GENE MAPPING STUDIES OF CHROMOSOME 8 IN MAN

A thesis submitted
to

THE FACULTY OF GRADUATE STUDIES
UNIVERSITY OF MANITOBA

In partial fulfillment of the requirements for the
DEGREE OF MASTER OF SCIENCE

by

EVE G. STRINGHAM, BSc.
DEPARTMENT OF HUMAN GENETICS

GENE MAPPING STUDIES OF CHROMOSOME 8 IN MAN

BY

EVE G. STRINGHAM

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

ACKNOWLEDGEMENTS

First and foremost, I wish to express my sincere gratitude to Dr. P.J. McAlpine for her excellent supervision and for giving me the opportunity to train and study under her guidance. I would also like to thank her for the kindness and understanding she showed me throughout this degree and acknowledge her impeccable taste in her appreciation of opera.

I would also like to thank the members of my committee, M. Lewis, R.A. Woods, and J.L. Hamerton for their helpful suggestions and valuable time.

My heartiest thanks go to the students, staff, and faculty of the Department of Human Genetics for their assistance, encouragement and for many unforgettable moments.

I would especially like to thank the following individuals who have been particularly helpful and supportive throughout this project:

Gwendda Coopland for her tremendous enthusiasm and expert advice on DNA techniques and "other things".

Linda Komarnicki for her perseverance in the face of computer program breakdown and starch gel catastrophe.

My honourary sixth brother, Mike Sutherland for his positive thinking, sympathetic ear and commiserative "baby" talk.

To my family I wish to express my warmest gratitude for all the support they have given me throughout this project and over the years, particularly:

My brothers, John, Michael, Peter, Richard and James for making their little sister feel special and for not treating her just like "one of the boys".

My mother, Gabrielle, for encouraging me to pursue my dreams and for teaching me patience and perseverance to realise them with.

My late faithful hound, Sammy, who was always ready to greet me at the door no matter how late I arrived home from the lab.

I wish to dedicate this thesis to a very special and unique individual, my father Elwood, who unknowingly stirred my interest in genetics and graduate studies. I only hope I can be as dedicated and principled a scientist, teacher, and humanitarian as he has been.

I wish to express my sincere appreciation to all the families who agreed to contribute blood samples for this study. Without their kind cooperation this research would not have been possible.

In addition, I would like to thank the following for providing samples, DNA probes and phenotypic data used in the present study:

1. M. Lewis, H. Kaita, and S. Philipps: Rh lab, University of Manitoba, Winnipeg.
2. P.W. Allerdice: Faculty of Medicine, Memorial University of Newfoundland, St. John's.
3. J. Winter: Department of Pediatrics and Pathology, University of Manitoba, Winnipeg.
4. F. Baas: Department of Experimental Medicine, Academic Medical Centre, Amsterdam, The Netherlands. (TG)
5. P. Venta: University of Michigan, Ann Arbor, USA. (CA2)
6. Celltech Ltd., Slough, U.K. (PLAT)

The work presented in this thesis was supported partially by a grant from SigmaXi, a grant (MT6112) from the Medical Research Council of Canada and funds from the Children's Hospital of Winnipeg Research Foundation.

ABSTRACT

The frequencies of carbonic anhydrase 2 (CA2), plasminogen activator-tissue type (PLAT), and thyroglobulin (TG) alleles estimated for a sample population were not significantly different from those reported in the literature, with the exception of the CA2 allele frequencies. Segregation of CA2 phenotypes in all mating types tested and TG phenotypes in two out of three mating types were consistent for a Mendelian system with two codominant alleles at a single autosomal locus.

Linkage analysis of TG data with MNG, a locus for multinodular goitre, suggested that the disease locus in the family studied was neither linked to nor the result of a mutation in TG. Close linkage between MNG and ABO was excluded while close linkage between MNG and RH or SE seemed unlikely. Positive lods for GPT:MNG suggest these loci could be linked.

Linkage studies of chromosome 8 marker loci in individuals not known to carry structurally rearranged chromosomes indicated that the loci, CA2, PLAT, TG and GPT did not appear to be closely linked. Segregation information for individuals carrying an inv(8)(p23q22) suggested that CA2 is closely linked to the 8q22 breakpoint and that PLAT is probably within measurable mapping distance of the 8p23 breakpoint. All the linkage data from the present study combined with the previously reported mapping data suggested

a gene order of 8pter:PLAT:cen:(GPT):CA2:(GPT): TG:qter.

Linkage analysis of TG and CA2 with loci not assigned to specific chromosomes implied that neither locus is probably closely linked to DO, JK or KEL. CA2 may be linked to F13B, CO or YT but negative lods suggest close linkage of the latter two loci to TG is unlikely.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.000 INTRODUCTION

1.100 The Human Gene Map

Over the last few years the development of the human gene map has become a subject of intense investigation. It has been estimated that there are 50,000 genes present in the human genome yet only about 3500 loci have been identified clearly or provisionally (McKusick,1986). Moreover, only ≈1479 genes, markers, cloned DNA segments and fragile sites have been mapped to specific human chromosomes. Many different approaches including classical linkage analysis, somatic cell hybridization, *in situ* DNA hybridization, comparative mapping, gene dosage and exclusion mapping have contributed to these data (HGM 8, 1985).

Construction of the human gene map has direct benefits for the identification of carrier status and for prenatal diagnosis of genetic disease, as well as for the detection of genetic heterogeneity and phenocopies. In the area of basic science, development of the human gene map may provide an improved understanding of gene interactions, position effects, factors affecting crossing over and recombination, the relation of the physical map to the genetic map and comparative mapping.

Note: All abbreviations not specified in text are described in Appendix.

The present study has employed the approach of classical linkage analysis using polymorphic DNA and enzyme markers in order to construct a genetic map of human chromosome 8.

1.200 Restriction Fragment Length Polymorphisms (RFLPs)

Restriction fragment length polymorphisms (RFLPs) are characterized by heritable differences in sizes of DNA molecules detected after digestion with DNA sequence-specific restriction endonucleases and hybridization to a cloned sequence or DNA probe. These differences in fragment sizes, which are the result of differences in restriction endonuclease recognition sites, are inherited as simple Mendelian codominant markers. The availability of this new class of genetic markers has allowed for the direct detection of inter-individual differences in DNA sequences, thus providing information on genetic variation at a much greater level of sensitivity than that detected at the protein product level. Furthermore, RFLPs can be identified without any requirement for gene expression or knowledge of the function of a particular DNA segment since the detection of genetic variation is at the level of the DNA itself. Theoretically, differences in the sizes of DNA sequences may be detected at any site in the human genome so anonymous fragments can act as markers for any chromosome region.

As the proportion of loci which show allelic differences in the sizes of DNA fragments released by restriction endonucleases appear to be an order of magnitude greater than that predicted from the electrophoretic analysis of proteins (Harris, 1974),loci defined by RFLPs are useful markers for studying linkage relationships in families. It now appears that a set of highly polymorphic loci, spaced out to cover a genome, can be identified without the requirement that these loci be whole genes or cistrons. In this connection, Botstein et al. (1980) suggested at least 150 linked markers, 20 cM apart, would be required to identify a two point linkage group involving any unmapped locus.

RFLP data can be used in conjunction with data for other types of segregating loci such as those coding for enzymes and cell surface antigens as well as with inherited chromosomal rearrangements to map loci. The current study demonstrates the use of RFLPs in combination with polymorphisms analysed by conventional technologies to study the arrangement of human genes on structurally normal and rearranged chromosomes.

1.300 Linkage Analysis

Linkage is the presence of two loci close enough together on a chromosome such that their alleles do not assort independently and should not be confused with synteny, the occurrence of two loci on the same chromosome. All linked loci

are syntenic but the converse is not necessarily true. It follows then that tightly linked loci would not only be syntenic but would also be expected to be mapped to the same region of a chromosome.

Recombination, is the observed result of crossing over at meiosis, the process which re-arranges alleles between homologues. Linkage is measured by the degree of recombination at meiosis between two loci of interest. In general the farther apart genes are on a chromosome, the greater the recombination value expected for them, as the value of the recombination fraction approaches 0.5, that of independent assortment.

Linkage analysis requires collecting data from family studies where the segregation of alleles at specific loci can be clearly observed in relation to segregation at other loci (i.e. families in which matings are informative). In each nuclear family to be considered for linkage analysis at least one parent must be heterozygous at two or more loci of interest. Individuals for which phase is known provide much more information than those for which phase is unknown (Maynard-Smith et al., 1961). In addition, testing of grandchildren and other children is essential for the determination of parental genotypes.

Morton (1955) presented a method to test for linkage between loci using "log odds" or lod scores (Morton, 1955) which combined the probability ratio test of Haldane and Smith (1947)

with the sequential analysis devised by Wald (1947). Morton's method is based upon calculating, for a pair of loci, the likelihood of obtaining the observed data at various values of the recombination fraction with those which would be obtained if allelic segregation followed independent assortment (i.e. theta= 0.50). The total relative probabilities at a particular theta for each nuclear family tested,expressed as a logarithm to the base 10, constitutes a lod score. With this method, families can be sequentially sampled without biasing the data, and the combined log 10 relative probabilities of all tested families easily calculated by summing the lod scores at any value of theta.

When no a priori information is available about the positions of two loci in the human genome, a lod score value of +3 at any value of theta indicates a probability ratio of 1000 to 1 that the two loci are linked. Alternatively, if the total lod score at any specific value of theta ever decreases below -2 (odds against linkage 100:1) linkage may be rejected. Lod score values between -2 and +3 are inconclusive indicating that more data are required before a definite conclusion can be made (Morton, 1955). However, when both loci have been independently mapped to the same chromosome, the prior probability that the two loci may be linked is increased since synteny has already been established. Conneally (1986, personal communication to P.J.McAlpine) has suggested that a lod score of +2 may be adequate for the acceptance of linkage between two syntenic loci.

1.400 Genetic Map of Chromosome 8

At HGM8 (1985) 22 markers were assigned to chromosome 8. Most of these assignments have been made to the physical map using somatic cell hybrids combined with isozyme or Southern analysis. In addition, cytogenetic studies have identified various fragile sites while deletion mapping was used to assign the loci for Langer-Giedion syndrome and Coagulation factor VII regulator to chromosome 8 (HGM8, 1985).

Only one locus, spherocytosis (SPH1) has been assigned to chromosome 8 by classical linkage analysis. In independent studies SPH1 was shown to be linked to a translocation (8;12)(p11;p13) by Kimberling et al. (1975), and to a translocation (3;8)(p21;p11) by Bass et al. (1983). As the 8p11 break point was common to both translocations, SPH1 has been inferred to lie in 8p11.

In spite of these assignments to the physical map, chromosome 8 remains a relatively uncharted chromosome genetically. In the present study, linkage data for three loci, carbonic anhydrase (CA2), thyroglobulin (TG) and plasminogen activator-tissue type (PLAT), which have been mapped to chromosome 8, and one locus glutamic pyruvate transaminase (GPT), which may be on chromosome 8, obtained from individuals carrying an inversion of chromosome 8 were compared to those from individuals

not known to carry structurally rearranged chromosomes. Using the break points of the inversion as additional loci for lod score analysis linkage data were obtained with the intent that this information could allow for the construction of a genetic map of chromosome 8.

1.410 Carbonic Anhydrase Two

Carbonic anhydrase is a zinc metalloenzyme which catalyzes the reversible hydration of carbon dioxide. To date, four genetically distinct carbonic anhydrases have been identified in human tissues. CA1 (formerly CAI) activity has been detected primarily in erythrocytes, CA2 in a wide variety of tissues, CA3 in red skeletal muscle and CA4 in lung. All of these enzymes are monomeric with a molecular weight of approximately 29,000 and, they all display amino acid sequence homology (Venta et al., 1983).

Although electrophoretically detectable variation of the CA1 and CA2 enzymes was established, mapping of these genes by interspecific somatic cell hybrids was unsuccesful because of the inability to detect the gene products in such hybrids (Tashian et al., 1980). The introduction of recombinant DNA techniques however has allowed CA2 to be localized to chromosome eight by Southern analysis of mouse - human somatic cell hybrids with a mouse cDNA probe. In addition, analysis of one of these somatic

cell hybrids (JSR-26C), which was positive for both CA2 DNA sequences and human GSR (glutathione reductase) activity but did not possess an intact chromosome 8, suggests that CA2 may be linked to GSR, which has been localized to 8p211 (Venta et al., 1983).

