Molecular Genetic Analysis of Autosomal Recessive Corneal Dystrophies

Ph.D. THESIS

By

AFIA SULTANA

Under the Supervision of

Dr. CHITRA KANNABIRAN

THE DEPARTMENT OF ANIMAL SCIENCES

SCHOOL OF LIFE SCIENCES

UNIVERSITY OF HYDERABAD

HYDERABAD

2007

Molecular Genetic Analysis of Autosomal Recessive Corneal Dystrophies

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

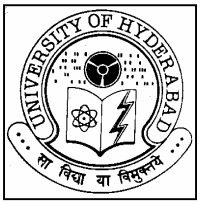
by

Ms. Afia Sultana

Department of Animal Sciences
School of Life Sciences
University of Hyderabad
Hyderabad, 500 046, INDIA

July - 2007

Enrollment No. 03LAPH04



University of Hyderabad
(A central university established in 1974 by an act of parliament)
Hyderabad- 500 046, INDIA

DECLARATION

The research work embodied in this thesis entitled, **"Molecular Genetic Analysis of Autosomal Recessive Corneal Dystrophies"**, has been carried out by me at the L.V. Prasad Eye Institute, Hyderabad, under the guidance of Dr. Chitra Kannabiran and Dr. Aparna Dutta Gupta. I hereby declare that this work is original and has not been submitted in part or full for any other degree or diploma of any other university.

Ms. Afia Sultana _____

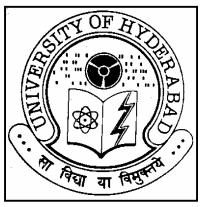
03LAPH04**,** (Ph.D. Student).

Dr. Chitra Kannabiran _____

Supervisor, Scientist, Kallam Anji Reddy Molecular Genetics Lab,
L. V. Prasad Eye Institute, Hyderabad.

Prof. Aparna Dutta Gupta _____

Co-supervisor, Department of Animal Sciences,
School of Life Sciences, University of Hyderabad.



University of Hyderabad
(A central university established in 1974 by an act of parliament)
Hyderabad- 500 046, INDIA

CERTIFICATE

This is to certify that this thesis entitled, **"Molecular Genetic Analysis of Autosomal Recessive Corneal Dystrophies"**, submitted by **Ms. Afia Sultana** for the degree of Doctor of Philosophy to the University of Hyderabad is based on the work carried out by her at the L. V. Prasad Eye Institute, Hyderabad, under our supervision. This work has not been submitted for any degree or diploma of any other University or Institution.

Dr. Chitra Kannabiran _____

Supervisor, Kallam Anji Reddy Molecular Genetics Lab,
L. V. Prasad Eye Institute, Hyderabad.

Prof. Aparna Dutta Gupta _____

Co-supervisor, Department of Animal Sciences,
School of Life Sciences, University of Hyderabad.

Head _____

Department of Animal Sciences

Dean _____

School of Life Sciences, University of Hyderabad.

ACKNOWLEDGMENTS

First and foremost I offer my sincere gratitude to my supervisor **Dr. Chitra Kannabiran** who has supported me throughout the entire study period with her knowledge and remarkable patience while allowing me the room to work in my own way. She taught me how to write academic papers, made me a better programmer, had confidence in me when I doubted myself, and brought out the good ideas in me. She has guided me in more than just my scientific endeavors, and I have had the opportunity to learn many valuable life-lessons at her hands.

I am greatly indebted to **Prof. D. Balasubramanian** who was source of constant encouragement, inspiration and moral support. **Prof. D. Balasubramanian** has a unique ability to awaken interests hidden within his students and help them bring their visions to life. I owe him the deepest of professional and academic debts.

I am thankful to **Prof. Aparna Dutta Gupta**, Department of Animal Sciences, School of life Science, University of Hyderabad, for accepting me as a student in her department and for the valuable suggestions, guidance and support I have received from her.

I would like to specially thank **Dr. G. N. Rao** Chairman, L.V. Prasad Eye Institute for allowing me to work with a great organization. He has been a great source of inspiration to me.

My special thanks to **Dr. Geeta K. Vemuganti** for helping me with the histopathology portion of my work, and in general for her constant encouragement and excellent advices. I thank **Aparna Jagannathan** who initiated this project and shared a great deal of wisdom in the short time I was able to study under her.

My sincere thanks to **Dr. Usha Gopinathan**, **Dr.Subhabrata Chakrabarti**, **Dr. Savitri Sharma**, **Dr. Inderjeet Kaur** and **Dr. Yashoda Ghanekar** for their valuable suggestions, encouragement and moral support.

I am grateful to cornea consultants **Dr. M. S. Sridhar**, **Dr. Prashant Garg**, **Dr. Balasubramanya Ramamurthy, Dr. Gangopadhyay Nibaran and Dr. Virender S. Sangwan** for helping me in the enrollment of patients for the present study and also for helping me to understand the clinical aspects of corneal dystrophies.

I am very thankful to all of my collaborators: **Dr. Srinivasan and Anuradha Chauhan** from IISc Bangalore, **Dr. G. K. Klintworth** from Duke Medical Center Durham and **Dr. J. F. Hejtmancik** from NEI Bethesda.

In my daily work I have been blessed with a friendly and cheerful group of fellow students. Special thanks to my friends **Anees, Guru, Hardeep, Kalyan, Kiran, Purshotham, Rajeshwari, Sagar** and **Venu** for helping me get through the difficult times, and for all the emotional support. My heartfelt thanks to all my colleagues: **Aparna, Avid, Joveeta, Madhavi, Nageshwar, Naidu, Naresh, Srilatha, Sridhar,**

Sankarathi, Saritha, Soundarya, Subhash, Surya and **Vidya** for the comraderie, entertainment, and caring they provided.

I thank **CSIR** for providing me the fellowship, which made my studies possible, and to **DBT** and **HERF** for funding the project.

I also gratefully acknowledge the support I received from the **staff** of the **Clinical Biochemistry, Library, MRD, ISD, Eye Bank, Photography, Stores and Housekeeping** departments.

I thank God for providing me with so many blessings and opportunities. I am indebted to my parents **Mr. Mohammed Zaheer** and **Mrs. Rafia Zaheer** for educating me, for their unconditional support and encouragement to pursue my interests. I am thankful to my **brother, sisters, and brothers- and sister-in-law** for listening to my complaints and frustrations, for believing in me and for their constant encouragement.

ABBREVIATIONS

μg	: Microgram
μl	: Microlitre
μM	: Micromolar
A	: Adenine
AR CHED	: Autosomal recessive congenital hereditary endothelial dystrophy
bp	: Base pair
C	: Cytosine
c.DNA	: Complementary DNA
DM	: Descemet's membrane
DMSO	: Dimethylsulphoxide
Del	: Deletion
dNTPs	: Deoxy nucleotide triphosphates
fs	: Frameshift
G	: Guanine
Ins	: Insertion
IHC	: Immunohistochemistry
KDa	: Kilo Dalton
MCD	: Macular corneal dystrophy
nm	: nanometer
PAPS	: 3-phosphoadenosine 5-phosphosulfate
PAS	: Periodic acid-Schiff
PBS	: Phosphate buffer saline
PAGE	: Polyacrylamide gel electrophoresis
PCR	: Polymerase chain reaction
RFLP	: Restriction fragment length polymorphism
rpm	: Revolutions per minute
SDS	: Sodium dodecylsulfate

SNP	: Single nucleotide polymorphism
SSCP	: Single strand conformation polymorphism
T	: Thymine
EDTA	: Ethylenediamine tetraacetic acid
TAE	: Tris (hydroxymethyl) aminomethane
VA	: Visual acuity

CONTENTS

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

CHAPTER 2: MATERIALS AND METHODS

1.1 CORNEA

The cornea is a thin transparent, avascular structure covering the front of the eye (Figure 1.2A). It is a powerful refracting surface, providing two-thirds of the eye's focusing power. The cornea is extremely sensitive due to presence of nerve endings in it. It comprises one-sixth of the anterior eyeball (Figure1.2A). The junction of the cornea with the sclera is known as the limbus. The adult cornea is about 0.5-0.6 mm thick at the center and 1.2 mm at the periphery.

The cornea consists of five layers (Figure 1.2B)

1. Epithelium
2. Bowman's layer
3. Stroma
4. Descemet's membrane
5. Endothelium

1.1.1 Epithelium

The epithelium originates from the surface ectoderm and constitutes about 10% of the total corneal thickness. The epithelium together with the tear film contributes to the maintenance of the optically smooth corneal surface. It acts as a barrier to external biological and chemical insults. It consists of 5 to 6 layers of three different types of epithelial cells. These are from posterior to anterior, a single layer of columnar basal cells, two to three layers of wing cells, and two to three layers of superficial squamous cells. The basal layer of the epithelium consists of pale tall polygonal cells containing oval nuclei oriented at right angles to the surface of the

cornea. Above the basal layer are three layers of polyhedral wing cells with convex anterior and concave posterior surfaces with tapering edges. Two layers of superficial or surface cells lie along the outer (anterior) aspect of the epithelium. Mitotic activity is greatest in the basal layer of cells and is virtually absent in the layers of surface cells, which normally are exfoliated and constantly replaced by cells from the next deeper layer of epithelium. Basal cells secrete the basement membrane, which is 40 to 60 nm thick, which can be identified easily with periodic acid schiff (PAS) staining. It is composed of type IV collagen and laminin. The presence of the basement membrane between the basal epithelium and the underlying stroma fixes the polarity of the epithelial cells and also provides a matrix on which cells can migrate which inturn is important for maintenance of stratified and well-organized corneal epithelium.

1.1.2 Bowman's layer

The Bowman's layer is an acellular membrane-like zone, which is present, between the corneal epithelium and corneal stroma. It is 12 μm thick and is composed of randomly arranged collagen fibers and proteoglycans. Collagen fibers in the Bowman's layer are primarily of type I and III, which are secreted by stromal keratocytes. The physiological role of Bowman's layer is unclear.

1.1.3 Stroma

The stroma originates from the neural crest and it forms approximately 90% of the thickness of the cornea. It consists of an extracellular matrix, keratocytes and nerve

fibers. Cellular components occupy only 2%-3% of total volume of the stroma where as the rest consists of the extracellular matrix, mainly comprised of collagen and proteoglycans. Keratocytes (stromal cells) are predominant cell type in the corneal stroma, which lie sandwiched between collagen fibrils. The keratocyte consists of prominent large nucleus, with endoplasmic reticulum and mitochondria, which increases significantly in a wounded cornea. Collagen occupies more than 70% of the dry weight of the cornea and is a major extracellular component in the stroma. Collagen in the stroma is mainly of type I with lesser amounts of type III, type V and type VI. Monomer units of tropocollagen are assembled end to end to form microfibrils which get cross-linked to form the collagen fibrils with diameter ranging from 22 to 32 nm. The diameter of collagen fibrils is more uniform in the center than at the limbus. Collagen fibrils group together to give rise to collagen lamellae. The direction of fibrils within lamellae is nearly orthogonal to the fibril direction in adjacent lamellae. The extremely uniform and constant arrangement of these collagen fibrils within the corneal stroma is important for corneal transparency. Any disturbance in this uniformity results in loss of transparency and hence loss of vision.

1.1.4 Proteoglycans

The hydrophilic properties of the stroma result from stromal proteoglycans that constitute the second most abundant biological component in stroma, after collagen (Funderburgh et al., 1991). Proteoglycans consist of central core protein linked covalently to glycosaminoglycans (GAGs). GAGs play an important role in stromal hydration. The normal corneal stroma is about 78% water by weight. Keratan sulfate

and chondroitin or dermatan sulfate are the major components of corneal proteoglycans. Keratan sulfate GAG contains repeating units of N-acetylglucosamine and galactose linked by β (1→4) linkage (Figure 1.1) while dermatan sulfate contains repeats of N-acetylgalactosamine and glucuronic acid linked by β (1→3) linkage.

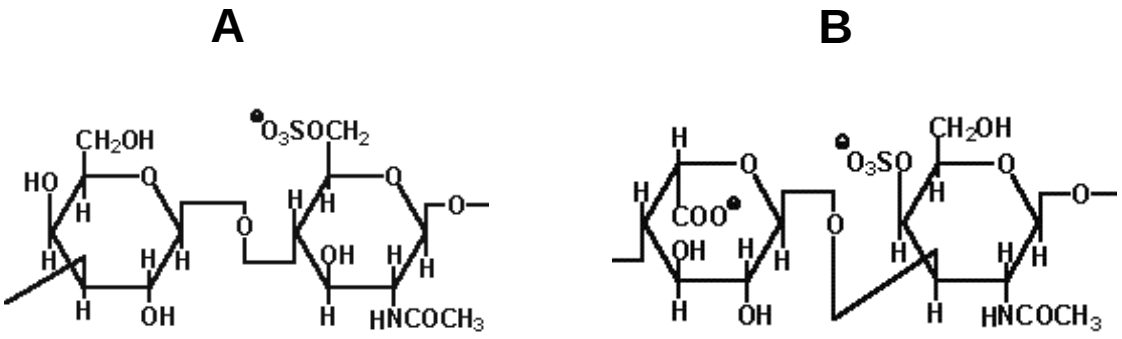


Figure 1.1: Disaccharide units of A: Keratan sulfate showing β (1→4) linkage between N-acetylglucosamine and galactose and B. Dermatan sulfate showing β (1→3) linkage N-acetylgalactosamine and glucuronic acid.

The keratan sulfate of cornea is unique in its abundance, high degree of polymerization, and sulfation. Unlike other GAGs keratan sulfate contains no uronic acids, so its anionic charge depends on sulfation. The three corneal KSPG core proteins that bind keratan sulfate are keratocan, lumican, and mimecan (osteoglycin). These three proteins belong to a family of small leucine-rich proteoglycan proteins (SLRP) (Funderburgh, 2000). These proteins are structurally and antigenically related, and each bears from one to three *N*-linked keratan sulfate chains in addition to several nonsulfated oligosaccharides (Corpuz et al., 1996, Funderburgh et al., 1993). KSPGs interact with the collagen fibrils and maintain the fibrils at a uniform diameter and spacing of 30 to 60 nm (Midura and Hascall, 1989).

Dermatan sulfate in cornea is present as decorin, a small widely distributed proteoglycan also known as PGII, PG40 or PGS2, (Bianco et. al., 1990). Corneal decorin is unusually low in sulfate and iduronic acid .The unique combination of highly sulfated KSPGs and of undersulfated dermatan sulfate proteoglycans (DSPG) are suggested to be important in maintenance of corneal transparency because of hydration properties of the two molecules (Funderburgh, 2000). Binding of water by KSPG is high but reversible. However DSPG exhibits a tight, essentially non-reversible binding to water (Bettelheim and Plessy, 1975).

Several metabolic-storage diseases, in which proteoglycans accumulate in tissues throughout the body, also result in corneal opacity at an early age. One such condition is mucopolysaccharidoses VI. It is caused due to the deficiency of lysosomal enzyme arylsulfatase B, which catalyzes the degradation of dermatan sulfate and chondroitin 4-sulfate (Dorfman and Matalon, 1972). Mutations in the gene encoding decorin cause congenital stromal corneal dystrophy, characterized by corneal opacities appearing as flakes and spots throughout all layers of the stroma (Bredrup et al., 2005). Macular corneal dystrophy (described in greater detail in section 1.3) is a condition in which there is an abnormality in the structure of proteoglycan due to a defect in sulfation of keratan sulfate. These genetic disorders involving alterations in proteoglycans illustrate their importance in corneal transparency. Various biochemical and molecular biological studies on proteoglycans have suggested that they have an important role in maintaining corneal transparency. In a biochemical study, corneas from 5 to 20 day chick embryos were labeled *in vitro* with D-(6-^{3}H) glucosamine and $H_2{}^{35}SO_4$. Measurement of the amount of label in each

glycosaminoglycan showed that a direct relationship exists between the rate of sulfation of KS and onset of corneal transparency (i.e. the rate of sulfation increases rapidly after day 6 and levels off on the onset of corneal transparency which occurs on day 14 (Hart, 1976). Studies on corneal wound healing in rabbits showed that, opaque corneal scars lack both keratan sulfate proteoglycan and uniform collagen fibril spacing. The restoration of transparency was associated with the return of both keratan sulfate proteoglycan and uniform collagen fibril spacing (Hassell et al., 1983). Cornuet and coworkers showed that during the acquisition of corneal transparency in developing chick embryos, the sulfation of lumican is essential for the development of corneal transparency (Cornuet et al., 1994). Transgenic mice with a null mutation in the gene for lumican, the KSPG core protein, develop stromal opacities with sulfated KS levels reduced by 25% (Chakravarti et al., 1998).

1.1.5 Descemet's membrane

The Descemet's membrane (DM) lies on the posterior aspect of the stroma and is secreted by corneal endothelial cells. It is mainly composed of collagen type IV and laminin. (Fitch et al., 1990). At its periphery DM terminates at the junction between corneal and trabecular endothelium. DM is thicker in adults than in children with a thickness of 3-4 μm at birth and 10-12 μm at age 50 years (Johnson et al., 1982). Normally the increase in thickness is uniform.

Transmission electron microscopy showed the normal Descemet's membrane to be composed of two morphologic components-1) the anterior portion secreted in utero that appears as a 110nm, vertical banded pattern and is approximately 3 μm

thick and 2) a posterior homogenous nonbanded layer that thickens with age. These two morphologic components are also called as anterior banded zone (ABZ) and posterior nonbanded zone (PBZ) respectively. The thickness of ABZ remains more or less unchanged throughout life (Johnson et al., 1982). All postnatal laminations have a nonbanded finely granular appearance. In adults there is an indistinct border between the anterior banded zone and the posterior nonbanded zone. The posterior nonbanded zone continues to increase in thickness throughout life. The average thickness of the posterior nonbanded zone at the age of 10 years is approximately 2 μm, which increases to 10 μm by the age of 80 years. It has been speculated that all postnatal acquired diseases spare the anterior banded zone and produce alterations only in the posterior nonbanded zone whereas all diseases arising in utero alter primarily the anterior banded zone (Johnson et al., 1982).

1.1.6 Endothelium

The corneal endothelium originates from the neural ectoderm and is a single layer of polygonal cells that extends over the inner surface of the DM. The endothelium is first detected during the sixth week of gestation. Endothelial cells have a large nucleus and contain abundant cytoplasmic organelles such as mitochondria, endoplasmic reticulum, free ribosomes, and Golgi apparatus. Presence of these cellular characteristics demonstrates that corneal endothelial cells are active in metabolism and secretion. The most important physiologic function of the corneal endothelium is to regulate the water content of the corneal stroma, which is normally 78% by weight (Waring et al., 1982). The gap junctions in the endothelial cells permit the

penetration of small molecules and electrolytes between endothelial cells. The interconnected endothelial cell layer provides a leaky barrier to the aqueous humor. There is slow decrease in the number and density of endothelial cells with advancing age. Approximately 3500 to 4000 cells per mm^2 are present at birth and by adult life almost 25% of these cells may be lost. This is accompanied by thinning and spreading of residual cells (Hiles et al., 1979). The adult human endothelium has a limited capacity to divide and replace aging or injured endothelial cells. However the endothelium has remarkable ability to enlarge, reorganize, migrate, and maintain tight apposition to neighboring cells, thereby maintaining an intact monolayer that regulates the passage of water from the aqueous humor into the stroma (Van and Hyndiuk, 1975). A minimum of 300-600 cells per mm^2 is required to ensure normal function of the endothelial layer. If the endothelial cell density drops below this, the functional reserve will be very minimal and hence corneal edema will occur (Landshman et al., 1988).

Transparency of the cornea is a pre-requisite for normal vision, and is maintained by an intact wet corneal epithelium, adequate number of functionally intact endothelial cells, appropriate hydration and proper orientation of the corneal stroma. Anything that affects this regularity results in loss of transparency.

1.2 CORNEAL DYSTROPHIES

Corneal dystrophies are a group of hereditary, bilateral, progressive, non-inflammatory disorders resulting in loss of transparency of the cornea and consequent loss of vision. These disorders may be present at birth or during

adolescence or adulthood. There are many different corneal dystrophies that are distinguished by the corneal layer(s) that they affect (i.e., epithelium, Bowman's layer, stroma, Descemet's membrane or endothelium), and by their mode of inheritance. Till date autosomal dominant, autosomal recessive and X-linked recessive modes of inheritance have been reported for corneal dystrophies (Klintworth, 2003). A summary of known corneal dystrophies is provided in Table 1.1.

Figure 1.2A

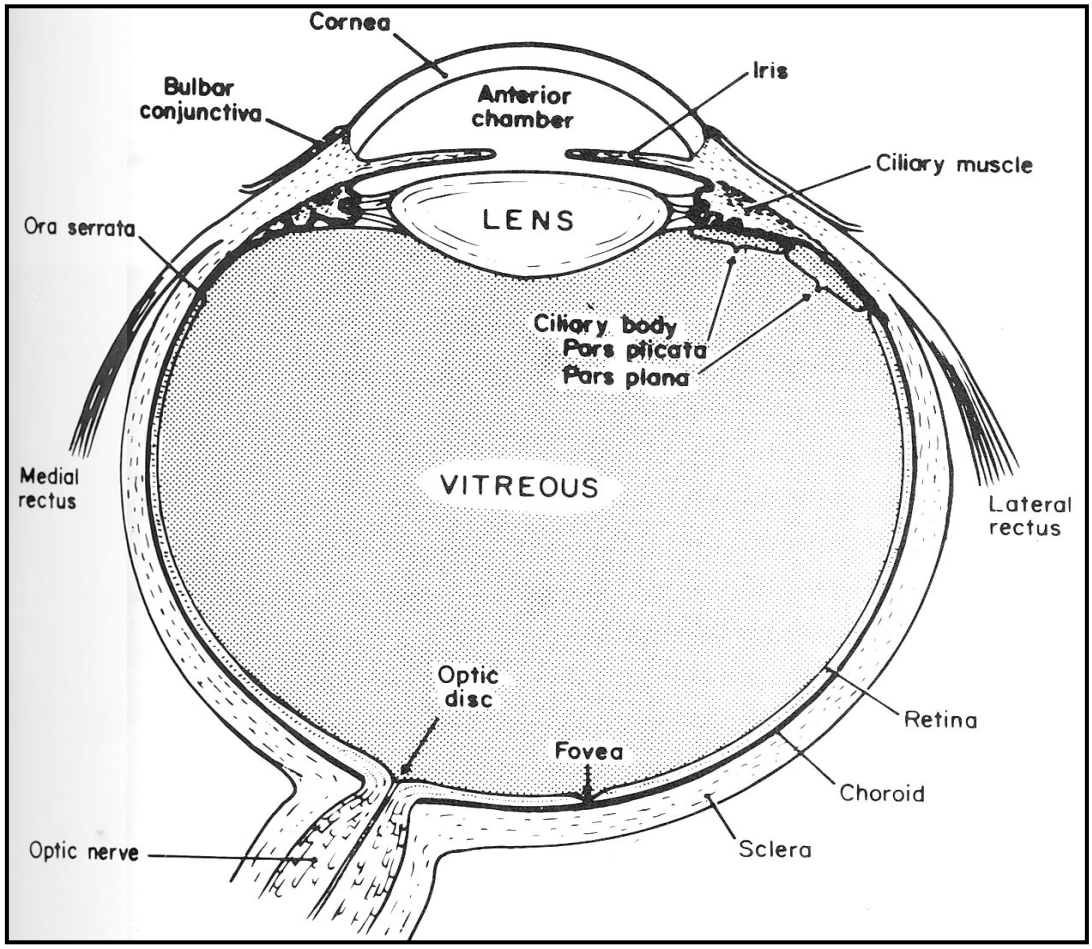


Figure1.2A: Cross-section of human eye (taken from home study course for ophthalmic medical assistants; published by American Academy of Ophthalmology 1983:p.23).

Figure 1.2B

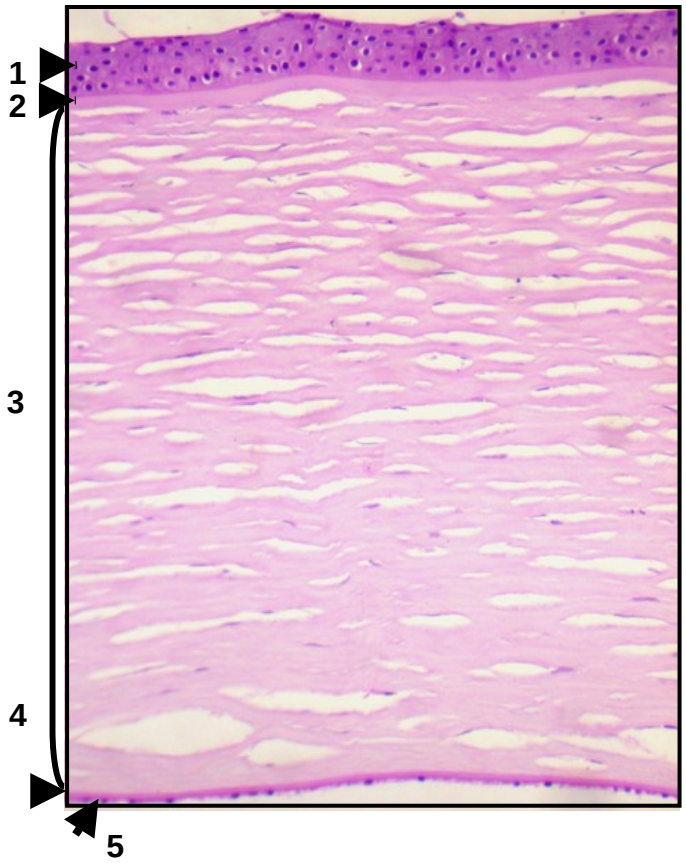


Figure1.2B: Section of normal human cornea (formalin-fixed, paraffin-embedded and stained with PAS. Magnification 100X) showing the different layers 1. Epithelium, 2. Bowman's membrane, 3. Stroma, 4. Descemet's membrane, 5. Endothelium. [Picture courtesy of Dr Geeta Kashyap Vemuganti, Ophthalmic Pathology Service, LVPEI, Hyderabad].

Table1.1. Corneal dystrophies, primary corneal layer involved and genetics.

Corneal dystrophies	Inheritance	Gene locus	Gene	Primary corneal layer affected	References
Fabry disease/Cornea verticillata	XR	Xq22	Alpha galactosidase (GLA)	Epithelium	Bernstein et al., 1989
Stocker-Holt dystrophy	AD	17q22	Cytokeratin12	Epithelium	Klintworth et al., 1999
Meesmann dystrophy	AD	12q13, 17q12	Cytokeratin 3, Cytokeratin12	Epithelium	Irvine et al., 1997
Gelatenous drop like dystrophy (familial subepithelial amyloidosis)	AR	1p32	Gastrointestinal tumor-associated antigen (M1S1)	Stroma	Tsujikawa et al., 1999
Bietti crystalline corneoretinal dystrophy	AR	4q35	Cytochrome P4V2	Stroma	Li et al., 2004
Congenital hereditary stromal dystrophy	AD	12q22	Decorin	Stroma	Bredrup et al., 2005
Central crystalline dystrophy, Schnyder dystrophy	AD	1p34.1-p36	Unknown	Stroma	Shearman et al., 1996
Fleck's dystrophy	AD	2q35	Phosphatidylinositol-3-phosphate-5-Kinase type III (PIP5K3)	Stroma	Li et al., 2005
Granular dystrophy	AD	5q31	Transforming growth factor beta induced (TGFβ1)	Stroma	Munier et al., 1997
Lattice dystrophy types I, IIIA,	AD	5q31	Transforming growth factor beta induced (TGFβ1)	Stroma	Munier et al., 1997; Yamamoto et al., 1998

Corneal dystrophies	Inheritance	Gene locus	Gene	Primary corneal layer affected	References
Lattice dystrophy type II	AD	9q34	Gelsolin	Stroma	Maury et al., 1990
Macular dystrophy	AR	16q22	Carbohydrate sulfotransferase-6 (CHST6)	Stroma	Akama et al., 2000
Cornea farinata and deep filiform dystrophy	XR	Xp22.32	Steroid sulfatase gene (STS)	Stroma	Basler et al., 1992
Congenital hereditary endothelial dystrophy type I	AD	20p11.2-q11.2	Unknown	Endothelium and Descemet's membrane	Toma et al., 1995
Congenital hereditary endothelial dystrophy II (CHED2)	AR	20p13	Sodium bicarbonate transporter like solute carrier 4 family member11 (*SLC4A11*)	Endothelium and Descemet's membrane	Vithana et al., 2006
Fuchs dystrophy	AD	1p34.3-p32	Alpha 2 chain type VIII collagen (COL8A2)	Endothelium and Descemet's membrane	Biswas et al., 2001,
Posterior polymorphous dystrophy	AD	1p34.3-p32, 20p11.2,10p11	Visual system homeobox1 homolog VSX1, Alpha 2 chain type VIII collagen (COL8A2), Transcription factor 8 (TCF8)	Endothelium	Heon et al., 2002, Biswas et al., 2001, Krafchak et al., 2005

Table1.1. Corneal dystrophies, their mode of inheritance, gene locus, gene and primary layer involved are shown. XR indicates X-linked recessive, AD indicates autosomal dominant and AR indicates autosomal recessive modes of inheritance.

The present study deals with molecular genetic analysis of two autosomal recessive corneal dystrophies, macular corneal dystrophy (MCD) and congenital hereditary endothelial dystrophy (CHED). MCD and AR CHED may be relatively more frequent than other types of corneal dystrophies in South India. A study from our institution has shown that MCD and CHED account for 29.3% and 34.8% respectively, of patients with corneal dystrophies undergoing keratoplasty (corneal grafting) (Pandrowala et al., 2004).

1.3 MACULAR CORNEAL DYSTROPHY

Macular corneal dystrophy (MCD; OMIM217800) is bilateral autosomal recessive disorder first described in 1890 by Groenouw (Groenouw et al., 1890).

1.3.1 Clinical features

Macular corneal dystrophy is characterized by corneal opacities resulting from intra- and extracellular deposits within the corneal stroma (SunderRaj et al., 1987) It is the most severe and one of the most common types of corneal dystrophies in South India. Symmetric changes are usually first noted between 3 and 9 years of age, characterized by a diffuse, fine, superficial clouding in the central stroma. Opacification extends to the periphery and usually involves the entire thickness of the cornea by the second decade of life. Multiple, irregular, dense, gray-white nodules with indistinct borders can protrude anteriorly, resulting in irregularity of the corneal surface (Figure1.3A). In the corneal periphery, deep posterior focal plaques grayness and a guttate appearance of DM may be seen (Francois et al., 1966). It has also

been associated with reduced central corneal thickness and corneal guttae (Quantock et al., 1990). Patients experience progressive loss of vision as well as attacks of irritation and photophobia and recurrent erosions (Jonasson et al., 1989). Vision gets severely affected by second or third decades.

1.3.2 Histopathology

Histologically MCD has been differentiated from the other two major stromal dystrophies known as granular and lattice corneal dystrophy by Jones and Zimmerman (Jones and Zimmerman, 1959). Histologically MCD is characterized by the accumulation of glycosaminoglycans (GAGs) between the stromal lamellae, underneath the epithelium, within keratocytes and endothelial cells which stain with Alcian blue, colloidal iron, metachromatic dyes and PAS (Figure1.3B) (Snip et al., 1973; Teng, 1996). Light microscopy demonstrated degeneration of the basal epithelial cells, and focal epithelial thinning over the accumulated material. Bowman's membrane may be irregular, thinned, or absent in some areas.

Figure 1.3

A

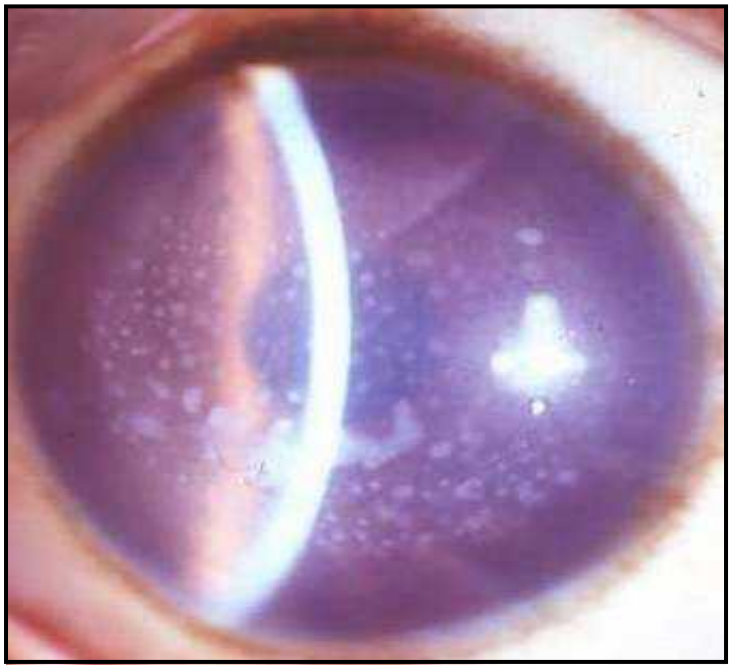


Figure1.3. A: Slit lamp view of cornea from patient with MCD showing rounded grayish-white stromal opacities [Picture courtesy of Dr. M.S. Sridhar from Cornea and Anterior Segment Services, LVPEI, Hyderabad]. B: Corneal section from patient with MCD stained with Alcian blue (counterstained with nuclear fast red solution) showing accumulation of unsulfated GAGs in subepithelial region, stroma, and endothelium (magnification 100X). [Picture courtesy of Dr Geeta K. Vemuganti, Ophthalmic Pathology Service, LVPEI, Hyderabad].

Figure 1.3

Initial investigation into the pathogenesis of MCD was done by light and electron microscopy along with histochemical procedures. Results of these studies suggested that MCD is a metabolic storage disease characterized by accumulation of mucopolysaccharide within stromal keratocytes (corneal fibroblasts) (Klintworth and Vogel, 1964). The anterior banded portion of the Descemet's membrane is normal, but the posterior, nonbanded portion is infiltrated by vesicular and granular material thought to be deposited by the abnormal endothelium (Francois et al., 1975; Garner, 1969).

1.3.3 Pathogenesis

Studies of sulfated glycosaminoglycans in corneal explant and confluent stromal cell cultures from MCD corneas showed abnormality in the synthesis of keratan sulfate. (Klintworth and Smith, 1977). Organ culture experiments with normal and MCD corneas using radioactive precursors of glycoproteins and proteoglycans done by Hassell and co-workers (Hassell 1980) showed that MCD-affected corneas fail to synthesize mature KS although they synthesize normal chondroitin sulfate proteoglycan. On the other hand, control corneas were found to synthesize both chondroitin sulfate and keratan sulfate proteoglycans. In addition, MCD corneas were found to synthesize a glycoprotein (slightly smaller than normal KS proteoglycan) with unusually large oligosaccharide side chains, which was not detected in normal control corneas (Hassell et al., 1980). Klintworth and coworkers found similar defects in the synthesis of keratan sulfate proteoglycans (Klintworth and Smith, 1983).

Studies by Nakazawa & coworkers showed that the defect in MCD corneas possibly involved sulfation of KS. MCD corneas were observed to synthesize an immunoreactive glycoprotein in nearly equal amounts as keratan sulfate proteoglycan synthesized by the control (keratoconus) cornea. The oligosaccharides on the immunoprecipitated macular glycoprotein appeared to be normal. However, the macromolecules contained an unsulfated glycoconjugate that was nearly as large as the normal KS chains. Further, the glycoconjugates from MCD corneas were resistant to digestion with keratanase indicating that MCD is caused by an error in the synthesis of keratan sulfate possibly involving the specific sulfotransferases involved in sulfation of the lactosaminoglycan backbone of the chains (Nakazawa et al., 1984). Altogether the results from organ culture studies by various groups showed that MCD is caused by a defect in the synthesis of mature keratan sulfate, possibly involving sulfation of lactosaminoglycan precursor.

Attempts to identify the sulfotransferase responsible for the defect in sulfation of KS were made biochemically, by measuring sulfation activity in both serum (Hassell and Klintworth, 1997; Hasegawa et al., 1999) and in corneal extracts from MCD patients (Hasegawa et al., 2000). No difference was found in the sulfotransferase activity between the serum from normal controls and MCD patients, suggesting that the defective sulfotransferase involved in the sulfation of keratan sulfate may not be secreted into the serum. The latter study based on the corneal extracts from normal and MCD corneas identified the activity that is defective in MCD as N-acetyl glucosamine sulfotransferase. (Hasegawa et al., 2000).

1.3.4 Immunophenotypes of MCD

Initially MCD was thought to be restricted to ocular tissue because its clinical manifestations were restricted to cornea. However the absence of antigenic keratan sulfate in the serum of a group of patients with MCD (see below) suggested that KS deficiency may be systemic, since serum KS is thought to originate from cartilage (Thonar et al., 1985; Klintworth et al., 1986). Direct evidence for absence of KS in cartilage from nose and ear in patients with MCD type I came from studies by Edward and coworkers (Edward et al., 1990).

Different immunophenotypes of MCD are recognized, based on the presence of antigenic KS in the patient's serum and cornea using a monoclonal antibody that specifically reacts with sulfated keratan sulfate (Caterson et al., 1983). In MCD type I, neither the serum nor the corneal tissue contain antigenic KS (AgKS) whereas in MCD type II, it is detectable in the corneal tissue and is present at variable levels in the serum. (Edward et al., 1988; Yang et al., 1988). Subsequently another type (MCD type IA) was identified in which sulfated KS is absent in the cornea and the serum but can be detected in the keratocytes (Klintworth et al., 1997; Cursiefen et al., 2001).

Structural differences between proteoglycans of MCD type I and II corneas have been noted. MCD type I corneas were found to contain KS-proteoglycan precursor (lactosaminoglycan) with defective sulfation of KS whereas MCD type II cornea produced a normal KSPG. The ratio of KSPG to dermatan sulfate-PG was similar to normal corneas, though the net synthesis of PGs was lower (Midura et al., 1990). A more recent study using fluorochrome-assisted carbohydrate

electrophoresis revealed that sizes of keratan sulfate side chains were smaller than in normal control corneas and sulfation was defective in both MCD types I and II. Hyaluronan which is absent in healthy adult corneas was detected in both MCD subtypes (Plaas et al., 2001). The use of lectin to detect poly N-acetyllactosamine (Lewis et al., 2000) and antibodies specific for unsulfated KS in combination with antibodies specific for sulfated epitopes demonstrated that MCD type I corneas have high levels of unsulfated KS with little or no sulfated KS.

1.3.5 Molecular genetics

Mapping of the locus for MCD on 16 American and Icelandic families (11 of which were type I and 5 were type II) showed linkage of disease to chromosome 16 suggesting that both MCD types were due to the same gene (Vance et al., 1996). The interval was narrowed down and two candidate genes (TAT and LCAT) were excluded from the critical region (Liu et al., 1998; 2000). Akama and coworkers (Akama et al., 2000) identified two expressed sequence tags (ESTs) for sulfotransferases that mapped within the MCD critical interval. One was the *CHST5* gene encoding intestinal N-acetyl glucosamine 5 sulfotransferase 1 and another highly homologous gene that they designated as carbohydrate sulfotransferase 6 (*CHST6*). (Akama et al., 2000; Hemmerich et al., 2001). They found mutations in *CHST6* in MCD patients but not in *CHST5*. The *CHST6* gene codes for corneal N-acetyl glucosamine 6-sulfotransferase (c-GlnNAc6ST), is expressed in the cornea and consists of four exons with the entire coding region present within the third exon (Figure1.4). Akama and co-workers reported mutations in the coding region in MCD

type I whereas for MCD type II, they found large deletions and or replacements in the upstream region of the *CHST6* gene (Akama et al., 2000). They proposed that the MCD type II mutations led to loss of cornea-specific expression of *CHST6*. Differences in the nature of *CHST6* mutations in MCD type1 versus type II have not been borne out by other studies (Liu et al., 2005, 2006; Warren et al., 2003). Further studies by Akama and co-workers (Akama et al., 2001) showed that c-GlnNAc6ST transfers the sulfate group specifically to C6 of GlcNAc and synthesizes highly sulfated keratan sulfate *in vitro*. Cloning and expression of missense mutations found in MCD *in vitro* using transfected Hela cells showed that mutant proteins were defective for enzymatic activity (Akama et al., 2001). Subsequent to the report by Akama et al., (2000) other studies reported *CHST6* mutations in MCD patients from various populations (Bao et al., 2001; El-Ashry et al., 2002; Liu et al., 2000; Niel et al., 2003).

Organization of *CHST6* gene

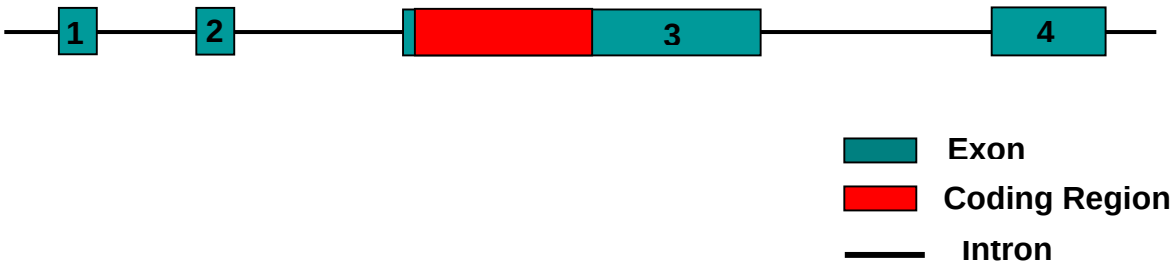


Figure1.4: Genomic organization of *CHST6* gene consisting of 4 exons. Part of exon 3 (red colour) codes for the protein.

1.3.6 Sulfotransferases

Sulfotransferases belong to a large family of enzymes which transfer the sulfate group from 3-phosphoadenosine 5-phosphosulfate (PAPS) to an acceptor group of numerous substrates. This reaction, which is observed widely from bacteria to humans, plays a key role in various biological processes such as cell communication, growth, development, and defense. Sulfotransferases (STs) are classified into two types: cytosolic sulfotransferases and membrane-bound sulfotransferases. Cytosolic sulfotransferases sulfonate steroids, drugs and environmental chemicals and are involved mainly in hormone homeostasis and metabolic detoxification or activation of xenochemicals. Golgi membrane-bound sulfotransferases sulfonate heparansulfates, GAGs, and glycoproteins (Lidholt and Lindahl, 1992; Falany, 1997). Although cytosolic and membrane-bound sulfotransferases lack overall sequence similarity, two motifs-one at the N-terminus called as 5'-phosphosulfate-binding domain (5-PSB) corresponding to five residues 47-54 in mouse estrogen sulfotransferase and the other in the middle of the protein called 3'-phosphate-binding domain (3'PB) corresponding to residues 121-141 in mouse estrogen sulfotransferase, are conserved among all the known sulfotransferases. These two regions are separated by a relatively fixed number of residues and are functionally important as they interact with 5'phosphate and 3'phosphate groups of PAPS (Kakuta et al., 1998). The sequence alignments and structural correlations by Kakuta and coworkers revealed structural similarity among sulfotransferases suggesting that they might have evolved from a common ancestral gene (Kakuta et al., 1998). The X-ray crystal structural studies of cytosolic and membrane bound sulfotransferases [mouse estrogen

sulfotranferase (mEST), human dopamine/catecholamine sulfotransferase (SULT1A3), human hydroxysteroid sulfotransferase (hHST), human estrogen sulfotransferase (hEST), the N-sulfotransferase domain of heparan sulfate N-deacetylase/N-sulfotransferase1 (NST1)] showed that sulfotransferases are globular proteins composed of $\propto$ & β domain with the β domain consisting of characteristic five-stranded parallel β-sheet. The β-sheet contains the PAPS-binding site and the core of the catalytic site, both of which are conserved in cytosolic as well as membrane STs. The core PAPS-binding site is made up of strand-loop-helix and strand-turn-helix motifs that contain the 5'PSB and 3'PBmotifs. The structure as well as sequence of the PAPS phosphate-binding sites of the sulfotransferases are conserved (Negishi et al., 2001).

1.4 CONGENITAL HEREDITARY ENDOTHELIAL DYSTROPHY

Congenital hereditary endothelial dystrophy (CHED) is a corneal disorder resulting from an abnormality of endothelial cells (Maumenee, 1960; Keates and Cvintal, 1964). CHED has both autosomal dominant (locus CHED-1 OMIM; 27100) as well as autosomal recessive (locus CHED-2 OMIM; 121700) modes of transmission. No consistent clinical differences between autosomal recessive CHED and autosomal dominant CHED have been found although recessive CHED is considered to manifest earlier (Judisch and Maumenee, 1978; Kirkness et al., 1987). Autosomal recessive CHED manifests at birth or within the first year of life.

1.4.1 Clinical features

Autosomal recessive CHED (AR CHED) manifests as bilateral, symmetric, noninflammatory corneal clouding without other anterior segment abnormalities, usually evident at birth or in the early postnatal period. The corneal opacification extends to the limbus without clear zones (Figure1.5A). It is associated with marked corneal thickening (two to three times that of the normal cornea). The corneal epithelium appears as irregular or roughened suggestive of edema. The stromal haziness involves all layers of the cornea, but is greater in the deeper layers with stromal dots and flakes (Kenyon and Maumenee, 1968). The Descemet's membrane appears thickened with endothelial cells decreased or absent. Nystagmus may be present in some children with reduced vision. Vascularization is absent in AR CHED except following ulceration in some older patients, after years of chronic corneal edema. The main differential diagnoses for AR CHED are congenital glaucoma, mucopolysaccharidoses and posterior polymorphous dystrophy (Pedersen et al., 1989; Mullaney et al., 1995). In a few cases, autosomal recessive CHED has been reported to be associated with teenage onset perceptive hearing loss, known as Harboyan syndrome (Harboyan et al., 1971; Judisch et al., 1978).

1.4.2 Histopathology

Light microscopic studies showed the epithelium in CHED as edematous, thin or atrophic with hydropic changes of the basal epithelium, vacuolization of intercellular spaces, and occasional microbullae. Subepithelial fibrosis was reported in a few cases (3%) (Antine, 1970; Kenyon and Antine, 1971; Kirkness et al., 1987). Changes

in the Bowman's membrane vary from minute disruptions, areas of partial loss, areas of irregular thickening to absence (Ehlers et al., 1998; McCartney et al., 1988). The stroma shows thickening, which is two to three times that of normal and is severely disorganized with disruption of lamellar pattern. The Descemet's membrane is intact but much thickened (Figure1.5B) with a few cases showing the presence of a posterior collagenous layer, which is less PAS-positive (Kirkness et al., 1987; Ehlers et al., 1998). The corneal endothelial cells are absent or markedly reduced and show highly abnormal morphology with irregular and multinucleated cells (Mullaney et al., 1995; Ehlers et al., 1998).

Ultrastructural studies showed the thickness of Descemet's membrane in AR CHED patients to be 18-20 μm thick while it was 6-7μm in normal age-matched normal corneas (normal corneas age range-12 weeks of gestation to 18 years) (Murphy et al., 1984; Ehlers et al., 1998). Descemet's membrane in CHED shows a normal anterior banded zone, which is approximately 3 μm thick. The posterior nonbanded zone is abnormal and is poorly demarcated with multiple focal areas of fibrillar deposits with thickness ranging from 20-24 μm (Ehlers et al., 1998).

Figure 1.5

A

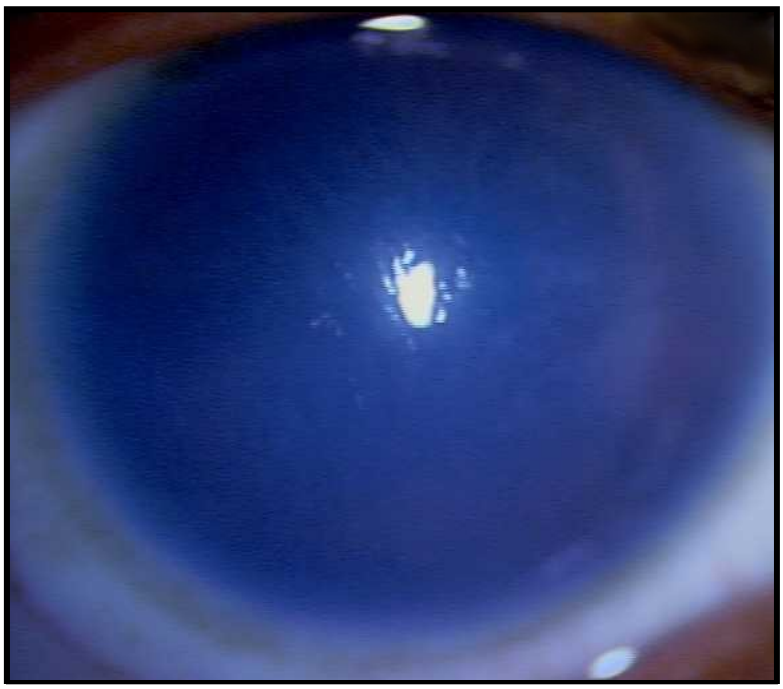
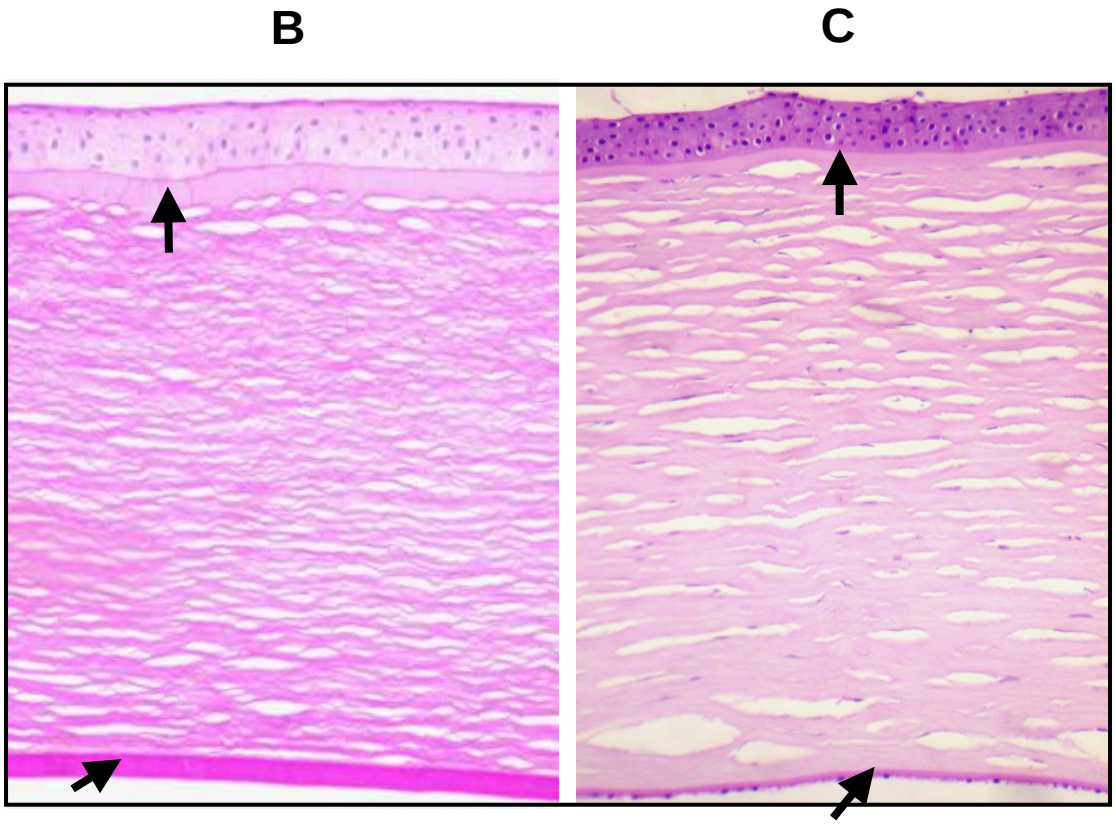


Figure1.5. A: Slit lamp view of cornea of a 3-year old patient with AR CHED showing bluish-white haze [Picture courtesy of Dr. Prashant Garg from Cornea and Anterior Segment Services, LVPEI, Hyderabad]. B: Corneal section from patient with AR CHED showing an absence of endothelial cells, a thickened Bowman's membrane and markedly thickened Descemet's membrane. C: Normal cornea of 8 year old patient with retinoblastoma for comparison (periodic-acid-shiff x100). Top and bottom arrows point to the Bowmans and Descemet's membranes, respectively [Picture courtesy of Dr Geeta K. Vemuganti, Ophthalmic Pathology Service, LVPEI, Hyderabad].

