevant tissue. Unfortunately this particular approach could not be used in this study because of the limited availability of human chorio-retinal RNA. An alternative strategy was devised to establish that the clones isolated did harbor expressed sequences.

The proposed strategy of analysis is summarized in Fig. 40. In brief, a specific fragment derived from the genomic clone is used to identify the cognate chorio-retinal cDNA clone. These cDNA clones are partially characterized by sequencing in order to search for unique stretches of nucleotide sequences. Oligonucleotides are then synthesized based on the stretches of unique sequences and used for S1 nuclease protection to establish whether these sequences are represented in the chorio-retinal RNA pool. Specific oligonucleotides which are indicative of an expressed sequence

FIG. 40 ALTERNATIVE SCHEME TO DEMONSTRATE THAT THE P+ HU+ HA- GENOMIC CLONES HARBOR EXPRESSED SEQUENCES.


P+ HU+ HA-
λ EMBL3 GENOMIC CLONE
P+ FRAGMENT
(MAPPED TO X-PANEL)
SUBCLONE
P+ FRAGMENT
INTO PTZ19R (PLASMID)
SCREEN THE
CHORIO-RETINAL
cDNA LIBRARY
IDENTIFICATION OF
COGNATE cDNA CLONES
PARTIAL SEQUENCING
IDENTIFICATION OF
LOW COPY SEQUENCES
LOCALIZATION BACK TO
SYNTHESIS OF
OLIGOMER
S1 NUCLEASE
PROTECTION ASSAY

are subsequently mapped back to the specific genomic clone to confirm the specificity of the sequence. These studies have been limited to the analysis of one of the genomic clones, R8. R8 was chosen primarily because of its map location, which placed it in a region of the X chromosome to which number of X-linked retinopathies have already been mapped.

9.100) CROSS SCREENING OF cDNA LIBRARY

The PTZ19R subclone of R8SF (R8(4)) was used to screen a portion of an unamplified chorio-retinal cDNA library. Clones that demonstrated positive hybridization were isolated and plaque purified. In total 7 independent R8 cDNA clones (cR8.P8.1, cR8.P8.2, cR8.P8.3, cR83.3, cR8.1.117, cR8.1.116, cR8SG21) were isolated from a screening of approximately 3000 cDNA clones. The frequency of R8 cross hybridizing cDNA clones in the library was calculated to be 0.2% which is 6 times less than the frequency of cDNA clones which cross hybridize to total human genomic DNA (1.2%). Screening duplicate lifts of the cDNA library with the PTZ19R subclone of R6SF (R6(8)) demonstrated that the two specific probes hybridized in a distinct manner, the majority of clones which hybridized with R8(4) did not hybridize with R6(8) and vice versa (Fig. 41).

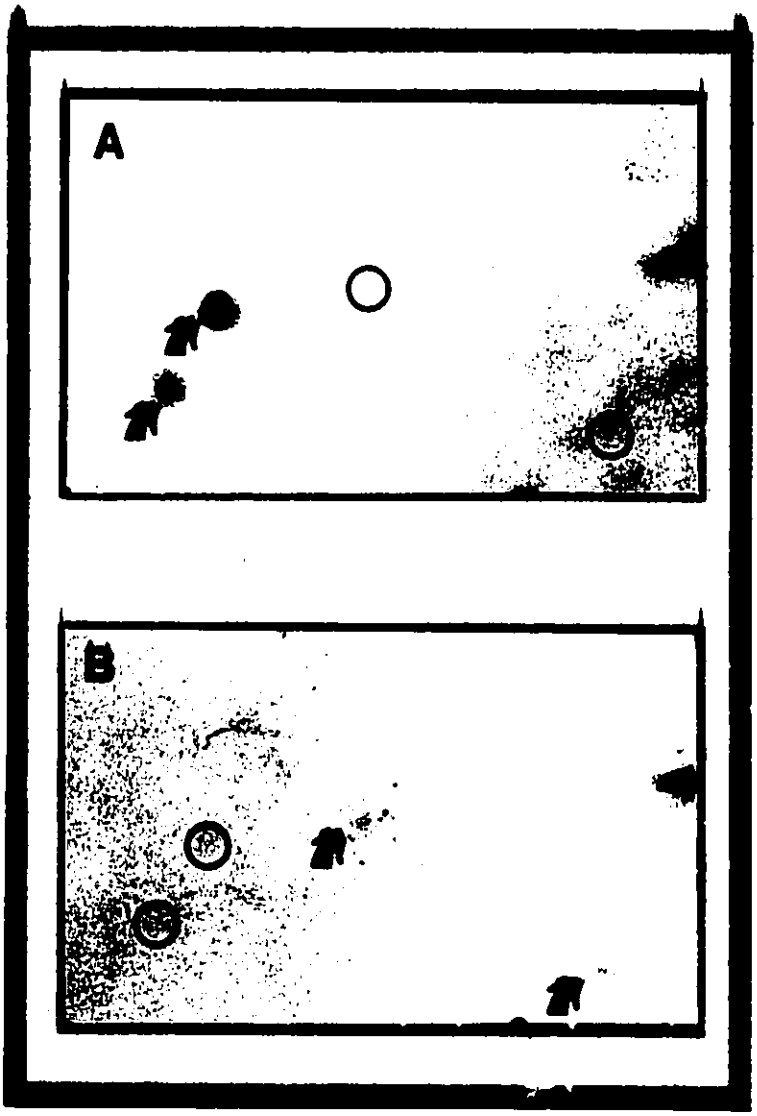
9.200) INITIAL ANALYSIS OF R8 cDNA CLONE SEQUENCES

The sequencing of the cDNA clones were performed as outlined in Methods. Single stranded DNA templates for sequencing were generated using one sided PCR

FIG. 41 ISOLATION OF R8 SPECIFIC HUMAN RETINAL cDNA CLONES.

A) Plaque lift probed with radiolabeled R8(4).
B) Duplicate plaque lift probed with radiolabeled R6(8).

↗: Indicates a plaque which shows hybridization with the respective probe.
O: Indicates a plaque which does not hybridize with the designated probe but does with the other probe.


A
B

of the original λgt10 cDNA clones (cR8.p8.1, cR8.p8.2, cR8.3.3) and by inducing single stranded synthesis from PTZ19R subclones of cDNA inserts isolated from the original λgt10 clones (cR8.1.116, cR8.1.117, cR8.01.04, cR8.SG.21).

Analysis of the sequences generated from each clone allowed the cDNA clones to be placed into 3 distinct groups, group 1(cR8.p8.1, cR8.p8.2, cR8.01.04, cR8.3.3), group 2(cR8.SG.21) and group 3 (cR8.1.116, cR8.1.117). In the case of group 1, the clones contained sufficient overlap and identity to generate a single consensus sequence. The group 1 consensus sequence and the sequences obtained for the remaining three clones is provided in Fig. 42.

All clones, with the exception of cR8.1.116 contained long stretches of sequence showing homologies to known Alu repetitive sequences. In the case of cR8.SG.21 this homology is specifically to the human Alu family interspersed repeat element, BLUR6. The remaining clones showed distinct homologies to the average human Alu family interspersed sequence. The isolation of a group of cDNA clones that share a common feature of repetitive sequences is of concern with respect to the specificity to which these clones have been isolated. However, if short non-repetitive sequences could be identified within these sequences then they could be used to verify the expressivity of the clone and to determine its association with the genomic clone used to isolate it. In both cR8.1.116 and cR8.1.117 there are regions of sequence which were found to be non-repetitive in nature.

FIG. 42 cR8 SEQUENCES

GATC : Sequences with a degree of homologies to known repetitive sequences (>85% homology).
GATC : Low copy (non-repetitive) sequences.

cR8.01-04 SEQUENCE

```
        10         20         30         40         50         60
GAATTCCGGA AATTGAGACC ATCCTGGTTA ACACGGTGAA ATCCCATCTC TACTAAAAAT
CTTAAGGCCT TTAACTCTGG TAGGACCAAT TGTGCCACTT TAGGGTAGAG ATGATTTTTA

        70         80         90        100        110        120
ACAAAAAATT AGCCAGGCAT GGTGGCAGGT GCTGTAGTCC AGCTACTCGG GAGGCTGAGG
TGTTTTTTAA TCGGTCCGTA CCACCGTCCA CGACATCAGG TCGATGAGCC CTCCGACTCC

       130        140        150        160        170        180
CAGGAGAATG GCGTGAACTC GGGAGGTGGA GCTTCGAGTG AGCCGAGATT GCACCACTGC
GTCCTCTTAC CGCACTTGAG CCCTCCACCT CGAAGCTCAC TCGGCTCTAA CGTGGTGACG

       190        200        210        220        230        240
ACTCCAGCCT GGATGACAGA GTGAGACTCT GTCTCAAAAA AAAAAAAAAA AAAAAAAAAA
TGAGGTCGGA CCTACTGTCT CACTCTGAGA CAGAGTTTTT TTTTTTTTTT TTTTTTTTTT

AAAGGATT
TTTCCTAA
```

cR8.8G 21 SEQUENCE

```
        10         20         30         40         50         60
CTCTCCGGAT CCAAGCTTAT CGATTTCGAA CCCGGGGTAC CGAATTCCCT CACACCTGTA
GAGAGGCCTA GGTTCGAATA GCTAAAGCTT GGGCCCCATG GCTTAAGGGA GTGTGGACAT

        70         80
ATCCCAGCAC TTTGGGAGGC TGAG
TAGGGTCGTG AAACCCTCCG ACTC
```

cR8.1.116 SEQUENCE

```
         10          20          30          40          50          60
PLASMID-GA ATTCCCCCTC GGCACCAGCT GCATATGTGT TCACTTTTAT TTAAATAAAC
PLTSMID-CT TAAGGGGGAG CCGTGGTCGA CGTATACACA AGTGAAAATA AATTTATTTG

         70          80          90         100         110         120
TTGTGTGGTA AAAGTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AACACACCAT TTTCATTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

        130         140         150
AGATCCTCAC ACTTGATCAG CACTGGAGCT AG
TCTAGGAGTG TGAACTAGTC GTGACCTCGA TC
```

cR8.1.117 SEQUENCE

```
         10          20          30          40          50          60
CTTTTTTTTT TTTTTTTTTT GAGACAGAGT CTTGCTCTGT TGCTGAGGCT GGAGTGCAGT
GAAAAAAAAA AAAAAAAAAA CTCTGTCTCA GAACGAGACA ACGACTCCGA CCTCACGTCA

         70          80          90         100         110         120
GGTGTGATCT TGGCTCACTG AAATCTCCGC CTCCGGGTTC AAGTGATTCT CCTGCCTCAG
CCACACTAGA ACCGAGTGAC TTTAGAGGCG GAGGCCCAAG TTCACTAAGA GGACGGAGTC

        130         140         150         160         170         180
CCTCCCGAGT AGCTGGGATT ACAGGCATGC ACCATCATAC TCAGCTAATT TTTGTATTTT
GGAGGGCTCA TCGACCCTAA TGTCCGTACG TGGTAGTATG AGTCGATTAA AAACATAAAA

        190         200         210         220         230         240
TAGTAGAGAC AGGTTCACAT ATGCAGCTGT CTGACTCGAC TCTAGTGATC ACCACTGCTC
ATCATCTCTG TCCAAGTGTA TACGTCGACA GACTGAGCTG AGATCACTAG TGGTGACGAG

        250         260         270         280         290         300
AAAATGCTAA GATTACAGCT AGAGCACCAC GCGATCTTTT TTTTTTTTTT TTTTTTTTTT
TTTTACGATT CTAATGTCGA TCTCGTGGTG CGCTAGAAAA AAAAAAAAAA AAAAAAAAAA

        310         320         330         340         350         360
TTTGAGACAG AGTCTGCTCT GTTGTCGAGG TCGGAGTGCA GTGGTGTGAT CTGGCTCACT
AAACTCTGTC TCAGACGAGA CAACAGCTCC AGCCTCACGT CACCACACTA GACCGAGTGA

        370         380         390
GAATCTCGCT CGGTTCAGTG ATCTCGCTCA
CTTAGAGCGA GCCAAGTCAC TAGAGCGAGT
```