Using a human genomic probe covering the 5' flanking region and the first two exons of CA2, Lee et al (1985) detected a two allele TaqI RFLP in human genomic DNA. Genomic DNA from unrelated individuals which was digested with TaqI and probed with H25-3.8 displayed invariant fragments of 2.1, 0.4, and 0.3 kb. The variant bands occurred in three patterns: a 5.4 kb band only, a 4.0 kb band together with a 1.4 kb band or 5.4,4.0 and 1.4 kb bands. Family studies indicated that the 5.4 kb fragment was allelic to the 4.0 kb and 1.4 kb bands and hence, the three phenotypes were attributed to homozygous and heterozygous combinations of the CA2*1 (5.4 kb) and CA2*2 (4.0 and 1.4 kb) alleles. Each allele in this system was estimated to have a frequency of 0.5 in a small sample of the Caucasian population. This high frequency of variation may prove useful for linkage studies and further mapping projects as well as for the detection of genetic diseases attributed to alleles at the CA2 locus or at a closely linked locus. CA2 deficiency has been reported to be the molecular defect of the autosomal recessive syndrome of osteopetrosis, which is characterized by renal tubular acidosis and cerebral calcification (Sly et al., 1983). Linkage analysis of the TaqI polymorphism of CA2 with this disease locus may prove informative in genetic counselling for families with affected individuals.

1.420 Plasminogen Activator - tissue type (PLAT)

Plasminogen activators are serine proteases that convert inactive plasminogen to plasmin; plasmin is itself a serine protease which degrades fibrin and resolves blood clots. Two immunologically distinct forms of plasminogen activators exist: tissue-type plasminogen activator, PLAT (formerly t-PA) and urokinase, PLAU (formerly u-PA). Mature PLAT, a single chain glycoprotein of 530 amino acids with a molecular weight of 70,000 (Degen et al., 1986), is synthesized in a wide variety of cell types and tumour cell lines and its expression is regulated in a tissue specific fashion by a wide assortment of hormones and hormone -like agents (Benham et al., 1984). As leukemic cells usually produce both PLAT and PLAU while acute myeloid leukemia resistant to chemotherapy secretes only PLAT (Yang-Feng et al., 1986), it has been postulated that PLAT activity is involved in the proteolysis of normal tissue surrounding malignancies.

An oligonucleotide constructed from the amino acid sequence of a tryptic peptide in PLAT (Benham et al., 1984) was used to screen cDNA libraries from the Bowes melanoma cell line (Rijken and Collen, 1981) to isolate useful probes for the PLAT gene

(Benham et al., 1984).

One of these clones, ptPA-4352, which consisted of a 2 kb BglII fragment of PLAT including all of the coding sequences and half of the 3' untranslated region of the gene, was used in the present study. The cloned nucleotide sequence containing PLAT consists of 32,720 bp from the transcription initiation site to the polyadenylation site divided into 14 exons and 13 introns, 3530 bp of 5' flanking sequences and 344 bp of 3' flanking sequences. Twenty eight copies of repeated Alu sequences, spread throughout the introns, account for 22% of the total nucleotide sequence of the gene (Degen et al., 1986). Clone ptPA-4352 was used to probe the DNA of 18 independently derived human-rodent somatic cell hybrids to localize PLAT to chromosome 8 (Benham et al., 1984). PLAT has been further localized to 8p12 -> q11.2 by in situ hybridization and Southern analysis of human-Chinese hamster somatic cell hybrids (Yang-Feng et al., 1986).

Southern blot analysis of the DNA of 38 unrelated individuals digested with EcoRI and probed with ptPA-4352 revealed a polymorphism of PLAT (Benham et al., 1984). In addition to commonly occurring fragments of 9.0, 3.5 and 1.7 kb, DNAs digested with EcoRI showed either a 2.9 kb fragment, a 2.5 kb fragment or both. The small number of hybridizing bands observed under low stringency conditions suggested that PLAT is probably present as a single copy in the human genome. Mendelian inheritance of the EcoRI polymorphism phenotypes was observed and the frequency of

each allele was estimated to be 0.5 in a sample of 38 Caucasian individuals (Benham et al., 1984).

The localization of PLAT to 8p12 -> q11.2 places this locus in the vicinity of the 8p breakpoint of t(8;9)(p11;q34), a translocation observed in the chromosomes of some patients with myeloproliferative disorders. The q34 break point on chromosome 9 corresponds to the position of ABL, the oncogene involved in the 9;22 translocations observed in patients with chronic myelogenous leukemia. Further elucidation of this 8;9 translocation may therefore be informative. Thus, the high frequency polymorphism may prove useful for the determination of linkage groups and as a marker in the study of PLAT's role in malignancy and other disorders mapped to this region of the genome.

1.430 Thyroglobulin (TG)

Thyroblobulin, with a molecular weight of 660,000, is the glycoprotein precursor of the thyroid hormones, 3,5,3'-triiodothyronine (T3) and thyroxine (T4). Five percent of cases of congenital hypothyroidism have been estimated to be due to inborn errors in the synthesis of thyroglobulin (Gons and deVijlder, unpublished observations cited by Baas et al., 1985); the specific molecular defect in the remaining 95% of cases is unknown. Thyroglobulin is synthesized by the thyroid cells and secreted into the colloid by the exocytosis of granules that also contain thyroid peroxidase. In the colloid, tyrosine residues bound in peptide linkage within the thyroglobulin molecule are iodinated. After the iodination of thyroglobulin, T3 and T4 are

released into the circulation by the digestion of thyroglobulin by lysosomal proteases (Ganong, 1981).

The thyroglobulin gene (TG) in humans was assigned to the q23q24 region of chromosome 8 by in situ hybridization using a rat cDNA probe (Avvedimento et al., 1985) and subsequently localized to q24 by Southern analysis of human-rodent somatic cell hybrids (Baas et al., 1985). Analysis of somatic cell hybrids derived from fusion of a Burkitt's lymphoma cell line carrying a t(8;14)(q24;q32) translocation with a Chinese hamster cell line, indicated that TG is located distal to MYC in 8q24 (Baas et al., 1985).

Characterization of cloned TG has suggested that it is extremely large, approximately 300 kb in length. Baas et al. (1985) screened 220 kb of the 3' end of the gene with 15 restriction enzymes for RFLPs but discovered only a single HindIII polymorphism with an estimated minor allele frequency of 0.02. This variant occurred in both normal and affected members of a family with congenital hypothyroidism due to a defect in thyroglobulin synthesis. Mendelian segregation implied that the low frequency variant was linked to the mutation responsible for the hypothyroidism in this family.

A more limited study ofthe 5' region of TG revealed two high frequency RFLPs (Baas et al., 1985). Probe pCHT16/0.9 detected an EcoRV polymorphism with a minor allele frequency of 0.15 while pCHT16/8.0 detected a TaqI polymorphism with a minor allele frequency of 0.20. These observations are consistent with

data from previous studies which revealed substantial differences in the structure and the coding sequences of TG between the 5' and 3' half of this gene (vanOmmen et al., 1983).

1.440 Glutamic Pyruvate Transaminase

Glutamic pyruvic transaminase (GPT), or alanine aminotransferase, is a dimeric enzyme responsible for the interconversion of L-alanine and alpha-ketoglutarate to pyruvate and L-glutamate. In humans, GPT activity is found in the cytoplasm (soluble) and possibly in the mitochondria (Ziegenbein, 1966) The cytoplasmic form is coded for by an autosomal locus, designated GPT which is expressed in liver and mature erythrocytes (Chen et al., 1972).

Analysis of human red cell lysates after starch gel electrophoresis revealed that the GPT locus was polymorphic (Chen and Giblett, 1971). The three commonly occurring phenotypes:GPT 1, GPT 2, and GPT 2-1, have been attributed to homozygous and heterozygous combinations of the GPT*1 and GPT*2 alleles segregating at a single autosomal locus (Chen and Giblett, 1971). However, in spite of the availability of this polymorphism for mapping studies, the chromosomal assignments for GPT based on the analysis of interspecific somatic cell hybrids have been conflicting.

GPT has been mapped to chromosome 8 from the analysis of somatic cell hybrids formed from the fusion of rat hepatoma and

human fetal liver fibroblasts (Astrin et al., 1981; Kielty et al., 1981). At the same time, Wijnen and Meera Khan (1981) assigned GPT to chromosome 16 from the analysis of Chinese hamster-human lymphocyte somatic cell hybrids. In the latter study of 39 independently derived primary and 30 secondary clones, these authors reported that expression of human GPT was most concordant (96%) with human PGP expression which is determined by a locus assigned to chromosome 16. Since the analysis in all these studies was of GPT gene products, and not the gene itself, it is possible that these groups were not observing products of the same gene.

From exclusion mapping it has been concluded that if GPT is on chromosome 8, it is in the region q13-qter whereas if it is on chromosome 16 it would have to be near the terminus of the short (p) arm (Cook et al., 1981). More recently Chodirker et al. (1986) have excluded the red cell GPT from the distal p13 region of chromosome 16 by starch gel electrophoresis of hemolysates from a girl with an abnormal karyotype, 46,XX,r(16)(p13q24) and her parents. In this particular study the sample from the affected child was determined to have a GPT 2-1 phenotype while her father and mother had GPT 2 and GPT 1 phenotypes respectively indicating normal Mendelian inheritance of GPT alleles from each parent.

1.500 Chromosome Inversions

Inversions occur when parts of chromosomes break, and the detached segments are re-inserted into the same chromosome with their genes in reversed order. If the inverted segment includes the centromere the result is a pericentric inversion;if the centromere is not involved, the inversion is described as being paracentric.

Hook and Hamerton (1977), in a summary evaluation of the information collected from six separate newborn studies, reported that eight balanced inversions and one unbalanced inversion were detected in 56,952 infants. In addition, 102 balanced and 11 unbalanced translocations, and 10 deletions were detected suggesting that in humans inversions are quite rare in comparison to translocations, but occur in frequencies comparable to deletions. It should be noted, however, that two of the six studies (London and New Haven) utilized "conventional" staining techniques (i.e. non-banded chromosomes) so only gross abnormalities in structure were detectable. Some rearrangements, particularly paracentric inversions may not have been recognized. It is interesting to note that neither the London nor the New Haven study detected an inversion although the population sizes of those studies, 1015 and 2177 respectively, were relatively small in comparison to the other studies of this composite survey and the total of 56,952 newborns for all studies. Moreover, these

studies only estimate the frequency of inversions in live births and do not adjust for the number of inversions, particularly of an unbalanced nature, that would not be ascertained because of pregnancy loss. So although inversions appear rare, 0.016% of live births in the six studies analyzed, this figure is probably an underestimate of the actual incidence of this class of chromosomal aberration (Hook and Hamerton, 1977).

Inversions were first shown to suppress crossing over in Drosophila when it was determined that single crossovers within an inverted segment resulted in gametes with duplications or deficiencies of genetic material (Sturtevant, 1926). These chromosomal imbalances were lethal so the frequency of recombinants was lowered in the offspring. Sturtevant and Beadle (1936) noted that reduction in the frequency of recombination was greater for shorter inversions than for longer ones and concluded that the risk of unbalanced recombinant chromosomes in offspring was related to the amount of genetic material involved in the inversion. In a subsequent study of humans, Sujansky et al. (1981) noted that the larger the inversion, the smaller the size of segments likely to be involved in duplication-deficient chromosomes and, therefore, a greater likelihood of viable offspring.

Inversions have played a significant role in the evolution and variability of Drosophila species. Dobzhansky and Spassky (1947) constructed a family tree of strains of Drosophila pseudoobscura based on successive overlapping inversions of the

third chromosome. They found that 16 different arrangements of the third chromosome existed in D. pseudoobscura , with each family being separated from the other by a paracentric inversion.

The X chromosomes of D. ananassae and D. melanogaster differ by a pericentric (which includes the centromere) inversion so the resulting chromosomes are morphologically very different from each other as well as from the original ancestral chromosome from which they were assumed to arise. For example, the X chromosome of D. ananassae is metacentric but a pericentric inversion transforms it into an acrocentric chromosome in D. melanogaster.

Gene displacement through inversions provides an unique opportunity for linkage studies. Genes which are normally far apart on a chromosome and display negative lod scores may demonstrate positive lod scores at small thetas, indicative of tight linkage, as the result of an inversion. Hence analysis of inversions may establish synteny for previously unmapped genes and may lead to the elucidation of the relative positions of loci.

