Linkage Analysis in Corneal Stromal Dystrophy Type 1

by
Reena Ray

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Human Genetics
University of Manitoba
Winnipeg, Manitoba

September, 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-612-16245-1

Canada

Name _____

Dissertation Abstracts International and *Masters Abstracts International* are arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation or thesis. Enter the corresponding four-digit code in the spaces provided.

SUBJECT TERM

BIOLOGY - GENETICS

SUBJECT CODE 0369 UMI

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Design and Decorative Arts 0389
Fine Arts 0357
Information Science 0723
Journalism 0391
Landscape Architecture 0390
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Educational Psychology 0525
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291
- Rhetoric and Composition 0681

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- Caribbean 0360
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Church 0330
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Fisheries and Aquaculture 0792
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Range Management 0777
- Soil Science 0481
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Animal Physiology 0433
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Plant Physiology 0817
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Dentistry 0567
- Education 0350
- Administration, Health Care 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Safety 0354
- Oncology 0992
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Rehabilitaion and Therapy 0382
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Condensed Matter 0611
- Electricity and Magnetism 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Environmental 0775
- Industrial 0546
- Marine and Ocean 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Cognitive 0633
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

LINKAGE ANALYSIS IN CORNEAL STROMAL DYSTROPHY TYPE 1

BY

REENA RAY

A Thesis/Practicum submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Abstract

Lattice corneal dystrophy type 1 (CSD1), an autosomal dominant disorder causes significant visual impairment. It is characterized histologically by the deposition of an, as yet undefined, amyloid specifically in the corneal stroma. By slit-lamp examination, refractile lines, dots, dashes or haze are characteristically evident. Treatment is symptomatic to provide relief from corneal erosions. Vision loss is restored by corneal transplantation. However, erosions and progressive visual loss reoccur, culminating in the need for re-transplantation.

A single large multigeneration CSD1 kindred of Belgian descent forms the basis of this study. All CSD1 patients and at-risk individuals are being monitored by one ophthalmologist. Those affected have been symptomatic and/or have demonstrated characteristic corneal changes by the age of 25 years. To date, none of the individuals assessed as unaffected at the age of 25 years have subsequently developed the signs and symptoms of CSD1 later on in life. There is no evidence of any systemic illnesses, specifically neurological, cardiac or renal in any affected individual. The 60 participating family members 25 years of age or older include 34 affected and 26 unaffected individuals as well as 13 unrelated spouses. An additional 15 family members between the ages of 15 and 24 years with a 50% risk of carrying the CSD1 gene were also studied.

Recently, *CSD1* was mapped to a 10 centiMorgan interval on chromosome 5q22-5q33.3 bound by *IL9* proximally and *D5S436* distally (Stone *et al*, 1994). This study was designed to confirm the mapping of *CSD1* and refine the candidate interval. Following a screen of 10 microsatellite markers in the candidate region, linkage analysis was performed using MARK III (Côté, 1975) and LINKAGE (Lathrop *et al*, 1984; Lathrop and Lalouel, 1984; Lathrop *et al*, 1986). The results generated by the two approaches were consistent and confirmed the mapping of *CSD1* to 5q. The highest lod score was observed between *CSD1* and *D5S414*. Our study provides sufficient evidence that *CSD1* maps to the same 2 cM interval defined by *D5S393* proximally and *D5S396* distally as reported by Gregory *et al* (1995). To date, none of the known 14 genes whose products are involved in amyloid deposition nor any potential candidate genes map to the candidate interval. All the known affected individuals share a common haplotype for the five most tightly linked markers which differs from the haplotypes reported by other groups. This finding implies different mutations have arisen independently at the CSD locus. The haplotype segregating with *CSD1*D* has been detected in 5 of 15 at-risk individuals screened, predicting they will, most likely, manifest the signs and symptoms of CSD1.

In screening the panel of PCR-formatted microsatellite markers, a non-amplifying allele at *D5S414* was observed to be segregating in one branch of the CSD1 kindred. A sequence variation at the primer annealing site was hypothesized to give rise to the non-amplifying allele. Molecular characterization revealed a two base pair change immediately adjacent to the 3'-end of the reverse primer. This sequence variation created a repetitive region of three tandemly repeated CG base pairs causing inefficient primer-template annealing and markedly reduced or no amplification. Failure to recognize non-amplifying alleles can result in the omission or misinterpretation of linkage data.

Acknowledgments

I would like to thank Dr. C. R. Greenberg for the opportunity to study in her laboratory.

I would also like to thank the members of my committee, Dr. P. J. McAlpine and Dr. T. Zelinski, for their advice and criticism of this thesis.

Thank you to Cheryl Taylor, Rupinder Singal and Dr. J. Haworth for their continued technical assistance as well as encouragement during my studies.

Special thank you to Edward Nylen for his expert technical assistance.

Many thanks to Sylvia Philipps for her help with the MARK III and PCMAP83 computer programs.

I would also like to thank Dr. K. Morgan and Tracey Weiler for their help with the LINKAGE computer program.

A very special thank you to Dr. M. Ekins and the members of this kindred whose interest and desire to participate in scientific research made this project possible.

List of Figures

List of Tables

List of Abbreviations

AA	amyloid A protein
ATP	adenosine 5'-triphosphate
bp	base pair
CHLC	Cooperative Human Linkage Centre
cM	centiMorgan
cR	centiray
CSD1	corneal stromal dystrophy type 1
CSD1*N	normal allele for CSD1
CSD1*D	disease allele for CSD1
CSD2	corneal stromal dystrophy type 2
CSD3	corneal stromal dystrophy type 3
DMSO	dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
FISH	fluorescence *in situ* hybridization
GDB	Genome Database
kb	kilobase
KCl	potassium chloride
kDa	kiloDalton
LCD1	lattice corneal dystrophy type 1
lod	log of the odds
M	molar; moles L^{-1}
$MgCl_2$	magnesium chloride
Mb	megabase
mL	millilitre
mM	millimolar
ng	nanogram
NH_4Cl	ammonium chloride
nm	nanometre
NR	nonrecombinant(s)
OMIM	Online Mendelian Inheritance in Man
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIC	polymorphic information content
R	recombinant(s)
rpm	revolutions per minute
SAP	serum amyloid P component
SDS	sodium dodecyl sulfate
SSM	slipped strand mispairing
STRP	short tandem repeat polymorphism
STS	sequenced tagged site
TAE	tris/acetate (buffer)
Taq	*Themus aquaticus* DNA (polymerase)

TBE	Tris/borate (buffer)
TE	Tris/EDTA (buffer)
TEMED	N,N,N',N'- tetramethylethylene diamine
Tm	melting temperature; thermal denaturation
Tris	Tris (hydroxymethyl) aminomethane
UV	ultraviolet
V	volts
YAC	yeast artificial chromosome
θ	theta; recombination fraction
μCi	microCurie
μL	microlitre
μM	micromolar
^{32}P	radioisotope of phosphorus

Table of Contents

1. Introduction

1.1. Anatomy of the Cornea

The cornea is transparent connective tissue which provides a physical barrier to protect the enclosed ocular tissues from infection, dust and other foreign material. The elasticity of the cornea further helps to protect it from damage by buffering external pressure or force (Hollwich, 1985). It is the primary refractive element of the eye with a refractive power equivalent to 43 diopters (Vaughan *et al*, 1992). The central thickness is 0.52 mm, increasing towards the periphery to 0.67 mm. The diameter of the cornea is 11.7 mm with a superficial area of 1.3 cm^2 (Hogan *et al*, 1971).

From the anterior to the posterior, the cornea is organized into five distinct layers: (1) epithelium, (2) Bowman layer, (3) stroma, (4) Descemet membrane and (5) endothelium (figure 1). The epithelium is made up of five to six layers of cells which occupy about ten percent of the total corneal thickness. The uniform thickness and regular arrangement of the cells results in a smooth anterior surface which contributes to corneal transparency and its maintenance. Bowman layer is a modified layer of the stroma which is both clear and acellular. The stroma constitutes ninety percent of the corneal thickness. It is composed almost entirely of collagenous lamellae which are arranged parallel to each other and to the surface of the cornea. The size and periodicity of the collagen fibrils is essential to the cornea's light-conducting properties. Descemet membrane is the basement membrane of the corneal endothelium. It is two to three times the thickness of the endothelium and characteristically clear and elastic. The endothelium is the innermost layer of single cells between the stroma and the aqueous humour. It is responsible for

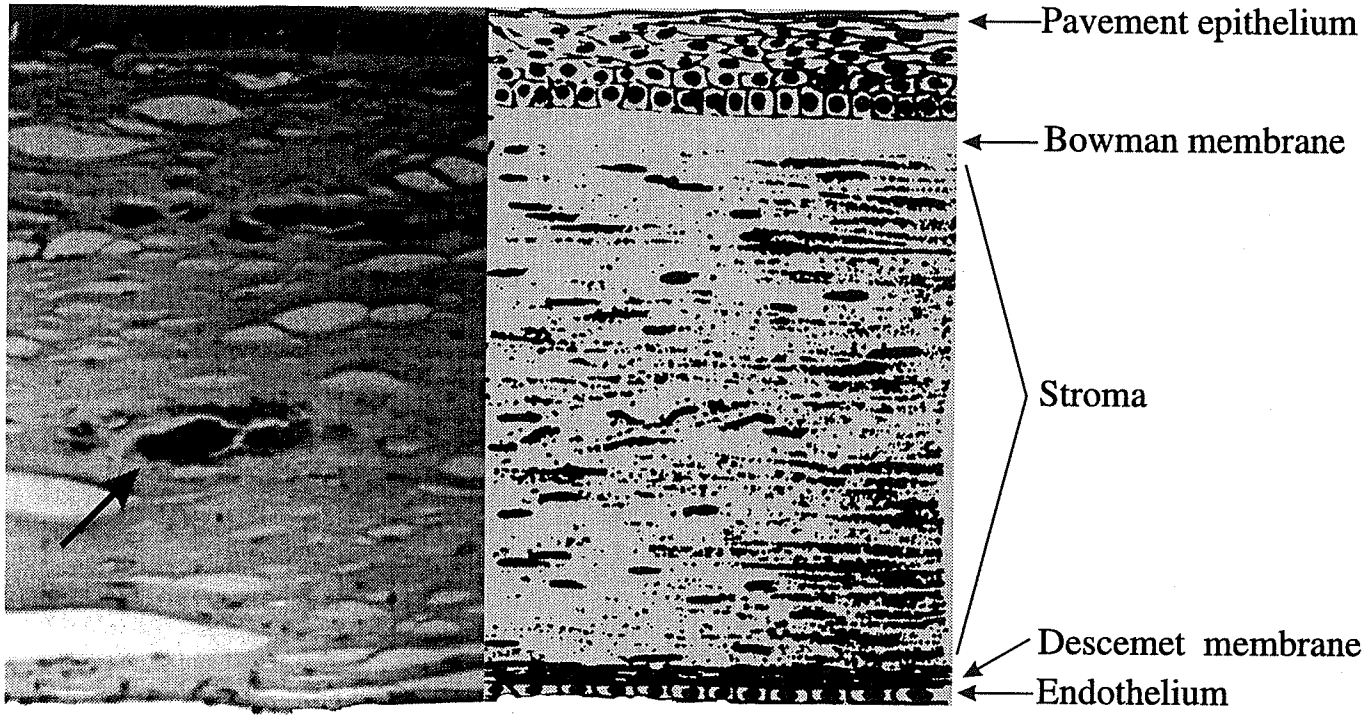


Figure 1: Photograph and line drawing of a corneal section illustrating the five layers of the cornea. The photograph depicts amyloid deposits in the centre of the stroma in an affected individual (→).

removing fluids from the stroma against the hydrostatic gradient of intraocular pressure, therefore preventing corneal edema (Hogan *et al*, 1971; Vaughan *et al*, 1992).

The avascular cornea receives nourishment from the limbal blood vessels, aqueous humour and the tears. Glucose and oxygen which diffuse through the epithelium and endothelium are of primary significance to the metabolism of the cornea. Superficially, corneal oxygen is also derived from the atmosphere. The cornea is innervated by the first ophthalmic division of the fifth (trigeminal) cranial nerve. These sensory nerves penetrate the epithelium, Bowman layer and the anterior of the stroma making the cornea extremely sensitive to pain from injury or inflammation. The three key factors contributing to and maintaining corneal transparency are its uniform structure, the absence of blood vessels and chemical composition and hydration of tissues (Hogan *et al*, 1971).

1.2. Hereditary Corneal Dystrophies

Heritable corneal dystrophies cause significant progressive visual impairment. As primary diseases of the cornea, they are characteristically bilateral and symmetric. However, unilateral and asymmetric cases have been described. In general, the dystrophies have a tendency to affect the central cornea more than the periphery. Corneal dystrophies are anatomically classified according to the affected corneal layer (Waring *et al,* 1978; Arffa, 1991). Transmission electron microscopy analysis has further aided in the classification process (Arffa, 1991). As specific genes causing corneal dystrophies are identified and responsible mutations determined, a gene-based classification system will be

most useful in distinguishing between the different dystrophies and improve our understanding of their pathogenesis (Folberg *et al,* 1994).

Treatment to relieve symptoms of CSD1 is generally symptomatic. Once vision is severely impaired, corneal transplantation is required. However with time, in several of the corneal dystrophies, the disease can recur in the corneal graft (Arffa, 1991; Krachmer and Palay, 1991). Although the frequency of corneal dystrophies is rare, the relative frequency varies with the geographic distribution of affected families (Waring *et al,* 1978). As in Manitoba, lattice corneal dystrophy type 1 has had a major impact on one large multigeneration kindred that has formed the basis of the study described herein.

1.2.1. Lattice Corneal Dystrophy

In 1890, Biber first described the corneal condition "gitterige keratisis". Today, it is referred to as lattice corneal dystrophy type 1. The hereditary nature was confirmed by Haab (1899) and Dimmer (1899) as cited by Ramsay (1958). Fuchs (1925) reexamined the original cases of Dimmer including other family members, in addition to new cases from which he was able to evaluate and describe the clinical features from early to late stages (cited by Ramsay, 1958). Klintworth (1967) provided evidence that lattice corneal dystrophy was an inherited amyloidosis specifically expressed in the cornea. In 1969, Meretoja described another type of lattice corneal dystrophy which was associated with systemic amyloidosis referred to as Finnish hereditary amyoloidosis. To distinguish between these two types, the classic inherited form of Biber-Haab-Dimmer corneal dystrophy was designated lattice corneal dystrophy type 1 (LCD1). The Meretoja form of

lattice corneal dystrophy was designated as type 2 (LCD2). A third, a Japanese type, was classified as type 3 (LCD3) (Gorevic *et al,* 1984). As a consequence of the recent mapping of LCD1 and two other corneal stromal dystrophies to chromosome 5q (Stone *et al,* 1994), it is now referred to as corneal stromal dystrophy type 1 (CSD1).

Lattice corneal dystrophy type II is inherited in an autosomal dominant manner. It presents with blepharochalasis, bilateral facial-nerve palsies, peripheral neuropathy and systemic amyloidosis (Krachmer and Palay, 1991). The characteristic lattice lines are less numerous, but coarser than those observed in CSD1. They are radially oriented and extend to the limbus. Although symptoms usually manifest by 20-35 years of age, vision acuity is generally preserved until the seventh decade of life. In the Meretoja form of lattice corneal dystrophy, LCD2 is associated with Finnish hereditary amyloidosis where the amyloid protein is derived from an abnormal variant of gelsolin, an actin binding protein (Haltia *et al,* 1990).

Lattice corneal dystrophy type III appears to be inherited in an autosomal recessive manner, presenting late in life, between 70 and 90 years of age with no previous history of corneal erosions or systemic amyloidosis. Amyloid deposition results in thick translucent lattice lines in the cornea (Hida *et al,* 1987). The deposits are composed of serum amyloid P component (SAP); therefore, the disease-causing gene may be located at 1q12-q32 where the SAP gene has been mapped (Mondino *et al,* 1981).

1.2.1.1. Corneal Stromal Dystrophy Type 1

CSD1 is inherited in an autosomal dominant manner with incomplete penetrance and variable expressivity. A disorder manifesting localized amyloidosis, CSD1 is limited to the stromal layer of the cornea with no relationship to any known primary or secondary systemic amyloidosis (Ramsay 1960; Waring *et al,* 1978; Krachmer and Palay 1991). The onset of CSD1 is generally in the first or second decade, although cases have been described where it was not observable until middle age. Early corneal lesions can initially involve the anterior stroma and Bowman layer appearing as small refractile lines, white dots and dashes or as diffuse haze in the central stroma by slit-lamp examination. With progression, the superficial lines become thicker, branching and radially oriented. The white dots and dashes appear as small nodules or thread like spicules. These findings are accompanied by recurrent epithelial erosions causing pain, photophobia and redness (Waring *et al,* 1978; Arffa, 1991). Some individuals with CSD1 experience corneal erosions without any evidence of stromal involvement. Yet others do not experience recurrent corneal erosions, but present with lattice lines and dots (Waring *et al,* 1978). With time, the central diffuse opacity becomes more dense due to recurrent corneal erosions or the process of amyloid accumulation and the lattice lines can extend into the deep stroma and epithelium. Later in life along with diminishing corneal sensation, the epithelial erosions become less frequent. The opacities coalesce by becoming more dense and extending to the corneal periphery which mask the lattice lines. Occasionally corneal vascularization develops secondary to the inflammation caused by corneal erosion (Waring *et al,* 1978; Arffa, 1991).

In general, by the third or fourth decade, there is marked visual impairment which becomes severe by the fifth or sixth decade (Ramsay, 1960; Waring *et al,* 1978). Treatment for recurrent corneal erosions is artificial tears, topical hypertonic solutions and ointments, pressure patching and soft contact lenses (Waring *et al,* 1978). Significant vision loss necessitates corneal transplantation. Yet, the dystrophy can reoccur 2 to 14 years postoperatively via infiltration from the rim of the original cornea which serves as an anchor for the new cornea. Recurrent erosions and progressive visual loss are again experienced culminating in the need for repeat corneal transplantation 5 to 15 years later (Ramsay, 1958; Waring *et al,* 1978). Two types of opacities appear in the corneal grafts: (1) dot-like, diffuse or filamentous subepithelial opacities or (2) faint peripheral stromal lattice lines and dots. These opacities do not reoccur centrally in the cornea and are not consistently composed of amyloid (Waring *et al,* 1978).

When examining isolated or atypical cases, CSD1 can occasionally be misdiagnosed as primary recurrent corneal erosions, herpes simplex, superficial punctate keratisis and keratoconjunctivitis eczematosa (Ramsay, 1958; Tsubota *et al,* 1987).

1.2.1.2. Pathology CSD1

Three components in the stroma which characterize CSD1 are refractile lines, white dots and central diffuse opacity. The characteristic lattice lines of CSD1 give a "lacework" appearance against red fundus reflection. Under direct illumination, the lines show irregular margins and occasionally assume a beaded appearance caused by the accumulation of small white dots within them (Waring *et al,* 1978). Under

retroillumination, the lines are typically refractile with a light and dark side and a clear core (Arffa, 1991). The lattice lines fluoresce in unstained and fluorochrome-stained histologic sections (Klintworth, 1967).

Early in disease, the intervening stroma is clear, but becomes progressively more hazy due to fine discrete white dots between the lattice lines. The opacities appear flake like, fluffy and slightly stellate under direct illumination and refractile under retroillumination. At this stage, it is often difficult to distinguish between CSD1 and granular corneal dystrophy. The use of electron microscopy techniques can reliably differentiate the deposits charactersitic of the two dystrophies, revealing the CSD1 deposits to be composed of fine, electron-dense, nonbranching fibrils 8 to 10 nm in diameter without periodicity. Many of the fibrils are randomly arranged, but most are systematically aligned to exhibit the phenomenon of birefringence and dichroism (Waring *et al*, 1978; Arffa, 1991).

Corneal lesions in the stroma viewed through the light microscope reveal a fusiform deposit of amyloid that pushes aside collagen lamellae. Bowman layer may be absent or partially replaced by the deposits and by irregular connective tissue. The deposits cause variations in thickness of the epithelial layer. The basal epithelial cells degenerate and the basement membrane is usually thickened and fragmented (Arffa, 1991). The population of stromal fibroblasts decreases due to degeneration of cellular components. Many fibroblasts appear metabolically active with increased amounts of rough endoplasmic reticulum and golgi apparatus. The corneal lesions are most concentrated anteriorly and may extend deeper into the stroma leaving the intervening

stroma clear. These findings correspond to the clinically observed lattice lines and dots (Waring *et al,* 1978).

1.2.2. Granular Corneal Dystrophy

Granular corneal dystrophy type 1 was originally described by Groenouw (1890, 1898, 1917). It is transmitted as an autosomal dominant trait with essentially full penetrance and variable expressivity. Initially, small discrete, sharply demarcated, grayish white opacities appear in the anterior axial stroma. At this stage, vision is not impaired as the stroma between the opacities remains clear and the epithelium remains regular. As the disease progresses, the opacities enlarge, increase in number, coalesce and extend deeper into the stroma (Waring *et al,* 1978). The opacities vary in size, shape and appearance: Some are round with solid or relatively clear centers, others resemble a "christmas tree" or "snowflake" like pattern or may look like a sponge imprint (Arffa, 1991). Visual acuity decreases gradually due to opacification of the stroma between granular spots and epithelial irregularity. The opacities can extend to the periphery of the cornea but not as far as the limbus. Painful corneal erosions do not commonly occur (Waring *et al,* 1978; Arffa, 1991).

Histopathologically, the eosinophilic hyaline deposits occur primarily in the central stroma and beneath the epithelium and stain a brilliant red with Masson's trichome. The peripheral regions of deposits have, in some cases, stained positive with Congo red (Brownstein *et al,* 1974). Electron microscopy reveals the extracellular deposits to be electron dense rod-shaped or trapezoid-shaped 100-500 μm wide. The morphology of the

randomly oriented rods are homogeneous, filamentous or moth-eaten in appearance. The stromal keratocytes show varying levels of degeneration with dilation of endoplasmic reticulum and golgi apparatus, swelling of mitochondria and cytoplasmic vacuoles. Histochemical characterization indicates the hyaline deposits to be non-collagenous protein composed primarily of tyrosine, tryptophan, arginine and sulphur- containing amino acids. The phospholipid, phosphatidylcholine, has been identified in the hyaline deposits. For most patients, treatment is not required. If vision is markedly reduced, surgery will restore vision. However, disease often reoccurs in corneal grafts. (Waring *et al,* 1978; Arffa, 1991; Krachmer and Palay, 1991).

1.2.3. Avellino Corneal Dystrophy

Folberg *et al* (1988) described a condition in four patients with histologic features of both lattice and granular corneal dystrophy. The three families originally studied and subsequent families trace their ancestry to the Italian province of Avellino (Holland *et al,* 1992; Rosenwasser *et al,* 1993). However, Avellino corneal dystrophy has been reported in a family of German and Swiss origin (Rosenwasser *et al,* 1993) as well as in an isolated case of Colombian origin (Dolmetsch *et al,* 1996), both with no known Italian ancestry. Avellino corneal dystrophy is inherited in an autosomal dominant manner with essentially full penetrance and moderate variable expressivity. Clinically, the characteristic signs present in the following sequential manner: first, discrete, gray-white granular deposits appear in the anterior stroma, followed by the formation of lattice lesions in the mid to

posterior stroma and lastly the emergence of anterior stromal haze (Folberg *et al,* 1988; Holland *et al,* 1992; Rosenwasser *et al,* 1993).

With increasing age, the granular lesions become larger and often coalesce to form linear opacities particularly in the inferior cornea. The lattice lines initially in the mid and deep stroma, begin to involve the entire stroma. As the disease progresses, stromal haze is only evident in patients with advanced stages of granular and lattice opacities. Under light microscopy, the superficial granular deposits stain red with Masson's trichrome stain. The deeper lattice-like deposits stain positive with Congo red and demonstrate birefringence and dichroism, characteristics of amyloid. Recurrent corneal erosions are experienced by Avellino patients, which are not commonly observed in typical granular corneal dystrophy cases (Folberg *et al,* 1988; Holland *et al,* 1992; Rosenwasser *et al,* 1993). Treatment is symptomatic. If corneal transplantation is required, in time the dystrophy reocurs with only granular deposits clinically evident (Holland *et al,* 1992).

1.2.4. Reis-Bücklers Corneal Dystrophy

There are many similarities in the clinical, histological and natural history of Reis-Bücklers corneal dystrophy and granular corneal dystrophy to suggest that they may be the same entity (Moller, 1989; 1991). However, consensus has not been reached in this respect. Traditionally, Reis-Bücklers corneal dystrophy is classified as Bowman membrane/anterior membrane dystrophy. It is inherited in an autosomal dominant manner with essentially full penetrance and variable expressivity. The initial manifestation involves recurrent erosions which cause photophobia, pain and ocular redness. The frequency of

these episodes decreases with increasing age. The earliest findings are reticular opacifications in the central cornea at the level of Bowman layer. As the disease progresses, the cornea becomes more clouded by superficial gray-white opacities. The density of the opacities is greatest in the central or midperipheral cornea, leaving a transparent zone at the extreme periphery. The opacities form ring-like, honey-combed, geographic or linear patterns localized to Bowman layer that focally project into the epithelial layer. This causes the corneal surface to become rough and irregularly thickened resulting in irregular astigmatism (Waring *et al,* 1978; Arffa, 1991). The abnormal epithelial architecture is evident by cytoplasmic vascularization, mitochondrial swelling and clumping of nuclear chromatin in many epithelial cells and especially the basal cells. In advanced cases, the anterior stroma has a fine granular appearance. Bowman layer is completely replaced by fibrocellular connective tissue consisting of a mixture of randomly arrayed collagen fibrils of 2.5 to 4 nm in diameter and short dense, curly fibrils 0.8 to 1.2 nm in diameter. The posterior stroma, Descemet membrane and endothelium remain unaffected (Waring *et al,* 1978). As there are no histochemical stains that characterize Reis-Bücklers corneal dystrophy, the composition and origin of these fibrils is yet to be determined (Arffa, 1991).

1.3. Amyloid

Studies by Virchow (1854) established amyloid as a pathological entity (cited by Benson, 1995). He chose the term amyloid based on his findings of cellulose-like deposits in tissues which gave a unique color reaction when stained with iodine and sulfuric acid.

Despite the demonstration by Friedrich and Kekule (1859) that amyloid deposits were composed of protein, the term "amyloid" has endured (cited by Benson, 1995). It is now thought the observations by Virchow had resulted from positive iodine staining of carbohydrates such as heparin sulfate proteoglycans which appear to be ubiquitous constituents of amyloid deposits (Snow *et al*, 1987).

Amyloid appears homogenous, eosinophilic and amorphous under the light microscope in tissue sections. Various staining techniques and polarization microscopy provide a definitive test for the presence of amyloid. In the past, metachromatic staining with aniline dyes such as methyl violet and crystal violet were used to identify amyloid. Presently, positive Congo red staining is the standard method for identification of amyloid deposits in tissue sections. With polarizing optics, unstained amyloid deposits exhibit birefringence, which implies the presence of a highly ordered structure. When stained with Congo red, the amyloid deposits take up the dye and emit a characteristic apple-green color when viewed with polarizing light. Under similar conditions, the increased intensity of birefringence displayed by amyloid deposits is not observed with other fibrillar structures such as collagen, elastin and fibrin. The increased intensity is explained by the linear dye molecules binding to the amyloid fibril axis in a parallel fashion. Collagen also appears birefringent in tissue sections. However, an apple-green birefringence is not observed because collagen does not bind to Congo red. Fluorescent dyes such as thioflavin are a sensitive staining method for detecting the presence of amyloid deposits in brain and other tissues. The presence of amyloid fibrils must be confirmed by Congo red staining (Sipe, 1992; 1994; Benson, 1995).

In 1959, Cohen and Calkin and Spiro characterized the amyloid deposits as non-branching fibrils of intermediate length with diameters of 0.75 to 1.00 nm. These non-branching fibrils appear to consist of two to several parallel subunit filaments measuring 25 to 35 nm in width. The fibrils often have a beaded appearance caused by the helical twisting about each other. Findings from X-ray diffraction studies and infrared spetroscopic analysis of amyloid fibrils reveal that the polypeptide backbone of amyloid fibril proteins is arranged in a cross-beta-pleated sheet conformation, is anti-parallel and is oriented perpendicularly to the long axis of the fibril. The amyloid fibril precursors are small, 3 to 30 kDa polyanionic proteins with high glutamic and aspartic acid and serine content. The polypeptide subunits are polymerized in each filament to force the adjacent polypeptide chains to assume the β-pleated sheet conformation. Folding is in an anti-parallel manner where the carboxyl and amino termini of adjacent polypeptide chains are oriented in opposite directions. The stability of the β-pleated sheet conformation not only imparts resistance to proteolytic digestion but is also insoluble under physiological conditions (Sipe, 1992; 1994; Benson, 1995).

The chemical characterization of amyloid was first elucidated in 1971 by Glenner *et al.* Here, amyloid fibrils were solubilized from tissues of patients presenting with primary amyloidosis. Following isolation and amino acid sequencing of the major subunit, results showed homology to the variable segment of immunoglobulin light chains. Subsequently, amyloid fibrils from patients with secondary amyloidosis were found to be composed of a previously undescribed protein, which was named "amyloid A protein" (AA) (Levin *et al*, 1972). Further studies revealed AA was derived from a circulating

precursor, serum amyloid A (SAA) shortened at the carboxyl terminus by proteolytic cleavage (Levin *et al*, 1973). The amyloidogenic protein first implicated in hereditary amyloidosis was identified by Costa *et al*, (1978) who showed the amyloid material from patients to react with antisera to plasma transthyretin (prealbumin). To date, over 45 mutations in the transthyretin gene have been identified and their involvement in hereditary amyloidosis described (Benson, 1995). Mutations in other plasma proteins including apolipoprotein A-I (Nichols *et al*, 1990), gelsolin (Maury, 1991) and α- fibrinogen (Benson *et al*, 1993) have also been found to be associated with hereditary amyloidosis.