TABLE 11 SUMMARY OF OLIGOMER SEQUENCES SYNTHESIZED

R	: Reverse strand
GATC	: Sequence derived from clone
<u>GATC</u>	: Mismatch sequence

LOCATION	Oligomer	SEQUENCE
cR8.1.116 120-148	A	AAGATCCTCA CACTTGATCA GCACTGGAGG TTTCGAACGG
151-121	AR	TAGCTCCAGT GCTGATCAAG TGTGAGGATC CACTATACCT A
cR8.1.116 41-74	J	TGTTCACTTT TATTTAAATA AACTTGTGTG GTAAAAGAGT GTTCACA
77-41	JR	TTACTTTTAC CACACAAGTT TATTTAAATA AAAGTGATAG TTCACT
cR8.1.117 196-238	H	CACATATGCA GCTGTCTGAC TCGACTCTAG TGATCACCAC TGCAAATAGC
235-194	HR	GTGGTGATCA CTAGAGTCGA GTCAGACAGC TGCATATGTG AATACATCAT
cR8.1.117 348-387	I	GATCTGGCTC ACTGAATCTC GCTCGGTTCA GTGATCTCGC ACGAGTTAGT AG
390-349	IR	TGAGCGAGAT CACTGAACCG AGCGAGATTC AGTGAGCCAG ATGGTCTGCT AC

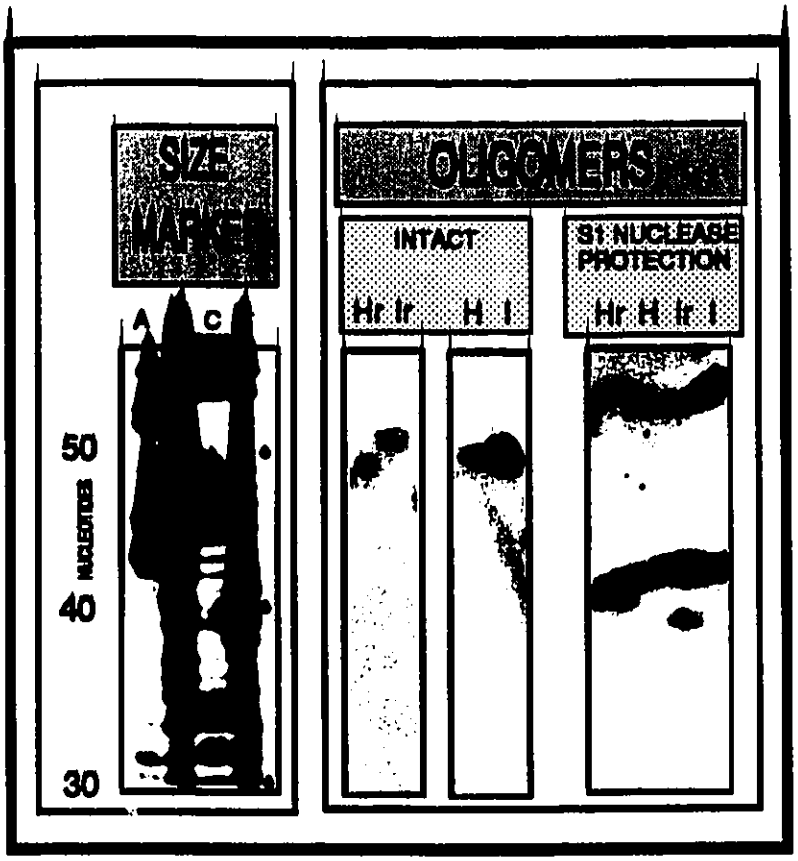
9.300) S1 NUCLEASE PROTECTION ASSAY

Based on the non-repetitive sequences defined above, synthetic oligonucleotides were made for use in the S1 nuclease protection assay (Table 11). Two oligonucleotides were made for each non-repetitive sequence, one representing the forward strand and the second representing the reverse nucleotide strand (r). In each case a 8-12 nucleotide extension, a stretch of sequence which is not present in the cDNA sequence, was added to the 3' ends of each oligomer to serve as an internal control to monitor the efficiency of the S1 digestion. The oligonucleotide sequences were 5' end labeled, hybridized to human chorio-retinal RNA, and subjected to S1 nuclease digestion.

The results of oligomers that demonstrated full S1 nuclease protection are shown in Fig. 43. Oligomers Hr and Ir both showed evidence of RNA protection with the appropriate decrease in band size. In each case where a specific oligomer showed evidence of RNA protection the complementary oligonucleotide did not. Oligo-A and oligo-Ar did not show evidence of RNA protection (not shown). Oligo-J also did not show evidence of RNA protection but Jr did show a partial protection in that there is a shift of approximately 7 nucleotides instead of the anticipated 9 nucleotides (not shown). The S1 nuclease protection results demonstrate that cR8.1.117 is part of an expressed gene sequence in the human choriod/retina and that the H and I sequences are derived from the sense strands. The expressivity of cR8.1.116 is not as clear cut although the results of the Jr oligo suggest that it to protects a specific RNA in the chorio-retinal expressed sequence pool.

FIG. 43 RESULTS FOR OLIGOMERS WHICH SHOWED RNA PROTECTION DURING THE S1 NUCLEASE PROTECTION ASSAY.

Two non-repetitive oligonucleotide sequences derived from clone cR8.1.117 were found to demonstrate chorio-retinal RNA protection. In each case the forward strands (H and I) did not show RNA protection. However, the corresponding reverse strand sequences (Hr and Ir) showed the complete loss of the mismatch region of the oligonucleotide sequence, thus resulting in a decreased band size.


SIZE MARKERS
A
C
50
40
30
NUCLEOTIDES
OLIGOMERS
INTACT
Hr Ir
H I
S1 NUCLEASE PROTECTION
Hr H Ir I

9.400) LOCALIZATION OF cR8 OLIGOMERS

Because of the presence of repetitive sequences in the original R8 cDNA clones it was not possible to use them as specific probes to map back to a specific genomic clone. Instead, the oligonucleotides synthesized for the S1 analysis, which represent a cDNA non-repetitive match sequence as well as an artificial mismatch region, were used in an attempt to define the genomic specificity of the individual cDNA clone. Because the oligonucleotide sequences are relatively short and because there is at most a 12 nucleotide mismatch to each of the probe sequences, a modified set of hybridization conditions was used (Fig. 44). The oligonucleotide probes (Hr and Ir) were used to screen the R8(4) genomic subclone and the R6(8) genomic subclone, which was used as a control. Hr and Ir showed hybridization to the R8(4) subclone but not to the R6(8) subclone, suggesting that the cDNA sequences are R8(4) specific. However, because of the decreased stringency of hybridization the possibility of non-specific hybridization cannot be precluded.

9.500) FURTHER SEQUENCE ANALYSIS

9.510) cR8.1.117 sequence analysis

In the initial screen of the data bank cR8.1.117 showed a 94.4% homology in a 15 nucleotide stretch with the anti-sense strand of the human dystrophin gene. This initial observation was interesting because dystrophin is known to be located on the short arm of the X chromosome just above the region to which clone R8 has been localized.

FIG. 44 LOCALIZATION OF OLIGOMERS DERIVED FROM cDNA SEQUENCE TO R8(4) AND R6(8) GENOMIC SUBCLONES.

Lane 1 : R6(8) has a 8.0 kbp human genomic insert.
Lane 2 : R8(4) has a 4.5 kbp human genomic insert.

Prehybridization and hybridization were performed as defined in materials and methods except that 0.5xSSC was used in the prehybidization and hybridization solutions and that hybridization was carried out at 39°C. Filter washes consisted of one fifteen min wash with 2xSSc at 39°C, followed by 4x 30 min washes with 2xSSC at room temperature, and a final 30 min wash in 1xSSC. 0.1% SDS at room temperature.

C oligo is a repetitive sequence oligo derived from a R6 cDNA sequence. This sequence does not overlap with any of the known R8 sequences.

G oligo is a repetitive oligo derived from the cR8 group 1 sequence.

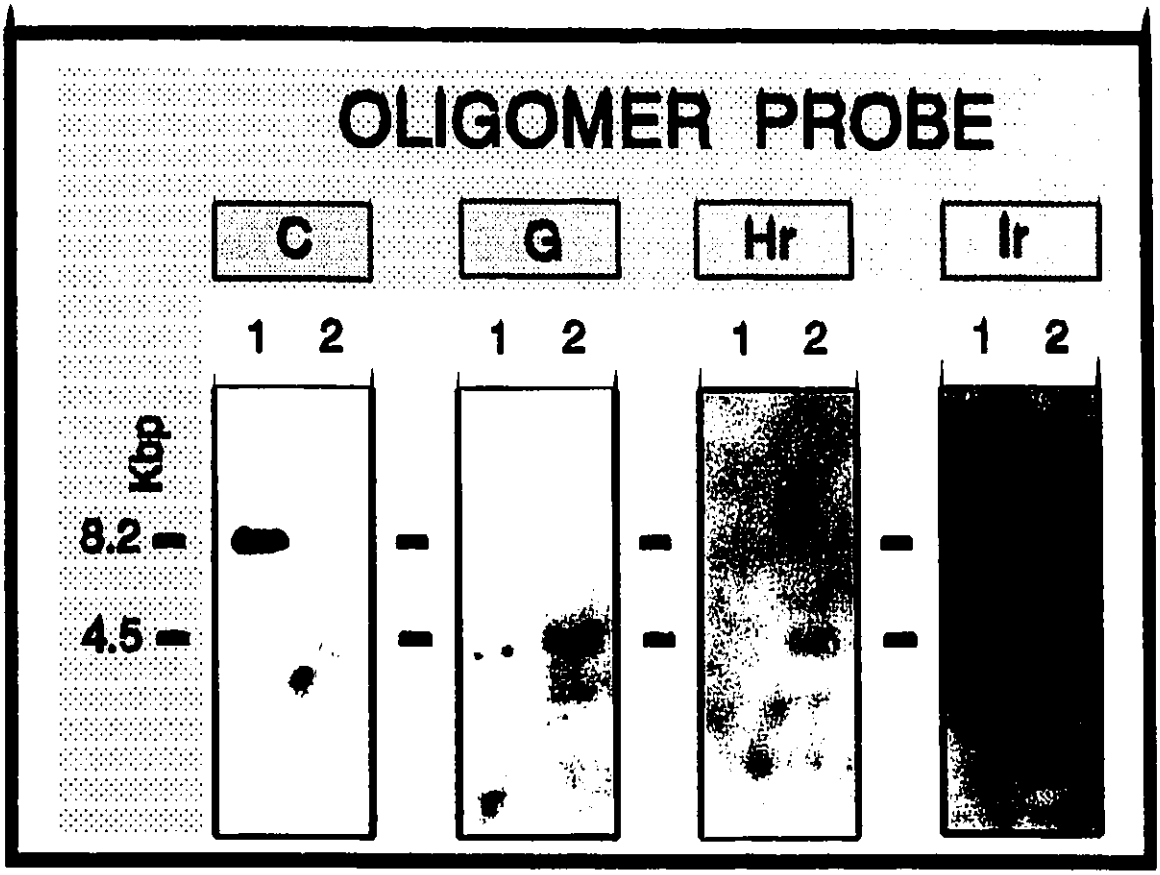
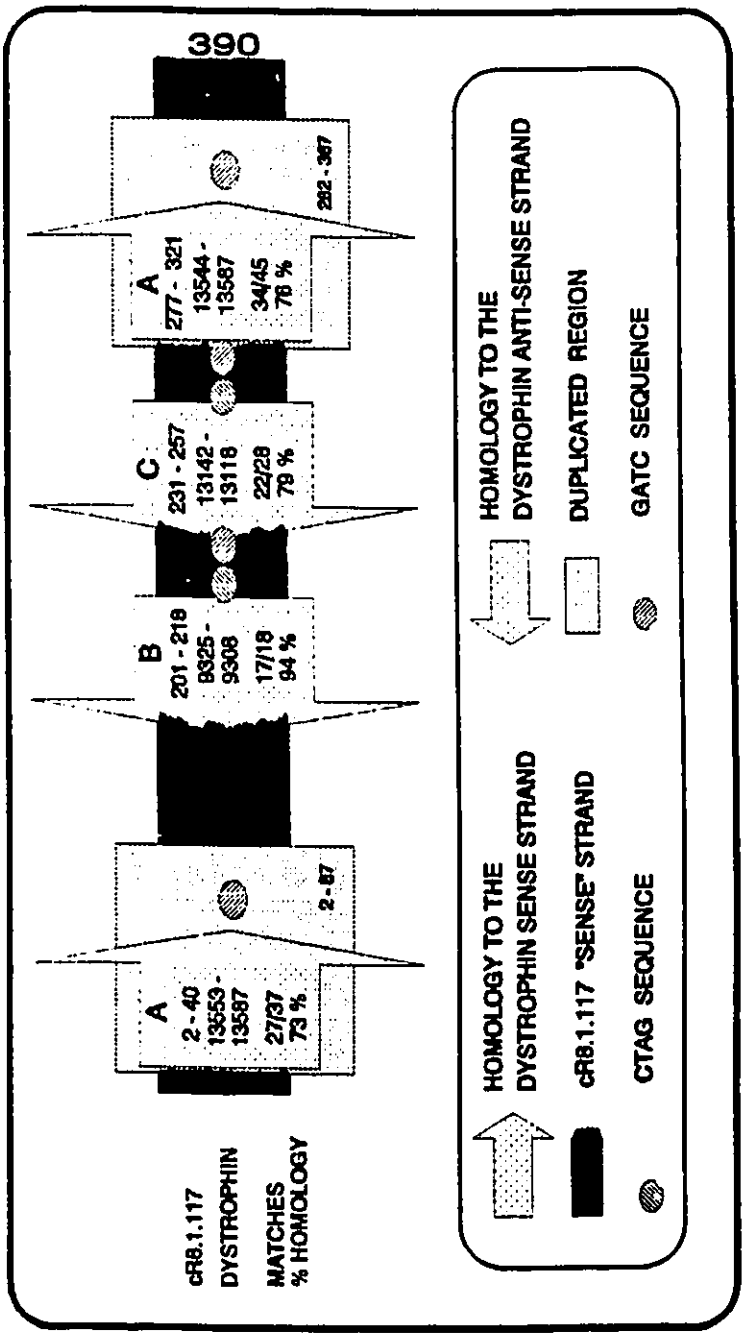