Figure 1.5



The exact pathophysiological mechanism(s) and the etiology of CHED are not fully understood but it has been hypothesized by Mullaney and coworkers based on light, transmission and scanning electron microscopy that the corneal endothelium in CHED is functionally normal during the first five months in utero, leading to deposition of a normal anterior part of Descemet's membrane. However after the fifth month in utero the endothelial cells undergo an abnormal differentiation and a degenerative transformation into "fibroblast-like cells" with accelerated and altered collagen secretion, leading to an abnormal posterior part of DM (Mullaney et al., 1995)

1.4.3 Molecular Genetics

Linkage analysis in large consanguineous Irish and Saudi Arabian pedigrees with AR CHED excluded the loci for AD CHED and posterior polymorphous dystrophy (PPMD) (Callaghan et al., 1999; Kanis et al., 1999). By homozygosity mapping Hand and coworkers localized autosomal recessive CHED to the short arm of chromosome 20 at 20p13 (CHED2). A region of homozygosity in all affected individuals was identified placing the disease gene locus within an 8-cM region flanked by markers D20S113 and D20S882 (Hand et al., 1999). This was subsequently confirmed in other study by Mohammed et al. (2001). Mutations in the sodium bicarbonate transporter-like solute carrier family 4 member 11 (*SLC4A11*) gene present in this locus were found to cause AR CHED (Vithana et al., 2006).

Sodium bicarbonate transporter-like solute carrier family 4 member 11 (SLC4A11)

The *SLC4A11* gene belongs to a superfamily of bicarbonate transporters. The *SLC4* genes (solute carrier family 4) code for integral membrane proteins with 10-14 transmembrane segments. These proteins function as Cl^--HCO_3^- exchangers, Na^+/HCO_3^- cotransporters or Na^+driven Cl^--HCO_3^- exchangers (Romero et al., 2005). *SLC4A11* is also known as *BTR1* (bicarbonate transporter related protein-1) or NaBC1 (sodium-coupled borate cotransporter). The *BTR1* cDNA was cloned from a kidney cDNA library by Parker and coworkers (Parker et al., 2001). The *SLC4A11* gene has 19 exons, which codes for a protein of 891 amino acids with a calculated molecular mass of 100 kDa. The *SLC4A11* protein has 14 transmembrane domains and intracellular N and C termini. It contains multiple intracellular phosphorylation sites and 2 extracellular N-glycosylation sites (Parker et al., 2001).

Expression of *BTR1/SLC4A11* gene is seen in kidney, salivary gland, testis, thyroid, trachea, and corneal endothelium (Parker et al., 2001; Gottsch et al., 2003). Vithana and coworkers by in situ hybridiztion showed its expression in the mouse cornea at embryonic day 8, which is equivalent to human gestational month 5, the time at which CHED2 pathology develops in humans (Vithana et al., 2006). *BTR1* is homologous to *BOR1*, a borate transporter in plants (Parker et al., 2001). It functions as a ubiquitous electrogenic sodium-coupled borate transporter in the presence of borate, while in the absence of borate it conducts Na+ and H+. In view of the requirement for borate in growth and development, *BTR1* may be a mediator for these processes (Park et al., 2004).

1.5 OBJECTIVES OF THE STUDY

MCD and CHED are the most frequent types of corneal dystrophies leading to visual impairment at an early stage of life. Autosomal recessive CHED manifests at birth or in the early neonatal period, where as MCD manifests by late first decade or early second decade. If not treated, bilateral corneal opacification in CHED may lead to deep and irreversible amblyopia or lazy eye which is an uncorrectable decrease in vision caused by either absence of or poor transmission of the visual image to the brain for a sustained period of dysfunction or disuse of eyes during early childhood. Both of these dystrophies require corneal grafting to restore vision, with its attendant risks and complications. Knowledge of the molecular genetic bases of these two dystrophies may provide valuable insights into the biology and normal functioning of the cornea. It may also contribute to an understanding the mechanisms of these disorders and potentially to development of newer therapies. When the present studies were undertaken, there were no published data on the molecular genetics of MCD from Indian populations and the gene for CHED2 had not been identified. This study was therefore designed with the following objectives.

MCD

1. To screen the *CHST6* gene in Indian families with MCD in order to identify the spectrum of pathogenic mutations.
2. Comparative three-dimensional modeling of *CHST6* protein to understand the impact of identified mutations on the structure of *CHST6* protein.

3. To determine the immunophenotypes of MCD in Indian patients and to determine whether there are correlations between types of mutation and immunophenotypes.

AR CHED

1. To determine whether the disease in Indian families with AR CHED maps to the CHED2 locus.
2. To refine the disease interval of the AR CHED locus by fine mapping.
3. To identify the gene responsible for AR CHED by screening candidate genes in the refined interval.
4. To establish the mutational spectrum of AR CHED.

2.1 ENROLLMENT OF PATIENTS AND CONTROLS

The study protocol was approved by the Institutional Review Board and adhered to the guidelines of the Declaration of Helsinki. Written consent for participation in the study was obtained from eligible patients who presented at the L. V. Prasad Eye Institute between June 2002 and July 2006. Family histories for the diseases in question were obtained from patients (or parents or guardians in case of minors). Pedigrees were drawn for each family based on information obtained using CYRILLIC software.

2.1.1 Clinical examinations

All patients and available family members underwent clinical examination by corneal specialists at the institute. Complete ophthalmic examination included visual acuity assessment, slit lamp bio-microscopy, pachymetry, fundus examination and measurement of intraocular pressure (IOP). The clinical data was documented in the patients' medical records by the concerned clinicians or optometrists during evaluation. Diagnostic criteria for inclusion in the study were as detailed below.

2.1.1A Inclusion criteria for MCD

Patients with two or more of the following characteristics were included in the study:

1. Bilateral full thickness corneal haze
2. Anterior nodular lesions
3. Clinically visible endothelial deposits with corneal thinning
4. Histopathological diagnosis of MCD

Patients with history of any prior surgery other than penetrating keratoplasty were excluded from the study.

2.1.1B Inclusion criteria for AR CHED patients

Patients with the following clinical and histopathological characteristics were included in the study:

Clinical characteristics

1. Cloudy cornea from birth to 10 years
2. Cornea thicker than normal
3. Corneal diameter appropriate for age (11-12mm)
4. IOP normal (< /=22mmHg)
5. Bilateral symmetrical or asymmetrical corneal edema

Histopathological characteristics

1. Thickening of cornea
2. Thickening of Descemets membrane
3. Normal or reduced endothelial cell count and changes in morphology

Patients with the following characteristics were excluded from the study:

1. Any prior surgery other than penetrating keratoplasty
2. Haabs striae or any sign of PCG

2.1.2 Sample collection

Peripheral blood samples (4-8ml) were collected in heparinized vaccutainer tubes (Vacuette, Greiner bio-one GmbH, Austria) from probands, affected and unaffected

family members available or willing to be recruited for the study. Blood samples were stored at –20º C until the DNA was extracted.

2.1.2A MCD

A total of 72 patients with MCD belonging to 59 unrelated families were enrolled in the study. In 56 patients, MCD was confirmed by the histopathological examination of the excised corneal buttons. Ages of onset of disease ranged between 7 to 49 years. Of the 72 patients studied, 24 were sporadic cases with no other affected relatives while the rest were familial cases. Consanguinity was present in 40 out of 59 families.

2.1.2B AR CHED

A total of 69 patients with AR CHED belonging to 56 unrelated families along with 108 unaffected family members were enrolled in the study. Ages of onset of disease ranged between 0 to 24 years. Of the 56 families studied, 39 families had only one affected, while the rest had more than one affected family member. Consanguinity was present in 39 out of 56 families.

2.1.2C Controls

Blood samples were drawn with informed consent from a total of 75 normal healthy unrelated individuals (without any corneal disorder by history or examination).

2.2 Molecular genetic analysis

2.2.1 Isolation of genomic DNA from peripheral blood

Genomic DNA was isolated from the peripheral blood using the phenol chloroform method with slight modifications (Sambrook et al*.,* 1989). Blood samples were mixed with an equal volume of 1X phosphate buffer saline (PBS) and centrifuged at 3500 rpm for 10 minutes. The supernatant was removed and the WBC pellet was washed 2-3 times with 8 ml of PBS. The pellet was then resuspended in 6 ml of lysis buffer [0.5%SDS, 0.1M EDTA and 10 mM Tris (pH 8.0)]. Proteinase K and RNAase A were added to a final concentration of 100 µg/ml and 20 µg/ml respectively. Samples were vortexed for 5-10 seconds and incubated overnight at 37^{O}C. Proteins and other non-nucleic acid organic molecules were removed first by extraction with an equal volume of phenol followed by two extractions with equal volumes of phenol-chloroform mix. DNA was then precipitated using 2 volumes of ethanol and 2 M ammonium acetate (1.5 ml of 10 M ammonium acetate stock). DNA was spooled out onto a pipette tip and transferred to a 1.5 ml microfuge tube containing 70% ethanol followed by centrifugation at 10,000 rpm for 30 seconds. The supernatant was removed and the pellet was air-dried and dissolved in TE buffer (10mM Tris-Hcl; 1mM EDTA; pH 8.0) with slow pipetting with a P-1000 and was stored at -20 ^{O}C till further application.

2.2.2 Quantification of genomic DNA

The concentration of DNA in stock solutions was determined by spectrophotometry. The absorbance was measured at wavelengths 260nm and 280nm. DNA samples were diluted 1:100 in water and absorbance was read in a spectrophotometer using deionised water as blank. Concentration of DNA was calculated taking A260 value of

1.0 as equal to 50 µg DNA/ml for double-stranded DNA. A260/280 ratios were measured to assess the purity of DNA. The concentration of DNA was calculated using the following formula.

Concentration of DNA = OD at 260 nm X 50 X 100 (dilution factor).

2.2.3 Polymerase chain reaction (PCR)

PCR of coding regions of the genes investigated was done for subsequent mutational screening. Primers for polymerase chain reaction were obtained commercially. Primers were designed using software freely available on the internet (see Appendix II) or sequences were obtained from published literature. All the primers designed were greater than 17 base pairs with a GC content range of 40-80%. Tm was calculated according to the formula Tm(C) = 2(A+T) + 4(G+C)-5. The annealing temperature for each PCR was typically estimated by Tm minus 5^{o}C and optimized experimentally.

PCR amplification was carried out with following reaction parameters

dNTPs	200µM
PCR reaction buffer	1X
Magnesium chloride	1-1.5 mM
Primer (forward)	5 pmoles
Primer (reverse)	5 pmoles
Template DNA	50 ng
Taq DNA polymerase	0.8-1.4U

The final reaction volumes were either 25 or 50 µl. Dimethyl sulphoxide (DMSO) was used at a concentration of 5% to 10% for GC-rich templates.

The thermocycler parameters used were as follows

Initial denaturation	94°C, 5 min
Denaturation	94°C, 30-60 sec
Annealing	55-65 °C, 30-40 sec
Elongation	72°C, 28-45 sec
Cycles	32-35
Final elongation	72°C, 5-7 min

(Tables listing sequences of primers used, annealing temperatures and concentrations of $MgCl_2$ for each primer pair are provided in Appendix I Table-1-9.)

2.2.4 Agarose gel electrophoresis

Agarose gel electrophoresis was used for analytical and preparative purposes. DNA fragments were detected by staining with ethidium bromide and visualised on a UV transilluminator (UVITec, Cambridge, England). The size of a DNA fragment was determined by comparison of its mobility to that of DNA size standards.

Agarose gels (usually in the range of 1-1.5% (w/v) agarose) were prepared by dissolving the required quantity of agarose in 1X Tris-Acetate-EDTA (TAE) electrophoresis buffer by heating in a microwave oven, followed by addition of 0.25 ug/ml ethidium bromide. The agarose solution was poured into a gel tray containing a comb, allowed to cool and solidify, and then placed in a electrophoresis tank and submerged in 1X TAE buffer. The DNA samples were mixed with DNA loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol 40% w/v sucrose). Samples were loaded on the gel along with 100bp ladder DNA size standard. Horizontal

electrophoresis was carried out at approximately 80-100V. The stained gel was photographed under UV light using the UVIDoc gel documentation system (UVITec, Cambridge, England).

2.2.5 Single strand conformation polymorphism (SSCP)

Single-strand conformational polymorphism (SSCP) analysis is a simple and sensitive technique for mutation detection. The sensitivity of SSCP may be up to 80% for fragments less then 300bp in size but decreases with increase in fragment size (Hayashi and Yandell, 1993). The principle of SSCP is based on the fact that single-stranded DNA assumes a conformation due to intra-strand base-pairing. A single base change in the sequence has the potential to lead to an altered conformation of the fragment and result in differential migration under conditions of non-denaturing electrophoresis. Therefore DNA samples having wild-type (normal) and mutant sequences display different band mobility patterns.

For SSCP analysis, 2 µl of PCR product was mixed with 4 µl of 95% formamide dye containing bromophenol blue and xylene cyanol. Samples were denatured at 95^{o}C for 5 min and chilled immediately on ice. Samples were then separated on 8% polyacrylamide gel (19.5:0.5 acrylamide to bisacrylamide) containing 0.5X TBE and 5% glycerol. All samples were electrophoresed at room temperature and at 4^{o}C. Gels were run at constant voltage of 75 V at room temperature and 120V at 4^{o}C. Gels were stained with 0.2% silver nitrate for the detection of DNA. This was done by fixing the gels in 10% ethanol: 0.5% acetic acid for 45 min. Gels were then washed three times with deionised water followed by

staining in 0.2% silver nitrate. They were then washed, developed in 1.5% sodium hydroxide, 0.4% formaldehyde until bands were visible at the desired intensity, washed in deionised water and photographed using a UVIDoc gel documentation system over white light (UVITec, Cambridge, England). Fragments showing altered mobility relative to controls were sequenced directly.

2.2.6 Sequencing

A. Purification of PCR products

PCR products were purified in order to remove primers and dNTPs prior to sequencing on Millipore Ultrafree spin columns (Millipore corporation, Billerica, MA) according to the protocol given by the manufacturer. PCR products (50 µl) were diluted to a final volume of 250 µl using autoclaved deionised water and loaded on the column, which was placed in a microfuge tube and centrifuged at 5000 rpm for 15 minutes. The columns were then placed in an inverted position in fresh microfuge tubes for elution of products. 20 µl of autoclaved deionised water was added to each column followed by centrifugation at 5000 rpm for 2 minutes. Products were checked by agarose gel electrophoresis after column purification.

B. Sequencing

Sequencing of PCR products was carried out by the dideoxy chain-termination method using fluorescent automated sequencing with the Big Dye version 3.1 cycle sequencing kit from Applied Biosystems Incorporated (ABI). Sequencing was

performed in a final reaction volume of 10µl. The reagents used in sequencing reaction are listed in Table 2.1. PCR conditions used for sequencing PCR were: 1 minute at 96°C for initial denaturation, followed by 25 cycles of 10 sec at 96°C, 5 sec at 50°C and 4 minutes at 60°C. This was followed by final hold at 4 °C for 15 minutes.

Products of sequencing reactions were purified by ethanol precipitation after dilution to a final volume of 100 µl with deionised water followed by the addition of 3 µl of 3M sodium acetate (0.09M final concentration) (pH 4.6) and 250 µl of ethanol (this protocol was as described in manual on automated DNA sequencing from ABI). Samples were centrifuged at 12,000 rpm for 20 minutes at room temperature. The pellets were then washed twice with 70% ethanol. Samples were dried and dissolved in 15 µl of formamide (Hi-dye formamide, ABI) and were analyzed on the ABI 310 genetic analyzer.

Table 2.1. List of reagents used for sequencing PCR.

S.No.	Reagent	Volume
1	Big Dye terminator	4.0 µl
2	Forward/Reverse primer	1.0 µl (3.2 pmols)
3	Template	2.0 µl (50 -100 ng)
4	MilliQ water	3.0 µl
5	Total reaction volume	10 µl

2.2.7 PCR-RFLP (Restriction Fragment Length Polymorphism) analysis

PCR-RFLP was used to test for the presence of sequence changes in normal controls and to ascertain segregation of mutations in family members of probands. Restriction enzymes that had recognition sites created or destroyed by mutations or polymorphisms were used for screening for the presence of these variations. Restriction enzyme sites were located on the relevant DNA fragment by using the web-based program from New England Biolabs NEBcutter V2.0 (see Appendix II for URL). PCR products were digested with 2 to 4 units of the restriction enzyme in a final volume of 20 µl using the recommended restriction enzyme buffer. All reactions were incubated overnight at appropriate temperatures. Following digestion, fragments were separated on an 8% non-denaturing polyacrylamide gels, stained with ethidium bromide (0.5µg/ml in 1XTBE) and DNA was visualized on a UV transilluminator (UVITec, Cambridge, England).

2.2.8 Mutational screening

A. Screening of *CHST6*

The coding region of *CHST6* was screened for mutations by direct sequencing of PCR products. Primers for amplification of the coding region and sequencing were described previously by Akama et. al. (2000). The sequences obtained were compared with the published c.DNA sequence of the *CHST6* gene (GenBank NM_021615). DNA samples from 75 unrelated normal control individuals were tested for the presence of identified mutations using restriction enzymes wherever

appropriate (details of restriction enzyme sites are in Table 3.1, 3.2, 3.3 & 3.4- Chapter 3 and 4.3, 4.4, 4.5- Chapters 4). For mutations that did not result in changes in restriction enzyme sites, single strand conformation polymorphism (SSCP) was performed. PCR products were amplified with primers such that they were <300 bp in size for SSCP analysis. In cases where SSCP did not detect the sequence change, direct sequencing was performed.

To evaluate the upstream region of the *CHST6* gene in patients with heterozygous or no mutations in the coding region, PCR was performed on genomic DNA according to the method of Akama et. al. (2000) with the following exceptions: primer F2M was replaced by 5'-CCACAGCCAATTCCA TCTTGGATTTTCTC-3' and primer R2 was replaced by 5'-CATTAGACACCTCACCTGCTTTGGC-3'. The annealing temperatures were adjusted to 58^0C and 62^0C for the primer pairs F2/R2 and F2M/R2, respectively. All amplicons indicating upstream DNA rearrangements were checked by direct sequencing. A new primer, R2M (annealing temperature 61^0C), was introduced to pair with F2M and produced an amplicon on the 3' side of region 'B' (Figure 2.1) (as defined by Akama et al., 2000). The 3' end of this amplicon was sequenced using a forward primer, 5'-GCAGAGGTTGCACACACCTGTC-3'. The details of primers used for screening of coding and upstream region of *CHST6* gene and sizes of PCR products are provided in the Table 1 and 2 of Appendix-I (screening of the *CHST6* gene upstream region was done in the laboratory of Dr. Gordon K. Klintworth).

Structure of *CHST6* gene

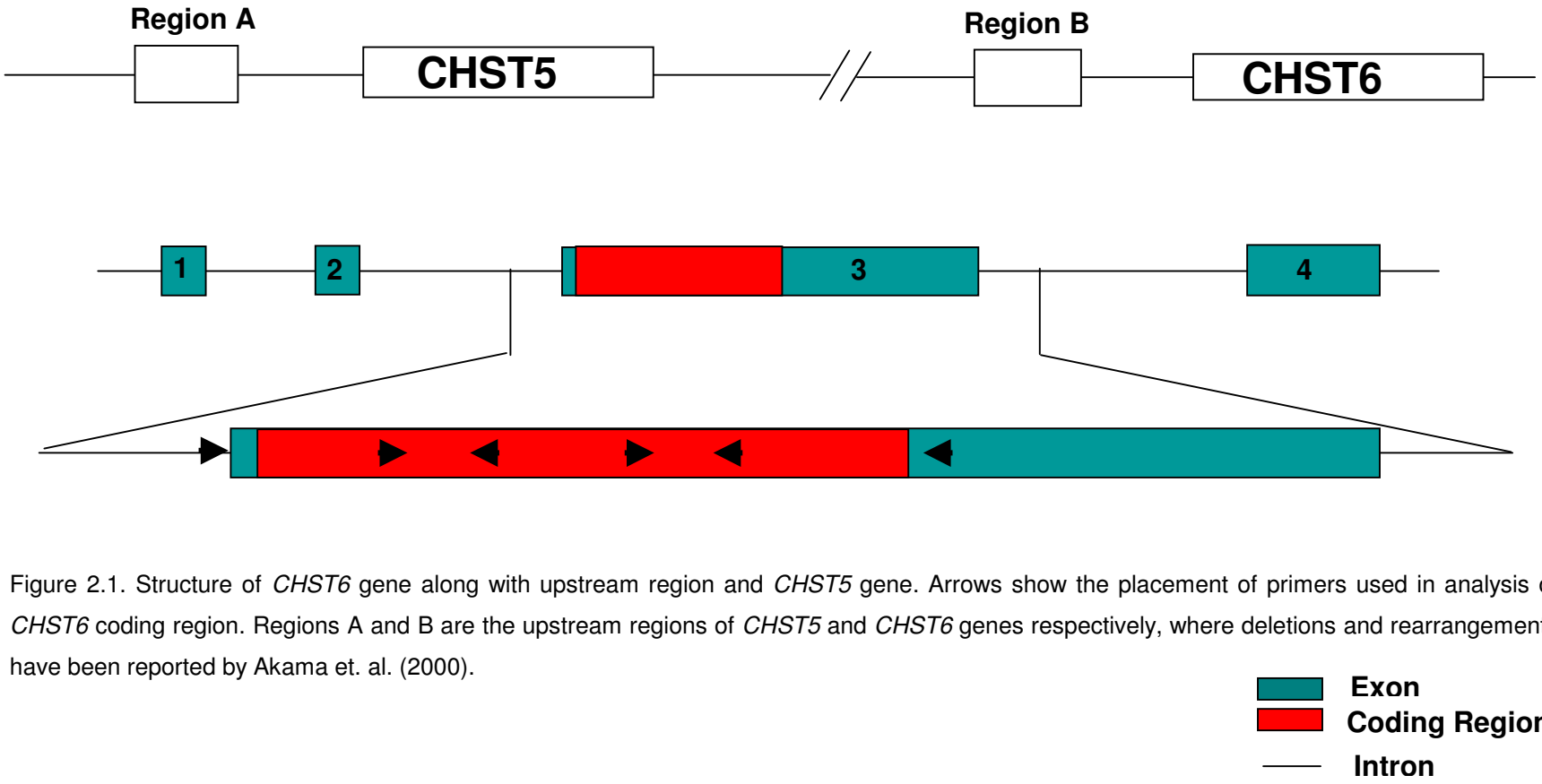


Figure 2.1. Structure of *CHST6* gene along with upstream region and *CHST5* gene. Arrows show the placement of primers used in analysis of *CHST6* coding region. Regions A and B are the upstream regions of *CHST5* and *CHST6* genes respectively, where deletions and rearrangements have been reported by Akama et. al. (2000).

B. Screening of candidate genes on chromosome 20

Thirteen candidate genes in the CHED2 critical region on chromosome 20p13 were screened as part of my work. These are listed below (Table 2.2) and details of primers used for amplification of these 13 genes are listed in Table 3-9, Appendix I. All primers were designed so as to amplify exons and flanking intronic sequences. Screening of the remaining candidate genes (of a total of 40 in the interval) was done by Dr. Hejtmancik and coworkers from National Eye Institute, Bethesda, MD, USA.

Table 2.2. Candidate genes screened in the CHED2 critical region

S.NO	Gene symbol	Gene description
1	TMC2	Transmembrane channel-like 2
2	KIAA1442	KIAA1442 protein
3	MRPS26	Mitochondrial ribosomal protein S26
4	OXT	Oxytocin, prepro- (neurophysin I)
5	AVP	Arginine vasopressin
6	ADRA1D	Alpha-1D adrenergic receptor
7	C20orf116	Chromosome 20 open reading frame 116
8	SLC4A11	Solute carrier family 4, sodium bicarbonate transporter-like, member 11
9	GFRA4	GDNF family receptor alpha 4
10	C20orf27	Chromosome 20 open reading frame 27
11	C20orf28	Chromosome 20 open reading frame 28
12	CENPB	Centromere protein B,
13	LOC440752	LOC440752 - homology to Alu repeat (predicted 175 amino acid peptide)

C. Screening of *SLC4A11*

Screening of exons and flanking intronic sequences of the *SLC4A11* gene was done by SSCP followed by sequencing of PCR amplified-products. Primer sequences and conditions used are listed in Table 9, Appendix I.

2.2.9 Allele-specific PCR

Allele-specific PCR refers to selective amplification of an allele while excluding other alleles differing even by a single base. Selective amplification is achieved by designing a primer such that the variant base is present at the 3' end of the primer.

Primers specific for mutant and wild-type sequences were designed for patient MCD-39 who had two consecutive heterozygous base changes [c.293C>T; c.294C>G]. Allele-specific PCR was performed to know whether the two sequence changes are in *cis* or *trans*. PCR was done with the patient and normal control DNAs at an annealing temperature of 57^0C. As an additional control, DNA from an MCD patient (MCD-10) having the change [c.293C>G; c294C>G] was used. This sequence differs from that of MCD-39 by one base in the region tested. Details of allele-specific primers used for mutant and wild type sequences of *CHST6* are provided in Table 2.3. The residues in bold at the 3' end of forward primers are the bases that are mutated.

Table 2.3. List of primers for allele specific PCR (*CHST6*).

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Name	**Sequence (5'-3')**			
1	CHST-ASOF1[Wild type forward]	GCGACCTGGTGCGCTCC	57	295	3
2	CHST-ASOF2 [Mutant forward]	GCGACCTGGTGCGCTTG			
3	CHST-ASOR [Common reverse]	GGGTCGCTGAGCAGCGG			

2.2.10 Genotyping

A total of 13 fluorescently labeled microsatellite markers (Table 2.3) spanning a region of 24 cM on chromosome 20p13 were selected for mapping. Primers for amplification of these microsatellites were commercially synthesized with fluorescent ABI-compatible dyes (Fam, Hex, Tet) attached to one primer of each pair. PCR reactions were carried out in 10µl volume containing 25ng of genomic DNA, 5 pmol each of forward and reverse primers, 1.0-1.5mM magnesium chloride, 1X PCR buffer (100mM Tris (pH9.0), 500mM potassium chloride, 0.1% gelatin), 200µM dNTPs and 1unit of AmpliTaq Gold DNA polymerase (Applied Biosystems Incorporated, California United States). Markers used, primer sequences and labels are detailed in table 2.4.

Table 2.4. Microsatellite markers and details of primers used in linkage and haplotype analysis.

S.NO	Microsatellite marker	Primer Sequence (5'-3')	Allele size range	Labeled with fluorescent dye
1	D20S117	GAACCAACCCTGCTGC TCCAGAGATCACCCCC	151-157	TET
2	D20S906	TGAGCCGAGATCGAGC CCACTTCAGAGTGACCCA	229-247	HEX
3	D20S198	AGTGAGCCCAGGTCGT TAGGAACATCATGTAACCCAGA	225-243	FAM
4	D20S193	GCATCCCTGGGCTAACTG GGAATCTTTTGGTGCGTG	136-146	TET
5	D20S889	GGTTTGGTGGAATCCTCTC CATCTTTCAAATGGGATAATGG	262-292	FAM
6	D20S97	GGGAGATGGAGAGTTGTTGC TATACGCACCCATACCACCA	262-268	FAM
7	D20S895	CCCAGGGAGGTAAAGGTT GTCAGGCTACATCAGCAAA	186-226	FAM
8	D20S835	ACAAACCTGCATGTCCT ATGGCTCAAGCCTTCA	196-222	TET
9	D20S882	ATGTAGGACAAATAGCCCG CTGTGCCTCTCAACAACTG	120-130	HEX
10	D20S905	AGCTTGAGGAGCAGTGTCTT TCAGCAGATCCCACCA	87-123	TET
11	D20S194	CCAGGATTCTTTCAGGCT GCCAGAGTCCAACGCT	178-224	HEX
12	D20S115	CCGGAGTTAGTTTTATTATCTTGC CCTGTCCAGAACTGTCATATTG	236-246	HEX
13	D20S177	AGCAATGAGTAAACCTGCCT CCACCTATCCACCTATGGTATG	94-102	FAM

The table summarizes the 13 markers used for mapping along with primer sequence, allele size range and the fluorescent dye used for labeling.

PCR conditions used were: 12 minutes at 95^{o}C for initial denaturation, followed by 10 cycles of 15 sec at 94^{o}C, 15 sec at 55^{o}C and 30 sec at 72^{o}C. This was followed by 20 cycles with 15 sec denaturation at 89^{o}C, 15 sec annealing at 55^{o}C , 30 sec extension at 72^{o}C and then one cycle of final extension at 72^{o}C for 10 minutes.

2.2.11 Linkage analysis

Two point linkage analyses were performed with the MLINK program of the LINKAGE program package (Cottingham et al., 1993; Lathrop and Lalouel, 1984) modeling CHED2 as a fully penetrant autosomal recessive disease. Disease gene frequency used was 1:100,000. Equal allele frequencies were used for all markers. Ten consanguineous families fulfilling all diagnostic criteria (13 affected and 24 unaffected individuals) were included in the linkage analysis.

2.3 IMMUNOPHENOTYPING

Immunophenotyping for MCD was done by estimating levels of antigenic keratan sulfate (AgKS) in corneal sections by immunohistochemistry and in serum by enzyme-linked immunosorbent assay (ELISA).

2.3.1 Immunohistochemistry

Immunohistochemistry is a technique for identifying antigens in cells or tissue sections by means of reactivity with a specific antibody.

Immunohistochemistry of corneas from MCD patients was performed on paraffin-embedded formalin-fixed corneal sections of four microns thickness. A monoclonal antibody directed against sulfated epitopes present on keratan sulfate (1/20/5-D-4 mouse IgG, ICN Pharmaceuticals, Inc, Irvine, California) was used for detection of AgKS in cornea. The secondary antibody used was biotinylated anti-mouse IgG (Dakocytomation Denmark). Detection was done by Avidin-biotin complex immunoperoxidase technique, in which avidin-conjugated peroxidase binds to

biotinylated antibody, followed by detection of peroxidase after addition of suitable substrate [3,3'-diaminobenzidine tetrahydrochloride, [DAB] (Sigma-Aldrich, USA)] (Hsu et al., 1981). The monoclonal antibody 1/20/5-D-4 is specific for sulfated keratan sulfate. The primary antibody was used at a dilution of 1:700, which was determined as the optimum dilution by testing different antibody dilutions using corneas from positive (normal) controls that included patients with AR CHED or GCD (granular corneal dystrophy), in which keratan sulfate is expected to be normal. For negative control, the primary antibody was omitted and phosphate-buffered saline was used. Interpretations of immunohistochemical assays were confirmed independently by three investigators, each masked to the findings of the others.

2.3.2 Estimation of serum AgKS

Serum antigenic keratan sulfate levels were determined in 57 MCD patients and two unaffected family members, by the method of ELISA using the 1/20/5-D4 antikeratan sulfate monoclonal antibody. This was done by Dr. E. J. Thonar and coworkers at Rush University Medical Center, Chicago, Illinois, USA.

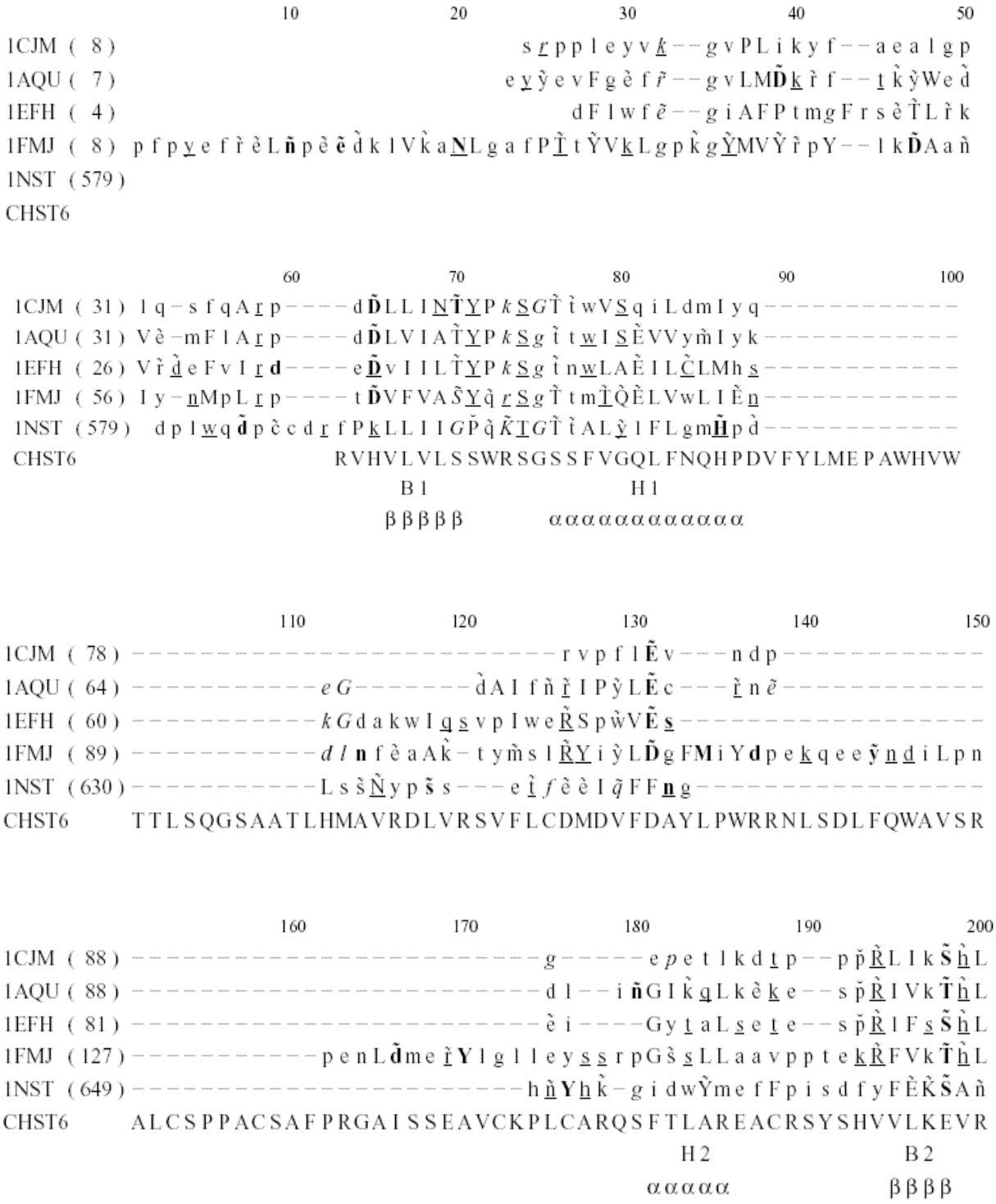
2.4 COMPARATIVE MODELING OF cGlnNAc6ST

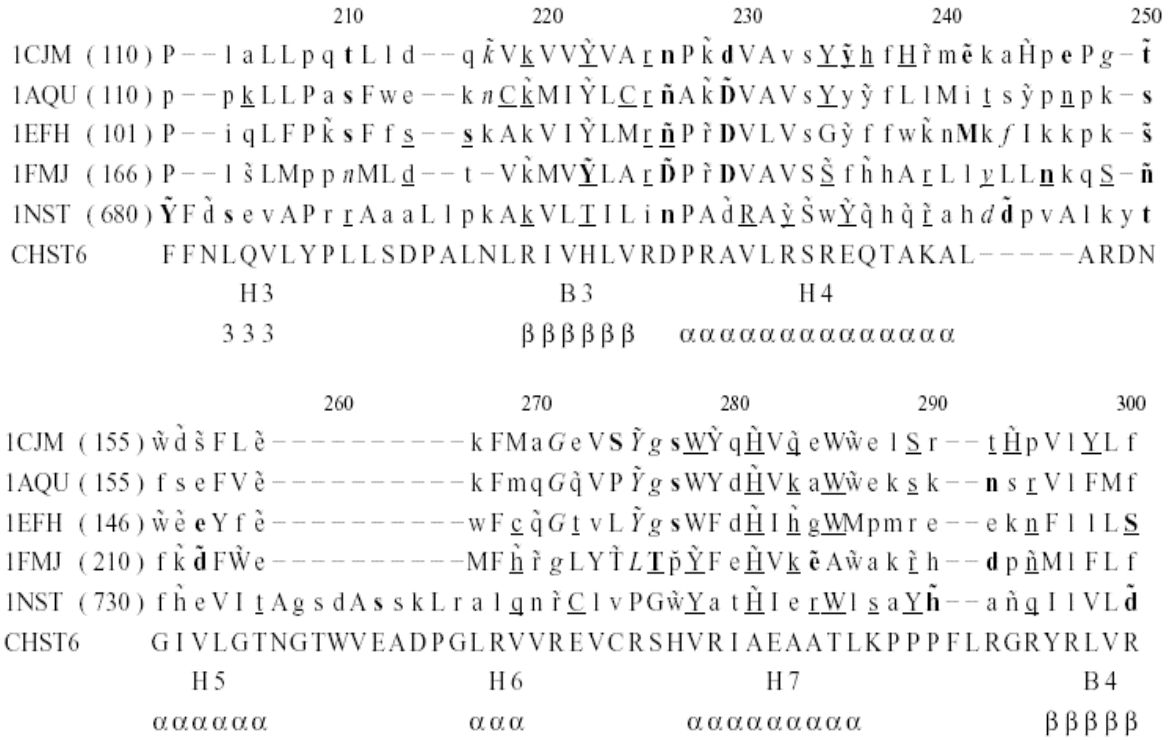
Comparative or homology protein structure modeling builds a three-dimensional model for a protein of unknown structure (the target) based on one or more related proteins of known structure (the templates) (Blundell et al., 1987; Johnson et al., 1994; Fiser et al., 2002).

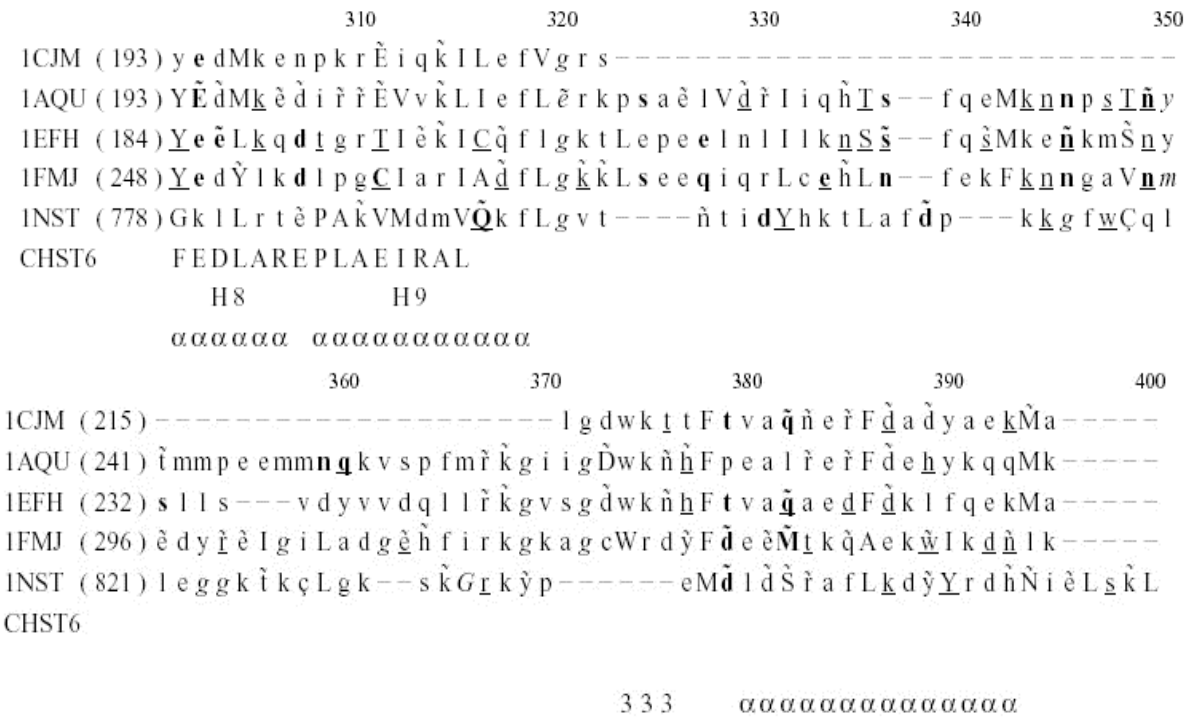
Homologues of corneal N-acetylglucosamine-6-sulfotransferase were obtained from Protein Data Bank (Berman et al., 2000) a database of proteins of known 3-D structures comprising about 40,000 structures, using BLAST with e-value cut-off of 0.01. No hits were picked up in this search, suggesting no closely related homologs of *CHST6* with known 3-D structures are present in the PDB. The nr database (a non-redundant database of all known proteins of known sequence) was then searched, using psiblast (BLASTP 2.2.4 [2002], (Altschul et al., 1997) with e-value and h-value cut off of 0.001. This search picked up homologs of known structure at round 4 of psiblast with e-values between 0.0001 and 0.003. Percentage identity between the query and the homologs of known structure were in the range of 14 to 18 percent. Relatedness of the putative homologous members of sulfotransferase family of known structure was ensured by consulting the database of Structural Classification of Proteins (SCOP) (Murzin et al., 1995). Multiple structural superposition of these homologs of known 3-D structure was represented using the program Joy (Overington et al., 1990) (Figure-2.2). Closely related homologs of *CHST6* with sequence identity greater than 35%, although they did not have known 3-D structures, were picked up from the 15th round of psiblast (BLASTP 2.2.4) (Altschul et al., 1997). A multiple sequence alignment of these homologs was created using ClustalW1.82 (Thompson et al., 1994). The secondary structure of *CHST6* protein was predicted by the automated e-mail server PHD (Rost et al., 1994). From structurally known homologs, retinol dehydratase, (Pakhomova et al., 2001) was chosen as the template for building a 3-D model of *CHST6* as it shows maximum coverage- (it has comparatively less insertions or deletions and hence covers

maximum sequence of *CHST6* protein as compared to other homologous protein sequences) of *CHST6* protein sequence. The suite of programs encoded in COMPOSER (Sutcliffe et al., 1987; Srinivasan and Blundell, 1993) of "SYBYL" was used to generate the 3D model of *CHST6*. The structures of the conserved regions of retinol dehydratase have been extrapolated to the equivalent regions of *CHST6*. The variable regions were modeled by identifying suitable segments from known structures. A template matching approach (Topham et al., 1993) to rank the candidate loops was also used. The best ranking loop with minimum short contacts with the rest of the protein was modeled. The COMPOSER generated models were energy minimized in SYBYL using the AMBER force field (Pearlman et al., 1991). During the initial cycles of energy minimization the backbone was kept rigid and side chains alone were moved. Subsequently all atoms in the structure were allowed to move during minimization. This approach creates least perturbation in the backbone structure while the short contacts and inconsistencies in geometry are rectified. The 3-D model of *CHST6* was superimposed on the structure of retinol dehydratase with topologically equivalent Cα atoms of functional residues and cofactor using the Super (B.S. Neela, unpublished). Super performs rigid body superposition of the user specified equivalent atoms of two molecules. The mutations on the 3-D model of *CHST6* with bound cofactor were further analyzed using interactive graphics program SETOR (Evans, 1993). Inferences on the probable affects of mutations were drawn by visual inspection by consideration of distances between putative interacting atoms.

Figure 2.2


10 20 30 40 50
1CJM (8)
1AQU (7)
1EFH (4)
1FMJ (8)
1NST (579)
CHST6
60 70 80 90 100
1CJM (31)
1AQU (31)
1EFH (26)
1FMJ (56)
1NST (579)
CHST6
B1
H1
110 120 130 140 150
1CJM (78)
1AQU (64)
1EFH (60)
1FMJ (89)
1NST (630)
CHST6
160 170 180 190 200
1CJM (88)
1AQU (88)
1EFH (81)
1FMJ (127)
1NST (649)
CHST6
H2
B2





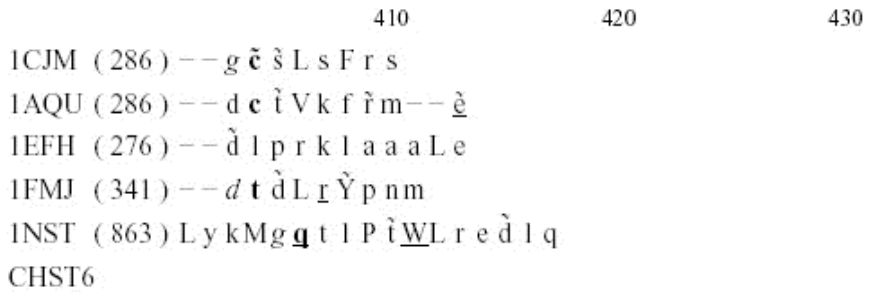


Figure 2.2. Multiple structural superposition of homologs of known 3-D structure using the programme JOY. Structure-based alignments of sulfotransferases of known structure integrated with the alignment of the sequence of *CHST6*. The first 5 proteins in an alignment block correspond to those with known 3-D structures and the last sequence corresponds to *CHST6* whose structural model has been generated by comparative modeling. H1, H2… and B1, B2… refer to the conserved helical and β-strand regions. The structural environments at the amino acid residues in the known structure are indicated. This figure has been generated using the program JOY. (CJM-Catecholamine sulfotransferase, AQU-Estrogen sulfotransferase, EFH-Human hydroxy steroid sulfotransferase, FMJ-Retinol dehydratase, NST-Human heparinsulfate-N-deacetylase/N-sulfotransferase).

Key to JOY alignments

solvent inaccessible	UPPER CASE	X
solvent accesible	lower case	x
positive ϕ	*italic*	*x*
cis−peptide	breve	x̆
hydrogen bond to other sidechain	tilde	x̃
hydrogen bond to mainchain amide	**bold**	**x**
hydrogen bond to mainchain carbonyl	underline	x
disulphide bond	cedilla	ç

2.5 MULTIPLE SEQUENCE ALIGNMENT

Multiple sequence alignment was done to compare protein sequences of different species to assess conservation of the residues identified to have mutations. Homologous protein sequences were obtained in the FASTA format from the NCBI database (see URL Appendix II). Multiple sequence alignment was performed using CLUSTALW software, available at the European Bio-informatics Institute (EBI URL, see Appendix II). Alignment was performed using score type as absolute and values of remaining parameters were default settings in the program.

2.6 CLINICAL AND HISTOPATHOLOGICAL CORRELATIONS

In order to look for genotype-phenotype correlations in AR CHED, clinical and histopathologic features were analyzed in probands and affected relatives in whom mutations were identified. The clinical features included age at onset of disease and presentation at our institution, pre-and post-operative visual acuities and presence of nystagmus. The histopathologic features included quantitative assessments of thickness of cornea, Descemet's membrane and endothelial cell counts. PAS-stained corneal sections of 33 patients were analyzed. Thickness of cornea, Descemet's membrane and endothelial cell counts were estimated on four micron thick corneal sections by image capture and analysis using Axiovision digital imaging software (Axiovision AC Rel4.5) from Carl-Zeiss AG (Hallbergmoos, Germany). Corneal thickness (at the center) was measured at 4X magnification, and the thickness of DM was measured at 100X magnification. For each specimen, thickness was measured at three points (for both cornea and DM) and the average reading was taken for each specimen. Endothelial cells were counted at a magnification 40X along a 6 mm

length of Descemet's membrane on one corneal section (of 4 μ thickness) from each patient.

Under identical conditions, quantitative assessments of thickness of cornea, Descemet's membrane and endothelial cell count was carried out on 20 age-matched normal corneas. The normal corneas with age groups 0-5 yrs, 5-10 yrs, 11-15 yrs, 16-20 yrs and >20 yrs were obtained from histopathology specimens of patients with retinoblastoma, uveal melanoma, malignant melanoma and cryptococcal granulomatous lesion of choroid.

3.1 INTRODUCTION

This chapter describes the molecular genetic analysis and immunophenotyping of patients with MCD, which is one of the most frequent corneal dystrophies seen at our institution, a tertiary care referral centre**.** The *CHST6* gene was identified as the cause of MCD by Akama et al (Akama et al., 2000). There was no reported data on molecular genetic aspects on MCD in India prior to this study and only few studies had been done in other populations. In order to understand the underlying molecular basis for MCD in Indian patients, the first aspect of this study was a mutational analysis of *CHST6* gene in MCD patients.

The second aspect covered in this chapter is the analysis of *CHST6* mutations through the approach of comparative three-dimensional modeling for wild type and mutant corneal N-acetyl glucosamine 6-sulfotransferase (CHST6; cGlcNAc6ST) proteins. This was done in order to assess the impact of missense mutations, which predominated the mutational pattern of *CHST6* in MCD.

Clinically although MCD is homogenous, three immunophenotypes I, IA, and II, are recognized based on the immunohistochemical reactivity of the patient's cornea and serum to antibody specific for antigenic KS (AgKS). The underlying molecular basis for the different immunophenotypes is not understood. A third aspect of this study was to determine immunophenotypes of MCD patients on whom genetic analyses were done, and to look for correlations between mutations and immunophenotypes. Immunophenotyping was done by estimating the levels of antigenic keratan sulfate in cornea and serum of MCD patients recruited.

3.2 SCREENING OF THE *CHST6* GENE

Fifty-nine probands with a diagnosis of MCD (72 patients and 6 unaffected family members) were screened for mutations in the *CHST6* gene as described in Chapter 2. Mutations were found in 57 probands (Tables 3.1, 3.2, 3.3 & 3.4). The mutations identified consisted of 25 missense, 6 nonsense, 5 deletion, 2 insertion and 2 complex mutations consisting of deletion and insertion. Sequence changes identified in patients or probands were tested for co-segregation in available family members (more than one was available and recruited for this study in 10 families). In addition, a control population of 75 unrelated individuals of Indian origin with no history of corneal disease was screened for each of the sequence changes found in patients in order to determine whether the identified variations are pathogenic mutations or polymorphisms. This was done by RFLP, SSCP, or direct sequencing as described in Chapter 2. Both mutant alleles were identified in 50 probands. Two probands from families MCD-35 and MCD-46 had compound heterozygous mutations. Probands from three families (MCD-10, MCD-20 and MCD-52) had two heterozygous mutations each. However it was not possible to determine whether the two alleles in these families were in *cis* or in *trans* due to unavailability of family members and lack of suitable RFLPs in the mutant or normal alleles. Two probands had single heterozygous mutations in the coding region of *CHST6*. These were MCD-9 & MCD-39. No mutations were found in two probands, MCD-58 & MCD-59.

3.2.1 Deletion mutations

Five deletion mutations were identified in 8 probands (Figure 3.1) of which 3 were not reported prior to this study. Three homozygous deletions found in 6 families were- a 25 bp deletion at cDNA positions c.16_40 (Table 3.1) predicting a frameshift at valine-6 followed by termination after 55 amino acids of an altered reading frame, a deletion of 7 bp (c.DNA position c.94_c.100) predicting a frameshift at serine-32 with termination after 36 amino acids, and deletion of a single base (c.180delC) predicting a frameshift at phenylalanine-60 with termination after 10 amino acids (Table 3.1).