Amyloidosis is the end result of a sequential process which begins with the polymerization of amyloid fibril precursor proteins which are assembled into filaments and eventually accumulate as bundles of fibrils and form deposits either localized or systematically distributed throughout the body. These deposits disrupt the architecture and function of normal tissue by invading extracellular space and ultimately causing cell death. Although the primary structure of amyloid precursors and fibrils differ widely, some common features are shared. First, the amyloid fibrils conform to the β-pleated sheet secondary structure. Second, the calcium dependent association of SAP which specifically binds to the fibrils. Third, the fibrils and deposits are intimately associated with sulfated glycosaminoglycans (Sipe, 1994).

There are several factors associated with amyloid formation. First, in the majority of cases, the amino acid sequence reflects the translated product of the normal gene. However, in some amyloid precursors, amino acid substitutions lead to self-aggregation. Second, both partially digested and intact precursor proteins have been found to coexist

within amyloid deposits, suggestive of possible defects in proteolytic processing as well as the need for a specific molecular mass of amyloid constituents. Third, the amyloid deposition process may be affected by the high concentration of precursor proteins either locally or in circulation. Furthermore, amyloid formation may be initiated by either early secondary and/or tertiary conformational changes of precursor molecules influenced by physiochemical components such as pH and the presence of free radicals and metal ions in the microenvironment where amyloid deposition occurs. Fourth, in these regions, undefined tissue-specific factors (cellular receptors, enzymes or extracellular matrix components) may account for the selectivity of amyloid localization. And lastly, the presence of amyloid enhancing factors (AEF), although uncharacterized biochemically are associated with several forms of amyloidosis and functionally described as accelerants of amyloidosis (Sipe, 1994; Castaño and Frangione, 1995).

The cellular origin and biochemical nature of selective amyloid deposition in CSD1 is not known. The earliest deposits are localized to the central region of the avascular cornea where there is a preponderance of fibroblasts, extracellular matrix precursors and collagen. Three hypothesis have been put forth to explain the origin of amyloid in CSD1. The first hypothesis suggests that fibroblasts which are known to synthesize and secrete filamentous precursors of collagen fibers, may also synthesize amyloid fibers. The observed close proximity and occasional intermingling of structurally intact collagen fibrils with extracellular deposits of amyloid led to the second hypothesis, that amyloid fibrils may arise from collagen degradation. However, both amyloid and collagen maintain their unique morphologic characteristics, thus implying that collagen degradation may not lead

directly to amyloid deposition. The third hypothesis speculates on the production of amyloid originating from other tissue sites and depositing in the cornea via the limbal blood vessels. However, the earliest corneal lesions occur furthest from the limbal blood vessels which suggests the amyloid does not reach the cornea via the bloodstream. Based on the complex chemical composition of the cornea, the involvement of other local factors in amyloid deposition cannot be excluded (Klintworth, 1971;Waring *et al*, 1978).

The reoccurrence of disease in corneal transplants begins at the periphery of the transplanted cornea, possibly from the rim of the original diseased cornea. These findings imply either the precursor proteins are derived from the circulation or fibroblasts from the rim of the original cornea secrete amyloid into localized regions of the transplant. The stromal lattice lines or white spots generally do not reoccur in the central region of the cornea. (Klintworth 1967; 1971; Waring *et al*, 1978).

The immunohistochemical characterization of amyloid deposits is based on the recognition of specific antisera to protein components. Attempts have been made to characterize the amyloid deposits specifically in CSD1. The SAP component, common to all forms of secondary and several localized amyloidosis was found in the corneal amyloid deposits (Mondino, 1980; Gorevic *et al*, 1984). An initial study detected protein AA reactivity in the stromal deposits of CSD1 (Mondino, 1980), which was not confirmed in , a subsequent detailed study (Gorevic *et al*, 1984). Protein AA is involved in reactive amyloidosis where a familial pattern of inheritance has only been observed in familial Mediterranean fever and Muckle-Wells syndrome (Benson, 1995).

The oligosaccharides associated with amyloid deposits in CSD1 were determined by screening a panel of lectins (Bishop *et al*, 1991). These are a specialized group of proteins or glycoproteins primarily derived from plants which have the ability to bind to specific classes of oligosaccharides. With the use of histochemical staining techniques, the distinct labeling patterns for each lectin reflects the distribution of specific carbohydrates (Freeman, 1992). Those with an affinity for corneal amyloid deposits have been identified and revealed the modifying oligosaccharides to consist of mannose/glucose, sialic acid/N-acetylglycosamine residues, terminal galactose residues and oligosaccharides with β-galactose-N-acetylgalactosamine disaccharides. The cellular mechanisms involving glycoproteins and the deposition of amyloid specifically to the stromal layer of the cornea in CSD1 is not clear and yet to be elucidated (Bishop *et al.* 1991).

1.4. Genome Maps

A genome map is defined as a linear representation of DNA which reflects the organization of genomic landmarks in the context of a co-ordinate system. The major classes of genome maps are physical maps and genetic maps. The co-ordinate system for each of these is experimentally established by measuring the distance separating the genomic landmarks. Each map has a characteristic resolution range which defines its utility in studying the human genome (Green *et al,* 1995).

1.4.1. Physical Maps

Physical maps of DNA specify the distance separating genomic landmarks from one another. These maps are constructed from either total genomic DNA or cloned fragments of the genome. However, in the latter case, the ordering of the DNA segments is lost. The collection of clones is comprised of numerous copies of each DNA fragment. Identical DNA segments may also be present on overlapping fragments of differing sizes. To order these clones, they are screened for the presence or absence of genomic landmarks. When two clones are matched at a genomic landmark, they overlap at this point to reconstruct a larger DNA segment. By continuing to screen and match genomic landmarks in a sequential manner, a set of ordered clones called a "contig" is obtained. A contig represents a contiguous segment of DNA which can be used to construct a contig map (Cooper, 1994; Green *et al,* 1995)

1.4.1.1. Restriction Map

Restriction sites are a class of genomic landmarks available for constructing physical DNA maps. The restriction site is a signal for cleavage by a specific restriction enzyme. Long-range restriction maps of uncloned, total genomic DNA are constructed by using pulsed field gel electrophoresis (PFGE) to delineate the order and distance between recognition sites for rare cutting enzymes. The resolution of this type of physical map is between 50, 000 to several million base pairs (Rose, 1991). Restriction sites can also be analyzed within individual clones to establish a fingerprint pattern which can be used to determine the overlapping relationship among different clones (Green *et al,* 1995).

1.4.1.2. Sequenced Tagged Site Maps

The most recent genomic landmark to emerge is the sequenced-tagged site (STS) (Olson *et al,*1989) which is defined as a ~60 to 1000 bp stretch of unique DNA specifically detectable by PCR. STS-based maps reflect the relative order and distance between STSs within a region of DNA. There are several features that make STSs suitable tools for physical mapping. First, PCR detection is most desirable because of its high sensitivity, specificity and efficiency. Secondly, all relevant information pertaining to each STS can be electronically stored and accessed by investigators. In this way, any DNA segment defined by a STS can be synthesized under the appropriate PCR conditions with the appropriate oligonucleotide primers (Green *et al,* 1995).

To construct physical maps of human chromosomes, STSs are the genomic landmarks and yeast artificial chromosomes (YACs) are the source of cloned DNA. YACs can accommodate a DNA fragment of 0.1 to 1 Mb (Green, 1990; Schlessinger, 1990). The magnitude of the insert size simplifies the process of assembling contigs within a chromosomal region. YAC clones are screened for the presence of STSs and sized by PFGE. Then, they are arranged in a sequential manner by overlapping the regions where clones contain one or more common STSs. Once the order of the clones is established, the resulting contig maps provide information about the relationship between the clones and relative order and distribution of STSs (Green *et al,* 1995).

YACs are an invaluable resource for covering large regions of the genome, however, there are some limitations. First, a high proportion of chimaeric clones containing non-contiguous genomic sequences have been observed. Additionally, the

presence of widely dispersed homologous sequences in the genome can lead to incorrect contig assembly. Furthermore, some human sequences are not amenable to cloning into yeast or exhibit instability once cloned. For these reasons, YAC contigs often contain deletions and rearrangements (Cooper, 1994).

1.4.1.3. Radiation Hybrid Maps

Radiation hybrid mapping is an alternative approach to order physically genomic landmarks along a chromosome. The success of this experimental technique relies on the ability to recover fragments of human chromosomes from rodent cell lines. Human-rodent hybrid cell cultures containing a single human chromosome are exposed to lethal X-ray irradiation. The resulting human chromosome fragments are rescued by fusion of the irradiated cells to another rodent cell line. Each of these cell lines contains fragments of specific human chromosomes which can be analyzed for the presence or absence of DNA markers such as STSs by PCR. The coexistence of markers within a collection of cell lines provide information about both their order and physical location. The closer two markers are on a chromosome, the less likely they will be separated by breakage caused by irradiation. The information obtained can be used to assemble STS-based radiation hybrid maps. In turn, these maps complement the YAC clones ordered by STS analysis, by helping to order disjointed YAC contigs and orient YAC contigs relative to the centromere and telomeres (Cooper, 1994; Green *et al,* 1995).

1.4.1.4. Sequence Maps

DNA sequence map reflects the precise order of the nucleotide sequence along a stretch of DNA and represents the highest resolution physical map. As the Human Genome Project advances, the complete sequence of the human genome will be eventually established. The resulting DNA sequence will be a composite of different individuals and reflect an individual's DNA for >99% of the genome (Green *et al,* 1995).

As less than 5% of the human genome is expressed, DNA sequence maps are being assembled which depict only the transcribed sequences. However, a map based only on sequenced cDNA has a number of limitations. First, this type of map provides no information about location, structure or regulation of the gene. Secondly, cDNA sequences do not give any insight into cases where multiple mRNAs are produced from the same gene or where one type of mRNA is produced from multiple genes. Lastly, a portion of transcribed sequences may go undetected due to low levels of expression, tissue-specific expression or species-specific expression (Rose, 1991)

1.4.2. Genetic Maps

Genetic maps also known as linkage maps reflect the relative location of genomic landmarks in an interval of DNA (Green *et al,* 1995). These maps are constructed by measuring the probability that two closely spaced genomic landmarks on a chromosome will be inherited as a unit and not be separated by recombination during meiosis. As the genetic distance separating the landmarks increases, the frequency of recombination increases. Likewise, the smaller the recombination fraction, the smaller the estimated

genetic distance. The unit of measure is the centiMorgan (cM), where 1 cM corresponds to a 1% frequency of recombination between two markers during meiosis. There is no direct relationship between genetic and physical distance due to sex-specific and regional-specific differences in the recombination frequency across the human genome (Renwick, 1969). It has been estimated that 1 cM on a linkage map on average corresponds to ~1 to 3 Mb on a physical map (Donis-Keller *et al*, 1987).

1.5. Linkage Analysis

Linkage analysis is an effective tool to identify and map genes responsible for inherited disease. Genetic linkage reflects the proximity of two genomic landmarks (genes, microsatellite loci) and their tendency to be inherited in a non-random manner in families. The closer two loci are situated on the same chromosome, the less likely their respective alleles will be separated by homologous recombination during meiosis. Loci that are situated physically close together along the chromosome are described as linked and the alleles will tend to be inherited together. However, if the loci are farther apart, linkage is disrupted by homologous recombination during meiosis resulting in new alleles at various loci on the chromosome pairs. As the distance between the loci increases, the frequency of recombination defined as the recombination fraction (θ), also increases. The recombination fraction varies in value from 0.0 which signifies no recombination to 0.5 which signifies independent assortment. When two loci on the same chromosome are located far enough apart that they assort independently, the probability of recombinant and non-recombinant chromosomes is equal (θ=0.5), similar to when loci reside on different

chromosomes (reviewed by Conneally and Rivas, 1980). The genetic distance between loci is estimated from the recombination fraction where 1 cM corresponds to 1% frequency of recombination. In other words, the recombination fraction represents the probability of 1 recombination event on average for every 100 meioses (Renwick, 1969).

Two assumptions are inherent to detecting linkage. First, the genetic mechanism of the disease must be known. However, for many complex diseases, the pattern of inheritance is not known and must be determined by performing linkage analysis under various models of inheritance. Second, the alleles at the disease and marker locus must be inherited in an autosomal codominant fashion. In an informative mating, one parent must be heterozygous at both loci under consideration to permit the segregation of alleles in the family to be followed unambiguously. If this is not the case, segregation can not be clearly defined and the mating is non-informative for the marker under study. Although family members, such as grandparents and grandchildren, are not used directly in the calculation of lod scores, they are helpful in assigning the parental genotypes in a nuclear family. The analysis of grandparents is particularly useful when the informative parent is a double heterozygote for the two loci being studied. In this case, two possible phases exist: if the marker allele and disease allele are on the same homologue, they are in coupling; alleles on opposing homologues are in repulsion. By determining the phase of the parent, it permits the classification of offspring as either recombinants or nonrecombinants. Phase-known families provide more linkage information than phase-unknown families. Thus, the strength of evidence in support of linkage is much greater in phase-known families compared to phase-unknown families (reviewed by Risch, 1992; Khoury *et al,* 1993).

Since human families tend to be relatively small and phase information is not always obtainable, Morton (1955) developed a sequential method for testing linkage based on two-point lod score analysis. Two-point lod scores can be calculated by inspection using lod score tables or generated with the aid of computer software packages such as MARK III (Côté, 1975) and LINKAGE (Lathrop *et al*, 1984; Lathrop and Lalouel, 1984; Lathrop *et al*, 1986). Regardless of the tools chosen, the maximum likelihood approach to linkage analysis measures the likelihood of the observed association between two traits in families under two hypotheses. Under the null hypothesis of no linkage, the two traits are not linked and have a recombination fraction equal to 0.5. The alternate hypothesis presumes the two loci are linked at a recombination fraction of <0.5. To conclude, statistically, linkage between two traits, an odds ratio, L, is determined by the ratio of the probability of observing the distributional pattern of two traits in a family under the alternate hypothesis of linkage at a given θ, to the same probability under the null hypothesis of no linkage. The lod score, the decimal logarithm of L, is calculated at several values of θ. The maximum lod score at its corresponding recombination fraction can be obtained graphically or with the aid of computer programs. If the meioses being scored are phase-known, the number of recombinants (R) and non-recombinants (NR) can be determined. If the mating is phase-unknown, then a Z score is used to represent children of each type. Sequential testing permits families to be sampled until conclusive evidence either for or against linkage is accumulated (reviewed by Conneally and Rivas, 1980; Khoury *et al,* 1993).

Morton (1955) established the conventional threshold for asserting linkage is a lod score ≥ 3. At this level, the observed segregation between the two traits is 1000 times more

likely to have arisen due to linkage than by chance alone. Thus, the posterior probability of linkage is 95% and the posterior probability of no linkage, or a probability of false positive is 5%. A lod score of -2.0 or less signifies a 100:1 odds against linkage at some specified value of θ. Lod scores between -2.0 and 3.0 summed over independent matings neither support linkage nor allow linkage to be excluded and require more data to be collected. When multiple loci are considered, if two are known to be located on the same chromosome, a lod score of 2.0 is sufficient to support linkage (Keats *et al*, 1987)

To study human disease under conventional methods of linkage analysis, it must be monogenetic and follow an autosomal or X-linked dominant or recessive pattern of inheritance (Risch, 1992). The primary assumption underlying this linkage study is that CSD1 conforms to a single locus, autosomal dominantly inherited Mendelian disorder. The CSD1 kindred shows vertical transmission with multiple affected generations. The offspring of affected parents, regardless of sex gender, have a 50% risk of developing the disease. The responsible gene appears to be 100% penetrant in family members 25 years or older. There is no evidence for reduced penetrance as each affected child has an affected parent.

1.6. Genetic Markers

1.6.1. Microsatellite Markers

Short tandem repeat polymorphisms (STRPs), commonly called microsatellites, are highly polymorphic genetic markers exhibiting variation in the number of di-, tri-, tetra- and pentanucleotide repeat motifs. These unique markers tend to have several alleles which are inherited in a co-dominant Mendelian manner (Weber and May, 1989; Sheffield *et al,* 1995).

As microsatellites occur abundantly and randomly throughout the human genome and are amenable to assay by PCR, they are used extensively in genome-wide searches for linkage analysis and gene mapping and for diagnostic tests.

The core repetitive region generally consists of 10-15 tandem repeats of 1-6 base pairs (Litt and Luty, 1989; Weber, 1990). All nucleotide combinations have been identified within the repetitive region. The abundance of different motifs is highly skewed and differs significantly from the predicted distribution based on the frequency of nucleotides in the human genome (Ashley and Dow, 1994). If spaced uniformly throughout the genome, tandem repeats of $(dC\text{-}dA)_n$-$(dG\text{-}dT)_n$ are the most commonly occurring at an average frequency of 30-60 kb (Litt and Luty, 1989).

The overall chromosomal distribution of microsatellite markers varies from chromosome to chromosome, with under representation on chromosomes 9, 19 and 21, for example, and over representation on chromosomes 14, 17 and 18. Cytogenetic correlations at the intrachromosomal level are inconclusive as no systematic trend has been observed for either G or R bands which might be rich or poor in (CA/GT)n microsatellite markers. The distribution of microsatellite markers in telomeric regions ranges from very poor to absent (Weissenbach *et al,* 1992; Weissenbach, 1993b). Tri- and tetranucleotide repeats occur much less frequently in the genome and differ considerably in their distribution from (CA/GT)n repeats (Gastier *et al,* 1995; Sheffield *et al,* 1995). The most abundant class of trinucleotide repeats in (AAT)n, occurring approximately every 500 kb. The least frequent class, (ACT)n occurs approximately every 25,000 kb (Gastier *et al,* 1995). Preliminary data suggest that tetranucleotide repeats of (GATA)n and (GGAA)n are relatively abundant in the genome. It is

thought the distribution of these markers will complement the distribution of dinucleotide repeat sequences in the genome (Sheffield *et al,* 1995).

The repetitive motif can be classified as perfect, imperfect, or compound (figure 2). Perfect repeats are uninterrupted tandem repeats without adjacent repeats of another sequence. Imperfect repeats are characterised by one or more interruptions of the core motif. Compound repeats are defined as different motifs located adjacent to each other. The length and type of the repetitive region is an indicator of the informativeness of the microsatellite marker. The informativeness of perfect repeats increases with increasing number of repeats, particularly in the range of 11-17 repeats. Imperfect repeats and compound repeats are generally less polymorphic than perfect repeats, thus less informative (Weber, 1990).

For microsatellite markers to be useful in linkage studies, they must fulfill the following criteria outlined by Conneally and Rivas (1980):

(1) The mode of inheritance of the marker must be clear. Microsatellite markers are inherited in a co-dominant Mendelian manner where the segregation of alleles can be followed in families.

(2) The phenotypic expression of each genotype must be evident. When screening microsatellite markers, individuals are characterized as homozygous or heterozygous at each locus on the basis of the inherited alleles. These alleles reflect the DNA sequence transmitted from parent to offspring.

(3) The locus must exhibit a high level of heterozygosity in the population. The term heterozygosity is defined as the chance an individual has of having two different alleles at a given locus. For a locus with only two alleles, the heterozygosity is generally 50%.

Perfect	TTTAGAAAAA**(AC)**$_{19}$CCCCAAAGCT
	CATGCACGTG**(CA)**$_{20}$ TACACCAGCT
Imperfect	TTGTTGATTT**(CA)**$_{11}$ CT**(CA)**$_4$ TACTGATGTG
	CTTTCTCAGGA**(CA)**$_7$GAG**(AC)**$_{14}$ AGGCAATGACA
Compound (perfect)	CCTTGTCTCT**(AC)**$_{16}$**(TC)**$_{10}$ AGCCAGGCAC
	GGCATGCATG**(CA)**$_9$**A(AC)**$_{19}$ A**(GA)**$_7$CATGCTGTTC

Figure 2: Examples of different categories of $(dC\text{-}dA)_n$-$(dG\text{-}dT)_n$ repeat sequences (Adapted from Weber, 1990).

In contrast, microsatellite markers with several alleles have heterozygosities in excess of 70% (Goodfellow, 1992). A measure of the probability of a marker locus being informative based on the number of alleles at a given locus and their frequency in the population is the polymorphic information content (PIC) (Botstein *et al*, 1980). The highest PIC scores are achieved for markers with many equally frequent alleles (Khoury *et al*, 1993). For linkage analysis, microsatellite markers with PIC > 0.7 are well suited because in these cases the parents are often heterozygous at the locus under study and the segregation of alleles can be followed unambiguously in the offspring. Genotyping highly polymorphic markers increases the informativeness of families (Hearne *et al*, 1992).

(4) It must be possible to detect phenotypic expression in tissues amenable to study. The polymorphic nature of microsatellites is expressed in DNA. The most readily available and accessible supply of genomic DNA is from white blood cells obtained by DNA extraction from whole blood. Thus, the tissue or organ affected by the disease in question does not need to be sampled. A permanent supply of genomic DNA can be established by *in vitro* transformation of lymphocytes. Other potential sources of genomic DNA suitable to microsatellite analysis include hair (Morin *et al*, 1992), cells from the buccal mucosa (Tautz, 1989) and fingernails (Kaneshige *et al*, 1992).

The highly polymorphic repetitive region characteristic of microsatellite markers is attributed to slipped strand mispairing (SSM). This mechanism of mutation has four main features. First, short simple tandem repeats that arise by chance in DNA sequences serve as the raw material for expansion by SSM of complementary bases at these short repeats during DNA

replication. Second, mutations such as base substitutions, insertions and deletions can give rise to new motifs which are further propagated by SSM events. This potentially gives rise to tandem or interspersed repeats of closely related motifs. The third feature, transitions (purines→purines or pyrimidines→pyrimidines) are more likely to be tolerated, thus occur with greater frequency than transversions (purine↔pyrimidine). As a result, there will be a tendency to maintain polypyrimidine and polypurine tracts within the repetitive region. Lastly, the abundance of methylated cytosine residues in the human genome at CpG islands may in fact favour TG repeats over GA repeats. At these CpG islands, the methylated cytosine can undergo deamination resulting in a transitions of C→T, thereby increasing the abundance of TpG regions (Levinson and Gutman, 1987; Tautz *et al,* 1986; Wright, 1994).

The repetitive regions characteristic of microsatellite markers are vulnerable to mutation. The mutation rate is not constant, varying from marker to marker. The published estimates of dinucleotide microsatellite markers is in the order of 10^{-3} to 10^{-5} (Weissenbach, 1993b). Trinucleotide repeats have a mutation rate approximately equivalent to dinucleotide markers. In contrast, tetranucleotide markers have a mutation rate estimated to be three to four times higher than that of dinucleotide markers (Gastier *et al,* 1995). It is necessary to consider the instability of microsatellite markers in studies such as linkage disequilibrium and other non-parametric genetic analyses where the difference between the mutated allele and a different allele is critical (Weissenbach, 1993a). In future, knowledge of mutation rates at specific loci could aid researchers in determining which genetic markers would be best suited to the study (Gastier *et al,* 1995). The high degree of variability of microsatellite loci may also be evident in genotypic differences between somatic and germinal cell lineages. For example, in

an apparently heterozygous individual (*a*,*b*), the allele *a* may be the result of a mutation of allele *b* in a portion of the parental germline. Artifacts which appear to distort the segregation of alleles in microsatellite analysis should be carefully examined by typing with adjacent markers. Although the mutation rate is an important consideration, in general, the alleles are inherited from generation to generation with high fidelity and can be reliably followed in family studies (Weissenbach, 1993a; Nelson, 1996).

Microsatellites are PCR-formatted such that the primers anneal to single-copy DNA flanking the repetitive element. The most commonly encountered pitfall genotyping dinucleotide markers are technical difficulties when scoring alleles. Most notably is the template switching phenomenon or strand slippage where mispairing of the template and newly synthesised strand during elongation across the dinucleotide tract results in a laddering effect. Tri- and tetranucleotide markers are less sensitive to these laddering artifacts and tend to exhibit better stability during PCR, thus, more easily interpretable results (The Utah Marker Development Group, 1995). The reason for the increased clarity of the banding pattern is not entirely clear. It may, in part, be due to the fact that a greater separation of alleles is achieved because they differ by increments of three or four bases. When genotyping dinucleotide repeats, it is useful to identify a heterozygote with two sets of multiple bands to distinguish the true alleles from the artifact bands. For each set, the band representing the true allele is more readily identifiable than the remaining "shadow bands". In general, the pattern of the extra bands is consistent within a given assay, therefore, the genotypes of different individuals can be deduced by only scoring the true alleles (Hudson, 1994).

There are five different factors which are potential causes of "shadow bands". The first three are technical and if appropriately controlled allow for dinucleotide repeats to be typed with relative efficiency and unambiguity. The first factor is the method of labelling the PCR product. Internal labelling is dependent upon the ratio of bases in the segment to be amplified. When a single radioactive nucleotide is used, if the two strands differ sufficiently in the ratio of the labelled nucleotide, strand separation does not interfere with scoring the alleles. If both strands are labelled with comparable intensity, then the true allele cannot be distinguished. This problem can be avoided by 5'-end labelling the forward or reverse PCR primer with polynucleotide kinase and (γ^{32}P) ATP to visualise only one of the two strands (Litt *et al*, 1993).

Secondly, the amplification of non-target regions of the genome can interfere with visualisation of the desired products. This problem is marker specific and can usually be minimised by appropriate primer design and sufficiently stringent annealing temperatures (Litt *et al*, 1993). Additionally, the concentration of Mg^{2+} in the reaction buffer can be increased and glycerol can be added to a maximum of 15% of the total reaction volume to improve amplification of the appropriately sized fragment. Another approach to eliminate non-specific amplification, yet facilitate efficient amplification of the target region is to digest the high molecular weight genomic DNA with an enzyme that cuts frequently, but not within the target region (Sharma *et al*, 1992).

The third potential cause of "shadow bands" is insufficient denaturing conditions during electrophoresis to separate the amplified fragments. Although 7-8.3M urea provides sufficient denaturing conditions for resolving PCR products of some dinucleotide markers, for others,

shadow bands and smears may still be evident even under these conditions. In these cases a formamide/urea system helps to eliminate these artifacts (Litt *et al,* 1993).

The final two factors responsible for shadow bands are: (1) the terminal transferase-like activity of *Taq* DNA polymerase which is responsible for the non-template dependent addition of a single nucleotide at the 3' end of the PCR product (Weber and May, 1989) and (2) the mutation phenomenon SSM described earlier, where the insertion or deletion of repeat units produces a 2 bp ladder of fragments (Hauge and Litt, 1993; Weber and May, 1989; Luty *et al,* 1990). From practical experience and evidence from the literature, the first three factors are technical in nature and can be controlled appropriately and the latter two factors do not prevent accurate genotyping of dinucleotide markers.

Additional pitfalls in genotyping PCR-microsatellite markers arise from minor to significant variations in the template DNA. The presence of an unidentified polymorphic motif nesting within the amplified region of interest (The Huntington Disease Collaborative Research Group, 1993) or duplication of the marker locus (Philips *et al,* 1993) can interfere with unambigous allele assignment. Secondary structures in the template DNA such as a stem and loop configuration have been observed to markedly reduce or inhibit PCR amplification (Oudet, 1991). Microdeletions encompassing the marker locus (Sander *et al,* 1994; Daniels *et al,* 1995; Taschner *et al,* 1995) or polymorphisms within the priming sequence for PCR amplification (Phillips *et al,* 1991; Weber *et al,* 1991; Callen *et al,* 1993) lead to reduced or complete loss of PCR product for a given marker resulting in a "non-amplifying" allele. Such a phenomenon is recognized when there is apparent non-inheritance of an allele from one or both parents (Daniels *et al,* 1995).

1.6.2. Trinucleotide Expansions and Disease

Several human genetic disorders have evolved from the expansion of trinucleotide repeats. This phenomenon represents a unique mechanism of mutagenesis which is often referred to as "allelic expansion" because the increase of repeats involves only one of two alleles carried by an individual. To date, X-linked spinobulbar muscular atrophy (SBMA) (LaSpada *et al*, 1991), fragile X (FRAXA) (Oberle *et al*, 1991), myotonic dystrophy (DM) (Harley *et al*, 1992), Huntington disease (HD) (The Huntington Disease Collaborative Research Group, 1993), spinocerebellar ataxia type 1 (SCA 1) (Orr *et al*, 1993), dentatorubral pallidoluysian atrophy (DRPLA) (Nagafuchi *et al*, 1994) and Machado-Joseph disease (MJD) (Kawaguchi *et al*,1994) are caused by trinucleotide expansion in the disease-causing gene. Two aspects common to these diseases are genetic instability and anticipation. The expanded region is unstable in parent to offspring transmission. Furthermore, the observed phenotypic variation is genearally related to the size of the expansion which has a tendency to increase with each successive generation and larger repeats are associated with more severe disease. This phenomenon is referred to as anticipation (LaSpada *et al,* 1994).

The instability of the trinucleotide repetitive region is generally related to the number of repeats, where an increase in the number of repeats increases the likelihood of change. In the normal range, these repeats are small, stable, polymorphisms with low mutation rates. However, when the repetitive region reaches a critical size threshold, the potential disease-causing gene becomes highly unstable. The likelihood and magnitude of expansion is related directly to maternal or paternal transmission (LaSpada *et al,* 1994;

Nelson 1996). In general, paternal transmission is more frequently associated with expansion because the paternal line is subject to more cell divisions per generation than the maternal line (Jennings 1995). There is apparent directionality in the instability of repeats, with expansion more likely than contraction. Furthermore, the limitation of expansion to only CGG and CAG trinucleotide repeats implies these sequences adopt some alternative DNA structures which contribute to the mechanism of instability (Nelson 1996).

The magnitude of instability in the repetitive region is divided into two classes. In the first class, the mutations are small scale changes demonstrating a two-fold increase in the original repeat length. Slipped strand mispairing may in part be responsible for these mutations. It is postulated that a defective mismatch repair system in families with trinucleotide expansion disorders may contribute to the expansion. Although no evidence has been found to support this hypothesis, the particular timing and tissue specific location of these changes may render the mismatch repair system ineffective during critical periods of development, such as in embryogenesis or gametogenesis. The second class of mutations is characterized by both small and large scale changes of up to 20-fold in a single transmission. Slipped strand mispairing is not believed to be a likely explanation for large scale type changes (Nelson, 1996).