1.510 Pericentric Inversions in Chromosome 8

At least 18 unrelated families in which pericentric inversions of chromosome 8 segregate have been described in the literature. In five of these families, with break points at (p23q11), (p11q24), and (p23q12), no unbalanced recombinant chromosomes were recovered in the offspring (Jacobs et al., 1974; Herva and de la Chapelle, 1976; Bui et al., 1982). In contrast,

break points at (p23q22) were observed in 13 kindreds where an unbalanced chromosome 8 was detected in offspring (Fujimoto et al., 1975; Moedjono and Sparkes, 1980; Peakman et al., 1977; Sujansky et al., 1981). Moreover, the recombinant chromosome in all offspring possessed the same nominal structural rearrangement, rec(8),dupq,inv(8)(p23q22) which was associated with developmental delay, congenital heart disease and unusual facies. Although these 13 families were presumably unrelated, at least seven were of Mexican-American ancestry and had migrated from the same geographic location in Mexico, suggesting to a possible common origin of the inversion in these families.

The occurrence of breakpoints at p23q22 in all cases where unbalanced recombinant 8 chromosomes were found in offspring insinuates that these break points may be critical for the formation of unbalanced recombinants. The distribution of inversion break points appears to be non-random with 8p23 considered to be a "hot-spot" (Yu et al., 1978). All of the inversion break points observed in chromosome 8 have occurred in R-bands (Yu et al., 1978) implying that in addition to the size of the inversion, factors such as the nucleotide sequence and the distribution of chiasmata along the chromosome arms affect formation of unbalanced gametes.

The large pericentric inv(8)(p23q22), segregating in the kindred A3:WPA included in this study, was first described by Allerdice (1974). To date no recombinant chromosomes have been detected in this family (Allerdice, personal communication)

suggesting that the large size of this inversion may facilitate the formation of double cross overs within the inverted segment at meiosis resulting in gametes with balanced chromosomes. Alternatively, the rearrangement may be more complex than the cytogenetic analysis to date has indicated, and may actually contain within it smaller inversions and/or duplications or deletions which suppress crossing over. Greater sensitivity in cytogenetic techniques and gene mapping studies of the A3:WPA inversion chromosome may shed some light on this anomaly.

1.600 Specific Aims

Mapping studies of human chromosome 8 were undertaken with the folllowing objectives in mind:

1. to estimate allele frequencies for the DNA marker loci CA2, TG, and PLAT in the families studied.

2. to determine the patterns of inheritance of CA2, TG and PLAT phenotypes in families

3. to determine if the multinodular goitre segregating in the kindred TG:HAR is due to or linked to a mutation in the thyroglobulin gene

4. to compare the linkage data obtained from individuals not known to carry structurally rearranged chromosomes with those from inversion 8 heterozygotes

5. to construct a genetic map of chromosome 8.

2.000 MATERIALS AND METHODS

2.100 Sample Selection

Samples were drawn from family collections gathered for gene mapping research. A kindred (A3:WPA) in which a pericentric inv(8)(p23q22), which will be referred to as inv(8), was segregating (Allerdice, 1974) was specifically selected for analysis. As well, a family segregating for an autosomal dominant form of multinodular goitre was referred for specific genetic studies of thyroglobulin by Dr. Jeremy Winter of the Department of Pediatrics and Child Health of the University of Manitoba. Other families studied were originally ascertained through variant blood group antigens or an interest in research.

None of these families, with the exception of A3:WPA, were known to possess any cytogenetically detectable rearrangement of chromosome eight. Segregation at a number of polymorphic loci was tested in all families to confirm the reported biological relationships of parents and children.

2.200 Preparation of DNA Samples

DNA was extracted from fresh white blood cell pellets, frozen white blood cell pellets or lymphoblastoid cell cultures.

Lymphoblast cultures were established by incubating approximately 1 million peripheral blood lymphocytes previously isolated on a Histopaque 1077 (Sigma) gradient, with 10^6 transforming units of Epstein Barr virus (Showa University Research Institute for Biomedicine in Florida) and RPMI medium supplemented with 15% fetal bovine serum to which cyclosporin at a concentration of 2 ug/ml was added according to the method of Buchwald (personal communication).

The method of DNA extraction was the same for cell pellets obtained from frozen cells, 5 ml of fresh heparinized blood, or 50 ml of cultured lymphoblasts. Cells were suspended in 1.6 ml of TNE2 buffer (10 mM Tris pH 8.0, 10mM NaCl, 2 mM EDTA) and then lysed for 5 hours to overnight at 60°C by the addition of 0.4 ml of 5X lysis buffer (1X TNE2, 5% sodium dodecyl sulphate, 1 mg/ml Proteinase K). Protein and insoluble contaminants were removed by two extractions with an equal volume of buffered phenol followed by two extractions with an equal volume of chloroform:isoamyl alcohol (24:1) (Maniatis et al., 1982). All extractions were centrifuged at 3200 rpm for 10 m (Damon IEC PR-6000 centrifuge, rotor #259). Subsequently 0.1 volume of 3 M sodium acetate buffer, pH 5.5 was added to the aqueous phase and the DNA was precipitated with one volume of cold 100 % isopropanol. High molecular weight DNA was pelleted by centrifugation at 10,000 rpm for 10 m, washed in 70 % ethanol and dried for 48 h in a vacuum dessicator. Finally the sample was resuspended in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

DNA was quantitated by spectrophotometric measurement of the amount of ultraviolet irradiation absorbed by the bases at wavelengths of 260 nm and 280 nm. The reading at 260 nm determines the concentration of nucleic acids present in the sample where an optical density (OD) of 1.0 is comparable to approximately 50 ug/ml of double stranded DNA. Pure DNA samples have an OD_{260}/OD_{280} ratio of 1.8 while contaminating protein or phenol will lower this ratio considerably. DNA samples were stored at 4°C until analysed.

2.300 Analysis of Samples

2.310 Restriction Endonuclease Digestions

DNA samples, one to three ug, were cleaved with the appropriate restriction endonuclease according to the conditions listed in Table 1. All reactions were terminated by the addition of 0.1 volume of 10X loading buffer (50 mM EDTA, 10% Ficoll, Orange G crystals to saturation), incubated at 65°C for 10 m and then plunged into ice water for 10 m before being loaded onto agarose gels for electrophoresis.

TABLE 1. CONDITIONS FOR DIGESTION OF DNA WITH RESTRICTION ENDONUCLEASES TAQI AND ECORI.

CONDITION VARIABLE	RESTRICTION ENDONUCLEASE	
	TaqI *	EcoRI *
10X Reaction Buffer	50 mM Tris-HCl, pH 8.0 10 mM MgCl 50 mM NaCl	100 mM Tris-HCl, pH 7.5 50 mM NaCl 10 mM MgCl
Reaction temperature	65°C	37°C
Units enzyme/ ug DNA	4 to 5	3
Duration of reaction	5 hours	16 hours

* obtained from Boehringer Mannheim or Pharmacia

2.320 Agarose Gel Electrophoresis

Gels were prepared by pouring 150 ml of 0.8 % agarose(w/v) in TAE buffer (0.04 M Tris, 0.001 M EDTA, brought to a pH of 8.0 with glacial acetic acid) containing 15 ul of 10 mg/ml ethidium bromide over 19.6 cm x 14.4 cm glass plates. While the gel was still liquid a plastic comb was inserted one mm above and 3 to 4 cm from the potential cathodal end of the plate to form sample slots. After the gel had solidified at room temperature the comb was removed and the gel submerged in an electrophoresis tank containing TAE buffer to a level of two to three cm above the top of the gel. Two hundred ng of lambda DNA digested with Hind III endonuclease were loaded into each of the two outer most sample slots on each gel to serve as a molecular weight standard. Digested samples were loaded into the remaining slots. Electrophoresis was carried out at 30 V/cm at room temperature until the orange G had run off the gel (approximately 16 h).

2.330 Southern Blot Protocols

DNA was transferred to nylon membranes ZETAPROBE (Pall Ultrafine Filtration Corporation) or BIODYNE (Bio-Rad), by Southern blotting (Southern, 1975), modified according to the manufacturers' instruction as outlined in Table 2. High molecular weight DNA fragments were nicked by irradiation of the gel with nominally 254 nm ultraviolet light from a transilluminator (Spectroline TR-254) for 30 s before transfer. DNA was transferred onto membranes overnight after which membranes were briefly rinsed in 2 X SSC and dried at room temperature. ZETAPROBE membranes were used without further treatment wheras BIODYNE membranes were baked for two hours at 80°C to bind the DNA permanently.

2.340 Isolation of Probes

Probes for thyroglobulin (TG), carbonic anhydrase 2 (CA2), and tissue specific plasminogen activator (PLAT), genes which have been assigned to chromosome 8, were obtained directly from investigators who had isolated the cloned sequences and are described in Table 3. After verification of the presence of the desired cloned sequence in the received samples according to the method of Rodriguez and Tait (1983), plasmids were isolated and purified by the two step cesium chloride gradient procedure described by Garger et al. (1983) with the following modifications.

TABLE 2. SOUTHERN TRANSFER PROTOCOLS WITH BIODYNE AND ZETAPROBE MEMBRANES.

TREATMENT	BIODYNE	ZETAPROBE
Denaturation	2 washes in 250 ml of 1.5 M NaCl, 0.5 M NaOH for 20 m at 23°C	Not required
Neutralization	2 washes in 250 ml of 3 M NaCl, 0.5 M Tris-HCl, pH 7.0 for 20 m at 23°C	Not required
Preparation of membrane for tranfer	Not required	rinsed briefly in dd H_2O
Transfer solution	20 X SSC (3 M NaCl, 0.3 M Na_3Citrate, pH 7.0)	0.4 N NaOH

Five hundred ml cultures of E.coli, harboring the desired plasmid, were grown overnight at 37°C on a shaker in Luria broth (10 g/l bactotryptone, 5 g/l bacto-yeast extract, 10g/l NaCl, pH 7.5) containing the antibiotics, as required by the antibiotic resistance genes of the plasmids. Cultures were harvested by centrifugation at 6000 rpm (Beckman J2-21 centrifuge, JS-7.5 rotor) for ten minutes followed by a wash of the cell pellet in 100 ml of TE buffer. The washed pellet was resuspended in 20 ml glucose buffer (25 mM Tris-HCl, pH 8.0, 10 mM EDTA, 50mM glucose). Four ml of lysozyme solution (20 mg lysozyme/ml glucose buffer) were added and the mixture incubated at room temperature for 10 m. Subsequently 50 ml of 1 % SDS in 0.2 N NaOH were added to lyse the cells after which the mixture was placed on ice for 5 m. Twenty - five ml of potassium acetate solution, 3M with respect to potassium and 5M with respect to acetate, were added and the mixture placed on ice for 15 m to precipitate protein, chromosomal DNA and RNA which were subsequently removed by centrifugation at 7500 rpm for 20 m (Beckman J2-21 centrifuge, JS-7.5 rotor). The supernatant was extracted once with an equal volume of phenol/chloroform/isoamyl alcohol (phenol:chloroform: isoamyl alcohol;25:24:1) and centrifuged for 10 m at 6000 rpm (Beckman J2-21 centrifuge, JS-7.5 rotor).

TABLE 3. CHARACTERISTICS OF PROBES FOR DNA MARKER LOCI ANALYSED.

	LOCUS		
PROBE	CA2	TG	PLAT
Name	HS25	pCHT16/8	ptPA-4352
Size(kb) of insert	3.8	8.0	2.0
Source of insert	genomicDNA	genomicDNA	cDNA
Vector	pBR325	pAT153	pCT37
Insertion site in vector	EcoRI	EcoRI	BglII
Antibiotic resistance	ampicillin, tetracycline	ampicillin, tetracycline	ampicillin
E. coli host	DH1	1046	not specified
Source	P. Venta	F. Baas	F. Benham

Nucleic acids were subsequently precipitated from the aqueous phase by the addition of one volume of cold isopropanol, pelleted by centrifugation at 10,000 rpm for 15 m (Sorvall RC2-B centrifuge), washed with 70% ethanol and dried before being resuspended in TE buffer. After plasmids were isolated from two step cesium chloride gradients as described by Garger et al. (1983), they were suspended in TE buffer and stored at 4°C until used in experimental protocols.

