MOLECULAR BIOLOGY OF X-CHROMOSOME DISEASES

Zheng-Yi Chen

A Thesis Submitted for the Degree of

Doctor of Philosophy

in the University of Oxford

Keble College

and

Genetics Laboratory

Department of Biochemistry

Oxford

October, 1992

ABSTRACT

MOLECULAR BIOLOGY OF X-CHROMOSOME DISEASES

Zheng-Yi Chen, Keble College, Oxford

Submitted for the degree of Doctor of Philosophy,

October, 1992

Genomic clones were isolated and characterized using the human monoamine oxidase A (MAOA) cDNA to screen a phage library, constructed from a human 4X cell line (48, XXXX). The genomic contig derived from overlapping phage clones showed that the size of the MAOA gene is over 80 kb. Exon-containing fragments from these phage clones were subcloned and sequenced. The data from this showed that the MAOA gene consists of 15 exons.

A YAC (yeast artificial chromosome) isolated using the MAOA cDNA was characterized. This YAC was found to contain both the MAOA and the MAOB genes. Using PFGE (pulsed-field gel electrophoresis) to investigate the YAC, it was found that the MAOA and the MAOB genes are located within 50 kb and adjacent to each other. The two genes are localized in a 3'-to-3' fashion, suggesting their expression may be regulated independently. The analysis of the homology shown by the two genes clearly demonstrated that they were derived from duplication of a common ancestral gene. A CpG island was discovered to be associated with the 5' end of both genes.

A restriction map of ~2.5 Mb of genomic DNA around the MAO genes was generated by PFGE. Long-range mapping defined the physical relationship between the marker L1.28 and the MAO genes as L1.28_MAOA_MAOB.

A number of genetic diseases have been linked to the Xp11.3 region. Strong linkage was known to exist between the Norrie disease locus and L1.28. Studies showed that some of the Norrie patients have deletions encompassing the region which contains L1.28 as well as the MAO genes. Another YAC isolated by using L1.28 as the probe was also characterized.

A phage library was constructed from the L1.28 YAC and the end clones were isolated. Studies on some of the Norrie deletion patients showed that the proximal end clone of the YAC was retained in one of the deletion patients. Previous studies had shown that the Norrie disease locus was also localized proximal to the 5' end of the MAOB gene. The combined information placed the disease locus to an interval of 240 kb within the YAC.

More phage clones were characterized in order to define further the region for the Norrie locus which was finally localized within 160 kb. A YAC fragment of 160 kb was isolated and used to screen two human retinal cDNA libraries. Among the cDNAs isolated, one group was found to be deleted in some of the Norrie patients previously without any known deletion, which established their candidacy as the transcripts of the Norrie disease locus. Further characterization of the candidate gene showed that it is conserved across species. The expression of the gene was detected in various tissues. The homology shared between the NDP gene and some of the growth factor binding proteins suggests its role in neural cell proliferation and differentiation.

ACKNOWLEDGEMENT

I would like to take the opportunity to thank all the people who helped me, directly or indirectly, over the past few years. All of the materials, cell lines and probes I received for my project have been acknowledged in the text. In addition my special thanks go to the following people:

Prof. J. H. Edwards: for bringing me to the lab (rightly or wrongly) and showing his constant interest in what I am doing, also for his great help in securing the funding.

David Hunter: for always getting me whatever I need on time and our fruitful collaboration for keeping the cells alive for my fellows in the genetics lab.

Gillian Oldfield: for her being extremely patient with me; also her help with my letters, phone calls and faxes.

Val Cooper: for always letting me have my oligos right on time, even without the order number.

Saiful Said: for asking questions.

Jo Poulton: for sharing her exciting results with me. I feel so thrilled every time when I was called "an absolute angel!".

Sue Riley: for sharing the bench.

Graeme Black: for advising me what to do if you are stung by a wasp!

Karl Morten: for organizing the lab activities (treasure-hunt; skittle competition).

Eli Hatchwell, Aarti Chand, Simon Fisher: for helpful discussions and sharing the lab.

Chris Porter: for his expert assistance with the computer.

Elinor Tolley: for allowing me to use her new sequencing kit when I couldn't see any hope.

David Roberts: for his helpful advice, telling me how lucky I was for staying so long in the same place. I could have been thrown out of here a little earlier.

William Brown: for his advice on why I should or shouldn't switch to a new field, and his supply for the important lab stuff.

Chris Tyler-Smith: for helpful discussions.

Kirsti Kvaloy: for her nice cakes.

Anil Day: for his support, enthusiasm and extremely helpful discussions, and being one of the best companions I can have in the lab.

John Powell: for his patience, help and willingness to receive a call every day from me; also for being able to spend a night with us in front of a computer!

Xandra Breakefield: for letting me do whatever I want and for taking care of me during my stay in Boston.

Sally Craig: for her effort in organizing the lab gatherings, the proof-readings and driving me to the IMM.

Anne Keen and Ali McGuire: for keeping my spirit high over Robert Burns' birthday!

Heather Hinds: for listening to me; her encouragement and help even in Boston.

Ian Craig: my supervisor, for his consistent support and interest, for being able to stand behind me whenever it's needed; especially for taking me on when the time was extremely difficult for me.

Beth Battinelli: for coming to join me in my last couple of months here and being able to go through every second of the hard time with me. I wouldn't have been able to see my thesis today if she wasn't around.

Rudi Hendriks: for making the mistake of choosing to join me when the time was really difficult; for his friendship and the effort to keep me going (either up or down!). Also for his initiative in organizing the weekly pub visit (hope it can become a daily event).

Lorenzo Giusti and Irina Antia, for their understanding, great help and being close to me as well as my family.

Mark Ross, Thomas Meitinger, and Mark Jobling: my greatest gratitude must go to them - my very special friends over the years. My life would have been completely different without them.

My family, the finest family I can imagine. They are always with me.

CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

This thesis concerns studies on the human monoamine oxidase and Norrie gene loci. These are closely linked on the human X chromosome and are co-deleted in some patients. Their significance in the context of human genetic disease and of their localization on the X chromosome make them particularly interesting targets for investigation.

The Biology of the Sex Chromosomes

The human X chromosome has been the subject of many intensive studies. It is one of the two chromosomes, X and Y, in human and many mammals that in different combination can determine sex and, are of pivotal importance in normal male and female development. In the human female, the genetic constitution is two X chromosomes plus 22 pairs of autosomes, while in the male the composition is an X and Y chromosome plus 22 pairs of autosomes. Although the Y chromosome appears to be male determining, the X chromosome has many properties necessary for the success of the sex determining process.

The human X chromosome is much larger (~ 150,000 kb) than the Y chromosome (~ 45,000 kb) and the latter is largely composed of repetitive sequences (typically ~ 60%), which form the characteristic heterochromatic block at the terminal part of the long arm (Lau, 1985; Cooke et al., 1982). Whereas on the X chromosome it is estimated that genes number a few thousand, very few genes have been found on the Y chromosome.

Evolution and homology of the sex chromosomes

The X and Y chromosomes are commonly believed to have evolved from a pair of homologous, isomorphic chromosomes which have undergone a series of rearrangements. In the eutherian mammals however, cytologically, there is a lack of intermediate states between the two extreme chromosomes. Nevertheless some believe that intermediate stages can be observed in the sex chromosomes of marsupial and monotreme mammals (Graves, 1987). Cytological and molecular studies of the human X and Y chromosomes have also provided some evidence to

support their presumed evolutionary relationship. It has been observed that during male meiosis the Y short arm pairs with the distal region of the X short arm (Moses et al., 1975; Solari, 1980). This pairing process may spread to the whole of the Y short arm and a third of the X (Chandley et al., 1984). This pairing is due to the existence of homologous sequences shared between the X and Y, suggesting a common ancestral origin of the sequences on the short arms of the sex-chromosomes. These sequences show partial sex linkage and behave in a "pseudoautosomal" fashion (Burgoyne, 1982). More recent studies have revealed the existence of a second pseudoautosomal region (Freije and Schlessinger, 1992) between Xq and Yq. Pairing has been observed in this region during male meiosis (Chandley et al., 1984; Batanian and Hulten, 1987); in addition restriction mapping of the region shows extensive homology over 500kb (W.R.A.Brown, personal communication), and preliminary evidence has been presented for meiotic exchange although the frequency is unknown. This to some extent parallels the behaviour observed between the X and Y in the short arm pseudoautosomal region.

As well as the pseudoautosomal sequences other classes of X and Y homology exist, these provide further evidence for the common origin of the sex chromosomes. Part of Xp shows homology to Yq. The steroid sulfatase (STS) and Kallmann gene (KAL-X), located just outside the pseudoautosomal region at the Xp22.3, have pseudogenes in Yq11.1(Yen et al., 1987; Ballabio et al., 1992; Foote et al.; 1992). These observations support the hypothesis that a pericentric inversion brought ancestrally pseudoautosomal sequences to their present positions on the Y long arm. (Fraser et al., 1987; Yen et al., 1987)

DXYS1, and several other X/Y homologous sequences have been reported to lie outside the pseudoautosomal region. DXYS1, which is typical of a large class of X and Y homologous sequences has been localized to Yp and Xq13.2-q21.2 (Page et al., 1982). For DXYS1, the divergence between the X and Y sequences is less than that between the X copies in human and chimpanzee genomes. The hypothesis that the existence of DXYS1 on both the X and Y must have arisen more recently than the divergence of man and chimp has been confirmed by primate studies (Page et al., 1984). The primate homologues were found to be only X-linked, suggesting in the human lineage a recent transposition of material from the X to the Y less than five million

years ago (Page, 1984).

As one of the consequences of heteromorphism of the mammalian sex chromosomes, dosage compensation brought about by X inactivation has become a necessity. Ohno hypothesized that the mammalian X chromosome would be highly conserved as a linkage group because of X inactivation (Ohno, 1967). Conservation of X-linked loci has been found to cross a wide range of species with some exceptions among marsupials and monotremes (Graves, 1987; for review see Grant and Chapman, 1988).

Sex determination

Elucidation of the mechanism of sex determination has presented a major challenge for developmental biologists and it has remained a topic of considerable interest in much wider circles. Studies on abnormal sex chromosomal constitutions led to the realization of the crucial role of the Y chromosome in sex determination (Jacobs and Strong, 1959; Welshons and Russell, 1959). Maleness is only a consequence when there is a Y chromosome present, regardless of the number of X chromosomes.

There are three major lines of work that have led to our current knowledge on the organization of the Y chromosome and its role in sex determination: a meiotic map of the pseudoautosomal region which is shared between X and Y chromosomes, deletion analysis of sex-reversed XX males and XY females, and a long-range map linking the first two maps. To map genetically most genes on the Y chromosome has been impossible because of the lack of recombination between the Y and any other chromosomes. However, the existence of the pseudoautosomal region between X and Y chromosomes enables Y chromosomes to undergo obligatory homologous recombination which has helped in the construction of a meiotic map in this region of the Y chromosome (Rouyer et al., 1986; Weissenbach et al., 1987). Clearly the testis determining factor, TDF, must lie proximal to the boundary of the pseudoautosomal region. Studies on sex-reversed XX males and XY females using Y-specific probes has shown the presence of Y material in many XX males (Guellaen et al., 1984) and the deletion of similar Y material in some XY females (Disteche et al.. 1986). This provided direct evidence for aberrant

exchange between the X and Y chromosomes at meiosis, and the apparent transfer of TDF to the X chromosome (Ferguson-Smith., 1966). Long-range physical mapping linked the meiotic map and the deletion map, and furthermore precisely placed TDF in a region of 140 kb immediately proximal to the pseudoautosomal region (Pritchard et al., 1987; Page et al., 1987). Although a candidate gene, ZFY, was isolated from this region and was found to be present in some XX males and absent from one of the XY females studied (Page et al., 1987), the later finding that ZFY has a homologue, ZFX, on the human X chromosome and the observation that it escapes X inactivation (Schneider-Gadicke et al., 1989; Palmer et al., 1990) threw doubt on the primary role of ZFY in sex determination. Furthermore, the observation that the ZFY-related sequences in marsupials were absent from the X or Y chromosomes but located on the autosomes (Sinclair et al., 1988), eliminated the possibility of ZFY being the TDF, as the Y is male determining in this group too. More recently, following the search for TDF, genes, SRY and Sry, (Sinclair et al., 1990; Gubbay et al., 1990) have been isolated and they have shown great promise as candidates for TDF in human and in mouse respectively.

SRY was detected by more careful deletion mapping and by the analysis of genomic clones from the relevant region, and appears to be absolutely necessary in normal sex determination. It is Y specific, and conserved across species, including marsupial. Most direct evidence of the role of SRY in human sex-determination comes from the studies of XY females: *de novo* mutations have been found in the SRY gene of two such individuals (Berta et al., 1990; Jager et al., 1990). More significantly, in recent experiments in mice, Sry was introduced into XX female embryos and some of the genotypically XX female mice were artificially sexually reversed to the male phenotype and were found to be Sry positive (Koopman et al., 1991)

There are still situations, however, in which SRY positive XY individuals fail to exhibit male phenotype without detectable mutation (Berta., 1990; Jager et al., 1990). In the mouse sex-reversal experiment, some of the XX mice with integrated Sry failed to exhibit male phenotype (Koopman et al., 1991). This may imply complexity in the process of sex determination, and many factors may be involved in this process. It is possible that a type of cascade leads to the final determined state, and that malfunction of any step in the pathway could affect the outcome. Also

the chromosomal position of Sry may play an important role in determining its function, so that, in some locations, even an intact copy of Sry will fail to function. Future study on the exact role of the SRY locus in human and its homologues in other mammals and the factors which regulate down-stream interaction should throw insight into the complexities of sex-determination in this group.

X inactivation

In females, one of the two X chromosomes is inactivated in the early stage of embryonic development thereby equalizing the dosage between the sexes. Inactivation occurs randomly and regardless of the maternal or paternal origin of the X chromosome. From the time of inactivation, the progeny of a particular cell will always have the same X chromosome inactivated (Lyon, 1961). Large numbers of studies have confirmed the essential premise of the theory. Cytogenetic evidence indicates that in the normal female, one of the two X chromosomes is always condensed, and can been seen as a darkly stained body (Barr body) in the interphase cell (Barr and Bertram., 1949). The condensed X chromosome is late-replicating, both in mouse and human, with the exception of the tip of the short arm in human, which shows a replication pattern which is synchronised with the active counterpart (Takagi, 1974; Weber et al., 1986). This condensation is not observed in single X chromosome of males. Also, in the case of XXX/XXXX/XXY/XYYYY individuals, the condensed numbers of X chromosome are two, three, one and none respectively. Here the number of inactivated X chromosomes follows an N-1 rule, where N is the total number of X chromosomes present. Condensation is equated with gene inactivation and this mechanism, understandably, is to ensure that there is only one set of genes on the X chromosome which are expressed regardless of the number of X chromosomes.

The random inactivation of the X chromosome was originally proposed by Lyon to explain the occurrence of the variegated-coat-colour in female mice (Lyon, 1961). More direct evidence came from molecular studies on X inactivation. It was first shown, directly, by the study of Davidson et al on the expression of X-linked glucose-6-phosphate dehydrogenase (G6PD) mutant alleles in human fibroblast cells in which it was demonstrated that the expression of G6PD is confined to one of the X chromosomes and is random with respect to its maternal or paternal

derivation (Davidson *et al.* 1963).

It appears that X inactivation does not happen simultaneously along the whole X chromosome, but rather proceeds in a cis-limited fashion in which the inactivaton event spreads from the X-inactivation centre (XIC) to the rest of the X chromosome (Cattanach,1975; Migeon et al, 1968; Mohandas et al., 1987). Studies in mice have led to the finding that in X-autosome translocations, only one of the translocated chromosomes is inactive and this suggests that there is a single locus in mice from where the initiation of X inactivation starts (Cattanach, 1975). This locus has been named the X chromosome-controlling element (Xce) (Cattanach, 1975; Keer et al., 1990). In the human the locus responsible for the spreading of X inactivation (XIC) has been mapped to Xq13 (Mattei et al., 1981) - this is also the site initiating Barr-body condensation (Therman et al., 1974) and the site of cytologically observed bend on inactive X chromosomes (Fiejter et al., 1984). It seems, therefore, that there is a direct correlation between XIC, the initiation site of Barr body condensation and the bending site on the inactive X chromosome.

The mechanism of X inactivation is still unknown. Investigation of hypoxanthine-guanine phosphoribosyltransferase (HPRT) activity in a man-mouse hybrid, which was HPRT deficient, but contained a structurally normal, inactive, human X chromosome, showed that the deficiency could be compensated by reactivating the human X chromosome by treating the hybrid with 5-azacytidine, a cytidine analogue. Incorporation of 5-azacytidine leads to hypomethylation of the DNA (Mohandas et al., 1981). Treated hybrids could be grown successfully in the medium HAT (hypoxanthine, aminopterin, thymidine), which selects for HPRT expressing clones. This indicated that methylation of the X chromosome may play a role in the X inactivation. However, the expression of G6PD and PGK in the same hybrid was very limited in comparison with that of HPRT (Mohandas et al., 1981). This casts doubt on a major role for methylation in the event of X inactivation, although local effects can not be excluded. Further investigation of the expression of HPRT in different stages of embryonic development in mice showed that in early development, X inactivation actually precedes methylation, therefore ruling out the possibility that methylation has a primary role in the events of X inactivation (Lock et al., 1987)

Inactivation of X-linked genes, however, is not without exception . Using somatic cell

hybrids of mouse background with the human X chromosome as the only human material allows the expression of the human X-linked genes to be studied. Hybrids which retain only the inactive human X chromosome are particularly good systems to investigate genes which can escape X inactivation (Migeon et al., 1982). Several genes which are known to have active Y chromosomal homologues escape inactivation; these are MIC2 (Goodfellow et al., 1984), GM-R (Disteche et al., 1992), RPS4X (Fisher et al., 1990), ZFX (Schneider-Gadicke et al., 1989), it is reasonable to assume that this escape from inactivation is to provide the required dosage balance between males and females. This, however, does not apply to the genes for STS (the level of STS in the female is 1.7 times the level in the male, suggesting one copy of the STS genes only partially escapes inactivation), KAL-X (Franco et al., 1991) and A1S9T (Brown and Willard, 1989), and this suggests that the dosage balance of these loci is not of great importance. The genes which escape inactivation are interspersed along the X chromosome, suggesting that the process of inactivation is of least in part locus or regional specific.

The recent finding that both the X-specific transcripts (XIST) gene (Brown et al., 1991) on the long arm of the human X chromosome near the X-inactivation center (XIC), and the Xist gene (Borsani et al., 1991; Brockdorff et al., 1991) on the mouse X chromosome near the X controlling element (Xce), exhibit a unique pattern of expression, has provided some new clues about the phenomenon of X inactivation.

The XIST gene is expressed exclusively by inactive X chromosomes. In cell lines which have different numbers of X chromosomes, the dosage of expression of XIST increases with the number of inactive X chromosomes present. Furthermore, the high degree of heterogeneity of the transcripts indicates that XIST may not encode a protein product but rather function directly through its transcriptional product to signal inactivation via a kind of cis-limited mechanism (Brown et al., 1991). Strong support for a possible primary role of XIST in inactivation comes from the study of the mouse homologue Xist (Borsani et al., 1991; Brockdroff et al., 1991). In the mouse, Xist exhibits almost identical properties to XIST in human. Especially important is its location in the vicinity of Xce, where inactivation starts. Recently, a study of mouse Xist indicated that, as in humans, it is unlikely there is a protein product from Xist. (Ballabio and Willard, 1992).

These findings provide the basis for some very preliminary suggestions concerning the mechanism of X inactivation. However many questions remained unanswered, such as " What is its primary role in the X inactivation process? What is the function of XIST? How is the inactivation spread to the whole X chromosome? How is it decided which X chromosome is going to be inactivated? What is the mechanism behind the selective escape from X inactivation for some genes?" The mechanisms underlying X inactivation in man and mouse may not be identical. If it is supposed that XIST regulates in a sequence-specific way, or in a regional fashion, which can be applied to the explanation of escape from inactivation of genes such as MIC2, ZFX and RPS4X, why then, in the mouse, do these genes respond to the inactivation process? It is possible that the explanation lies in the observed heterogeneity between XIST and Xist transcripts which, for XIST, presents as a smear hybridization pattern on Northern blots in the range from over 200 base pairs to over 10 kb, whereas Xist gives more distinct major bands of 11.3 and 12.6 kb. It is possible that human and mouse transcripts differ in function. Much exciting and challenging research awaits those interested in unravelling the complexities of the inactivation process.

Human Genetic Disease and the Human Genome

It is estimated that the size of the human haploid genome comprises $3x10^9$ base pairs(bp) which is distributed between the 23 chromosomes. The size of each chromosome varies greatly, from the biggest, chromosome one - 263 Mb (million base pairs) to that with smallest unique DNA content, chromosome Y - 45Mb (Morton, 1991). It is also estimated that there are 30,000 - 100,000 genes encoded in the human genome. Although the accuracy of these figures is still far from clear, progress from studies on other organisms including yeast (Oliver et al., 1992), *C. elegans* (Waterston et al., 1992) *Drosophila* (Hoheisel et al., 1992) and mice (Dracopoli et al., 1992) suggest they are broadly correct. The consequences of the human genome initiative (Jordan, 1992; Chumakov et al., 1992; Foote et al., 1992) are providing, with unprecedented speed, a better understanding of the human genome and its relationship to those of other organisms.

In *Mendelian Inheritance in Man*, over 5000 single gene disorders have been described

(McKusick, 1991). With increasing numbers of genes being identified and with much clearer clinical classification for disorders which may differ subtly, more genetic disorders are likely to be identified.

Although it would be foolish to underestimate the contribution of earlier work, it can be said that human genetics has been revolutionized by the advent of molecular techniques - the so called "new genetics". (Weatherall, 1990). Following the discovery of 46 human chromosomes (Tijo & Levan, 1956), the development of human genetics has progressed through a series of leaps forward, often preceded by the development of new techniques and enhanced understanding of biological phenomena. Starting with the advent of reliable karyotyping of human chromosomes, numerical chromosomal abnormalities could easily be recognized. This brought about the recognition of Down's syndrome (due to trisomy 21, Lejeune, 1953), Turner syndrome (due to 45, XO) and numerous other cases of aneuploidy. Also, as discussed previously, it was realized that the presence of the Y chromosome is essential for sex determination, following the discovery of 47, XXY Klinefelter syndrome individuals who were phenotypically close to normal males, and the finding that mice with only a single X chromosome were female (Jacobs and Strong, 1959; Welshons and Russell, 1959). With the emergence of chromosome banding techniques (Arrighi and Hsu, 1971) in 1970, a wide range of structural abnormalities of chromosomes became apparent. This led to the discovery that chronic myelocytic leukemias were caused by a reciprocal translocation between chromosomes 22 and 9; the translocated chromosome 22 was named Philadelphia chromosome. The advent of somatic cell genetics, developing at about the same time, has provided a powerful tool for the study of human genetics. Routine production of somatic hybrids from the fusion of man and mouse cells became reality. This development greatly enhanced the work of gene mapping, one of the first localisations being that of thymidine kinase gene to chromosome 17 (Weiss & Green, 1967). *In situ* hybridization (Pardue and Gall, 1969) has now developed to a stage where, for many purposes, it is the perfect means to localize genes. These approaches have formed a firm base for the application of genetics to human genetic diseases (McKusick, 1992).

Among all of the human genetic diseases, the best characterized are the single gene

disorders; this is the direct consequence of the applicability of modern molecular genetic approaches. Depending on their pattern of inheritance, they can be characterized as autosomal dominant, autosomal recessive and X-linked. Autosomal dominant and recessive disorders can be distinguished on the basis that, whereas the dominant diseases are manifested in heterozygotes, the recessive manifestation can only be found in individuals homozygous for the mutant allele. In reality, drawing simple conclusions may not be so straightforward due to, in some cases, decreased penetrance. The term penetrance refers to the clinical expression, or lack of it, of the mutant gene.(Gelehrter and Collins, 1990). The X-linked traits and their associated diseases are of special interest because of their transmission pattern as well as their association with X-inactivation. The unique pattern of inheritance of X-linked recessive disorders is one in which the female carrier passes her defective gene to only half of her sons who will be affected and to half of her daughters who will be carriers in turn. Observation of this pattern has made it possible to assign a large number of disease genes to the X chromosome without exhaustive linkage analysis.

Another interesting aspect associated with X chromosomal disorders is X-inactivation (for details, see section on X-inactivation). Due to the random nature of X-inactivation, recessive alleles can be expressed in certain population of cells even in heterozygotes. Non-random X-inactivation has been observed in some disorders. This can be caused by some chromosomal rearrangements as in the case for Duchenne muscular dystrophy (DMD). Many female DMD patients have been found to have balanced translocations involving the X chromosome at Xp21 and an autosome. However, the translocated X chromosome is generally preferentially active and the normal X chromosome undergoes inactivation (Verellen-Pumoulin et al., 1984). This observation provided compelling evidence to support the localization of DMD to the short arm of the X chromosome at Xp21 (Monaco et al., 1985), also it provides an example of a recessive disorder manifesting in females. For X-linked immunodeficiency diseases, it is found that female carriers undergo an unilateral X chromosome inactivation in the primary affected cell lineages which can be used for carrier detection and prenatal diagnosis (Hendriks, 1991).

Recent cloning of some disease genes has provided insight into the mechanism of

previously inexplicable biological phenomena, promises to increase our understanding of genome biology as a whole.

One of the most unexpected developments came from cloning of the fragile-X gene. This has been shown to contain a stretch of unstable trinucleotide repeats (CCG)n. The amplification of the copy numbers of (CCG)n which do not follow Mendelian features in their inheritance pattern, implies a novel mechanism underlying the cause of fragile-X syndrome. Following this discovery, a similar type of amplification of triplets has also been observed to be the cause in some other genetic disorders such as myotonic dystrophy (DM) and Kennedy disease (Davies, 1992). The amplification of such unstable DNA sequences has helped to elucidate a long standing problem in the field relating to anticipation. "Anticipation" is the term describing the apparent occurrence of an inherited disorder with a progressively earlier age of onset in successive generations. In the case of fragile-X, this is referred to as the Sherman paradox (Sherman et al., 1984; 1985). Normal X chromosomes have between 6 and 60 copies of the repeat. For the fragile-X syndrome, carrier males have an increased copy number (between 60 and 200.); affected individuals have more than 200 copies (see review, Richards and Sutherland, 1992). This may suggest that the amplification of certain triplets may be a rather general mechanism underlying some genetic defects (Davies, 1992). Also, equally importantly, the isolation of the fragile-X gene has started to bear fruit in the context of its application to medicine. It offers an opportunity for accurate, efficient and cost-effective clinical diagnosis of a syndrome which is the second largest cause of mental retardation.

Identification and isolation of human disease genes

Progress in the isolation of genes responsible for genetic disorders has been considerable over the past few years, aided both by our understanding of the human genome and those of other organisms, and by the employment of new techniques. Until recently, very few options were available for the identification and isolation of disease genes. Most of the diseases whose genes were cloned in this early period were the consequences of previously known biochemical lesions. The approach adopted would involve characterization of the biochemical defects and purification and characterization of protein products. This would be followed by partial

determination of the sequence of the protein which could be used to deduce its DNA sequence, thereby leading to the isolation of the corresponding gene by the use of a hybridization probe. This approach has been successfully used in the study of diseases such as Christmas disease (Anson et al., 1984).

This type of direct strategy, however, is limited by the fact that the majority of single gene disorders are without any known biochemical defect. In facing the increasing importance attributed to human genetic diseases in medicine, effective alternatives are required to improve and complement those methods already existing.

The "reverse genetic" approach basically consists of two main areas, namely the "positional cloning" and the "candidate gene" approach. The "positional cloning" strategy is to clone a gene from knowledge of its location in the genome only. To use this approach successfully, it is essential to map a disease locus as finely and precisely as possible. This normally starts with the use of linkage analysis to map a disease locus segregating within a collection of disease families using a series of informative probes (Ott, 1986). Linkage analysis is performed by calculating the likelihood that two loci are linked. A likelihood ratio is established:

Likelihood ratio (θ)= likelihood (data/θ) / likelihood (data/ no linkage)

This ratio is expressed logarithmically: the lod score (logarithm of the odds). The most likely recombination fraction ($\hat{\theta}$) is the value of θ which gives the highest positive lod score ($\hat{Z}$). Positive lod scores suggest the presence of linkage, whereas negative lod scores indicate linkage is unlikely. Normally a lod score of +3, or greater, is considered definite evidence of linkage, indicating 1000 to 1 odds that the link pattern of segregation observed did not occur by chance. A lod score of <-2 (indicating 100 to 1 or greater odds against linkage) is taken as evidence of exclusion of linkage for this value of θ. The use of linkage analysis has been greatly facilitated through the advent of RFLPs (restriction fragment length polymorphism) and other polymorphic markers (such as microsatellite (AC)n repeats).

An RFLP is a variation in DNA sequence that alters the length either through presence or absence of certain restriction sites, or as a result of variation in the copy number of short tandem

repeats within the fragment (Solomon and Bodmer, 1979; Botstein et al., 1980). Microsatellite (AC)n repeats are stretches of dinucleotide sequences which are abundantly distributed in the mammalian genome (in the human genome, once every 50 kb on average). The copy numbers of the repeats vary between individuals and number variants follow Mendelian inheritance within families.

By examining the co-segregation of a series of polymorphic markers and an unidentified genetic disease, the genetic map can be established until the linkage between the disease locus and one or more of the markers is evident. Practically, however, it is unusual to have more than 100 meioses in the pedigree material, and linkage analysis therefore is limited to placing a locus in an interval of 1 centimorgan (cM) which corresponds, roughly, to 1 million bases (with some variations).

Following linkage analysis, physical mapping in order to narrow down the region further usually can be achieved through long-range mapping of genomic DNA using pulsed-field gel electrophoresis (PFGE) and by characterizing genomic clones isolated from that region. Since its first description (Schwartz et al., 1984; Carle and Olson, 1984). PFGE has been used for a variety of purposes. In particular, it has played a significant role in the identification of disease genes. PFGE in conjunction with the use of the "CpG sensitive" restriction enzymes to digest the human genomic DNA will allow long-range physical mapping. In the mammalian genome, the frequency of the CpG dinucleotide is only a fifth of that expected from considerations of C+G content alone. This is due to the high mutation rate of cytosine to thymine as the consequence of methylation of cytosine and its subsequent deamination (for review, see Bird, 1987). Most of the CpGs are clustered close to the 5' end promoter region of "housekeeping genes" and they are not methylated (with the exception of inactive X-chromosomes). These clusters are called "CpG islands" and can provide direct indication of the location of those island-associated genes (Bird, 1986); Also, due to the scarcity of unmethylated CpG in the genome, use of "CpG sensitive" enzymes tends to generate very large fragments. Two or more markers relatively far apart in the genome can be linked physically in this way. PFGE can then be used to resolve these fragments and display them for further analysis (such as Southern blotting). The uses of the PFGE to link

markers (see chapter 2), to detect deletions distant from the available markers and to directly locate "CpG" islands and therefore their associated genes (Ross et al, 1990) have all been very successful approaches

In most cases, however, the linking of flanking markers close to a disease locus marks only the end of first stage. The next step will be to isolate the genomic sequences from the target region for detailed analysis. Although various cloning systems exist for this purpose such as phage and cosmids which have to be propagated in *E.coli*, the most powerful is the yeast artificial chromosomes (YACs) system (Burke et al., 1987; Little et al., 1989). These are plasmid vectors which contain sequences allowing them to function as a linear chromosomes in yeast cells. They are capable of carrying inserts larger in size to any other cloning system; The size of fragments accommodated ranges from 50kb to a few Mb, which makes it easier to analyse the physical relationship of markers at great distances. It is also important that some classes of human DNA sequences which undergo rearrangement in conventional cloning systems (normally due to some repeats which can form a secondary structure leading to instability during propagation) can be cloned into YAC vectors. This is due to the fact that the yeast is itself a eukaryote, and that its genome contains both palindromic (Klein & Welch, 1980) and other repeated DNA (Cramer et al., 1976) so it is more tolerant of cloned eukaryote DNA than is *E.coli*. To cover the target region as fully as possible, more than one YAC is normally needed in order to build a genomic contig. Walking from one YAC to another can be achieved through generating STSs (sequence-tagged sites. Olson et al., 1989) from the ends of the YAC which in turn can be used to isolate other overlapping YACs. Since any genomic contig built on the basis of using flanking markers still tends to be very big, any gross chromosomal rearrangements within the region which are disease associated are of considerable importance in providing a short cut to identifying candidate genes. This was demonstrated in the search for Neurofibromatosis type I gene where two translocations helped to defined the region encompassing the NF gene (Cawthon et al., 1990).

Expressed sequences can subsequently be isolated from the region of interest through various means, such as by the search for conserved sequences across species (Franco et al; 1991); by the use of exon-trapping systems on genomic clones (Buckler et al., 1991), the use of

large genomic clones to screen appropriate cDNA libraries (Groden et al., 1991), or by sequencing the whole region and then subjecting the results to computer analysis (Legouis et al., 1991). Transcripts derived from the region can be analysed for alterations in individuals who have the disease. Tissue expression patterns can also be used to provide circumstantial support for the candidacy of the cloned gene. The ultimate proof of identification of a candidate with the disease locus lies with the characterization of mutations and information from the sequence analysis. Ample evidence for the success of the approach comes from the isolation of such genes as cystic fibrosis (Riordan et al., 1989), neurofibromatosis typel (Cawthon et al., 1990) and Wilms' tumor (Rose et al., 1990). Although the cloning of NF and Wilm's tumor genes was greatly facilitated by the chromosomal rearrangements including the gene, the isolation of CF gene was based purely on its location and the identification of mutations.

Another increasingly important strategy to isolate disease genes is the candidate gene approach. This is based on the co-localization on a chromosome of a disease locus and a gene for a protein plausibly involved in the disorder. With increasing numbers of transcripts being identified for each chromosome, together with the progress from the Human Genome Project it will, conceivably, become one of the dominant approaches. Its applicability has been enhanced by the studies of genomes from other organisms, especially that of the mouse. The identification of the mouse mutant *Splotch*, due to the mutation of the Pax-3 gene, has led the isolation of the gene for its human homologue which underlies Waardenburg's syndrome - an autosomal dominant combination of deafness and pigment defects (Tassabehji et al., 1992 Baldwin et al., 1992). In the early-onset version of Alzheimer's disease, the disorder was mapped to chromosome 21. Amyloid precursor protein (APP), which is believed to be involved in synaptic contact and acts as a growth regulating factor and is the product of the B-amyloid gene, was also localized to the same region. Mutations identified from APP in some of the Alzheimer patients demonstrated APP's probable etiological involvement in this type of Alzheimer's disease (Hardy,1992).