Small scale instability gives rise to expansion of CAG repeats in five dominantly inherited neurodegenerative diseases. These are SBMA, HD, SCA 1, DRPLA and MJD. There is a greater likelihood of juvenile-onset cases of HD, SCA1, and DRPLA when the expansion is inherited from affected fathers rather than affected mothers. The alterations in the expansion region are much larger in paternal transmission versus maternal

transmission. Consequently, affected males are more likely to transmit larger repeats to their offspring, thus a greater chance of early onset of disease (LaSpada *et al,* 1994; Nelson, 1996). Sperm analysis shows the male bias is primarily attributed to the tendency of the repetitive region to expand during spermatogenesis (Jennings, 1995).

Current findings suggest that the expansion of CAG/polyglutamine tracts in the disease-causing gene leads to a gain-of-function mutation (Jennings, 1995). Studies in available models have found the loss-of-function phenotype does not resemble the disease pathology. For example, complete loss of the Huntington gene product, huntingtin, causes embryonic lethality (Duyao *et al,* 1995; Nasir *at al,* 1995). The gene involved in SBMA encodes the androgen receptor whose absence or defective function leads to complete or partial androgen insensitivity without any indication of neurodegeneration. In transgenic mice with the expanded SCA1 gene, Purkinje cell degeneration is observed (Burright *et al,* 1995). Throughout evolution, across species, glutamine rich domains and polyglutamine tracts have been conserved and identified in transcription factors as well as other developmentally important genes. At present, many suspect that the expansion of the polyglutamine tracts results in a novel, pathological interaction with a target protein which is yet to be determined (LaSpada *et al,* 1994; Jennings, 1995).

The expansion for the five CAG-repeat diseases occurs within the protein-coding region which suggests that expansion beyond a given number of repeats may abolish the functioning protein. To gain further insight into this trend, Stallings (1994) surveyed the distribution of trinucleotide microsatellite sequences in the GenBank database (available at http://www.ncbi.nlm.nih.gov/) The motif (CAG)n was found to be the most abundant

trinucleotide microsatellite in the database occurring almost exclusively within exons which, in part explains their involvement in the pathogenesis of trinucleotide expansion disorders. The exclusion of (CAG)n from intronic regions in a strand specific manner suggests a potential mechanism to differentiate from the similar, but highly conserved CAGG sequence of the 3'-acceptor consensus splice site (Shapiro and Senapathy, 1987). As most proteins with polyglutamine tracts are transcription factors, the presence of (CAG)n repetitive region within intronic regions would not be amenable to binding other proteins (Richards *et al,* 1993).

In fragile site related mutations such as in FRAXA and in DM, large expansions are evident. In both of these disorders, the expansion occurs in the untranslated region of the gene. Unlike the five other expansion disorders where beyond a given size threshold the gene becomes unstable, the untranslated, non-CAG repeat disorders must overcome a second expansion size threshold. Mutations which evolve beyond the second threshold lead to greater expansion and full phenotypic expression of disease. Carrier individuals have repeats falling between the two size thresholds and present with minimal to undetectable disease. These may lead to full expression in subsequent generations. In FRAXA, a parental bias is observed where expansion occurs through maternal transmission. The male germline does not undergo a similar expansion event. In contrast, both maternal and paternal transmissions can lead to large expansions in myotonic dystrophy. Reductions of the expansion region rarely occur in either of the disorders (LaSpada *et al,* 1994; Nelson 1996).

Most recently, Campuzana *et al* (1996) identified an intronic GAA trinucleotide expansion to cause Friedreich's ataxia (FRDA), an autosomal recessive degenerative disorder. This trinucleotide expansion accounts for 98% of *FRDA*D* (mutation containing) chromosomes studied to date. The GAA repeat is polymorphic in the general population, varying from 7-22 repeat units. However, the disease-causing allele has from 200 to 900 GAA repeat units, with most alleles containing 700 to 800 units. The instability of the repetitive region has been clearly demonstrated in parent-to-offspring transmission. The phenomenon of anticipation does not characterize the inheritance of FRDA. Heterozygous carriers of expanded alleles are asymptomatic and, therefore, cannot be identified by physical examination. Consequently, the allele is maintained in the population at a much higher frequency as compared to the frequency of the mutant alleles in the other trinucleotide expansion disorders.

As investigations into the human genome continue, additional inherited disorders with trinucleotide expansions will be identified. Trinucleotide expansions may be responsible for other disease-causing fragile sites. Furthermore, the (CAG)n expansions may be involved in other dominantly inherited neurodegenerative disease or perhaps certain neurological or psychiatric disorders. Nevertheless, the most definitive answers will be found when both the normal and the pathogenic function of the trinucleotide expansion in disease-causing genes are defined (LaSpada *et al,* 1994).

1.7. Gene Mapping

1.7.1. Functional Cloning Approach

There are an estimated 50, 000 to 100, 000 genes in the human genome.Yet, to date, there is some information only on >8000 genes. Of these, >3700 have been assigned to individual chromosomes and the proportion known to harbour mutations disrupting normal function is continually increasing. The classical approach to identify and isolate human disease genes is termed functional cloning. in which the disease determining gene is isolated based on pre-existing knowledge about the function of the defective gene or its protein product. Once sufficient information about its function is acquired or the amino acid sequence of the purified protein product is available, the responsible gene can be mapped and cloned (Collins, 1992; 1995). Some well known examples of disease determining genes identified by a functional cloning strategy include β-thalassemia (Orkin and Nathan 1981), Lesch-Nyhan syndrome (Jolly *et al,* 1982), phenylketonuria (Kwok *et al,* 1985) and G6PD deficiency (Persico *et al,* 1986).

1.7.2. Candidate Gene Approach

The candidate gene approach relies on partial information about the biological or biochemical basis of the disease. It is an intermediate strategy between functional cloning in which mapping is directed by prior knowledge about the biochemical defect and positional cloning in which the gene is mapped and then cloned without prior knowledge of function (Collins, 1995; Green *et al,* 1995). A candidate gene is selected because of its known involvement in disease pathogenesis, lies within the appropriate interval on the

chromosome or appears to be physiologically relevant. The initial steps in the candidate gene approach involve testing for linkage between the putative candidate gene or closely flanking DNA markers and the locus for the disease under study. Once linkage is confirmed, the gene must be scanned for disease-causing mutations (Gilbert, 1994).

1.7.3. Positional Cloning Approach

In instances where there is limited knowledge about gene function and its role in disease, an alternative approach must be taken to identify the disease-causing gene. A positional cloning strategy is used to establish the "position" of the responsible gene within the genome by utilizing genetic and physical mapping techniques. Genetic mapping tools are used to narrow the region of the genome to a stage which allows for molecular cloning of the interval, scanning for expressed genes and eventual identification of the causative mutation(s) (Collins, 1995; Green *et al,* 1995).

The positional cloning approach involves a series of sequential steps to identify and isolate the disease gene. First, the disease phenotype must be clearly defined to determine if the disease has a significant genetic component and to ensure consistent diagnosis in family ascertainment and data analysis. Pedigrees are then collected where the defective gene is segregating in multiple affected individuals. DNA samples are collected from both affected and unaffected family members (Gilbert, 1994; Green *et al,* 1995).

The next stage is to localize the mutant gene to a region of the genome. Cytogenetic abnormalities such as deletions, translocations and fragile sites associated with the disease help considerably to define the interval where the gene is situated.

However, the majority of genetic diseases are not associated with cytogenetic abnormalities (Collins, 1992). In these cases, the family members are genotyped using a set of markers that span the entire genome to search for a marker which segregates with the disease in the families being studied. Once, the disease-causing gene has been assigned a chromosomal location, additional genetic markers can be typed to refine its position in the genome. The extent to which the region of interest can be narrowed depends on family resources such as the number and size of families and availability of biological samples. In addition, factors such as the number, distribution and informativeness of the genetic markers in the region of interest must be considered. When all the genotype data are collected, various computational tools are utilized to determine if the data collected are sufficient to confidently establish linkage (Gilbert, 1994; Green *et al,* 1995).

Ideally, genetic mapping allows for the candidate region to be limited to approximately 1 cM. However, genetic and physical distances are not linearly related due to sex-specific and regional-specific differences in recombination rates across the genome (Renwick, 1969). Typically, a 1 cM genetic distance corresponds to 1 to 3 Mb (Donis-Keller *et al,* 1987) which can be physically mapped by the use of overlapping clones. The starting point for isolating clones are the closest distal and proximal genetic markers. From these isolated clones, new markers are derived which serve to identify new, overlapping clones. This process, commonly referred to as "chromosome walking" is repeated systematically until the entire DNA interval defined by the genetic markers is represented by overlapping clones. The collection of overlapping clones can be used to

develop additional genetic markers for further fine mapping of the candidate interval and later for gene isolation and mutation detection (Rommens *et al,* 1989).

The greatest challenge in the positional cloning strategy is to identify coding genes in the candidate region. The three major approaches available are: (1) DNA-DNA hybridization (2) functional analysis and (3) DNA sequence analysis. DNA-DNA hybridization methods include zoo blots in which DNA fragments from larger clones are used to probe a panel of genomic DNAs from a collection of organisms to identify the degree of evolutionary conservation of a gene (Watson *et al,* 1992). These same probes can be used in Northern blot techniques to hybridize mRNA derived from various human tissues. A third technique relies on digesting the probes with specific enzymes to identify CpG islands which often are indicators of the 5'-ends of genes. Alternatively, large genomic clones such as yeast artificial chromosomes (YACs) and cosmids can be used to directly probe cDNA libraries derived from tissue specific mRNA (Green *et al,* 1995).

The second group of available methods are function based in which a particular gene function or structure is exploited. For example, in "exon trapping", the unique sequences for splicing introns are used to isolate adjacent exons. Methods to trap the most 3'-end exon in a gene have been developed using the characteristic poly A tract of mRNA as a signal. Another method is to transfer the cloned genomic segment into an appropriate mammalian expression system and select for the desired gene function (Green *et al,* 1995).

The last approach to identifying genes is by systematically analyzing DNA sequence data from the candidate region. The most simple analysis is to translate these

data into all six possible reading frames to identify the longest open reading frame uninterrupted by a stop codon. Further, computer analysis can be used to detect possible exons, known protein motifs and other structural features of genes. In addition, the DNA sequence can be compared across species to detect motifs that have been conserved through evolution (Watson *et al,* 1992).

Any genes identified in the region of interest are potential candidates for the disease. To prove the candidate gene is indeed the correct one, mutation(s) in the gene must be associated with the disease. The mutation(s) must demonstrate the same pattern of inheritance, such as autosomal dominant or recessive, as the disease. The candidate gene is scanned for DNA alterations using techniques such as denaturing gradient gel electrophoresis, RNase or chemical cleavage of mismatches, single-strand conformation polymorphism analysis, heteroduplex analysis and direct sequencing (Gilbert, 1994; Green *et al,* 1995). In the course of detecting mutations, it is important to discriminate between benign sequence polymorphisms part of the disease determining gene and pathogenic mutations. Generally, an amino acid substitution in a predicted functional domain of the protein provides strong evidence for its involvement in disease. Yet, the most convincing evidence that a candidate gene is indeed involved in disease comes from expression studies demonstrating the mutant gene to cause the abnormal phenotype and the normal gene to correct the abnormal phenotype (Green *et al,* 1995).

Since 1986, in excess of 42 disease-causing genes have been identified using a positional cloning approach. Over 50% of these genes were mapped and cloned using knowledge of an associated cytogenetic rearrangement (Collins, 1995). Some of the

earliest genes identified are responsible for chronic granulomatous disease (Royer-Pokora *et al*, 1986), Duchenne muscular dystrophy (Monaco *et al*, 1986), retinoblastoma (Friend *et al*, 1986), the six previously described trinucleotide expansion disorders (FRAXA, DM, HD, SCA1, DRPLA, MJD) and most recently spinal muscular atrophy (Lefebvre *et al*, 1995) in addition to others.

A more streamlined and efficient method of identifying a disease gene is the "positional-candidate" approach. This strategy relies on the initial correct chromosomal localization of the disease locus followed by a survey of the interval of interest for suitable candidate genes. Once a candidate gene is identified, it must be scanned for potential mutations which correlate with features of the disease phenotype (Collins, 1995). This approach has been used successfully to identify greater than 19 disease-causing genes such as those responsible for some forms of retinitis pigmentosa (Dryja *et al*, 1990), malignant hyperthermia (MacLennan *et al*, 1990), Li-Fraumeni syndrome (Malkin *et al*, 1990), early onset Alzheimer disease (Goate *et al*, 1991), Marfan syndrome (Lee *et al*, 1991), Charcot-Marie-Tooth disease (Patel *et al*, 1992), Waardenburg syndrome (Tassabehji *et al*, 1992), the three allelic sydromes, Crouzon (Reardon *et al*, 1994), Jackson-Weiss (Jabs *et al*, 1994) and Pfeiffer (Lajeunie *et al*, 1995) and other diseases.

1.8. Mapping Hereditary Corneal Dystrophies

1.8.1. Previous Mapping Efforts of *CSD1*

As early as 1975, linkage studies between *CSD1* and various genomic markers on chromosomes 1, 2, 4, 6, 9, 13, 14, 16, 19, and 20 had been conducted. Hammerstein and

Scholz (1975) tested four CSD1 families and observed low positive lod scores with the ACPL (0.211), MNS blood group system (0.385), HP (0.402) and GC (0.602) loci. There was no evidence for linkage to ABO blood group markers. Kivlin *et al* (1984) studied two CSD1 kindreds consisting of 27 affected and 24 unaffected individuals. In this study, tight linkage to 15 informative blood group markers was excluded. The most positive lod score, 0.56 at $\hat{\Theta}$=0.17, was obtained between haptoglobin (*HP*) and *CSD1*. When the results between *CSD1* and *HP* from these two studies were combined, a weakly positive lod score of 0.96 at $\hat{\Theta}$=0.17 resulted. To test the potential chromosome 16 localization of *CSD1*, Marles (1994) studied the CSD1 kindred described herein. Similar to previous studies, linkage to 14 blood group markers was excluded. A total of 23 informative meioses were scored for linkage between *CSD1* and *HP*. A maximum lod score of 1.555 at $\hat{\Theta}$=0.10 was obtained. When all the *CSD1:HP* data from the three studies were combined, the maximum lod score, 2.46 at $\hat{\Theta}$=0.20, was not sufficient to demonstrate confidently linkage to chromosome 16q.

A candidate gene approach was also under taken to identify the gene(s) responsible for CSD1. Although the nature of the amyloid deposits in CSD1 is yet to be characterized, several precursor proteins are likely involved in the formation and deposition of amyloid specifically to the corneal stroma. Two proteins involved in systemic amyloidoses, gelsolin (GSN) and transthyretin (prealbumin) (TTR) were investigated further for the possibility of harbouring mutation(s) responsible for CSD1. GSN is involved in the Finnish type IV amyloidosis, a disorder that presents with corneal dystrophy and other manisfestations. Thirty-two informative meioses scored for linkage in

the family described herein resulted in strongly negative lod scores (Z=-6.29 at $\hat{\Theta}$=0.05) between a highly polymorphic, intronic microsatellite of the gelsolin gene and *CSD1*. This finding provided sufficient genetic evidence to exclude *GSN* as a candidate gene for CSD1 (Wiens *et al*, 1992).

Mutations in the TTR gene have been implicated in causing both familial amyloid polyneuropathies types I and II which include corneal dystrophy among the phenotypic features. Single strand conformational analysis and direct sequencing of coding regions of the TTR gene showed no sequence differences in genomic DNA from control individuals and those affected with CSD1 (Marles, 1994).

These two findings reflect the absence of linkage between *CSD1* and *ABO* whose locus maps to 9q34.1-q34.2 in the same general location as *GSN* (9q33) and between *CSD1* and the *JK*, Kidd blood group locus, which maps to 18q11-q12 in the same general vicinity as *TTR* (18q12.1).

1.8.2. Corneal Dystrophies Mapping to 5q

A genome-wide linkage analysis of microsatellite markers enabled investigators to map the loci for four phenotypically distinct corneal dystrophies to 5q22-33.3 over the course of the last three years (figure 3 and table 1). In 1993, Eiberg *et al* screened genomic markers in a large granular corneal dystrophy kindred of 124 individuals (52 affected and 72 unaffected). Highly positive lod scores were observed with 5q microsatellite markers, specifically to the interval defined by *IL9* proximally and *D5S436* distally. The highest lod score,15.96 at a $\hat{\Theta}$=0.00, was observed between granular corneal dystrophy and *IL9* (Eiberg *et al*, 1994). Stone *et al* (1994) likewise mapped granular corneal dystrophy to the identical 10 cM interval

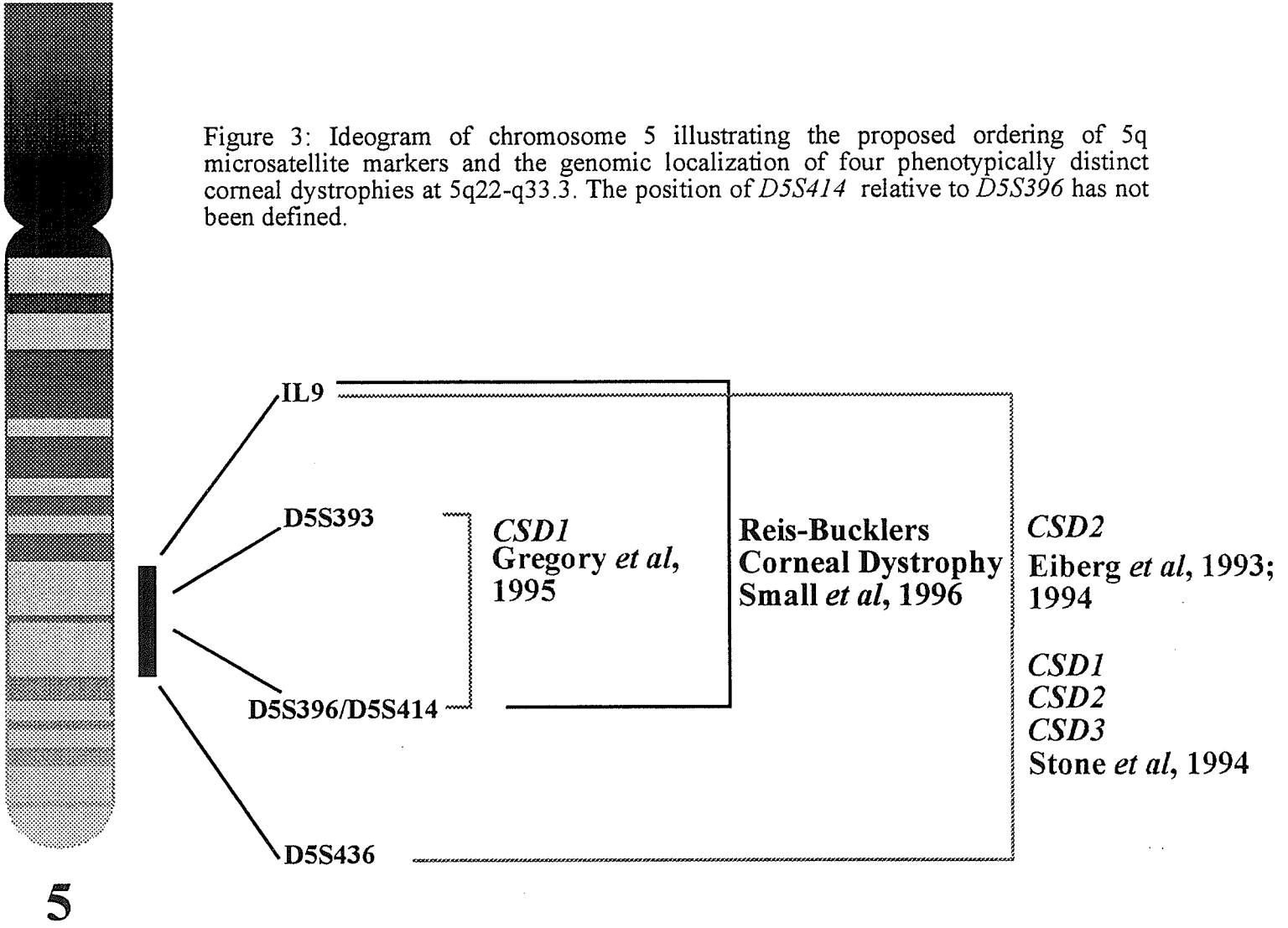


Figure 3: Ideogram of chromosome 5 illustrating the proposed ordering of 5q microsatellite markers and the genomic localization of four phenotypically distinct corneal dystrophies at 5q22-q33.3. The position of *D5S414* relative to *D5S396* has not been defined.

Table 1: Summary of closest linked markers at their corresponding lod scores for corneal dystrophies mapping to 5q22-q33.3

Research Group	Type of Corneal Dystrophy	Closest Linked Markers	Lod Score	
			Zmax	$\hat{\Theta}$
Eiberg *et al*, 1993	CSD2	*IL9*	15.96	0.00
Stone *et al*, 1994	CSD1	*D5S393*	10.70	0.00
	CSD2	*D5S178*	6.90	0.00
	CSD3	*D5S393 & D5S414*	11.30	0.00
Gregory *et al*, 1995	CSD1	*D5S393*	7.51	0.03
Small *et al*, 1996	Reis-Bücklers	*D5S414*	3.54	0.00

on 5q as well as two other clinically and histologically distinct corneal stromal dystrophies, lattice corneal dystrophy type 1 and Avellino corneal dystrophy. As these three stromal dystrophies, lattice type 1, granular and Avellino map to the same locus, they have been renamed corneal stromal dystrophy type 1 (CSD1), type 2 (CSD2) and type 3 (CSD3) respectively. The highest lod scores for *CSD1*, *CSD2* and *CSD3* were observed between different microsatellite loci. The maximum lod score for *CSD1* was obtained with marker *D5S393* (Z =10.70, $\hat{\Theta}$=0.0), for *CSD2* with marker *D5S178* (Z=6.90, $\hat{\Theta}$=0.00) and *CSD3* with both markers *D5S393* and *D5S414* (Z=11.30, $\hat{\Theta}$=0.00). The combined maximum lod score using all 114 affected individuals from CSD1, CSD2 and CSD3 kindreds was 28.60 at $\hat{\Theta}$=0.00 between *CSD* and *D5S393*. The candidate interval has since been refined further to 2 cM by Gregory *et al* (1995) who conducted a linkage study of a CSD1 kindred of 46 family members (24 affected, 11 unaffected and 11 unrelated spouses). This study positioned *CSD1* between *D5S393* (distal to *IL9*) proximally and *D5S369* (proximal to *D5S436*) distally. Most recently, Reis-Bücklers corneal dystrophy has also been mapped to 5q in the interval *IL9* proximally and *D5S414* distally. Although, it is a non-amyloid type of corneal dystrophy which is clinically and histopathologically distinct from CSD1 and CSD3, it does share some common clinical features with CSD2 (Moller, 1989). The highest lod score, 3.54 at $\hat{\Theta}$=0.00, was obtained between the Reis-Bucklers corneal dystrophy locus and *D5S414* in a study of 22 individuals (11 affected and 11 unaffected) of a single five generation kindred (Small *et al,* 1996).

1.8.3. Genes Mapping to Candidate Region

The region 5q22-q33.3 encodes several disease-causing genes as well as some important gene families. Those genes involved in disease cause hereditary startle disease, diastrophic dysplasia, Treacher Collins syndrome, an autosomal dominant form of deafness, one of several autosomal forms of limb girdle muscular dystrophy, a myeloid disorder associated with 5q- syndrome and tumour development in specific types of colorectal cancer. Various genes from the interleukin family (IL3, IL4, IL5, IL9, and IL13) reside within this region. In addition, genes for growth factor receptors such as glucocorticoid receptor (GRL) and adrenergic receptor (ADRB2) and genes for growth factors such as early growth response (EGR1), fibroblast growth factor 1 (FGF1) are also located at 5q22-33.3. Some regulatory function genes found, include transcription factor 7 (TCF7) and transcription elongation factor 13 (TCEB1L). A number of genes involved in calcium binding such as osteonectin (SPARC), annexin IV (ANX6) and calnexin (CANX) are present on 5q as well. A more comprehensive list is available from Online Mendelian Inheritance in Man (OMIM) (available at http://www3.ncbi.nlm.gov/omim/). To date, the genes mapped or cloned to the candidate interval do not appear as obvious candidates for CSD1 or the other three corneal dystrophies, but none have been systematically tested. Eiberg *et al* (1994) speculated the gene encoding SPARC as a potential candidate for CSD2. Small *et al* (1996) tested for linkage between the loci for Reis-Bücklers corneal dystrophy and SPARC and obtained significantly negative lod scores.

The evolutionary conservation of linkage groups between humans and mice on the autosomes has been well documented. Many genes or disease loci have been identified by

comparative mapping methods. Here, predictions are made regarding the chromosomal position of gene(s) in the human genome once their homologues have been mapped to conserved regions in the mouse genome (Lalley *et al*, 1989; Searle *et al*, 1994). The human chromosomal region 5q22-q35 is conserved on mouse chromosomes 11, 13, and 18. To date, there are approximately 36 human genes within this region that have homologous loci on mouse chromosome 11, 13, and 18. For example, on mouse chromosome 11 at position 28-29 cM, homologues of human interleukin genes, IL3, IL4, IL5, and IL13 can be found. At 29 cM on mouse chromosome 11, the human homologue of ANX6 (5q32-q34) is present. The homologues of the SPARC (5q31-q33) and CANX (5q35) genes can be found at 30 cM on mouse chromosome 11. The human IL9 gene (5q22-q32) is positioned at a 33 cM on mouse chromosome 13. Homologues of a number of the growth factors and growth factor receptors are on mouse chromosome 18. A more detailed list can be obtained from the Mouse Genome Database (http://www.informatics.jax.org/mgd.html) where tools to perform mammalian homology searches are also available.

Since the candidate interval of interest on human chromosome 5q is not conserved on a known syntenic segment of the mouse genome, comparative mapping does not offer further insight to a potential candidate gene. It appears that evolutionary changes have dispersed segments homologous to human chromosome 5 to four different mouse chromosomes 11, 13, 15 and 18. Nevertheless, by locating a candidate loci in either mammalian genome can lead to the search for its homologue in the other species. The

conservation of genes across species is a clear indicator of the important biologic function of a given gene.

2. Project Objectives

The linkage analysis study of a CSD1 kindred of Belgian descent was carried out with the following objectives:

(1) to confirm the mapping of *CSD1* to the 10 cM interval defined by *IL9* proximally and *D5S436* distally at 5q22-q33.3.

(2) to refine the mapping of *CSD1* to a candidate interval smaller than 10 cM at 5q22-q33.3.

In pursuing the two objectives outlined above, technical difficulties encountered with genotyping PCR-formatted microsatellite loci led to an additional objective: to determine the cause for the unusual pattern of segregation of alleles observed at microsatellite locus D5S414 in one branch of the CSD1 kindred.

3. Materials and Methods

3.1. Clinical Studies

3.1.1. CSD1 Kindred

A single large multigeneration CSD1 family of Belgian descent is the focus of this study (figure 4). This kindred was first described in 1958 by Dr. R. Ramsay (Ramsay, 1958). The mode of transmission of CSD1 is consistent with autosomal dominant

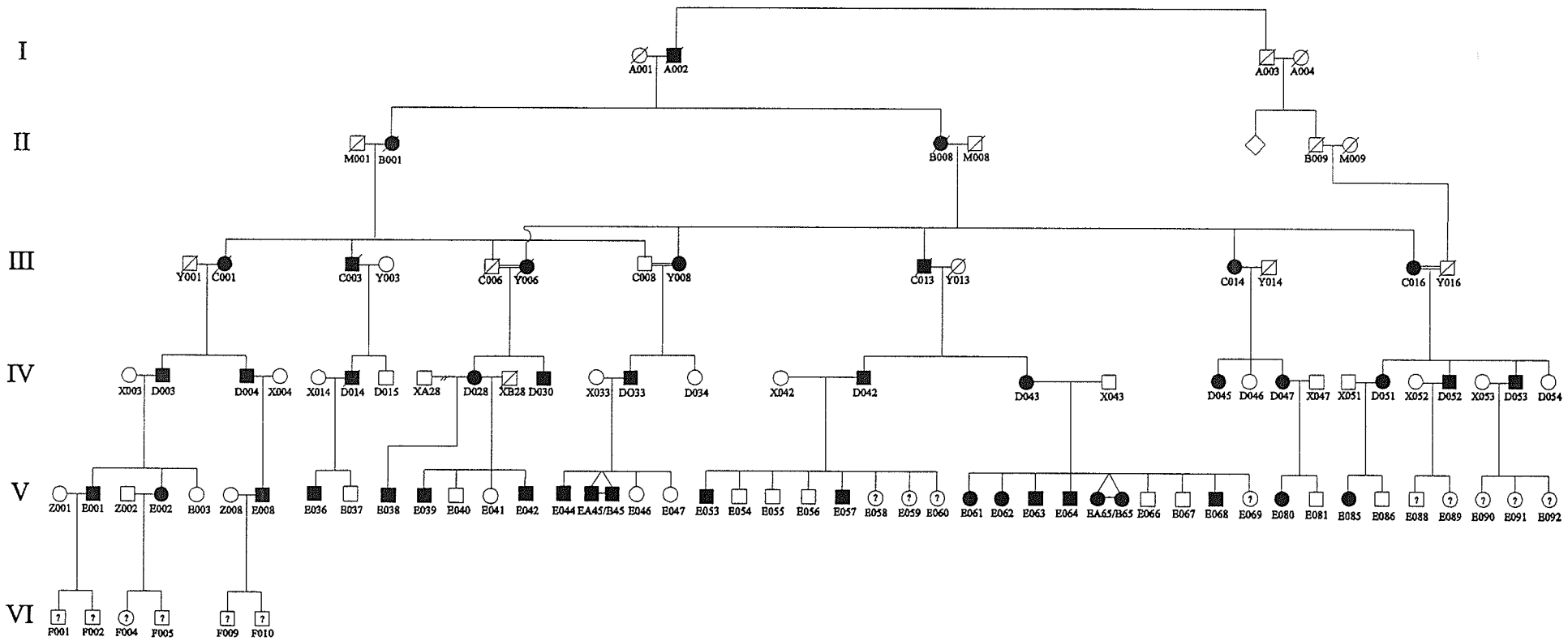


Figure 4: Corneal stromal dystrophy type 1 kindred of Belgian descent under study. Affected individuals are represented by filled-in symbols; unaffected individuals are represented by open symbols. Fifteen individuals at fifty percent risk of being affected are represented by open symbols with question mark.

inheritance with essentially full penetrance affecting males and females equally. There are no apparent clinical differences among affected individuals born to affected mothers versus affected fathers. All CSD1 patients and at-risk individuals are being closely followed by one ophthalmologist, Dr. M. Ekins (Department of Ophthalomology, University of Manitoba). Those affected have been symptomatic and/or have demonstrated characteristic corneal changes by the age of 25 years. To date, none of the individuals assessed as unaffected at the age of 25 years have subsequently gone on to develop the signs and symptoms of CSD1 later in life. There is no evidence of unusual systemic illnesses, specifically neurological, cardiac or renal in any affected individual. Presently, those at-risk are 15-24 years of age. At their last eye examination by Dr. M. Ekins, all showed no signs or symptoms of CSD1 (table 2). Over the past decade, 26 patients have received a total of 42 corneal grafts and several more are awaiting transplantation. Characteristic amyloid deposits have been confirmed histopathologically by Dr. J. Safneck (Department of Pathology, University of Manitoba) in all the corneas studied.