OLIGOMER PROBE
C
G
Hr
Ir
1 2
1 2
1 2
1 2
Kbp
8.2
4.5

FIG. 45 A COMPARISON OF THE cR8.1.117 SEQUENCE TO THE mRNA SEQUENCE OF HUMAN DYSTROPHIN.


390
cR8.1.117
DYSTROPHIN
MATCHES
% HOMOLOGY
A
2 - 40
13553 -
13587
27/37
73 %
2 - 67
B
201 - 218
8325 -
9308
17/18
94 %
C
231 - 257
13142 -
13116
22/28
79 %
A
277 - 321
13544 -
13587
34/45
76 %
282 - 367
HOMOLOGY TO THE
DYSTROPHIN SENSE STRAND
cR8.1.117 "SENSE" STRAND
CTAG SEQUENCE
HOMOLOGY TO THE
DYSTROPHIN ANTI-SENSE STRAND
DUPLICATED REGION
GATC SEQUENCE

Additional homology screens specifically between cR8.1.117 and the human dystrophin gene have revealed additional regions which show homology to the dystrophin gene (Fig. 45). The observed homology regions (A, B, C, A) correspond to the regions spanning nucleotides 13544-13587, 9325-9308, 13142-13118, 13544-13587, respectively, of the dystrophin cDNA sequence. What becomes apparent from this analysis is that R8.1.117 shows a high degree of homology to dystrophin but, because the identity is not perfect, is not dystrophin. It does however appear to consist of cassettes of sequences that may have been derived from the dystrophin gene. Also noted in cR8.1.117 is the short sequence CTAG and its mirror image GATC which appears just before and just after a dystrophin homology region. In cR8.1.117, a pair of these 4-mers are present in stretches bridging 2 dystrophin homology regions. The significance of these sequences, if any, is not known.

The cR8.1.117 coding strand was determined by the S1 protection results. Both the Hr and Ir oligonucleotides are derived from the lower strand and both show the anticipated pattern of protection, thus the coding strand must be the upper strand. As such the proposed coding strand has been translated in all three possible reading frames, one of which demonstrates a long open reading frame (Fig. 46). In this case there is a clear open reading frame from nucleotide position 1 to 302, ending with a stop codon at position 303.

FIG. 46 COMPUTER TRANSLATION OF THE cR8.1.117 SEQUENCE IN THE THIRD READING FRAME.

End TGA (bolded)

```
        10         20         30         40         50         60
CTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTTGCTGAGGCTGGAGTGCAGT
  PhePhePhePhePhePheGluThrGluSerCysSerValAlaGluAlaGlyValGlnT

        70         80         90        100        110        120
GGTGTGATCTTGGCTCACTGAAATCTCCGCCTCCGGGTTCAAGTGATTCTCCTGCCTCAG
rpCysAspLeuGlySerLeuLysSerProProProGlySerSerAspSerProAlaSerA

       130        140        150        160        170        180
CCTCCCGAGTAGCTGGGATTACAGGCATGCACCATCATACTCAGCTAATTTTTGTATTTT
laSerArgValAlaGlyIleThrGlyMetHisHisHisThrGlnLeuIlePheValPheL

       190        200        210        220        230        240
TAGTAGAGACAGGTTCACATATGCAGCTGTCTGACTCGACTCTAGTGATCACCACTGCTC
euValGluThrGlySerHisMetGlnLeuSerAspSerThrLeuValIleThrThrAlaG

       250        260        270        280        290        300
AAAATGCTAAGATTACAGCTAGAGCACCACGCGATCTTTTTTTTTTTTTTTTTTTTTTTT
lnAsnAlaLysIleThrAlaArgAlaProArgAspLeuPhePhePhePhePhePhePheP

       310        320        330        340        350        360
TTTGAGACAGAGTCTGCTCTGTTGTCGAGGTCGGAGTGCAGTGGTGTGATCTGGCTCACT
heEndAspArgValCysSerValValGluValGlyValGlnTrpCysAspLeuAlaHisE

       370        380        390
GAATCTCGCTCGGTTCAGTGATCTCGCTCA
ndIleSerLeuGlySerValIleSerLeu
```

9.520) cR8.1.116 sequence analysis

The cR8.1.116 sequence does not possess any long stretches of homology to the Alu consensus sequence. However, there are a number of small stretches (10-15 nucleotides) which show homology to some parts of the Alu consensus sequence. Between these small Alu runs there exists a stretch of sequence (nucleotide 40-70) that represents a unique nucleotide stretch. In a data base search this non-repetitive stretch of cR8.1.116 did not show absolute identity to other sequences in the Genbank database, however a number of homologies (>85%) to a number of known gene sequences were found (Table 12). In all cases homologies seem to be to genes which code for extracellular proteins.

The initial search of the data bank with non-repetitive regions of cR8.1.116 did not identify any long regions of homology to the human dystrophin gene. However in the visual analysis of the cR8.1.116 sequence data it was noted that there is a short 9 nucleotide homology region between the 3'end of the upper cR8.1.117 strand and the 5'end of the lower cR8.1.116 strand. Because of this observation a homology analysis between cR8.1.116 and dystrophin was performed.

The results of the homology search revealed that there exists many 10-15 nucleotide stretches which showed homology of greater than 80% to dystrophin sequences (Table 13). Moreover these cR8.1.116 sequences appear in multiple copies within the dystrophin sequence. All possible homology regions appear to fall within

TABLE 12 SUMMARY OF HOMOLOGIES (>85%) BETWEEN cR8.1.116 (NUCLEOTIDES 40-75) AND KNOWN GENE SEQUENCES.

cR81.116 HOMOLOGY REGION	STRETCHES OF HOMOLOGY (> 85%) TO	REGION	MATCHES	NUCLEOTIDES	% HOMOLOGY
45-59	Human bone marrow serine protease gene mRNA	3611-3625	15	15	100
41-74	Human complement protein C8 beta subunit mRNA *	1984-1950	27	35	77.1
47-65	Human CR2/CD21/C3d/Epstein-Barr virus receptor mRNA	3784-3766	17	19	89.5
47-65	Human Epstein-Barr virus complement receptor type II	3941-3923	17	19	89.5
42-63	Human platelet glycoprotein IIIa, exon 12	913-892	20	22	90.9
42-60	Human T-cell receptor active alpha-chain V-region (V-J-C) mRNA	286-268	17	19	89.5

* Included because it covers the full cR8.1.116 region considered.

TABLE 13 SUMMARY OF HOMOLOGIES (>80%) BETWEEN cR8.1.116 (NUCLEOTIDES 40-75) AND THE HUMAN DYSTROPHIN GENE (mRNA SEQUENCE).

cR81.116 HOMOLOGY REGION	DYSTROPHIN HOMOLOGY REGIONS (>80 %)	MATCHES	NUCLEOTIDES	% HOMOLOGY
43-55	13503-13489	12	15	80
45-57	13617-13605	11	13	85
47-58	13172-13161	10	12	83
47-62	13501-13486	13	16	81
50-62	13514-13500	12	15	80
52-66	13090-13075	13	16	81
52-66	13125-13110	13	16	81
59-48	13616-13605	10	12	83
59-48	13940-13929	10	12	83
60-48	12907-12895	11	13	85
64-52	13548-13536	11	13	85
73-59	921-936	13	16	81
75-66	161-170	10	10	100

the first kb or the last 2 kb of the human dystrophin cDNA sequence. The majority of the homologies with greater than 80% homology are limited to the far 3' end of the dystrophin sequence. These findings together with the identification of one poly adenylation motif in each strand (54-59; 55-50) of the cR8.1.116 sequence suggest that it might be derived from the 3' end of a gene.

9.600) FUTURE STUDIES

Since the mapping of R8 to Xp11.4-Xq11 it has received considerable interest. This interest comes primarily from the fact that a number of X-linked eye disorder genes, including constant stationary night blindness, Norrie's disease, retinitis pigmentosa 2 and retinitis pigmentosa 3, have also been localized to the vicinity of this X region.

R8 has recently been registered with the Human Gene Mapping organization as X.EH.8 (DXS542). At the present time two additional studies have been planned for R8. The first line of study therefore is to test R8 as a candidate gene for these known eye disorders. A preliminary study, in collaboration with Dr. T. Bech-Hansen (Calgary), to examine R8 as a candidate gene for the genetic eye disorder congenital stationary night blindness (CSNB) has already commenced. The second direction that the R8 research will take is to define DNA polymorphisms using R8 sequences. In this way R8 can be used in linkage and other genetic studies to more finely define

the R8 locus relative to known X markers and to be used in future studies to help define the map location of other unmapped loci.

9.700) SUMMARY

The data presented in this chapter is suggestive that R8 harbors some chorio-retinal expressed sequence (cR8.1.117). Additional experiments will be required to confirm this and are discussed in chapter ten. What is apparent however is that the potential usefulness of R8 (as discussed above) does not necessarily depend on a confirmation as to whether it does or does not harbor an expressed gene.

CHAPTER TEN

A CRITIQUE

10.000) LIBRARY TO LIBRARY CROSS SCREENING

In the present body of work a lateral approach has been taken to isolate genomic clones which should have a preponderance for harboring choroid/retina expressed sequences as well as being X-linked. The approach developed represents a simple method of performing library to library cross screening. In the present study, a cDNA library is used to screen a genomic library, however the general approach of library to library cross screening can be applied to any combination of library types (cDNA library to genomic library, cDNA library to cDNA library, and genomic library to cDNA library). The only requirements are that the size of the inserts in the library to be used as the probe does not exceed the limitations imposed by the PCR step and that the cloning sites of the two vectors used to prepare the respective libraries are different from each other. The advantages of this generalized approach include the development of a source of probe which is independent of tissue supply (which may be limited), the use of common vectors so that existing libraries can be used if desired, and in the present case an association between tissue expression and chromosome specificity. A discussion of the proposed lateral system in the light of the present data is presented below.