Both strategies, are not without restraints. For positional cloning, without gross alteration of the gene, it is often difficult to find the real disease gene from among the many transcripts,

This has been clearly demonstrated in the pursuit of the Huntington's gene. Huntington's disease is an autosomal dominant degenerative neurological disorder characterized by abnormal movements and progressive loss of mental function. Linkage data has placed the gene to a region of roughly around 2 Mb on the short arm of chromosome 4 (Bates et al., 1992). This region may contain over 100 genes, therefore to find out which one of them is the disease gene is a formidable task.

The "candidate gene" alternative can be limited in its present form because our understanding of the possible functions of many protein products is very scant and the number of transcripts available for analysis is still extremely low. With the rapid development of this field, further approaches may emerge as useful alternatives, such as the functional complementation employed in isolating the gene for Fanconi's anaemia (Strathdee et al., 1992).

If it can be said that recent development has been very impressive on single gene disorders, the progress on multifactorial disorders has been extremely slow in comparison. New concepts and new approaches are both required to ease the situation. There are already some signs that both may come sooner rather than later. (Jacob et al., 1991).

The biology of monoamine oxidase(MAO)

Human monoamine oxidase has long been studied because of its function in oxidising aromatic amines, dietary as well as tertiary amines. The discovery of an enzyme in mammalian liver which oxidised tyramine with the evolution of hydrogen peroxide and ammonia (Hare, 1928) marked the beginning of research into momoamine oxidase.

Monoamine oxidase is the principal enzyme catalyzing the oxidative degradation of biogenic amines, including neurotransmitters such as catecholamines and serotonin, in the central nervous system and in peripheral tissues (Murphy, 1978; Glover and Sandler, 1986; Weyler et al., 1990). It was evident from early studies that there are at least two forms of MAO: MAO-A and MAO-B. They can be distinguished on the basis of substrate specificity and sensitivity to different drugs: MAO-A has a higher affinity for catecholamines and serotonin and is more sensitive to inhibition by clorgyline (Johnston, 1968), whereas MAO-B has a higher affinity

for the dietary amines such as phenylethylamine and is selectively inhibited by deprenyl (Knoll and Magyar, 1972). Also they were shown to have different molecular weights, peptide proteolytic digest products and antigenic determinants (McCauley and Racker, 1973; Cawthon and Breakefield, 1979; Denney and Denney, 1985; Brown et al., 1980). The enzymes are located at outer membrane of mitochondria (Schnaitman et al., 1967) and present in a wide range of tissues, with different levels in different cells. Both forms can be found in the liver in the human, for example; whereas in human brain it is almost exclusively the B form (80-95%). The two forms also exhibit different activities during development as seen for example in both human and rodent nervous system, where the A form appears first , followed by the B form (Mantle et al., 1976; Lewinsohn et al., 1980; and Tsang et al., 1986).

Historically, the main interest in MAOs came from the biochemical and pharmacological studies demonstrating that inhibitors of the enzymes can be used as antidepressants (Murphy et al., 1984). Disruption in catecholaminergic and serotonergic pathways in the brain have been hypothesized to have a critical role in many psychiatric and neurologic diseases, including schizophrenia, manic-depressive illness, autism, alcoholism and Parkinson disease. The evidence to support such theories came from the use of drugs which interfere with aminergic transmission in treatment of these diseases, from findings of altered amine metabolite levels in plasma and CSF of patients and from the involvement of these transmitter systems in limbic system functions. Recently cDNAs for both forms of MAO were isolated (Bach et al., 1988; Hsu et al., 1988). The two genes have extensive homology, both at DNA and protein levels, with average homology over 70%, whereas in some parts of the flavin-binding domain the homology is greater than 90%, which suggests strongly that they were derived from duplication of a common ancestral gene. Genes for both forms were mapped to human X chromosome Xp11.3 (Pintar et al., 1981; Levy et al., 1989; Ozelius et al., 1988) and also to the mouse X chromosome (Laval et al., 1991). Because of the possible association between altered MAO activities and neuropsychiatric disorders (Hsu et al., 1989) studies on the genetics of MAO have been intensified, especially in the light of linkage with other diseases. Recombination studies showed that MAOs are linked to a few important genetic diseases such as Norrie disease (Lan et al., 1989; Sims et al., 1989 a,b), X-linked retinitis pigmentosa 2 (Riley et al., 1991), Leber's hereditary

optic neuroretinopathy (LHON; Vilkki et al., 1991). It has been shown that statistically significant differences of MAO-A activities are associated with a different RFLP allele caused by substitution of the third base of a triplet codon. This suggests that the MAO-A gene itself is a major determinant of activity levels.(Hotamisligil and Breakefield, 1991). Overexpressed MAO-B gene in transgenic mice leads to profound neurological changes and behaviour problems as well as early death (Aderson, personal communication). Complete loss of MAO genes has also been observed in some atypical Norrie deletion patients (see below) who have disrupted amine metabolism together with sever mental retardation. Further studies on metabolite levels in various disorders and the regulation of the genes should help to determine the extent to which MAO aberrations underlie human genetic disease.

Norrie disease

Norrie disease (MuKusick 310600, 1990) is a severe neurological disorder. The characterization of Norrie disease is based on the blindness from infancy due to retinal dysplasia followed by retinal detachment and bulbar atrophy. From family studies of the disease, it is clear that it is an X-linked recesssive disorder (Warburg, 1966). In typical Norrie patients (who have exclusively been males from reported studies), two thirds of them have some degree of mental retardation and one third of them have progressive sensorineural hearing loss. Marked variability in the severity of retardation and the onset and severity of hearing loss within the same families has also been observed (Collins et al., 1992).

The gene for Norrie disease was mapped to Xp11.3-Xp11.4 through linkage study of a number of Norrie families (Gal et al., 1985; Bleeker-Wagemakers., 1985). Strong linkage was observed between the disease locus and the probe L1.28 (DXS7), which is located on Xp11.4. Support for the localization of the gene also came from a study on a chromosome abnormality involving a translocation. Ohba and Yamashita (1986) described a female patient who manifests the disease with reciprocal translocation between the X chromosome and chromosome 10. The position where the translocation occurs on the X chromosome, Xp11.3, is consistent with the linkage data. More direct evidence to support the localization of Norrie locus came from studies on Norrie deletion patients. Gal et al (1986) reported a Norrie family in which a patient was found

to have submicroscopic deletion encompassing L1.28. Apart from classic manifestations such as retinal detachment, mental retardation and gradual loss of hearing, this patient also exhibits other series of symptoms including hypogonadism, growth failure, severe mental retardation and increased susceptibility to infection. More cases of Norrie patients having submicroscopic deletion for L1.28 were reported later (de la Chapelle et al, 1985; Zhu et al., 1989 and Donnai et al., 1988) All of the deletion patients described suffer from a wide range of abnormalities with different degrees of severity. These include microcephaly, seizures, altered peripheral autonomic function, sleep disturbance, myoclonus, delayed sexual maturation.

These deletion patients were later found to have MAO genes (both A and B forms) deleted as well. Because of the importance of MAO in neurological metabolism, it is conceivable that the deficiency in MAO may have the causal role in the manifestation of Norrie disease. Sims (1989) et al studied the MAO levels in those classic Norrie patients and showed that their MAO levels are within the normal range of variation as the controls. Their studies excluded MAO as the possible gene responsible for simple NDP. The manifestation of complex symptoms in atypical Norrie deletion patients probably represents the consequences of co-deletion of more than just the Norrie gene, representing cases of a "contiguous gene syndrome"(Emanuel, 1988).

Diergaarde et al. (1990) mapped the deletion breakpoints in one patient to the region between probes px59 (recognised by DXS77) and L1,28. Further study by Lan et al (1989) showed in one Norrie family, a single recombination event between L1.28 and NDP, but not between OAT (located proximal to L1.28) and NDP. This result is of considerable importance since it places NDP locus proximal to L1.28, which has accelerated the search for the NDP gene (see chapter 4 and 5).

Outline of the research

The aim of this project was to study the MAO and Norrie genes on the short arm of the X chromosome, thereby providing information on the genomic organization as well as genetic aspects of the genes which may be of value in disease etiology.

Chapter 2 describes the work on the characterization of the MAO-A genomic

organization. This provided reagents with which to study variation of the gene in different populations as part of a strategy to understand its etiological role.

Chapter 3 describes the characterization of a YAC encompassing both MAO genes. It was possible to demonstrate that the genes were derived from duplication of a common ancestor. The two MAO genes are adjacent to each other and organized in a tail-to-tail fashion. This observation is relevant to an understanding of the mechanism underlying gene duplication. A long range map surrounding the MAO genes was constructed, providing a starting point to localize the gene for Norrie disease. Certain features of the genes such as their CpG islands and the possible escape from inactivation for the MAO-B gene are discussed.

In Chapter 4, another YAC of 650 kb, isolated using the probe L1.28, is described. The YAC provided very important information for pinpointing the region for the Norrie disease locus. The mapping data derived from this YAC confirmed the data derived from the MAO YAC in chapter 3, as well as some of the data from long-range mapping of genomic DNA. More importantly, due to the retention of one of the YAC ends in one of the Norrie deletion patients, the proximal end of a region containing the Norrie gene was defined.

Chapter 5 presents work to isolate a candidate gene for Norrie disease via a positional cloning approach. Based on the information from Chapter 4, a YAC restriction fragment of 160kb, which is the obligatory region for the Norrie locus, was used to screen two human retinal cDNA libraries. Three groups of cDNAs were isolated by using the fragment. Further characterization of these cDNAs led to identification of one group representing the candidate locus for the disease gene, i.e., they were deleted from those classic Norrie patients without previously known deletions. The open reading frame for the gene is very small (400 bp) encoding a 133 amino acid peptide which is highly charged, suggesting a secretory protein. No homology with other proteins could be identified during this stage.

The final experimental chapter of the thesis is concentrated on the further characterization of the NDP gene. This demonstrated that NDP gene is 28 kb in size and consists of three exons with the first not used in translation. Detailed mapping of deletion patients shows

that there is a heterogeneity in deletion size and position. Also, further search for homology reveals that the C-terminal part of NDP protein is very cysteine rich and the position for these residues is very conserved among mucin, Van Willebrand factor and growth factor binding proteins (cyr-61 in mouse and CEF-10 in chick). The N-terminal part of the protein has a signal peptide suggesting post-translation modification of the protein, supporting the rationale of the gene product being a secretory protein. Because of the potential for cysteine to form di-sulphide bonds, it is postulated that the function of the protein may involve interaction with other growth factors which are important to the development of the eye as well as the brain.

Part of the project has involved cloning of the mouse homologue for the MAO-A gene which was mapped to the mouse X chromosome in relation to other markers from the region. To be able to provide more reagents for linkage analysis for other genetic diseases, one simple microsatellite (CA)n from the second intron of the MAO-A was characterized. Another microsatellite repeat together with a VNTR in the first intron of the MAO-A gene was also isolated and characterized. Finally, by coupling the VNTR with the determination of the methylation status of the CpG island at the promoter region of the MAO-A gene, an assay for X-inactivation status has been developed.

Nucleic Acids Research, Vol. 19, No. 16 4537—4541

Chapter 2. Structure of the human gene for monoamine oxidase type A

Zheng-Yi Chen, Gokhan S.Hotamisligil, Jenq-Kuen Huang, Lisa Wen, Diala Ezzeddine, Nese Aydin-Muderrisoglu, John F.Powell, Rosa H.Huang, Xandra O.Breakefield, Ian Craig and Yun-Pung Paul Hsu

ABSTRACT

Monoamine oxidases, type A and type B, are principal enzymes for the degradation of biogenic amines, including catecholamines and serotonin. These isozymes have been implicated in neuropsychiatric disorders. Previously, cDNA clones for both MAO-A and MAO-B have been sequenced and the genes encoding them have been localized to human chromosome Xp11.23 - Xp11.4. In this work, we isolated human genomic clones spanning almost all the MAOA gene from cosmid and phage libraries using a cDNA probe for MAO-A. Restriction mapping and sequencing show that the human MAOA gene extends over 70 kb and is composed of 15 exons. The exon structure of human MAOA is similar to that described by others for human MAOB. Exon 12 (bearing the codon for cysteine, which carries the covalently bound FAD cofactor) and exon 13 are highly conserved between human MAOA and MAOB genes (92% at the amino acid level). Earlier work revealed two species of MAO-A mRNA, 2.1 kb and 4.5 - 5.5 kb. We now report on further cDNA isolation and sequencing, which demonstrates that the longer message has an extension of 2.2 kb in the 3' noncoding region. This extended region is contained entirely within exon 15. The two messages therefore appear to be generated by the use of two alternative polyadenylation sites. Results from the present work should facilitate the mutational analysis of functional domains of MAO-A and MAO-B. Knowledge of the gene structure will also help in evaluating the role of genetic variations in MAO-A in human disease through the use of genomic DNA, which is more accessible than the RNA, as a template for PCR-amplification and sequencing.

INTRODUCTION

Monoamine oxidases (MAO; monoamine: O_2 oxidoreductase; EC1.4.3.4.), type A and type B, are principal enzymes for the degradation of biogenic amines, (for review see 1,2,3). They catalyze the oxidative deamination of amine neurotransmitters such as dopamine, norepinephrine, epinephrine and serotonin, as well as dietary amines such as tryptamine and phenylethylamine. These enzymes are located in the outer membrane of mitochondria and are present in all types of cells, albeit at widely varying levels. Intensive biochemical and pharmacological studies of MAO have been stimulated from findings that inhibitors of the enzymes can be used as antidepressants (4). These studies have led to the characterization of two forms of the enzyme, MAO-A and MAO-B, which differ in specificity for substrates, sensitivity to inhibitors, tissue distribution, antigenic determinants, protein size and peptide maps (for review see 1, 2, 3). It is now clear from cDNA clones that these MAO isozymes are encoded by separate genes, which share about 70% overall homology in amino acid sequence (5—9).

Both human MAO genes have been mapped to the chromosome X in the p11.23-11.4 region (10—14). Absence of both genes has been described in atypical, male Norrie disease patients from four different families, who have a submicroscopic deletion in this region of the X chromosome (13, 15—18; Z.Chen and I.Craig, unpublished data; F.Collins and S.Antonarakis, manuscript submitted). It is conceivable that some features of their disease phenotype, such as microcephaly, mental retardation, seizures, and sleep disturbances, may result from the complete absence of MAO enzymatic activities in these patients throughout development and in later life. Other studies have shown that variant activities of platelet MAO-B are statistically associated with a number of neuropsychiatric diseases including affective disorders (19, 20), schizophrenia (21) and alcoholism (22, 23). A potential role of MAO in diseases of the nervous system is further suggested by findings that MAO-B can convert MPTP into a neurotoxin that causes neurodegeneration similar to Parkinson disease (24). In addition, deprenyl, an inhibitor of MAO-B, appears to be effective in slowing down the progression of Parkinson disease (25, 26). Allele association studies of MAO-A indicate that genetic variations at the structural locus are major determinants of activity levels in human skin fibroblasts (27). The structure of the human *MAOA*

gene is reported here as a basis for studying the regulation of enzyme activity and its etiological role in human diseases.

MATERIAL AND METHODS

Isolation of genomic DNA and cDNA clones:

Genomic clone A2 was isolated from a cosmid library prepared from a Sau IIIA partial digest of human genomic DNA (gift from Dr. Patricia Watkins, Integrated Genetics, Framingham, MA) established in the vector c2XB (28). A full length human cDNA clone for MAO-A, HMII (6) was used as a probe, unless otherwise indicated. Genomic clone 5-5C (gift from Ms. Laurie Ozelius, Massachusetts General Hospital) was isolated from a cosmid library prepared from a Sau IIIA partial digest of genomic DNA from a homozygotic patient of Huntington disease (gift from Dr. James Gusella, prepared by Stratagene, La Jolla, CA) cloned in the vector pWE15 (29). Genomic clones 6.12, common, 4.15, 1.23 and 6.1 were isolated from an amplified human library prepared from an EcoRI partial digest of genomic DNA from a 4X female (gift of Dr. Dereck Blake, Oxford, U.K.) cloned in EMBL vectors (Promega). Genomic clone gMAO-AI was isolated from a human lymphocyte lambda DASH genomic library (gift of Dr. Lily Hsu, Beckman Res. Inst., City of Hope, Duarte, CA, prepared by Stratagene) and probed with a 2.8 kb cDNA clone, pMAOAI, from a human placental cDNA library (prepared by Clontech). Two cDNA clones, CP221 and CP223, were isolated from a human placental cDNA library in lambda gt10 (Clontech Laboratories, Inc., Palo Alto, CA). Colony or plaque screening was done according to standard procedures (30).

Southern hybridization and restriction mapping:

Procedures for blotting of DNA onto nylon filters, hybridization and autoradiography have been described previously (12). DNA probes were labeled with ^{32}p dCTP by random hexamer priming (31). The orientation of the MAO-A coding region relative to the vector was determined by probing with fragments from the 5' and 3' ends of HM 11, a 0.15 kb AvaI fragment and a 0.52 kb AvaI fragment, respectively. Restriction mapping of clones 6.12, common, 4.15, 1.23 and 6.1 was done by single and double digestions with restriction enzymes, EcoRI, HindIII, SalI and BglII. Fragments were resolved by agarose gel electrophoresis, and the orientation was determined by probing with gel-purified fragments of SacI digested HMII cDNA, which yielded three fragments of 528, 648 and 788 bp in the 5' to 3' direction. The sizes of introns 1, 3 and 5 were determined by hybridizing oligoprimers (Milligen-BioSearch/Cyclone DNA synthesizer) derived from exon 2, 4

and 5 to a Southern blot of DNA from a human MAO-A-bearing YAC (pYAC4 from Center for Genetics in Medicine, Washington Univ., St. Louis, MO) partially digested with rare cutting enzymes and resolved by pulse-field gel electrophoresis (Z.-Y.Chen and I.Craig, manuscript in preparation).

Analysis of DNA sequence:

DNA sequences near intron-exon junctions were determined either by subcloning genomic fragments into appropriate vectors followed by sequencing; or by PCR-amplification of total genomic DNA and direct sequencing. These two approaches are briefly described as follows: (A) DNA fragments from restriction digestion were cloned into M13 (32), Bluescript KS (Stratagene, La Jolla, CA) pUC9 or pUCl9 (33) vectors and sequenced by the chain-termination method (34) using modified T7 DNA polymerase (35; Sequenase Version 2.0 sequencing kit, United States Biochemical Co., Cleveland, OH) and [alpha-^{35}S]-dATP (36) with chain elongation on ice. Double stranded DNA sequencing was done by the alkaline-denaturation method (37). (B) Genomic DNA was amplified by polymerase chain reactions (38) using thermostable *Taq* DNA polymerase (GeneAmp kit, Perkin-Elmer Cetus, CA) according to manufacturer's instruction. Typically the reactions were done in 50 ml total volume in the presence of 200 pmoles of primers in 30 cycles (1 min at 94°C, 2 min at 55°C and 3 min at 72°C) with a final extension for 10 min. The amplified DNA fragments were separated by electrophoresis on a 2% low melting temperature agarose (SeaPlaque or NuSieve GTG agarose from FMC). Desired DNA bands were electroeluted onto DEAE membrane (Schleicher & Schuell), as described (39); or excised, melted and used subsequently for direct sequencing according to a simplified method (40) with minor modifications using Taq polymerase and [alpha-^{32}P]d-CTP, or Sequenase (41) with minor modifications and [alpha-^{35}S]dATP.

RESULTS

We have isolated genomic clones from three libraries (see Material and Methods). Clones 6.12, common, 4.15, 1.23 and 6.1 are from a phage library and cover most of the MAOA gene. All clones, together with those isolated from other libraries, including A2, gMAO-A1 and *5-5C* were aligned to yield a region for MAOA covering about 70 kb. These clones were analyzed by Southern analysis of restriction digests using various fragments from the cDNA, HM 11, as probes. Detailed alignment was done by superimposition of restriction sites for EcoRI, HindIII, BglII and SalI. A composite restriction map for these enzymes is shown in Figure 2.1. Exon 1 contains 5'-noncoding sequences corresponding to the cDNA clone hMAO A-7 (5) (Figure 2.2). The 5' end of the first exon has not been defined precisely, but a potential TATA box appears 194 bp from the 5' end of the published cDNA sequence, which may serve as a promoter element. A comparison between the two published cDNA sequences, HM11 (6) and hMAO A-7 (5), revealed that a 35 nucleotide segment in the 5'-noncoding region of HMII (from nucleotide no. 8 to no. 42 in that reference) is an inverted repeat of a segment in hMAO A-7 (from nucleotide no. 31 to 65 in that reference). Since sequences from hMAO A-7 have also been found in genomic clones (Figure 2.2), it appears those in HM11 are a cloning artifact.

We have determined the exon structure of *MAOA* by two approaches. (A) Genomic clones were subcloned into sequencing vectors and sequenced. (B) Genomic DNA fragments were amplified using the polymerase chain reaction with appropriate primers from the cDNA sequence to generate DNA segments for direct sequencing. The latter approach was feasible only for segments containing exons and introns totalling less than 3 kb. The structure of the exons and nucleotide sequences near the intron/exon junctions are shown in figure 2.3. All exon-intron junctions showed appropriate consensus signals for splicing; GT at the 5 ' end of introns and AG at the 3 ' end. The gene contains 15 exons. All but exon 4 were analyzed in this study. The approximate size of the introns 1, 3 and 5 were determined by Southern blot hybridization to fragments generated by pulsed field gel electrophoresis of a YAC clone containing human MAOA (provided by Dr. Bernard Brownstein, Washington Univ.) digested with rare cutting enzymes (Z.-Y.Chen and I.Craig, manuscript in preparation). The size of other introns was evaluated by PCR

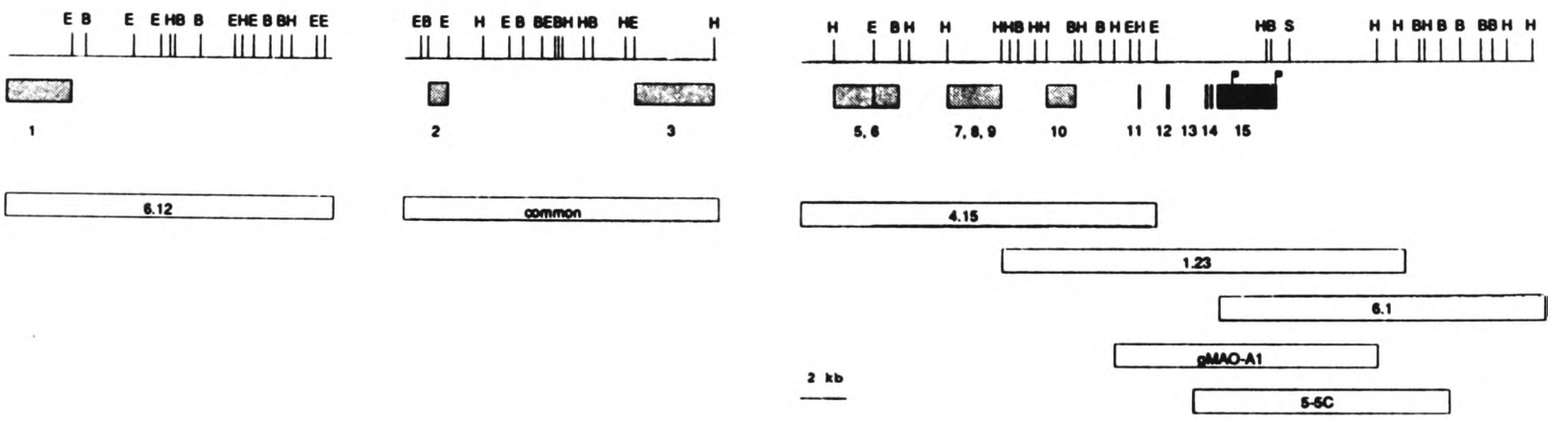
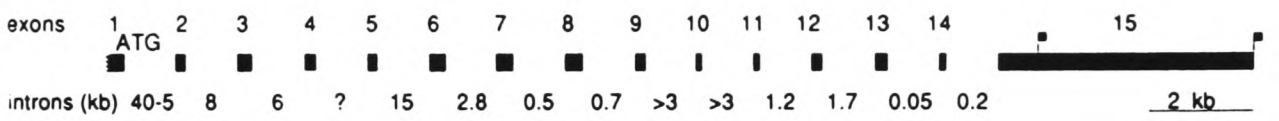


Figure 2.1 Structure of the human MAOA gene. Recognition sites for restriction endonucleases HindIII, EcoRI, BglII, and SalI are represented by letters H, E, B, and S, respectively. The solid horizontal bars beneath the restriction map indicate the position and extent of some exons, as determined by sequencing; the solid flags indicate polyadenylation sites. Open dotted bars indicate restriction fragments within which exons lie, as is determined by hybridisation of cDNA fragments to restriction digested genomic DNA clones. The open boxes below the exons show the genomic clones of MAOA used, with the names inside the bars.



exon	3′ end of exon		5′splice site	intron (kb)	3′ splice site		5′ end of exon	exon
1	...ATT TCA G ...Ile Ser Gly	123 25	gtcagtgtgg...	(40-5)	...tttcctttag	124	GA CTA TCT GCT... Leu Ser Ala...	2
2	...ACT ATA AGG ...Thr Ile Arg	218 56	gtaagtgatt...	(~8)	...ttttttgcag	219	AAT GAG CAT... Asn Glu His...	3
3	...TAT GTC AAG ...Tyr Val Lys	356 102	gtaaggctga...	(6)		357	GGG AAA ACA... Gly Lys Thr...	4
4	...GGG AAG GAG ...Gly Lys Glu	461 137		(?)	...ttctctacag	462	ATT CCA ACT... Ile Pro Thr...	5
5	...TGG ACA AA ...Trp Thr Lys	553 168	gtaagtagac...	(15)	...ctatgaacag	554	G ACT GCT AGG... Thr Ala Arg...	6
6	...GGT GGC CAG ...Gly Gly Gln	695 215	gtatggt?t?...	(2.8)	...cctcctgtag	696	GAA CGG AAG... Glu Arg Lys...	7
7	...CAT TAT GAG ...His Tyr Glu	845 265	gtaactca??...	(.47)	...tgtgttctag	846	TGC AAA TAC... Cys Lys Tyr...	8
8	...AAG AAG G ...Lys Lys Asp	1005 319	gtaggctgct...	(.7)	...cctggttaag	1006	AT TAC TGT GGC... Tyr Cys Gly...	9
9	...ATC ATG GG ...Ile Met Gly	1102 351	gtaggttaga...	(>3)+	...ctctctccag	1103	C TTC ATT CTT... Phe Ile Leu...	10
10	...GAA ATA AG ...Glu Ile Arg	1156 369	gtaagaattt...	(>3)+	...tgttttatag	1157	G AAG AAG AAA... Lys Lys Lys...	11
11	...GCT TTA CAT ...Ala Leu His	1214 388	gtaagaaact...	(1.2)	...tgaactgcag	1215	CCA GTG CAT... Pro Val His...	12
12	...TAT GGA AG ...Tyr Gly Arg	1312 421	gtattacgca...	(1.7)	...tcccttgaag	1313	G GTG ATT CGT... Val Ile Arg...	13
13	...GCT AGG GAG ...Ala Arg Glu	1424 458	gtaagcagga...	(.05)	...tttctttcag	1425	GTC TTA AAT... Val Leu Asn...	14
14	...GAA TCA AAG ...Glu Ser Lys	1487 479	gtaagtttgg...	(.19)	...ccttatctag	1488	GAC GTT CCA... Asp Val Pro...	15
15	...TTGTTAATAA...	3917						

Figure 2.3 Exon structure and exon-intron junctions of MAOA gene. Solid flags above exon No. 15 indicate sites for polyadenylation signals. Numbers to the right or left of nucleotide and amino acid sequences correspond to position numbers reported previously for cDNA clone HM11 (7). ? = not completely sure; > = could not PCR across the intron using primers in flanking exons.

amplification across them and/or by sequence analysis.

The cDNA clones from human placenta, CP223 and CP221, are 2.3 kb and 2.8 kb in length, respectively (6, 7). Partial sequencing was done to align these clones with HMI 1. The 5' end of CP223 starts about 100 bp downstream from the 5' end of HM11, and contains about 200 bp of additional sequence at the 3' end. The 5' end of CP221 starts about 700 bp upstream from the 3' end of HM11, goes beyond the 3' end of HM11 for an additional 2.1 kb. The combined sequences of these three cDNA clones covers 4.0 kb of the longer mRNA (7). The sequence AATAAA, which is a consensus signal for polyadenylation, is found 194 bases upstream from the 3' end of CP221. A cDNA clone similar to CP221 has been obtained by Huang and co-workers (42). The 3' non-coding sequences in CP221 are also found in genomic clones 5-5C and gMAOAI. This 3' end, including the other AATAAA polyadenylation signal found in HM11, plus all additional sequences of CP221 homologous to HM11 are contained within exon 15, which is 2.4 kb in length. Previously observed RNA species, 2.1 kb and 4.2 kb, for MAO-A from human placenta and liver (5,6), are similar in size to HM11 and to HM11-CP221 combined sequences, respectively, assuming that poly-A of tails about

200 bp are added to the RNA species *in vivo*. These results are consistent with the notion that the two messages for MAO-A arise from alternative use of two polyadenylation sites, which are present in the same exon. The role of this alternative termination of messages in regulation of cellular MAO activities or expression remains to be investigated.

TCTAAACCTA ATAACTCTCG CCGAGTGTCA GTACAAGGGT CGCCCCGCTC
TCAGTGCCCA. GCTCCCCCGG GTATCAGCTG AAACATCAGC TGCCCCTGGG
TACGCTCCCG GAGTATCAGC AAAAGGTTCG CCCCGCCCAC AGTGCGGCTC
CCCCGGGTAT CAAAAGAAGG ATCGGCTCCG CCGCCGGGTC CCCGGGGGAG
TTGATAGAAG GGTCCTTCCC ACCCTTTGCC CTGTGCCTAC GACCCAGGAG
CGTGTCAGCCAAAGCATG

Figure 2.2. 5'-noncoding sequence of MAO-A. The underlines indicate sequences corresponding to published MAO-A cDNA sequence hMAO A-7 (5), starting from nucleotide no. 1 of the cDNA. The potential TATA box and start site of translation are boxed in.

DISCUSSION

This study has demonstrated that the human *MAOA* gene comprises 15 exons distributed over 70 kb of the X chromosome. The exon structure of the human *MAOB* gene, which is located very close to the *MAOA* gene on the same 270 kb Sfi fragment (Z.-Y.Chen, unpublished data) is identical to it (43). Comparative gene mapping studies show that, although X linked in mammals, MAOA is autosomally located in marsupials and monotremes (44). It is therefore of interest that activity levels in human females are consistently higher (about 20%) than those observed in males (45). There is the possibility that the MAO gene(s) are not completely inactivated on the inactive X in female cells, as other loci in this region of Xp have been found to resist this inactivation (46).

The *MAOB* gene is also divided into 15 exons with identical exon boundaries, at least for exons 2—14 (43). Both nucleotide and amino acid sequences of the exons corresponding to the coding regions of the cDNAs for the two genes show extensive homology (5,6). Similarities in amino acids for exon [l] through [15] are: 60% [1], 72% [2], 74% [3], 77% [4], 70% [5], 75% [6], 74% [7], 68% [8], 66% [9], 67% [10], 80% [11], 94% [12], 89% [13], 48% [14], and 60% [15]. Exons with the highest similarities, numbers 11, 12, and 13, are centered around the FAD-covalent binding site which resides in exon 12 (Cys 406 and Cys 397, for *MAOA* and *MAOB,* respectively). This region is immediately followed by two exons with the lowest similarity, number 14 and number 15 at the carboxyl terminus. Exons 11, 12, and 13 may have originated from sequences encoding ancestor flavin-binding polypeptide, while exon 14 and 15 may have evolved to confer different affinities for monoamine substrates between these two isoforms of MAO. It should also be pointed out that, although exon 14, and to a lesser extent, exon 10, are highly divergent between the two forms of MAO, they are highly conserved for each form among different species (7). Evidence exists for the presence, in amphibia, of two forms of monoamine oxidase corresponding in substrate specificity and inhibition properties to the A and B forms observed in mammals (47). It appears that the two closely linked genes in humans represent the products of duplication event occurring > 500 MY ago (48). Further studies on the conservation of sequences between A and B forms and between similar forms in different species may enable the

elucidation of those functional domains which are important in determining the differences in their substrate specificity and kinetic behavior. Studies of other duplicated genes which have remained very tightly linked over long evolutionary periods have generally established that, where the loci share similar functions, the genes are expressed in different tissues and/or at different developmental times.

There is considerable interest in establishing the extent of genetic variation in MAO genes for two reasons. First, humans inherit widely differing levels of MAO-A (49, 50) and MAO-B (45) activities, and these variations are thought to contribute to predispositions to certain neurological and psychiatric diseases. Second, male adults have been described who carry deletions of both *MAOA* and *MAOB* loci and who exhibit severe neurological problems, as well as manifesting Norrie disease (an X-linked syndrome characterized by congenital blindness) (15—18). These combined symptoms, which are more severe than those reported for typical Norrie disease, are thought to represent a contiguous gene syndrome and suggest that, although complete loss of MAO-A and MAO-B activity is compatible with life, can cause other neurologic problems.

Knowledge of the organization of the MAOA gene will enable the construction of oligonucleotide primers that can be employed in the PCR amplification of coding sequences from genomic DNA for analysis of gene mutations and for sequencing from individuals with variations and deficiencies in enzyme activity. Such reagents will also simplify the detection of RFLPs, several of which have been reported for the MAOA gene. These include those detected by EcoRV (12, 17), MspI (51) and Fnu4HI (27). In addition, an informative (CA)n repeat polymorphism has been described within the MAOA gene (52). Using RFLP analysis of DNA from 40 males whose MAO-A activity measured in cultured skin fibroblasts varied over 200-fold, a strong correlation was observed between allelic status and activity level (27), indicating that the structural locus is a strong determinant of activity levels. There is clearly much to be learned at a variety of levels concerning the relationship between the two forms of the enzymeand in the differences and similarities in their functional domains.The elucidation of their detailed genomic organization also provides a basis for exploring variations in gene structure associated with disease states.