3.2. Sample Collection

Heparinized venous blood samples were collected for previous studies (Wiens *et al,* 1992, Marles *et al,* 1994) from consenting familiy members. The 60 participating family members 25 years of age or older included 34 affected individuals, 26 unaffected individuals and 13 unrelated spouses. Additionally, 15 samples were collected and banked

Table 2: Summary of clinical status of at-risk individuals

Individual	Age of Ascertainment (1991)	Present Age (1996)	Last Eye Examination	Corneal Status
F001	13	18	Aug/96	clear
F002	12	17	Aug/96	clear
F004	18	23	Mar/96	clear
F005	14	19	Mar/96	clear
F009	16	21	Mar/96	clear
F010	13	18	Aug/90	clear
E058	21	26	Oct/94	clear
E059	16	21	Oct/94	clear
E060	13	18	Nov/93	clear
E069	17	22	Nov/90	clear
E088	18	23	Jul/93	clear
E089	13	18	Jun/94	clear
E090	18	23	NA	NA
E091	16	21	NA	NA
E092	14	19	NA	NA

from "at-risk" individuals younger than 25 years of age all of whom have an affected parent.

Figure 5 illustrates the nuclear family studied to characterize the segregation of a non-amplifying allele at *D5S414*. DNA sample from unrelated spouse X045 was previously banked. DNA samples were obtained from the two unaffected offspring, E076 and E077.

3.3. Molecular Methods

3.3.1. DNA Isolation

DNA was extracted from whole blood following the method outlined below. First, five volumes of lysis buffer (0.155 M NH_4Cl, 0.170 M Tris, pH 7.65) warmed to 37°C were added to one volume of whole blood to lyse the non-nucleated red blood cells. Following incubation for five minutes at 37°C, the blood and lysis buffer were centrifuged at 2000 rpm for ten minutes (IEC HN-SII Centrifuge) to pellet the nucleated cells. The supernatant was carefully aspirated and discarded leaving behind the white blood cell pellet and the residual loose red blood cell pellet. The pellet was resuspended in 20 mL saline solution (0.85% NaCl) and centrifuged at 2000 rpm for ten minutes. The supernatant was then aspirated leaving the white blood cell pellet. This pellet was resuspended in 2 mL high TE (100 mM Tris, 40 mM EDTA, pH 8.0), then immediately lysed by the addition of 2 mL white blood cell lysis buffer (100 mM Tris, 40 mM EDTA, 1 M NaCl, 0.2% SDS, pH 8.0) using a syringe attached to a 16 gauge needle. To extract the DNA, an equal volume of phenol was added to the lysed cells. After gently mixing for ten

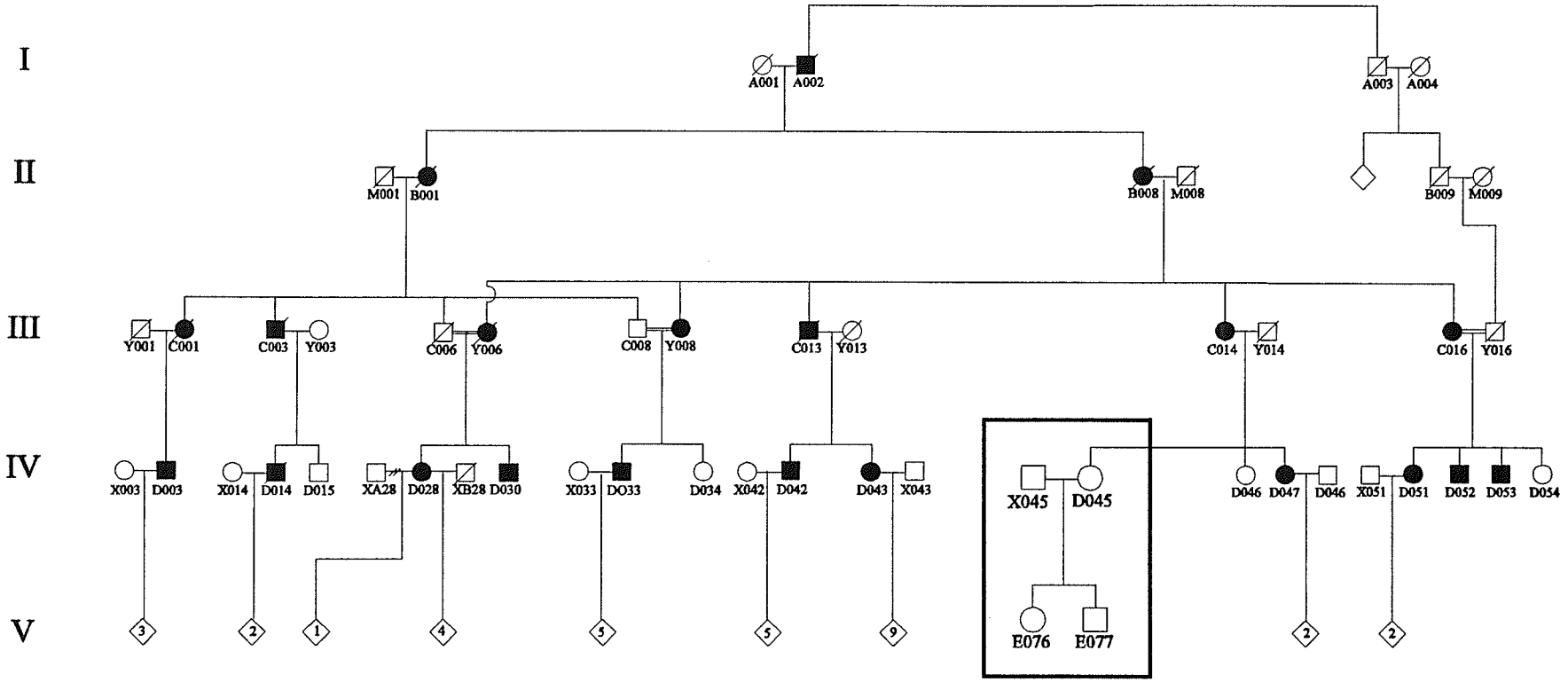


Figure 5: A branch of the CSD1 kindred where segregation of a non-amplifying allele at *D5S414* was detected (boxed area).

minutes by inversion, the milky white suspension was centrifuged at 2000 rpm for five minutes. The upper aqueous phase was removed and the extraction repeated by adding phenol followed by mixing and centrifugation under the same conditions. The aqueous layer was then extracted with an equal volume of chloroform:isoamyl alcohol (24:1) and centrifuged at 2000 rpm for five minutes. To the upper aqueous layer, 1/10 volume of 4 M ammonium acetate and an equal volume of isopropanol was added. After inverting several times, the high molecular weight DNA precipitated out. The isolated DNA was washed first with 70% isopropanol, followed by a final wash of absolute isopropanol to remove any traces of water and further condense the DNA. The pellet was allowed to dry in a speed vacuum (Speed Vac Concentrator, Savant) for fifteen minutes and reconstituted in the appropriate volume of low TE (10 mM Tris, 1 mM EDTA, pH 8.0). The concentration of DNA was determined spectrophotometrically at 260 nm. The purity was estimated by measuring the ratio of absorbance at 260 nm and 280 nm. The DNA samples were stored at 4°C.

3.3.2. Genotyping Microsatellite Markers

A series of ten microsatellite markers in the candidate region was selected for study from the published literature (Stone *et al,* 1994) and personal communication with Dr. A. Kanis (Department of Pediatrics, University of Iowa, Iowa City, Iowa). Information on the oligonucleotide primer sequences and their chromosomal location was obtained from the Genome Database (GDB) (http://gdbwww.gdb.org) and the Cooperative Human Linkage Centre (CHLC) (http://www.chlc.org). Seven of the selected

markers were (CA)n dinucleotide repeats and the remaining three were tetranucleotide repeats (table 3).

Radioactive PCR was performed in a total volume of 25 μL. The forward primer was labelled by the following reaction mixture: 0.3 μL sterile nanopure water, 0.2 μL 5x kinase buffer (300 mM Tris-HCl pH 7.8, 50 mM $MgCl_2$, 1 M KCl, Gibco BRL), 4 μM appropriate primer, 1 μL (γ32P)dATP (10.0 mCi/mL Dupont) and 1/15 units T_4 polynucloetide kinase (Gibco BRL). The labelling mix was incubated at 37°C for thirty minutes and the reaction stopped by incubation at 55°C for five minutes. The PCR reaction was set up as follows: 100 ng of each individual's DNA was used as template in a reaction containing 4 μM of non-labeled and ^{32}P end-labeled of the appropriate primer, 5mM of each dATP, dCTP, dGTP and dTTP, 2.5 μL 10x buffer (100mM TRIS, 500 mM KCl, 15 mM $MgCl_2$, 0.01w/v gelatin) and 0.2μl 5000 units/mL *Taq* DNA polymerase (Gibco BRL). Samples were overlaid with mineral oil and incubated in a DNA thermocycler (Perkin Elmer Cetus) under standard conditions with an initial denaturation at 95°C for 3 minutes, followed by 30 cycles of 95°C for 1 minute, annealing 55°C for 1 minute and extension at 72°C for 2 minutes and a final extension of 10 minutes at 72°C. The annealing temperature and number of cycles were often adjusted to optimize the PCR conditions for a given primer pair. In addition, glycerol was added up to a maximum of 15% of the total volume to the PCR reaction mix to reduce non-specific amplification. Following amplification, 12 μL of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) were added to the PCR products and then stored at -70°C. PCR products were heated at 95°C for 5 minutes and immediately placed on

Table 3: Series of chromosome 5q microsatellite markers screened

Locus	Chromosomal Location	Allele Sizes (bp)	Annealing Temperature (°C)	Repeat Motif	Max. Het.	*Source
D5S818	5q	141-145	57	AGAT	-	CHLC
IL9	5q22-q32	123-139	58	CA	-	GDB
D5S816	5q	223-251	57	ATCT	0.95	CHLC
D5S393	5q	162-182	55	CA	0.84	GDB
D5S399	5q	116-134	62	CA	0.80	GDB
D5S414	5q	186-206	53	CA	0.83	GDB
D5S396	5q	120-128	57	CA	0.66	GDB
D5S436	5q	234-254	62	CA	0.87	GDB
D5S178	5q31	98-120	57	CA	0.93	GDB
D5S2508	5q	161-189	57	ATAC	0.67	CHLC

*Cooperative Human Linkage Centre (CHLC) available at http://www.chlc.org

Genome Database (GDB) available at http://gdbwww.gdb.org

ice. Samples were loaded along side an M13 sequencing ladder (Sequenase® Version 2.0 United States Biochemical) on 7M urea, 6% polyacrylamide sequencing gels to separate the amplified fragments by electrophoresis for 2-5 hours at 2000 V. The M13 DNA was supplied in the Sequenase® Version 2.0 kit and sequenced according to the protocol provided without modification. The gels were dried and exposed to X-ray film (DuPont) for 4-24 hours. The film was developed in an AFP Imaging 14XL X-ray film processor. For each locus studied, fragment sizes were determined relative to the M13 sequencing ladder and control DNA from other family members and a CEPH individual (classification number 134702) with previously assigned allele sizes. Each nuclear family was genotyped as a unit to facilitate unambiguous allele assignment. Allele designations were initially based on the number of base pairs and later arbitrarily assigned consecutive numbers starting with the minimum number of repeats (table 4). For example, "1" represented the allele with the smallest number of repeat units and "10" represented the allele with the maximum number of repeat units.

3.3.3. Cloning PCR Products

In genotyping *D5S414*, it was brought to our attention by our collaborator Dr. E. Stone (Department of Ophthalmology, University of Iowa, College of Medicine, Iowa City, Iowa) of the possible segregation of a non-amplifying allele at this locus. By inspection of the genotype data obtained in our kindred, a non-amplifying allele appeared to be segregating in one nuclear family. The published oligonucleotide primer sequences that define *D5S414* are shown in figure 6. The published primer pairs designated A and B

Table 4: Allele coding scheme for a series of 5q microsatellite markers screened

	Coded allele designations for this study							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Alleles at *D5S818* (bp)	146	150	154	158				
Published coded allele designations	3	4	5	6				
Alleles at *IL9* (bp)	119	123	125	127	129	131	133	106
Published coded allele designations	*NA	NA	NA	7	6	5	4	NA
Alleles at *D5S816* (bp)	225	233	237	241	245	249	253	
Published coded allele designations	1	2	3	4	5	6	7	
Alleles at *D5S393* (bp)	160	162	164	166	168	170	172	174
Published coded allele designations	NA	1	2	3	4	5	6	7
Alleles at *D5S399* (bp)	116	118	124	126	128	130	132	134
Published coded allele designations	1	2	NA	3	4	5	6	NA
Alleles at *D5S414* (bp)	190	192	194	196	198	200	204	206
Published coded allele designations	2	3	4	5	6	7	9	10
Alleles at *D5S396* (bp)	120	122	124	126	128			
Published coded allele designations	NA	1	2	3	4			
Alleles at *D5S436* (bp)	236	238	240	242	244	246	248	
Published coded allele designations	2	3	4	NA	5	6	7	
Alleles at *D5S178* (bp)	98	108	110	112	114	116	120	
Published coded allele designations	11	7	6	5	4	3	1	
Alleles at *D5S2508* (bp)	163	167	175	179	183	187	191	
Published coded allele designations	1	2	NA	3	4	5	6	

*(NA) These alleles are not listed in GDB or CHLC

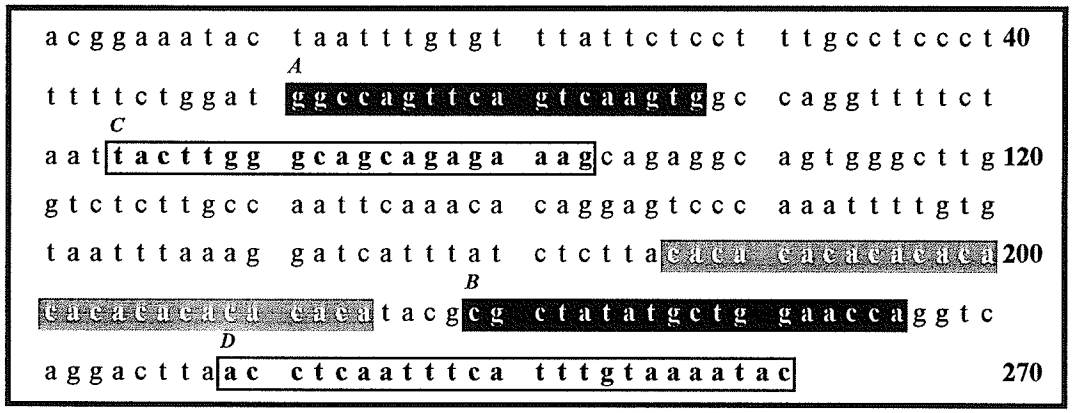


Figure 6: The published oligonucleotide sequence at locus D5S414. The blacked areas indicate the published sequence of forward primer A and reverse primer B initially used for genotyping. The boxed areas represent the forward primer C and reverse primer D designed to amplify the identical region of (CA)n repeats. The grey region indicates the dinucleotide repetitive region.

and a newly designed primer pair, C and D were used to amplify this locus. To determine the cause of nonamplification, two separate PCR reactions were set up, first using primers A and D and the second using primers C and B.

A 220 bp fragment was generated by non-radioactive PCR with primers A and D by omitting the primer labelling step from the procedure outlined earlier for D045 and a control individual. The PCR products were then pooled and gel purified. Samples were subjected to electrophoresis on 5% acrylamide gels composed of 30% acrylamide solution, 1x TBE (89 mM Tris-borate, 2mM EDTA, pH 8.3) and sufficient amount of Temed (Sigma) and 25% ammonium persulfate for polymerization. The samples were electrophoresed with 1 μL of standard DNA marker (1000 μg/mL pGem® DNA Markers Promega) at 270 V for 30 minutes. The gels were stained for 10 minutes in a 10% solution of ethidium bromide. The bands were visualized under longwave length ultraviolet light. The appropriate bands of 220 bp were cut from the gels and incubated overnight with shaking in 0.5 mL low TE at 37°C. Following centrifugation at 10,000 rpm (International Clinical Centrifuge ICE) for 5 minutes, the liquid portion was collected. An equal volume of phenol chloroform (1:1) was added and mixed for 30 seconds. After a five minute incubation at room temperature and centrifugation at 10,000 rpm (IEC Micro-MB Centrifuge) for 5 minutes, the upper aqueous layer was extracted again as above with an equal volume of chloroform:isoamyl (24:1). To the aqueous layer, 1/10 volume of 3 M ammonium acetate and double the volume of absolute alcohol was added. After mixing well, it was incubated at -70°C for 2-3 hours to precipitate the DNA. The samples were centrifuged at 13,000 rpm (Baxter Canlab Biofuge 13) for 20 minutes at 4°C to isolate the

pellet. After removing the supernatant, the pellet was washed with 70% alcohol, then centrifuged at 13,000 rpm (Baxter Canlab Biofuge 13) for 15 minutes at 4°C. The pellet was allowed to air dry for 1 hour and resuspended in 15 μL sterile nanopure water. The purity of the PCR products was confirmed by gel electrophoresis of an aliquot of each sample along side the standard DNA marker as described above. The gel was photographed to determine product purity and as a reference for future quantification.

The TA Cloning ™ System from Invitrogen Corporation was used without modification for direct insertion of gel purified PCR products into the plasmid vector, pCR™II, and subsequent transformation into competent *Escherichia coli* cells. The vector supplied in this kit is specifically designed to take advantage of the non-template dependent activity of *Taq* DNA polymerase which adds a single deoxyadenosine to the 3'-end of the double stranded PCR products. The vector, initially linear, has 3' T-overhangs at the insertion point which can anneal with the PCR product and then be circularized.

Following transformation, twelve colonies were selected on the basis of their white colour, as these were indicative of the incorporation of the insert into the plasmid and disruption of the lacZ gene marker. Three blue colonies were selected as control plasmids indicative of an intact lacZ gene and no insert. Aliquots of 2.5 mL Luria Bertani (LB) broth made according to standard protocol containing 50 μL/mL ampicillan (Sigma) were inoculated with each individual colony and incubated at 37°C overnight in a shaking incubator.

3.3.4. Plasmid Isolation and Purification

A small scale plasmid purification was carried out following the alkali method (Sambrock *et al,* 1989) with some modification. The overnight cultures were centrifuged for five minutes at 4°C at 13,000 rpm (Baxter Canlab Biofuge 13). The pellet was suspended in 100 μL solution of glucose buffer (100 mM glucose, 0.5 M EDTA, 1 M TRIS, pH 8.0) and left at room temperature for five minutes. The cells were lysed by adding 200 μL of freshly prepared SDS-NaOH solution (10% SDS, 5M NaOH) and chilled on ice for five minutes. To precipitate chromosomal DNA, high molecular weight RNA and protein, 150 μL of potassium acetate (3 M potassium, 5 M acetate, pH 4.8) were added, after which the reaction mixture was incubated on ice for thirty minutes and then centrifuged at 13,000 rpm (Baxter Canlab Biofuge 13) for 15 minutes at 4°C. The recovered supernatant was extracted with 500 μL phenol. Following one minute of vigorous mixing, the aqueous layer was separated by centrifugation at 10,000 rpm (IEC Micro-MB Centrifuge) for 5 minutes at room temperature. The upper aqueous layer was extracted again with 500 μL of phenol-chloroform (1:1), followed by incubation at room temperature for five minutes and centrifugation at 13,000 rpm (Baxter Canlab Biofuge 13) for 5 minutes at 4°C. The supernatant was removed and 1 mL of 95% ethanol was added. The contents were left at room temperature for 5 minutes and then centrifuged for 15 minutes at room temperature. The pellet was washed a second time with 70% alcohol. After discarding the ethanol, the pellet was allowed to air dry for 1 hour and dissolved in 50 μL sterile nanopure water.

The sequence for *D5S414* was analysed using the computer program Gene Runner (version 3.0) to determine all possible restriction sites present in the fragment. The *EcoR1* restriction site was selected as it was absent from the insert but present within the plasmid on either side of the insert. *EcoR1* restriction digests were carried out for the 15 plasmid preparations to confirm the presence/absence of insert. The reaction mixture was made up of 2 µl required 10x buffer, 2 µl 10mg/mL RNase (Boehringer Mannheim), 2 µl *EcoR1* (25,000 units/mL Pharmacia), 7 µl sterile nanopure water and 7 µl plasmid preparation. Following incubation at 37°C for three hours, the samples were subjected to electrophorese on 1% agarose gel (electrophoretic grade agarose, 1x TBE) at 90 V for 45 minutes. A 250 ng DNA Molecular-weight Marker II (Boehringer Mannheim) was run along side the digested samples to determine fragment size. A photograph was taken for a permanent record of clones containing the appropriately sized insert.

Two clones with the appropriate insert were selected for large scale plasmid preparation. Aliquots of 2 mL LB broth containing 50 µL/mL ampicillin (Sigma) were innoculated with a loopful of cells. These initiating cultures were incubated with shaking at 37°C for two hours. The initiating cultures were added to 100 mL aliquots of LB broth containing 50 µL/mL ampicillin (Sigma) and incubated overnight at 37°C with shaking. The plasmid was purified using the QIAGEN Plasmid Kit (QIAGEN Inc.) according to protocol without modification. An *EcoR1* digest as outlined above was repeated on these purified samples to determine quality of the purification procedure and yield of the insert. A photograph was taken for permanent record of the two clones containing the appropriately sized insert.

3.3.5. DNA Sequencing

To determine the sequence variation suspected to give rise to a non-amplifying allele, cycle sequencing of double stranded purified PCR products was carried out according to protocol supplied by dsDNA Cycle Sequencing System (Gibco BRL). As the sequencing data were difficult to interpret, several aliquots of cell culture containing the insert-carrying plasmid for D045 and a control individual were sent to the Sequencing Core Facility of the Canada Genetics Disease Network of Centres of Excellence directed by Dr. H. S. Chen (Department of Genetics, Hospital for Sick Children, Toronto). Here, a small scale plasmid DNA preparation was recovered using the alkali method with an additional RNase digestion step. The double stranded DNA was sequenced on the ABI PRISM 377 sequencer by the Dye Terminator Cycle Sequencing method. The cycle sequencing program was 96°C for 30 seconds, 50°C for 15 seconds and 60°C for 4 minutes for 25 cycles.

3.4. Data Analysis

3.4.1. Haplotype Construction

Haplotypes were manually constructed for each nuclear family based on all 10 microsatellite markers genotyped to reflect the minimum number of recombination events assuming that no mutations have occurred. Haplotypes were also constructed for at-risk individuals in the context of their affected/unaffected siblings and parents. The genetic risk to at-risk individuals was predicted by inspection of their haplotype data.

3.4.2. Linkage Analysis

Two-point lod scores were calculated assuming that CSD1 is inherited in an autosomal dominant manner with essentially full penetrance. To facilitate the linkage analysis the kindred was divided into 8 nuclear families defined by a sibship with one or more parents and up to four grandparents (Côté, 1975). Separate maternal and paternal lod scores were calculated manually for fully informative meioses using published lod score tables at the following values of θ: 0.00, 0.05, 0.10, 0.20, 0.30 and 0.40. Total lod score were obtained by summing the total maternal and paternal lod scores.

The coded genotype data were entered into the computer in the correct format to perform linkage analysis using both the MARK III (Côté, 1975) and LINKAGE (Lathrop *et al*, 1984; Lathrop and Lalouel, 1984; Lathrop *et al*, 1986) computer programs. The MARK III linkage analysis was performed by Sylvia Philipps (Rh Laboratory, University of Manitoba). MARK III is designed to calculate two-point lod scores on the assumption that each nuclear family in the kindred is independent (Côté, 1975). Similar to the manual calculations, maternal and paternal lod scores were generated at θ: 0.10, 0.20, 0.30 and 0.40 for each nuclear family and then summed to obtain the total lod score. The lod scores at θ=0.0, 0.01 and 0.05 were determined from published lod score tables. For each marker, the maximum lod score at its corresponding value of θ was determined using the PCMAP83 computer program (Sherman *et al*, 1984).

Linkage analysis of the complete pedigree was performed using the LINKAGE programs (version 5.1 and 5.2) (Lathop *et al*, 1984; Lathop and Lalouel, 1984; Lathrop *et*

al, 1993; Schaffer *et al*, 1994; Schaffer *et al*, 1996). Two consanguineous loops in this CSD1 kindred were broken by inserting a genetically identical person into the pedigree as described by Terwilliger and Ott (1994). MLINK was used to calculate two-point lod scores between *CSD1* and each microsatellite locus at θ: 0.00, 0.05, 0.10, 0.20, 0.30 and 0.40, assuming an autosomal dominant mode of inheritance with 100% penetrance and a disease gene frequency of $1x10^{-5}$. The microsatellite marker allele frequencies were assumed to be equal in the absence of published allele frequencies for all the alleles observed in our kindred. The linkage analysis was repeated using a modified pedigree where all the unaffected individuals were removed to avoid the confounding possibility of incomplete penetrance.

3.4.3. PCMAP83 Analysis

To test various orders of the microsatellite loci and to determine the most likely position of *CSD1*, the PCMAP83 computer program (Sherman *et al*, 1984) was used with the two-point lod score data generated by MARK III (Côté, 1975). The best ordering of the loci including *CSD1* was determined from the goodness of fit evaluated by a p-value as close to or equal to 1.00 for the combined maternal and paternal lod score data.

3.4.4. Compilation of Genome Maps

Chromosome 5 maps of microsatellite markers and genes, specifically at 5q22-q33.3 were obtained from the published literature and via the following electronic information sources on the World Wide Web (table 5).

Table 5: Electronic information sources on the World Wide Web for human genetic maps

Source	Access
Genome Database (GDB)	http://gdbwww.gdb.org
Cooperative Human Linkage Center (CHLC)	http://www.chlc.org
Online Mendelian Inheritane in Man (OMIM)	http://www3.ncbi.nlm.gov/omim/
An STS-Based Map of the Human Genome	http://www-genome.wi.mit.edu/

4. Results

4.1. Genotyping Microsatellite Markers

In this CSD1 kindred of Belgian descent, a total of 60 individuals were genotyped using 10 selected microsatellite markers (table 3). It was necessary to optimise the PCR conditions for each marker studied to facilitate unambiguous allele assignments. For dinucleotide markers, *IL9* and *D5S178*, optimisation failed to reduce substantially laddering artifacts. However, phenotyping nuclear families as a unit along side an M13 sequencing ladder and samples of other affected and unaffected family members enabled individuals to be scored accurately. Repeat analyses were performed as required. Figure 7 illustrates the phenomenon of "shadow bands" which complicated the scoring of alleles and increased the time needed to genotype the kindred. For example, if individual 1 is considered in isolation, it is difficult to discern if he is homozygous or heterozygous at *D5S436*. However, by scoring the alleles in the context of a family unit, it is apparent that individual 1 is indeed heterozygous because his children have inherited two different paternal alleles.

In comparison, tetranucleotide microsatellite markers (*D5S816*, *D5S818* and *D5S2508*) were less sensitive to laddering artifacts which significantly increased readability of the autoradigrams and reduced the time needed to screen the kindred. Figure 8 illustrates genotyping with the tetranucleotide marker *D5S816*. The allele assignment for each family member is unambiguous. The pattern of segregation is clear showing individual 4 to be heterozygous at this locus and one of her offspring, individual 6, to be homozygous.

For each microsatellite locus, the alleles were scored and arbitrarily assigned consecutive numbers starting with the minimum number of repeats. Both GDB and CHLC

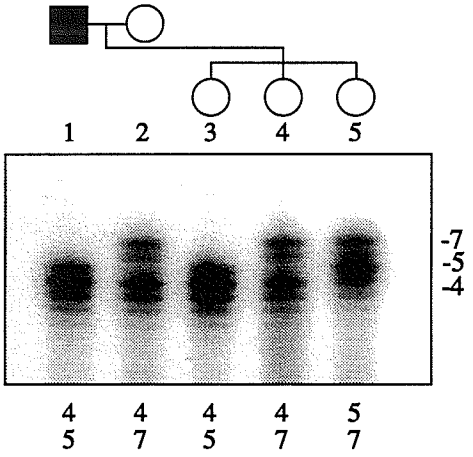


Figure 7: Example of shadow bands when genotyping *D5S436* in one nuclear family of the CSD1 kindred. The coded genotypes are indicated below.

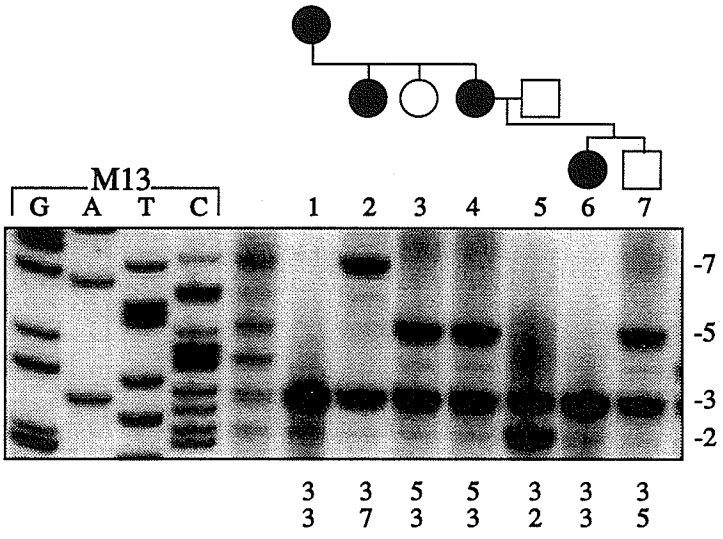


Figure 8: Example of the segregation of alleles when typing with a tetranucleotide marker *D5S816.* The sequence on the left is an M13 sequencing ladder used to determine allele sizes. The coded genotypes are indicated below.

employ the same coding system, however these designations were not used because in genotyping this CSD1 kindred microsatellite loci *IL9*, *D5S393*, *D5S399*, *D5S396*, *D5S436* and *D5S2508* revealed alleles not listed in the databases (table 4). The remaining microsatellite loci *D5S818*, *D5S816*, *D5S414* and *D5S178* showed alleles consistent with those listed in GDB or CHLC.