10.100) POTENTIAL LIMITATIONS

10.110) cDNA library probe

The issue of what the P probe should be capable of detecting has already been discussed in section 5.500. Briefly, sequences that exist at roughly 10-20 copies

in the initial cDNA library pool should provide significant levels of detection of cognate clones. If the identification of less abundant expressed sequences is desired then the P probe should be equalized and subsequently reamplified for use as a probe (Ko, 1990).

The use of a pool of expressed sequences as a mixed probe does not ensure the isolation of potential candidate clones for every given disease affecting a particular tissue. Theoretically it will only lead to the isolation of genes that are active in the tissue at the specific stage of differentiation at the time the RNA is isolated. Since the long term goal was to develop a means of identifying genes underlying adult degenerative disorders, the use of an adult choroid/retina cDNA library P probe was appropriate for the isolation of candidate clones. However, to identify genes underlying a disease that is manifested during development (such as Norrie's disease), it is necessary to use cDNA libraries from the appropriate developmental period. In the case of human research, this is difficult at the present time.

10.120) False negatives

In some cases it is conceivable that the actual size and distribution of introns and exons might contribute to the lack or faint hybridization of a given cDNA based sequence to its cognate genomic clone. This is especially the case if the exons of a given gene are relatively small and its introns are relatively large.

10.130) False positives

The number of P+ HU+ HA- genomic clones isolated (0.30 % of the clones screened) is higher than the number anticipated (0.03%). Based on the estimate made in section 7.300 (0.03%) it would appear that 90% of the clones isolated represent false positives. The nature of the isolation of these inherent background clones is not clearly understood. There are at least two possible explanations for this (1. the estimated expected frequency of positive clones; 2. the nature of hybridization of the P probe).

10.131) Estimation of false positives

The estimation of the expected number of positive clones is based on rather vague estimates of the size of the human genome, on the total number of genes in the genome, and the total number of genes that might be active in a given cell (see section 7.300). At the present time there is reason to believe that the human genome may in fact harbor more than a 100,000 genes (Dr. A. Mackenzie, personal communication). If this is true then the estimation of expected clones made here based on the previously accepted information would represent a underestimation. As a consequence, the number of P+ HU+ HA- clones observed may in fact not represent a 10 fold higher value than should be expected, and may represent a relatively accurate estimate.

10.132) The nature of hybridization

It appears from the initial cross screening data (chapter 6) that the P probe hybridized in a semi-independent manner relative to total human genomic DNA. 0.41 % of total genomic clones screened were found to be positive with the P probe in a primary screen and positive with a total human genomic probe in a secondary screen. In a separate experiment, 0.64% of total genomic clones screened were positive with the human genomic probe in a primary screen (Table 6). If the P probe were clearly a human genomic repetitive probe, the same frequency of positive clones should have been observed in both cases.

The data presented in the characterization of independent clones (chapter 7) also supports this view. The hybridization of the P probe and the human genomic probe to individual genomic clones (Fig. 27) indicates that the P probe hybridizes to very specific bands in the restriction profile as opposed to the human genomic probe which hybridizes to all bands. The interpretation of this data is that the two probes are distinct, each hybridizing to a distinct class of sequences. Subsequently the screening of the choroid/retina cDNA library with a human genomic probe yielded very few positive clones (approximately 1.4%), indicating that only a few of the cDNA clones in the library possess sequences which are highly repetitive in the genome (section 7.510). Taken together, these findings suggest that the sequences that are abundant in the genome and present in the choroid/retinal cDNA library do not play a major role in the hybridization of the P probe to specific genomic clones.

In addition, on rescreening the P+ HU+ HA- clones with a P probe that had been depleted for highly repetitive sequences (chapter 7). 50 % of the clones showed a decrease in relative hybridization intensity but none showed an absolute absence of hybridization. This indicates that the clones had not been identified on the basis of sequences which were highly repetitive in the genome.

The data obtained support the hypothesis that the isolation of genomic clones with the P probe is not achieved on the basis of sequences that are highly repetitive in the genome. In retrospect, however, this does not preclude sequences that are slightly repetitive in the genome. Sequences that appear at a high frequency in the P probe, due to high levels of expression of specific sequences, but are present at lower copy numbers in the genome may contribute to the isolation of false positive clones. The level of these types of sequences in the genome would not be sufficient to deplete the P probe. It would be predicted that these sequences would exist at high enough levels in the P probe to detect specific clones but exist at levels in the genome insufficient to detect the same clones.

If the major factor in the isolation of false positive clones are repetitive sequences then it would be advantageous to use a P probe derived from a cDNA library constructed using random primers rather than with oligo(dT) primers since expressed repetitive sequences are found predominantly in the 3' untranslated regions of mRNA sequences. This would ensure that the sequences used to generate the P

probe would not be biased to the 3' end of mRNA and would therefore possess a lower proportion of repetitive sequences.

10.200) MAPPING STUDIES

Since the human somatic cell line 82082a has a single X-chromosome as its only human component, the human clones isolated from the genomic library originate from the X-chromosome. The mapping of specific clones to refined regions of the X-chromosome is an important part of the candidate gene approach and an important initial step in determining if a given clone should be pursued in the elucidation of a specific genetic disorder.

In the present study, a mapping panel that effectively dissects the X-chromosome into 4 regions was used (Fig. 29). To obtain a more accurate map position, a more complete X-panel could be used. In addition the efficiency of characterizing clones and Southern analysis should be simplified and optimized. At the present time, entire phage clones as opposed to distinct fragments are being used to probe the X-panel.

The faint hybridization on Southern analysis to the X-panel in chapter 8 was initially thought to be due to the low specific activity of the probes. Recent changes in the Southern methodology that we have adopted has shown that the hybridization conditions and the particular nylon membrane used were probably the more critical

determinants of the observed results. A marked improvement in hybridization to background signal has been observed when a charged nylon membrane (Genescreen Plus) is used and when Denhardt's solution is eliminated from both the prehybridization and hybridization solutions.

10.300) EXPRESSION

For a clone to be considered as a potential candidate gene clone it should harbor an expressed sequence. The most straightforward way to demonstrate that a particular clone harbors an expressed sequence is to use the genomic clone (or a small restriction fragment of the clone) to directly probe a Northern blot of choroid/retinal RNA. Unfortunately the factor that led to the development of the present lateral strategy, the lack of sufficient amounts of choroid/retina tissue, also limits the capability of performing this analysis. As a consequence, an alternative approach to examine the expression of a single clone R8 was taken (section 9.000).

In the case of the analysis of R8, the compounding problem in the isolation of cDNA clones by the proposed scheme (Fig. 40) was the finding of a number of different cDNA clones whose only similarity is that they contain some derivation of the human Alu sequence (section 9.200). It would appear that the reassociation of the R8 genomic insert with human genomic DNA was not efficient enough to block hybridization on the basis of Alu repetitive elements. In retrospect, it would have perhaps been advantageous to have first attempted to further dissect the R8 4.5 Kb

fragment in order to derive a smaller fragment free of Alu sequences which still hybridized with the P probe before proceeding to the screening of the cDNA library.

Sequencing of the R8 cDNA clones allowed the synthesis of specific non-repetitive oligonucleotide sequences (section 9.300; Table 11). These oligonucleotides (which consisted of a cDNA clone match region and a mismatch region) were used in the S1 nuclease protection assay to verify that they correspond to expressed mRNA, as well as to determine the potential sense strand (Fig. 43). The same oligonucleotides were used to localize back to the R8 genomic 4.5 kb fragment by Southern analysis (Fig. 44). Because of the mismatch area of the specific oligonucleotides the hybridization stringency was brought down to accommodate this factor. Undoubtedly the data would have been more conclusive if the hybridization was carried out at higher stringency to eliminate the possibility of non specific hybridization. As it stands, the observed hybridization data (Fig. 44) suggests, but not conclusively, that an association between the R8.1.117 cDNA derived non-repetitive sequence (Oligo H and Oligo I) and the R8 4.5 Kb fragment exists. The hybridization data of the S1 oligonucleotides to the respective R8 fragment could have been strengthened by direct sequencing of the R8 genomic fragment and the subsequent comparison of the cDNA to genomic sequence. Alternatively, primers derived from the H and Ir match region could have been used in a PCR reaction to see if they could amplify a specific fragment from the R8 genomic clone. In either case a positive finding would have provided additional support for a direct relationship between the cDNA clone and the R8 genomic fragment.

In hindsight, a more practical approach would have been the generation and screening of different organ RNA blots, irrespective of the availability of human choroid/retina RNA, in order to obtain evidence of expression in any human tissue. This strategy was initially considered but again the inability to obtain human tissue of any kind other than blood, placenta, or cultured fibroblasts led us to abandon this approach. Recently, however, there appears to have been a growth of commercially available preparations of RNA from different types of human tissue as well as an emergence of tissue banks who will supply tissues for scientific research, thus making this avenue one which might now be further pursued.

10.400) COMPLICATIONS

The lateral approach defined in the present work leads to the isolation of human genomic clones which are derived from the human X-chromosome and which harbor choroid/retinal expressed sequences. On the basis of the observed data analysis, the nature of these choroid/retinal sequences can either be unique sequences or repetitive elements which are present in the P pool of sequences. This of course complicates matters in that the detection of a specific human genomic clone with the P probe does not necessarily define that clone as harboring a specific gene. If however, as discussed previously in this section, the repetitive elements in the cDNA are of low abundant sequences in the human genome, the clones isolated with the P probe should have a relative higher probability for harboring gene sequences than human X-clones taken at random from the library. Due to this

complication one of two things needs to be developed. Either it will be necessary to develop an efficient way of differentiating clones which harbor genes and those which do not or it will be necessary to streamline the detection system so that the background of false positives is significantly decreased.

10.500) CANDIDATE GENE APPROACH

The initial goal of the present lateral approach was to isolate a subset of clones which would have the potential to become a candidate gene clone for X-linked eye disorders due to their X-specificity and expression in the choroid of retina. Given the time that it has taken to partially characterize the R8 clone, it would be unrealistic to expect the systematic characterization of all clones isolated in the defined screenings in any expedient manner. It is therefore important to place a priority on those clones which should first be analyzed.

The most useful information for identifying clones that should be pursued is the location of the clone on the map of the X-chromosome. Clones that co-localize to the same regions as known ophthalmic disorders should be examined first. Mapping of a clone to a specific region by DNA cell panels provides at best an approximate localization of that sequence. Because the genetic location defined by a map panel is usually very broad more than one expressed gene will reside within a given map region. Thus in addition to a partial characterization of the candidate

clone, linkage studies to establish if the clone and disease locus co-segregate should be done in parallel. This would require the development of a polymorphic marker system with the clone in question and would require having the DNA from affected individuals and their families for analysis. Clones showing complete linkage to the disease locus should then be characterized by sequencing and for expressed sequences. The expressed sequences should then be analyzed and compared to DNA sequences isolated from patient DNA to identify differences between the normal and disease gene.

CONCLUSION

The work presented in the present thesis has focused on the development of a lateral screening strategy. A choroid/retina cDNA library has been used to cross screen a hamster/human X-specific genomic library in an attempt to isolate human X-specific genomic clones which harbor choroid/retinal expressed sequences.

