LINKAGE STUDIES OF X-LINKED CLEFT PALATE AND ANKYLOGLOSSIA IN A BRITISH COLUMBIA NATIVE KINDRED

by

SHARON M. GORSKI

B.Sc., Simon Fraser University, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

GENETICS PROGRAMME

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1992

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature) _____

Department of MEDICAL GENETICS (GENETICS PROGRAMME)

The University of British Columbia
Vancouver, Canada

Date Dec. 7, 1992

DE-6 (2/88)

ABSTRACT

Human craniofacial malformations are a class of common congenital anomalies. Their etiology is heterogeneous and often poorly understood. To elucidate the nature of craniofacial defects at the molecular level, one approach is to study the exceptional examples of malformations which segregate in families as single gene disorders. This study employs such an approach: it is directed toward the isolation of a locus responsible for cleft palate and ankyloglossia which segregate as a single X-linked trait (*CPX*) in a British Columbia (B.C.) Native kindred.

The original description (Lowry 1970) of the clefting defect in the B.C kindred included submucous cleft palate and bifid or absent uvula. Sixty-three of the B.C. family members were clinically reevaluated and it was observed that some of the affected males and carrier females also present with ankyloglossia (tongue-tie). Ankyloglossia previously has been associated with X-linked cleft palate in an Icelandic kindred in which a locus responsible for cleft palate was provisionally assigned to the Xq21.3-q22 region (Moore et al. 1987; Ivens et al. 1988). This thesis describes linkage analyses in the B.C. kindred which were initiated with DNA markers from the Xq21-q22 region and were later expanded to include markers from the Xq13 region. No recombination was observed between *CPX* and the DNA markers *DXS447* (peak lod score [Z_{max}] = 9.38), *DXS72* (Z_{max} = 7.74; Gorski et al. 1992) and *DXS326* (Z_{max} = 2.27). A new polymorphic DNA marker, X850A/L-7, was generated and mapped to the Xq21.1-q21.3 region. X850A/L-7 was found to be partially informative in the B.C. family and also nonrecombinant with respect to *CPX* (Z_{max} = 1.91). Recombination was observed between *CPX* and *PGK1* (Z_{max} = 7.63 at recombination fraction [$\hat{\theta}$] = 0.03) and between *CPX* and *DXYS1* (Z_{max} = 5.59 at $\hat{\theta}$ = 0.04). These results localize B.C. *CPX* between *PGK1* and *DXYS1* in the Xq13.3-q21.31 region.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I thank my supervisor, Dr. Paul Goodfellow, for providing guidance, enthusiasm, support and opportunity. I would also like to thank the members of my supervisory committee, Dr. J. Friedman, Dr. D. Juriloff and Dr. R. McMaster, for their encouragement, helpful comments and discussions. I am grateful to Patricia Birch for updating the B.C. family pedigree, and performing blood sampling and clinical diagnoses, Dr. Fred Dill and Dr. Dagmar Kalousek for performing the cytogenetic analyses, Dr. Steve Wood for providing helpful assistance with the LINKAGE programs, and Dr. Virginia Diewert for sharing her time for interesting discussions.

Thank-you to the members of the laboratory, Karen Adams, Angie Brooks-Wilson, Helen McDonald, Diane Miller, and Duane Smailus who all provided expert technical advice and much moral support during the past two years. Special gratitude to Karen Adams who performed many of the DNA sample preparations and DNA genotypings, and to Diane Miller who shared with me, step by step, the experiences of graduate school. I am grateful to my family and friends for their encouragement and support throughout. Special thanks to my husband, Marco Marra, for sharing interest in my work, providing patient support, love, and friendship. Thank-you also for never failing to make me laugh.

I thank the B.C. family members for their participation which has made this study possible. Financial assistance was provided by a Medical Research Council of Canada studentship.

1. INTRODUCTION

1.1 General introduction

Cleft palate is a common congenital malformation in humans. It may be a feature of malformation syndromes which are caused by chromosomal, Mendelian or environmental factors (reviewed in Melnick et al. 1980; Sperber 1989). Alternatively, cleft palate may occur in isolation where there is evidence for involvement of both genetic and environmental factors, but underlying mechanisms are unknown (reviewed in Melnick et al. 1980). The complex etiology of cleft palate makes its study particularly difficult. One approach to the investigation of cleft palate is the study of the exceptional examples of cleft palate which segregate in families as single gene disorders. Such forms of cleft palate are amenable to analysis by virtue of the chromosomal location of the underlying gene. The process of the isolation of a gene based on the map position of a mutation in that gene is a strategy termed positional cloning. The isolation and characterization of a locus responsible for a rare Mendelian form of cleft palate may provide insight into the pathogenesis of the more common multifactorial forms of cleft palate.

The main objective of this study was to map the locus responsible for an X-linked form of cleft palate (*CPX*) associated with ankyloglossia (tongue-tie) which segregates in a British Columbia (B.C.) Native family. To better understand the nature of the molecules that may be involved in the etiology of *CPX*, the normal development of the human palate and tongue is reviewed. Some anomalies of palate and tongue development are subsequently described, followed by a description of the positional cloning method and its applications to a subset of palatal anomalies.

1.2 Development of the palate in humans

A complex sequence of events is involved in the development of the human palate. Ectomesenchymal tissue arises from the crests of the neural fold by about three weeks postconception (PC). In the cranial region, neural crest cells migrate to the six pairs of branchial arches and form the major source of skeletal and connective tissues for the face and anterior neck (Sperber 1989). In the first branchial arch, in association with pharyngeal ectoderm, neural crest cells form the bilateral maxillary prominences. Fusion of the maxillary prominences and the medial nasal and lateral nasal prominences, at about five to six weeks PC, forms the primary palate and establishes the initial separation between oral and nasal cavities (Diewert 1985; Johnston and Bronsky 1991). The primary palate gives rise to the upper lip, the alveolus, and anterior palatal region around the incisive foramen (Diewert 1985).

Extensions arise from the maxillary prominences at about six and a half weeks PC. These extensions become the bilateral palatal shelves which initially grow vertically down the sides of the intervening tongue. At approximately eight weeks, the oronasal cavity expands vertically, Meckel's cartilage and the tongue extend forward beneath the primary palate, and the palatal shelves elevate rapidly to a horizontal position above the tongue (Diewert 1983). Concomitant with these events are lifting of the upper face away from the thorax and reflex opening of the mouth (Diewert 1985). Fusion of the palatal shelves in the midline and anteriorly with the primary palate forms the secondary palate, completing the separation of the oral and nasal cavities. At the same time, fusion occurs between the downward growing nasal septum and the upper surface of the future hard palate region. The regions that will form the soft palate and uvula remain unattached to the nasal septum (Sperber 1989).

Palatal shelf midline fusion initiates in the anterior third of the palate and proceeds bidirectionally (Ferguson 1988). Fusion occurs by formation of an epithelial seam in the hard palate region, and merging of the soft palate is subsequent to

closure of the hard palate (Burdi and Faist 1967). The epithelial seam is comprised of medial edge epithelial cells of the opposing palatal shelves which adhere to each other by means of cell surface glycoproteins and desmosomes (Ferguson 1988). The epithelial cells of this seam either degenerate or transform into mesenchymal connective tissue cells, resulting in mesenchymal continuity across the hard palate (Ferguson 1988; Fitchett and Hay 1989). Ossification of the anterior two-thirds of the secondary palate proceeds during the eighth week, giving rise to the hard palate. Myogenic mesenchyme of the first, second, and fourth branchial arches migrates into the posterior third of the secondary palate, giving rise to the muscle of the soft palate (Sperber 1989).

1.3 Development of the tongue in humans

Formation of the human tongue initiates in the floor of the pharynx during the fourth week PC. Proliferation of mesenchyme from the first branchial arches results in the two lateral lingual swellings which subsequently enlarge and fuse with each other. The lingual swellings also fuse with a median elevation, the tuberculum impar, to form the mucosa of the body (anterior two-thirds) of the tongue (Pansky 1982; Sperber 1989). The mucosa of the posterior third, or root of the tongue, is provided by the copula (connector) derived from the second, third, and , to a lesser extent, the fourth branchial arches. The anterior body and posterior root of the tongue are demarcated by the V-shaped sulcus terminalis in the adult (Corliss 1976; Pansky 1982).

Underneath the fused lingual swellings and tuberculum impar (i.e. anterior body of the tongue), there is an accumulation of epithelial cells (Corliss 1976). These epithelial cells proliferate into the underlying mesenchyme of the floor of the primitive oral cavity and anchor the developing tongue to that floor. Beginning in week five, most of these cells degenerate to free the tongue body from the floor of the mouth.

The midline cell mass that remains is the lingual frenulum (Corliss 1976; Sperber 1989).

The muscles of the tongue, thought to be derived from occipital myotomes, push forward from the floor of the pharynx and underneath the mucosal layer of the tongue. The hypoglossal nerve follows along with this muscle mass, while the nerve supply of the other tongue components is derived from their respective diverse branchial arch origins (Pansky 1982; Sperber 1989).

1.4 Congenital malformations of the palate in humans

A disruption in any of the processes of normal palate development could lead to palatal clefting defects. For example, failure of the medial prominences to merge results in median cleft lip. Failure of the lateral, medial and maxillary prominences to fuse results in unilateral or bilateral cleft lip. Severity varies from a small notch of the lip to a complete cleft of the primary palate (still referred to as "cleft lip") extending into the floor of the nostril and through the alveolar process. Cleft lip may be associated with cleft palate and, in this situation, clefting of the palate occurs secondarily to the lip clefting. Nonsyndromic cleft lip with or without cleft palate has a frequency of 0.1% in European white populations (Farrall and Holder 1992), but is about one half that in black populations (Corliss 1976) and twice that in the Japanese population (reviewed in Natsume et al. 1989).

Clefts of the secondary palate range in severity to include clefts of the uvula, soft palate, and hard palate. Cleft or bifid uvula is a split in the midline of the uvula and is considered, by some, to be the least severe form of cleft palate (Pansky 1982; Sperber 1989). The prevalence of bifid uvula is approximately 1% to 3% in a mixed racial population (Weatherley-White et al. 1972; Shprintzen et al. 1985; Wharton and Mowrer 1992), and as frequent as 10% and 11% in the British Columbia Native and Navajo Indian populations, respectively (Lowry 1970; Jaffe and De Blanc 1970). Bifid

uvula is often indicative of submucous cleft palate, a condition where the soft palate is composed of mucosa with little underlying muscle (Lowry et al. 1973; Shprintzen et al. 1985). Other indicators of submucous cleft palate are a notch in the posterior edge of the hard palate and muscle separation in the midline of the soft palate (Calnan 1954). The effects of submucous clefts may include nasality, speech difficulties, and minor swallowing problems (Lowry et al. 1973). It is possible that submucous clefts are due to deficient movement of mesenchyme into the soft palate region (V. Diewert, personal communication). Another suggestion is that some submucous clefts are the result of cell death in the posterior regions of the maxillary prominence (reported in Johnston and Bronsky 1991).

The frequency of submucous cleft palate is approximately 0.08% in a mixed racial population (Weatherley-White et al. 1972). In a survey of 868 B.C. Native school children, Lowry (1979) found the frequency of submucous cleft palate to be 0.3%. Many submucous clefts remain undiagnosed even throughout adulthood (McWilliams 1991) and varying criteria for the diagnosis of submucous cleft palate is evident in the literature (reviewed in Peterson-Falzone 1991). It appears that clefts of the submucous variety, therefore, are often not included in estimates of isolated cleft palate frequencies.

Clefts of the hard palate, which almost always include soft palate clefts, are the most severe of the secondary palate clefts. The frequency of isolated clefts of the hard palate is approximately 0.07% and varies little between populations (Leck 1984; Lowry et al. 1989). Such clefts of the secondary palate are caused by a failure of the lateral palatal shelves to meet and/or fuse with each other and the nasal septum. Studies in the mouse indicate that the failure is attributable to two major factors: one is reduced palatal shelf size, and the other is delayed movement of the palatal shelves from a vertical to a horizontal position (Diewert and Pratt 1981). Similar

alterations in development are thought to be responsible for isolated hard palate clefts in humans (Diewert 1986; Johnston and Bronsky 1991).

Clefts which include the hard palate region occur twice as frequently in females as in males (reviewed in Burdi and Silvey 1969). Observations of human embryos have indicated that closure of the palatal shelves takes place approximately one week later in females than in males, and, therefore, females may be susceptible to any disruptive factors for a longer period of time (Burdi and Silvey 1969). Gender difference in the incidence of cleft palate is, however, dependent on cleft severity. The data reviewed by Burdi and Silvey (1969) indicated a female to male ratio of only 1.2:1 for clefts of the soft palate. Similar observations were made in the Japanese population where the female to male ratio was as high as 3:1 for complete secondary palate clefts but only 1:1 for bifid uvula (Natsume et al. 1989). In the Danish population, the female to male ratio was 1.1:1 in surgically treated cases of cleft palate and 0.7:1 in nonoperated cases (Christensen et al. 1992).

1.5 Congenital malformations of the tongue in humans

Anomalies of tongue development in humans include aglossia, macroglossia, microglossia, cleft tongue, bifid tongue, and ankyloglossia. Aglossia occurs when the tongue fails to develop, and is very rare. Alterations in the normal growth rate of the tongue result in macroglossia, an excessively large tongue, and microglossia, an abnormally small tongue. Incomplete fusion and absence of fusion of the lateral lingual swellings lead to cleft tongue and bifid tongue, respectively (Pansky 1982; Sperber 1989).

Ankyloglossia is manifested by an abnormally short and often thick lingual frenulum (Warden 1991). Complete ankyloglossia occurs when there is a total fusion between the tongue and the floor of the mouth. Partial ankyloglossia, or tongue-tie, varies in degree from those individuals with a shortened lingual frenulum to those in

which there is a marked fibrosis of both the lingual frenulum and the underlying genioglossus muscle (major extrinsic muscle of the tongue). Ankyloglossia limits the normal range of motion of the tongue and can contribute to speech difficulties (Horton et al. 1969; Warden 1991).

The prevalence of partial ankyloglossia in neonates has been reported as 1.7% to 4.4% with no population (black and white) predilection but a significant 3:1 (male:female) gender difference (Friend et al. 1990; Harris et al. 1992; Jorgenson et al. 1982). Earlier studies reported partial ankyloglossia frequencies of only 0.04% to 0.1% and, in one report, a 1:1 male to female ratio (McEnery and Gaines 1941; reviewed in Warden 1991). These differences may be attributed to the fact that the earlier studies were based on diagnoses in children and adults. At birth, the tongue is often short with the frenulum extending to the tip. During the early weeks to the first two years of life, the tongue grows longer and the frenulum stretches, often alleviating "neonatal ankyloglossia" (Horton et al. 1969; Sanders 1979; Warden 1991).

1.6 Positional cloning

The malformations of the palate and tongue described above are not well understood at the molecular level. One way to identify the molecules involved in these malformations is to identify families in which palate and/or tongue defects segregate as single gene disorders. Genes responsible for heritable disorders can be isolated, without prior knowledge of their function, by the positional cloning method. Positional cloning is the process of the isolation of a gene (generally a form of a gene that causes a disease) based on its position in the genome. An outline of the steps involved in the positional cloning strategy is as follows: 1) localization of the disease gene to a chromosomal region, 2) identification of nearest flanking markers to define a candidate region, 3) physical mapping of the candidate region, 4) isolation and mapping of markers across the candidate region, 5) identification and cloning of

candidate gene sequences, and 6) mutation searching (Wicking and Williamson 1991; Collins 1992). This outline is a generalized one, dependent primarily on the resources available for any given mapping project. The first two stages of positional cloning, detailed below, were employed in this study directed ultimately toward the isolation and characterization of an X-linked cleft palate and ankyloglossia locus (*CPX*).

1.6.1 Regional localization

The subchromosomal localization of a disease gene may be accomplished by the identification of cytogenetically detectable chromosomal rearrangements (deletions, duplications, inversions, or translocations) in affected individuals. For example, the localization of the Duchenne muscular dystrophy gene (*DMD*) to the Xp21 region was made possible by the detection of X/autosome translocations in affected females (Verellen et al. 1978; Jacobs et al. 1981; Boyd et al. 1986), and a deletion in an affected male (Francke et al. 1985). The assignment of *DMD* to Xp21 was subsequently confirmed by using polymorphic DNA probes from the X chromosome short arm in family linkage studies (Goodfellow et al. 1985). In the absence of any cytogenetically detectable chromosomal abnormalities, regional localization of a disease gene by positional cloning depends solely on linkage mapping. Regional localization by linkage mapping may be expedited by testing for linkage between the disease locus and previously mapped genes with functions and/or expression patterns that make them reasonable candidates for the disease.

1.6.1.1 Linkage mapping

A linkage map is a linear array of markers (genes or anonymous DNA sequences, each with at least two alleles - N.B. it is now convention in human genetics to use the term alleles to describe alternative forms of a locus, not only of a

gene) where each marker is genetically linked to at least one other marker. Two loci are said to be genetically linked when they exist on the same chromosome and the recombination fraction between them is less than 0.5. The extent of genetic linkage, then, is measured by the recombination fraction, defined as the probability that a gamete produced by a parent is a recombinant (Ott 1991). In organisms such as *Caenorhabditis elegans* and *Drosophila melanogaster*, it is relatively easy, due to short generation times and large numbers of progeny, to determine the recombination fraction between two loci. An unlimited number of planned matings can be set up and the number of resultant nonparental (i.e. recombinant) progeny is divided by the total number of progeny to obtain the fraction of recombinants. A prerequisite to accurate linkage mapping in humans is the availability of large multigeneration families in which the segregation of loci can be observed. Even then, the number of meioses is limited so mathematical models have been developed to obtain the maximum amount of information from segregation data.

The lod score method of linkage analysis, commonly used in human genetics, is based on Morton's (1955) sequential tests procedure. The lod score method determines an odds ratio for observing the segregation pattern of two loci in a family on the assumption that the two loci are linked at a given recombination fraction (θ) value (eg. between 0.0 and 0.5) compared to the assumption that the two loci are unlinked. The likelihood that two loci are linked at a given θ value [$L(\theta)$]is divided by the likelihood that the two loci are unlinked [$L(0.5)$] to give a likelihood ratio (L^*), or odds for linkage: $L^*(\theta) = L(\theta)/L(0.5)$. For any given value of θ, the corresponding likelihood ratio can be calculated and that θ value associated with the highest likelihood ratio is deemed the maximum likelihood estimate of the recombination fraction between the two loci in question. The likelihood ratio is conveniently expressed as its logarithm to the base ten, or lod (logarithm of the odds) score (Barnard 1949). Lod (Z) scores at the same θ values but generated from different

families can thus be simply added since the summation of logarithms is equivalent to the multiplication of independent likelihood ratios. It is advantageous that likelihood ratios or Z scores from different families can be combined because it is often the smaller or the less informative two-generation families that are more frequently available for analysis. Traditionally (Morton 1955), a lod score of 3 (1000:1 odds in favour of linkage) indicates that there is sufficient support for linkage of autosomal loci, a lod score of 2 indicates that there is sufficient support for linkage of loci known to be X-linked, and a lod score of -2 is considered sufficient for the absence of linkage. These values were determined by the aim that the false positive probability stay below 5 percent (Ott 1991; Risch 1992). The prior probability that two randomly selected loci are linked (eg. for $\theta < 0.3$) is approximately 2 percent (Elston and Lange 1975). At a lod score of 3, the posterior odds for linkage is the product of the prior odds and the odds from the data, or 1/50 x 1000/1 = 20/1. The posterior probability of linkage is thus 20/21 or 95.2 percent; the posterior probability of no linkage (i.e. false positive probability) is 1/21 or less than 5 percent (Risch 1992). It appears that if two loci are known to be X-linked, then the prior probability that they will be linked (for $\theta <$ 0.3) is at least 20 percent. Accordingly, the posterior odds for linkage would be 20:1 at a lod score of just 2 (1/5 x 100/1 = 20/1), and the false positive probability would be below 5 percent.

The large increase in the number of polymorphic markers on the human genetic linkage map has made possible the routine use of multipoint linkage analysis. Multipoint linkage analysis is the simultaneous analysis of several linked loci (Ott 1991). It accounts for the nonindependence of recombination estimates and the result is that multipoint analysis is often more efficient than two-point analysis (Lathrop et al. 1985; Ott 1991). In human linkage studies, for example, some families or portions of families are often informative for different subsets of marker loci. Multipoint analysis allows the incorporation of information from all available data and can

increase the number of families informative for a given linkage study (Lathrop et al. 1985).

In the positional cloning approach, multipoint analysis is often used to determine the support for various positions, x, of a disease locus relative to a map of fixed marker loci. A likelihood ratio (R) is calculated as follows: $R = L(x, \theta_{23}, \theta_{34},...) / L(\infty, \theta_{23}, \theta_{34},...)$ where x specifies a location of the disease locus in a specified interval on the fixed map and ∞ indicates that the disease locus is off the map (Keats et al. 1989). Likelihood ratios (R) are often expressed as location scores, S, where $S = -2\ln(R)$, or as multipoint lod (Z) scores, where multipoint $Z = \log_{10}(R)$ (Ott 1991).

Calculation of lod scores can now be accomplished by a number of computer programs such as LIPED (Ott 1974) and LINKAGE (Lathrop et al. 1984, 1985). The LINKAGE package of programs was used for the linkage analyses described in this thesis. The input into the LINKAGE programs is divided into a "pedfile" containing pedigree and genotype data, and a "datafile" containing loci descriptions, allele frequencies, penetrance values, recombination rates and gene order. Prior to analysis, input data is pre-processed by a program called Unknown. Unknown infers possible genotypes for parents with unspecified genotypes, and reports any inconsistencies in the data (Linkage Analysis Package User's Guide). The MLINK program performs likelihood calculations for specified incremental values of the recombination rate between two loci. Likelihood values (i.e. for $\theta = 0.5$ and the specified θ) and lod scores are reported. The ILINK program determines, by iteration, the maximum likelihood estimate of recombination fractions for a designated number of markers. When only two loci are considered, the maximum lod score is also determined (Lathrop et al. 1985; Linkage Analysis Package User's Guide). In multipoint analysis by the LINKMAP program, the location of a test locus is estimated relative to a map of known loci whose order and interlocus recombination fractions are specified and assumed to be fixed. A likelihood ratio is then calculated, for a

number of specified points within each interval on the fixed map, by comparing the likelihood of each location to the likelihood that the test locus is unlinked to the multilocus map. The resultant likelihood ratios are reported as location scores. In addition, the support for one order over another is quantified by odds ratios (Ott, 1991).