2.350 Labelling of Probes

DNA probes were labelled by nick translation using ^{32}P dATP and ^{32}P dCTP according to a method modified after Maniatis et al., (1982). Components for this procedure were purchased as a kit (Amersham N5500) or prepared locally. Reaction mixtures consisting of 300 to 600 ng of plasmid, combined with unlabelled dTTP and dGTP, each at a final concentration of luM, 4 to 5 ul of enzyme solution (containing 0.5 ug of DNaseI and 5 units of E. coli DNA Polymerase I) and 50uCi each of ^{32}P dATP and ^{32}P dCTP were incubated for two h at 15°C after which an equal volume of 2X "stop" buffer (bromophenol blue, glycerol, EDTA) was added. Unincorporated nucleotides were then separated from the nick translated DNA by centrifugation of the reaction mixture through 1 ml of Sephadex G-50 saturated with TE pH8.0, 0.1 % SDS in a 1 ml tuberculin syringe at 2300 rpm for 1s (Damon IEC PR-6000 centrifuge, rotor #269). Three hundred ul of formamide and 150,000 to 200,000 cpm

of labelled lambda/HindIII marker were added to the eluate before the probe was denatured (65°C for 10 m). Subsequently, the probe was cooled on ice for 5 m.

2.360 Hybridization of DNA filters

Before hybridization, BIODYNE and ZETAPROBE membranes were incubated in 20 ml of the appropriate prehybridization solution overnight at 42°C in a shaking water bath as described in Table 4. After removal of the prehybridization solution, 10 ml of hybridization solution and the denatured probe were added to the membrane which was then sealed in a plastic bag and incubated for 36 h at 42°C in a shaking water bath.

2.370 Washing of Membranes after Hybridization

After hybridization, background radiation was washed off the filters as described in Table 5. Washes were carried out in 250 ml of the indicated solution for 15 m until the background radiation detected by Geiger counter readings was low.

TABLE 4. HYBRIDIZATION PROTOCOLS WITH BIODYNE AND ZETAPROBE MEMBRANES.

TREATMENT	BIODYNE	ZETAPROBE
Prehybridization solution	50 % formamide 0.5 mg/ml sssDNA 5X SSC 40 mM NaH_2PO_4 8 mM Na_2HPO_4 0.1 % Denhardt solution	50 % formamide 0.5 mg/ml sssDNA 4X SSPE 1 % SDS 0.5 % Blotto
Hybridization solution	50 % formamide 0.1 mg/ml sssDNA 5X SSC 40 mM NaH_2PO_4 8 mM Na_2HPO_4 0.1 % Denhardt solution 10 % dextran sulphate	47 % formamide 0.5 mg/ml sssDNA 3X SSPE 1 % SDS 0.5 % Blotto 10 % dextran sulphate

TABLE 5. POST HYBRIDIZATION WASHES OF BIODYNE AND ZETAPROBE MEMBRANES.

BIODYNE	ZETAPROBE
1. 2 X SSC/ 0.1 % SDS at room temperature	1. 2 X SSC/ 0.1 % SDS at room temperature
2. 2 X SSC/ 0.1 % SDS at 50°C	2. 0.1 X SSC/ 0.1 % SDS at 50°C
3. 0.1 X SSC/ 0.1 % SDS at 50°C	3. 0.1 X SSC/ 1 % SDS at 50°C

2.380 Autoradiography

After the final wash membranes were rinsed briefly in 2X SSC, drained of excess liquid and wrapped in plastic film. The filter was then placed in an X-ray cassette sandwiched between two sheets of Kodak XAR - 5 film and two Cronex intensifying screens and left overnight to four days at -70°C. Films were developed according to manufacturer's specifications.

2.390 Re-probing of Filters

Filters of TaqI digests were initially probed with HS25 to detect the CA2 polymorphism and then stripped and re-probed with pCHT16 to detect TG alleles. Filters were normally stripped by washing three times in 0.1 x SSC/0.1 % SDS at 100°C for 20 m. After the final wash, filters were briefly rinsed in 2X SSC, wrapped in plastic film and exposed to X-ray film for two days to check for complete removal of the original probe from the membrane. Subsequent prehybridization and hybridization of the membrane with pCHT16 were carried out as previously described.

2.400 Detection of GPT Gene Variants

GPT phenotypes were determined after starch gel electrophoresis of washed and packed red blood cells as described by Chen et al., (1972).

2.500 Linkage Analysis

Segregation of CA2, TG, PLAT and the inv(8),(designated 8INV), was tested for linkage in pairwise combination using Morton's lod score analysis (1955) performed by the Mark III program of Côté (1975). Lod scores were separated according to the parental origin of the segregation information and the presence or absence of the inv(8) in the informative parents. Members of families for whom chromosome analyses were not available were considered, for the purpose of this study, to have normal karyotypes as no specific information indicative of a segregating chromosome rearrangement was obtained when these families were ascertained.

2.600 Nomenclature

The genetic notation used in the present study followed the International Systems for Human Gene Nomenclature (HGM8, 1985). Genetic notations used which are not found in HGM8 (1985) are listed in the appendix.

3.000 RESULTS

3.100 Description of Phenotypes

3.110 CA2

The hybridization patterns of human DNA samples restricted with TaqI endonuclease and probed with HS25, a 3.8 kb human genomic clone which detects CA2 sequences, were similar to those described by Lee et al. (1985). The CA2 1 phenotype is characterized by a band of 5.4 kb while an extra TaqI restriction site in this fragment generates the 4.0 and 1.4 kb fragments representative of the CA2 2 phenotype. Heterozygotes (CA2 2-1) possess all three bands. Bands at 2.1, 0.4 and 0.3 kb are common to all phenotypes (Fig.1).

3.120 PLAT

The DNA hybridization patterns of EcoRI digests probed with ptPA4352, a cDNA clone for PLAT showed variant bands at 2.9 and 2.5 kb. PLAT 1 homozygotes possessed the 2.9 kb band, PLAT 2 homozygotes the 2.5 kb band and PLAT 2-1 heterozygotes both. Bands of approximately 9.5, 7, 3.5, and 1 kb were observed in all phenotypes although the 7 kb fragment has not been previously

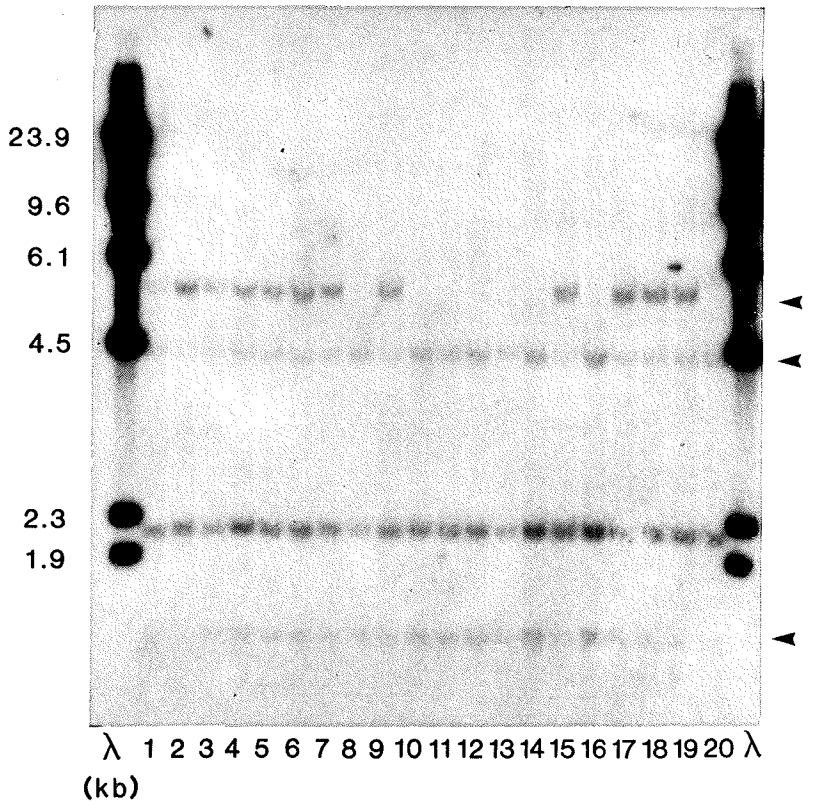

Fig. 1 Photograph showing the Taq1 RFLP phenotypes of CA2

Lanes 1-7, 9, 15 and 17-19: CA2 2-1;

Lanes 8, 10-14, 16 and 20: CA2 2

described in the literature (Fig.2). Additional bands of approximately 4.6 kb in lane 1 and 4.3 kb in lane 2 are presumably due to incomplete digestion of the samples with EcoRI endonuclease.

3.130 TG

Southern blots of TaqI digested DNA samples probed with pCHT16/8.0 which detects TG, demonstrated band patterns similar to those previously described by Baas et al., (1985). TG 1 phenotypes presented with a 5.8 kb fragment while TG 2 phenotypes possessed a 5.2 kb fragment and heterozygotes had both. In addition, fragments of 3 and 1 kb were observed in all phenotypes (Fig.3).

3.200 Estimation of Allele Frequencies for CA2, PLAT and TG

The distributions of CA2, PLAT, and TG phenotypes in a sample of the individuals studied are presented in Table 6. Following the precedence of Neel (1964), the individuals selected for the sample population included all of the parents tested in the present study and any children who were tested without any prior knowledge of parental phenotypes. Allele frequencies for PLAT and TG were estimated from these data (Table 6) and not found to deviate significantly from the allele frequency estimates previously reported in the literature (Benham et al., 1984; Baas et al., 1985), PLAT:X^2 = 0.6, df 1, $p > 0.25$; TG:X^2 = 0.3, df 1, p>

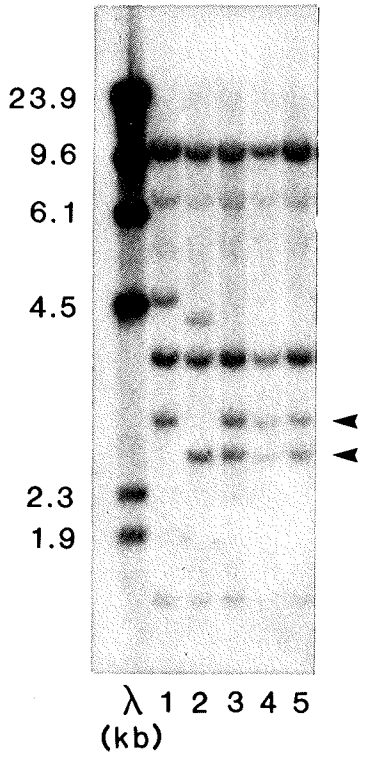

Fig. 2 Photograph showing the EcoRI RFLP phenotypes of PLAT

Lane 1: PLAT 1; Lane 2: PLAT 2; Lanes 3-5: PLAT 2-1

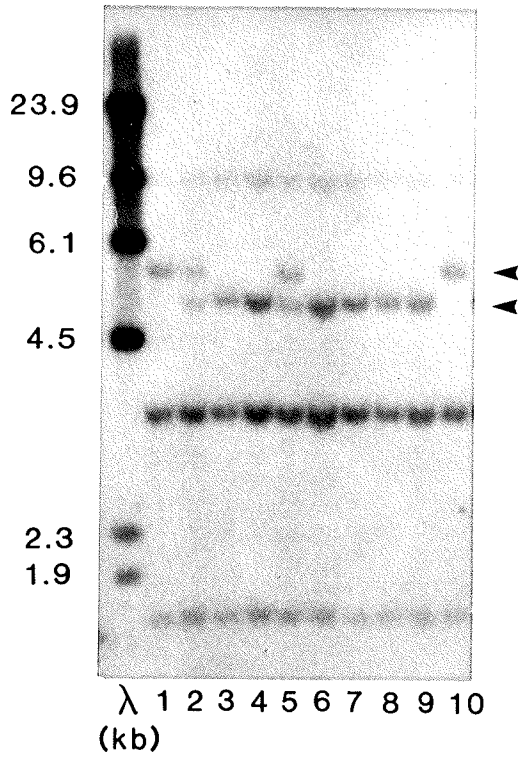

Fig. 3 Photograph showing the Taq1 RFLP phenotypes of TG

Lanes 1 and 11: TG 1; Lanes 2 and 5: TG 2-1;
Lanes 3, 4 and 6-9: TG 2

TABLE 6. THE DISTRIBUTIONS OF CA2, PLAT AND TG PHENOTYPES IN SELECTED SAMPLE POPULATIONS OF THE PRESENT STUDY.