One heterozygous deletion of a single base (c.198delC) predicting a frameshift at phenylalanine-67 and immediate termination of protein at amino acid 70 was found in family MCD-9. In this family, the second mutation has not been identified upon sequencing of the coding and upstream region of *CHST6* gene.

One compound heterozygous mutation consisting of a deletion and a missense mutation was found in MCD-46 (Table 3.1). The first mutant allele consists of a deletion of single base (c.545delA) leading to frameshift at glutamine-182 and termination after 199 amino acids. The second mutant allele consists of a missense mutation resulting in substitution of leucine-276 for proline. These two mutations were found to be in *trans* by segregation analysis of these mutant alleles in the proband's mother (unaffected) and maternal cousin (affected). The proband's mother was heterozygous for leucine-276 to proline mutation. The deletion mutation (c.545delA) was absent in the mother and would therefore have come from the father. There was no consanguinity in this family. The proband's cousin sister (the proband's maternal

uncle's daughter) also an MCD patient, showed a homozygous missense mutation of leucine-276 to proline. Based on the analysis of these two family members it can be concluded that the two mutant alleles were in *trans* and the proband of MCD-46 family had a compound heterozygous mutation.

Table 3.1. Deletion mutations identified in *CHST6* gene among patients with MCD.

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
1	MCD-1	c.16_40del	p.Val 6 ProfsX55	-	This study
2	MCD-33	c.16_40del	p.Val 6 ProfsX55	-	This study
3	MCD-3	c.94_100del	p.Ser32GlnfsX36	-	This study
4	MCD-8	c.180delC	p.Phe60LeufsX10	*Mse*1 (+)	This study
5	MCD-36	c.180delC	p.Phe60LeufsX10	*Mse1 (+)*	This study
6	MCD-37	c.180delC	p.Phe60LeufsX10	*Mse1 (+)*	This study
7	MCD-9*****	c.[198delC]*****+[?]	p.[Phe67SerfsX3]+[?]	-	Warren et al., 2003
8	MCD-46[#]	c. [545delA]+[827T>C]	p. [Gln182ArgfsX199]+[Leu276Pro]	*Pst* I (-)	Warren et. al., 2003; Aldave et al., 2004

This table summarizes the deletion mutations identified in *CHST6* gene in probands from 8 families. Mutation nomenclature is according to the current recommendations in http://www.hgvs.org/mutnomen/recs.html. Numbering is according to the cDNA (GenBank NM_021615) with A of the ATG translation initiation codon of *CHST6* as +1. *Family MCD-9 had a single heterozygous mutant allele. [#]Family MCD-46 had compound heterozygous mutation. The signs (+) and (-) indicate the gain or loss of restriction sites respectively.

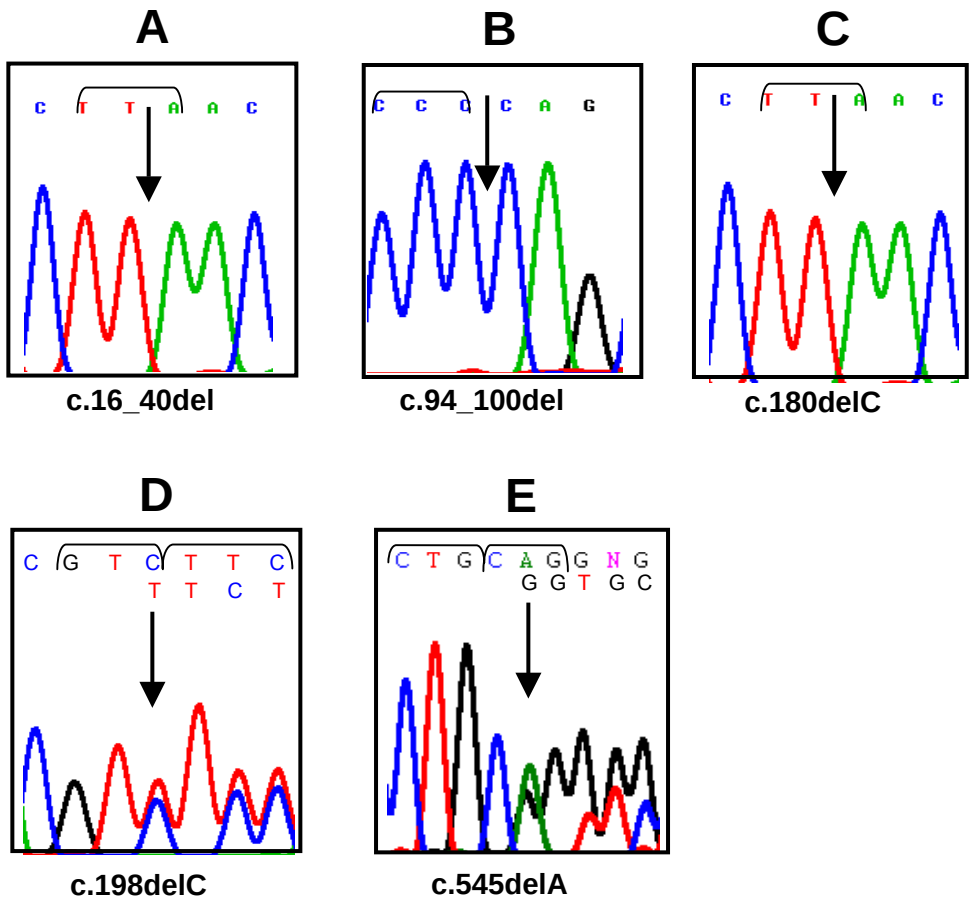


Figure 3.1. Sequence electropherograms of deletion mutations. Figures show homozygous deletions identified in families MCD-1 & MCD-33 (A), MCD-3 (B), MCD-8, MCD-36 & MCD-37 (C). Figures D & E show heterozygous 1 base deletions identified in families MCD-9 and MCD-46 respectively. Sequence of the normal allele is shown below that of the mutant allele. Arrows indicate the sites of deletion. Codons are marked by bracketts. Details are in Table 3.1.

3.2.2 Complex mutations

Two complex mutations consisting of deletion plus insertion, both homozygous, were identified in 3 probands (Figure 3.2 A & B). The first mutation was a deletion of GCG at nucleotide c.612 and an insertion of AT, seen in MCD-19, which is predicted to result in a frameshift at arginine-205, with 176 amino acids of altered reading frame. A second homozygous complex mutation consisting of a deletion of 10 bp and insertion of TTG at nucleotide c.1002 was found in two probands from MCD-56 & MCD-57. This mutation is predicted to produce a frameshift at histidine-335 followed by altered reading frame of 27 amino acids.

3.2.3 Insertion mutations

Two in-frame insertions of a single amino acid were found in 7 probands (Figure 3.2C & D). A homozygous insertion of CTG at nucleotide c.656 corresponding to an insertion of tyrosine after alanine-219 (Figure 3.2C) was identified in 6 families (MCD-22, MCD-23, MCD-24, MCD-25, MCD-50 and MCD-51). The second insertion was c.587insACG predicting the insertion of an arginine residue after leucine-195 in proband MCD-16 (Figure 3.2D). Nucleotides 585-587 of the wild type *CHST6* cDNA have the sequence ACG, and hence this insertion could have occurred at either position 584 or at 587, with the same consequence.

Table 3.2. Complex and insertion mutations identified in *CHST6* gene among patients with MCD.

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
1	MCD-19	c.612_614del3insAT	p.Arg205TrpfsX176	*Cfr421 (-)*	Warren et al., 2003
2	MCD-56	c.1002_1012delinsTTG	p.His335CysfsX27	*Nla* III (-)	This study
3	MCD-57	c.1002_1012delinsTTG	p.His335CysfsX27	*Nla* III (-)	This study
4	MCD-16	c.587insACG	p.Arg195-196ins	-	This study
5	MCD-22	c.656insCTG	p.Trp219-220ins	-	This study
6	MCD-23	c.656insCTG	p.Trp219-220ins	-	This study
7	MCD-24	c.656insCTG	p.Trp219-220ins	-	This study
8	MCD-25	c.656insCTG	p.Trp219-220ins	-	This study
9	MCD-50	c.656insCTG	p.Trp219-220ins	-	This study
10	MCD-51	c.656insCTG	p.Trp219-220ins	-	This study

The table summarizes the homozygous complex and insertion mutations identified in *CHST6* gene in probands from 10 families. Restriction enzymes with sites altered as a result of mutation are shown with minus sign indicating the loss of restriction site.

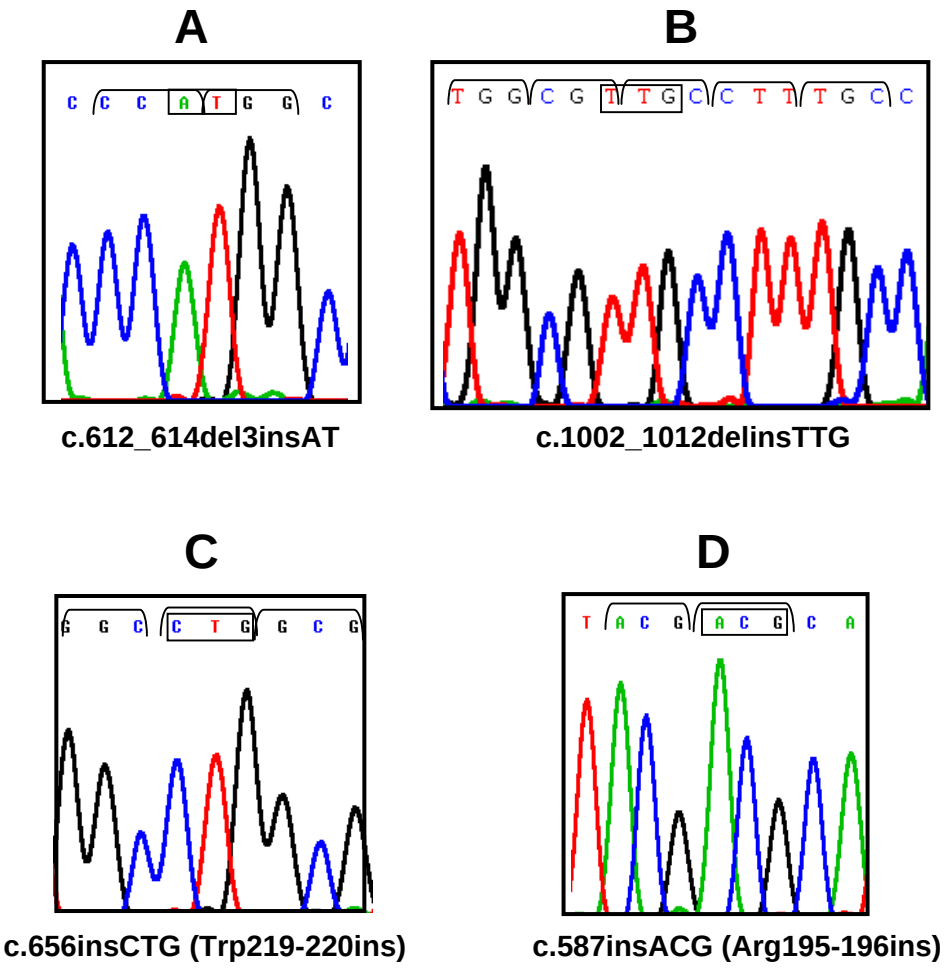


Figure 3.2. Sequence electropherograms of homozygous complex and insertion mutations. Complex mutations identified in families MCD-19 (A), MCD-56 & MCD-57 (B) are shown. In-frame insertion of CTG identified in families MCD-22, MCD-23, MCD-24, MCD-25, MCD-50 and MCD-51 (C) and ACG in family MCD-16 (D). Boxes indicate the residues inserted. Codons are bracketted. Details are in Table 3.2.

3.2.4 Nonsense mutations

Six homozygous nonsense mutations were identified in 9 probands (Figure 3.3). A homozygous nonsense mutation at tryptophan-2 (Trp2X) was found in patient from family MCD-32 (Table 3.3). In addition to the nonsense mutation, this patient showed an additional homozygous missense mutation at cDNA position c.699C>A resulting in the substitution of leucine-3 for methionine. The presence of the Leu3Met change raises the possibility that it creates an initiation site after the truncation at Trp2 in the same reading frame as the wild type *CHST6* protein, giving rise to mutant protein that lacks the first three amino acids.

The remaining five nonsense mutations (Table 3.3) involve the formation of stop codons at glutamine-18 (Gln18X) in one proband (MCD-2), tryptophan-123 (Trp123X) in two probands (MCD-14 and MCD-40), cysteine-153 (Cys153X) in two probands (MCD-15 and MCD-42), glycine-309 (Gly309X) in one proband (MCD-29) and glutamic acid-347 (Glu347X) in two probands (MCD-30 and MCD-31).

Table 3.3. Nonsense mutations identified in *CHST6* gene among patients with MCD.

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
1	MCD-32	c. [6G>A; 7C>A] +[6G>A; 7C>A]	p.[Trp2X; Leu3Met]+ .[Trp2X; Leu3Met	-	This study
2	MCD-2	c.52C>T	p.Gln18X	*Xmi*1 (+)	This study
3	MCD-14	c.369G>A	p.Trp123X	*Cfr*131 (-)	This study
4	MCD-40	c.369G>A	p.Trp123X	*Cfr*131 (-)	This study
5	MCD-15	c.459C>A	p.Cys153X	-	This study
6	MCD-42	c.459C>A	p.Cys153X	-	This study
7	MCD-29	c.925G>T	p.Gly309X	*Tspr*1 (+)	This study
8	MCD-30	c.1039G>T	p.Glu347X	-	This study
9	MCD-31	c.1039G>T	p.Glu347X	-	This study

The table summarizes the nonsense mutations identified in the *CHST6* gene in 9 families with MCD. Restriction enzyme sites created or destroyed as a result of the mutations are shown with the signs (+) and (-) indicating the gain or loss of restriction sites respectively.

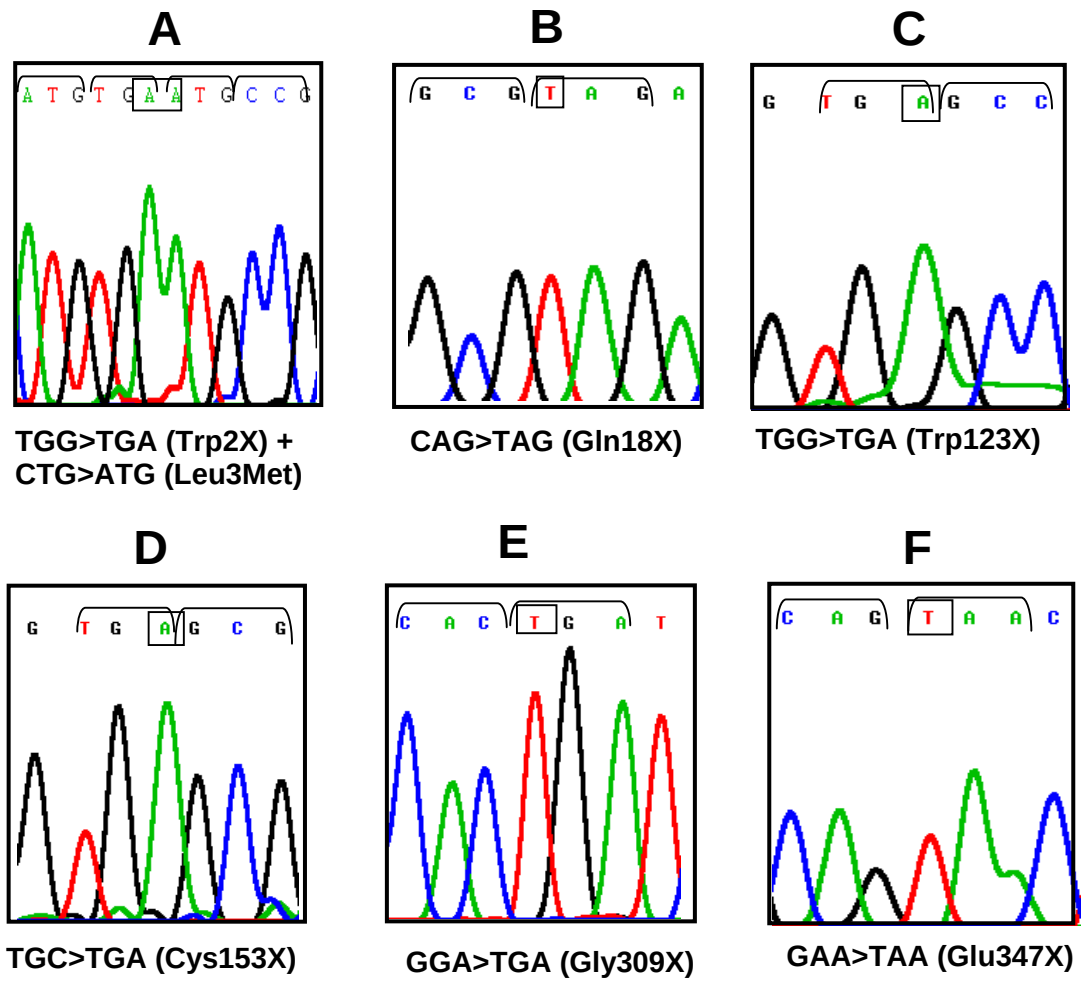


Figure 3.3. Sequence electropherograms of homozygous nonsense mutations with codon and amino acid changes shown below each panel. A**:** Nonsense & missense mutations identified in family MCD-32. Nonsense mutations identified in families MCD-2 (B), MCD-14 & MCD-40 (C), MCD-15 & MCD-42 (D), MCD-29 (E), MCD-30 & MCD-31 (F). Boxes indicate the mutated residues. Codons are marked by bracketts. Details are in Table 3.3.

3.2.5 Missense mutations

A total of 25 missense mutations were identified in 31 probands. Mutations are listed in Table 3.4. Twenty-six probands were homozygous, one was heterozygous for one mutant allele, and one was compound heterozygous for two mutant alleles. In the remaining three probands, two heterozygous mutant alleles were identified but the status of alleles was unknown due to inability to determine whether they were in *cis* or in *trans*.

Nineteen missense mutations were not reported previously and sequences of these are shown in Figure 3.3A and 3.4. The most frequently occurring missense mutation was Asp221Glu, which was found in five families. Other mutations found in 2 or more families were Gly52Asp, Ser53Leu, Phe107Ser, Arg202Ser, Ser210Phe and Asp221Tyr. The remaining mutations shown in Table 3.4 were found in one family each.

Table 3.4. Missense mutations identified in *CHST6* gene among patients with MCD.

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
1	MCD-4	c.155G>A	p.Gly52Asp	*Hin*f1 (+)	This study
2	MCD-5	c.155G>A	p.Gly52Asp	*Hin*f1 (+)	This study
3	MCD-6	c.158C>T	p.Ser53Leu		Warren et al., 2003
4	MCD-7	c.158C>T	p.Ser53Leu		Warren et al., 2004
5	MCD-34	c.161C>T	p.Ser54Phe	*Mbo* II (+)	This study
6	MCD-35	c. [166G>A;167T>G] + [500C>T]	p.[Val56Arg]+ [Ser167Phe]	*Mbi* I (-)+ *Eco*57I (+)	This study
7	MCD-38	c.217G>A	p.Ala73Thr	*Bsh*1236I (-)	This study
8	MCD-10[#]	c. [293C>G;294C>G(+) 320T>C]	p.[Ser98Trp(+) Phe107Ser]		This study
9	MCD-39*	c. [293C>T;294C>G]+ [?]	p.[Ser98Leu]+[?]		This study
10	MCD-11	c.320T>C	p.Phe107Ser		El-Ashry et al., 2002
11	MCD-12	c.320T>C	p.Phe107Ser		El-Ashry et al., 2003
12	MCD-13	c.363C>G	p.Phe121Leu	*Bse*N1 (-)	This study
13	MCD-41	c.391T>C	p.Ser131Pro		Niel et al., 2003

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
14	MCD-43	c. [494G>C;495C>T]	p.Cys165Ser		Bao et al., 2001
15	MCD-44	c.495C>G	p.Cys165Trp		This study
16	MCD-45	c.533T>G	p.Phe178Cys		This study
17	MCD-47	c.578T>C	p.Leu193Pro		This study
18	MCD-48	c.604 C>A	p.Arg202Ser	*Hph1 (+)*	This study
19	MCD-49	c.611C>G	p.Pro204Arg	*Cfr*421 (-)	This study
20	MCD-18	c.611C>A	p.Pro204Gln	*Cfr*421 (-)	Niel et al., 2003
21	MCD-52[#]	c. [629C>T(+)663C>G]	p. [Ser210Phe (+)Asp221Glu]		This study
22	MCD-21	c.661G>T	p.Asp221Tyr	*Nmuc1* (-)	This study
23	MCD-53	c.661G>T	p.Asp221Tyr	*Nmuc1* (-)	This study
24	MCD-54	c.814C>A	p.Arg272Ser	*Alu* I (+)	This study
25	MCD-55	c.1000C>T	p.Arg334Cys	*Hae*II (-)	This study
26	MCD-20[#]	c. [629C>T(+)663C>G]	p. [Ser210Phe (+)Asp221Glu]		This study
27	MCD-26	c.663C>G	p.Asp221Glu		This study
28	MCD-27	c.663C>G	p.Asp221Glu		This study
29	MCD-28	c.663C>G	p.Asp221Glu		This study

S.N0	Family#	Mutation in cDNA	Mutation in protein	Restriction site change	Reported by
30	MCD-17	c.604 C>A	p.Arg202Ser	*Hph*1 (+)	This study
31	MCD-32	c. [6G>A;7C>A] +[6G>A;7C>A]	p. [Trp2X;Leu3Met]+ [Trp2X; Leu3Met]		This study

This table shows the missense mutations identified in the *CHST6* gene in probands from 31 families. The signs (+) and (-) indicate the gain or loss of restriction sites respectively. # Families with two mutations but with unknown alleles. *Families with single heterozygous changes. Family MCD-32 has homozygous missense mutation in addition to a nonsense mutation and hence is also included in this table.

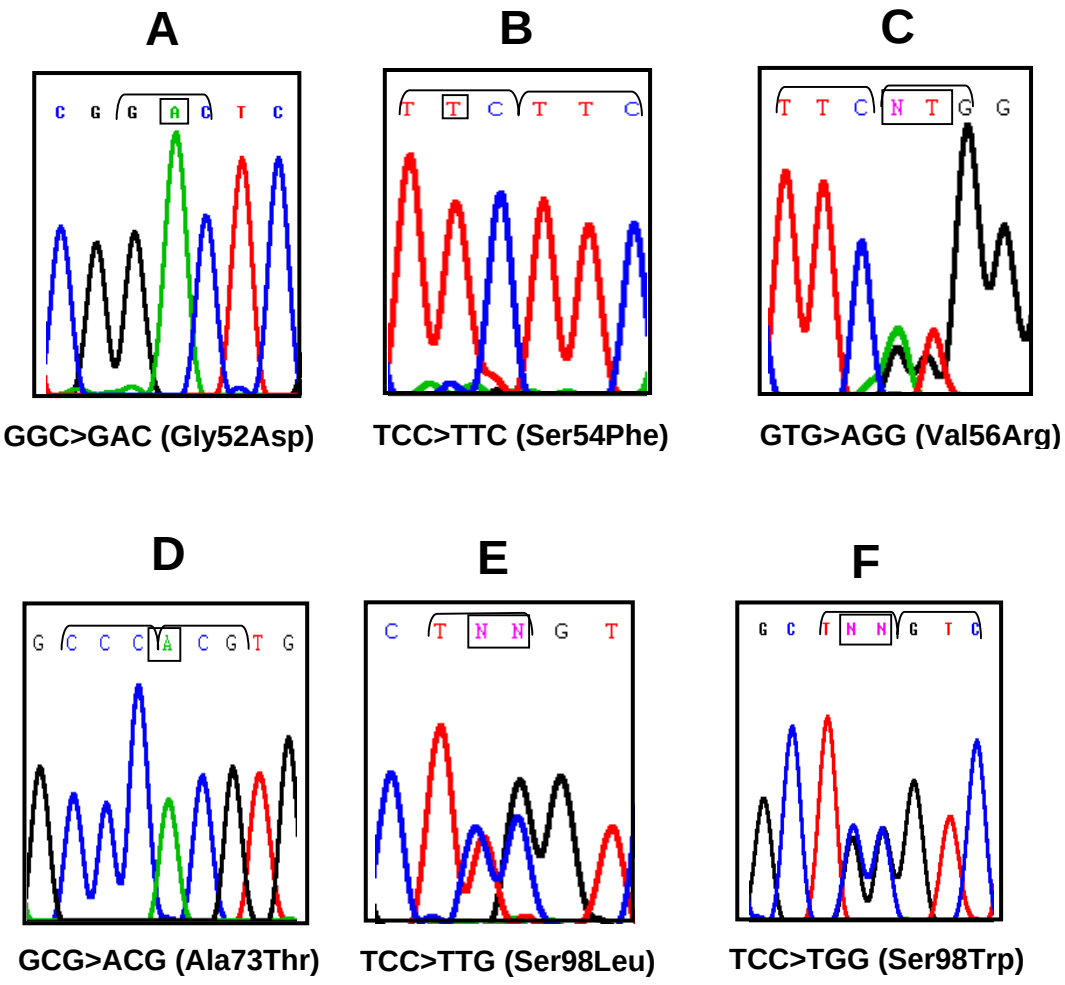


Figure 3.4. Sequence electropherograms of novel missense mutations with codon and amino acid changes shown below each panel. Details are in table 3.4. A: Homozygous mutations identified in families MCD-4 & MCD-5 (A), MCD-34 (B) MCD-38 (D), MCD-17 & MCD-48 (L), MCD-21 & MCD-53 (N), MCD-13 (G), MCD-44 (H), MCD-45 (J), MCD-47 (K), MCD-49 (M), MCD-54 (P), MCD-55 (Q). Heterozygous mutations identified in families MCD-35 (C&I) & MCD-39 (E), MCD-10 (F), MCD-20 (O). Boxes indicate the mutated residues and codons are marked by bracketts.

Figure 3.4

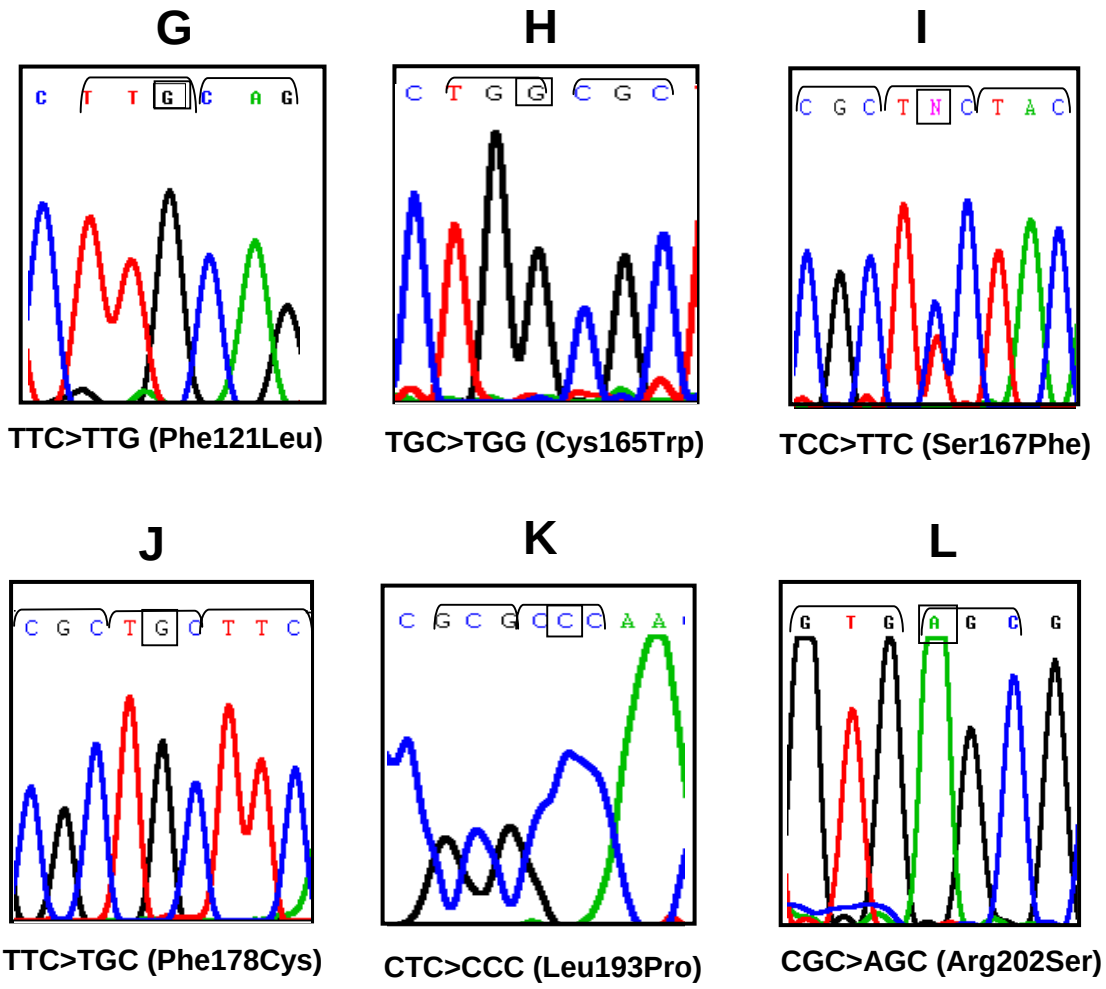
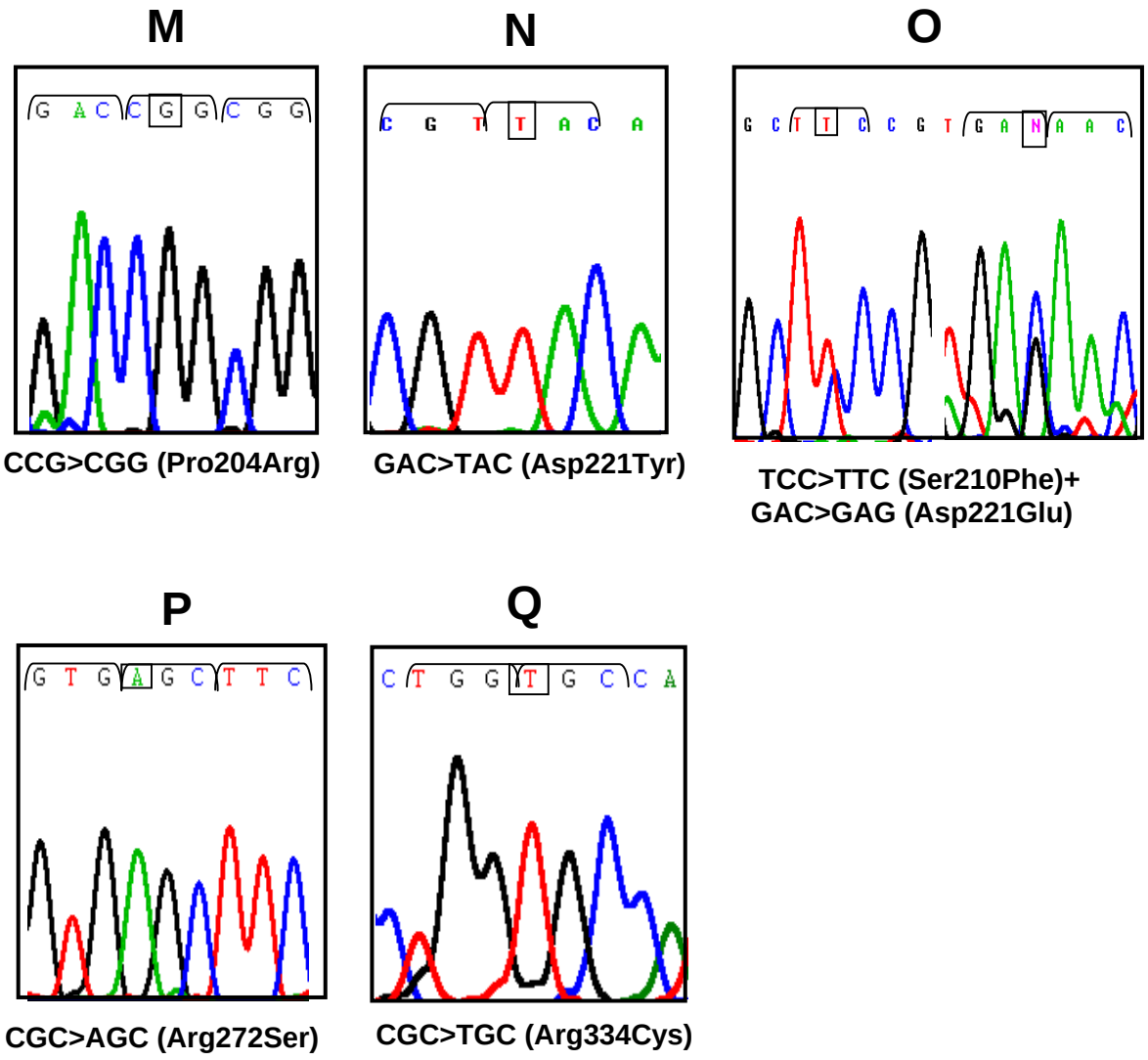

G
H
I
TTC>TTG (Phe121Leu)
TGC>TGG (Cys165Trp)
TCC>TTC (Ser167Phe)
J
K
L
TTC>TGC (Phe178Cys)
CTC>CCC (Leu193Pro)
CGC>AGC (Arg202Ser)

Figure 3.4


M
N
O
CCG>CGG (Pro204Arg)
GAC>TAC (Asp221Tyr)
TCC>TTC (Ser210Phe)+
GAC>GAG (Asp221Glu)
P
Q
CGC>AGC (Arg272Ser)
CGC>TGC (Arg334Cys)

Three probands, from families MCD-10, MCD-20 and MCD-52 showed two heterozygous missense mutations each. In MCD-10, changes of serine-98 to tryptophan (Ser98Trp) and phenylalanine-107 to serine (Phe107Ser) were found in the proband and an affected sibling. Changes serine-210 to phenylalanine (Ser210Phe) and aspartic acid-221 to glutamic acid (Asp221Glu) were found in proband of MCD-52 and in proband and affected sibling from MCD-20. The compound or simple heterozygosity of these mutations in the three families could not be confirmed because of the unavailability of other family members and also due to lack of suitable RFLPs in the normal or mutant sequences.

Compound heterozygosity of two missense mutations i.e., valine-56 to arginine and serine-167 to phenylalanine, found in MCD-35, was established by RFLP analysis. These two mutations (Val56Arg+Ser167Phe) result in the loss and gain of *MbiI* and *Eco57I* restriction sites respectively. In order to determine whether these two mutations are in *cis* or in *trans,* double digestion of the PCR product containing these residues was carried out with *MbiI* and *Eco57I* enzymes. The primerCK1h-intrn (Table 1, Appendix 1) was used as the forward primer and is located at positions g.15155 to g.15175 of *CHST6* (within intron 2 and 30 bases upstream of exon 3). The reverse primer CK71h R1674 (Table 1 Appendix I) is located within exon 3 and extends from positions g.16130 to g.16147 of *CHST6* genomic sequence (945 bases from the start of the exon 3) (Accession no NW_926528). The use of these primers for amplification would be expected to generate a product of 993bp.

The PCR product from normal genomic DNA has a restriction site only for *MbiI* enzyme at position 546 of the PCR product. The expected pattern of fragments from double digestion of the PCR product from normal genomic DNA is 546 and 447 bp (shown in lane 2 Figure 3.6). In the presence of mutation c.500C>T, the restriction site for *MbiI* would be abolished. For enzyme *Eco57I*, normal DNA has no restriction site whereas in the presence of mutation c.166G>A at position 213 of the PCR product, restriction site for *Eco57I* would be created leading to fragments of 780 and 213bp. If both the above mutations were present in *cis* (i.e., present on the same allele, with the remaining allele being normal) the expected pattern after double digestion is 780bp, 546bp, 447bp and 213bp (Figure 3.5A). If these mutations are present in *trans* (i.e., the mutations are present in two different alleles) then the expected pattern of digestion after double digestion is 993bp, 447bp, 333bp, and 213bp (Figure 3.5B). The PCR-amplified product from proband of MCD-35 upon double digestion *MbiI* and *Eco57I* gave a band pattern of 993bp, 447bp, 333bp, and 213bp (lane 3, Figure 3.6). Hence it was concluded that the two heterozygous missense mutations identified in MCD-35 proband were present on two different alleles and are therefore in *trans*.

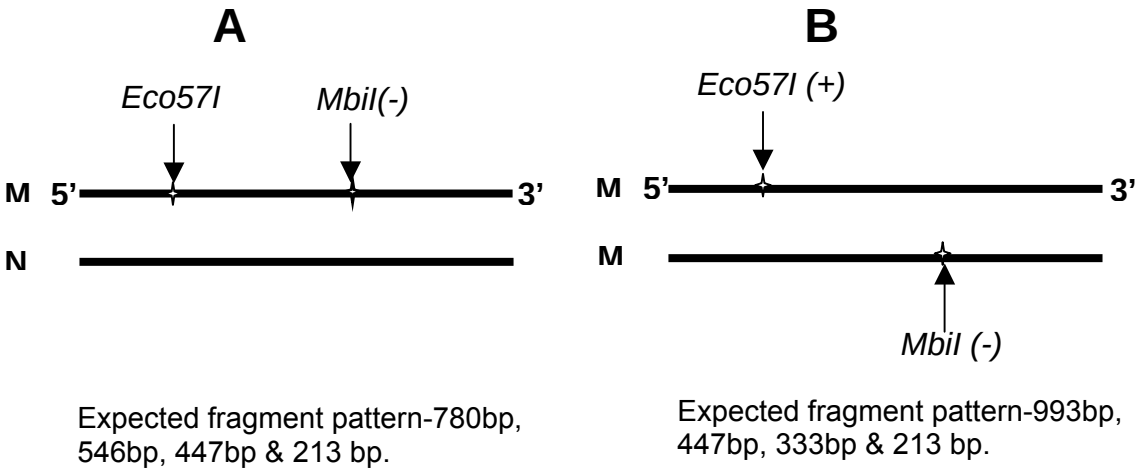


Figure 3.5. Schematic diagram showing positions of restriction enzyme sites altered by heterozygous mutations of Val56Arg and Ser167Phe identified in family MCD-39. RFLP patterns if the two changes are in *cis* (A) and in *trans* (B) are shown. 'M' indicates mutant and 'N' indicates normal allele. Restriction enzyme cuts are marked by arrows and mutations by stars. The sign (+) and (-) indicate gain and loss of restriction site due to mutation. Details are in text.

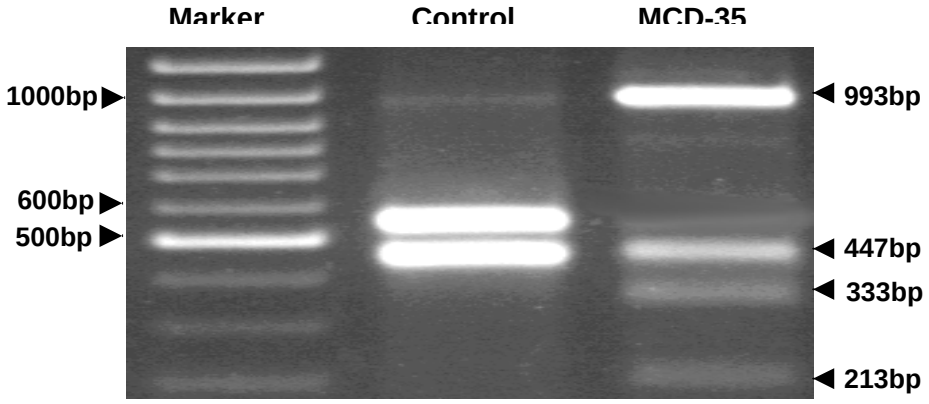


Figure 3.6. Double digestion of PCR products from control and MCD-35 with *MbiI* and *Eco57I* enzymes shows bands corresponding to fragments of 546 and 447 base-pairs in the control and 993, 447,333 and 213 base-pairs in MCD-35. Marker-100 bp DNA ladder

To determine the alleles of proband from MCD-39 with two consecutive heterozygous base changes [c.293C > T; c.294C > G], allele-specific PCR was carried out (shown in Figure 3.7). Primers used (described in Chapter 2, Table 2.2) consisted of CHST-ASOF1 [wild type (forward)], CHST-ASOF2 [mutant (forward)], and CHST-ASOR [common (reverse)]. PCR conditions were optimized such that genomic DNA from normal control amplified with the ASOF1 allele-specific primer (complementary to the normal sequence), but did not amplify with the mutant allele-specific primer ASOF2. The mutant primer ASOF2 differed from the wild type primer (ASOF1) by 2 bases at the 3' end (Chapter 2, Table 2.2). An additional control that was used was DNA from MCD-10, a patient identified to have two consecutive mutations [c.293C>G; c.294C>G]. The sequence in MCD-10 differs by one base in this region from the mutant primer ASOF2. Under the PCR conditions used, MCD-10 showed no amplification with ASOF2 but did amplify with wild type primer ASOF1 (lanes 2, 6, Figure 3.7). If the 2 changes in MCD-39 were in *trans* then it was expected that there would be no amplification with either of the primers [ASOF1 (normal) or ASOF2 (mutant)]. If the 2 changes were in *cis* then it would be expected to amplify with ASOF1 (normal) as well as ASOF2 (mutant) primer. The DNA from MCD-39 (Figure 3.7, lanes 3 and 5) amplified with allele-specific primers ASOF1 and ASOF2 for both normal and mutant sequences, respectively. Based on these results it can be concluded that the two base changes at c.293 and c.294 are in *cis* in proband MCD-39. This would predict a heterozygous mutation at the protein level of Ser98Leu.

In addition to 19 missense mutations not reported prior to this study, six more missense mutations were identified which were reported previously. These include

Ser53Leu (Warren etal. 2003), Phe107Ser (El-Ashry et al., 2002), Ser131Pro (Niel et al., 2003), Cys165Ser (Bao et al., 2001), Pro204Gln and Leu276Pro (Aldave et al., 2004).

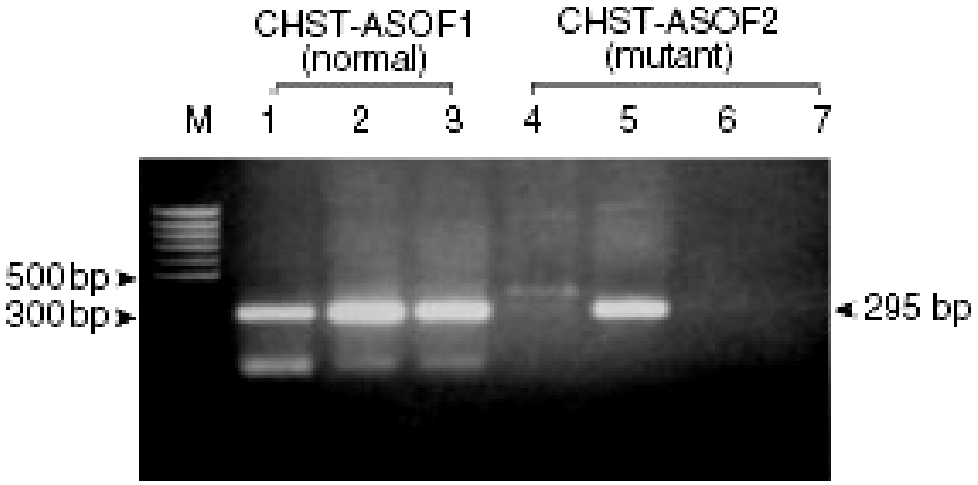


Figure 3.7. Results of allele-specific polymerase chain reaction. Lanes 1 and 4: DNA from normal control. Lanes 2 and 6: DNA from MCD-10. Lanes 3 and 5: DNA from proband of Family 39. Lane 7: Negative control. M:100 bp DNA ladder.

3.3 MULTIPLE SEQUENCE ALIGNMENT

Missense mutations were the most common type of mutation identified in *CHST6*. A multiple sequence alignment of protein sequences of different sulfotransferases was built using CLUSTALW software to look at evolutionary conservation of residues mutated. Alignment of six sulfotransferases (human intestinal GlcNAc 6-O-sulfotransferase, human high-endothelial-cell N-acetylglucosamine 6-O-sulfotransferase, human chondroitin 6-sulfotransferase-2, mouse L-selectin ligand sulfotransferase, mouse N-acetylglucosamine 6-*O*-sulfotransferase and mouse chondroitin 6-sulfotransferase) of human and mouse origin showed that all the residues with missense mutations were highly conserved (Tables 3.5 & 3.6). Eight mutations involved conserved residues required for interacting with the sulfate donor 3'-phosphoadenosine 5'-phosphosulfate (PAPS). Of these mutations Gly52Asp, Ser53Leu, Ser54Phe, and Val56Arg are located within the 5'-PSB (5'-phosphosulfate-binding) motif (Table 3.5), which consists of 8 amino acids (5'-Pro Lys Ser Gly Thr Thr Trp Ile 3' (sequence of mouse estrogen sulfotransferase). Mutations Arg202Ser, Pro204Gln, Pro204Arg and Ser210Phe involve residues within the 3'-PB (3'-phosphate-binding) motif (Table 3.6). This region consists of 16 amino acids which are highly conserved among all known sulfotransferases (5' Ile Tyr Leu Cys Arg Asn Ala Lys Asp Val Ala Val Ser Tyr Tyr Tyr 3' [sequence of mouse estrogen sulfotransferase]) that interact with the 3'-phosphate group of PAPS (Kakuta et al., 1998).

Table 3.5. Sequence alignment of sulfotransferases showing conservation of amino acids mutated in MCD.

Sequence	Gly52Asp, Ser53Leu	Ser54Thr, Val56Arg	Ala73Thr	Ser98Trp/ Leu	Phe107Ser	Phe121Leu	Cys165Trp, Ser167Phe
Human corneal GlcNAc 6-sulfotransferase (AF219990)	S **G** **S** S	**S** F **V** G	E P **A** W	R **S** V F	D V **F** D	L **F** Q W	A **C** R **S**
Human intestinal GlcNAc 6-*O* sulfotransferase (AF246718)	S **G** **S** S	**S** F **V** G	E P **A** W	R **S** I F	D V **F** D	F **F** N W	A **C** R **S**
Human high-endothelial-cell N-acetylglucosamine 6-*O*-sulfotransferase (AF131235)	S **G** **S** S	**S** F **V** G	E P **A** W	R **A** V F	S V **F** D	L **F** Q W	A **C** R **S**
Human chondroitin 6-sulfotransferase-2 (AB037187)	T **G** **S** S	**S** F **L** G	E P **M** W	R **S** L F	S V **L** R	L **F** R W	E **C** R **K**
Mouse L-selectin ligand sulfotransferase (AF109155)	S **G** **S** S	**S** F **L** G	E P **A** W	R **S** V F	S V **F** D	L **F** Q W	A **C** R **S**
Mouse N-acetylglucosamine 6-*O*-sulfotransferase (AF176841)	S **G** **S** S	**S** F **V** G	E P **A** W	R **S** V F	D V **F** D	L **F** Q W	A **C** S **S**
Mouse chondroitin 6-sulfotransferase (AB008937)	T **G** **S** S	**S** F **V** G	E P **L** W	K **Q** L L	Y V **L** E	L **F** R R	A **C** R **R**

The protein sequences of the sulfotransferases shown were aligned using the software CLUSTALW. Amino acid sequences flanking each mutation are shown in this table. Mutations are indicated at the top and the corresponding residues highlighted in bold. GenBank accession nos. of the corresponding mRNA sequences are given in parentheses.

Table 3.6. Sequence alignment of sulfotransferases showing conservation of amino acids mutated in MCD.

Sequence	Phe178Cys	Leu193Pro	Arg202Ser, Pro204Gln	Ser210Phe	Asp221Glu/ Tyr	Arg272Ser	Arg334Cys
Human corneal GlcNAc 6-*O*-sulfotransferase (AF219990)	- V R **F** F	- P A **L** N	- R D **P** R	- L R **S** R	- A R **D** N	- V **R** F E	- A W **R** H
Human intestinal GlcNAc 6-*O*-sulfotransferase (AF246718)	- V R **F** F	- P A **L** N	- R D **P** R	- L R **S** R	- A R **D** N	- V **R** F E	- A W **R** H
Human high-endothelial-cell N-acetylglucosamine 6-*O*-sulfotransferase (AF131235)	- V R **F** F	- P S **L** N	- R D **P** R	- F R **S** R	- M I **D** S	- V **R** Y E	- A W **R** W
Human chondroitin 6-sulfotransferase-2 (AB037187)	- V R **L** L	- P G **L** N	- R D **P** R	- H N **S** R	- L R **E** S	- L **R** Y E	- A W **R** E
Mouse L-selectin ligand sulfotransferase (AF131235)	- V R **F** L	- P S **L** N	- R D **P** R	- F R **S** R	- V V **D** S	- L **R** Y E	- A W **R** W
Mouse N-acetylglucosamine 6-*O*-sulfotransferase (AF176841)	- V R **F** F	- P A **L** N	- R D **P** R	- L R **S** R	- A R **D** N	- V **R** Y E	- A W **R** H
Mouse chondroitin 6-sulfotransferase (AB008937)	- V R **I** R	- L R **L** D	- R D **P** R	- L A **S** R	- Y E **N** W	- V **R** Y E	- K W **R** F

The protein sequences of the sulfotransferases shown were aligned using the software CLUSTALW. Amino acid residues flanking each of the mutated residues only are shown above. Mutations are indicated at the top and the residues undergoing mutation are highlighted in bold. GenBank accession nos. of the mRNA sequences are given in parentheses.

3.4 STRUCTURAL ANALYSIS OF CHST6 MUTANT PROTEINS

Sixty two percent (25/40) of the mutations identified in the present study are missense mutations. Unlike nonsense and frameshift mutations, missense mutations involve the substitution of one amino acid by another. Hence it was tempting to study the impact of missense mutations on the structure and function of corneal N-acetylglucosamine-6-sulfotransferase protein in order to determine their significance. This was done by comparative three-dimensional modeling.

3.4.1 Modeling of CHST6 protein and its mutants

A three dimensional model of corneal N-acetylglucosamine-6-sulfotransferase protein was built using retinol dehydratase as the template. Retinol dehydratase was chosen as the template because it shows maximum coverage, with comparatively less insertions or deletions compared to other structurally known homologues. The sequence alignment used for model building is shown in Figure 2.2 Chapter 2. The alignment is characterized by insertions and deletions at various places. Since the first 39 residues from the N-terminal end and the last 114 residues at the C-terminal in CHST6 did not have corresponding equivalent regions in retinol dehydratase, the modeling of CHST6 protein was carried out from 40 to the 287th residue followed by a rigorous refinement of the model by means of energy minimization. The modeled wild type cGlnNAc6ST consist of 8 α helices and 4 β strands. The model was used to identify the structural parts, which make up the cofactor and substrate-binding site. Mutants were modeled by replacing the side chains of the wild type cGlnNAc6ST with the side chains found in the various mutant forms.

As implied firmly by the homology with retinol dehydratase and its homologues, and by the virtue of comparative modeling, the structural model of cGlnNAc6ST corresponds to the α/β fold of P-loop containing nucleoside triphosphate hydrolases (Figure 3.8). This fold corresponds to three layers with middle layer of parallel β-sheet sandwiched by layers of α-helical regions. However the C-terminal region of the domain, encompassing the last helix and part of the preceding helix, was not modeled owing to the poor sequence similarity between cGlnNAc6ST and retinal dehydratase in that region. Otherwise the modeled structure of cGlnNAc6ST contains all the essential structural elements as in the members of the superfamily of P-loop containing nucleoside triphosphate hydrolases. It has been observed that the surface of the modeled structure is predominantly made up of polar residues whereas the residues in the interior of the modeled structure are mostly apolar. This observation supports further the validity of the overall structural model of cGlnNAc6ST.