Note. All intronic sequences obtained during this analysis, as well as the 3 ' sequences for the longer form of the MAO-A mRNA encoded in exon 15, have been deposited in EMBL (accession nos X60517-X60541). Following review of this paper, another paper was published describing the exon structure of the human *MAOA* gene (Grimsby et al, 1991). These two papers agree exactly on exon structure, with minor differences in intron sequences, which may reflect polymorphisms.

ACKNOWLEDGEMENTS

We are especially thankful to Ms. Laurie Ozelius for isolating 30 the *5-5C* clone; Ms. Margot Utterback, Ms. Hanna Konieczki, and Dr. Julie Andersen for help with DNA analyses; and to Ms. Andrea McClatchey and Ms. Donna Romano for synthesizing 31oligonucleotide primers. This work was supported by NIH grant 32NS21921 (Senator Jacob Javits Award) and ADAMHA grant A08683 to X.O.B. The cloning and sequencing work of gMAOAI is supported by the grant NIAAA ROI AA 07643; L.W. is supported by a grant from Western Illinois University, Research Council (J.K.H. was a member of L.W.'s laboratory). We thank the research program of the VA Medical Center, Brockton/West Roxbury, for support during the final phase of this project.

REFERENCES

1. Murphy, D.L. (1978) Biochem. Pharmacol., 27, 1889-1893.

2. Glover, V., Sandler, M. (1986) Cell. Biochem. Func., 4, 89-97.

3. Weyler, W., Hsu, Y.-P.P. and Breakefield, X.O. (1990) *J. Pharmacol. Therapeut.,* 47, 391- 417.

4. Murphy, D.L., Garrick, N.A., Aulakh, C.S. and Cohen, R.M. (1984) *J. Clin. Psychiatry, 45,* 37-3.

5. Bach, A.W.J., Lan, N.C., Johnson, D.L., Abell, C.W., Bembenek, M.E., Kwan, S.-W., Seeburg, P.H. and Shih, J.C. (1988) *Proc. Natl. Acad. Sci. U.S.A., 85,* 4934-4938.

6. Hsu, Y.-P.P., Weyler, W., Chen, S., Sims, K.B., Rinehart, W.B., Utterback, Powell, J.F. and Breakefield, X.O. (1988) *J. Neurochem., 51,,* 1321-1324.

7. Hsu, Y.P., Powell, J.F., Sims, K.B. and Breakefield, X.O. (1989) *J. Neurochem.,* 53, 12-18.

8. Powell, J.F., Hsu, Y.P., Weyler, W., Chen, S.A.,Salach, J., Andrikopoulos, K., Mallet, J. and Breakefield, X.O. (1989) *Biochem. J.,* 259, 407-413.

9. Ito, A., Kuwahaia, T., Inadome, S. and Sapara, Y. (1988) *Biochem. Biophys. Res. Comm.,* 157, 970-976.

10. Pintar, J.E., Barbosa, J., Francke, U., Castiglione, C.M., Hawkins, M., Jr. and Breakefield, X.O. (1981) J. Neurosci., 1, 166-175.

11. Kochersperger, L.M., Parker, E.L., Siciliano, M., Darlington, G.J. and Denney, R.M. (1986) J. Neurosci. Res., 16, 601-616.

12. Ozelius, L., Hsu, Y.P., Bruns, G., Powell, J.F., Chen, S., Weyler, W., Utterback, M., Zucker, D., Haines, J., Trofatter, J.A. et al. (1988) *Genomics,* 3, 53-58.

13. Lan, N.C., Heinzmann, C., Gal, A., Klisak, 1., Orth, U., Lai, E., Grimsby, J., Sparkes, R.S., Mohandas, T. and Shih, J.C. (1989) *Genomics, 4,* 552 -559.

14. Levy, E.R., Powell, J.F., Buckle, V.J., Hsu, Y.P., Breakefield, X.O. and Craig, I.W. (1989) *Genomics, 5,* 368-370.

15. de la Chapelle, A., Sankila, E.-M., Lindof, M., Aula, P. and Norio, R. (1985) Clin. Genet., 28, 317-320.

16. Gal, A., Wieringa, B., Smeets, D.F., Bleeker-Wagemakers, L.M. and Ropers, H.H. (1986) *Cytogenet. Cell. Genet.,* 42, 219-224.

17. Donnai, D., Mountford, R.C. and Read, A.P. (1988) J. Med. Genet., 25, 73 -78.

18. Sims, K.B., de la Chapelle, A., Norio, R., Sankila, E.-M., Hsu, Y.-P. P., Rinehatt, W.B., Corey, T.J., Ozelius, L., Powell, J.F., Bruns, G., Gusella, J.F., Murphy, D.L. and Breakefield, X.O. (1989) *Neuron,* 2, 1069-1076.

19. Samson, J.A., Gudeman, J.E., Schatzberg, A.F., Kizuka, P.P., Orsulak, P.J., Cole, J.O. and Schildkraut, J.J. (1985) *J. Psychuzt. Res.* ,19, 547-555.

20. von Knorring, L., Perris, C., Oreland, L., Eisemann, M., Holmgren, S. and Perris, H. (1985) *Psychiat. Res. 15,* 271-279.

21. Zureick, J.L. and Meltzer, H.Y. (1988) *Biol. Psychiatry,* 24, 63-78.

22. Tabakoff, B., Hoffman, P.L., Lee, J.M., Saito, T., Willard, B.,De, L. and Jones, F. (1988) M *EngL J. Med,* 318, 134-139.

23. Sullivan, J.L., Baenziger, J.C., Wagner, D.L., Rauscher, F.P., Numberger, J.I. and Holmes, J.S. (1990) *Biol. Psych.,* 27, 911-922.

24. Singer, T.P., Castagnoli, N., Ramsay, R.R. and Trevor, A.J. (1987) *J. •Neurochem.,* 49, 1-8.

25. Tetrud, J.W. and Langston, J.W. (1989) *Science,* 245, 519-522.

26. Parkinson Study Group (1989) New Eng. J. Med., 321, 1364-1371.

27. Hotamisligil, G.S. and Breakefield, X.O. (1991) *Am. J. Hum. Genet.,* in press.

28. Bates, P. (1987) *Method EnzymoL,* 153: 82-94.

29. Evans, G.A. and Wahl, G.M. (1987) *Methods Enzymol.,* 152, 604-610.

30. Maniatis, T., Sambrook, J. and Fritsch, E.F. (1982) *Molecular Cloning, A laboratory Manual* Cold Spring Harbor University Press. Cold Spring Harbor.

31. Feinberg, A.P. and Volgelstein, B. (1984) *Anal. Biochem.,* 137, 266-267.

32. Messing, J. (1983) *Methods Enzymol.,* 101, 20-78.

33. Yanisch, P.C., Vieira, J. and Messing, J. (1985) *Gene,* 33, 103-119.

34. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acod. Sci. U.S.A.,* 74, 5463-5467.

35. Tabor, S. and Richardson, C.C. (1987) *Proc. Natl. Acod. Sci. U.S.A.,* 84, 4767 -477 1 .

36. Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Nad Acad. Sci. U.S.A.,* 80, 3963-3965.

37. ToMequzzo, F., GlyM, S., Levi, E., MJolsness, S. and Hayday, A. (1988) *Biotechniques,* 6, 460-469.

38. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science,* 239, 487-491.

39. Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981) *Anal. Biochem.,* 112, 295—298.

40. Gorman, K.B. and Steinberg, R.A. (1989) BiotecMiques, 7, 326-329.

41. Winship, P.R. (1989) *Nucleic Acids Res.* 17, 1266.

42. Huang, J., Huang, C., Wen, L. and Huang, R. (1989) Neurosci. Abst. 15, 986.

43. Grimsby, J., Chen, K., Wang, L.-J., Lan, N. and Shih, J.C. (1991) Proc. Natl. Acad. Sci. USA 88, 3637-3641..

44. O'Brien, S.J. and Graves, J.A.M. (1991) Cytogenet. Cell Genet. 55, 406-433.

45. Murphy, D.L., Wright, C., Buchsbaum, M., Nichols, A., Costa, J.L. and Wyatt, R.J. (1976) Biochem. Med., 16, 254-265.

46. Brown, C.J. and Willard, H.F. (1990) Am. J. Hum. Genet. 46, 273-279.

47. Nicotra, A. and Senatori, O. (1984) Comp. Biochem. Physiol., 78C, 353 - 356.

48. Jeffreys, A.J. (1984) In Dover, G.A. and Flavell, R.B. (eds.) Genome Evolution. Academic Press, New York, pp. 157-176.

49. Breakefield, X.O., Giller, E.L., Jr., Nurnberger, J.I., Castiglione, C.M., Buchsbaum, M.S. and Gershon, E.S. (1980) Psychiat. Res., 2, 307-314.

50. Costa, M.R.C., Edelstein, S.B., Castiglione, C.M., Chao, H. and Breakefield, X.O. (1980) Biochem. Genet., 18, 577-589.

51. Ozelius, L., Gusella, J.F. and Breakefield, X.O. (1989) Nucleic Acids Res., 17, 10516.

52. Black, G.C.M., Chen, Z.-Y., Craig, I.W. and Powell, J.F. (1991) Nucleic Acids Res., 19, 689.

GENOMICS 14, 75-82 (1992)

Chapter 3 Organization of the Human Monoamine Oxidase Genes and Long-Range Physical Mapping around Them

Z-Y. Chen, J. F. Powell, Y-P. P. Hsu, X. 0. Breakefield and I. W. Craig

SUMMARY

A 265-kb yeast artificial chromosome containing sequences for human monoamine oxidase A and B (MAOA and MAO-B) genes has been characterized. These two genes are localized within a region of about 240 kb and are arranged in a tail-to-tail configuration, with the 3' coding sequences separated by about 50 kb. A region about 2.5 Mb around the MAO loci was mapped by pulsed-field gel electrophoresis (PFGE). Comparisons between the restriction maps derived from the YAC and the long-range map derived from genomic digestions were in general agreement. The important features identified include a CpG island at the 5' end of the MAO-A and MAO-B genes, respectively. The combined information supports the order of markers within this region to be *DXS77-DXS7-MAOA-MAOB*.

INTRODUCTION

Monoamine oxidase A and B (MAO-A and MAO-B) are the principal enzymeg catalyzing the oxidative degradation of biogenic amines (including catecholamines and serotonin) in the central nervous system and in peripheral tissues (Murphy, 1978; Glover and Sandler, 1986; Weyler et *al.,* 1990). These enzymes are located in the outer membrane of mitochondria (Greenawalt and Schnaitman, 1970) and are present in a variety of cell types in many different species (Hsu et *al.,* 1989). The two types of MAO were originally distinguished on the basis of their substrate affinity, inhibitor sensitivity, and tissue distribution. They were subsequently shown to have different molecular weights, peptide proteolytic digest products, and antigenic determinants (McCauley and Racker, 1973; Cawthon and Breakefield, 1979; Denney and Denney, 1985; Brown et *al.,* 1980). MAO-A preferentially deaminates serotonin and can be irreversibly inactivated by the acetylenic inhibitor clorgyline; phenylethylamine is a preferred substrate for MAO-B and can be irreversibly inactivated by the inhibitor deprenyl (Johnston, 1968; Knoll and Magyar, 1972).

MAO genes have been mapped to the X chromosome both in humans (Pintar et *al.,* 1981; Ozelius et *al.,* 1988; Levy et *al.,* 1989; Lan et *al.,* 1989) and in mice (Laval et *al.,* 1991). In humans, MAO genes have been localized to the Xp11.23-Xp114 region (Levy et *al.,* 1989). The cDNA and genomic sequences for both confirm the existence of separate genes encoding these two forms of the enzyme (Chen et *al.,* 1991; Grimsby et *al.,* 1991; Bach et *al.,* 1989; Hsu et *al.,* 1989; Powell et *al.,* 1989). The detailed genomic structure of the *MAOA* gene shows it to have a total of 15 exons, with the first exon separated by more than 30 kb from the rest of the exons and the last 5 exons clustered together (Chen et *al.,* 1991; Grimsby et *al.,* 1991). The *MAOB* gene has a similar overall organization and identical exon-intron boundaries for exons 2-14 compared with the *MAOA* gene (Grimsby et *al.,* 1991). These isozymes show extensive homology at both the protein and DNA levels, with an average overall homology at the amino acid level of 70% and with the flavin binding domain having homology of over 90% (Hsu et *al.,* 1989; Chen et *al.,* 1991;

Bach et *al.,* 1989). Given the high degree of conservation of the two genes, it is presumed that they share a common origin through duplication of a common ancestral gene (Nicotra and Senatori, 1988).

Interest in the genetics of MAO arises, in part, from the possible association between altered MAO activities and a number of neuropsychiatric diseases (for review, see Hsu et *al.,* 1989). Furthermore, the MAO loci have been shown, through recombination studies, to be linked to several important genetic disorders including Norrie disease (Lan *et al.,* 1989; Diergaarde *et al.,* 1989; Sims *et al.,* 1989a,b), X-linked retinitis pigmentosa 2 (XLRP; Riley *et al.,* 1991), and Leber's hereditary optic neuroretinopathy (LHON; Vilkki *et al.,* 1991). Complete deletion of MAO genes and the surrounding region causes a "contiguous gene syndrome" (Emanuel, 1988) in some atypical Norrie patients who also manifest severe mental retardation, growth failure, atonic and myoclonic seizures, microcephaly, autonomic dysfunction, and sleep disturbance (Sims *et al.,* 1989a; Lan *et al.,* 1989; Gal *et al.,* 1986; de la Chapelle *et al.,* 1985; Donnai *et al.,* 1988; Collins *et al.,* 1992). The long-range map around the MAO loci can serve as a starting point toward identifying disease genes near the MAO loci.

Pulsed-field gel electrophoresis (PFGE) has provided a powerful tool for physical mapping (Smith and Cantor, 1987) and, together with the isolation of YAC clones containing large genomic fragments (Burke *et al.,* 1987), can provide an efficient approach to the detailed characterization of linked loci. Furthermore, complementary studies of the distribution of rare cutting sites in YAC and genomic DNA can provide important information on the presence of CpG islands, which appear to mark the 5' end of some "housekeeping" genes (Bird, 1986; Gardiner-Garden and Frommer, 1987). These and other regulatory elements may control the tissue- and cellspecific pattern of MAO-A and -B expression. For example, within the nervous system, MAO-A is very high in catecholaminergic neurons, and MAO-B predominates in serotonergic neurons and astrocytes (Levitt *et al.,* 1982; Westlund *et al.,* 1985). It would be of considerable interest to establish whether there is a CpG island associated with either or both of the MAO loci, as is generally the case with housekeeping genes (Bird, 1986; Gardiner-Garden and Frommer, 1987). A detailed survey of the distribution of rare cutting sites will be important in the interpretation of

PFGE data for the inactive and active X chromosomes, which are differentially affected hy methylation (Toniolo *et al.,* 1988) and in the determination of whether the MAO loci are inactivated.

We report here the characterization of a 265-kb YAC containing sequence for the human MAO genes and the construction of the physical map for this region of the Xp chromosome based on the data obtained by PFGE.

MATERIALS AND METHODS

MAO genomic clones. A yeast clone containing a 265-kb yeast artificial chromosome, YMAO, was identified by polymerase chain reaction (PCR) using primers derived from the human MAO-A cDNA to screen a pYAC4 vector library (Center for Genetics in Medicine, Washington I University, St. LOUIS, MO) (Brownstein *et al.,* 1989). The PCR primers corresponded to nucleotides 1428-1446 and 1866-1885 in exon 14 and exon 15 ot HM11, respectively (Hsu *et al.,* 1988; Chen *et al.,* 1991). Agarose plugs containing DNA from the YAC host was prepared according to the protocol of Anand and Southern (1990). For restriction analysis, the plugs were washed twice in IX TE at 50°C for 30 min. The plugs were then equilibrated in appropriate restriction buffer on ice for 30 min before being subject to digestion; a third of the DNA from each plug was digested. Digested DNA was separated by conventional gel electrophoresis and the products were transferred to Hybond-N $^{+}$ membanes (Amersham) and hybridized to radiolabeled probes according to Ross *et al.* (1990). For long-range restriction mapping, the YAC was digested with a range of rare cutting enzymes (either completely or partially) and the products were fractionated by pulsed-field gel electrophoresis on a "Waltzer" apparatus (Southern *et al.,* 1987) in 0.5X TAE, 1..5% agarose (Sigma type 11), employing 18-s switch time at 150V for 36 h. Conditions for blotting and hybridization following PFGE were the same as those for conventional gels (Ross *et al.,* 1990) except that the gel was exposed to UV for 4 min before blotting. The enzymes used for YAC digests were *Eco*RI, *HindIII, and* Bg*l*II for conventional gels, and BssHII, *ClaI, MluI, SalI,* ,SfiI, *SmaI, Sfi*I *Sst*II, and *Xho*I (BRL) for pulsed-field gels.

Preparation and digestion of high-molecular-weight genomic DNA. DNA plugs were made from peripheral blood leukocytes of normal males and a normal female. Cells were embedded in 0. 5% lowgelling-temperature agarose (Sea Plaque, FMC Bioproducts) in phosphate-buffered saline (PBS) at a concentration of 1.5 x 107 cells/ml. Plugs containing l00 ml cell suspension were set in Perspex molds (Ross *et al., 1990).* For double digests, each sample was washed, following the first digestion, in 1X TE for 1 h to elute the enzyme completely. The sample was

then reequilibrated in an appropriate buffer. The enzymes used for genomic digestion were BssHII, *MluI, SalI,* SfiI, and *SstII* and the hybridizations were performed as was described (Ross *et al., 199()).*

Probes. Two full-length cDNAs of MAO were used in this study, MAO-A cDNA designated HM11 (Hsu *et al.,* 1988) and MAO-B cDNA designated BSMAOB (a gift from Dr. Julie Andersen) (Titlow *et al., 1992).* (genomic plasmid clones containing the first (3-1), the second (exon-2), and the last exon (exon-15) of the *MAOA* gene and the last exon of *MAOB* (6-B) were derived from the phage clones isolated from a human 4X-amplified FMBL-3 phage library (a gift from Dr. Derek Blake, Oxford). Fragment BMAOBI, containing the first exon of *MAoB,* was prepared hy digestion of an MAO-B cDNA with PstI and Tth 1111, which releases a sequence covering this region. The probes used to detect the ends of the YAC were prepared from BamHI/PvuII-digested pBR322 plasmid. The derived fragments of 1.7 kb (PBR-R) and 2.7 kb (PBR-L), which correspond to the right and the left arms of the YAC, respectively, were used to hybridize to filters with partial digests of the YAC. Probes pX59 (for locus DXS77; Willard et *al.,* 1983) and L1.28 (for locus DXS7; Wieacker *et al.,* 1983) were obtained from Dr. Hans-Hilger Ropers (University of Nijmegen, The Netherlands).

RESULTS

To determine of the MAO YAC contained all of the sequences for the MAOA gene, YAC DNA was digested with *E*coRI, *Hind*III, and *BglII,* and fragments resolved hy conventional gel electrophoresis were hyhridized to the MAO-A cDNA, HMII. A comparison of its hybridization pattern with that derived from human genomic DNA digested with the same enzymes was identical except for the absence of one major hand from the YAC DNA in the EcoRI digest, suggesting that a part of the *MAoA* gene was missing (Figure. 3.1). A similar comparison of the fragments of YAC and genomic DNA hybridizing to the MAO-B cDNA showed no differences (data not shown) .

Further investigation into the representation of MAO-A and MAO-B coding sequences was undertaken through the long-range mapping of the entire YAC. Partial digests were made with the rare cutting enzymes *BssHII, ClaI, MluI, SalI, SmaI, SfiI,* SstII, and *XhoI,* which were probed with end-labeled PBR-R and PBR-L fragments (representing the right and left ends of the YAC, respectively) following PFGE. This enabled a long-range restriction map of the YAC to be constructed (Figure. 3.2) and revealed single restriction sites for SstII and *MluI,* two SfiI sites, and three *BssHII* sites, together with numerous sites for *SmaI, ClaI, SalI,* and *XhoI.* A CpG(I island containing recognition sequences for enzymes *SmaI, ClaI, XhoI, BssHII,* and *MluI* was identified 20 kb from right end of the YAC (Figure. 3.2).

The physical relation of the two MAO genes was investigated by hybridizing the MAO-A- and MAO-B-specific 5' and 3' genomic clones to the same PFGE filter. Whereas the 5'-specific sequence of MAO-A (3-1), which contains the first exon, failed to detect any signal, another clone of MAO-A (exon-2), containing the second exon and its surrounding sequences, hybridized to fragments in a pattern suggesting a location about 30 kb from the left end of the YAC, between a *XhoI* site and a *SfiI/XhoI* site. It appears then that the first and second exons are separated by at least 30 kb. By applying the same strategy, the 3' end of *MAOA* (exon-15) was localized to a region about 95 kb from the left end of the YAC, between a *SalI* site and an *XhoI* site (Figure.

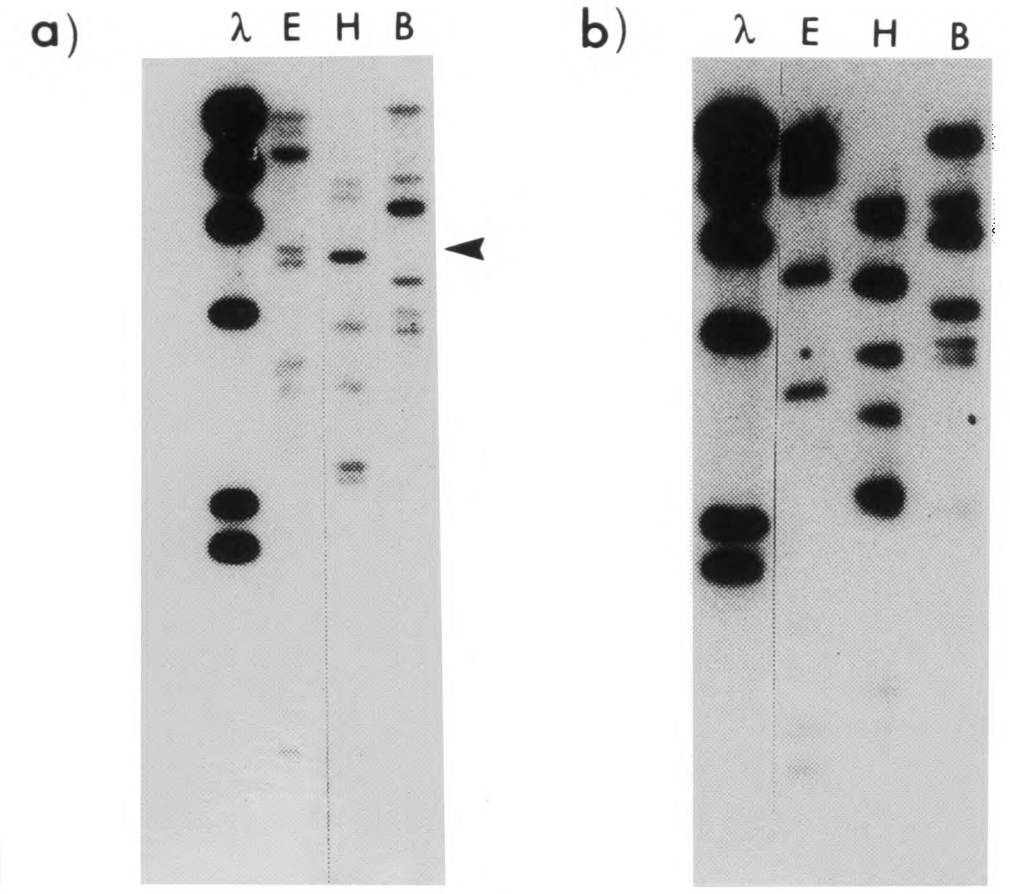


Figure 3.1 Southern blot of human genomic and MAO-YAC DNA hybridized to a full-length cDNA for MAO-A. Five-microgram samples of DNA were digested with different restriction enzymes according to the manufacturers instructions. Fragments were resolved by electrophoresis in 0.8% agarose gels and blotted onto nylon membranes. Blots were hybridised under conditions of high stringency (see Materials and Methods) to a cDNA probe for MAO-A, HM11, labeled with [^{32}P] dCTP, (a) Genomic DNA ; (b) MAO-YAC DNA.λ,HindIII-digested lambda DNA ; E, EcoRI- digested DNA; H,HindIII-digested DNA; B, BglII-digested DNA. The arrowhead indicates the bottom band in the doublet in (a) in the EcoRI digest that is missing in the (b) EcoRI digest.

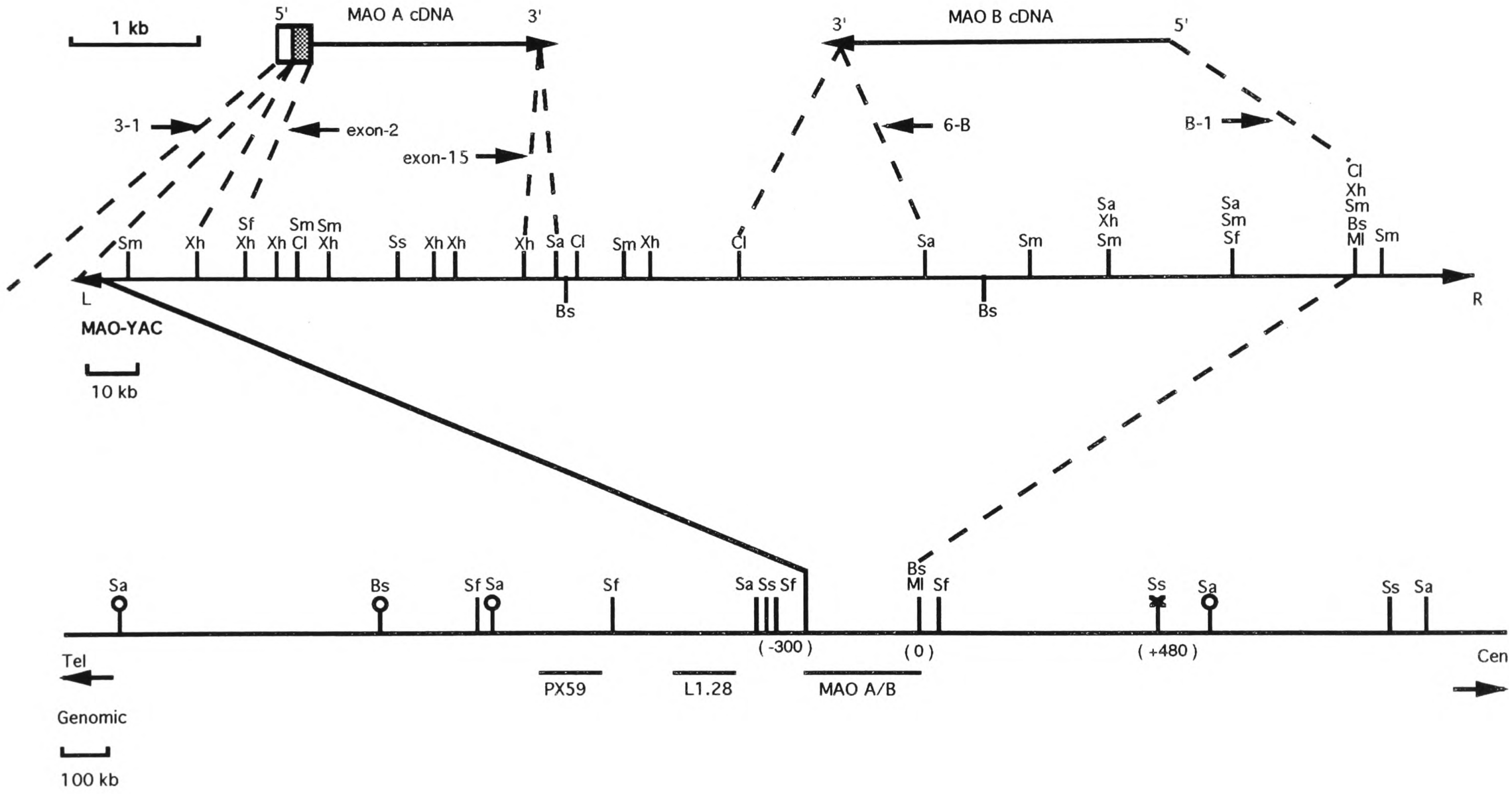


Figure 3.2 Long-range restriction maps of the MAO-YAC and the genomic region including the MAO loci. The top line shows the size and orientation of cDNAs for MAO-A and MAO-B. The arrows on these cDNAs indicate the directions of transcription. The open box indicates the first exon of MAO-A, which is not in the MAO-YAC, and the stippled box the second exon, which is in the YAC. The arrows under the t line indicate the positions of the MAO genomic clones within the YAC. The middle line shows the restriction map of the MAO-YAC. The left and right arms of the YAC are markes L and R, respectively. The bottom line shows the restriction map of the human genomic DNA, indicating the relative positions of loci *DXS77* (pX59), *DXS7* (L1.28), and MAOA/MAOB. The dashed line between the YAC and genome indicates the restriction sites in the YAC that are also cleavable in the genome. The solid line between the YAC and genome indicates the position of the end of the YAC in the genome. The presumed chromo somal orientation of the genomic map is also indicated, Tel, telomere of parm; Cen, centromere. The *Bss*HII (Bs) and *Sal*I (Sa) sites in the genome which are cleavable only on the inactive X chromosome are indicated by circles; the *Sst*II site that is cleaved approximately 50% on the active X chromosome is indicated by an X. Enzymes used: Bs, *Bss*HII; Cl, *Cla*I; Ml, *Mlu*I; Sa, *Sal*I; Sf, *Sfi*I; Sm, *Sma*I; Ss, *Sst*II and Xh, *Xho*I.

3.2). From phage mapping data, it is known that the last coding exon of MAO-A lies 6 kb from a *SalI* site in the 3' untranslated region (Chen et *al.,* 1991). The only candidate for this *SalI* site is one lying 95 kb from the left end of the YAC (Figure 3.2). The YAC therefore contains all the *MAOA* gene except for the first exon. In similar experiments, the last exon of the *MAoB* gene (6-B) was shown to lie between a *ClaI* site and a *SalI* site, 100-130 kb from the right end of the YAC. The first exon of *MAOB* (BMAOBI) was localized close to the CpG island within a 20-kb region from the right end of the insert. It can therefore be concluded that the YAC contains all of the MAO-B gene coding sequences (Figure 3.2).

The exact location of the sequence detected by subclone 6-B, which contains the last exon of the MAO-B gene, within the 30-kb *ClaI-SalI* fragment to which it hybridizes, could not be determined. However, by comparison with restriction map data from a phage clone containing the last exon of the MAOB gene (Z.-Y. Chen, unpublished data), it appears that the last exon is less than 15 kb from this *ClaI* site. This also confirms that the two MAO genes must lie in a tail-to-tail arrangement (Sims *et al.,* 1992). The size of the MAOB gene is therefore around 95 kb, and the sizes of the *MAOA* and *MAOB* genes are thus quite similar (Grimsby *et al.,* 1991).

Long-range mapping of human genomic DNA data were obtained from probing peripheral blood lymphocyte DNA from males and females digested with rare cutting enzymes. Previous observations had suggested that the *MAOA* and *MAOB* genes were physically very close (Levy *et al.,* 1989; Lan *et al.,* 1989; Sims *et al.,* 1989a; Diergaarde *et al.,* 1989). The PFGE data here revealed that probes for *MAOA* and *MAOB* genes gave identical hybridization patterns for all enzymes tested. Table 3.1 lists the sizes of bands obtained by using MAOA and MAO-B cDNA as probes. To construct a map, double digestions were also performed. *SfiI* gives the smallest fragment, 340 kb, that contains both MAO genes. Although there are three BssHII sites and one *MluI* site in the YAC, the sites lying 20 kb away from the right end of the insert are the only ones that correspond to those cleaved in the genomic DNA. For simplicity, the coincident BssHII and *MluI* site on the map is designated *BssHII/MluI* (0 in Figure 3.2). The positions to the right of this site are designated (+) kb and those to the left (—) kb. The reason for suggesting that the *BssHII/MluI* sites in the YAC are the only ones cut in the genome are as follows. In the double digests of *BssHII + SfiI* and *MluI + SfiI,* the fragments detected by MAO-A and MAO-B cDNAs, were of the

same size (300 kb), which is 40 kb smaller than the single *SfiI* fragment, suggesting that there are two *SfiI* sites at (-300) and (+40) on either side of the reference *BssHII/MluI* site (0 in Figure 3.2). This is supported through the observation of similar results with double digests of *BssHII/MluI together* with *SalI* or *SstII*. By comparing the double- and single-digest products, two *SstII* sites were mapped at (+480) and (-320) and two *SalI* sites were localized to (+600) and (-350) (see Figure 3.2).