To place loci on a genetic linkage map, it is necessary to transform recombination fractions (θ) to map distances, measured in map units or centimorgans (cM). The genetic map distance (in cM) between two loci is the expected percentage of crossovers occurring in the interval between them. For closely linked loci, 1 cM is equivalent to a recombination fraction of .01 (or 1%). For loci further apart, map distance values (cM) are greater than their corresponding recombination values (%) due to the occurrence of double crossovers between two loci which are not observable as recombinants. The occurrence of double crossovers, however, is also proportional to the level of interference between two loci. Interference is the nonindependence of crossovers, or the observation that one crossover "interferes" with the occurrence of another crossover nearby. Some conversions of recombination fraction to map distance (mapping functions) incorporate different levels of interference along with the probability of double crossover events into their conversion formulas. Several mapping functions have been developed for different organisms (reviewed in Ott 1991). For human data, the Haldane (1919) and Kosambi (1944) mapping functions are most commonly used.

1.6.1.2 DNA polymorphisms

For a DNA locus to be useful in linkage studies, it must be polymorphic. In human genetics, a mating is termed potentially informative for linkage between two loci when at least one of the two parents is a double heterozygote. The degree of usefulness of a locus, therefore, depends on the number and frequency of its alleles.

Measures used to express the usefulness, or degree of polymorphism, of a locus are the heterozygosity value and the PIC (polymorphism information content) value (see section 3.2.6). Heterozygosity is the probability that a random individual is heterozygous for any two alleles at a given locus (Ott 1991); PIC is the probability that a meiotic event will be informative for a given locus (Botstein et al. 1980). Polymorphic loci may be genes or anonymous (of unknown function) DNA markers. In the case of genes, the variant may be detected as the expression of one particular genotype or another (i.e. as a phenotype), or it may be detected, like the anonymous DNA markers, at the level of DNA sequence.

Variation in DNA sequence is detected in the form of restriction fragment length polymorphisms (RFLPs). RFLPs can be attributed to DNA insertions or deletions. Alternatively, RFLPs can result from the occurrence of a point mutation which either creates or destroys the recognition sequence of a restriction endonuclease. The presence or absence of the recognition site can be distinguished as a difference in length of DNA restriction fragments by Southern blot analysis. Single nucleotide mutations are common in the human genome with a disproportionately high occurrence at cytosine residues in CpG dinucleotides (Bird 1987). The increased mutation rate is due to an increased methylation of such cytosines (5-methylcytosine) followed by deamination. If not repaired, the result is a cytosine to thymine transition (Bird 1987). Therefore, restriction enzymes that contain CpG in their recognition sequence (eg. *Msp*I, *Taq*I) are especially useful for detecting RFLPs. Most marker loci detected as RFLPs have only two alleles (i.e. a restriction site is either present or absent) and thus are often of limited use in a given linkage study; the maximum probability that an individual will be heterozygous is 50%.

Tandemly repeated sequences frequently result in more informative polymorphisms because they usually have several alleles corresponding to different numbers of repeating units. Minisatellites are short 11 to 60 bp core sequences

(Jeffreys et al. 1985) which are present as variable numbers of tandem repeats (VNTRs) (Nakamura et al. 1987). Differences in repeat number can be detected by restriction enzyme digestion and Southern blot analysis (i.e. as RFLPs) or by the polymerase chain reaction (PCR). Microsatellites consist of 1 to 6 bp motifs iterated 10 - 50 times in tandem arrays which occur randomly throughout most eukaryotic genomes (Hamada et al. 1982; Weber 1990). The type of microsatellite most often used for linkage studies in humans is the CA dinucleotide. The CA dinucleotide microsatellite is present approximately 50,000 to 100,000 times in the human genome (Hamada and Kakunaga 1982). Blocks of (dC-dA)n · (dG-dT)n (i.e. CA repeats) often vary by multiples of 2 bp between individuals and have been demonstrated to be highly polymorphic (Weber and May 1989; Litt and Luty 1989). Microsatellite polymorphisms are detected by PCR using primers flanking the CA repeats followed by polyacrylamide gel electrophoresis.

Anonymous polymorphic loci are assigned names according to guidelines set by members of the DNA Committee of the Human Gene Mapping Workshops (Williamson et al. 1991). Each locus name begins with the letter D (for DNA) followed by the chromosome number (or X, Y), a symbol indicating DNA complexity (eg. S for single copy), and a number to individually identify different loci on the same chromosome.

1.6.2 Identification of nearest flanking markers

Once a gene has been mapped to a chromosomal region, additional polymorphic DNA markers from that region are used in linkage mapping to determine the nearest markers flanking (i.e. proximal and distal to) the disease gene. This stage identifies critical recombinant individuals and defines the candidate region.

Refinement of the candidate region may be possible by identifying additional individuals who are recombinant with respect to markers within the candidate region.

Alternatively, additional polymorphic markers from within the candidate region may be isolated to more precisely localize the crossovers in the recombinant individuals already identified. The second approach often entails isolating the chromosomal region of interest and may employ the use of flow sorted human chromosomes, somatic cell hybrids (i.e. human and rodent cells fused in culture) containing single human chromosomes, somatic cell hybrids containing portions of human chromosomes, or microdissected human metaphase chromosomes. DNA fragments can be isolated from these chromosomal sources by constructing recombinant DNA libraries which are subsequently screened with total human DNA, or by utilizing Alu element-mediated PCR (Nelson et al. 1989; Brooks-Wilson et al. 1990). Various applications of the latter technique have been used to isolate human-specific DNA markers from a subchromosomal region of interest (eg. Ledbetter et al. 1990; Bernard et al. 1991).

In Alu element-mediated PCR (Brooks-Wilson et al. 1990), a primer homologous to the consensus sequence of the extreme 3' end of the human Alu family is used to amplify human DNA found between two adjacent elements in opposite orientations. Alu elements are members of the SINE (short interspersed repeat) family of mammalian repetitive elements and occur on average every 4 kb or about 900,000 times in the human genome (Britten et al. 1988). They are approximately 300 bp long and near their middle most contain a single restriction site for the enzyme *AluI*. There is sufficient sequence divergence between human Alu elements and their rodent homologs to allow human specific inter-Alu amplification by PCR from somatic cell human/hamster hybrid DNA.

Alu element-mediated PCR has been extended to include the use of primers directed to the 3' end of L1 elements (Ledbetter et al. 1990; Brooks-Wilson et al. 1992). L1 elements are members of the LINE (long interspersed repeat) family of mammalian repetitive elements, occur about 10,000 times in the human genome,

have an approximate size of 6 kb, contain a *Kpn*I restriction enzyme site, and display species-specific nucleotide sequence conservation (reviewed in Scott et al. 1987). The rationale behind the inclusion of L1 element homologous primers in what is now called human repeat element-mediated PCR, is that the SINE and LINE families of repeat elements are inversely distributed in the human genome (Korenberg and Rykowski 1988; Bickmore and Sumner 1989). The LINE family dominates the G positive bands of Giemsa stained human chromosomes while the SINE family dominates the G negative bands. Depending on the subchromosomal region of interest, primers homologous to human Alu and L1 elements would be expected to generate different amounts and varying ratios of inter-Alu, inter-L1,and inter-Alu/L1 PCR products.

Enrichment for a subchromosomal region is possible by restriction enzyme digestion and preparative pulsed field gel electrophoresis of DNA from a somatic cell hybrid containing the chromosome (or chromosome portion) of interest. Pulsed field gel electrophoresis (PFGE) utilizes non-uniform, alternately pulsed electrical fields to facilitate the separation of very large fragments (up to about 6,000 kb) of DNA (Schwartz and Cantor 1984; Q-life systems Inc. 1991). When PFGE is carried out with a low melting temperature agarose (preparative) gel, a DNA fragment of desired size can be excised and used as a source for cloning. Typically, the DNA is purified from the agarose gel, restriction enzyme digested, and used to construct a recombinant DNA library which is subsequently screened with total human DNA to recover human-specific clones from the starting mixed DNA source (Anand et al. 1988; van de Pol et al. 1990). Alternatively, an isolated and purified DNA fragment can be used as a template for Alu-PCR (Burright et al. 1991). Both approaches have demonstrated enrichment for clones from a defined subchromosomal region.

Once markers are generated from a subchromosomal region, it is necessary to confirm their localization in the genome. Physical mapping can be done by *in situ*

hybridization to chromosomes, or by hybridization to Southern blots or PCR amplification of DNA from somatic cell hybrids containing portions of human chromosomes or DNA from human cell lines carrying deletions or duplications. On a smaller scale, new markers can be hybridized to Southern blots of human or hybrid DNA separated by PFGE to look for the comigration of new markers with any existing markers already mapped. Markers that are demonstrated to be from the subchromosomal region of interest are subsequently screened for polymorphisms (section 1.6.1.2) for use in linkage mapping. If informative in the disease family and recombinant with respect to the disease gene, a new marker may enable a decrease in the size of the candidate region. Such a new marker would then be identified as a nearest flanking marker.

Given the current state of the human genetic map, further positional cloning experiments would likely include cloning the candidate interval for physical mapping, isolation of new markers, and identification of gene sequences (i.e. steps 3 to 5 of the positional cloning method). The final objective of positional cloning is the demonstration that an isolated gene is, in fact, the disease gene of interest. An interim goal of the Human Genome Project is to construct a genetic linkage map of polymorphic markers spaced 2 cM apart (Roberts 1990; Caskey and Rossiter 1992). A map of that resolution would ideally allow the localization of a disease gene to a candidate interval of 2 cM, or within approximately 2 megabases (Mb) of DNA (Renwick 1969). A subsequent plan of the Human Genome Project is to sequence the human genome; realization of that goal will allow exploitation of the candidate gene approach and thus increase the efficiency of the positional cloning process.

1.7 Genetic analysis of isolated cleft palate in humans

The first genetic analysis of cleft palate was carried out by Fogh-Anderson (1942) who used family recurrence data to show that isolated nonsyndromic cleft

palate is etiologically distinct from cleft lip and cleft palate. While family recurrence data suggest the involvement of genetic factors in isolated cleft palate, a simple Mendelian pattern of inheritance is not usually observed. Instead, the multifactorial (i.e. combination of several genetic and environmental factors) threshold theory is postulated for most cases of isolated cleft palate but neither the genetic nor the environmental components are well understood (Fraser 1980a; Nora and Fraser 1989). There are some exceptional examples of families, however, in which the segregation of isolated cleft palate does suggest a major gene determinant (Lowry 1970; Rushton 1979; Rollnick and Kaye 1986; Bixler 1987; Hall 1987; Moore et al. 1987; Gorski et al. 1992). Such forms of cleft palate can be investigated by the positional cloning approach, provided that adequate family material is available. Isolation and characterization of a gene involved in one of the rare Mendelian forms of cleft palate may provide insight into the nature of genes and gene products involved in the more frequently occurring, non-Mendelian cleft palate in the general population.

There exist reports of two families exhibiting autosomal dominant inheritance of isolated cleft palate. In one report, in which the ethnic background of the family is not stated, the palatal defect is described as being a cleft of the soft palate (Jenkins and Stady 1980). In a second study, Rollnick and Kaye (1986) describe a German/Dutch family with palatal anomalies varying from bifid uvula to a cleft of the posterior half of the hard palate. Another autosomal dominant form of cleft palate occurs as part of a clefting syndrome, the Van der Woude syndrome (VWS). In VWS families, an affected individual may have only an isolated cleft palate, but his/her family pedigree will also include paramedian lower-lip pits and/or cleft lip and/or hypodontia (Murray et al. 1990). Linkage analyses based on a combined candidate gene and candidate cytogenetic region approach were performed in six three-generation VWS families and resulted in the localization of the VWS gene to 1q32-1q41 (Murray et al. 1990).

X-linked inheritance of isolated cleft palate has been documented in at least six families (Lowry 1970; Rushton 1979; Rollnick and Kaye 1986; Bixler 1987; Hall 1987; Moore et al. 1987; Gorski et al. 1992). In one of the first reports of X-linked cleft palate, Lowry (1970) presented a large Native kindred from British Columbia (B.C.), Canada. The cleft palate in this family was described as being inherited as an X-linked trait and characterized by submucous cleft palate and bifid or absent uvula (Lowry 1970, 1971). This thesis includes a clinical reevaluation of a portion of the B.C. Native family which led to the observation that some of the affected males and carrier females present with ankyloglossia. X-linked cleft palate has been associated with ankyloglossia in at least three other families. The first family is German (Rollnick and Kaye 1986), the second is from Kentucky (Hall 1987), and the third is Icelandic (Moore et al. 1987). The cleft palate in these families is characterized by hard palate clefts (Icelandic family and, possibly, Kentucky family), soft palate clefts, submucous clefts and bifid or absent uvula (Rollnick and Kaye 1986; Hall 1987; Bjornsson et al. 1989). Linkage studies in the Icelandic kindred resulted in a provisional assignment of an X-linked locus responsible for cleft palate (*CPX*) to the Xq21.3-q22 region between *DXYS12* and *DXS17* (Moore et al. 1987; Ivens et al. 1988; Mandel et al. 1989; Stanier et al. 1991). Linkage studies suggest that the cleft palate locus in the German family maps to the same region (reported in Moore et al. 1990).

The large size of the B.C. Native family, the availability of a large number of polymorphic markers on the X chromosome, and the provisional assignment of the Icelandic *CPX*, provide an ideal framework for mapping the B.C. cleft palate and ankyloglossia locus (B.C. *CPX*). Directed toward this goal, this thesis has two main objectives. The first objective is the regional localization of B.C. *CPX* on the X chromosome. This goal was accomplished by performing linkage studies in the B.C. family with existing DNA polymorphic markers from the Xq21-q22 region (i.e. the Icelandic *CPX* candidate region). Linkage analyses were later expanded to include

markers from the Xq13 region. The second main objective of this thesis was to refine the B.C. *CPX* candidate region to the maximum resolution possible given the available family material. To address this objective, new DNA markers were generated from the *CPX* candidate region to facilitate the precise localization of two critical crossovers.

2. MATERIALS AND METHODS

2.1 Pedigree and clinical evaluation of family members

The B.C. family consists of more than 160 living individuals in 4 generations. Observations described in this thesis and previous records establish classical X-linked inheritance over 5 generations. In addition to the submucous cleft palate and bifid or absent uvula already reported (Lowry 1970), ankyloglossia was observed in a significant number of affected males and carrier females. Individuals were considered to be affected if they exhibited submucous cleft palate and/or ankyloglossia. Bifid or absent uvula and/or a notch in the posterior edge of the hard palate were used as indicators of submucous cleft palate (Lowry et al. 1973; McWilliams 1991). Bifid uvula alone was not considered as adequate for assigning disease status in light of its high frequency (10%) in the British Columbia Native population (Lowry 1970).

Sixty-three family members were examined clinically (P. Birch, J. Friedman, and P. Goodfellow, personal communication; for details of the family structure, see pedigree in fig.1). Ten of the 16 affected males have both submucous cleft palate and ankyloglossia. Four affected males have submucous cleft palate alone, and one affected male has ankyloglossia alone (III-5). We were unable to examine personally one of the 15 affected males (II-7), but it was observed and reported elsewhere that he has a visible incomplete cleft (Lowry 1970). It is unknown whether he has ankyloglossia. Females were considered carriers if they are clinically affected, have affected sons, or have an affected father. Six (75%) of the eight carrier females are affected. Five of the affected females exhibit ankyloglossia, and one female (II-4) has both ankyloglossia and submucous cleft palate. The disease phenotype of one individual (IV-8) is uncertain at present.

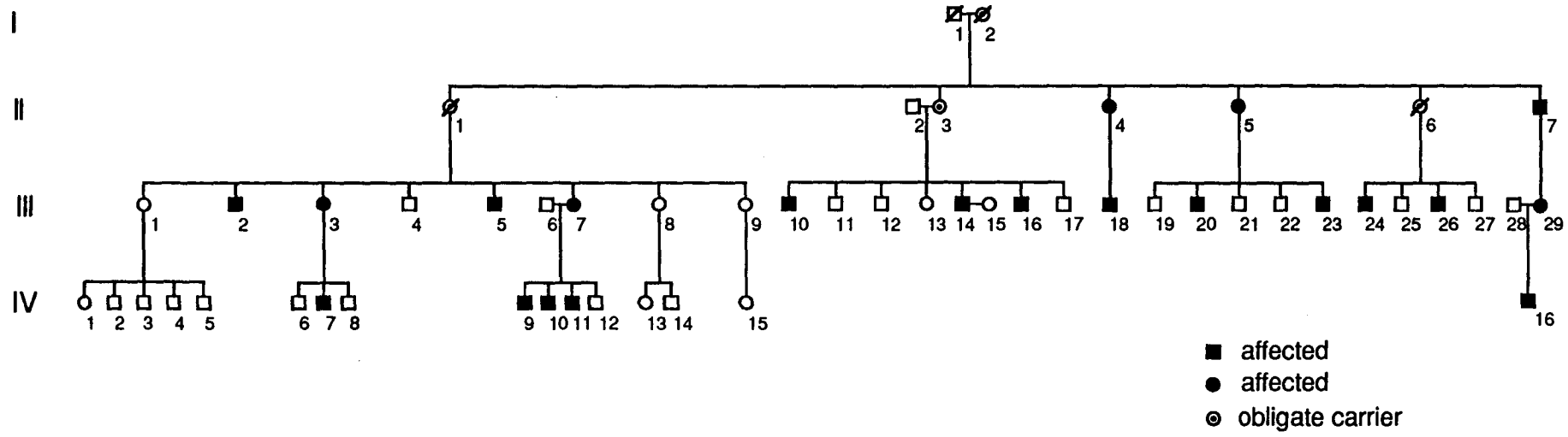


Figure 1. Pedigree of B.C. *CPX* kindred members used for linkage analysis. Genotype data were obtained for all living individuals shown. The disease phenotype of individual IV-8 is uncertain at present. A slash (/) through a symbol indicates that the individual is deceased.

2.2 Cytogenetic analysis

High-resolution analysis of X chromosomes was requested for two B.C. family members, an affected female (III-7) and her affected son (IV-11). The GTG (G banding using Trypsin and Giemsa) technique was performed on prophase chromosome preparations from blood lymphocyte cultures by D. Kalousek (Cytogenetics Laboratory, B.C. Children's Hospital). Interpretation of the cytogenetic analysis was confirmed by F. Dill.

2.3 DNA preparation, Southern transfer, and hybridization

Blood samples (10 to 30 ml) were obtained from 50 family members (fig.1). Genomic DNA was extracted using standard SDS (sodium dodecyl sulphate) / proteinase K digestion and phenol extraction protocols (Sambrook et al. 1989). Aliquots of 5 µg of genomic DNA were digested with appropriate restriction enzymes for Southern blot analyses. The DNA was separated by gel electrophoresis on 0.7-1.0% agarose gels and transferred by the method of Southern (1975) to GeneScreenPlus (Du Pont) membrane. Prehybridization, hybridization, and washes were carried out according to the manufacturer's recommendations. In general, prehybridization was performed for at least four hours at 65°C in GeneScreenPlus hybridization solution (1% SDS, 1 M sodium chloride, 10% dextran sulphate). Hybridization was carried out overnight at 65°C in the same solution with the addition of denatured salmon sperm DNA (100 µg/ml hybridization solution). Membranes were washed for 10 to 20 min at RT in 2 X SSC, and 15 to 60 min at 65°C in 0.2 X SSC and 0.2% SDS (1 X SSC is 150 mM sodium chloride, 15 mM tri-sodium citrate). Autoradiography was at -70°C for 1-10 days with intensifying screens.

Blots were stripped in 0.4N NaOH for 40 min at 43°C, followed by neutralization in 0.2M Tris, 0.2 X SSC, and 0.2% SDS for 30 min at 43°C. When not in use, blots were stored in sealed plastic bags at 4°C.

2.4 DNA polymorphisms

The B.C. family was tested for DNA polymorphisms at twenty-five loci. Table 1 lists the twenty-five DNA marker loci, names of probes used, X chromosome regional assignments, enzymes with which variants are detected, allele sizes and frequencies, and genotypes for the markers uninformative in the B.C. family. The latter is reported for possible future reference. Information in table 1 is from D. Barker (personal communication), P. Bridge (personal communication), Browne et al. (1991), Fain et al. (1991), S. Gorski (sections 3.1.2 and 3.2.5), Lafrenière et al. (1991), Molloy et al. (1992), M-A. O'Reilly (personal communication) and Williamson et al. (1991).

To facilitate DNA manipulations, the inserts for pXG7a (*DXS95*) and pDP34 (*DXYS1*) were recloned into pBluescriptIIKS+ (Stratagene). The pBluescript DNA and clone DNA were digested with the appropriate restriction enzymes. Digested DNAs in a vector/insert molar ratio of 1:2 were combined in a 10 µl ligation reaction with 1 X T4 DNA ligase buffer (5 X DNA Ligase Reaction Buffer = 250 mM Tris, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25% polyethylene glycol-8000; GIBCO BRL) and 1 U of T4 DNA ligase (GIBCO BRL). Ligation was carried out overnight at 16°C. The ligation reaction was diluted 5 X with TE prior to transformation into competent *E. coli* strain DH5α (prepared by K. Adams). Transformants were selected on ampicillin (50 µg/ml) and X-Gal/IPTG containing LB agar plates (Sambrook et al. 1989). The resultant pBluescript clones were named pXG7a-Bls and pDP34-Bls (table 1). Plasmid DNA was prepared according to a standard minipreparation alkaline lysis protocol (Sambrook et al. 1989).

Table 1. X chromosome markers used in present study.