As in 3-D structures of homologues, the functional site in the modeled structure of cGlnNAc6ST is formed by the crevice created by crossover connection of loop-helix region. This region is responsible for switching of lining-up of helical regions from one side of the β-sheet to the other side resulting in the open β-sandwich structure.

Figure 3.8. Model of wild type CHST6 protein. Helices are shown in green colour and β-strands are shown in blue colour.

A total of 23 missense mutations and one insertion mutation were studied for their predicted effects on the structure of the wild type CHST6 protein. Three major effects were identified by which the enzyme activity was hampered. These include disruption of hydrophobic cluster, loss of interaction with cofactor and loss of disulphide bond.

3.4.2 Disruption of hydrophobic cluster

Mutations in this group included Gly52Asp, Val56Arg, Ala73Thr, Phe107Ser and Phe178Cys. The residue Gly-52 occurred just before helix H1 (Chapter 2; Figure 2.2). The helix H1 is buried in the model and adopts a positive ϕ conformation in the homologues of known structure. This conformation is unfavorable for residues such as aspartic acid with a side-chain. Glycine-52 is situated in the middle of strong hydrophobic environment provided by Val56 and Phe273 (Figure 3.9A). Replacement of Gly-52 by aspartic acid brings in a negatively charged, bulky amino acid in the middle of an apolar environment which is likely to be unfavorable to the stability of the native fold. Also, there is not enough space available in the 3-D structure to accommodate the side chain of aspartic acid. Hence it could be predicted that the 3-D structural features of Gly52Asp mutant were likely to be significantly different compared to that of the wild type, and possibly hampering the enzymatic function.

In Val56Arg, Valine-56 is in the middle of the helix H1 (Chapter 2; Figure-2.2.), a part of a hydrophobic cluster that enables packing among two helices and a β-strand at the core of the protein structure. The other residues that contributed to the formation of the hydrophobic cluster were Val45, Leu200, Val198, Ile197, Ile284 and Leu281 (Figure 3.9B). Introduction of Arg in the place of Val56 is expected to

disintegrate the cluster of hydrophobic residues, due to introduction of charge in the apolar environment and also due to probable lack of space to accommodate a large residue such as Arg. Thus this change is likely to affect the overall fold of the protein and its enzymatic function.

Alanine-73 is located in the middle of 21 amino acid insertion (after helix H1, (Chapter 2; Figure 2.2), compared to the sulfotransferases of known 3-D structure. The insertion is characterized by a substantial proportion of hydrophobic residues and is predicted to have two helices and an extended β-strand. It is possible that the alanine residue is involved in hydrophobic interaction with other hydrophobic residues present in insertion region such as Trp, Leu Val and Met. Disruption of this hydrophobic interaction could occur in presence of threonine, which is a polar residue. However the expected accuracy of the 3-D model in this region is not high enough to provide a more concrete prediction.

Phenylalanines 107 and 178 are located in an aromatic cluster along with Phe158 (Figure 3.9C) surrounded by the hydrophobic environment provided by some local aliphatic amino acid such as Val67, Val137, Val144, Leu145. Replacement of phenylalanine by a non-aromatic amino acid in an aromatic cluster would be expected to be unfavorable for the integrity of the structure. Replacement of Phe by smaller residues would result in adjustments in packing of the side chains in order to avoid the void volume created by replacement of a bulky residue by a smaller residue

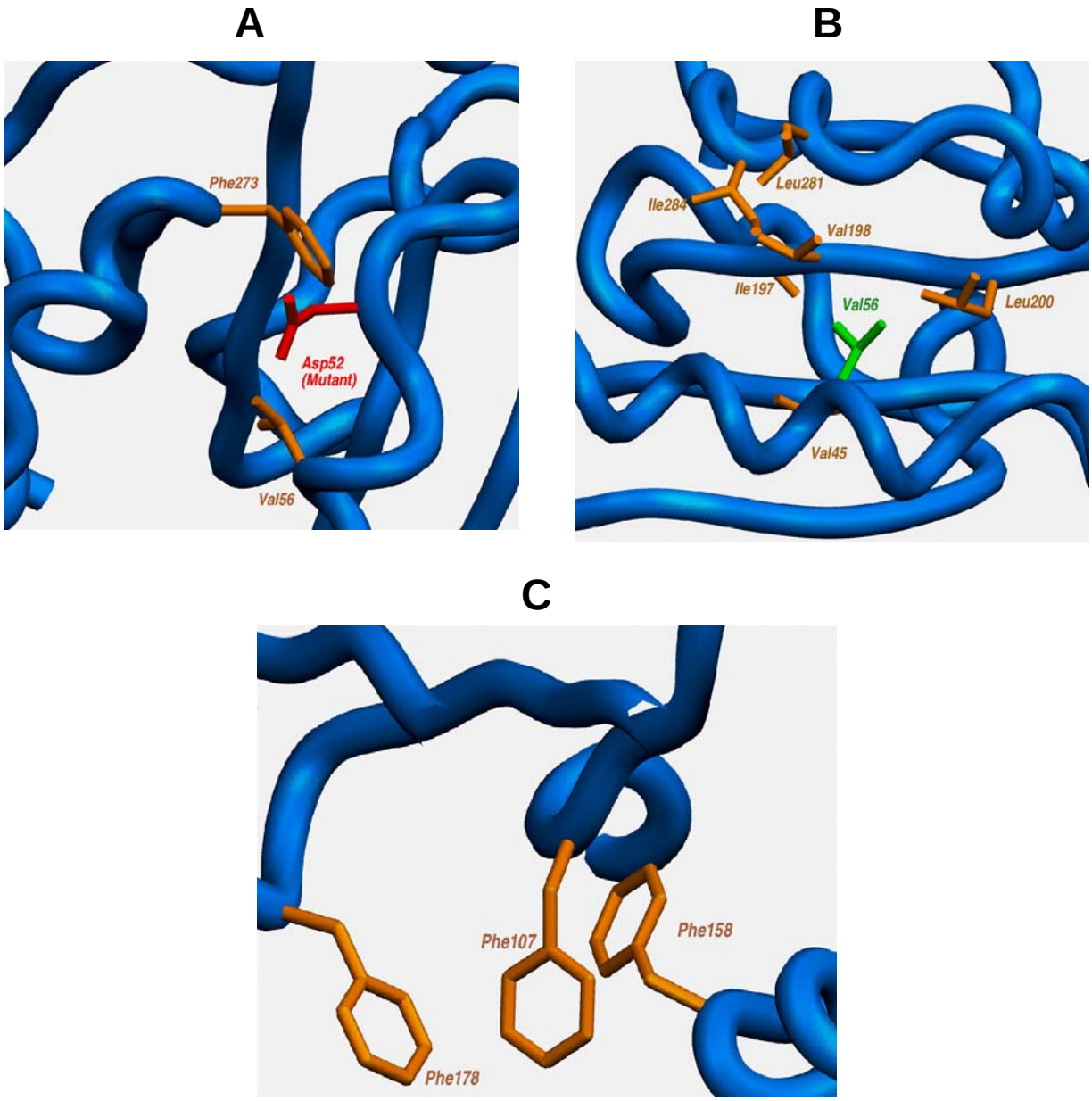


Figure 3.9. Close-up of the structural model of CHST6 protein. The backbone is represented in blue and the proximal hydrophobic residues in brown. A: Asp52 surrounded by hydrophobic residues. B: Location of valine-56 (green colour) in an apolar environment (brown coloured side-chains). C: Cluster of aromatic residues in the core of the model of the wild type cGlnNAc6ST (CHST6) protein.

3.4.3 Loss of interaction with cofactor

Mutations in this group included Ser53Leu, Ser54Phe, Ser98Trp/Leu, Pro204Glu/Arg, Ser210Phe, Asp221Glu/Tyr & insertion of Trp between 219_220, Arg272Ser and Leu276Pro. In the crystal structure of retinol dehydratase (Pakhomova et al., 2001) serine-53, a functional residue, is found to be involved in hydrogen bond interaction with the co-factor PAPS (Figure 3.9D) (the distance between hydroxyl group of serine side-chain and phosphate group of PAPS was 2.75 Å). In the Ser53Leu mutant this critical interaction would be affected and the side chain of the leucine residue would create a hydrophobic environment proximal to the polar group of PAPS. In serine-54 the hydroxyl group of its side chain was predicted to form a hydrogen bond with the cofactor PAPS (Figure 3.9D), based on the observation of similar interaction observed in the crystal structure of retinol dehydratase involving the topologically equivalent threonine side chain. Substitution of this serine by phenylalanine, a bulky aromatic amino acid, would be expected to prevent such an interaction with the cofactor, and also displace the cofactor from its binding site.

Serine-98 interacts with the carboxyl-terminus of a helix by forming a hydrogen bond with the main chain (a hydrogen bond distance of about 3.3Å (Figure 3.9E). Thus serine-98 acts as a capping residue stabilizing the termination of the helix. Replacement of serine by any of the non-polar amino acids (tryptophan or leucine) is predicted to result in loss of this interaction and destabilization of the helix termini. This could result in an alteration of the length of the helix.

Proline-204 is located at the tip of a short loop of irregular conformation (between β-strand B3 and α-helix H4) and flanked by arginine-202 and serine-210. Both arginine-202 and serine-210 are predicted to be interacting with the cofactor. Proline-204 may contribute to maintaining the structure by providing rigidity to the otherwise flexible loop and thereby secure the interaction between arginine, serine and PAPS. Occurrence of a non-polar residue (glutamine or arginine) at 204 can increase the flexibility of the loop and destabilize the interaction between arginine, serine and PAPS. Serine-210 directly interacts with the phosphate group of PAPS and the occurrence of phenylalanine at this position would be expected to have similar effects as substitution of serine-53, 54 and result in the loss of interaction, as well as expulsion of the cofactor from the binding site.

Residues 221, 219 and 220 are located in a region, which is unique to the CHST6 protein, compared to homologous sulfotransferases of known three-dimensional structure. Hence it was difficult to model this region reliably. However based on the modeled 3-D structure, this region is located close to the cofactor-binding site and is possibly important for function of the protein. Therefore mutations Asp221Glu or Tyr and insertion of Tryptophan between 219 and 220 amino acid are expected to interfere with binding of cofactor PAPS.

Arginine-272 is located at the C-terminal end of β-structure and potentially interacts with two acidic residues, glutamic acid-274 and aspartic acid-275, in the subsequent α-helix (H8) (Figure 2.2 Chapter 2) which packs with the N-terminal α-helix containing cofactor binding residues. The replacement of arginine by serine

would result in loss of interaction between these residues and acidic residues in the α-helix. This will affect the packing between the β-strand and the following α-helix.

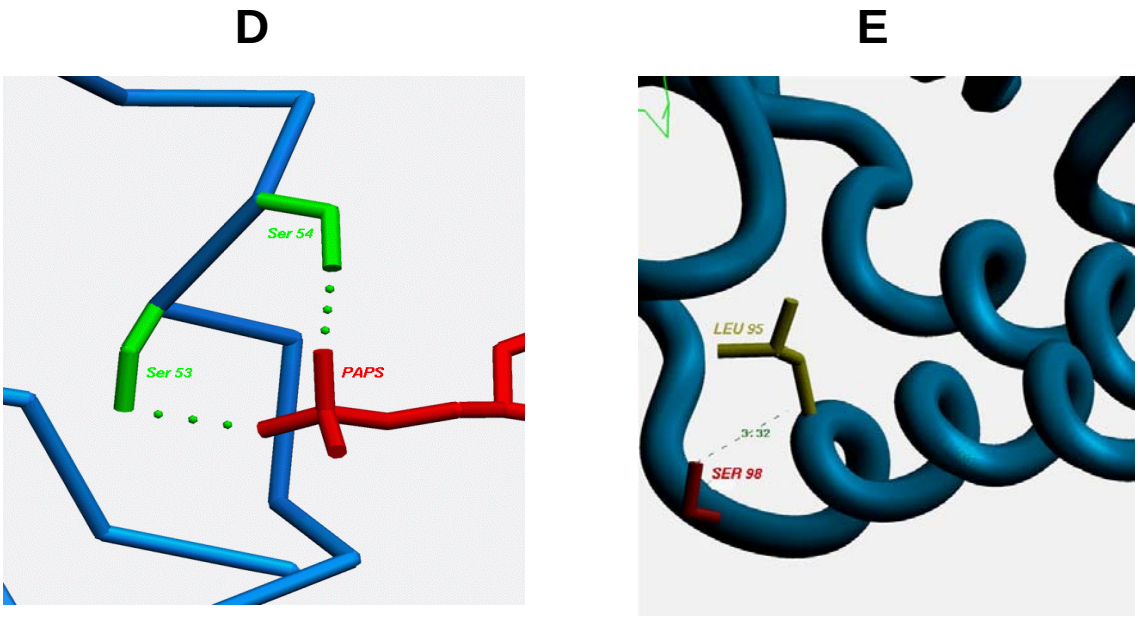


Figure 3.9D: Interaction of PAPS with Ser53 and Ser54 in the structural model of CHST6 protein. PAPS is shown in red, side-chains of Ser in green and the Cα trajectory in blue. E: Interaction of the side-chain of Ser98 with the helix termini in the model of the wild type CHST6 protein. The hydrogen bonding distance is 3.32 Å

Leucine-276 is located in a region showing high similarity with the corresponding region of retinol dehydratase. This segment is predicted to adopt α-helical structure (α helix H8 Figure 2.2; Chapter 2) with the leucine located in the middle of the helix. This helix packs with the helix containing serine-53 and other residues that bind with the cofactor. When proline is substituted for leucine it is expected to introduce a kink into the helix, and thereby alter its interaction with the N-terminal helix. This could affect the spatial orientation of the cofactor-binding site in mutant.

3.4.4 Loss of disulfide bond

Mutations in this group include Cys165Ser and Cys165Trp. Cys-165 lies in a region (before H2 Figure 2.2 Chapter 2.2) that is a highly variable region among sulfotransferases of known three-dimensional structures. However cysteine-165 is absolutely conserved in sulfotransferases, which were closely related to CHST6 such as human chondroitin-6-sulfotransferase, human keratan sulfate galactose-6-sulfotransferase etc. Three-dimensional modeling suggested proximity of this cysteine to cysteine-102 (Figure 3.9F). Interestingly, cysteine-102 is also conserved in close homologues of CHST6. The spatial proximity of two conserved cysteine residues suggests the possibility of disulfide bond formation. This could contribute to the three-dimensional stability of the protein and would predict that substitution of either cysteine-165 or cysteine-102 would lead to loss of stability.

Serine-167 lies in a structurally variable region far away from the functional site of the enzyme. Hence, it is unlikely to participate in protein function. However,

among the homologous proteins, this residue is substituted only by polar amino acid residues and, in the three-dimensional model, it is predicted to be highly solvent exposed. Hence, replacement of this residue by phenylalanine, a hydrophobic amino acid, may affect the stability of structure, as phenylalanine would be exposed to the solvent.

F

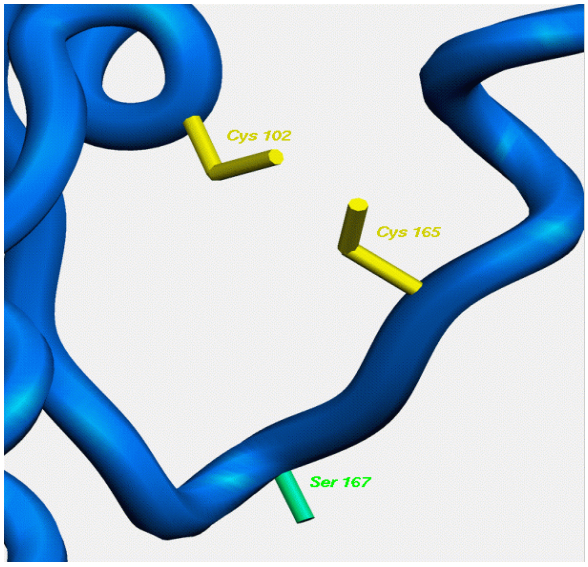


Figure 3.9F: Spatially proximal cysteine residues (Cys102 & Cys165) forming a disulphide bridge.

3.5 IMMUNOPHENOTYPING

Immunophenotypes of MCD patients included in this study were determined by estimation of levels of antigenic KS in serum and in cornea. Serum AgKS levels in these patients were estimated in the laboratory of E. J. Thonar and AgKS in cornea was evaluated by immunohistochemical analyses of corneal sections (details in Chapter 2).

3.5.1 Pattern of AgKS in cornea

IHC was standardized by testing dilutions of keratan sulfate-specific monoclonal antibody 1/20/5-D-4 ranging from 1:5000 to1:500 on normal corneal sections. Test corneas were subjected to IHC under optimized conditions along with negative and positive controls (as detailed in Chapter 2). Results of IHC on positive and negative control specimens are shown in Figures 3.10 A & B. Of the total of 48 patients evaluated for AgKS in cornea, 33 patients showed no reactivity in any of the corneal layers (Table 3.7). An example of a corneal section in this group is shown in Figure 3.11A. Corneal sections from 12 patients showed reactivity for AgKS mainly in stromal keratocytes with very little or no positive reaction in the stroma. An example of immunohistochemical result from this group is shown in Figure 3.11B. The cornea from three patients showed positive reactivity throughout the entire stroma, keratocytes and endothelium (Figure 3.11C).

3.5.2 Determination of MCD sub-types

Among the 48 patients whose corneas were evaluated, serum levels of AgKS could be done in 35 patients. The values ranged from <4 to 388 ng/ml (Table 3.7). The results presented in Table 3.7 show that 22 patients with AgKS determined in both cornea and serum had AgKS pattern consistent with MCD type I (AgKS –ve in cornea, <10 ng/ml in serum) (Figure 3.11A); five patients had a pattern consistent to that reported for MCD type IA (AgKS +ve in corneal keratocytes, <10 ng/ml in serum) (Figure 3.11B); one patient had a pattern consistent with that reported for MCD type II (AgKS +ve in cornea, 101 ng/ml in serum) (Figure 3.11C). Apart from these, there were seven patients whose AgKS patterns did not conform to those described above for the known MCD sub-types. These were classified as 'atypical'. These include families MCD-2, MCD-20, MCD-45, & MCD-49 that had corneal AgKS patterns consistent with MCD I, and families MCD-28, MCD-39, and MCD-52 that had corneal AgKS patterns consistent with MCD IA. However, the serum AgKS levels in these patients were not compatible with types I or IA in which AgKS levels have been reported as <10ng/ml (Klintworth et al., 1997). AgKS in serum of all these patients ranged from 19 to 388 ng/ml.

An additional 16 patients evaluated for AgKS only in serum showed serum AgKS levels of <8ng/ml, consistent with MCD type I although corneal AgKS was not determined. The two unaffected family members evaluated for AgKS in serum showed AgKS levels of 166 and 287 ng/ml respectively.

3.5.3 Correlation between immunophenotypes and mutations

Mutations associated with each immunophenotype were: 6 missense, 2 nonsense, 5 frameshift, and 2 insertion in MCD type I, 4 missense and one insertion in MCD type IA and one missense mutation in MCD type II.

Patients with atypical immunophenotypes showed missense and nonsense mutations. Five patients from 4 MCD families (MCD-26, MCD-27, MCD-28 & MCD-52) had the same mutation Asp221Glu and showed MCD type IA immunophenotype. Serum levels of AgKS however, were not consistent with type IA for all, being atypical for two patients (from families MCD-28 & MCD-52, shown in Table 3.7). Another example of diverse immunophenotypes in association with the same mutation was found in family MCD-21 with three affected members having the same mutation, Asp221Tyr. In this family, the proband showed type II immunophenotype, with AgKS +ve in cornea, and 101 ng/ml AgKS in serum. His wife showed type IA immunophenotype with positive reactivity for AgKS in stromal keratocytes and <4ng/ml in serum (Figures 3.12 A & B). His son showed MCD type I immunophenotype with absence of AgKS in cornea and <4ng/ml in serum (shown in Table 3.7).

Two patients (MCD-9 & MCD-32) showed corneal immunophenotypes corresponding to MCD type II. One (MCD9) had a single heterozygous mutation (c.198delC) and the other had 2 homozygous mutations viz., c.698G>A (Trp2X) and c.699C>A (Leu3Met). However serum was not available for the analysis in these two patients. Based on all the above results it was concluded that there are no evident correlations between the immunophenotypes and mutations in the *CHST6* gene.

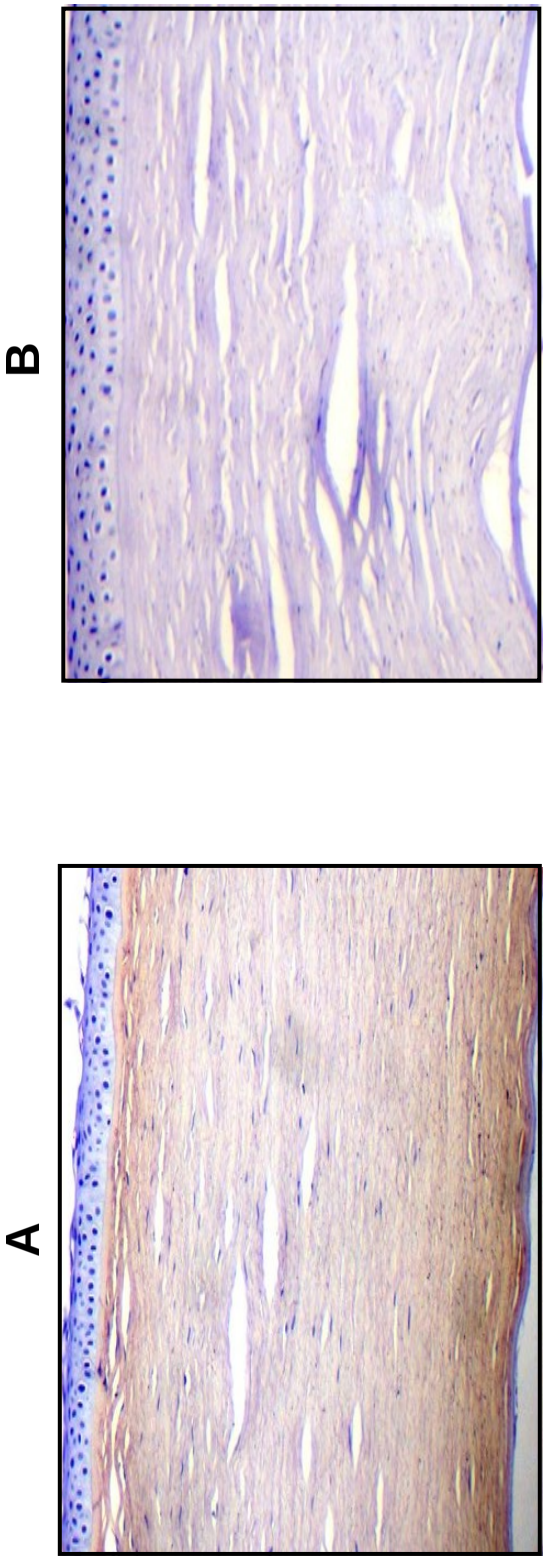


Figure 3.10. Immunohistochemical staining of normal corneas (from patient with AR CHED) with I/20/5-D4 antibody (counterstained with hematoxylin eosin). A: Positive control. B: Negative control, same specimen as in A with primary antibody omitted (x100).

Figure 3.11 Light microscopic appearances of different immunophenotypes of MCD (counterstained with hematoxylin eosin, X100). A: Cornea with absence of AgKS in keratocytes and stroma (MCD type I). B: Cornea with reactivity to AgKS only in stromal keratocytes (MCD type IA). C: Cornea with AgKS detected in stroma, keratocytes, Descemet's membrane and endothelium (MCD type II). Arrows point to the deposits in stromal keratocytes.

Figure 3.12. Different immunophenotypes identified in corneal sections from family MCD-21 (counterstained with hematoxylin eosin and PAS, X400). A: Proband (MCD21A) with MCD type II immunophenotype identified in proband MCD21A. B: The proband's spouse (MCD-21B) with MCD type IA immunophenotype. Arrows point to the deposits in stromal keratocytes.

Table 3.7. Correlations between Immunophenotypes and mutations.

MCD Family#	Corneal IHC (AgKS +ve/-ve in)		Serum KS ng/ml	Immuno-phenotypes	Mutation
	Stroma	Keratocytes			
MCD-2	-ve	-ve	63	AT	p.Gln18X
MCD-3	-ve	-ve	<4	I	p.Ser32GlnfsX36
MCD-4	-ve	-ve	<2	I	p.Gly52Asp
MCD-5	Not done	Not done	<4		p.Gly52Asp
MCD-6	-ve	-ve		I	p.Ser53Leu
MCD-8	Not done	Not done	<4		p.Phe60LeufsX10
MCD-9	+ve	+ve		II	p.[Phe67SerfsX3]+[?]
MCD-10A	-ve	+ve	<4	IA	p.[Ser98Trp(+) Phe107Ser]
MCD-10B	-ve	+ve	<4	IA	p.[Ser98Trp(+) Phe107Ser]
MCD-11	-ve	-ve	<4	I	p.Phe107Ser
MCD-12	-ve	-ve	<4	I	p.Phe107Ser
MCD-13	-ve	+ve	<4	IA	p.Phe121Ser
MCD-14	-ve	-ve		I	p.Trp123X
MCD-15	-ve	-ve	<4	I	p.Cys153X
MCD-16*	-ve	-ve	<2	I, I	p.Arg195-196ins
MCD-18	-ve	-ve	<4	I	p.Pro204Gln
MCD-19*	-ve	-ve	<2	I, I	p.Arg205TrpfsX176
MCD-20A	-ve	-ve	61	AT	p. [Ser210Phe (+)Asp221Glu]
MCD-20B	-ve	+ve		IA	p. [Ser210Phe (+)Asp221Glu]
MCD-21A	+ve	+ve	101	II	p.Asp221Tyr
MCD-21B	-ve	+ve	<4	IA	p.Asp221Tyr
MCD-21C	-ve	-ve	<4	I	p.Asp221Tyr
MCD-22*	-ve	+ve	<2	IA, IA	p.Trp219-220ins
MCD-23	-ve	-ve	6	I	p.Trp219-220ins

MCD Family#	Corneal IHC (AgKS +ve/-ve in)		Serum KS ng/ml	Immuno-phenotypes	Mutation
	Stroma	Keratocytes			
MCD-24	-ve	-ve	<4	I	p.Trp219-220ins
MCD-25	-ve	-ve	<4	I	p.Trp219-220ins
MCD-26	-ve	+ve		IA	p.Asp221Glu
MCD-27*	-ve	+ve		IA, IA	p.Asp221Glu
MCD-27-B	-ve	+ve		IA	p.Asp221Glu
MCD-28	-ve	+ve	95	AT	p.Asp221Glu
MCD-29	Not done	Not done	<4		p.Gly309X
MCD-31	-ve	-ve	<4	I	p.Glu347X
MCD-32	+ve	+ve		II	p.[Trp2X;Leu3Met]+ [Trp2X;Leu3Met]
MCD-33	-ve	-ve	5	I	p.Val 6 ProfsX55
MCD-34	-ve	-ve	6	I	p.Ser54Phe
MCD-35	Not done	Not done	<4		p.[Val56Arg]+ [Ser167Phe]
MCD-36	-ve	-ve	7	I	p.Phe60LeufsX10
MCD-37	-ve	-ve	8	I	p.Phe60LeufsX10
MCD-38	Not done	Not done	<4		p.Ala73Thr
MCD-39*	-ve	+ve	388	AT	p.[Ser98Leu]+[?]
MCD-40	Not done	Not done	8		p.Trp123X
MCD-41	"	"	<4		p.Ser131Pro
MCD-43	-ve	-ve		I	p.Cys165Ser
MCD-44	-ve	-ve		I	p.Cys165Trp
MCD-45	-ve	-ve	33	AT	p.Phe178Cys
MCD-46A	Not done	Not done	<4		p. [Gln182ArgfsX199]+ [Leu276Pro]
MCD-46B	"	"	<4		p.Leu276Pro
MCD-47	"	"	<2		p.Leu193Pro
MCD-48	-ve	-ve	19	AT	p.Arg202Ser

MCD Family#	Corneal IHC (AgKS +ve/-ve in)		Serum KS ng/ml	Immuno-phenotypes	Mutation
	Stroma	**Keratocytes**			
MCD-49A	-ve	-ve	<2	I	p.Pro204Arg
MCD-49-B	Not done	Not done	<2		p.Pro204Arg
MCD-50	-ve	-ve	<2	I	p.Trp219-220ins
MCD-51*	-ve	-ve		I, I	p.Trp219-220ins
MCD-52	-ve	+ve	142	AT	p.Asp221Glu
MCD-53*	-ve	-ve	<2	I, I	p.Asp221Tyr
MCD-54	Not done	Not done	<2		p.Arg272Ser
MCD-55	-ve	-ve		I	p.Arg334Cys
MCD-56	-ve	-ve		I	p.His335CysfsX27
MCD-57A	-ve	-ve	<4	I	p.His335CysfsX27
MCD-57B	-ve	-ve	<2	I	p.His335CysfsX27
MCD-57C	Not done	Not done	<2		p.His335CysfsX27
MCD-57D	"	"	<2		p.His335CysfsX27
MCD-59	"	"	105		No mutation
MCD-59B	"	"	42		No mutation

This table summarizes the immunophenotypes of MCD patients along with mutations found in each patient. AT indicates atypical immunophenotype, -ve &+ve indicate negative and positive reactivity with antikeratan sulfate antibody respectively. *Families in which corneas from both the eyes of proband were available.

3.6 DISCUSSION

In the present study, a total of 72 patients (from 59 families) with MCD were screened in order to identify pathogenic mutations in the *CHST6* gene in Indian patients. Fourty different mutations were identified. Majority of the mutations 25/40 (62.5%) identified in the present study were missense mutations followed by frameshift 7/40 (17.5%), nonsense 6/40 (15%) and insertion mutations 2/40 (5%). As shown in tables 3.1, 3.2, 3.3 and 3.4, the majority of probands (50) had homozygous pathogenic mutations within the coding region of *CHST6* gene. Two probands were compound heterozygous.

Thirty-one mutations identified in the present study were not reported previously. These 31 mutations, identified in 48 probands, included 6 nonsense, 4 frameshift, 2 insertion and 19 missense mutations.

Mutations identified were distributed throughout the length of protein. Recurrent mutations identified in the present study included c.1348insCTG, giving rise to an in frame insertion of tryptophan at position 219 (Trp219-220ins), and a missense mutation leading to change of aspartic acid-221 to glutamic acid / tyrosine (Asp221Glu/Tyr) (Table 3.2 &3.4). Both of these mutations were each found in 6 and 7 unrelated patients respectively. A deletion of C at position 872 of the cDNA, producing a frameshift at phenylalanine-60, was also recurrent in the patients studied, and it was found in 3/57 patients (Table 3.1).

Nine of the mutations identified in the present study were previously reported; c.158C>T (Ser53Leu), c.198delC, c.545delA, and c.612_614 del3insAT were reported in Indian patients (Warren et al., 2003), the Phe107Ser has been reported

previously in British family (El-Ashry et al., 2002), Ser131Pro and Pro204Gln were previously reported in French patients (Niel et al., 2003), Cys165Ser was reported in Saudi Arabian family (Bao et al., 2000), Leu276Pro was reported in American family (Aldave et al., 2004).

Akama and coworkers (Akama et al., 2000) identified deletions or rearrangements mutation in the region between *CHST5* and *CHST6* in cases of MCD type II. Hence in the 4 families in which either one mutant allele or no mutant allele was detected, analyses of the region upstream of *CHST6* (between *CHST5* and *CHST6*) was carried out in order to look for deletions or rearrangements. However no upstream deletion or rearrangements were identified in these patients (data not shown).

Since MCD appears to be a monogenic disorder it is possible that the cases with only one or no mutant alleles may have changes in the non-coding regions of *CHST6* that were not tested in this study. Sequences involved may include unidentified regulatory elements of the *CHST6* gene or other genes that affect the expression of the *CHST6* gene. The 5' UTR contributes to the specificity and overall efficiency of translation initiation in various genes (mammalian β2 adrenergic receptor, mouse retinoic acid receptor β2 and human cytomegalovirus gp48 transcript etc) (Gray et al., 1998). The 5'-untranslated region (5'-UTR) of the *CHST6* gene has not been characterized fully and unidentified mutations may reside in these regions.

A comparative model of cGlnNAc6ST was generated based on the homologues of known structure, in order to understand the possible impact of

missense and insertion mutations on the protein structure. Analysis of structural model of CHST6 protein revealed that all mutations studied are likely to hamper the activity of CHST6 protein by three major effects i.e, disruption of hydrophobic cluster, loss of interaction with cofactor and loss of disulphide bond.

The only substitution that involved replacement with a chemically similar amino acid was Asp221Glu (Table 3.4). The occurrence of this change in probands from 3 families (MCD-26, MCD-27 & MCD-28) as a homozygous change and in two families (MCD-20 & MCD-52) as a heterozygous change, together with its absence in 100 unrelated control individuals supports the conclusion that it is pathogenic. The mutation Asp221Glu is located in a region that is unique to cGlnNAc6ST protein compared to homologous sulfotransferases of known three–dimensional structure and hence was difficult to model. However based on the modeled 3-D structure, this region appears to be located close to the cofactor-binding site.

Immunophenotyping in MCD patients showed that MCD type I 63% (22/35) is the most frequent type followed by MCDIA 14% (5/35) and MCDII 3% (1/35). The remaining 20% of patients showed atypical immunophenotypes (i.e not compatible with patterns reported for three immunophenotypes) and serum AgKS ranged from >19-388 ng/ml in these patients. Identical immunophenotypes in both eyes of an individual were found in seven patients for whom both the corneas were available. No correlations were found between *CHST6* mutations and immunophenotypes of MCD.

4.1 INTRODUCTION

This chapter describes the molecular genetic analysis of AR CHED. Hand and coworkers localized the autosomal recessive form of CHED to an 8 cM interval on chromosome 20p13 (Hand et al., 1999). The same locus was found by other studies (Mohammed et al., 2001) though the CHED2 gene was not yet identified. Since recessive CHED is one of the more frequent corneal dystrophies presenting at our institution, suitable consanguineous families were available for study. Families with AR CHED were recruited and used for mapping, candidate gene analysis and subsequent mutational screening. The strategy used was to first test whether AR CHED in families of Indian origin also maps to the previously identified CHED2 locus on chromosome 20p13. This was followed by haplotype analysis to narrow down the critical interval and then, screening of genes in the interval for mutations. Identification of *SLC4A11* as the CHED2 gene in this study closely followed another study by Vithana and co-workers (Vithana et al., 2006). Further mutational analysis of *SLC4A11* was carried out on a larger group of families with AR CHED and genotype-phenotype correlations were investigated.

4.2 LINKAGE ANALYSIS

Linkage analysis was carried out on ten consanguineous AR CHED families (13 affected and 24 unaffected family members). These families fulfilled all the diagnostic criteria. Thirteen markers (Table 2.4, Chapter 2) present in the CHED2 locus on chromosome 20p13 (Hand et al., 1999) were genotyped. Two-point linkage analyses showed evidence of significant linkage with lod scores of 9.5, 9.6, and 9.7 with

markers D20S906, D20S193 and D20S97 (Table 4.1). The maximum lod score obtained was 11.1 with marker D20S117 at θ=0.0. Haplotype analysis indicated that all the markers telomeric of D20S895 (Figure 4.1) were homozygous in affected members and heterozygous in unaffected members of 8 families. In family CH-15 (Figure 4.2) the affected member (CH15-4) showed recombination with marker D20S889 indicating this as the centromeric boundary of the CHED2 locus. The proband from family CH-37 was heterozygous at the marker D20S198 (Figure 4.2) suggesting that this is the telomeric boundary. The lod score for this marker was 2.6 at recombination fraction [θ] of 0.0 which increased to 6.5 at θ = 0.01. Based on these observations the AR CHED critical region was narrowed down to a region of 2.2cM (1.3Mb) between markers D20S889 and D20S198 (Figure 4.1).

Table 4.1. Linkage analysis of CHED2.

Marker	**Lod score at θ =**							**Z_{max}**	**θ_{max}**
	0	**0.01**	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**		
D20S117	**11.1**	**10.8**	**9.7**	**8.3**	**5.6**	**3.2**	**1.3**	**11.1**	**0.00**
D20S906	**9.5**	**9.2**	**8.2**	**6.9**	**4.5**	**2.4**	**0.8**	**9.5**	**0.00**
D20S198	**2.6**	**6.5**	**6.3**	**5.5**	**3.6**	**1.9**	**0.6**	**6.6**	**0.02**
D20S193	**9.6**	**9.4**	**8.4**	**7.1**	**4.7**	**2.5**	**0.9**	**9.6**	**0.00**
D20S889	**-∞**	**8.9**	**8.5**	**7.4**	**5.0**	**2.7**	**1.0**	**8.9**	**0.01**
D20S97	**7.6**	**7.4**	**6.5**	**5.5**	**3.5**	**1.8**	**0.6**	**7.6**	**0.00**
D20S895	**-1.1**	**4.0**	**4.7**	**4.3**	**2.9**	**1.5**	**0.5**	**4.7**	**0.04**
D20S835	**-5.4**	**1.7**	**3.1**	**3.1**	**2.3**	**1.3**	**0.5**	**3.2**	**0.07**
D20S882	**-12.7**	**-1.2**	**1.5**	**2.0**	**1.6**	**0.8**	**0.2**	**2.0**	**0.09**
D20S905	**-9.2**	**1.2**	**3.1**	**3.3**	**2.6**	**1.5**	**0.6**	**3.3**	**0.09**
D20S194	**-18.0**	**-2.6**	**0.8**	**1.7**	**1.7**	**1.0**	**0.3**	**1.7**	**0.11**
D20S115	**-∞**	**0.1**	**2.7**	**3.1**	**2.6**	**1.6**	**0.6**	**3.1**	**0.09**
D20S177	**-∞**	**-0.4**	**2.2**	**2.7**	**2.3**	**1.5**	**0.6**	**2.7**	**0.10**

The table summarizes the lod scores of the thirteen markers genotyped (in the region of CHED2 locus on chromosome 20p13) at recombination fraction (θ) ranging from 0.0 to 0.4. Z_{max} indicates the maximum lod score and θ_{max} indicates the value of θ at which lod score is maximum.

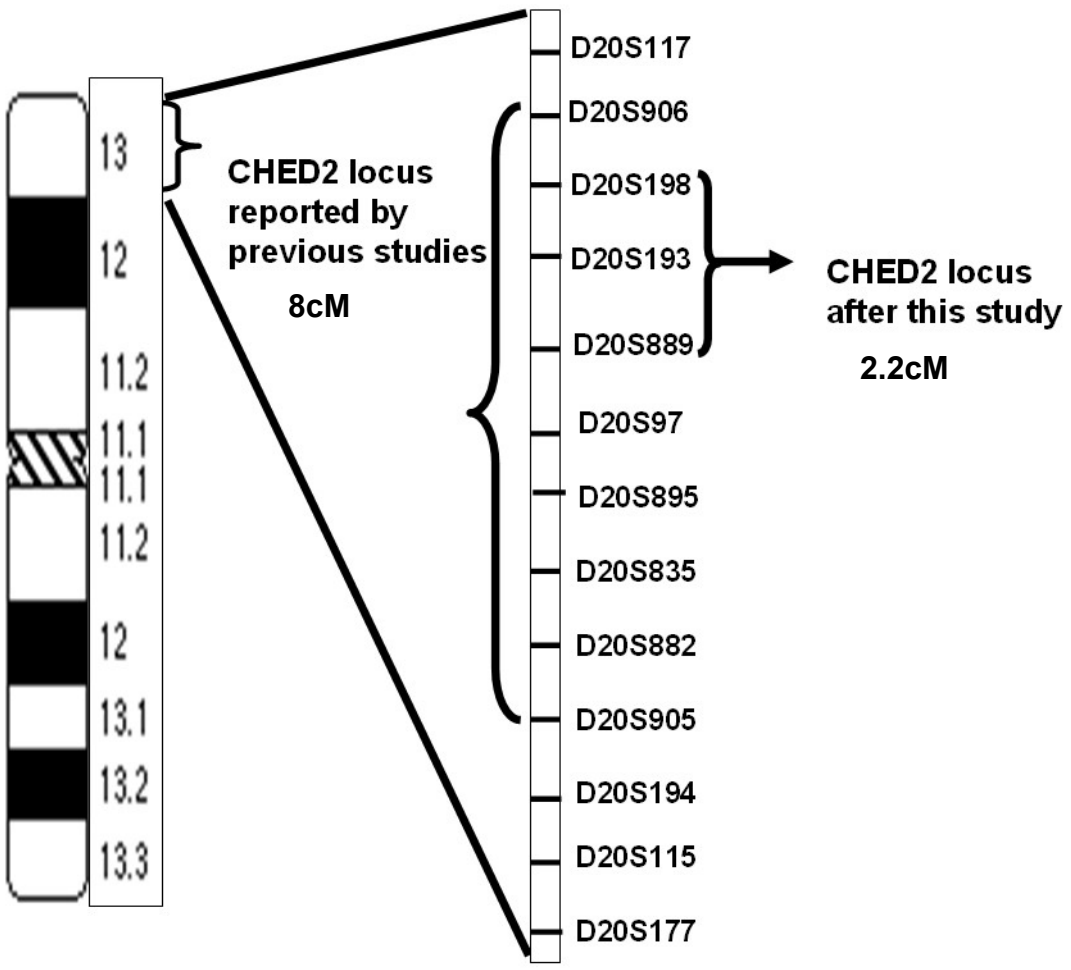


Figure 4.1. Ideogram of the pericentric region of chromosome 20 showing relative location of microsatellite markers used in linkage.

4.3 SCREENING OF CANDIDATE GENES

Thirteen genes located within the critical interval were screened for mutations in 8 probands and 8 unaffected members (one from each family). The genes that were screened are listed in Table 2.1 Chapter 2. A total of 11 SNPs were identified in five genes (TMC2, KIAA1442, C20orf116, C20orf28 and OXT; shown in Table 4.2). Of the 11 SNPs identified, four were found in the coding region. One SNP found in TMC2 in exon 5 (CAA>CGA, Gln205Arg) was non-synonymous. Synonymous SNPs identified consists of UGA>UAA (X163X) identified in C20orf28 in one patient, GAC>GAU (Asp527Asp) identified in the TMC2 gene in 2 patients, CCG>CCA (Pro648Pro) identified in KIAA1442 in 1 patient. All the SNPs identified are previously reported except one in KIAA1442 in intron 15 (g.60055G>A), which was found in three patients and one normal control. Screening of the remaining 27 genes within the 1.3 Mb critical interval for CHED2 was carried out by Dr. J. F. Hejtmancik and co-workers at the National Eye Institute, MD, USA, as part of a collaboration, and these are not detailed in this thesis. Analysis on 16 families with AR CHED (including the 10 families used in linkage analysis) revealed pathogenic alterations in the sodium bicarbonate transporter-like solute carrier family 4 member 11 (*SLC4A11)* gene in 12 out of 16 families. Nine different mutations were identified in 12 families (CH-4, CH-13, CH-14, CH-15, CH-24, CH-26, CH-29, CH-35, CH-37, CH-39, CH-43, CH-44) including 1 deletion, 2 nonsense, 2 complex, and 4 missense mutations. These mutations are detailed in Tables 4.3, 4.4 and 4.5, along with other families screened subsequently and described in section 4.4. Affected individuals were homozygous for

the mutations. One non-pathogenic homozygous variant, Ala160Thr, was found in probands of two families CH-26 and CH-15 and in one unaffected family member.

Table 4.2. SNPs identified in candidate genes.

Gene	Change in DNA	SNP (dbSNP id)	Intron / Exon	Change in protein if any	No of patients/ families
TMC2	g.26017 C>T	rs7270277	Intron 4		2
	g.36232 A>G	rs11908093	Exon 5	Gln205Arg	1
	g.36274 T>C	rs1883980	Intron 5		2
	g.44064 T>G	rs2422794	Intron 7		1
	g.74580 C>T	rs6050576	Exon 12	Asp527Asp	2
	g.80110 G>A	rs6050622	Intron 14		1
KIAA1442	g.60023 G>A	rs2235813	Exon 15	Pro648Pro	1
	g.60055 G>A	Novel	Intron 15	-	3
C20orf116	g.9906 A>G	rs2295552	Intron 4	-	2
C20orf28	g.3156 G>A	ENSSNP1894 814	Exon 4	X163X	1
OXT	g.1288 G>/A (in-del)	rs34097556	Intron 2	-	2

Table summarizes the SNPs identified in candidate genes screened in the CHED2 locus. Genes listed are transmembrane channel-like 2 (TMC2), early B-cell factor 4 transcription factor, (KIAA1442) chromosome 20 open reading frame 116 (C20orf116), chromosome 20 open reading frame 28 (C20orf28) and oxytocin (OXT).

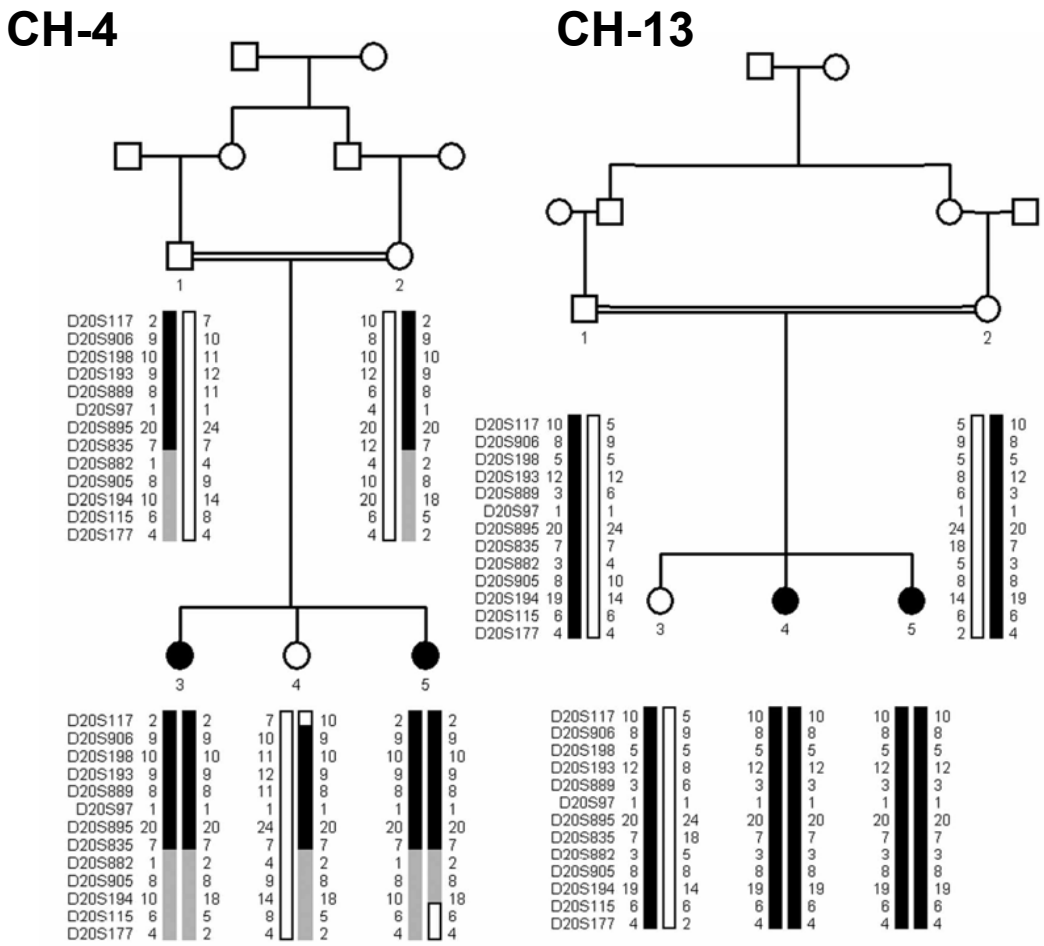


Figure 4.2. Pedigrees of 10 consanguineous families (CH-4, CH-13, CH-15, CH-22, CH-24, CH-26, CH-37, CH-43, CH-44 and CH-49) used in linkage and haplotype analysis. The dark bars correspond to affected haplotypes including alleles, which co-segregate with the disease and are homozygous in affected individuals. Grey bars indicate alleles not homozygous in affected individuals while white bars indicate recombinant alleles.