The work has centred on the synthesis of a representative human choroid/retinal cDNA library, the synthesis of a random human X-specific genomic library, and the development of a general library-to-library cross screening procedure. The verification of the results of the proposed cross screen with respect to the isolation of human X genomic clones which harbor choroid/retinal expressed sequences has been partly successful. There does not seem to be a problem in showing whether a specific clone is or is not an X-specific clone by direct Southern analysis of X-specific DNA cell panels. Problems have arisen, however, in attempts to demonstrate that the isolated clones harbor sequences which are expressed in the choroid or retina. In the case of R8 the existing data, while suggestive, does not conclusively establish that the clone harbors an expressed sequence. Conclusive demonstration that the R8 clone harbors an expressed sequence has been confounded by the possible presence of repetitive sequences within the cDNA clones, leading to a background of false positives. From the overall analysis of the data it appears that the high background may be due to repetitive elements which have a relatively high abundance in the cDNA population (probably due to high expression of a select number of specific genes) but exist at a sufficiently low abundance in the

genome such that total genomic DNA does not effectively deplete these sequences. In both cases, suggestions to overcome the problem of validation of expressivity and to decrease the proportion of false positives have been presented.

In the light of these refinements a modified lateral cross screening strategy should offer a unique avenue in which to isolate tissue expressed genes with a predefined chromosome specificity. Within the realm of the choroid/retina project the refined strategy and the products of the defined work will represent a continuous well from which additional clones can be drawn. In this way a bank of human X-genomic clones which harbor choriod/retinal expressed sequences (and which possess a higher relative probability of harboring an active gene) can be built and used in future candidate gene and genetic mapping studies. The proposed strategy should fully augment and complement the candidate gene approach in the isolation of disease genes.

REFERENCES

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1983). Molecular Biology of the Cell. (Garland Publishing, Inc., New York) p409.

Alitalo, T., Karna, J., Forsius, H., and de la Chapelle, A. (1989). X-linked retinoschisis is closely linked to DXS41 and DXS16 but not DXS85. Clin. Genet. 32: 192-195.

Allday, M.J. and Jones, M.D. (1987). Rapid processing of nitrocellulose filter lifts of bacteriophage lambda libraries. Nucleic Acids Res. 15: 10592.

Auch, D. and Reth,M. (1991). Exon trap cloning: using PCR to rapidly detect and clone exons from genomic DNA fragments. Nucleic Acids Res. 22: 6743-6744.

Auffray, C., and Rougeon, F. (1979). Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. 107: 303-314.

Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA. 69: 1408-1412.

Bascom, R.A., Connell, G., Garcia-Heras, J., Collins, L., Ledbetter, D., Molday, R.S., Kalnins, V., and McInnes, R.R. (1990). Molecular and ultrastuctrual characterization of the products of the human retinopathy candidate genes ROM1 and RDS. Am. J. Hum. Genet. 47: A101.

Bascom, R., Manara, S., Gallie, B., Willard, H., Kalnins, V., and McInnes, R.R. (1989). Identification of a new mammalian photoreceptor-specific gene family. Am. J. Hum. Genet. 45: A172.

Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M., and Rodier, F. (1985). The mosaic genome of warm blooded vertebrates. Science 228: 953-958.

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

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

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

Bleeker-Wagemaker, L.M., Friedrich, U., Gal, A., Wienker, T.F., Warburg, M., and Ropers, H.H. (1985). Close linkage between Norrie's disease, a cloned DNA sequence from the proximal short arm, and the centromere of the X chromosome. Hum. Genet. 71: 211-214.

Boernke, W.E., Stevens, F.J., and Peraino, C. (1981). Effects of self-association of ornithine aminotransferase on its physiochemical characteristics. Biochemistry 20: 115-121.

Bonner, W.M., and Laskey, R.A. (1974). A film detection for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. 46: 83-88.

Brison, O., Ardeshir, F., and Stark, G.R. (1982). General method for cloning amplified DNA by differential screening with genomic probes. Mol. Cell. Biol. 2: 578-587.

Brooks, J.E. (1987). Properties and uses of restriction endonucleases. Methods in Enzymol. 152:113-129.

Bugawan, T.L., Saiki, R.K., Levenson, C.H., Watson, R.M. and Erlich, H.A. (1988). The use of non-radioactive oligonucleotide probes to analyze enzymatically amplified DNA for prenatal diagnosis and forensic HLA typing. Bio/Technology 6:943-947.

Cantor, C.R. (1984). Huntington's disease: Charting the path to the gene. Nature 308:404-405).

Chelley, J., Kaplan, J.-C., Maire, P., Gautron, S. and Kahn, A. (1988). Transcription of the dystrophin gene in human muscle and non-muscle tissues. Nature 333: 858-860.

Clarke, L. and Carbon, J. (1976). A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire E. coli genome. Cell 9:91-99.

Colantuoni, V., Rocchi, M., Roncuzzi, L., and Romeok G. (1986). Mapping of human cellular retinol-binding protein to chromosome 3. Cytogenet. Cell Genet. 43:221-222.

Collins, F.S., Drumm, M.L., Cole, J.L., Lockwood, W.K., Vande Woude, G.F., and Iannuzzi, M.C. (1984). Construction of a general human chromosome jumping library, with application to cystic fibrosis. Science 235:1046-1049.

Corbo, L., Maley, J.A., Nelson, D.L., and Caskey, C.T. (1990). Direct cloning of human transcripts with hnRNA from hybrid cell lines. Science 249: 652-655.

Crampton, J.M., Davies, K.E., and Knapp, T.F. (1981) The occurrence of families of repetitive sequences in a library of cloned cDNA from human lymphocytes. Nucleic Acids Res. 9: 3821-2834.

Cremers , F.P.M., Brunsmann, F., van del Pol, T.J.R., Pawiowitzki, I.H., Paulsen, K., Wieringa, B., and Ropers, H.H. (1987). Deletion of the DXS165 locus in patients with classical choroideremia. Clin. Genet. 32:421-423.

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

Cremers., F.P.M., van de Pol, J.R., Diergarre, P.J., Wieringa, Be., Nussbaum, R.L., Schwartz, M., and Ropers, H.H. (1989b). Physical fine mapping of the choroideremia locus using Xq21 deletions associated with complex syndromes. Genomics 4:41-46.

Cremers, F.P.M., van de Pol, T.J.R., Wieringa, B., Collins, F.S., Sankila, E.-M., Siu, V.M., Flintoff, W.F., Brunsmann, F., Blonden, L.A.J., and Ropers, H.H. (1989a). Chromosomal jumping from the DXYS165 locus allows a molecular characterization of four microdeletions and a de novo chromosome x/13 translocation associated with choroideremia. Proc. Nat. Acad. Sci. USA 86:7510-7514.

Cremers, F.P.M., van de Pol, T.J.R., Wieringa, B., Hofker, M.H., Pearson, P.L., Pfeiffer, R.A., Mikkelsen, M., Tabor, A., and Ropers, H.H. (1988). Molecular analysis of male-viable deletions and duplications allows ordering of 52 DNA probes on proximal Xq. Am. J. Hum. Genet. 43: 452-462.

de Saint-Basile, G., Bohler, M.C., Fischer, A., Cartron, J., Dufier, J.L., Griscelli, C., and Orkin, S.H. (1989). Xq21 DNA microdeletion in a patient with chronic granulomatous disease, retinitis pigmentosa, and McLeod phenotype. Hum. Genet. 80: 85-89.

De Chen, J., Ploder, L., Collins, L., Thorner, P., Kalnins, V., Duncan, A., Taylor, B., and McInnes, R.R. (1990). Chromosomal sublocalization and cellular expression of the retinal homeobox gene HOX10. Am. J. Hum. Genet. 47: A102

De Chen, J., Bapat, B., Bascom, R., Wilard, H., Gallie, B., and McInnes, R.R. (1989). Identification of a developmentally regulated human retinal homeobox gene. Am. J. Hum. Genet. 45: A111.

De La Chapelle , A., Sankila, E.-M., Lindlof, M., Aula, P., and Norio, R. (1985). Norrie's disease caused by a gene deletion allowing carrier detection and prenatal diagnosis. Clin. Genet. 28: 317-320.

Demmer, L.A., Birkenmeier, E.H., Sweetser, D.A., Levin, M.S., Zollman, S., Sparkes, R.S., Mohandas, T., Lusis, A.J., and Gordon, J.I. (1987). The cellular retinol binding protein II gene. J. Biol. Chem. 262: 2458-2467.

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

Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981). A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. 112: 295-298.

Dryja, T.P., McGee, T.L., Reichel, E., Hahn, L.B., Cowley, G.S., Yandell, D.W., Sandberg, M.A., and Berson, E.L. (1990). A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. Nature 343: 264-366.

Duyk, G.M., Kim, S., Pritchard, C., Bull, L., Zuo, J., Myers, R.M., and Cox, D.R. (1990). Isolation of Huntington disease candidate genes by exon trapping. Am. J. Hum. Genet. 47: A249.

Duyk, G.M., Peterson, A., Kim, S., Myers, R.M., and Cox, D.R. (1989). Exon trapping: a genetic screen for coding sequences within cloned anonymous genomic DNA. Am. J. Hum. Genet. 45: A184.

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

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

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323: 643-646.

Frischauf, A-M. (1987). Construction and characterization of a genomic library. Methods in Enzymol. 152:190-199.

Gal, A., Brunsmann, Hogenkamp, D., Ruther, K., Ahlert, D., Wienker, T.F., Hammerstein, W., and Pawlositzki, I.H. (1986). Choroideremia maps between DXS3 and DXS11 on Xq. Hum. Genet. 73: 123-126.

Gross-Bellard, M., Oudet, P., and Chambon, P. (1973). Isolation of high-molecular-weight DNA from mammalian cells. Eur. J. Biochem. 36: 32-38.

Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. Gene 25: 263-269.

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

Helms, C., Dutchik, J.E., and Olson, M.V. (1987). A lambda DNA protocol based on purification of phage on DEAE-cellulose. Methods in Enzymol. 153:69-82.

Hochgeschwender, U., Sutcliffe, J.G., and Brennan, M.B. (1989). Construction and screening of a genomic library specific for mouse chromosome 16. Proc. Natl. Acad. Sci. USA 86:8482-8486.

Hodgson, S.V., Heckmatt, J.Z., Hughes, E., Crolla, J.A., Dubowitz, V., Bobrow, M. (1986). A balanced *de novo* X/autosome translocation in a girl with manifestations of Lowe syndrome. Am. J. Med. Genet. 23: 837-847.

Hodgson, S.V., Robertson, M.E., Fear, C.N., Goodship, J., Malcom, S., Jay, B., Bobrow, M., and Pembrey, M.E. (1987). Prenatal diagnosis of X-linked choroideremia with mental retardation, associated with a cytologically detectable X-chromosome deletion. Hum. Genet. 75: 286-290.

Huynh, T.V., Young, R.A., and Davies, R.W. (1985). Construction and screening cDNA libraries in lambda gt10 and lambda gt11. In DNA cloning techniques, a practical approach. Ed: Glover, D. (IRL Press, Oxford, England) pp49-78.

Ikemura, T., Wada, K., and Aota, S. (1990). Giant G+C mosaic structures of the human genome found by arrangement of Genbank human DNA sequnces according to genetic positions. Genomics 8: 207-216.

Inglehearn, C.F., Papiha, S.S., Jay, M., Wright, A.F., Moore, A.T., and Bhattacharya. (1990). Linkage of internal minisatellite loci on chromosome 1 and exclusion of autosomal dominant retinitis pigmentosa proximal to rhesus. J. Med. Genet. 27:14-15.

Innana, G., Totsuka, S., Redmond, M., Dougherty, T., Nagle, J., Shiono, T., Ohura, T., Kominami, E., and Katunuma, N. (1986). Molecular cloning of human ornithine aminotransferase mRNA. Proc. Natl. Acad. Sci. USA. 83: 1203-1207.

Kerem, B.-S., Rommens, J.M., Buchanan, J.A., Markiewixz, D., Cox, T.K., Chakravarti, A., Buchwald, M., Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: genetic analysis. Science 245: 1073-1080.

Ko, M.S.H. (1990). An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. Nucleic Acids Res. 19: 5705-5711.

Korenberg, J.R., Croyle, M.L., and Cox, D.R. (1987). Isolation and regional mapping of DNA sequences unique to human chromosome 21. Am. J. Hum. Genet. 41: 963-978.