Locus	Probe	Regional Assignment	Enzyme	Alleles (kb)	Frequencies	*Genotype if uninformative
DXS1	p8	Xq11-q12	*Taq*I	15.0/9.0	.84/.16	1
DXS159	cpX289	Xq12	*Pst*I	5.5/1.6	.67/.33	1
DXS106	cpX203	Xq12	*Bgl*II	1.0/5.8	.64/.36	1
DXS441	pRX214H1(c)	Xq13.2-q13.3	*Taq*I	2.7/2.5	.64/.36	2
PGK1	pXPGK-RI0.9	Xq13.3	*Pst*I	5.2/1.8	.85/.15	-
DXS447	pRX404E2(b)	Xq21.1	*Bgl*II	4.4/3.6	.65/.35	-
DXS355	pRX167R3	Xq13-q21.2	*Hin*dIII	7.5/6.5	.83/.17	2
DXS349	pRX98H3	Xq13-q21.2	*Msp*I	14/7.2/7.0/5.2	.10/.67/.13/.10	2 or 3
DXS364	pRX272E2	Xq21	*Taq*I	6.5/2.6	.60/.40	1
DXS346	pRX86R2(a)	Xq21.1	*Bgl*II	2.4/2.1	.38/.62	2
DXS72	pX65H7	Xq21.1	*Hin*dIII	7.2/0.7	.45/.55	-
DXS367	pCH1	Xq21.1-q21.2	*Eco*RI	7.8/5.6	.47/.53	2
DXS540	pz11a pz11a pz11c	Xq21.2	*Eco*RV *Msp*I *Eco*RI	7.3/5.3 7.5/16.5 4.5/4.2	.53/.47 .69/.31 .64/.36	1 1 1
DXS326	pQST38M1(c)	Xq21.1-q21.3	*Msp*I	3.7/2.7,1.0	.62/.38	-
DXS95	pXG7a pXG7a-Bls	Xq21.2-q21.3	*Taq*I	10.5/9.0	.90/.10	1
DXS262	pKZ040H2	Xq21.2-q22.1	*Bgl*II	14.5/8.3	.36/.64	1
DXYS1	pDP34 pDP34-Bls primers: DXYS1/4-1A DXYS1/4-1B	Xq21.31	*Taq*I (CA)n	10.6/11.8 .174/.172/ .170	.60/.40 .54/.42/.04	- **
DXYS12	St25-1	Xq21.31-q21.33	*Taq*I	2.1/1.6	.23/.77	-
DXS3	p19-2	Xq21.3	*Taq*I *Msp*	3.0,2.0/5.0 4.4/12	.62/.38 .75/.25	1 -
PLP	primers: 4 5	Xq21.33-q22	*Aha*II (*Acy*I)	.333/.235 + .098	.74/.26	-

Table 1. contd.

Locus	Probe	Regional Assignment	Enzyme	Alleles (kb) 1/2/3/4/5	Frequencies	Genotype if uninformative
DXS178	p212/9	Xq21.33-q22	*TaqI*	3.2/1.8	.70/.30	-
DXS94	pXG-12	Xq22	*PstI*	6.5/7.2	.52/.48	1
DXS101	cX52.5	Xq22	*MspI*	7.7/7.5	.35/.65	2
DXS17	S21	Xq22	*TaqI*	2.2/2.0	.65/.35	1
DXS456	primers: XG30BL XG30BR	Xq21-q22	(CA)n	.156/.148/ .146/.154/ .144	***.20/.20/ .20/.20/.20	-

*Genotypes given are those observed in the B.C. *CPX* kindred. Genotype numbers correspond to the order of the alleles (kb) given for each locus. Generally, the larger sized restriction fragment is designated 1 and the smaller sized restriction fragment is designated 2.

**Uninformative but allele size not determined.

***Allele population frequencies unknown. Frequencies given are those used for the MLINK program.

NOTE.-HGM11 reference markers are underlined. Other marker loci are listed primarily in physical map order (D. Barker, personal communication; S. Gorski, section 3.2.5; Lafreniere et al. 1991; Williamson et al. 1991).

DNA probes for all loci, with the exception of *PLP* and *DXS456* described below, detected RFLPs on Southern blots. Probes were labeled using the Bethesda Research Laboratories Nick Translation System with [^{32}P]dCTP. Alternatively, probes were gel isolated and labeled by the random primer method of Feinberg and Vogelstein (1984). Gel isolation was generally carried out in 1 X TBE buffer (0.1 M Tris, 0.1 M boric acid, 2 mM EDTA) and 1% agarose succeeded by a second isolation in 1 X TBE and 0.8% low-melting-point agarose (GIBCO BRL). Any probes containing repetitive sequences were preassociated (15 ng) with 0.5 mg of denatured sheared human placental DNA in 1 ml of GeneScreenPlus hybridization solution for 1 hr at 65°C. Hybridization was then carried out at 70°C.

The RFLP at the *PLP* locus was detected by PCR amplification with 0.4 μM each of primers 4 and 5 (table 1; P. Bridge, personal communication). Reaction conditions were altered to consist of 50-100 ng genomic DNA, 200 μM of each dNTP, 1.1U Promega Taq polymerase, and 1 x Promega reaction buffer. Thirty cycles of 2 min at 95°C, 2 min at 55°C and 3 min at 72°C, followed by 10 min at 72°C, were performed with a Perkin-Elmer Cetus thermal cycler. The PCR product was digested with *Acy*I (Promega) and electrophoresed on a 3% NuSieve (FMC) + 1% agarose gel.

The *DXS456* locus represents a microsatellite $(CA)_n$ polymorphism. Primers and reaction conditions used were as described elsewhere (Luty et al. 1990) with the following modifications: Promega reaction buffer and Taq polymerase (1.1U) were used with the addition of 0.25 mM spermidine. Thirty cycles of 2 min at 95°C, 1 min at 47°C and 2 min at 72°C, with a final 72°C incubation for 10 min, were performed with a Perkin-Elmer Cetus thermal cycler.

2.5 Linkage analyses

Linkage analyses were performed using the MLINK, ILINK, and LINKMAP programs of the LINKAGE package (v5.10) (Lathrop et al. 1984, 1985). Two-point linkage analyses between *CPX* and marker loci were conducted using MLINK. Lod score (Z) values were calculated at various recombination fraction (θ) values (table 3). The maximum two-point Z's (Z_{max}) and corresponding $\hat{\theta}$'s were determined using ILINK (table 3). The frequency of the *CPX* mutant allele was set at 0.0005. The penetrance of *CPX* was set at 1.0 in males and 0.75 in heterozygous females. Females were scored as affected or unaffected according to clinical diagnosis. All deceased individuals were scored as unknown.

Multipoint linkage analysis was performed using LINKMAP. Location scores (twice the natural logarithm-of-the-odds ratio) were calculated for all possible positions (10 points within each interval) of B.C. *CPX* relative to a map of six of the most closely linked marker loci deduced from the two-point analyses. The order of the marker loci and the distances between them were assumed to be fixed; the locus order used was *PGK1-DXS447-DXS72-DXYS1-DXYS12-DXS3*, with respective θ's of .001, .039, .044, .048 and .027. The θ's are approximations based on combined data from Keats et al. (1989, 1990) and Puck et al. (1991). *DXS326* was not included in the multipoint analysis because there are no previous reports of its map distance relative to the other markers used. Genetic distances were calculated using Haldane's (1919) mapping function: $d = -1/2 \ln (1-2\theta)$ where d = genetic distance (in M) and θ = recombination fraction.

2.6 Cosmid library screen

The pWE15 cosmid library is a commercially available library (Stratagene) made with human male lymphocyte DNA in the pWE15 vector. The library was plated by A. Brooks-Wilson and duplicate filter lifts (with Hybond-N [Amersham]

membrane) were prepared by D. Smailus. To enable repeated use, the library "plates" were stored at -70°C and duplicate filters were stored at 4°C (A. Brooks-Wilson and D. Smailus). I screened approximately 7.5 X 10^5 cosmids with 200 ng of gel isolated pXG7a-BIs insert radiolabeled by the random primer method (section 2.4). Prehybridization and hybridization were carried out in Denhardt's hybridization solution (5 X SSPE [0.9 M NaCl, 50 mM sodium phosphate, 5 mM EDTA], 5 X Denhardt's solution [100 X Denhardt's = 2% bovine serum albumin, 2% Ficoll, 2% polyvinylpyrrolidone], 0.5% SDS) with the addition of 100 µg/ml denatured salmon sperm DNA during the hybridization period. Duplicate filters were washed for 2 X 10 min at RT in 2 X SSC, and 3 X 20 min at 65°C in 0.2 X SSC and 0.2% SDS. Autoradiography was for 8 days at -70°C with intensifying screens.

2.7 Cell lines

Cl2D is a fibroblast hybrid cell line with a single human X chromosome in a Chinese hamster background (Goss and Harris 1975). Cl2D cells were grown under HAT (sodium hypoxanthine, aminopterin, thymidine) selection in Dulbecco's Modified Eagle Medium (D-MEM) supplemented with 10% fetal calf serum and 1 X antibiotic-antimycotic (all from GIBCO).

The lymphoblast cell line A.S. (used for the *Sfi*I control blot) was kindly established by A. Junker. Cell line A.S. was derived from an affected member (IV-16) of the B.C. CPX kindred. The lymphoblast Xq21 deletion cell lines, XL45, CM, XL62 and NP, were a gift from F. Cremers. The extent of the deletion in each cell line (Cremers et al. 1989) is indicated in fig.8 (section 3.2.3). All lymphoblast cell lines were expanded in culture in RPMI 1640 medium (GIBCO) supplemented with 10% fetal calf serum and 1 X antibiotic-antimycotic.

Preparation of genomic DNA from the lymphoblast cell lines was essentially the same as described in section 2.3 except that resuspension and SDS/proteinase

K digestion were initiated from a cell pellet derived from a dense 100 ml culture of growing cells (Sambrook et al. 1989). Genomic DNA from CI2D (described above), WT49 (human lymphoblast cell line), and CHOK1 (hamster fibroblast cell line; described in Brooks-Wilson et al. 1992) was prepared previously by P. Goodfellow.

2.8 Field inversion gel electrophoresis (SfiI test blot)

Agarose blocks containing approximately 5 X 10^5 cells (= 3.5 µg DNA) were prepared (Herrmann et al. 1987) from the CI2D and A.S. cell lines. Agarose cell blocks from four male controls were kindly provided by A. Brooks-Wilson. Each block was preincubated overnight at 4°C in 20 U *Sfi*I (BRL) and 1 X M buffer (Maniatis et al. 1982) in a total volume of 200 µl to allow diffusion of the buffer/restriction enzyme into the block. Digestion was performed for 5 h at 50°C. DNA samples were electrophoresed in 0.8% agarose in 1 X TBE using ROM card 5 (100 - 1100 kbp) of the AutoBase electrophoresis system (Q-liFE Systems Inc.). The AutoBase system uses computer generated electrophoresis parameters optimized for DNA separation in pulsed fields with a one dimensional electrode configuration (i.e. field inversion). Individual memory (ROM) cards store the protocols to facilitate standardization of experimental conditions (AutoBase Instruction Manual). Prior to the usual Southern blot procedure (section 2.3), the field inversion agarose gel was soaked in 0.25N HCl for 20 min to facilitate transfer of large DNA fragments (exposure to 0.25N HCl causes partial depurination of DNA; subsequent exposure of the DNA to strong base causes hydrolysis of the phosphodiester backbone at the sites of depurination). Hybridization was carried out as described in section 2.3.

2.9 Preparative field inversion gel electrophoresis to enrich for the DXS95 region

Twenty-six CI2D cell blocks were digested with *Sfi*I (section 2.8) and transferred to a 2 mm wide and 130 mm long slot in a 1% low melting point agarose

gel. An additional three *Sfi*I digested CI2D blocks and a DNA size marker (*Saccharomyces cerevisiae* strain YNN295 chromosomes, BioRad) were loaded into separate wells. Field inversion gel electrophoresis was performed with ROM card 5 of the AutoBase electrophoresis system. After electrophoresis, the lanes containing the separate CI2D samples and size marker were cut off and Southern blotted (section 2.3). The remainder of the gel was cut into 2 mm wide horizontal strips (section 3.2.1, fig.6) which were individually stored in 0.5 M EDTA, pH8 at 4°C.

The Southern blot of the 3 CI2D samples and size marker was probed with pXG7a-BIs insert. The resultant hybridization signal corresponded to several of the agarose gel horizontal strips (fig.6). Segments (3-lane-wide) from each of these strips were electrophoresed in 0.8% agarose (ROM card 5), Southern blotted, and probed with pXG7a-BIs insert. The agarose gel strip with the greatest hybridization signal was designated X850 (X chromosome 850 kb *Sfi*I fragment).

2.10 Human repeat element-mediated PCR

Human repeat element-mediated PCR was performed using the primers and reaction conditions described by Brooks-Wilson et al. (1990, 1992). The primers A1B and L1S are homologous to the consensus sequence of the extreme 3' end of the human Alu and L1 repeat elements, respectively. To facilitate cloning of PCR products, the primers have been modified by the addition of restriction enzyme sites. The A1B primer contains a *Bam*HI recognition site and the L1S primer contains a *Sal*I recognition site (Brooks-Wilson et al. 1992). Reaction templates included 100 ng of WT49 (total human XX), CHOK1 (total hamster XX) and CI2D DNAs, and 5 µl and 15 µl aliquots of X850 gel slice (washed 2 X 30 min in TE; TE = 10 mM Tris, 1 mM EDTA). For each template, PCR amplification was performed with the A1B primer, L1S primer, and the A1B and L1S primers together. For the X850 template (15 µl aliquot/high concentration reaction) two rounds of amplification were

performed. Following the first reaction, 1 μl aliquots (A1B, L1S and A1B plus L1S) were used as templates for a second amplification reaction. Reaction products were visualized on an ethidium bromide (EtBr) stained 1.2% agarose gel.

The PCR amplification products from each of the X850 reactions were used as probes for Southern blot analysis. They were extracted twice with an equal volume of chloroform/isoamyl alcohol (24:1), precipitated with 1/10 X volume sodium acetate [2.5M] and 2.5 X volume 95% ethanol, and resuspended in 1 X TE. Approximately 15 ng of products from each of the three X850 reactions were then radiolabeled by the random primer method and preassociated as described in section 2.4. Southern blots were also preassociated or "flooded" with denatured salmon sperm DNA (200 μg/ml hybridization solution) and denatured sheared human placental DNA (100 μg/ml hybridization solution) during the prehybridization and hybridization periods.

Individual A1B/L1S PCR products were cloned by ligating 25 ng of *Bam*HI/*Sal*I digested A1B/L1S product pool DNA with 5 to 40 ng of *Bam*HI/*Sal*I digested pUC18. Ligation, transformation, and selection were carried out as described in section 2.4 except that Library Efficiency DH5α™ competent cells (GIBCO BRL) were used. Positive transformants (white colonies) were tested for the presence of DNA plasmids with inserts by using small portions of the bacterial colonies as templates in Alu/L1 PCR. Thirty colonies were tested and approximately one-third of each PCR mixture was electrophoresed in a 1.2% agarose gel and transferred to a Hybond-N (Amersham) nylon membrane. The remaining two-thirds of select PCR mixtures was gel isolated, radiolabeled (as above), and hybridized (solutions as in section 2.6) to the membrane to determine the number and identities of distinct products.

3. RESULTS

3.1 Regional localization of CPX

In an effort to define the region of the X chromosome in which *CPX* is located, two approaches were taken. First, X chromosome cytogenetic analysis was requested to look for any gross structural abnormalities that might be causally associated with *CPX*. Second, linkage analyses were performed to look for cosegregation of the cleft palate and ankyloglossia phenotype (CPX) with polymorphic DNA markers already regionally mapped on the X chromosome.

3.1.1 Cytogenetics

GTG (G banding using Trypsin and Giemsa) analysis was performed on prophase chromosome preparations from blood lymphocyte cultures of an affected female (III-7) and her affected son (IV-11) (D. Kalousek, Cytogenetics Laboratory, B.C. Children's Hospital). High resolution analysis of the X chromosomes, at the 850 band level or greater, revealed no abnormalities (D. Kalousek and F. Dill, personal communication). The cytogenetics reports are appended (Appendix #1).

3.1.2 Genotypes

Genotypes for twenty-five polymorphic markers from the Xq11-q22 region were determined for the purpose of linkage analyses. Fifteen markers were uninformative. These markers and their corresponding B.C. kindred genotypes are included in table 1. The ten remaining markers were fully or partially informative and their genotype data for each living B.C. family member is listed in table 2. Individual ID is the family member's blood sample code number; B.C. pedigree ID number corresponds to the numbering in the pedigree (fig.1); allele numbers for each locus correspond to those

given (in kb) in table 1 (- = not determined); order of markers is chromosomal order, proximal to distal; the positions of *DXS326* and *PLP* are provisional.

All genotypes, except for the markers *PLP* and *DXS456* (described in 2.4), were determined by conventional RFLP analyses. An example of an autoradiograph from Southern blot analysis with the dimorphic marker *DXS447* is shown in figure 2. Deduced haplotypes for a portion of the family, which includes three recombinant individuals (III-10, IV-6 and IV-10), are presented in figure 3.

Table 2. B.C. kindred genotypes for informative markers from Xq13-q22

indiv. ID	B.C. pedigree ID	*PGK1*	*DXS447*	*DXS72*	*DXS326*	*DXYS1*	*DXYS12*	*DXS3*	*PLP*	*DXS178*	*DXS456*
1	IV-2	2	2	2	1	1	2	2	2	1	1
2	III-9	2 2	2 2	1 2	1 2	1 2	2 2	1 2	1 2	1 2	1 2
3	III-4	2	2	2	1	1	2	2	1	2	1
4	III-15	1 2	1 2	1 2	1 2	1 1	- -	1 1	1 1	1 2	1 5
5	III-10	1	1	1	1	1	2	2	1	2	1
6	III-5	1	1	1	1	2	1	1	1	2	1
7	III-12	2	2	2	1	1	2	2	1	2	1
8	IV-15	2 2	1 2	1 2	1 2	1 2	2 2	1 1	1 1	2 2	1 3
9	III-17	2	2	2	1	1	2	2	2	1	1
10	III-2	1	1	1	1	2	1	1	1	2	1
11	III-14	1	1	1	1	2	1	1	2	1	1
12	IV-3	2	2	2	1	1	2	2	2	1	1
13	III-16	1	1	1	1	2	1	1	2	1	1
14	II-2	2	2	1	2	2	2	1	2	1	2
15	III-11	2	2	2	1	-	2	2	1	2	1
16	II-3	1 2	1 2	1 2	1 1	1 2	1 2	1 2	1 2	1 2	1 1
17	III-24	1	1	1	1	2	1	1	2	1	1
18	III-26	1	1	1	1	2	1	1	2	1	1
19	III-25	2	2	2	1	1	2	2	1	2	1
20	IV-12	2	2	1	2	-	2	1	2	1	2
21	IV-11	1	1	1	1	2	1	1	2	1	1
22	III-8	2 2	2 2	1 2	1 2	1 2	2 2	1 2	1 2	1 2	1 2
23	III-7	1 2	1 2	1 1	1 2	2 2	1 2	1 1	2 2	1 1	1 2
24	IV-10	1	1	1	1	2	2	1	2	1	2
25	III-6	2	1	-	1	2	2	2	2	1	1
26	IV-9	1	1	-	1	-	1	1	2	1	1
27	IV-13	1 2	2 2	1 2	1 2	1 2	2 2	1 1	1 2	1 1	2 4
28	III-1	2 2	2 2	1 2	1 2	1 2	2 2	1 2	2 2	1 1	1 2
29	IV-4	2	2	-	2	2	2	1	2	1	2
30	III-29	1 2	- -	1 2	1 2	- -	1 1	- -	2 2	1 1	1 1
31	III-28	2	2	1	1	-	-	1	2	2	1
32	IV-14	2	2	2	1	1	2	2	1	2	1
33	III-13	2 2	2 2	- -	1 2	1 2	2 2	1 2	1 2	1 2	1 2
34	IV-1	2 2	- -	- -	1 2	- -	- -	- -	1 2	1 1	2 5
35	IV-5	2	2	2	1	1	-	2	2	1	1
36	III-18	1	1	1	1	2	1	1	2	1	1
37	II-4	1 2	1 2	1 2	1 1	- -	1 2	1 2	1 2	1 2	1 1
38	II-5	1 2	1 2	1 2	1 1	1 2	1 2	1 2	1 2	1 2	1 1
39	IV-6	1	2	1	2	2	2	1	2	1	2
40	IV-7	1	1	1	1	2	1	1	2	1	2
41	IV-8	1	1	1	1	2	1	1	2	1	1

Table 2. contd.

indiv. ID	B.C. pedigree ID	*PGK1*	*DXS447*	*DXS72*	*DXS326*	*DXYS1*	*DXYS12*	*DXS3*	*PLP*	*DXS178*	*DXS456*
42	III-3	1 2	1 2	1 1	1 2	2 2	1 2	1 1	2 2	1 1	1 2
43	III-21	2	2	2	1	1	2	2	1	2	1
44	III-22	2	2	2	1	1	2	2	1	2	1
45	III-19	2	2	2	1	1	2	2	1	2	1
46	III-23	1	1	1	1	2	1	1	2	1	1
47	III-20	1	1	1	1	2	1	1	2	1	1
48	III-27	2	2	2	1	1	2	2	1	2	-
49	II-7	1	1	1	1	2	1	1	2	1	1
50	IV-16	1	-	1	1	-	-	1	-	-	-

NOTE.-Allele numbers correspond to those given in Table 1; - indicates that the genotype is not determined.

Order of markers is chromosomal order; left to right is proximal to distal.

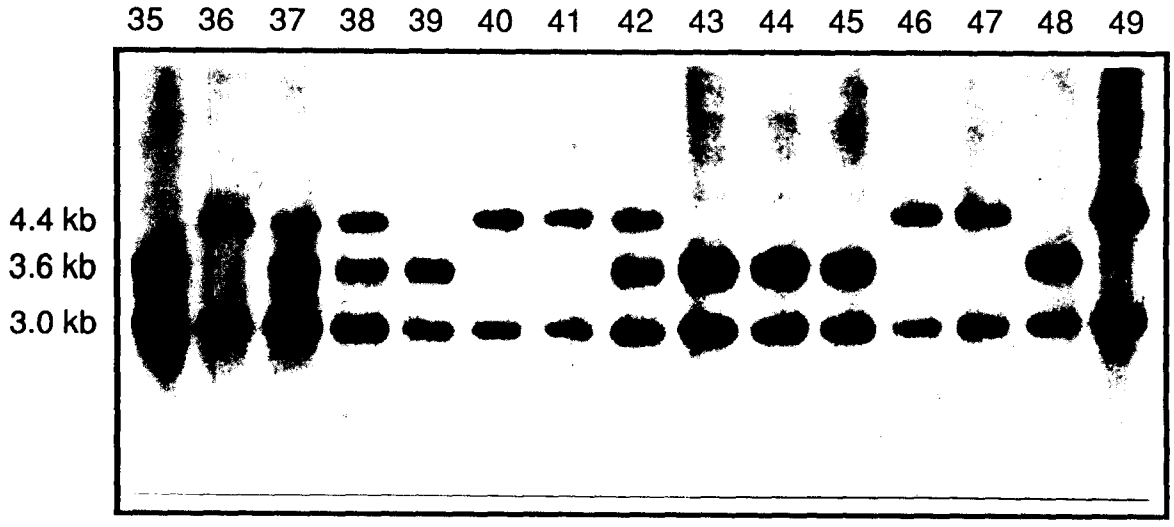


Figure 2. Autoradiograph of Southern blot illustrating *DXS447* genotypes for a portion of the B.C. *CPX* family. Lane numbers correspond to individual IDs in table 2. The *DXS447* probe (pRX404E2) detects a *Bgl*II constant band of 3.0 kb and alleles of 4.4 kb and 3.6 kb.