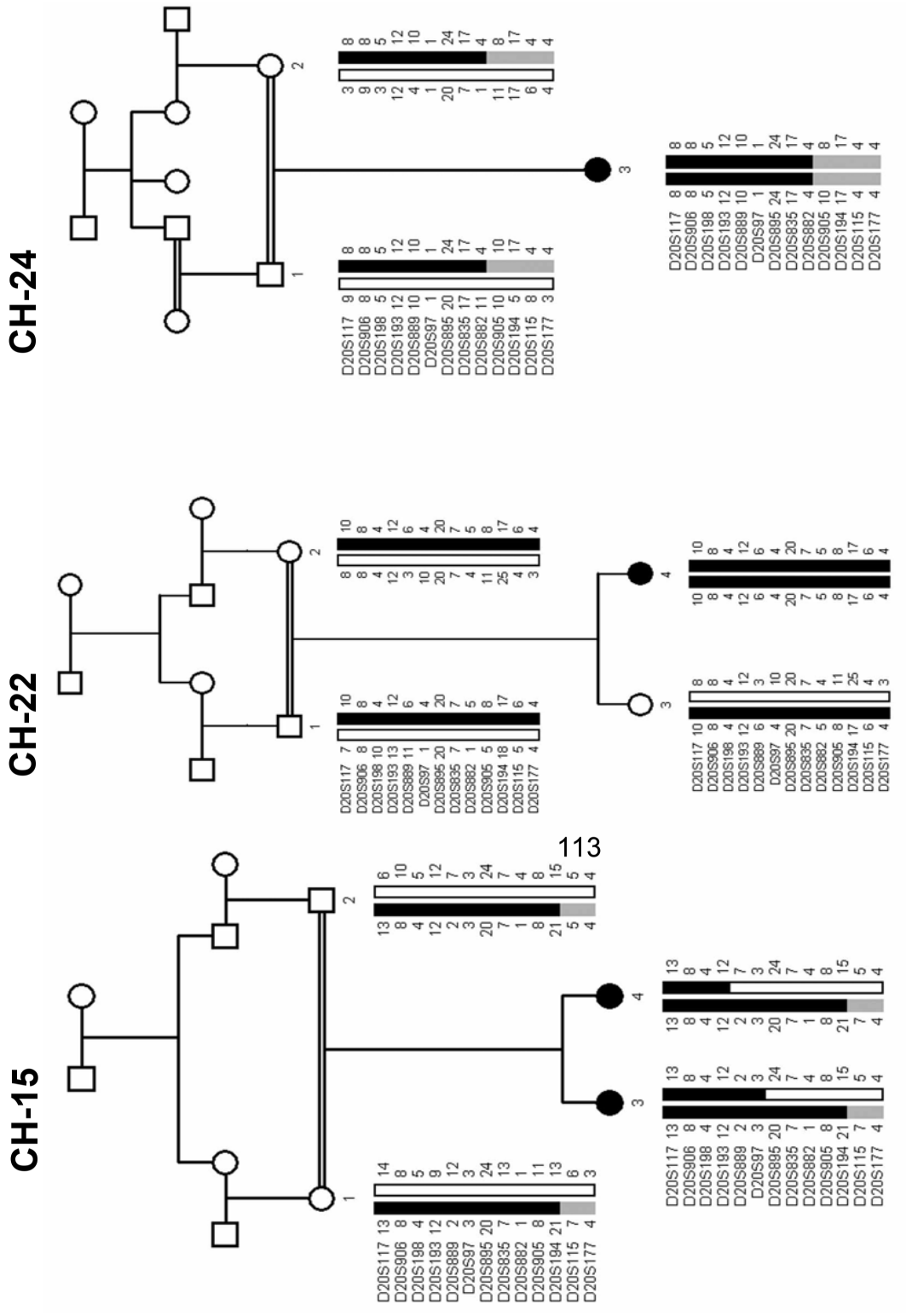


Figure 4.2

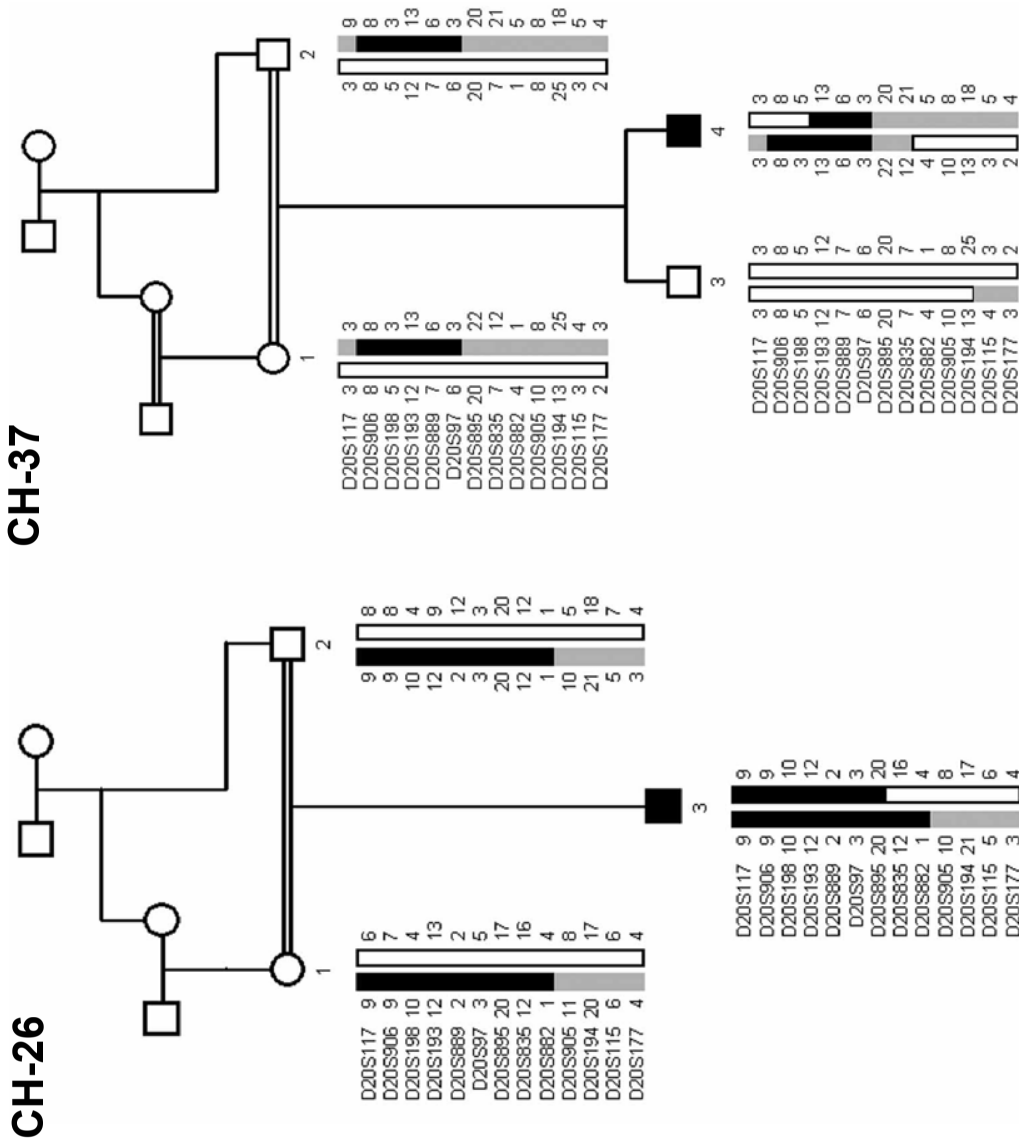

CH-37
CH-26
D20S117
D20S906
D20S198
D20S193
D20S889
D20S97
D20S895
D20S835
D20S882
D20S905
D20S194
D20S115
D20S177

Figure 4.2

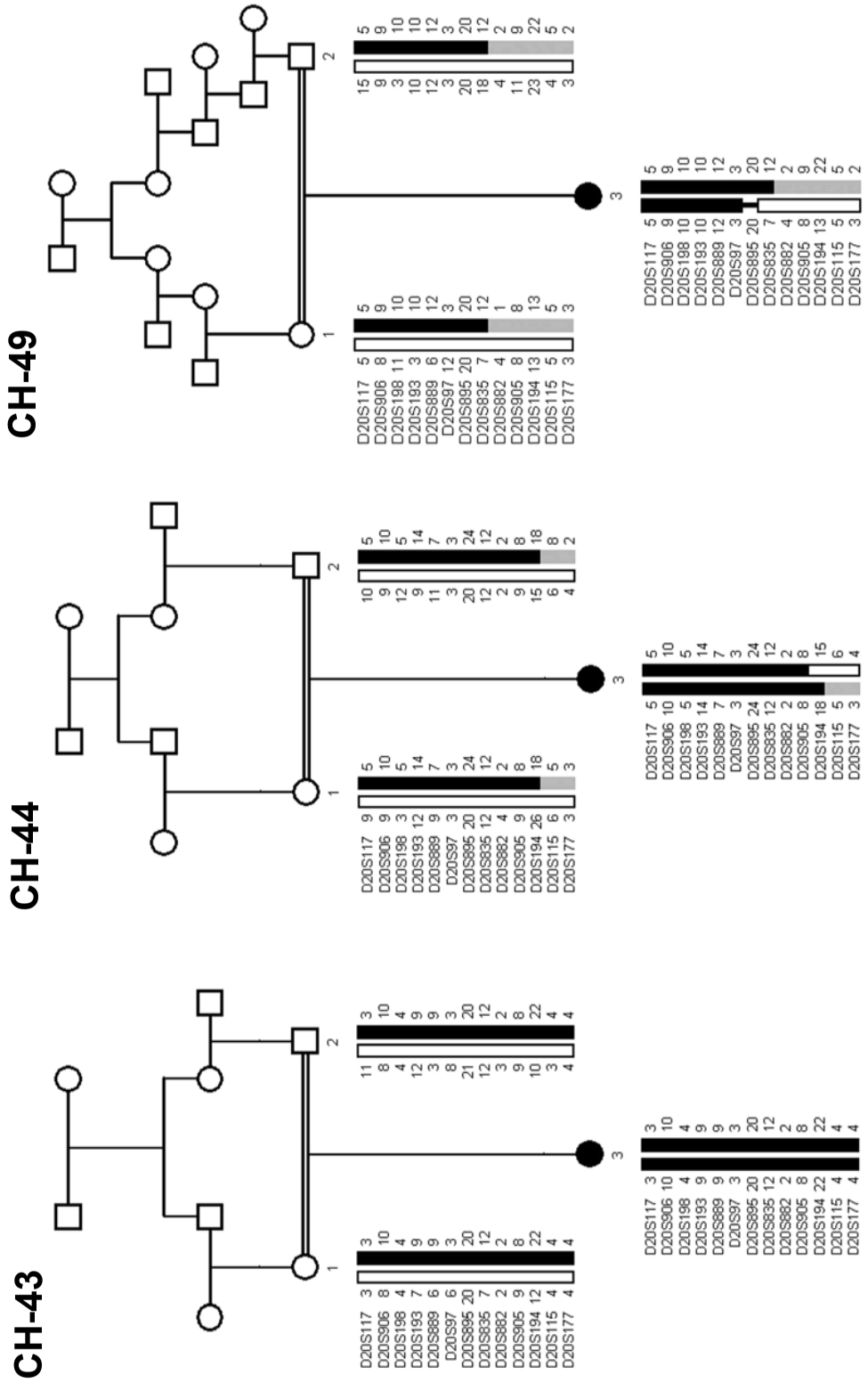

CH-49
CH-44
CH-43
D20S117
D20S906
D20S198
D20S193
D20S889
D20S97
D20S895
D20S835
D20S882
D20S905
D20S194
D20S115
D20S177

Figure 4.2

4.4 FURTHER SCREENING OF THE *SLC4A11* GENE

We recruited a relatively large cohort of patients and family members with recessive CHED. We wanted to characterize the mutations in these families in order to know the extent of allelic heterogeneity and to determine if there was any locus heterogeneity in AR CHED. 44 unrelated families (including fifty-three affected and 76 unaffected family members) were enrolled. Families CH-19, CH-20, CH-49, and CH-45 were screened in the initial analysis described in the preceding section (section 4.3). They were again included in this analysis since all exons could not be screened completely in the initial phase due to technical flaws in amplification and we could not completely rule out the presence of mutations in them. Screening of *SLC4A11* was carried out by the method of SSCP followed by sequencing of PCR products. A total of 28 different mutations were identified in 36 families with AR CHED. Out of the 28 mutations identified, there were 8 deletions, 1 complex mutation consisting of deletion and insertion, 5 nonsense, 13 missense mutations and one splice site mutation (Table 4.3). No mutations were identified in 8 families by the method of SSCP.

Of the 36 families with mutations identified, cosegregation was tested by sequencing or RFLP or SSCP analyses on family members from 31 families. Fourteen families had both the parents with or without other unaffected family members. Fifteen families had only one parent. Two families had unaffected family members other than parents. The remaining five families had no available family members. All of these mutations were tested for presence or absence in a control

population of 50 unrelated normal individuals by restriction enzyme digestion or SSCP analysis.

Details of mutations in *SLC4A11* described in the following sections pertain to the results on screening of 44 families mentioned above as well as the 12 families screened previously during identification of the CHED2 gene (described in section 4.3).

4.4.1 Deletion mutations

Seven different deletion mutations were found in nine families (Figure 4.3), with 8 families being homozygous for the deletions. Five deletions occurring in 7 families were predicted to cause frameshifts (detailed in Table 4.3). One of these five deletions was a single heterozygous deletion of c.306 del C found in CH-42 (Table 4.3). In this family the second mutation was not identified by SSCP analysis or by direct sequencing of 15 exons. Two in-frame deletions were found in two families CH-50 and CH-31 (Table 4.3), consisting of deletions of one and four amino acids respectively.

4.4.2 Complex mutations

Three different complex mutations consisting of deletion plus insertion were found in three families (CH-12, CH-24, CH-4; Figure 4.4; Table 4.3), all homozygous, and potentially leading to frameshifts.

Table 4.3. Deletion and complex mutations identified in *SLC4A11* among patients with AR CHED.

S.NO	Family#	Affected	Unaffected	Mutation in c.DNA	Mutation in protein	Exon/Intron	Restriction site change if any	Ref if reported
1	CH-34	1	0	c.140delA	p.Tyr47SerfsX69	2		This study
2	CH-42	1	2 [F, M]	c.[306delC]+[?] *	p.[Gly103ValfsX13]+[?]	3	*MspI* (-)	This study
3	CH-2	2	2 [F, U]	c.473_480del8	p.Arg158GlnfsX4	4	*ItaI* (-)	Desir et al., 2007
4	CH-5	1	1 [M]	c.473_480del8	p.Arg158GlnfsX4	4	*ItaI* (-)	Desir et al., 2008
5	CH-19	2	2 [F, M]	c.618_619delAG	p.Val208AlafsX38	5		This study
6	CH-45	1	4 [F, M, S]	c.618_619delAG	p.Val208AlafsX38	5		This study
7	CH-31	1	1 [F]	c.878_889del12bp	p.Glu293_Glu296del	7	*EcoRI* (-)	This study
8	CH-13	2	3 [F, M, S]	c.1704_1705delCT	p.Ser569ArgfsX177	13		This study
9	CH-50	1	2 [F, M]	c.2389_2391delGAT	p.Asp797del	17	*TaqI* (-)	This study
10	CH-24	1	3 [F, M, S]	c.246_247delTTinsA	p.Phe84LeufsX33	2		This study
11	CH-12	1	3 [F, M, S]	c.1317_1322delGCTCTA ins AGTGGGTG	p.Leu440ValfsX6	10	*HaeII* (-)	This study
12	CH-4	2	4 [F, M, S]	c.2420delTinsGG	p.Leu807ArgfsX71	17		This study

Families with AR CHED in which deletion and complex mutations were identified are shown with numbers of affected and unaffected members available for the study. F, M, S, U & G refer to father, mother, sibling uncle and grandparent, respectively. The sign (-) indicates loss of restriction site. Mutation nomenclature is according to the recommendations in: http://www.hgvs.org/mutnomen/recs.html. *Family CH-42 had a single heterozygous mutant allele.

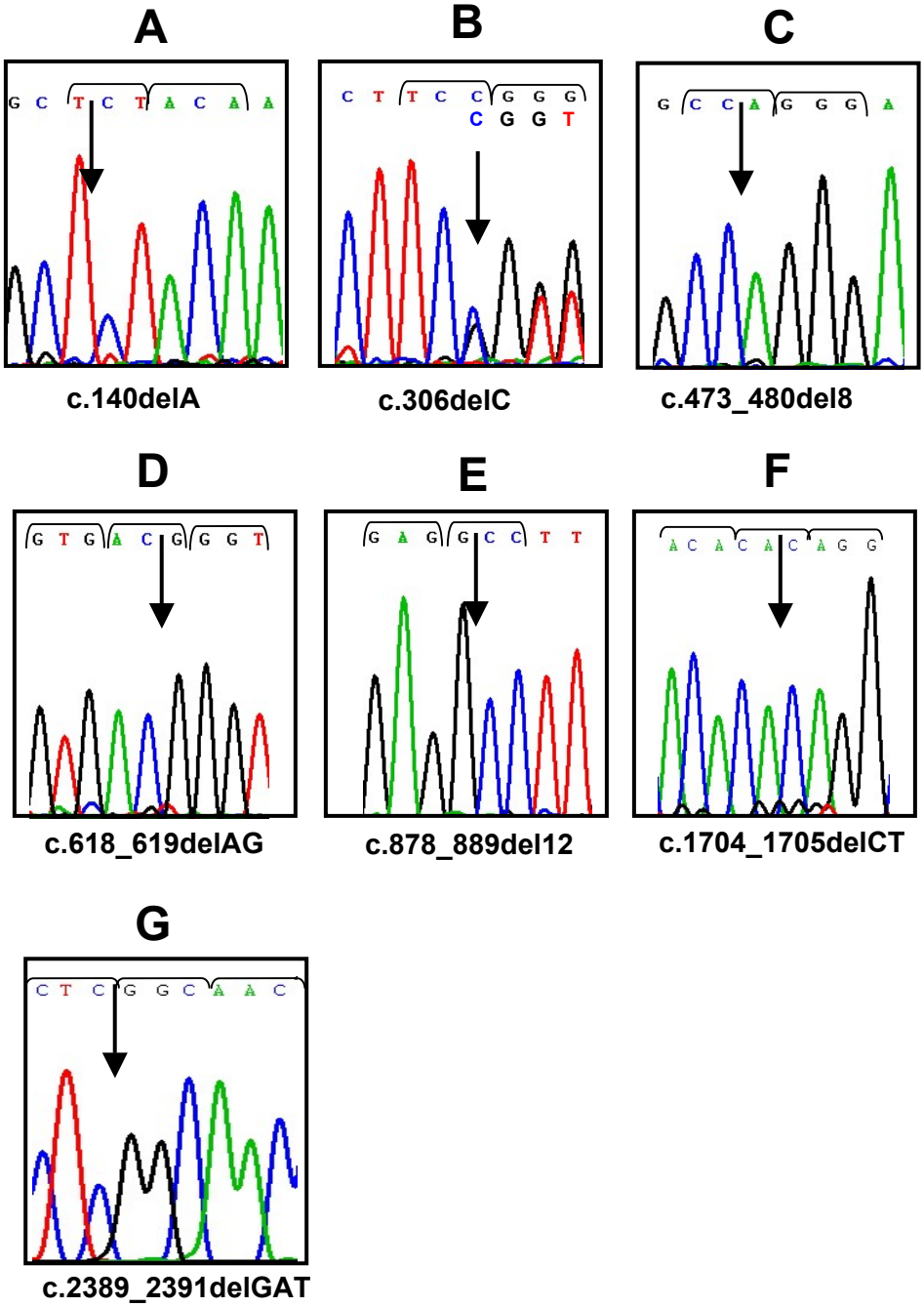


Figure 4.3. Sequence electropherograms of deletion mutations. Figures show homozygous mutations identified in families CH-34 (A), CH-2 & CH-5 (C), CH-19 & CH-45 (D), CH-31 (E), CH-13 (F) and CH-50 (G). B: Heterozygous mutation identified in CH-42 with sequence of normal allele shown below that of the mutant allele. Codons are marked by bracketts; Arrows show the sites of deletion. Details are in Table 4.3.

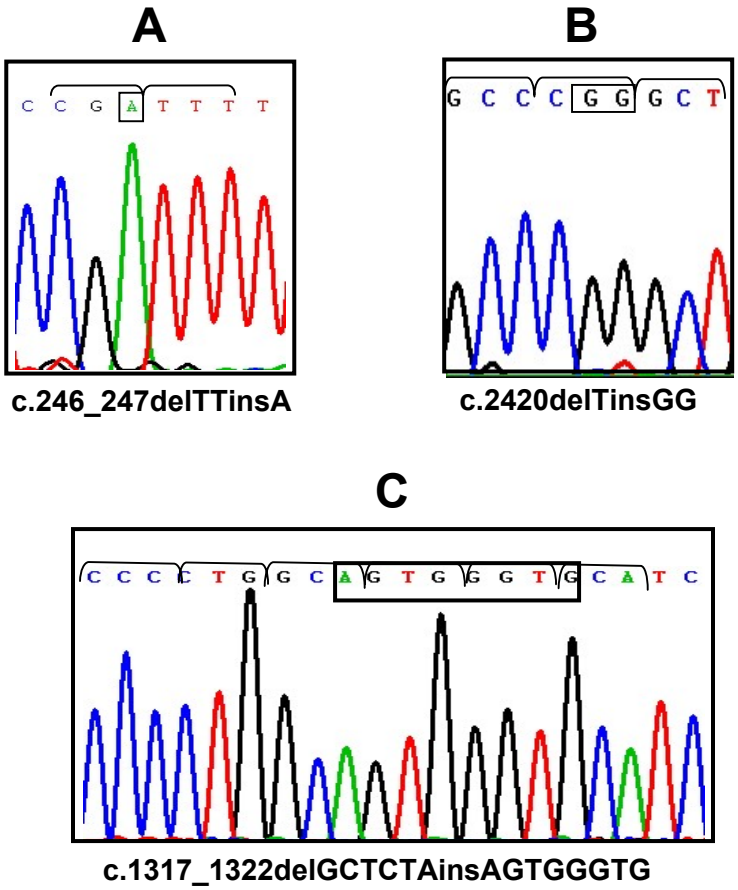


Figure 4.4. Sequence electropherograms of homozygous complex mutations. Complex mutations identified in families CH-24 (A), CH-4 (B) and CH-12 (C) are shown. Residues inserted are marked by boxes. Codons are bracketted. Details are in Table 4.3.

4.4.3 Nonsense mutations

Five different nonsense mutations were found in 12 families (Figure 4.5; Table 4.4). Four of the five mutations (Arg112X, Arg605X, Glu632X, Gln803X) were homozygous in 8 families. Two probands showed compound heterozygosity for Arg112X and a second missense change. Family CH-28 had a single heterozygous mutation of Arg605X. In this family the other heterozygous mutant allele was not identified by SSCP analysis /direct sequencing of the entire coding region. A fifth nonsense mutation, Arg875X occurred as a compound heterozygous mutation with a missense mutant allele in one family CH-30 (details in Table 4.4). Compound heterozygousity of mutations in the three families mentioned was confirmed by the analysis of unaffected family members who had a single heterozygous mutation each (both the parents from CH-40 and CH-30 and only the father from family CH-21 [Table 4.4]).

Table 4.4. Nonsense mutations identified in *SLC4A11* among patients with AR CHED.

S.NO	Family#	Affected	Unaffected	Mutation in c.DNA	Mutation in protein	Exon/Intron	Restriction site change if any	Ref if reported
1	CH-52	1	3 [F, M, S]	c.334C>T	p.Arg112X	3		This study
2	CH-21	1	1 [F]	c.[334C>T]+ [1751C>A]**	p.[Arg112X]+ [Thr584Lys]	3+13	*AluI* (+)	This study
3	CH-40	1	4 [F, M, G, U]	c.[334C>T]+ [2318C>T]**	p.[Arg112X]+ [Pro773Leu]	3+17	*HapII* (-)	This study
4	CH-15	2	2 [F, M]	c.1813C>T	p.Arg605X	14		Vithana et al., 2006
5	CH-26	1	2 [F, M]	c.1813C>T	p.Arg605X	14		Vithana et al., 2007
6	CH-28	1	0	c.[1813C>T]+[?] *	p.[Arg605X]+[?]	14		Vithana et al., 2008
7	CH-63	1	0	c.1813C>T	p.Arg605X	14		Vithana et al., 2009
8	CH-65	1	0	c.1813C>T	p.Arg605X	14		Vithana et al., 2010
9	CH-35	1	3 [F, M, U.]	c.1894G>T	p.Glu632X	14		This study
10	CH-38	1	1 [F]	c.2407C>T	p.Gln803X	18	*MaeI* (+)	This study
11	CH-48	1	1 [G]	c.1894G>T	p.Glu632X	14		This study
12	CH-30	1	2 [F, M]	c.[2264G>A]+ [2623C>T]**	p.[Arg755Gln]+ [Arg875X]	17+19		Vithana et al., 2006+This study

Table summarizes the nonsense mutations identified in AR CHED families. The signs (+) or (-) indicates the gain or loss of restriction site. *Family CH-28 had a single heterozygous mutant allele. **Families CH-20, CH-40 and CH-30 had compound heterozygous mutations.

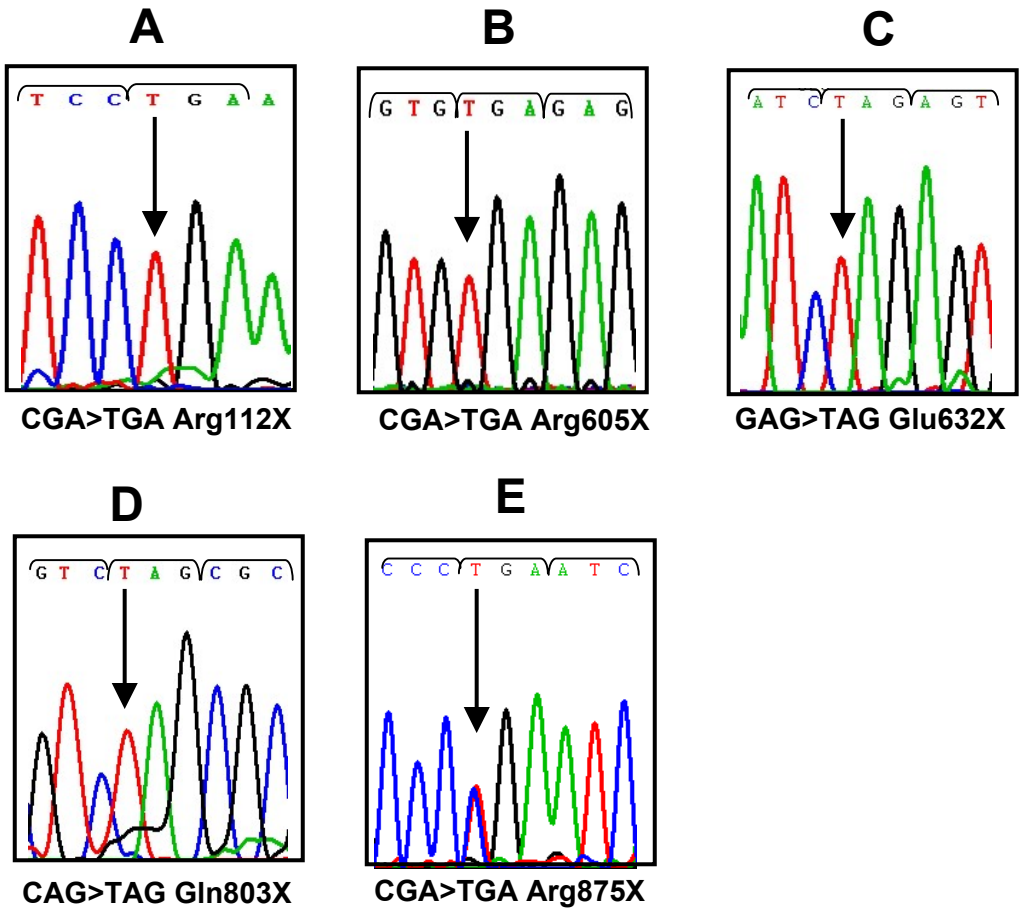


Figure 4.5. Sequence electropherograms of nonsense mutations with codon and amino acid changes shown below each panel. Homozygous nonsense mutations identified in families CH-52 (A), CH-15, CH-63, CH-65, CH-26 (B), CH-48 & CH-35 (C), CH-38 (D). E: Heterozygous nonsense mutation identified in family CH-30. Mutated residues are marked by arrows. Codons are marked by bracketts. Details are in Table 4.4.

4.4.4 Missense mutations

A total of 17 different missense mutations were identified in 21 families. (listed in Table 4.5; Figure 4.6). Twelve mutations that are novel from this study are shown in Figure 4.6. Fourteen mutations identified in 19 families were homozygous. In family CH-17 a compound heterozygous mutation was identified (Figure 4.6) consisting of leucine-473 to arginine (Leu473Arg) and threonine-401 to lysine (Thr401Lys). In this family the unaffected father carried a single heterozygous mutation of Thr401Lys only, suggesting that the two alleles in the proband were indeed in *trans*. A heterozygous missense change of glutamine-199 to glutamic acid (Gln199Glu) was found in CH-59. The second mutated allele was not identified in this family by SSCP analysis.

4.4.5 Splice site and intronic mutations

A homozygous substitution involving the splice acceptor dinucleotide at the intron 8-exon 9 junction (c.1091-1G>C) was identified (Figure 4.7A; Table 4.5) in family CH-53. This is predicted to inactivate the acceptor site based on the high degree of conservation of the acceptor dinucleotide. Analysis of the splice potential of the mutant using neural network–based method for splice site detection (URL is provided in Appendix 2) showed complete loss of acceptor site in the mutant as compared to wild type sequence at a threshold score of > 85% and with false positive rate of 1.4%.

An intronic homozygous deletion of 19 bp was identified in 2 families, CH-36 and CH-47, encompassing residues +26 to +44 of intron 7 (Figure 4.7B). This

mutation cosegregated with disease (unaffected parents of family CH-36 and the unaffected father of family CH-47 were heterozygous for this mutation). It was not detected by 8% polyacrylamide gel electrophoresis in 50 unrelated normal controls, suggesting that it may be pathogenic.

Table 4.5. Missense and intronic mutations identified in *SLC4A11* among patients with AR CHED.

S.NO	Family#	Affected	Unaffected	Mutation in c.DNA	Mutation in protein	Exon/Intron	Restriction site change if any	Ref if reported
1	CH-59	2	0	[c.595C>G]+[?] *	p. [Gln199Glu]+[?]	3	*TaqI* (+)	This study
2	CH-3	2	6 [F, M, S, U]	c.625C>T	p.Arg209Trp	5		This study
3	CH-20	1	4 [F, M, S]	c.625C>T	p.Arg209Trp	5		This study
4	CH-1	1	2 [F, M]	c.638C>T	p.Ser213Leu	5		This study
5	CH-27	1	1 [U]	c.697C>T	p.Arg233Cys	6	*BbvI* (+)	This study
6	CH-61	1	1 [M]	c.1253G>A	p.Gly418Asp	10	*BveI* (+)	This study
7	CH-17	1	1 [F]	c.[1202C>A]+ [1418T>G]**	p.[Thr401Lys]+ [Leu473Arg]	9+11	*AccII* (+)	This study
8	CH-49	1	2 [F, M]	c.1466C>T	p.Ser489Leu	12		Vithana et al., 2006
9	CH-16	2	1 [M]	c.1751C>A	p.Thr584Lys	13	*AluI* (+)	This study
10	CH-11	1	3 [F, M, S]	c.2263C>T	p.Arg755Trp	17		Vithana et al., 2006
11	CH-57	2	1 [F]	c.2318C>T	p.Pro773Leu	17	*HapII* (-)	This study
12	CH-14	2	4 [F, M, S]	c.2411G>A	p.Arg804His	17		This study
13	CH-8	1	1 [F]	c.2470G>A	p.Val824Met	18	*BshNI* (-)	Desir et al., 2007

S.NO	Family#	Affected	Unaffected	Mutation in c.DNA	Mutation in protein	Exon/Intron	Restriction site change if any	Ref if reported
14	CH-33	1	1 [M]	c.2470G>A	p.Val824Met	18	*BshNI* (-)	Desir et al., 2008
15	CH-29	1	2 [F, M]	c.2498C>T	p.Thr833Met	18		This study
16	CH-37	1	3 [F, M, S]	c.2264G>T	p.Arg755Gln	17		Vithana et al., 2006
17	CH-32	1	1 [M]	c.2605C>T	p.Arg869Cys	18		Vithana et al., 2006
18	CH-58	2	1 [F]	c.2605C>T	p.Arg869Cys	18		Vithana et al., 2007
19	CH-39	1	2 [F, M]	c.2606G>A	p.Arg869His	18		This study
20	CH-43	1	2 [F, M]	c.2606G>A	p.Arg869His	18		This study
21	CH-44	1	2 [F, M]	c.2498C>T	p.Thr833Met	18		This study
22	CH-36	1	2 [F, M]	c.996+26C_+44Cdel19	Not known	Intron 7	PAGE	This study
23	CH-47	1	1 [F]	c.996+26C_+44Cdel19	Not known	Intron 7	PAGE	This study
24	CH-53	1	1 [M]	c.1091-1G>C	Not known	Intron 8	*MaeI* (-)	This study

Table summarizes the missense and intronic mutations identified in AR CHED families. The signs (+) or (-) indicates the gain or loss of restriction site. *Family CH-59 had a single heterozygous mutant allele. **Family CH-17 had a compound heterozygous mutation.

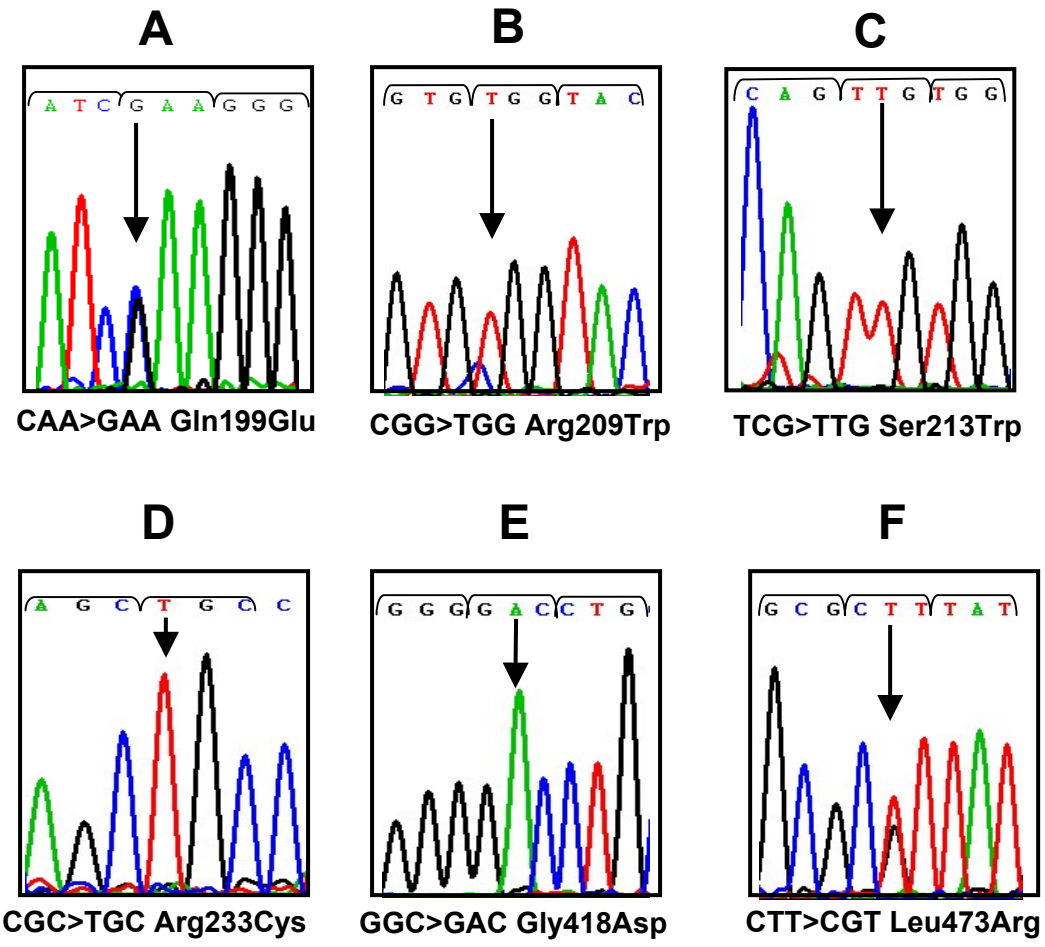
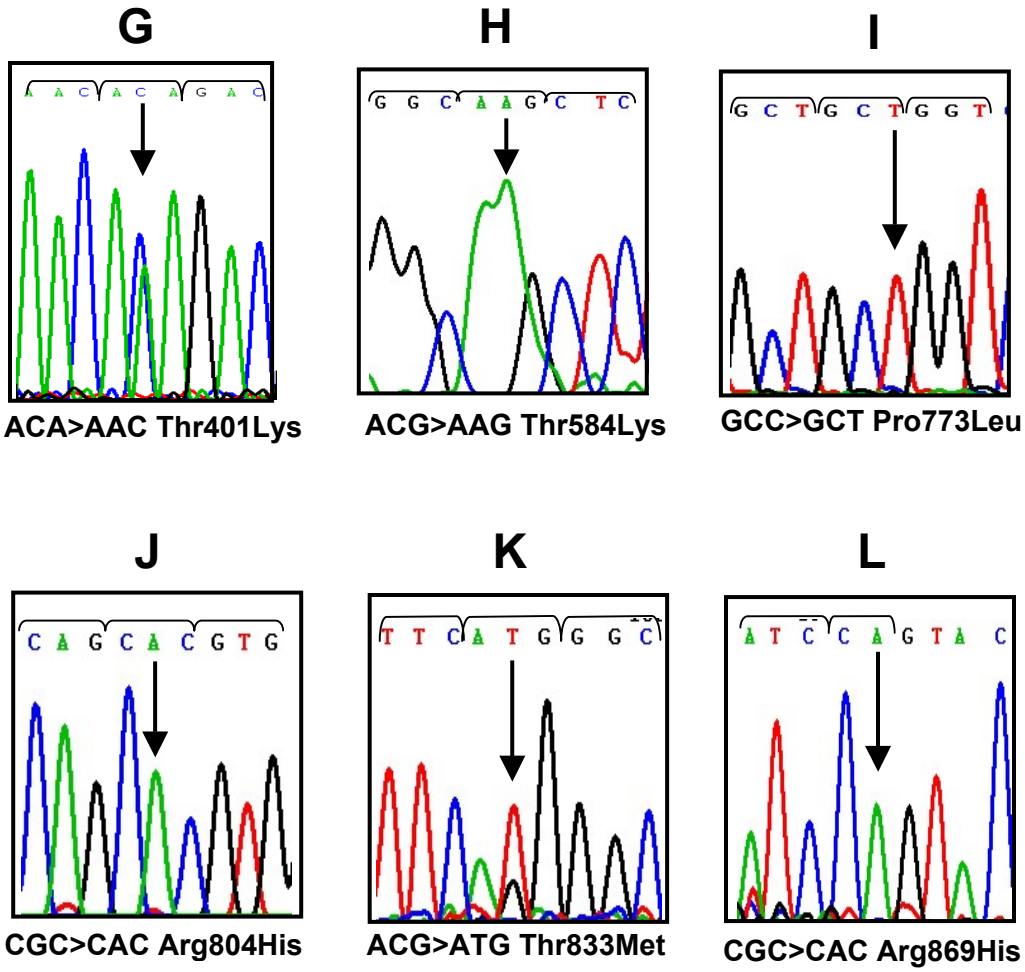


Figure 4.6. Sequence electropherograms of novel missense mutations with codon and amino acid changes shown below each panel. Homozygous mutations identified in families CH-3 & CH-20 (B), CH-1 (C), CH-27 (D), CH-61 (E), CH-16 (H) CH-57 (I), CH-14 (J), CH-29 & CH-44 (K) and CH-39 & 43 (L). Heterozygous mutation identified in families CH-59 (A) and CH-17 (F&G). Mutated residues are marked by arrows. Codons are marked by bracketts. Details are in Table 4.5.

Figure 4.6


G
H
I
ACA>AAC Thr401Lys
ACG>AAG Thr584Lys
GCC>GCT Pro773Leu
J
K
L
CGC>CAC Arg804His
ACG>ATG Thr833Met
CGC>CAC Arg869His

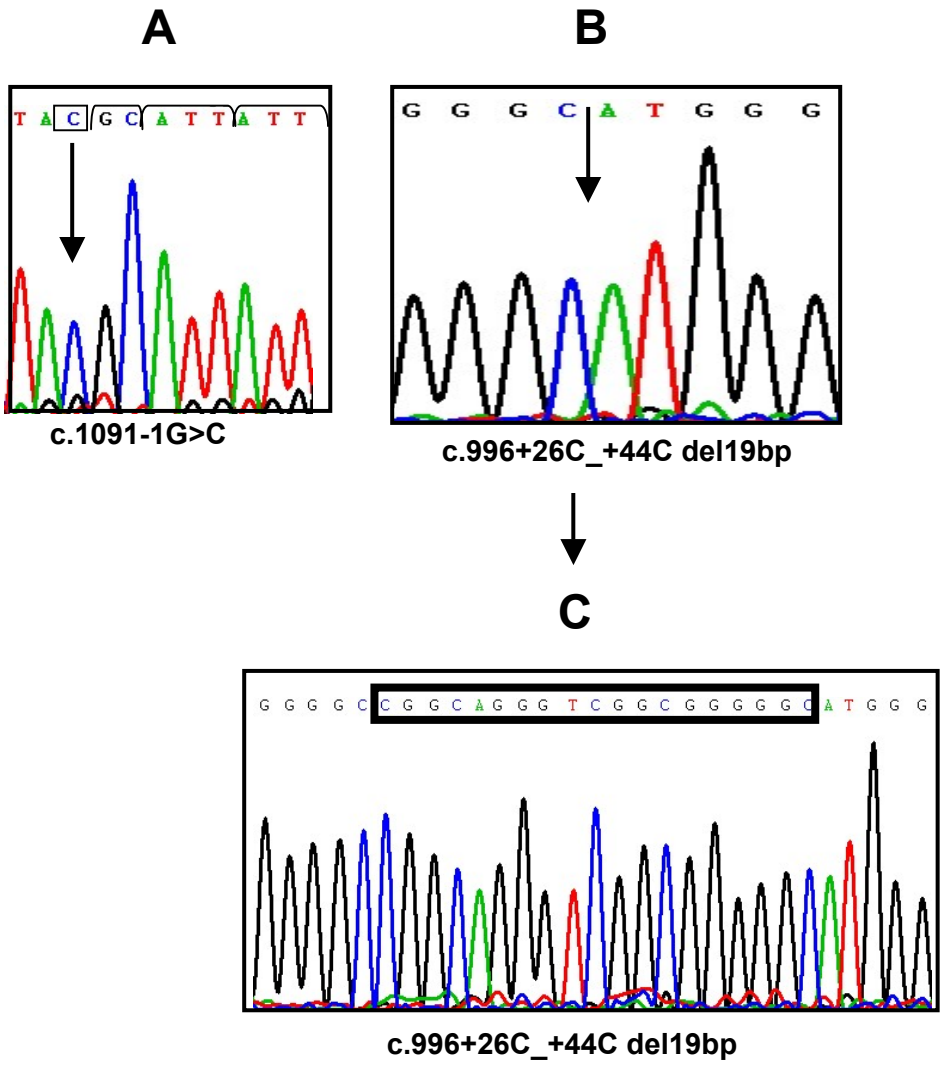


Figure 4.7. A: Sequence electropherogram of mutation involving splice acceptor site A**G**>A**C** (mutated residue C is boxed) identified in family CH-53. Sequence electropherograms of proband of family CH-36 with homozygous deletion of 19 bp in intron 7 (B) and control (C). The sequence deleted is boxed in C. Arrow shows the site of deletion. Details are in Table 4.5.

4.5 MULTIPLE SEQUENCE ALIGNMENT

Multiple sequence alignment was carried out on the *SLC4A11* proteins from different species (xenopus, chick, mouse and human) and representative members of the human SLC4 bicarbonate family (*SLC4A1, SLC4A4, SLC4A7, SLC4A8, SLC4A9*). This was done in order to look at evolutionary conservation of residues involved in missense changes identified in patients. Twelve residues (Gln199Glu, Arg209Trp, Ser213Leu, Arg233Cys, Thr401Lys, Gly418Asp, Leu473, Thr584Leu, Pro773Leu, Arg804His, Thr833Met, Arg869His) found to have novel missense changes were examined. Alignments are shown in Table 4.6A &B. Amino acid residues with the missense mutations Thr401Lys, Gly418Asp, Thr584Leu, Pro773Leu, Arg804His, Thr833Met and Arg869His were highly conserved across *SLC4A11* proteins of different species and also across representative members of the human SLC4 family (Table 4.6A & B). Amino acid residues involved in the missense mutations Gln199Glu, Ser213Leu, Arg233Cys and Leu473Arg were highly conserved across *SLC4A11* proteins of different species (xenopus, chick, mouse and human). These residues were not conserved across members of human SLC4 family. However the Blossum-80 scores for substitutions in Ser213Leu, Arg209Trp, Arg233Cys and Leu473Arg were -4, -5, -6 and -4 respectively. Such high negative scores of Blossum-80 suggest that these substitutions are unlikely to occur. The Gln to Glu change in Gln199Glu has a positive Blossum-80 score of 2. However the replacement of glutamine (polar neutral amino acid) by glutamic acid (polar negatively charged amino acid) may have an adverse effect by introducing a charge into the protein.

Table 4.6A. Sequence alignment of SLC4A showing conservation of amino acids mutated in AR CHED

SLC4A Protein sequence	Gln199Glu	Arg209Trp	Ser213Leu	Arg233Cyc	Thr401Lys	Gly418Asp	Leu473Arg
SLC4A11 Xenopus	T I **Q** G V	G V **Q** Y Q	Q Q **S** - L	I S **R** L E	E S **T** A G	I G **G** V I	L I **L** Y S
SLC4A11 Chick	T I **Q** G V	G V **Q** Y Q	Q Q **S** - L	I S **R** L E	E N **T** R G	I G **G** L L	L V **M** Y S
SLC4A11 Human	T I **Q** G V	G V **R** Y Q	Q Q **S** - L	I S **R** L V	E N **T** D G	I G **G** L L	L A **L** Y A
SLC4A11 Mouse	T I **Q** G V	G V **Q** Y E	E Q **S** - L	I S **R** L V	E N **T** N G	I G **G** L L	L A **L** Y A
SLC4A11 Dog	T I **Q** G V	G V **Q** Y Q	Q Q **S** - L	I S **R** L V	E N **T** N G	I G **G** L L	L V **L** Y A
SLC4A7 Human	D M **N** F M	G A **E** A S	S N **V** - V	L A **P** A V	E A **T** E G	L T **G** I A	C I **V** L V
SLC4A8 Human	D L **H** F M	G A **E** A S	S N **V** L V	L S **P** A V	E A **T** E G	M T **G** I A	C I **V** L V
SLC4A4 Human	- - **K** F M	D A **E** A S	S N **V** L V	L Q **Q** A V	D A **T** D N	V S **G** A I	C L **I** L V
SLC4A9 Human	- - **P** L R	G A **E** A G	G T **V** L A	L R **N** P V	D A **T** D G	M A **G** A A	C L **V** L V
SLC4A1 Human	- - **G** I L	D S **E** A T	T L **V** L V	L Q **E** A A	E K **T** R N	V Q **G** I L	V V **L** V V

Multiple sequence alignment of SLC4A from different species and representative members of human SLC4 bicarbonate family members showing the conservation of mutated residues. Multiple sequence alignment was built using the software CLUSTALW. GenBank accession nos. are as follows; Xenopus (BJO39327.1), Chick (XP_420881.2), Human (AF336127), Mouse (AAI11885.1), Dog (XP_542919.2), SLC4A7 (NP_003606.2), SLC4A8 (NP_004849.2), SLC4A4 (NP_003750.1), SLC4A9 (NP_113655.1) and SLC4A1 (P02730).

Table 4.6B. Sequence alignment of SLC4A showing conservation of amino acids mutated in AR CHED

SLC4A Protein sequences	Pro773Leu						Arg804His						Thr833Met						Arg869His						Thr584Leu				
SLC4A11 Xenopus	L	L	**P**	Y	P	-	F	E	**R**	V	A	-	Y	F	**T**	G	L	-	P	I	**R**	Y	N	-	F	G	**T**	L	W
SLC4A11 Chick	L	L	**P**	F	P	-	F	E	**R**	V	A	-	Y	F	**T**	G	L	-	P	I	**R**	Y	N	-	L	G	**T**	L	W
SLC4A11 Human	L	L	**P**	F	P	-	V	Q	**R**	V	A	-	Y	F	**T**	G	L	-	P	I	**R**	Y	I	-	L	G	**T**	L	W
SLC4A11 Mouse	L	L	**P**	F	P	-	F	S	**R**	V	A	-	Y	F	**T**	G	L	-	P	I	**R**	Y	N	-	L	G	**T**	L	W
SLC4A11 Dog	L	L	**P**	F	P	-	F	E	**R**	V	A	-	Y	F	**T**	G	L	-	P	I	**R**	Y	N	-	L	G	**T**	L	W
SLC4A7 Human	M	T	**S**	V	-	-	F	D	**R**	I	K	-	I	F	**T**	V	I	-	F	V	**R**	K	L	-	F	T	**T**	F	F
SLC4A8 Human	M	T	**A**	I	-	-	F	D	**R**	L	K	-	L	F	**T**	L	I	-	F	V	**R**	K	V	-	F	T	**T**	F	I
SLC4A4 Human	M	A	**P**	I	-	-	M	D	**R**	L	K	-	L	F	**T**	F	L	-	A	V	**R**	K	G	-	L	G	**T**	Y	T
SLC4A9 Human	L	A	**P**	V	-	-	T	N	**R**	V	K	-	L	F	**T**	A	I	-	G	V	**R**	K	A	-	L	T	**S**	F	F
SLC4A1 Human	M	E	**P**	I	-	-	F	D	**R**	I	L	-	L	F	**T**	G	I	-	P	L	**R**	R	V	-	A	G	**T**	F	F

Multiple sequence alignment of SLC4A from different species and representative members of human SLC4 bicarbonate family members showing the conservation of mutated residues. Multiple sequence alignment was built using the software CLUSTALW. GenBank accession nos. are as follows; Xenopus (BJO39327.1), Chick (XP_420881.2), Human (AF336127), Mouse (AAI11885.1), Dog (XP_542919.2), SLC4A7 (NP_003606.2), SLC4A8 (NP_004849.2), SLC4A4 (NP_003750.1), SLC4A9 (NP_113655.1) and SLC4A1 (P02730).

4.6 CLINICAL AND HISTOPATHOLOGICAL CORRELATIONS

In order to look for genotype-phenotype correlations, clinical and histopathlogic features were examined in 33 probands with identified mutations. Clinical parameters analyzed were age at onset and presentation, pre- and post-operative visual acuities and presence of nystagmus. Histopathologic parameters as assessed on formalin-fixed, PAS-stained corneal sections were thickness of cornea, thickness of Descemet's membrane and endothelial cell count. Values for thickness of cornea, Descemet's membrane and endothelial cell count were compared with those obtained for age-matched normal control corneas as detailed in Chapter 2. These data are shown for 33 probands studied along with mutations in Table 4.8. The mean values for thickness of Descemet's membrane and endothelial cell counts for control corneas are shown in Table 4.7.

Mean endothelial cell counts were reduced in patients as compared with age-matched controls. The mean endothelial cell count in controls was 327 (range 211-481 Table 4.7) whereas in patients it was 62 (range 0-257 Table 4.8). The maximum endothelial cell count of 257 was found in proband from family CH-40 who underwent the first corneal graft at <1 year of age. The endothelial cell count of age-matched normal cornea was 481. Endothelial cell counts of 100 or above were found in four of the five patients who underwent corneal grafts at 1 yr or less. These four patients (CH-36, CH-40, CH-49, CH-61) showed normal or borderline thickening of Descemet's membrane. This could probably be due to a smaller time period between onset of disease and corneal grafting (done within one year of age) so that there was a lesser extent of abnormal deposition of the DM. The maximum thickness of grade

4+ was seen in seven probands (CH-34, CH-2, CH-12, CH-48, CH-38, CH-3 and CH-20; Table 4.8). The age at corneal graft for these patients ranged from 7 to 20 years. Endothelial cells were not detectable in 5 probands belonging to families CH-2, CH-50, CH-28, CH-11 and CH-33. Overall all patients showed much reduced endothelial counts as compared to age-matched controls.

Corneal thickness was measured on formalin-fixed sections from patients and age-matched controls. However no significant differences were found between corneal thickness of patients and controls, possibly due to artefactual changes in thickness during processing of the tissue. The thickness of the DM of AR CHED corneas was graded for each age group assessed as follows: N=normal; +/- = borderline; 1+ to 4+ = increasing thickness. Thickness of the DM in age-matched normal controls as measured under identical conditions are detailed in Table 4.7 and ranged from 2 μm to 8.7μm. Thickness of the DM in patients studied varied from 3 μm to 30 μm. Post-operative visual acuity assessed for 33 probands, ranged from 20/20 (CH-8) to counting fingers (CF) at 1.5 meters (CH-45). Post-operative visual acuity of 20/50 or better was found in 11 probands (Table 4.8). Visual acuities (VA) of 20/100 or worse were found in ten probands. Nystagmus was present in 12 of the 33 patients.

Upon examining the phenotypic features described above in relation to types of mutations, it was found that thickness of DM and visual acuities could be graded into less or more severe while the other features could not be categorized based on extent of disease or severity. Groups of patients having lesser or greater thickening of the DM as well as those with poorer or better visual acuities showed no evident

correlations between their phenotypes and types of mutations in *SLC4A11*. Hence based on the above-mentioned clinical and histopathological data, it is concluded that there were no correlations that were detectable between mutations in *SLC4A11* and clinical or histopathological features.

Table 4.7. Corneal thickness, Descemet's membrane (DM) thickness and endothelial cell counts in normal control corneas.

