Genetic Analysis of Corneal Dystrophies

Salina Siddiqui

MBChB, BSc (Hons), FRCOphth

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Medicine

April 2016

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Acknowledgements

I would firstly like to thank the patients and families who kindly agreed to be involved in this study, without whom this work would not have been achievable. This PhD was generously funded by the Medical Research Council and this support is acknowledged with gratitude.

I am grateful to many people for supporting me during my PhD. I would like to thank my main supervisor Dr Manir Ali. Without his unstinting support and guidance, the data presented in this thesis could not have been produced. I would also like to thank my two other supervisors Dr Carmel Toomes and Professor Chris Inglehearn for their help and expertise. I am certain that the knowledge and skills that I have learned from them will guide me throughout my clinical and academic career.

Many of my colleagues on level 8, past and present, have helped me during this PhD. I am indebted to my friend and colleague James Poulter for his sound advice and good company. I would also like to thank Kamron Khan for setting up the BOSU project, Clare Logan and Layal Abi Farraj for teaching me library preparation for whole exome sequencing, Dave Parry and Ian Carr for their bioinformatic assistance, as well as Claire Smith, Gabrielle Wheway, Jose Ivorra, Evi Panagiotou and Kasia Szymanska.

Dr Aine Rice initiated the corneal dystrophy ethics application and recruitment of the initial patients in the cohort, for which I am grateful. I have been fortunate to work with a fantastic team of research nurses Frances Cassidy, Charmain Tidswell and Alice van Lare. Mr Martin McKibbin leads the Academic Unit of Ophthalmology and has always supported the recruitment of patients for this project.

My thanks also go to all of the project's clinical collaborators Mr Nigel James, Mr James Ball, Mr Andrew Morrell, Mr Andrew Chung and Mr John Buchan.

Dr Nick Thyer, Associate Professor in Audiology who assisted me with the interpretation of audiograms is also acknowledged.

The support of my family and friends has been vital during this PhD, in particular from my parents whose help with childcare has assisted me enormously.

Finally, I owe a huge debt of gratitude to my wonderfully kind husband Niki, for his unwavering support throughout this PhD, his Excel expertise, for looking after our daughters while I was writing and generally for his good humour in the weeks just prior to submitting this thesis. Without his help, this would not have been possible.

Abstract

Corneal disease is a major cause of global blindness accounting for around 2% of severe visual impairment in the UK. Corneal dystrophies are a group of rare, bilateral conditions with a genetic basis.

In conjunction with the British Ophthalmic Surveillance Unit (BOSU), a national incidence for new cases of corneal dystrophy in patients aged below 40 years was identified. 73 cases were reported to BOSU with 27 cases (42%) returned by questionnaire. There was a positive family history in 48% of cases. A minimum UK incidence for new cases per annum of 6.7 cases per 10 000 000 population was calculated.

To investigate the link between Congenital Hereditary Endothelial Dystrophy (CHED), Harboyan syndrome and Fuchs Endothelial Corneal Dystrophy (FECD), a longitudinal observational study was performed. CHED and Harboyan syndrome (CHED with sensorineural hearing loss) are both caused by biallelic mutations in *SLC4A11*. All four of the CHED patients examined had varying degrees of hearing loss at high frequencies, suggesting that CHED and Harboyan syndrome are the same condition at different developmental stages. In addition, two of the four parents of CHED patients examined had guttata, suggesting that the parents are at risk of developing FECD.

FECD is a common, complex corneal endothelial disease. The relative contributions of the *TCF4* SNP rs613872, the intronic *TCF4* CTG18.1 trinucleotide expansion and *LOXHD1* variants in a UK Caucasian FECD cohort ethnically-matched controls were compared. The results of segregation of the CTG18.1 expansion and whole exome sequencing in three local FECD families indicated that the CTG18.1 expansion was causative for the FECD in two of the three families. This indicated that the *TCF4* expansion is a major contributor to the pathogenesis of FECD. Whole exome sequencing in the third family revealed some good gene candidates, which were considered for further screening in the FECD cohort.

Contents

List of Figures

List of Tables

Abbreviations

AGBL1	ATP/GTP-Binding Protein-Like 1
ARMD	Age-related macular degeneration
ARSNHL	Autosomal Recessive Non-syndromic Hearing Loss
BAM	Binary Alignment Map
BCL	Bandage Contact Lens
BDGP	Berkeley Drosophila Genome Project
BF	Bayes Factor
BM	Basement membrane
BOSU	British Ophthalmic Surveillance Unit
BRI	Bradford Royal Infirmary
CADD	Combined Annotation Dependent Depletion
CEPH	Centre d'Étude du Polymorphism Humain
CHED	Congenital Hereditary Endothelial Syndrome
CHST6	*carbohydrate sulphotransferase 6*
CNV	Copy Number Variation
COL8A2	Collagen, type VIII, alpha 2
CRYZL1	Crystallin, zeta(quinone reductase)-like 1
CV	Coefficient of Variation
CVI	Certificate of Visual Impairment
dbSNP	Database SNP
DM	Descemet's Membrane
DMEK	Descemet's membrane endothelial keratoplasty
DSEK	Descemet's stripping endothelial keratoplasty
EBMD	Epithelial Basement Membrane Dystrophy
EDTA	Ethylenediaminetetraacetic acid
ECM	Extracellular Matrix
EMT	Epithelial-to-Mesenchymal Transition
EVS	Exome Variant Server
ExAC	Exome Aggregation Consortium
FECD	Fuchs Endothelial Corneal Dystrophy

FEVR	Familial Exudative Vitreoretinopathy
GATK	Genome Analysis Toolkit
GCD1	Granular Dystrophy Type I
GCD2	Granular Dystrophy Type II
GCP	Good Clinical Practice
GPRIN2	G protein regulated inducer of neurite outgrowth 2
GWAS	Genome Wide Association Study
HM	Hand Movements
HSPG	Heparin Sulphate Proteoglycans
HWE	Hardy Weinberg Equilibrium
IC3D	International Classification of Corneal Dystrophies
IOP	Intraocular Pressure
LASIK	Laser-assisted In-Situ Keratomileusis
LCD1	Lattice Corneal Dystrophy Type I
LIMM	Leeds Institute of Molecular Medicine
Lod	Logarithm of the odds
logMAR	Logarithm of the Minimum Angle of Resolution
LOXHD1	Lipoxygenase Homology Domain 1
LRP5	Low-density lipoprotein receptor-related protein 5
MCD	Macular Corneal Dystrophy
MECD	Meesman Corneal Dystrophy
mRNA	messenger Ribonucleic Acid
mtDNA	mitochondrial DNA
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NIHR	National Institute of Health Research
OD	Ocular Dextra
OMIM®	Online Mendelian Inheritance in Man®
ONH	Optic Nerve Head
OS	Ocular Sinistra
OSBPL1A	oxysterol binding protein-like 1A
OVOL2	Ovo-Like 2
PBS	Phosphate Buffered Saline

PCG	primary congenital glaucoma
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PIC	Patient Identification Centre
PL	Perception of Light
PLAT	polycystin/lipoxygenase/alpha-toxin
PROVEAN	Protein Variation Effect Analyser
PPCD	Posterior Polymorphous Corneal Dystrophy
RBCD	Reis-Bucklers Corneal Dystrophy
R&D	Research and Development
RPM	Revolutions per minute
SAGE	Serial Analysis of Gene Expression
SAM	Sequence/Alignment Map
SIFT	Sorting Intolerant From Tolerant
SLC4A11	solute carrier family 4, sodium borate transporter, member 11
SM	Specular Microscopy
SMRT	Single-molecule real-time
SNV	Single Nucleotide Variant
SJUH	St James's University Hospital
SNP	Single Nucleotide Polymorphism
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
STR	Short tandem repeat
SVI	Severe Visual Impairment
SYT15	Synaptotagmin 15
TAE	Tris-acetate-EDTA
TBCD	Thiel-Behnke corneal dystrophy
TE	Tris-EDTA
TBE	Tris Borate Electrophoresis
TCF4	Transcription Factor 4
TCF8	Transcription Factor 8
TC2N	Tandem C2 domains, nuclear

TGFβI	Transforming Growth Factor β Induced
TP-PCR	Triplet-Primed PCR
UKCRN	United Kingdom Clinical Research Network
URB1	ribosome biogenesis 1 homolog
VA	Visual Acuity
VCF	Variant Call Format
VEP	Variant Effect Predictor
VSX1	Visual System Homeobox 1
WES	Whole Exome Sequencing
WHO	World Health Organisation
XECD	X-Linked Endothelial Dystrophy
ZEB1	Zinc finger E-box-binding Homeobox 1

List of Publications

Publications from this thesis

Siddiqui S, Zenteno JC, Rice A, Chacón-Camacho O, Naylor SG, Rivera-de la Parra D, Spokes DM, James N, Toomes C, Inglehearn CF, Ali M. Congenital hereditary endothelial dystrophy caused by *SLC4A11* mutations progresses to Harboyan syndrome. *Cornea* 2014 **33**(3):247-251.

Publications in preparation from this thesis

Siddiqui S, Ali M, Inglehearn CF, Corneal Dystrophy BOSU Study Team, Foot B, Khan K. A study of corneal dystrophies in conjunction with the British Ophthalmic Surveillance Unit (BOSU).

Siddiqui S, Rice A, Poulter J, Parry D, Varga Z, Ball JL, Morrell AJ, James N, Chung A, Ivorra JL, Buchan JC, Johnson C, Toomes C, Inglehearn CF, Ali M. The inheritance of the *TCF4* intronic trinucleotide repeat in families with Fuchs Endothelial Dystrophy Cohort and in a UK cohort.

Publications related to other work done during the study

Khan K, Al-Maskari A, McKibbin M, Carr IM, Booth A, Mohamed M, **Siddiqui S**, Poulter JA, Hashmi A, Sahi T, Jafri H, Raashid Y, Markham AF, Toomes C, Rice A, Sheridan E, Inglehearn CF and Ali M. Genetic Heterogeneity for Recessively-Inherited Congenital Cataract- Microcornea with Corneal Opacity: New Loci at Chromosomes 10cen and 20p. *Invest Ophthalmol Vis Sci* 2011; **52**(7): 4294-4299

Lechner J, Bae HA, Guduric-Fuchs J, Rice A, Govindarajan G, **Siddiqui S** Farraj A, Yip SP, Yap MK, Das MR, Souzeau E, Coster D, Mills RA, Lindsay R, Phillips T, Mitchell, Ali M, Inglehearn CF, Sundaresan P, Craig JE, Simpson DA, Burdon KP, Willoughby CE. Mutational analysis of MIR184 in sporadic keratoconus and myopia. *Invest Ophthalmol Vis Sci 2013* **54**(8): 5266-72.

1 Introduction

1.1 Gross Anatomy and Function of the Human Eye

The human eye comprises the anterior segment that includes the cornea, iris, trabecular meshwork, aqueous humour and the lens, and the posterior segment, consisting of the vitreous gel, retina and the optic nerve. The retina consists of a neural layer and pigmented retinal pigment epithelium (RPE). The macula is a specialised area at the centre of the posterior retina (Snell and Lemp, 1997). A schematic diagram of the eye in sagittal cross-section is shown in Figure 1.1.

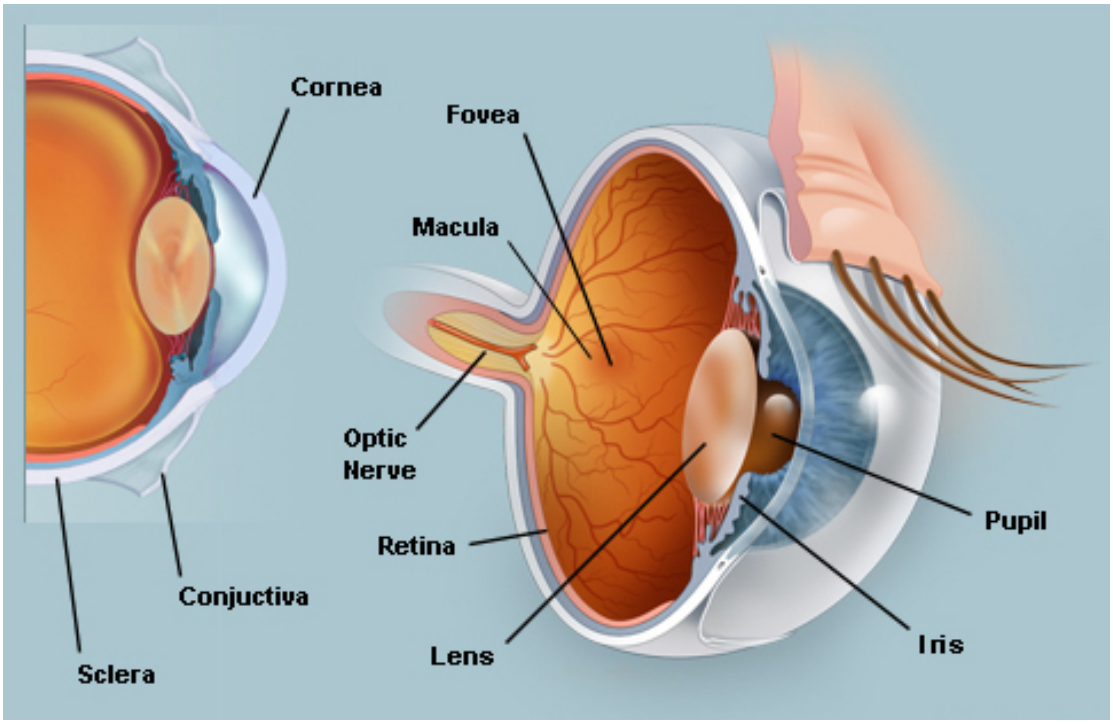


Figure 1.1 Diagram of the human eye in sagittal view, indicating the gross structures. Reproduced with permission from WebMD (http://www.webmd.com/eye-health/picture-of-the-eyes).

The optical components of the eye are the cornea, aqueous humour, lens and vitreous body. The speed with which light travels through these structures is

inversely proportional to it density. Therefore light waves striking the cornea are slowed differentially, as the refractive index of aqueous humour is greater than that of air (Elkington et al., 1999). Light traverses these structures and focuses onto the retina. This process relies on the cornea and lens transparency, enabling the light to focus on a single point on the retina where the light signal is converted into a neural signal by the phototransduction cascade. This neural signal is transferred through the optic nerve eventually forming an image at the brain visual cortex. The biology of the anterior segment is orientated towards achieving clear transmission and sharp focusing of light on the retina (Hejtmancik and Nickerson, 2015). The opaque scleral coat forms the posterior five-sixths of the eyeball. The sclera itself is a relatively avascular structure, however the anterior ciliary arteries form a dense episcleral plexus. The tough fibrous structure protects the intraocular contents from trauma and mechanical displacement. The intervening layer between the retina and the sclera is the choroid, a soft brown, vascular coat lining the inner surface of the sclera. This layer provides nourishment for the retina as well as for absorbing excess light (Snell and Lemp, 1997).

1.2 Embryological Development of the Human Eye

The eye develops from several embryonic layers. The ectoderm gives rise to the lens and the corneal epithelium. The neuroectoderm forms the pigmented epithelium and the neural retina. The neural crest cells develop into the corneal stroma, the ciliary and iris muscles and the vascular choroid layer together with the fibrous sclera. The mesoderm contributes to the cornea and forms the angioblasts of the choroid layer. The first evidence of eye formation is the formation of the optic sulcus at 22 days of human development. At day 24, the cranial neuropore closes, by which time the optic stalk is evident (Larsen et al., 2009). By the sixth week, the rudimentary eye including the optic cup and partially encapsulated lens vesicle are visible. The bilayered optic cups, partially encapsulating the lens vesicles have formed (Figure 1.2)

(Ali and Sowden, 2011). As the subject of this thesis is the cornea, its embryogenesis and development are covered in greater detail below.

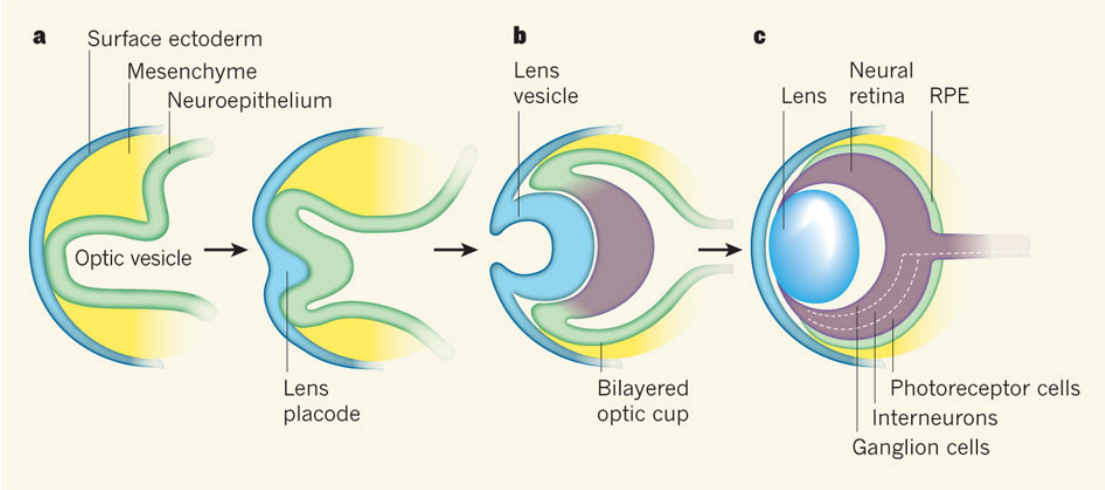


Figure 1.2 The embryonic development of the human eye a. At early stages of eye development, the surface ectoderm thickens and invaginates together with the underlying neuroepithelium of the optic vesicle. **b.** The inner layer of the bilayered optic cup gives rise to neural retina and the outer layer gives rise to the retinal pigmented epithelium (RPE) **c.** The mature neural retina. Reproduced with permission (Ali and Sowden, 2011).

1.2.1 Corneal Embryogenesis

The formation of the cornea is induced by the lens and the optic cup, and is the last series of major inductive events during eye development at around 5 to 6 weeks of human gestation (18mm stage), when the surface ectoderm interacts with the lens vesicle (Figure 1.3). When completely separated, the space between them is filled with perinuclear mesenchyme cells from the neural crest. The mesenchyme condenses and forms several layers separated by extracellular matrix. The mesenchymal cells closer to the lens become the endothelium and the surface ectoderm on the anterior surface, become the corneal epithelium (Larsen et al., 2009, Zavala et al., 2013). Therefore, the corneal layers are derived from differing embryonic origins.