It has been shown that probes L1.28 and pX59 detect sequences close to the MAO loci with which they also share a 1150-kb *BssHII* fragment (Diergaarde *et al.,* 1989). These two probes were therefore used to probe the YAC and genomic PFGE filters. Neither probe hybridized to digests of the YAC. The hybridization patterns for these two probes in genomic DNA digested with various rare cutting enzyme were almost identical, but very different from that observed with MAO probes (Figure 3.3, Table 3.1). The only difference observed between the hybridization patterns of the two probes was that L1.28 has a slightly larger *SfiI* fragment (340 kb) compared with 310 kb for pX59. In the digests of female DNA, however, they shared the smallest *SalI* fragment of 580 kb, which confirmed that they recognize two adjacent *SfiI* fragments but the same *SalI* fragment (Diergaarde *et al.,* 1989). In digests of female DNA, L1.28, pX59, and MAO probes all hybridized to the same 1150-kb *BssHII* fragment (Figure 3.4). Comparison of MAO and L1.28 hybridization patterns showed that both probes hybridized to a 340-kb *SfiI* fragment, but the double-digestion products from *BssHII/MluI + SfiI gave* different fragments, with L1.28 unaltered and MAO now 300 kb (i.e., 40 kb smaller). This is because the *BssHII/MluI* (O in Figure 3.2) site is within the 340-kb *SfiI* fragment detected by MAO and is not present on the one recognized by L1.28. Also *SalI(*—350) and SstII(—320) sites are within the 340-kb *SfiI* fragment detected by L1.28, but not on that detected by MAO. The double digests of *SalI + SfiI* and SstII + *SfiI* gave altered fragments of 290 and 320 kb when probed with DXS7 (Fig. 3). It is apparent, therefore, that L1.28 and the MAO genes lie in two different, but adjacent, *SfiI* fragments of the same size (Figure 3.2). A comparison of PFGE maps for the genomic and YAC DNA in this region shows that of the rare cutting sites tested, only the *BssHII* and *MluI* sites are in the YAC (Figure 3.2). Alignment of the PFGE data obtained from the four probes allows the conclusion that, in a region of about 2.5 Mb around MAO loci, the order for the loci is *DXS77-DXS7-MAOA-MAOB,* and they

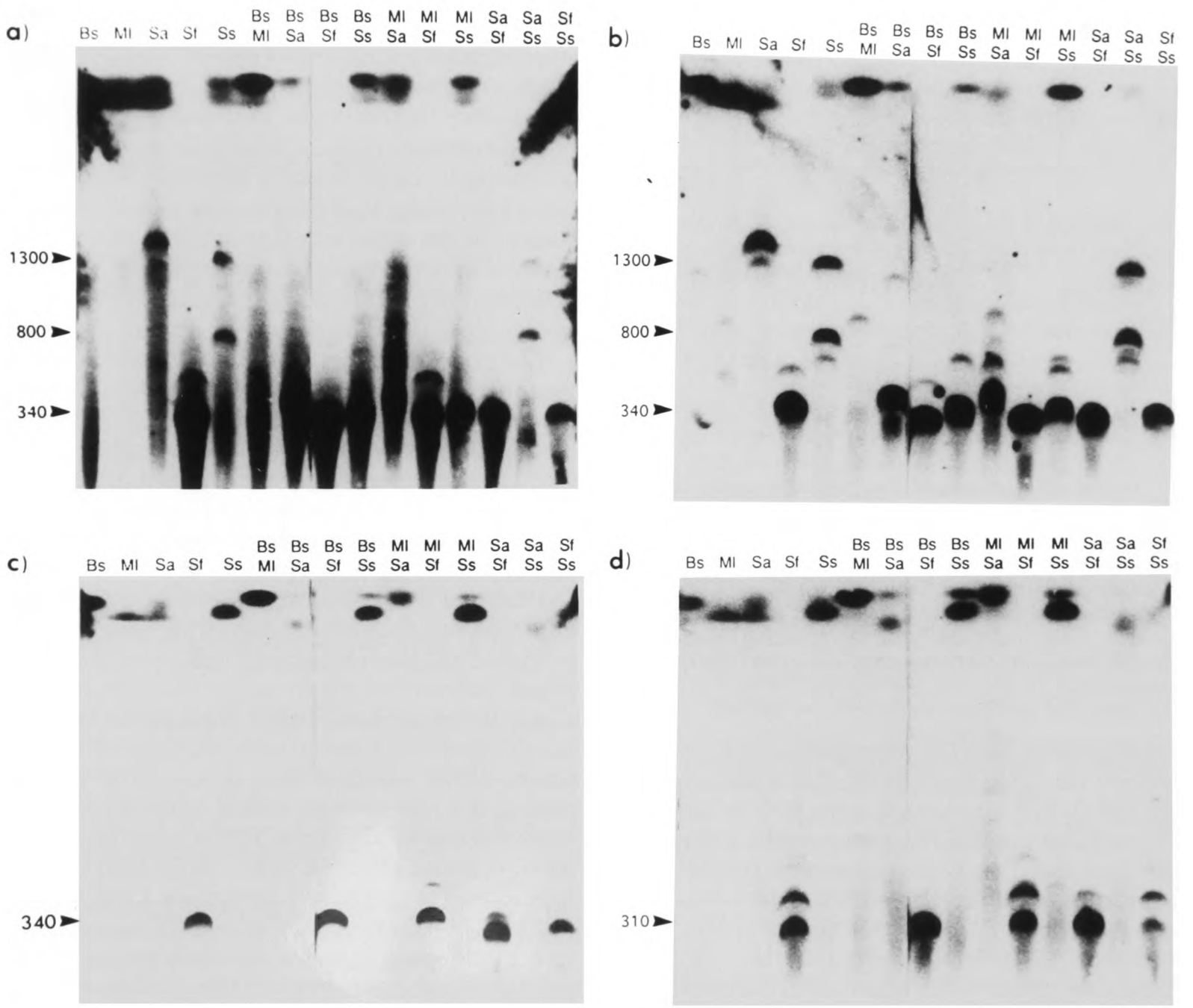


Figure 3.3 Pulsed-field gels of genomic DNA hybridised to different probes in the Xp11.3 region. Human genomic DNA was digested in blocks and resolved by pulsed-field gel electrophoresis as described under Materials and Methods. Fragment sizes were determined by parallel electrophoresis of lambda oligomers and *Saccharomyces cervisiae* (Ross et al., 1990). Gels were blotted onto nylon membranes and hybridised under conditions of high strigency to ^{32}P-labeled probes (see Materials and Methods). Restriction enzymes used for digestion are indicated above the lanes. Probes used for hybridisation were : (a) HM11, (b) BSMAOB, (c) L1.28, and (d) pX59. Enzymes used : Bs, BssHII ; Ml, MluI,: Sa, SalI; Sf, SfiI; St, SstII.

TABLE 3.1

PULSED FIELD GEL FRAGMENTS DETECTED BY VARIOUS PROBES

	Probe			
Restriction enzymes	pX59	L1.28	HM11	BSMAOB
BssHII	LM M 1115 F	LM M 1115 F	LM M 1115 F	LM M 1115 F
MluI	>2000	>2000	>2000	>2000
SalI	LM M 580 F 1200 F	LM M 580 F 1200 F	950 F 1400	950 F 1400
SfiI	310	340	340	340
SstII	>2000	>2000	800 1300	800 1300
BssHII/MluI	LM	LM	LM	LM
BssHII/SalI	1800	1800	350	350
BssHII/SfiI	310	340	300	300
BssHII/SstII	1900	1900	320	320
MluI/SalI	LM	LM	350	350
MluI/SfiI	310	340	300	300
MluI/SstII	1900	1900	320	320
SalI/SfiI	310	300	340	340
SalI/SstII	LM	LM	800 1300	800 1300
SfiI/SstII	310	320	340	340

Note. Values are given kilobases. LM-Limited mobility of fragment(s); M- Fragment only observed in male genomic DNA; F- fragment only observed in female genomic DNA.

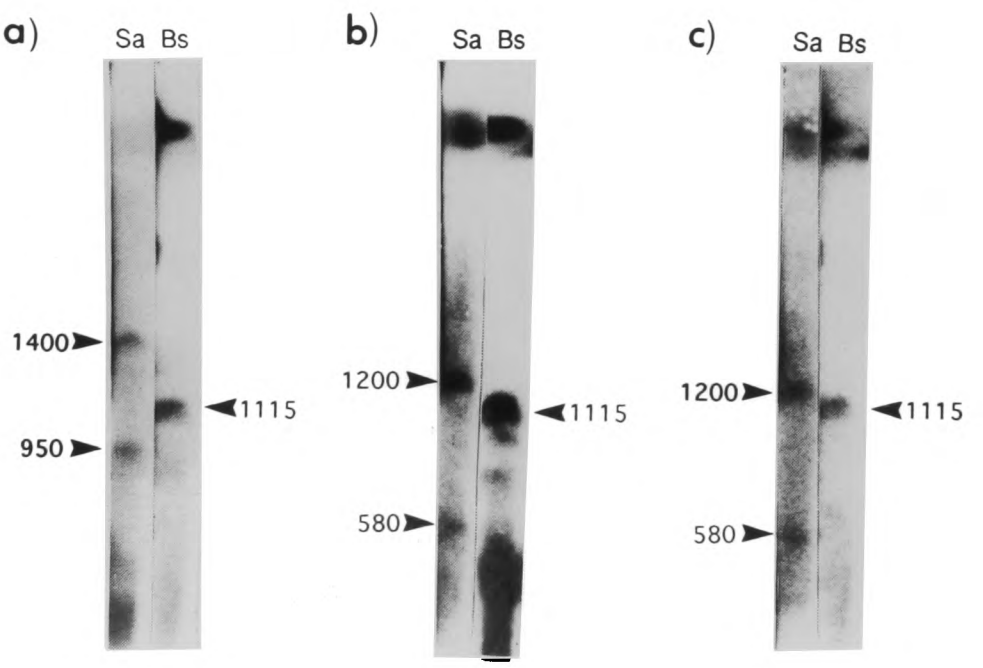


Figure 3.4 Pulsed-field gel of female genomic DNA hybridised to several probes. Female genomic DNA was digested, resolved, and hybridised as in the legend to Figure 3.3. Restriction enzymes used for hybridisation are indicated above the lanes; Sa, SalI; Bs, BssHII. Probes used for hybridisation were: (a) HM11 (identical bands were seen using probe BMAOB, data not shown); (b) L1.28; (c) pX59. Arrows indicate bands of interest. Note: pX59 and L1.28 hybridised to the same size SalI fragment, whereas MAO probes recognised two different SalI fragments.

are localized in three adjacent SfiI fragments (Figure 3.2).

As noted above, a 1150-kb *BssHII* fragment together with a *BssHII* compression band is detected in female DNA probed with L1.28, pX59, or MAO clones; however, in the male, hybridization is to the compression band only (Fig. 3), indicating that one of the BssHII(-1150) sites is methylated on the active X chromosome, but unmethylated on the inactive X chromosome in females. A similar result was observed in *SalI* digests, where, in the female sample, two bands of 950 and 1400 kb are detected by MAO-A and MAO-B cDNAs, respectively (Figure 3.4). In the male, however, only a 1400-kb fragment was present. This is consistent with reports suggesting that non-island-associated CpG dinucleotides may show methylation on the active but not the inactive X (Boyd and Fraser, 1990; Lock et *al.,* 1987).

Although there is no evidence from PFGE data for a CpG island associated with the *MAOA* gene, a 1.4-kb region containing the first exon of *MAOA* and 1.1 kb of upstream noncoding sequence has been sequenced (M. Utterback and Z-Y. Chen, unpublished data) and found to be relatively rich in G + C (60%) and *HpaII* sites (10); this fits the criteria for a CpG island (Bird, 1987). Examination of the 5' end of *MAOB* gene shows a region that recognizes rare cutting enzymes, *SmaI, XhoI,* BssHII, *ClaI,* and *MluI,* and is very likely to be a CpG island. Sequence data for the putative MAOB-associated island are not yet available and the exact location of the first exon of MAO-B with respect to this island has yet to be determined. The YAC data, however, suggest that the first exon of *MAOB* is actually very close to the island.

DISCUSSION

The characterization of two distinct genes for *MAOA* and *MAOB* has been of considerable significance in the context of understanding the neurogenetics of catecholamine metabolism *(Hsu et al.,* 1989). The detailed analysis of the MAO YAC presented here and its observed colinearity with genomic DNA supports the integrity of this YAC and the 3'-to-3' configuration of the MAO genes in a region of 240 kb. These data and complementary studies of other YACs in this region (Chen *et al.,* 1992; Sims *et al.,* 1992) demonstrate that the *MAOA* and *MAOB* genes are about 50 kb apart, with each spanning around 95 kb. There are many other examples of genes or gene families duplicated from ancient progenitor genes, such as globin and collagen (Jeffreys, 1984; Poschl *et al.,* 1988). The physical relationship between such duplicated genes has been observed to be very variable, commonly including tandem arrays (such as globin genes), but also head-to-head orientations in collagen and histone genes (Hentschel and Birnstiel, 1981; Wells, 1986). The tail-to-tail configuration of the MAO genes observed has also been reported in mouse Clr and Cls genes (Kusumoto *et al.,* 1988). The mechanism responsible for producing this latter type of organization is unknown. One possibility would be the formation of a looped structure prior to an unequal crossing over event. A comparison of MAO gene organization across species may assist in understanding this event. It seems likely that the duplication occurred at some stage in amphibian evolution as it has been reported that teleosts and some amphibia possess only type A MAO (Hall and Uruena, 1983), whereas in toad both forms of MAO are present (Nicotra and Senatori, 1988).

An important consequence of the configuration of the two MAO genes is that they must be transcribed from different strands. This confirms that the two MAO genes are encoded in different genes and precludes the possibility that differential splicing could result in multiple variant forms of the enzyme. This orientation would also allow for a differential evolution of 5' regulatory elements for these two genes. Both the *MAOA* and *MAOB* genes have CpG islands, indicating that they are both "housekeeping" genes. Differential regulation is apparent, as there are profoundly different levels and distribution of activities in different tissues and during different stages of development (for review, see Weyler *et al.,* 1990). This would not, however, eliminate

the possibility of some type of transcriptional interference or exclusion operating between the A and B genes.

The restriction maps obtained here for the YAC and genomic DNA are in general agreement. A preliminary study, on the methylation status of the *HpaII* sites of the 5' region of MAOA gene in human lymphocytes indicates that they are completely methylated on the inactive X chromosome, but unmethylated on the active one (Hendriks and Chen, unpublished data). In the case of the 5' end of *MAOB* gene, the present study indicates that the BssHII and *MluI* sites are cleavable on both the active and inactive X chromosomes, which stands in contrast to most CpG islands, which tend to be methylated, and hence resistant to cleavage, on the inactive X chromosome (Toniolo *et al.,* 1988). This preliminary evidence for the lack of methylation of the *MAOB* CpG island in female DNA suggests that it would be worthwhile to pursue the inactivation status of the gene, particularly since females tend to have higher levels of MAO-B activity than males (Murphy *et al.,* 1976). MAO genes in marsupials and monotremes are not sex linked and not subject to X inactivation (O'Brien and Graves, 1991); useful information may therefore be forthcoming through comparison of the gene organizations in those species.

The hybridization patterns of PFGE-digested genomic DNA detected with MAO cDNA probes are almost identical. There are, however, some faint bands only detected with MAO-B cDNA which may represent a subset of homologous sequences in the genome. They could be derived from pseudogenes, genes for other enzymes with conserved domains, or other genes for MAO, since in some tissues the enzyme activity detected is typical of neither MAO-A nor MAO-B (Youdim, 1972; Mantle and Tipton, 1978).

Data presented here show clearly that probes for MAO and L1.28 hybridize to different *SfiI* fragments which are coincidentally of similar size and adjacent to each other in the genome. This observation supports the finding of Diergaarde *et al.* (1989) and may explain the discrepant findings reported by Lan *et al.* (1989), which could result from partial *SfiI* digestion or differences in running conditions for PFGE.

Tight linkage between markers within and flanking this MAO-YAC and a number of disease genes, such as Norrie and LHON, have been reported. Three highly polymorphic (AC)n repeats have been described in the MAO region (Black *et al.,* 1991; Konradi *et al.,* 1992; Chen,

manuscript in preparation), as have three restriction fragment length polymorphisms (Ozelius *et al.,* 1988, 1989; Hotamisligil and Breakefield, 1991). The availability of the long-range map together with additional polymorphic markers should lead to further definition of the location of other disease genes in this region.

ACKNOWLEDGEMENTS

We thank Dr. Julie Andersen for providing the MAO-B cDNA, Dr. Bernard Brownstein for helping with the YAC isolation, Dr. Mark Jobling for his critical reading of the manuscripts, Dr. Katherine Sims and Ms. Margot Utterback for sharing their unpublished findings and insights with us, and Ms. Valerie Abrahamsen and Ms. Suzanne McDavitt for skilled assistance in preparing this manuscript. This work was supported by NIH Grant NS-21921 (Senator Jacob Javits Award) to X.O.B.

REFERENCES

Anand, R., and Southern, E. M. (1990). *In* "Gel Electrophoresis of Nucleic Acids" (D. Rickwood and B. D. Hames, Eds.), pp. 101-123, IRL Press, Oxford.

Bach, A. W. J., Lan, N. C., Johnson, D. L., Abell, C. W., Bembenek, M. E., Kwan, S-W., Seeburg, P. H., and Shih, J. C. (1989). cDNA cloning of human liver monoamine oxidase A and B: Molecular basis of differences in enzymatic properties. *Proc. Natl. Acad. Sci. USA* 85: 4934-4938.

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

Bird, A. P. (1987). CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* 3: 342-347.

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

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

Brown, G. K., Powell, J. F., and Craig, I. W. (1980). Molecular weight differences between human platelet and placental monoamine oxidase. *Biochem. Pharmacol.* 29: 2595-2603.

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

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

Cawthon, R. M., and Breakefield, X. O. (1979). Differences in the A and B forms of monoamine oxidase revealed by limited proteolysis and peptide mapping. *Nature* 281: 692-694.

Chen, Z.-Y., Hendriks, R. W., Jobling, M. A., Powell, J.-F., Breakefield, X. O., Sims, K. B., and Craig, I. W. (1992). Isolation and characterisation of a gene implicated in Norrie disease. *Nature Genet. 1:* 204-208.

Chen, Z-Y., Hotamisligil, G. S., Huang, J-K., Wen, L., Ezzeddine, D., Aydin Muderrisoglu,N., Powell,J. F., Huang, R. H., Breakefield, X. 0., Craig, I. W., and Hsu, Y-P.P. (1991).

Structure of the human gene for monoamine oxidase type A. *Nucl. Acids Res. 19:* 4537-4541.

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

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

Denney, R. M., and Denney, C. B. (1985). An update on the identity crisis of monoamine oxidase: New and old evidence for the independence of MAOA and MAOB. *Pharmacol. Ther.* 30: 227-259.

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

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.

Emanuel, B. S. (1988). Molecular cytogenetics: Toward dissection of the contiguous gene syndromes. *Am. J. Hum. Genet.* 43: 575-578.

Gal, A., Wieringa, B., Smeets, D. F., Bleeker-Wagemakers, L. M., and Ropers, H. H. (1986). Submicroscopic interstitial deletion of the X chromosome explains a complex genetic syndrome dominated by Norrie disease. *Cytogenet. Cell Genet.* 40: 219-224.

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. J. *Mol. Biol. l9ff:* 261-282.

Glover, V., and Sandler, M. (1986). Clinical chemistry of monoamine oxidase. *Cell Biochem. Funct.* 4: 87-97.

Greenawalt, J. W., and Schnaitman, C. (1970). An appraisal of the use of monoamine oxidase as an enzyme marker for the outer membrane of rat liver mitochondria. J. *Cell Biol.* 46:173-179.

Grimsby, J., Chen, K., Wang, L-J., Lan, N. C., and Shih, J. C. (1991). Human monoamine oxidase A and B genes exhibit identical exonintron organization. *Proc. Natl. Acad. Sci. USA* 88: 3637-3641.

Hall, T. R., and Uruena, G. (1983). Pharmacology and physiology of monoamine oxidase activity in vertebrates—A comparative study. *Comp. Biochem. Physiol. B* 7ff(3): 393-397.

Hentschel, C. C., and Birnstiel, M. L. (1981). The organization and expression of histone gene families. *Cell* 25: 301-313.

Hotamisligil, G. S., and Breakefield, X. O. (1991). Human monoamine oxidase A gene determines levels of enzyme activity. *Am.* J. *Hum. Genet.* 49: 383-392.

Hsu, Y. P., Powell, J. F., Sims, K. B., and Breakefield, X. O. (1989). Molecular genetics of the monoamine oxidases. J. *Neurochem.* 53: 12-18.

Hsu, Y-P.P., Weyler, W., Chen, S., Sims, K. B., Rinehart, W. B., Utterback, M. C., Powell, J. F., and Breakefield, X. O. (1988). Structure features of human monoamine oxidase A elucidated from cDNA and peptide sequence. J. *Neurochem.* 51:1321-1324.

Jeffreys, A. J. (1984). *In* "Genome Evolution" (G. A. Dover and R. B. Flavell, eds.), pp. 157-176, Academic Press, New York.

Johnston, J. P. (1968). Some observations upon a new inhibitor of monoamine oxidase in human brain. *Biochem. Pharmacol.* 17: 1285-1297.

Knoll, J. and Magyar, K. (1972) Some puzzling pharmacological effects of monoamine oxidase inhibitors. *Adu. Biochem. Psychopharmacol.* 5: 393-408.

Konradi, C., Ozelius, L., and Breakefield, X. O. (1992). Highly polymorphic (GT),, repeat sequence in intron II of the human MAOB gene. *Genomics* 12: 176-177.

Kusumoto, H., Hirosawa, S., Salier, J. P., Hagen, F. S., and Kurachi, K. (1988). Human genes for complement components Clr and Cls in a close tail-to-tail arrangements. *Proc. Natl. Acad. Sci. USA* 85: 7307-7311.

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

Laval, S. H., Chen, Z.-Y., and Boyd, Y. (1991). The properdin structural locus (Pfc) lies close to

the locus for tissue inhibitor of metallothionein proteases (TIMP) on the mouse X chromosome. *Genomics* IO: 1030-1034.

Levitt, P., Pintar, J. E., and Breakefield, X. O. (1982). Immunocytochemical demonstration of monoamine oxidase B in brain astrocytes and serotonergic neurons. *Proc. Natl. Acad. Sci. USA* 79: 6385-6389.

Levy, E. R., Powell, J. F., Buckle, V. J., Hsu, Y. P., Breakefield, X. O., and Craig, I. W. (1989). Localization of human monoamine oxidaseA gene to Xpll.23-11.4 by in situ hybridization: Implications for Norrie disease. *Genomics* 5: 368-370.

Lock, L. F., Takagi, N., and Martin, G. R. (1987). Methylation of the HPRT gene on the inactive X occurs after chromosome inactivation. *Cell* 48: 39-46.

Mantle, T. J., and Tipton, K. F. (1978). Monoamine oxidase A and B: A useful concept? *Biochem. Pharmacol.* 27: 97-101.

McCauley, R., and Racker, E. (1973). Separation of two monoamine oxidases from bovine brain. *Mol. Cell Biochem. 1:* 73-81.

Murphy, D. L. (1978). Substrate-selective monoamine oxidases: Inhibitor, tissue, species and functional differences. *Biochem. Pharmacol.* 27: 1889-1893.

Murphy, D. L., Wright, C., Buchsbaum, M., Nichols, A., Costa, J. L., and Wyatt, R. J. (1976). Platelet and plasma amine oxidase activity in 680 normals: Sex and age differences and stability over time. *Biochem. Med.* 16: 254-265.

Nicotra, A. and Senatori, O. (1988). Changes in monoamine oxidase activity by mitochondria isolated from late embryos of *Bufo bufo. Comp. Biochem. Physiol. C* 89(1): 5-9.

O'Brien, S. J., and Graves, J. A. M. (1991). Report of the Committee on Comparative Gene Mapping. *Cytogenet. Cell Genet.* 55: 406-433.

Ozelius, L., Gusella, J. F., and Breakefield, X. O. (1989). MspI RFLP for human MAOA gene. *Nucleic Acids Res.* 17:10516.

Ozelius, L., Hsu, Y. P., Bruns, G., Powell, J. F., Chen, S., Weyler, W., Utterback, M., Zucker, D., Haines, J., Trofatter, J. A., Conneally, P. M., Gusella, J. F., and Breakefield, X. O. (1988). Human monoamine oxidase gene (MAOA): Chromosome position (Xp21-pll) and DNA polymorphism. *Genomics 3:* 53-58.

Pintar, J. E., Barbosa, J., Francke, U., Castiglione, C. M., Hawkins, M., Jr., and Breakefield, X. O. (1981). Gene for monoamine oxidase type A assigned to the human X chromosome. *J. Neurosci.* 1:166175.

Poschl, E., Pollner, R., and Kuhn, K. (1988). The genes for the alpha I(IV) and alpha 2(IV) chains of human basement membrane collagen type IV are arranged head-to-headand separated by a bidirectional promoter of unique structure. *EMBO J.* 7: 2687-2695.

Powell, J. F., Hsu, Y.-P. P., Weyler, W., Chen, S., Salach, J., Andrikopoulos, K., Mallet, J., and Breakefield, X. O. (1989). The primary structure of bovine monoamine oxidase type A: Comparison with peptide sequences of bovine monoamine oxidase type B and other flavoenzymes. *Biochem. J.* 259: 407-413.

Riley, S., Black, G., Chen, Z.-Y., Hatchwell, E., Lorenz, B., Meitinger, T., Powell, J., Wittwer, B., and Craig, I. (1991). *In* "Degenerative Retinopathies: Advances in Clinical and Genetic Research" (P. Humphries, S. Bhattacharya, and A. Bird, Eds.), pp. 119-128, CRC Press, Boca Raton, FL.

Ross, M. T., Ballabio, A., and Craig, I. W. (1990). Long-range physical mapping around the human steroid sulfatase locus. *Genomics* 6: 528-539

Sims, K. B., de la Chapelle, A., Norio, R., Sankila, E.-M., Hsu, Y.-P. P., Rinehart, W. B., Corey, T. J., Ozelius, L., Powell, J. F., Bruns, G., Gusella, J. F., Murphy, D. L., and Breakefield, X. O. (1989a). Monoamine oxidase deficiency in males with an X chromosome deletion. *Neuron* 2: 1069-1076.

Sims, K. B., Lebo, R. V., Benson, G., Shalish, C., Chen, Z-Y., Bruns, G., Golbus, M. S., Craig, I. W., and Breakefield, X. O. (1992). The Norrie disease gene maps to a 150 kb region chromosome Xpll.3. *Hum. Mol. Genet. 1:* 83-89.

Sims, K. B., Ozelius, L., Corey, T., Rinehart, W. B., Liberfarb, R., Haines, J., Chen, W. J., Norio, R., Sankila, E., de la Chapelle, A., Murphy, D. L., Gusella, J., and Breakefield, X. O. (1989b). The Norrie disease gene is distinct from the monoamine oxidase genes. *Am. J. Hum. Genet.* 45: 424-434.

Smith, C. L., and Cantor, C. R. (1987). Purification, specific fragmentation, and separation of l arge DNA molecules. *In* "Methods in Enzymology" (R. Wu, Ed.), Vol. 155, pp. 449-467,

Academic Press, New York.

Southern, E. M., Anand, R., Brown, W. R. A., and Fletcher, D. S. (1987). A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* 15: 5925-5943.

Titlow, C. C., Andersen, J. K., Trofatter, J. A., and Breakefield, X. O. (1992). In *vitro* translation of truncated proteins using RNA generated by SP6 RNA polymerase transcription from PCR products. *PCR Methods,* in press.

Toniolo, D., Martini, G., Migeon, B. R., and Dono, R. (1988). Expression of the G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA. *EMBO J.* 7: 401-406.

Vilkki, J., Ott, J., Savontaus, M.-L., Aula, P., and Nikoskelainen, E. K. (1991) . Optic atrophy in Leber hereditary optic neuroretinopathy is probably determined by an X-chromosomal gene closely linked to DXS7. *Am. J. Hum. Genet.* 48: 486-491.

Wells, D. E. (1986). Compilation analysis of histones and histone genes. *Nucleic Acids Res.* 14(Suppl): rli9-rl49.

Westlund, K. N., Denney, R. M., Kochersperger, L. M., Rose, R. M., and Abell, C. W. (1985). Distinct monoamine oxidase A and B populations in primate brain. *Science* 230: 181-183.

Weyler, W., Hsu, Y.-P. P., and Breakefield, X. O. (1990). Biochemistry and genetics of monoamine oxidase. *J. Pharmacol. Ther.* 47: 391-417.

Wieacker, P., Davis, K. E., Cooke, H. J., Pearson, P. L., Williamson, R., Battacharya, S., Zimmer, J., and Ropers, H. H. (1984). Towards a complete linkage map of the human X chromosome: Regional assignment of 16 cloned single copy DNA sequences employing a panel of somatic cell hybrids. *Am. J. Hum. Genet.* 36: 265-276.

Willard, H. E., Schmeckpeper, B. J., Smith, K. D., and Goss, S. J. (1983). High resolution regional mapping of human X-chromosome by radiation induced segregation analysis of X-linked DNA probes. *Cytogenet. Cell Genet.* 37: 609-610.

Youdim, M. B. H. (1972). Multiple forms of mitochondrial monoamine oxidase. *Br. Med. Bull. 29:* 120-122.

Human Molecular Genetics, Vol. 1, No. 3 161-164

Chapter 4. Characterisation of a YAC containing part or all of the Norrie disease locus

Chen Z-Y., Sims, K.B., Coleman, M., Donnai, D., Monaco, A., Breakefield, X.O., Davies, K.E., and Craig, I.W.

ABSTRACT

It has been shown from pulsed-field gel electrophoresis (PFGE) that the monoamine oxidase genes A and B (MAOA & MAOB) and DXS7 loci are physically very close. We have therefore extended studies on their relationship through the characterisation of a 650 kb YAC isolated using L1.28 (recognising the DXS7 locus) as a probe. Restriction mapping of the YAC indicates that it contains both MAOA and MAOB genes in addition to the DXS7 locus. The map derived from the YL1.28-YAC is compatible both with the map from an independently derived YAC carrying MAOA and B genes and with the long range genomic map for the region .

A series of subclones prepared from a 'phage library (lambda DASH II) of the YAC have been characterised and have been employed to determine the end point of the deletion of a Norrie disease (NDP) patient who has been shown to lack both DXS7 and MAO coding sequences. The pattern of retention of subclones in the deletion patient place the end point of the deletion within 30 - 130 kb of the proximal end of the YAC. By combining the data with established recombination analysis, we provide evidence that all or part of the NDP lies in the interval of approximately 250kb within the YAC.

INTRODUCTION

Norrie disease (pseudoglioma, NDP) is a severe, X-linked, recessive neurological disorder of unknown pathogenesis and is characterised by congenital blindness, sensory neural deafness and mental retardation.

Its localisation to Xp11.4-p11.3 was established by virtue of its manifestation in patients from a Finnish kindred showing an atypical complex phenotype comprising microcephaly, atonic seizures, myoclonus and somatic growth failure. It was deduced that patients from this kindred had a submicroscopic deletion of the X chromosomal region embracing the disease locus and which included the anonymous probe DXS7 (1). Subsequently, both monoamine oxidase A and B genes (MAOA and MAOB) were shown to be deleted in affected individuals in this kindred resulting in a deficiency for monoamine oxidase enzyme activity (2). Three other atypical Norrie disease patients with submicroscopic deletions, including the DXS7 locus, have been described (3,4,5); all of whom lack the MAOA and B genes (6,7; Chen, Z-Y., Powell, J.F. and Craig, I.W.-unpublished data). Physical mapping of the Norrie disease locus in deletion patients has identified a distal flanking marker DXS77, which is recognised by the probe pX59 (8).

Recombination has been observed between DXS7 (L1.28) and NDP (9,10); furthermore, additional observations on one of the individuals revealed that no recombination had occurred between the disease locus and the OAT complex - which is physically located proximal to DXS7 (11). It is therefore suggested that NDP lies proximal to DXS7. This information has been recently extended by the demonstration by Lindsay et al (12) that the cross-over is also above the microsatellite marker, DXS426, which maps proximal to DXS7 at Xp11.4-p11.23.

Because of the close proximity between DXS7, MAOA, MAOB and Norrie disease, we have undertaken a detailed investigation of the physical map in the region. This analysis, together with further analysis of a Norrie patient with a deletion and additional recombination data now available, has enabled us to determine that sequences important in the manifestation of Norrie disease are localized to an interval of about 250kb within a YAC of total size 650kb.

RESULTS

The 650kb YAC we studied, was obtained by screening a YAC library for the anonymous probe DXS7 (see Methods Section); it is designated YL1.28. Detailed restriction mapping of the YAC indicates that it contains both MAOA and MAOB genes in addition to the DXS7 locus (Figure 4.1). DXS7 was mapped to a 30kb ClaI fragment, at the left end of the YAC. DXS7 and MAO genes are on adjacent SfiI fragments within the YAC and separated by about 140 kb. MAOA and MAOB genes are found in a tail-to-tail configuration with the 3' coding sequences separated by about 40kb. The map derived from the YL1.28-YAC is compatible both with the information obtained with the YMAOA-YAC and the long range genomic map for the region (13).

A phage library in lambda DASH II has been constructed from the YL1.28-YAC and single copy sub-clones have been derived from several of the individual 'phage isolates. Additionally, clones corresponding to both ends of the YAC have been isolated by screening with left arm- and right arm-specific sequences. The end clone, designated Yz, which is located 5' to the MAOB gene and is presumed to be centromeric (see below) has been employed to extend the long range map obtained from previous work (Figure 4.1).

The new mapping data integrate perfectly with and further extend those previously established(13) - indicating that the YAC is not a co-ligation product. The location of some of the sub-clones isolated from 'phage mapping between the 5' end of the MAOB gene and the YAC end clone, Yz, has been established. These sub-clones have been employed subsequently in an attempt to determine the end point of the deletion of the Manchester patient who has been shown to lack MAO coding sequences.

The map of the clones employed is illustrated in Figure 4.1 and Figure 4.2 shows the pattern of their retention in the deletion patient. The deletion in the Manchester patient (4) appears to end between the MAOB gene and the YAC end clone, Yz, which is clearly present in this individual - although absent from a second deletion patient described by de la Chapelle et al (1) - (Figure 4.3). Analysis of other sub-clones which have been mapped between the MAOB gene and the YAC end clone (Figures 4.1 and 4.2) shows that, in addition, the Manchester

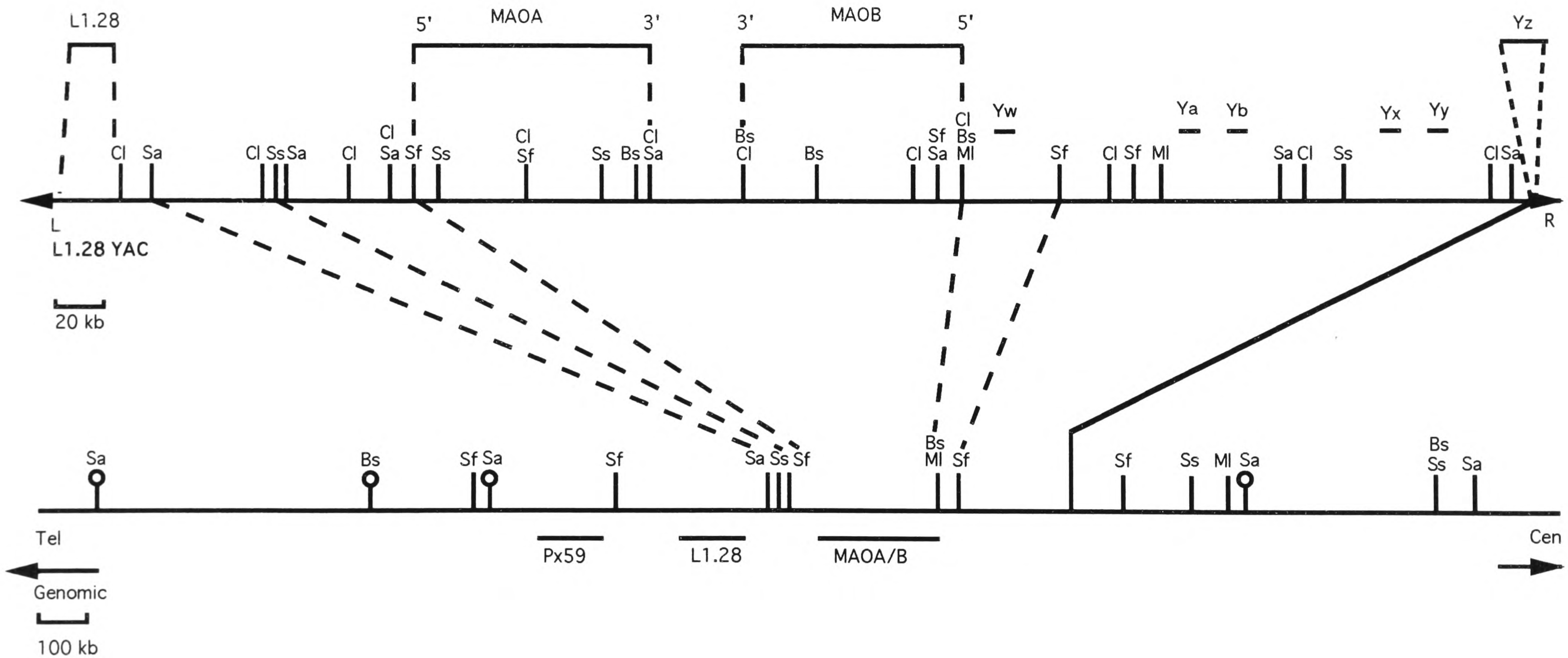


Figure 4.1 Restriction map of the L1.28 YAC indicating the location of probes employed in mapping. The series of lambda sub-clones (lambda DASHII) are indicated Ya, Yb etc.. Long-range restriction map for the region (orientation based on (8, 14)). The dashed line between the YAC and genomic DNA indicates the restriction sites in the YAC which are cleavable in the genome. ‡ represents sites which are methylated on the active X chromosome (see (13)).