Age group (no of specimens included)	Corneal thickness (4X) Mean (Range)	DM thickness (100X) Mean (Range)	Endothelial cell count (40X) Mean (Range)
Group I-0-5 yrs (10)	601 (424.24-791.22)	2.9 (1.95-3.64)	379.8 (211-481)
Group II-6-10 yrs (5)	470.28 (264.99-637.83)	3.91 (2.39-4.22)	249.2 (290-339)
Group-III11-15 yrs (1)	330.39	2.78	307
Group-IV 20 and above yrs (4)	550.8(343.57-798.76)	5.75(3.28-8.67)	301.5(203-407)

Quantitative assessment of corneal, DM thickness and endothelial cell count in 20 age-matched normal controls. Normal corneas were divided into 4 groups (I to IV) according to age.

Table 4.8. Mutations and phenotypic features of patients with AR CHED.

S.NO.	Family#	Mutation in SLC4A11 protein	Age at onset (age at first graft, yrs)	Thickness of:		Endothelial cell count	Post-op VA (RE, LE)	Nystagmus
				Cornea	DM			
1	CH-34	p. Tyr47SerfsX69	At birth (20)	+	4+	34	20/125, 20/100	Y
2	CH-42	p. [Gly103ValfsX13]+[?]*	NA (7)	+	1+	11	20/125, 20/200	Y
3	CH-2	p. Arg158GlnfsX4	At birth (9)	+	4+	0	20/60, 20/30	-
4	CH-5	p. Arg158GlnfsX4	NA (13)	+	+/-	17	20/40	-
5	CH-19	p. Val208AlafsX38	At birth (5)	+	2+	150	20/125	Y
6	CH-45	p. Val208AlafsX38	At birth (20)	+	1+	9	CF1.5M	Y
7	CH-31	p. Glu293_Glu296del	At birth (8)	+	N	21	20/25, 20/40	-
8	CH-50	p. Asp797del	At birth (6)	+	1+	0	20/50	-
9	CH-12	p. Leu440ValfsX6	At birth (9)	+	4+	20	20/70, 20/60	-
11	CH-40	p. [Arg112X]+[Pro773Leu]**	At birth (<1)	+	N	257	NA	-
12	CH-21	p. [Arg112X]+[Thr584Lys]**	At birth (4)	+	1+	58	20/80	Y
13	CH-63	p. Arg605X	At birth (<1)	+	2+	4	NA	-
14	CH-65	p. Arg605X	At birth (4)	+	2+	95	20/40, 20/50	-
15	CH-28	p. [Arg605X]+[?]*	At birth (9)	+	N	0	CF1M, 20/70	-
16	CH-48	p. Glu632X	At birth (12)	+	4+	21	20/80	-
17	CH-38	p. Gln803X	At birth (7)	+	4+	3	NA	-
18	CH-30	p. [Arg755Gln]+[Arg875X]	At birth (2)	+	2+	98	20/40	-
19	CH-3	p. Arg209Trp	NA (11)	+	4+	40	20/50, 20/40	-

S.NO.	Family#	Mutation in SLC4A11 protein	Age at onset (age at first graft, yrs)	Thickness of:		Endothelial cell count	Post-op VA (RE, LE)	Nystagmus
				Cornea	DM			
20	CH-20	p. Arg209Trp	At birth (18)	+	4+	30	20/400	Y
21	CH-1	p. Ser213Leu	At birth (10)	+	2+	9	20/400, 20/125	Y
22	CH-27	p. Arg233Cys	At birth (6)	+	2+	22	20/50, 20/40	-
23	CH-61	p. Gly418Asp	At birth (<1)	+	N	160	NA	Y
24	CH-17	p. [Thr401Lys]+[Leu473Arg]	At birth (18)	+	2+	8	NA, 20/100	-
25	CH-49	p. Ser489Leu	At birth (<1)	+	+/-	230	FFL	-
26	CH-16	p. Thr584Lys	At birth (2)	+	+/-	90	20/100, 20/400	Y
27	CH-11	p. Arg755Trp	6 yrs (10)	+	2+	0	20/50, 20/80	-
29	CH-8	p. Val824Met	At birth (9)	+	3+	78	20/30, 20/20	-
30	CH-33	p. Val824Met	NA (16)	+	N	5	20/50, NA	-
31	CH-58	p. Arg869Cys	At birth (5)	+	1-2+	77	20/400	Y
32	CH-32	p. Arg869Cys	NA (13)	+	1+	17	NA	-
33	CH-36	c.996+26C_+44Cdel19bp	At birth (<1)	+	N	110	20/200, NA	Y
34	CH-47	c.996+26C_+44Cdel19bp	At birth (6)	+	N-1+	6	20/50, 20/100	-
35	CH-53	c.1091-1G>C	NA (24)	+	N	24	20/200, 20/160	Y

Shown above are mutations in 33 probands along with corresponding clinical and histopathologic features. 'NA' indicates not available. Corneal thickening is indicated by '+'. For DM thickness, N indicates normal, +/- indicates borderline thickening, and grades 1+ to 4+ represent increasing thickness as compared with corneas from age-matched normal controls. Visual acuities (VA) listed correspond to the best VA obtained after the first corneal graft. RE -right eye, LE-left eye. Nystagmus is shown as present (Y) or absent (-). *Families with a single heterozygous mutant allele. ** Families with compound heterozygous mutations.

4.7 DISCUSSION

Linkage and haplotype analysis on ten consanguineous Indian families showed that the disease in these families maps to the previously reported CHED2 locus on chromosome 20p13 (Callaghan et al., 1999; Hand et al., 1999; Kanis et al., 1999). By identification of critical recombinations, the interval was further refined to 2.2 cM as compared with the 8cM region mapped by Hand and co-workers (Hand et al., 1999) .In consanguineous pedigrees segregating with an autosomal recessive disorder, it would be expected that affected individuals would have genetically linked segments of the genome identical by descent. Genotypes of polymorphic markers in these segments would be expected to show homozygosity (Sheffield et al., 1994). Of the 10 families enrolled for linkage analysis, affected individuals from eight families showed homozygous genotypes at the CHED2 locus, which cosegregated with disease. Recombination events and lack of homozygosity in affected individuals in the families studied placed the interval between the markers D20S889 and D20S198. Screening of candidate genes within the region between D20S889 and D20S198 identified *SLC4A11* as the gene responsible for AR CHED. The *SLC4A11* is a membrane protein with 14 transmembrane domains (Parker et al., 2001). *SLC4A11* is thought to function as a Na/borate co-transporter (also known as Na^{+-} coupled borate cotransporter NaBC1) and may have an important role in borate homeostasis (Parker et al., 2001). The role of SLC4A11/NaBC1 in modulating the proliferative effect of borate was suggested by studies using gene-silencing techniques (Park et al., 2004; 2005).

Overall a total of 56 unrelated AR CHED families were screened for mutations in the *SLC4A11* gene. Mutations were identified in 48 families. 41/48 families had homozygous mutations (Table 4.3, 4.4, 4.5) and four families showed compound heterozygous mutations. Thirty-five of 41 families with homozygous changes were consanguineous. Only one mutant allele was identified in 3 families and no mutations were identified in eight families by SSCP analysis. It is possible that in the families with no mutation/single mutant allele detected, mutations may be located in the coding regions but not picked up by SSCP analysis, or that mutations are located in the promoter or in a non-coding region of the gene. Mutations identified were distributed throughout the length of protein. It is unlikely that these families represent autosomal dominant CHED, due to the fact that none of the probands had affected parents or affected relatives in previous generations upon history and/or examination.

Out of the 34 mutations identified in this study (detailed in sections 4.3 and 4.4), twenty-seven were not reported previously. Fourteen of 34 mutations identified are predicted to result in premature termination of the encoded protein (5 deletions, 5 nonsense mutations, 3 complex mutations and one splice site mutation) and may result in functionally null alleles.

A mutation of unknown consequence identified in this study was a deletion of 19 bp in intron 7 in two families, which cosegregated with disease as analysed in three unaffected family members available and was absent from 100 normal chromosomes. Its effect on the mRNA is not clear at present as the deleted sequence lies internally within the intron and does not affect known splice signals.

Qualitative and quantitative analysis of *SLC4A11* mRNA from individuals having this deletion would be required to determine its effects.

The present study covered aspects of AR CHED that were not reported earlier in the literature at the time of undertaking this study. These include mapping of the locus in Indian families, screening of candidate genes and identification of the gene for AR CHED, and characterization of the mutation spectrum of *SLC4A11* gene in Indian families. In addition, genotype-phenotype correlations were investigated using a range of clinical and histopathologic features of AR CHED patients.

DISCUSSION

Data from our tertiary eye care referral centre indicate that MCD and AR CHED are the major types of corneal dystrophy (especially in South India). The higher frequency of these dystrophies, which are both autosomal recessive, may be attributable to a higher prevalence of consanguinity in this region. Similar high frequencies (74.4%) of these two disorders have been reported from Saudi Arabia (Table 5.1) where consanguinity is common.

Table 5.1. Frequency of MCD and CHED corneal dystrophies.

S.NO	Country	Study period (years)	No of corneal buttons studied	Corneal dystrophies (%)		Reference
				MCD	CHED	
1	Saudi Arabia	6	86	61.6	12.8	Al Faran et al., 1991
2	India	6	181	29.3	34.8	Pandrowala et al., 2004
3	Japan	34	159	21	1.2	Ruth et al., 1994
4	United States	32	143	11.2	3.5	Arentsen et al., 1976
4	Germany	11	185	16	1.6	Lang et al., 1987

The table shows the frequency of corneal dystrophies MCD and CHED among patients undergoing corneal grafts for various corneal dystrophies from different countries over different study periods.

A total of 125 mutations and 88 single nucleotide polymorphisms have been reported in *CHST6* gene from the 300 families screened till date from Britain (El-Ashry et al., 2002), France (Niel et al., 2003), Iceland (Liu et al., 2000), India (Sultana

et al., 2003; 2005; Warren et al., 2003), Italy (Abbruzzese et al., 2004), Japan (Akama et al., 2000; Ida-Hasegawa et al., 2003), Saudi Arabia (Bao et al., 2001), United States (Aldave et al., 2004; Klintworth et al., 2003, 2006; Liu et al., 2005, 2006), and Vietnam (Ha et al., 2003; 2003).

The mutational spectrum overall as depicted in Table 5.2 mirrors the pattern obtained in this study, with missense mutations being the most frequent cause of disease among MCD patients, with 66.3% (199/300) followed by frameshift 15% (45/300), nonsense 5.3% (16/300), insertion 5.3% (15/300) and complex mutations 2.6% (8/100). In 5.5% (17/300) of all families screened, either a single mutant allele or no mutant alleles were identifiable (Table-5.2).

Table 5.2. Summary of different mutations identified in the *CHST6* gene in MCD patients.

Type of Mutation	No of different mutations	No of families
Missense	90	199
Nonsense	12	16
Deletion in coding region	9	35
Deletion in the upstream region	1	7
Rearrangements in the upstream region	1	3
Complex	4	8
Insertion mutations	8	15
No mutation in one / both alleles	0	17

Table shows the number of patients having each of the different types of mutations

It is possible that in these 5% of MCD cases, upstream or non-coding regions of *CHST6* or another locus is involved in the disease.

Similar to the results of this study, studies on MCD patients of other populations have found MCD type I to be predominant such as in Saudi Arabian and North American patients (Klintworth et al., 1997; Klintworth et al., 2006). A study on German patients (Cursiefen et al., 2001) showed the frequency of MCD types I and IA to be fairly similar (46% and 43% respectively). Similar to the present study, Bao and coworkers in Saudi Arabian patients also reported different immunophenotypes in patients with identical mutations (Bao et al., 2001).

Akama and coworkers in their study showed that MCD type II immunophenotype results from mutations in the upstream region of *CHST6*, which is expected to regulate gene expression. They proposed that the MCD type II mutations led to loss of cornea-specific expression of *CHST6* and hence might cause the milder immunophenotype of MCD characterized by AgKS in the cornea and serum. Results of the present study and those of other groups (Liu et al., 2005, 2006) have shown not only the presence of mutations in the coding region of *CHST6* in patients with MCD type II but also failed to show defects in the upstream region of *CHST6* gene.

Overall the present molecular genetic study on MCD contributed to the mutation spectrum of the *CHST6* gene in MCD patients in a large study population of Indian origin and provided the largest series of patients so far reported in whom analysis of MCD immunophenotypes was carried out. Data show a high degree of allelic heterogeneity in MCD and suggest that immunophenotypes of MCD may be determined by factors other than *CHST6* gene mutations.

The molecular genetic analysis of AR CHED in Indian patients resulted in identification of *SLC4A11* as the causative gene for AR CHED. While this work was being done, another study by Vithana and co-workers (Vithana et al., 2006) reported the CHED2 gene to be *SLC4A11*. These results confirm that *SLC4A11* is the cause of AR CHED. Further, the present study established the mutation spectrum of *SLC4A11* in CHED and also showed lack of obvious genotype-phenotype correlations in AR CHED.

Mutations in the *SLC4A11* gene have been reported in eighty-three families including all studies reported till date, of which 68 families are of Indian origin. Fifty-six of these families are those described in this thesis (Jiao et al., 2006; Sultana et al., 2007). Studies from the other groups include those from Vithana and coworkers who screened Mayanmar, Pakistan, and Indian families (Vithana et al., 2006), Kumar and coworkers who reported mutations in two Indian families (Kumar et al., 2007), and Ramprasad and coworkers who screened families from United Kingdom and India (Ramprasad et al., 2007). In addition, the *SLC4A11* gene was found to have mutations in patients with CDPD as reported by Desir and coworkers (Desir et al., 2007).

Recurrent mutations identified in the present study include Arg605Stop found in 5/56, Arg869His/Cys in 4/56 and Arg755Gln/Trp in 3/56 families screened. The total number of mutations identified in the *SLC4A11* gene with the present study reaches 52, including 23 missense, 8 nonsense, 10 deletion, 5 complex, 1 duplication, and 5 splice site mutations.

Functional studies of a few missense and nonsense mutations by Vithana and coworkers showed that mutant *SLC4A11* proteins do not localize to the plasma membrane and are consequently degraded due to lack of appropriate processing through endoplasmic reticulum (Vithana et al., 2006).

The knowledge obtained from the present molecular genetic analysis of autosomal recessive corneal dystrophies could be utilized for genetic counseling of families with affected members, although the allelic heterogeneity found in *SLC4A11* renders mutational screening time-consuming and expensive. These data together with future studies on the function of *SLC4A11* in the corneal endothelium may provide clues to treatments for endothelial disorders.

SUMMARY

Chapter 1: Introduction and literature review

This chapter gives a brief description of the structure of the cornea and its different layers followed by a review of macular corneal dystrophy and autosomal recessive congenital hereditary endothelial dystrophy. The aspects of these disorders that are discussed include the clinical and histopathological characteristics, as well as the underlying biochemical and molecular genetic bases. The background for the present study in this chapter is followed by statement of the aims and objectives of the study.

Chapter 2: Materials and Methods

This chapter provides the details of methods employed including enrollment and sample collection for patients and controls, and principles and procedures of the various molecular genetic techniques employed, comparative three-dimensional modeling of the CHST6 protein, immunohistochemistry, collection of clinical data, and morphometric evaluation of AR CHED-affected corneas for histopathological correlations.

Chapter 3: Mutations and immunophenotypes in MCD

This chapter presents the results of molecular genetic analysis of macular corneal dystrophy, including screening of the *CHST6* gene, determination of

immunophenotypes of MCD, correlations of immunophenotypes with mutations, and structural analysis of CHST6 mutants by comparative modeling.

Chapter 4: Mapping, gene identification and mutational analysis of AR CHED

This chapter presents the results of molecular genetic analysis of autosomal recessive congenital hereditary endothelial dystrophy. This includes linkage and haplotype analysis, and screening of candidate genes in the critical interval for identification of the CHED2 gene. Further screening of a cohort of patients with AR CHED for mutations in the *SLC4A11* gene and evauation of genotype-phenotype correlations are also presented.

Chapter 5: Discussion

This chapter discusses the results obtained in the present study in the background of available literature.

Important findings of the present study

MCD

This study describes mutational screening and immunophenotyping of the largest series of MCD patients of the same ethnic origin reported so far. This data enabled an examination of correlations between CHST6 mutations and immunophenotypes in MCD. The data on *CHST6* mutations in MCD confirmed a high degree of allelic heterogeneity noted in studies done elsewhere and found missense mutations to be the most frequent type of mutations identified in the *CHST6* gene. Structural modeling of missense and in-frame mutant proteins predicted possible effects of the mutations that may be a basis for their pathogenicity. Evaluation of MCD immunophenotypes showed MCD type I to be the most frequent immunophenotype followed by MCD type IA. MCD type II was found to be of very low frequency. No correlations were detected between *CHST6* mutations and immunophenotypes.

AR CHED

Linkage analysis mapped the disease in families with AR CHED to the CHED2 locus on chromosome 20p13. Using consanguineous families for linkage and haplotype analysis, the CHED2 interval was refined to 2.2cM (1.3Mb), thus facilitating candidate gene screening of the disease interval. Screening of candidate genes in the refined interval (1.3Mb) identified *SLC4A11* as the gene to be defective in AR

CHED. Mutational heterogeneity of the *SLC4A11* gene in AR CHED was established. There was no evidence for genotype-phenotype correlations in AR CHED.

REFERENCES

- Abbruzzese C, Kuhn U, Molina F, Rama P, De Luca M. Novel mutations in the CHST6 gene causing macular corneal dystrophy. *Clin Genet.* 2004;65:120-125.

- Abramowicz MJ, Albuquerque-Silva J, Zanen A. Corneal dystrophy and perceptive deafness (Harboyan syndrome):CDPD1 maps to 20p13. *J Med Genet*. 2002;39:110-112.

- Akama TO, Nishida K, Nakayama J, Watanabe H, Ozaki K, Nakamura T, Dota A, Kawasaki S, Inoue Y, Maeda N, Yamamoto S, Fujiwara T, Thonar EJ, Shimomura Y, Kinoshita S, Tanigami A, Fukuda MN. Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulphotransferase gene. *Nat Genet*. 2000;26:237-241.

- Akama TO, Nakayama J, Nishida K, Hiraoka N, Suzuki M, McAuliffe J, Hindsgaul O, Fukuda M, Fukuda MN. Human corneal GlcNac 6-O-sulfotransferase and mouse intestinal GlcNac 6-O-sulfotransferase both produce keratan sulfate. *J Biol Chem*. 2001;276:16271-16280.

- Akama TO, Misra AK, Hindsgaul O, Fukuda MN. Enzymatic synthesis in vitro of the disulfated disaccharide unit of corneal keratan sulfate. *J Biol Chem*. 2002;277:42505-42513.

- Aldave AJ, Yellore VS, Thonar EJ, Udar N, Warren JF, Yoon MK, Cohen EJ, Rapuano CJ, Laibson PR, Margolis TP, Small K. Novel mutations in the

carbohydrate sulfotransferase gene (CHST6) in American patients with macular corneal dystrophy. *Am J Ophthalmol.* 2004;137:465-473.

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". *Nucleic Acids Res*. 1997;25:3389-3402.

- Antine B. Histology of congenital hereditary corneal dystrophy. *Am J Ophthalmol*.1970;69:964-969.

- Bao W, Smith CF, al-Rajhi A, Chandler JW, Karcioglu ZA, Akama TO, Fukuda MN, Klintworth, GK. Novel mutations in the CHST6 gene in Saudi Arabic patients with macular corneal dystrophy. *ARVO Annual Meeting*. 2001; April 29-May 4: Fort Lauderdale, FL.

- Basler E, Grompe M, Parenti G, Yates J, Ballabio A. Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis. *Am J Hum Genet*. 1992;50:483-491.

- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. *Nucleic Acids Res.* 2000;28:235–242.

- Bernstein HS, Bishop DF, Astrin KH, Kornreich R, Eng CM, Sakuraba H, Desnick RJ. Fabry disease: six gene rearrangements and an exonic point mutation in the alpha-galactosidase gene. *J Clin Invest*. 1989;83:1390-1399.

- Bettelheim FA, Plessy B. The hydration of proteoglycans of bovine cornea. *Biochim Biophys Acta*. 1975;381:203-214.

- Bianco P, Fisher LW, Young MF, Termine JD, Robey PG. Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues. *J Histochem Cytochem*. 1990;38:1549-1563.

- Biswas S, Munier FL, Yardley J, Hart-Holden N, Perveen R, Cousin P, Sutphin JE, Noble B, Batterbury M, Kielty C, Hackett A, Bonshek R, Ridgway A, McLeod D, Sheffield VC, Stone EM, Schorderet DF, Black GC. Missense mutations in COL8A2, the gene encoding the alpha2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy. *Hum Mol Genet*. 2001;10:2415-2423.

- Blundell TL, Sibanda BL, Sternberg MJ, Thornton JM. Knowledge-based prediction of protein structures and the design of novel molecules. *Nature.* 1987;*326*:347-352.

- Bredrup C, Knappskog PM, Majewski J, Rodahl E, Boman H. Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. *Invest Ophthalmol Vis Sci*. 2005;46:420-426.

- Callaghan M, Hand CK, Kennedy SM, FitzSimon JS, Collum LM, Parfrey NA. Homozygosity mapping and linkage analysis demonstrate that autosomal recessive congenital hereditary endothelial dystrophy (CHED) and autosomal dominant CHED are genetically distinct. *Br J Ophthalmol*. 1999;83:115-119.

- Caterson B, Christner JE, Baker JR. Identification of a monoclonal antibody that specifically recognizes corneal and skeletal keratan sulfate. Monoclonal antibodies to cartilage proteoglycan. *J Biol Chem*. 1983;258:8848-8854.

- Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C, Carroll H. Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol*. 1998;141:1277-1286.

- Chan CC, Green WR, Barraquer-Somers E, de la Cruz ZC. Similarities between posterior polymorphous and congenital hereditary endothelial dystrophies: a study of 14 buttons of 11 cases. *Cornea*. 1982;1:155-172.

- Cornuet PK, Blochberger TC, Hassell JR. Molecular polymorphism of lumican during corneal development. *Invest Ophthalmol Vis Sci*. 1994;35:870-877.

- Corpuz LM, Funderburgh JL, Funderburgh ML, Bottomley GS, Prakash S, Conrad GW. Molecular cloning and tissue distribution of keratocan. Bovine corneal keratan sulfate proteoglycan 37A. *J Biol Chem*. 1996;271:9759-9763.

- Cursiefen C, Hofmann-Rummelt C, Schlotzer-Schrehardt U, Fischer DC, Haubeck HD, Kuchle M, Naumann GO. Immunohistochemical classification of primary and recurrent macular corneal dystrophy in Germany: subclassification of immunophenotype I A using a novel keratan sulfate antibody. *Exp Eye Res*. 2001;73:593-600.

- Cottingham RW, Idury RM, and Schaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet.* 1993;53:252-263.

- Donnenfeld ED, Cohen EJ, Ingraham HJ, Poleski SA, Goldsmith E, Laibson PR. Corneal thinning in macular corneal dystrophy. *Am J Ophthalmol.* 1986;101:112-113.

- Dorfman A and Matalon R. The metabolic basis of inherited diseases. [Eds Stanbury JB, Wyngaarden JB, Fredrickson DS), p1218-1272. McGraw-Hill, New York.1972.

- Edward DP, Yue BY, Sugar J, Thonar EJ, SunderRaj N, Stock EL, Tso MO. Heterogeneity in macular corneal dystrophy. *Arch Ophthalmol.* 1988;106:1579-1583.

- Edward DP, Thonar EJ, Srinivasan M, Yue BJ, Tso MO. Macular dystrophy of the cornea. A systemic disorder of keratan sulfate metabolism. *Ophthalmology.* 1990;97:1194-2000.

- Ehlers N, Modis L, Moller-Pedersen T. A morphological and functional study of congenital hereditary endothelial dystrophy. *Acta Ophthalmol Scand*. 1998;76:314-318.

- El-Ashry MF, El-Aziz MM, Wilkins S, Cheetham ME, Wilkie SE, Hardcastle AJ, Halford S, Bayoumi AY, Ficker LA, Tuft S, Bhattacharya SS, Ebenezer ND. Identification of novel mutations in the carbohydrate sulfotransferase gene (CHST6) causing macular corneal dystrophy. *Invest Ophthalmol Vis Sci***.** 2002;43:377-382.

- Evans SV. SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J Mol Graph*. 1993;11:134-138,127-128.

- Falany CN. Sulfation and sulfotransferases introduction: Changing view of sulfation and the cytosolic sulfotransferases. *FASEB J*. 1997;11:1-2.

- Fiser A, Feig M, Brooks CL, III, Sali A. Evolution and Physics in Comparative Protein Structure Modeling. *Acc Chem Res.* 2002;*35*:413-421.

- Fitch JM, Birk DE, Linsenmayer C, Linsenmayer TF. The spatial organization of Descemet's membrane-associated type IV collage in the avian cornea. *J Cell Biol*.1990;110:1457-1468.

- Francois J, Hanssens M, Teuchy H, Sbruyns M.Ultrastructural findings in corneal macular dystrophy (Groenouw II type). *Ophthalmic Res*. 1975;7:80-98.

- Funderburgh JL, Funderburgh ML, Man MM, Conrad GW. Physical and biological properties of keratan sulphate proteoglycan. *Biochem Soc Trans* 1991;19:871-876.

- Funderburgh JL, Funderburgh ML, Brown SJ, Vergnes JP, Hassell JR, Mann MM, Conrad GW. Sequence and structural implications of a bovine corneal keratan sulfate proteoglycan core protein. Protein 37B represents bovine lumican and proteins 37A and 25 are unique. *J Biol Chem*. 1993;268:11874-11880.

- Funderburgh JL. Keratan sulfate: structure, biosynthesis, and function. *Glycobiology*. 2000;10:951-958.

- Garner A. Histochemistry of corneal macular dystrophy. *Invest Ophthalmol*.1969;8:475-483.

- Gottsch JD, Bowers AL, Margulies EH, Seitzman GD, Kim SW, Saha S, Jun AS, Stark WJ, Liu SH. Serial analysis of gene expression in the corneal endothelium of Fuchs' dystrophy. *Invest Ophthalmol Vis Sci*. 2003;44:594-599.

- Gray NK, Wickens M. Control of translation initiation in animals. *Annu Rev Cell Dev Biol*. 1998;14:399-458.

- Groenouw A. Knötchenförmige Hornhauttrübungen (Nodulicorneae). *Archiv für Augenheilkunde.* 1890;21:281-289.

- Hand CK, Harmon DL, Kennedy SM, Fitz Simon JS, Collum LM, Parfery NA.. Localization of the gene for autosomal recessive congenital hereditary endothelial dystrophy (CHED2) to chromosome 20 by homozygosity mapping. *Genomics*. 1999;61:1-4.

- Ha NT, Chau HM, Cung le X, Thanh TK, Fujiki K, Murakami A, Hiratsuka Y, Hasegawa N, Kanai A. Identification of novel mutations of the CHST6 gene in Vietnamese families affected with macular corneal dystrophy in two generations. *Cornea.* 2003;22:508-511.

- Ha NT, Chau HM, Cung le X, Thanh TK, Fujiki K, Murakami A, Hiratsuka Y, Kanai A. Mutation analysis of the carbohydrate sulfotransferase gene in Vietnamese with macular corneal dystrophy. *Invest Ophthalmol Vis Sci.* 2003;44:3310-3316.

- Harboyan G, Mamo J, Kaloustian VD, Karam F. Congenital corneal dystrophy.Progressive sensorineural deafness in a family. *Arch Ophthalmol*. 1971;85:27-32.

- Hart GW. Biosynthesis of glycosaminoglycans during corneal development. *J Biol Chem.* 1976;251:6513-6521.

- Hasegawa N, Torii T, Nagaoka I, Nakayasu K, Miyajima H, Habuchi O. Measurement of activities of human serum sulfotransferases which transfer sulfate to the galactose residues of keratan sulfate and to the nonreducing end N acetylglucosamine residues of N-acetyllactosamine trisaccharide: comparison between normal controls and patients with macular corneal dystrophy. *J Biochem.* 1999;125:245-252.

- Hasegawa N, Torii T, Kato T, Miyajima H, Furuhata A, Nakayasu K, Kanai A, Habuchi O. Decreased GlcNAc 6-O-sulfotransferase activity in the cornea with macular corneal dystrophy. *Invest Ophthalmol Vis Sci.* 2000;41:3670-3677.

- Hassell JR, Newsome DA, Krachmer JH, Rodrigues MM. Macular corneal dystrophy: failure to synthesize a mature keratan sulfate proteoglycan. *Proc Natl Acad Sci.* 1980;77:3705-3709.

- Hassell JR, Cintron C, Kublin C, Newsome DA. Proteoglycan changes during restoration of transparency in corneal scars. *Arch Biochem Biophys.* 1983;222:362-369.

- Hassell JR, Klintworth GK. Serum sulfotransferase levels in patients with macular corneal dystrophy type I. *Arch Ophthalmol.* 1997;115:1419-1421.

- Hayashi K, Yandell DW. How sensitive is PCR-SSCP? *Hum Mutat.*1993;2:338-346.

- Hemmerich S, Lee JK, Bhakta S, Bistrup A, Ruddle NR, Rosen SD. Chromosomal localization and genomic organization for the galactose/ N-acetylgalactosamine/N-acetylglucosamine 6-O-sulfotransferase gene family. *Glycobiology.* 2001;11:75-87.

- Heon E, Mathers WD, Alward WL, Weisenthal RW, Sunden SL, Fishbaugh JA, Taylor CM, Krachmer JH, Sheffield VC, Stone EM. Linkage of posterior polymorphous corneal dystrophy to 20q11. *Hum Mol Genet.* 1995;4:485-488.

- Hiles DA, Biglan AW, Fetherolf EC. Central endothelial counts in children. *J Am Intraoc Implant Soc.* 1979;5:292-300.

- Hsu SM, Raine L, Franger H. Use of avidin-biotin-peroxidase complex (ABC) in immunoperoxidase technique: A comparison between ABC and unlabeled antibody (PAP) procedures. *J Histochem Cytochem.* 1981;27:577-580.

- Irvine AD, Corden LD, Swensson O, Swensson B, Moore JE, Frazer DG, Smith FJ,Knowlton RG, Christophers E, Rochels R, Uitto J, McLean WH. Mutations in cornea-specific keratin K3 or K12 genes cause Meesmann's corneal dystrophy. *Nat Genet.* 1997;16:184-187.

- Jiao X, Sultana A, Garg P, Ramamurthy B, Vemuganti GK, Gangopadhyay N, Hejtmancik JF, Kannabiran C. Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11. *J Med Genet*. 2007;44:64-68.

- Johnson DH, Bourne WM, Campbell RJ. The ultrastructure of Descemet's membrane. I. Changes with age in normal corneas. *Arch Ophthalmol.* 1982;100:1942-1947.

- Jonasson F, Johannsson JH, Garner Rice NS. Macular corneal dystrophy in Iceland. *Eye.* 1989;3:446-454.

- Jonasson F, Oshima E, Thonar EJ, Smith CF, Johannsson JH, Klintworth GK. Macular corneal dystrophy in Iceland. A clinical, genealogic, and immunohistochemical study of 28 patients. *Ophthalmology.* 1996;103:1111-1117.

- Johnson MS, Srinivasan N, Sowdhamini R, Blundell TL. Knowledge-based protein modelling. *Crit Rev Biochem Mol Biol* .1994;*29*:1-68.

- Jones ST, Zimmerman LE. Macular dystrophy of the cornea (Groenouw type II); clinicopathologic report of two cases with comments concerning its differential diagnosis from lattice dystrophy (Biber-Haab-Dimmer). *Am J Ophthalmol.* 1959;47:1-16.

- Judisch GF, Maumenee IH. Clinical differentiation of recessive congenital hereditary endothelial dystrophy and dominant hereditary endothelial dystrophy. *Am J Ophthalmol.* 1978;85:606-612.

- Kakuta Y, Pedersen LG, Pedersen LC, Negishi M. Conserved structural motifs in the sulfotranferase family. *Trends Biochem Sci.* 1998;23:129-130.

- Kanis AB, Al-Rajhi AA, Taylor CM, Mathers WD, Folberg RY, Nishimura DY, Sheffield VC, Stone EM. Exclusion of ARCHED from the chromosome 20 region containing the PPMD and AD-CHED loci. *Ophthalmic Genet.* 1999;20:243-249.

- Keates RH, Cvintal T. Congenital hereditary corneal dystrophy. *Am J Ophthalmol.* 1964;60:892-894.

- Kenyon KR, Maumenee AE. The histological and ultrastructural pathology of congenital hereditary corneal dystrophy. A case report. *Invest Ophthalmol***.** 1968;7:475-500.

- Kenyon KR, Antine B. The pathogenesis of congenital hereditary endothelial dystrophy of the cornea. *Am J Ophthalmol.* 1971;72:787-795.

- Kirkness CM, McCartney A, Rice NS, Garner A, Steele AD. Congenital hereditary corneal oedema of Maumenee. Its clinical features, management and pathology. *Br J Ophthalmol***.** 1987;71:130-144.

- Klintworth GK, Vogel FS. Macular corneal dystrophy. An inherited acid mucopolysaccharide storage disease of the corneal fibroblast. *Am J Pathol.* 1964;45:565-586.

- Klintworth GK, Smith CF. Macular corneal dystrophy. Studies of sulfated glycosaminoglycans in corneal explant and confluent stromal cell cultures. *Am J Pathol.* 1977;89:167-182.

- Klintworth GK, Smith CF. Abnormalities of proteoglycans and glycoproteins synthesized by corneal organ cultures derived from patients with macular corneal dystrophy. *Lab Invest.* 1983;48:603-612.

- Klintworth GK, Meyer R, Dennis R, Hewitt AT, Stock EL, Lenz ME, Hassell JR, Stark WJ Jr, Kuettner KE, Thonar EJ. Macular corneal dystrophy. Lack of keratan sulfate in serum and cornea. *Ophthalmic Pediatric Genet.* 1986;7:139-143.

- Klintworth GK, Oshima E, al-Rajhi A, al-Saif A, Thonar EJ, Karcioglu ZA. Macular corneal dystrophy in Saudi Arabia: a study of 56 cases and recognition of a new immunophenotype. *Am J Ophthalmol*. 1997;124:9-18.

- Klintworth GK, Sommer JR , Karolak LA, Reed JW. Identification of a new keratin K12 mutations associated with Stocker-Holt corneal dystrophy that differs from mutations found in Meesmann corneal dystrophy. *Invest Ophthalmol Vis Sci.* 1999;40:S563.

- Klintworth GK. The molecular genetics of the corneal dystrophies-current status. *Front Biosci*. 2003;8:687-713.

- Krafchak CM, Pawar H, Moroi SE, Sugar A, Lichter PR, Mackey DA, Mian S, Nairus T, Elner V, Schteingart MT, Downs CA, Kijek TG, Johnson JM, Trager EH, Rozsa FW, Mandal MN, Epstein MP, Vollrath D, Ayyagari R, Boehnke M, Richards JE. Mutations in TCF8 cause posterior polymorphous corneal dystrophy and ectopic expression of COL4A3 by corneal endothelial cells. *Am J Hum Genet*. 2005;77:694-708.

- Landshman N, Ben-Hanan I, Assia E, Ben-Chaim O, Belkin M. Relationship between morphology and functional ability of regenerated corneal endothelium. *Invest Ophthalmol Vis Sci*. 1988;29:1100-1109.

- Lathrop GM and Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. *Am.J.Hum.Genet.* 1984;36:460-465.

- Lewis D, Davies Y, Nieduszynki IA, Lawrence F, Quantock AJ, Bonshek R, Fullwood NJ. Ultrastructural localization of sulfated and unsulfated keratan sulfate in normal and macular corneal dystrophy type I. *Glycobiology.* 2000;10:305-312.

- Iida-Hasegawa N, Furuhata A, Hayatsu H, Murakami A, Fujiki K, Nakayasu K, Kanai A. Mutations in the CHST6 gene in patients with macular corneal dystrophy: immunohistochemical evidence of heterogeneity. *Invest Ophthalmol Vis Sci*. 2003;44:3272-3277.

- Li A, Jiao X, Munier FL, Schorderet DF, Yao W, Iwata F, Hayakawa M, Kanai A, Shy Chen M, Alan Lewis R, Heckenlively J, Weleber RG, Traboulsi EI, Zhang Q, Xiao X, Kaiser-Kupfer M, Sergeev YV, Hejtmancik JF. Bietti crystalline corneoretinal dystrophy is caused by mutations in the novel gene CYP4V2. *Am J Hum Genet.* 2004;74:817-826.

- Lidholt K, Lindahl V. Conserved structural motifs in the sulfotranferase family. *Biochem J*. 1992;287:21-29.

- Liu NP, Baldwin J, Jonasson F, Dew-Knight S, Stajich JM, Lennon F, Pericak-Vance MA, Klintworth GK, Vance JM. Haplotype analysis in Icelandic families defines a minimal interval for the macular corneal dystrophy type I gene. *Am. J Hum Genet.* 1998;63:912-917.

- Liu NP, Dew-Knight S, Jonasson F, Gilbert JR, Klintworth GK, Vance JM. Physical and genetic mapping of the macular corneal dystrophy locus on chromosome 16q and exclusion of TAT and LCAT as candidate genes. *Mol Vis*. 2000;6:95-100.

- Liu NP, Dew-Knight S, Rayner M, Jonasson F, Akama TO, Fukuda MN, Bao W, Gilbert JR, Vance JM, Klintworth GK. Mutations in corneal carbohydrate sulfotransferase 6 gene (CHST6) cause macular corneal dystrophy in Iceland. *Mol Vis.* 2000;6:261-264.

- Liu NP, Bao W, Smith CF, Vance JM, Klintworth GK. Different mutations in carbohydrate sulfotransferase 6 (CHST6) gene cause macular corneal

dystrophy types I and II in a single sibship. *Am J Ophthalmol.* 2005;139:1118-1120.

- Liu NP, Smith CF, Bowling BL, Jonasson F, Klintworth GK. Macular corneal dystrophy types I and II are caused by distinct mutations in the CHST6 gene in Iceland. *Mol Vis.* 2006;212:1148-1152.

- Marti-Renom MA, Madhusudhan MS, Fiser A, Rost B, Sali A. Reliability of assessment of protein structure prediction methods. *Structure.* 2002;*10*:435-440.

- Maumenee AE. Congenital hereditary corneal dystrophy. *Am J Ophthalmol.* 1960;60:1114-1124.

- Maury CP, Kere J, Tolvanen R, de la Chapelle A. Finnish hereditary amyloidosis is caused by a single nucleotide substitution in the gelsolin gene. *FEBS Lett*. 1990;276:75-77.

- Midura RJ, Hascall VC. Analysis of the proteoglycans synthesized by corneal explants from embryonic chicken. II. Structural characterization of the keratan sulfate and dermatan sulfate proteoglycans from corneal stroma. *J Biol Chem.* 1989;264:1423-1430.

- Midura RJ, Hascall VC, MacCallum DK, Meyer RF, Thonar EJ, Hassell JR, Smith CF, Klintworth GK. Proteoglycan biosynthesis by human corneas from patients with types 1 and 2 macular corneal dystrophy. *J Biol Chem.* 1990;265:15947-15955.

- Mohammed MD, Mckibbin M, Jafri H, Rashed Y.A new pedigree with recessive CHED mapping to the CHED2 locus on 20p13. *Br J Ophthalmol.* 2000;85:758-759.

- Mullaney PB, Risco JM, Teichmann K, Millar L. Congenital hereditary endothelial dystrophy associated with glaucoma. *Ophthalmology.* 1995;102: 186-192.

- Munier FL, Korvatska E, Djemai A, Le Paslier D, Zografos L, Pescia G, Schorderet DF. Kerato-epithelin mutations in four 5q31-linked corneal dystrophies*. Nat Genet.* 1997;15:224-225.

- Murphy C, Alvarado J, Juster R. Prenatal and postnatal growth of the human Descemets membrane*. Invest Ophthalmol Vis Sci.* 1984;25:1402-1415.

- Murzin AG, Bernner SE, Hubbard T, Chothia C. SCOP:A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol*. 1995;247:536-540.

- Nakazawa K, Hassell JR, Hascall VC, Lohmander LS, Newsome DA, Krachmer J. Defective processing of keratan sulfate in macular corneal dystrophy. *J Biol Chem*. 1984;259:13751-13757.

- Negishi M, Pedersen LG, Petrotchenko E, Shevtsov S, Gorokhov A, Kakuta Y, Pedersen LC. Structure and function of sulfotransferases. *Arch Biochem Biophys.* 2001;390:149-157.

- Newsome DA, Hassell JR, Rodrigues MM, Rahe AE, Krachmer JH. Biochemical and histological analysis of recurrent macular corneal dystrophy. *Arch Ophthal.* 1982;100:1125-1131.

- Niel F, Ellies P, Dighiero P, Soria J, Sabbagh C, San C, Renard G, Delpech M, Valleix S. Truncating mutations in the carbohydrate sulfotransferase 6 gene (CHST6) result in macular corneal dystrophy. *Invest Ophthalmol Vis Sci.* 2003; 44:2949-2953.

- Okada M, Yamamoto S, Tsujikawa M, Watanabe H, Inoue Y, Maeda N, Shimomura Y, Nishida K, Quantock AJ, Kinoshita S, Tano Y. Two distinct kerato-epithelin mutations in Reis-Bucklers corneal dystrophy. *Am J Ophthalmol.* 1998;126:535-542.

- Overington J, Johnson MS, Sali A, Blundell TL. Tertiary structural constraints on protein evolutionary diversity: templates,key residues and structure prediction. *Proc Biol Sci*. 1990;241:132-145.

- Pakhomova S, Kobayashi M, Buck J, Newcomer ME A helical lid converts a sulfotransferase to a dehydratase. *Nature Strl Biol.* 2001;8:447-451.

- Pandrowala H, Bansal A, Vemuganti GK, Rao GN. Frequency, distribution, and outcome of keratoplasty for corneal dystrophies at a tertiary eye care center in South India. *Cornea.* 2004;23:541-546.

- Park M, Li Q, Shcheynikov N, Zeng W, Muallem S. NaBC1 is a ubiquitous electrogenic Na+ -coupled borate transporter essential for cellular boron homeostasis and cell growth and proliferation. *Mol Cell.* 2004;16:331-341.

- Park M, Li Q, Shcheynikov N, Muallem S, Zeng W. Borate transport and cell growth and proliferation. Not only in plants. *Cell Cycle.* 2005;4:24-26.

- Parker, MD, Ourmozdi, EP, Tanner, MJ. Human BTR1, a new bicarbonate transporter superfamily member and human AE4 from kidney. *Biochem Biophys Res Commun*. 2001;282:1103-1109.

- Pearlman DA, David A. Case DA, Caldwell JC, Seibel GL Singh UC, Weiner PK, Kollman PA. 1991; AMBER 4.0: University of California, San Francisco.

- Pedersen O, Rushood A, Olesen EG. Anterior mesenchymal dysgenesis of the eye. Congenital hereditary endothelial dystrophy and congenital glaucoma. *Acta Ophthalmol.* 1989;67:470-476.

- Plaas AH, West LA, Thonar EJ, Karcioglu ZA, Smith CJ, Klintworth GK, Hascall VC. Altered fine structures of corneal and skeletal keratan sulfate and chondroitin/dermatan sulfate in macular corneal dystrophy. *J Biol Chem.* 2001;276:39788-39796.

- Quantock AJ, Meek KM, Ridgway AE, Bron AJ, Thonar EJ. Macular corneal dystrophy: reduction in both corneal thickness and collagen interfibrillar spacing. *Curr Eye Res*. 1990;9:393-398.

- Ramprasad VL, Ebenezer ND, Aung T, Rajagopal R, Yong VH, Tuft SJ, Viswanathan D, El-Ashry MF, Liskova P, Tan DT, Bhattacharya SS, Kumaramanickavel G, Vithana EN. Novel SLC4A11 mutations in patients with recessive congenital hereditaryendothelial dystrophy (CHED2). Mutation in brief #958. Online. *Hum Mutat*. 2007;28:522-530.

- Romero MF, Fulton CM, Boron WF. The SLC4 family of HCO3- transporters. *Pflugers Arch*. 2004;5:495-509.

- Rost B, Sander C, Schneider R. PHD-an automatic mail server for protein secondary structure prediction. *Comput Appl Biosci*. 1994;10:53-60.

- Sambrook J, Fritsch EF, Maniatis T. Molecular cloning. A laboratory manual. 2nd ed. New York, Cold Spring Harbor Laboratory Press. 1989.

- Santo RM, Yamaguchi T, Kanai A, Okisaka S, Nakajima A. Clinical and histopathologic features of corneal dystrophies in Japan. *Ophthalmology* 1995;102:557-567.

- Shearman AM, Hudson TJ, Andresen JM, Wu X, Sohn RL, Haluska F, Housman DE, WeissJS. The gene for schnyder's crystalline corneal dystrophy maps to human chromosome1p34.1-p36. *Hum Mol Genet.* 1996;5:1667-1672.

- Sheffield VC, Carmi R, Kwitek-Black A, Rokhlina T, Nishimura D, Duyk GM, Elbedour K, Sunden SL, Stone EM. Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. *Hum Mol Genet.* 1994;3:1331-1335.

- Snip RC, Kenyon KR, Green WR. Macular corneal dystrophy. Ultrastructural pathology of corneal endothelium and Descemet's membrane. *Invest Ophthalmol.* 1973;12:88.

- Sony P, Sharma N, Sen S, Vajpayee RB. Indications of penetrating keratoplasty in northern India. *Cornea.* 2005;24:989-991.

- Srinivasan N, Blundell TL. An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure. *Prot Engng.* 1993;6:501-512.

- Sultana A, Sridhar MS, Jagannathan A, Balasubramanian D, Kannabiran C, Klintworth GK. Novel mutations of the carbohydrate sulfotransferase-6 (CHST6) gene causing macular corneal dystrophy in India. *Mol Vis.* 2003;9:730-734.

- Sultana A, Sridhar MS, Klintworth GK, Balasubramanian D, Kannabiran C. Allelic heterogeneity of the carbohydrate sulfotransferase-6 gene in patients with macular corneal dystrophy. *Clin Genet.* 2005;68:454-460.

- Sultana A, Garg P, Ramamurthy B, Vemuganti GK, Kannabiran C. Mutational spectrum of the SLC4A11 gene in autosomal recessive congenital hereditary endothelial dystrophy. *Mol Vis.* 2007; In press.

- SunderRaj N, Barbacci-Tobin E, Howe WE, Robertson SM, Limetti G.Macular corneal dystrophy: Immunochemical characterization using monoclonal antibodies. *Invest Ophthalmol Vis Sci*. 1987;28:167-186.

- Sutcliffe MJ, Haneef I, Carney D, Blundell TL. Knowledge based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng*. 1987;1:377-384.

- Sutcliffe MJ, Hayes FR, Blundell TL. Knowledge based modelling of homologous proteins, Part II: Rules for the conformations of substituted sidechains. *Protein Eng*. 1987;1:385-392.

- Teng CC. Macular dystrophy of the cornea. A histochemical and electron microscopic study. *Am J Ophthalmol*. 1966;62:436-454

- Thompson JD, Higgins DG, et al. CLUSTAL Wimproving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research.* 1994;22:4673-4680.

- Thonar EJ, Lenz ME, Klintworth GK, Caterson B, Pachman LM, Glickman P, Katz R, Huff J, Kuettner KE. Quantification of keratan sulfate in blood as a marker of cartilage catabolism. *Arthritis Rheum*. 1985;28:1367-1376.

- Toma NM, Ebenezer ND, Inglehearn CF, Plant C, Ficker LA, Bhattacharya SS. Linkage of congenital hereditary endothelial dystrophy to chromosome 20. *Hum Mol Genet*. 1995;4:2395-2398.

- Topham CM, McLeod A, Eisenmenger F, Overington JP, Johnson MS, Blundell TL. Fragment ranking in modelling of protein structure.

Conformationally constrained environmental amino acid substitution tables. *J Mol Biol.* 1993;229:194-220.

- Tsujikawa M, Kurahashi H, Tanaka T, Nishida K, Shimomura Y, Tano Y, Nakamura Y. Identification of the gene responsible for gelatinous drop-like corneal dystrophy. *Nat Genet.* 1999;21:420-423.

- Van Horn DL, Hyndiuk RA. Endothelial wound repair in primate cornea. *Exp Eye Res.* 1975;21:113-124.

- Vance JM, Jonasson F, Lennon F, Sarrica J, Damji KF, Stauffer J, Pericak-Vance MA, Klintworth GK. Linkage of a gene for macular corneal dystrophy to chromosome 16. *Am J Hum Genet*. 1996;58:757-762.

- Vithana EN, Morgan P, Sundaresan P, Ebenezer ND, Tan DT, Mohamed MD, Anand S, Khine KO, Venkataraman D, Yong VH, Salto-Tellez M, Venkatraman A, Guo K, Hemadevi B, Srinivasan M, Prajna V, Khine M, Casey JR, Inglehearn CF, Aung T. Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2). *Nat Genet.* 2006;38:755-757.

- Waring GO 3rd, Bourne WM, Edelhauser HF, Kenyon KR. The corneal endothelium. Normal and pathologic structure and function. *Ophthalmology.* 1982;89:531-590.

- Warren JF, Aldave AJ, Srinivasan M, Thonar EJ, Kumar AB, Cevallos V, Whitcher JP, Margolis TP. Novel mutations in the CHST6 gene associated

with macular corneal dystrophy in southern India. *Arch Ophthalmol.* 2003;121:1608-1612.

- Yamamoto S, Okada M, Tsujikawa M, Shimomura Y, Nishida K, Inoue Y, Watanabe H, Maeda N, Kurahashi H, Kinoshita S, Nakamura Y, Tano Y. A kerato-epithelin (betaig-h3) mutation in lattice corneal dystrophy type IIIA. *Am J Hum Genet*. 1998;62:719-722.

- Yang CJ, SundarRaj N, Thonar EJ, Klintworth GK. Immunohistochemical evidence of heterogeneity in macular corneal dystrophy. *Am J Ophthalmol.* 1988;106:65-71.

- Young RD, Akama TO, Liskova P, Ebenezer ND, Allan B, Kerr B, Caterson B, Fukuda MN, Quantock AJ. Differential immunogold localisation of sulphated and unsulphated keratan sulphate proteoglycans in normal and macular dystrophy cornea using sulphation motif-specific antibodies. *Histochem Cell Biol.* 2006;127:115-120.

APPENDIX-I

Table 1. List of primers for carbohydrate sulfotransferase-6 gene (*CHST6*).

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Upstream/Exon
	Primer name	Primer sequence (5'-3')			
1	RIM	TGCTGAATGGCTAACTGAAGGAATACTATAC	62	1,502	Upstream-Region A
2	F2M	GGCCAAGTTCAGGTCAGCTTCCA			
3	F2M' (Replaced primer)	CCACAGCCAATTCCATCTTGGATTTTCTC			
4	F1	CCACAGAAGGAAGGACAGAGTAAATGAA	55	956	Upstream-Region B
5	R1	TTCCCTTTACTATTATAAAAATGCTGCTAATG			
6	F2	CATATCCTGTCTGGCCTAAACCTTAGTTTAC	58	966	Upstream-Region B
7	R2	GGGCACAGACAGAGGGAAAAACC			
8	R2' (Replaced primer)	CATTAGACACCTCACCTGCTTTGGC			
9	CK71h-intrn	GCCCCTAACCGCTGCGCTCTC	60	498	Exon 4 / 5' coding region
10	CK71h-R1180	GGCTTGCACACGGCCTCGCT			
11	CK71h- F1041	GACGTGTTTGATGCCTATCTGCCTTG	57	635	Exon 4 / Middle coding region
12	CK71h- R1674	CGGCGCGCACCAGGTCCA			
13	CK71h-F1355	CTCCCGGGAGCAGACAGCCAA	60	599	Exon 4 / 3' coding region
14	CK71h-R1953	CTCCCGGGCCTAGCGCCT			

This table list the primers used for screening of *CHST6* upstream and coding region by direct sequencing.

Table 2. List of primers for *CHST6* gene.