Lesko, J.G., Lewis, R.A., and Nussbaum, R.L. (1987). Multipoint linkage analysis of loci in the proximal long arm of the human X chromosome: Application to mapping the choroideremia locus. Am. J. Hum. Genet. 40:303-311.

Lewis, R.A., Holcomb, J.D., Bromley, W.C., Wilson, M.C., Roderick, T.H., and Hejtmancik, J.F. (1987). Mapping X-linked ophthalmic diseases III. Provisional assignment of the locus for blue cone monochromacy to Xq28. Arch. Ophthalmol. 105: 1055-1059.

Lewis, R.A., Nussbaum, R.L., and Ferrell, R. (1985). Mapping X-linked ophthalmic diseases. Ophthalmol. 92: 800-805.

Lewis, R.A., Nussbaum,R.L., and Stambolian, D. (1990). Mapping X-linked ophthalmic diseases IV. Provisional assignment of the locus for X-linked congenital cataracts and microcornea (the Nance-Horan syndrome) to Xp22.2-22.3. Ophthalmol. 97: 110-121.

Lindsay, S., and Bird, A.P. (1987). Use of restiction enzymes to detect potential gene sequences in mammalian DNA. Nature 327: 336-338.

Lindsay, S.J., Bower, D.J., Sealey, P.G., Ackford, H. and Bhattacharya, S.S. (1988). Screening of human X-specific genomic clones for CpG rich islands. J. Cell. Biochem. 12B(supp):238.

Liu, P., Legerski, R., and Siciliano, M.J. (1989). Isolation of human transcribed sequences from human-rodent somatic cell hybrids. Science 246: 813-815.

Lu, X. and Werner, D. (1988). Construction and quality of cDNA libraries prepared from cytoplasmic RNA not enriched in poly(A)+RNA. Gene 71: 157-164.

Lyons, R.T. and Pitot, H.C. (1976). The regulation of ornithine aminotransferase synthesis by glucagon in the rat. Arch. Biochem. Biophys. 174: 262-272.

MacDonald, I.M., Sandre, R.M., Hunter, A.G.W., and Tenniswood, M.P.R. (1987a). Gene mapping of X-linked choroideremia with restriction fragment length polymorphisms. Can. J. Ophthalmol. 22:310-315.

MacDonald, I.M., Sandre, R.M., Wong, P., Hunter, A.G.W., and Tenniswood, M.P.R. (1987b). Linkage relationships of X-linked choroideremia to DXYS1 and DXS3. Hum. Genet. 77: 233-235.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning, A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maynard-Smith, S., Penrose, L.S., and Smith, C.A.B. (1961). Mathematical tables for research workers in human genetics. (J. and A. Churchill Ltd., London) 74pgs.

McInnes, R.R., Manara, S., Sexsmith, E., Warsh, J., Sadler, S., and Taylor, R.G. (1988). The isolation of abundant retinal-specific cDNAs: candidate genes for retinal diseases and probes for the study of retinal development. Am. J. Hum. Genet. 43:A195.

McWilliam, P., Farrar, G.J., Kenna, P., Bradley, D.G., Humphries, M.M., Sharp, E.M., McConnell, D.J., Lawler, M., Sheils, D., Ryan, C., Stevens, K., Daiger, S.P., and Humphries, P. (1989). Autosomal dominant retinitis pigmentosa (ADRP): localization of an ADRP gene to the long arm of chromsome 3. Genomics 5:619-622.

Merry, D.E., Lesko, J.G., Siu, V., Flintoff, W.F., Collins, F., Lewis, R.A., and Nussbaum, R.L. (1990). DXS165 detects a translocation breakpoint in a woman with choroideremia and a de novo X:13 translocation. Genomics 6:609-615.

Merry, D.E., Lesko, J.G., Sosnoski, D.M., Lewis, R.A., Lubinsky, M., Trask, B., van den Engh, G., Collins, F.S., and Nussbaum, R.L. (1989). Choroideremia and deafness with stapes fixation: a contiguous gene deletion syndrome in Xq21. Am. J. Hum. Genet. 45: 530-540.

Mitchell, G.A., Looney, J.E., Brody, L.C., Stel, G., Suchanek, M., Engelhardt, J.F., Willard, H., and Valle, D. (1988). Human ornithine-o-aminotransferase cDNA cloning and analysis of the structural gene. J. Biol. Chem. 263: 14288-14295.

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

Mullis, K.B. and Faloona,F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods in Enzymol. 155: 335-350.

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 blindnesss (CSNB1) to Xp11.3. Genomics 5: 727-737.

Myal, Y., Robinson, D.B., Iwasiow, B., Tsuyuki, D., Wong, P., and Shiu, P.C. (1991). Mechanisms of action of prolactin and androgen in gene expresion revealed by studying the human prolactin-inducible protein (PIP)/gross cystic disease fluid protein (GCDFP-15) gene. Gene (submitted).

Nathans, J., Davenport, C.M., Maumenee, I.H., Lewis, R.A., Hejtmancik, J.F., Litt, M., Lovrien, E., Weleber, R., Bachynski, B., Zwas, F., Klingaman, R., and Fishman, G. (1989). Molecular genetics of human blue cone monchromacy. Science 245: 831-838.

Nathans, J. and Hogness, D.S. (1983). Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. Cell 34: 807-814.

Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T. B., and Hogness, D.S. (1986b). Molecular genetics of inherited variation in human color vision. Science 232: 203-210.

Nathans, J., Thomas, D., and Hogness, D.S. (1986a). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. Science 232:193-202.

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

Neve, R.L., Stewart, G.D., Newcomb, P., Van Keuren, M.-L., Patterson, D., Drabkin, H.A., and Kurnit, D.M. (1986). Human chromosome 21-encoded cDNA clones. Gene 49: 361-369.

Newbold, C.I., Boyle, D.B., Smith, C.C., and Brown, K.N. (1982). Stage specific protein and nucleic acid synthesis during the asexual cycle of rodent malaria *Plasmodium chabaudi*. Mol. Biochem. Parasitol. 5: 33-44.

Nilsson, M.H.L., Spurr, N.K., Lundvall, J., Rask, L., and Peterson, P.A. (1988). Human cellular retinol-binding protein gene organization and chromosmal location. Eur. J. Biochem. 173: 35-44.

Noremore, W.M., Shapiro H.S., and Setlow, P. (1976). in CRC handbook of biochemistry and molecular biology. (G.D. Fasman, ed), (CRC press, Cleveland, Ohio) p65.

Nussbaum, R.L., Lesko, J.G., Lewis, R.A., Ledbetter, S.A., and Ledbetter, D.H. (1987). Isolation of anonymous DNA sequences from within a submicroscopic X chromosomal deletion in a patient with choroideremia, deafness and mental retardation. Proc. Natl. Acad. Sci. USA. 84:6521-6525.

Nussbaum, R.L., Lewis, R.A., Lesko, J.G., and Ferrill, R. (1985). Choroideremia is linked to the restriction fragment length polymorphism DXYS1 at Xq13-21. Am. J. Hum. Genet. 37: 473-481.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250: 4007-4021.

Ohba, N and Yamashita, T. (1986). Primary vitreoretinal dysplasia resembling Norrie's disease in a female: association with X autosome chromosomal translocation. Br. J. Ophthalmol. 70: 64-71.

Paulson, K., Forrest, S., Scherer, G., Ropers, H.H., and Davies, K. (1986). Regional localization of X chromosome short arm probes. Human Genet. 74:155-159.

Pelham, H.R.B., and Jackson, K.J. (1976). An efficient mRNA-dependent translation system from reticulocyte lysate. Eur. J. Biochem. 67: 247-256.

Polomeno, R.C., Zeesman, S., MacDonald, I.M., Crozier, D.G., Tenniswood, M.P.R. and Kaplan, P. (1987). Norrie's disease in a French Canadian kindred: attempt to detect carriers by DNA analysis. Can. J. Ophthalmol. 22:21-23.

Queen, C., and Korn, L.J. (1984). A comprehensive sequence analysis program for the IBM personal computer. Nucl. Acids Res. 12:581-599.

Riddell, D.C., Wang, H.S., Beckett, J., Chan, A., Holden, J.J.A., Mulligan, L.M., Phillips, M.A., Simpson, N.E., Wrogemann, K., Hamerton, J.L., and White, B.N. (1986). Regional localization of 18 human X-linked DNA sequences. Cytogenet Cell Genet 42:123-128.

Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specfic activity *in vitro* by nick translation with DNA Polymerase I. J. Mol. Biol. 113: 32-37.

Riordan, J.R., Rommens, J.M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavisic, N., Chou, J.-L. Drumm, M.L., Iannuzzi, M.C., Collins, F.S., and Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245: 1066-1073.

Rommens, J.M., Iannuzzi, M.C., Keren, B.-S., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N,., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.-C., Collins, F.S. (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065.

Rosenberg, T, Schwartz, M., Niebuhr, E., Yang, H-M., Sardemann, H., Andersen, O., and Lundsteen, C. (1986). Choroideremia in interstitial deletion of the X-chromosome. Ophthalmic Pediatr.Genet. 7:205-210.

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

Rutledge, R.G., Seligy, V.L., Cote, M.-J., Dimock, K., Lewin, L.L., and Tenniswood, M. (1988). Rapid synthesis and cloning of complementary DNA from any RNA molecule into plasmid and M13 vectors. Gene 68: 151-158.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-490.

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

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning, a laboratory manual, second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanada, Y., Shiotani, T., Okuno, E., and Katunuma, N. (1976). Coenzyme-dependent conformational properties of rat liver ornithine aminotransferase. Eur. J. Biochem. 69: 507-515.

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

Sankila, E-M., de la Chapelle, A., Karna, J., Forsius, H., Frants, R., and Eriksson, A. (1987). Choroideremia: close linkage to DXYS1 and DXYS12 demonstrated by segregation analysis and historical-genealogical evidence. Clin. Genet. 31: 315-322

Sarkar, G. and Sommer, S.S. (1988). RNA amplification with transcript sequencing (RAWTS). Nucleic Acids Res. 16: 5197.

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

Scherer, G., Telford, J., Baldari, C., and Pirrotta, V. (1981). Isolation of cloned genes differentially expressed at early and late stages of Drosophila embryonic development. Develop. Biol. 86: 438-447.

Schnur, R.E., Trask, B.J., van den Engh, G., Punnet, H.H., Kistenmacher, M., Tomeo, M.A., Naids, R.E., and Nussbaum, R.L. (1989). An Xp22 microdeletion associated with ocular albinism and ichthyosis: approximation of breakpoints and estimation of deletion size by using cloned DNA probes and flow cytometry. Am. J. Hum. Genet. 45: 706-720.

Schwartz, M., Rosenberg, T., Niebuhr, E., Lundsteen, C., Sardemann, H., Andersen, O., Yang, H-M., Lamm, L.U. (1986). Choroideremia: further evidence for assignment of the locus to Xq21. Hum. Genet. 74:449-452.

Schwartz M., Yang, H-M., Niebuhr, E., Rosenberg, T., and Page, D.C. (1988). Regional localization of polymorphic DNA loci on the proximal long arm of the X chromosome using deletions associated with choroideremia. Hum. Genet. 78: 156-160.

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

Shih, V.E. (1981). Regulation of ornithine metabolism. Enzyme (basel) 26:254-258.

Silver, D.N., Lewis, R.A., and Nussbaum, R.L. (1986). Mapping the Lowe oculocerebrorenal syndrome to Xq24-q26 by use of restriction fragment length polymorphisms. J. Clin. Invest. 79:282-285.

Simell, O. and Takki, K. (1973). Raised plasma-ornithine and gyrate atrophy of the choroid and retina. Lancet 1:1031-1033.

Siu, V.M., Gonder, J.R., Sergovich, F.R., Flintoff, W.F. (1988). Choroideremia associated with an X-autosomal translocation. Am. J. Hum. Genet. 43:A96.