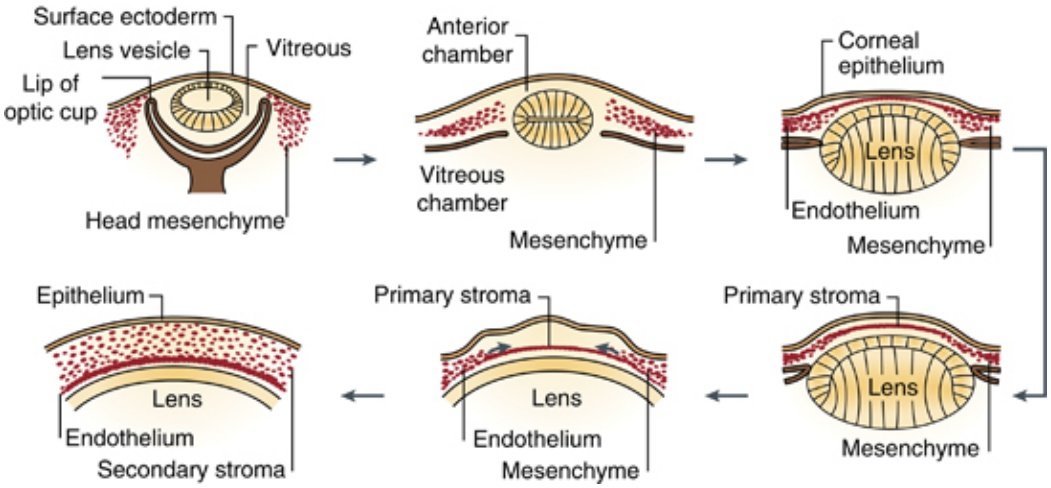


Figure 1.3 Formation of the cornea. The cornea begins to develop when the surface ectoderm closes after the formation of the lens vesicle and its detachment from the surface ectoderm. Mesenchymal cells (neural crest cells) invade the cornea and form the corneal stroma after condensation. Reproduced with permission (Zavala et al., 2013).

1.2.2 Development of the Corneal Endothelium and Descemet's Membrane

The neural crest cells undergo epithelial-to-mesenchymal transition (EMT) and form a cell monolayer that occupies the posterior surface of the cornea. The presumptive corneal endothelium begins as a loosely arranged monolayer at about 8 weeks gestation. The zonula occludens (tight junctions in between the endothelial cells) are present by week 17, although the endothelial pump function (Section 1.3.6) is not fully established. The early endothelium is evident during the fourth month of human gestation. It secretes the basement membrane, Descemet's membrane (DM), as a multilaminated layer between the endothelial cell layer and the posterior stroma. It is a cell-free matrix consisting predominantly of collagens (Waring et al., 1982) (Lwigale, 2015).

1.3 Detailed Anatomy of the Adult Human Cornea

The cornea, the major refracting structure of the human eye, occupies one third of the external eyeball. It consists of six layers: the epithelium, epithelial basement membrane, Bowman's layer, the stroma, DM and the endothelium at the posterior surface of the cornea. These are numbered sequentially in Figure 1.4.

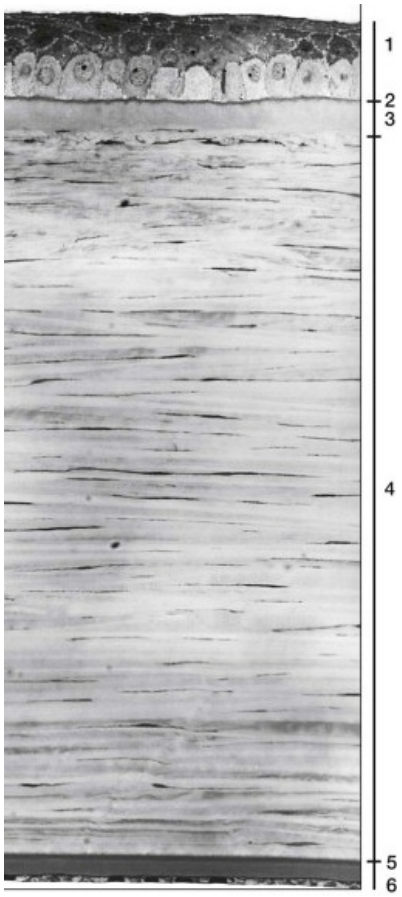


Figure 1.4 Histology of the cornea. The layers are numbered sequentially; the epithelium (*1*), epithelial basement membrane (*2*), Bowman's layer (*3*), corneal stroma (*4*), DM (*5*), and endothelium (*6*). Reproduced with permission (Krachmer et al., 2011).

1.3.1 The corneal epithelium

The epithelium comprises four to six layers of non-keratinised stratified squamous epithelial cells, and is 50µm in thickness. The superficial two to

three layers consist of flat, polygonal cells which have microvillae and micropliae on the apical surface. The cell periphery exhibits tight junctions which prevent the entry of microorganisms into the cornea. Basal epithelial cells are the posterior-most layer of the corneal epithelium and are responsible for laying down the basement membrane. Perilimbal basal cells differentiate and migrate anteriorly to regenerate the cornea. They utilise hemidesmosomes to adhere to the underlying basement membrane (Eghrari et al., 2015b). There is little variation in cell density and morphology of basal epithelial cells amongst individuals (Harrison et al., 2003).

1.3.2 Basement Membrane

The basal cells of the corneal epithelium are anchored to the basement membrane (BM) which is 40-60nm thick (Krachmer et al., 2011). The BM is composed of four primary components – collagens (predominantly collagen type IV), laminins, heparin sulphate proteoglycans (HSPG) and nidogens. Laminins are the most abundant non-collagenous proteins in the BM, and interact with collagen networks via nidogens. They have the unique ability to self-assemble into sheet-like structures. The most prevalent HSPG is perlecan, a multidomain protein which mediates the migration, proliferation and differentiation of a variety of cells by mediating cell signalling events (Torricelli et al., 2013).

1.3.3 Bowman's Membrane

Bowman's layer is an acellular, non-regenerating layer that is posterior to the epithelial BM. It is approximately 8-12µm in depth and decreases in thickness with age. It consists of collagen fibrils, which are two-thirds the thickness of those of the stroma, and which merge with those of the anterior stroma (Eghrari et al., 2015b).

1.3.4 Stroma or Substantia Propria

This layer forms 90% of the corneal thickness (Figure 1.5). It is transparent, fibrous and compact. Stromal fibrils consist mainly of type I collagen, with smaller amounts of types III, V and VI (Snell and Lemp, 1997). The stroma comprises over 200 sheets of lamellae each 1-2μm thick of collagen fibrils approximately 36nm in diameter. This arrangement is highly ordered such that they lie in parallel, an arrangement essential for maintaining corneal clarity. Keratocytes are interspersed between the lamellae forming an interlinking network. The main glycosaminoglycan present is keratin sulphate. The stroma has a natural tendency to absorb water and swell due to the hydrophilic nature of the proteoglycan matrix surrounding the collagen fibrils. The mechanism by which the stroma of the cornea remains relatively dehydrated is deturgescence (Freegard, 1997).

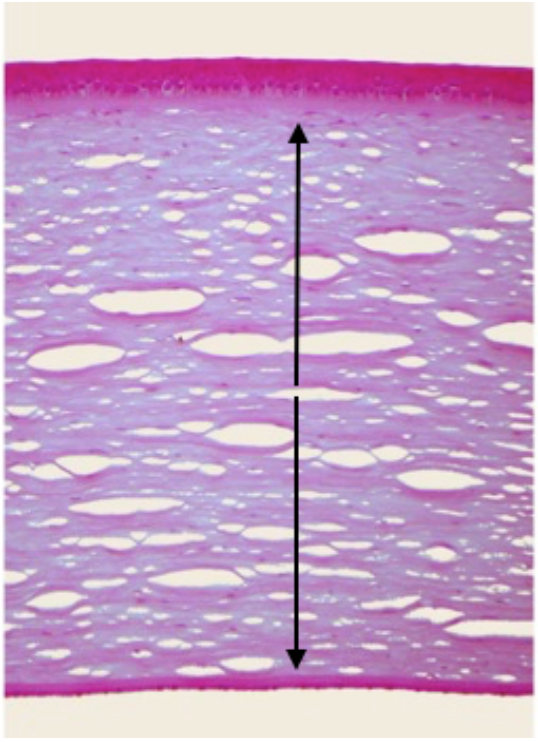


Figure 1.5 Corneal histology image of the human cornea. The stroma occupies 90% of the corneal thickness, as indicated by the two black arrows (courtesy of Dr Hardeep S Mudhar, Consultant Ophthalmic Histopathologist, Royal Hallamshire Hospital, Sheffield, UK).

1.3.5 Descemet's Membrane

DM is a basement membrane of the corneal endothelium and measures 3µm in children and 8-10µm in adults. Transmission electron microscopy reveals the anterior banded layer, approximately 3µm thick, which consists of a latticework of collagen fibrils with periodic banding at 110nm intervals. On tangential section, the layer reveals a hexagonal formation of collagen fibrils. After birth the endothelial cells secrete the relatively homogeneous posterior non-banded layer which has a fine granular appearance, and which thickens with age (Johnson et al., 1982).

DM contains collagen types IV and VIII fibrils, type VIII being specific to DM. Similar to the stroma the ECM protein fibronectin, laminin, keratin sulphate, herparin sulphate and dermatin sulphate are also present (Eghrari et al., 2015b). Fibronectin may play a role in the adhesion of the endothelial cells to the DM (Waring et al., 1982). The DM is shown by the black arrow in 1.6A and is also labelled in 1.6B.

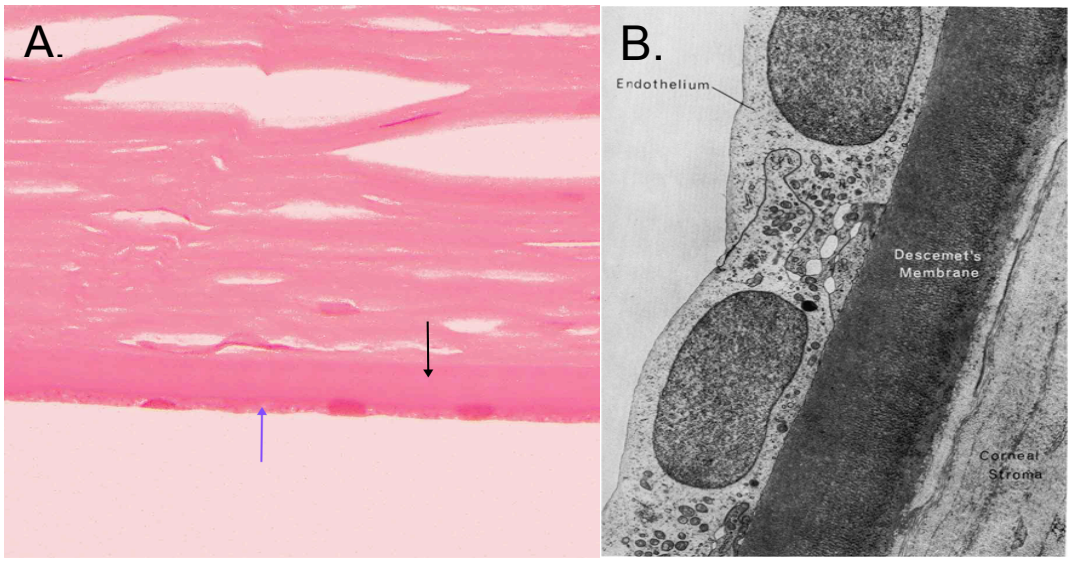


Figure 1.6 The posterior aspect of the human cornea. A. Histopathology image of the showing the endothelium and DM. The black arrow indicates DM. The endothelial cell layer is indicated by the purple arrow (courtesy of Dr Hardeep S Mudhar, Consultant Ophthalmic Histopathologist, Royal Hallamshire Hospital, Sheffield, UK). B. Electron micrograph of corneal endothelium underlying DM. Reproduced with permission (Zavala et al., 2013).

1.3.6 Corneal Endothelium

The human corneal endothelial surface (Figure 1.6) is 130mm^2 and comprises a monolayer of polygonal corneal endothelial cells, which cover the posterior surface of DM in a mosaic pattern (Figure 1.7). These cells are 5μm in thickness, 20μm in width and are polygonal, mostly hexagonal, in shape. The cell density in a 3 to 6-year-old child is 3500–4000 cells per mm^2 but this value decreases with age. The normal endothelial cell count in a 30 year old ranges from between 2700 to 2900 per mm^2 and that in an adult over 75 years old ranges from 2400 to 2600 per mm^2 (McCarey et al., 2008). The corneal endothelium maintains corneal clarity by keeping the stroma in a state of relative dehydration, as well as providing a barrier and pump function. The active transport of bicarbonate ions into the aqueous humour is thought to be a major role of the endothelial pump function, although the transport of sodium and potassium ions also plays a role (Waring et al., 1982). Endothelial cells contain a large nucleus and abundant cytoplasmic organelles including mitochondria, endoplasmic reticulum, free ribosomes, and Golgi apparatus, suggesting that they are highly metabolically active (Krachmer et al., 2011). When the endothelial cell counts drops to 400-700 cells per mm^2 or less, corneal decompensation (corneal oedema resulting from failure of the endothelium to maintain deturgescence) occurs, which suggests that there is a substantial reserve (Edelhauser, 2006). When the endothelial cell count is low, the loss of zonula occludens allows more fluid to enter the stroma, thus disrupting the parallel arrangement of the stromal collagen fibrils and compromising corneal clarity.

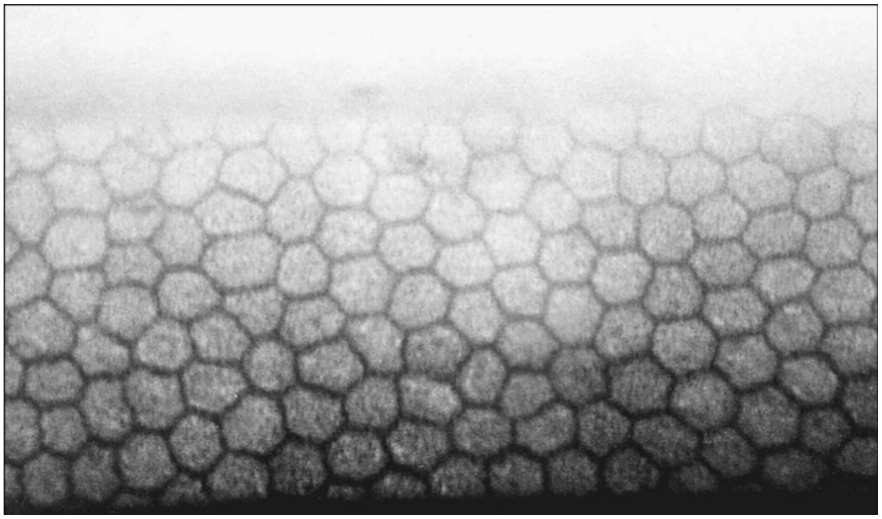


Figure 1.7 Normal corneal endothelium as photographed by specular microscopy. The regular array of hexagonal cells, all having nearly the same area, is seen. Reproduced with permission (Krachmer et al., 2011).

1.3.7 Regenerative Capacity of the Corneal Endothelium

Individuals are born with differing numbers of endothelial cells (Yuen et al., 2005). Traditionally, it was thought that endothelial cells do not divide and are arrested in the G1 phase of the cell cycle (Bourne, 2003). Therefore, as the endothelial cells die with age, the cells that remain have a limited regenerative capacity (Waring et al., 1982). However, recently this theory has been challenged by a study, which suggests that the corneal periphery contains a reservoir of stem-like cells that replace damaged or dead endothelium (He et al., 2012). The authors evaluated the microanatomy of 88 whole endothelia. In 61% of the corneas, they noted that the cells located at the extreme periphery (the peripheral 200µm of the endothelium) were organized in small clusters, estimated to occupy one third of the circumference of the cornea. Additionally, in 68% of the corneas, peripheral endothelium formed radial rows, which were variable in length but occupied around 86% of the corneal circumference. On staining, these cells were less differentiated but expressed stem cell specific markers.

1.4 Epidemiology of Eye Disease

1.4.1 Global Burden of Eye Disease

There are approximately 161 million people worldwide with low vision, 37 million of whom are severely sight-impaired. This figure excludes those with uncorrected refractive error (Resnikoff et al., 2004). The main causes of global blindness are cataract, glaucoma, corneal scarring, age-related macular degeneration (ARMD) and diabetic retinopathy. All of these are considered avoidable except for ARMD. The World Health Organisation (WHO) Prevention of Blindness programme and the International Agency for the Prevention of Blindness set up the VISION 2020 - Right to Sight initiative in 1999 to eliminate 80% of avoidable blindness by the year 2020. It sought to do this by focusing on the diseases, which are the main causes of blindness and for which proven cost-effective interventions are available. The most recent global action plan for the period 2014-2019 aims to reduce visual impairment as a global public health problem and secure access to rehabilitation for visually impaired services (http://www.iapb.org/vision-2020). It has been calculated using economic and epidemiological modelling that without global initiatives the number of blind individuals would increase to 76 million in 2020. The cost impact of blindness can be assessed in “blind person-years” defined as one year of blindness for one individual. A successful VISION 2020 initiative would avoid 429 million blind person-years (Frick and Foster, 2003). The impact of this initiative was assessed in 2005. The estimates of global blindness were 15 million blind person-years less than projected and indicated that VISION 2020 was reducing global blindness (Foster and Resnikoff, 2005, Foster et al., 2008).

Visual acuity is the ability to resolve detail, for instance to be able to detect a gap between two objects in space. It is a commonly used proxy measure for the degree of visual acuity as it is easy to measure and is understood by the lay public (https://www.rcophth.ac.uk/professional-resources/revalidation/clinical-sub-specialties/cataract/visual-acuity/).

Snellen acuity (Figure 1.8A) has been used to measure visual acuity since 1862, however this has more recently been superseded by its assessment using a logarithmic method, LogMAR visual acuity (Figure 1.8B) (https://www.rcophth.ac.uk/wp-content/uploads/2015/11/LogMAR-vs-Snellen.pdf).

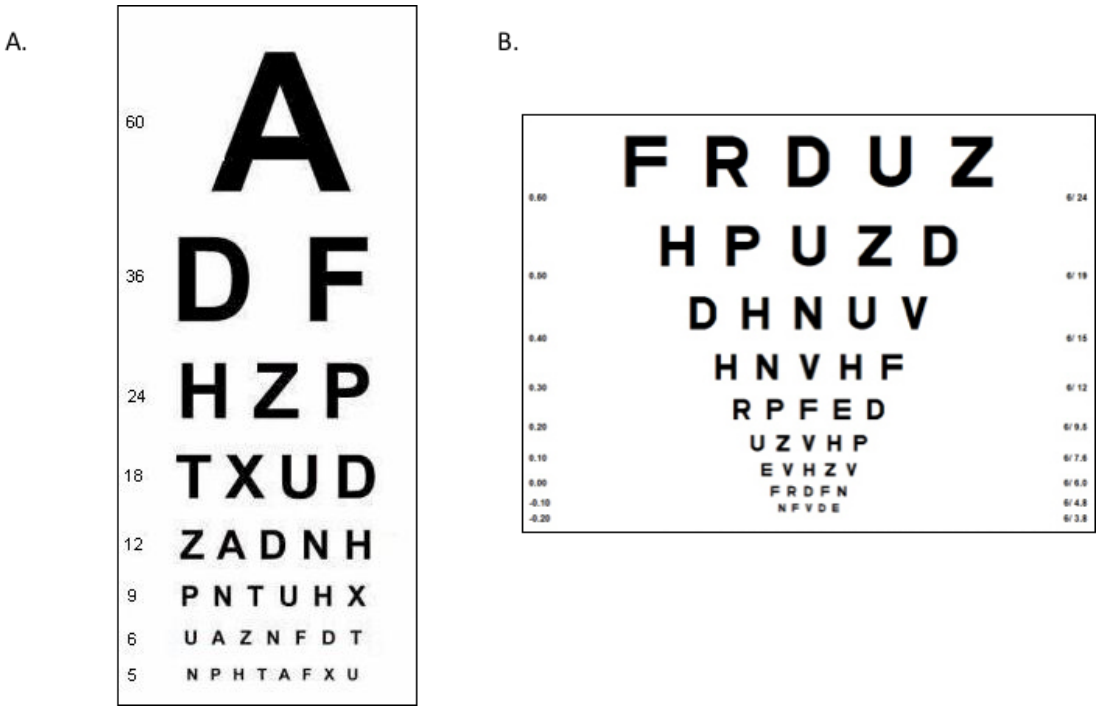


Figure 1.8 An example of A. a Snellen chart B. LogMAR for testing visual acuity.