4.2. Characterization of Non-amplifying Allele

The segregation of alleles at *D5S414* in one branch of the kindred followed a non-Mendelian pattern of inheritance when using published primers A and B to amplify the locus (figures 6 and 9). Individuals E076 and E077 have apparently inherited one band of 200 bp from their father, individual X045. No maternal contribution from individual D045 is evident in the offspring. Identical results were observed when genotyping was repeated on separate DNA samples of family members. Although a very faint band in the region of 206 bp could be visualized in individuals D045, E076 and E077 with either long exposure of autoradiograms (36 to 72 hours) or by reducing the annealing temperature (figure 10), it could not be definitively scored as an allele. A second set of primers (figure 6) was designed to amplify the identical (CA)n repetitive motif at *D5S414*. The forward primer C is situated 15 bp downstream of the original primer A. Similarly, reverse primer D is located 12 bp downstream from primer B. Genotyping with this second set of primers (C and D) demonstrated segregation of alleles in a classical Mendelian co-dominant manner with an allele of 210 bp clearly evident in individual D045 and her children, E076, and E077 (figure 11).

To determine at which published primer site (A and/or B), the template sequence variation was present, two other PCR analyses were performed. The results from primer pair A and D show clear, unambiguous segregation of maternal and paternal alleles (figure 12a). Primer pair C and B yielded results identical to those observed with the original primer pair (A and B) where no maternal allele was apparent in E076 and E077 (figure 12b). These results indicated the sequence variation in individual D045 and inherited by E076 and E077 must lie within the template DNA complementary to reverse primer B.

The D5S414 locus, as defined by the forward and reverse primers, was sequenced in genomic DNA from D045 and a control individual to identify any sequence variation that could lead to a PCR-refractory allele. In the control DNA sequence, TACG directly followed the (CA)n repetitive motif and immediately preceded the 3'-end of reverse primer B. In individual D045, the DNA sequence at this position had a two base pair change of TACG to CGCG immediately preceding the 3'-end of primer B (figure 13). In this region, a six base pair repetitive region is created because the 3'-end of reverse primer B begins with CG. No other sequence variation including, the primer annealing sites, were identified at *D5S414* in individual D045 and the control.

4.3. Haplotype Analysis

4.3.1. CSD1 Kindred

Haplotypes were constructed for each family member in the eight nuclear families based on all 10 microsatellite loci genotyped (figures 14 to 21). The greater the distance between these loci, the greater the chance of recombination. For example, the distance between

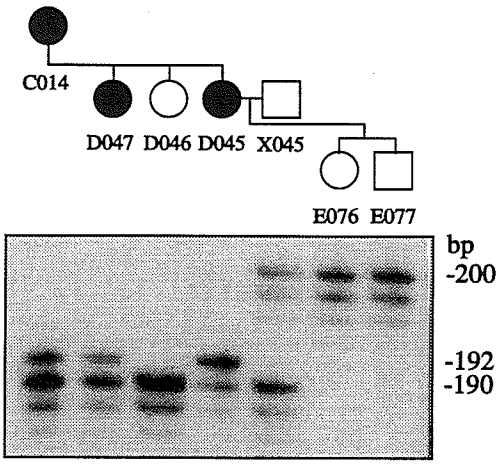


Figure 9: A three generation nuclear family demonstrates non-Mendelian segregation of alleles at *D5S414* using primers A and B. Individuals E076 and E077 appear to have inherited only a paternal allele of 200 bp from X045. No maternal contribution from D045 is evident.

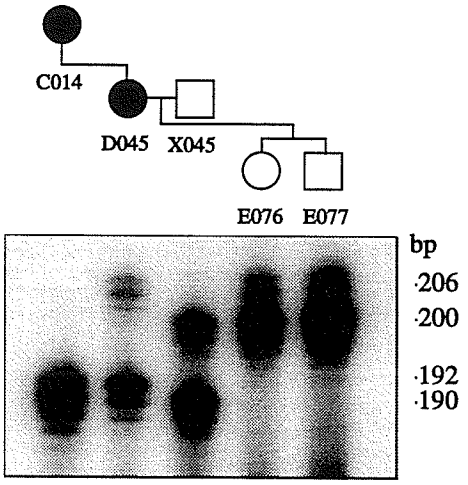


Figure 10: Reducing the annealing temperature during PCR permitted a faint second allele at 206 bp to be visualized in individuals D045, E076 and E077. Similarly, on long exposed autoradiograms this band was occasionally observed.

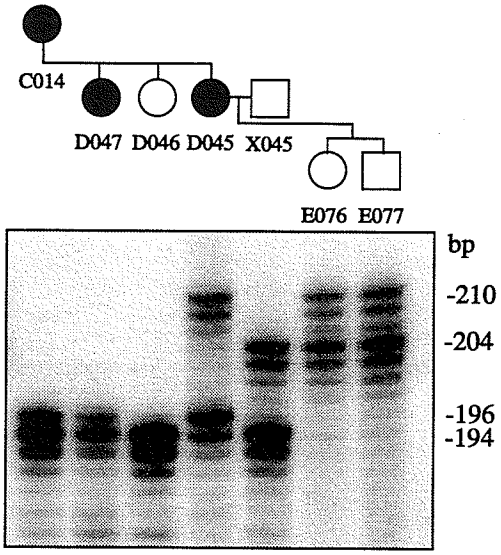


Figure 11: The same three generation nuclear family was genotyped with alternate primers C and D at *D5S414*. The segregation of alleles follows Mendelian codominant inheritance with a maternal allele of 210 bp clearly evident in E076 and E077.

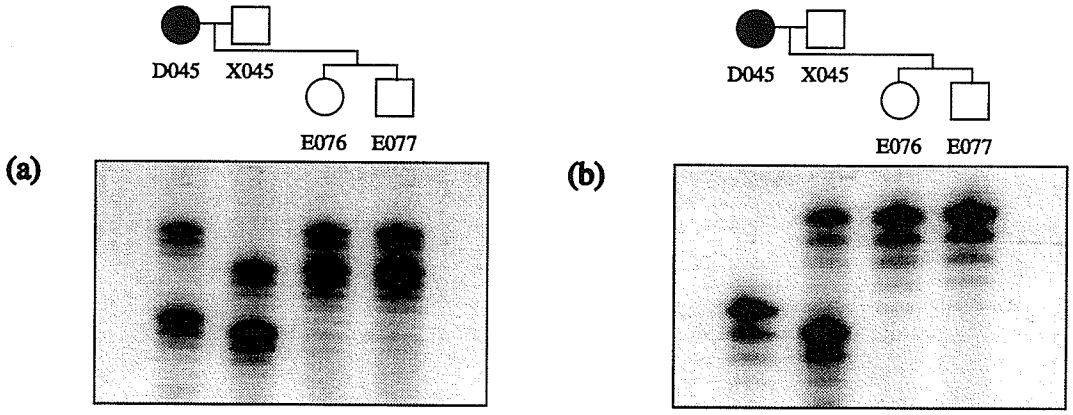


Figure 12: (a) Genotyping with primer pair A and D demonstrate Mendelian codominant segregation of alleles in the four member nuclear family. (b) In contrast, genotyping with primer pair C and D shows apparent segregation of a non-amplifying allele. No maternal contribution is evident in E076 and E077.

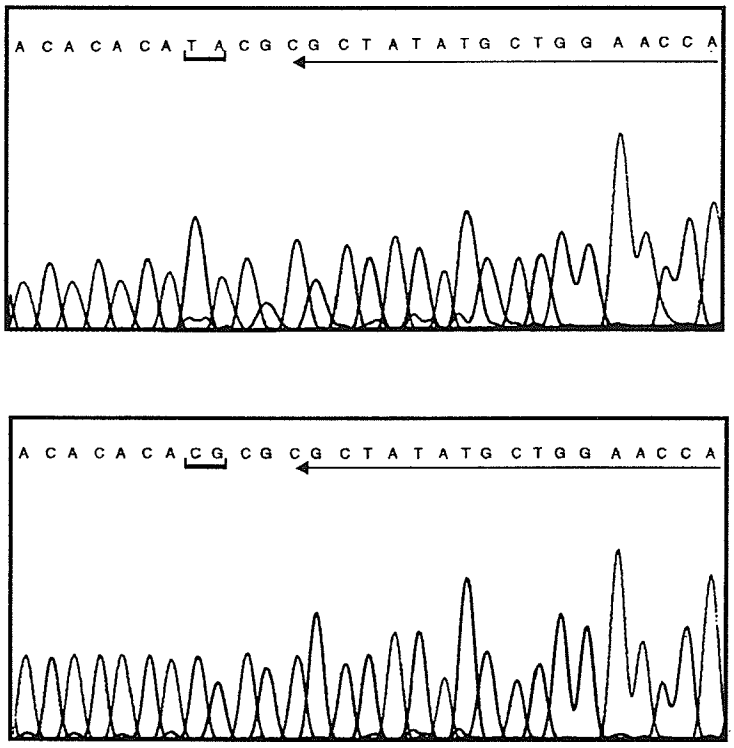


Figure 13: In the control DNA sequence (upper panel), TACG directly follows the (CA)n repetitive motif and immediately precedes the 3'-end of reverse primer B (underlined). The DNA sequence in individual D045 shows a two base pair change of TA→CG (⌊⌋) immediately preceding the 3'-end of the reverse primer B, thus creating a six base pair repetitive region (lower panel).

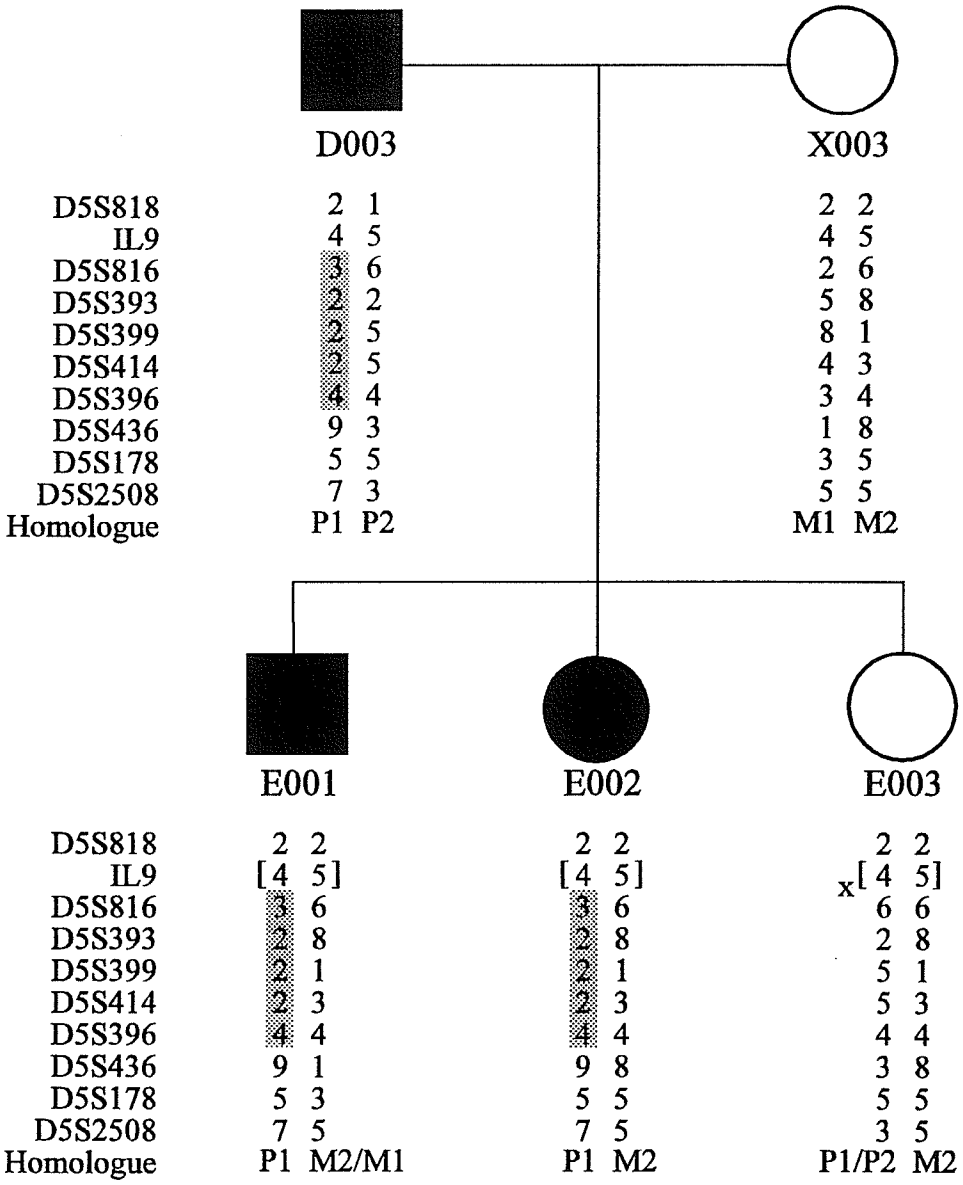


Figure 14: Haplotype constructed for nuclear family 1 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.

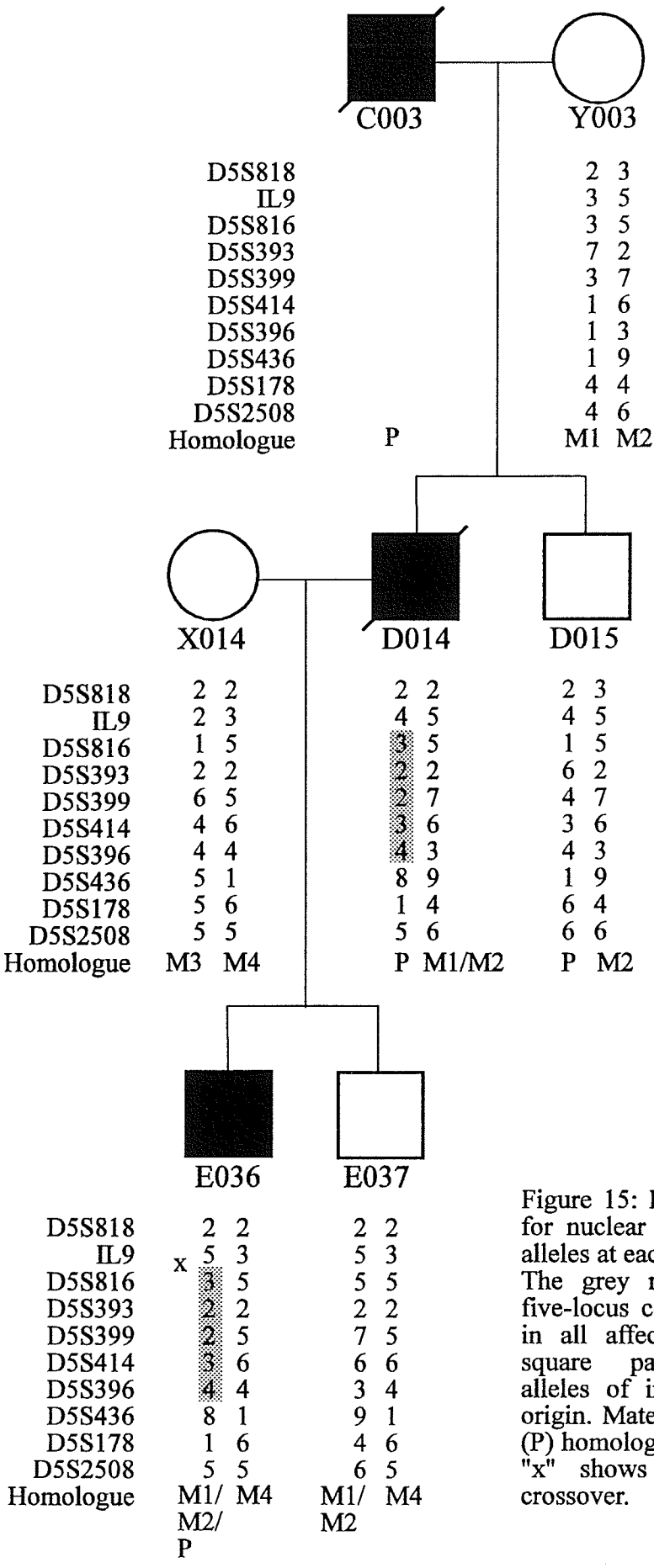


Figure 15: Haplotype constructed for nuclear family 2 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.

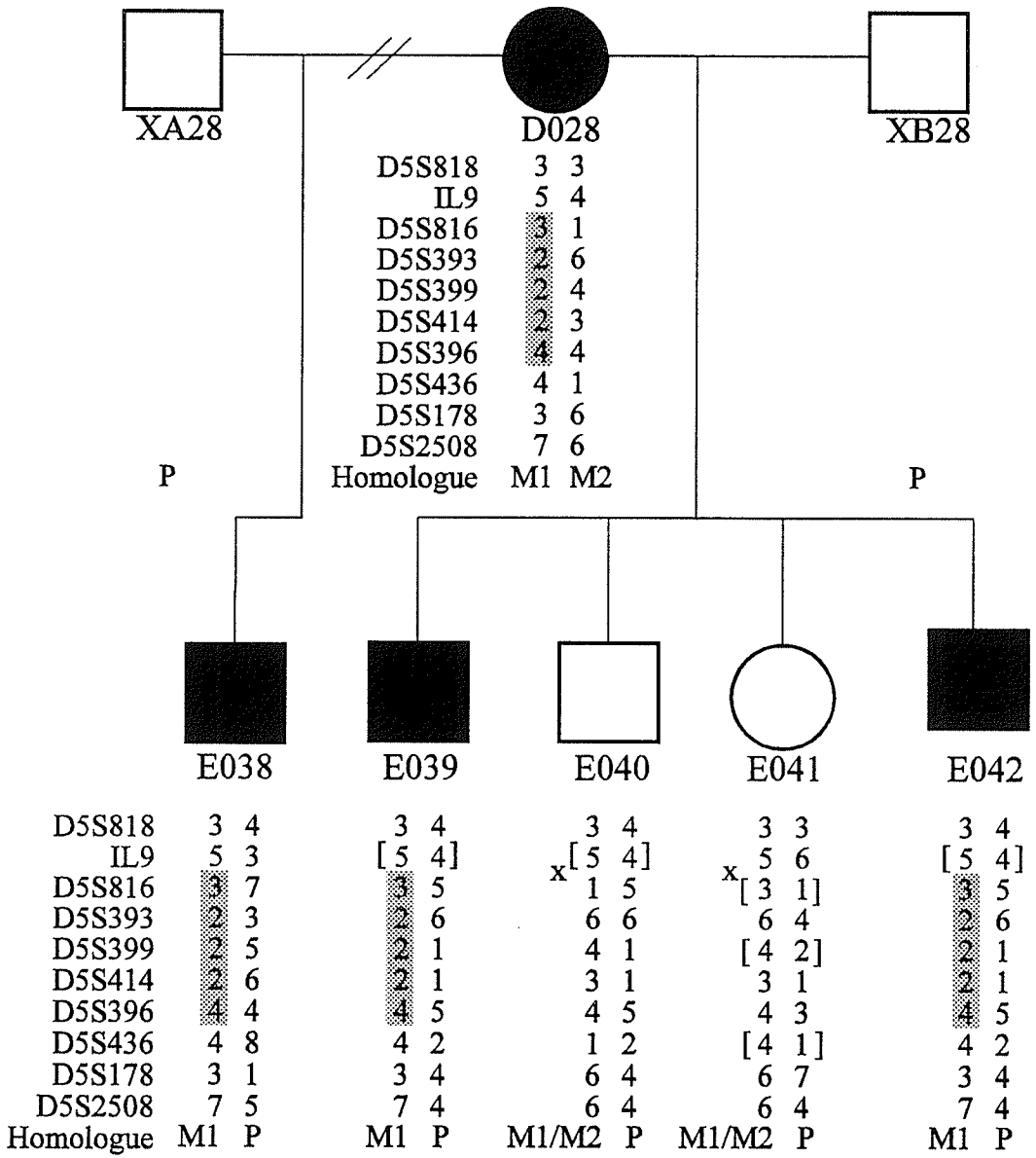
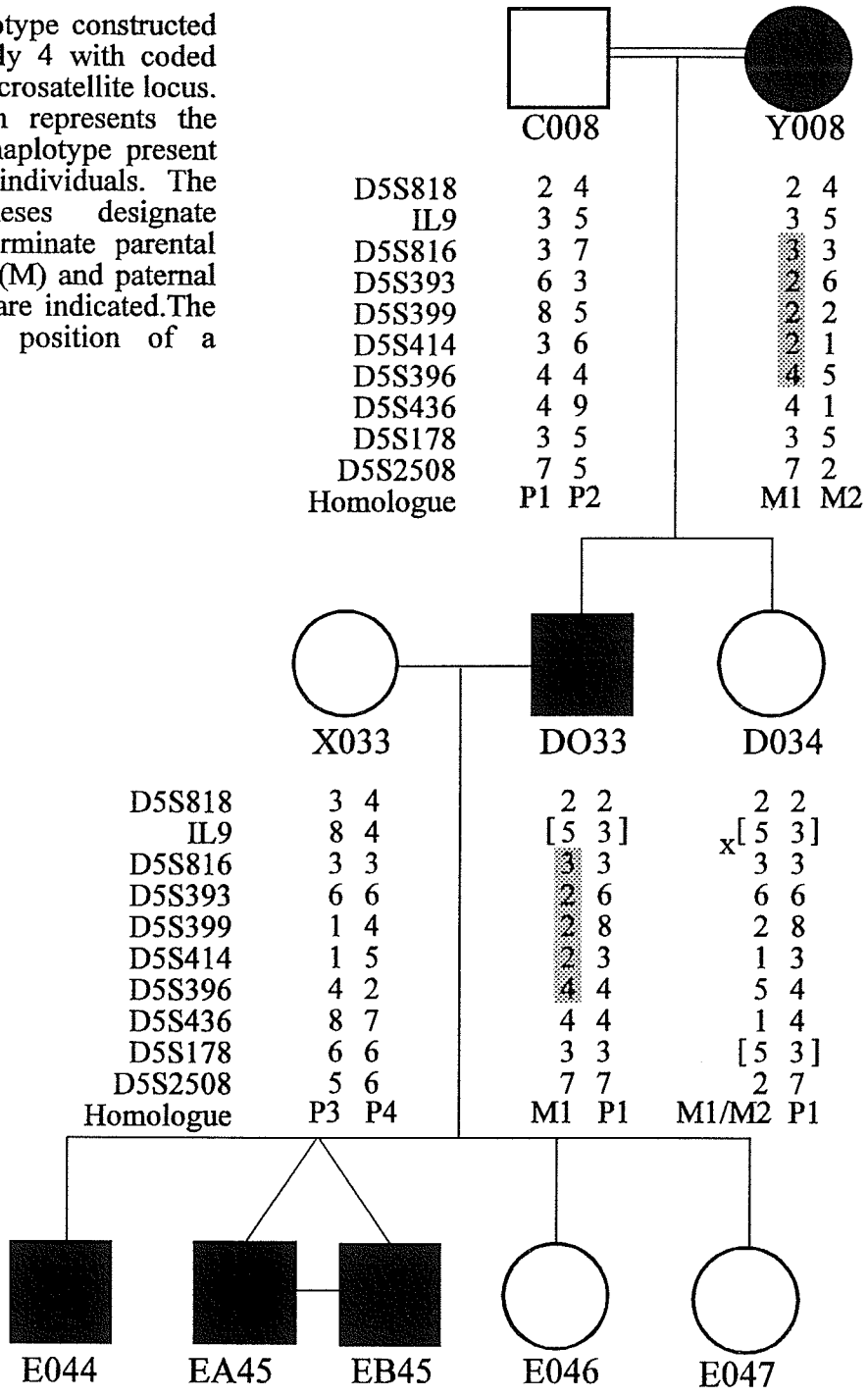


Figure 16: Haplotype constructed for nuclear family 3 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.

Figure 17: Haplotype constructed for nuclear family 4 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated.The "x" shows the position of a crossover.



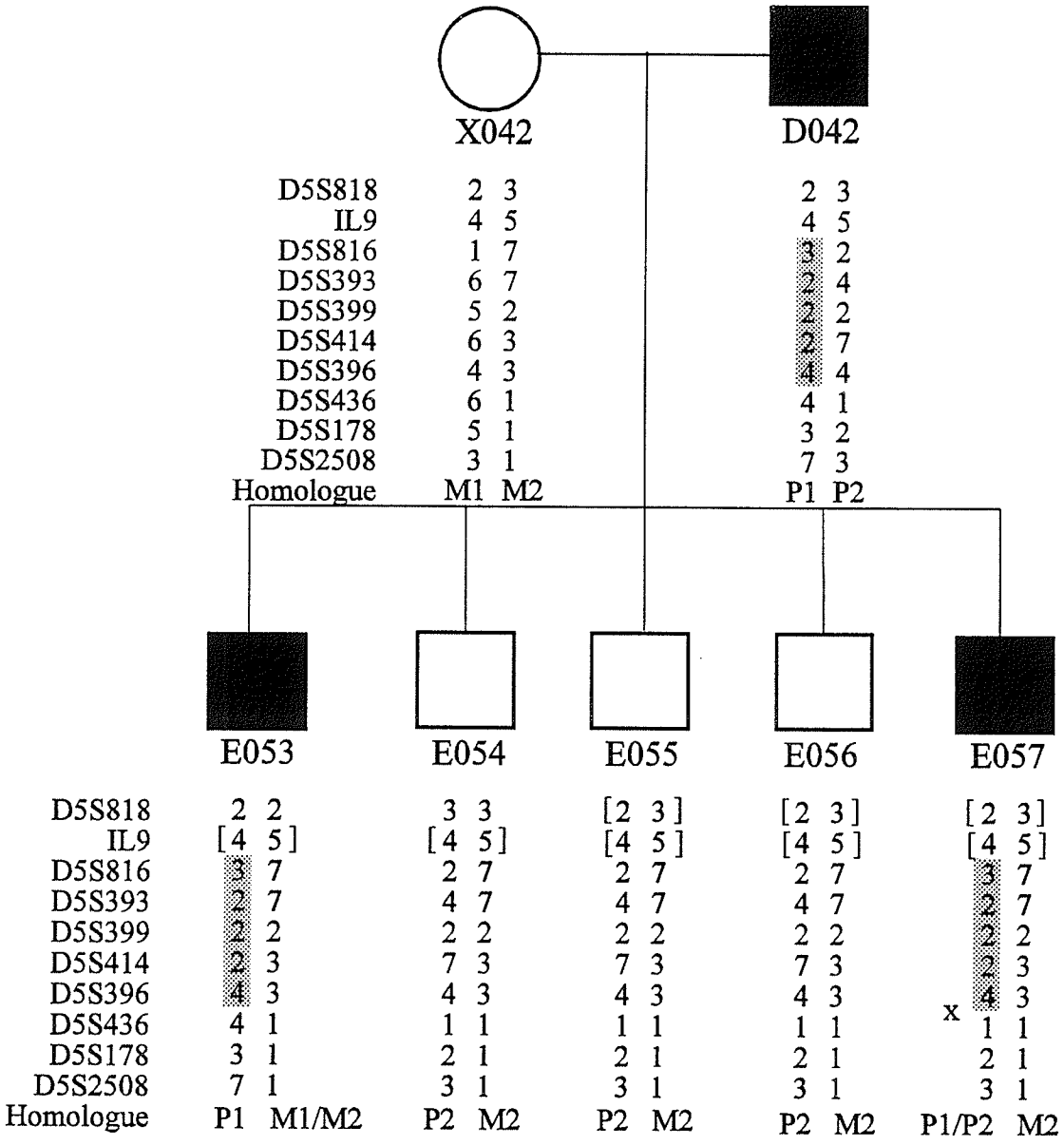
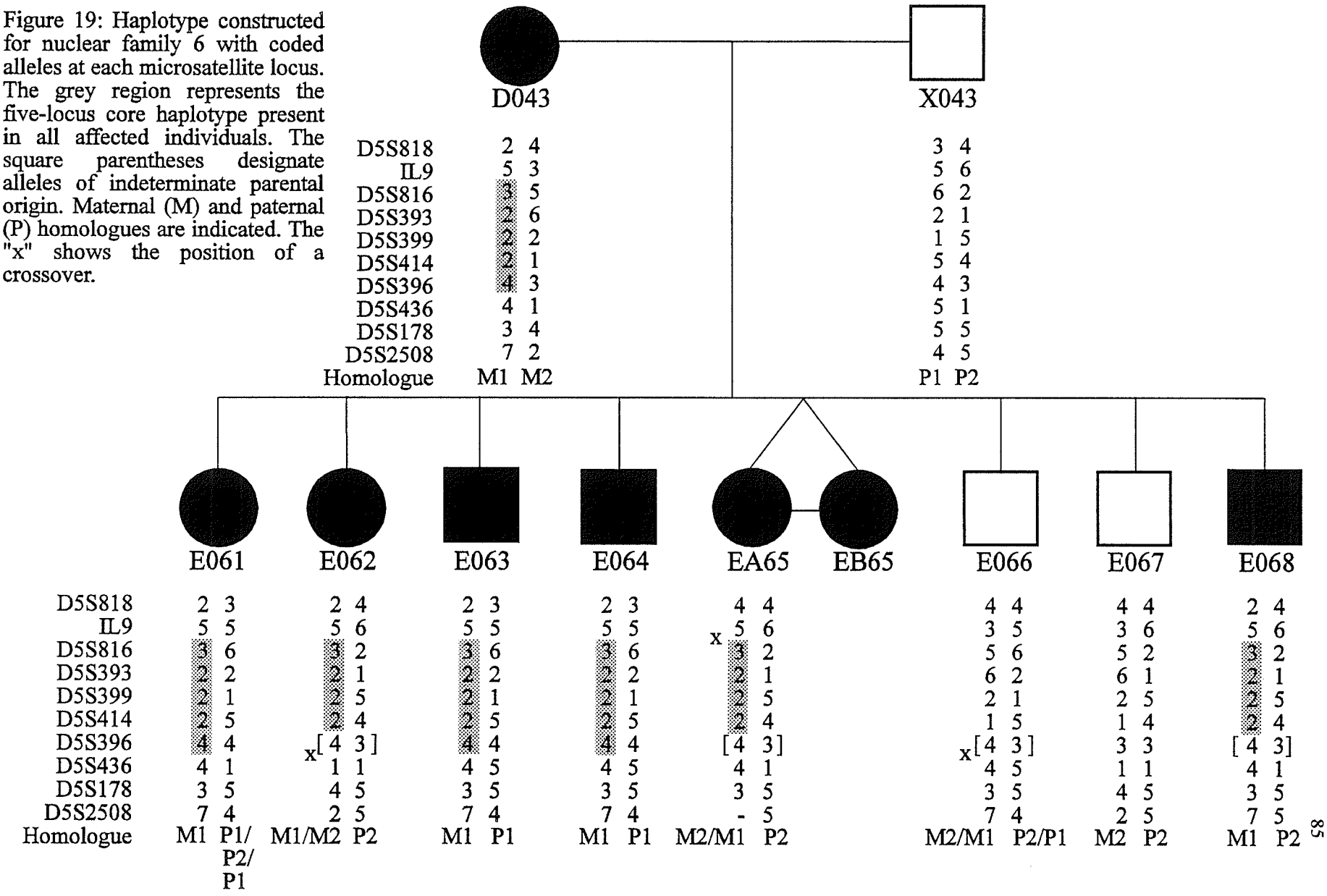


Figure 18: Haplotype constructed for nuclear family 5 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.