PHENOTYPE	CA2	PLAT	TG
1	3	6	0
2-1	29	6	20
2	22	3	25
TOTAL	54	15	45

ESTIMATED ALLELE FREQUENCIES:

CA2*1: 0.3 ± 0.04	CA2*2: 0.7 ± 0.04
PLAT*1: 0.6 ± 0.09	PLAT*2: 0.4 ± 0.09
TG*1: 0.2 ± 0.04	TG*2: 0.8 ± 0.04

0.25. The estimated allele frequencies for CA2 on the other hand did deviate significantly from those estimated by Lee et al. (1985), $X^2 = 13.4$, df 1, $p < 0.005$.

3.300 Segregation Analysis of DNA Marker Loci

The distribution of CA2 and TG phenotypes in 19 families (18 of which are in common) are presented in Tables 7 and 8. In accordance with a Mendelian system where two codominant alleles segregate at a single autosomal locus, the segregation ratios for CA2 2-1 x CA2 2-1 ($X^2 = 1.435$, df 2, $p > 0.10$), CA2 2-1 x CA2 2 ($X^2 = 3.457$, df 1, $p > 0.05$), TG 2-1 x TG 2-1 ($X^2 = 5.609$, df 2, $p > 0.025$), did not deviate significantly from the expected values. However, the TG 2-1 x TG 2 mating did show a statistically significant variation ($X^2 = 4.829$, df 1, $p < 0.025$).

The segregation of PLAT phenotypes in seven families is presented in Table 9. These data were not analyzed for statistically significant deviation from expected ratios since the expected numbers in all classes were less than five.

TABLE 7. THE DISTRIBUTION OF CA2 PHENOTYPES IN 19 FAMILIES

PARENTAL PHENOTYPES		NUMBER OF FAMILIES	OFFSPRING PHENOTYPES 1	2-1	2	TOTAL
1	1	–	–	–	–	–
1	2-1	1	0	3	0	3
1	2	–	–	–	–	–
2-1	2-1	7	4	11	8	23
2-1	2	8	0	23	12	35
2	2	3	0	0	10	10
TOTAL		19	4	37	30	71

TABLE 8. THE DISTRIBUTION OF TG PHENOTYPES IN 19 FAMILIES

PARENTAL PHENOTYPES		NUMBER OF FAMILIES	OFFSPRING PHENOTYPES 1	2-1	2	TOTAL
1	1	-	-	-	-	-
1	2-1	-	-	-	-	-
1	2	-	-	-	-	-
2-1	2-1	6	2	11	10	23
2-1	2	10	0	11	24	35
2	2	3	0	0	14	14
TOTAL		19	2	22	48	72

TABLE 9. THE DISTRIBUTION OF PLAT PHENOTYPES IN SEVEN FAMILIES.

PARENTAL PHENOTYPES		NUMBER OF FAMILIES	OFFSPRING PHENOTYPES 1	2-1	2	TOTAL
1	1	–	–	–	–	–
1	2-1	4	4	4	0	8
1	2	2	0	5	0	5
2-1	2-1	1	1	3	0	4
2-1	2	–	–	–	–	–
2	2	–	–	–	–	–
TOTAL		7	5	12	0	17

3.400 Thyroglobulin and Linkage Studies with TG:HAR kindred

After DNA samples from members of a family (TG:HAR) in which an autosomal dominant form of multinodular goitre (MNG) occurs were examined for TG size variation, TG segregation data were analyzed for linkage to the goitre disease locus (MNG). A 4:2 phase unknown z score was obtained corresponding to lods of -1.185 and -0.018 at theta values of 0.05 and 0.40 respectively. These results indicated that alleles at TG and the disease locus in this family segregated independently.

Lod scores for MNG versus other marker loci are presented in Table 10. Close linkage between MNG and ABO, a chromosome 9 locus, is unlikely in females as the lod score was less than -2, the formal level for exclusion of linkage, at a theta value of 0.05. Close linkage between MNG and RH, a chromosome 1 marker locus, is also unlikely ($z = -1.185$ at theta= 0.05) in females although the data do not reach a formal level of significance. Positive lod scores ($z = 0.646$ at theta= 0.20) were obtained for GPT:MNG. Slightly positive lods at large theta values ($z= 0.005$ at theta = 0.40) were also achieved between SE, a chromosome 19 marker locus, and MNG. Since these lod scores did not reach formal levels of significance, further testing of other nuclear families in this kindred would be necessary to determine if MNG is possibly linked to GPT or SE.

TABLE 10

LOD SCORES FOR LINKAGE BETWEEN MNG AND OTHER LOCI FOR THE TG:HAR KINDRED

C	LOCUS	SEGREGATION INFORMATION MT	R	NR	NUMBER OF F	C	ASSUMED RECOMBINATION FRACTION 0.05	0.10	0.20	0.30	0.40
8	TG	M	-	-	1	6	-1.185	-0.673	-0.254	-0.087	-0.018
8 or 16	GPT	M	-	-	1	6	0.093	0.276	0.323	0.222	0.076
1	RH	M	-	-	1	6	-1.185	-0.673	-0.254	-0.087	-0.018
9	ABO	M	-	-	1	6	-2.164	-1.331	-0.581	-0.227	-0.053
19	SE	M	-	-	1	6	-0.398	-0.172	-0.022	0.009	0.005

C: CHROMOSOME; MT: MATING TYPE; R: NUMBER OF RECOMBINANTS; NR: NUMBER OF NONRECOMBINANTS; F: FAMILIES; C: CHILDREN.

3.500 Linkage Studies

Lod scores for chromosome 8 marker loci obtained from the "chromosomally normal" families are presented in Table 11. The CA2:GPT paternal lods were slightly positive at larger theta values (z= 0.018, theta= 0.30) while the paternally derived CA2:PLAT and CA2:TG lods were negative at all values of theta tested. The maternal CA2:GPT lods were slightly positive at larger theta values (z= 0.387, theta= 0.20) wheras the CA2:TG lods were negative at all values of theta tested and exclude linkage between CA2 and TG at theta $\leq$ 0.05 in females.

Only a single 2:0 phase unknown score, which occurred in the paternal segregation data, was obtained for GPT and PLAT. For GPT and TG, the single segregating sibship in the male data gave negative lods at all values of theta tested while the maternally derived lods were positive at larger values of theta ($\hat{z}$ = 0.215, theta= 0.25). The single sibship segregating for PLAT and TG occurred in the male segregation data and gave lods which were negative at all values of theta tested. Taken together these lods imply that the four marker loci: CA2, GPT, PLAT and TG, are not closely linked in "chromosomally normal" males. The slightly positive CA2:GPT and GPT:TG lods in the female segregation data imply that the members of each of these pairs may be loosely linked.

TABLE 11

LOD SCORES FOR CA2, GPT, PLAT AND TG FROM INDIVIDUALS NOT KNOWN TO CARRY STRUCTURALLY REARRANGED CHROMOSOMES.

LOCUS PAIR	SEGREGATION INFORMATION MT	R	NR	NUMBER OF F	C	ASSUMED RECOMBINATION FRACTION 0.05	0.10	0.20	0.30	0.40
CA2:GPT	P	-	-	3	11	-0.907	-0.423	-0.070	0.018	0.013
	M	-	-	4	19	-0.557	0.068	0.387	0.305	0.106
	I	-	-	1	4	-0.927	-0.458	-0.121	-0.023	-0.001
	T			8	34	-2.391	-0.813	0.196	0.300	0.118
CA2:PLAT	P	-	-	1	3	-0.721	-0.444	-0.194	-0.076	-0.018
CA2:TG	P	-	-	3	12	-1.184	-0.673	-0.254	-0.088	-0.019
	M	-	-	2	10	-2.163	-1.331	-0.582	-0.227	-0.053
	T			5	22	-3.347	-2.004	-0.836	-0.315	-0.072
GPT:PLAT	P	-	-	1	3	0.258	0.215	0.134	0.064	0.017
GPT:TG	P	-	-	1	3	-0.721	-0.444	-0.194	-0.076	-0.018
	M	-	-	4	18	-0.834	-0.182	0.203	0.199	0.074
	I	-	-	1	6	-1.687	-0.967	-0.395	-0.152	-0.035
	T			6	27	-3.242	-1.593	-0.386	-0.029	0.021
PLAT:TG	P	-	-	1	3	-0.721	-0.444	-0.194	-0.076	-0.018

MT: MATING TYPE; R: NUMBER OF RECOMBINANTS; NR: NUMBER OF NONRECOMBINANTS; F: FAMILIES; C: CHILDREN

lod scores for chromosome 8 marker loci obtained from the inv(8)(p23q22) heterozygotes are presented in Table 12. Paternally and maternally derived CA2:8INV lods were positive at all values of theta tested reaching a peak total lod score of 1.58 at a theta of 0.00. Similarly, CA2:PLAT paternal and maternal lods were positive at all values of theta tested, (z= 0.279 in males; 0.258 in females, theta= 0.05). A single maternally derived 2:0 phase unknown score was obtained for CA2 and TG.

The GPT:TG paternal lods were positive at all values of theta tested (z= 0.793, theta= 0.05) while GPT:8INV paternal lods were slightly positive at larger theta values ($\hat{z}$= 0.17, theta= 0.28). In contrast, GPT:8INV maternal lods were negative at all thetas tested and exclude linkage between GPT and 8INV at theta 0.05 in females.

Paternally derived PLAT:8INV lods were positive at all values of theta tested (z= 0.537, theta= 0.05) while paternal TG:8INV lods were negative at all values tested. Maternally derived TG:8INV and PLAT:8INV lods were also negative at all values of theta tested. A single 2:0 phase unknown score in female segregation data was obtained for PLAT and TG.

These results although inconclusive suggest that CA2 may be linked to 8INV, PLAT and TG in inversion heterozygotes. Similarly paternal GPT:TG, PLAT:8INV and maternal PLAT:TG scores imply that these loci pairs may also be linked. In contrast, TG and GPT are

TABLE 12

LOD SCORES FOR CA2, GPT, PLAT, TG AND INV(8)(p23q22) FROM INVERSION HETEROZYGOTES

LOCUS PAIR	SEGREGATION INFORMATION MT	R	NR	NUMBER OF F	C	ASSUMED RECOMBINATION FRACTION 0.05	0.10	0.20	0.30	0.40
CA2:GPT										
CA2:PLAT	P	-	1	1	1	0.279	0.255	0.204	0.146	0.079
	M	-	-	1	3	0.258	0.215	0.134	0.064	0.017
	T		1	2	4	0.537	0.470	0.338	0.210	0.096
CA2:TG	M	-	-	1	3	0.258	0.215	0.134	0.064	0.017
CA2:8INV	P	0	1	1	1	0.279	0.255	0.204	0.146	0.079
	M	0	3	2	5	1.177	1.068	0.839	0.588	0.310
	T		4	3	6	1.456	1.323	1.043	0.734	0.389
GPT:PLAT										
GPT:TG	P	-	-	2	8	0.793	0.680	0.452	0.234	0.066
GPT:8INV	P	0	1	3	11	-0.628	-0.167	0.134	0.165	0.092
	M	-	-	1	7	-2.164	-1.331	-0.581	-0.227	-0.053
	T			4	18	-2.792	-1.498	-0.447	-0.062	0.039
PLAT:TG	M	-	-	1	3	0.258	0.215	0.134	0.064	0.017
PLAT:8INV	P	0	1	2	3	0.537	0.470	0.338	0.210	0.096
	M	1	0	3	6	-0.878	-0.588	-0.319	-0.182	-0.086
	T			5	9	-0.341	-0.118	0.021	0.028	0.010
TG:8INV	P	-	-	2	8	-1.185	-0.673	-0.254	-0.087	-0.019
	M	-	-	3	11	-1.604	-1.054	-0.483	-0.143	-0.046
	T			5	19	-2.789	-1.727	-0.737	-0.280	-0.065

MT: MATING TYPE; R: NUMBER OF RECOMBINANTS; NR: NUMBER OF NONRECOMBINANTS; F: FAMILIES; C: CHILDREN

probably not closely linked to 8INV.

Lod scores for CA2 and unassigned loci obtained from the "chromosomally normal" families are presented in Table 13. The CA2:YT paternally derived lods were positive at all values of theta tested ($\hat{z}$= 0.61, theta= 0.00) while the maternally derived lods for this locus pair were positive at larger theta values (z= 0.013, theta= 0.40).

Paternally derived CA2:JK and CA2:DO lods were negative at all values of theta tested and exclude linkage between CA2 and DO in males at theta 0.05. Maternal CA2:DO and CA2:JK lods were also negative at all values of theta tested and exclude linkage between CA2 and JK in females at theta $\leq$ 0.10.