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	CHST3-1F	CCCCTAACCGCTGCGCTCTC	60	267	3
2	CHST3-1R	CCACGCGGGCTCCATTAGGTA			
3	CHST3-2F	CTCGGGCTCGTCCTTCGTGGG	58	269	3
4	CHST3-2R	GGGGAAAGGCACTGCAGGCG			
5	CHST3-3F	TGTCCGACCTCTTCCAGTGGG	60	263	3
6	CHST3-3R	CGGGTCGCGCACCAGGT			
7	CHST3-4F	GGTGCTCTACCCGCTGCTCAG	55.4	269	3
8	`CHST3-4R	GCACCAGGCGGTAGCGGC			
9	CHST3-5F	TAGCCACGTACGCATCGCCG	62	279	3
10	CHST3-5R	GCAAAGGGCAGCGCATGG			
11	CHST3-6F	ACTTCGTCCAGGAATGCGCTC	62	279	3
12	CHST3-6R	CACCATGCACTCTCCTCCCGG`			

This table list the primers used for screening of *CHST6* coding region by SSCP analysis.

Table 3. List of primers for early B-cell factor 4-transcription factor (KIAA1442).

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	Ki1F	ATGGCCCGGGCCCGAGG	60	388	1
2	Ki1R	GCCCTTTCCCCCTCGCCGC			
3	Ki2F	GCTGAGCTAGACGCCCGCA	60	292	2
4	Ki2R	CTCCAGCAGCTCCAGCAAGC			
5	Ki3F	AGGCGCTGGAGTCTTTCTCAC	60	281	3
6	Ki3R	TGTGTGTGGTTGGGGGTTCC			
7	Ki4-5F	GGGAAAGAAGTTGGGGTTAGG	62	460	4 & 5
8	Ki4-5R	GGAGGCTCACTCAGACCAG			
9	Ki6F	TGCCACCTGCACCTCAGAGG	60.6	205	6
10	Ki6R	CAGGGCCTGGGTAGAGTCCT			
11	Ki7F	CATCAGCACCCTCAGCTCAGG	63	221	7
12	Ki7R	CCCTCCCCCACTCTCCTCTG			
13	Ki8F	TCATCTGTTTCCAGAACCCCC	60	218	8
14	Ki8R	GTCACAGCCTTCCCTTCCCC			
15	Ki9-11F	CCCACCACATTCCCCTCCCG	62	720	9, 10 &11

16	Ki9-11R	CACCAAGAACGGAGACCAGTGC			
17	Ki12F	GCGTCTCACCCTGGGAGTGG	60	287	12
18	Ki12R	GAGAGACAGCGGGACGGA			
19	Ki13-14F	TCCCGAGGGCCCCTTGGTC	57	523	13 & 14
20	Ki13-14R	TGCTGCAGCTGCGCGCGTAG			
21	Ki15F	CAGCCCGCTGGCCATCGCC	58	396	15
22	Ki15R	GGCTGAGTCCCGTTCCGCT			
23	Ki16F	CCCCACCACCAGCGTGTTC	60	227	16
24	Ki16R	AGATGTGGACAGAGCAGGCCC			

Table 4. List of primers for chromosome 20 open reading frame 116 and 28 (C20ORF116 & C20ORF28).

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	116-1F	GTTCTGCACTGAGGCCCTCG	58	187	1
2	116-1R	GCGCGACGGTCCACAAAG			
3	116-2F	GCCGCTCCTCCTGCCATG	58.7	345	2
3	116-2R	CACTTAGCTTTGGGGGCCC			
5	116-3F	ACGAGGGGTTGGGGATCCAT	58	272	3
6	116-3R	ACAAGCCCTCAGTCTGCCTTA			
7	116-4F	TAGTCTGTCTGGGCGGTGGG	60	252	4
8	116-4R	AGGTGCCAGGGAGCTGCATA			
9	116-5F	AGGGGTAGCAGCAACCAGAAC	60	283	5
10	116-5R	CTGGGATGGGGTGGCTAGC			

11	116-6F	GAACCCTGTGGGCTTGCTAGG	63	209	6
12	116-6R	CCCCTGTCCACTCCTGCATG			
13	116-7F	GAATGCAAACCTGGAGGGTGC	63	217	7
14	116-7R	ATCTTTAGGGCAGGTCCTCTCTG			
15	116-8F	GGTTTCAGTGCTTTTTCTGGG	58	201	8
16	116-8R	TCCTGCCTCTGCAGCACATC			
17	116-9F	GCTGTGAAGGAGTAGGTGGGC	60.6	345	9
18	116-9R	GCCACACCATCACTTCCCC			
19	28-1F	TGCAGTGGCGAGAAAGGAAGC	62	274	1
20	28-1R	GTTTGCACAGAGACCCGTTACAC			
21	28-2F	CACCCCCACCTTCTCTCACC	60.6	272	2
22	28-2R	TCTGCGTTGACCCTCCCCAC			
23	28-3-4F	AACCACCATGCCCCCTTCCC	62	629	3&4
24	28-3-4R	GGGCCTTCTGGGATACACCTGA			

Table 5. List of primers for arginine vasopressin (AVP) and oxytocin neurophysin I (OXT) genes.

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	AVP1F	CCAGCAGAGGCAGCAGCACAG	62	282	1
2	AVP1R	CCACCACCCATGACTTCCCTC			
3	AVP2F	GCCTCGCTGCGTTCCCCTC	59	337	2
4	AVP2R	GCGTCCCCCCCACCCAAGC			
5	AVP3F	CGGCAGGGAGGGTGTGGGC	64	310	3
6	AVP3R	TATTGTCCGTGCTGCAGGGGC			
7	OXT1F	CATAAAAAGGCCAGGCCGGAG	61	242	1
8	OXT1R	TGTGGCTGCGGGTCCCTCC			
9	OXT2-3F	TCCGAGCGAGTCCCCAGCG	61	545	2 & 3
10	OXT2-3R	CAATGCCCCCTCCCCGTCC			

Table 6. List of primers for centromere protein B (CENPB) and adrenergic alpha 1D receptor (ARA1D) genes.

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	CEN1F	GGGGGCGCGCCGGGATGG	61.5	459	1
2	CEN1R	GGGGGGCAGCGTTTCGCG			
3	CEN2F	GCTGGACCGCTTCCGCCG	59.5	457	2
4	CEN2R	CAGGGCCTGGGTGGTGACA			
5	CEN3F	CCCGCGCAGGCCAAGCCG	61.5	427	3
6	CEN3R	CCACCCCCAAAGCCAGCCTCA			
7	CEN4F	TCACGGAGGCCCTGCACTTTG	62.5	541	4
8	CEN4R	ACAGGCACCTCATCACCATCCT			
9	CEN5F	CAGACAGTGAGGAAGAGGACGA	61.5	295	5
10	CEN5R	AGGTTGGGGGCACAGCTAGG			
11	ADR1F	CCCTGCCGGCCGCTCGTT	60	531	1
12	ADR1R	GAACCTCCATGGTGGCCGAG			
13	ADR2F	CATCGTGAACCTGGCCGTGG	60	450	2
14	ADR2R	GTGCTGCGCGCGACCACG			
15	ADR3F	CTCCGTGTGCTCCTTCTACC	60	437	3
16	ADR3R	TCTAGGTTCAGGTGGGGGTC			
17	ADR4F	CCTTTGCGCCCAGAGCCG	59.5	508	4
18	ADR4R	CTGGAGACTTTGGCGCGCA			
19	ADR5F	GCCACCCAGCGCCTTCCG	63	325	5
20	ADR5R	GAACACCAGCCCGCCTCTCTG			

Table 7. List of primers for gonadotropin releasing hormone-2 (GNRH2), mitochondrial ribosomal protein (MRSP26), bipartite nuclear localization protein (NC) and 175 amino acid unknown peptide (Q8N1Y6).

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	GR-1F	GCTCTGCCAGCCATCCTTCT	59.5	241	1
2	GR-1R	GCTCAGCACATAGTGGGGTCAA			
3	GR-2F	TGAAGTGTGCATGCGGGGC	55	652	2
4	GR-2R	CCACCCGGGCGGAGATATC			
5	GR-3-4F	GGTGGGGACGGGGCCTCTCTC	53	474	3&4
6	GR-3-4R	CCCTCTAAAATGGCTCCCAGC			
7	GR-5-6F	CTGTGCTAGGGTTTCCCAGC	59.5	476	5&6
8	GR-5-6R	CTTCTCAAGGGGCCAGTAGG			
9	MR1-2F	CTACAGGTCCCAAGGTTCCCC	63	615	1&2
10	MR1-2R	CAAAGGGCCGAGGCTGAGC			
11	MR-3F	GGAGAAGCCCGGGCCCCG	59.5	267	3
12	MR-3R	TGCACATAGCCAGCGCCCGC			
13	MR-4F	TGGGCTGGCTGGCATGTGCCC	59.5	260	4
14	MR-4R	AGGTCAGGGCCAACTCTTCCG			
15	NC-1F	AGTGCCCAATCCCAAAGCCAG	59.5	459	1
16	NC-1R	CCAGCGGCGCGTACTCATC			
17	NC-2F	GGGGTCGGTGATGAAGCGC	55	442	2

18	NC-2R	CGGAGGTGCAGGGTGTGAAG			
19	Q8-1F	CGTCAGTGGGGACTGTGCTG	59.5	600	1
20	Q8-1R	AGATCTGGGCCAGAGGGACC			

Table 8. List of primers for transmembrane channel-like protein-2 (TMC2).

S.NO	Primer		Annealing (^{0}C)	Product Size (bp)	Exon
	Primer name	Primer sequence (5'- 3')			
1	TMC1F	CCACCATGAGGGCAGGAAAGC	68	193	1
2	TMC1R	CCCATCCCCTGCTGCTGCC			
3	TMC2F	GTCTGATTTCTC ACGGCCGG	68	467	2
4	TMC2R	AAGGGGAAGGGCCGCGAAG			
5	TMC3F	GCCCCTCCTCATTCCACAGC	66	354	3
6	TMC3R	CCTTTGGGGAAATGTGGAGGG			
7	TMC4F	TTCACACATGTGTTTGAGGCTTC	59	242	4
8	TMC4R	GGTCGCCAGATTCCAAAGCT			
9	TMC5F	CCTGTCGGCCTCCTCGTGC	62	230	5
10	TMC5R	GCTCTTGGGAAACGCTTTAGGA			
11	TMC6F	GTGTCCATTTTTTCCTTCTGCA	67	301	6
12	TMC6R	ACAATTACTTTTGCACCAGCCTA			
13	TMC7F	TCTGAGAAGGAAAGCATTTACCC	59	258	7
14	TMC7R	TTATCCATGCAATCTTTAGCCTT			
15	TMC8F	GGTAACCACCAGCTCTGAGCTT	68.5	280	8

16	TMC8R	AGGGAGGTAGGCATAGAGTGGC			
17	TMC9F	CAA GGTTCCCTGGGGGTGA	64	291	9
18	TMC9R	CCAGCTGGGAGAAGTCTCTCA			
19	TMC10F	ATGGTTCTGG CATTTCTGGC	64	321	10
20	TMC10R	TGAAATGAATGGCATAGGAACC			
21	TMC11F	GGATGGCCATGGGAGTGCA	58	315	11
22	TMC11R	CTATGGGCACCAGGAATTGCA			
23	TMC12F	GCATTTGCTGGCTTTCACTGT	58	332	12
24	TMC12R	TTCAGGGTTATTCCTCAAATGG			
25	TMC13F	CAAACTCTCAGGGAGGGTGGG	63	297	13
26	TMC13R	AAATGAAATGCCCAAGAATGC			
27	TMC14F	GGCTGAATTCACCAAACGTGC	61	258	14
28	TMC14R	GGAGTTCAGAGAGGCAAGAGGA			
29	TMC15F	CCCTAACCCATCCGTCCCATT	61	417	15
30	TMC15R	AACTGGAATTCTGGGCCGAGG			
31	TMC16F	CATGGACACAGCTGTGAATGC	58	285	16
32	TMC16R	GGCAAGATGAAAAGGACCTGC			
33	TMC17F	GGCGCCCAGCTCCTATCATCT	61	226	17
34	TMC17R	TGGGCTCACACAGCAAATTAGC			
35	TMC18F	CCTCTGAGCCTCAAAGACCC	58	236	18
36	TMC18R	TTTACACCTTTCCATGTGGCC			
37	TMC19F	AGGGCGACTCCAGAGAGCTCC	63	388	19
38	TMC19R	GGGGAGAGAACCTTGGTTTGG			

Table 9. List of primers for sodium bicarbonate transporter-like solute carrier family 4 member 11 (SLC4A11)

S.NO	Primer		Annealing (^{0}C)	Product size (bp)	Exon
	Primer name	Primer sequence (5'-3')			
1	SLC4A111F	GGAGACAGGTGCACACAGGA	57	197	1
2	SLC4A111R	TCTGGTGGGTAGGCTATGCA			
3	SLC4A112F	CTTCCTGTGTGTGGCACTTT	54.4	274	2
4	SLC4A112R	TAAGGCGAGTCACACCTGC			
5	SLC4A113F	AGACTGGGGGAAGGCACG	54	269	3
6	SLC4A113R	TGTTGAGCTGCTCCTGGA			
7	SLC4A114F	ACCAGGCAGTGACAGCATCT	55.2	343	4
8	SLC4A114R	TCAACAGCCCCTCCCAAC			
9	SLC4A115F	CGTTGGGAGGGGCTGTTGA	60	205	5
10	SLC4A115R	TGCAGGGCACAGGGGACAT			
11	SLC4A116F	TCAAGGTCGAGGGGGTTCT	54.4	269	6
12	SLC4A116R	CAAGCAGAGGGCGGGTAA			
13	SLC4A117F	GCAGGGCCTCCTCTGTTTCT	60	272	7
14	SLC4A117R	TGTACCTCTGGGTGTCTGTGG			
15	SLC4A118F	AGAACGAAGGAACGCAGCA	54.4	286	8
16	SLC4A118R	ATGCAGGACAGGCACACGT			
17	SLC4A119F	TCCTGCATCCCCAGCAAA	55	295	9
18	SLC4A119R	TCTGCCCGGCTATGGTCTT			
19	SLC4A1110F	CGGTGGGTGGGGGTATGCTG	61.7	232	10
20	SLC4A1110R	CAGGAACCAGGGGTCTCAGGC			
21	SLC4A1111F	AGATGGTGCCTGAGACCCCT	57	247	11

22	SLC4A1111R	AACCAGATCCCAAGCCTTGA			
23	SLC4A1112F	ACTGTGCAGGGGCTCAGGGA	57	245	12
24	SLC4A1112R	CAGTCCTATGTGCCCCCA			
25	SLC4A1113F	ACACGCTGGAGAGGGCTGA	55	297	13
26	SLC4A1113R	ATTGGTAGAGGGTGTAGCCCA			
27	SLC4A1114F	TGGGCTACACCCTCTACCAA	57	274	14
28	SLC4A1114R	TACTGCTATGCCTGCAGCG			
29	SLC4A1115F	GTGGGTGACGTGGGGTAGCTC	61.7	329	15
30	SLC4A1115R	GCCCTTCACCAGCCTGCAGC			
31	SLC4A1116F	TGTGGAGGGTGGGTGAGGGT	61	307	16
32	SLC4A1116R	AGAGGCTCCCCACTCCTCAG			
33	SLC4A1117F	GGAGTGAGGCCCTGTGGACA	59.5	358	17
34	SLC4A1117R	TGTGTCGGGGGGTACGCAGT			
35	SLC4A1118F	GTGGGGACACAGCCCCAT	57.4	265	18
36	SLC4A1118R	ACACCTAGACTGGGCCCCT			
37	SLC4A1119F	TAGCCTGGCCCTGTGGGT	54.4	293	19
38	SLC4A1119R	TGAGAAGGCGCAGCACAGA			

APPENDIX-II

URLs of electronic databases and web-sites cited in the present study

1. NCBI protein data base:
   http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=protein
2. Mutation nomenclature: http://www.hgvs.org/mutnomen/recs.html
3. NEBcutter V2.0, New England Biolabs:
   http://tools.neb.com/NEBcutter2/index.php
4. Neural network-based program for splice site prediction at the web-site for Berkeley Drosophila Genome Project: http://www.fruitfly.org/cgi-bin/seq_tools/splice.pl
5. Primer design software: http://seq.yeastgenome.org/cgi-bin/webprimer
6. Protein Data Bank: http://www.rcsb.org/pdb/home/home.do

LIST OF PUBLICATIONS

1. **Sultana A**, Sridhar MS, Jagannathan A, Balasubramanian D, Kannabiran C, Klintworth GK. Novel mutations of the carbohydrate sulfotransferase-6 (CHST6) gene causing macular corneal dystrophy in India. ***Mol Vis*** 2003;9:730-4

2. **Sultana A**, Sridhar MS, Klintworth GK, Balasubramanian D, Kannabiran C. Allelic heterogeneity of the carbohydrate sulfotransferase-6 gene in patients with macular corneal dystrophy. ***Clin Genet***. 2005;68:454-60.

3. Jiao X, **Sultana A**, Garg P, Ramamurthy B, Vemuganti GK, Gangopadhyay N, Hejtmancik JF, Kannabiran C. Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11. ***J Med Genet***. 2007 Jan;44(1):64-8.

4. **Sultana A**, Garg P, Ramamurthy B, Vemuganti GK, Gangopadhyay N, Kannabiran C. Identification of novel mutations in SLC4A11 gene causing autosomal recessive corneal endothelial dystrophy2. 2007 (In press ***Mol Vis***).

LIST OF PRESENTATIONS

1. Presented a poster titled " **Identification of novel mutations in patients with macular corneal dystrophy"** at 10th FAOBMB, Indian Institute of Science**,** Bangalore-2003**.**

2. Presented poster titled **"Allelic heterogeneity in carbohydrate sulfotransferase-6 gene associated with macular corneal dystrophy"** at 14th annual meeting of Indian eye research group, Sankara Nethralaya, Chennai-2004.

3. Presented poster titled **"Structural interpretation of genetic mutations in carbohydrate sulfotransferase-6 using comparative modeling"** at 15th annual meeting of Indian eye research group, LVPEI, Hyderabad-2005.

4. Presented a paper titled **"Molecular genetic analysis of autosomal recessive corneal dystrophies"** at 31st annual conference of Indian society of human genetics**,** Jawaharlal Nehru University**,** New Delhi-2006.

5. Presented poster titled **"Identification of novel mutations in the *SLC4A11* gene in autosomal recessive corneal endothelial dystrophy"** at Asia ARVO Singapore-2007.

Molecular Vision 2003; 9:730-4 <http://www.molvis.org/molvis/v9/a87>
Received 17 October 2003 | Accepted 22 December 2003 | Published 22 December 2003

Novel mutations of the carbohydrate sulfotransferase-6 (*CHST6*) gene causing macular corneal dystrophy in India

Afia Sultana,[1] Mittanamalli S. Sridhar,[2] Aparna Jagannathan,[1] Dorairajan Balasubramanian,[1] Chitra Kannabiran,[1] Gordon K. Klintworth[3,4]

[1]Kallam Anji Reddy Molecular Genetics Laboratory, Professor Brien Holden Eye Research Centre and [2]Cornea and Anterior Segment Services, L. V. Prasad Eye Institute, L. V. Prasad Marg, Banjara Hills, Hyderabad, India; Departments of [3]Ophthalmology and [4]Pathology, Duke University Medical Center, Durham, NC

Purpose: Macular corneal dystrophy (MCD) is an autosomal recessive disorder characterized by progressive central haze, confluent punctate opacities and abnormal deposits in the cornea. It is caused by mutations in the carbohydrate sulfotransferase-6 (*CHST6*) gene, encoding corneal N-acetyl glucosamine-6-O-sulfotransferase (C-GlcNAc-6-ST). We screened the *CHST6* gene for mutations in Indian families with MCD, in order to determine the range of pathogenic mutations.

Methods: Genomic DNA was isolated from peripheral blood leukocytes of patients with MCD and normal controls. The coding regions of the *CHST6* gene were amplified using three pairs of primers and amplified products were directly sequenced.

Results: We identified 22 (5 nonsense, 5 frameshift, 2 insertion, and 10 missense) mutations in 36 patients from 31 families with MCD, supporting the conclusion that loss of function of this gene is responsible for this corneal disease. Seventeen of these mutations are novel.

Conclusions: These data highlight the allelic heterogeneity of macular corneal dystrophy in Indian patients.

The corneal dystrophies are a heterogeneous group of disorders that may lead to severe visual impairment [1]. Macular corneal dystrophy (MCD, MCDC1 [OMIM 217800]) is an autosomal recessive disorder clinically characterized by bilateral corneal opacification. Initially the patients have diffuse, fine superficial clouding in the central stroma. The opacities extend through the entire thickness of the cornea and involve the central and peripheral corneal stroma. The corneal stroma is often thinner than normal [2-4]. The corneal endothelium is involved and guttae form on the Descemet's membrane [5]. The prevalence of MCD varies immensely in different parts of the world but in most populations the condition is rare. In some countries MCD accounts for 10-75% of the corneal dystrophies requiring corneal grafting [6,7].

Keratan sulfate (KS) is the major corneal glycosaminoglycan [8] and is a component of three corneal proteoglycans (lumican, keratocan, and mimecan). Sulfate ions contribute significantly to the negative charge of proteoglycans and in the cornea this highly anionic charge on KS is believed to contribute to its ability to imbibe water and influence corneal hydration, which must be critically controlled for the maintenance of corneal transparency [9]. During corneal development, sulfation of polylactosamine (non-sulfated precursor of KS) occurs at later stages and is required for attainment of full corneal transparency [10]. Macular corneal dystrophy arises due to a failure to synthesize normal keratan sulfate proteoglycan within the cornea [11-13] and corneas with this disorder accumulate a glycoaminoglycan within the keratocytes, corneal endothelium, Bowman's layer, Descemet's membrane, and extracellularly in the stroma.

MCD is divided into three immunophenotypes (MCD types I, IA, and II) based on the reactivity of the patient's serum and corneal tissue to an antibody that recognizes sulfated KS, although these subtypes are clinically indistinguishable from each other [14-16]. In MCD type I neither the serum nor the corneal tissue contain antigenic KS (AgKS). In MCD type IA, sulfated KS is absent in the cornea and the serum but can be detected in the keratocytes [16]. MCD type II is characterized by the presence of AgKS in the corneal tissue and normal levels of AgKS in the serum. After identifying the locus for MCD on chromosome 16 [17] and fine mapping the gene [18,19], mutations in the carbohydrate sulfotransferase gene (*CHST6*) encoding corneal N-acetyl glucosamine-6-O-sulfotransferase (C-GlcNAc-6-ST) were identified as the cause of MCD types I and II [20]. While mutations of the coding region of *CHST6* were found in MCD types I, IA, and II, deletions and rearrangements of the upstream region of *CHST6* as well as missense mutations have been reported in MCD type II [20-22].

MCD is one of the most common types of corneal dystrophy in India and constitutes about 30% of all dystrophies requiring penetrating keratoplasty (G.K. Vemuganti, personal communication). We screened 36 patients with MCD belonging to 31 families for mutations in the *CHST6* gene. Our study revealed 22 mutations 17 of which are novel, consisting of deletions, insertions, missense, and nonsense mutations.

Correspondence to: Dr. Chitra Kannabiran, Kallam Anji Reddy Molecular Genetics Laboratory, Professor Brien Holden Eye Research Centre, L. V. Prasad Eye Institute, L. V. Prasad Marg, Banjara Hills, Hyderabad 500 034, India; Phone: 91-40-2354 8267; FAX: 91-40-2354 8271; email: chitra@lvpei.org

METHODS

Clinical evaluation and patient selection: The study had the approval of Institutional Review Boards of the L. V. Prasad Eye Institute and Duke University Medical Center and conformed to the tenets of the declaration of Helsinki. Informed consent was obtained from all participants for clinical and molecular genetic studies. Patients were evaluated clinically by a corneal specialist (M.S.S.). Patients were diagnosed as having MCD based on typical clinical features. All patients had a characteristic haze and whitish nodular lesions involving the central cornea. According to the patients, clinical symptoms first became apparent at 13 to 49 years of age (Table 1). Decreased visual acuity was the common presenting complaint (found in 85% of patients) followed by irritation and tearing (45% of patients) and the presence of a white opacity (20% of patients). Among patients who did not undergo penetrating keratoplasty, the visual acuity ranged from counting fingers to 20/30. Climatic droplet keratopathy (spheroidal degeneration) was associated in 2 patients (5%), who not only had the typical features of MCD, but also golden-colored spherical deposits in the superficial cornea. None of the patients had raised intraocular pressure or posterior segment abnormalities. Out of 17 patients who underwent penetrating keratoplasty (PK), the grafts remained clear in all except 5 cases for the period of follow up, which ranged from 0 to 14 years. The grafts failed in five cases and in all of these instances the grafted tissue became edematous, without clinical evidence of recurrent MCD. In 17 patients who had undergone keratoplasty (8 unilateral grafts, 9 bilateral grafts) diagnosis was confirmed by histopathologic examination. Of the 36 patients studied, 14 were sporadic cases with no other affected relatives while the rest had familial disease. Consanguinity was known to be present in 20 out of 31 families.

PCR amplification and sequence analysis: Genomic DNA was extracted from 8-10 ml of blood samples obtained from all the patients involved in the study and also from the normal controls without any corneal diseases. The *CHST6* gene is 16.9 kb in length and consists of 4 exons, of which only exon 4 contains the coding sequence of 1,189 bp. For amplification of the *CHST6* coding region, 3 primer pairs were used for 5'-, middle, and 3'-segments as described by Akama et al. [18]. All PCR reactions were done (thermal cycler PTC 200 MJ Research, Watertown, MA) using 75 ng genomic DNA in 25 μl reaction containing 1X PCR buffer (200 μM of dNTPs, 0.5 μM of each primer, 4% dimethyl sulfoxide [DMSO] and 1 U *Taq* polymerase). PCR products were purified on Amicon columns (Microcon PCR; Millipore, Bedford, MA) and directly sequenced using fluorescent dideoxynucleotides (Big Dye

Table 1. Mutations of the *CHST6* gene among patients with macular corneal dystrophy

Family	Mutation	Mutation type	Consequence in protein	Restriction site change	Age (yrs) of onset	Consanguinity	Corneal graft
1	c.708-732del	Deletion (25 bp)	FS at R5	-	14	N	N
2	c.744C>T	Nonsense	Q18X	Xmi1	26	N	Bilateral
3	c.786-792del	Deletion (7 bp)	FS at P31	-	26	Y	N
4	c.847G>A	Missense	G52D	Hinf1	13	Y	N
5	c.847G>A	Missense	G52D	Hinf1	40	Y	Bilateral
6	c.850C>T*	Missense	S53L	Eco241	23	Y	Unilateral
7	c.850C>T*	Missense	S53L	Eco241	20	N	N
8	c.872delC	Deletion (1 bp)	FS at F60	Mse1	15	Y	N
9	c.890delC*	Deletion (1 bp)	FS at V66	-	43	N	N
10	[c.985C>G, c.986C>G] + c.1012T>C	Missense	S98W+F107S	-	14	N	Bilateral
11	c.1012T>C*	Missense	F107S	-	10	Y	Unilateral
12	c.1012T>C*	Missense	F107S	-	15	Y	N
13	c.1055C>G	Missense	F121L	BseN1	23	Y	N
14	c.1061G>A	Nonsense	W123X	Cfr131	16	N	Bilateral
15	c.1151C>A	Nonsense	C153X	-	34	N	Unilateral
16	c.1279insACG	Insertion	R195-196ins	-	10	Y	Bilateral
17	c.1296C>A	Missense	R202S	Hph1	21	Y	Bilateral
18	c.1303C>A*	Missense	P204Q	Cfr421	49	Y	N
19	c.1304-1306 del3insAT*	Deletion + insertion	FS at P204	Cfr421	25	N	N
20	[c.1321C>T + c.1355C>G]	Missense	S210F+D221E	-	29	Y	N
21	c.1353G>T	Missense	D221Y	Nmuc1	19	Y	Bilateral
22	c.1348insCTG	Insertion (in frame)	W219-220ins	-	17	N	Unilateral
23	c.1348insCTG	Insertion (in frame)	W219-220ins	-	15	Y	Unilateral
24	c.1348insCTG	Insertion (in frame)	W219-220ins	-	12	Y	N
25	c.1348insCTG	Insertion (in frame)	W219-220ins	-	14	Y	N
26	c.1355C>G	Missense	D221E	Nmuc1	13	Y	Unilateral
27	c.1355C>G	Missense	D221E	Nmuc1	childhood	N	Bilateral
28	c.1355C>G	Missense	D221E	Nmuc1	35	Y	Unilateral
29	c.1617G>T	Nonsense	G309X	Tspr1	22	Y	N
30	c.1731G>T	Nonsense	E347X	-	34	N	Unilateral
31	c.1731G>T	Nonsense	E347X	-	15	Y	Bilateral

Shown are details of probands of the families studied and mutations identified. The asterisks ("*") designate mutations reported previously. FS indicates a frameshift. The last column indicates whether corneal grafts were performed and, if yes, whether they were bilateral or unilateral.

Terminator Kit; Applied Biosystems Incorporated, Foster City, CA) on an automated DNA sequencer (ABI Prism 377). The sequences were compared with the published cDNA sequence of *CHST6* (AF219990) and mutations identified were excluded from 75 control DNA samples using restriction enzymes where appropriate. For mutations without changes in restriction enzyme sites, single strand conformation polymorphism (SSCP) using fragments <300 bp in length amplified with overlapping sets of primers or direct sequencing was performed.

RESULTS

In the population analyzed, mutations in the coding region of *CHST6* were identified in all cases. We found 22 different mutations and none of them were detected in a control population of the same ethnic origin as determined by restriction enzyme digests or SSCP analysis on at least 150 chromosomes, suggesting that they are likely to be pathogenic mutations.

Out of the 22 identified mutations, there were 4 deletions, 2 insertions, 1 complex mutation consisting of deletion and insertion, 5 nonsense and 10 missense mutations (Table 1). A homozygous 25 bp deletion was identified (family 1), resulting in a frameshift at arginine-5, followed by 55 amino acids of altered reading frame before a stop codon. A second homozygous deletion of 7 bp was found in family 3, producing a frameshift at proline-31, with termination after 36 amino acids. In family 8, homozygous deletion of a single base C-872 was found, leading to a frameshift at phenylalanine-60, with termination after eight amino acids. A heterozygous deletion of C-890 predicting a frameshift at valine-66 and immediate termination of protein at amino acid 70 was found in family 9. In this family, the second mutation has not been identified upon sequencing of the coding region. A complex homozygous mutation, consisting of a deletion of GCG at nucleotide 1304 and an insertion of AT was seen in family 19, which resulted in a frameshift at proline-204. A 174 amino acid polypeptide of altered reading frame is predicted after the frameshift.

Five nonsense mutations were found in 6 families, involving the formation of stop codons at glutamine-18 (Q18X) in family 2, tryptophan-123 (W123X) in family 14, cysteine-153 (C153X) in family 15, glycine-309 (G309X) in family 29 and glutamic acid-347 (E347X) in families 30 and 31.

Two insertions resulting in an in-frame insertion of a single amino acid were found. A recurrent homozygous insertion of CTG at nucleotide 1348 in families 22-25 resulted in the insertion of an extra amino acid tyrosine, after alanine-219. Homozygous insertion of ACG in family 16 resulted in the insertion of an arginine residue after leucine-195. Since there is an "ACG" sequence from nucleotides 1277-1279 of the *CHST6* cDNA, the insertion of an ACG either at position 1276 or at 1279 would produce the same result. Hence the insertion could have occurred at either of these positions with the resulting amino acid sequence being identical in both cases.

Apart from these, we identified 10 missense mutations in 16 families (Table 1). Homozygous mutations identified were glycine-52 to aspartic acid (G52D) in 2 families, serine-53 to leucine (S53L) in 2 families, phenylalanine-107 to serine (F107S) in 2 families, phenylalanine-121 to leucine (F121L) in one family, arginine-202 to serine (R202S) in one family, proline-204 to glutamine (P204Q) in one family, aspartic acid-221 to glutamic acid (D221E) in 3 families, and aspartic acid-221 to tyrosine (D221Y) in one family. Two novel compound heterozygous missense mutations were identified in families 10 and 20. In family 10, changes of serine-98 to tryptophan (S98W) and phenylalanine-107 to serine (F107S) were found in two affected members of the family. In family 20, the proband was a compound heterozygote for serine-210 to phenylalanine (S210F) and aspartic acid-221 to glutamic acid (D221E) missense mutations.

We failed to detect any apparent relationship between *CHST6* mutations and the clinical manifestations in the patients with MCD. Families with the same mutations had variations in the clinical features. For example, age of onset as well as presence of deposits up to the corneoscleral limbus varied among members of families 26-28 which had the D221E change (Table 1). Truncating and non-truncating mutations did not cause apparent differences in the age of onset of the cases.

DISCUSSION

Corneal N-acetyl glucosamine-6-O-sulfotransferase catalyzes the transfer of a sulfate group to C6 of N-acetyl glucosamine in keratan sulfate, a common component of corneal proteoglycans. Structurally, this enzyme shares conserved sequences with many sulfotransferases (NCBI Conserved Domain Database), consisting of a single α/β domain [23]. It is presumed to play a key role in the production of corneal keratan sulfate [24,25].

We studied a cohort of 36 patients (31 unrelated) with MCD, which is a common corneal dystrophy in India. Twenty-two mutations including 17 novel mutations were identified. A notable feature of the mutations in this series is that 10 are predicted to result in a premature termination of the encoded protein and would be expected to produce functionally null alleles. As shown in Table 1 the majority of the unrelated patients (28) had homozygous pathogenic mutations within the coding region of *CHST6*. Two patients (from families 10 and 20), were compound heterozygotes for two missense mutations. However, in 1 family (family 9), only a single heterozygous change was detected within the coding region of *CHST6*. Because the second mutation was not identified in the coding region, it is possible that it is located in the promoter or in a non-coding region of the gene.

Overall, the data indicate a high degree of heterogeneity among the patients studied. Similar heterogeneity has been a feature in other studies carried out on individuals with MCD in other populations including those from Britain [22], France [26], Iceland [21], India [27], Japan [28], Saudi Arabia [29], United States of America [30], and Vietnam [31]. Warren et al. [27] in their study on MCD in a Southern Indian population also noted a high degree of mutational heterogeneity. We found only three of the mutations that were reported by them in our patient population (see below). The spectrum of mutations described in their study comprised missense, deletions,

and complex mutations. In contrast to our study they did not observe any nonsense mutations [27]. Since our patients represent diverse regions, our data do not permit conclusions as to whether any of the mutations are specific to an ethnic subgroup. An analysis of a larger cohort of patients would be required to determine if nonsense mutations are a significant type of mutation in MCD patients.

To assess the significance of the identified missense mutations, an alignment of protein sequences of the conserved domains from several sulfotransferases of human and mouse origin was examined. As can be seen from Table 2, all residues mutated are highly conserved, suggesting that these amino acids are essential for the structure and function of the protein. Five mutations involve the conserved domains required for interacting with the sulfate donor 3'-phospho adenosine 5'-phosphosulfate (PAPS) [32]. Mutations G52D and S53L are located within the 5'-PSB (5'-phosphosulfate-binding) domain which interacts with the 5'-phosphate group of PAPS and is highly conserved among all sulfotransferases known to date [23]. The R202S, P204Q, and S210F involve the 3'-PB (3'-phosphate-binding) domain, the highly conserved domain interacting with the 3'-phosphate of PAPS. Five of the mutations identified in the present study were previously reported; c.850C>T (S53L), c.890delC, and del1304-1306insAT were reported in Indian patients [27], the F107S mutation has been reported previously in a British family [22] and the P204Q mutation has been recently reported in Japanese patients [28].

The only substitution that involved replacement with a chemically similar amino acid is D221E (Table 1). The occurrence of this change in probands from 3 families (Table 1, families 26-28) as a homozygous change and in one family (family 20) as a heterozygous change, together with its absence in 75 unrelated control individuals supports the conclusion that it is pathogenic. Structural modeling of normal and mutant proteins as well as in vitro studies on the properties of the wild type and mutant proteins may be required to determine the exact consequences of this change. It is possible that a mutation of aspartic acid-221 causes changes in local structure or alters protein stability or processing, thus leading to deficient enzyme activity.

Analysis of clinical parameters indicated that most patients in our study had similar features. There were no consistent differences in phenotype between patients with the various amino acid substitutions and truncating mutations. Biochemical studies on the enzymatic activity of the mutant proteins identified in our study may help elucidate the role of specific residues on the function of the *CHST6* gene product.

ACKNOWLEDGEMENTS

This work was supported by the Hyderabad Eye Research Foundation and in part by a research grant from the National Eye Institute (RO1-EY08249).

Table 2. Sequence alignment of sulfotransferases showing conservation of amino acids mutated in MCD patients

	G52D, S53L	S98W	F107S	F121L	R202S, P204Q	D221E/Y	S210F
Human corneal GlcNAc 6-O-sulfotransferase (AF219990)	S G S S	R S V F	D V F D	L F Q W	V R D P	A R D N	L R S R
Human intestinal GlcNAc 6-O-sulfotransferase (AF246718)	S G S S	R S I F	D V F D	F F N W	V R D P	A R D N	L R S R
N-acetylglucosamine-6-O-sulfotransferase (human) (AF131235)	S G S S	R A V F	S V F D	L F Q W	V R D P	M I D S	F R S R
Human chondroitin 6-sulfotransferase-2 (AB037187)	T G S S	R S L F	S V L R	L F R W	F R D P	L R E S	H N S R
Mouse L-selectin ligand sulfotransferase (AF131235)	S G S S	R S V F	S V F D	L F Q W	V R D P	V V D S	F R S R
N-acetylglucosamine-6-O-sulfotransferase (Mouse) (AF176841)	S G S S	R S V F	D V F D	L F Q W	V R D P	A R D N	L R S R
Mouse chondroitin 6-sulfotransferase (AB008937)	T G S S	K Q L L	Y V L E	L F R R	V R D P	Y E N W	L A S R

The protein sequences of the sulfotransferases shown were aligned using the software Omiga version 2.0 (Oxford Molecular Ltd., Cambridge, England). Residues mutated along with flanking amino acids are shown in this table. Mutations are indicated at the top of each column. GenBank accession numbers of the sequences are given in parentheses.

REFERENCES

1. Klintworth GK. The molecular genetics of the corneal dystrophies—current status. Front Biosci 2003; 8:687-713.
2. Ehlers N, Bramsen T. Central thickness in corneal disorders. Acta Ophthalmol (Copenh) 1978; 56:412-6.
3. Donnenfeld ED, Cohen EJ, Ingraham HJ, Poleski SA, Goldsmith E, Laibson PR. Corneal thinning in macular corneal dystrophy. Am J Ophthalmol 1986; 101:112-3.
4. Quantock AJ, Meek KM, Ridgway AE, Bron AJ, Thonar EJ. Macular corneal dystrophy: reduction in both corneal thickness and collagen interfibrillar spacing. Curr Eye Res 1990; 9:393-8.
5. Francois J. Heredo-familial corneal dystrophies. Trans Ophthalmol Soc U K 1966; 86:367-416.
6. Jonasson F, Johannsson JH, Garner A, Rice NS. Macular corneal dystrophy in Iceland. Eye 1989; 3 (Pt 4):446-54.
7. Santo RM, Yamaguchi T, Kanai A, Okisaka S, Nakajima A. Clinical and histopathologic features of corneal dystrophies in Japan. Ophthalmology 1995; 102:557-67.
8. Funderburgh JL. Keratan sulfate: structure, biosynthesis, and function. Glycobiology 2000; 10:951-8.
9. Funderburgh JL, Funderburgh ML, Mann MM, Conrad GW. Physical and biological properties of keratan sulphate proteoglycan. Biochem Soc Trans 1991; 19:871-6.
10. Cornuet PK, Blochberger TC, Hassell JR. Molecular polymorphism of lumican during corneal development. Invest Ophthalmol Vis Sci 1994; 35:870-7.
11. Klintworth GK, Smith CF. Macular corneal dystrophy. Studies of sulfated glycosaminoglycans in corneal explant and confluent stromal cell cultures. Am J Pathol 1977; 89:167-82.
12. Hassell JR, Newsome DA, Krachmer JH, Rodrigues MM. Macular corneal dystrophy: failure to synthesize a mature keratan sulfate proteoglycan. Proc Natl Acad Sci U S A 1980; 77:3705-9.
13. Nakazawa K, Hassell JR, Hascall VC, Lohmander LS, Newsome DA, Krachmer J. Defective processing of keratan sulfate in macular corneal dystrophy. J Biol Chem 1984; 259:13751-7.
14. Yang CJ, SundarRaj N, Thonar EJ, Klintworth GK. Immunohistochemical evidence of heterogeneity in macular corneal dystrophy. Am J Ophthalmol 1988; 106:65-71.
15. Edward DP, Yue BY, Sugar J, Thonar EJ, SunderRaj N, Stock EL, Tso MO. Heterogeneity in macular corneal dystrophy. Arch Ophthalmol 1988; 106:1579-83.
16. Klintworth GK, Oshima E, al-Rajhi A, al-Saif A, Thonar EJ, Karcioglu ZA. Macular corneal dystrophy in Saudi Arabia: a study of 56 cases and recognition of a new immunophenotype. Am J Ophthalmol 1997; 124:9-18.
17. Vance JM, Jonasson F, Lennon F, Sarrica J, Damji KF, Stauffer J, Pericak-Vance MA, Klintworth GK. Linkage of a gene for macular corneal dystrophy to chromosome 16. Am J Hum Genet 1996; 58:757-62.
18. Liu NP, Baldwin J, Jonasson F, Dew-Knight S, Stajich JM, Lennon F, Pericak-Vance MA, Klintworth GK, Vance JM. Haplotype analysis in Icelandic families defines a minimal interval for the macular corneal dystrophy type I gene. Am J Hum Genet 1998; 63:912-7.
19. Liu NP, Dew-Knight S, Jonasson F, Gilbert JR, Klintworth GK, Vance JM. Physical and genetic mapping of the macular corneal dystrophy locus on chromosome 16q and exclusion of TAT and LCAT as candidate genes. Mol Vis 2000; 6:95-100 .
20. Akama TO, Nishida K, Nakayama J, Watanabe H, Ozaki K, Nakamura T, Dota A, Kawasaki S, Inoue Y, Maeda N, Yamamoto S, Fujiwara T, Thonar EJ, Shimomura Y, Kinoshita S, Tanigami A, Fukuda MN. Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulphotransferase gene. Nat Genet 2000; 26:237-41.
21. Liu NP, Dew-Knight S, Rayner M, Jonasson F, Akama TO, Fukuda MN, Bao W, Gilbert JR, Vance JM, Klintworth GK. Mutations in corneal carbohydrate sulfotransferase 6 gene (CHST6) cause macular corneal dystrophy in Iceland. Mol Vis 2000; 6:261-4 .
22. El-Ashry MF, El-Aziz MM, Wilkins S, Cheetham ME, Wilkie SE, Hardcastle AJ, Halford S, Bayoumi AY, Ficker LA, Tuft S, Bhattacharya SS, Ebenezer ND. Identification of novel mutations in the carbohydrate sulfotransferase gene (CHST6) causing macular corneal dystrophy. Invest Ophthalmol Vis Sci 2002; 43:377-82.
23. Negishi M, Pedersen LG, Petrotchenko E, Shevtsov S, Gorokhov A, Kakuta Y, Pedersen LC. Structure and function of sulfotransferases. Arch Biochem Biophys 2001; 390:149-57.
24. Nakazawa K, Takahashi I, Yamamoto Y. Glycosyltransferase and sulfotransferase activities in chick corneal stromal cells before and after in vitro culture. Arch Biochem Biophys 1998; 359:269-82.
25. Akama TO, Nakayama J, Nishida K, Hiraoka N, Suzuki M, McAuliffe J, Hindsgaul O, Fukuda M, Fukuda MN. Human corneal GlcNac 6-O-sulfotransferase and mouse intestinal GlcNac 6-O-sulfotransferase both produce keratan sulfate. J Biol Chem 2001; 276:16271-8.
26. Niel F, Ellies P, Dighiero P, Soria J, Sabbagh C, San C, Renard G, Delpech M, Valleix S. Truncating mutations in the carbohydrate sulfotransferase 6 gene (CHST6) result in macular corneal dystrophy. Invest Ophthalmol Vis Sci 2003; 44:2949-53.
27. Warren JF, Aldave AJ, Srinivasan M, Thonar EJ, Kumar AB, Cevallos V, Whitcher JP, Margolis TP. Novel mutations in the CHST6 gene associated with macular corneal dystrophy in southern India. Arch Ophthalmol 2003; 121:1608-12.
28. Iida-Hasegawa N, Furuhata A, Hayatsu H, Murakami A, Fujiki K, Nakayasu K, Kanai A. Mutations in the CHST6 gene in patients with macular corneal dystrophy: immunohistochemical evidence of heterogeneity. Invest Ophthalmol Vis Sci 2003; 44:3272-7.
29. Bao W, Smith CF, al-Rajhi A, Chandler JW, Karcioglu ZA, Akama TO, Fukuda MN, Klintworth, GK. Novel mutations in the CHST6 gene in Saudi Arabic patients with macular corneal dystrophy. ARVO Annual Meeting; 2001 April 29-May 4; Fort Lauderdale, FL.
30. Aldave AJ, Thonar EJ, Warren JF, Self CA, Margolis TP. Novel mutations in the CHST6 gene associated with macular corneal dystrophy in American patients. ARVO Annual Meeting; 2003 May 4-May 9; Fort Lauderdale, FL.
31. Ha NT, Chau HM, Cung le X, Thanh TK, Fujiki K, Murakami A, Hiratsuka Y, Kanai A. Mutation analysis of the carbohydrate sulfotransferase gene in Vietnamese with macular corneal dystrophy. Invest Ophthalmol Vis Sci 2003; 44:3310-6.
32. Kakuta Y, Pedersen LG, Pedersen LC, Negishi M. Conserved structural motifs in the sulfotransferase family. Trends Biochem Sci 1998; 23:129-30.

The print version of this article was created on 22 Dec 2003. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.

Clin Genet 2005: 68: 454–460
Printed in Singapore. All rights reserved

CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2005.00517.x

Short Report

Allelic heterogeneity of the carbohydrate sulfotransferase-6 gene in patients with macular corneal dystrophy

Sultana A, Sridhar MS, Klintworth GK, Balasubramanian D, Kannabiran C. Allelic heterogeneity of the carbohydrate sulfotransferase-6 gene in patients with macular corneal dystrophy.
Clin Genet 2005: 68: 454–460. © Blackwell Munksgaard, 2005

Macular corneal dystrophy (MCD) is an autosomal recessive disorder characterized by grayish white opacities in the cornea. It is caused by mutations in the carbohydrate sulfotransferase-6 (*CHST6*) gene, which codes for the enzyme corneal *N*-acetylglucosamine-6-sulfotransferase. This enzyme catalyzes the sulfation of keratan sulfate, an important component of corneal proteoglycans. We screened 31 patients from 26 families with MCD for mutations in the coding region of the *CHST6* gene. Twenty-six different mutations were identified, of which 14 mutations are novel. The novel mutations are one nonsense mutation found in one patient (Trp2Ter), one frameshift (insertion plus deletion) mutation in two patients (His335fs), and 12 missense mutations (Leu3Met, Ser54Phe, Val56Arg, Ala73Thr, Ser98Leu, Cys165Trp, Ser167Phe, Phe178Cys, Leu193Pro, Pro204Arg, Arg272Ser, and Arg334Cys) in 11 patients. These data demonstrate a high degree of allelic heterogeneity of the *CHST6* gene in patient populations with MCD from Southern India, where this disease may have a relatively higher prevalence than in outbred communities.

A Sultana[a], M S Sridhar[b], G K Klintworth[c], D Balasubramanian[a] and C Kannabiran[a]

[a]Kallam Anji Reddy Molecular Genetics Laboratory; [b]Cornea and Anterior Segment Services, L.V. Prasad Eye Institute, L.V. Prasad Marg, Banjara Hills, Hyderabad, India; and [C]Departments of Ophthalmology and Pathology, Duke University Medical Center, Durham, NC, USA

Key words: carbohydrate sulfotransferase – CHST6 – macular corneal dystrophy – mutations

Corresponding author: Chitra Kannabiran, L.V. Prasad Eye Institute, L.V. Prasad Marg, Banjara Hills, Hyderabad 500 034, India.
Tel.: +91 40 30612 507;
fax: +91 40 2354 8271;
e-mail: chitra@lvpei.org

Received 3 May 2005, revised and accepted for publication 10 August 2005

Macular corneal dystrophy (MCD; MCDC1, MIM217800) is an autosomal recessive disorder characterized clinically by the formation of opacities in the corneal stroma that usually become evident in childhood or adolescence. A diffuse fine symmetric clouding in the central corneal stroma and discrete white opacities develop that extend to the periphery and eventually involve the entire thickness of the cornea, leading to visual impairment (Fig. 1). Corneas with MCD are distinguished histochemically by the presence of intra- and extracellular deposits that stain positively for glycosaminoglycans in the Bowman layer, stroma, Descemet membrane, and endothelium (1). Biochemical studies on surgically excised corneal tissues revealed a defect in the sulfation of keratan sulfate (KS) moieties in the KS-containing proteoglycans (2–4). Although MCD is clinically homogenous, it can be divided into three immunophenotypes, MCD types I, IA, and II, based on the immunohistochemical reactivity to anti-KS antibody in the cornea and serum (5–7). The gene responsible for MCD was mapped to chromosome 16q22.1 (8), and several years later, after fine mapping of the gene (9), mutations in the carbohydrate sulfotransferase-6 (*CHST6*) gene were found in patients with MCD (9, 10). The *CHST6* gene codes for corneal *N*-acetyl glucosamine sulfotransferase, an enzyme that specifically transfers the sulfate group on to the unsulfated KS proteoglycan. In the absence of this sulfation reaction, unsulfated proteoglycan precursors presumably accumulate in the cornea and give rise to the characteristic deposits seen in MCD.

We screened the *CHST6* gene for mutations in 31 patients from 26 Indian families with MCD, and herein, we report 14 novel mutations in this gene.

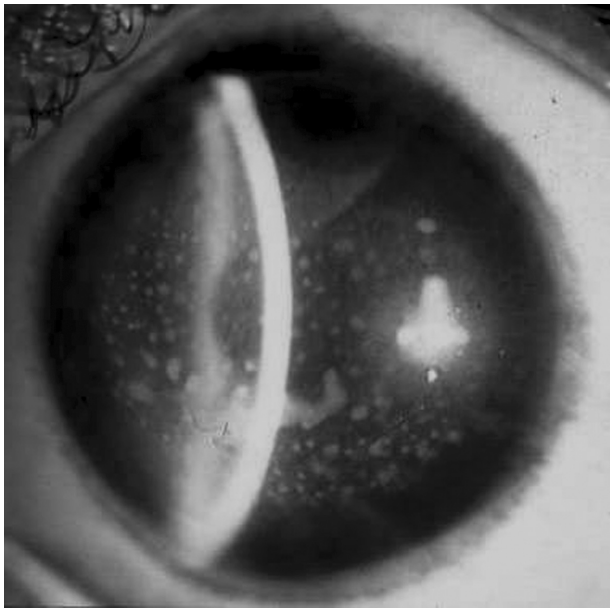


Fig. 1. Clinical photo of patient with macular corneal dystrophy (MCD). Slit lamp view of cornea of patient with MCD showing rounded stromal opacities.