Figure 19: Haplotype constructed for nuclear family 6 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.



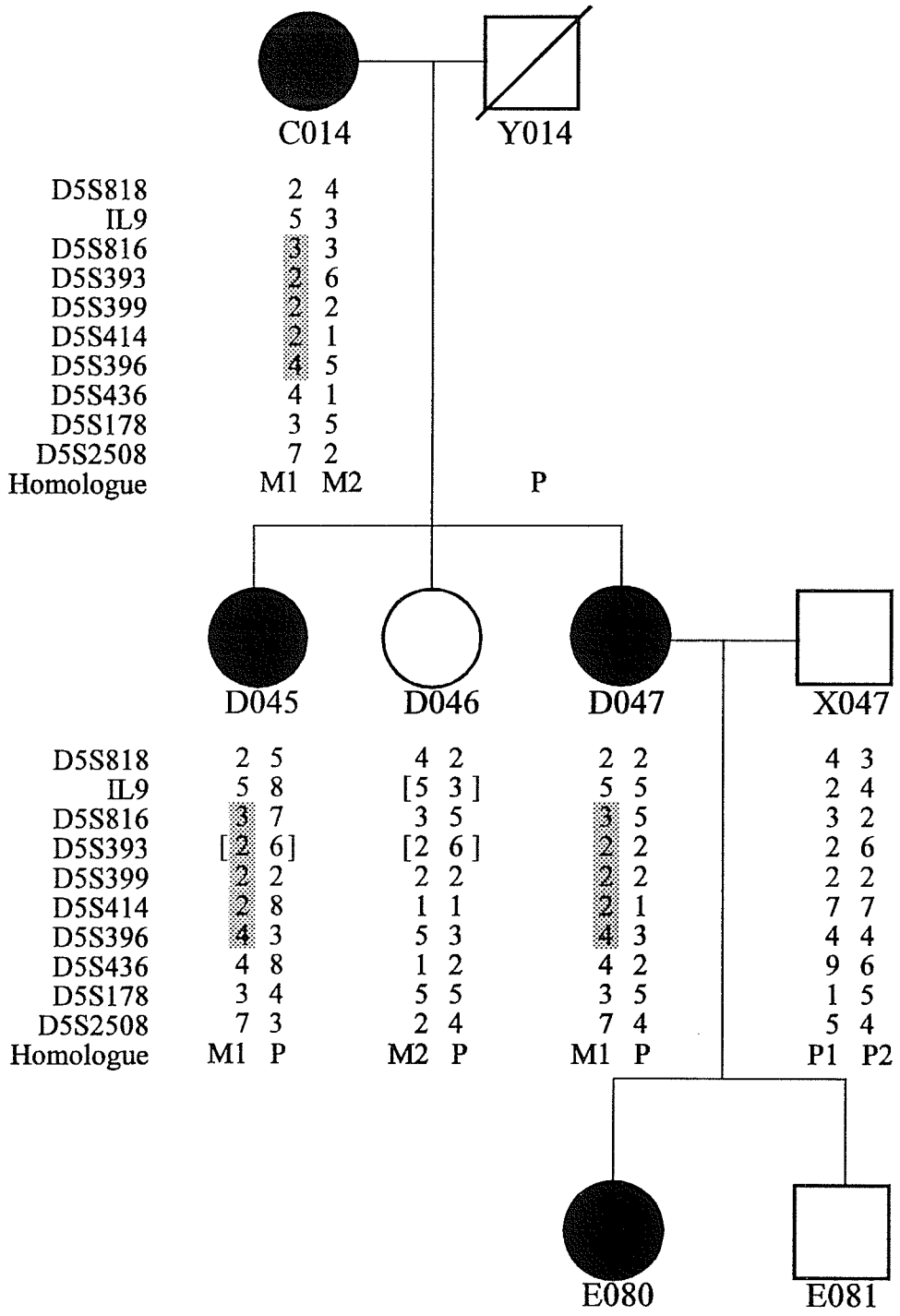


Figure 20: Haplotype constructed for nuclear family 7 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated.

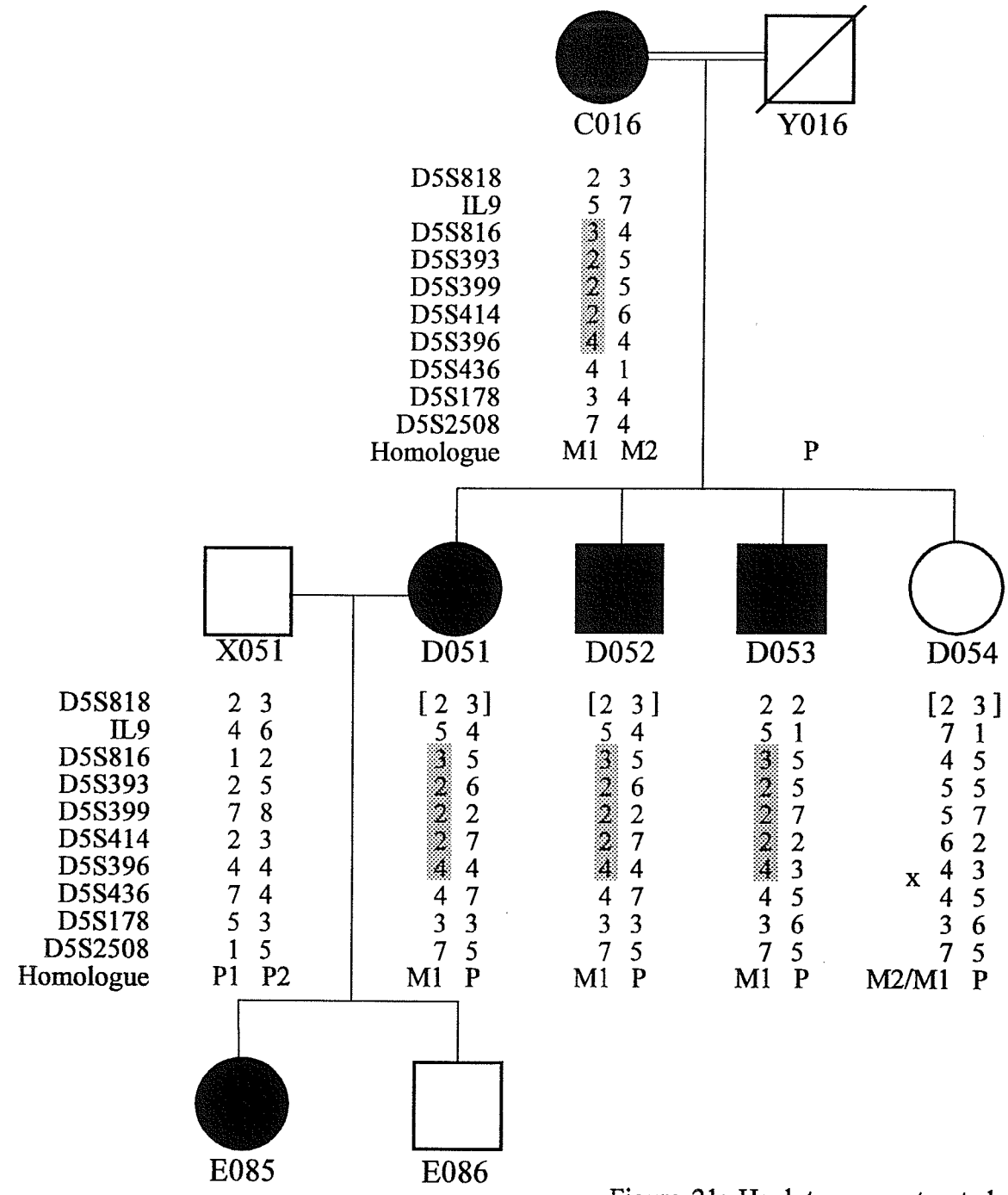


Figure 21: Haplotype constructed for nuclear family 8 with coded alleles at each microsatellite locus. The grey region represents the five-locus core haplotype present in all affected individuals. The square parentheses designate alleles of indeterminate parental origin. Maternal (M) and paternal (P) homologues are indicated. The "x" shows the position of a crossover.

IL9 and *D5S436* is 10 cM, thus, recombination was expected and observed. By analysis of the haplotype data, the 10 microsatellite loci were divided into three subgroups. Group 1 consists of *D5S818* and *IL9*. The central five-locus core consists of *D5S816, D5S393, D5S399, D5S396* and *D5S414*. And group 3 consists of *D5S436, D5S178* and *D5S2508*. Of the 84 haplotypes (42 meioses), 65 haplotypes were inherited intact and 19 haplotypes were derived from parental recombination. The parental origin of these 19 recombinant haplotypes is as follows: 1 haplotype derived from a single maternal crossover between group 1 and the five-locus core; 5 haplotypes (3 maternal and 2 paternal) derived from a single crossover between the five-locus core and group 3; 5 haplotypes (3 maternal and 2 paternal) derived from a single crossover within group 1; and 1 haplotype derived from a double paternal crossover within the third marker group. Although the 10 microsatellite markers studied have maximum heterozygosities ranging from 0.66 to 0.95, there were instances where the parental origin of an allele could not be determined. Therefore, in the 7 remaining haplotypes (5 maternal and 2 paternal), the position of the crossover event within group 1 or between group 1 and the five-locus core could not be determined.

There was no evidence of recombination between the markers of the five-locus core and *CSD1*D* and *CSD1*N*, confirming that these loci are tightly linked and positioned in close proximity to each other in the interval 5q22-q33.3. All 34 affected individuals, except for two, inherited the haplotype, "3, 2, 2, 2, 4", for this five-locus core from the affected parent. The exception was noted in family members, D014 and his son, E036 who inherited the haplotype "3, 2, 2, 3, 4" at the five-locus core. In this family, a 194 bp allele (coded 3) was segregating

with *CSD1* at *D5S414* rather than a 192 bp allele (coded 2) which was present in all of the other affected family members (figure 15).

From the haplotype analysis, there is evidence of recombination only within marker groups 1 and 3 and between these latter two groups and the five-locus core. No recombination was observed between the five markers, *D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414*. Thus the haplotype data places *CSD1* within the five-locus core.

4.3.2. At-risk Individuals

Haplotypes were constructed for all 15 at-risk family members in 7 nuclear families based on all 10 microsatellite loci genotyped. For reasons of confidentiality, the haplotype data are not included. Of the 30 at-risk haplotypes (15 meioses), 23 (9 maternal and 14 paternal) were inherited intact and 7 haplotypes were derived from parental recombination. These recombinant haplotypes consisted of 3 (2 maternal and 1 paternal) derived from a crossover between group one and the five-locus core; 2 (1 maternal and 1 paternal) derived from a crossover between the five-locus core and group three; and the remaining 2 haplotypes derived from maternal crossover events within marker group three. Similar to the haplotype analysis in affected and unaffected family members 25 years of age or older, there was no apparent recombination between either *CSD1*D* or *CSD1*N* and the five-locus core. Of the 15 at-risk individuals genotyped, haplotype analysis revealed that five individuals have inherited the identical five-locus core haplotype "3, 2, 2, 2, 4" observed to be segregating with *CSD1*D* in all affected family members. These individuals are predicted to have a higher risk of becoming affected.

4.4. Linkage Analysis

In this large CSD1 kindred, there are 42 possible meioses to be scored. The maternal and paternal two-point lod scores generated by MARK III (Côté, 1975) with each of the 10 microsatellite marker loci genotyped are summarised in tables 6 to 15. The most informative marker was *D5S414* for which 40 meioses were scored. The least informative marker was *D5S396* for which only 13 meioses were scored. Among the five markers, *D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414* that showed no apparent recombination with *CSD1*, *D5S414* yielded the most positive lod score, 9.93 at $\hat{\Theta}$=0.00 in the MARK III analysis. In this cluster of markers, *CSD1:D5S396* yielded a lod score of 2.71 at $\hat{\Theta}$=0.00. As this lod score lies between -2 and 3, linkage cannot be unequivocally excluded or declared. The three other markers, *D5S816*, *D5S393* and *D5S399*, demonstrated peak lod scores of 7.53, 7.83 and 4.52 at $\hat{\Theta}$=0.00, respectively, with *CSD1*.

By inspection of the data for each nuclear family, recombination was evident between *CSD1* and *D5S818*, *IL9*, *D5S436*, *D5S178* and *D5S2508*. Three (2 affected and 1 unaffected) of sixty family members genotyped showed recombination between *CSD1* and *D5S818*. Recombination between *CSD1* and *IL9* was observed among four (1 affected and 3 unaffected) family members. Among another four individuals (2 affected and 2 unaffected) recombination was observed between *CSD1* and each of *D5S436*, *D5S178* and *D5S2508*.

Inspection of the genotype data for the complete kindred reveals two different alleles at *D5S414* to be segregating with *CSD1* (figure 22). Among seven of the eight nuclear families, an allele of 192 bp is segregating with *CSD1*D*. An alternate allele of 194 bp is segregating with *CSD1*D* in one branch of the kindred in affected individual D014 and inherited by his

Table 6: Summary of maternal and paternal lod scores between *CSD1* and *D5S818*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z2:1	$-\infty$	-1.40	-0.72	-0.44	-0.19	-0.08	-0.02
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Total	$-\infty$	-1.11	-0.46	-0.23	-0.06	-0.01	0.00
Maternal							
Z1:1	$-\infty$	-1.40	-0.72	-0.44	-0.19	-0.08	-0.02
Z7:1	$-\infty$	0.08	0.65	0.79	0.73	0.50	0.19
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Total	$-\infty$	-0.74	0.46	0.81	0.85	0.60	0.22
Total	$-\infty$	**-1.85**	**0.00**	**0.58**	**0.79**	**0.59**	**0.22**
PCMAP83 Estimate	Zmax=0.80 at $\hat{\Theta}$=0.19						

Table 7: Summary of maternal and paternal lod scores between *CSD1* and *IL9*

	Recombination Fraction (Θ)						
Segregation Information	**0.00**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z1:1	$-\infty$	-1.40	-0.72	-0.44	-0.19	-0.08	-0.02
Z2:2	$-\infty$	-2.81	-1.44	-0.89	-0.39	-0.15	-0.04
Total	$-\infty$	-4.21	-2.16	-1.33	-0.58	-0.23	-0.05
Maternal							
Z1:1	$-\infty$	-1.40	-0.72	-0.44	-0.19	-0.08	-0.02
Z8:0	2.11	2.07	1.93	1.74	1.33	0.87	0.35
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	4.21	2.74	3.11	2.99	2.38	1.55	0.63
Total	$-\infty$	**-1.47**	**0.95**	**1.66**	**1.80**	**1.33**	**0.58**
PCMAP83 Estimate	Zmax=1.83 at $\hat{\Theta}$=0.17						

Table 8: Summary of maternal and paternal lod scores between *CSD1* and *D5S816*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Total	2.41	2.36	2.14	1.87	1.31	0.73	0.23
Maternal							
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
Z8:0	2.11	2.07	1.93	1.74	1.33	0.87	0.35
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	5.12	5.03	4.67	4.20	3.18	2.05	0.85
Total	**7.53**	**7.39**	**6.82**	**6.07**	**4.49**	**2.78**	**1.09**
PCMAP83 Estimate	Zmax=7.20 at $\hat{\Theta}$=0.03						

Table 9: Summary of maternal and paternal lod scores between *CSD1* and *D5S393*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
4NR:0R	1.20	1.19	1.12	1.02	0.82	0.58	0.32
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Total	2.71	2.66	2.47	2.21	1.67	1.08	0.48
Maternal							
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z8:0	2.11	2.07	1.93	1.74	1.33	0.87	0.35
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	5.12	5.03	4.65	4.16	3.11	1.96	0.77
Total	**7.83**	**7.69**	**7.12**	**6.37**	**4.78**	**3.04**	**1.25**
PCMAP83 Estimate	Zmax=7.52 at $\hat{\Theta}$=0.03						

Table 10: Summary of maternal and paternal lod scores between *CSD1* and *D5S399*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
4NR:0R	1.20	1.19	1.12	1.02	0.82	0.58	0.32
Total	2.71	2.66	2.47	2.21	1.68	1.11	0.54
Maternal							
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
Total	1.81	1.77	1.63	1.44	1.03	0.60	0.19
Total	**4.52**	**4.43**	**4.09**	**3.65**	**2.71**	**1.71**	**0.73**
PCMAP83 Estimate	Zmax=4.32 at $\hat{\Theta}$=0.03						

Table 11: Summary of maternal and paternal lod scores between *CSD1* and *D5S396*

	Recombination Fraction (Θ)						
Segregation Information	**0.00**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Maternal							
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Total	2.11	2.06	1.86	1.62	1.10	0.60	0.18
Total	**2.71**	**2.65**	**2.42**	**2.13**	**1.51**	**0.89**	**0.34**
PCMAP83 Estimate	Zmax=2.56 at $\hat{\Theta}$=0.03						

Table 12: Summary of maternal and paternal lod scores between *CSD1* and *D5S414*

	Recombination Fraction (Θ)						
Segregation Information	**0.00**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
4NR:0R	1.20	1.19	1.12	1.02	0.82	0.58	0.32
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Total	3.61	3.55	3.30	2.97	2.26	1.48	0.67
Maternal							
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z8:0	2.11	2.07	1.93	1.74	1.33	0.87	0.35
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	6.32	6.21	5.74	5.14	3.84	2.42	0.97
Total	**9.93**	**9.76**	**9.04**	**8.11**	**6.10**	**3.90**	**1.65**
PCMAP83 Estimate	Zmax=9.55 at $\hat{\Theta}$=0.03						

Table 13: Summary of maternal and paternal lod scores between *CSD1* and *D5S436*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z4:1	$-\infty$	-0.81	-0.19	0.02	0.12	0.10	0.03
Total	$-\infty$	0.66	1.16	1.21	0.98	0.62	0.26
Maternal							
Z4:0	0.90	0.89	0.81	0.72	0.52	0.30	0.09
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z6:2	$-\infty$	-1.92	-0.63	-0.17	0.13	0.15	0.06
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z3:1	$-\infty$	-1.11	-0.46	-0.23	-0.06	-0.01	0.00
1NR:0R	0.30	0.30	0.28	0.26	0.20	0.15	0.08
Total	2.71	-0.37	1.35	1.77	1.65	1.11	0.46
Total	$-\infty$	**0.29**	**2.51**	**2.98**	**2.64**	**1.73**	**0.71**
PCMAP83 Estimate	Zmax=2.99 at $\hat{\Theta}$=0.12						

Table 14: Summary of maternal and paternal lod scores between *CSD1* and *D5S178*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z4:1	$-\infty$	-0.81	-0.19	0.02	0.12	0.10	0.03
Total	$-\infty$	0.07	0.63	0.75	0.67	0.45	0.21
Maternal							
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Z6:2	$-\infty$	-1.92	-0.63	-0.17	0.13	0.15	0.06
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z3:1	$-\infty$	-1.11	-0.46	-0.23	-0.06	-0.01	0.00
Total	2.41	-0.67	1.09	1.56	1.52	1.03	0.41
Total	$-\infty$	**-0.59**	**1.72**	**2.30**	**2.18**	**1.48**	**0.62**
PCMAP83 Estimate	Zmax=2.43 at $\hat{\Theta}$=0.14						

Table 15: Summary of maternal and paternal lod scores between *CSD1* and *D5S2508*

Segregation Information	Recombination Fraction (Θ)						
	0.00	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**
Paternal							
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z4:1	$-\infty$	-0.81	-0.19	0.02	0.12	0.10	0.03
Total	$-\infty$	0.66	1.16	1.21	0.98	0.62	0.26
Maternal							
Z5:0	1.20	1.18	1.09	0.98	0.72	0.44	0.15
Z2:0	0.30	0.29	0.26	0.22	0.13	0.06	0.02
Z5:2	$-\infty$	-2.22	-0.91	-0.42	-0.07	-0.02	0.01
Z3:0	0.60	0.59	0.53	0.47	0.32	0.17	0.05
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Z3:1	$-\infty$	-1.11	-0.46	-0.23	-0.06	-0.01	0.00
2NR:0R	0.60	0.59	0.56	0.51	0.41	0.29	0.16
Total	$-\infty$	-0.08	1.63	2.03	1.86	1.22	0.54
Total	$-\infty$	**0.58**	**2.79**	**3.24**	**2.84**	**1.85**	**0.80**
PCMAP83 Estimate	Zmax=3.24 at $\hat{\Theta}$=0.11						

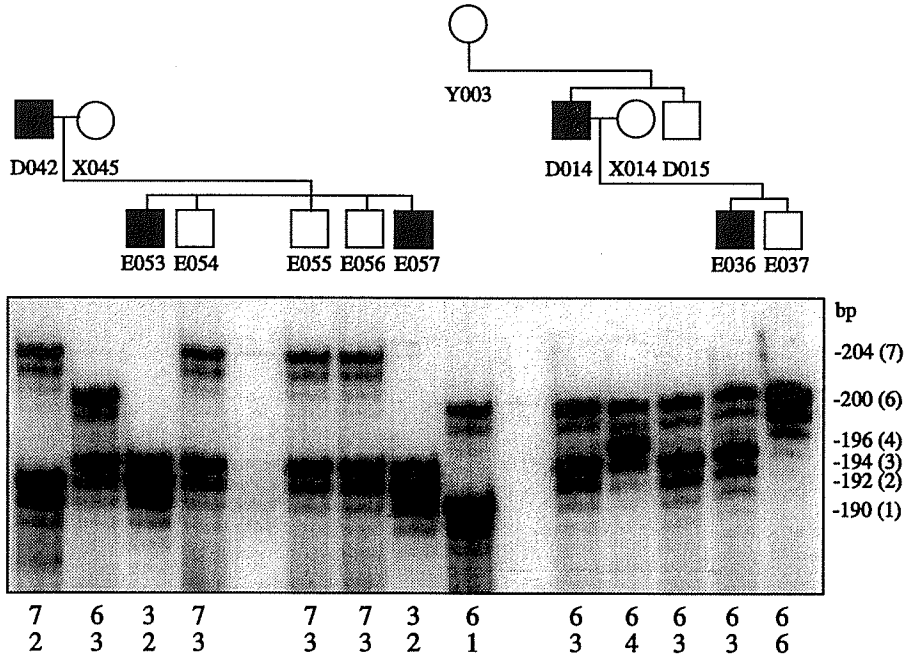


Figure 22: Two nuclear families genotyped at *D5S414*. An allele of 192 bp (coded 2) is segregating with *CSD1**D in the first nuclear family. In the second, a 194 bp allele (coded 3) is segregating with *CSD1**D in individuals D014 and E036. The coded genotypes are indicated below.

affected son E036. This finding was not reflected in the MARK III analysis because the lod scores were determined independently for each family unit.

Data analysis using the LINKAGE computer program package (Lathrop *et al*, 1984; Lathrop and Lalouel, 1984; Lathrop *et al*, 1986) consistently generated higher two-point lod scores than the MARK III and PCMAP83 analyses (table 16 and 17). The most significant two-point lod score, 12.57 at $\hat{\Theta}$=0.02, was observed with *CSD1:D5S414*. Under this method of analysis, linkage between *CSD1* and *D5S396* (Z=4.49 at $\hat{\Theta}$=0.00) meets the requirements to declare linkage. Between *CSD1* and the three other loci, *D5S816*, *D5S393* and *D5S399*, peak lod scores of 10.66, 11.76 and 7.35 respectively were generated at $\hat{\Theta}$=0.00 as well. Based on the trend observed in the MARK III and PCMAP83 analyses, the peak lod scores were predicted to occur at similar recombination fractions for all the markers in the five-locus core. The larger *CSD1:D5S414* recombination fraction returned by LINKAGE was most likely attributed to the alternate allele observed to be segregating at *D5S414* as previously described.

The LINKAGE analysis was repeated on a modified dataset for which all 13 unaffected siblings were removed to evaluate the confounding possibility of incomplete penetrance. As expected, the lod scores generated between *CSD1* and each microsatellite marker loci were lower but consistent with the initial LINKAGE analysis (tables 17 and 18). The most significant two-point lod score, 7.74 at $\hat{\Theta}$=0.03, was observed with *CSD1:D5S414*. At the four other loci, *D5S816*, *D5S393*, *D5S399* and *D5S396* peak lod scores of 6.87, 7.11, 5.15 and 3.11, respectively were generated at $\hat{\Theta}$=0.00 with *CSD1*.

Table 16: Summary lod scores for *CSD1* and 5q microsatellite markers using LINKAGE

	Recombination fraction (Θ)								
Locus	**0.00**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**	**Zmax**	**$\hat{\Theta}$**
D5S818	$-\infty$	-1.55	0.84	1.53	1.65	1.20	0.50	1.71	0.16
IL9	$-\infty$	-2.46	1.26	2.42	2.77	2.16	1.05	2.80	0.17
D5S816	10.66	10.50	9.83	8.94	7.00	4.85	2.50	10.66	0.00
D5S393	11.76	11.59	10.87	9.87	7.62	5.06	2.27	11.76	0.00
D5S399	7.35	7.21	6.65	5.92	4.38	2.73	1.15	7.35	0.00
D5S414	$-\infty$	12.50	12.25	11.30	8.90	6.07	2.91	12.57	0.02
D5S396	4.49	4.41	4.10	3.70	2.86	1.95	0.97	4.49	0.00
D5S436	$-\infty$	-0.37	3.11	3.99	3.79	2.64	1.11	4.09	0.13
D5S178	$-\infty$	-0.46	3.06	3.99	3.92	2.94	1.49	4.14	0.14
D5S2508	$-\infty$	2.61	5.31	5.80	5.18	3.79	1.98	5.86	0.11

Table 17: Summary of peak lod scores at their corresponding recombination fraction generated by PCMAP83 and LINKAGE between *CSD1* and a series of 5q microsatellite markers

	PCMAP83 estimates		LINKAGE estimates		*LINKAGE estimates	
LOCUS	**Zmax**	**$\hat{\Theta}$**	**Zmax**	**$\hat{\Theta}$**	**Zmax**	**$\hat{\Theta}$**
D5S818	0.80	0.19	1.71	0.16	1.94	0.10
IL9	1.83	0.17	2.80	0.17	3.06	0.08
D5S816	7.20	0.03	10.66	0.00	6.87	0.00
D5S393	7.52	0.03	11.76	0.00	7.11	0.00
D5S399	4.32	0.03	7.35	0.00	5.15	0.00
D5S414	9.55	0.03	12.57	0.02	7.74	0.03
D5S396	2.56	0.03	4.49	0.00	3.11	0.00
D5S436	2.99	0.12	4.09	0.13	2.44	0.12
D5S178	2.43	0.14	4.14	0.14	2.59	0.14
D5S2508	3.24	0.11	5.86	0.11	3.98	0.10

*All unaffected siblings excluded from analysis.

Table 18: Summary lod scores for *CSD1* and 5q microsatellite markers using LINKAGE*

	Recombination fraction (Θ)								
Locus	**0.00**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**	**Zmax**	**$\hat{\Theta}$**
D5S818	$-\infty$	0.63	1.74	1.94	1.65	1.07	0.42	1.94	0.10
IL9	$-\infty$	2.40	3.51	3.61	2.99	1.96	0.80	3.06	0.08
D5S816	6.87	6.79	6.40	5.83	4.51	3.06	1.53	6.87	0.00
D5S393	7.11	6.99	6.51	5.84	4.35	2.73	1.14	7.11	0.00
D5S399	5.15	5.05	4.61	4.05	2.87	1.67	0.62	5.15	0.00
D5S414	$-\infty$	7.60	7.60	7.01	5.40	3.52	1.55	7.74	0.03
D5S396	3.11	3.05	2.82	2.51	1.89	1.25	0.60	3.11	0.00
D5S436	$-\infty$	-0.44	1.86	2.40	2.17	1.34	0.50	2.44	0.12
D5S178	$-\infty$	-0.42	1.91	2.51	2.41	1.71	0.77	2.59	0.14
D5S2508	$-\infty$	2.02	3.68	3.98	3.52	2.53	1.27	3.98	0.10

*All unaffected siblings excluded from analysis.