Smith, R.J., Downing, S.J., and Phang, J.M. (1977). Enzymatic synthesis and purification of L-pyrroline-5-carboxylic acid. Anal. Biochem. 82: 170-176.

Stoflet, E.S., Koeberi, D.D., Sarkar, G., and Sommer, S.S. (1988). Genomic amplification with transcript sequencing. Science 239: 491-494.

Tabor, S., and Richardson, C.C. (1987). DNA sequence analysis with a modified bacteriophage T& DNA polymerase. Proc. Natl. Acad. Sci. USA 84:4767-4771.

Tartof, K.,D. (1975). Redundant genes. Annu. Rev. Genet. 9:355-385.

Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maizel, A., and Saunders, G.F. (1981). Presence of a highly repetitive and widely dispersed DNA sequence in the human genome. Proc. Natl. Acad. Sci. USA. 78: 1508-1512.

Tenniswood, M., and Simpson, A.J.G. (1982). The extraction, characterization and *in vitro* translation of RNA from adult *Schistosoma mansoni*. Parasitology 84: 253-261.

Tokarskaya, O.N., Tchurikov, NA., Ivanov, P.L., Kramerov, D.A., and Ryskov, A.P. (1980). Sequences hybridizing to mRNA, oligo(dT) and dsRNA from pre-mRNA are contiguous in the cloned mouse DNA fragments. Nucleic Acids Res. 8:425-440.

van de Pol, T.J.R., Cremers, F.P.M., Brohet, R.M., and Ropers, H.H. (1990). Derivation of clones from the choroideremia locus by preparative field inversion gel electrophoresis. Nucleic Acids Res. 18:725-731.

Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA. 76: 615-619.

Volpe, P., Sawamura, R., and Strecker, H.J. (1969). Control of ornithine o-transaminase in rat liver and kidney. J. Biol. Chem. 244: 719-726.

Weiss, M.C. and Green, H. (1967). Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. Genetics 58: 1104-1111.

Williams, J.G. (1981). The preparation and screening of a cDNA clone bank in Genetic engineering 1, Ed Williamson, R. (Academic Press, London) pp 1-59.

Williams, J.G. and Lloyd, M.M. (1979). Changes in the abundance of polyadenylated RNA during slime mould development measured using cloned molecular hybridization probes. J. Mol. Biol. 129:19-35.

Wong, C., Djowling, C.E., Saiki, R.K., Higuchi, R.G., Erlich, H.A., and Kazazian Jr., H.H. (1987). Characterization of ß-thalassemia mutations using direct genomic sequencing of amplified single copy DNA. Nature 330: 384-386.

Wong, P., MacDonald, I., and Tenniswood,M. (1989). Use of the polymerase chain reaction for the differential cross screening of libraries cloned into phage-lambda-based vectors. Gene 85: 59-65.

Wong, P., Myal, Y., Shiu, R.P.C., and Tenniswood, M.P.R. (1991). A simple method for the generation of more representative genomic libraries (in preparation).

Wong, P., Tenniswood, M. and MacDonald, I. (1988). DXYS1 and DXS3 do not flank the gene locus for X-linked Choroideremia (TCD). J. Cell. Biochem. 12B(sup):243.

Wymann, A.R., Wertman, K.F., Barker, D., Helms, C., and Petri, W.H. (1986). Factors which equalize the representation of genome segments in recombinant libraries. Gene 49:263-271.

Zhu, D., Antonarakis, S.E., Schmeckpeper, B.J., Diergaarde, P.J., Greb, A.E., and Maumenee, I.H. (1989). Microdeletion in the X-chromosome and prenatal diagnosis in a family with Norrie's disease. Amer. J. Med. Genet. 33:485-488.

APPENDIX ONE

ABBREVIATIONS

AMV	avian myeloblastosis virus
bp	nucleotide base pair
BSA	bovine serum albumin
cDNA	complementary DNA
chorio-retina	choroid and retina
cpm	counts per minute
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
DEAE	diethylaminoethyl
dGTP	deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide 5'-triphosphate
ds-cDNA	double stranded cDNA
DTT	dithiothrietol
EDTA	ethylenediaminetetra-acetic acid
G-50	Sephadex G-50
gm	gram
h	hour(s)
HA	Hamster
HA+	Positive hybridization to a hamster genomic probe
HA-	Negative hybridization to a hamster genomic probe
HMW	high molecular weight
hnRNA	heteronuclear RNA
HTF	Hpa II tiny fragments
HU	Human

HU+	Positive hybridization to a human genomic probe
HU-	Negative hybridization to a human genomic probe
kb	kilobase
kbp	kilobase pair
LMA	low melting agarose
min	minute
mRNA	messenger RNA
O.D.	optical density
oligo-(dT)	oligodeoxythimidylate
p	short arm of a chromosome
P probe	Probe derived from the PCR products of the amplification of cDNA library
P+	positive hybridization with the P probe
P-	negative hybridization with the P probe
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	poly ethylene glycol
pfu	plaque forming unit
pH	$-\log_{10}[H^+]$
pI	isoelectric point
PIPES	1,4-piperazinediethanesulfonic Acid
poly(A)+	RNA retained on oligo(dT)-cellulose column
poly(A)-	RNA not retained on oligo(dT)-cellulose column
q	long arm of a chromosome
RNA	ribonucleic acid
RNase	ribonuclease

RP	Retinitis Pigmentosa
rpm	revolutions per minute
rRNA	ribosomal RNA
S	Svedberg unit
SDS	sodium dodecyl sulfate
SF	subfragment
ss-cDNA	single-stranded cDNA
t	translocation
Taq	*Thermus aquaticus*
TCA	trichloroacetic acid
TCD	Choroideremia gene
Tris	Tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
TTP	thymidine 5'-triphosphate
U	unit of enzyme activity
UV	ultra violet
V	volts

APPENDIX TWO

SUPPLIES AND GENERAL SOLUTIONS

Amersham (Oakville, Ontario): [^{35}S]-methionine (800 Ci/mmole), [^{32}P]-dCTP (3000 Ci/mmole), [^{32}P]-dATP (3000 Ci/mmole), nick translation kits, multiprime kits, λgt10 cloning kit, Hybond™ nylon membranes, Hybond™ N+ nylon membranes, restriction endonucleases.

BDH Canada (Toronto, Ontario): Ficoll, polyvinylpyrrolidone.

Beckman Canada (Mississauga, Ontario): Ready-Solv ET Scintillation fluid.

Bethesda Research Laboratories (Burlington, Ontario): Ampholytes, RNAse H, restriction endonucleases, Hind III digested λ DNA.

Bill's Camera Craft (Ottawa, Ontario) & Henry's (Toronto, Ontario): Cronex X-ray film, Illford films, Ilford and Kodak paper and other dark room chemicals.

Bio-Can Scientific Inc (Mississauga, Ontario): Gene clean™ kits

Bio-Rad Laboratories (Canada) Ltd (Mississauga, Ontario): Acrylamide, N'N'-methylene-bisacrylamide, N,N,N,N,-tetramethylenediamine, Coomassie blue, protein molecular weight standards, Bovine serum albumin, AG 501-X8D mixed bed resin, Penicillin, streptomycin.

Boehringer-Mannheim Canada (Dorval, Quebec): Ampicillin, T4 ligase, DNA polymerase 1, RNAse A, yeast tRNA, calf thymus DNA, Tris, oligo(dT-cellulose), polynucleotide kinase, RNAse H, restriction endonucleases.

Dupont (Lachine, Quebec): [^{32}P]-dCTP (3000 Ci/mmole), [^{32}P]-dATP (3000 Ci/mmole), [^{35}S]-dATP (1045 Ci/mmole), NEN™ labeling kits.

International Biotechnologies Incorporated (Toronto, Ontario): Polyacrylamide, SDS, N',N'-methylene-bisacrylamide, urea, agarose, N,N,N',N'-tetra-methylendiamine, ammonium persulfate, TEMED.

Life Sciences (St. Petersburg, FL): AMV reverse transcriptase

Oxoid Canada (nepean, Ontario): Agar, tryptone, bacto-yeast extract.

Perkin-Elmer Cetus (Nepean, Ontario): Gene amplification-DNA amplification kit, Taq Polymerase.

Pharmacia (Dorval, Quebec): random-sequence hexadeoxyribonucleotides, oligo-(dT)cellulose, Sephadex G-50, Sepharose 4B, nucleotides (dNTPs), ampholytes (LKB), restriction endonucleases.

Promega Biotec (Madison, Wisconsin): Rabbit reticulocyte lysate, EMBL3 cloning kit.

Sigma (St. Louis, Mo): Oligo dT_{12-18}

Universtiy of Ottawa Biotechnology Research Institute (Ottawa, Ontario): Synthetic oligonucleotides and λgt10 primers.

US Biochemicals Corp.(Cleveland, Ohio): Sequenase™ DNA sequencing kits.

Whatman BioSystems Inc.(Clifton, New Jersey): DE52-preswollen DEAE cellulose.

All other chemicals, of reagent grade, were purchased from Fisher Scientific Co. Ltd (Ottawa, Ontario) or BDH (Toronto, Ontario).

BACTERIAL CELL LINES

NM514: selective bacterial host for the vector λgt10 (Amersham, Oakville, Ontario).

L87 : non-selective bacterial host for the vector λgt10 (Amersham, Oakville, Ontario).

MB406: bacterial host for the vector EMBL3 (Promega Biotec, Madison, Wisconsin).

DH5α : bacterial host for the vector PTZ19R (Promega Biotec, Madison, Wisconsin).

GENERAL SOLUTIONS

2YT broth	1.6% bactotryptone, 1% bacto-yeast extract, 0.5% NaCl
Denhardt's	0.02% BSA, 0.02% polyvinyl pyrrolidone, 0.02% Ficoll
High TE	100 mM Tris-HCl pH 8.0, 1 mM EDTA
L-agar	1% bactotryptone, 0.5% bacto-yeast extract, 1% NaCl, 1.5% bacto-agar, pH 7
L-broth	1% bactotryptone, 0.5% bacto-yeast extract, 1% NaCl, pH 7
MOPS	20 mM morpholonopropanesulfonic acid, 5 mM NaOAc, 1 mM EDTA, pH 7.0
SM buffer	50 mM Tris, 100 mM NaCl, 10 mM $MgSO_4$, 0.1% gelatin, pH 7.5
SSC	0.15 M NaCl, 0.015 M sodium citrate, pH 7
STE	100 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 7
TAE loading buffer	10x TAE buffer, 50% glycerol, 0.5 mg Xylene Cyanol/ml, 0.25 mg Bromophenol Blue
TAE buffer	40 mM Tris, 20 mM EDTA, 0.23 % glacial acetic acid, pH 7.7
TBE buffer	44 mM boric acid, 44 mM Tris, 20 mM EDTA, pH 8.3

TBE loading buffer	20 mM EDTA, 96% deionized formamide, 0.5 mg Xylene Cyanol/ml, 0.25 mg Bromophenol Blue
TE	10 mM Tris-HCl pH 7.5 or 8.0, 1 mM EDTA
Top agar	1% bactotryptone, 0.5% bacto-yeast extract, 0.5% NaCl, 0.25% $MgSO_4$, 1% bact-agar, pH 7

APPENDIX THREE

CURRICULUM VITAE

PERSONAL DATA

NAME	:	PAUL W. WONG
PERMANENT ADDRESS	:	54 Calder Bay Winnipeg, Manitoba, Canada R3T 5L6 Telephone: (204)269-0964
CURRENT ADDRESS	:	202-1725 Riverside Drive Ottawa, Ontario, Canada K1G OE6 Telephone: (613)523-6339
CITIZENSHIP	:	CANADIAN

PRESENT STATUS

PRESENT STATUS	:	Graduate Student
PROGRAM	:	Ph.D.
DATE OF COMMENCEMENT	:	September, 1986
UNIVERSITY	:	University of Ottawa
DEPARTMENT	:	Department of Biochemistry
SUPERVISOR	:	Dr.M.P.R.Tenniswood
ADDRESS	:	Department of Biochemistry University of Ottawa 451 Smyth Road Ottawa, Ontario K1H 8M5 Telephone: (613)787-6419 FAX : (613)738-3191