The CA2:CO and CA2:F13B maternal lod scores were positive at all values of theta tested (z= 0.535 and 0.258 respectively, theta= 0.05) while maternally derived lods for CA2 and KEL were negative at all thetas tested and exclude linkage between these two loci at theta = 0.05.

The lod scores imply that CA2 may be linked to F13B, CO or YT but is probably not closely linked to DO, JK or KEL, although more data are required to make formal conclusions.

Table 14 presents the lod scores for TG and unassigned loci for "chromosomally normal" families. Paternally derived lods for TG and CO were negative at all values of theta tested (z= -0.464,

TABLE 13

LOD SCORES FOR CA2 AND UNASSIGNED LOCI FROM INDIVIDUALS NOT KNOWN TO CARRY STRUCTURALLY REARRANGED CHROMOSOMES

LOCUS	SEGREGATION INFORMATION MT	R	NR	NUMBER OF F	C	ASSUMED RECOMBINATION FRACTION 0.05	0.10	0.20	0.30	0.40
CO	M	-	-	1	5	0.535	0.465	0.318	0.170	0.049
DO	P	-	-	2	9	-2.163	-1.331	-0.582	-0.227	-0.053
	M	-	-	1	6	-0.219	-0.168	-0.091	-0.039	-0.010
	I	-	-	1	4	0.742	0.630	0.407	0.205	0.056
	T			4	19	-1.640	-0.869	-0.266	-0.061	-0.007
JK	P	-	-	1	4	-0.721	-0.444	-0.194	-0.076	-0.018
	M	-	-	2	11	-3.606	-2.218	-0.969	-0.378	-0.088
	T			3	15	-4.327	-2.662	-1.163	-0.454	-0.106
KEL	M	-	-	2	11	-2.627	-1.560	-0.642	-0.238	-0.053
F13B	M	-	-	1	2	0.258	0.215	0.134	0.064	0.017
YT	P	-	-	1	3	0.535	0.465	0.318	0.170	0.049
	M	-	-	2	11	-1.371	-0.651	-0.130	0.008	0.013
	T			3	14	-0.836	-0.186	0.188	0.178	0.062

MT: MATING TYPE; R: NUMBER OF RECOMBINANTS; NR: NUMBER OF NONRECOMBINANTS; F: FAMILIES; C: CHILDREN

TABLE 14

LOD SCORES FOR TG AND UNASSIGNED LOCI FROM INDIVIDUALS NOT KNOWN TO CARRY STRUCTURALLY REARRANGED CHROMOSOMES

LOCUS	SEGREGATION INFORMATION MT	R	NR	NUMBER OF F	C	ASSUMED RECOMBINATION FRACTION 0.05	0.10	0.20	0.30	0.40
CO	P	-	-	1	6	-0.464	-0.229	-0.060	-0.011	-0.001
	M	-	-	1	5	-1.442	-0.887	-0.388	-0.151	-0.035
	T			2	11	-1.906	-1.116	-0.448	-0.162	-0.036
DO	M	-	-	1	3	-0.721	-0.444	-0.194	-0.076	-0.018
	I	-	-	1	6	-1.382	-0.838	-0.358	-0.137	-0.032
	T			2	9	-2.103	-1.282	-0.552	-0.213	-0.050
JK	M	-	-	3	14	-3.605	-2.218	-0.970	-0.378	-0.088
KEL	M	-	-	2	11	-1.906	-1.116	-0.448	-0.162	-0.036
YT	M	-	-	1	5	-1.442	-0.887	-0.388	-0.151	-0.035

MT: MATING TYPE; R: NUMBER OF RECOMBINANTS; NR: NUMBER OF NONRECOMBINANTS; F: FAMILIES; C: CHILDREN

theta= 0.05) as were the maternal TG:CO lods (z= -1.442, theta= 0.05). Similarly maternal lods between TG and DO, JK, KEL and YT were negative at all values of theta tested and exclude linkage between TG and JK in females at theta $\leq$0.10.

These results suggests that TG is probably not closely linked to DO, JK, KEL or YT in females or closely linked to CO in either males or females.

4.000 DISCUSSION

4.100 Phenotypic Variation

The commonly occurring phenotypes for CA2, TG and PLAT were detected as described in the literature. However, an additional 7 kb fragment common to all PLAT phenotypes, not previously described, was also observed. This band was much fainter than any of the other fragments common to all phenotypes and may, therefore, reflect the sensitivity of the specific filter used in the Southern transfer. It is possible that the BIODYNE membrane used in this study binds DNA more efficiently than the nitrocellulose filters used in the original study thereby allowing detection of this 7 kb fragment.

The 7 kb fragment itself might be homologous to other serine protease genes, all of which have probably evolved from the same ancestral gene as PLAT (Rajput et al., 1985). Homologies in the distribution of Alu sequences and in the nucleotide sequences of the PLAT and PLAU genes support the hypothesis of a common origin for the serine proteases (Degen et al., 1986). Alternatively, the clone used in this study might not be exactly the same as the clone used in the original study.

4.200 Segregation Analysis

The distributions of phenotypes in the offspring of CA2 matings were characteristic of a Mendelian system with two codominant alleles segregating at a single autosomal locus. Similar results were obtained for TG with the exception of the TG 2-1 x TG 2 mating type where the distribution of phenotypes observed in the offspring was statistically significant from the expected distribution. Such deviation probably reflects random fluctuation due to the small number of families analyzed. Alternatively, it is possible that sample degradation has generated breakdown of the 5.8 kb fragment, corresponding to the TG*1 allele, into a 5.2 kb fragment, characteristic of the TG*2 allele so that true TG heterozygotes would then be interpreted as TG 2 homozygotes. The latter possibility is unlikely, however, as analysis of the pre-restricted DNA samples gave no particular evidence of degradation of high molecular weight DNA and it is unlikely that such a presumably random process would generate the exact restriction site necessary to create a 5.2 kb fragment.

4.300 Thyroglobulin and Linkage Studies in TG:HAR

Analysis of the TG data in the TG:HAR kindred suggests that the gene determining multinodular goitre segregating in this family is neither linked to nor due to a mutation in the TG locus. Although the inheritance of goitre in this family is autosomal dominant, reduced penetrance in males has been noted in other nuclear families of this kindred (Couch et al, 1986). There was no evidence that segregation of the disease phenotype in the nuclear family studied was distorted, as the mother and four of the six children had previously presented with goitre. Two of the affected children were males so lod scores were determined assuming complete penetrance of the gene. However, one or both of the normal sons could theoretically be carrying the gene for goitre but not expressing it. Under the premise that the MNG mutation is non-penetrant in one of these unaffected sons, a 5:1 phase unknown TG:MNG score would be obtained and if both unaffected sons carried the MNG mutation, a 4:2 score would be obtained. The observation of at least one recombinant among the offspring, however, reduces the probability that the MNG defect is in the TG gene itself or closely linked to it.

All of the patients in TG:HAR with goitre were euthyroid and demonstrated normal thyroidal uptake eliminating an iodide trapping defect as a possible cause of the disease. Moreover,

histological examinations of surgically removed thyroid nodules suggested that dyshormonogenesis was the cause of goitre in this family. Although these patients were all euthyroid, no anti-thyroid microsomal and thyroglobulin antibodies were present in their sera (Couch et al, 1986). These observations, combined with the absence of close linkage between TG and MNG are consistent with the hypothesis that the abnormality in this family may be a defect of transport or processing of thyroglobulin mRNA, with another gene responsible for the regulation of these processes in the thyroglobulin molecule. Defects in the processing of TG mRNA have been proposed to explain the minute amounts of thyroglobulin related antigens in a strain of hypothyroid goats (vanVoorthuizen et al, 1978). Similar investigations of the mRNA in goitres of subsequent patients in this kindred may determine the nature of the molecular defect.

Linkage studies suggested that MNG is probably not closely linked to ABO or RH. Although MNG:GPT and MNG:SE lod scores were slightly positive, these locus pairs are probably not closely linked. The possibility of loose linkage between either pair of loci is possible.

4.400 Linkage Studies

An hypothesized map of chromosome 8 with the positions of chromosome 8 marker loci relative to each other is presented in Fig. 4. PLAT has been localized to 8p12 (Yang-Feng et al, 1986) (i.e. within the inversion(8)(p23q22)) while TG has been localized to 8q24 (Baas et al, 1985) (i.e. outside the inversion). However, in inv(8)(p23q22) chromosomes PLAT is physically transposed such that it is much closer to TG than in structurally normal chromosomes. The negative lods obtained between TG and PLAT in "chromosomally normal" individuals and the positive lods for these loci in inversion 8 heterozygotes are consistent with the regional localisations of these loci.

Although the PLAT locus must be located within the inversion, the lod scores indicated that crossing over occurred between this locus and the inversion breakpoints without resulting in a duplication/deficient chromosome 8. These data suggest that the inv(8) chromosome involved underwent an even number of cross overs within the inversion. However, the mother in whom these events occurred was not informative for other chromosome 8 marker loci so the relative positions of these crossovers cannot be determined.

The most promising linkage relationship is between CA2 and 8INV, as indicated by the positive lod scores obtained in both male and female data at all values of theta, with the

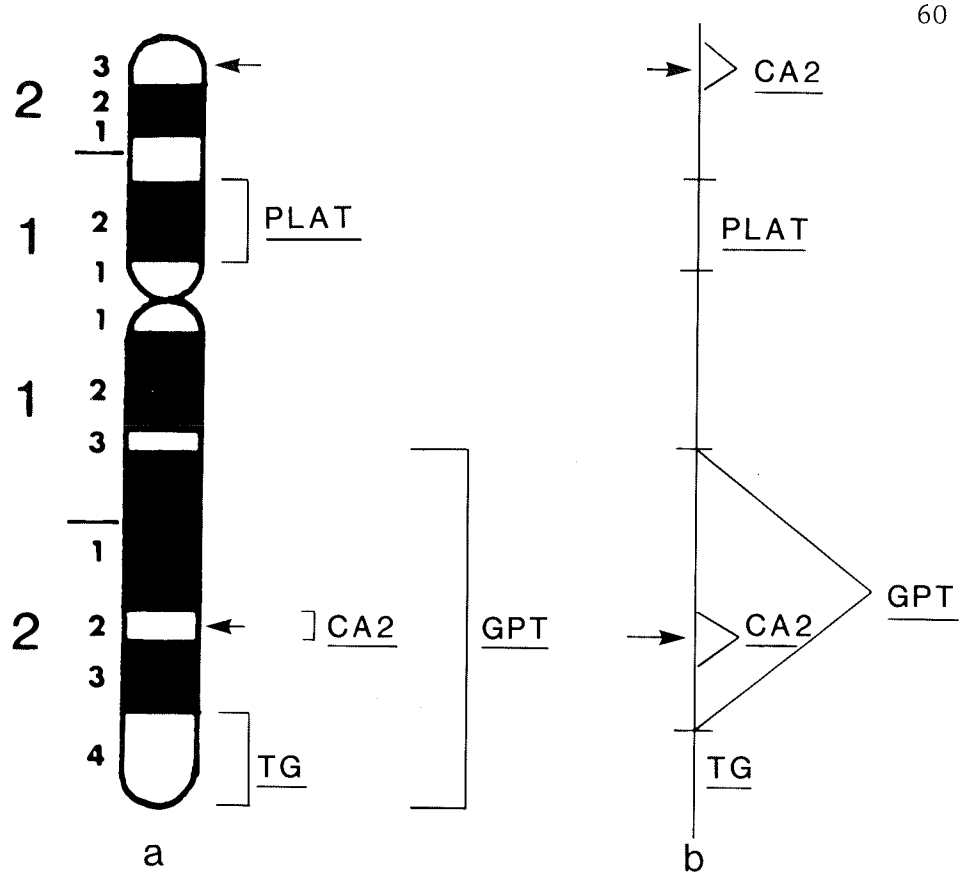

Fig. 4 a) Physical map of chromosome 8 showing regional localisations of loci with confirmed or provisional assignments, examined in this study.

b) Genetic map of chromosome 8 showing relative positions of gene loci and chromosome break points studied.

Note: Only the most likely position of CA2 on the physical map is indicated: two relative positions are possible for CA2 and GPT on the genetic map.