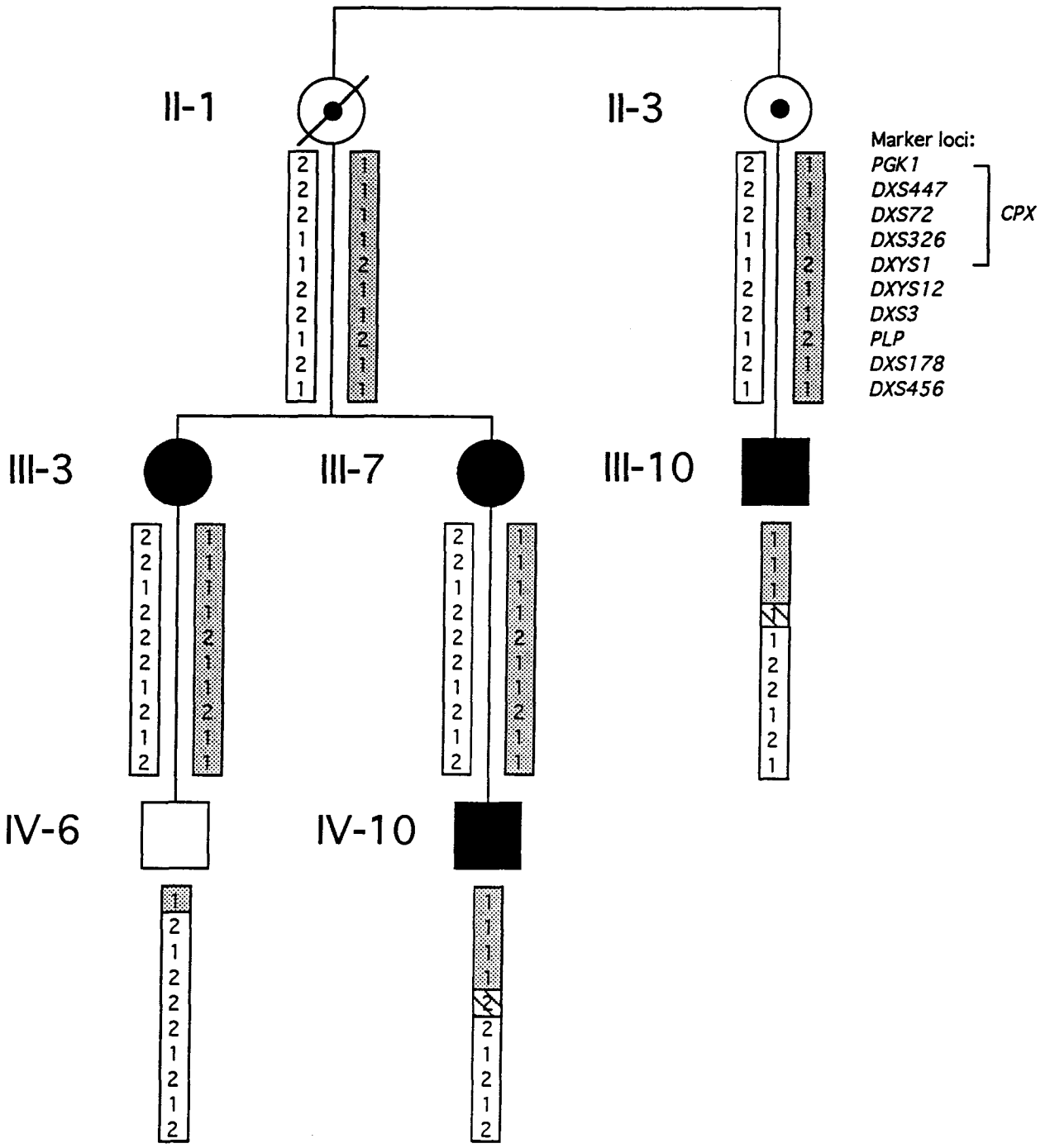


Figure 3. Demonstration of DNA marker haplotypes for a portion of the B.C. family (II-1, II-3, III-3, III-7, III-10, IV-6, and IV-10 in fig.1). Phase was inferred so as to minimize the number of crossovers. The recombination events detected in IV-6 and III-10 localize *CPX* distal to *PGK1* and proximal to *DXYS1*, respectively. The genotypes segregating with *CPX* are indicated by shading (▨). Genotypes indicated by diagonal lines (▧) are uninformative with respect to *CPX* recombination. Other symbols are as in fig.1.

3.1.3 Two-point linkage analysis

Lod score (Z) values for two-point disease-to-marker analysis are summarized in table 3. The number of informative meioses for each marker is indicated. No recombination was observed between *CPX* and *DXS447*, *DXS72*, and *DXS326* (Z_{max} = 9.38, 7.74, and 2.27, respectively). A single recombination event between *CPX* and the marker *PGK1* (Z_{max} = 7.63 at $\hat{\theta}$ = 0.03) places *CPX* distal to *PGK1* (fig.3). This critical crossover was detected in an unaffected male (IV-6). Seven recombination events were observed between *CPX* and markers distal to *DXS326*. One crossover was detected between *CPX* and *DXYS1* (Z_{max} = 5.59 at $\hat{\theta}$ = 0.04) and places *CPX* proximal to *DXYS1* (fig.3). This critical crossover was detected in an affected male (III-10). A second distal recombination event was observed between *CPX* and *DXYS12* (Z_{max} = 6.46 at $\hat{\theta}$ = 0.06) (fig.3). This recombination event was also detected in an affected male (IV-10), who was uninformative for *DXYS1*. Four of the recombination events occurred between *CPX* and both *PLP* and *DXS178*, and the final event occurred between *CPX* and *DXS456*. No multiple crossover events were observed in the interval examined. The combined results of the two-point linkage analyses, then, localize B.C. *CPX* between *PGK1* and *DXYS1* in the Xq13.3-q21.31 region (fig. 4). The genetic distance between these two loci is approximately 8.8 cM (Keats et al. 1990).

Table 3. Two-point lod scores between *CPX* and X chromosome marker loci.

	Z at θ of												
Locus	0.00	.001	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	Z_{max}	$\hat{\theta}$
PGK1	-∞	6.56	7.57	7.14	6.54	5.85	5.08	4.24	3.32	2.31	1.20	7.63	0.03
DXS447	9.38	9.37	8.66	7.90	7.11	6.28	5.40	4.47	3.48	2.41	1.25	9.38	0.00
DXS72	7.74	7.73	7.12	6.48	5.81	5.11	4.39	3.62	2.81	1.94	1.00	7.74	0.00
DXS326	2.27	2.27	2.08	1.88	1.67	1.47	1.25	1.03	0.80	0.55	0.29	2.27	0.00
DXYS1	-∞	4.44	5.56	5.26	4.82	4.29	3.71	3.08	2.40	1.66	0.85	5.59	0.04
DXYS12	-∞	3.78	6.46	6.29	5.84	5.25	4.58	3.82	2.99	2.08	1.08	6.46	0.06
DXS3	-∞	4.56	5.71	5.43	4.99	4.47	3.89	3.25	2.54	1.77	0.92	5.73	0.04
PLP	-∞	-6.30	0.75	1.70	2.03	2.09	1.98	1.77	1.45	1.05	0.56	2.09	0.19
DXS178	-∞	-6.43	0.64	1.61	1.96	2.04	1.96	1.75	1.44	1.05	0.56	2.04	0.20
DXS456	-∞	-3.64	-0.42	0.00	0.17	0.23	0.25	0.23	0.19	0.14	0.08	0.25	0.24

NOTE.-Linkage data for *PGK1*, *DXS447*, *DXS72*, *DXS326*, *DXYS1*, *DXYS12*, *DXS3*, *PLP*, *DXS178*, and *DXS456* were contributed primarily by 31, 30, 24, 7, 23, 30, 24, 24, 24, and 6 meioses, respectively.

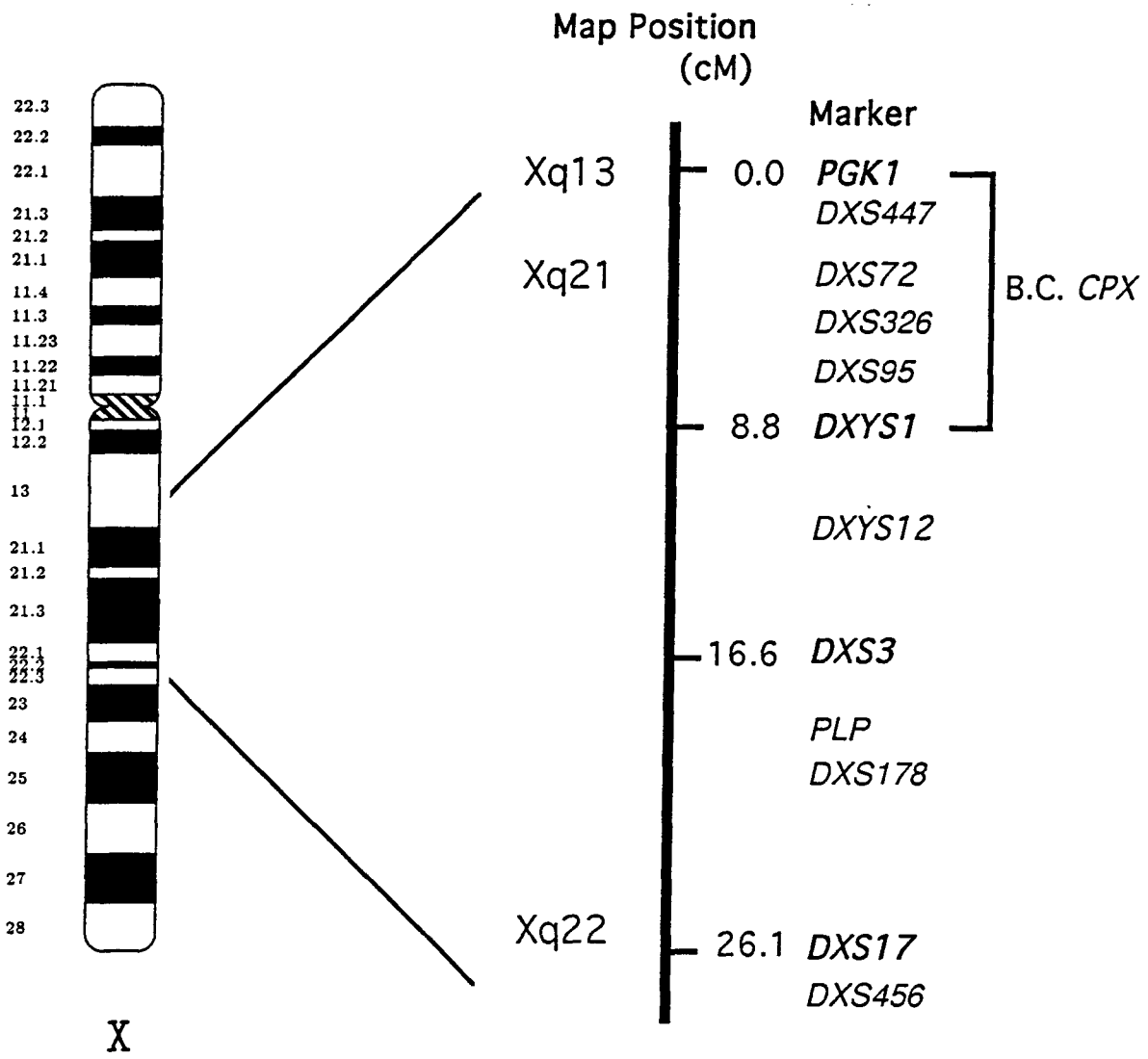


Figure 4. X chromosome map illustrating the B.C. *CPX* candidate region (Xq13.3-q21.31). Linkage relationships of the DNA markers informative for linkage analyses in the B.C. kindred are shown. Two uninformative markers, *DXS95* and *DXS17*, are also indicated. Map positions (in cM) are from HGM10.5 (Keats et al. 1990). *PGK1* was arbitrarily set at 0.0 cM on the genetic map.

3.1.4 Multipoint linkage analysis

Multipoint linkage analysis supports the *CPX* localization suggested by the recombination events described in section 3.1.3. The *CPX* location score results and fixed loci map are presented in figure 5. The location score, S (S = -2ln [likelihood ratio]), is maximum at the position of *DXS447* (S = 49.0). This value corresponds to a multipoint Z score (= log10 [likelihood ratio] or S / 2ln10]) of 10.64. The support interval (1 lod difference from the maximum-likelihood estimate of location) for *CPX* at *DXS447* is from 0.01 cM to 7.8 cM and is indicated by the dashed line in figure 5. The odds against *CPX* localization are greater than 100:1 (i.e. the value generally considered sufficient for exclusion of an interval) in all intervals tested outside of the support interval. These results indicate that *CPX* can be excluded from regions proximal to *PGK1* and distal to *DXYS1*.

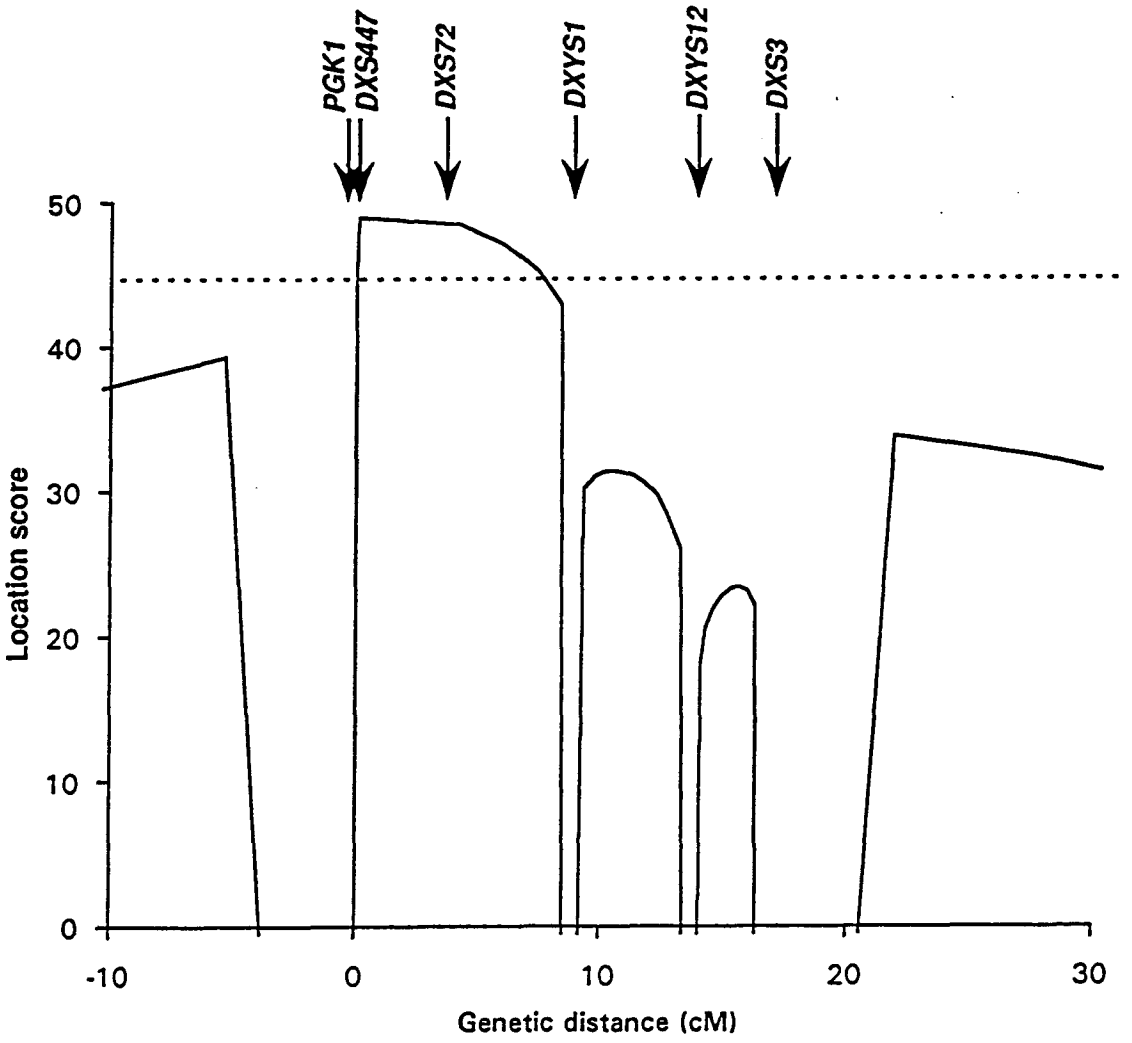


Figure 5. Multipoint linkage analysis of *CPX* vs. a fixed loci map. The position of *PGK1* was arbitrarily set at 0 cM. The overall maximum location score (= -2ln[likelihood ratio]) occurs at the *DXS447* position. The dashed line is drawn 1 lod interval below the maximum-likelihood estimate of *CPX* location.

3.2 Refinement of crossover map positions

Two-point and multipoint linkage analysis served to assign B.C. *CPX* to the interval between *PGK1* and *DXYS1*. Two, and possibly three, recombination events within this region were identified (fig.3). One approach to refining the localization of B.C. *CPX* is to precisely localize these crossovers by observing the segregation pattern of *CPX* with additional markers from within the candidate region. All reported and several as yet unpublished polymorphic markers were tested for informativeness in the B.C. kindred. In an effort to refine the localization of the disease locus, my efforts were directed to screening for additional polymorphisms and generating new markers from the *DXS95* region. This region was chosen for the following reasons:
i) The proximal crossover (detected in IV-6) had already been localized to the region between *PGK1* and *DXS447*. The distance between these loci is relatively small according to previously published data (Puck et al. 1991).
ii) In the current X chromosome map, *DXS95* is the closest marker proximal to *DXYS1* (*CPX* distal flanking marker) and, if informative, would have the potential to refine the location of both distal crossovers.
iii) A cloning source that enriches for the DXS95 locus (described below) had already been prepared in response to an unconfirmed report that *DXS95* was the closest proximal flanking marker for a cleft palate and ankyloglossia locus in another kindred. That report was later learned to be erroneous. The cloning source had, however, been prepared and represented a potentially efficient means of generating new markers from the *DXS95* region.

3.2.1 The DXS95 region: screening for additional polymorphisms and generation of a cloning source

The existing *DXS95* probe (pXG7a-Bls) was tested for its ability to detect polymorphic restriction sites in the B.C. family. Southern blot analysis of relevant B.C.

CPX family member DNAs restricted with enzymes (see section 3.2.6 below) failed to reveal additional RFLPs. Probe pXG7a-BIs was then used to screen a total human (male) genomic DNA pWE15 cosmid library in an effort to expand the locus and to develop a highly informative microsatellite repeat marker system. No *DXS95*-positive cosmids were obtained. Finally, pXG7a-BIs was used to identify a cloning source (described below) that enriches for *DXS95*. The cloning source was prepared for the generation of region-specific markers which subsequently could be used to screen Southern blots for RFLPs and ultimately to expand the locus if required.