The International Statistical Classification of Diseases and Related Health Problems 10th Revision 2015 (http://apps.who.int/classifications/icd10/browse/2015/en#!/H53) currently classifies visual impairment as outlined in Table 1.1.

Category	Presenting distance visual acuity	
	Worse than:	Equal to or better than:
0 Mild or no visual impairment		6/18 3/10 (0.3) 20/70
1 Moderate visual impairment	6/18 3/10 (0.3) 20/70	6/60 1/10 (0.1) 20/200
2 Severe visual impairment	6/60 1/10 (0.1) 20/200	3/60 1/20 (0.05) 20/400
3 Blindness	3/60 1/20 (0.05) 20/400	1/60* 1/50 (0.02) 5/300 (20/1200)
4 Blindness	1/60* 1/50 (0.02) 5/300 (20/1200)	Light perception
5 Blindness	No light perception	
9	Undetermined or unspecified	
	* or counts fingers (CF) at 1 metre.	

Table 1.1 Categories of visual impairment (ICD-10 Version 2016) (http://apps.who.int/classifications/icd10/browse/2016/en#/H53). Visual acuity is shown in both Snellen and LogMAR forms.

1.4.2 Childhood Visual Impairment Epidemiology and Case Identification

Childhood blindness is the second largest category of blind-person years. Globally about 70 million blind person years are caused by childhood blindness (Yorston, 1999). In the UK, an epidemiological study of 439 children with severe visual impairment (SVI) was carried out. The results indicated that SVI had an incidence of 4 per 100,000 and occurred more frequently in conjunction with co-existing non-ophthalmic impairments. The authors found that children at the most risk were those of low-birth weight and from ethnic minority groups. Cerebral visual impairment was the most common cause of SVI in children. SVI caused by sclerocornea and corneal opacity accounted for 2% of patients (Rahi and Cable, 2003). Cases of visual impairment may be identified by a parent, carer or teacher, through perinatal or pre-school screening. Cases may also be identified through assessment following a family history of sight loss, or a medical condition with an ophthalmic manifestation (https://www.rcophth.ac.uk/professional-resources/information-from-the-paediatric-sub-committee-for-healthcare-professionals/).

1.4.3 Visual Impairment in Adults

Researchers examining the causes of certifiable visual impairment (CVI) in working age adults recently found for the first time in five decades the leading cause of certifiable blindness was not diabetic retinopathy/maculopathy but hereditary retinal disorders (20.2%). This was followed by diabetic retinopathy/maculopathy (14.4%), optic atrophy (14.1%) and corneal disease which accounted for 45/1756 (2.6%) (Liew et al., 2014). Overall in adults macular degeneration was the most common cause of CVI with corneal disorders accounting for 100/7437 (2.1%) of the total number of CVI cases (Bunce et al., 2010). A form of corneal disease is the focus of this thesis.

1.5 Overview of Corneal Dystrophies

Corneal dystrophies are defined as a group of bilateral, genetically determined, non-inflammatory diseases that are, in the majority of cases, limited to the cornea (Klintworth, 2009). There are however exceptions to every part of this definition (Weiss et al., 2015). They are both clinically and genetically heterogeneous. Many of these conditions are important causes of congenital corneal opacity (Nischal, 2015). They can occur as isolated conditions or with systemic manifestations.

The International Corneal Dystrophy (IC3D) committee was formed in 2005 with the purpose of reducing confusion regarding corneal dystrophy nomenclature, to critically evaluate the literature and remove outdated information. The committee held its first meeting in October 2005 and proposed a new classification of corneal dystrophies in 2008 (Weiss et al., 2008) which consisted of five categories: Epithelial and Subepithelial Dystrophies, Bowman Layer Dystrophies, Stromal Dystrophies and Descemet Membrane and Endothelial Dystrophies. Additionally, a series of descriptive evidential categories indicating the level of evidence supporting the existence of each given dystrophy was outlined (Table 1.2). It was postulated that with increased knowledge about a dystrophy, its category should progress over

time from 4 to 1, and that dystrophies that remain in category 4 should be eventually removed from the nomenclature.

Category 1	A well-defined corneal dystrophy in which the gene has been mapped and identified and specific mutations are known.
Category 2	A well-defined corneal dystrophy that has been mapped to 1 or more specific chromosomal loci, but the gene(s) remains to be identified.
Category 3	A well-defined corneal dystrophy in which the disorder has not yet been mapped to a chromosomal locus.
Category 4	This category is reserved for a suspected new, or previously documented, corneal dystrophy, although the evidence for it, being a distinct entity, is not yet convincing.

Table 1.2 Levels of evidence supporting the existence of corneal dystrophies, as described by the IC3D Committee (Weiss et al., 2008).

The traditional anatomic classification of corneal dystrophies described (Weiss et al., 2008) was thought to have limitations and in 2015 corneal dystrophies were reclassified into four groups; Epithelial and subepithelial, Epithelial-stromal *TGFβI*, Stromal and Endothelial dystrophies (Weiss et al., 2015). These four classes of corneal dystrophy are summarised below, with examples of each given.

Class of Corneal Dystrophy	Dystrophy	MIM	Category of Evidence	Mode of Inheritance	Locus	Gene
Epithelial and Subepithelial Dystrophies	Epithelial Basement membrane dystrophy (EBMD) majority degenerative	MIM #121820	rarely C1	Unknown	5q31	*TGFβI*
	Epithelial recurrent erosion dystrophies (EREDs)	MIM #122400				
	Franceschetti corneal dystrophy (FRCD)		C3	Autosomal Dominant	Unknown	Unknown
	Dystrophia Smolandiensis (DS)		C3	Autosomal Dominant	Unknown	Unknown
	Dystrophia Helsinglandica (DH)		C3	Autosomal Dominant	Unknown	Unknown
	Subepithelial mucinous corneal dystrophy (SMCD)	MIM #612867	C4	Autosomal Dominant (but X-linked not excluded)	Unknown	Unknown
	Meesmann corneal dystrophy (MECD)	MIM #122100	C1	Autosomal Dominant	12q13 (KRT3)	*KRT3*
	- Stocker-Holt variant		C1	Autosomal Dominant	17q12 (KRT12)	*KRT12*
	Lisch epithelial corneal dystrophy (LECD)	MIM #300778	C2	X-chromosomal Dominant	Xp22.3	Unknown
	Gelatinous drop-like corneal dystrophy (GDLD)	MIM #204870	C1	Autosomal Recessive	1p32	*TACSTD2*
Epithelial-stromal *TGFβI*	Reis–Bücklers corneal dystrophy (RBCD)	MIM #608470	C1	Autosomal Dominant	5q31	*TGFβI*
	Thiel-Behnke corneal dystrophy (TBCD)	MIM #602082	C1	Autosomal Dominant	5q31	*TGFβI*
	Lattice corneal dystrophy, type 1 (LCD1), variants (III, IIIA, I/IIIA,IV) of lattice corneal dystrophy	MIM #122200	C1	Autosomal Dominant	5q31	*TGFβI*
	Granular corneal dystrophy, type 1 (GCD1)	MIM #121900	C1	Autosomal Dominant	5q31	*TGFβI*
	Granular corneal dystrophy, type 2 (GCD2)	MIM #607541	C1	Autosomal Dominant	5q31	*TGFβI*
Stromal Dystrophies	Macular corneal dystrophy (MCD)	MIM #217800	C1	Autosomal Recessive	16q22	*CHST6*
	Schnyder corneal dystrophy (SCD)	MIM #21800	C1	Autosomal Dominant	1p36	*UBIAD1*
	Congenital stromal corneal dystrophy (CSCD)	MIM #610048	C1	Autosomal Dominant	12q21.33	*DCN*
	Fleck corneal dystrophy (FCD)	MIM #121850	C1	Autosomal Dominant	2q34	*PIKFYVE*
	Posterior amorphous corneal dystrophy (PACD)	MIM #612868	C1	Autosomal Dominant	12q21.33	*KERA, LUM, DCN, EPYC*
	Central cloudy dystrophy of François (CCDF)	MIM #217600	C1	Unknown	Unknown	Unknown
	Pre-Descemet corneal dystrophy (PDCD)	none				
	- isolated PDCD		C4	Unknown	Unknown	Unknown
	-associated with X-linked ichthyosis		C1	Autosomal Dominant	Xp22.31	*STS*
Endothelial Dystrophies	Fuchs endothelial corneal dystrophy (FECD)					
	- Early-onset FECD	MIM #136800 (FECD1)	C1	Autosomal Dominant	1p34.3–p32 (FECD1)	*COL8A2*
	- Late-onset FECD	MIM #610158 (FECD2) MIM #613267 (FECD3) MIM #613268 (FECD4) MIM #613269 (FECD5) MIM #613270 (FECD6) MIM #613271 (FECD7) MIM #615523 (FECD8)	C2 in patients with defined genetic loci C3 in patients without known inheritance	Cases without known inheritance are most common. Some cases with Autosomal Dominant inheritance. Complex and heterogeneous, demonstrating variable expressivity and incomplete penetrance	13pter-q12.13 (FECD2) 18q21.2-q21.3 (FECD3) 20p13-p12 (FECD4) 5q33.1-q35.2 (FECD5) 10p11.2 (FECD6) 9p24.1-p22.1 (FECD7) 15q25 (FECD8)	*TCF4, LOXHD1* *TCF8* *AGBL1*
	Posterior polymorphous corneal dystrophy (PPCD)	MIM #122000 (PPCD1) MIM #609140 (PPCD2) MIM #609141 (PPCD3)	C1 - PPCD 2 and PPCD 3 C2 - PPCD 1	Autosomal Dominant Isolated cases with no known inheritance pattern	PPCD 1: 20p11.2-q11.2 PPCD 2: 1p34.3-p32.3 PPCD 3: 10p11.22	PPCD1: unknown PPCD2: *COL8A2* PPCD 3: *ZEB1*
	Congenital hereditary endothelial dystrophy (CHED)	MIM #217700	C1	Autosomal Recessive	20p13	*SLC4A11*
	X-linked endothelial corneal dystrophy (XECD)	none	C2	X-chromosomal Dominant	Xq25	Unknown

Table 1.3 Table outlining each corneal dystrophy, its category of supporting evidence (Table 1.2), mode of inheritance, locus and gene if known.

1.5.1 Epithelial and Subepithelial Dystrophies

Epithelial Basement Membrane Dystrophy (EBMD) (MIM#121820), the most common anterior corneal dystrophy, was first described by Cogan (Cogan et al., 1964) and is characterised by the presence of subepithelial dots, map-like changes (so-called as they are shaped like maps) and blebs Figure 1.9. The patient may be asymptomatic or may present with recurrent corneal erosions (Schorderet, 2015). Mutations in the Transforming Growth Factor β Induced gene (*TGFβI*) (Section 1.5.2) have been reported in two families with autosomal dominant inheritance as well as sporadic cases (Boutboul et al., 2006).

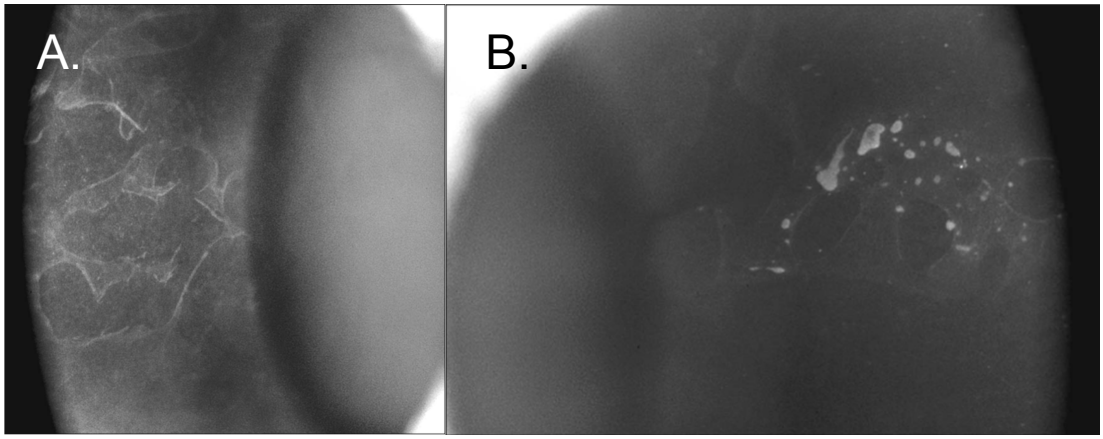


Figure 1.9 Epithelial Basement Membrane Dystrophy A. Map-like changes in EBMD. B. Both microcystic changes and map-like changes together. Reproduced with permission (Krachmer et al., 2011).

Meesman Corneal Dystrophy (MECD) MCD, (MIM#122100), another condition affecting the corneal epithelium, is characterised by distinct tiny bubble-like, round to-oval punctate opacities which present during infancy (Klintworth, 2009). Dominant mutations in both Keratin 3 (*KRT3)* (MIM*148043) on chromosome 12 and Keratin 12 (*KRT12)* (MIM*601687) on chromosome 17 have been implicated (Irvine et al., 1997). Lisch Epithelial Corneal Dystrophy (LECD) (MIM%300778) is characterised by feather shaped opacities and microcysts in the corneal epithelium (Lisch et al., 1992). Painless blurred vision may begin after 60 years (Klintworth, 2009). Linkage

analysis in a large German multi-generational family mapped the disease to a locus on chromosome Xp22.3 (Lisch et al., 2000), although the LOD score was below 3 (2.93) therefore the linkage to this region was tentative.

1.5.2 Epithelial-stromal Transforming Growth Factor beta Induced (*TGFβI*) Dystrophies

The corneal dystrophies described below are all caused by heterozygous mutations in the *TGFβI* gene on chromosome 5 and are inherited as dominant conditions (Lisch et al., 2000, Munier et al., 2002).

Reis-Bucklers Corneal Dystrophy (RBCD) (MIM#608470) presents with bilateral symmetrical irregular ring-shaped subepithelial opacities, which become evident in the first few years of life and are associated with pain, redness and photophobia (Waring et al., 1978a). Thiel-Behnke Corneal Dystrophy (TBCD) (MIM#602082) is similar to RBCD in that it is characterised by subepithelial corneal opacities. However clinically the lesions can be differentiated by their honeycomb pattern and a clear area near the corneo-scleral limbus (Krachmer et al., 2011). Although most causative mutations are in the *TGFβI* gene, the disease has been also been mapped to chromosome 10q23-q24 (Yee et al., 1997). However, when the original linkage data was reviewed (Jonsson et al., 2015) a synonymous variant in *COL17A1* c.3156C>T was identified. This variant introduced a cryptic donor site and resulting in aberrant pre-mRNA splicing, and therefore was likely to be pathogenic.

Lattice Corneal Dystrophy Type I (LCD1) (MIM#122200) presents during the first decade. Clinically a network of interdigitating branching filaments is seen within the corneal stroma (Klintworth, 2009). Accumulations of amyloid material in the stroma distort the architecture of the corneal lamellae (Figure 1.10A and B). Granular Dystrophy Type I (GCD1) (121900) is a disease of the corneal stroma. Characteristic bilateral breadcrumb-like corneal deposits in the anterior stroma give rise to glare, reduced vision and photophobia

(Krachmer et al., 2011) (Figure 1.10C). Granular Dystrophy Type II (GCD2) (MIM#607541) shares some features with Lattice Dystrophy. Clinically and histopathologically GCD2 it can be considered to have features of both GCD1 and LCD1 (Schorderet, 2015) (Figure 1.10D).

TGFβI encodes the transforming growth factor β protein (TGFβI), also known as keratoepithelin, a 68-kDa protein composed of 683 amino acid residues. It contains a secretory signal peptide sequence, a cysteine-rich EMI domain, four homologous fasciclin 1 (FAS1) domains which each contain 140 amino acid residues at the N-terminus, and an arginine-glycine-aspartate (RGD) motif, which binds to integrin at the C-terminus. It is known to be expressed in other organs including the heart, liver, pancreas, bone, tendon, endometrium and kidney (Han et al., 2016). TGFβpI is a member of the TGFβ family of cytokines, which regulate diverse cellular processes including proliferation, apoptosis, differentiation and ECM homeostasis (Kim et al., 2015a).

1.5.3 Stromal Dystrophies

These dystrophies are those that affect the corneal stroma but are not caused by dominant mutations in the *TGFβI* gene. One such condition, Macular Corneal Dystrophy (MCD) (MIM#217800), is an autosomal recessive disease. Faint stromal white opacities are seen in the first decade. The opacities progress over time and a grainy haze develops (Yanoff and Duker 3rd, 2009). MCD is not entirely stromal, as over time the deposits extend to DM and the endothelium, giving rise to endothelial decompensation and corneal oedema. The disease is caused by recessive mutations in Carbohydrate Sulfotransferase 6 (*CHST6)* on chromosome 16. The encoded protein is an enzyme which catalyses the transfer of a sulphate group to the N-Acetylglucosamine amino acid residue of keratan (Akama et al., 2000). Keratan sulphate is a major component of the corneal stroma and helps to maintain corneal clarity (Section 1.3.4).

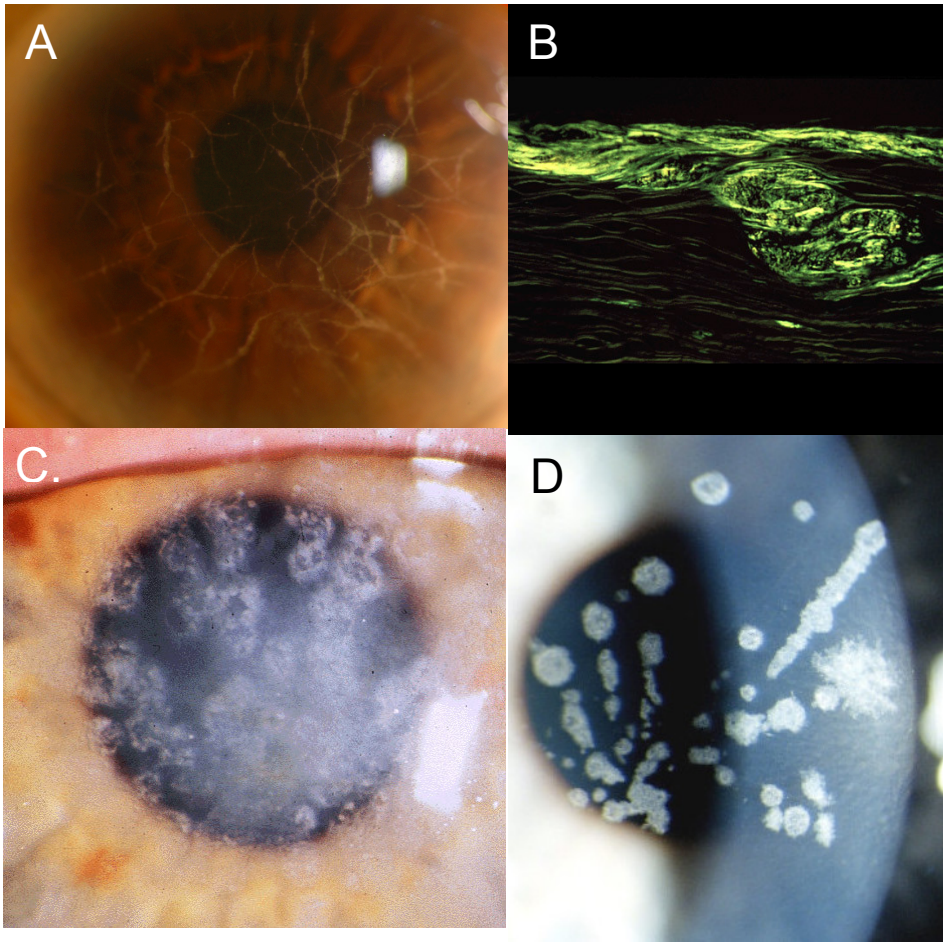


Figure 1.10 Corneal photographs of two *TGFβI* stromal dystrophies. A. Lattice Corneal Dystrophy Type I (courtesy of Mr James Ball, Consultant Ophthalmic Surgeon, St James's University Hospital, Leeds). B. Lattice corneal dystrophy type I variant. The amyloid within the corneal stroma viewed under ultraviolet light after staining the fluorescent dye Thioflavin T. C. Granular corneal dystrophy type I. Numerous irregular shaped discrete crumb-like corneal opacities. D. Granular corneal dystrophy type II. Variable sized crumb-like opacities in the corneal stroma that have become fused in areas giving rise to elongated and stellate shapes. Figure 1.10 B-D reproduced with permission (Klintworth, 2009).