4.5. PCMAP83 Analysis

The PCMAP83 (Sherman *et al,* 1984) computer program was used to test the most likely order of the microsatellite loci *D5S818*, *IL9*, *D5S436*, *D5S178* and *D5S2508* relative to *CSD1*. Recombination between these five markers and *CSD1* was observed. The best possible order of markers was determined from a Chi-square test on separate maternal, paternal and combined lod score data, where the p-value was as close to or equal to 1.0. The results for the combined data as well as for male and female segrgation data analysed separately indicate the most likely order is: *IL9-D5S818-CSD1-D5S2508-D5S178-D5S436* at varying levels of interference from 0.0 to 0.5 (table 19).

The markers from the five-locus core (*D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414*) were not included in the initial analysis. These markers showed no apparent recombination with *CSD1*, suggesting that *CSD1* is positioned within the five-locus core. Therefore, the disease locus was used in the initial analysis as a representative of the core. The analysis was repeated using *D5S414* to represent the core because the greatest number of meioses were scored between *CSD1:D5S414* and *D5S414* and other microsatellite loci. To include the remaining core markers in the analysis would not be meaningful because marker-marker comparisons between the core markers and those flanking are each informative for different children. The results at varying levels of interference from 0.0 to 0.5 for both males and females are consistent with the initial analysis where *IL9-D5S818-D5S414-D5S2508-D5S178-D5S436* is predicted to be the most likely order. Here, *D5S414* is in the same sequential position as *CSD1* relative to the

Table 19: Summary of PCMAP83 analysis at varying levels of interference for the most likely ordering of five microsatellite markers

	Levels of interference for data analysis								
	Combined male and female			Male			Female		
Proposed map orde	**0.00**	**0.351**	**0.50**	**0.00**	**0.351**	**0.50**	**0.00**	**0.351**	**0.50**
IL9									
	5.48	5.82	6.17	13.96	14.35	12.20	0.55	0.58	0.64
D5S818									
	7.19	7.42	7.75	0.00	0.00	4.32	10.21	10.44	10.69
CSD1									
	7.78	7.97	8.26	5.99	6.20	4.52	12.51	12.83	13.34
D5S2508									
	2.40	2.43	2.47	1.30	1.32	1.37	3.05	3.09	3.15
D5S178									
	2.71	2.72	2.75	2.15	2.14	2.05	3.10	3.12	3.13
D5S436									
Total Length (cM)	25.56	26.36	27.40	23.41	24.01	24.46	29.43	30.10	30.95
Chi-Square Test	16.47	16.19	15.87	9.83	9.68	7.96	2.82	2.84	2.86
p-value	0.88	0.89	0.90	0.48	0.49	0.57	0.99	0.99	0.98

flanking markers. The order of the flanking markers is identical to the initial analysis performed to position the disease locus.

In general, increasing the level of interference for both males and females increased the p-value for the Chi-square test. When testing potential orders of loci, exceptions were noted when the disease locus was excluded or a representative of the core was included in the analysis. In these instances, the p-value for the Chi-square test for separate and combined maternal and paternal lod score data remained unchanged at varying levels of interference.

5. Discussion

The search for a disease-causing gene and its underlying genetic abnormality can be pursued via two approaches. The "candidate gene" approach to identifying the CSD1 gene(s) in itself did not implicate any particular gene(s), as the nature of the amyloid deposit remains uncharacterized. Since, there are no cytogenetic rearrangements associated with the CSD1 phenotype, a positional cloning approach was initiated by a genome-wide scan of microsatellite markers to chromosomally map *CSD1*. At the onset of this study, a 10 cM interval at 5q22-5q33.3 defined the candidate interval for CSD1 (Stone *et al*, 1994). Our linkage analysis confirmed the mapping of *CSD1* to 5q in a large multigenerational kindred. Furthermore, it confirmed the localization of *CSD1* to the identical 2 cM interval defined by *D5S393* (proximally) and *D5S396* (distally) as reported by Gregory *et al* (1995). The series of 5q microsatellite markers screened in this kindred could not be definitively ordered due to the limitations of the available resources and the

number of informative meioses. Electronically available genetic and physical maps will provide further information about the potential ordering of the microsatellite markers, particularly in the five-locus core as well as other available markers to screen within the candidate region. Together, these resources will help in directing the search for the gene(s) responsible for CSD1.

5.1. Linkage Analysis

Linkage analysis was performed using three well established methods. The first two methods, manual lod score calculations using lod score tables and MARK III generated lod scores require a large kindred to be broken into nuclear family units before analysis. Lod scores are then calculated separately for each nuclear family. These can be summed to obtain separate maternal and paternal lod scores and combined to obtain an overall total lod score. The results generated by either method are expected to be in agreement. In our analysis, lod scores calculated manually between *CSD1* and microsatellite loci are in complete agreement with those generated by the MARK III analysis (Côté, 1975). Both of these approaches generate lod scores independently for each nuclear family and exclude individuals linking the family units together. In this particular kindred, there are two kinship loops defined by two brothers (C006 and C008) marrying two sisters (Y006 and Y008). The exclusion of this familial relationship in linkage analysis potentially results in a substantial loss of information.

The third method to perform linkage analysis utilizes the LINKAGE program (Lathrop *et al*, 1984; Lathrop and Lalouel, 1984; Lathrop *et al*, 1986). LINKAGE is

designed to incorporate various parameters involved in human genetic disease into the analysis. Although the complete kindred is considered, the kinship loops must be broken prior to the analysis by inserting a genetically identical individual into the pedigree (Terwilliger and Ott, 1994). Parameters such as the mode of inheritance, gene frequency, allele frequency and penetrance must be stipulated in the analysis and can be varied to test the data under different models (Khoury *et al*, 1993).

The two-point lod scores calculated by LINKAGE (assuming an autosomal dominant mode of inheritance with complete penetrance, a disease gene frequency of $1x10^{-5}$ and equal microsatellite marker allele frequencies) between *CSD1* and 5q microsatellite loci were consistently higher than those generated by MARK III. These differences are attributed primarily to the manner in which each computer program utilizes the genotype data and familial relationships, as well as the inclusion or omission of information such as allele frequencies and penetrance. Regardless of the differences in the lod scores obtained by either computer method of analyses, linkage between *CSD1* and a series of 5q microsatellite loci confirms the mapping of *CSD1* to 5q22-q33.3 in our kindred.

5.2. Haplotype Analysis

The five-locus core (*D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414*) spans approximately 2 cM and was inherited intact in both affected and unaffected family members. Recombination was only evident among the markers proximal and distal to the five-locus core. The haplotype observed in to be segregating with *CSD1*D* in the kindred

described herein differs from the haplotypes presented by Stone *et al* (1994) and Gregory *et al* (1995), suggesting the responsible mutation(s) of lattice corneal dystrophy type 1 in the CSD gene(s) must have arisen independently on different chromosomal backgrounds.

All but two of 34 affected family members inherited the identical haplotype at the five-locus core. In the two exceptional individuals, D014 and his son E036, an alternate allele at *D5S414* was consistently observed to be segregating with *CSD1*D* (figures 15 and 22). The template DNA used to genotype family members was extracted from whole blood and not from lymphoblastoid cell lines which have been shown to contribute to somatic mutations at repetitive regions of microsatellite markers (Weber and Wong, 1993; Banchs *et al*, 1994). There are two possible explanations for this finding. First, a recombination event between *CSD1:D5S414* may have occurred in previous generations not available for study and inherited by the succeeding generations under study. Secondly, a possible mutation event may have occurred on the *CSD1*D* chromosome preceding generation IV and subsequently has been inherited by individual, D014 and his son, E036 (figure 4). This cannot be confirmed because both individuals C003, the affected parent of D014 and his affected sibling, C001 are deceased. Furthermore, D014 has no other available affected sibling(s) who could be genotyped to yield further information about the origin of the alternate allele. The latter explanation seems more likely, as the alternate allele does not segregate with *CSD1*D* in the affected son, D003 and affected grandchildren, E001 and E002 of C001.

5.3. Genome Maps

In surveying the published literature and the electronically accessible databases a variety of maps are available spanning the candidate region 5q22-5q33.3 (table 20). Specifically, the maps obtained through GDB were examined. All of the 10 microsatellite loci studied have been localized cytogenetically on the HUGO Cytogenetic Map and Public Chromosome 5 Cytogenetic Map. Although, this information does not give any indication of the potential ordering of the microsatellite markers, it confirms they are present on chromosome 5.

One linkage map positioning all 10 microsatellite loci has not been established. The best available genetic map is from Genethon which has mapped markers *D5S816*, *D5S393*, *D5S399*, *D5S396*, *D5S414* and *D5S436*. The first five markers are those specific to the five-locus core, the region where *CSD1* has been localized. The sex-averaged length of chromosome 5 is 197.6 cM, which corresponds to approximately 194 Mb of DNA (Dib *et al*, 1996). From the top of the chromosome 5 linkage group, *D5S816* has been mapped between 148-153 cM. *D5S393* has been mapped at 150 cM. The three markers *D5S399*, *D5S396* and *D5S414* form a cluster of unordered markers at 152 cM. The locus *D5S436* is positioned 7 cM distal to this cluster at 159 cM from the top of the chromosome 5 linkage group (figure 23).

Both *D5S396* and *D5S399* have been localized on a radiation hybrid map at 425 and 430 centiRay (cR) respectively. Presently, the remaining markers genotyped in this study are being mapped to a radiation hybrid panel by Dr. K. Morgan (personal communication, Centre for Human Genetics, McGill University, Montreal).

The STS-based maps of the human genome demonstrate markers *D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414* to be all singly-linked to YACs representing the contiguous

Table 20: Summary of electronically available maps for chromosome 5 containing the 10 microsatellite loci studied

Genome Maps	Locus									
	D5S818	IL9	D5S816	D5S393	D5S399	D5S396	D5S414	D5S436	D5S178	D5S2508
*HUGO5 Cytogenetic Map	X	X	X	X	X	X	X	X	X	X
*Public Chromosome 5 Cytogenetic Map	X	X	X	X	X	X	X	X	X	X
*CEPH/Genethon Chromosome 5 Linkage Map				X	X	X		X		
*Genethon Chromosome 5 (March 1996)			X	X	X	X	X	X		
*Whitehead DR9 (Dec 1995) Chromosome 5					X	X				
†Singly-linked Contig WC5.10			X	X	X	X	X			
†Doubly-linked Contig WC-1131			X	X	X	X				

*Genome Database (GDB) available at http://gdbwww.gdb.org

†An STS-Based Map of the Human Genome available at http://www-genome.wi.mit.edu/

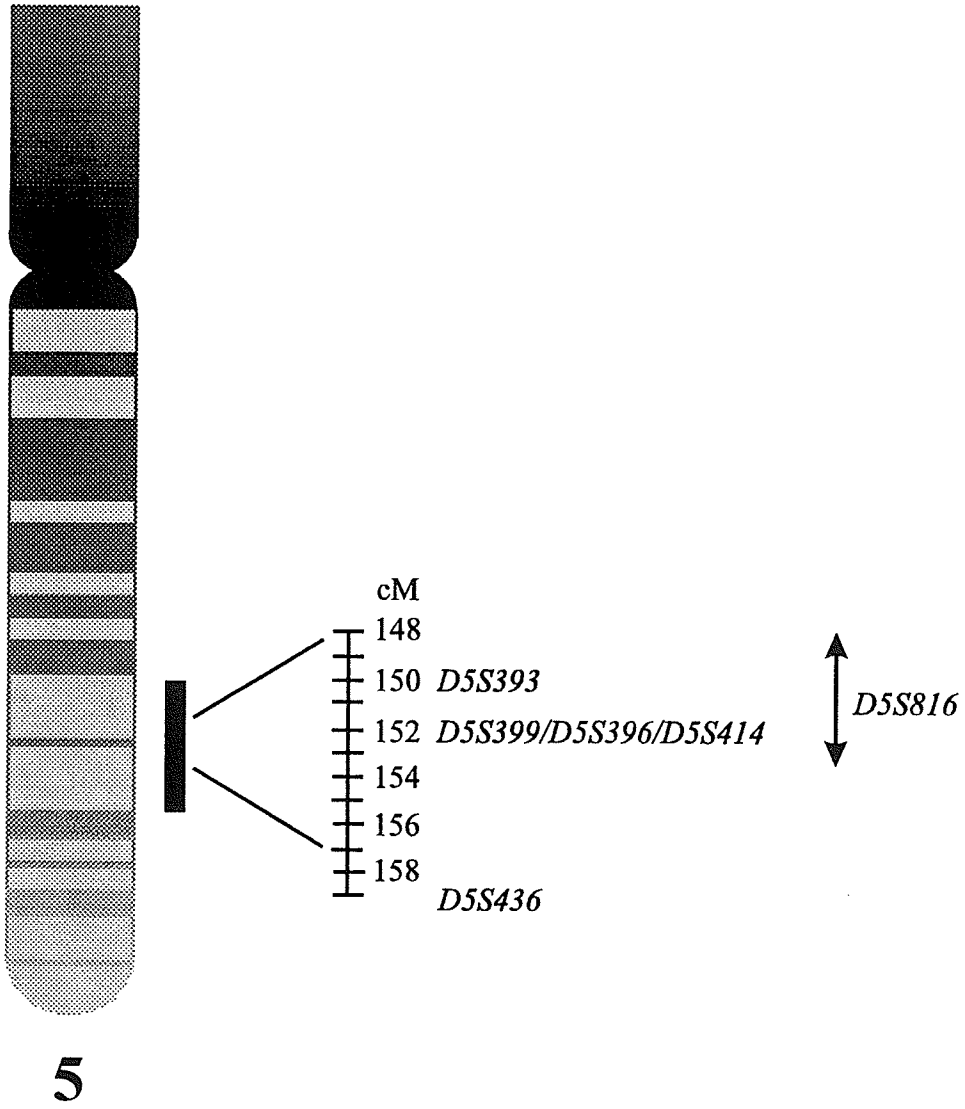


Figure 23: Ideogram of chromosome 5 illustrating the proposed ordering of microsatellite markers in the five-locus core on the Genethon Chromosome 5 linkage map.

segment, WC5.10. All of these markers excluding *D5S414* are doubly-linked to contig WC-1131. Two genomic markers are defined as "singly-linked" if they are both positive on the same YAC and "doubly-linked" if they are positive on two different YACs. Due to the chimeric nature of YACs, it is preferable to have doubly-linked loci when constructing a contig map. Fairman *et al* (1995) constructed a YAC contig spanning from *IL9* to *EGR1*. The loci IL9, D5S393, D5S399, D5S396, and D5S414 have been physically mapped to two overlapping YACs. The YAC 745f8 containing a 1300 Kb insert tested positive for *IL9*, *D5S393* and *D5S399*. The second YAC, 737b9 containing a 1100 Kb insert tested positive for *D5S399*, *D5S396*, *D5S414* and *EGR1*. The combined distance covered by these two overlapping YACs is estimated to be <2.4 Mb with a proposed order of: *IL9-D5S393-D5S399-D5S396-D5S414*.

5.4. Ordering of Microsatellite Loci

The series of 5q microsatellite markers used in this study could not be definitively ordered from the available resources. From the haplotype data, it was possible to subdivide the markers into three subgroups, where the order between the centromere and telomere is as follows: group 1 (*IL9*, *D5S818*), group 2, the five-locus core (*D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414*) and group 3 (*D5S436*, *D5S178* and *D5S2508*). The best possible ordering of loci obtainable from the PCMAP83 (Sherman *et al*, 1984) analysis was as follows: *IL9-D5S818-CSD1-D5S2508-D5S178-D5S436*. As of yet, this order cannot be confirmed on any available physical or genetic map. The candidate interval for *CSD1* is limited to the five-locus core. Therefore, the ordering of the flanking markers in group 1 and group 3 does not provide any further information. Since, there was

no apparent recombination between the five-locus core markers and between *CSD1* and the core markers, these loci could not be ordered using PCMAP83. Genethon presents the best linkage map of the region of interest, however the distal boundary has not been defined because *D5S399*, *D5S396* and *D5S414* remain unordered (figure 23).

5.5. Genotyping Errors: non-amplifying alleles

A non-amplifying allele in one branch of the CSD1 kindred was observed to be segregating at *D5S414*. We hypothesized that sequence variation within the template DNA at the PCR priming site led to the failure of primer annealing and gave rise to a non-amplifying allele. In fact, molecular characterization identified a two base pair change immediately following the (CA)n repetitive tract and adjacent to the 3'-end of reverse primer B. This created a six base pair repetitive region of (CG) and three potential annealing sites for primer B. Only one of these sites allows for the full length primer B to anneal. Thus, under standard PCR conditions, primer B and template annealing was very inefficient. Under less stringent PCR conditions, this sequence variation was tolerated and permited significantly more amplification.

The problem of non-amplification was overcome by designing an alternate reverse primer which allowed the identical $(CA)_n$ repeat at *D5S414* to be amplified. Mendelian co-dominant inheritance was now observed confirming the presence of a non-amplifying allele when using primer B. When a reverse primer extended to incorporate the TA→CG change at the 3'-end was used to amplify *D5S414*, the predicted allele-specific products were generated only in those individuals carrying the two base pair change (personal

communication, Dr. A. Kanis, Department of Pediatrics, University of Iowa, Iowa City, Iowa).

Non-amplifying alleles attributed to mutations within the primer annealing sites, have been previously observed when microsatellite markers and linkage maps were generated. Such markers have generally been excluded from linkage maps (Weissenbach *et al*, 1992; Gyapay *et al*, 1994) and careful attention to primer design has been emphasized (Weber *et al*, 1991; Callen *et al*, 1993; Phillips *et al*, 1993). There is little evidence for any relationship between the sequence or position of primers flanking the repetitive tract and the occurrence of alleles refractory to PCR amplification (Callen *et al*, 1993). Non-amplifying alleles have also been reported in diagnostic testing for Huntington disease (de Rooij *et al*, 1993; Cross *et al*, 1994), cystic fibrosis (Fujimura *et al*, 1990) and Duchenne muscular dystrophy (Clark *et al*, 1990; Oudet *et al*, 1991). However, the molecular basis for such nonamplification has not been described. Although the spontaneous mutation rate for microsatellites is cited to be low in the range of 10^{-3} to 10^{-5} (Weissenbach, 1993a), nonamplification of alleles due to inherited sequence variation in the template DNA must occur more commonly than has been cited in the literature to date. Failure to recognize the segregation of a non-amplifying allele especially within pedigree data, can result in the omission of or misinterpretation of linkage data.

5.6. Implication of Four Corneal Dystrophies Mapping to 5q

To date, linkage of *CSD1* to 5q22-q33.3 has been described in a total of four kindreds (Stone *et al*, 1994, 2 kindreds; Gregory *et al*, 1995, 1 kindred and our Manitoba

kindred). Although CSD2, CSD3 and Reis-Bücklers corneal dystrophy are clinically and histologically distinct from CSD1, they have been mapped to the same interval (Eiberg *et al*, 1993; 1994; Stone *et al*, 1994; Small *et al*, 1996). As described in section 1.2, the corneal dystrophies are inherited in an autosomal dominant manner and pathogenesis is limited to the cornea. However, the composition and pattern of deposits differ among the four dystrophies. Furthermore, CSD1, CSD2 and CSD3 are traditionally classified as stromal dystrophies where as Reis-Bücklers is primarily a dystrophy of Bowman layer (Arffa, 1991).

The mapping of four distinct corneal dystrophies to 5q22-q33.3 suggests the importance of this region in corneal function. The clinical and histological differences may represent a family of closely linked genes or arise from the expression of different mutations in a single gene (Folberg *et al*, 1994; Stone *et al*, 1994; Gregory *et al*, 1995; Small *et al*, 1996). The latter explanation seems more likely based on the following observations. Histological studies of corneal specimens from patients with CSD3 show signs of both characteristic granular-type deposits of CSD2 and fusiform amyloid deposits of CSD1. In patients with typical CSD2, amyloid deposits have been described adjacent to and separate from the granular deposits. Conversely, in families with typical CSD1, histological findings of granular deposits have been noted. There is evidence demonstrating that the granular deposits and amyloid deposits may in fact not be separated spacially in the cornea but juxtaposed (Brownstein *et al*, 1974; Folberg *et al*, 1994). From observations, Brownstein *et al* (1974) speculates that a morphological sequence of events begins with the accumulation of macromolecules gradually forming the "homogenous"

appearing granules. Furthermore, granular deposits surrounded by amyloid suggests these may be the early stages of the amyloid deposits, thus, demonstrative of a biochemical relationship between CSD1 and CSD2 (Brownstein *et al,* 1974). Although the specific nature and location of the primary abnormality of Reis-Bücklers dystrophy is unclear, it has been reported in association with CSD1-type deposits (Malbran *et al,* 1988) as well as CSD2-type deposits (Moller, 1989). Moller (1989) suggests that Reis-Bücklers dystrophy may not be distinctive but represent a variant in the wide clinical spectrum of granular corneal dystrophy.

It is now apparent that at least two genes are responsible for the pathogenesis of lattice-type corneal dystrophies. Lattice corneal dystrophy type 2 has many similarities to lattice type 1. However, lattice corneal dystrophy type 2 has been shown to be caused by mutations in the gelsolin gene on chromosome 9. Although granular, Avellino and Reis-Bücklers are clinically dissimilar, they seem to be caused by mutations in the same gene giving rise to CSD1. A similar situation exists for retinitis pigmentosa and pattern dystrophy of the macula. Clinically, indistinguishable retinitis pigmentosa can be caused by mutations in various genes, including rhodopsin and RDS/peripherin (summary available from OMIM). However, mutations in RDS/peripherin gene create allelic variants capable of causing retinitis pigmentosa or retinal disorders such as pattern dystrophy, fundus flavimaculatus and cone-rod dystrophy (Nichols *et al,* 1993; Weleber *et al,* 1993; Folberg *et al,* 1994).

If these four clinically distinct dystrophies are indeed the consequence of mutations in the same gene, this is not unusual. There are numerous examples illustrating the diverse

clinical presentation of human disease arising from alterations within a single gene. For example, allelic variants of fibroblast growth factor receptor 3 (FGFR3) give rise to achondroplasia (Shiang *et al*, 1994), thanatophoric dysplasia type I and II (Tavormina *et al*, 1995) as well as Crouzon craniosynostosis syndrome with acanthosis nigricans (Meyers *et al*, 1995). Similarly allelic variants of fibroblast growth factor receptor 2 (FGFR2) account for four clinically distinct syndromes: Crouzon (Reardon *et al*, 1994), Jackson-Weiss (Jabs *et al*, 1994), Apert (Wilkie *et al*, 1995) and Pfeiffer (LaJeunie *et al*, 1995). Another example of allelic diversity involves the locus encoding the peripheral myelin protein 22 (PMP22), where a duplication of the gene is the basis for Charcot-Marie-Tooth type 1a (CMT1A) (Lupski *et al*, 1991; Patel *et al*, 1992). Other allelic variants of *PMP22* are responsible for Dejerine-Sottas syndrome, which is clinically similar to, but more severe than, CMT1A (Roa *et al*, 1993), whereas a deletion of the PMP22 gene has been identified in hereditary neuropathy with liability to pressure palsies (HNPP) (Chance *et al*, 1993).

The phenotypic variation observed in other human amyloidoses due to different mutations within a single gene or due to the effect of other genes on the expression of a single mutation, may in part explain the diversity of phenotypes in CSD1, CSD2, CSD3 and Reis-Bücklers corneal dystrophy. For example, the most common type of hereditary amyloidosis reported is a methione to valine substitution at position 30 in the transthyretin protein. This particular mutation is associated with variable age of onset with a spectrum of clinical disease among different ethnic populations, thus, implying the involvement of other modulating factors (Benson, 1995). Three prion diseases, familial Creutzfeldt-Jakob

disease (CJD) (Hsiao *et al*, 1991) Gerstmann-Sträussler-Scheinker syndrome (GSS) (Hsiao *et al*, 1992) and fatal familial insomnia (Medori *et al*, 1992) are all caused by dominantly inherited mutations of the prion protein gene (*PRNP*) encoded in the genome (Prusiner, 1991). Specifically, an aspartic acid to asparagine substitution at position 178 of the *PRNP*, results in fatal familial insomnia when the amino acid at position 129 is methionine and CJD when the amino acid at position 129 is valine (Goldfarb *et al*, 1992). Mutations in the amyloid precursor protein gene (*APP*) have been found to be associated with both hereditary cerebral hemorrhage and Alzheimer disease (Benson, 1995). In patients with Alzheimer disease, Tagliavini *et al* (1991) noted that both amyloid and non-amyloid deposits of APP were present. This type of co-existence is analogous to lattice-type deposits and granular type deposits found in Avellino corneal dystrophy (Stone *et al*, 1994). None of the 14 human gene products known to be involved in amyloid deposition map to chromosome 5, thus confirming that the previously characterized human amyloidoses are not allelic with the corneal dystrophies described herein.

5.7. Future Directions

Linkage analysis represents the preliminary steps of a positional cloning approach towards identifying the CSD1 gene. This task has been successfully completed by this and previous studies (Stone *et al*, 1994; Gregory *et al*, 1995). It is apparent from our linkage study there no apparent recombination between *CSD1* and the markers of the five-locus core. On this basis, screening other microsatellite markers from the candidate region will unlikely reveal any key recombinants, however, they will confirm the haplotype

segregating with *CSD1*D* and *CSD1*N* to be inherited intact. Furthermore, the presence of other microsatellite markers in addition to those screened will aid in confirming the presence of the loci in YACs of interest and raise the level of confidence to assemble overlapping YACs linked by more than two microsatellite markers. Similarly, the presence or absence of the most tightly linked microsatellite marker loci can be determined to isolate the chromosomal fragment containing the candidate interval in a radiation hybrid panel.

The CSD1 gene has been localized to a 2 cM candidate interval on the Genethon linkage map and physically by two overlapping YACs. Presently, the 5q microsatellite markers screened in our CSD1 kindred and others are being ordered on a radiation hybrid map by Dr. K. Morgan (personal communication, Centre for Human Genetics, McGill University, Montreal). With these resources, further refinement of the candidate interval will be possible. From here, it will be necessary to utilize various methods to identify encoded gene sequences, determine their significance and screen them for mutations associated with CSD1 as well as demonstrate an autosomal dominant pattern of inheritance. In the final steps of identifying the disease-causing gene, numerous encoded gene sequences are isolated from the candidate region and each must be systematically studied until the correct one is found. Throughout this process, the literature and the electronically accessible databases will continually be surveyed to identify any potential candidate genes which may help to accelerate the search for the CSD1 gene(s). In addition, further linkage studies in kindreds with CSD2, CSD3 and Reis-Bücklers corneal dystrophy will help to determine, if they too map to the same 2 cM interval defined for

CSD1. The findings from such studies will elucidate if a family of closely linked genes is responsible for the four autosomal dominant corneal dystrophies or if in fact they are allelic disorders with different mutations in a single gene giving rise to the phenotypic variation.

The purposes for identifying the gene(s) responsible for CSD1 as well as the other corneal dystrophies are two-fold. First, cloning and identifying disease-causing mutations permits presymptomatic diagnosis in at-risk families. Information of this type is important to provide accurate genetic counseling as well as help to alleviate concerns regarding adult onset diseases, such as CSD1, and to consider carefully lifestyle decisions and career planning. Such information also presents new dilemmas for physicians and at-risk family members regarding the appropriate age to offer preclinical testing and the potential impact the knowledge of test results on those who opt to be tested.

Prior to cloning the actual gene, results from haplotype analysis of tightly linked markers within a family can predict those among at-risk individuals who are likely to become affected. In the CSD1 kindred under study, at-risk families have been contacted regarding the availability of presymptomatic testing. Five of the fifteen at-risk individuals available for genotyping were found to have inherited the haplotype segrgating with *CSD1*D*. These results will only be available upon request by family members.

Second, cloning the gene will help towards understanding the pathogenesis of amyloid and non-amyloid type corneal dystrophies. The biochemical and functional knowledge about the gene product and its role in localized amyloidosis will provide insight into the relationship between CSD1, CSD2, CSD3 and Reis-Bucklers as well as other

diseases characterized by localized and systemic amyloidosis. In the short term, it will help to develop non-invasive treatment strategies which can arrest the progression of disease and break-down the accumulated deposits. This type of treatment is particularly important because corneal dystrophies have been well documented to reoccur in corneal grafts. In the long term, the cornea and pathogenesis of amyloid deposition serve as a potential model system for developing therapies for systemic amyloidosis where amyloid deposits cannot be easily monitored.

6. Summary

1. A large multigenerational kindred, one of four studied since 1994, formed the basis of this linkage study which confirmed the mapping of *CSD1* to 5q22-5q33.3. The candidate region harbouring the CSD1 gene was estimated to be approximately 2 cM in length between *D5S393* proximally and *D5S399/D5S396/D5S414* distally. The five-locus core (*D5S816*, *D5S393*, *D5S399*, *D5S396* and *D5S414*) haplotype segregated with *CSD1*D* in 32 of 34 affected family members genotyped.
2. In screening a panel of 5q microsatellite markers, a non-amplifying allele was observed to be segregating at *D5S414*. Molecular characterization revealed a template sequence variation outside, yet in close proximity to the primer annealing site. The problem of non-amplifying alleles was easily solved by designing a new primer pair to amplify the region of interest. The occurrence of non-amplifying alleles is an important consideration when screening microsatellite markers as non-amplifying alleles can significantly affect genotype data and the outcome of linkage analyses.

7. References

Arffa RC (1991): Dystrophies of the epithelium, Bowman's layer, and stroma. In Grayson's Disease of the Cornea. 3rd ed. Mosby Year Book. St. Louis. Chapter 17, pp 364-416.

Ashley MV, Dow BD (1994): The use of microsatellite analysis in population biology: background, methods and potential applications. EXS 69:185-201

Banchs I, Bosch A, Guimera J, Lazaro C, Puig A, Estivill X (1994): New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines. Hum Mutat 3:365-372

Benson MD (1995): Amyloidosis. In The Metabolic and Molecular bases of Inherited Disease. 7th ed. Eds. Scriver CR, Beaudet AL, Sly WS, Valle D. McGraw-Hill Inc. New York, Chapter 139, pp 4159-4191.

Benson MD, Liepnieks J, Uemichi T, Wheeler G, Correa R (1993): Hereditary renal amyloidosis associated with a mutant fibrinogen alpha-chain. Nat Genet 3:252-255.

Biber H (1890): Uber einigs seltehere hornhauterkrankungen: Die oberflachliche gittrige keratitis. Diss Zurich pp 35-42.