EDUCATION

1) **Bachelor of Science (First Class Honors)**, Department of Zoology, University of Manitoba, May 25, 1983

2) **Masters of Science**, Department of Anatomy (Genetics Option), University of Manitoba, February 5, 1986

SCHOLARSHIPS AND ACADEMIC AWARDS

1.	University of Winnipeg Board of Reagents special entrance scholarship.(declined).	University of Winnipeg	1979-1980
2.	University of Manitoba Entrance scholarship.	University of Manitoba	1979-1980
3.	University of Manitoba (faculty of Science) Dean's honour list.	University of Manitoba	1979-1980 1981-1982 1982-1983
4.	Manitoba health research council graduate studentship.	University of Manitoba	1984-1985
5.	University of Ottawa graduate research scholarship.	University of Ottawa	1986-1987
6.	Ontario Graduate Scholarship	University of Ottawa	1987-1988
7.	Ontario Graduate Scholarship (declined)	University of Ottawa	1988-1989
8.	University of Ottawa Research Excellence Scholarships in the Sciences and Engineering	University of Ottawa	1988-1990
9.	Medical Research Council Studentship (Canada)	University of Ottawa	1988-1990
10.	University of Ottawa Travel Award	University of Ottawa	1990

RESEARCH AND RELATED WORK EXPERIENCES

1. Dates : Summer 1982/ Summer 1983
   Appointment: Research Technician, Division of Human Genetics, University of Manitoba
   Supervisor : Dr.L.J.Donald

2. Dates : Sept.1983- Feb.1986
   Status : M.Sc. Graduate Student, Department of Anatomy (genetics option),University of Manitoba.
   Supervisor : Dr.P.J.McAlpine
   Thesis : Genetic studies of human blood coagulation factor XIII.

3. Dates : 1985
   Appointment: Teacher's Assistant
   Supervisor : Dr.P.J.McAlpine
   Course : Med 1-Genetics, Division of Human Genetic, University of Manitoba.

4. Dates : Feb.1986-Aug.1986
   Appointment: Lab Technician (tech 4,step 1)
   Department of Human Genetics, University of Manitoba
   Supervisor : Dr.P.J.McAlpine.

5. Dates : Jan.1987- 1989
   Status : Research Consultant
   Supervisor : Dr. RJR.McKendry, Division of Rheumatology, Ottawa General Hospital
   Project : Linkage Analysis of Human Nodal Osteoarthritis.

6. Dates : Sept.1986-Present
   Status : Ph.D Graduate Student, Department of Biochemistry, University of Ottawa.
   Supervisor : Dr.M.P.R.Tenniswood
   Project : Lateral genetics: An approach to the cloning of potential human choriod/retinal X-linked genes.

7. Dates : Sept.1987-1991
   Appointment: Teacher's Assistant
   Supervisor : Dr.P.Anderson
   Course : Med1-Biochemistry, Department of Biochemistry, University of Ottawa

THESES

1. **Wong P.** 1986. Genetic studies of blood coagulation factor XIII. M.Sc. Thesis, University of Manitoba, Winnipeg, Manitoba, Canada.

FULL MANUSCRIPTS

1. Allderdice PW, Kaita H, Lewis M, McAlpine PJ, **Wong P**, and Onyett H. (1986) Segregation of marker loci in families with an inherited paracentric insertion of chromosome 9. Am. J. Hum. Genet. 39:612-617.

2. Lewis M, Kaita H, Philipps S, McAlpine PJ, **Wong P**, Giblett ER, and Anderson J. (1987) The Yt blood group system (ISBT # 011):Genetic studies. Vox Sanguinis 53:52-56.

3. MacDonald IM, Sandre RM, **Wong P**, Hunter AGW, and Tenniswood MPR. (1987) Linkage relationships of X-Linked Choroideremia to DXYS1 and DXS3. Human Genetics 77:233-235.

4. **Wong P**, Komarniki L,Schroeder ML, Lewis M, Kaita H, Philipps S, Stranc L, and McAlpine PJ. (1988) Analysis for linkage between F13A and three chromosome 6 marker loci: evidence for 6pter:F13A: HLA:GLO:cen gene order. Human Genetics 79:228-230.

5. Zelinski T, Kaita H, Lewis M, Coghlan G, Philipps S, Belcher E, McAlpine P, Coopland G, **Wong P**. (1988) The Colton blood group locus: a linkage analysis. Transfusion 28:435-438.

6. Rodriguez de Cordoba S, Rey-Campos J, Dykes DD, McAlpine PJ, **Wong P**, and Rubinstein P. (1988) Coagulation Factor XIII B Subunit is encoded by a gene linked to the regulator of complement activation (RCA) Gene cluster in Man. Immunogenetics 28:452-454.

7. Lewis M, Kaita H, Philipps S, McAlpine PJ, **Wong P**. (1989). Exclusion of unassigned blood group system loci from the regulator of complement activation (RCA) gene cluster on chromosome 1q32. Vox Sanguinis 57:210-212.

8. **Wong P**, MacDonald IM, and Tenniswood M. (1989). The use of polymerase chain reaction for the differential cross screening of libraries cloned into Lambda based vectors. Gene:85:59-65.

9. Tenniswood MP, Montpetit ML, Leger JG, **Wong P**, Pineault J, and Rouleau M. (1990). Epithelial-Stromal Interactions and Cell Death in the Prostate. In The Prostate as an Endocrine Gland. eds. RJ Ablin and WE Farnsworth, CRC Press, Boca Raton Florida. Pg 187-204.

10. McAlpine PJ., Feasby TE., Hahn AF., Komarnicki L., James S., Guy C., Dixon M., Qayyum S., Wright J., Coopland G., Lewis M., Kaita H., Philipps S., **Wong P.**, Koopman W., Cox DW and Yee WC. (1989). Localization of a locus for Charot-Marie-Tooth Neuropathy Type Ia (CMT1A) to chromosome 17. Genomics 7:408-415.

11. Myal Y., Robinson D.B., Iwasiow B., Tsuyuki D., **Wong P.**, and Shiu R.P.C. (1991). Mechanisms of action of prolactin and androgen in gene expression revealed by studying the human prolactin-inducible protein (PIP)/ gross cystic disease fluid protein (GCDFP-15) gene. Molecular and Cellular Endocrinology 80:165-175.

MANUSCRIPTS IN PREPARATION

1. **Wong, P.**, Myal, Y., Shiu, R.P.C., and Tenniswood, M.P.R. (1991). A simple method for the generation of more representative genomic libraries.

2. **Wong, P.**, MacDonald,I.M., and Tenniswood, M.P.R. (1991). Identification of a potential retinopathy candidate disease gene by lateral genetics.

3. Guenette, R.S., **Wong, P.**, Daehlin, L., and Tenniswood, M.P.R. (1991). Characterization of a novel gene expressed in the epithelium of regressing hormone dependent tissues.

ABSTRACTS

1. **Wong P**, Wallace S, Nickel BE, Cox DW, Hughes H, Allderdice PW, Lewis M, Kaita H, Philipps S and McAlpine PJ. Mapping the coagulation factor XIIIB locus(F13B) in man. (Abstract of an oral presentation at the Annual Congress of the genetic society of Canada, Ste Foy, Quebec, June 1986).

2. Feasby TE, Hahn AF, Koopman WJ, McAlpine PJ, **Wong P**, Lewis M, Kaita H, and Philipps S. Gene mapping studies of hereditary motor-sensory neuropathy type 1 mutations. (Abstract of presentation at the VIth International Congress on Neuromuscular Disease, Los Angeles, Calif, 1986). J. of Amer Ass'n Electromyography and Electrodiagnosis 9(supp):232.

3. **Wong P**, MacDonald I, and Tenniswood M. Human chorioretinal RNA: characterization and cloning. (Abstract of poster presentation at the 30th Annual meeting of the Canadian Federation of Biological Societies, Winnipeg, Manitoba, June 1987).

4. MacDonald I, **Wong P**, and Tenniswood M. Characterization and cloning of RNA from human chorioretinal tissue. (Abstract of poster presentation at the 50th Annual meeting of the Canadian Ophthalmological Society, Montreal, Quebec, June 1987).

5. **Wong P**, Komarniki L, Schroeder M, Lewis M, Kaita H, Philipps S, McAlpine PJ. Linkage analysis giving 6pter:F13A:HLA: GLO:cen order. (Abstract of a poster presentation at Human Gene Mapping 9: Ninth International Workshop on Human Gene Mapping, Paris, France, September 1987). Cytogenet Cell Genet

6. McKendry RJR, Sengar D, and **Wong P**. Linkage and association between Nodal OsteoArthritis (NOA) and HLA haplotypes. (Abstract of oral presentation at the Annual Meeting of the Canadian Rheumatological Association, Winnipeg, Manitoba, September 1987).

7. **Wong P**, Tenniswood M. & MacDonald I. DXYS1 and DXS3 do not flank the gene locus for X-linked Choroideremia (TCD). (Abstract of poster presentation at the UCLA Symposia: Molecular Biology of the Eye: Genes, Vision, and Ocular Disease, Santa Fe, New Mexico, February 1988). Journal of Cellular Biochemistry 12B(sup):243.

8. Zelinski T, Kaita H, Philipps S, Coghlan G, Belcher E, Lewis M, McAlpine P, Komarnicki L, **Wong P**, and Schwartz J. The froese blood group locus (FR). (Abstract of presentation at the annual meeting of the International Society for Blood Transfusion, London, England, July 1988).

9. **Wong P**, MacDonald I, Tenniswood M. A rapid and novel approach in the identifiation of human X-specific sequences expressed in the human retina. (Abstract of a poster presentation at the Annual Congress of the Genetic Society of Canada, Calgary, Alberta, June 1989). Genetics Society of Canada Builetin 20(sup):56

10. Coghlan G., Kaita H., Belcher E., Philipps S., **Wong P.**, McAlpine PJ., Zelinski T., Lewis M. Genetic linkage between the Kell and YT blood group loci. (Abstract of a poster presentation at Human Gene Mapping 10: Tenth International Workshop on Human Gene Mapping, New Haven, USA, June 1989). Cytogenet Cell Genet. 51:978

11. **Wong P.**, MacDonald I., and Tenniswood M. Selective isolation of human X-specific sequences expressed in the human retina. (Abstract of a poster presentation at the American society of human genetics annual meeting, Baltimore, November 1989). Am. J. Hum. Genet. 45(4):A230.

12. **Wong P.**, MacDonald I., and Tenniswood M. Identification of candidate genes for X-linked eye disorders; An application of lateral genetics. (Abstract of a poster presentation presented at the 6th world congress of the International Retinitis Pigmentosa Association, Dublin, Ireland, July 1990).

13. **Wong P.**, and Tenniswood M. A new cross screening methodology in the identification of human X-specific genomic clones that harbor retinal expressed sequences. (Abstract of a poster presentation at the 20th Annual FEBS meeting, Budapest, Hungary, August 1990).

14. **Wong P.**, MacDonald I., and Tenniswood M. Genetic mapping and characterization of putative X-specific genomic clones which harbor chorioretinal expressed sequences: cloning by a lateral genetics approach. (Abstract at the American society of human genetics annual meeting, Cincinatti, October, 1990). Am. J.Hum. Genet. 47(3): A265.

15. **Wong P.** Lateral genetics, an alternative strategy for the isolation of candidate genes for X-linked ophthalmic disorders. (Abstract of an oral presentation at the National Retinitis Pigmentosa Foundation workshop on molecular genetics of retinal degeneration, Cincinnati, October, 1990).

16. Tenniswood, M., **Wong, P.**, and Guenette, R.S. Co-ordinate regulation of gene expression during apoptosis (programmed cell death) in the rat ventral prostrate. (Abstract of a presentation at the annual meeting of the Society of Basic Urological Research (Society of Basic Urological Research), Toronto, June, 1991).