1.5.4 Endothelial Dystrophies

Several dystrophies that primarily affect the corneal endothelium have been described. These are Congenital Hereditary Endothelial Dystrophy (CHED), Posterior Polymorphous Corneal Endothelial Dystrophy (PPCD), X-Linked Endothelial Dystrophy (XECD) and Fuchs Endothelial Corneal Dystrophy (FECD). These conditions are all outlined in more detail in separate sections below.

1.6 Congenital Hereditary Endothelial Dystrophy

1.6.1 Clinical Features

Congenital Hereditary Endothelial Dystrophy (CHED) (MIM#217700) was first described by Maumenee (Maumenee, 1960) who recognised that the primary pathology originated in the corneal endothelium. The hallmark of CHED is corneal oedema and opacification presenting at birth or shortly after birth (Patel and Parker, 2015). Kirkness et al reported a large case series of 23 patients which were divided into recessive and dominant, providing a detailed description of clinical features, management and pathology (Kirkness et al., 1987). They noted that progression of corneal signs, including development of corneal stromal scarring and deposition of plaques, was observed in four of the recessive and three of the dominant cases. Additionally, they examined 26 corneal buttons, noting thickening of DM in the majority of cases, and described more marked thickening in the recessive group. The endothelium was discernable in most cases but loss of cells and degeneration was a common finding.

The original author's subsequent report (Judisch and Maumenee, 1978) attempted to differentiate the clinical appearance of CHED according to its recessive or dominant inheritance. They described what they thought was an autosomal dominant CHED patient as having clear corneas at birth with slowly progressive corneal opacification and felt that "infantile" might be a more descriptive name for the dominant variant. It was therefore previously thought that CHED may be inherited in a dominant (CHED1, MIM#121700) or a recessive manner (CHED2, MIM#217700) (Callaghan et al., 1999). CHED2 was thought to be more common and characterised by a more severe phenotype than CHED1. However, CHED1 is insufficiently distinct to be considered a unique dystrophy (Nischal, 2015). In the most recent corneal dystrophy classification (Weiss et al., 2015), CHED1 was reclassified following a review of the original clinical and pathological descriptions, concluding that all the families could have PPCD (Section 1.7). Consequently, CHED1 has been eliminated from the corneal dystrophy classification and the

previously described CHED2 is now simply referred to as CHED, and will be throughout this thesis.

1.6.2 Association with Glaucoma

The normal intraocular pressure (IOP) is 12-22mmHg. The upper limit of 22mmHg is the traditional cut-off for European-derived populations, however some Asian populations have a lower mean IOP and a pressure above 19mmHg might be considered abnormal in such individuals (Yanoff and Duker 3rd, 2009). The association of raised intraocular pressure (IOP) with CHED has been documented (Pedersen et al., 1989, Keates and Cvintal, 1965), and has the potential to pose a diagnostic and therapeutic dilemma as it may mimic primary congenital glaucoma (PCG). Additionally, stromal oedema can give rise to a false artificially raised IOP. PCG may present with additional signs of increased corneal diameter, Haab striae (horizontal striae seen on the corneal endothelium) and enlarged axial length (Ko et al., 2015). Kirkness et al revealed that 5/23 of their case series of CHED patients had undergone a surgical glaucoma procedure. It has also been suggested in retrospective analysis of 10 CHED patients that CHED and congenital glaucoma may co-exist in some patients (Ramamurthy et al., 2007).

1.6.3 The genetic basis of CHED

CHED was originally mapped to chromosome 20p13 (Toma et al., 1995). Subsequently recessive mutations in solute carrier family 4, sodium borate transporter, member 11 (*SLC4A11*) (MIM*610206) were found to be the cause of CHED (Vithana et al., 2006). The authors described 10 families of Myanmar, Pakistani and Indian origin. Missense mutations accounted for the condition in 8 out of 10 families. In one family, of Indian origin, there was a frameshift mutation with 11 amino acid residues followed by a premature stop codon, indicating that CHED was cause by a loss-of-function mutation. Mutation screening of the coding exons and promoter of *SLC4A11* in a cohort of 20 Indian CHED families only identified mutations in 11 families leading the

authors to speculate that genetic heterogeneity may exist for CHED (Hemadevi et al., 2008).

1.6.4 Harboyan Syndrome

Harboyan syndrome (MIM#217400) is characterised by CHED in association with progressive sensorineural hearing loss. Typically, the hearing loss does not present at birth but within the age range of 10-25 years, affecting mainly the 20-50db range (mild to moderate in severity) and affecting the higher frequencies (Desir et al., 2007). Recessive mutations in *SLC4A11* also account for this phenotype, and the authors, who studied six families with Harboyan syndrome and one with CHED and possible Harboyan syndrome from varying ethnic backgrounds, concluded that some *SLC4A11* mutations caused CHED while others caused Harboyan syndrome. The reason for which remained unclear. Over 60 different homozygous or compound heterozygous mutations in *SLC4A11* have been reported to cause either CHED or Harboyan syndrome with little evidence to support a genetic basis for the difference between these phenotypes (Desir and Abramowicz, 2008, Mehta et al., 2010).

1.6.5 Solute Carrier Family 4, Member 11

Much of the previous work on the human Solute Carrier Family 4, Member 11 gene *(SLC4A11)* was performed on the NM_032034 transcript, which consists of 19 exons spanning 11,774 base pairs (bp) of genomic DNA and encodes an 891 amino acid protein (Vilas et al., 2011). However, the largest mRNA transcript described NM_001174090 consists of 20 exons and encodes a 918 amino acid protein. Expression of *SLC4A11* has been demonstrated in the human corneal endothelium on transcriptome analysis (Chng et al., 2013). The *Slc4a11* mouse knockout was shown to exhibit increased endothelial cell size and decreased endothelial cell density with increasing age (Han et al., 2013). The gene is also expressed in the fibrocytes of the stria vascularis in the inner ear in mice (Lopez et al., 2009, Groger et al., 2010) and in the

kidney (Groger et al., 2010, Han et al., 2013). It has been postulated that as the intermediate cells of the striae vascularis share a common neural crest origin with corneal endothelial cells, the absence of a borate dependent effect on the proliferation of these cells might give rise to sensorineural hearing loss, or alternatively that this might result from fluid imbalance in the inner ear (Desir et al., 2007).

The SLC4A11 protein is expressed on the cell membrane and has 14 transmembrane domains. It was previously thought to be a sodium borate transporter and was originally named Bicarbonate Transporter-related Protein 1 (BRT1). When *SLC4A11* expression constructs were transfected into human embryonic kidney cells (HEK293) with the mutant and wild-type *SLC4A11* cDNAs, immunoblots indicated that there was little or no expression on the cell surface of the predicted 120kDa mutant protein compared to the wild-type BRT1. The mutant protein therefore appeared not to be processed through the endoplasmic reticulum and subsequently failed to reach the cell membrane to fulfil its role as an ion transporter (Vithana et al., 2006, Vilas et al., 2012b). More recent studies have indicated that the SLC4A11 protein is a Na+:OH- transporter with no significant affinity to borate or bicarbonate ions, but with a role in transporting Na+ coupled to OH contributing to pH regulation in the corneal endothelium (Ogando et al., 2013, Jalimarada et al., 2013, Kao et al., 2015).

1.7 Posterior Polymorphous Corneal Dystrophy

1.7.1 Clinical Features

Posterior Polymorphous Dystrophy (PPCD) (PPCD1 MIM#122000, PPCD2 MIM#609140, PPCD3 MIM#609141) is an autosomal dominant condition characterised by deep corneal lesions of varying shape including nodular, vesicular and blister-like lesions. It is rarely present in the immediate post-natal period and clinically shows a slow or non-progressive course (Schorderet, 2015). PPCD shares some of the corneal features with the non-

inherited condition iridocorneal endothelial syndrome. The clinical history (in particular the presence of a family history) is useful in distinguishing the two conditions, and specular microscopy has also been shown to play a useful diagnostic role (Laganowski et al., 1991).

1.7.2 The genetic basis of PPCD

Dominant mutations in collagen type VIII, alpha-2 (*COL8A2*) (MIM*120252) have been implicated in PPCD in the same study that found *COL8A2* mutations to be causative for early-onset FECD (Section 1.10.2). The authors sequenced the *COL8A2* gene in one a PPCD family with two affected family members (the structure of the pedigree was not shown) both of whom had undergone previous penetrating keratoplasty, and ultrastructural analysis of the corneal button had confirmed PPCD. A c.1364C>A;p.Gln455Lys mutation was detected in the affected family members but not in any of the 15 unrelated sporadic cases of PPCD, nor was it present in ethnically-matched controls (Biswas et al., 2001). Mutations in *COL8A2* have not been identified in PPCD in any other study (Kobayashi et al., 2004, Yellore et al., 2005) suggests the possibility that this variant may be a rare polymorphism in the PPCD family described (Biswas et al., 2001).

The evidence for mutations in Transcription Factor 8 (*TCF8*) (MIM*189909) (Section 1.10.4) also known as Zinc Finger E Box-binding Homeobox 1 (*ZEB1*), as causative for PPCD is more convincing. Screening of 14 affected family members in three generations of a large PPCD family revealed a 2bp deletion 2916_2917delTG in the last exon of *TCF8*. This was present in all family members, however there were also unaffected family members who harboured the mutation. Frameshift mutations in *TCF8* were found in 4 other PPCD families of the 10 that were screened (Krafchak et al., 2005). Functional studies involving the transfection of mutant constructs corresponding to *TCF8* truncating mutations into a human corneal endothelial cell line indicated that certain truncating mutations were associated with altered nuclear localisation, whereas *TCF8* missense changes did not. The

resulting reduction of protein abundance in truncating mutations compared with the *TCF8* wild type construct led the authors to postulate that the observed decrease in protein levels was likely the result of nonsense-mediated decay, which in turn would lead to haploinsufficiency (Chung et al., 2014). Mutational analysis of 18 unrelated patients with PPCD in another study further supported the finding that truncating mutations in *TCF8* give rise to PPCD. By contrast, missense mutations caused two other progressive eye conditions with a genetic basis, Fuchs Endothelial Corneal Dystrophy (described fully in Section 1.9) and Keratoconus (Lechner et al., 2013). Liskova et al performed whole exome sequencing, SNP arrays and long PCR to ascertain the chromosomal deletion breakpoints in the 3 patients of a PPCD family, along with a cohort of 14 unrelated PPCD patients previously found to be negative for *TCF8* mutations. A large heterozygous deletion of around 3.3Mb encompassing the *TCF8* gene was found all affected patients in the PPCD family. Two additional deletions with different breakpoints were identified in two other unrelated patients (Liskova et al., 2015).

Causative mutations in the Visual System Homeobox Gene 1 (*VSX1)* (MIM*605020) gene on chromosome 20 have been detected in both PPCD and Keratoconus. *VSX1* is a member of the Vsx1 group of vertebrate paired-like homeodomain transcription factors. These transcription factors are distinguished by the presence of the CVC domain, a highly conserved region of unknown function, which lies C-terminal of the homeodomain (Héon et al., 2002). Recently, non-coding mutations in the promoter of Ovo-Like 2 (*OVOL2*) (MIM*616441) on Chromosome 20p have been implicated in PPCD. *OVOL2* encodes ovo-like zinc finger 2, a C2H2 zinc-finger transcription factor which induces epithelial-to-mesenchymal transition (EMT) via direct repression of *TCF8* expression (Davidson et al., 2016).

1.8 X-Linked Endothelial Dystrophy

A large multi-generational pedigree from Western Austria was described with a total of 60 family members of whom 35 were individuals affected with a

corneal endothelial dystrophy which followed X-linked inheritance (MIM%300779) (Schmid et al., 2006). Males were more severely affected than females and there were no examples of male-to-male transmission. The one-year-old proband presented with a bilateral, milky ground-glass corneal opacification. Other males had abnormalities of the endothelium resembling moon craters, complicated by severe band keratopathy. Females presented with moon crater-like appearance on slit-lamp examination of the corneal endothelium. Ten eyes from six patients underwent penetrating keratoplasty. Light microscopy of the corneal button of one of the affected males revealed irregular thinning of the epithelium and Bowman's layer. DM was irregularly thickened with occasional excavations. Endothelial cells were atypical and multi-layered, with areas of DM devoid of endothelial cells. Linkage analysis utilised a panel of 25 microsatellite markers covering the X chromosome in 50 family members. On multipoint analysis a maximum LOD score of 10.90 was obtained, indicating strong evidence of linkage in this family. The 14.79Mb interval between markers DXS8057 and DXS1192 on chromosome predicted to contain 72 genes. However, the mutation, gene and protein involved have yet to be identified.

1.9 Fuchs Endothelial Dystrophy

Prior to the advent of slit-lamp biomicroscopy, Fuchs endothelial corneal dystrophy (FECD) was originally described by an Austrian ophthalmologist, Ernst Fuchs, as "dystrophia epithelialis" (Fuchs, 1910) due to epithelial involvement at the advanced stages. The primary pathology however is located in the corneal endothelium. FECD is the most common endothelial dystrophy and tends to occur as a late-onset disease. FECD accounts for 22% of corneal transplants in the UK (Keenan et al., 2012). Its genetic basis is complex and heterogeneous, demonstrating variable expressivity and incomplete penetrance.

1.9.1 Epidemiology of FECD

FECD is more common in women than men (Krachmer et al., 1978). There are no prevalence studies for FECD in the UK and few large-scale studies world-wide. Many of the studies published examine their study populations for the prevalence of guttata (Section 1.9.2.1). Lorenzetta et al examined a multicultural cohort of 1016 individuals (783 females and 233 males) in a US outpatient clinic. The author stratified the recruited patients by age, sex and race, and concluded that 3.9% of individuals over the age of 40 years old exhibited confluent guttata and found the sex preponderance of corneal guttata to be equal between males and females (Lorenzetti et al., 1967). The Reykjavik Eye Study assessed the prevalence of corneal guttata in 774 study in subjects 55 years or older. Guttata were seen in 11% of females and 7% of males (Zoega et al., 2006). By contrast a Japanese study found the prevalence of guttata to be 3.8% in patients aged over 56 years old (Nagaki et al., 1996).

1.9.2 Clinical Features of FECD

1.9.2.1 Corneal Guttata

FECD is characterised by focal accumulations of collagen posterior to the DM, known as guttata, and by loss of endothelial cells. The remaining endothelial cells exhibit altered morphology and degeneration, and there are wide intercellular spaces and an absence of intercellular junctions (Waring et al., 1978b). A longitudinal study of 4 FECD patients monitored by corneal photographs taken up to 30 months apart revealed that the relative positions of individual guttata remained unchanged, very few guttata disappeared and the emergence of many new guttata was observed (Gottsch et al., 2006). Increased numbers of guttata have been correlated with a statistically significant reduction in endothelial cells counts (Jackson et al., 1999). This causes an influx of fluid into the stroma, leading to oedema and loss of corneal clarity followed by disrupted vision. End-stage disease is typified by painful epithelial bullae (small vesicles) (Figure 1.11D) as the cornea

decompensates in its ability to maintain stromal hydration (Eghrari et al., 2015a).

1.9.2.2 Clinical Presentation and Course

FECD exists in early-onset and late-onset forms, the former being very rare. Early-onset FECD represents a distinct phenotypic variant, typically with smaller guttata than are seen in late-onset FECD, and may present as early as the first decade. Late-onset FECD generally presents in the fifth or sixth decade (Weiss et al., 2015). The patient may complain of glare, diminished vision and discomfort, which is typically worse on awakening due to increased epithelial and stromal oedema. Corneal examination reveals endothelial guttata, as shown in Figure 1.11A. These are seen on retroillumination prior to the patient becoming symptomatic. As the disease progresses, guttata coalesce and the endothelial cell numbers are reduced giving rise to stromal oedema and full corneal oedema with epithelial involvement (Krachmer et al., 2011). Subepithelial fibrous scarring and superficial vascularisation may also be seen in advanced cases. In 1979, Dr Jay Krachmer and colleagues proposed a scale for grading FECD which is shown in Table 1.4.

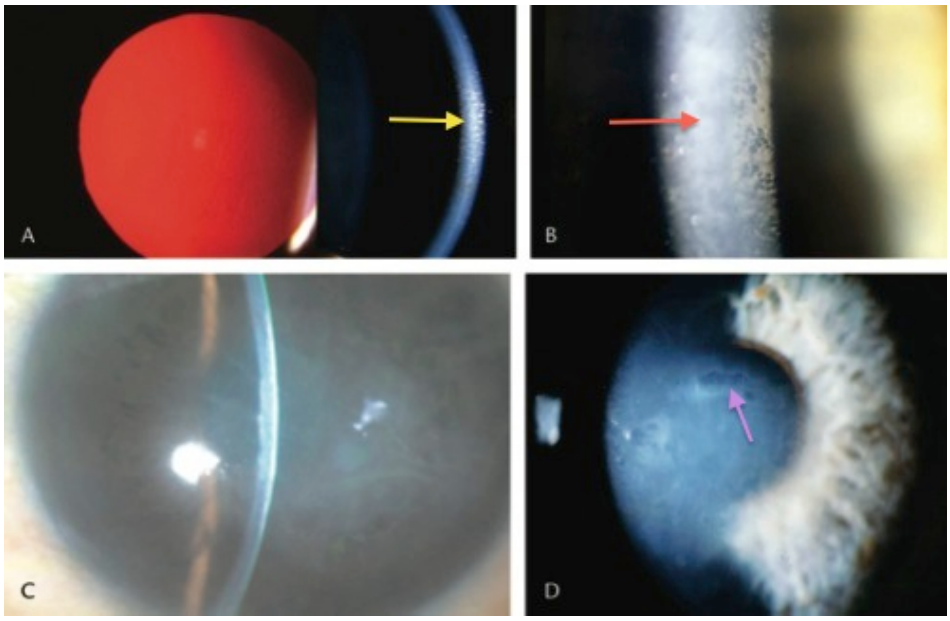


Figure 1.11 Fuchs endothelial corneal dystrophy. A. Central guttata viewed in retroillumination and in the slit beam (shown with the yellow arrow) B. Cornea guttata as seen in specular reflection (red arrow). C. Advanced stromal oedema. D. Advanced endothelial decompensation with epithelial microcyst (purple arrow) and bullous oedema. Reproduced with permission (Weiss et al., 2008).