Patients and methods

Patients and sample collection

Thirty-one patients from 26 families were evaluated clinically by a corneal specialist (MSS). The diagnosis of MCD was based upon the distinctive clinical features, and in 20 families, this was confirmed by a histopathologic examination of the excised corneal buttons. Patients underwent complete eye examination. This included visual acuity assessment, slit-lamp examination, intraocular pressure, and fundus examination. In patients where fundus details were not visible, ultrasound B-scan examination was performed. The various slit lamp findings seen in patients with MCD were full thickness corneal haze from limbus to limbus, anterior stromal nodular deposits, corneal thinning, and endothelial deposits.

Age of onset according to the patients ranged from 7 to 33 years. Of the 31 patients studied, nine were sporadic cases with no affected relatives, while 22 patients from 17 families had a positive family history. Consanguinity was known to be present in 20 of 26 families. Peripheral blood samples were collected from 31 patients after obtaining informed consent, and genomic DNA was extracted from the leucocytes. The study protocol adhered to the tenets of the declaration of Helsinki, and the research was done after the prior approval by the institutional review board of L.V. Prasad Eye Institute.

Polymerase chain reaction amplification and sequence analysis

The *CHST6* gene is 16.9 kb in length and consists of four exons, of which only third exon contains the coding region. The coding region was amplified using polymerase chain reaction (PCR) with the primers described previously (9). Conditions for PCR amplification and sequencing were as reported previously (11). The nucleotide sequences were compared with the published cDNA sequence of the *CHST6* gene (GenBank NM_021615).

DNA from 75 unrelated healthy Indian individuals, without a history of MCD, was screened for the sequence changes identified, using restriction enzymes whenever appropriate to confirm pathogenicity. For mutations that did not change any restriction site, direct sequencing or single-strand conformation polymorphism analysis was carried out using primers designed to amplify overlapping fragments of <300 base pairs in length for amplification as described previously (12).

For one patient (no. 39, Table 1) having the two consecutive heterozygous base changes [c.293C > T; c.294C > G], we designed primers specific for mutant and wild-type sequences. We performed allele-specific PCR to know whether the two sequence changes are in *cis* or *trans*. PCR was done with the patient and normal control DNAs at an annealing temperature of 57 °C. As an additional control, we used DNA from an MCD patient having the change [c.293C > G; c294C > G] identified in a previous study (11). This sample differs from the one in the present study by one base in the region tested. Sequences of oligonucleotides were

CHST-ASOF1 [wild type (forward)]: GCGAC CTGGTGCGCT**CC**

CHST-ASOF2 [Mutant (forward)]: GCGACC TGGTGCGCT**TG**

CHST-ASOR [Common (reverse)]: GGGTCG CTGAGCAGCGG

The residues in bold at the 3′ end of the forward primers are the bases that are mutated.

To evaluate the upstream region of the *CHST6* gene in patients with heterozygous mutations in the coding region, we performed PCR on genomic DNA according to the method of Akama et al. (9) using the primers and conditions described on the Nature Genetics website (http://www.nature.com/ng/supplementary_info/ng1000_237/ng1000_237_S1.doc), with the following exceptions: primer F2M was replaced by 5′-CCACAGCCAATTCCATCTTGGATTTTC TC-3′ and primer R2 was replaced by 5′-CATTA

Table 1. Mutations in *CHST6* gene among patients with macular corneal dystrophy

Family/Patient number	Mutation[a]	Mutation type	Consequence in protein	Restriction site change (if any)	Novel/previously reported[b]
32	[c.6G > A; c.7C > A] + [c.6G > A; c.7C > A]	Nonsense + missense	Trp2X + Leu3Met		Novel
33	c.16_40del	Deletion (25 bp)	Val6fs		11
34	c.161C > T	Missense	Ser54Phe	*Mbo*II	Novel
35	[c.166G > A; c.167T > G][c.500C > T] + ;[=]	Missense	Val56Arg + Ser167Phe	*Mbi*I + *Eco*57I	Novel
36	c.180delC	Deletion (1 bp)	Phe60fs		11
37	c.180delC	Deletion (1 bp)	Phe60fs		11
38	c.217G > A	Missense	Ala73Thr	*Bsh*1236I	Novel
39	[c.293C > T; c.294C > G] + [=]	Missense	Ser98Leu		Novel
40	c.369G > A	Nonsense	Trp123X	*Cfr*13I	11
41	c.391T > C	Missense	Ser131Pro		18
42	c.459C > A	Nonsense	Cys153X		11
43	c.494G > C; c.495C > T	Missense	Cys165Ser		16
44	c.495C > G	Missense	Cys165Trp		Novel
45	c.533T > G	Missense	Phe178Cys		Novel
46	[c.545delA] + [c.827 T > C]	Deletion Missense	Gln182fs + Leu276Pro	*Pst*I	15, 21
47	c.578T > C	Missense	Leu193Pro		Novel
48	c.604C > A	Missense	Arg202Ser		11
49	c.611C > G	Missense	Pro204Arg		Novel
50	c.656_657insCTG	Insertion (in frame)	Ala219_Arg220insTrp		11
51	c.656_657insCTG	Insertion (in frame)	Ala219_Arg220insTrp		11
52	c.663C > G	Missense	Asp221Glu		11
53	c.661G > T	Missense	Asp221Tyr		11
54	c.814C > A	Missense	Arg272Ser	*Alu*I	Novel
55	c.1000C > T	Missense	Arg334Cys	*Hae*II	Novel
56	c.1002_1012delinsTTG	Frameshift	His335fs	*Nla*III	Novel
57	c.1002_1012delinsTTG	Frameshift	His335fs	*Nla*III	Novel

This table summarizes the *CHST6* gene mutations identified in probands from each of the 26 families. The family/patient numbers are a continuation of cases after those previously reported (11).

[a]Mutation nomenclature is according to the current recommendations in http://www.hgvs.org/mutnomen/recs.html#DNA.

[b]Whether mutations are novel or reported and the reference of the first report, is shown.

GACACCTCACCTGCTTTGGC-3′. In numerous analyses the original primers (9) did not yield consistent results, perhaps because of mispairing due to the high degree of similarity of CHST5 and *CHST6*. We have hence modified the method using new primers that cover areas where there is less homology between CHST5 and *CHST6*. The annealing temperatures were adjusted to 58 and 62 °C for the primer pairs F2/R2 and F2M/R2, respectively. All amplicons indicating upstream DNA rearrangements were checked using sequencing methods. A new primer, R2M (annealing temperature 61 °C), was introduced to pair with F2M and produce an amplicon on the 3′ side of region 'B' [as defined by Akama et al. (9)]. The 3′ end of this amplicon was sequenced using a forward primer, 5′-GCAGAGGTTGCACACACCTGTC-3′. In our experience, these changes produced a more robust method that yielded more consistent results. Additionally, because two of the resulting amplicons were extended into less homologous areas, their sequences could be more readily verified. All PCR amplicons were electrophoresed on 2% agarose gels, and the gels were documented using the BIOCHEMI IMAGE ACQUISITION AND ANALYSIS Software (UVP BioImaging Systems, Upland, CA).

Results and discussion

Screening of *CHST6* gene identified 26 mutations in 26 families. Two probands had single heterozygous mutations, one proband had a compound heterozygous mutation and all the remaining 23 probands had homozygous mutations. Mutations identified (details in Table 1) consisted of 18 missense mutations (17 patients; three heterozygous), three nonsense mutations (three patients), three deletions (one heterozygous, two homozygous in four patients), one complex mutation (deletion + insertion; two patients), and one insertion (two patients). These sequence changes were absent in 75 unrelated normal individuals.

Novel mutations identified in 13 unrelated patients consisted of one nonsense mutation, one frameshift, and 12 missense mutations. A novel homozygous nonsense mutation predicting a truncation at tryptophan-2 was found in one patient (Table 1). Interestingly, this individual was found to have two homozygous changes, i.e. Trp2Ter and Leu3Met (Fig. 1B). The second mutation is predicted to create a methionine codon just downstream of the termination codon, thus raising the possibility of a re-initiation of translation at this position, causing a protein product that lacks the first three amino acids. Two factors make the synthesis of such a protein unlikely. The initiation codon created as a result of this mutation does not conform to the Kozak sequence because it lacks the highly conserved purine residues located three and six bases upstream of the initiation site (13) and would therefore be expected to function weakly or not at all. In addition, the existence of the nonsense codon may target the mRNA for degradation.

Apart from the Leu3Met mutation discussed above, there were 11 novel missense mutations, three heterozygous and eight homozygous, in 10 patients (Table 1). These were Ser54Phe, Val56Arg, Ala73Thr, Ser98Leu, Ser167Phe, Cys165Trp, Phe178Cys, Leu193Pro, Pro204Arg, Arg272Ser, and Arg334Cys. One complex heterozygous allele with two changes, i.e. Val56Arg and Ser167Phe, was found in one patient (Family 35), and one heterozygous allele with a change of Ser98Leu was found in another patient (Family 39). The mutations were in Family 35, the Val56Arg and Ser167Phe changes, was confirmed to be in *cis* using RFLP analysis as the two mutations resulted in changes in restriction sites for *Mbi*I and *Eco*57I (data not shown). The Ser98Leu mutation involved two consecutive heterozygous base changes [c.293C > T; c.294C > G]. These were tested for being in *cis* or *trans* using allele-specific PCR (Fig. 2). Normal control DNA gave the expected amplified product with the ASOF1 allele-specific primer complementary to the normal sequence but did not show amplification with the mutant allele-specific primer ASOF2 (Fig. 2, lanes 1 and 4). The DNA from Family 39 (Fig. 2, lanes 3 and 5) amplified with allele-specific primers ASOF1 and ASOF2 for both

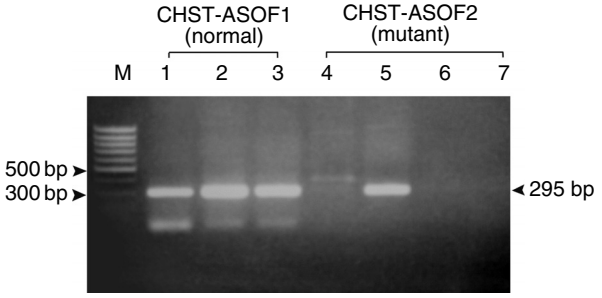


Fig. 2. Results of allele-specific polymerase chain reaction. Details of the assay are in the text. Lanes 1 and 4: DNA from normal control. Lanes 2 and 6: DNA from macular corneal dystrophy family 10 [11]. Lanes 3 and 5: DNA from proband of Family 39. Lane 7: negative control. CHST-ASOF1 forward allele-specific primer for normal sequence. CHST-ASOF2 forward allele-specific primer for mutant sequence. Reverse primer used was common to both sets of reactions. M-DNA size standards.

normal and mutant sequences, respectively (see *Methods*). DNA from MCD Family 10 having the [c.293C > G; c294C > G] mutation identified in a previous study (11) amplified with the ASOF1 primer but not with ASOF2 (Fig. 2, lanes 2 and 6). These results indicate that the two base changes at c.293 and c.294 are in *cis* in Family 39.

The residues involved in missense mutations are fairly well-conserved among various sulfotransferases of human and mouse origin as shown in the partial sequence alignment in Table 2, suggesting that mutations at these sites may be deleterious to the function of the enzyme. Ser54Phe and Val56Arg are located in the conserved 5′ phosphosulfate-binding loop, while Arg202Ser and Pro204Arg are present in the 3′-phosphate-binding domain (14). These domains are required for binding of the sulfate donor 3′-phosphoadenosine-5′-phosphosulfate. Scores from the Blosum 62 substitution matrix for the various amino acid missense mutations in Table 2 have negative values suggesting that these substitutions are unlikely to occur and are therefore possibly unfavorable.

A complex mutation consisting of a deletion plus an insertion resulting in a frameshift at histidine-335 was identified in two patients (Table 1).

The remaining 12 mutations identified in 13 unrelated patients in the present study have been reported previously (Table 1), mostly in Indian patients (11, 15). Consistent with our earlier observations (11), data from the present study demonstrate further allelic heterogeneity in MCD in an Indian patient population. Similar allelic heterogeneity has been observed among patients from several populations including Saudi Arabian (16), British (17), French (18), Vietnamese (19) Southern Indian (15), Icelandic (10), and American (20, 21).

We identified both mutant alleles in all except two patients who had single heterozygous changes (Table 1). Analysis of the upstream regions of CHST5 and *CHST6* in these patients for deletions/rearrangements (9) revealed no alterations (data not shown).

A major proportion of the data available on mutations in MCD so far is from the Indian population. We have analyzed 57 unrelated Indian probands with MCD for mutations in *CHST6* (this study; 11). Of the 47 mutations that we identified in *CHST6* in India in this study and an earlier one (11), 30 were not previously reported. Recurrent mutations in the population studied include c.656_657insCTG, giving rise to an in-frame insertion of tryptophan at position 219 (Ala219_Arg220insTrp), and a missense mutation leading to change of aspartic acid-221 to glutamic acid (Asp221Glu). Both these mutations were each found in 6/57 patients. A deletion of C at position 872 of the cDNA, producing a frameshift at phenylalanine-60, was also recurrent in our patient population, and it was found in 3/57 patients [Table 1; (11)]. MCD represents one of the major inherited corneal disorders among patients presenting at our tertiary care institution in Southern India, accounting for at least one-third of all corneal grafts performed on patients with corneal dystrophies. In general, autosomal recessive corneal disorders appear to predominate over the autosomal dominant types in this patient population, possibly reflecting the high prevalence of consanguinity found in communities from Southern India.

Mutations are distributed throughout the length of the protein, which shares blocks of conserved sequences with other sulfotransferases and is predicted to form a tertiary structural domain that is conserved among sulfotransferase enzymes (NCBI Conserved Domain Database, URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = cdd) (22). A comparison of the mutational spectrum that we obtained with that reported in other studies is shown in Table 3. Missense mutations appear to be the most frequent cause of MCD, with 136/210 (64%) patients studied from different populations having missense mutations (Table 3). Also, of 99 different mutations in *CHST6* identified so far in patients with MCD, 68 are missense. In Indian patients, null mutations (nonsense and frameshift) account for the remaining one-third to one-half of patients (Table 3). The smaller number of patients studied in other ethnic groups precludes any analysis of the relative frequencies of the other categories of mutation. The vast number of missense mutations in *CHST6* all resulting in a fairly uniform clinical picture seem to suggest that none of the resultant mutant proteins found in MCD patients would have any residual activity or that the activity is impaired below a critical level sufficient for disease. Our further studies are aimed at understanding the impact of these pathogenic alterations on the structure of corneal *N*-acetyl glucoamine 6-*O*-sulfotransferase enzyme using structural modeling.

Acknowledgements

This work was supported by a grant from the Department of Biotechnology, Government of India and by the Hyderabad Eye Research Foundation. AS was supported by a fellowship from the Council for Scientific and Industrial Research, India. GKK is supported in part by a research grant (RO1-EY08249) from the National Eye Institute.

Table 2. Sequence alignment of sulfotransferases showing conservation of amino acids mutated in macular corneal dystrophy patients

Sequence	Ser54Phe Val56Arg	Ala73Thr	Ser 98 Leu	Cys165Trp Ser167Phe	Phe178Cys	Leu193Pro	Pro204Arg	Arg272Ser	Arg334Cys
Human corneal GlcNAc 6-*O*-sulfotransferase (NM_021615)	**S** F **V**	G – E P **A**	W – R **S**	V F – A **C** R **S**	– V R **F**	F – P A **L**	N – R D **P**	R – V **R**	F E – A W **R** H
Human intestinal GlcNAc 6-*O*-sulfotransferase (AF246718)	**S** F **V**	G – E P **A**	**W** – R **A**	V F – A **C** R **S**	– V R **F**	F – P A **L**	N – R D **P**	R – V **R**	F E – A W **R** H
N-acetylglucosamine 6-*O*-sulfotransferase (human)(AF131235)	**S** F **V**	G – E P **A**	W – R **S**	V F – A **C** R **S**	– V R **F**	F – P S **L**	N – R D **P**	R – V **R**	Y E – A W **R** W
Human chondroitin 6-sulfotransferase-2 (AB037187)	**S** F **L**	G – E P **M**	W – R **S**	L F – E **C** R **K**	– V R **L**	L – P G **L**	N – R D **P**	R – L **R**	Y E – A W **R** E
Mouse$_L$-selectin ligand sulfotransferase (AF109155)	**S** F **L**	G – E P **A**	W – R **S**	I F – A **C** R **S**	– V R **F**	L – P S **L**	N – R D **P**	R – L **R**	Y E – A W **R** W
N-acetylglucosamine 6-*O*-sulfotransferase (Mouse) (AF176841)	**S** F **V**	G – E P **A**	W – R **S**	V F – A **C** S **S**	– V R **F**	F – P A **L**	N – R D **P**	R – V **R**	Y E – A W **R** H
Mouse chondroitin 6-sulfotransferase (AB008937)	**S** F **V**	G – E P **L**	W – K **Q**	L L – A **C** R **R**	– V R **I**	R – L R **L**	D – R D **P**	R – V **R**	Y E – K W **R** F
Blosum62 substitution score[a]	−2 −3	0	−2	−2 −2	−2	−3	−2	−1	−3

The protein sequences of the sulfotransferases (GenBank accession numbers of the sequences in parentheses) were aligned using the software OMIGA ver 2.0 (Oxford Molecular Ltd, Cambridge, England). Shown above are sequence alignments surrounding the site of novel missense mutations identified in this study. Only partial sequences consisting of four amino acids surrounding the mutant residues are shown for clarity. Residues mutated are highlighted in bold, and mutations are indicated at the top of each column.
[a]Substitution scores on the Blosum-62 matrix for each of the mutations. The values were obtained from http://www.expasy.org/cgi-bin/blosum.pl.

Table 3. CHST6 gene mutations in macular corneal dystrophy patients from different populations

Country of origin of patients	Number of unrelated patients/ families	Missense mutations	Nonsense mutations	Deletion/ Deletion + insertion	Insertion	Compound heterozygous mutations[a]	Upstream deletion/ re-arrangement	No mutation[b]	Reference
India	108	52	9	35	7	1	0	3	[11, 15] present study
America	24	17	1	0	0	1	0	1	[21]
Iceland	6	5	0	0	0	1	NT	0	[10]
Britain	5	5	0	0	0	0	NT	0	[17]
Japan	20	13	0	1	1	1	4	0	[9, 23]
Saudi Arabia	12	12	0	0	0	0	NT	0	[16]
Italy	3	2	0	0	0	0	1	0	[24]
France	11	7	1	0	1	2	0	0	[18]
Vietnam	21	14	1	0	3	3	0	0	[19]

NT, not tested. The table shows the total number of patients screened and the numbers of patients having each of the different types of mutations.
[a]Patients who were heterozygous for two different types of mutation.
[b]No mutations were detected in the regions screened.

References

1. Klintworth GK. Disorders of glycosaminoglycans (mucopolysaccharides) and proteoglycans. In: Garner A, Klintworth GK, eds. Pathobiology of Ocular Disease: A Dynamic Approach, 2nd edn. New York: Marcel Dekker Inc. 1994, 855–892.
2. Klintworth GK, Smith CF. Macular corneal dystrophy. Studies of sulfated glycosaminoglycans in corneal explant and confluent stromal cell cultures. Am J Pathol 1977: 89: 167–182.
3. Hassell JR, Newsome DA, Krachmer JH, Rodrigues MM. Macular corneal dystrophy: failure to synthesize a mature keratan sulfate proteoglycan. Proc Natl Acad Sci USA 1980: 77: 3705–3709.
4. Nakazawa K, Hassell JR, Hascall VC et al. Defective processing of keratan sulfate in macular corneal dystrophy. J Biol Chem 1984: 259: 13751–13757.
5. Edward DP, Yue BY, Sugar J et al. Heterogeneity in macular corneal dystrophy. Arch Ophthalmol 1988: 106: 1579–1583.
6. Yang CJ, SundarRaj N, Thonar EJ, Klintworth GK. Immunohistochemical evidence of heterogeneity in macular corneal dystrophy. Am J Ophthalmol 1988: 106: 65–71.
7. Klintworth GK, Oshima E, al-Rajhi A et al. Macular corneal dystrophy in Saudi Arabia: a study of 56 cases and recognition of a new immunophenotype. Am J Ophthalmol 1997: 124: 9–18.
8. Vance JM, Jonasson F, Lennon F et al. Linkage of a gene for macular corneal dystrophy to chromosome 16. Am J Hum Genet 1996: 58: 757–762.
9. Akama TO, Nishida K, Nakayama J et al. Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulfotransferase gene. Nat Genet 2000: 26: 237–241.
10. Liu NP, Dew-Knight S, Rayner M et al. Mutations in corneal carbohydrate sulfotransferase 6 gene (*CHST6*) cause macular corneal dystrophy in Iceland Mol Vis 2000: 6: 261–264. Retrieved from http://www.molvis.org/molvis/v6/a35.
11. Sultana A, Sridhar MS, Jagannathan A et al. Novel mutations of the carbohydrate sulfotransferase-6 (*CHST6*) gene causing macular corneal dystrophy in India. Mol Vis 2003: 22: 730–734.
12. Kiran VS, Kannabiran C, Chakravarthi K et al. Mutational screening of the RB1 gene in Indian patients with retinoblastoma reveals 8 novel and several recurrent mutations. Hum Mutat 2003: 22 (4): 339.
13. Kozak M. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 1987: 15: 8125–8148.
14. Kakuta Y, Pedersen LG, Pedersen LC, Negishi M. Conserved structural motifs in the sulfotransferase family. Trends Biochem Sci 1998: 23: 129–130.
15. Warren J, Aldave AJ, Srinivasan M et al. Novel mutations in the *CHST6* gene associated with macular corneal dystrophy in Southern India. Arch Ophthalmol 2003: 121: 1608–1612.
16. Bao W, Smith CF, al-Rajhi A et al. Novel mutations in the *CHST6* gene in Saudi Arabic patients with macular corneal dystrophy. Invest Ophthalmol Vis Sci (Suppl) 2001: 42: S483.
17. El Ashry MF, El Aziz MM, Wilkins S et al. Identification of novel mutations in the carbohydrate sulfotransferase gene (*CHST6*) causing macular corneal dystrophy. Invest Ophthalmol Vis Sci 2002: 43: 377–382.
18. Niel F, Ellies P, Dighiero P et al. Truncating mutations in the carbohydrate sulfotransferase 6 gene *CHST6* result in macular corneal dystrophy. Invest Ophthalmol Vis Sci 2003: 44: 2949–2953.
19. Ha NT, Chau HM, Cung LX et al. Mutation analysis in the carbohydrate sulfotransferase gene in Vietnamese with macular corneal dystrophy. Invest Ophthalmol Vis Sci 2003: 44: 3310–3316.
20. Klintworth GK. The molecular genetics of the corneal dystrophies-current status. Front Biosci 2003: 8: d687–783.
21. Aldave AJ, Yellore VS, Thonar EJ et al. Novel mutations in the carbohydrate sulfotransferase gene (*CHST6*) in American patients with macular corneal dystrophy. Am J Opthalmol 2004: 137: 465–473.
22. Marchler-Bauer A, Anderson JB, DeWeese-Scott C et al. 'CDD: a curated Entrez database of conserved domain alignments'. Nucleic Acids Res 2003: 31: 383–387.
23. Abbruzzese C, Kuhn U, Molina F et al. Novel mutations in the *CHST6* gene causing macular corneal dystrophy. Clin Genet 2004: 65: 120–125.
24. Iida-Hasegawa N, Furuhata A, Hayatsu H et al. Mutations in the *CHST6* gene in patients with macular corneal dystrophy: immunohistochemical evidence of heterogeneity. Invest Ophthalmol Vis Sci 2003: 44: 3272–3277.

Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11

Xiaodong Jiao, Afia Sultana, Prashant Garg, Balasubramanya Ramamurthy, Geeta K Vemuganti, Nibaran Gangopadhyay, J Fielding Hejtmancik and Chitra Kannabiran

J. Med. Genet. 2007;44;64-68; originally published online 6 Jul 2006; doi:10.1136/jmg.2006.044644

LETTER TO JMG

Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11

Xiaodong Jiao, Afia Sultana, Prashant Garg, Balasubramanya Ramamurthy, Geeta K Vemuganti, Nibaran Gangopadhyay, J Fielding Hejtmancik, Chitra Kannabiran

J Med Genet 2007;**44**:64–68. doi: 10.1136/jmg.2006.044644

Objective: To map and identify the gene for autosomal recessive congenital hereditary endothelial dystrophy (CHED2, OMIM 217700), a disorder characterised by diffuse bilateral corneal clouding that may lead to visual impairment and requiring corneal transplantation.

Methods: Members of 16 families with autosomal recessive CHED were genotyped for 13 microsatellite markers at the CHED2 locus on chromosome 20p13-12. Two-point linkage analysis was carried out using the FASTLINK version of the MLINK program. Mutation screening was carried out by amplification of exons and flanking regions by polymerase chain reaction, followed by direct automated sequencing.

Results: Linkage and haplotype analysis placed the disease locus within a 2.2 cM (1.3 Mb) interval flanked by D20S198 and D20S889, including SLC4A11. The maximum limit of detection score of 11.1 was obtained with D20S117 at $\theta = 0$. Sequencing of SLC4A11 showed homozygotic mutations in affected members from 12 of 16 families.

Conclusion: These results confirm that mutations in the SLC4A11 gene cause autosomal recessive CHED.

Autosomal recessive corneal endothelial dystrophy, also called congenital hereditary endothelial dystrophy (CHED2, MIM 217700), is an uncommon hereditary corneal disorder clinically presenting at birth or in early childhood. It commonly presents as a corneal haze that has a ground-glass appearance. The cornea is also thickened overall and shows epithelial and stromal oedema. It is clinically similar to autosomal dominant corneal endothelial dystrophy (CHED1, MIM 121700), although it may present earlier.[1] The age of onset and progression of the disease and the degree of visual impairment can be variable, although some children with CHED can develop nystagmus.[2 3] Histologically, corneal endothelial dystrophy is characterised by irregular and multinucleated corneal endothelial cells with a normal or reduced density and markedly thickened Descemet's membrane showing a trilaminar structure with an anterior banded zone, a posterior non-banded zone and a posterior collagenous layer having multiple focal areas of abnormal fibrillar deposits in the posterior half.[4] These fibrillary changes are more common in autosomal dominant CHED, whereas recessive CHED shows failure of growth regulation characterised by a more marked thickening of the anterior banded zone.[5] The disorder is bilateral and has a high degree of penetrance. The preferred treatment is penetrating keratoplasty, which can provide marked visual improvement even in advanced stages of the disease. CHED2 has been mapped to an 8 cM region on chromosome 20p12 flanked by D20S113 and D20S882.[6] This region is distinct from the CHED1 locus in the pericentromeric region of chromosome 20,[7] which colocalises with VSX1.[8]

SLC4A11, sodium bicarbonate transporter-like solute carrier family 4 member 11, is a member of the SLC4 family of bicarbonate transporters, which, along with the SLC26 family, are the main bicarbonate transporter proteins in humans.[9] The SLC4 gene family classically includes three types of bicarbonate exchangers: Cl-HCO_3 exchangers, Na/HCO_3 cotransporters and Na-driver Cl-HCO_3 exchangers. SLC4A11, also called BTR1 or NaBC1, is phylogenetically the most distant member of the SLC4 gene family.[10] Mutations in SLCA1 and SLCA4 result in distal and proximal renal tubular acidosis, respectively. However, SLC4A11 is an electrogenic Na/borate cotransporter that stimulates cell growth and proliferation by increasing intracellular borate and activating the MAPK pathway.[11 12] The SLC4A11 gene on chromosome 20p12 has 19 exons, all containing coding sequences and encoding 891 amino acids. As shown in the National Center for Biotechnology Information UniGene expression profile, it is highly expressed in the eye, blood cells, ovary, tongue, lung, skin and colon, and to a lesser extent in the brain, pancreas, kidney and skin. It is also expressed in some tumours, including gastrointestinal, oral, ovarian, respiratory and skin tumours, as well as in leukaemia and retinoblastoma.

We studied 16 families with autosomal recessive CHED by mapping and candidate gene screening to identify the gene associated with the disease. Our results confirm those of a recent study,[13] and show mutations in the SLC4A11 gene in 12 of 16 families.

METHODS

A diagnosis of CHED was made in all patients by clinical and histopathological evaluation. Patients were examined using indirect ophthalmoscopy and slit lamp examination. Affected members were diagnosed at ages <1–10 years. Inclusion criteria for diagnosis of CHED were the presence of a cloudy cornea from birth to 10 years of age, with increased corneal thickness and bilateral corneal oedema in the presence of a normal corneal diameter and normal intraocular pressure. Exclusion criteria were the presence of Haab's striae or any sign of primary congenital glaucoma, and a history of prior ocular surgery. Histopathological criteria for CHED were thickening of the cornea, thickening of Descemet's membrane and normal/reduced endothelial cell count and morphology. This study was approved by the institutional review boards of The LV Prasad Eye Institute, Hyderabad, India, and the National Eye Institute, Bethesda, Maryland, USA, and informed consent was obtained from each person studied, consistent with the tenets of the Declaration of Helsinki. DNA samples were collected from 16 families, all of which were consanguineous. Twelve families fulfilled all criteria listed earlier. Four families, 73019, 73029, 73035 and 73089, were excluded from linkage analysis because previous surgery or signs of primary congenital glaucoma

Abbreviation: CHED, congenital hereditary endothelial dystrophy

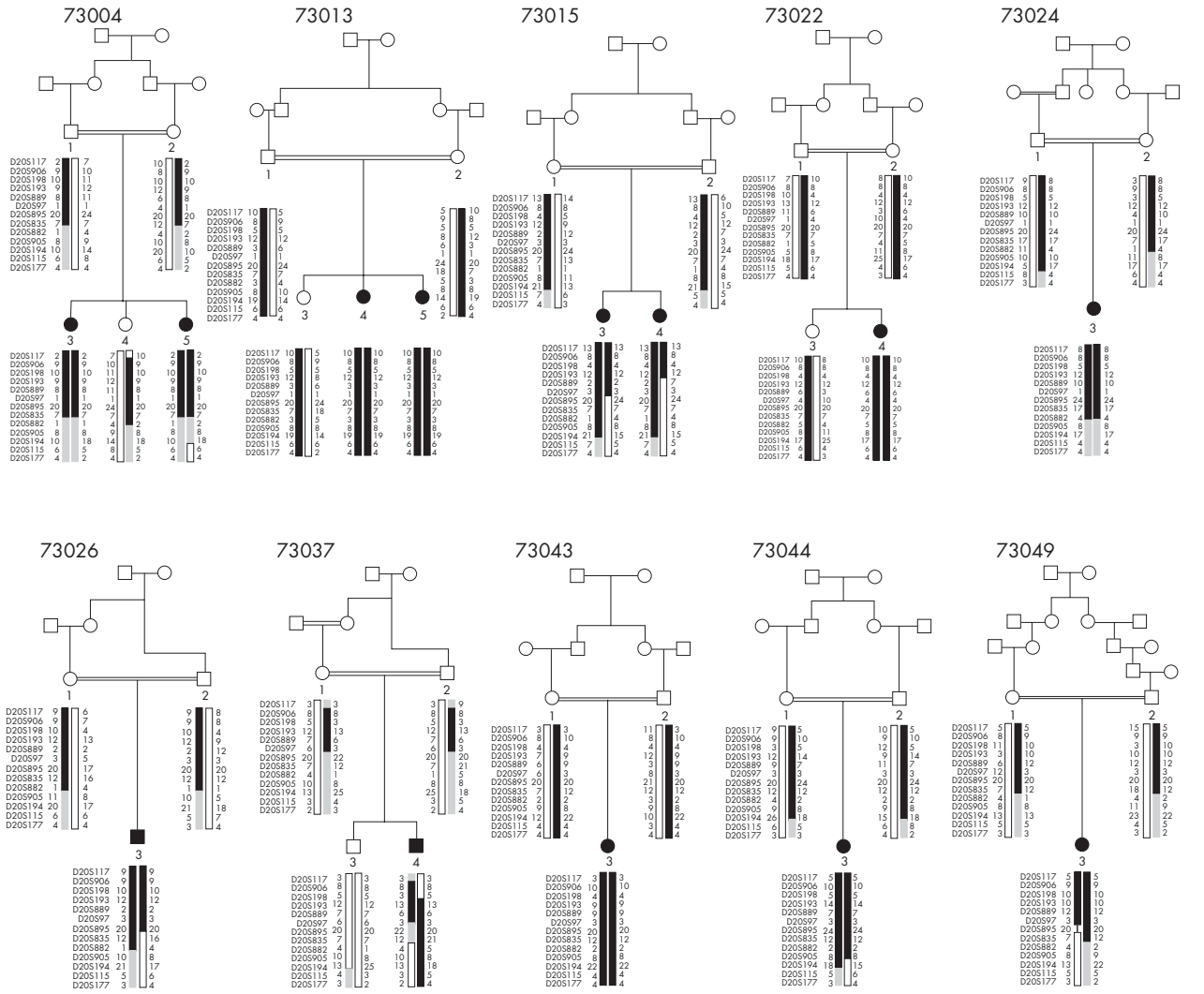


Figure 1 Pedigrees of families 73004, 73013, 73015, 73022, 73024, 73026, 73037, 73043, 73044 and 73049, used in linkage and haplotype analysis. The dark bars correspond to affected haplotypes, including alleles that cosegregate with the disease and are homozygotic in affected members. Grey bars indicate alleles not homozygotic in affected members and white bars indicate recombinant alleles.

prevented diagnostic criteria from being met, but were included in mutation screening. Data from 10 families (13 affected and 24 unaffected members) were used in linkage analysis; pedigrees of these are shown in fig 1.

For linkage analysis, DNA was extracted directly from blood by standard phenol–chloroform protocols.[14] Linkage analysis was carried out with 13 markers in the region of chromosome 20p flanked by D209S113 and D20S882, to which CHED2 had been mapped by Hand *et al*.[6] Markers in this region were genotyped using fluorescent-labelled microsatellite markers (ABI Linkage Mapping Set MD-10, Foster City, California, USA). Multiplexed polymerase chain reaction was carried out as previously described.[15] Two-point linkage analyses were performed with the FASTLINK implementation of the MLINK program of the LINKAGE program package,[16 17] modelling CHED2 as a fully penetrant autosomal recessive disease. Equal allele frequencies arbitrarily set at 0.05 were used for all markers.

The reference cDNA (NM 032034) and genomic sequences (NT 011387) for the SLC4A11 gene are available from the National Center for Biotechnology Information. For mutation screening, coding exons and adjacent intronic sequences of candidate genes in the critical interval were amplified from genomic DNA of 16 affected patients and two unaffected members. Table 1 shows the primers used for amplification and sequencing of the 19 exons of SLC4A11. Exons 18 and 19 were amplified using an Invitrogen High GC Kit (Carlsbad, California, USA). Products of the sequencing reactions were purified using the Edge Biosystem Performa TM DTR Gel Filtration System (Gaithersburg, Maryland, USA) or AmPure and CleanSeq reagents (Agencourt, Boston, Massachusetts, USA) on a Beckman Biomek NX Laboratory Automation Workstation (Fullerton, California, USA).

RESULTS AND DISCUSSION

Evidence of significant linkage was obtained with markers in the chromosome 20p12 region. All markers telomeric of D20S835 gave LOD scores >3, the highest LOD score being observed with D20S117, which gave 11.1 at $\theta = 0$ (table 2). Examination of the haplotypes did indicate a probable recombination with D20S835. D20S889 shows an obligate recombination, although the maximum LOD score at $\theta = 0.01$ is 8.9. D20S115 and DD20S177, which lie centromeric to D20S889, both show obligate recombinations, as do markers

Table 1 Primers used for amplification of exons of SLC4A11 by polymerase chain reaction

Primer	Sequence	Annealing temperature (°C)
Exon 1	5′ CTAGGGTGGCGTGGGTTG 5′ AGCACTAGAGTGGCCCAGAT	60
Exon 2–3	5′ GATGGCCTCTCCCACCAC 5′ CTCCCTGTTGAGCTGCTCCT	60
Exon 4–5	5′ TCCAGGAGCAGCTCAACAG 5′ TCTTCTCCCAAGTTGGTTGG	60
Exon 6	5′ CAAGGTCGAGGGGGTTCT 5′ GTTTCTGACACACCCACAGG	60
Exon 7–8	5′ AGCCTGGGTGACAGTGAGAC 5′ ACAGCCTTGTTTTTCCCAAT	60
Exon 9–10	5′ ACTGATGGTACGTGGCCTCT 5′ CGTCCATGCGTAGAAGGAGT	60
Exon 11–12	5′ CATTGGTGATTCTGCTGACC 5′ ACTCAGCTTGAGCCAGTCCT	60
Exon 13–14	5′ GAGCCCTTTCTCCCTGAGAT 5′ GGTTGTAGCGGAACTTGCTC	60
Exon 15*	5′ GCCTTCTCCCTCATCAGCTC 5′ GTAGGCAGTGCCCTTCACC	60
Exon 16*	5′ AATGCACCGGAGAACAGGT 5′ CCGCGAGTGTCACCTCTG	60
Exon 17*	5′ CGTGGACCCTGAGGAGTG 5′ CCCTCCGGATGTAGTGTGTC	60
Exon 18*	5′ CTCGATGGCAACCAGCTC 5′ CTAGGCAGGACCCCTCCTC	66
Exon 19*	5′ CAGGAGGGGCTCCAGTCTA 5′ ACAGAGCAGTCACCCACACA	66

*Exons 15–19 were amplified using an Invitrogen GC-Rich Kit (Carlsbad).

centromeric to this region. D20S906 and D20S193, both within the included region, gave maximum LOD scores of 9.5 and 9.6, respectively, at $\theta = 0$. D20S198, while not showing an obligate recombination, shows a lack of homozygosity in family 73037, which decreases the LOD score to 2.6 at $\theta = 0$, rising to 6.6 at $\theta = 0.02$. D20S835, D20S882, D20S905 and D20S194 all show negative LOD scores < -2. Thus, using a LOD score of $\leqslant -2$ for exclusion, linkage analysis alone seems to place the CHED2 locus telomeric of D20S835.

Examination of the haplotypes cosegregating with CHED2 in 5 of the 16 families clarifies and extends the linkage results if the simplifying assumption that the rare CHED2 disease allele and alleles of closely linked markers are probably identical by descent in these consanguineous pedigrees is made (fig 1). An obligate recombination event between members 3 and 4 of family 73015 with D20S889 provides the centromeric boundary to the linked region. Additional recombination events and lack of homozygosity in affected members from multiple families for markers centromeric to D20S889 support this interpretation. As this region lies near the telomere of chromosome 20, probable telomeric recombination events are less common, although a lack of homozygosity at D20S198 in member 4 of family 73037 suggests that this is the telomeric boundary. Conservation of the tightly linked haplotypes suggests that D20S117 also recombines, although this is not apparent from the two-point linkage data. Thus, examination of haplotypes in these families places the CHED2 locus between D20S198 and D20S889, a 1.3 Mb (2.2 cM) interval that contains the SLC4A11 gene.

Significant mutations identified in SLC4A11 as described later and in table 3 were not seen in 50 controls (100 chromosomes) of southern Indian ethnicity. All mutations identified in SLC4A11 cosegregated with the disease allele, except the polymorphism identified in family 73026.

CHED2 is a recessive disease, suggesting that causative mutations might result in a loss of SLC4A11 function. Mutation analysis in this set of families is simplified by the occurrence of all mutations in a homozygotic form owing to consanguinity. Some of the observed mutations would be expected to have dramatic effects on the structure or expression of the protein. These include two nonsense mutations: p.Arg605X (g.8298C→T, families 73015 and 73026), and p.Glu632X (g.8379G→T, family 73035). Three frameshift mutations (p.ArgR82ArgfsX33 (g.2943delTTinsA, family 73024), p.His568HisfsX177 (g.8118delCT, family 73013) and p.Leu807ArgfsX71 (g.9200delTinsGG, family 73004)) would also result in loss of a large number of amino acids from the carboxy terminus of the protein and would also replace these sequences with long novel amino acid sequences encoded by the shifted reading frame. In addition, four of these cause premature termination in an internal exon and are likely to be subject to nonsense-mediated decay.[18] The last frameshift mutation, p.Leu807ArgfsX71 (g.9200delTinsGG, family 73004), causes a termination in the final exon and would not be expected to cause nonsense-mediated decay. However, it would remove the final 85 amino acids, replacing them with 71 novel amino acids, and would be expected to cause major, but difficult-to-predict, changes in the SLC4A11 structure. In addition, the deletion mutations p.Arg82ArgfsX33s, p.His568HisfsX177, p.Arg605X and p.Glu632X are expected to change the internal symmetry of the transmembrane channel that is necessary for the cotransport of sodium and boron hydroxide through the membrane, as suggested for the anion exchanger AE11.[11]

Four missense mutations occurred in highly conserved residues in six families (table 3): Arg755Gln (family 73037),

Table 2 LOD scores of markers on chromosome 20p

Marker	Mb*	cM†	LOD score at θ							z_{max}	θ_{max}
			0	0.01	0.05	0.1	0.2	0.3	0.4		
D20S117	0.60	2.9	11.1	10.8	9.7	8.3	5.6	3.2	1.3	11.1	0.00
D20S906	1.45	7.5	9.5	9.2	8.2	6.9	4.5	2.4	0.8	9.5	0.00
D20S198	2.59	8.8	2.6	6.5	6.3	5.5	3.6	1.9	0.6	6.6	0.02
D20S193	3.26	9.4	9.6	9.4	8.4	7.1	4.7	2.5	0.9	9.6	0.00
D20S889	3.89	11.0	$-\infty$	8.9	8.5	7.4	5.0	2.7	1.0	8.9	0.01
D20S97	4.59	11.0	7.6	7.4	6.5	5.5	3.5	1.8	0.6	7.6	0.00
D20S895	5.03	13.7	−1.1	4.0	4.7	4.3	2.9	1.5	0.5	4.7	0.04
D20S835	5.26	14.8	−5.4	1.7	3.1	3.1	2.3	1.3	0.5	3.2	0.07
D20S882	5.58	14.8	−12.7	−1.2	1.5	2.0	1.6	0.8	0.2	2.0	0.09
D20S905	5.81	16.9	−9.2	1.2	3.1	3.3	2.6	1.5	0.6	3.3	0.09
D20S194	6.09	18.0	−18.0	−2.6	0.8	1.7	1.7	1.0	0.3	1.7	0.11
D20S115	7.73	20.9	$-\infty$	0.1	2.7	3.1	2.6	1.6	0.6	3.1	0.09
D20S177	8.86	26.9	$-\infty$	−0.4	2.2	2.7	2.3	1.5	0.6	2.7	0.10

*From the National Center for Biotechnology Information (NCBI) UniSTS.
†From Genethon map, NCBI UniSTS.

Table 3 Summary of SLC4A11 mutations found in families with CHED2

Family	Nucleotide change	Exon	Protein mutation (Blosum 80 Score; % residue conservation*)	Location in protein
73024	g.2943delTTinsA†	2	p. Arg82ArgfsX33	Amino-terminal soluble polypeptide with a new sequence FLSMSTLRCRPPTL added after residue Arg82. Most of the cytoplasmic soluble domain and the transmembrane domain are truncated
73026, 73015‡	g.3552G→A§	4	p. Ala160Thr (0; 3%)	Located in the loop connecting two α-helices from opposite sides of the αβα-sandwich
73013	g.8118delCT†	13	p. His568HisfsX177	Truncates the transmembrane domain after residue His568 and adds a hydrophilic peptide of 28 residues at the C terminus (RPGDRRAQPPHHAGHALAGLHPLPIQEE)
73026, 73015	g.8298C→T†	14	p. Arg605X	Truncates transmembrane domain at residue Arg605
73035‡	g.8379G→T†	14	p. Glu632X	Truncates transmembrane domain at residue Glu632
73037	g.9044G→A	17	p. Arg755Gln (1; 97%)	Located at the surface of transmembrane helix 9. Might affect interaction of transmembrane helix 9 and cytoplasmic membrane
73014‡	g.9191G→A	17	p. Arg804His (0; 100%)	Located in the loop connecting helices 11 and 12 in the predicted transmembrane domain structure. Mutation changes hydrophobic interaction of methyl groups located in ARG stem with those of P640 (distance increases by 1.2 A). This might change the loop stability
73004	g.9200delTinsGG	17	p. Leu807ArgfsX71	Truncation of the C-terminal part after residue L807 and after addition of the six-residue peptide (RAAQGA)
73044, 73029‡	g.9361C→T	18	p. Thr833Met (−1; 91%)	Predicted to be located in transmembrane helix 11. The mutation might destabilise the polar cluster formed by residues Thr833, Gln826, Arg827 and Lys828, and affect the loop stability
7039‡, 7043	g.9469G→A	18	p. Arg869His (0; 94%)	Located at the surface of transmembrane helix 12. Might affect interaction of transmembrane helix 12 and cytoplasmic membrane

*The Blosum 80 substitution matrix is based on a threshold of 80% sequence identity. The percentage residue conservation was calculated by comparing corresponding amino acids in all available members (35) of the SLC4A protein family from the UniProtKB/Swiss-Prot database.
†Subject to potential nonsense-mediated decay.
‡These families were not used in linkage analysis.
§A polymorphism that occurs with g.8298C→T (p. Arg605X) in families 73026 and 73015 and was found in an unaffected relative in family 73026.
Note: Mutations were not observed in families 73022 and 73049, both used in linkage analysis.

Arg804His (family 73014), Thr833Met (families 73029 and 73044) and Arg869His (families 73039 and 73043). A fifth missense change of Ala160Thr was found to be homozygotic in affected members of two families (73026 and 73015). This change is unlikely to have pathogenic significance, as it also occurred in homozygotic form in an unaffected member in one family (73026), and the residue is not conserved among SLC4A11 family members. Affected members in these two families also are homozygotic for the nonsense mutation Arg605X, which did not occur in homozygous form in any unaffected family members. It seems likely that loss of SLC4A11 would not only cause loss of sodium and bicarbonate transport with resultant oedema but also abnormalities in cell division by virtue of its function as a borate transporter. Both these functions are consistent with the clinical phenotype of CHED2.

In summary, CHED2 has been mapped to a 2.2 cM (1.3 Mb) interval flanked by D20S198 and D20S889, including SLC4A11. This region includes SLC4A11, mutations which are shown to be associated with CHED2 in 12 of 16 families examined.

Key points

- Linkage and haplotype analysis of families with autosomal recessive congenital hereditary endothelial dystrophy mapped the disease to a 2.2 cM interval flanked by D20S198 and D20S889 on chromosome 20p12. The maximum limit of detection score of 11.1 was obtained with D20S117 at $\theta = 0$.
- Evaluation of candidate genes in this interval showed mutations in the sodium bicarbonate transporter-like solute carrier family 4 member 11 (SLC4A11) gene in 12 of 16 families.
- Mutations consisted of three frameshift, two nonsense and four missense mutations.

ACKNOWLEDGEMENTS

We thank Professors D Balasubramanian and Ravi Thomas for their help and suggestions and the members of the families for their participation in this study.

Authors' affiliations
X Jiao, J F Hejtmancik, Ophthalmic Genetics and Visual Function Branch, National Eye Institute, NIH, Bethesda, Maryland, USA

Electronic database information

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/
Genethon, http://www.genethon.fr
NCBI, http://www.ncbi.nlm.nih.gov/

A Sultana, P Garg, B Ramamurthy, G K Vemuganti, N Gangopadhyay, C Kannabiran, LV Prasad Eye Institute, Hyderabad, India

Funding: AS was supported by a fellowship from the Council for Scientific and Industrial Research, India. This study was supported by a grant from the Department of Biotechnology, Government of India (#BT/PR4774/MED/12/181/04) and by the Hyderabad Eye Research Foundation.

Correspondence to: Dr C Kannabiran, LV Prasad Eye Institute, Banjara Hills, Hyderabad 500 034, India; chitra@lvpei.org

Received 9 June 2006
Revised 26 June 2006
Accepted 1 July 2006
Published Online First 6 July 2006

REFERENCES

1 **Kirkness CM**, McCartney A, Rice NS, Garner A, Steele AD. Congenital hereditary corneal oedema of Maumenee: its clinical features, management, and pathology. *Br J Ophthalmol* 1987;**71**:130–44.
2 **Weisenthal RW**, Streeten BW. Posterior membrane dystrophies. In: Krachmer JH, Mannis MJ, Holland EJ, eds. *Cornea*. Philadelphia: Elsevier Mosby, 2005:929–54.
3 **Maumenee AE**. Congenital hereditary corneal dystrophy. *Am J Ophthalmol* 1960;**50**:1114–24.
4 **Ehlers N**, Modis L, Moller-Pedersen T. A morphological and functional study of congenital hereditary endothelial dystrophy. *Acta Ophthalmol Scand* 1998;**76**:314–18.
5 **McCartney AC**, Kirkness CM. Comparison between posterior polymorphous dystrophy and congenital hereditary endothelial dystrophy of the cornea. *Eye* 1988;**2**:63–70.
6 **Hand CK**, Harmon DL, Kennedy SM, FitzSimon JS, Collum LM, Parfrey NA. Localization of the gene for autosomal recessive congenital hereditary endothelial dystrophy (CHED2) to chromosome 20 by homozygosity mapping. *Genomics* 1999;**61**:1–4.
7 **Toma NM**, Ebenezer ND, Inglehearn CF, Plant C, Ficker LA, Bhattacharya SS. Linkage of congenital hereditary endothelial dystrophy to chromosome 20. *Hum Mol Genet* 1995;**4**:2395–8.
8 **Heon E**, Greenberg A, Kopp KK, Rootman D, Vincent AL, Billingsley G, Priston M, Dorval KM, Chow RL, McInnes RR, Heathcote G, Westall C, Sutphin JE, Semina E, Bremner R, Stone EM. VSX1: a gene for posterior polymorphous dystrophy and keratoconus. *Hum Mol Genet* 2002;**11**:1029–36.
9 **Romero MF**. Molecular pathophysiology of SLC4 bicarbonate transporters. *Curr Opin Nephrol Hypertens* 2005;**14**:495–501.
10 **Romero MF**, Fulton CM, Boron WF. The SLC4 family of HCO_3-transporters. *Pflugers Arch* 2004;**447**:495–509.
11 **Park M**, Li Q, Shcheynikov N, Zeng W, Muallem S. NaBC1 is a ubiquitous electrogenic Na^+-coupled borate transporter essential for cellular boron homeostasis and cell growth and proliferation. *Mol Cell* 2004;**16**:331–41.
12 **Park M**, Li Q, Shcheynikov N, Muallem S, Zeng W. Borate transport and cell growth and proliferation. Not only in plants. *Cell Cycle* 2005;**4**:24–6.
13 **Vithana EN**, Morgan P, Sundaresan P, Ebenezer ND, Tan DT, Mohamed MD, Anand S, Khine KO, Venkataraman D, Yong VH, Salto-Tellez M, Venkatraman A, Guo K, Hemadevi B, Srinivasan M, Prajna V, Khine M, Casey JR, Inglehearn CF, Aung T. Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2). *Nat Genet* 2006;**38**:755–7.
14 **Smith RJ**, Holcomb JD, Daiger SP, Caskey CT, Pelias MZ, Alford BR, Fontenot DD, Hejtmancik JF. Exclusion of Usher syndrome gene from much of chromosome 4. *Cytogenet Cell Genet* 1989;**50**:102–6.
15 **Jiao X**, Munier FL, Iwata F, Hayakawa M, Kanai A, Lee J, Schorderet DF, Chen MS, Kaiser-Kupfer M, Hejtmancik JF. Genetic linkage of Bietti crystalline corneoretinal dystrophy to chromosome 4q35. *Am J Hum Genet* 2000;**67**:1309–13.
16 **Cottingham RW**, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet* 1993;**53**:252–63.
17 **Lathrop GM**, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet* 1984;**36**:460–5.
18 **Hentze MW**, Kulozik AE. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell* 1999;**96**:307–10.