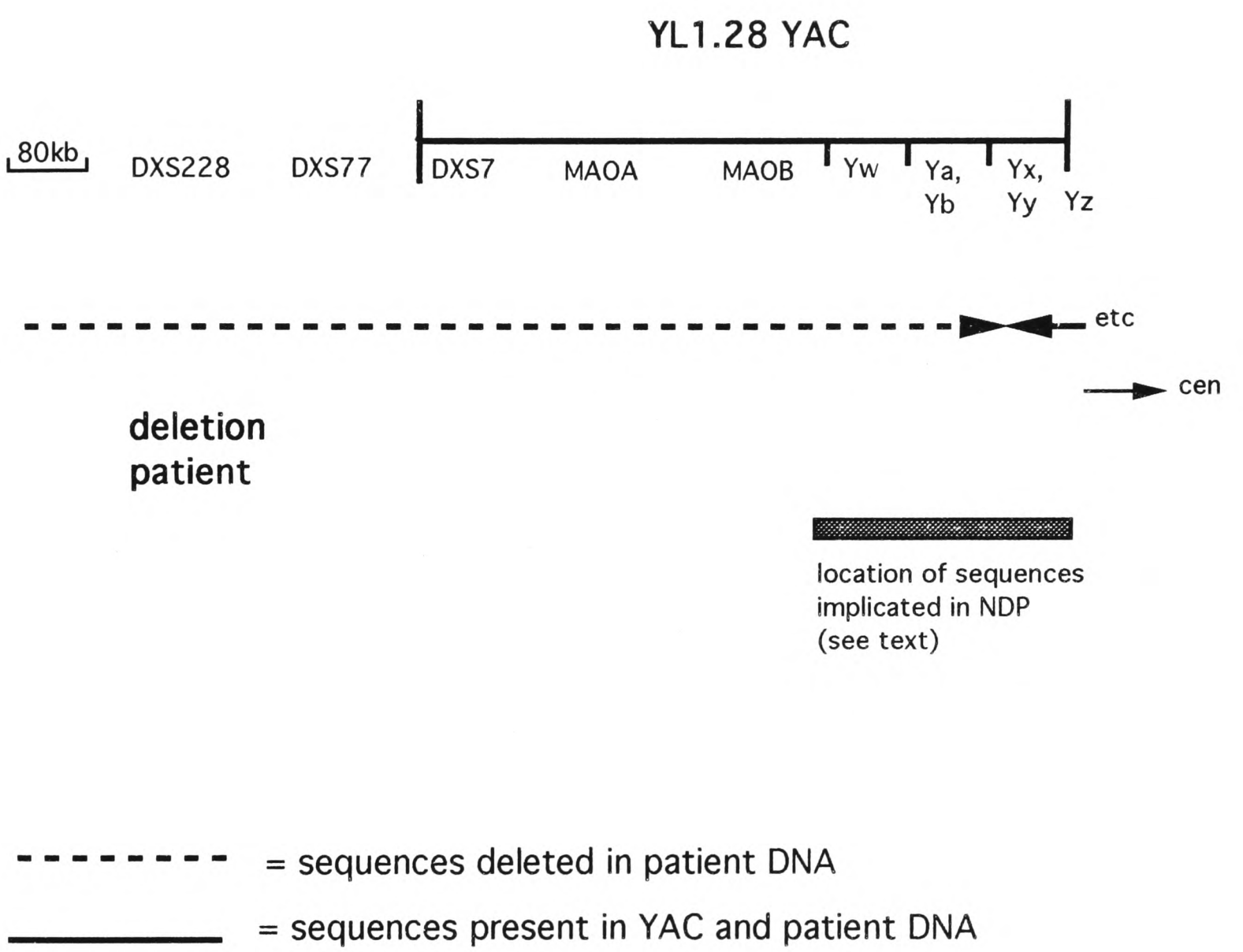


Figure 4.2 The physical order of probes in the region of Norrie disease including those represented in or derived from the YAC and the position determined for the proximal end point of the deletion. The distances represented are approximate.

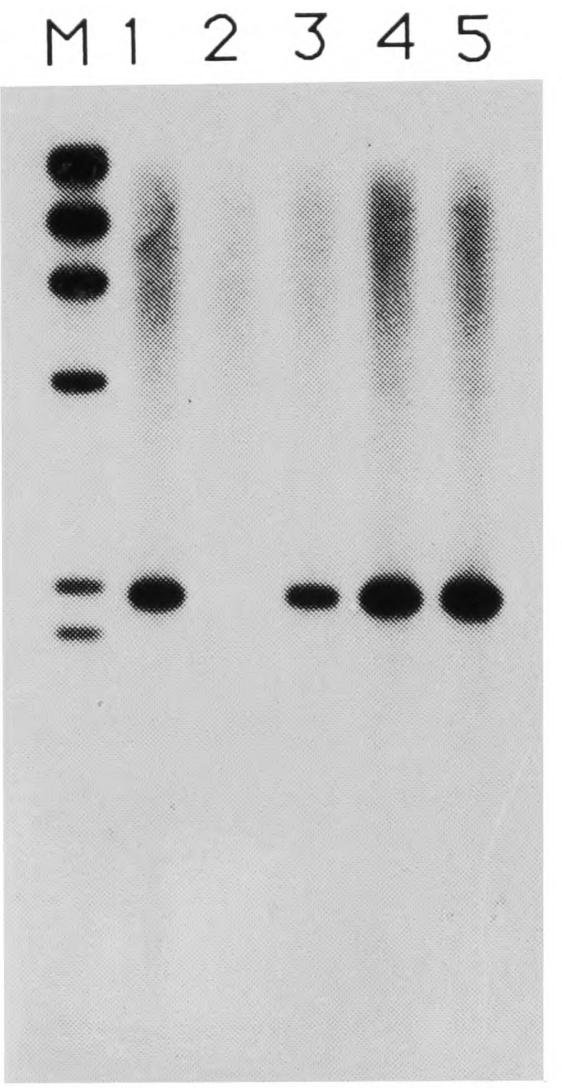


Figure 4.3 *Hind*III digest of patient and control DNA probed with the YL1.28 YAC end clone. M = Marker HindIII digest of lambda phage. 1, deletion patient (Donnai et al, 1988); 2, deletion patient (de la Chapelle et al, 1985); 3, normal male; 4, normal female: 5, normal female.

patient retains the sequences Yx and Yy, but is deleted for Ya,Yb and Yw thereby positioning the proximal breakpoint of the deletion between about 30kb and about 130kb from the right hand end of the YAC. This is surmised because the probes Ya and Yb are from independent 'phages with inserts of about 15kb, both of which lie in a MluI/SstII fragment of 80kb, which is 80kb distant from Yz. The probes Yx and Yy have been mapped to the 80kb interval between an SstII site (Figure 4.1) and the right hand end of the YAC and, because both are present in the Manchester patient in addition to the end clone, the minimum distance of the proximal breakpoint from the terminus of the YAC has therefore been estimated at 30kb. In contrast, none of the sequences isolated from the YL1.28 YAC, are present in the patient described by de la Chapelle et al (1).

DISCUSSION

Our studies have attempted to refine the position of the Norrie disease locus by mapping one end of a submicroscopic deletion present in a patient with respect to a series of probes - whose physical linkage has been established. The probes that we have examined are present in a 650kb YAC which also contains the DXS7, MAOA and MAOB loci. The deletion patient chosen for study here was previously known to lack DXS7, DXS77 and DXS228 (4). Two of these probes, DXS77 and DXS228, are not, however, deleted in the patient described by Diergaarde et al, (8); nor are they represented in the 650kb YAC.

The physical map for the YAC and the pulsed field electrophoretic analysis which we have established confirms the arrangement of the MAO genes which we described previously (13, 14) and provides physical distances between the DXS7, MAOA and MAOB loci. The map for the region overlaps and extends in one direction that given previously by Diergaarde et al (8) based on analysis with DXS228, DXS77, DXS7 and MAO. The orientation which they provide is with DXS228 telomeric and, while there are only few data available which support this, we have made independent observations which confirm this configuration (14).

The ordering of NDP with respect to DXS7 and the centromere is obviously of considerable importance to the construction of the map for the region and the localisation of the disease locus. Previous studies have shown recombination between DXS7 and the Norrie disease locus (9,10); however, in the one informative DXS7-recombinant individual, no recombination was observed between NDP and OAT (11), or between NDP and DXS426 (12). This suggests that the probable location of the disease locus is between DXS7 and the centromere. Recent recombination analysis of the same individual employing a microsatellite marker (15) has enabled the Norrie disease locus to be positioned proximal to the second intron of the MAOB gene (14). Given this, together with the observation that the proximal end of the deletion in the patient studied here lies close to the end of the YL1.28-YAC, but does not include the probes Yx,Yy or Yz, it is most likely that a region involved in the expression of the disease is located in the interval of about 250kb between the second intron of the MAOB gene and the minimum distance estimated for the end point of the Manchester patient's deletion (about 30kb from the YAC end).

The order with respect to MAO is therefore: tel_DXS7_MAOA_MAOB_NDP_Yx_cen.

There is no information on the nature of the gene product presumed to be deficient in Norrie disease and the region of the YAC identified as having importance may contain the entire locus or a small part of the relevant structural and/or regulatory elements. Nevertheless, the narrowing down of the interval which contains sequences whose deletion results in Norrie disease provides the basis for a strategy to isolate the locus, both through the isolation of conserved sequences for the region and the screening of appropriate cDNA libraries.

MATERIALS AND METHODS

L1.28 Yeast artificial chromosome

A yeast clone containing a 650kb yeast artificial chromosome, was identified by screening a 48, XXXX cell line YAC library (16) with the probe L1.28 which reconizes the DXS7 locus. Agarose plugs containing DNA from the YAC host were prepared according to the protocol of Anand and Southern (17). For restriction analysis, the plugs were washed twice in 1 x TE at 50°C for 30'. The plugs were then equilibrated in appropriate restriction buffer on ice for 30' before being subject to digestion, for which a third of each plug was used. For the long range restriction map, the YAC was digested with various rare cutting enzymes (either completely or partially) and the products were fractionated by pulsed-field gel electrophoresis on a "Waltzer" apparatus (17) in 0.5 x TAE, 1.5% agarose (Sigma type II), employing 45 sec. switch time at 150V for 36 hours. Conditions for blotting and hybridization following PFGE were the same as those detailed previously (18) except that the gel was exposed to UV for 4' before blotting. The enzymes used for pulsed field gel analysis were BssHII, ClaI, MluI, SalI, SfiI and SstII.

Preparation and digestion of high-molecular-weight genomic DNA

DNA plugs were made from peripheral blood leukocytes of several normal males and a normal female. Cells were embedded in 0.5% PBS low-gelling-temperature agarose (Sea Plaque, FMC Bioproducts) at a concentration of 1.5 x 10^7 cells/ml. Plugs containing 100ul cell suspension were set in Perspex moulds. For double digests, each sample was washed, following the first digestion, in 1 x TE for an hour in order to elute any trace of enzyme. The sample was then re-equilibrated in an appropriate buffer. The enzymes used for genomic digestion were BssHII, MluI, SalI, SfiI and SstII and the hybridizations were performed as was described for the YAC PFGE filters.

Probes

Two full-length cDNAs of MAO were used in this study, MAOA cDNA designated HM11 (19)

and MAOB cDNA designated BSMAOB (a gift from Dr. Julie Andersen, Massachusetts General Hospital). Genomic plasmid clones containing the first (3-1), the second (A2) and the last exon (exon-15) of the MAOA gene and the last (15th) exon of MAOB (6-B) were derived from the phage clones isolated from a human 4X amplified EMBL-3 phage library (a gift from Dr. Derek Blake, Oxford). MAO-B1, the first exon of MAOB, was prepared by digestion of MAOB cDNA with PstI and Tth111I which releases a sequence covering the region of the first exon. The probes used to detect either end of the YAC were prepared from BamHI/PvuII digested pBR322 plasmid. The derived fragments of 1.7 kb (PBR-R) and 2.7 kb (PBR-L), which hybridize specifically to the right and the left arms of the YAC respectively, were used to hybridize to filters with partial digests of the YAC. Probes pX59 (for locus DXS77), L1.28 (for locus DXS7) and 1aA6 (for locus DXS228) have been described previously (20).

Norrie disease patient

A cell-line has been established from this patient described before(4). To avoid confusion with other deletion patients this is referred to in the text as the Manchester patient.

Construction and screening of a bacteriophage library derived from the YL1.28-YAC.

A 'phage library was constructed in a lambda Dash II vector (Stratagene) as described in Sambrook, Fritsch and Maniatis (21). Phage clones containing the left and right ends of the YAC insert sequences were isolated by screening the library using YAC left and right arm-specific pBR322-plasmid sequences. The 2.2kb HindIII fragment from the right end of the YAC was subcloned in pUC9 .

REFERENCES

1.de la Chapelle, A., Sankila, E.-M., Lindof, M., Aula, P. and Norio, R. (1985).*Clin. Genett.* **28**, 317-320.

2.Sims, K. B., Ozelius, L., Corey, T., Rinehart, W. B., Liberfarb, R., Hanies, J., Chen, W. J., Norio, R., Sankila, E.-M de la Chapelle, A., Murphy, D. L., Gusella, J., and Breakefield, X. O. (1989) *Am. J. Hum. Genet.* **45**, 424-434.

3.Gal, A., Wieringa, B., Smeets, D.F., Bleeker-Wagemakers, L.M. and Ropers, H.H. (1986) *Cytogenet. Cell Genet.* **40**, 219-224.

4.Donnai, D., Mountford, R.C. and Read, A.P. (1988). *J. Med. Genet.* **43**, 575-578.

5. Zhu, D., Antonarakis, S.E., Scheckpepper, B.J., Diergaarde, P.J., Greb, A.E. and Manmenee, I.H. (1989) *Am. J. Med. Genet.* **33**, 485-488.

6.Lan, N.C., Heinzmann, C., Gal, A., Klisak, I., Ooth, U., Lai, E., Grimbsby, J., Sparkes, R.S., Mohandas, T. and Shih, J.C.(1989). *Genomics* **4**, 552-559.

7.Collins , F.A. Murphy, D.L., Reiss, A., Sims, K.B., Lewis, J. and Antonarakis, S.E. (1992) *Am. J. Med. Genet.* **42**, 127-134.

8.Diergaarde, P.J., Wieringa, B., Bleeker-Wagemakers, L.M., Sims, K.B., Breakefield, X.O. and Ropers, H.H. (1989).*Hum. Genet.* **84**, 22-26

9.Katayama, S., Wohlferd, M. and Golbus, M.S. (1988).*Am. J. Med. Genet.* **30**, 967-970.

10.Ngo, J. T., Spence, M.A.,. Cortessis, V., Sparkes, R. S., and Bateman, J.B. (1988) *Clin. Genet.* **34**, 43-47.

11.Ngo, J. T., Bateman, J. B., Cortessis, V., Sparkes, R. S., Mohandas, T., Inana, G., and Spence, M. A. (1989) *Genomics* **4**, 539-545.

12.Lindsay, S., Thiselton, D.L., Bateman, J.B., Ngo, J.T., Sparkes, R.S., Coleman, M. , Davies, K.E. and Bhattacharya, S.S. (1992) *Hum Genet.* **88**, 349-350.

13.Chen, Z-Y., Powell, J.F., Hsu, Y.-P.P., Breakefield, X.O. and Craig, I.W. (1992) *Genomics* in press

14.Sims, K. B., Lebo, R.V., Benson, G., Shalish, C., Chen, Z-Y., Bruns, G., Golbus, M.S., Craig, I.W.

and Breakefield, X. O. (1992).*Hum. Molec. Genet.* in press

15.Konradi, C., Ozelius, L. and Breakefield, X.O. (1992).*Genomics* **12**, 176-177.

16.Larin, Z., Monaco, A.P. and Lehrach, H. (1991).*Proc. Natl. Acad. Sci. USA* **88**, 4123- 4217.

17.Anand, R. and Southern, E,M. (1990) In Rickwood, D. and Hames, B.D. eds, "Gel electrophoresis of Nucleic Acids" IRL Press, Oxford. pp 101-123.

18.Ross M.T., Ballabio, A., and Craig, I.W. (1990) *Genomics* **6**, 528-539.

19.Hsu, Y.-P. P., Weyler, W., Chen, S., Sims, K. B., Rinehart, W. B., Utterback, M. C., Powell, J. F., and Breakefield, X. O. (1988) *J. Neurochem.* **51,** 1321-1324.

20.Davies, K.E., Mandel, J-L., Monaco, A.P,. Nussbaum, R.L. and Willard, H.F. (1991) *Cytogenet. Cell Genet.* **58**, 853-966.

21.Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual 2nd edition. Cold Spring Harbor Laboratory

Nature genetics 1, 204-208 (1992)

Chapter 5. Isolation and characterisation of a candidate gene for Norrie disease

Chen, Z.-Y., Hendriks, R.W., Jobling, M.A., Powell, J.F., Breakefield, X.O., Sims, K.B.and Craig, I.W.

SUMMARY

Previous analysis has refined the location of the gene for Norrie disease, a severe X-linked neurodevelopmental disorder, to a YAC sub-fragment of 160kb. This fragment was used to screen cDNA libraries from human fetal and adult retina. We have identified an evolutionarily conserved cDNA, which is expressed in fetal and adult brain, and encodes a predicted protein of 133 amino acids. The cDNA detects genomic sequences which span a maximum of 50kb, and which are partly deleted in several typical Norrie disease patients. An *EcoR*I polymorphism with a calculated heterozygosity value of 43% was observed. The locus identified is a strong candidate for the Norrie disease gene.

INTRODUCTION

Norrie disease (Pseudoglioma, NDP, McKusick 31060) is a severe, X-linked recessive neurodevelopmental disorder of unknown pathogenesis[1,2]. In typical affected males, congenital blindness is accompanied by a retinal dysplasia with pseudoglioma formation and often associated with mental retardation and progressive sensorineural deafness. Numerous examples exist of intrafamilial variability in cognitive function and in the onset or severity of hearing loss in affected males (K.B.S., unpublished data). NDP, although an uncommon disorder, about 1 in 100,000 is of considerable interest because of its severity and its significance in the context of brain and eye development.

Deletion mapping and recombination studies have led to a refined localisation of the NDP locus on the X chromosome. DNA analysis on four Norrie patients from unrelated kindreds[3-6] displaying complex and atypical phenotypes, including microcephaly, atonic seizures, myoclonus and somatic growth failure, indicated that these affected males have submicroscopic deletions in Xp11.2-p11.3 which include the disease locus. These deletions encompass the locus DXS7, detected by the anonymous probe L1.28, as well as the monoamine oxidase A and B genes (MAOA and MAOB)[7,8]. MAOA and MAOB have been excluded as NDP candidate genes by DNA and enzyme studies[9]. Linkage data[10-13] favour the order Xpter - DXS7 - NDP - DXS426 - Xcen, and place the disease locus in an interval of approximately 12cM.

A more precise localisation of NDP has relied on long range genomic restriction mapping and the analysis of two yeast artificial chromosome (YAC) clones, both of which contain MAOA and MAOB, and one of which (YL1.28) also contains DXS7[14-16]. Subclones isolated from a lambda phage library, constructed from partial *Mbo*I digests of YL1.28, were mapped against the DNA from two microdeletion patients; these studies have determined the position of the proximal ends of the deletion and indicated that all or part of the NDP locus was contained within YL1.28[16]. A single recombination event between a microsatellite marker within the MAOB gene and NDP[17] confined the disease locus to an interval of about 160kb. We have used a 160kb sub-fragment isolated from YL1.28 to screen cDNA libraries from human retina. Clones identified in these experiments detect genomic sequences which are deleted in several typical Norrie disease patients. These clones derive

from a gene that is a strong candidate for the NDP locus.

RESULTS

Isolation of candidate cDNAs from the NDP region

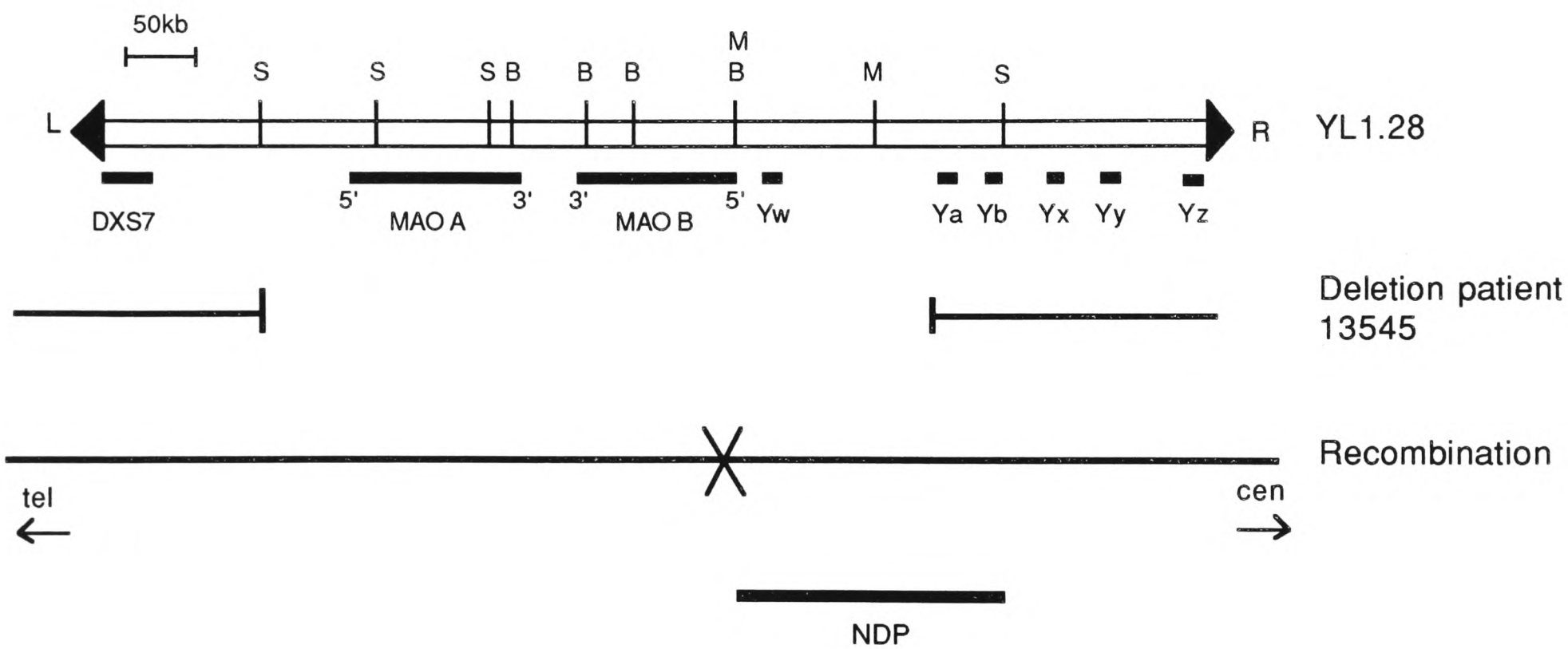
Previous experiments have demonstrated that all or part of the Norrie disease locus must lie within a 160kb *Bss*HII-*Sst*II YAC fragment (Figure 5.1a) of a 650kb YAC clone (YL1.28), isolated from the library of Larin et al.[18], by using L1.28 as a probe[14,16].

This 160 kb *BssH*II-*Sst*II fragment was separated from a double digest of YL1.28 DNA by preparative pulsed-field gel electrophoresis, purified and used as a hybridisation probe to screen two human cDNA libraries, one from fetal and the other from adult retina. Five clones from the fetal retinal library (designated FR.1, FR.2, FR.7, FR.12, and FR.14), and two from the adult library (AR.4 and AR.13) were isolated, and mapped within the YAC by hybridisation to *Bss*HII, *Sst*II and *Mlu*I digests. The patterns of hybridisation of the cDNAs to *Eco*RI, *Hin*dIII and *Pst*I digests of genomic DNA suggested that the clones FR.2, FR.7, AR.4 and AR.13 were related in sequence, as were clones FR.12 and FR.14 (data not shown).

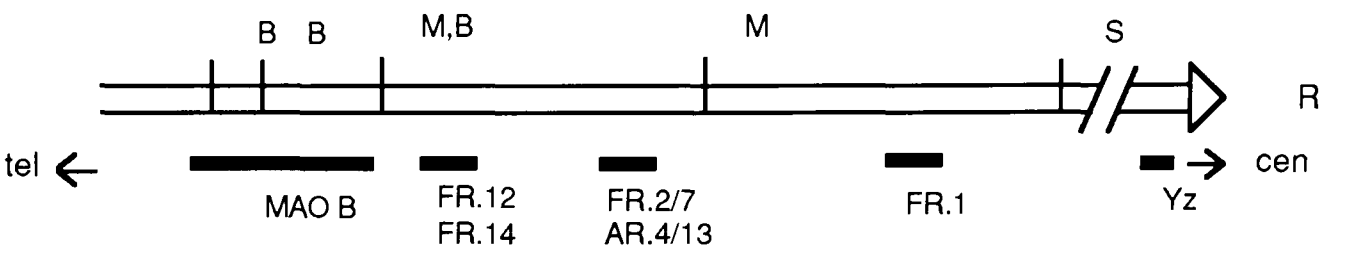
These cDNA clones were used to probe a panel of DNA digests from five microdeletion patients (Figure 5.1b). The presence of the full sequence detected by FR.1 in two of the five patients excludes this cDNA as a likely candidate. Sequences corresponding to the related cDNAs FR.12 and FR.14 were found to be absent in all microdeletion patients. Several lines of evidence indicate that these two clones are unlikely to derive from the NDP gene. Of the fourteen human genomic *Hin*dIII fragments detected by the 3.5kb cDNA, FR.12, only one band, of 1.8kb, is present in the human X-only hybrid ThyB-X. Analysis of mouse/hamster hybrids indicates the absence of any sequences homologous to FR.12/FR.14 on the mouse X chromosome. These data suggest that the FR.12/FR.14 cDNAs originate from an autosomal gene with an X-linked pseudogene homologue in human.

In contrast, genomic sequences detected by FR.2, FR.7, AR.4, and AR.13 were partly or completely deleted in all five microdeletion patients (see Figure 5.1b). The 1.7kb cDNA FR.2 detects three *Hin*dIII fragments of 4.7, 3.6 and 1.3 kb in genomic DNA (Figure 5.2a). All of these fragments are X-linked on the basis of dosage in samples with varying X chromosome copy number and presence in human X-only hybrids; all three are also present in YL1.28 (Figure 5.2a), and are contained within a single *Mlu*I fragment of 80kb (Figure 5.2b). The murine homologue of FR.2 was

5.1A.


50kb
S
S
S B
B
B
M
B
M
S
L
R
YL1.28
DXS7
5'
MAO A
3'
3'
MAO B
5'
Yw
Ya
Yb
Yx
Yy
Yz
Deletion patient
13545
Recombination
tel
cen
NDP

5.1 B



Deletion Patient					
13545 (17)	-	-	-/+	+	+
10017 (*)	-	-	-/+	+	+
Manchester(5)	-	-	-	-	+
LGL515 (3)	-	-	-	-	-
NHF88 (6)	-	-	-	-	ND

Figure 5.1 Physical map of the *NDP* region. a, Restriction map of YL1.28 indicating locations of monoamine oxidase genes MAOA and MAOB, DXS7, and various λ phage subclones (Ya to Yz) used in mapping the end points of the deletion patient 13545. The most distal position for a single recombination event reported which orients MAOB and NDP with respect top the centromere, is designated by a cross. Left and right ends of the YAC are indicated by L and R respectively. M, *Mlu*I, S, *Sst*II, and B, *Bss*HII. The lower solid bar represents the 160 kb region identified as important in the aetiology of NDP.b Deletion mapping of cDNA clones. Map positions of MAOB, cDNA clones isolated and the λ phage clone (Yz) are indicated below the restriction map of the YAC. cDNA clones below the same solid line are related. +, present; -, absent; -/+, partially deleted; ND not done. * This study. M, *Mlu*I; S, *Sst*II and B, *Bss*HII.

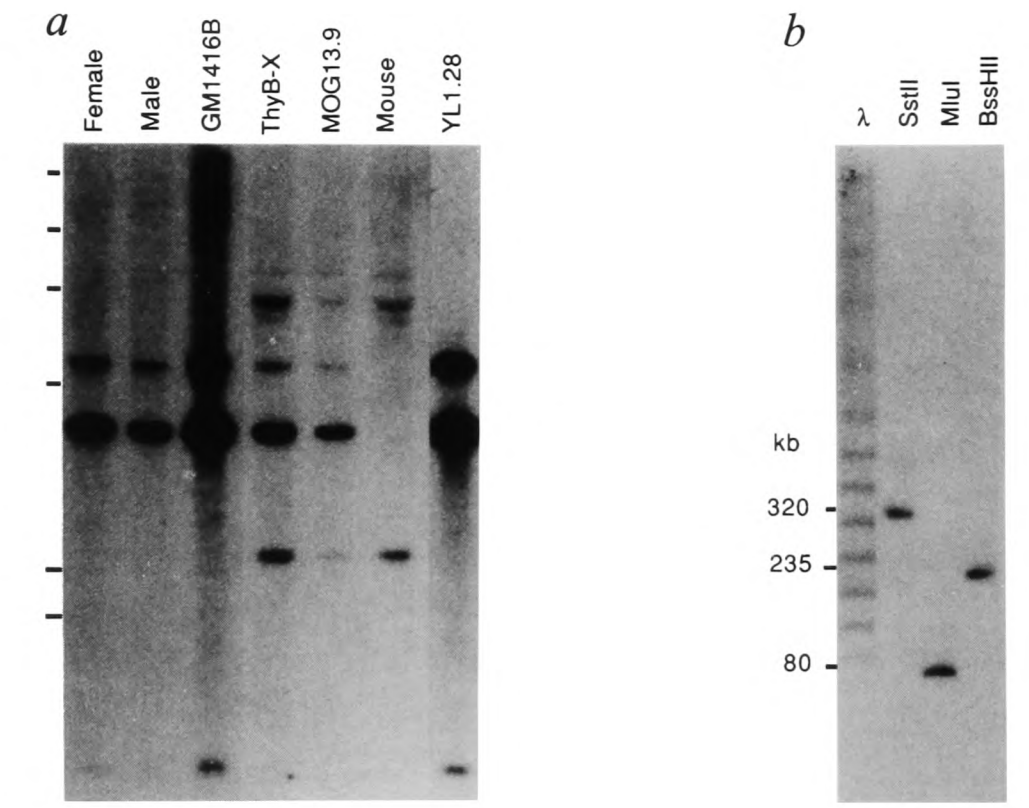


Figure 5.2 a: Hybridisation of cDNA clone FR.2 to *Hin*dIII digests of control male, female, GM1416B (48,XXXX), human X-only hybrids (ThyB-X and MOG13.9), mouse and the YAC (YL1.28). Positions of *Hin*dIII-digested lambda DNA markers are indicated.b: Hybridisation of FR.2 to indicated digests of YL1.28, showing that all of the sequences detected are contained within a single *Mlu*I fragment of 80kb. λ : oligomers of phage lambda DNA.

found to be confined entirely to the X chromosome (data not shown). It was concluded that these clones were possible NDP candidates.

FR.2 detects small deletions in typical NDP patients

FR.2 was used to screen panels of *Eco*RI, *Hin*dIII, and *Pst*I digests of DNA from twenty typical Norrie patients, not previously known to have deletions, and the five microdeletion patients, as well as normal males and females.

In genomic *Eco*RI digests of DNA from normal individuals, FR.2 detects two invariant fragments (20kb and 5kb), and one polymorphic fragment (9 or 10kb). The calculated heterozygosity for the RFLP was 43%, based on observations of fifty-two X chromosomes (twenty males and sixteen females). Three microdeletion patients are deleted for all three *Eco*RI fragments, whereas two lack only the 20kb fragment (13545 and 10017; Figure 5.3). Five out of twenty typical Norrie patients were found to have abnormal hybridisation patterns. Two patients (3883 and 12316) lack the 9/10kb *Eco*RI band and one (6446) demonstrates considerable reduction in hybridisation to the 20kb band. In patient 8810, the 5kb *Eco*RI fragment is slightly reduced in size (Figure 5.3). In *Hin*dIII digests of DNA from patients 3883 and 12316 the 4.7kb fragment was absent, whereas in patient 6446 an altered fragment of 3.8kb instead of 3.5kb was detected (data not shown). In a fifth case (12315), showing normal *Eco*RI and *Hin*dIII patterns, an altered *Pst*I fragment was observed. Analysis of the group of normal males and females showed no evidence of deletions or rearrangements.

Properties of the candidate cDNAs

Hybridisation to genomic digests of distantly related species, including *Drosophila* (Figure 5.4), indicates that the FR.2 sequence is highly conserved. In Northern analysis of poly-A-enriched RNA from rat tissues, strong hybridisation was seen in brain, but not in liver. Two transcripts (in the range of 1.7 to 2kb) were detected in fetal brain, only the smaller of which was observed in adult brain (Figure 5.5).