Grade 0	No apparent disease. Up to 11 central guttata.
Grade 1	Definitive onset of the disease. Twelve or more central, non-confluent guttata in at least one eye.
Grade 2	A zone of confluent central guttata 1 to 2 mm in horizontal width.
Grade 3	A zone of confluent central guttata 2 to 5 mm wide.
Grade 4	A zone of confluent central guttata greater than 5 mm wide.
Grade 5	A zone of confluent central guttata greater than 5 mm wide plus oedema of the corneal stroma and/or corneal epithelium.

Table 1.4 The FECD Krachmer Grading Scale (Krachmer et al., 2011).

1.9.3 FECD Associations

A link between FECD and angle closure glaucoma has been described, but with conflicting results. One study described a series of 24 patients with FECD and found a statistically significant difference between axial hypermetropia/shallow anterior chamber and FECD (Pitts and Jay, 1990). However, a further series of 23 FECD patients found no association between the FECD and angle closure glaucoma (Loewenstein et al., 1991).

One retrospective analysis of 257 patients with FECD compared with 584 controls indicated no statistically significant association in open angle glaucoma between cases and controls (Rice et al., 2014). 107 of the 1610 eyes (6.6%) enrolled as part of the FECD Multicentre Study were noted to have glaucoma or ocular hypertension whereas the prevalence in the control group was 6.0%. Although there was little overall difference in the prevalence between the controls and the FECD cohort, the prevalence of glaucoma or ocular hypertension was higher in index cases (11.2%) with a FECD grading of 4-6 (Table 1.4) compared with controls (Nagarsheth et al., 2012). The authors concluded that patients with more clinically severe FECD should be monitored for the development of glaucoma.

Given the observed association between endothelial dystrophies and hearing loss (Desir et al., 2007, Desir and Abramowicz, 2008) and the overlap in genetic aetiology between CHED and FECD (Vithana et al., 2006, Vithana et al., 2008), the possibility of hearing impairment in FECD patients has been explored. A cross-sectional observational case-control study was carried in the Netherlands (Stehouwer et al., 2011). A cohort of 72 FECD patients and 180 matched controls were interviewed by means of a telephone questionnaire about their hearing. A higher percentage of the FECD group (45.8%) reported hearing disability compared with the control group (34.7%) (odds ratio of 1.59).

An analysis of risk factors from the FECD Multicenter Study found that female sex increased the odds of developing FECD by 34%. Smoking increased the odds of developing FECD by 30% (Zhang et al., 2013).

1.10 FECD Genetics

1.10.1 Summary of FECD Genetics

FECD appears to exist most commonly in the absence of family history but familial forms with dominant inheritance have been documented (Krachmer et al., 1978, Rosenblum et al., 1980). Studies of such families have implicated a number of genes in FECD. Two studies have identified dominant mutations in *COL8A2* (MIM*120252) (p.Gly455Lys and p.Leu450Trp) on chromosome 1 as a cause of early-onset FECD (Biswas et al., 2001) (Gottsch et al., 2005). Dominant mutations causing late-onset FECD have been identified in five other genes implicating them in late-onset FECD, four of which (*SLC4A11* (MIM*610206) (Vithana et al., 2008), *ZEB1*/*TCF8* (MIM*189909) (Riazuddin et al., 2010c), *LOXHD1* (MIM*613072) (Riazuddin et al., 2012) and *AGBL1* (MIM*615496) (Riazuddin et al., 2013)), appear to account for only a small proportion of cases. The fifth was identified by Baratz et al, who performed a relatively small-scale genome-wide association study (GWAS) using only 130 FECD cases and 260 controls. Their results showed a strong association between FECD and common non-coding variants in the gene encoding transcription factor 4 (*TCF4*) (MIM*602272) on chromosome 18 (Baratz et al., 2010). Subsequently a trinucleotide expansion in *TCF4* (Baratz et al., 2010, Wieben et al., 2012) was found in 79% of cases compared to only 3% of controls, which suggests that this expansion account for most cases of the condition.

The FECD Genetics Multicentre Study Group which ultimately aims to map genes for FECD published the baseline characteristics of recruited cases and controls, and calculated heritability estimates. The group investigated 322 families, 650 sibling pairs and 304 controls and found heritability estimates of

0.304 for severe disease and concluded that the clinical phenotype of FECD was highly heritable in a general Caucasian population. They also showed that central corneal thickness was more strongly heritable than FECD, with an overall heritability measure of 0.466 (Louttit et al., 2012).

1.10.2 Collagen, Type VIII, Alpha-2

Family members of a large Caucasian multi-generational pedigree from the north-east of England identified has having early-onset FECD were independently examined by three ophthalmologists. All 15 affected individuals and unaffected individuals over the age of 50 were selected for linkage analysis. The authors demonstrated linkage to chromosome 1p34.3-p32, and felt that Collagen, Type VIII, Alpha-2 (*COL8A2)* represented a strong candidate gene. *COL8A2* is a large 2-exon gene, which encodes the alpha 2 chain of the type VIII collagens. Type VIII collagens are heterotrimeric proteins composed of alpha-1(VIII) (COL8A1) and alpha-2(VIII) (COL8A2) polypeptides in a 2:1 ratio. They are members of the 'short chain collagen' subfamily. *COL8A2* encodes a 644-amino acid protein with a distinct N-terminal domain, a central triple-helical domain and C-terminal non-triple helical domain. Transcriptome analysis showed that the *COL8A2* mRNA is present in the corneal endothelial transcriptome (Chng et al., 2013) and is a major component of DM (Muragaki et al., 1991). Sanger sequencing of all family members revealed a mutation in *COL8A2,* c.1364C>A;p.Gln455Lys. This mutation was also found to segregate with the phenotype in two additional early-onset FECD families and a PPCD family (Biswas et al., 2001). *COL8A2* mutations have also been implicated in two further reported early-onset FECD families (Gottsch et al., 2005, Liskova et al., 2007) and in 15 early-onset FECD patients from six pedigrees of Korean origin (Mok et al., 2009). However, the lack of *COL8A2* mutations in early-onset FECD cohorts in other studies (Aldave et al., 2006, Kobayashi et al., 2004, Kuot et al., 2013) suggests mutations in this gene are a rare cause of this disease.

1.10.3 Solute Carrier Family 4, Member 11

Recessive mutations in Solute Carrier Family 4, Member 11 (*SLC4A11)* cause CHED (Vithana et al., 2006) (Section 1.6.5). Dominant mutations in *SLC4A11* are also a rare cause of FECD (Vithana et al., 2008). In Vithana et al's 2008 study, the authors recruited 89 patients to their study, 64 of whom were of Chinese ethnicity from Singapore and Hong Kong. On Sanger sequencing of all 19 exons of the *SLC4A11* gene in these patients, four heterozygous mutations were identified that consisted of 3 missense mutations and one 2bp deletion, all of which were absent from 354 ethnically-matched controls. The paper also showed that combinations of mutant (CHED and FECD) and wild-type constructs gave differing amounts of expression on the cell surface. It postulates a dominant negative mechanism with the FECD missense mutations. A full description of the *SLC4A11* gene is provided in Section 1.6.5**.**

1.10.4 Transcription Factor 8

Transcription Factor 8 *(TCF8)*, located on chromosome 10, encodes a zinc-finger homeodomain protein, a transcription factor that binds to DNA at a conserved sequence (CACCTG). It plays a critically important role both in development and disease through the repression of transcription of genes important for maintaining the epithelial phenotype (Vandewalle et al., 2009). Heterozygous frameshift mutations in the *TCF8* gene were previously identified as causing PPCD (Krafchak et al., 2005) (section 1.7.2). Subsequently, heterozygous missense mutations in *TCF8* were identified as causing FECD (Riazuddin et al., 2010a). As *SLC4A11* mutations are causative for CHED (Vithana et al., 2006) and a small proportion of FECD cases (Vithana et al., 2008), and *TCF8* mutations had been identified as causative for PPCD (Krafchak et al., 2005) the authors described this as their rationale for screening a large FECD pedigree with 12 affected patients in two generations for *TCF8* mutations. They found a heterozygous missense mutation c.2519A>C; p.Q840P which was present in 5/12 affected family

members, and additionally found a further 4 missense mutations on screening 192 unrelated FECD patients. They performed a functional assessment by suppressing *TCF8* expression in zebrafish embryos, which gave rise to shortened body axis and pronounced detachment of cells along the dorsal axis. These phenotypes were then rescued with constructs containing the WT *TCF8* mRNA. Although the authors were cautious in extrapolating the results of a developmental assay to late-onset corneal condition, they postulated that the TCF8 mutations they identified had a loss-of-function effect.

1.10.5 Lipoxygenase Homology Domain-containing 1

In 2012, mutations in Lipoxygenase Homology Domain-containing 1 (*LOXHD1)* (MIM*613072) were implicated in FECD (Riazuddin et al., 2012). They analysed three families previously linked to a locus on chromosome 18q (Sundin et al., 2006a). Whole exome sequencing was carried out on one affected and one unaffected individual from each family, however in two of the families no causal mutation was found. They identified a c.1639C>T, p.Arg547Cys variant in the third FECD family then went on to identify an additional 14 predicted pathogenic variants which they proposed accounted for 7.2% of FECD cases of their cohort of 207 unrelated FECD patients. The authors approached their analysis as a complex disorder not as a traditional Mendelian disorder. Out of 288 controls (576 chromosomes) only 8 potentially pathogenic alleles were observed (1.4% of alleles), suggesting significant enrichment of putative pathogenic variants in FECD cases over controls.

The authors then went on to detect the distribution of LOXHD1 in mouse corneas using a rabbit polyclonal antibody, detecting it in the epithelium and the endothelium. They then assessed the LOXHD1 distribution in a corneal button from an affected individual of one of their FECD pedigrees found to harbour a *LOXHD1* mutation, and compared this with two control corneal buttons. Examination of the corneal sections, showed an increase in staining in the endothelium and DM of the proband with the c.1639C>T; p.Arg547Cys *LOXHD1* mutation compared with both controls.

LOXHD1 is a large gene consisting of 40 exons with a main transcript which is 179,774bp in length. The LOXHD1 protein consists of 15 PLAT (polycystin/lipooxygenase/alpha-toxin) domains, 120 amino acid domains which form two opposing anti-parallel beta sheets to make a structures known as a beta-sandwich. LOXHD1 is thought to be involved in targeting proteins to the plasma membrane. It is expressed in the mechanosensory hair cells in the inner ear, and recessive mutations in this gene lead to auditory deafness (Grillet et al., 2009) (Edvardson et al., 2011) (MIM#613079). Although the exact pathogenic mechanism of *LOXHD1* in FECD is not known, it has been postulated that, due to the presence of marked precipitates in the corneas of the index family, a proportion of FECD cases might be caused by aggregation defects similar to those seen in other late-onset diseases such as Parkinson's disease. As LOXHD1 is present in low abundances in the normal cornea, the authors (Riazuddin et al., 2012) speculated that cytotoxic effects of increased LOXHD1 concentrations might have a role in the pathogenesis of FECD in these cases.

1.10.6 Transcription Factor 4

The role of Transcription Factor 4 (*TCF4)* in FECD was first described by Baratz et al (Baratz et al., 2010), who performed a small-scale GWAS using 130 FECD cases and 260 controls. Control subjects were 60-years-old or over and had no guttata observed. They identified one SNP, rs613872, in intron 3 of *TCF4*, that reached genome-wide significance, with minor allele G enriched in FECD cases compared to controls. The odds ratio for one copy of the risk allele was 5.47.

Following the publication of this GWAS (Baratz et al., 2010), there was speculation as to potential mechanisms by which *TCF4* gives rise to FECD. The largest *TCF4* transcript is 21 exons in length, however multiple alternatively spliced transcripts exist (https://genome.ucsc.edu/index.html). Heterozygous missense mutations cause Pitt-Hopkins syndrome (MIM #610954) (Sweatt, 2013). The TCF4 protein is a member of the E-protein

family of class I basic helix-loop-helix (bHLH) transcription factors and is known to upregulate *TCF8*, which in turn has a role in the repressing of E-cadherin, which can be associated with epithelial-to-mesenchymal transition (EMT). Although little is known about EMT in the cornea, one hypothesis proposed before the start of this PhD was that corneal stem cells that are located in a niche at the corneal endothelial periphery require EMT in order to migrate toward and replace damaged endothelium (Wright and Dhillon, 2010). It was postulated that mutations in *TCF4* might reduce this protective process.

The association of the SNP rs613872 has since been replicated in many studies in Caucasians (Igo et al., 2012, Li et al., 2011, Eghrari et al., 2012, Kuot et al., 2012, Stamler et al., 2013). The rs613872 minor G allele was not found in one study of Chinese patients (Wang et al., 2013). In another study of a Chinese population, the rs6137872 SNP was also not found to be polymorphic, but the authors found two other *TCF4* SNPs (rs17089887 and rs17089925) to be associated with FECD, conferring a >2.3-fold increase in disease risk per copy of risk allele compared to the wild-type (Thalamuthu et al., 2011).

Subsequently, Wieben et al (Wieben et al., 2012) found that a known intronic trinucleotide repeat in *TCF4*, located 43 kilobases away from rs613872, was even more strongly associated with FECD than rs613872. Of their cohort, 79% of cases carried at least one *TCF4* allele with more than 50 trinucleotide repeats, compared with only 3% in controls. This was independently replicated by Mootha et al (Mootha et al., 2014) who used a PCR-based method, previously used to test for repeats in Myotonic Dystrophy (Warner et al., 1996), to confirm the association of the expanded CTG18.1 allele with FECD.

1.10.7 ATP/GTP-Binding Protein-like 1

Riazuddin et al reported that dominant mutations in ATP/GTP-Binding Protein-like 1 (*AGBL1)*, a 24-exon gene on chromosome 15q, were causative of FECD. The authors described a 3-generational FECD family and initially performed linkage analysis but the resulting lod scores were not significant. As they did note some signal from chromosome 3p and 15q, they performed WES of one unaffected and two affected individuals to identify a heterozygous nonsense mutation on chromosome 15q. The variant c.3082C>T;p.Arg1028* was not seen in 384 ethnically-matched controls, but was found on the Exome Variant Server at a low frequency of 0.35%. The variant did not however segregate fully with the disease in the family with multiple FECD affecteds. Screening of their FECD cohort revealed two additional unrelated FECD individuals who harboured this mutation. In addition, a further heterozygous missense mutation c.2969G>C;p.>Cys990Ser was also seen in another case from their FECD cohort. They suggested that *AGBL1* mutations accounted for 1-2% of the genetic burden of FECD (Riazuddin et al., 2013). The gene encodes a member of the cytosolic carboxypeptidase family, which catalyses the deglutamylation of polyglutamated proteins (Rogowski et al., 2010). The authors (Riazuddin et al., 2013) also concluded that *AGBL1* altered protein-protein interactions with TCF4, although given that now the pathogenic effect of *TCF4* is thought to be due to the toxic effect of the *TCF4* trinucleotide expansion on the cells, their proposed mechanism seems less plausible.

1.10.8 Other FECD Loci

In addition to the mutations in the genes described above, 3 dominant loci on chromosomes 5 (FECD5, MIM%613269) (Riazuddin et al., 2009), 9 (FECD7, MIM%613271) (Riazuddin et al., 2010a) and 13 (FECD2, MIM%610158) (Sundin et al., 2006b) have been reported. One locus for FECD was mapped to on chromosome 18q21.2–21.32 (FECD3) (Sundin et al., 2006a). The authors identified three large late-onset pedigrees and established linkage to chromosome 18 (maximum LOD score 3.41) in one of the families and

tentative linkage (maximum LOD scores 2.89 and 2.45) in the other two families. There were, however 8 phenocopies in 36 individuals with no disease overall in the three families and one example of non-penetrance. The causative gene in this locus was subsequently identified as *LOXHD1*, although *TCF4* also lies within this locus and may be the reason for the positive LOD scores in this linkage study.

1.11 Mitochondrial disease and Oxidative Stress in FECD

Theories of oxidative stress have been proposed in FECD. One case report described a 48-year-old woman who had FECD as well as sensorineural hearing loss, diabetes, cardiac conduction defects, ataxia and hyperreflexia (defects usually associated with another eye condition Leber's Hereditary optic neuropathy which is strongly associated with mitochondrial missense mutations). Lymphocyte mitochondrial DNA (mtDNA) studies were carried out and showed missense mutations at mt15257 (G to A, aspartate to asparagine in cytochrome b) and mt4216 (T to C, tyrosine to histidine in ND1), suggesting that mitochondrial defects might play a role in the pathogenesis of FECD (Albin, 1998). Another study conducted in Poland (Wojcik et al., 2015) examined the relationship between 5 polymorphisms in base excision repair genes in FECD cases and controls. The c.1196A>G polymorphism of the *XRCC1* gene was positively correlated with FECD. Other mitochondrial polymorphisms have also been reported to be associated with FECD. Ten mtDNA variants were identified as part of a study of GWAS of 530 FECD cases and 498 controls of European descent. Many of the variants identified were also associated with other human diseases including Alzheimers disease, Parkinson disease, as well as other ocular diseases ARMD, POAG and Keratoconus (Li et al., 2014). Furthermore, it has been demonstrated that when FECD lymphocytes are subject to oxidative stress, the damage to mtDNA is not fully repaired and the number of lesions remaining is higher when compared with controls (Czarny et al., 2014), thus supporting the hypothesis that mutagenesis of mtDNA might cause susceptibility to FECD. As the cornea is part of the anterior segment, it is potentially exposed to

greater levels of UV light than other parts of the eye. One of the main inducers of apoptosis is oxidative stress, and FECD endothelia are more susceptible to oxidative stress compared with that of normal controls (Jurkunas et al., 2010, Azizi et al., 2011). It is therefore feasible that FECD patients have less efficient systems of repair, and that mtDNA mutations could feasibly reduce the cells' capacity to repair DNA, leaving it susceptible to oxidative stress and apoptosis.

1.12 Gene expression in the corneal endothelium

Gene expression in FECD patients has been carried out using serial analysis of gene expression (SAGE), which looked at the upregulation and downregulation of RNA transcripts compared with normal human endothelium. The study identified 9,530 tags from normal endothelium and 9,606 from FECD endothelium. The expression of 18 transcripts were upregulated and 36 that were down-regulated compared with control tissue (Gottsch et al., 2003). Expression analysis of cultured corneal endothelial cells from both young and old donors has been carried out using high throughput RNA sequencing methods (Chng et al., 2013). The study involved the dissection of corneal endothelial cells from donor corneas, RNA sequencing and annotation of the variants, and reported a comprehensive expression profile for the ageing human corneal endothelium.

1.13 Imaging of the Corneal Endothelium

Although the measurement of central corneal thickness allows the pump function of the endothelium to be inferred, endothelial cells can be directly visualised using a non-contact method known as specular microscopy (SM) (Maurice, 1974) which provides a high magnification view of specular light reflected from the corneal endothelium. It has the disadvantage of giving poor image quality in the presence of severe corneal oedema. This can be overcome by the use of confocal microscopy (Kaufman et al., 2004), although

in general the method of choice is SM as it is relatively user friendly and widely available (Chiou et al., 1999).