Bishop PN, Bonshek RE, Jones CJ, Ridgway AE, Stoddart RW (1991): Lectin binding sites in normal, scarred, and lattice dystrophy corneas. Br J Ophthalmol 75:22-27

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

Brownstein S, Fine BS, Sherman ME, Zimmerman LE (1974): Granular dystrophy of the cornea: light and eletron microscopic confirmation of recurrence in a graft. Am J Ophthalmol 77:701-710.

Burright EN, Clark HB, Servadio A, Matilla T, Feddersen RM, Yunis WS, Duvick LA, Zoghbi HY, Orr HT (1995): SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. Cell 82:937-948.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR (1993): Incidence and origin of "null" alleles in the (AC)n microsatellite markers. Am J Hum Genet 52:922-927.

Campuzano V, Montermini L, Molto MD, Pianese L, Cossee M, Cavalcanti F, Monros E, Rodius F, Duclos F, Monticelli A, Zara F, Canizares J, Koutnikova H, Bidichandani SI, Gellera C, Brice A, Trouillas P, De Michele G, Filla A, De Frutos R, Palau F, Patel PI, Di Donato S, Mandel JL, Pandolfo M (1996): Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 271:1423-1427.

Castano EM, Frangione B (1995): Non-Alzheimer's disease amyloidoses of the nervous system. Curr Opin Neurol 8:279-285.

Chance PF, Alderson MK, Leppig KA, Lensch MW, Matsunami N, Smith B, Swanson PD, Odelberg SJ, Disteche CM, Bird TD (1993): DNA deletion associated with hereditary neuropathy with liability to pressure palsies. Cell 72: 143-151 1993.

Clark S, Yau SC, Abbs S, Roberts RG, Bentley DR, Mathew CG, Bobrow M (1990): Experience of the polymerase chain reaction in the diagnosis of Duchenne and Becker Muscular Dystrophy: Benefits and Problems Encountered. J Med Genet 27:656A.

Cohen AS, Calkin E (1959): Electron microscopic observation on a fibrous component in amyloid of diverse origins. Nature 183-1202.

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

Collins FS (1995): Positional cloning moves from perditional to traditional. Nat Genet 9:347-350

Conneally PM, Rivas ML (1980): Linkage analysis in man. In Advances in Human Genetics. Volume 10. Eds. Harris H, Hirschhorn K. Plenum Press. New York. Chapter 3, pp. 209-266.

Cooper DN (1994): Mapping the human genome. Ann Genet 37:101-116

Costa PP, Figuera AS, Bravo FR (1978): Amyloid fibril protein related to prealbumin in familial amyloidotic polyneuropathy. Proc Natl Acad Sci USA 75:4499.

Côté GB (1975): Centromeric linkage in man. PhD thesis, University of Birmingham, Birmingham, England.

Cottingham RWJ, Idury RM, Schaffer AA (1993): Faster sequential genetic linkage computations. Am J Hum Genet 53:252-263.

Cross G, Pitt T, Sharif A, Bates G, Lehrach H (1994): False-negative result for Huntington's disease mutation. Lancet 343:1232

Daniels RJ, Campbell L, Rodrigues NR, Francis MJ, Morrison KE, McLean M, MacKenzie A, Ignatius J, Dubowitz V, Davies KE (1995):Genomic rearrangements in childhood spinal muscular atrophy: linkage disequilibrium with a null allele. J Med Genet 32:93-96

de Rooij KE, de Koning Gans PA, Losekoot M, Bakker E, den Dunnen JT, Vegter van der Vlis M, Roos RA, van Ommen GJ (1993): Borderline repeat expansion in Huntington's disease. Lancet 342:1491-1492.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996): A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152-154.

Dimmer F (1899): Ube oberflachiche gittrige gornhauttr ubung. Z Augenhei LK. 2:354-361.

Dolmetsch AM, Stockl FA, Folberg R, Gensini I, Burnier MN (1996): Combined granular-lattice corneal dystrophy (Avellino) in a patient with no known Italian ancestry. Can J Ophthalmol 31:29-31.

Donis-Keller HP, Green C, Helms S, Cartinhour B, Weiffenbach B, Stephens K, Keith TP, Bowden DW, Smith DR, Lander ES, Botstein D, Akots G, Rediker KS, Gravius T, Brown VA, Rising MB, Parker C, Powers JA, Watt DE, Kauffman ER, Bricker A, Phipps P, Muller-Kahle H, Fulton TR, Ng S, Schumm JW, Braman JC, Knowlton RG, Barker DF, Crooks SM, Lincoln SE, Daly MJ, Abrahamson (1987): A genetic linkage map of the human genome. Cell 51:319-337.

Dryja TP, McGee TL, Reichel E, Hahn LB, Cowley GS, Yandell DW, Sandberg MA, Berson EL (1990): A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. Nature 343:364-366.

Duyao MP, Auerbach AB, Ryan A, Persichetti F, Barnes GT, McNeil SM, Ge P, Vonsattel JP, Gusella JF, Joyner AL, MacDonald ME (1995): Inactivation of the mouse Huntington's disease gene homolog Hdh. Science 269:407-410.

Eiberg H, Moller HU, Berendt I, Mohr J (1993): Assignment of granular corneal dystrophy Groenouw type I (CDGG1) to chromosome 5q: close linkage to IL9 and D5S210. (Abstract) Human Genome Mapping Workshop 93.

Eiberg H, Moller HU, Berendt I, Mohr J (1994): Assignment of granular corneal dystrophy Groenouw type I (CDGG1) to chromosome 5q. Eur J Hum Genet 2:132-138.

Fairman J, Chumakov I, Chinault AC, Nowell PC, Nagarajan L (1995): Physical mapping of the minimal region of loss in 5q- chromosome. Proc Natl Acad Sci U S A 92:7406-7410.

Folberg R, Alfonso E, Croxatto JO, Driezen NG, Panjwani N, Laibson PR, Boruchoff SA, Baum J, Malbran ES, Fernandez Meijide R, Morrison JA, Bernardino VB, Arbizo VV, Albert DM (1988): Clinically atypical granular corneal dystrophy with pathologic features of lattice-like amyloid deposits. A study of these families. Ophthalmology 95:46-51.

Folberg R, Stone EM, Sheffield VC, Mathers WD (1994): The relationship between granular, lattice type 1, and Avellino corneal dystrophies. A histopathologic study. Arch Ophthalmol 112:1080-1085.

Freeman HJ (1992): Topography of lectin binding sites in celiac spruce. Can J Gastroenterol 6:271-276.

Friedrich N, Kekule A (1859): Zur amyloidfrage. Virch Arch Path Anaf Physiol. 16:50-65.

Friend SH, Bernards R, Rogelj S, Weinberg RA, Rapaport JM, Albert DM, Dryja TP (1986): A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323:643-646.

Fuchs A (1925): Zur kenntnis der gitterigen hornhautentartung: Haab-Dimmer. Ztschr f Augenh 57:159-186.

Fujimura FK, Northrup H, Beaudet AL, O'Brien WE (1990): Genotyping errors with the polymerase chain reaction. N Engl J Med 322:61.

Gastier JM, Pulido JC, Sunden S, Brody T, Buetow KH, Murray JC, Weber JL, Hudson TJ, Sheffield VC, Duyk GM (1995): Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. Hum Mol Genet 4:1829-1836.

Gene Runner (1994) version 3.0 Hastings Software Inc.

Gilbert J (1994): Strategies for genotype generation. In Current Protocols in Human Genetics Eds. Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR. Current Protocols USA. Chapter 1, pp 1.31.-1.3.10.

Glenner. GG, Terry W, Harada M, Isersky C, Page D (1971): Amyloid fibril proteins: Proof homology with immunoglobulin light chains by sequence analysis. Science 172:1150.

Goate A, Chartier Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, Mant R, Newton P, Rooke K, Roques P, Talbot C, Pericak-Vance M, Roses A, Williamson R, Rossor M, Owen M, Hardy J (1991): Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature 349:704-706.

Goldfarb LG, Peterse RB, Tabaton M, Brown P, LeBlanc AC, Montagna P, Cortelli P, Julien J, Vital C, Pendelbury WW, Haltia M, Wills PR, Hauw JJ, McKeever PE, Monari L, Schrank B, Swergold GD, Gambetti-Autilio L, Gajdusek CD, Lugaresi E, Gambetti P (1992): Fatal familial insomnia and familial Creutzfeldt-Jakob disease: disease phenotype determined by a DNA polymorphism. Science 258:806-808.

Goodfellow PN (1992): Human genome project. Variation is now the theme. Nature 359:777-778.

Gorevic PD, Rodrigues MM, Krachmer JH, Green C, Fujihara S, Glenner GG (1984): Lack of evidence for protein AA reactivity in amyloid deposits of lattice corneal dystrophy and amyloid corneal degeneration. Am J Ophthalmol 98:216-224 .

Green ED, Cox DR, Myers RM (1995): The Human Genome Project and its impact on the study of human disease. In The Metabolic and Molecular bases of Inherited Disease. 7th ed. Eds. Scriver CR, Beaudet AL, Sly WS, Valle D. McGraw-Hill Inc. New York, Chapter 6, pp 401-436.

Green ED, Olson MV (1990): Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: A model for human genome mapping. Science 250:94.

Gregory CY, Evans K, Bhattacharya SS (1995): Genetic refinement of the chromosome 5q lattice corneal dystrophy type I locus to within a 2 cM interval. J Med Genet 32:224-226.

Groenouw A (1890): Knoetchenfoermige hornhauttruebungen (Noduli corneae). Arch Augenheilk 21:281-289.

Groenouw A (1898): Knoetchenfoermigen hornhauttruebungen. Graefe Arch Ophthal 46:85-102.

Groenouw A (1917): Knoetchenfoermige hornhauttruebungen, vererbt durch drei generationen. Klin Mbl Augenheilk 58:411-420.

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994): The 1993-94 Genethon human genetic linkage map. Nat Genet 7:246-339.

Haab O (1899): Die gittrige keratitis. Zie Augenheilk 2:235-246.

Haltia M, Prelli F, Ghiso J, Kiuru S, Somer H, Palo J, Frangione B (1990): Amyloid protein in familial amyloidosis (Finnish type) is homologous to gelsolin, an actin-binding protein. Biochem Biophys Res Commun 167:927-932.

Hammerstein W, Scholz W (1975): Klinik und genetik der gittrigen hornhautdystrophie. Graefes Arch Ophthal 193: 109-120.

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, Housman DE, Shaw DJ (1992): Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355:545-546.

Hauge XY, Litt M (1993): A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet 2:411-415

Hearne CM, Ghosh S, Todd JA (1992): Microsatellites for linkage analysis of genetic traits. Trends Genet 8:288-294.

Hida T, Proia AD, Kigasawa K, Sanfilippo FP, Burchette JL, Jr., Akiya S, Klintworth GK (1987): Histopathologic and immunochemical features of lattice corneal dystrophy type III. Am J Ophthalmol 104:249-254.

Hogan MJ, Alvardo JA, Weddell JE (1971). The cornea. In: Histology of the Human Eye. W. B. Saunders Company. Philadelphia. Chapter 3, pp 55-111.

Holland EJ, Daya SM, Stone EM, Folberg R, Dobler AA, Cameron JD, Doughman DJ (1992): Avellino corneal dystrophy. Clinical manifestations and natural history. Ophthalmology 99:1564-1568.

Hollwich F (1985): The cornea. In: Ophthalmology: a short textbook. Georg Thierne Verlag. Stuttgart. Chapter 5, pp 74-99.

Hsiao K, Dlouhy SR, Farlow MR, Cass C, Da Costa M, Conneally PM, Hodes ME, Ghetti B, Prusiner SB (1992): Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with neurofibrillary tangles. Nat Genet 1:68-71.

Hsiao K, Meiner Z, Kahana E, Cass C, Kahana I, Avrahami D, Scarlato G, Abramsky O, Prusiner SB, Gabizon R (1991): Mutation of the prion protein in Libyan Jews with Creutzfeldt-Jakob disease. N Engl J Med 324:1091-1097.

Hubert R, Weber JL, Schmitt K, Zhang L, Arnheim N (1992): A new source of polymorphic DNA markers for sperm typing: analysis of microsatellite repeats in single cells. Am J Hum Genet 51:985-991.

Hudson TJ (1994): PCR methods of genotyping. In Current Protocols in Human Genetics. Eds. Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR. Current Protocols. U.S.A. Chapter 2. pp 2.5.5-2.510.

The Huntington Disease Collaborative Research Group (1993): A novel gene containing a trinucleotide repeat that is expanded and unstable in Huntington's disease chromosomes. Cell 72:971-983.

An International System for Human Cytogenetic Nomenclature (1985) ISCN 1985. Report of the Standing Committee on Human Cytogenetic Nomenclature. Birth Defects 21:1-117.

Jabs EW, Li X, Scott AF, Meyers G, Chen W, Eccles M, Mao J, Charnas LR, Jackson CE, Jaye M (1994): Jackson-Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor 2. Nature Genet 8:275-279.

Jennings C (1995): How trinucleotide repeats may function. Nature 378:127.

Jolly DJ, Esty AC, Bernard HU, Freidman T (1982): Isolation of a genomic clone partially encoding human hypoxanthine phophoribosyltransferase. Proc Natl Acad Sci USA. 79:5038.

Kaneshige T, Takagi K, Nakamura S, Hirasawa T, Sada M, Uchida K (1992): Genetic analysis using fingernail DNA. Nucleic Acids Res 20:5489-5490.

Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiguchi I, Kimura J, Narumiya S, Kakizuka A (1994): CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. Nat Genet 8:221-228.

Keats BJ, Conneally PM, Lalouel JM (1987): Report of the committee on linkage data and gene order. Cytogenet Cell Genet 46:339-343.

Khoury MJ, Beaty TH, Cohen BH (1993): Genetic Approaches to Familial Aggregation. In Fundamentals of Genetic Epidemiology. Oxford University Press. Oxford. Chapter 9, pp 285-311.

Klintworth GK (1967): Lattice corneal dystrophy. An inherited variety of amyloidosis restricted to the cornea. Am J Path 50:371-399.

Klintworth GK (1971): Current concepts in the ultrastructural pathogenesis of macular and lattice corneal dystrophies. Birth Defects: Original Article Series (VII) 3:27-31.

Krachmer JH, Palay DA (1991): Corneal disease. N Engl J Med 325:1804-1806

Kwok SCM, Ledley FD, DiLella AG, Robson KJH, Woo SLC (1985): Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase. Biochemistry 24:556.

Lalley PA, Davisson MT, Graves JA, Obrien SJ, Womack JE, Roderick TH, Creau-Goldberg N, Hillyard AL, Doolittle DP, Rogers JA (1989): Report of the committee on comparative mapping, Human Gene Mapping 10: Tenth international Workshop on Human Gene Mapping. Cytogen Cell Genet 51:503-532.

La Spada AR, Paulson HL, Fischbeck KH (1994): Trinucleotide repeat expansion in neurological disease. Ann Neurol 36:814-822.

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH (1991): Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352:77-79.

Lajeunie E, Ma HW, Bonaventure J, Munnich A, Le Merrer M, Renier D (1995): FGFR2 mutations in Pfeiffer syndrome. Nature Genet 9:108.

Lathrop GM, Lalouel JM (1984): Easy calculatlions of lod scores and genetic risks on small computers. Am J Hum Genet 36:460-465.

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

Lathrop GM, Lalouel JM, White RL (1986): Construction of human linkage maps: Likelihood calculations for multilocus linkage analysis. Genet Epidemiol 3:39-52.

Lee B, Godfrey M, Vitale E, Hori H, Mattei MG, Sarfarazi M, Tsipouras P, Ramirez F, Hollister DW (1991): Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. Nature 352:330-334.

Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, Le Paslier D, Frezal J, Cohen D, Weissenbach J, Munnich A, Melki J (1995): Identification and characterization of a spinal muscular atrophy-determining gene. Cell 80:155-165.

Levin M, Franklin EC, Frangione B, Pras M (1972): The amino acid sequence of a major nonimmunoglobulin component of some amyloid fibrils. J Clin Invest 51:2772.

Levin M, Pras M, Franklin EC (1973): Immunologic studies of the major non-immunoglobulin protein of amyloid I. Identification and partial characterization of a related serum component. J Exp Med 138:373.

Levinson G, Gutman GA (1987): Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203-221.

Litt M, Hauge X, Sharma V (1993): Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and cures. Biotechniques 15:280-284.

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

Lupski JR, de Oca Luna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ, Saucedo Cardenas O, Killian JM, Garcia CA, Chkravarti A, Patel PI (1991): DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell 66:219-232.

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

MacLennan DH, Duff C, Zorzato F, Fujii J, Phillips M, Korneluk RG, Frodis W, Britt BA, Worton RG (1990): Ryanodine receptor gene is a candidate for predisposition to malignant hyperthermia. Nature 343:559-561.

Malbran ES, Meijide RF, Croxatto JO (1988): Atypical corneal dystrophy with stromal amyloid deposits. Cornea 7:210-213.

Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, Friend SH (1990): Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-1238.

Marles S (1994): Linkage studies of lattice corneal dystrophy type 1. M.Sc. thesis, University of Manitoba, Manitoba, Canada.

Maury CP (1991): Gelsolin-related amyloidosis. Identification of the amyloid protein in Finnish hereditary amyloidosis as a fragment of variant gelsolin. J Clin Invest 87:1195-1199.

Medori R, Montagna P, Tritschler HJ, LeBlanc A, Cortelli P, Tinuper P, Lugaresi E, Gambetti P (1992): Fatal familial insomnia: a second kindred with mutation of prion protein gene at codon 178. Neurology 42:669-670.

Meretoja J (1969): Familial systemic paramyloidosis with lattice dystrophy of the cornea, progressive cranial neuropathy, skin changes and various internal symptoms. Ann Clin Res 1:314.

Meyers GA, Orlow SJ, Munro IR, Przylepa KA, Jabs EW (1995): Fibroblast growth factor receptor 3 (FGFR3) transmembrane mutation in Crouzon syndrome with acanthosis nigricans. Nature Genet 11:462-464.

Moller HU (1989): Granular corneal dystrophy Groenouw type I (GrI) and Reis-Bucklers' corneal dystrophy (R-B). One entity? Acta Ophthalmol Copenh 67:678-684.

Moller HU (1991): Granular corneal dystrophy Groenouw type I. Clinical and genetic aspects. Acta Ophthalmol Suppl 1-40.

Monaco AP, Neve RL, Colletti Feener C, Bertelson CJ, Kurnit DM, Kunkel LM (1986): Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323:646-650.

Mondino BJ, Rabb MF, Sugar J, Sundar Raj CV, Brown SI (1981): Primary familial amyloidosis of the cornea. Am J Ophthal 92-732-736.

Mondino BJ, Raj CV, Skinner M, Cohen AS, Brown SI (1980): Protein AA and lattice corneal dystrophy. Am J Ophthal 89:377-381.

Morin PA, Moore JJ, Woodruff DS (1992): Identification of chimpanzee subspecies with DNA from hair and allele-specific probes. Proc R Soc Lond B Biol Sci 249:293-297.

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

Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, Tadokoro K, Kondo I, Murayama N, Tanaka Y, Kikushima H, Umino K, Kurosawa H, furukawa T, Nihei K, Inoue T, Sano A, Komure O, Takahashi M, Yoshizawa T, Kanazawa I, Yamada M (1994): Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. Nat Genet 6:14-18.

Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, Zeisler J, Borowski A, Marth JD, Phillips AG, Hayden MR (1995): Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell 81:811-823.

Nelson DL (1996): Allelic expansion underlies many genetic diseases. Growth: Genetics and Hormones 12:1-4.

Nichols BE, Sheffield VC, Vandenburgh K, Drack AV, Kimura AE, Stone EM (1993): Butterfly-shaped pigment dystrophy of the fovea caused by a point mutation in codon 167 of the RDS gene. Nat Genet 3:202-207.

Nichols WC, Gregg RE, Brewer HB, Jr., Benson MD (1990): A mutation in apolipoprotein A-I in the Iowa type of familial amyloidotic polyneuropathy. Genomics 8:318-323.

Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boue J, Bertheas MF, Mandel JL (1991): Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 252:1097-1102.

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

Orkin SH, Nathan DG (1981): The molecular genetics of thalassemia. In Advances in Human Genetics. Volume 11. Eds. Harris H, Hirschhorn K. Plenum Press. New York. pp 233-280.

Orr HT, Chung MY, Banfi S, Kwiatkowski TJ, Jr., Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY (1993): Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat Genet 4:221-226.

Oudet C, Heilig R, Hanauer A, Mandel JL (1991): Nonradioactive assay for new microsatellite polymorphisms at the 5' end of the dystrophin gene, and estimation of intragenic recombination. Am J Hum Genet 49:311-319.

Patel PI, Roa BB, Welcher AA, Schoener Scott R, Trask BJ, Pentao L, Snipes GJ, Garcia CA, Francke U, Shooter EM, Lupski JR, Suter U (1992): The gene for the peripheral myelin protein PMP-22 is a candidate for Charcot-Marie-Tooth disease type 1A. Nat Genet 1:159-165.

Persico MG, Viglietto G, Martini G, Toniolo D, Paonessa G, Moscatelli C, Dono R (1986): Isolation of human glucose-6-phosphate dehydrogenase (G6PD) cDNA clones: Primary structure of the protein and unusual 5' non-coding region. Nucleic Acids Res 14:2511.

Phillips HA, Hyland VJ, Holman K, Callen DF, Richards RI, Mulley JC (1991): Dinucleotide repeat polymorphism at D16S287. Nucleic Acids Res 19:6664.

Phillips HA, Thompson AD, Kozman HM, Sutherland GR, Mulley JC (1993): A microsatellite marker within the duplicated D16S79 locus has a null allele; significance for linkage mapping. Cytogenet Cell Genet 64:131-132.

Prusiner SB (1991): Molecular biology of prion diseases. Science 252:1515.

Ramsay RM (1958): Familial corneal dystrophy of the lattice type. Trans Am Path Soc 55:710-739.

Ramsay RM (1960): Familial corneal dystrophy-lattice type. Trans Can Ophtha Soc 23: 222-229.

Reardon W, Winter RM, Rutland P, Pulleyn LJ, Jones BM, Malcolm S (1994): Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. Nature Genet 8:98-103.

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

Richards RI, Holman K, Yu S, Sutherland GR (1993): Fragile X syndrome unstable element, p(CCG)n, and other simple tandem repeat sequences are binding sites for specific nuclear proteins. Hum Mol Genet 2:1429-1435.

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

Roa BB, Dyck PJ, Marks HG, Chance PF, Lupski JR (1993): Dejerine-Sottas syndrome associated with point mutation in the peripheral myelin protein 22 (PMP22) gene. Nature Genet 5:269-273.

Rommens JM, Iannuzzi MC, Kerem NS, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989): Identification of the cystic fibrosis: chromosome walking and jumping. Science 245:1059-1065.

Rose EA (1991): Applications of the polymerase chain reaction to genome analysis. FASEB J 5:46-54.

Rosenwasser GO, Sucheski BM, Rosa N, Pastena B, Sebastiani A, Sassani JW, Perry HD (1993): Phenotypic variation in combined granular-lattice (Avellino) corneal dystrophy. Arch Ophthalmol 111:1546-1552.

Royer Pokora B, Kunkel LM, Monaco AP, Goff SC, Newburger PE, Baehner RL, Cole FS, Curnutte JT, Orkin SH (1986): Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location. Nature 322:32-38.

Sambrook J, Fritsch EF, Maniatis T (1989): Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Harbour Laboratory Press. Cold Spring Harbour. pp 1.25-1.28.

Sander A, Schmelzle R, Murray J (1994): Evidence for a microdeletion in 1q32-41 involving the gene responsible for Van der Woude syndrome. Hum Mol Genet 3:575-578.

Schaffer AA (1996): Faster linkage analysis computations for pedigrees with loops or unused alleles. human Heredity 46:226-235.

Schaffer AA, Gupta SK, Shriram K and Cottingham RWJ (1994): Avoiding recomputation in linkage analysis. Hum Hered 44:225-237.

Schlessinger D (1990): Yeat artificial chromosomes: tools for mapping and anlysis of complex genomes. Trends Genet 6:248-258.

Searle AG, Edwards JH, Hall JG (1994): Mouse homologuess of human heriditary disease. J Med Genet 31:1-19.

Shapiro MB, Senapathy P (1987): RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15:7155-7174.

Sharma JK, Gopalkrishna V, Das BC (1992): A simple method for elimination of unspecific amplifications in polymerase chain reaction. Nucleic Acids Res 20:6117-6118.

Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier JM, Pulido JC, Yandava C, Sunden SL, Mattes G, Businga T, McClain A, Beck J, Scherpier T, Gilliam J, Zhong J, Duyk GM (1995):A collection of tri- and tetranucleotide repeat markers used to generate high quality, high resolution human genome-wide linkage maps. Hum Mol Genet 4:1837-1844.

Sherman SL, King J, Robson EB, Yee S (1984): A Revised Map of Chromosome 1. Ann Hum Genet 48:243-251.

Shiang R, Thompson L.M, Zhu YZ, Church DM, Fielder TJ, Bocian M, Winokur ST, Wasmuth JJ (1994): Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. Cell 78:335-342.

Sipe JD (1992): Amyloidosis. Annu Rev Biochem 61:947-975.

Sipe JD (1994): Amyloidosis. Crit Rev Clin Lab Sci 31:325-354.

Small KW, Mullen L, Barletta J, Graham K, Glasgow B, Stern G, Yee R (1996): Mapping of Reis-Bucklers' corneal dystrophy to chromosome 5q. Am J Ophthalmol 121:384-390.

Snow AD, Willmer J, Kisilevsky R (1987): Sulfated glycosaminoglycans: a common constituent of all amyloids? Lab Invest 56:120-123.

Spiro D (1959): The structural basis of proteinuria in man. Electron microscopic studies of renal biopsy specimens from patients with lipid nephrosis, amyloidosis and subacute and chronic glomerulonephritis. Am J Pathol 35:47.

Stallings RL (1994): Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequence: implications for human genetic diseases. Genomics 21:116-121.

Stone EM, Mathers WD, Rosenwasser GO, Holland EJ, Folberg R, Krachmer JH, Nichols BE, Gorevic PD, Taylor CM, Streb LM, Fishbaugh JA, Daley TE, sucheski BM, Sheffield VC (1994): Three autosomal dominant corneal dystrophies map to chromosome 5q. Nat Genet 6:47-51.

Tagliavini F, Giaccone G, Verga L, Ghiso J, Frangione B, Bugiani O (1991): Alzheimer patients: preamyloid deposits are immunoreactive with antibodies to extracellular domains of the amyloid precursor protein. Neurosci Lett 128:117-120.

Taschner PE, de Vos N, Thompson AD, Callen DF, Doggett N, Mole SE, Dooley TP, Barth PG, Breuning MH (1995): Chromosome 16 microdeletion in a patient with juvenile neuronal ceroid lipofuscinosis (Batten disease). Am J Hum Genet 56:663-668.

Tassabehji M, Read AP, Newton VE, Harris R, Balling R, Gruss P, Strachan T (1992): Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. Nature 355:635-636.

Tautz D (1989): Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463-6471.

Tautz D, Trick M, Dover GA (1986): Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656.

Tavormina PL, Shiang R, Thompson L.M, Zhu YZ, Wilkin DJ, Lachman RS, Wilcox WR (1995): Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. Nature Genet 9: 321-328.

Terwilliger JD, Ott J (1994): Hanbook of Human Genetic Linkage. The Johns Hopkins University Press, Baltimore.

Thompson MW, McInnes RR, Willard HF (1991): The human gene map: Gene mapping and linkage analysis. In Genetics in Medicine. 5th ed. W.B. Saunders Company. Philadelphia. Chapter 8, pp 167-199.

Tsubota K, Hida T, Murata H, Akiya S, Shinji T, Kimura W (1987): Lattice dystrophy type 1: a report of 8 families. Ophthalmologica 194:71-76.

The Utah Marker Development Group (1995): A collection of ordered tetranucleotide-repeat markers from the human genome. Am J Hum Genet 57:619-628

Vaughan D, Asbury T, Riordan-Eva P (1992): The cornea. In General Ophthalmology. 13th ed. Applton and Lange. Norwalk. pp 7.

Virchow R (1854): On a new substance found in the human brain and spinal cord which reacts chemically like cellulose. Virch Arch Path Anat Physiol, 6:135-137.

Waring GO, Rodrigues MM, Laibson PR (1978): Corneal dystrophies I. Dystropies of epithelium, Bowman's layer and stroma. Surv Ophthal 23:71-122.

Watson JD, Gilman M, Witkowski J, Zoller M (1992): Recombinant DNA. 3rd ed. W.H. Freeman and Company. New York.

Weber JL (1990): Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics 7:524-530.

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

Weber JL, Polymeropoulos MH, May PE, Kwitek AE, Xiao H, McPherson JD, Wasmuth JJ (1991):Mapping of human chromosome 5 microsatellite DNA polymorphisms. Genomics 11:695-700.

Weber JL, Wong C (1993): Mutation of human short tandem repeats. Hum Mol Genet 2:1123-1128.

Weissenbach J (1993a): Microsatellite polymorphisms and the genetic linkage map of the human genome. Curr Opin Genet Dev 3:414-417.

Weissenbach J (1993b): A second generation linkage map of the human genome based on highly informative microsatellite loci. Gene 135:275-278.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992): A second-generation linkage map of the human genome. Nature 359:794-801.

Weleber RG, Carr RE, Murphey WH, Sheffield VC, Stone EM (1993): Phenotypic variation including retinitis pigmentosa, pattern dystrophy, and fundus flavimaculatus in a single family with a deletion of codon 153 or 154 of the peripherin/RDS gene. Arch Ophthalmol 111:1531-1542.

Wiens A, Marles S, Safneck J, Kwiatkowski DJ, Maury CP, Zelinski T, Philipps S, Ekins MB, Greenberg CR (1992): Exclusion of the gelsolin gene on 9q32-34 as the cause of familial lattice corneal dystrophy type I. Am J Hum Genet 51:156-160.

Wilkie AO M, Slaney SF, Oldridge M, Poole MD, Ashworth GJ, Hockley AD, Hayward RD, David DJ, Pulleyn LJ, Rutland P, Malcolm S, Winter RM, Reardon W (1995): Apert syndrome results from localized mutations of FGFR2 and is allelic with Crouzon syndrome. Nature Genet 9:165-172.

Wright J (1994): Mutation at VNTRs: Are minisatellites the evolutionary progeny of microsatellites? Genome 37:345-347.