Arrows indicate breakpoints of inv(8)(p23q22).

total lod score peaking to 1.58 at theta= 0.00. Moreover, only six children from three families were scored for this locus pair and in the four instances where phase could be established, all offspring were CA2:8INV non-recombinants. Considering the small number of children scored, these positive lods are strongly suggestive of close linkage between CA2 and 8INV but the question then remains "to which 8INV break point is CA2 probably linked?" In "chromosomally normal" individuals the CA2:PLAT paternally derived lods were negative although positive paternal and maternal lods were obtained from inv(8)(p23q22) heterozygotes. Considering the location of PLAT on 8p these results suggest that CA2 is linked to the q22 break point of the inversion.

On the other hand CA2:TG lod scores were low enough to exclude formally linkage between these loci in "chromosomally normal" people These results suggest that either CA2 and TG are both on 8q but not closely linked or that CA2 is on 8p. The limited CA2:TG segregation information from the inv(8) heterozygotes (z= 2:0 for one female) does not help to distinguish the location of CA2 on chromosome 8 as the phase unknown score observed may just as easily represent two recombinants as non-recombinants. Collectively, the CA2 mapping data from this study then favour close linkage of CA2 to the q22 break point of 8INV but do not determine whether the locus is inside or outside the inversion. These results support the recent regional

localisation of CA2 to 8q22 by in situ hybridization (Nakai et al., in press).

The negative paternal and maternal lod scores obtained in this study indicate that TG is probably not closely linked to 8INV. Moreover, these results together with the localization of TG distal to MYC in 8q24 (Baas et al., 1985), suggest that TG may be positioned close to the telomere on 8q. However, the genetic distance between TG and MYC remains to be determined pending identification of a MYC RFLP. Although TG and 8q22 may appear to be relatively close to each other on the physical map of chromosome 8, the distance between these two points on the genetic map may be considerably greater since crossing over is more frequent in distal, as opposed to proximal, regions of chromosomes. (Cook et al., 1974).

The negative lod scores obtained for GPT:CA2 and GPT:TG suggest that if GPT is on chromosome 8 it is not located near CA2 or TG. Moreover linkage between GPT and 8INV was excluded at theta= 0.05 for maternal lod scores. Assuming GPT is on chromosome 8, the independent segregation of GPT alleles with the inversion suggests that GPT is either located outside of the inversion, or inside the inversion such that an even number of crossovers may occur without the formation of a duplication/deficient chromosome. Positive paternal lods were obtained for GPT:TG for inversion 8 heterozygotes. Although more

data are required before a firm conclusion can be made the data from this study hint that GPT may be located near PLAT on the genetic map, with its position being transposed downward towards TG in inv(8)(p23q22) chromosomes.

Although only one "chromosomally normal" family was informative for GPT and PLAT, the positive lod scores for this locus pair support the hypothesis outlined above that GPT may be linked to PLAT. No data were available for GPT versus PLAT in inv(8)(p23q22) heterozygotes although under this hypothesis, the same results as observed for "chromosomally normal" individuals could be possible. Exclusion mapping has shown that GPT may be located on chromosome 8, anywhere from q13 -> qter (Cook et al., 1981), so it may be close enough to show linkage to PLAT and yet far enough from CA2, TG and the 8INV break points for its alleles to segregate independently of these sites.

DO has been tentatively assigned to chromosome 1 while CO has been postulated to be on chromosome 2 (HGM8, 1985). Although the present study cannot confirm these assignments the results suggest exclusion of DO from a considerable portion of 8q since negative paternal and maternal lod scores were obtained for DO with CA2 and TG both of which have already been mapped to chromosome 8 though they themselves are apparently not closely linked. Close linkage of CO with TG is unlikely since negative

lod scores were obtained at all theta values in both males and females. Positive maternal lods were observed however between CA2 and CO for one family (z= 0.535, theta = 0.05) but more data are required to determine if these loci are linked.

4.500 Frequency of Multiple Heterozygotes and Allele Frequencies

In order to test loci for linkage it is necessary to study families in which at least one parent is doubly heterozygous at the two loci of interest. The frequency of double heterozygotes in a population is dependent on the frequencies of the variant alleles at each test locus in the population.

Consider the following example. Locus A and locus B are to be tested for linkage. Two common alleles whose frequencies are p and q segregate at A while r and s are the frequencies of two commonly occurring alleles at the B locus. For population N, p + q = 1 and r + s = 1. Say also that p = q = 0.5 and r = s = 0.5.

The frequency of heterozygotes at A in population N equals 2pq, which is numerically 2x0.5x0.5= 0.5 in this example. Similarly the frequency of heterozygotes for B in population N is 2rs or 2x0.5x0.5= 0.5. The frequency of double heterozygotes in

population N may then be derived as follows,

$$(2pq)(2rs) = (2x0.5x0.5)(2x0.5x0.5)$$

$$= 0.5x0.5 = 0.25,$$

one half the frequency of heterozygotes.

If another locus, C, with two common alleles at frequencies of t and u, is added to the equation then in population N: t + u= 1 and if t=u=0.5 then 2tu= 0.5. The frequency of triple heterozygotes in population N will be (2pq)(2rs)(2tu)= 0.5x0.5x0.5= 0.125; one half the frequency of double heterozygotes and one quarter the frequency of single heterozygotes.

It is apparent that as more loci are studied, the frequency of multiple heterozygotes decreases.

From studies of small samples of the Caucasian populations, the following frequencies for alleles at CA2, TG and PLAT loci have been estimated.

CA2: p= 0.5, q= 0.5 (Lee et al., 1985)

TG : p= 0.2, q= 0.8 (Baas et al., 1985)

PLAT: p= 0.5, q= 0.5 (Benham et al., 1984)

Using these estimates the expected frequency of CA2:TG:PLAT triple heterozygotes in a randomly selected sample of the Caucasian population may be estimated to be,

(2x0.5x0.5)(2x0.2x0.8)(2x0.5x0.5)= 0.5x0.32x0.5= 0.08

or 8 in 100.

Using the allele frequencies estimated from the present study the expected frequency of triple heterozygotes may be estimated to be,

(2x0.3x0.7)(2x0.6x0.4)(2x0.2x0.8)=0.065, or 6.5 in 100,

less than that predicted by using the estimated allele frequencies for these loci previously presented in the literature. Regardless of the precise frequency of CA2:TG:PLAT heterozygotes in the general population, large numbers of individuals will have to be screened to identify these triple heterozygotes. Consequently the collection of significant amounts of highly informative multipoint linkage data will be very labour intensive.

The scarcity of multiple heterozygotes as experienced in this study, suggests that although multipoint linkage information is valuable, two point linkage data will contribute more to the development of the human genetic map. This situation casts doubt on the validity of the hypothesis that the map of the human genome would be constructed largely from multipoint linkage data, which was proposed when the first autosomal linkage triplet of the loci for ABO, nail-patella syndrome and adenylate kinase was identified by Rapley et al. (1968).

4.600 Isolates and Genetic Drift

The estimated allele frequencies for CA2 in the sample population of the present study were shown to deviate significantly from the estimates calculated by Lee et al. (1985). Since the CA2 allele frequencies were consistent between families of different ethnic origins in this study, the likelihood that the observed deviation is due to genetic drift is low. However, the sample sizes for estimating alleles in each ethnic group of the present study and in Lee's study (Lee et al., 1985) were comparatively small so that any possible bias of ascertainment or other deviations from a truly homogeneous group of randomly selected unrelated individuals might have affected tha allele frequencies estimated.

Gene frequencies in populations may be altered by deterministic (directed) processes such as selection and by stochastic (random) processes such as genetic drift. The founder effect, and fluctuations in mutation and migration rates are also random processes. When gene pools are evolving, it should be recognized that stochastic processes do not necessarily act in isolation from deterministic proceeses (Vogel and Motulsky,1982).

Pan-neutralists do not believe that natural selection is the driving force of evolution but rather that genetic drift and neutral mutation play an immense role in genetic variation and evolutionary processes (Dobzhansky, 1977). All the polymorphisms

included in this study have been presumed to be benign, with individual phenotypes conferring no known advantages or disadvantages.

The GPT alleles expressed in red cells, as detected by starch gel electrophoresis, code for qualitatively different gene products each of which is associated with a characteristic level of enzyme activity (Chen et al., 1972). The restriction enzyme sites which generate the RFLPs studied occur in the coding sequences of PLAT, the 5' flanking sequences of CA2 and the 5' end of TG. At present it is unknown whether the TG RFLP sequences occur in an intron or exon. It is possible that the PLAT alleles and perhaps the TG alleles, like those of GPT, code for qualitatively different gene products. The redundancy of the genetic code should not be discounted however. Point mutations which produce new restriction sites in genes do not necessarily affect amino acid sequence in gene products. If, however, the PLAT and TG alleles do code for different gene products, the assumption that these polymorphisms are benign seems reasonable since the majority of amino acids in a typical protein are not necessary for the proper function of that protein. Hence, most mutations that occur probably are neither selected for or against. Their presence or absence in a population is dependent largely upon stochastic processes such as genetic drift.

The magnitude and effect of genetic drift in turn depends on population size. For infinitely large populations, random changes in gene frequency are negligible. However, in isolated and inbred populations such as the A3:WPA kindred, chance fluctuations can have a dramatic impact on gene frequencies, particularly over a number of successive generations. If the apparent predominance of CA2*2 alleles in the families studied is real, the most likely explanations are random genetic drift and founder effect.

4.700 A3:WPA Kindred

Twenty-one parents in the kindred A3:WPA carried the pericentric inversion of interest, inv(8)(p23q22). However, four of these individuals were deceased which left only seventeen to test; of these only twelve individuals consented to provide blood samples for analysis. These 12 parents had 51 offspring of whom only 32 consented to participate in this study. Consequently only about half of the originally observed potential of the kindred could be realised.

Moreover, phase could be determined for only 16 children from seven sibships, so lod scores were usually calculated with phase unknown data. In addition, five of the tested inversion 8 parents were widowed leaving only seven inversion (8)(p23q22) heterozygotes with tested spouses.

The inability to obtain samples for missing spouses and children proved to be most troublesome. When variant alleles are rare, genotypes of untested parents can usually be assigned reliably from the marker data for the rest of the family whereas with commonly occurring alleles it is more difficult to determine genotypes of untested parents and, hence, missing data become problematic.

To compound the problem most spouses in this kindred were probably related to some degree as a consequence of the geographical isolation of this kindred for several generations. The incidence of incomplete families in A3:WPA drastically lowered the amount of data obtainable and ultimately the amount of information deducible from this study.

Despite the limitations encountered in this study, a data base of information pertaining to the segregation of chromosome 8 marker loci has been initiated which may prove useful for future studies investigating the molecular anatomy and meiotic behaviour of the inversion 8 in the large Newfoundland kindred A3:WPA.

5.000 SUMMARY

1. The distributions of CA2, PLAT, and TG phenotypes in a sample population were studied. Allele frequencies for these loci were estimated and found not to deviate significantly from the estimates recorded in the literature with the exception of the CA2*1 and CA2*2 allele frequencies which were 0.3± 0.03 and 0.7 ± 0.02 respectively. (X^2 = 13.4, df 1, p>0.05).

2. Segregation analysis of CA2 and TG phenotypes indicated that their mode of inheritance followed the pattern expected for a Mendelian system where two codominant alleles segregate at a single autosomal locus. A statistically significant deviation in the distribution of offspring was observed however for the TG2-1 x TG2 mating where relatively few TG2 offspring phenotypes were identified.

3. Linkage studies of TG in the kindred TG:HAR showed that the goitre phenotype (MNG) segregating in this family is neither linked to or due to a mutation in the TG gene. Linkage studies of MNG with other loci suggested that the disease locus is not linked to ABO or RH. Slightly positive lods obtained for MNG:GPT and MNG:SE suggest that these loci pairs may be loosely linked.

4. Linkage studies of chromosome 8 marker loci in individuals not known to carry structurally rearranged chromosomes suggested that probably none of the four marker loci:CA2, GPT, PLAT and TG are closely linked in males. In contrast, CA2:PLAT, CA2:TG, GPT:TG and PLAT:TG lod scores were suggestive of loose linkage for these loci pairs in inversion heterozygotes although formal levels of significance were not reached.

5. The CA2:8INV lods were strongly suggestive of linkage and peaked at + 1.58 at a theta value of 0.00. In the four instances where phase was established all of the offspring were non-recombinants giving further credence to a CA2:8INV linkage relationship. In addition, slightly positive paternal lods indicate PLAT may also be loosely linked to 8INV. TG and GPT lods suggested that neither of these loci are probably closely linked to 8INV.

6. Lod scores for CA2 and unassigned loci in "chromosomally normal" individuals suggested possible linkage relationships with 13B, CO and YT. Negative lod scores between TG and unassigned loci suggested that TG is probably not closely linked to CO, DO, JK, KL or YT.