Sequence analysis of FR.2, FR.7, and AR.13 showed that the three clones were identical, except for the extent of their 5' ends; the entire 1846bp sequence of the longest clone, FR.7, is shown in Figure 5.6. It contains an open reading frame, starting with an ATG at position 409, which fulfils the

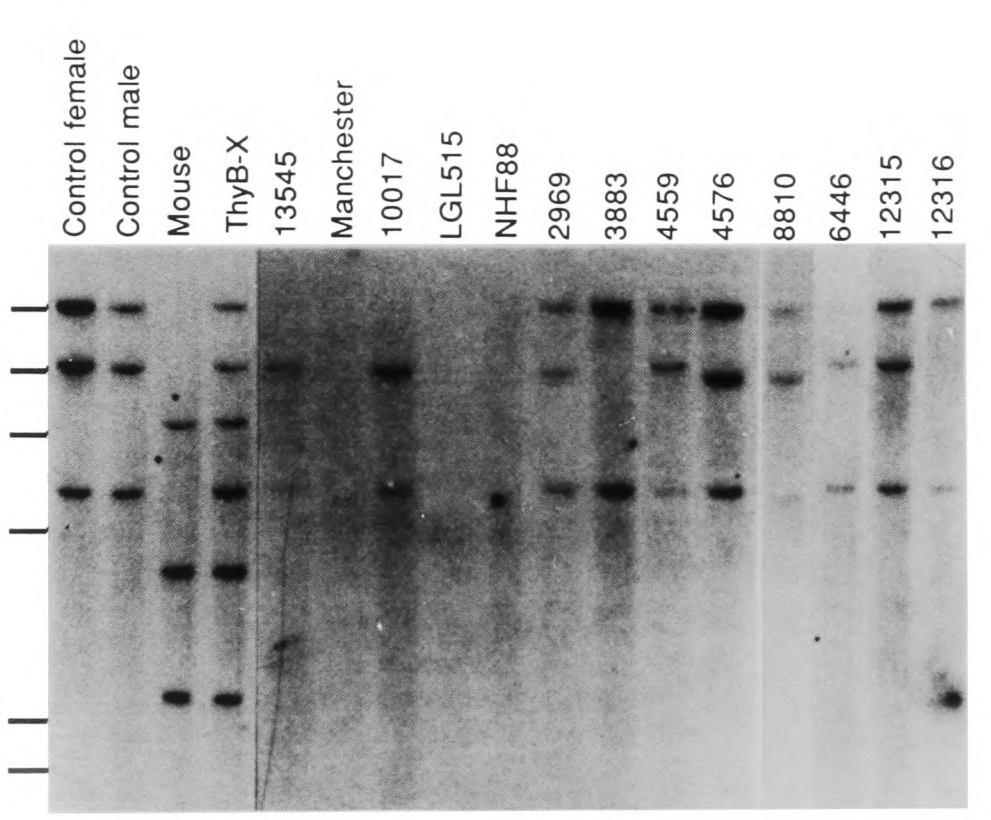


Figure 5.3 Hybridisation of FR.2 to *Eco*RI digests of DNA from control male and female, mouse and human X-only hybrid, ThyB-X. Other lanes contain samples from various Norrie patients as described in the text. Positions of *Hin*dIII-digested lambda DNA markers are indicated.

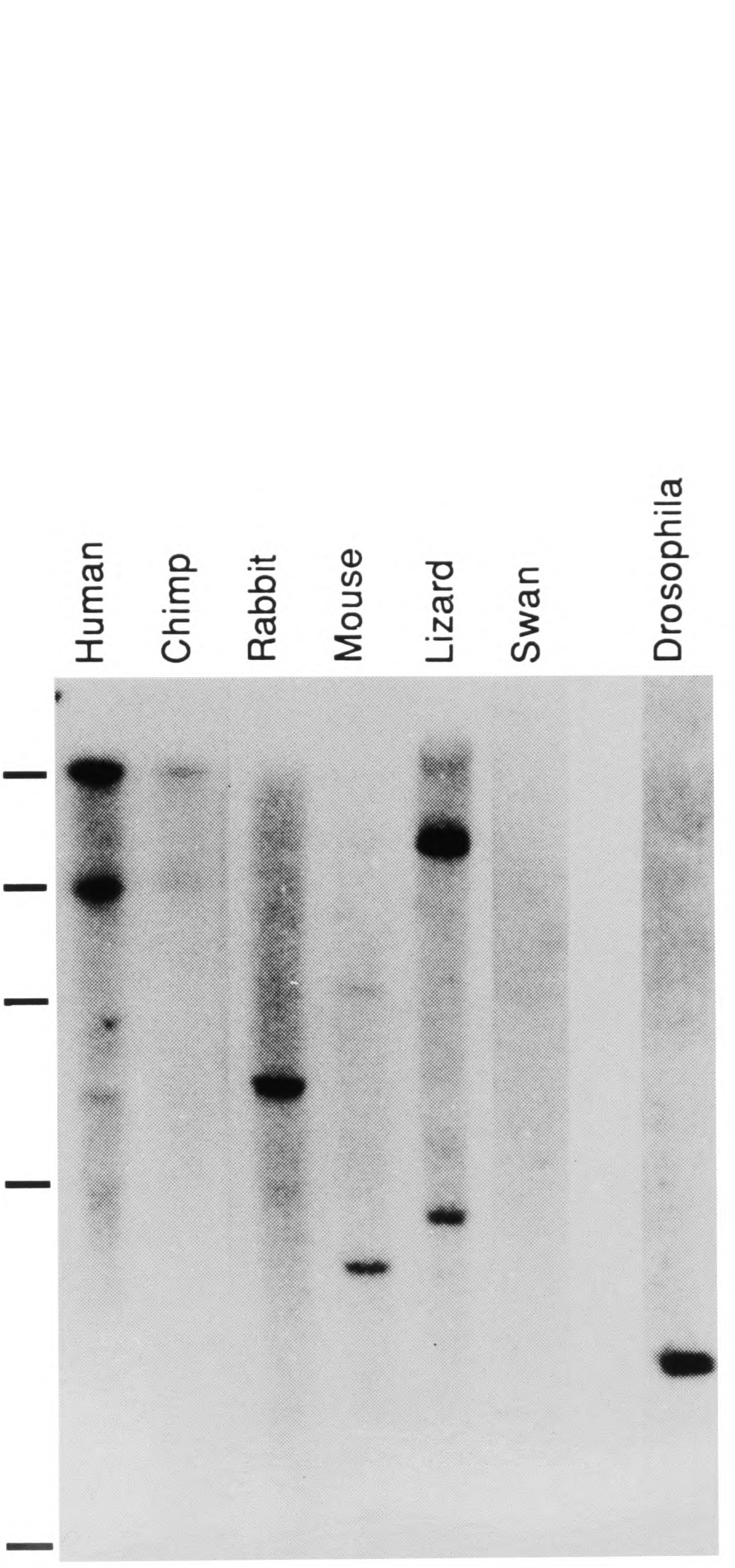


Figure 5.4 *Eco*RI digests of DNA from a range of species probed with the cDNA, FR.2. Washing was as described in Methodology, except for the *Drosophila* filter, which was washed to 400mM sodium phosphate, pH7.2 / 0.1% SDS at 65°C. Positions of *Hin*dIII-digested lambda DNA markers are indicated.

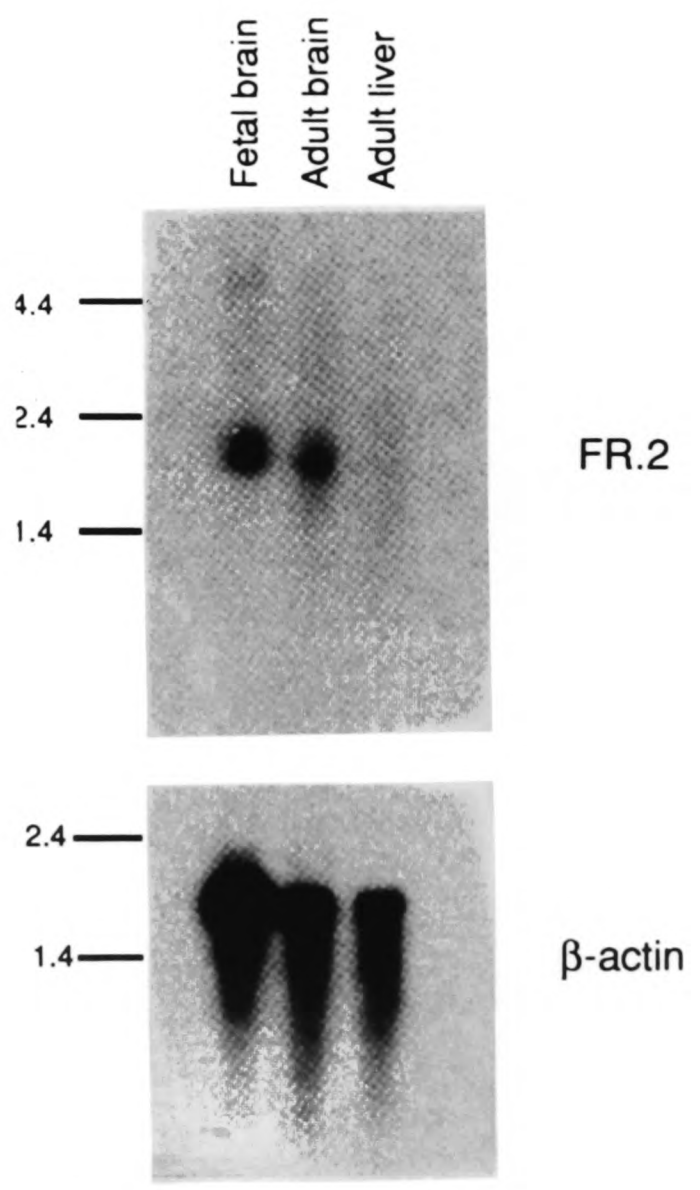


Figure 5.5 Northern blot of poly-A-enriched mRNA from various rat tissues probed with FR.2, and subsequently with human β-actin. Positions of relevant RNA molecular weight markers (Gibco BRL) are indicated.

```
   1 CCTCTCTCTCTCTCCCTCTCTCTCTCCCTCTCTCTCTCTCCCTGTGTCGCTTAAACAACA
  61 GTCCTAACTTTTGTGTGTTGCAAATATAAAAGGCAAGCCATGTGACAGAGGGACAGAAGA
 121 ACAAAAGCATTTGGAAGTAACAGGACCTCTTTCTAGCTCTCAGAAAAGTCTGAGAAGAAA
 181 GGAGCCCTGCGTTCCCCTAAGCTGTGCAGCAGATACTGTGATGATGGATTGCAAGTGCAA
 241 AGAGTAAGACAAAACTCCAGCACATAAAGGACAATGACAACCAGAAAGCTTCAGCCCGAT
                              BamHI
 301 CCTGCCCTTTCCTTGAACGGGACTGGATCCTAGGAGGTGAAGCCATTTCCAATTTTTTGT

 361 CCTCTGCCTCCCTCTGCTGTTCTTCTAGAGAAGTTTTTCCTTACAACAATGAGAAAACAT
 (1)                                                 MetArgLysHis

 421 GTACTAGCTGCATCCTTTTCTATGCTCTCCCTGCTGGTGATAATGGGAGATACAGACAGT
 (5) ValLeuAlaAlaSerPheSerMetLeuSerLeuLeuValIleMetGlyAspThrAspSer

 481 AAAACGGACAGCTCATTCATAATGGACTCGGACCCTCGACGCTGCATGAGGCACCACTAT
(25) LysThrAspSerSerPheIleMetAspSerAspProArgArgCysMetArgHisHisTyr

 541 GTGGATTCTATCAGTCACCCATTGTACAAGTGTAGCTCAAAGATGGTGCTCCTGGCCAGG
(45) ValAspSerIleSerHisProLeuTyrLysCysSerSerLysMetValLeuLeuAlaArg
               PstI
 601 TGCGAGGGGCACTGCAGCCAGGCGTCACGCTCCGAGCCTTTGGTGTCGTTCAGCACTGTC
(65) CysGluGlyHisCysSerGlnAlaSerArgSerGluProLeuValSerPheSerThrVal

 661 CTCAAGCAACCCTTCCGTTCCTCCTGTCACTGCTGCCGGCCCCAGACTTCCAAGCTGAAG
(85) LeuLysGlnProPheArgSerSerCysHisCysCysArgProGlnThrSerLysLeuLys

 721 GCACTGCGGCTGCGATGCTCAGGGGGCATGCGACTCACTGCCACCTACCGGTACATCCTC
(105) AlaLeuArgLeuArgCysSerGlyGlyMetArgLeuThrAlaThrTyrArgTyrIleLeu

 781 TCCTGTCACTGCGAGGAATGCAATTCCTGAGGCCCGCTGCTGTGTGTGGCTTCTGGATGG
(125) SerCysHisCysGluGluCysAsnSer***

 841 GACAACTGTAGAGGCAGTTCGACCAGCCAGGGAAAGACTGGCAAGAAAAGAGTTAAGGCA
 901 AAAAAGGATGCAACAATTCTCCCGGGACTCTGCATATTCTAGTAATAAAGACTCTACATG
 961 CTTGTTGACAGAGAGAGATACTCTGGGAACTTCTTTGCAGTTCCCATCTCCTTTCTCTGG
1021 TACAATTTCTTTTGGTTCATTTTCAGATTCAGGCATTTTCCCCCTTGGCTCTCAATGCTG
1081 TTTGGGTTTCCAACAATTCAGCATTAGTGGGAAAAAGTGGGCCCTCATACACAAGCGTGT
1141 CAGGCTGTCAGTGTTTGGTGCACGCTGGGGAAGAATTTACTTTGGAAAGTAGAAAAGCCC
1201 AGCTTTTCCTGGGACATCTTCTGTTATTGTTGATGTTTTTTTTTACCTTGTCATTTTGGT
1261 CTAAGGTTGCCATTGCTGCTAAAGGTTACCGATTTCAAAGTCCAGATACCAAGCATGTGG
1321 ATATGTTTAGCTACGTTTACTCACAGCCAGCGAACTGACATTAAAATAACTAACAAACAG
1381 ATTCTTTTATGTGATGCTGGAACTCTTGACAGCTATAATTATTATTCAGAAATGACTTTT
1441 TGAAAGTAAAAGCAGCATAAAGAATTTGTCACAGGAAGGCTGTCTCAGATAAATTATGGT
1501 AAAATTTTGTAAGGGAGCAGACTTTTAAAGACTTGCACAAATACGGATCCTGCACTGACT
1561 CTGGAAAAGGCATATATGTACTAGTGGCATGGAGAATGCACCATACTCATGCATGCAAAT
1621 TAGACAACCAAGTATGAATCTATTTGTGGGTGTGCTATAGCTTTAGCGTGTCACGGGCAT
1681 CATTCTCTAATATCCACTTGTCCATGTGAAACATGTTGCCAAAATGGTGGCCTGGCTTGT
1741 CTTCTGAACGTTTGGTTCAAATGTGTTTTGGTCCTGGAGGCTCAAATTTTGAGTTATTCC
1801 CACGTTTTGAAATAAAAAGAGTATATTCAAAAAAAAAAAAAAAAAA
```

Figure 5.6 Complete nucleotide and predicted amino acid sequence of NDP candidate clone FR.7. Two polyadenylation sites and a putative TATA box are underlined, and the 5' termini of clones FR.2 and AR.13 indicated by dots. Positions of the BamHI and PstI sites illustrated in Fig.5.7 are also indicated. Nucleotide positions are numbered; amino acid positions are numbered in parentheses. A second potential open reading frame, of 282 bp, begins at position 908 of the sequence; however, Fickett's analysis indicates that this is very unlikely to be used.

sequence context criteria for an initiation codon in that it has an A at position -3[19]. The reading frame extends for 399bp, and terminates with a TGA at nucleotide 808; there are two canonical polyadenylation signals, at positions 944 and 1811. The sequence ends with a poly-A tail.

The 399bp open reading frame (ORF) coincides with a coding region identified by Fickett's analysis[22], and predicts a protein of 133 amino acids of molecular weight 15,043. This protein shows no regions of extended homology to sequences in the Swiss Protein Database. There are no detectable domains of known biological function, and no glycosylation sites.

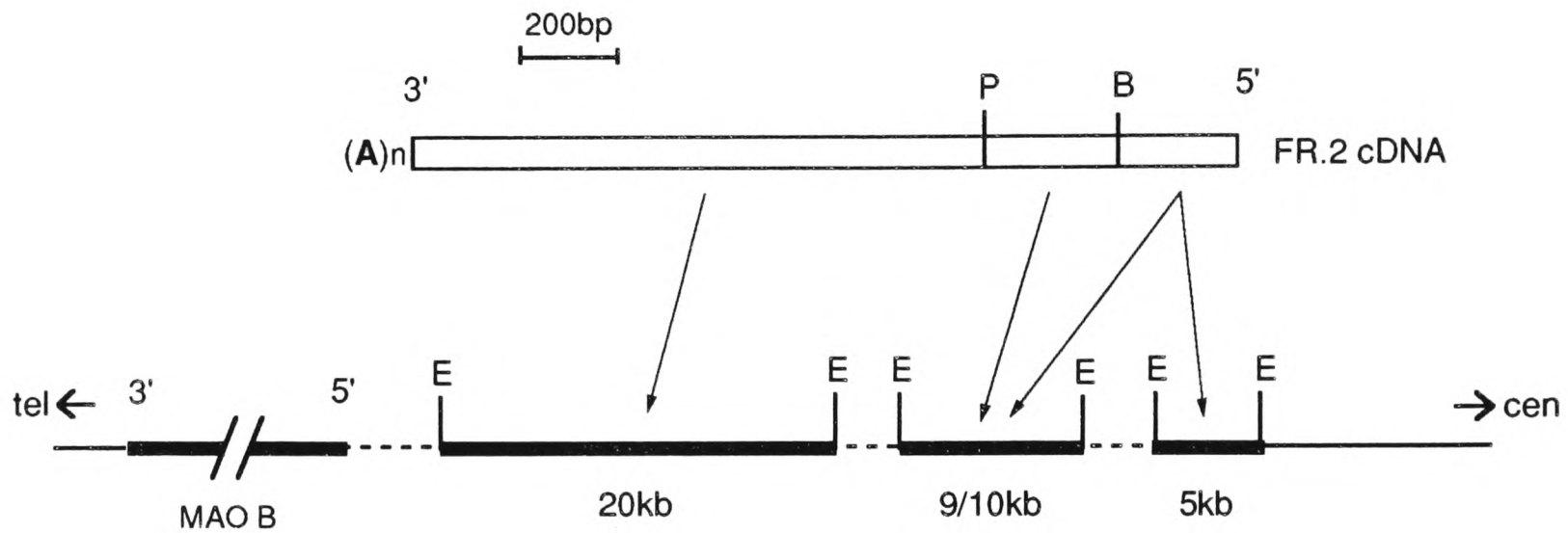
This predicted protein is rather small, and has a high proportion of polar (62%), and, specifically, basic (19%) and cysteine (8%) residues. There is no evidence for hydrophobic domains, or a signal peptide sequence.

The 5' untranslated region contains several pyrimidine-rich tracts which have also been observed in other genes, and which may have some role in regulation of expression[20,21]. Included in the 5' untranslated region is a potential TATA box[20] at position 85, 24bp upstream from the start of clone FR2.

Orientation of the candidate gene

FR.2 detects three genomic *Eco*RI fragments, but itself contains no *Eco*RI sites, indicating a minimum of three exons for the NDP candidate gene. The order of these three fragments with respect to FR.2 was determined by hybridisation of cDNA sub-fragments to genomic *Eco*RI digests (Figure 5.7). Patients 13545 and 10017, who are deleted for MAOB, also lack the 20kb *Eco*RI fragment detected by the 3' probe; thus the gene is oriented in the same direction as MAOB: 3' towards the telomere.

Pulsed-field analysis of YL1.28 shows that the smallest restriction fragment which completely contains the gene is a 50kb *Sfi*I fragment (data not shown), itself contained within the 80kb *Mlu*I fragment (Figure 5.2b).



Patient				
13545	-	-	+	+
10017	-	-	+	+
6446	+	R	+	+
3883	+	+	-	+
12316	+	+	-	+
8810	+	+	+	A

Figure 5.7. Schematic diagram of the structure and organization of the NDP candidate gene. Arrows show the order of genomic *Eco*RI fragments with respect to cDNA sub-fragments. The distances between the different *Eco*RI fragments are not known. *Eco*RI hybridization patterns of selected patients are illustrated below. P, *Pst*I; B, *Bam*HI; E, *Eco*RI. -, signal absent; +, signal present; R, signal with reduced intensity; A, altered fragment.

DISCUSSION

We have isolated a candidate Norrie disease gene using a positional cloning strategy. The use of the 160kb YAC sub-fragment to screen retinal cDNA libraries provided a sensitive and effective means of identifying transcripts from this region of the X chromosome.

The strongest evidence that one group of these transcripts derives from the NDP locus comes from a consideration of the aberrant hybridisation patterns in some patients (Figures 5.3 and 5.7). Patients 13545 and 10017, who are deleted for MAOB, also lack the 20kb *Eco*RI fragment detected by the 3' probe. Patients 3883 and 12316 show a deletion of the internal 9/10kb *Eco*RI fragment. In addition, an altered 5' 5kb fragment is seen in patient 8810; thus, alterations or deletions are observed for all three *Eco*RI fragments. Since the locus has a maximum size of 50kb, little space exists for a further unidentified gene. Moreover, there is no region within the locus which overlaps all detected deletions or alterations, indicating it is very unlikely that the Norrie gene is located within a single intron of the isolated gene. These findings argue strongly that the isolated cDNA is indeed the NDP gene. The observation of frequent deletions and an informative RFLP at the locus will provide the basis for improved prenatal diagnosis for Norrie disease.

The 5' cDNA sub-fragment detects two genomic *Eco*RI fragments (Figure 5.7), but contains no *Eco*RI sites, implying the occurrence of a splicing event. Since this 220bp region does not include any part of the proposed open reading frame, splicing must have taken place in the 5'-untranslated sequence. This phenomenon has been observed previously, for example in the mouse a-amylase gene, where differential splicing in this region confers different tissue specificities[23].

Because of the lack of homology between the predicted product of the candidate gene and sequences in the database, few conclusions can be drawn about its function. Its small size, high proportion of polar residues, and lack of a signal peptide sequence together indicate that it is likely to be a soluble protein. Tissue *in situ* hybridisation will elucidate its pattern and timing of expression, and the cloning and analysis of the mouse homologue may allow the construction of transgenic mouse models for Norrie disease. A study of the homologous *Drosophila* sequence may also be illuminating.

Since the Xp11.2-Xp11.3 region is very rich for eye disease loci (such as RP2, congenital stationary night blindness and Aland island eye disease)[10], further analysis of the NDP gene and

other cDNAs isolated from adjacent regions will help us to identify those eye disease loci. Ultimately, knowledge of the function of the NDP gene product will help in the understanding of the phenomenology of the disorder, and the role of the protein in normal neurodevelopment.

METHODOLOGY

General procedures

Conventional gel electrophoresis was in 0.8% (w/v) agarose (Sigma Type II-A) / 1 x TBE gels at 1.5V/cm. Gels were transferred to Hybond-N+ (Amersham) by alkaline blotting, and filters neutralised in 0.2M Tris-Cl (pH 7.5), 2 x SSC. Plasmids were prepared by standard methods[24], and inserts purified by the Geneclean (Bio 101) procedure, or labelled directly in Seaplaque (FMC Bioproducts) agarose. Labelling was by random priming with [a-^{32}P]-dCTP[25], and probes were pre-reassociated where necessary, as described[26]. Hybridisation was at 65°C in the buffer of Church and Gilbert[27], and washing to a final stringency of 100mM sodium phosphate, pH 7.2 / 0.1% SDS, at the same temperature. Filters were exposed to Kodak XAR film with Kodak intensifying screens at -70°C, and stripped by washing for 30 minutes in 2mM Tris-Cl, pH7.5 / 1mM EDTA / 1% SDS at 65°C.

Preparation and analysis of YAC DNA

Agarose plugs containing DNA from the YAC YL1.28, isolated from the 48,XXXX library of Larin et al[18], were prepared by standard methods[28]. YAC DNA digests were separated by pulsed field gel electrophoresis in a "Waltzer" apparatus [29], using the following parameters: 1.5% (w/v) agarose (Sigma Type I) / 0.5 x TAE; 16°C; 3.6V/cm; 45" switch time; 30h run time. Gels were exposed to UV light for 4 minutes prior to transfer.

Purification of YAC sub-fragment

Eighteen agarose plugs were digested to completion with *Sst*II and *Bss*HII as described elsewhere[30]. The 160kb *Bss*HII-*Sst*II fragment was separated by PFGE in a 1.5% (w/v) agarose (Sigma Type I), 16°C; 3.6V/cm; 13" switch time; 24h run time. The gel was visualised under UV, the fragment was cut out and purified by the Geneclean procedure. One half of the total yield of DNA (approximately 100ng) was used for labelling.

Screening of cDNA libraries

cDNA libraries were from fetal (Stratagene) and adult [31] retinal sources.

4×10^5 recombinants from each library were plated on two 22 x 22cm plates, and two replica plaque lifts on Hybond-N+ prepared according to manufacturer's instructions (Amersham). YAC sub-fragment DNA was labelled for six hours at 37°C using 160µCi [a-^{32}P]-dCTP. Pre-reassociation was carried out in a final volume of 350µl 5 x SSC for 3 hours using 10µg/µl denatured sonicated human placental DNA at 65°C; this corresponds to a C_0t of 180. Filters were prehybridised overnight in Church buffer[27] in the presence of 100µg/ml sonicated human placental DNA. Hybridisation was carried out overnight in fresh buffer lacking placental DNA. Washing was as described above, to a stringency of 100mM phosphate.

Phage DNA isolation was by standard methods[24], and *in vivo* excision of pBluescript SK- plasmids from the fetal retinal library according to manufacturer's instructions (Stratagene).

DNA sequencing

Clones were sequenced on both strands by the dideoxy method[32] with Sequenase Version 2.0 (USB), using universal and reverse primers, as well as primers derived from the obtained sequence. dITP was used to resolve compressions (USB).

Northern analysis

Total RNA was isolated from rat fetal (E19) and adult brain and liver, using standard methods[24] and poly-A-enriched mRNA prepared. Samples (2µg), together with RNA markers (Gibco BRL), were separated on formaldehyde denaturing agarose gels (1%), and alkali blotted onto Hybond-N+ according to manufacturer's instructions (Amersham).

Norrie patients

Patients with microdeletions 13545, Manchester, LGL515 and NHF88 have been described before[3,5,6,17]. Microdeletion patient 10017 is a 36 year old affected male with an affected brother. Both have moderate progressive hearing loss, but no mental retardation or any other neurological abnormalities. The panel of twenty typical Norrie patients was collected at the Massachusetts General Hospital. They were unrelated males, and diagnosed on the basis of congential blindness with typical pseudoglioma. Some have additional hearing loss or mental retardation. Several families showed

multiple affected males with characteristic X-linked inheritance patterns.

Cell-lines, genomic DNA samples, and probes

Genomic DNA was prepared by standard methods from the lymphoblastoid cell-line GM1416B (Human Genetic Mutant Repository, Camden, New Jersey) and a cell-line established from the patient of Donnai et al.[5] and from the mouse-human X-only hybrids ThyB-X[33] and MOG 13.9[34]. DNA from the lung of a male 3H1 mouse (MRC Radiobiology Unit, Harwell) was a gift from Yvonne Boyd, and DNA from the mouse-hamster mouse-X-only hybrid CAK was a gift from Phil Avner.

Zoo-blot DNAs were from the Genetics Laboratory collection.

The MAOB cDNA, BSMAOB, was a gift from Dr Julie Andersen, MGH. Human β–actin was obtained from Clontech.

REFERENCES

1. Norrie, G. (1927) Causes of blindness in children. *Acta Ophthalmol.* **5**, 357-386.
2. Warburg, M. (1966) Norrie's disease: a congenital progressive oculo-acoustico-cerebral degeneration. *Acta Ophthalmol.* (Suppl.) **89**, 1-47.
3. de la Chapelle, A., Sankila, E.-M., Lindof, M., Aula, P. & Norio, R. (1985) Norrie disease caused by a gene deletion allowing carrier detection and prenatal diagnosis. *Clin. Genet.* **28**, 317-320.
4. Gal, A., Wieringa, B., Smeets, D.F., Bleeker-Wagemakers, L.M., & Ropers, H.H. (1986) Submicroscopic interstitial deletion of the X chromosome explains a complex genetic syndrome dominated by Norrie disease. *Cytogenet. Cell Genet.* **40**, 219-224.
5. Donnai, D., Mountford, R.C., & Read, A.P. (1988) Norrie disease resulting from a gene deletion: clinical features and DNA studies. *J. Med. Genet.* **25**, 73-78.
6. Zhu, D., Antonarakis, S.E., Schmeckpeper, B.J., Diergaarde, P.J., Greb, A.E., & Maumenee, I.H. (1989) Microdeletion in the X chromosome and prenatal diagnosis in a family with Norrie disease. *Am. J. Med. Genet.* **33**, 485-488.
7. Sims, K.B., de la Chapelle, A., Norio, R., Sankila, E.-M., Hsu, Y.-P.P., Rinehart, W.B., Corey, T.J., Ozelius, L., Powell, J.F., Bruns, G., Gusella, J.F., Murphy, D.L., & Breakefield, X.O. (1989) Monoamine oxidase deficiency in males with an X chromosome deletion. *Neuron* **2**, 1069-1076.
8. Lan, N.C., Heinzmann, C., Gal, A., Klisak, I., Orth, U., Lai, E., Grimsby, J., Sparkes, R.S., Mohandas, T. & Shih, J.C. (1989) Human monoamine oxidase A and B genes map to Xp11.23 and are deleted in a patient with Norrie disease. *Genomics* **4**, 552-559.
9. Sims, K.B., Ozelius, I., Corey, T., Rinehart, W.B., Liberfarb, R., Haines, J., Chen, W.J., Norio, R., Sankila, E., de la Chapelle, A., Murphy, D.L., Gusella, J., & Breakefield, X.O. (1989) The Norrie disease gene is distinct from the monoamine oxidase genes. *Am. J. Hum. Genet.* **45**, 424-434.
10. Davies, K.E., Mandel, J.-L., Monaco, A.P., Nussbaum, R.L., & Willard, H.F. (1991) Report of the committee on the genetic constitution of the X chromosome. *Cytogenet. Cell Genet.* **58**, 853-66.
11. Katayama, S., Wohlferd, M., & Golbus, M.S. (1988) First demonstration of recombination between the gene for Norrie disease and the probe L1.28. *Am.J. Med. Genet.* **30**, 967-970.

12. Ngo, J.T., Spence, A., Cortessis, V., Sparkes, R.S., & Bateman, J.B. (1988) Recombinational event between Norrie disease and DXS7 loci. *Clin. Genet.* **34**, 43-47.

13. Lindsay, S., Thiselton, D.L., Bateman, J.B., Ngo, J.T., Sparkes, R.S., Coleman, M., Davies, K.E., & Bhattacharya, S.S. (1992) Localisation of the gene for Norrie disease to between DXS7 and DXS426 on Xp. *Hum. Genet.* **88**, 349-350.

14. Chen, Z.-Y., Meindl, A., de la Chapelle, A., Sankila, E.-M., Monaco, A., Powell, J.F., Craig, I.W., & Meitinger, T. (1991) Characterisation of a 650kb yeast artificial chromosome (YAC) containing DXS7, MAOA and MAOB loci - implications for the isolation of the Norrie disease locus. *Cytogenet. Cell Genet.* **58**, 2060.

15. Chen, Z.-Y., Hsu, P. Y.-P., Powell, J.F., Hendriks, R.W., Breakefield, X.O., & Craig, I.W. (1991) Long range physical mapping and organisation of the region including the genes for monoamine oxidase A and B (MAOA & MAOB).*Cytogenet. Cell Genet.* **58**, 2059.

16. Chen, Z-Y., Sims, K.B., Coleman, M., Donnai, D., Monaco, A., Breakefield, X.O., Davies, K.E., and Craig, I.W. (1992) Characterisation of a YAC containing part or all of the Norrie disease locus. *Hum. Molec. Genet.* in press

17. Sims, K.B., Lebo, R.V., Benson, G., Shalish, C., Chen, Z.-Y., Bruns, G., Golbus, M.S., Craig, I.W., & Breakefield, X.O (1992). The Norrie disease gene maps to a 150kb region on chromosome Xp11.3. *Hum. Molec. Genet.* in press.

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

19. Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8148

20. Breathnach, R., & Chambon, P. (1981) Organization and expression of eukaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**, 349-383.

21. Dudov, K.P., & Perry, R.P. (1984) The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. *Cell* **37**, 457-468.

22. Fickett, J.W. (1982) Recognition of protein coding regions in DNA sequences. *Nucleic Acids Res.* **10**, 5311-5318

23. Leff, S.E., Rosenfeld, M.G., & Evans, R.M. (1986) Complex transcriptional units: diversity in gene expression by alternative RNA processing. *Ann. Rev. Biochem.* **55**, 1092-1117.

24. Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) *Molecular Cloning: a laboratory manual.* 2nd edition (Cold Spring Harbor Laboratory Press, New York).

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

26. Sealey, P.G., Whittaker, P.A., & Southern, E.M. (1985) Removal of repeated sequencesfrom hybridisation probes. *Nucleic Acids Res.* **13**, 1905-1922.

27. Church, G.M., & Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991-1995.

28. Anand, R., Riley, J.H., Butler, R., Smith, J.C., & Markham, A.F. (1990) A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *Nucleic Acids Res.* **18**, 1951-1956.

29. Southern, E.M., Anand, R., Brown, W.R.A., & Fletcher, D.S. (1987) A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* **15**, 5925-5943.

30. Ross, M.T., Ballabio, A., & Craig, I.W. (1990) Long-range physical mapping around the human steroid sulfatase locus. *Genomics* **6**, 528-539.

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

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

33. Lund, E., Bostock, C., Robertson, M., Christie, S., Mitchen, J.L., & Dahlberg, J.E. (1983) U1 small nuclear RNA genes are located on human chromosome 1 and are expressed in mouse-human hybrid cells. *Mol. Cell. Biol.* **3**, 2211-2220.