From the SM images several morphological variables can be calculated, including the mean cell density, mean cell area, the coefficient of variation of mean cell area (CV) (a measure of variation in cell size, polymegathism). In the healthy cornea, 70-80% of endothelial cells are hexagonal. Deviation from hexagonality is referred to as pleomorphism (Bourne and McLaren, 2004).

1.14 Management of Endothelial Dystrophies

The initial treatment for FECD is temporizing and supportive. Medical management includes the use of hypertonic saline solutions and of a hair-dryer to increase tear evaporation (Yanoff and Duker 3rd, 2009).

The mainstay of surgical management for endothelial disease remains corneal graft surgery (keratoplasty), which involves replacing the diseased corneal tissue with donor tissue. Penetrating Keratoplasty (PK) (full-thickness graft) is still the primary surgical procedure for CHED. A large case series suggested that delayed keratoplasty, even in the presence of nystagmus, seemed to offer more favourable outcome compared with those operated on early (before the age of 12 years) (Ozdemir et al., 2012).

Corneal graft techniques have evolved in recent years (Guell et al., 2014). Corneal graft surgery performed in the UK (Keenan et al., 2012) and the US (Park et al., 2015) for FECD and PPCD is increasingly carried out using a lamellar technique, Descemet's stripping endothelial keratoplasty (DSEK), a method described originally by Melles (Melles et al., 1998). This involves replacing only the defective endothelial layer of the cornea with donor DM and a posterior stroma, leaving behind the healthy epithelium and anterior stroma. Further refinement of DSEK has taken place with the development of Descemet's Membrane Endothelial Keratoplasty (DMEK) where the donor tissue is composed solely of DM and endothelium (Melles et al., 2006).

Comparison of the two lamellar techniques has indicated faster and more complete visual rehabilitation with the use of DMEK (Tourtas et al., 2012). In CHED patients, successful DSEK has been reported (Mittal et al., 2011). A paired-eye comparison of PK and DSEK in a series of CHED patients indicated earlier visual rehabilitation and stabilisation of refraction in the eyes managed with DSEK (Ashar et al., 2013).

The regenerative capacity of the corneal endothelium, as discussed in Section 1.3.7, might ultimately prove more forgiving than some of the current surgical approaches. Minimizing surgical intervention and thus avoiding surgical complications may greatly benefit endothelial dystrophy patients (Bruinsma et al., 2013) but before this point can be reached further knowledge about disease pathogenesis and new treatments are required. As the cornea is an easily visualized and accessible tissue, it is potentially an ideal target for future gene therapies (Williams and Klebe, 2012). A sound understanding of the genetics of the disease may well lead to the successful tailoring of treatments in endothelial dystrophy.

1.15 Objectives of this Study

The aims of this study are:

1. A British Ophthalmic Survey Unit (BOSU) study carried out in conjunction with the Royal College of Ophthalmologists is presented in Chapter III. It examines the epidemiology of corneal dystrophies in patients of 40 years or less.

2. The second aim is to determine whether recessively inherited CHED, Harboyan syndrome and late-onset FECD may all coexist over time within one family, as the result of the same mutation. Two previously reported CHED families (Vithana et al., 2006) are revisited. A newly-recruited CHED patient is also analysed along with one of his parents who was available for examination. This work is presented in Chapter IV.

3. The genetics of FECD are explored in Chapter V. The relative contributions of the *TCF4* SNP rs613872, the intronic CTG18.1 trinucleotide expansion and *LOXHD1* variants in a UK Caucasian FECD cohort are compared in order to clarify the significance of the original findings at the FECD3 locus. The results of both segregation of the *TCF4* CTG18.1 expansion and whole exome sequencing in three local FECD families are also presented.

2 Materials and Methods

2.1 Solutions

1 x Tris-EDTA (TE) Buffer (pH8.0)

10mM Tris (pH 8.0)
1mM ethylenediaminetetraacetic acid (EDTA)
Adjust to pH 8.0 with concentrated hydrochloric acid

Tris Borate-EDTA (TBE) Electrophoresis Buffer (10x)

0.89M Tris
0.89M Orthoboric acid ($B(OH)_3$)
25mM EDTA (pH 8.0)

TAE Tris-acetate-EDTA Electrophoresis Buffer (1x)

40 mM Tris base
20mM Glacial acetic acid $C_2H_40_2$
1mM EDTA (pH8.0)

10x Gel Loading Dye

3x TBE
20%[w/v] Ficoll 400
0.1% [w/v] Bromophenol blue
0.2% [w/v] Xylene cyanol

Red Cell Lysis Solution

155mM Ammonium Chloride (NH_4Cl)

10mM	Potassium Bicarbonate ($KHCO_3$)
1mM	Ethylenediaminetetraacetic Acid (EDTA)

White Cell Lysis Solution

25mM	EDTA (pH 8.0)
2% [v/v]	Sodium dodecyl sulphate (SDS)

Cell Lysis Buffer – DNA

10mM	Tris pH 8.0
100mM	EDTA
20ug/ml	RNAase A
0.25% [v/v]	SDS

Make up to 1ml with deionized water.

2.2 Patient Recruitment

The project was approved by Leeds East Research Ethics Committee (reference 10/H1306/63) under the title “Investigation of human inherited corneal dystrophies: genetic, tissue and transplant analyses” and adhered to the tenets of the Declaration of Helsinki. This application was initiated by Dr Aine Rice (AR) then completed by the author (SS) during the PhD after AR left Leeds, and all project management was handled by SS. All researchers involved in patient recruitment underwent training in taking informed consent for research, Good Clinical Practice (GCP) and NHS information governance. Local FECD patients were identified by AR, SS and other members of the corneal clinical team at St James’s University Hospital (SJUH). The project was adopted onto the National Institute of Health Research (NIHR) portfolio (UKCRN portfolio no. 11297) in December 2011. Pinderfields Hospital,

Wakefield was changed from a Patient Identification Centre to a main recruiting site in August 2013.

FECD patients were ascertained in outpatient and operating theatre lists at SJUH, Leeds by AR, SS and research nurses Frances Cassidy (FC), Charmain Tidswell (CT) and Alice Van Lare (AVL) and from Bradford Royal Infirmary and Pinderfields Hospital, Wakefield, UK. When SS took over the role of PI, all patients on forthcoming corneal operating theatre lists who were identified as having endothelial dystrophies as documented in electronic casenote letters and the Medisoft database (http://www.medisoft.co.uk) were sent a written invitation with the relevant patient information leaflet, and were then subsequently approached about joining the study. These patients had undergone a detailed slit-lamp examination and specular microscopy analysis (Tomey EM-3000 specular microscope, Tomey GmbH, Erlangen-Tennenlohe, Germany) at the time of listing and satisfied the criteria of FECD Grade 1 or above were selected for recruitment following their informed consent. Peripheral blood (2-6ml) was collected by venepunture into BD vacutainer® EDTA blood collection tubes (BD Biosciences, Oxford, UK) and genomic DNA was extracted by Yorkshire Regional Genetics (SJUH) from blood leukocytes according to standard procedures (Section 2.7).

Amendments to the research protocol were sought through the Leeds East Research Ethics Committee to enable blood samples to be collected from controls individuals. Control patients were subject to slit lamp examination at NHS cataract clinics and were identified as suitable with normal endothelia and with no family history of eye disease by Consultant Ophthalmologist John Buchan. They were recruited following their informed consent.

Human random control panels, HRC-1 and HRC-3, (Public Health England, Porton Downs, UK) consisting of 192 genomic DNA samples extracted from EBV-immortalised single donor lymphoblastoid cell lines were obtained commercially (Sigma-Aldrich).

2.3 British Ophthalmic Survey Unit (BOSU) - Incidence of Young Onset Corneal Dystrophy

The BOSU study was set-up by Kamron Khan. Cases were defined as a new-onset corneal dystrophy in a patient aged below 40 years. New cases were ascertained using population based active surveillance through the BOSU at the Royal College of Ophthalmologists. All permanently employed consultant and associate specialist ophthalmologists in the UK received a monthly reporting card for a 24-month period commencing November 2011. Respondents were asked to indicate any new cases of young-onset corneal dystrophy or confirm that there were no new cases to report. Following notification, information on patients' ophthalmic history, examination findings, further treatment and follow-up was sought by questionnaire. Questionnaires were collected and the results analysed by SS. The questionnaire is shown in Appendix I.

2.4 Audiometry

CHED patients underwent audiometry screening using either the Madsen Aurical audiometer (GN Otometrics A/S, Taastrup, Denmark) at Bradford Royal Infirmary or an AD229 diagnostic audiometer (Interacoustics A/S, Assens, Denmark) Institute of Ophthalmology, Mexico City according to the manufacturers' instructions.

2.5 Corneal Tissue

2.5.1 Collection of Human Corneal Tissue

Corneal tissue was retrieved from SJUH Ophthalmology operating theatres from patients undergoing graft surgery who had given their informed consent for their tissue to be collected for research.

Amendments to the research protocol were sought through the local research ethics committee to enable control corneal tissue to be collected. Therefore, corneal tissue both from corneal endothelial dystrophy patients as well as donor tissue not utilised in lamellar surgery was collected from SJUH operating theatre. Control tissue was only collected where the donor had consented for their tissue to be used for medical research. All tissue was stored in RNAlater (Thermo Fisher Scientific) immediately following dissection in theatre.

Control tissue was processed, stored, assessed and packaged according to Manchester Eye Bank standards (Appendix II), and was deemed fit for clinical use. All eyes had been enucleated as soon as possible post mortem and always within 24 hours of death. On the intended day of grafting, corneal graft tissue stored in Eagles Minimal Essential Medium containing 2% Foetal Bovine Serum and antibiotic preparations (Appendix II) was couriered to SJUH operating theatre. Following dissection under sterile operating conditions all tissue was stored in RNAlater.

2.5.2 Collection of Bovine Tissue

Bovine eyes were enucleated from freshly culled cows by Professor Chris Inglehearn (CFI) at Dunbia Abattoir, Sawley, then brought back to the LIMM on ice. Ocular structures were dissected immediately by SS. Corneal tissue was grossly dissected from the eyes then whole corneal buttons were removed using a 9mm corneal graft trephine. Corneal tissue was either stored as a whole corneal button or further dissected into endothelium and stroma. Cornea (and other ocular structures such as lens, iris and retina) were dissected and preserved in formalin, snap frozen in liquid nitrogen or stored in RNAlater.

2.6 Collection of Human Enucleated Eyes

Pairs of eyes donated for medical research were enucleated by Aidan Hindley, the Leeds GIFT Tissue Bank Coordinator (http://www.gift.leeds.ac.uk/). The donor's gender, age, and cause of death were recorded.

The eyes were transported in a container with moist cotton wool soaked in saline by SS. Corneal debris was removed under magnification using forceps. Two openings were created using a 25 Gauge needle 3mm posterior to the limbus. One eye was embedded in paraffin and the other prepared for cryopreservation.

2.6.1 Eye Fixed for Embedding in Paraffin

The eye was placed in 4% paraformaldeyde (PFA), ensuring that it was completely submerged. After 48 hours, the eye was washed in Phosphate Buffered Saline (PBS) and stored in 70% ethanol for 48 hours.

After processing both eyes were then passed to LIMM pathologist Mike Shires for paraffin embedding using the Leica ASP200 Fully Enclosed Tissue Processor (Leica, Milton Keynes, UK) and then subject to the conditions described in Table 2.1.

	Time	Temperature (°C)
70% ethanol	30 min	37
80% ethanol	30 min	37
90% ethanol	30 min	37
95% ethanol	30 min	37
100% ethanol	1:00h	37
100% ethanol	1:00h	37
100% ethanol	1:30h	37
Xylene	1:00h	37
Xylene	1:30h	37
Xylene	1:30h	37
Wax	1:00h	65
Wax	1:00h	65
Wax	1:00h	65

Table 2.1 The paraffin embedding conditions that sectioned tissue was subject to. (Ethanol (Sigma-Aldrich), Xylene (Fisher Scientific UK, Loughborough, UK), Wax (Cellpath, Powys, Wales)).

2.6.2 Eye for Cryopreservation

The eye was placed in 4% PFA, ensuring that it was completely submerged for 2 hours. It was then stored in 10% w/v sucrose in PBS for 12 hours, 20% w/v sucrose in PBS for 12 hours then 30% w/v sucrose in PBS for 12 hours.

After processing both eyes were then passed to LIMM pathologist Mike Shires for cryopreservation using the following protocol:

Fresh tissue was carefully dissected with a sharp blade and handled with care. Fresh Sterile saline was used to prevent tissue dehydration.

1. Freshly excised tissue blocks approximately 5mm^3 was placed in a drop of Optimal Cutting Temperature (OCT) medium on a cork disc. The tissue was quickly covered in OCT and frozen at the earliest opportunity.

2. A beaker of liquid isopentane was carefully placed in liquid nitrogen until a slush of liquid and frozen isopentane is present.

3. The tissue was immersed in the isopentane slush for 30 seconds using long forceps.

4. The tissue was wrapped in aluminium foil and place in a labelled air tight container.

5. The tissue was stored at -70°C.

2.7 DNA Extraction Protocol

2.7.1 Salt Precipitation

In general, a salt precipitation technique was performed to extract DNA from blood samples that had not been frozen. Briefly, 3ml of whole blood was aliquoted into polypropylene tubes and 9ml of red cell lysis solution (Section 2.1) was added. After shaking for 10 minutes, samples were spun at 2000 x g for a further 10 minutes and the supernatant removed, leaving the white cell pellet containing DNA. The pellet was resuspended in 3ml white cell lysis solution (Section 2.1) and cells lysed by pipetting. To remove any contaminating protein, 1ml of protein precipitation solution (10M ammonium acetate) was added and samples mixed well for 20 seconds before centrifuging for 10 minutes at 2000 x g. The supernatant, containing the DNA, was aliquoted into fresh tubes and DNA was precipitated using isopropanol at 2000 x g, followed by two washes in 70% ethanol. The precipitate was air dried and redissolved in TE buffer (Section 2.1).

2.7.2 Phenol-Chloroform Extraction

To extract DNA from blood samples that had been frozen, a phenol-chloroform extraction procedure was performed. Briefly, 3ml of whole blood was aliquoted into polypropylene tubes and 9ml of red cell lysis solution (Section 2.1) was added. After shaking for 10 minutes, samples were spun at 2000 x g for 10 minutes. The supernatant was removed, leaving the white cell pellet, to which 500µl of cell lysis buffer - DNA (Section 2.1) was added. After a 1-hour incubation at 37°C, proteinase K was added to a final concentration of 100µg/ml followed by an incubation at 55°C for 1 hour. 500µl of 1:1 phenol:chloroform was subsequently added and the tube inverted several times. The tube was then mixed well until a milky solution formed and centrifuged for 10 minutes at 3100 x g. The upper aqueous phase was removed into a fresh tube and an equal volume of chloroform added to this. The tube was inverted several times and centrifuged at 3100 x g for 5 minutes. After aliquoting the upper aqueous phase into another fresh tube, sodium chloride (NaCl) was added to a final concentration of 0.2M and a further 2 volumes of 100% ethanol added before mixing and centrifuging at 3100 x g for 5 minutes. The supernatant was discarded and 75% ethanol was added to the pellet, followed by a further centrifugation at 3100 x g for 5 minutes. After removal of the supernatant, the resulting pellet was dried and redissolved in deionised water.

2.8 Primer Design and Optimisation

Oligonucleotide primer pairs were designed using either Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi), ExonPrimer through the UCSC Genome Browser (https://ihg.helmholtz-muenchen.de/cgi-bin/primer/ExonPrimerUCSC.pl?db=hg19&acc=uc010zqf.2) or Autoprimer 3 software (http://sourceforge.net/projects/autoprimer3). The different primer design software used reflected a change in preference as new software was developed during the fellowship. The reference sequence for all genes was downloaded from UCSC Genome Browser (http://genome.ucsc.edu) (Section

2.16). Details of the primer sequences and PCR conditions can be found in Appendices IIII and IV.

2.9 DNA Polymerase Chain Reaction (PCR) Amplification

For standard PCR, DNA was amplified in a final reaction volume of 10µl using 50ng of genomic DNA (2µl of DNA and 8µl of DNA mastermix). Mastermix comprised PCR buffer (Invitrogen), 1.5mM $MgCl_2$, 200µM of each dNTP (Invitrogen), 10 pmol/µl primers, and 1.0 unit of *Taq* polymerase. A standard PCR cycle, Touchdown and Hotshot cycles are shown.

Standard PCR cycle

95°C for 2 minutes (Denaturation)
40 Cycles of:
94°C for 30 seconds (Denaturation)
53 - 65°C for 45 seconds (Annealing)
72°C for 45 seconds (Extension)
72°C for 5 minutes (Final Extension)

Touch Down

98°C for 30 seconds (Denaturation)
45 cycles of:
98°C for 10 seconds (Denaturation)
70°C* for 20 seconds (Annealing)
72°C for 90 seconds (Extension)
*Reduce temperature by one degree each cycle for 5 cycles then complete 40 cycles at 64°C
Final Extension step 72°C for 10 minutes (Final Extension)

Hotshot

95°C for 10 minutes (Denaturation)
40 Cycles of:

94°C for 30 seconds (Denaturation)
53 - 65°C for 45 seconds (Annealing)
72°C for 45 seconds (Extension)
72°C for 5 minutes (Final Extension)

Reactions using HotShot mastermix (Client Life Science, Stourbridge, UK) were carried out in a 10µl volume using the manufacturer's instructions.

2.10 Whole Genome Amplification

Whole genome amplification was performed using the Illustra GenomiPhi V2 DNA Amplification kit (GE Healthcare Life Sciences, Little Chalfont, UK) according to the manufacturer's instructions. After amplification, samples were diluted 1/30 in sterile distilled water and a test PCR was performed to ensure DNA was at a concentration sufficient for PCR amplification. The standard PCR conditions utilised are outlined in Section 2.9.

2.11 DNA Visualisation and Size Fractionation using Agarose Gel electrophoresis

PCR products were mixed with gel loading dye (solution from Section 2.1) and loaded using a pipette into the wells of a 1.5% agarose gel. The 1.5% agarose gel was made up with 0.5x TBE or 1xTAE and ethidium bromide at a final concentration of 0.5µg/ml. DNA products were quantified and sized using Easy Ladder 1 (Bioline Reagents Limited, London, UK). The agarose gels were visualized on a Bio-Rad UV transilluminator and displayed using Image Lab Software (Bio-Rad Laboratories Limited, Hemel Hempstead, UK).

2.12 Genotyping and Linkage Analysis

Genotyping was performed using fluorescently tagged microsatellite markers. The chromosomal position of the markers was located against the hg19 version of the Genome Browser. PCR was carried out in a 10µl volume (as described in Section 1.7). Following the amplification step, PCR products were size fractionated. Samples were denatured and run on an ABI Genetic Analyser (Life Technologies Ltd, Paisley, UK) using polymer POP-7 and the FragmentAnalysis36_POP7-1 module 3130xl using a 500 ROX size standard (Life Technologies Ltd) and subsequently analysed using GeneMapper v4.0 software (Life Technologies Ltd). Linkage analysis was performed under the assumption of a dominant inheritance model with penetrance of 99% and zero phenocopy rate (disease allele frequency = 0.001). Multipoint LOD scores were calculated using Superlink (http://bioinfo.cs.technion.ac.il/superlink/) (Fishelson and Geiger, 2002). Marker allele frequencies were estimated on the basis of data from CEPH (http://www.cephb.fr). All microsatellite markers used can be found in Appendix IV.