The *DXS95* cloning source was prepared as follows. The probe pXG7a-BIs was hybridized to a Southern blot of FIGE size separated *Sfi*I digested DNA from the somatic cell hybrid CI2D (human X chromosome only). A single *Sfi*I restriction fragment of approximately 850 kb was detected, consistent with previously published reports (Cremers et al. 1989; Merry et al. 1989). The FIGE blot containing the *Sfi*I digested CI2D DNA was generated from a portion of a larger preparative gel, the remainder of which had been cut into horizontal strips and stored (section 2.9 and figure 6). The pXG7a-BIs hybridization signal corresponded to several of these agarose strips. To increase *DXS95* enrichment, segments of each strip were electrophoresed again on a field inversion agarose gel. Observation of the photograph (not shown) of the ethidium bromide stained gel indicated that during the second FIGE, the DNA in each agarose strip was further resolved. The agarose gel was Southern blotted, probed with pXG7a-BIs, and the strip with the greatest hybridization signal was identified. This agarose gel strip was designated X850 (for X chromosome 850 kb *Sfi*I fragment). The hybridization signal in X850 corresponded to a band size of approximately 850 kb. The gel photograph indicated that DNA in this size range comprised at least 50% of the DNA in the X850 gel strip while the remainder of the DNA was observed as a smear ranging from 650 kb to 900 kb in size. Despite containing comigrating SfiI fragments, X850 represented the cloning

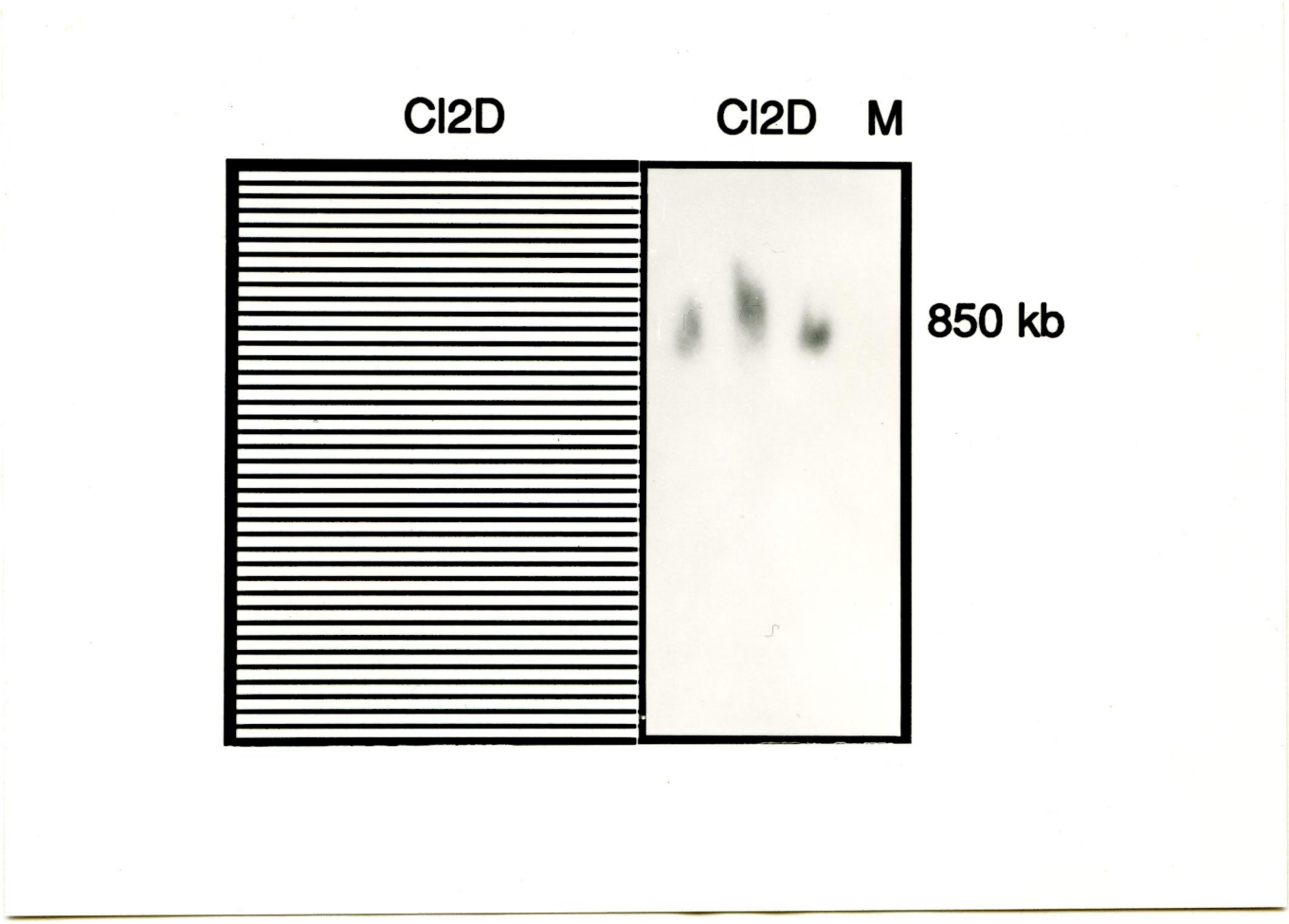


Figure 6. Schematic representation of preparative FIGE gel size fractionation (left) and Southern blot analysis of *Sfi*I digested Cl2D (human X only hybrid) DNA (right). The autoradiograph shows the results of the pXG7a-BIs hybridization (to three separate lanes containing Cl2D DNA) used to determine which 2 mm wide preparative gel slices contained the *DXS95* sequences. M = Bio-Rad DNA size standard - Yeast chromosomes, *Saccharomyces cerevisiae*.

source most enriched for *DXS95* and was subsequently used for generating new markers by human repeat element-mediated PCR.

3.2.2 Human repeat element-mediated PCR

As a means of obtaining human clones from the CI2D 850 kb *Sfi*I agarose gel fragment (X850) identified above, aliquots of the agarose gel strip were used directly as templates in human repeat element-mediated PCR. As expected (Brooks-Wilson et al. 1992), a total human genomic DNA control template (WT49) gave rise to smears on an ethidium bromide stained agarose gel with the A1B primer and the A1B/L1S primers (figure 7). A series of more discrete bands was observed with the L1S primer alone. The CI2D genomic DNA control template yielded a distinct pattern of discrete bands with each of the three primer combinations (fig.7), consistent with a reduction in human template complexity and varied distributions of Alu and L1 elements in the human genome. The total hamster genomic DNA control template (CHOK1; fig.7) and the X850 template did not produce any visible PCR products with any of the primers (not shown). When aliquots of each of the three (A1B, L1S, A1B/L1S) X850 PCR mixtures were used in second PCR runs, several distinct products and a background smear were visible in each reaction (fig.7). The products ranged in size from approximately 0.2 kb to 1.4 kb. Hamster DNA subjected to a second round of amplification yielded no visible products.

3.2.3 Mapping human repeat element-mediated PCR products by Southern blot hybridization: localization to the DXS72 - DXS95 region

The PCR products present in each second X850 reaction mixture could originate from (i) the human *DXS95* 850 kb SfiI fragment, (ii) other human X chromosome 850 kb *Sfi*I fragments, (iii) any hamster chromosome 850 kb *Sfi*I fragments, and/or (iv) human and hamster *Sfi*I fragments of a different size that have

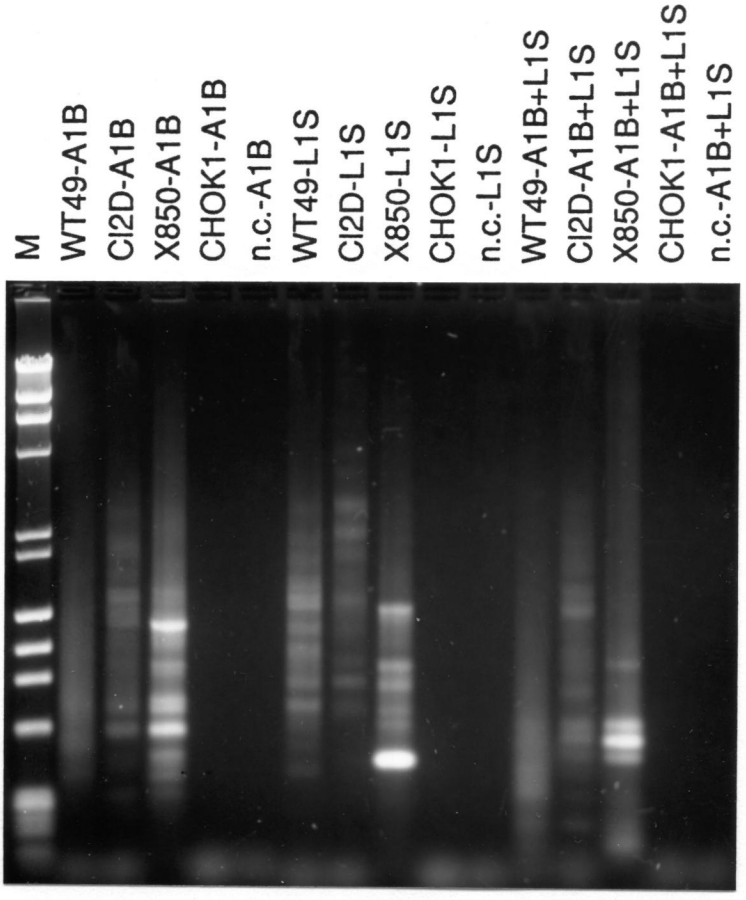


Figure 7. Human repeat element-mediated PCR products. Amplification products were fractionated on a 1.2% agarose gel and stained with ethidium bromide. Templates and primers are as indicated. The X850 amplification products are those generated after two rounds of PCR amplification. M = *Hin*dIII digested lambda DNA and *Hae*III digested ΦX174 DNA (BRL); WT49 = total human genomic DNA; CHOK1 = total hamster genomic DNA; Cl2D = somatic cell hybrid DNA (human X chromosome only); X850 = Cl2D 850 kb *Sfi*I agarose gel strip; n.c. = negative control (no DNA)

comigrated with X850. To help distinguish between these possibilities, pools of the X850 PCR products from each second reaction were hybridized to Southern blots containing *Eco*R1 digested genomic DNA from the following: WT49, a human cell line control, CHOK1, a hamster cell line control, and four male Xq21 deletion human cell lines (see section 2.7). The deletion cell line DNAs have in common a deletion of the region between *DXS72* and *DXS95* (figure 8). The smallest of the deletions, present in the XL45 cell line, breaks within the *DXS95* locus. PCR products generated from the *DXS95* region, therefore, are expected to be absent from the DNA of CM, XL62, NP, and may or may not be absent from XL45.

The A1B PCR products hybridized to five distinct *Eco*RI fragments in WT49 and all four male Xq21 deletion cell line DNAs (figure 9A). Faint hybridization to two *Eco*RI fragments in hamster DNA was detected. The L1S PCR products hybridized to at least nine *Eco*R1 fragments in all the human DNAs (figure 9B). Hybridization to hamster DNA was also significant; seven *Eco*R1 fragments, most of a different size than those observed in human DNA, were detected. The intensity of hybridization of the L1S PCR products was greater in hamster than human DNA. This result suggests that the X850 cloning source contains a region of the hamster genome rich in L1-like sequences, or at least enriched for L1S oligonucleotide sequences. The A1B/L1S PCR products hybridized to four discrete *Eco*R1 fragments in both WT49 and the four Xq21 deletion cell lines (figure 9C). A 3 kb band was detected in WT49 DNA but not in any of the four male Xq21 deletion cell line DNAs, suggesting that a portion of the A1B/L1S PCR products was derived from the *DXS72* - *DXS95* region. An alternate, but less likely, explanation is that the 3 kb band represents one allele of an *Eco*RI polymorphism with the alternate allele represented by one of the larger bands observed on the autoradiograph (N.B. WT49 DNA is XX).

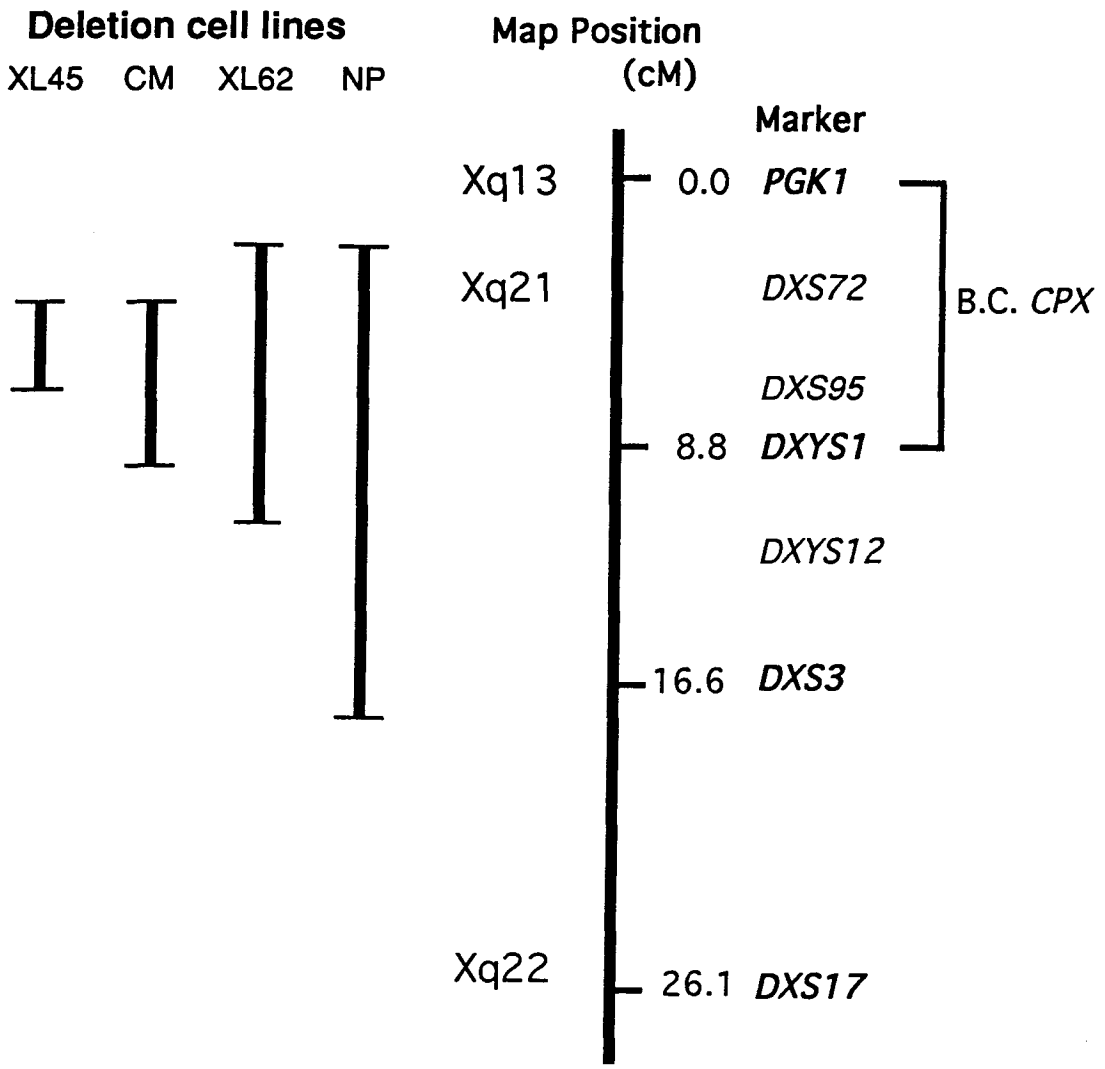
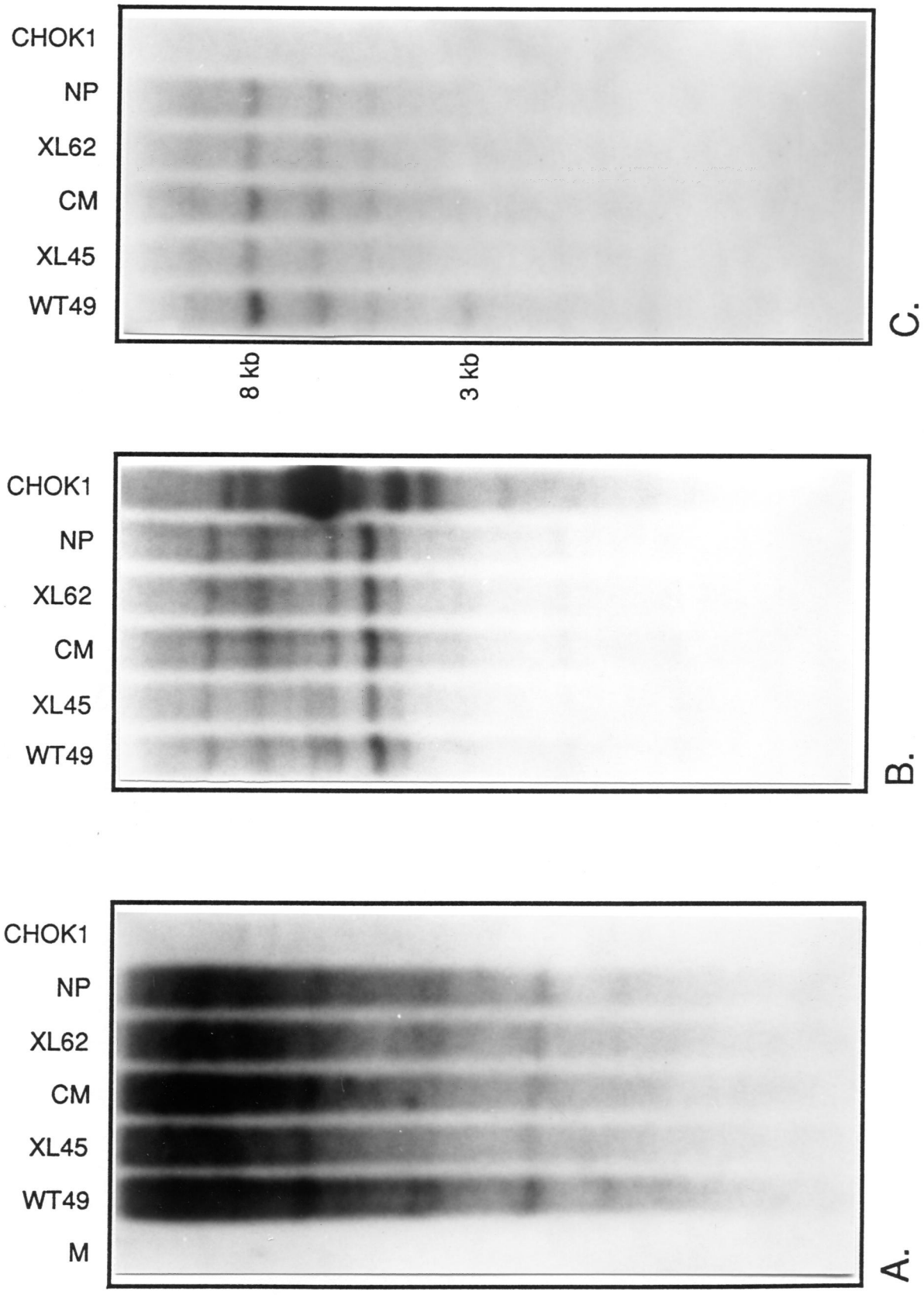


Figure 8. Deletion breakpoints of the four male Xq21 deletion cell lines used in this study (Cremers *et al.* 1989). Deleted regions are represented by vertical lines and approximate breakpoints are indicated by horizontal lines. The distal breakpoint in the XL45 cell line is within the *DXS95* locus. The X chromosome map on the right shows the marker constitution of the deletion cell lines. Genetic distances (in cM) are from Keats et al. 1990. The B.C. *CPX* candidate region is indicated.

Figure 9. Mapping human repeat element-mediated PCR products by Southern blot hybridization: localization to the *DXS72* - *DXS95* region. Pools of the X850 PCR products generated using the A1B, L1S, and A1B/L1S primers (panels A, B, and C, respectively) were used to probe Southern blots containing *Eco*RI digested genomic DNA from WT49 (human cell line), XL45, CM, XL62, NP (male Xq21 deletion cell lines), and CHOK1 (hamster cell line). One fragment detected by the A1B/L1S PCR products appears to be from the *DXS72* - *DXS95* region as a 3 kb band is detected in WT49 but not in any of the deletion cell lines (panel C). Lane marked M contains *Hin*dIII digested lambda DNA and *Hae*III digested ΦX174 DNA (BRL) as size standards.

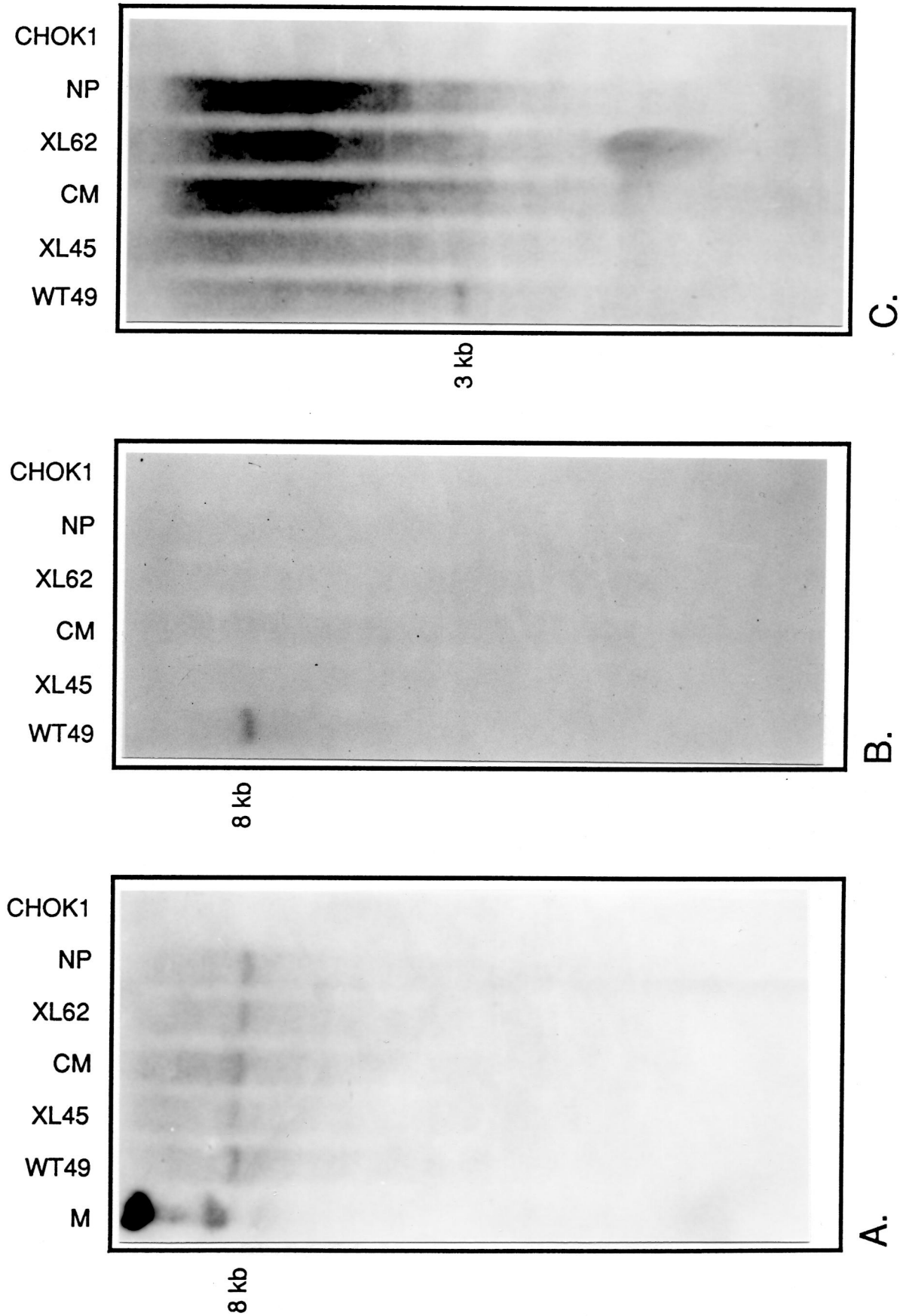

CHOK1
NP
XL62
CM
XL45
WT49
8 kb
3 kb
C.
CHOK1
NP
XL62
CM
XL45
WT49
B.
CHOK1
NP
XL62
CM
XL45
WT49
M
A.