34. Povey, S., Jeremiah, S.J., Barker, R.F., Hopkinson, D.A., Robson, E.B., Cook, P.J.L., Solomon, E., Bobrow, M., Carritt, B., & Buckton, K.E. (1980) Assignment of the human locus determining phosphoglycolate phosphatase (PGP) to chromosome 16. *Ann. Hum. Genet.* **43**, 241-248.

Acknowledgements

We are grateful to the patients and their clinicians who provided valuable material and in particular to A. de la Chapelle, D. Donnai, M. Golbus, H. Jackson, R. Lebo, and H. Middleton-Price. We especially thank Antony Monaco (Institute of Molecular Medicine, Oxford) for providing us with the YL1.28 YAC clone. We also thank Mick Coleman and Kay Davies (Institute of Molecular Medicine, Oxford); Anil Day, Chris Porter, Ruth Brown, Janet Mellor and Val Cooper (Genetics Laboratory, Oxford); Deborah Schuback and Alan Buckler (MGH); Michel de Weers (University Hospital, Leiden) who provided advice, assistance, or material at various stages of the project. The work has been supported by the British Retinitis Pigmentosa Society (IWC), Niels Stensen Stichting (RWH), EPA Cephalosporin Fund (MAJ), The Wellcome Trust (IWC), and NIH grants: RCDA Award (HD 00824) to KBS; Senator Jacob Javits Award (NS21921) to XOB.

Chapter 6. STRUCTURAL CHARACTERIZATION OF THE NORRIE DISEASE GENE

Z.-Y. Chen, E.M. Battinelli, R.W. Hendriks, J.F. Powell, X.O. Breakefield, K.B. Sims and I.W. Craig

SUMMARY

Positional cloning experiments recently resulted in the isolation of a candidate gene for Norrie disease (Pseudoglioma, NDP), a severe X-linked neuro-developmental disorder. Here we report the isolation and analysis of human genomic DNA clones encompassing the NDP gene. The gene spans 28 kb and consists of 3 exons, the first of which is entirely contained within the 5' untranslated region. Detailed analysis of genomic deletions in Norrie patients show that they are heterogenous, both in size and in position. By PCR analysis, we found that expression of the NDP gene was not confined to the eye or brain. An extensive DNA and protein sequence comparison between the human NDP gene and related genes from the database, revealed homology with cysteine-rich protein-binding domains of immediate-early genes implicated in the regulation of cell proliferation. We propose that NDP is a molecule related in function to these genes and may be involved in a pathway that regulates neural cell differentiation and proliferation.

INTRODUCTION

Norrie disease (NDP, McKusick 31060) is a severe, X-linked, recessive neurodevelopmental disorder of unknown pathogenesis (1,2). The disease is characterized by dysplasia of the retina and consequent retinal detachment leading to congenital blindness. The presence of poorly differentiated and malformed retinas in the patients, suggests an early failure in retinal development involving sensory cells, ganglion cells and the inner neuroblastic layer. In typical affected males congenital blindness is accompanied by the growth of pseudoglioma in the eye socket, often associated with mental retardation and progressive sensorineural deafness. The expression of the NDP phenotype in males is heterogeneous with intrafamilial variability of cognitive function and in the onset or severity of hearing loss.

Deletion mapping and recombination studies have led to a refined localization of the NDP locus on the short arm of the X chromosome at Xp11.2-Xp11.3 (10, 11). NDP was found to be closely linked to the DXS7 locus, detected by the L1.28 probe, and to the genes for the monoamine oxidases (MAO) A and B (3-9). Recombination between a GT-repeat polymorphism within the MAOB gene and the NDP locus, delineated this locus a telomeric flanking marker (10). Long range genomic restriction mapping and the analysis of yeast artificial chromosome (YAC) clones further refined the location of the NDP gene to an interval of approximately 160 kb, immediately centromeric to the MOAB gene (11,12). By screening of cDNA libraries from human retina with cosmid clones or YAC fragments within this region, a candidate gene for Norrie disease was isolated (13, 14). The gene was found to be partly deleted in several typical Norrie patients. Expression of the gene was detected in fetal and adult brain and retina. DNA sequence analysis of the isolated cDNA clones revealed an open reading frame with a predicted protein of approximately 15kD.

For a structural analysis of the NDP locus, genomic DNA clones comprising the entire gene were isolated. The genomic map obtained from this study has enabled a detailed investigation of the genomic deletions present in Norrie patients. To gain insight into the function of the NDP gene, we have performed a further protein sequence comparison

RESULTS

Genomic organization of the NDP gene

To investigate the structural organization of the NDP gene, genomic clones spanning the gene were isolated from a lambda phage library, constructed from partial MboI digests of the YAC clone YL1.28 (11).

The entire NDP gene was isolated in three overlapping phage clones (128.ND.4, 128.ND.8 and 128.ND.11; Figure 6.1). Exon sequences were identified by restriction mapping and Southern blot hybridization of phage DNA digests, using the NDP cDNA probe FR.2 (13) and different oligonucleotides derived from FR.2. The precise locations of each of the 5' and 3' exon-intron boundaries were subsequently determined by sequencing the appropriate regions of the cloned genomic DNA. The analysis established that the NDP gene is 28 kb long and divided into 3 exons of 201 bp, 380 bp and 1257 bp in size, separated by introns of approximately 16 and 9 kb, respectively (Figure 6.1).

The sequences at the 5' and 3' boundaries of each intron (Figure 6.2) are in agreement with the consensus sequence for exon-intron boundaries of other eukaryotic genes (15). The 5' untranslated region of FR.7, our longest NDP specific cDNA clone, extends 408 bp upstream of the translation initiator methionine codon (13). In genomic DNA, this region is interrupted by the first intron, with the first exon being entirely untranslated. A potential TATA box (15) appears 218 bp from the 5' end of the FR.7 cDNA sequence (Fig. 6.2). A CCAAT sequence was found at position of 246 bp located in the opposite orientation. The 200 bp genomic sequences immediately 5' to the start of the FR.7 cDNA is relatively C+G rich (55%) with 15 CpG and 22 GpC dinucleotides, but does not contain Sp-1 binding sites. Within the first intron segment, a canonical TATAAAT sequence was identified 185 bp 5' from the start of the second exon.

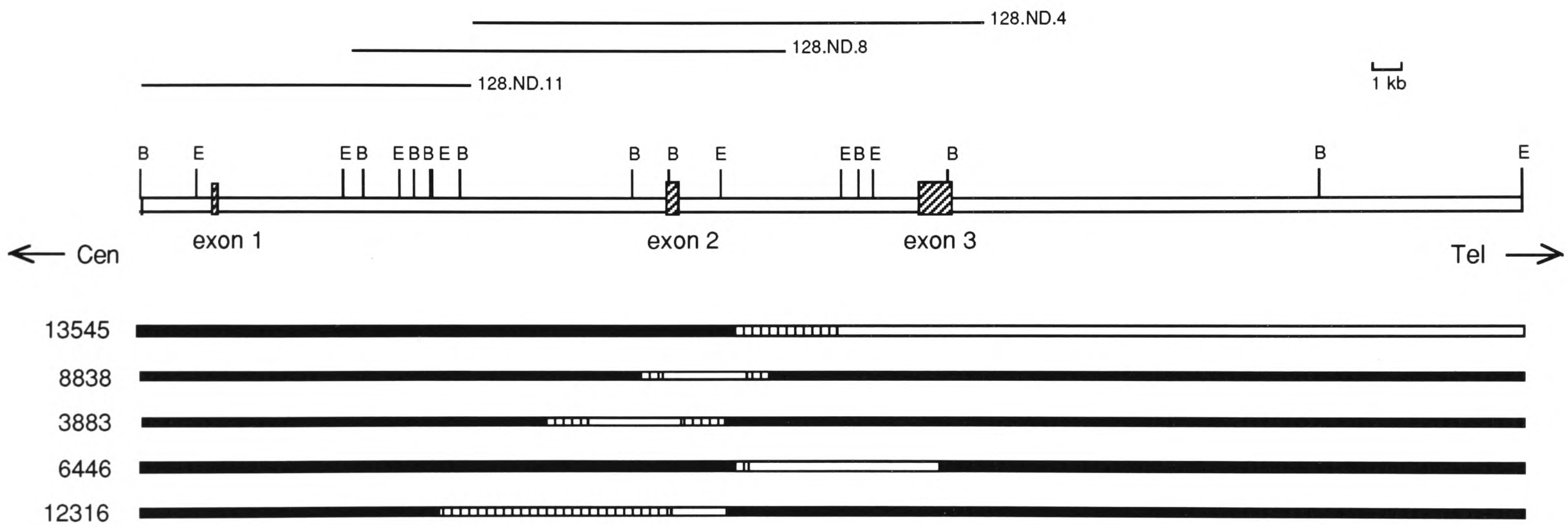


Figure 6.1 Genomic restriction map of the NDP gene and the locations of the three exons. The overlapping lambda phage clones containing the NDP gene are given above the map. Below the map the microdeletions characterized in the patients are indicated (open bars: regions deleted; Filled bars: regions present; hatched bars: the extent of deletion unknown). The orientation of the gene is also gven. E=*Eco*RI; B=*Bam*HI.

AAATTCAGAT ATTCATTGGC CTCTTATTAG TTCCATAATA CCATTAAAAA
AGAAAGAAAG AAAGAAACTT CCTCGCCCTT GTTCTCGTAC GCTGTTCCCA
TCGTAAGATG CTCCGTGGAA GGGAGCCGAG CGGTGGGCAG AGGCTGAGTC
CCCGATAACG AGCGCCTCAC ATTTCCGTGG CATTCCCATT TGCTAGTGCG
CTGCTGCGGC CGCACGCCTG ATTGGATATA TGACTGCAAT GGCACTTTTC

CATTTGACAT TCTC-----exon 1 (201bp)-----TAAG GTAAGCAGGA

AAACAAGCGC TGCCGGCTTG ---16kb--- ACCCAATTCG GTTACGTTGT
TGCCAGAACA ACATGTTTAA TCTTTAACAT GGGTTCAAAC TATTCTTGGC
CCTAGGAACA TGGACTTCAG CAATTAAAGT CAACATGTGC TTCCATTAAA
CCATTGTGTC CACCTCCAAA TGGTTATAAA TATGTAGCAT AAGCTATGGG
AGTTGGGGTG GAATGGATGA CAGCCTTTGC TAATGACGCT CTAGAAACCA
ATATTCTCCT CTCAAAATAA CATGGAAAAA TTCTACTTAA TACCTCCTGG
GTTCCATTAG TGGTTCTGGG TAAATAATTC TGGGGAAAGT AATTTCTGTT

TTCATTCCAG CTGT-----exon 2 (380bp)-----AAAG GTAAGACCAA

GGTCTCTGTG AGGAGAGCAT ----8kb--- CCCACAGATA ACCTCCTTTT

CTTTCCCCAG ATGG-----exon 3 (1257bp)------TC AAATTTGACT

GTTAGTATTT

Figure 6.2. Genomic structure of the NDP gene. The nucleotide sequences of the regions flanking the three exons (over and underlined) is shown. Potential TATA-boxes and the CCAAT sequence are underlined.

Characterization of deletions in Norrie patients

In a previous study (13), we reported aberrant cDNA hybridization patterns in several typical Norrie patients. The obtained genomic map of the NDP gene enabled a more detailed investigation of these genomic deletions present in five Norrie patients. Using subcloned fragments from the 128.ND.4, 128.ND.8 and 128.ND.11 phage clones as probes in Southern blotting experiments, the position of the breakpoints and the sizes of the deletions were determined (Figure 6.1, lower half). Three patients showed deletions encompassing the second exon (8838, 3883 and 12316) , which were 4 kb, 5 kb and 2-10 kb in size, respectively. Patient 13545 manifested a larger deletion, extending downstream from the second intron of the NDP gene and also encompassing the entire MAOB gene (13). In patient 6446 part of the third exon was absent due to a deletion of approximately 7 kb.

Using subcloned fragments of the three genomic clones in Southern blotting experiments, we could not detect any additional deletions within the two intron segments in the panel of twenty patients that was previously analyzed using the cDNA probe FR.2 (13).

Expression of the NDP gene

To analyze the expression profile of the NDP gene, PCR primers were designed for specific amplification of cDNA synthesized from various tissues. Using NDP38, located within the first exon, and NDP72, located within the 3' untranslated region, the expected product of 722 bp was detected in all samples of fetal eye, and fetal and adult brain (Table 6.1). Expression was also found in fetal lung and adult muscle. In these experiments, we could not detect transcripts in fetal or adult liver and cultured liver cells.

In northern analysis of mRNA derived from human glioma (IN358; IN157), glioblastoma (A172) and neural (SK.N.SH) cells strong hybridization was observed to transcripts in the 1.8-2 kb size range. In addition, several transcripts of lower density and with larger sizes (3 kb and 6 .kb) were detected.

TABLE 6.1. Expression of the NDP gene

Tissues [1]	NDP expression
Fetal tissues:	
Eye (n=3)[1]	+
Brain (n=5)[1]	+
Lung	+
Liver	-
Amnionic and chorionic membrane	+
Adult tissues:	
Brain (n=3)[1]	+
Glioma cell lines IN358/IN157[2]	+
Glioblastoma A172[2]	+
Neural cell line SK.N.SH[2]	+
Liver	-
Hepatoma cell line HEPG2	-
Fibroblasts	+
Muscle	+

[1]) In cases where more than one sample was analyzed, the sample sizes are given in parentheses.

[2]) These samples were investigated by Northern blotting, all other tissues by PCR analysis.

Protein and DNA sequence comparison

An extensive DNA and protein sequence comparison between the human NDP gene and related genes from the NBRF and Swiss protein database, revealed a significant homology with C-terminal cysteine-rich domains shared by a family of mucin-like proteins and von Willebrand factor (16-17) (Fig. 6.3). The same region also exhibited homology with two immediate-early genes: the phorbol ester-repressive and v-src -inducible CEF-10 gene, identified in chicken embryo fibroblasts (18) and the growth factor- and phorbol ester-inducible Cyr61 gene, identified in murine fibroblasts (19).

```
VWF-human       2773    CCSPTRTEPMQVALHCTNGSVVYHEVLNAMECKCS  2807
                        ** *          * *  *              * *
NDP-human         95    CCRPQTSKLKALRLRCSGGMRLTATYRYILSCHCE   129
                        ** **        * **  *   *     * ** *
CEF10-chick      316    CCTPQQTRTVKIRFRCDDGETFTKSVMMIQSCRCN   351
                        ** *         * **  *         * ** *
CYR6-mouse       320    CCTPLQTRTVKMRFRCEDGEMFSKNVMMIQSCKCN   354
```

Figure 6.3 Protein sequence comparison between the C-terminal cysteine-rich domains of the human genes for von Willebrand factor and NDP, as well as the chicken CEF-10 and the murine Cyr61 gene. Identical amino acid residues are indicated.

DISCUSSION

We have isolated and analyzed human genomic DNA clones encompassing the NDP gene, which span 28 kb and contain 3 identifiable exons.

The finding that the first exon is entirely contained within the 5' untranslated region of the gene, confirmed our previous evidence for a putative splicing event within the 5' untranslated sequence, based on Southern blot hybridizations of *Eco*RI genomic digests with a 5' cDNA sub-fragment (13). There are two potential TATA box sequences located 36 bp and 214 bp, respectively, 5' to the beginning of the longest cDNA FR.7. None of them fits the criteria for the TATA box strictly (15). However the combination of the positions for the GGTTA (CCAAT box in the reversed orientation which can be functional. 20) and the ATTAAAA makes the latter more likely be the site for the TATA binding protein. Splicing events outside the open reading frame have been observed previously, for example in the mouse alpha-amylase and the human G6PD genes (21,22). For the NDP gene, differential splicing may confer its tissue specific or developmental stage-dependent expression. Its structural organization, in particular the presence of a consensus TATAAAT sequence just 5' of the second exon, would allow an alternative start of transcription. Obviously, the expression of the NDP gene could then be regulated by the activity of two independent promoter regions. Further study is needed to confirm the pattern and regulation of alternate splicing events at the 5' end of the NDP gene.

The investigation of NDP expression in different tissues showed that expression is not confined to the eye or brain.; however, it is unclear from the clinical phenotype of the Norrie patients what the possible function of the gene in tissues like lungs or muscle might be.

The protein sequence comparisons reveal a significant homology with the C-terminal cysteine-rich domains shared by a family of mucin-like proteins and the von Willebrand factor. From the known function of these proteins and from the pattern of homology, no obvious conclusions can be drawn for a possible function of the NDP gene in neurodevelopment. The same region also exhibits homology with two immediate-early genes, the chicken CEF-10 gene, and the murine Cyr61 gene. In addition to the observed protein sequence homology, some

structural features of the NDP gene show additional resemblance to these immediate-early genes and to other genes implied in the regulation of cell proliferation, like insulin-like growth factor-binding proteins (IBP) (23).

Firstly, they are proteins lacking N-linked glycosylation but have a secretory signal peptide. The N-terminus of the NDP gene contains the three structurally distinct regions of a signal sequence: a basic region (amino acids 1-4), a central hydrophobic region (amino acids 5-19) and a more polar region (amino acids 21-30). A potential cleavage site is predicted between serine-24 and lysine-25 by Von Heijne's weighted matrix method (24). Secondly they contain large numbers of cysteine residues that are not clustered, but are absent from the central part of the residue NDP molecule, from amino acid residues 70 to 92. Cysteine-rich regions are frequently found in the binding portions of proteins and are implicated in confering ligand specificity (25-27).

Thirdly, their 3' untranslated regions contain sequences that have been proposed to destabilize mRNA molecules and which are also frequently found in other transiently expressed genes, such as lymphokines, cytokines, and proto-oncogenes. The NDP 3' untranslated region is relative AT rich (60%) and contains 10 copies of the common motif of these sequences, a single adenosine, followed by three or more thymidines (28).

In contrast, there are several features of the NDP gene which are not shared by the immediate-early gene class. For example, the N-terminal region of the NDP region is less rich in cysteine (only two in 60 residues), and the central part of the molecule does not contain a highly acidic region, as found in CEF-10 and Cyr61. Also, regarding the differences in tissue-specificity, it is very unlikely that the NDP gene is the human homologue of these genes. Therefore we propose that NDP is a molecule related in function, and which may be involved in a pathway that regulates neural cell differentiation and proliferation.

MATERIALS AND METHODS

General procedures

Genomic DNA was prepared by standard methods from peripheral blood samples obtained from a panel of patients collected at the Massachusetts General Hospital. Agarose gel electrophoresis, northern and Southern blotting procedures were as described previously (13). Labelling of the cDNA clone FR.2 and various restriction fragments from phage clones (128.ND) was by random priming with [a-32P]-dCTP (29), and probes were prereassociated where necessary, as described (30). Oliogonucleotides were end-labelled with T4 polynucleotide kinase (New England Biolabs, Beverley, MA) using [gamma-32P]-ATP (31).

Construction of a lambda phage clone contig

The lambda phage library constructed from a partial MboI digest of DNA from the L1.28 YAC (32) has been described before (11). Phage clones containing exon sequences of the NDP gene were isolated by screening the library using the FR.2 cDNA probe (13). For restriction mapping single and double digests with *Eco*RI and *Bam*HI were resolved on agarose gels and, after Southern blotting, were probed with the FR.2 cDNA probe, various 17-mer oligonucleotides derived from FR.2, as well as EcoRI or BamHI restriction fragments from the phage clones.

Identification of intron/exon boundaries

DNA from BamHI digests of the phage clones 128.ND.4, 128.ND.8 and 128.ND.11 was cloned into Bluescript KS (Stratagene, La Jolla, CA) pUC9 vectors. Double strand sequencing was performed by the chain-termination method (32) using the sequenase version 2.0 kit (United States Biochemical Co., Cleveland, OH) following the manufacturer's instructions. The sequence primers used were: M13 universal and M13 reversal primers, as well as 17-mer oligonucleotides

designed from the obtained DNA sequence. For exon 1: NDP76 (5'-ATGGCTTGCCTTTTATA-3') and NDP 35 (5'-TATAAAAGGCAAGCCAT-3'), for exon 2: NDP31 (5'-CTGTTCTTCTAGAGAAG-3') and for exon 3: NDP73 (5'-AAAGGATGCAACAATTC-3').

Expression analysis by PCR amplification of cDNA

Double-stranded cDNA was prepared from RNA by using cDNA Synthesis system (31). Specific primers, NDP38 (5'-GAAAAGTCTGAGAAGAAA-3') and NDP72 (5'-TTG-CCATTGCTGCTAAA-3'), in the first exon and the 3' untranslated region of the NDP gene, respectively, were designed and used in PCR amplification (34). The synthesized cDNA samples were amplified for 30 cycles of 0.5 min at 94^{o}C, 0.5 min at 44^{o}C, 0.5 min at 75^{o}C, with a final extension of 10 min at 75^{o}C in 50 µl of reaction mixture containing 2-4 U Taq DNA polymerase (Boehringer Mannheim, Germany), 0.2 mM of each dNTP, 1µM of each primer, using the buffers supplied by the manufacturer.

Acknowledgements

We would like to acknowledge the patients and their clinicians for providing valuable material. We thank D. Blake, M. Moscoso and Val Cooper for advice and assistance at various stages of the project. The work was partly supported by the British Retinitis Pigmentosa Society (I.W.C), Niels Stensen Stichting (R.W.H.), The Welcome Trust (I.W.C.), and NIH grants: RCDA Award (K.B.S.) and a Senator Jacob Javits Award (X.O.B.).

References

1. Norrie, G. (1927) *Acta Ophthalmol.* **5,** 357-386.

2. Warburg, M. (1966) *Acta Ophthalmol.* (Suppl.) **89,** 1-47.

3. de la Chapelle, A., Sankila, E.-M., Lindof, M., Aula, P. & Norio, R. (1985) *Clin. Genet.* **28**, 317-320.

4. Donnai, D., Mountford, R.C., & Read, A.P. (1988) *J. Med. Genet.* **25**, 73-78.

5. Zhu, D., Antonarakis, S.E., Schmeckpeper, B.J., Diergaarde, P.J., Greb, A.E., & Maumenee, I.H. (1989) *Am. J. Med. Genet.* **33**, 485-488.

6. Sims, K.B., de la Chapelle, A., Norio, R., Sankila, E.-M., Hsu, Y.-P.P., Rinehart, W.B., Corey, T.J., Ozelius, L., Powell, J.F., Bruns, G., Gusella, J.F., Murphy, D.L., & Breakefield, X.O. (1989) *Neuron* **2**, 1069-1076.

7. Katayama, S., Wohlferd, M., & Golbus, M.S. (1988) *Am.J. Med. Genet.* **30**, 967-970.

8. Sims, K.B., Ozelius, I., Corey, T., Rinehart, W.B., Liberfarb, R., Haines, J., Chen, W.J., Norio, R., Sankila, E., de la Chapelle, A., Murphy, D.L., Gusella, J., & Breakefield, X.O. (1989) *Am. J. Hum. Genet.* **45**, 424-434.

9. Lindsay, S., Thiselton, D.L., Bateman, J.B., Ngo, J.T., Sparkes, R.S., Coleman, M., Davies, K.E., & Bhattacharya, S.S. (1992) *Hum. Genet.* **88**, 349-350.

10. Sims, K.B., Lebo, R.V., Benson, G., Shalish, C., Chen, Z.-Y., Bruns, G., Golbus, M.S., Craig, I.W., & Breakefield, X.O. (1992) *Hum. Molec. Genet.* **1,** 83-89.

11. Chen, Z.-Y., Sims, K.B., Coleman, M., Donnai, D., Monaco, A., Breakefield, X.O., Davies, K.E. & Craig, I.W. (1992) *Hum. Molec. Genet.* **1**, 161-164.

12. Chen, Z.-Y., Powell, J.F., Hsu, Y.-P.P., Breakefield, X.O. & Craig, I.W. (1992). *Genomics.* **14**, 75-82

13. Chen, Z.-Y., Hendriks, R.W., Jobling, M.A., Powell, J.F., Breakefield, X.O., Sims, K.B. & Craig, I.W. (1992) *Nature Genet.* b, 204-208.

14. Berger, W., Meindl, A., Van de Pol, T.J.R., Cremers, F.P.M., Ropers, H.H., Doerner, C., Monaco, A., Bergen, A.A.B., Lebo, R., Warburg, M., Zergollern, L., Lorenz, B., Gal., A., Bleeker-Wagemakers, E.M. & Meitinger, T. (1992) *Nature Genet.* **1,** 199-203.

15. Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* **50**, 349-383.

16. Titani, K., Kumar, S., Takio, K., Ericsson, L.H., Wade, R.D., Ashida, K., Walsh, K.A., Chopek, M.W., Sadler, J.E., Fujikawa, K. (1986) *Biochemistry* **25**,3171-3184.

17.Probst, J.C., Gertzen, E.-M. and Hoffmann, W. (1990) Biochemistry **29**, 2640-6244.

18. Simmons, D., Levy, D.B., Yannoni, Y. and Erikson, R.L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1178-1182.

19. O'Brien, T., Yang, G.P., Saunders, L. and Lau, L.F. (1990) *Molec. Cell. Biol.* **10**, 3569-3577.

20. Hames, B.D. and Glover, D.M. (1988) Transcription and Splicing. IRL Press Limited, Oxford.

21. Leff, S.E., Rosenfeld, M.G., and Evans, R.M. (1986) Ann. Rev.Biochem.**55**, 1092-1117.

22. Martini, G., Toniolo, D., Vulliamy, T., Luzzatto, L., Dono, R., Viglietto, G., Paonessa, G., D'Urso, M., and Persico, M.G. (1986) *EMBO J.* **5,** 1849-1855.

23. Brinkman, A., Groffen, C., Kortleve, D.J., Geurts van Kessel, A. and Drop, S.L.S. (1988) *EMBO J.* **7**, 2417-2423.

24. Von Heijne, G. (1986) *Nucl. Acids Res.* **14,** 4683-3690.

25. Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yarden, Y., Libermann, T.A., Schlessinger, J., Downward, J., Mayes, E.L.V., Whittle, N., Waterfield, M.D. and Seeburg, P.H. (1984). *Nature* **309,**418-425.

26. Ullrich, A., Gray, A., Tam, A.W., Yang-Feng, T., Tsubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E., Jacobs, S., Franke, U., Ramachandran, J. and Fujita-Yamaguchi, Y. (1986).*EMBO J.* **5**,2503-2512.

27. Yamamoto, Davis, C.G., Brown, M.S., Schneider, W.J., Casey, M.L., Goldstein, J.L., and Russell, D.W. (1984) *Cell* **39**, 27-38.

28 Shaw, G. and Kamen, R. (1986) *Cell* **46,** 659-667.

29 Feinberg, A.P., & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13.

30. Sealey, P.G., Whittaker, P.A., & Southern, E.M. (1985) *Nucleic Acids Res.* **13**, 1905-1922.

31. Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) Molecular Cloning: a laboratory manual, 2nd edition (Cold Spring Harbor Laboratory Press, New York).

32. Larin, Z., Monaco, A.P., & Lehrach, H. (1991) *Proc. Natl. Acad. Sci. U.S A.* **88**, 4123-4127.

33 Sanger, F., Nicklen, S., & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S. A.* **74**, 5463-5467.

34 Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Ehrlich, H.A. (1988) *Science* **239**, 487-491.

CHAPTER 7: GENERAL CONCLUSIONS

At the start of the investigations monoamine oxidase A and B were regarded as enzymes of considerable pharmacological importance and whose biochemical properties had been extensively investigated. In addition, considerable progress towards the cloning of both genes had been made; however, although there had been many suggestions as to their possible involvement in a variety of disorders, no specific diseases had been unabiguously associated with mutations affecting these loci. In contrast, Norrie disease is a profound disorder affecting eye development, resulting in congenital blindness but the nature of the biochemical defect and the gene reponsible were undefined. Gene mapping placed both NDP and MAO loci in close proximity and the possibility that MAO genes may have been involved in the aetiology of NDP provided a focus for the detailed investigation of the region in the hope that it might provide both insights into the evolutionary relationship of the MAO genes and allow the development of genetic markers for diseases analysis.

Characterisation of the genomic structure of MAOA gene, described in chapter 2 has shown it to comprise 15 exons. Because the sequences containing the boundaries of the 4th exon were difficult to identify, an accurate estimate of the size of the gene was not posssible. An additional complicating factor was the relatively big intron sequence (>40kb) between exon 1 and exon 2. It was apparent that the approach based on phage libraries had limitations (at least two clones would be needed to fill in the intron gap). As YAC libraries were becoming increasingly available the next logical step was, therefore, to isolate and characterize YAC clones isolated with MAOA cDNA.

In Chapter 3 long-range mapping of genomic DNA around the MAO genes was undertaken as well as the characterisation of a YAC containing the majority of sequence from both loci. One serious concern working with YAC clones is the possibility that they may consist of two or more pieces of DNA which are from non contiguous regions of the genome. To verify the authenticity of a

YAC clone one has to compare the YAC restriction map with the genomic map (it may also be valuable to compare the map with that for other YAC clone(s) representing overlapping regions). The map established for the MAO-YAC clone showed that the A and B genes are in a tai-to-tail arrangement. The PFGE map of the region was based on rare cutting enzyme digests of male and female genomic DNA. Double digests were, however, generally carried out on male DNA. It is, therefore, possible that the methylation pattern of the male samples may have excluded the identification of sites which are detectable only on the inactive X in females. There are two CpG islands within the genomic region. The first containing a cluster of rare-cutting enzyme sites is found in the genomic DNA and in the YAC (position 0 - Figure 3.2). The other CpG island was not detected in genomic digests and, because the relevant region was missing from the YAC (first exon of MAOA gene) it remained undiscovered. That there is a CpG island associated with the MAOA gene came to light eventually from detailed sequence analysis of 5' end of the gene (Hendriks, 1991). It was found the sequence upstream the first exon of MAOA gene is very rich both in C+G content and CpGs, fitting criteria for a CpG island (Bird, 1987). This result indicated that it is possible to miss a CpG island by simply looking at the pattern generated by rare-cutting enzyme digests. More detailed study on cloned genomic sequence may be required and should be accompanied by sequence analysis whenever possible.

Extension of the investigation to flanking regions and confirmation of the results obtained with the MAO-YAC are presented in Chapter4. A futher YAC clone was made available (A. Monaco, Institute of Molecular Medicine, Oxford) which was isolated with another probe L1.28 (detecting the DXS7 locus which PFGE analysis indicated to be adjacent to the MAO loci). The long range maps derived from this YAC clone was found to be in general agreement with that examined in Chapter 3 and confirmed the physical relationship between the two MAO genes. The work described in this chapter also extended the PFGE map suggesting that the previous genomic mapping around the MAO loci was largely correct. The importance of undertaking a detailed characterisation of the genomic organisation, in this case of the MAO loci, is illustrated in Chapter 4 where it served as the basis for undertaking the isolation of the Norrie disease gene, NDP. The work described also provided one of the two key pieces of information allowing the location of NDP to a segment of the LI.28 YAC by establishing that subclones from one end of the YAC were present in the Manchester

deletion patient. The second critical observation was that a recombination event had occurred in a Norrie pedigree placing the NDP locus proximal (centromeric) to MAOB (Sims et al, 1992).

Work reported in chapter 5 may have made it appear that the isolation of a disease gene is straightforward once its location is established. In practice, progress was complicated by the isolation of three loci for candidate cDNAs in the obligatory region of the Norrie locus. The locus for one group of cDNAs (clones FR12 and 14) appeared on the basis of its presence and absence in some deletion patients to be more a likely candidate than that encoding the group of cDNAs (FR2 and 7) which was eventually shown to be the NDP gene. However, the failure to detect any homologue corresponding to FR12/14 on the mouse X chromosome indicated that the corresponding locus is very unlikely to be the locus involved in the disease. One obvious oversight in the work reported in this chapter was the failure to detect the existence of a signal peptide sequence in the predicted Norrie gene protein. The main reason for this was that the work was carried out in an extremely short period and under considerable pressure to publish rapidly the results. It was very difficult to complete a thorough analysis of the protein sequence in the time available. The same reason also contributed to the fact that extensive examination of protein homology was not completed at this stage.

A more detailed examination of deletion mutations and of the nature of the gene product and the pattern of its expression is provided chapter 6. This work revealed that, at the N-terminal of the protein, there is a signal peptide defined by three structurally distinct regions: a basic region (amino acids 1-4), a central hydrophobic region (amino acids 5-19) and a more polar region (amino acids 21-30). An extensive protein sequence comparison between the human NDP gene and the related genes from the NBRF and Swiss protein database reveal some homology with C-termial cysteine-rich domains by a family of mucin-like protein, von Willebrand factor, the immediate-early genes (chicken CEF-10, murine Cyr61) and *Drosophila slit* protien. All of the homologies identified between the Norrie gene product and those proteins are statistically significant, e.g. comparing NDP to bovine mucin gave a BLAST score of 70 over 43 amino acids and a Poisson Probability (comparing the observed to a random alignment) of 0.0033 and for NDP vs chick CEF-10, a score of 67 over 36 amino acids, with a Poisson P of 0.0070. The significance of the homology is

highlighted by the observations that not only are 8 of the 10 cysteines strictly conserved, but so are numerous hydrophobic and turn-like motifs. Hydrophobicity patterns and secondary structure predictions indicate an all-beta topology for all family members examined (P. Bork, EMBL, Heidleberg - personal communication). Additional support for homology comes from the known genomic structures of the von Willebrand factor, members of this new growth regulator family (the immediate early genes referred to above) and the NDP gene in which this domain (defined now as the "CT" domain - P. Bork personal communication) is encoded by a single exon.

Expression studies(Table6.1) were carried out by both PCR and Northern analysis. . PCR product from the NDP gene transcript was detected in samples from of fetal eye and brain, adult brain, fetal lung and adult muscle. These studies must be regarded as preliminary and a more accurate expression pattern should be sought by quantitative PCR and Northerns and by *in situ* study of the mRNA.