Genotyping of trinucleotide repeat expansions was carried out using the method previously described by Warner et al and adapted by Mootha et al (Warner et al., 1996, Mootha et al., 2014) (Figure 2.1).

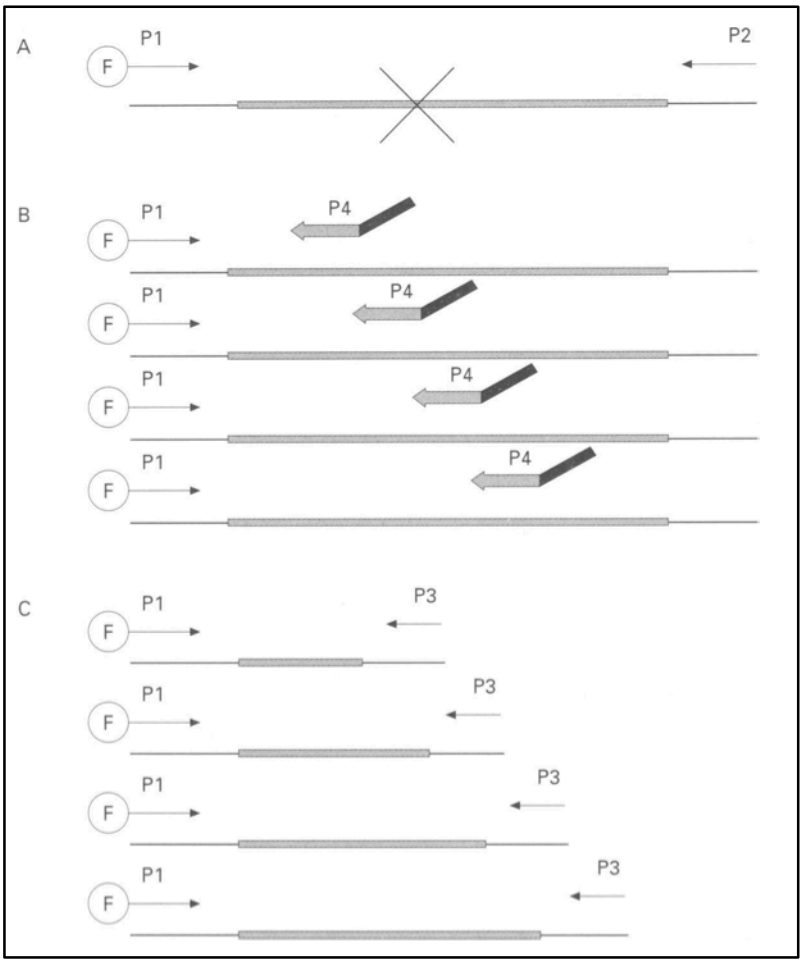


Figure 2.1 Overview of the TP PCR to genotype the CTG18.1. A general method for the detection of large repeat expansions by fluorescent PCR described by Warner et al. A. the STR assay uses primers P1 and P2 that flank the CTG repeat, but the allele fails to genotype expanded alleles. B. The specific 3' end of the P4 binds at numerous sites within the CTG repeat within the early rounds of amplification, resulting in a mixture of products. P4 is quickly consumed due to the 33:1 molar concentration of P3 to P4. C. P3 amplifies from the end of the mixture of products of the prior cycles. Reproduced with permission (Warner et al., 1996).

2.13 Sanger Sequencing

PCR products were treated with 1.5µl Exo-SAP-IT (GE Healthcare Life Sciences) then place in the thermal cycler for 37°C for 30 minutes, then 80°C for 15 minutes Purified products were subsequently sequenced using the BigDye terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems) with one of the original PCR primers according to the manufacturer's instructions. The samples were subject to an initial denaturation step of 96°C for 10 seconds followed by 25 cycles of 96°C for 10 seconds, 50°C for 5

seconds and 60°C for 4 minutes. After the sequencing reaction, 5µl of 125mM EDTA was added and the PCR products were precipitated by adding 60µl of 100% ethanol, mixing and centrifuging at 3900 revolutions per minute (rpm) for 30 minutes at 22°C using the Eppendorf Centrifuge 5180R (Fisher Scientific UK, Loughborough, UK). The supernatant was carefully removed and the pellet was washed with 70% ethanol and centrifuged at 2000rpm for 15 minutes at 4°C. After resuspending the pellet in Hi-Di Formamide (Applied Biosystems), the sequencing reactions were analysed on the ABI 3130xl Genetic Analyser using polymer POP-7 and the default RapidSeq36POP7 module. SeqScape v2.5 and Sequencing Analysis version 5.2 (Life Technologies Ltd) software were used to analyse the resulting sequence data.

2.14 RNA extraction and Reverse Transcription PCR

Patient and bovine corneal tissues stored in RNAlater were thawed on ice, homogenized using the Ribolyser (Hybaid Ltd, Teddington, UK) and RNA extracted using TRIzol Reagent (Sigma-Aldrich) and the RNAeasy Plus Universal Minikit (Qiagen) according to the manufacturer's instructions.

RNA was treated with Ambion DNAase treatment kit to prevent degradation of RNA by nucleases. cDNA synthesis was carried out using reverse transcriptase (RT) PCR. Samples were thawed on ice and 1mM Oligo $(dT)_{18}$ primers (Thermo Scientific), Random Hexamer primers (Thermo Scientific) and Diethylpyrocarbonate (DEPC) water were added. Samples were incubated at 70°C for 5 minutes then chilled on ice. A mastermix was prepared of 4mM dNTP, Bioscript™ RT (Bioline™), 2x RT Buffer (Bioline™), 40mM DTT and DEPC-treated water. 10 µl of the mastermix was added to the reaction and the samples incubated at 42°C for 30 minutes. The reaction was then terminated by incubating at 85°C and the cDNA purified using the QIAquick™ PCR Purification columns (Qiagen Ltd, Manchester, UK) according to the manufacturer's instructions. The samples then underwent a standard PCR reaction for *p53* as a house-keeping gene. The sample products were size fractionated on a 1.5% agarose gel as described before

(Section 1.10). Additionally, the samples were quantified and assessed for purity on the Agilent 2100 Bioanalyser (Agilent Technologies LDA UK Ltd, Stockport, UK) using the Agilent RNA 6000 Nano Guide (Agilent Technologies).

2.15 Whole Exome Sequencing (WES)

Genomic DNA sample concentrations were measured using a Qubit fluorometric quantification assay (ThermoFisher Scientific). Whole exome sequencing (WES) was carried out either on a HiSeq2000 using the Nimblegen v2 chip by Otogenetics Corporation (Norcross, GA), or on an Illumina Genome Analyser HiSeq2500 in the Leeds University Next Generation Sequencing Facility following library preparation using the Agilent SureSelect All Exon v4 exome enrichment kit for target capture. Libraries were prepared either by SS, Layal Abi Farraj or Clare Logan. The sequencing output consisted of 80 base pair reads. An overview of the sample preparation work flow is given in Figure 2.2.

2.16 Bioinformatics

The UCSC Genome Browser (https://genome.ucsc.edu) contains the reference sequence and working draft assemblies for a large collection of genomes. Bioinformatic searches of genomic regions and initial information about genes of interest were obtained using this website, including exon and intron sequences, mRNA sequences, protein sequences and the location of polymorphisms. Literature searches for information on known genes and loci were performed using Pubmed (http://www.ncbi.nlm.nih.gov/pubmed). Variants identified by Sanger sequencing were annotated using Mutation Mapper (http://sourceforge.net/projects/mutationmapper). The reverse compliment of an oligonucleotide sequence was calculated using Reverse Complement (http://sourceforge.net/projects/revcomp/).

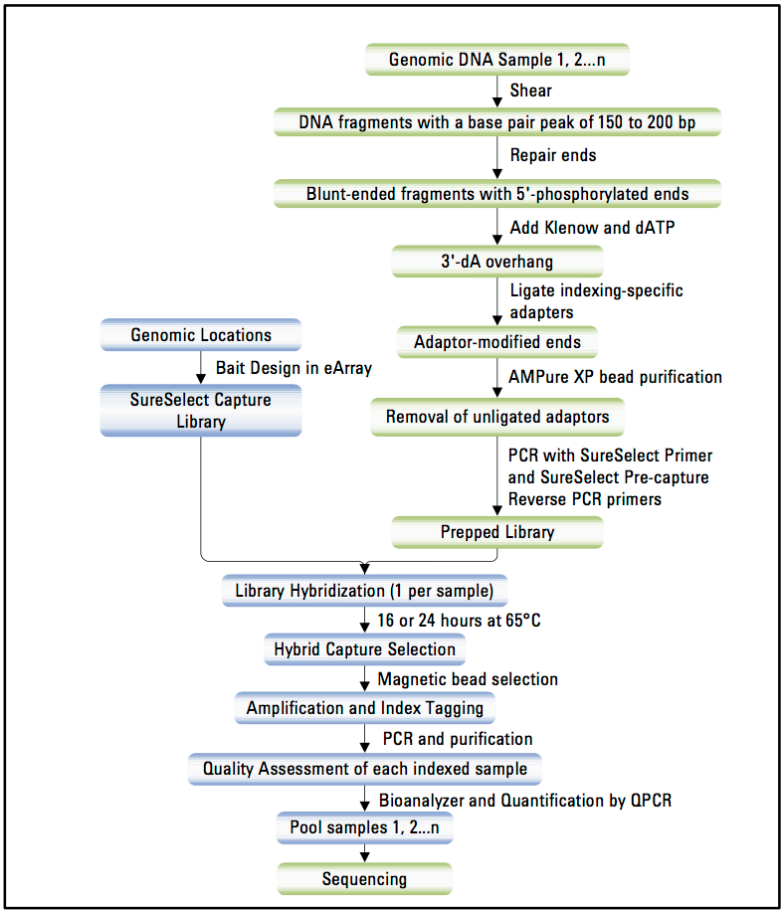


Figure 2.2 Overview of sequencing sample preparation using the Agilent SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library.

2.17 Assessment of Variant Pathogenicity

In order to determine the potential pathogenicity of missense mutations, variants were assessed using a selection of online matrices. A description of these is given below. Given that a single database could grade a variant incorrectly, these were used in combination with each other to prioritise variants.

2.17.1 BLOSUM62

The BLOSUM62 substitution matrix (Henikoff and Henikoff, 1992) (http://www.uky.edu/Classes/BIO/520/BIO520WWW/blosum62.htm) was developed from 2000 blocks of aligned sequence segments characterising more than 500 groups of related proteins. The authors calculated a log-odds score for each of the 210 possible substitution pairs of the 20 standard amino acids. The scores range from -4 to +3 for a non-synonymous amino acid substitution, with -4 meaning that the change is highly unlikely to be benign and a +3 that the change was highly likely to be benign.

2.17.2 Polyphen2

Polyphen2 (Polymorphism Phenotyping) (Adzhubei et al., 2010) (http://genetics.bwh.harvard.edu/pph2) uses eight sequence based and three structure based predictive features as part of its algorithm for determining variant pathogenicity. The programme compares features of the wild type and mutant protein and characterises how well the two human alleles fit into the pattern of multiple sequence alignment of homologous proteins. Two datasets have been used to develop Polyphen, HumDiv and HumVar. In the most recent version of Polyphen the user can choose which dataset can be used for analysis. The authors recommend the use of HumVar for distinguishing Mendelian mutations with drastic effects from normal human variation, and HumDiv for evaluating rare alleles in complex diseases. The output gives a score of between 0 and 1, 0 being Benign and 1 Damaging. A score of below 0.2 is scored as Benign, 0.2-0.85 as Possibly Damaging and above 0.85 is Probably Damaging.

2.17.3 SIFT

SIFT (Sorting Intolerant from Tolerant) (http://sift.jcvi.org) is a programme which uses sequence homology to predict whether an amino acid affects

protein function (Ng and Henikoff, 2001). Like Polyphen, it can only be used for the assessment of missense variants. SIFT differentiates variants which are “Tolerated” and “Damaging”. The score ranges from 0-1, <=0.05 predicted as damaging and >0.05 predicted as tolerated.

2.17.4 PROVEAN

PROVEAN (http://provean.jcvi.org/index.php) (Protein Variation Effect Analyser) is a software tool which predicts whether a variant has an impact on the biological function of the protein (Choi et al., 2012). It can be used to assess nonsynonymous single nucleotide variants and indels. A clustering of BLAST hits is performed with a parameter of global sequence identity. The top 30 clusters of closely related sequences form a supporting sequence set which is used to predict the effect of the variant. An alignment score is calculated for the variant and the scores averaged to generate a final PROVEAN score. If the PROVEAN score is equal to or below the threshold (the default threshold is -2.5), the protein variant is predicted to have a "deleterious" effect. If the PROVEAN score is above the threshold, the variant is predicted to have a "neutral" effect.

2.17.5 Mutation Taster

Mutation taster (http://www.mutationtaster.org) employs a Bayes Factor (BF) classifier to predict the disease potential of a variant (Schwarz et al., 2010b). Analyses consist of evolutionary conservation, splice-site changes, loss of protein features and changes that could affect mRNA levels. There are four possible outputs:

1. Disease-causing - i.e. probably deleterious
2. Disease-causing automatic - i.e. known to be deleterious
3. Polymorphism - i.e. probably harmless
4. Polymorphism automatic - i.e. known to be harmless

The output score for amino acid substitutions also gives a prediction probability value (ranging from 0-1). A value closer to 1 indicates a high certainty of the prediction. No numerical score is provided for amino acid insertions and deletions. Like PROVEAN, Mutation Taster has the ability to process indels.

2.17.6 Combined Annotation Dependent Depletion (CADD)

Many of the databases described above utilise a single type of information, for example conservation or protein- based metrics. CADD scores (Kircher et al., 2014) integrate diverse genome annotations and score any single nucleotide variant (SNV) or small Indel (Kircher et al., 2014). This can be used to grade nonsense, missense and splice site variants as well as those in intronic regions. A variant with a CADD score greater than or equal to 10 is predicted to be within the 10% most deleterious substitutions in the human genome. A score of greater or equal to 20 indicates the 1% most deleterious.

2.17.7 Splice prediction tools

The pathogenicity of potential splice site variants was assessed using Berkeley Drosophila Genome Project (BDGP) splice prediction site (http://www.fruitfly.org/seq_tools/splice.html). This is based on a neural network, which is trained to recognise splice donor or acceptor sites using a set of known sequences. The input sequence is compared with the training sets, and the probability of the input containing a splice site is calculated.

2.18 Assessment of Conservation

To assess whether the normal amino acid residue was evolutionarily conserved, protein sequences from different species were downloaded from the NCBI (http://www.ncbi.nlm.nih.gov) and aligned using ClustalW

(http://www.ebi.ac.uk/Tools/msa/clustalw2) (Larkin et al., 2007) or from NCBI Homologene (http://www.ncbi.nlm.nih.gov/homologene).

Clustal W is a multiple sequence alignment programme for DNA or proteins. Homologous DNA sequences from multiple species can be pasted into the document and are then aligned as closely as possible. This was used to align orthologous proteins of interest, allowing the conservation of a particular amino acid sequence to be determined.

NCBI Homologene database is an automated system for constructing putative homology groups from a wide range of eukaryotic species. The programme also aligns protein sequences of interest from different organisms and compares these with one another.

2.19 Tools to Aid Candidate Gene Prioritisation

Candidate genes were prioritized using several tools. Mouse Genome Informatics (MGI) (http://www.informatics.jax.org) is an international database of genetic and biological data from the laboratory mouse. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) (http://string-db.org) is a database of known and predicted protein interactions. Published human corneal endothelial cells RNA sequencing data (Chng et al., 2013) was also utilsed to prioritise candidates.

2.20 Public Databases

The 1000 Genomes (http://www.1000genomes.org), Exome Variant Server (EVS) (http://evs.gs.washington.edu/EVS) and Exome Aggregation Consortium (http://exac.broadinstitute.org) databases of samples that have been whole exome sequenced were utilized to exclude common polymorphisms prior to the comparison of missense changes in FECD cases versus controls. The frequency of 1000 Genomes SNPs was viewed in

SNPedia (http://www.snpedia.com/index.php/SNPedia).

2.21 Next Generation Sequencing (NGS) Bioinformatic Analysis Pipeline

Whole exome data was processed using two pipelines, summarised in Table 2.2. Briefly, reads were aligned to hg19 using Bowtie version 2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) (Langmead and Salzberg, 2012). The resulting SAM files were sorted and indexed using Samtools (http://samtools.sourceforge.net) and processed using the Genome Analysis Toolkit (GATK) (https://www.broadinstitute.org/gatk/) (McKenna et al., 2010) with Picard (http://broadinstitute.github.io/picard/) to perform indel realignment and duplicate removal. SNVs and indel variants were called using the UnifiedGenotyper feature of GATK. Variants were annotated using Annovar (http://annovar.openbioinformatics.org/en/latest/) (Wang et al., 2010).

Step	Programme	Website
Alignment	Bowtie2	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Removal of non-unique reads	N/A	
Sorting and indexing to the reference file	Samtools	(http://samtools.sourceforge.net)
Local realignment around indel: Create list of dubious alignments Realign over dubious regions	GATK	https://www.broadinstitute.org/gatk/
Removal of duplicate reads	Picard	http://broadinstitute.github.io/picard/
Variant Calling (SNPs and Indels)	GATK	as above
Recalibration of SNPs	GATK	as above
Apply recalibration of SNPs	GATK	as above
Combine list of recalibrated SNPs and Indels	GATK	as above
Annotation of variants	Annovar	http://annovar.openbioinformatics.org/en/latest/

Table 2.2 Bioinfomatic Pipeline 1 for the processing of NGS FASTQ files

A second pipeline was utilised which incorporated the use of CADD scores at the annotation stage (Table 2.3). For the second pipeline reads were aligned using Novolign (http://www.novocraft.com). SAM files were then sorted using Samtools and duplicates marked using Picard. Indel realignment and SNPs recalibrated took place using GATK. Following variant calling in GATK,

variants were then annotated using Variant Effect Predictor (VEP) (http://www.ensembl.org/info/docs/tools/vep/index.html).

Step	Programme	Website
Alignment	Novolign	http://www.novocraft.com
Sorting and indexing to the reference file	Samtools	(http://samtools.sourceforge.net)
Removal of duplicate reads	Picard	http://broadinstitute.github.io/picard/
Local realignment around indel: Create list of dubious alignments Realign over dubious regions	GATK	https://www.broadinstitute.org/gatk/
Recalibration of base quality scores	GATK	as above
Variant Calling	GATK	as above
Annotation of Variants	VEP	http://www.ensembl.org/info/docs/tools/vep/index.html

Table 2.3 Bioinfomatic Pipeline 2 for the processing of NGS FASTQ files

2.22 Filtering

Following annotation, Variant Call Format (VCF) files from individual affected family members were merged into a single file of shared variants using in-house perl scripts and the Agile Variant Selector software. For those families that had data from an unaffected individual, two files were created; one with just affected patients and one with affected patients excluding variants from the unaffected family member.