3.2.4. Cloning and mapping of A1B/L1S PCR products

The A1B/L1S products were restriction digested with *Sal*I and *Bam*HI and cloned into pUC18. Seven different clones were identified from among thirty which were analyzed by PCR amplification of inserts, gel electrophoresis, and Southern blot analysis (see section 2.10). Sizes of the amplified inserts ranged from approximately 0.15 kb to 0.7 kb, similar to those seen in figure 7. The PCR amplification products from three randomly chosen clones, X850A/L-4 (0.5 kb insert), X850A/L-7 (0.55 kb insert), and X850A/L-26 (0.4 kb insert), were individually hybridized to the Southern blots containing the Xq21 deletion cell line DNAs. Clone X850A/L-4 detected an *Eco*R1 fragment of 8 kb in WT49 and all four Xq21 deletion cell lines, but not in CHOK1 (figure 10A). Clone X850A/L-4, thus, is derived from a human X chromosome *Sfi*I fragment located outside of the *DXS72-DXS95* region. The second clone, X850A/L-7 also detected an 8 kb *Eco*R1 fragment, but it was present in WT49 DNA and not present in any of the Xq21 deletion cell line DNAs (figure 10B). Reexamination of the PCR pool preliminary mapping autoradiograph (figure 9C) revealed a band of this size with increased signal intensity in WT49 compared to the Xq21 deletion cell lines. The third clone, X850A/L-26 detected the 3 kb *Eco*R1 fragment, present only in WT49, that was observed in the preliminary mapping (figure 10C). X850A/L-7 and X850A/L-26, therefore, are both human derived clones which map to the *DXS72-DXS95* region, and demonstrate the human specific regional enrichment of the combined preparative FIGE and human repeat element-mediated PCR methods.

To determine if X850A/L-7 and X850A/L-26 originated from an X chromosome 850 kb *Sfi*I fragment, each was hybridized against a Southern blot containing *Sfi*I digested genomic DNA from Cl2D and five different male controls (*Sfi*I test blot; hybridization results not shown). For accurate size comparisons, the *DXS95* probe was first hybridized to the *Sfi*I test blot so that the resultant autoradiograph could be

Figure 10. Mapping individual A1B/L1S PCR products by Southern blot hybridization. Inserts from three of the A1B/L1S PCR product clones were hybridized to the same blots as shown in fig. 9. The resultant autoradiographs are as follows: A. clone X850A/L-4, B. clone X850A/L-7, and C. clone X850A/L-26. Clones X850A/L-7 and X850A/L-26 map to the *DXS72* - *DXS95* region.


CHOK1
NP
XL62
CM
XL45
WT49
3 kb
C.
CHOK1
NP
XL62
CM
XL45
WT49
8 kb
B.
CHOK1
NP
XL62
CM
XL45
WT49
M
8 kb
A.

overlayed with those from X850A/L-7 and X850A/L-26. The *DXS95* probe was found to detect an *Sfi*I fragment of approximately 785 kb in all samples and an additional 430 kb fragment in two of the human cell line samples. The size difference (i.e. 785 kb instead of 850 kb) can be attributed to differences in agarose, buffer, and DNA concentration. The altered *Sfi*I restriction pattern in the two individuals is believed to represent a normal variant. The clone X850A/L-7 hybridized to 320 kb and 650 kb *Sfi*I fragments. X850A/L-26 hybridized to a 650 kb fragment and also hybridized weakly to a 710 kb *Sfi*I fragment. X850A/L-7 was then used to probe the Southern blot used to identify X850. The greatest hybridization signal was detected in two agarose gel strips containing the majority of DNA in the 650 kb size range, and a faint hybridization signal was detected in the same size range in the X850 gel strip. It appears that X850A/L-7 was derived from a 650 kb partial *Sfi*I digestion product that comigrated in the X850 cloning source. X850A/L-26 is likely derived from a 650 kb *Sfi*I fragment or 710 kb partial *Sfi*I digestion fragment that also comigrated in the X850 cloning source. Because X850A/L-7 and X850A/L-26 produced distinct nybridization patterns on the same *Sfi*I test blot, the clones must be derived from different *Sfi*I fragments.

A 625 kb *Sfi*I fragment containing the polymorphic marker *DXS540* (table 1) is located within the *DXS72* - *DXS95* region (Cremers et al. 1990). To test the possibility whether either X850A/L-7 or X850A/L-26 could be derived from the *DXS540 Sfi*I fragment, the probe (pZ11c) for *DXS540* was hybridized to the *Sfi*I test blot. A single *Sfi*I fragment of 600 kb was detected, indicating that neither X850A/L-7 nor X850A/L-26 is located on the same *Sfi*I fragment as *DXS540*.

3.2.5 Physical mapping of DXS326

The deletion cell lines used for mapping the A1B/L1S PCR products also proved useful in refining the location of an existing marker, *DXS326*. The *DXS326*

locus was partially informative in the B.C. family and nonrecombinant with respect to *CPX*. *DXS326* was known to be located between *DXS447* and *DXYS1* (D. Barker, personal communication), but had not been mapped relative to the other markers in this interval. The probe for *DXS326* was hybridized to the Southern blot that includes DNA from the male Xq21 deletion cell lines. No bands were detected in any of the four deletion cell line DNAs, indicating that *DXS326* lies between *DXS72* and *DXS95*, or at *DXS95*. When the *DXS326* probe was subsequently hybridized to the *Sfi*I test blot, a single band size of 320 kb was observed. Since *DXS326* does not lie on the same *Sfi*I fragment as does *DXS95*, it can be concluded that *DXS326* maps between *DXS72* and *DXS95*. This result served to refine the crossover position in individual IV-10 (fig.3).

3.2.6 Screening for polymorphisms with new markers from Xq21

To facilitate the genetic mapping of *CPX* as well as the mapping efforts of others working in the Xq21 region, the X850A/L-7 and X850A/L-26 clones were screened to determine if they detect polymorphisms. To screen for conventional RFLPs, the clones were hybridized to Southern blots of relevant B.C. family members' DNAs digested with the following restriction enzymes: *Bgl*II, *Eco*RI, *Hin*dIII, *Msp*I, *Pst*I, *Pvu*II, and *Taq*I. Clone X850A/L-26 did not detect any RFLPs. A *Taq*I RFLP was detected, however, with X850A/L-7 (K. Adams, personal communication). The RFLP screening filter indicated that the *Taq*I polymorphism was informative in some of the B.C. family members, including individual III-7. The RFLP, therefore, had the potential to localize the crossover observed in individual IV-10. Like *DXS326*, the X850A/L-7 clone did not recombine with *CPX* (fig. 11) and thus localized the crossover distal to X850A/L-7. The relative order of X850A/L-7 and *DXS326* is not known.

The *Taq*I RFLP detected by X850A/L-7 is dimorphic with allele sizes of 5.0 kb (allele 1) and 4.7 kb (allele 2). DNA samples from unrelated Native individuals were

not available. DNA samples from 12 unrelated Caucasian individuals (8 females, 4 males; i.e. 20 different X chromosomes) were digested with *Taq*I, Southern blotted (H. Jenkins, personal communication) and probed with X850A/L-7. Frequencies of 0.30 and 0.70 were observed for allele 1 and allele 2, respectively. The heterozygosity value (H) for X850A/L-7 is thus 0.42 ($H = 1 - \sum p_i^2$ where p_i is the population frequency of the ith allele). The PIC value (Botstein et al. 1980) for X850A/L-7 is equivalent to the heterozygosity value since, for X-linked loci, a meiotic event in a heterozygous female will always be informative (see section 1.6.1.2).

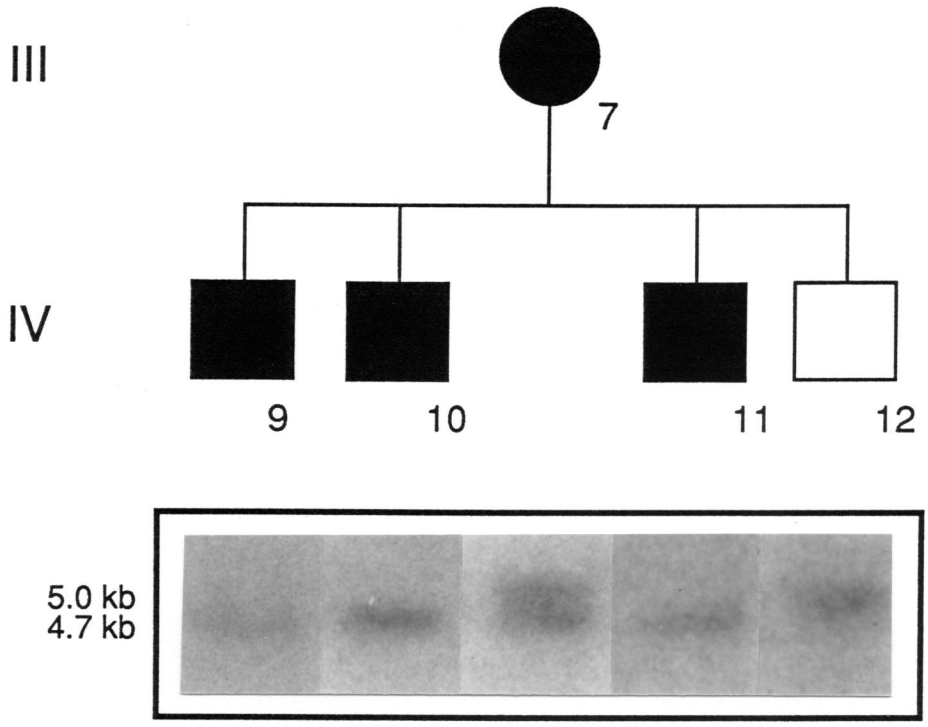


Figure 11. X850A/L-7 *Taq*I polymorphism. Segregation of the 5.0 kb and 4.7 kb alleles of X850A/L-7 in a portion of the B.C. *CPX* kindred (III-7, IV-9 to IV-12 in fig.1) is illustrated. No recombination was detected between *CPX* and X850A/L-7. Symbols are as in fig.1.

4. DISCUSSION

4.1 Regional localization of CPX

The human X chromosome is estimated to consist of 150 to 200 Mb of DNA containing 2500 genes (Keats et al. 1989; Caskey and Rossiter 1992). The gene map of the X chromosome is relatively dense, with at least 160 genes assigned to it on the basis of X-linked inheritance (McKusick et al. 1990). DNA marker linkage studies have been successful in regionally localizing a large number of these X-linked genes (Mandel et al. 1989; Davies et al. 1991). This thesis describes the localization, by linkage analyses, of an X-linked locus responsible for cleft palate and/or ankyloglossia (*CPX*) in a B.C. Native kindred. Both two-point and multipoint linkage analyses resulted in localization of B.C. *CPX* to the Xq13.3-q21.31 region between *PGK1* and *DXYS1*. The intervals proximal to *PGK1* and distal to *DXYS1* were excluded as candidate locations for *CPX* with odds of at least 100:1 by multipoint analyses. Linkage between *CPX* and markers in the Xq13-q21.3 region is supported by high lod score values. The designation of *PGK1* and *DXYS1* as flanking markers is dependent on critical recombination events.

The crossovers localizing B.C. *CPX* proximal to *DXYS1* and distal to *PGK1* were detected in an affected male (III-10) and an unaffected male (IV-6), respectively. Individual III-10 is affected, having ankyloglossia and a short bifid uvula typical of those seen in other affected members of the B.C. family. Individual IV-6 shows no indications of being affected; there is no notch in the posterior edge of the hard palate, the soft palate appears intact, and the uvula and tongue appear normal. Penetrance in the male members of the B.C. family was assumed to be 1.0 for the linkage analyses. In no instance was the complete CPX-associated haplotype observed in an unaffected male. Nonpenetrance has been reported in three obligate affected males, one in each of the three previously reported families with X-linked cleft palate and

ankyloglossia (Rollnick and Kaye 1986; Hall 1987; Bjornsson et al. 1989). However, two of these males were unexamined and nonpenetrance was assumed on the basis of past written records or from interviews with family members (Hall 1987; Bjornsson et al. 1989). Experience with the B.C. family indicates that these are unreliable diagnostic methods due to the sometimes subtle nature of the *CPX* defect. The third obligate affected male was examined and no palatal anomaly was detected (Rollnick and Kaye 1986). It was not reported whether he has ankyloglossia. A total of over 100 males were clinically examined in the four reported families with cleft palate and ankyloglossia (Lowry 1970; Rollnick and Kaye 1986; Hall 1987; Bjornsson et al. 1989; Gorski et al. 1992); there are no other reports of putative nonexpressing obligate affected males and no reports of unaffected fathers with affected or carrier daughters. Given these data, it appears unlikely that individual IV-6 in the B.C. kindred is a nonexpressing carrier of *CPX*.

Other values assumed for the linkage analyses, besides penetrance, were the *CPX* mutant allele frequency and the DNA marker allele frequencies. The *CPX* mutant allele frequency was set at 0.0005 for the linkage analyses reported in this thesis. Lowry (1979) found the frequency of submucous cleft palate in the B.C. Native population to be 0.003. This value (0.003) represents a possible upper limit for the frequency of the *CPX* mutant allele in the B.C. Native population (no ankyloglossia was noted), and is probably an over-estimate as the affected individuals likely present with forms of submucous cleft palate other than CPX. Linkage analyses carried out with the *CPX* mutant allele frequency set at 0.003 resulted in lod score values very close to those reported in table 3; the greatest difference was a decrease in Z_{max} of 0.02 (data not shown). Similarly, calculations at a range of marker allele frequencies made little difference to the lod score results. The marker allele frequencies used for the analyses in this thesis were determined primarily in Caucasian populations (table 2) but could be different in the B.C. Native population (eg. by founder effects and

inbreeding). However, the structure of the B.C. pedigree is such that very few genotypes were inferred for the linkage analyses: the sibships are large with many males; there are few deceased individuals; disease status is known in all cases but one (fig.1).

The B.C. *CPX* localization is inconsistent with that proposed for Icelandic *CPX*. The initial report of the localization of *CPX* in the Icelandic kindred included clinical examinations of 182 individuals from four generations. Blood was collected from 82 individuals, including nine males with cleft palate and ankyloglossia, ten females with ankyloglossia alone, one female with cleft palate alone and one male with a high-vaulted palate (Moore et al. 1987). Clefts ranged in severity from hard palate clefts to bifid uvula (Bjornsson et al. 1989). Individuals with any form of cleft and/or ankyloglossia were considered affected. Linkage data was contributed primarily by eight meioses and indicated linkage between *CPX* and *DXYS1* with a Z_{max} of 3.07 at $\theta = 0.0$ (Moore et al. 1987). Further analyses with additional distal markers identified four recombinant individuals, two each delineated by the probes for *DXYS12* (proximal flanking marker) and *DXS17* (distal flanking marker) (see fig.4 for marker positions) (Ivens et al. 1988; Stanier et al. 1991). Both recombination events placing Icelandic *CPX* proximal to *DXS17* were detected in affected males and recombination events placing *CPX* distal to *DXYS12* were detected in an affected female and a carrier/affected (i.e. not reported) female (Ivens et al. 1988; Stanier et al. 1991). A multipoint lod score of 4.1 was obtained for Icelandic *CPX* mapping between *DXYS12* and *DXS17* (Ivens et al. 1988).

In the B.C. kindred, two recombination events observed in affected males place B.C. *CPX* proximal to *DXYS12* (fig.3). One crossover event was between *CPX* and all informative markers distal to, and including, *DXYS1* (detected in individual III-10 discussed above). The second crossover event occurred between *CPX* and all informative markers distal to, and including, *DXYS12*. In the second individual (IV-

10), *DXYS1* was uninformative. Individual IV-10 was diagnosed as having a submucous cleft palate and ankyloglossia. In the B.C. kindred multipoint analysis, the intervals distal to *DXYS12* were excluded with odds of greater than 1,000:1.

There are several possible explanations for the apparent inconsistency in the localization of the *CPX* gene(s) in the Icelandic and B.C. kindreds. First, CPX in the two families may be due to mutations in different genes. Genetic heterogeneity seems unlikely, however, in light of the similar phenotypes (see section 4.4) and the proximity of B.C. *CPX* and Icelandic *CPX* on the X chromosome. Second, it is possible that blood or DNA sample mix-up occurred in either family and was not detected. A third possibility is that DNA marker genotypes were assigned incorrectly. A fourth explanation for the inconsistent localization of B.C. and Icelandic *CPX* is that disease status was assigned incorrectly in one or more individuals. Both *DXYS12* recombinant individuals in the B.C. family are affected sons of obligate carrier females. The *DXYS12* recombinant individuals in the Icelandic kindred are both daughters of carrier females. Detailed clinical descriptions and pedigree information were not reported for these recombinant Icelandic females. One is known to be affected, but it is uncertain whether she has ankyloglossia or is the female with cleft palate only (and what the nature of that cleft is). It is possible that one or more of the affected *DXYS12* recombinant individuals in the B.C. and Icelandic kindreds are phenocopies. For example, an individual in the Icelandic family with a cleft of the hard palate could have the common multifactorial form of cleft palate. Isolated submucous cleft palate and isolated ankyloglossia are also both common birth defects (Weatherley-White et al. 1972; Lowry 1979; Warden 1991). Since there are two *DXYS12* recombinants in both the B.C. and Icelandic families, it is likely that the apparent inconsistency in *CPX* regional assignments is the result of more than one of the above proposals.

4.2 The Xq13.3-q21.31 region of the human genome

Linkage analyses defined the B.C. *CPX* candidate region as the interval between the polymorphic loci *PGK1* and *DXYS1* in the Xq13.3-q21.31 region. It is estimated that *PGK1* and *DXYS1* are separated by 8.8 Mb of DNA (Keats et al. 1990). The *PGK1* marker is from the gene for phosphoglycerate kinase (PGK), an enzyme involved in the glycolytic pathway. *DXYS1* is an anonymous DNA marker, the *XY* indicating that it detects homologous sequences on the Y chromosome. The *PGK1* - *DXYS1* region is also the candidate region, at least in part, for a number of other loci responsible for human genetic disorders. These loci include the gene(s) for 1) X-linked dominant Charcot-Marie-Tooth (CMT) disease, a progressive motor and sensory neuropathy (Ionasescu et al. 1992), 2) a rare X-linked form of dystonia-parkinsonism syndrome (XPD) (Kupke et al. 1992), 3) Aland Island eye disease (AIED) (Alitalo et al. 1991), 4) X-linked severe combined immunodeficiency (SCIDX1) (Puck et al. 1991), 5) Allan-Herndon Syndrome (AHS) type of mental retardation (Schwartz et al. 1990), 6) nonsyndromic deafness (Reardon et al. 1991), and 7) Simpson-Golabi-Behmel (SGB) syndrome, an overgrowth dysplasia syndrome (Hughes-Benzie et al. 1992).

The localization of *SGB* in a Dutch-Canadian family to the Xqcen-Xq21.3 region is of particular interest because submucous cleft palate and ankyloglossia are found in some SGB pedigrees (Hughes-Benzie et al. 1992). Other features of SGB syndrome include coarse face, macrosomia, visceromegaly, renal dysplasia, hernias, midline groove of lower lip, grooved tongue, congenital heart defects and skeletal abnormalities. Heterozygous females show partial expression of *SGB* which, in the Dutch-Canadian family, includes ankyloglossia. Also in that family is one SGB male with a unilateral cleft lip and palate. Because of its complex phenotype, it is possible that SGB syndrome is due to a chromosomal rearrangement involving several genes, one of which might be *CPX*. However, recent linkage analyses including a second

family (British) suggest that the SGB locus may map distal to the B.C. *CPX* candidate region (A. Mackenzie, personal communication).

The only known gene that maps between *PGK1* and *DXYS1* and has been cloned (Cremers et al. 1990) is that for choroideremia (CHM) also called tapeto-choroideal dystrophy (TCD). Choroideremia is an eye disorder characterized by progressive dystrophy of the choroid, retinal pigment epithelium, and retina in affected males. The locus for choroideremia was mapped to the Xq13-q22 region by linkage analyses and the identification, by cytogenetic analyses, of affected individuals with deletions and translocations in Xq21 (reviewed in Cremers et al. 1989 and Merry et al. 1989). Clinical symptoms associated with the Xq21 deletions can include choroideremia, mental retardation, and deafness. Female carriers of the deletions are asymptomatic, with the exception of retinal changes characteristic of the choroideremia carrier state (Tabor et al. 1983; Hodgson et al. 1987; Nussbaum et al. 1987; Cremers et al. 1989; Merry et al. 1989; Wells et al. 1991). The only documented Xq21 deletion individual (N.P.) (Tabor et al. 1983) with cleft palate has unilateral cleft lip and palate. Cleft palate in association with cleft lip is believed to be etiologically distinct from isolated cleft palate (Fogh-Anderson 1942; Fraser 1980b) and has not been observed in any of the kindreds with cleft palate and ankyloglossia (Lowry 1970; Rollnick and Kaye 1986; Hall 1987; Bjornsson et al. 1989; Gorski et al. 1992).

There are several possible explanations for the absence of cleft palate in males with Xq21-region deletions. It is possible that the *CPX* mutation is not an amorph (loss-of-function mutation). This suggestion is consistent with the apparent dominant nature of the mutation. The cleft palate and/or ankyloglossia phenotype may be the result of an alteration in the *CPX* product. Another possibility is that *CPX* is located outside the region encompassed by the Xq21 deletions. The distal breakpoint of the largest male viable deletion (individual RvD) is just centromeric to

DXS17, and the proximal breakpoint is between *DXS72* and *PGK1* (Cremers et al. 1989). It would follow, then, that *CPX* is located proximal to *DXS72*.

If *CPX* is a dominant mutation, the CPX phenotype may be due to an increase in the amount or activity of the CPX gene product (i.e. a hypermorph). An extra copy of *CPX*, therefore, might give rise to cleft palate and ankyloglossia. There exist reports of at least five viable males with duplications which, together, encompass the entire *CPX* candidate region (Steinbach et al. 1980; Vejerslev et al. 1985; Schwartz et al. 1986; Cremers et al. 1988; Muscatelli et al. 1992) Anomalies associated with the duplications include growth retardation, psychomotor retardation, cryptorchidism, and hypotonia but not cleft palate or ankyloglossia. One male with a duplication including Xq21-q24 has a high arched palate (Schwartz et al. 1986), a feature noted in the Icelandic *CPX* family (Bjornsson et al. 1989). The absence of CPX, however, suggests that the *CPX* mutation may not be a hypermorph. It is also possible that the observed anomalies in Xq21 duplication individuals are not due to gene dosage but, instead, result from position effects or interruption of coding sequences at the duplication breakpoints.