The cloning of the Norrie gene has opened up opportunities to examine its function. The gradual emergence of similarity between the NDP gene product and other proteins has led to suggestions that its role may be achieved through interaction with other proteins (presumably with specific growth factors). This function appears to be confined to the C-terminal part of the protein. The function of the other part of the protein is still unknown and may not be forthcoming simply by searching through data banks. More studies are needed to allow further analysis of its function. Three main areas have to be explored. Firstly, its expression pattern in different tissues at different times of development must be investigated. From the DNA structural analysis, it is suggested that there might be a second promoter for the expression of the gene. Northern blots also revealed that more than one transcript can be detected. Genes having 5' untranslated exons generally have more than just one transcript; the various forms being associated either with tissue specificity or different developmental stages (Young et al., 1981). Useful information may also come through tissue *in situ* studies, and the collection of further mutation data. Secondly, "knock-out" technique will be used to produce mouse models (eg. with gene deficiency). Studies with mice also allow the relevant tissues to be examined at early stages during development, thereby improving the chance of defining the primary target of the gene and the critical timing period for its expression (Lee *et al.*, 1992). Thirdly, a similar approach can be applied to the study of *Drosophila*. As described in Chapter 5, it has been shown that the NDP gene is conserved in *Drosophila*, and a study of the NDP homologue in this easily manipulated experimental organism

highlighted by the observations that not only are 8 of the 10 cysteines strictly conserved, but so are numerous hydrophobic and turn-like motifs. Hydrophobicity patterns and secondary structure predictions indicate an all-beta topology for all family members examined (P. Bork, EMBL, Heidleberg - personal communication). Additional support for homology comes from the known genomic structures of the von Willebrand factor, members of this new growth regulator family (the immediate early genes referred to above) and the NDP gene in which this domain (defined now as the "CT" domain - P. Bork personal communication) is encoded by a single exon.

Expression studies(Table6.1) were carried out by both PCR and Northern analysis. . PCR product from the NDP gene transcript was detected in samples from of fetal eye and brain, adult brain, fetal lung and adult muscle. These studies must be regarded as preliminary and a more accurate expression pattern should be sought by quantitative PCR and Northerns and by *in situ* study of the mRNA.

The cloning of the Norrie gene has opened up opportunities to examine its function. The gradual emergence of similarity between the NDP gene product and other proteins has led to suggestions that its role may be achieved through interaction with other proteins (presumably with specific growth factors). This function appears to be confined to the C-terminal part of the protein. The function of the other part of the protein is still unknown and may not be forthcoming simply by searching through data banks. More studies are needed to allow further analysis of its function. Three main areas have to be explored. Firstly, its expression pattern in different tissues at different times of development must be investigated. From the DNA structural analysis, it is suggested that there might be a second promoter for the expression of the gene. Northern blots also revealed that more than one transcript can be detected. Genes having 5' untranslated exons generally have more than just one transcript; the various forms being associated either with tissue specificity or different developmental stages (Young et al., 1981). Useful information may also come through tissue *in situ* studies, and the collection of further mutation data. Secondly, "knock-out" technique will be used to produce mouse models (eg. with gene deficiency). Studies with mice also allow the relevant tissues to be examined at early stages during development, thereby improving the chance of defining the primary target of the gene and the critical timing period for its expression (Lee *et al.*, 1992). Thirdly, a similar approach can be applied to the study of *Drosophila.* As described in Chapter 5, it has been shown that the NDP gene is conserved in *Drosophila,* and a study of the NDP homologue in this easily manipulated experimental organism

may provide more information on its role in the development of the eye, brain and nervous system as a whole.

Recent progress in all aspects of human genetics has been very impressive. The number of genes and markers being assigned to the individual chromosomes is expanding exponentially (Goodfellow and Sefton, 1991). The construction of YAC contigs of human chromosomes 21 and Y (Chumakov *et al.*, 1992; Foote *et al.*, 1992) has been a landmark in the physical characterization of the human genome and it may not be too long before contigs of all of the human chromosome can be constructed. The consequences of this work are of considerable importance for the isolation of human disease genes, forming a firm base on which the deciphering of the human genome can proceed. As more is learned about the detailed map, further insight will be gained with regard to the number of genes and their natures; together with this, more will be learned about the mechanisms underlying their functions, and more importantly the functions of their products. Knowledge accumulated through this will eventually change the way that medicine is practised today and help to promote a healthy society (hopefully both physiologically and psychologically).

REFERENCES

ARRIGHI, F.E., AND HSU, T.C. (1971) Localisation of heterochromatin in human chromosomes *Cytogenetics* 10: 81-86.

ANSON, D.S., CHOO, KH., REES, D.J.G. et al.(1984) The gene structure of human anti-haemophilic factor IX. *EMBL. J.* 3: 1053-1060.

BACH, A. W. J., LAN, N. C., JOHNSON, D. L., ABELL, C. W., BEMBENEK, M. E., KWAN, S-W., SEEBURG, P. H., AND SHIH, J. C. (1989). cDNA cloning of human liver monoamine oxidase A and B: Molecular basis of differences in enzymatic properties. *Proc. Natl. Acad. Sci. USA* 85: 4934-4938.

BALDWIN, C.T., HOTH, C.F., AMOS, J.A., DA SILVA, E.O., AND MILUNSKY, A. (1992) An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* 355:637-638.

BALLABIO A. and CAVERINO (1992) The Gene for X-Linked Kallmann syndrome: a neuronal migration defect.*Current opinion in genet and develop* 2:417-421.

BALLABIO A. AND WILLARD H. (1992) Mammalian X-Chromosome inactivation and the XIST gene. *Current opinion in genet and develop* 2: 439-447.

BATANIAN, J. AND HULTEN, M. (1987) First meiotic pairing of the X and Y in the human male. *Development.* (Suppl): 197.

BARR, M.L., AND BERTRAM, E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* 163: 676.

BATES, G.P., VALDES, J., HAMMERICH, H. et al. (1992) Characterisation of a yeast artificial chromosome contig spanning the Huntington's disease candidate region.*Nature Genet.* 1: 180-187.

BERGER W., MEINDL, A., VAN DE POL, T.J.R., CREMERS, F.P.M., ROPERS, H.H., DOEMER, C., MONACO, A.,BERGEN, A.A.B., LEBO, R., WARBURG, M., ZERGOLLEM, L., LORENZ, B., GAL, A., BLEEKER-WAGEMAKERS, E.M. & MEITINGER, T. (1992) Isolation of a candidate gene for Norrie Disease using positional cloning. *Nature Genet.* 1, 199-203.

BERTA, P., HAWKINS, J.R., SINCLAIR, A.H., TAYLOR, A., GRIFFITHS, B.L., GOODFELLOW, P.N., AND FELLOUS, M. (1990) Genetic evidence equating *SRY* and the testis-determining factor. *Nature* 348: 448-450.

BIRD, A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature* 3 21: 209-213.

BIRD, A.P. (1987) CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* 3: 342-347.

BLEEKER-WAGEMAKERS , E.M. FRIEDRICH, U., GAL, A., WIENKER, T.F., WARBURG, M., AND ROPERS, H.H. (1985) Norrie disease-close linkage between Norrie disease, a cloned DNA sequence from the proximal short arm, and the centromere of X chromosome. *Human Genet.* 71: 211-214.

BORSANI, G., TONLORENZI, R., SIMMLER, M.C., DANDOLO, L., ARNAUD, D., CAPRA, V., GROMPE, M., PIZZUTI, A., MUZNY, D., LAWRENCE, C., WILLARD, H.F., AVNER, P., BALLABIO, A. (1991) Characterization of a murine gene expressed from the inactive X chromosome *Nature* 351: 325-327.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, M., AND DAVIS, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 32: 314-331.

BROCKDORFF, N., ASHWORTH, A., KAY, G.F., COOPER, P., SMITH, S., McCABE, V.M., NORRIS, D.P., PENNY, G.D., PATEL, D., RASTAN, S. (1991)Conservation of postition and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351: 329-331.

BROWN, C.J., LAFRENIERE, R.G., POWERS, V.E., SEBASTIO, G., BALLABIO, A., PETTIGREW, A.L., LEDBETTER, D.H., LEVY, E., CRAIG, I.W., AND WILLARD, H.F. (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* 349: 82-84.

BROWN W.R.A., MOCKINNON P.J., VILLANSANTE, A., SPURR, N., BUCKLE, V.J., AND DOBSON, M.J. (1990) Structure and polymorphism of human telomere-associated DNA. *Cell* 63: 119-132.

BROWN, G. K., POWELLI, J. F., and CRAIG, I. W. (1980). Molecular weight differences between human platelet and placental monoamine oxidase. *Biochem. Pharmacol.* 29: 2595-2603.

BROWN, C.J., AND WILLARD, H.F., (1989) Non-inactivation of a selectable human X-linked gene that complements a murine temperature-sensitive cell cycle defect. *Am. J. Hum. Genet.* 45: 592.

BUCKLER, A.J., CHANG, D.D., GRAW, S.L., BROOK, D., HABER, D.A., SHARP, P. A., AND HOUSMAN, D.E. (1991) Exon amplification: A strategy to isolate mammalian genes based on RNA splicing *Proc. Natl. Acad. Sci USA* 88: 4005-4009.

BURGOYNE, P.,S,(1982) Genetic Homology and crossing over in the X and Y chromosomes of mammals *Hum. Genet.* 61: 85-90.

BURKE, D.T., CARLE, G.F., AND OLSON, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236: 806-812.

CARLE, G.F., AND OLSON, M.V. (1984) Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* 14: 5647-5664.

CATTANACH, B. (1975) Control of Chromosome inactivation. *Ann. Rev. Genet.* 9: 1-8.

CAWTHON, R., WEISS, R., XU, G-F., et a. (1990) A major segment of the Neurofibrotosis Type I Gene: cDNA sequence, Genomic Structure, and Point Mutation. *Cell* 62: 193-201.

CAWTHON, R. M., AND BREAKEFIELD, X. O. (1979). Differences in the A and B forms of monoamine oxidase revealed by limited proteolysis and peptide mapping. *Nature* 281: 692-694.

CHANDLEY, A.C., AMBROS, P., McBEATH, S., HARGREAVE, T.B., KILANOWSKI, F., AND SPOWART, G. (1986) Short arm dicentric Y chromosome with associated statural defects in a sterile man. *Hum. Genet.* 73: 350-353.

CHUMAKOV, I., RIGAULT, P., GUILLON, S. et al. (1992) Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359: 380-387.

COLLINS, F.A., MURPHY, D.L., REISS, A., SIMS, K.B., LEWIS,J., and ANTONARAKIS, S.E. (1992). Clinical, biochemical and neuropsychiatric evaluation of a patient with a contiguous gene syndrome due to a microdeletion of Xp11.3 including the Norrie disease locus and monoamine oxidase (MAOA and MAOB genes. *Am. J. Med.Genet* in Press

COOKE, H.J., SCHMIDTKE, J., AND GOSDEN, J.R. (1982) Characterisation of a human Y chromosome repeated sequence and related sequences in higher primates. *Chromosoma* 87: 491-502.

CRAMER, J.H., FARELLY, F.W., AND ROUND, R.H. (1976) Restriction enzyme analysis of ribosomal DNA from *S. cerevisiae*. *Mol. Gen. Genet.* 148: 233-241.

DAVIDSON, R.G., NITOWSKY, H.N., AND CHILDS, B. (1963) Demonstration of two populations of cells in the human female heterozygous for glucose-6-phoshate -dehydrogenase variants.

Proc. Natl. Acad. Sci. USA 50: 481.

DAVIES, K. (1991) The costs of instability. *Nature* 356:15.

DE LA CHAPELLE, A., SANKILA, E.-M., LINDOF, M., AULA, P. & NORIO, R. (1985) Norrie disease caused by a gene deletion allowing carrier detection and prenatal diagnosis. *Clin. Genet.* 28, 317-320.

DENNEY, R. M., and DENNEY, C. B. (1985). An update on the identity crisis of monoamine oxidase: New and old evidence for the independence of MAOA and MAOB. *Pharmacol. Ther.* 30: 227-259.

DIERGAARDE, P.J., WIERINGA, B., BLEEKER-WAGEMAKERS, L.M., SIMS, K.B., BREAKFIELD, X.O. and ROPERS, H.H. (1989). Physical fine mapping of a deletion spanning the Norrie gene. *Hum. Genet.* 84, 22-26

DISTECHE, C.M., BRANNAN, C.I., LARSEN, A., ADLER, D.A. SCHORDERET, D.F., GEARING, D., COPELAND, N.G, JENKINS, N.A. (1992) The human pseudoautosomal GM-CSF receptor a subunit gene is autosomal in mouse. *Nature Genet.* 1: 333-336.

DISTECHE, C.M., CASANOVA, M., SAAL, H., FRIEDMAN, C., SYBERT, V., GRAHAM, J., THULINE, H., PAGE, D.C., AND FELLOUS, M. (1986) Small deletions of the short arm of the Y chromosome in 46,XY females. *Proc. Natl. Acad. Sci. USA* 83: 7841-7844.

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

DRACOPOLI, N.C., MILLER, J., JOYCE, D., STEEN,R. et al. (1992) Progress towards a complete map of the mouse genome. Genomic Mapping and Sequencing *Cold Springs Harbor* Abstract.

ELLIS, N., KIDD, J., GOODFELLOW, P.J., KIDD, K., AND GOODFELLOW, P.N. (1990a) Strong linkage disequilibrium between the *XY274* polymorphism and the pseudoautosomal boundary. *Am. J. Hum. Genet.* 46: 950-955.

FERGUSON-SMITH, M.A. (1966) X-Y chromosomal interchange in the etiology of true hermaphroditism and of XX Klinefelter's syndrome. *Lancet* ii: 475-476.

FIEJTER (1984) *Am. J. Hum. Genet.* 36: 218-226.

FISHER, E.M.C., BEER-ROMERO, P., BROWN, L.G., RIDLEY, A., McNEIL, J.A., LAWRENCE, J.B., WILLARD, H.F., BIEBER, F.R., AND PAGE, D.C. (1990b) Homologous ribosomal protein genes on the human X and Y chromosomes: escape from inactivation and possible implications for Turner syndrome. *Cell* 63: 1205-1218.

FOOTE, S., VOLLRATH, D., HILTON, A., PAGE, D.C. (1992) The human Y chromosome overlapping DNA clones spanning the eukaryotic region. *Science* 258: 60-66.

FRASER, N., BALLABIO, A., ZOLLO, M., PERSICO, G., AND CRAIG, I. (1987) Identification of incomplete coding sequences for steroid sulphatase on the human Y chromosome: evidence for an ancestral pseudoautosomal gene? *Development* 101 (Suppl.): 127-132.

FREIJE, D., SCHLESSINGER, D., (1992) A 1.6-Mb Contig of Yeast Artificial chromosome around the Human Factor VIII Genes reveals three regions homologous to probes for the DXS115 locus and two for the DXYS64 locus.*Am. J. Genet.* 51: 66-80.

GAL, A., WIERINGA, B., SMEETS, D.F., BLEEKER-WAGEMAKERS, L.M., & ROPERS, H.H. (1986) Submicroscopic interstitial deletion of the X chromosome explains a complex genetic syndrome dominated by Norrie disease. *Cytogenet. Cell Genet.* 40, 219-224.

GELEHRTER AND COLLINS (1990) Principles of Medical Genetics *Williams and Wilkins*, Baltimore.

GLOVER, V., and SANDLER, M. (1986). Clinical chemistry of monoamine oxidase. *Cell Biochem. Funct.* 4: 87-97.

GOODFELLOW, P.N., PYM,B.,MOHANDAS, T., AND SHAPIRO, L.J. (1984) The MIC2X locus escapes X-inactivation. *Am. J. Hum. Genet.* 36: 777-782.

GOODFELLOW, P.J., MONDELLO, C., DARLING, S.M., PYM, B., LITTLE, P., AND GOODFELLOW, P.N. (1988) Absense of methylation of a CpG-rich region at the 5' end of the MIC2 gene on the active X, the inactive X, and the Y chromosome. *Proc. Natl. Acad. Sci. USA* 85: 5605-5609.

GOODFELLOW P.N. AND SEFTON, L. (1992) Language of the Genome *Nature* 353: 117-118.

GRANT, S.G. AND CHAPMAN, V.M. (1988) Mechanisms of X-chromosome regulation. *Ann. Rev. Genet.* 22: 199-233.

GRAVES, J.A.M. (1987) The evolution of mammalian sex chromosomes and dosage compensation: clues from marsupials and monotremes. *Trends Genet.* 3: 252-256.

GRODEN, J., Thliveris, A., Samowitz, W., *et al.* (1991) Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 66: 589-600.

GUBBAY, J., COLLIGNON, J., KOOPMAN, P., CAPEL, B., ECONOMOU, A., MÜNSTERBERG, A., VIVIAN, N., GOODFELLOW, P., AND LOVELL-BADGE, R. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 346: 245-250.

GUELLÄEN, G., CASANOVA, M., BISHOP, C., GELDWERTH, D., ANDRE, G., FELLOUS, M., AND WEISSENBACH, J. (1984) Human XX males with Y single-copy DNA fragments. *Nature* 307: 172-173.

HARDY, J. (1992) Framing B-amyloid. *Nature Genet.* 1: 233-234.

HARE, M.L.C. (1928) Tyramine oxidase A new enzyme system in liver. *Biochem J.* 22:968-979.

HENDRIKS, R.W. (1991) X Chromosome Inactivation in Human X linked Immunodeficiency diseases. Ph.D Thesis.

HOHEISEL, J.D., GRESS, T., AND LEHRACH, H. (1992) Partial sequencing by hybridisation of transcribed sequences of Drosophila melanogaster. Genomic Mapping and sequencing *Cold Springs Harbor* Abstract.

HOHEISEL, J.D., LEHRACH, H. (1992) Relational genomic analysis of Drosophila melanogaster. Genomic Mapping and sequencing *Cold Springs Harbor* Abstract.

HOTARMISLIGIL, G. S., and BREAKEFIELD, X. O. (1991). Human monoamine oxidase A gene determines levels of enzyme activity. *Am.* J. *Hum. Genet.* 49: 383-392.

HSU,Y-P., POWELL, J. F., SIMS, K. B., and BREAKEFIELD, X. O. (1989). Molecular genetics of the monoamine oxidases. J. *Neurochem.* 53: 12-18.

HSU, Y-P.P., WEYLER, W., CHEN, S., SIMS, K. B., RINEHARTT, W. B., UTTERBACK, M. C., POWELL J. F., AND BREAKEFIELD, X. O. (1988). Structure features of human monoamine oxidase A elucidated from cDNA and peptide sequence. J. *Neurochem.* 51:1321-1324.

JACOBS, H.J., LINDPAINTER, K., LINCOLN, S.E., et al. (1991) Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive Rat. *Cell* 67: 213-224.

JACOBS, P.A., AND STRONG, J.A. (1959) A case of human intersexuality having a possible *XXY* sex-determining mechanism. *Nature* 183: 302-303.

JÄGER, R.J., ANVRET, M., HALL, K., AND SCHERER, G. (1990) A human XY female with a frame shift mutation in the candidate testis-determining gene *SRY*. *Nature* 3 48: 452-454.

JORDAN, E. (1992) The human genome project: where did it come from, where is it going? *Am. J. Hum. Gnet.* 51:1-6.

KLEIN, H.L., AND WELCH, S.K. (1980) Inverted repeat sequences in yeast nuclear DNA. *Nucleic Acids Res.* 8: 4651-4669

KNOLL, J. and MAGYAR, K. (1972) Some puzzling pharmacological effects of monoamine oxidase inhibitors. *Anu . Biochem. Psychopharmacol.* 5: 393-408.

KOOPMAN, P., ASHWORTH, A., AND LOVELL-BADGE, R. (1991) The *ZFY* gene family in humans and mice. *Trends Genet.* 7: 132-136.

LAN, N. C., HEINZMANN, C., GAL, A., KLISAK, I., ORTH, U., LAI, E., GRIMSBY, J., SPARKES, R. S.,MOHANDAS T., and SHIH J. C. (1989). Human monoamine oxidase A and B genes map to Xp11.23 and are deleted in a patient with Norrie disease. *Genomics* 4: 552-559.

LAVAL, S. H., CHEN, Z.-Y., and BOYD, Y. (1991). The properdin structural locus (Pfc) lies close to the locus for tissue inhibitor of metallothionein proteases (TIMP) on the mouse X chromosome. *Genomics* I0: 1030-1034.

LAU, Y.-F. (1985) Organization of the human Y-specific *Hae*III 3.4kb repeat sequences and their application in clinical diagnosis. pp.177-192, in The Y chromosome, Part A: Basic characteristics of the Y chromosome, ed. SANDBERG, A.A., A.R.Liss, New York.

LEE, E. Y-H, CHANG, C-Y, HU, N. et al. (1992) Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. *Nature* 359: 288-294.

LEGIOUS, R., HARDELIN, J-P., LEVILLIERS, J., CLAVERIE, J-M., MILASSEAU, P., LE PASLIER, D., COHEN, D., CATERINA, D., BOUGUELERET, L., DELEMARRE-VAN DE WAAL, H., LUTFALLA, G., WEISSENBACH, J., AND PETIT, C. (1991) The Candidate Gene for the X-linked Kallman Syndrome Encodes a Protein Related to adhesion molecules *Cell* 67: 423-435.

LEVY, E. R., POWELL, J. F., BUCKLE, V. J., HSU, Y. P., BREAKEFIELD, X. O., and CRAIG, I. W. (1989).Localization of human monoamine oxidaseA gene to Xp11.23-11.4 by in situ hybridization: Implications for Norrie disease. *Genomics* 5: 368-370.

LEWISOHN, R., GLOVER, V., AND SANDLER, M. (1980) Development of benzylamine oxidase and monoamine oxidase A and B in man. *Biochem. Pharmacol.* 29: 1221-1230.

LITTLE, R.D., PORTA, G., CARLE, G.F., SCHLESSINGER, D., AND D'URSO, M. (1989) Yeast artificial chromosomes with 200-to 800-kilobase inserts of human DNA containing HLA, Vk, 5S, and Xq24-Xq28 sequences. *Proc. Natl. Acad. Sci. USA* 86: 1598-1602.

LOCK, L. F., TAKAGI, N., and MARTIN, G. R. (1987). Methylation of the HPRT gene on the inactive X occurs after chromosome inactivation. *Cell* 48: 39-46.

LYKKESFELDT, G., LYKKESFELDT, A.E., AND SKAKKEBÆK, N.E. (1984) Steroid sulphatase in man: a non inactivated X-locus with partial gene dosage compensation. *Hum. Genet.* 65: 355-357.

LYON, M.F. (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190: 372-373.

MANTLE, T. J., and TIPTON, K. F. (1978). Monoamine oxidase A and B: A useful concept? *Biochem. Pharmacol.* 27: 97-101.

MCCAULEY, R., and RACKER, E. (1973). Separation of two monoamine oxidases from bovine brain. *Mol. Cell Biochem. 1:* 73-81.

McCOMBIE, W.R., ADAMS, M.D., KELLEY, J.M. et al. (1992) Caenorhabditis elegans expressed sequence types identify gene families and potential disease gene homologues. *Nature Genet.* 1: 124-131.

McKUSICK, V.A. (1990) Mendelian inheritance in man, 8th ed. Johns Hopkins University Press.

McKUSICK V. (1992) Presidential Address, Eighth International Congress of Human Genetics: Human Genetics: The Last 35 Years, the Present, and the Future. *Am. J. Hum. Genet.* 50: 661-670.

MIGEON, B.R., SHAPIRO, L.J., NORUM, R.A., MOHANDAS, T., AXELMAN, J., AND DABORA, R.C. (1982) Differential expression of the steroid sulphatase locus on the active and inactive human X chromosome. *Nature* 299: 838-840.

MOHANDAS T. (1987) Cytogenetic and Molecular studies on a recombinant Human X chromosome:implications for the spreading of X chromosome inactivation.*Proc Natl Acad Sc USA* 84: 4954-4958.

MOHANDAS, T., SPARKES, R.S., AND SHAPIRO, L.J. (1981) Reactivation of an inactive human X chromosome: evidence for X-inactivation by DNA methylation *Science* 211: 383-396.

MONACO, A.P., BERTELSON, C.J., MIDDLEWORTH, W. et al. (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA sequence. *Nature* 316: 842.

MORTON N. (1991) Parameters of the human genome. *Proc Natl Acad Sci. USA* 88: 7474-7476.

MOSES, M.J., COUNCE, S.J., AND PAULSON, D.F. (1975) Synatonemal complex complement of man in spreads of spermatocytes, with details of the sex chromosome pair. *Science* 187: 363-365.

MÜLLER, C.R., BEATE, M., TRAUPE, H., AND ROPERS, H.-H. (1980) X-linked steroid sulfatase: evidence for different gene-dosage in males and females. *Hum. Genet.* 54: 197-199.

MURPHY, D. L. (1978). Substrate-selective monoamine oxidases: Inhibitor, tissue, species and functional differences. *Biochem. Pharmacol.* 27: 1889-1893.

OHNO, S. (1967) Sex chromosomes and sex-linked genes. Springer-Verlag, Berlin.

ORKIN, S.H., ALTER, B., ALTAY, C., MAHONEY, M.J. et al. (1978) Application of endonuclease mapping to the analysis and prenatal diagnosis of thalassaemia by globin gene deletion. *N. E. J. Med.* 299:166.

OTT,J. (1986) A short guide to linkage analysis in "Human genetic disease: A practical Approach. *IRL Press* Oxford.

OLIVER, S.G. AND CO-AUTHORS (1992) The Complete Sequence of Yeast Chromosome III. *Nature* 357: 38-47.

OZELIUS, L., HSU, Y. P., BRUNS, G., POWELL, J. F., CHEN, S., WEYLER, W., UTTERBACK, M., ZUCKER, D., HAINES, J., TROFALTER, J. A., CONNEALY, P. M., GUSELLA, J. F., and BREAKEFIELD, X. O. (1988). Human monoamine oxidase gene (MAOA): Chromosome position (Xp21-pll) and DNA polymorphism. *Genomics 3:* 53-58.

PAGE, D.C., deMARTINVILLE, B., BARKAR, B., WYMAN, E. WHITE, R. FRANCKE U. BOTSTEIN, B. (1982) Single copy sequence hybridisaes to polymorphic and homologous loci on humna X and Y chromosomes. *Proc Natl. Acad. Sci . USA* 79: 5352-5256.

PAGE, D.C., BIEKER, K., BROWN, L.G., HINTON, S., LEPPERT, M., LALOUEL, J.-M., LATHROP, M., NYSTRÖM-LAHTI, M., de la CHAPELLE, A., AND WHITE, R. (1987a) Linkage, physical mapping, and DNA sequence analysis of pseudoautosomal loci on the human X and Y chromosomes. *Genomics* 1: 243-256.

PAGE, D.C., HARPER, M.E., LOVE, J., AND BOTSTEIN, D. (1984) Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. *Nature* 311: 119-122.

PALMER, M.S., SINCLAIR, A.H., BERTA, P., ELLIS, N.A., GOODFELLOW, P.N., ABBAS, N.E., AND FELLOUS, M. (1989) Genetic evidence that *ZFY* is not the testis-determining factor. *Nature* 342: 937-939.

PARDUE, M.L.AND GALL, J.G. (1969) Molecular hybridisation of radioactive DNA to the DNA of cytological preparation. *Proc. Natl. Acad. Sci. USA* 64:600-604.

PINTAR, J. E., BARBOSA, J., FRANCKE, U., CASTIGLIONE, C. M., HAWKINS, M., Jr., and BREAKEFIELD, X. O. (1981). Gene for monoamine oxidase type A assigned to the human X chromosome. *J. Neurosci.* 1:166175.

POSCHL, E., POLLNER, R., and KUHN, K. (1988). The genes for the alpha l(IV) and alpha 2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure. *EMBO J.* 7: 2687-2695.

PRITCHARD, C.A., GOODFELLOW, P.J., AND GOODFELLOW, P.N. (1987) Mapping the limits of the human pseudoautosomal region and a candidate sequence for the male-determining gene. *Nature* 328: 273-275.

RICHARDS, R. AND SUTHERLAND, G.R. (1992) Heritable unstable DNA sequences *Nature Genet.* 1: 7-9.

RILEY, S., BLACK, G., CHEN, Z.-Y., HATCHWELL, E., LORENZ, B., MEITNIGER, T., POWELL, J., WITTWER,B., and CRAIG, I. (1991). *In* "Degenerative Retinopathies: Advances in Clinical and Genetic Research" (P. Humphries, S. Bhattacharya, and A. Bird, Eds.), pp. 119-128, CRC Press, Boca Raton, FL.

RIODAN, J.R., ROMMENS, J.M., KEREM, B.S., et al. (1989) Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. *Science* 245: 1066-1073.

ROSE, E.A., JONES, C., SMITH, C.L. et al. (1990) Complete physical map of the WAGR region of 11p13 localizes a candidate Wilm's Tumor Gene. *Cell* 60: 495-508.

ROSS, M.T., BALLABIO, A., AND CRAIG, I.W. (1990) Long-range physical mapping around the human steroid sulfatase locus. *Genomics* 6: 528-539.

SCHNAITMAN, C., ERWIN, V., AND GREENAWALT, J.W. (1967) The submitochondrial localization of monoamine oxidase. *J. Cell Biol.* 32: 719-735.

SCHNEIDER-GÄDICKE, A., BEER-ROMERO, P., BROWN, L.G., NUSSBAUM, R., AND PAGE, D.C. (1989) *ZFX* has a gene structure similar to *ZFY*, the putative human sex determinant, and escapes X inactivation. *Cell* 57: 1247-1258.

SCHWARTZ, D.C., AND CANTOR, C.R. (1984) Separation of yeast chromosome-sized DNAs by pulsed-field gradient gel electrophoresis. *Cell* 3 7: 67-75.

SHERMAN, S.L. MORTON, N.E., JACOBS, P.A., TURNER, G. (1984) The marker (X) syndrome: a cytogenetic and genetic analysis. *Ann. Hum. Genet.* 48: 21-37.

SIMS, K.B., DE LA CHAPELLE, A., NORIO, R., SANKILA, E.-M., HSU, Y.-P.P., RINEHART, W.B., COREY, T.J.,OZELIUS, L., POWELL, J.F., BRUNS, G., GUSELLA, J.F., MURPHY, D.L., & BREAKFIELD, X.O. (1989a) Monoamine oxidase deficiency in males with an X chromosome deletion. *Neuron* **2**, 1069-1076.

SIMS, K. B., OZELIUS, L., COREY, T., RINEHART, W. B., LIBERFARB, R., HAINES, J., CHEN, W. J.,NORIO, R., SANKILA, E., de la CHAPELLE, A., MURPHY, D. L., GUSELLA, J., and BREAKEFIELD, X. O. (1989b). The Norrie disease gene is distinct from the monoamine oxidase genes. *Am. J. Hum. Genet.* 45: 424-434.

SINCLAIR, A.H., FOSTER, J.W., SPENCER, J.A., PAGE, D.C., PALMER, M., GOODFELLOW, P.N., AND GRAVES, J.A.M. (1988) Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. *Nature* 336: 780-783.

SINCLAIR, A.H., BERTA, P., PALMER, M.S., HAWKINS, J.R., GRIFFITHS, B., SMITH, M.J., FOSTER, J.W., FRISCHAUF, A.-M., LOVELL-BADGE, R., AND GOODFELLOW, P.N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346: 240-244.

SOLARI, A.J. (1980) Synaptonemal complexes and associated structures in microspread human spermatocytes. *Chromosoma* 81: 315-337.

SOLOMON, E., AND BODMER, W.F. (1979) Evolution of sickle variant gene. *Lancet* i: 923.

STRATHDEE, C.A., GAVISH, H., SHANNON, W.R., AND BUCKWALD, M. (1992) Cloning of cDNAs for Fanconi's anaemia by functional conplementation. *Nature* 356: 763-767.

TAGAKI, N. (1974) Differentiation of X chromosomes in early female mouse embryos. *Exp. Cell. Res.* 86: 127-135.

TAKAGI, N., SUGAWARA, O., SASAKI, M. (1982) Regional and temporal changes in the pattern of X chromosome replication during the early postimplantation development of the female mouse. *Crhromosoma* 85:275.

TASSABEHJI, M., READ, A., NEWTON, V. E., HARRIS, R., BALLING, R., GRUSS, P., AND STRACHAN, T. (1992) Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene *Nature* 355:635-636.

TIJO, H.J., AND LEVEN, A., et al. (1956) The chromosome number of man *Heriditas* 4 2:1-6.

THERMAN, E., AND PATAU, K. (1974) Abnormal X chromosomes in man: origin, behaviour and effects. *Hum Genet.* 25:1.

TSANG, D., HO, K.P., AND WEN, H.L. (1986) Ontogenesis of multiple forms of monoamine oxidase in rat brain regions and liver. *Develop. Neurosci.* 8: 243-250.

VERELLEN-PUMOULIN, C., FREUND, M., DEMEYER, R. et al. (1986) Expression of a X-linked muscular dystrophy in a female due to translaocation involving Xp21 and non-random inactivation of the normal X chromosome. *Human Genet.* 67: 115.

VILKKI, J., OTT, J., SAVONTAUS, M.-L., AULA, P., and NIKOSKELAININ, E. K. (1991) . Optic atrophy in Leber hereditary optic neuroretinopathy is probably determined by an X-chromosomal gene closely linked to DXS7. *Am. J. Hum. Genet.* 48: 486-491.

WARBURG, M. (1966) Norrie's disease: a congenital progressive oculo-acoustico-cerebral degeneration. *Acta Ophthalmol.* (Suppl.) **89**, 1-47.

WATERSON, R., MARTIN, C., CRAXTON, M. et al. (1992) A survey of expressed gene in Caenorhabditis elegans. *Nature Genet.* 1: 114-123.

WEBER, B., SCHEMPP, W., AND WIESNER, H. (1986) An evolutionarily conserved early replicating segment on the sex chromosomes of man and the great apes. *Cytogenet. Cell Genet.* 43: 72-78.

WEATHERALL, D. (1990) The New Genetics and Clinical Practice. *Oxford University Press*Oxford.

WEISS, M.C AND GREEN, H (1967) Human-mouse hybride cell lines containing partial complements of human chromosomes and functioning human genes.*Proc Natl. Acad. Sci . USA .* 45: 1104-1111.

WEISSENBACH, J., GOODFELLOW, P.N., AND SMITH, K.D. (1989) Report of the committee on the genetic constitution of the Y chromosome. *Cytogenet. Cell Genet.* 51: 438-449.

WEYLERR, W., HSU, Y.-P. P., and BREAKEFIELD, X. O. (1990). Biochemistry and genetics of monoamine oxidase. *J. Pharmacol. Ther.* 47: 391-417.

YEN, P.H., ALLEN, E., MARSH, B., MOHANDAS, T., WANG, N., TAGGART, R.T., AND SHAPIRO, L.J. (1987) Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y interchange. *Cell* 49: 443-454.

YOUNG, R.A., HAGENBUCHLE, O., SCHIBLER, U./ (1981) A single mouse a-amylase gene specifies two different tissues specific mRNAs. *Cell* 23: 451-458.

ZHU, D., ANTONARAKIS, S.E., SCHMECKPEPER, B.J., DIERGAARDE, P.J., GREB, A.E., & MAUMENEE, I.H. (1989) Microdeletion in the X chromosome and prenatal diagnosis in a family with Norrie disease. *Am. J. Med. Genet.* **33**, 485-488.