6.000 REFERENCES

Allerdice, P.W. (1987). personal communication.

Allerdice, P.W. (1974). A large kindred with segregation of pericentric inversion heterozygotes inv(3)(p25q21) and inv(8)(p23q22). Am. J. Hum. Genet. 26:9A.

Astrin, K.H., Arredondo-Vega, F.X., Desnick, R.J., and Smith, M. (1981). Expression of differentiated liver functions in rat-human hybrids. HGM6. Cytogenet. Cell Genet. 32:249-250.

Avvedimento, V.E., diLauro, R., Monticella, A., Bernardi, F.,Patracchini, P., Calzolari, E., Martini, G., and Varrone, S. (1985). Mapping of human thyroglobulin gene on the long arm of chromosome 8 by in situ hybridization. Hum. Genet. 71:163-166.

Baas, F., Bikker, H., van Kessel, A.G., Melsert, R., Pearson, P.L., deVijlder, J.J.M., and van Ommen, G.J.B. (1985). The human thyroglobulin gene: a polymorphic marker localized distal to c-myc on chromosome 8 band q24. Hum. Genet. 69:138-143.

Bass, E.B., Smith, S.W., Stevenson, R.E., and Rosse, W.F. (1983). Further evidence for location of the spherocytosis gene on chromosome 8. Ann. Int. Med. 99:192-193.

Benham, F.J., Spurr, N., Povey, S., Brinton, B.T., Goodfellow, P.N., Solomon, E., and Harris, T.J.R. (1984). Assignment of tissue-type plasminogen activator to chromosome 8 in man and identification of a common restriction length polymorphism within the gene. Mol. Biol. Med. 2:251-259.

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

Buchwald, M. (1984). personal communication.

Bui, T.H., Sichong, Z., and Castro, I. (1982). A familial pericentric inversion of chromosome 8 analysed with a high resolution chromosome banding technique. Clin. Genet. 21:266-271.

Chen, S.H., and Giblett, E.R. (1971). Polymorphism of soluble glutamic-pyruvic transaminase:a new genetic marker in man. Science. 173:148-149.

Chen, S., Giblett, E.R., Anderson, J.E., and Fossum, B.L.G. (1972) Genetics of glutamic-pyruvic transaminase:its inheritance, common and rare variants, population distribution and differences in catalytic activity. Ann. Hum. Genet. 35:401-409.

Chodirker, B., McAlpine. P.J., Ray, M., Vust, A., Pugh, D., and Chudley, A.E. (1986). Exclusion of red cell glutamate-pyruvate transaminase (GPT) locus from the distal p13 region of chromosome 16 in a patient with a ring chromosome 16. Amer. J. Hum. Genet. 39:A150.

Conneally, P.M. (1986). personal communication.

Cook, P.J.L., Jeremiah, S.J., and Buckton, K.E. (1981). Exclusion mapping of GPT. Cytogenet. Cell Genet. 32:258.

Cook, P.J.L., Robson, E.B., Buckton, K.E., Jacobs, P.A., and Polani, P.E. (1974). Segregation of genetic markers in families with chromosome polymorphisms and structural rearrangements involving chromosome 1. Am. J. Hum. Genet. 37:261-274.

Cote, G.B. (1975). Centromeric linkage in man. Ph.D. thesis, University of Birmingham.

Couch, R.M., Hughes, M.B., deSa, D.J., Schiffrin, A., Guyda, H., and Winter, J.S.D. (1986). An autosomal dominant form of adolescent multinodular goitre. Am. J. Hum. Genet. 39:811-816.

Degen, S.J.F., Rajput, B. and Reich, E. (1986). The human tissue plasminogen activator gene. J. Biol. Chem. 261:6972-6985.

Dobzhansky, T., and Spassky, B. (1947). Evolutionary changes in laboratory cultures of D. pseudoobscura. Evolution. 1:191-216.

Dobzhansky, T. (1977). Populations, races, subspecies in Evolution. (pp128-163). W.H. Freeman and Company, San Francisco.

Fujimoto, A., Wilson, M.G., and Towner, J.W. (1975). Familial inversion of chromosome no. 8:an affected child and a carrier fetus. Hum. Genet. 27:67-73.

Ganong, W.F. (1981). The thyroid gland. in Review of medical physiology. (pp250-263). Lange Medical Publications, Los Altos, California.

Garger, S.J., Griffith, O.M., and Grill,L.K. (1983). Rapid purification of plasmid DNA by a single centrifugation in a two-step cesium chloride-ethidium bromide gradient. Biochem. Biophys. Res. Comm. 117:835-842.

Haldane, J.B.S., and Smith, C.A.B. (1947). A new estimate of the linkage between the genes for colour-blindness and hemophilia in man. Ann. Eugen. 14:10-31.

Harris, H. (1974). Enzyme and protein diversity in human populations. in The Principles of Human Biochemical Genetics. (pp278-338). American Elsevier, New York.

Herva, R., and de la Chapelle, A. (1976). A large pericentric inversion of chromosome 8 analysed with a high resolution chromosome banding technique. Hum. Genet. 21:266-271.

Hook, E.B., Hamerton, J.L. (1977). The frequency of chromosome abnormalities detected in consecutive newborn studies-differences between studies-results by sex and by severity of phenotypic involvement. in Population cytogenetics-Studies in humans.

Human gene mapping 8. (1985). Cytogenet. Cell. Genet. 40:1.

Jacobs, P.A., Buckton, K.E., Cunningham, C., and Newton,M. (1974). An analysis of break points of structural rearrangements in man. J. Med. Genet. 11:50-64.

Kielty, C.M., and Povey,S. (1981). Mapping of liver-specific enzymes. I. Glutamate-pyruvate transaminase (GPT). HGM6. Cytogenet. Cell Genet. 32:289-290.

Kimberling, W.J., Fulbeck, T., Dixon, L., Lubs, H.A. (1975). Localization of spherocytosis to chromosome 8 or 12 and report of a family with spherocytosis and a reciprocal translocation. Am. J. Hum. Genet. 27:586-594.

Lee, B.L., Venta, P.J., and Tashian,R.E. (1985). DNA polymorphism on the 5' flanking region of the human carbonic anhydrase II gene on chromosome 8. Hum. Genet. 69:337-339.

Maniatis, T., Fritsch, E,F., Sambrook, J. (1982). Molecular Cloning. (ppl09-112). Cold Spring Harbor Laboratory, New York.

Maynard-Smith, S., Penrose, L.S. and Smith, C.A.B. (1961). Mathematical tables for research workers in human genetics. (pp74). Churchill,London.

McKusick, V.A. (1986). Mendelian Inheritance in Man. (p xi). John Hopkins University Press, Baltimore.

Moedjono, S.J., and Sparkes, R.S. (1980). Familial pericentric inversion of chromosome 8:is breakpoint p23q22 important in the formation of unbalanced recombinants? Ann. Genet. 23:235-237.
Morton, N.E. (1955). Sequential tests for the detection of linkage. Am. J. Hum. Genet. 7:277-318.

Nakai, H., Byers, M.G., Venta, P.J., Tashian, R.E., and Shows, T.B. (1987). The gene for human carbonic anhydrase 2 is located at chromosome 8q22. Cytogenet. Cell Genet. in press.

Neel, J.V., Salzano, F.M., Junqueira, P.C., Keiter,F., and Maybury-Lewis, D. (1964). Studies on the Xavante indians of the Brazilian Mato Grosso. Am. J. Hum. Genet. 16:52-140.

Peakman, D.C., Smith, A.C.M., and Sujansky,E. (1977). Familial inversion of chromosome number 8. Amer. J. Hum. Genet. 29:87A.

Rajput, B., Degen, S.F. Reich, E., Waller, E.K., Axelrod, J., Eddy, R.L., and Shows, T.B. (1985). Chromosomal locations of human tissue plasminogen activator and urokinase genes. Science. 230:672-674.

Rapley, S., Robson, E.B., Harris, H., and Smith, S.M. (1968). Data on the incidence, segregation and linkage relationships of the adenylate kinase (AK) polymorphism. Ann. Hum. Genet. 31:237-242.

Rijken, D.C., and Collen,D. (1981). Purification and characterisation of the plasminogen activator secreted by human melanoma cells in culture. J. Biol. Chem. 256:7035-7041.

Rodriguez, R.L., and Tait, R.C. (1983). Rapid isolation of plasmid DNA. in Recombinant DNA techniques. (pp50-51). The Benjamin/Cummings Publishing Company, Don Mills, Ontario.

Sly, W.S., Hewett-Emmett, D., Whyte, M.P., Yu Ysl, and Tashian, R.E. (1983). Carbonic anhydrase II deficiency identified as the primary defect in the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. Proc. Natl. Acad. Sci., USA. 80:2752-2756.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Sturtevant, A.H. (1926). A crossover reducer in Drosophila melanogaster due to inversion of a section of the third chromosome. Biol. Zbl. 46:697.

Sturtevant, A.H., and Beadle,G.W. (1936). The relations of inversions in the X chromosome of Drosophila melanogaster to crossing over and disjunction. Genetics. 21:554-602.

Sujansky, E., Smith, A.C.M., Peakman, D.C., McConnell, T.S., Baca,P., and Robinson, A. (1981). Familial pericentric inversion of chromosome 8. Amer. J. Med. Genet. 10:229-235.

Tashian, R.E., Kendall, A., and Carter, N.D. (1980). Inherited variants of human red cell carbonic anhydrases. Hemoglobin. 4:635-651.

van Ommen, G.J.B., Arnberg, A.C., Baas, F., Brocas, H., Sterk, A., Tegelaers, H.H., Vassart, G., and de Vijlder, J.J.M. (1983). The human thyroglobulin gene contains two 15-17 kb introns near its 3'-end. Nuc. Ac. Res. 11:2273-2285.

van Voorthuizen, W.F., Dinsart,C., Flavell, R.A., deVijlder, J.J.M., and Vassart, G. (1978). Abnormal cellular localization of thyroglobulin mRNA associated with hereditary congenital goitre and thyroglobulin deficiency. Proc. Natl. Acad. Sci. USA. 75:74-78.

Venta, P.J., Shows,T.B., Curtis, P.J. and Tashian, R.E. (1983). Polymorphic gene for human carbonic anhydrase II:a molecular disease marker located on chromosome 8. Proc. Natl. Acad. Sci., USA. 80:4437-4440.

Vogel, F., and Motulsky, A.G. (1982). Human Genetics. (pp282-283). Springer-Verlag, New York.

Wald, A. (1947). Sequential Analysis. (pp212). Wiley, New York.

Wijnen, L.M.M., and Meera Khan, P. (1981). Assignment of GPT1 to human chromosome 16. Cytogenet. Cell. Genet. 32:327.

Yang-Feng, T.L., Opdenakker, G., Volckaert, G., and Francke, U. (1986). Human tissue-type plasminogen activator gene located near chromosomal breakpoint in myeloproliferative disorder. Am. J. Hum. Genet. 39:79-87.

Yu, C.W., Borgaonkar, D.S., and Bolling, D.R. (1978). Break points in human chromosomes. Hum. Hered. 28:210-225.

Ziegenbein, R. (1966). Two different forms of glutamic pyruvic transaminase in rat heart and their intracellular localization. Nature. 212:935.

7.000 APPENDIX

7.100 List of Reagents

10% Blotto: 0.1 g/ml skim milk powder
0.2 % sodium azide

1 % Denhardt solution: 0.01 g/ml Bovine serum albumin
0.01 g/ml Ficoll
0.01 g/ml polyvinyl pyrolidine
0.01 g/ml sodium dodecyl sulphate

20 X SSC: 3 M NaCl
0.3 M sodium citrate (pH 7.0)

20 X SSPE: 3.6 M NaCl
0.2 M sodium phosphate (pH 7.0)
0.2 M EDTA

7.200 List of Abbreviations

bp = base pairs
kb = kilobase pairs
cM = centimorgans
cpm = counts per minute
uCi = microCurie
g = gram
ug = microgram
ng = nanogram
h = hour
m = minute
s = second
l = litre
ml = millilitre
ul = microlitre
M = molar
mM = millimolar
ddH_2O = deionized distilled water
SDS = sodium dodecyl sulphate
YT = locus symbol for the Yt blood group
sssDNA = sonicated salmon sperm DNA
nm = nanometer

Note: All chemicals used in this study were of the highest purity. All solutions were made with ddH_2O.