4.3 Generation of DNA markers from the CPX candidate region

At the present time, one of the steps often necessary in the positional cloning approach is the generation of DNA probes from the candidate region. In this study, efforts were directed to the derivation of polymorphic markers from the *DXS95* region. The *DXS95* locus is located within the *CPX* candidate region, approximately one to two cM proximal to *DXYS1* (F. Cremers, personal communication). If a polymorphic marker identified in the *DXS95* region was recombinant with respect to *CPX* (in III-10 or IV-10, fig.3), it would be the nearest distal flanking marker and thus refine the candidate region. If a *DXS95*-region polymorphic marker was nonrecombinant with respect to *CPX* (in III-10 or IV-10, fig.3), it could further localize two of the previously

identified crossovers and thus more clearly define the genetic limitations of the available B.C. family material. A combined preparative FIGE and human repeat element-mediated PCR strategy was employed with the intention of generating new markers from the *DXS95* locus. It was known beforehand that the *DXS95* probe was located on an approximately 850-950 kb *Sfi*I restriction fragment (Merry et al. 1989; Cremers et al. 1989). Anand et al. (1988) used densitometric scans of ethidium bromide-stained gel photographs to show that the bulk of *Sfi*I DNA fragments from a human X chromosome only hybrid is between 100-500 kb in size and a relatively low amount of DNA is present in the 850 kb size range. This information indicated that a human X chromosome only hybrid (CI2D) 850 kb *Sfi*I agarose gel fragment (X850) isolated previously (by S. Gorski) using FIGE would be enriched for the *DXS95* locus.

The efficacy of deriving a clone from the *DXS95* region is considerably greater using a *DXS95*-containing 850 kb *Sfi*I fragment as a cloning source compared to using the entire X chromosome as a cloning source. Assuming a total length of the human X chromosome of 200 Mb, an 850 kb fragment represents approximately 0.43% (1/233) of its total length. The use of a single *DXS95*-containing 850 kb *Sfi*I fragment rather than the entire X chromosome as a cloning source would thus represent an enrichment of 233-fold. However, in the preparative FIGE method, multiple *Sfi*I fragments of 850 kb as well as comigrating *Sfi*I fragments of other sizes could be contained in the cloning source. For this reason, an enrichment of less than 233-fold would be expected. Also, it is known that not all regions of the human genome can be cloned with equal efficiency. In a study by Anand et al. (1988) in which an X only hybrid 840 kb *Sfi*I fragment isolated using preparative PFGE was used as a cloning source, 14% of the human clones isolated were from the X chromosome 840 kb fragment of interest. This value represents an enrichment of 33-fold in comparison to the frequency of clones expected to be derived from a single X chromosome 840 kb fragment using the entire X chromosome as a cloning source

(14% [frequency of human clones isolated] x 238 [enrichment expected if 100% of human clones were derived from the X chromosome 840 kb fragment of interest] = 33; assumptions being that the X chromosome is 200 Mb in length and that clones are generated randomly along that length). It is reasonable to assume that a similar level of enrichment for clones from the *DXS95* 850 kb *Sfi*I fragment is possible.

Human repeat element-mediated PCR (Ledbetter et al. 1990; Brooks-Wilson et al. 1992) was chosen as a relatively quick method of obtaining human-specific clones from X850. Human specific amplification of hybrid DNA fractionated by pulsed-field gel electrophoresis has been described previously (Burright et al. 1991) but was modified to eliminate the time-consuming DNA recovery (electroelute from agarose, dialyze against TE, sequentially extract with phenol, chloroform, and ether, precipitate and resuspend) step. Rather, an aliquot of the X850 gel slice was used directly in one round of PCR amplification, and then an aliquot of that PCR reaction mixture was used as a template for a second amplification reaction. It is not known whether the two methods result in different distributions or yields of PCR products from a given gel slice.

Because the distribution of repeat sequences in the human genome is region dependent (Bickmore and Sumner 1989), a combination of Alu, L1, and Alu/L1 PCR amplification was used. The cytogenetic location of *DXS95* is bands Xq21.2-21.3 (Davies et al. 1991); band Xq21.2 is an Alu-rich Giemsa (G) -negative band and Xq21.3 is a L1-rich G-positive band.

The identification of relevant clones from X850 was determined by Southern blot analysis using DNA from a human cell line (WT49) and four male cell lines with deletions encompassing *DXS95*. When an enriched cloning source such as X850 is used as a template in human repeat element-mediated PCR, the pattern and number of products generated is highly specific - individual products may not be similarly represented in the human repeat element-mediated PCR products derived from total

human DNA (K. Adams, personal communication). For this reason, total *Eco*R1 digested genomic DNA from WT49 and the deletion cell lines was used for the Southern blot analysis instead of the more standard approach of using repeat element-mediated PCR products from the mapping resource (Brooks-Wilson et al. 1992). The human specificity of the X850 PCR products was tested concomitantly with map position by including a *Eco*R1 digested hamster cell line DNA sample next to the WT49 and Xq21 deletion cell line DNAs on the Southern blot.

The size of the *DXS72* - *DXS95* region is estimated to be approximately 2.6 to 3.6 Mb in size (Keats et al. 1989,1990; F. Cremers, personal communication). A 625 kb *Sfi*I fragment containing part of the *CHM* locus and the polymorphic marker *DXS540* (table 1) is known to be located within the *DXS72* - *DXS95* region (Cremers et al. 1990). To test the possibility whether either X850A/L-7 or X850A/L-26 could be derived from that same *Sfi*I fragment, the probe for *DXS540* was hybridized to the *Sfi*I test blot. A single *Sfi*I fragment of 600 kb was detected, indicating that neither of the X850A/L probes are physically linked to *DXS540* on the same *Sfi*I fragment. My mapping data thus suggest that the region deleted in the XL45 cell line (i.e approximately equivalent to the interval between *DXS72* and *DXS95*) contains different *Sfi*I fragments of 650 kb (partial SfiI digestion fragment containing X850A/L-7), 710 kb (partial SfiI digestion fragment containing X850A/L-26), 600 kb (containing *DXS540*), 320 kb (containing *DXS326*; see section 3.2.5), and 785 kb (containing *DXS95*). The distal breakpoint of the deletion in the XL45 cell line is within the *DXS95* locus but it is not known what fraction of the 785 kb fragment is located within the deletion region. The summed sizes of all the different *Sfi*I fragments listed above is 3065 kb. This value is consistent with the size estimates for the *DXS72* - *DXS95* region.

The Alu/L1 PCR clone mapping results suggest that the X850 cloning source contains comigrating 650 kb and, possibly, 710 kb *Sfi*I fragments. PFGE is very

sensitive to DNA concentration and thus DNA overloading may have resulted in inadequate separation of *Sfi*I fragments. This possibility was further substantiated upon observation of the photograph (not shown) of the ethidium bromide stained agarose gel which was used in Southern blot analysis to identify which single agarose gel strip (i.e. X850) would be the optimal *DXS95* cloning source (section 3.2.1). The photograph indicated that during the second FIGE, the DNA in X850 was further resolved than in the first FIGE (fig. 5). It appeared that at least 50% of the DNA corresponded to a fragment size of 850 kb but there was a smear of additional DNA fragments ranging from 650 kb to 900 kb in size. When X850A/L-7 was used to probe the Southern blot used to identify X850, the agarose gel strips containing DNA predominantly in the 650 kb size range gave the greatest hybridization signal. Faint hybridization was detected in the 650 kb region of X850. These results suggest that the clone X850A/L-7 is derived from a comigrating 650 kb partial *Sfi*I digestion product. Similarly, X850A/L-26 is likely derived from a 650 kb *Sfi*I fragment or the 710 kb partial *Sfi*I digestion fragment.

The Alu/L1 PCR clone mapping results were unexpected in that two of three clones mapped to the *DXS72* - *DXS95* region. Furthermore, the two clones mapped to that region despite not being derived from the size selected *DXS95* 850 kb *Sfi*I fragment. Enrichment for the *DXS95 Sfi*I fragment comparable to that obtained by Anand et al. (1988) for a 840 kb *Sfi*I fragment of interest might still have been achieved but, unlike Anand's method (preparation of a recombinant library), the derivation of clones from X850 was nonrandom. Rather, clone generation was dependent on the distribution of repeat sequences contained in X850 DNA, suggesting that relatively Alu-L1 rich 650 kb and, possibly, 710 kb *Sfi*I fragments are contained therein. It would be inaccurate to make conclusions regarding the possible enrichment for clones from the *DXS95* 850 kb *Sfi*I fragment because of the small sample size of clones that were investigated. However, the apparent failure to

generate Alu/L1 clones from the *DXS95 Sfi*I fragment could be explained by a scarcity of either Alu or L1 elements, or a distribution of Alu and L1 elements not amenable to Alu/L1 PCR. It is possible that either the inter-Alu or inter-L1 PCR product pools are more greatly enriched for clones from the *DXS95* 850 kb *Sfi*I fragment. Despite the apparent failure of clone derivation from *DXS95*, the Xq21 deletion mapping filter allowed relevant clones to be detected. A polymorphic clone derived from anywhere between *DXS72* and *DXS95* still has the potential to refine two of the previously identified distal crossovers (fig. 3).

The X850A/L-7 clone was found to detect a *Taq*I RFLP which was partially informative in the B.C. family and nonrecombinant with respect to *CPX*. This result served to localize the crossover observed in individual IV-10 to the region distal to X850A/L-7. The relative order of X850A/L-7 and *DXS326*, also in the *DXS72* - *DXS95* region and nonrecombinant with *CPX*, is not known. The X850A/L-7 result, however, does illustrate how the generation of polymorphic markers from the candidate region is used to define the limitations of the genetic information obtainable from a given individual.

Besides the availability of genetic material, another limitation of the positional cloning approach is illustrated by individual III-10 and his mother II-3. The carrier female II-3 is uninformative for both X850A/L-7 and *DXS326* and, therefore, the location of the crossover detected in III-10 cannot be further refined. The paucity of highly informative markers in a relevant region is thus a second limitation of the positional cloning approach. It would be useful, at this point, to use X850A/L-7, X850A/L-26, and/or the *DXS326* probe(s) to screen a cosmid library in an effort to expand the loci and to develop a highly informative microsatellite repeat marker system. Other limitations or factors influencing the success of any positional cloning project include the amount of penetrance, heterozygote detection (in the case of

recessive alleles), time of onset of clinical manifestations, diagnosis of clinical manifestations, and genetic heterogeneity (Nora and Fraser 1989) .

4.4 The B.C. CPX phenotype

The cleft palate and/or ankyloglossia phenotype appeared to be fully penetrant in the male members of the B.C. family. In no instance was the *CPX*-associated haplotype observed in an unaffected male, and in no instance is there an unaffected male with an affected or obligate carrier daughter. In carrier females, cleft palate and/or ankyloglossia was 75% penetrant. This value is similar to the 82% penetrance obtained for carrier females in the Icelandic family (Moore et al. 1987).

The expressivity of cleft palate and ankyloglossia is highly variable in both males and females but, generally, is greater in males. These observations are consistent with the phenotypic variability described in the three previously reported X-linked cleft palate and ankyloglossia families. Differences in severity of *CPX* expression could be due to environmental factors, other genetic factors, and/or, in carrier females, random X inactivation.

Bifid or absent uvula was not used as a sufficient indicator of affected status in the B.C. kindred because of the high frequency (10%) of such anomalies in the B.C. Native population (Lowry 1970). However, bifid or absent uvula in the B.C. kindred was observed in 63% of affected males and in 38% of carrier females, compared with 6% of unaffected males and 0% of noncarrier females. The affected individuals with bifid uvula have relatively short uvulas in comparison to the unaffected individual with bifid uvula. Bifid or absent uvula has also been observed and reported in the three other families with X-linked cleft palate and ankyloglossia (Rollnick and Kaye 1986; Hall 1987; Bjornsson et al. 1989).

As suggested by Hall (1987), ankyloglossia may be a useful diagnostic marker for X-linked cleft palate. In the B.C. kindred, ankyloglossia was present in 69% of

affected males and in 75% of carrier females. Ankyloglossia is not usually associated with multifactorial forms of isolated cleft palate. In two studies describing the clinical findings in the oral cavity of neonates, ankyloglossia was found in approximately 2.2% of the 2,758 neonates examined (Jorgenson et al. 1982; Friend et al. 1990). There was only one instance of absent uvula reported. However, both study samples consisted of neonates from well baby nurseries and it is possible that neonates with submucous cleft palate would require special care (Lowry et al. 1973) and thus not be included in the study sample.

At this stage in the present study, insight into the etiology of CPX remains limited to observations of the phenotype. Reduction in the amount of soft palate muscle, separation of soft palate muscle, and/or absent or short bifid uvulas are observed in some affected *CPX* individuals. These findings suggest that the amount of mesenchyme which migrates into the posterior third of the secondary palate may be insufficient and/or that the mesenchyme of the palatal shelves may fail to merge properly. Normal mesenchyme migration and merging are thought to include the action of extracellular matrix (ECM) molecules, growth factors and growth factor receptors (Ferguson 1987,1988). The locus for Stickler syndrome, a Mendelian syndrome associated with cleft palate, has been linked to the gene for type II collagen (i.e. an ECM molecule) on chromosome 12 (Francomano et al. 1987). In the Report of the Second X Chromosome Workshop (Davies and Craig 1991), is mention of the isolation and mapping to Xq21-q22 of a new gene or pseudogene related to the epidermal growth factor family (TDGF3). If the report is true, the TDGF3 locus could be a candidate for *CPX*. A mesenchymal protein essential for epithelial cell organization in a variety of epithelial tissues has been identified recently (Hirai et al. 1992). A molecule of this type might also be involved in palatal epithelial-mesenchymal interactions.

Cleft palate and ankyloglossia may be caused by a mutation in the same gene, mutations in two independent genes, or one may cause the other (eg. ankyloglossia causes cleft palate). It has been suggested that one of the events most subject to error in the development of the human secondary palate is removal of the tongue from between the palatal shelves (Johnston and Bronsky 1991). Failure of tongue removal could prevent or delay palatal shelf elevation and contact (Diewert 1986; Johnston and Bronsky 1991). It is possible, then, that the ankyloglossia observed in the B.C. family causes cleft palate by delaying shelf elevation and thus mesenchymal merging in the posterior soft palate region. Sexual differences in the timing of palatal shelf elevation (one week earlier in males) might account for the increased severity of *CPX* in males; assuming no sexual differences in the timing of tongue development, females would have a greater time period to alleviate or reduce the severity of ankyloglossia before shelf elevation occurred. Ankyloglossia is not observed in all *CPX* affected individuals but may have been present earlier in development at the critical period of shelf elevation. If ankyloglossia is responsible for the cleft palate, then a mutation which prevents degeneration of the cells attaching the tongue to the floor of the mouth might cause CPX. A mutation which results in absent or reduced amounts of an epidermal growth factor (EGF)-like molecule is a good candidate: EGF in mice has been shown to initiate the breakdown of the fusion of the eyelid in new-born mice (Cohen 1962; D. Juriloff, personal communication). A question that remains, however, is why ankyloglossia is not more commonly associated with cleft palate. It is possible that careful examination for microforms of cleft palate has not been conducted in most ankyloglossia cases.

4.5 Conclusions

This study was directed toward the identification of the locus responsible for cleft palate and ankyloglossia in a B.C. Native kindred. The main objectives were

twofold. The first objective was to map the locus to a chromosomal region. The second goal was to refine the regional map position to the maximum resolution possible given the available B.C. family material.

Clinical investigations of the B.C. kindred and update of the B.C. pedigree confirmed that the palatal defect is X-linked as originally proposed by B. Lowry (1970). In addition, the clinical reevaluations revealed that some of the affected males and carrier females have ankyloglossia. This observation indicated that the phenotype in the B.C. kindred is similar to the X-linked cleft palate and ankyloglossia phenotype reported in three other families (Rollnick and Kaye 1986; Hall 1987; Moore et al. 1987). In one of these other families (Icelandic; Moore et al. 1987), the locus responsible for X-linked cleft palate and ankyloglossia (*CPX*) was provisionally assigned to the Xq21.3-q22 region (Moore et al. 1987; Ivens et al. 1988; Mandel et al. 1989) and, therefore, provided a starting point for mapping B.C. *CPX*. Regional localization of B.C. *CPX* and identification of a distal flanking marker were accomplished by linkage analysis using polymorphic markers from the Xq21-q22 region. Linkage analyses were then extended to include markers from the Xq13 region and resulted in the identification of a proximal flanking marker for B.C. *CPX*. Results of both two-point and multipoint linkage analysis support the localization of B.C. *CPX* to the Xq13.3-q21.31 region between the markers *PGK1* and *DXYS1*. The distance between these two B.C. *CPX* flanking markers is approximately 8.8 cM (Keats et al. 1990). The position of B.C. *CPX* is further proximal on the X chromosome long arm than the location proposed for *CPX* based on recombination events detected in the Icelandic kindred. Further studies are required to determine if the apparent inconsistency is due to genetic heterogeneity or procedural error(s).

The second objective of this research was to refine the B.C. *CPX* candidate region by localizing more precisely two crossovers identified during the B.C. kindred linkage analyses. New X chromosome markers were generated and one of these,

X850A/L-7, was found to detect an RFLP and map to the Xq21.1-q21.3 region. An existing polymorphic marker, *DXS326*, was also mapped to the Xq21.1-q21.3 region. The physical mapping of these two polymorphic markers may be useful for other linkage studies involving genes localized to the Xq21 region. In the B.C. kindred linkage studies, both markers were partially informative and nonrecombinant with respect to *CPX*. These results did not narrow the *CPX* candidate region but did further refine the map position of one of the crossovers. Refinement of the position of the second crossover was limited by marker informativeness. Based on the knowledge that the second crossover occurs within the interval between *DXS72* and *DXYS1*, it can be estimated that the size of the candidate region can potentially be decreased by a maximum of 4.6 cM. Still, the remaining candidate region would be at least 4.2 cM; fine mapping of *CPX* in the B.C. kindred is, therefore, limited by the number of meioses available. Further strategies for cloning *CPX* should thus include an expansion of the genetic resources. Alternately, the Human Genome Project may allow future studies to exploit the candidate gene approach for identifying *CPX*. It is possible that the eventual elucidation of the molecular defect underlying cleft palate and ankyloglossia in the B.C. family will further the understanding of at least one of the factors involved in craniofacial development.

4.6 Summary

1. Sixty-three members of the B.C. Native family including sixteen affected males and eight carrier females were examined clinically. An important observation was made - some of the affected males and carrier females present with ankyloglossia. Penetrance of cleft palate and/or ankyloglossia appears to be complete in males and 75% in females. Expressivity of *CPX* is highly variable in both sexes but, generally, is greater in males.

2. Cytogenetic analyses detected no chromosomal abnormalities associated with CPX .

3. Two-point linkage analyses revealed no recombination between *CPX* and the polymorphic loci *DXS447*, *DXS72*, and *DXS326* (Z_{max} = 9.38, 7.74, and 2.27, respectively). A single recombination event between *CPX* and *PGK1* (Z_{max} = 7.63 at θ = 0.03) places *CPX* distal to *PGK1*. A single recombination event between *CPX* and *DXYS1* (Z_{max} = 5.59 at θ = 0.04) places *CPX* proximal to *DXYS1*.

4. Multipoint linkage analysis resulted in a maximum likelihood estimate of location for *CPX* at the position of *DXS447* (S = 49.0). The odds against *CPX* localization are greater than 100:1 in the regions proximal to *PGK1* and distal to *DXYS1*.

5. New markers were generated from a *DXS95*-enriched cloning source by a combined preparative FIGE and human repeat element-mediated PCR method. Seven PCR products were cloned and three were characterized in more detail.

6. Two new markers mapped within the *DXS72* - *DXS95* region. The marker designated X850A/L-7 hybridized to 320 kb and 650 kb *Sfi*I fragments. The clone X850A/L-26 hybridized to 650 kb and 710 kb *Sfi*I fragments.

7. An existing marker, *DXS326*, was mapped between *DXS72* and *DXS95*. *DXS326* hybridizes to a 320 kb *Sfi*I fragment. The fine mapping of *DXS326* allowed refinement of the position of the crossover detected in IV-10 to the region distal to *DXS326*.

8. Clone X850A/L-7 was found to detect a *Taq*I RFLP. Clone X850A/L-7 is partially informative in the B.C. family, nonrecombinant with respect to *CPX*, and thus localizes the crossover detected in IV-10 distal to X850A/L-7.

4.7 Proposals for further research

1. Attempt to physically link X850A/L-7 and X850A/L-26 with existing markers by long range restriction mapping using additional restriction enzymes.

2. Clone and map the PCR products from the A1B and L1S X850 product pools. Compare the amount and distribution of repeat element-mediated PCR products derived from X850, the *DXS95* 850 kb *Sfi*I fragment, and other *Sfi*I fragments.

3. Develop a highly informative marker system for probes X850A/L-7, X850A/L-26 and/or *DXS326* for the purpose of refining the position of the crossover detected in individual III-10.

4. Continue to monitor the *PGK1 - DXYS1* region in humans for the mapping and cloning of any putative *CPX* candidate genes. In addition, continue to monitor the *PGK1 - DXYS1* homologous region in the mouse for the mapping and cloning of any putative *CPX* candidate gene homologues (i.e. comparative mapping).

5. Expand the genetic resources.

a) In the original report of the B.C. Native family, Lowry (1970) presented a pedigree indicating that individual I-2 (of the B.C. *CPX* pedigree, Fig.1) had a sister who was an obligate *CPX* carrier. It might be possible to locate descendants of that carrier female.

b) A large Manitoba Mennonite family with a similar X-linked cleft palate and ankyloglossia phenotype has been identified (C. Greenberg, personal communication). This family presents an opportunity to address the question of *CPX* heterogeneity that was raised by this study.

c) Another Native kindred with an apparently similar X-linked cleft palate and ankyloglossia phenotype has been identified and DNA linkage studies have been carried out (C. Johnson, personal communication). It is possible that this Native kindred may share a common ancestral *CPX* mutation with the B.C. Native kindred. If that is the case, one strategy for potentially narrowing the *CPX* candidate region would be to identify a common core haplotype for DNA polymorphisms in the region.

REFERENCES

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

Anand R, Honeycombe J, Whittaker PA, Elder JK, Southern EM (1988) Clones from an 840-kb fragment containing the 5' region of the DMD locus enriched by pulsed field gel electrophoresis. Genomics 3:177-186

Barnard GA (1949) Statistical inference. J R Statist Soc B11:115-139

Bernard LE, Brooks-Wilson AR, Wood S (1991) Isolation of DNA fragments from a human chromosomal subregion by Alu PCR differential hybridization. Genomics 9:241-246

Bickmore WA, Sumner AT (1989) Mammalian chromosome banding - an expression of genome organization. TIG 5:144-148

Bird AP (1987) CpG islands as gene markers in the vertebrate nucleus. TIG 3:342-347

Bixler D (1987) Letter to the editor: X-linked cleft palate. Am J Med Genet 28:503-505

Bjornsson A, Arnason A, Tippet P (1989) X-linked cleft palate and ankyloglossia in an Icelandic family. Cleft Palate J 26:3-8

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314-331

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

Britten RJ, Baron WF, Stout DB, Davidson EH (1988) Sources and evolution of human Alu repeated sequences. Proc Natl Acad Sci USA 85:4770-4774

Brooks-Wilson AR, Goodfellow PN, Povey S, Nevanlinna HA, de Jong PJ, Goodfellow PJ (1990) Rapid cloning and characterization of new chromosome 10 DNA markers by Alu element-mediated PCR. Genomics 7:614-620

Brooks-Wilson AR, Smailus DE, Weier H-UG, Goodfellow PJ (1992) Human repeat element-mediated PCR: cloning and mapping of chromosome 10 DNA markers. Genomics 13:409-414

Browne DL, Zonana J, Litt M (1991) Dinucleotide repeat polymorphism at the DXYS1X locus. Nucleic Acids Res 19:1721

Burdi AR, Faist K (1967) Morphogenesis of the palate in normal human embryos with special emphasis on the mechanisms involved. Am J Anat 120:149-159

Burdi AR, Silvey RG (1969) Sexual differences in closure of the human palatal shelves. Cleft Palate J 6:1-7

Burright EN, Reyner EL, Gorski JL (1991) Human-specific amplification of radiation hybrid DNA fractionated by pulsed-field gel electrophoresis. Nucleic Acids Res 19:401-402

Calnan J (1954) Submucous cleft palate. Br J Plast Surg 6:264-272

Caskey CT, Rossiter BJF (1992) The Human Genome Project. Purpose and potential. J Pharm Pharmacol 44 (Suppl 1):198-204

Christensen K, Holm NV, Olsen J, Kock K, Fogh-Andersen P (1992) Selection bias in genetic-epidemiological studies of cleft lip and palate. Am J Hum Genet 51:654-659

Cohen S (1962) Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. J of Biol Chem 237:1555-1562

Collins FS (1992) Positional cloning: let's not call it reverse anymore. Nature genetics 1:3-6

Corliss CE (1976) Patten's human embryology: elements of clinical development. McGraw-Hill Book company, New York

Cremers FPM, van de Pol TJR, Wieringa B, Hofker MH, Pearson PL, Pfeiffer RA, Mikkelsen M, Tabor A, Ropers HH (1988) Molecular analysis of male-viable deletions and duplications allows ordering of 52 DNA probes on proximal Xq. Am J Hum Genet 43:452-461

Cremers FPM, van de Pol DJR, Diergaarde PJ, Wieringa B, Nussbaum RL, Schwartz M, Ropers H-H (1989) Physical fine mapping of the choroideremia locus using Xq21 deletions associated with complex syndromes. Genomics 4:41-46

Cremers FPM, van de Pol DJR, van Kerkhoff LPM, Wieringa B, Ropers H-H (1990) Cloning of a gene that is rearranged in patients with choroideraemia. Nature 347:674-677

Davies KE, Mandel J-L, Monaco AP, Nussbaum RL, Willard HF (1991) Report of the committee on the genetic constitution of the X chromosome. Cytogenet Cell Genet 58:853-966

Davies KE, Craig IW (1991) Report of the second X chromosome workshop. Genomics 10:842-846

Diewert VM, Pratt RM (1981) Cortisone-induced cleft palate in A/J mice: failure of palatal shelf contact. Teratology 24:149-162

Diewert VM (1983) A morphometric analysis of craniofacial growth and changes in spatial relations during secondary palatal development in human embryos and fetuses. Am J Anat 167:495-522

Diewert VM (1985) Development of human craniofacial morphology during the late embryonic and early fetal periods. Am J Orthod 88:64-76

Diewert VM (1986) Craniofacial growth during human secondary palate formation and potential relevance of experimental cleft palate observations. J Craniofac Genet Dev Biol (Suppl) 2:267-276

Elston RC, Lange K (1975) The prior probability of autosomal linkage. Ann Hum Genet 38:341-350

Fain PR, Luty JA, Guo Z, Nguyen K, Barker DF, Litt M (1991) Localization of the highly polymorphic microsatellite DXS456 on the genetic linkage map of the human X chromosome. Genomics 11:1155-1157

Farrall M, Holder S (1992) Familial recurrence-pattern analysis of cleft lip with or without cleft palate. Am J Hum Genet 50:270-277

Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity: addendum. Anal Biochem 137:266-267

Ferguson MWJ (1988) Palate development. Development 103 (Suppl):41-60

Ferguson MWJ (1987) Palate development: mechanisms and malformations. Irish J Med Sci 156:309-315

Fitchett JE, Hay ED (1989) Medial edge epithelium transforms to mesenchyme after embryonic palatal shelves fuse. Dev Biol 131:455-474

Fogh-Andersen P (1942) Inheritance of hare lip and cleft palate. Arnold Busck, Copenhagen

Francke U, Ochs HD, de Martinville B, Giacalone J, Lindgren V, Disteche C, Pagon RA, Hofker MG, van Ommen G-JB, Pearson PL, Wedgwood RJ (1985) Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa, and McLeod syndrome. Am J Hum Genet 37:250-267

Francomano CA, Liberfarb RM, Tatsuo H, Maumenee IH, Strecter EA, Meyers DA, Pyeritz RE (1987) The Stickler syndrome: evidence for close linkage to the structural gene for type II collagen. Genomics 1:293-296

Fraser FC (1980a) The William Allan memorial award address: evolution of a palatable multifactorial threshold model. Am J Hum Genet 32:796-813

Fraser FC (1980b) The genetics of cleft lip and palate: yet another look. In: Pratt RM, Christiansen KL (eds) Current research trends in prenatal craniofacial development, Elsevier, New York, pp357-366

Friend GW, Harris EF, Mincer HH, Fong TL, Carruth KR (1990) Oral anomalies in the neonate, by race and gender, in an urban setting. Ped Dent 12:157-161

Goodfellow PN, Davies KE, Ropers HH (1985) Report of the committee on the genetic constitution of the X and Y chromosomes. Cytogenet Cell Genet 40:296-352

Gorski SM, Adams KJ, Birch PH, Friedman JM, Goodfellow PJ (1992) The gene responsible for X-linked cleft palate (CPX) in a British Columbia Native kindred is localized between PGK1 and DXYS1. Am J Hum Genet 50:1129-1136

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

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 8:299-309

Hall BD (1987) Letter to the editor: A further X-linked isolated nonsyndromic cleft palate family with a nonexpressing obligate affected male. Am J Med Genet 26:239-240

Hamada H, Kakunaga T (1982) Potential Z-DNa forming sequences are highly dispersed in the human genome. Nature 298:396-398

Hamada H, Petrino MG, Kakunaga T (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Natl Acad Sci USA 79:6465-6469

Harris EF, Friend GW, Tolley EA (1992) Enhanced prevalence of ankyloglossia with maternal cocaine use. Cleft Palate-Craniofac J 29:72-76

Herrmann BG, Barlow DP, Lehrach H (1987) A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal t complex inversion. Cell 48:813-825

Hirai Y, Takebe K, Takashina M, Kobayashi S, Takeichi M (1992) Epimorphin: a mesenchymal protein essential for epithelial morphogenesis. Cell 69:471-481

Hodgson SV, Robertson ME, Fear CN, Goodship J, Malcolm S, Jay B, Bobrow M, Pembrey ME (1987) Prenatal diagnosis of X-linked choroideremia with mental retardation, associated with a cytologically detectable X-chromosome deletion. Hum Genet 75:286-290

Horton CE, Crawford HH, Adamson JE, Ashbell TS (1969) Tongue-tie. Cleft Palate J 6:8-23

Hughes-Benzie R, Hunter AGW, Allanson JE, Mackenzie AE (1992) Simpson-Golabi-Behmel syndrome associated with renal dysplasia and embryonal tumor: localization of the gene to Xqcen-q21. Am J Med Genet 43:428-435

Ionasescu VV, Trofatter J, Haines JL, Ionasescu R, Searby C (1992) Mapping of the gene for X-linked dominant Charcot-Marie-Tooth neuropathy. Neurology 42:903-908

Ivens A, Moore GE, Chambers J, Arnason A, Jensson O, Bjornsson A, R. Williamson (1988) X-linked cleft palate: the gene is localized between polymorphic DNA markers DXYS12 and DXS17. Hum Genet 78:356-358

Jacobs PA, Hunt PA, Mayer M, Bart RD (1981) Duchenne muscular dystrophy (DMD) in a female with an X/autosome translocation: further evidence that the DMD locus is at Xp21. Am J Hum Genet 33:513-518

Jaffe BF, De Blanc GB (1970) Cleft palate, cleft lip, and cleft uvula in Navajo Indians: incidence and otorhinolaryngologic problems. Cleft palate J 7:300-305

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67-73

Jenkins M, Stady C (1980) Dominant inheritance of cleft of the soft palate. Hum Genet 53:341-342

Johnston MC, Bronsky PT (1991) Animal models for human craniofacial malformations J Craniofac Genet Dev Biol 11:277-291

Jorgenson RJ, Shapiro SD, Salinas CF, Levin LS (1982) Intraoral findings and anomalies in neonates. Pediatrics 69:577-582

Keats B, Ott J, Conneally M (1989) Report of the committee on linkage and gene order. Cytogenet Cell Genet 51:459-502

Keats BJB, Sherman SL, Ott J (1990) Report of the committee on linkage and gene order. Cytogenet Cell Genet 55:387-394

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

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172-175

Kupke KG, Graeber MB, Muller U (1992) Dystonia-Parkinsonism syndrome (XDP) locus: flanking markers in Xq12-q21.1. Am J Hum Genet 50:808-815

Lafrenière RG, Brown CJ, Powers VE, Carrel L, Davies KE, Barker DF, Willard HF (1991) Physical mapping of 60 DNA markers in the p21.1-q21.3 region of the human X chromosome. Genomics 11:352-363

Lathrop GM, Lalouel JM, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81:3443-3446

Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet 37:482-498

Leck I (1984) The geographical distribution of neural tube defects and oral clefts. Brit Med Bull 40:390-395

Ledbetter SA, Nelson DL, Warren ST, Ledbetter DH (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. Genomics 6:475-481

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397-401

Lowry (1970) Sex-linked cleft palate in a British Columbia Indian Family. Pediatrics 46(1):123-128

Lowry RB (1971) X-linked cleft palate. Birth Defects: Original Article Series VII:76-79

Lowry RB, Courtemanche AD, MacDonald C (1973) Submucous cleft palate and the general practitioner. CMA Journal 109:995-997

Lowry RB (1979) Genetic studies of cleft lip and cleft palate in the North American Indians of British Columbia. Thesis, Doctor of Medicine, Queen's University, Belfast

Lowry RB, Thunem NY, Uh SH (1989) Birth prevalence of cleft lip and palate in British Columbia between 1952 and 1986: stability of rates. Can Med Assoc J 140:1167-70

Luty JA, Guo Z, Willard HF, Ledbetter DH, Ledbetter S, Litt M (1990) Five polymorphic microsatellite VNTRs on the human X chromosome. Am J Hum Genet 46:776-783

Mandel JL, Willard HF, Nussbaum RL, Romeo G, Puck JM, Davies KE (1989) Report of the committee on the genetic constitution of the X chromosome. Cytogenet Cell Genet 51:384-437

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, USA

McEnery E, Gaines F (1941) Tongue-tie in infants and children. J Pediatr 18:252-255

McKusick VA, Francomano CA, Antonarakis SE (1990) Mendelian inheritance in man. The Johns Hopkins University Press, Baltimore

McWilliams BJ (1991) Submucous clefts of the palate: how likely are they to be symptomatic? Cleft Palate-Craniofacial J 28:247-249

Melnick M, Bixler D, Shields ED (1980) Etiology of cleft lip and cleft palate. Alan R Liss, Inc, New York

Merry DE, Lesko JG, Sosnoski DM, Lewis RA, Lubinsky M, Trask B, van den Engh G, Collins FS, Nussbaum RL (1989) Choroideremia and deafness with stapes fixation: a contiguous gene deletion syndrome in Xq21. Am J Hum Genet 45:530-540

Molloy CM, van de Pol TJR, Brohet RM, Ropers H-H, Cremers FPM (1992) Three RFLPs for pZ11 (DXS540) in the choroideremia gene at Xq21.2. Nucleic Acids Res 20:1434

Moore GE, Ivens A, Chambers J, Farrall M, Williamson R, Page DC, Bjornsson A, Arnason A, Jensson O (1987) Linkage of an X-chromosome cleft palate gene. Nature 326:91-92

Moore GE, Ivens A, Newton R, Balacs MA, Henderson DJ, Jensson O (1990) X chromosome genes involved in the regulation of facial clefting and spina bifida. Cleft Palate J 27:131-135

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7:277-318

Murray JC, Nishimura DY, Buetow KH, Ardinger HH, Spence MA, Sparkes RS, Falk RE, Falk PM, Gardner RJM, Harkness EM, Glinski LP, Pauli RM, Nakamura Y, Green PP, Schinzel A (1990) Linkage of an autosomal dominant clefting syndrome (Van der Woude) to loci on chromosome 1q. Am J Hum Genet 46:486-491

Muscatelli F, Verna JM, Philip N, Moncla A, Mattei MG, Mattei JF, Fontes M (1992) Physical mapping of an Xq-proximal interstitial duplication in a male. Hum Genet 88:691-694

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616-1622

Natsume N, Miura S, Kawai T (1989) Effect of sexual differences on the development of cleft palate in the human. Plast Reconstr Surg 84:854-855

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 86:6686-6690

Nora JJ, Fraser FC (1989) Medical genetics: principles and practice, 3rd ed. Lea and Febiger, Philadelphia

Nussbaum RL, Lesko JG, Alan Lewis R, Ledbetter SA, Ledbetter DH (1987) Isolation of anonymous DNA sequences from within a submicroscopic X chromosomal deletion in a patient with choroideremia, deafness, and mental retardation. Proc Natl Acad Sci USA 84:6521-6525

Ott J (1974) Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. Am J Hum Genet 26:588-597

Ott J (1991) Analysis of human genetic linkage, rev ed. The Johns Hopkins University Press, Baltimore

Pansky B (1982) Review of medical embryology. Macmillan Publishing Co Inc, New York

Peterson-Falzone SJ (1991) Commentary. Cleft Palate-Craniofacial J 28:250-251

Puck JM, Bailey LC, Conley ME (1991) Update on linkage of X-linked severe combined immunodeficiency (SCIDX1) to loci in Xq13. Human Gene Mapping 11. Cytogenet Cell Genet 58:2082-2083

Reardon W, Middleton-Price HR, Sandkuijl L, Phelps P, Bellman S, Luxon L, Pembrey ME, Malcolm S (1991) A multipedigree linkage study of X-linked deafness: linkage to Xq13-q21 and evidence for genetic heterogeneity. Genomics 11:885-894

Renwick JH (1969) Progress in mapping human autosomes. Br Med Bull 25:65-73

Risch N (1992) Genetic linkage: interpreting lod scores. Science 255:803-804

Roberts L (1990) Whatever happened to the genetic map? Science 247:281-282

Rollnick BR, Kaye CI (1986) Mendelian inheritance of isolated nonsyndromic cleft palate. Am J Med Genet 24:465-473

Rushton AR (1979) Sex-linked inheritance of cleft palate. Hum Genet 48:179-181

Sanders B (1979) Pediatric oral and maxillofacial surgery. The C.V. Mosby Company, St. Louis

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, USA

Sanders, B. (1979) Pediatric oral and maxillofacial surgery. The C.V. Mosby Company, St. Louis

Schwartz CE, Ulmer J, Brown A, Pancoast I, Goodman HO, Stevenson RE (1990) Allan-Herndon Syndrome. II. Linkage to DNA markers in Xq21. Am J Hum Genet 47:454-458

Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37:67-75

Schwartz S, Schwartz MF, Panny SR, Peterson CJ, Waters E, Cohen MM (1986) Inherited X-chromosome inverted tandem duplication in a male traced to a grandparental mitotic error. Am J Hum Genet 38:741-750

Scott AF, Schmeckpeper BJ, Abdelrazik M, Theisen Comey C, O'Hara B, Pratt Rossiter J, Cooley T, Heath P, Smith KD, Margolet L (1987) Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1:113-125

Shprintzen RJ, Schwartz RH, Daniller A, Hoch L (1985) Morphologic significance of bifid uvula. Pediatrics 75:553-561

Sperber GH (1989) Craniofacial Embryology, 4th ed. Butterworths, Cambridge

Stanier PM, Newton R, Ivens A, Arnason A, Bjornsson A, Moore GE (1991) New polymorphisms in the region of X-linked cleft palate and ankyloglossia. Human Gene Mapping 11. Cytogenet Cell Genet 58:2086

Steinbach P, Horstmann W, Scholz W (1980) Tandem duplication dup(X)(q13q22) in a male proband inherited from the mother showing mosaicism of X-inactivation. Hum Genet 54:309-313

Tabor A, Andersen O, Lundsteen C, Niebuhr E, Sardemann H (1983) Interstitial deletion in the "critical region" of the long arm of the X chromosome in a mentally retarded boy and his normal mother. Hum Genet 64:196-199

van de Pol TJR, Cremers FPM, Brohet RM, Wieringa B, Ropers H-H (1990) Derivation of clones from the choroideremia locus by preparative field inversion gel electrophoresis. Nucleic Acids Res 18:725-731

Vejerslev LO, Rix M, Jespersen B (1985) Inherited tandem duplication dup(X) (q131-q212) in a male proband. Clin Genet 27:276-281

Verellen C, Markovic V, DeMeyer R, Freund M, Laterre C, Worton R (1978) Expression of an X-linked recessive disease in a female due to non-random inactivation of the X chromosome. Am J Hum Genet 30:97A

Warden PJ (1991) Ankyloglossia: a review of the literature. Gen Dent 39:252-253

Weatherley-White R, Sakura C, Brenner L, Stewart J, Ott J (1972) Submucous cleft palate. Plast Reconstr Surg 49:297-304

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388-396

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n\cdot(dG\text{-}dT)_n$ polymorphisms. Genomics 7:524-530

Wells S, Mould S, Robins D, Robinson D, Jacobs P (1991) Molecular and cytogenetic analysis of a familial microdeletion of Xq. J Med Genet 28:163-166

Wharton P, Mowrer DE (1992) Prevalence of cleft uvula among school children in kindergarten through grade five. Cleft Palate-Craniofac J 29:10-12

Wicking C, Williamson B (1991) From linked marker to gene. TIG 7:288-293

Williamson R, Bowcock A, Kidd K, Pearson P, Schmidtke J, Ceverha P, Chipperfield M, et al (1991) Report of the DNA committee and catalogues of cloned and mapped genes, markers formatted for PCR and DNA polymorphisms. Cytogenet Cell Genet 58:1190-1832

APPENDIX 1. Cytogenetics reports. High-resolution X chromosome analysis was performed for two members of the B.C. *CPX* family: an affected female (III-7 in fig.1; p. 89) and her affected son (IV-11 in fig.1; p.90) (F. Dill and D. Kalousek, personal communication).

CYTOGENETICS LABORATORY

British Columbia's Children's Hospital
DEPARTMENT OF PATHOLOGY

RUN: 22 APR 1991 1149 — B.C.'S CHILDREN'S HOSPITAL CYTOGENETICS REPORT — PAGE 1 FOR: 22 APR 1991

[redacted] (U:7502758)
9 MAY 1949 F OUT (17 DEC)
DR. JAN M. FRIEDMAN

90-CG02388R COLL: 17/12/90 LOG: 17/12/90 1037

PHYSICIAN: JAN M. FRIEDMAN
COMMENT: cc: Dr.P.Goodfellow
MG#
CLINICAL HISTORY: Son ([redacted]): 90-CG2388
X-linked cleft palate.

SURGICAL PROCEDURE(S): ANONEA
TISSUE(S): CHROMOSOME PAIR 23, BLOOD LYMPHOCYTE

** FINAL CYTOGENETIC DIAGNOSIS **

High resolution cytogenetic analysis of the X chromosomes is normal (> 850 band level).

=> SIGNED OUT BY: D.E. MCFADDEN- 22/4/91

[IYX123 H00150]
[I9X220 IL1100 PP3150 PP3174 PP3188 Z01072]

91472

CYTOGENETICS LABORATORY

British Columbia's Children's Hospital
DEPARTMENT OF PATHOLOGY

RUN: 22 APR 1991 1149 | B.C.'S CHILDREN'S HOSPITAL CYTOGENETICS REPORT | PAGE 1 FOR: 22 APR 1991

[redacted] (U:0036953)
31 MAY 1974 M OUT (17 DEC)
DR. JAN M. FRIEDMAN

90-CG02288R COLL: 17/12/90 LOG: 17/12/90 1034

PHYSICIAN: JAN M. FRIEDMAN
COMMENT: cc: CH-MR,Dr.P.Goodfellow
MG4
CLINICAL HISTORY: X-linked cleft palate.
Mother ([redacted]: 90-CG2289
SURGICAL PROCEDURE(S): *NONE*
TISSUE(S): CHROMOSOME PAIR 23, BLOOD LYMPHOCYTE

** FINAL CYTOGENETIC DIAGNOSIS **

High resolution cytogenetic analysis of the X chromosome is normal (850 band level).

=> SIGNED OUT BY: B.E. MCFADDEN [redacted] 22/4/91

[TYX123 M00150]
[TOX220 IL1100 PP3160 PP3174 PP3188 Z01073]

91472