Homozygous variants were filtered out in the first instance, then all variants in known published FECD genes were analysed. Synonymous and intronic variants were then filtered out. Splice site variants in positions 1 to 5 were retained along with exons and exon/splicing (variants within the exon but close to the intron/exon boundary) variants. Variants that had a frequency of more than 2% on the EVS and 1000 Genomes databases were excluded. Variants were then prioritised on the basis of variant pathogenicity scores (Section 2.17, the RNAseq expression data and MGI (Section 2.19)). This process is summarised in Figure 2.3.

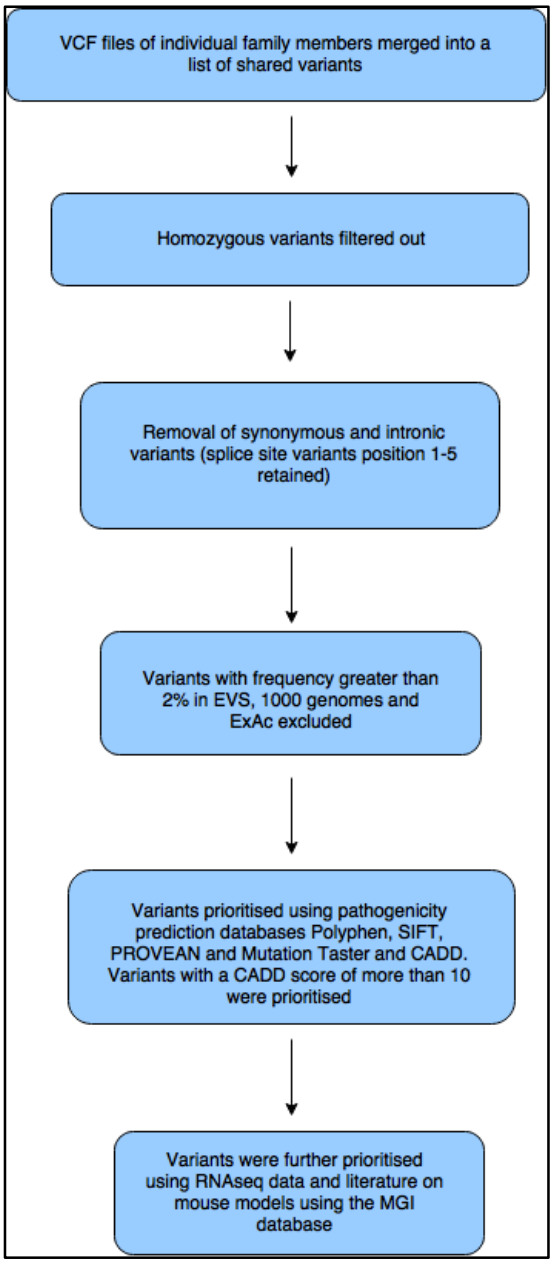


Figure 2.3 Flowchart summarising the process of filtering and prioritisation of variants generated from WES

2.23 CNV Analysis

2.23.1 ExomeDepth

The ExomeDepth software uses read depth to call CNVs from exome sequencing data (Plagnol et al., 2012). (https://cran.rproject.org/web/packages/ExomeDepth/index.html). It compares this data to an aggregate reference set generated from exomes run at the

same time on the same lane of the sequencer. The programme ranks the CNV calls using the BF which quantifies the statistical support for each CNV. This is the log10 of the likelihood ratio of data for the CNV called divided by the null (normal copy number). Therefore, the higher the BF value, the greater the confidence regarding the presence of a CNV. The type of CNV can be indicated by the ratio of the observed and expected reads:

0	homozygous deletion
0.5	heterozygous deletion
1.5	heterozygous duplication
2.0	homozygous duplication

The programme is able to identify and annotate common CNV as identified in the Conrad database (Conrad et al., 2010). FASTQ files for cases and controls were processed using Pipeline 1 described in Table 2.2, up to the variant-calling step. The resulting BAM files were run on the ExomeDepth programme by Evi Panagiotou.

2.24 Statistical analysis

PLINK 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007) was used to analyse the genotyped alleles and examine Hardy-Weinberg equilibrium (HWE) with Fischer's exact test but also to calculate odds ratios (OR) with standard error confidence limits (CI). This was carried out by Dr Jose Ivorra. The degree of linkage disequilibrium (r^2) between rs613872 and the CTG18.1 polymorphism was calculated using Haploview (Barrett et al., 2005).

3 A study of young-onset corneal dystrophies in conjunction with the British Ophthalmic Surveillance Unit (BOSU)

3.1 Introduction

Corneal dystrophies are a heterogeneous group of non-inflammatory diseases that are, in the majority of cases, limited to the cornea (Klintworth, 2009). They can occur as isolated cases but most are familial with a genetic basis and affected patients can manifest symptoms from birth, during childhood or in adult life. The traditional anatomic classification of dystrophies was thought to have limitations and in the 2015 International Corneal Dystrophy (IC3D) classification (Weiss et al., 2015), corneal dystrophies were reclassified as Epithelial and subepithelial, Epithelial-stromal *TGFβI*, Stromal and Endothelial Dystrophies. Most of these dystrophies follow a dominant inheritance pattern, but documented examples of autosomal recessive corneal dystrophies also exist (Akama et al., 2000, Vithana et al., 2006).

Although corneal dystrophies are rare, the commonest, Fuchs endothelial corneal dystrophy (FECD), is adult onset and accounts for 22% of corneal transplants performed in the UK (Keenan et al., 2012). FECD has been estimated to affect 4% of the US population over 40 (Lorenzetti et al., 1967). The prevalence of the remaining corneal dystrophies is essentially unknown and varies worldwide, with the highest numbers found in countries associated with a common founder mutation. For example, gelatinous drop-like corneal dystrophy (GDCD) is prevalent in Japan, where the incidence is estimated to be 1 in 300,000 people (Tsujikawa et al., 1998), whereas X-linked endothelial dystrophy has only been reported once in a multigenerational family from Western Austria (Schmid et al., 2006).

Most corneal dystrophies cause corneal opacity and result in visual impairment or even blindness (Nischal, 2015). Severe visual impairment (SVI)

caused by corneal opacity accounts for 2% of patients (Rahi and Cable, 2003). Corneal disease as a whole accounted for 45/1756 (2.6%) of Certificates of Visual Impairment (CVI) in working age adults in the UK in the year 2009/10 (Liew et al., 2014). Previous BOSU studies of rare ocular disease in the UK population (Papadopoulos et al., 2007), have suggested a higher prevalence in certain communities where consanguineous marriages are the cultural norm. Advances in gene mapping and the advent of next generation sequencing have, in the last decade, led to significant improvement in our ability to characterise the genotypes that are associated with these phenotypes at a molecular level. Given the rarity of many of these corneal dystrophies, there is great potential benefit to patients in studying them collaboratively. There is however a lack of studies examining the incidence of corneal dystrophies.

This chapter presents data that aims to advance our knowledge of the epidemiological aspects of young onset corneal dystrophies (presenting before the age of 40 years) by establishing an incidence rate in the UK, as there are currently no national studies in this field.

3.2 Results

The British Ophthalmological Surveillance Unit (BOSU) runs a nation-wide surveillance system across the United Kingdom for the epidemiological investigation of the incidence and clinical features of rare eye conditions of public health or scientific importance (https://www.rcophth.ac.uk/standards-publications-research/the-british-ophthalmological-surveillance-unit-bosu/).

The author, in collaboration with Mr Kamron Khan, set up a BOSU epidemiological survey for the reporting of young onset corneal dystrophies. In total 73 cases of new-onset corneal dystrophy were reported to the BOSU in response to this survey. The notifying clinicians were sent a questionnaire about the case, a sample of which is shown in Appendix 1. There were 31

respondents (a response rate of 42%). Four questionnaires were incompletely filled in, leaving a total of 27 questionnaires that were analysed.

3.2.1 Source of Patient Referral and Demographics

Patient referrals were from six different sources. 8/27 (29.6%) were referred by an Optometrist, 8/27 (29.6%) by their Ophthalmologist and 7/27 (25.9%) by their GP. 2/27 (7.4%) were reviewed due to an affected parent/relative, 1/27 (3.7%) as an acute referral for another presentation and 1/27 (3.7%) via orthoptic vision screening. The patients' ethnicities are summarised in Figure 3.1.

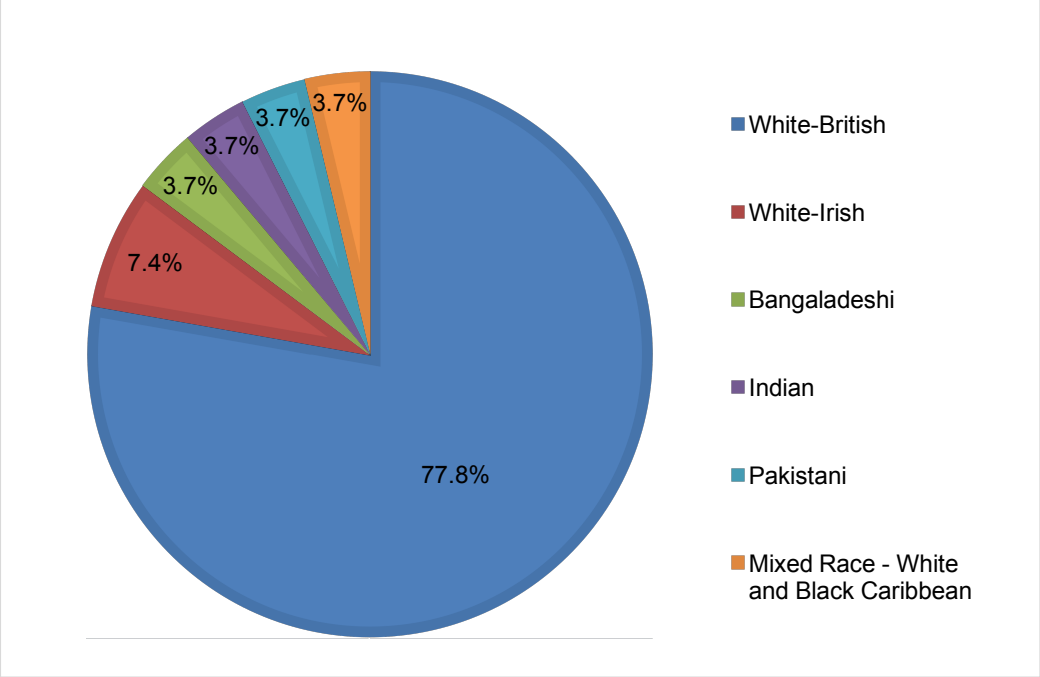


Figure 3.1 Pie chart indicating the ethnicity of young-onset corneal dystrophy patients.

The mean age at presentation was 12 years. 48% were female. There was a positive family history in 48% of cases. A documented history of consanguinity was found in 4% of cases, and in 26% of cases this was unknown.

3.2.2 Clinical Presentation

Best-corrected visual acuity (BCVA) was assessed using LogMAR visual acuity in 25.9% of patients. BCVA at presentation in all patients ranged from LogMAR 0.0 to Perception of light (PL), with a mean and median of 0.2. Corneal sensation was normal in 6 cases (22%) and not documented in 21 cases (78%). Symptoms at presentation are summarised in Table 3.1.

Symptom(s)	No. of Patients
None	4
Glare; surface irritation	2
Reduced photopic/scotopic VA	7
Reduced photopic/scotopic VA; glare	3
Reduced photopic/scotopic VA; glare; surface irritation	1
Reduced photopic/scotopic VA; surface irritation	1
Surface irritation	8
Not documented	1

Table 3.1 Symptoms of young-onset corneal dystrophy patients at presentation

The extent of corneal disease was found to be symmetrical in 81.4% of cases. Examples of dystrophies which were reported as having markedly unilateral signs included Meesman corneal dystrophy, LCD, RBCD, PPCD and Lisch Dystrophy. Some degree of corneal opacity was present in 96% of patients. In 52% of cases this involved the epithelial or subepithelial layers. Epitheliopathy was present bilaterally in 22%, unilaterally in 7% and not documented due to the patient's phobia of drops in 4%. Corneal oedema was present diffusely and bilaterally in 11% (3/27) and absent in 78% (21/27) of the patients. All cases which exhibited endothelial signs had a diagnosis of an endothelial dystrophy (either FECD or PPCD). Focal corneal oedema was present unilaterally in 7.5% (2/27) of cases.

3.2.3 Investigations, Genetic Testing and Diagnoses

The investigations utilised in all cases are outlined in Figure 3.2. In 74% (20/27) of cases slit lamp ophthalmoscopy was sufficient for diagnosis.

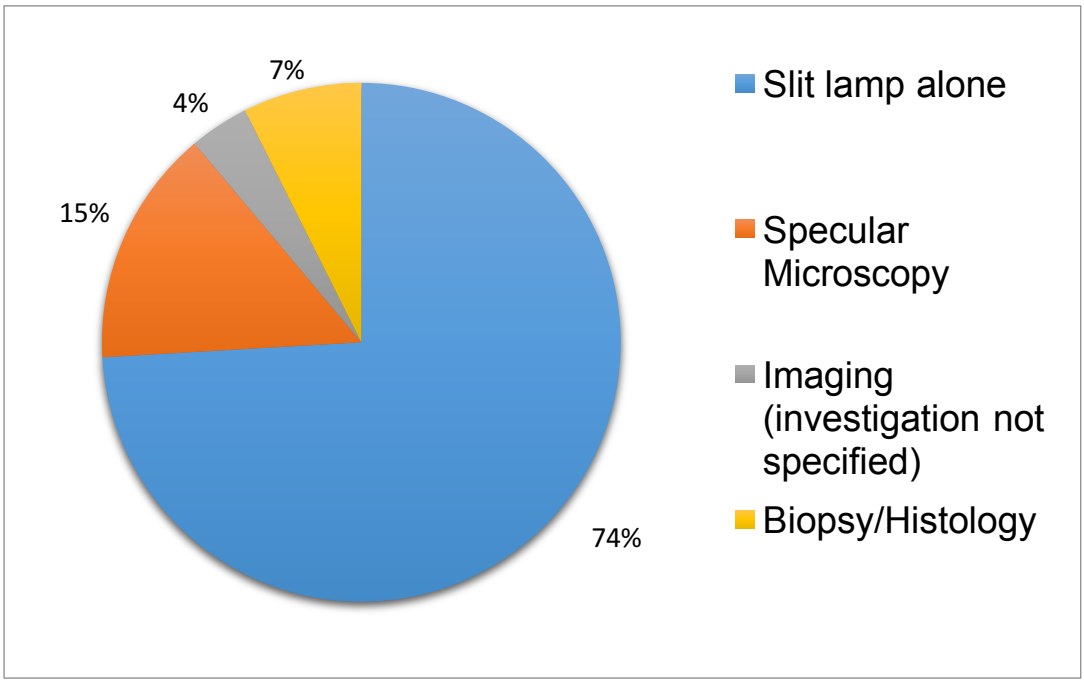


Figure 3.2 Pie chart indicating the investigations performed on newly-identified young-onset corneal dystrophy patients.

A summary of all diagnoses is given in Table 3.2. In two cases of stromal dystrophy a clinical diagnosis could not be made. In one of these cases whole exome next generation sequencing was requested. In the other case, a metabolic cause was suspected and a referral to medical genetics was made. Genetic tests were also requested for three other cases; one clinical case of Thiel-Bhenke dystrophy with affected parents which was confirmed on genetic testing, one case of posterior polymorphous dystrophy (PPCD) which was found to have a *ZEB1* mutation and an additional case of PPCD whose results were unknown.

Corneal Dystrophy Subtype	Diagnosis	No. of Patients
Epithelial and Subepithelial	Epithelial basement membrane dystrophy	4
	Meesman Dystrophy	2
	Lisch Dystrophy	1
Stromal	Macular Dystrophy	1
	Unknown	2
Epithelial-stromal TGFβ1	Reis-Bucklers Dystrophy	3
	Thiel-Bhenke Dystrophy	2
	Granular Dystrophy	2
	Lattice Dystrophy	1
Endothelial	Posterior Polymorphous Corneal Dystrophy	5
	Fuchs Endothelial Dystrophy	4

Table 3.2 Diagnoses of newly-identified young onset corneal dystrophy patients

3.2.4 Clinical Management and Follow-up

The most frequent form of management was observation with provision of artificial lubricants. Two cases required bandage contact lens use. Surgical intervention was required in three cases of endothelial dystrophy (Descemet's Stripping Endothelial Keratoplasty, Descemet's Stripping Automated Endothelial Keratoplasty and Penetrating Keratoplasty) (Figure 3.3).

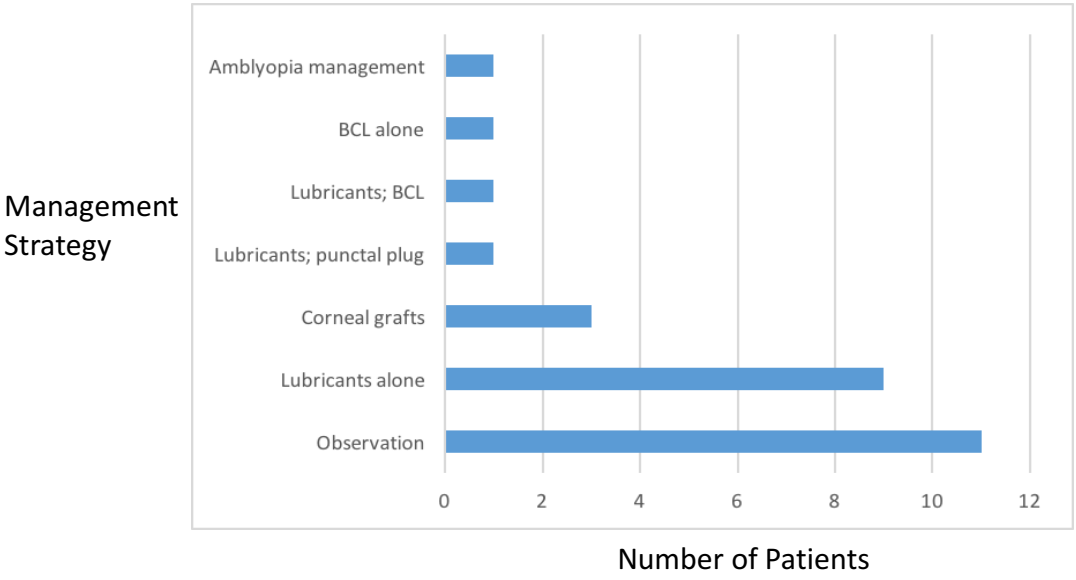


Figure 3.3 Management strategies utilised in young-onset corneal dystrophy patients. BCL = Bandage contact lens (a special type of contact lens used to protect the ocular surface).

85% of patients were followed up subsequently at six months in the presenting clinic, 7% were referred to another ophthalmologist with a specialist interest, 4% were referred to another clinic for another ophthalmological problem and 4% were discharged. One patient, who suffered from PPCD and had undergone penetrating keratoplasty had vision sufficiently reduced to be eligible for CVI.

Comparing the incidence of corneal dystrophy with the 2011 UK census data, young-onset corneal dystrophies are extremely rare and this study would suggest a minimum UK incidence of 6.7 newly-diagnosed cases per 10 000 000 population aged below 40 years per annum.

3.3 Discussion

Corneal dystrophies are mostly bilateral and are often inherited. There is however a lack of studies examining the incidence of corneal dystrophies in the UK. This chapter describes a study, conducted in conjunction with BOSU, aiming to assess the incidence of new onset corneal dystrophies in patients below 40 years of age. The results show that corneal dystrophies occur rarely in young people, thus highlighting the need to study them collaboratively. However, the incidence of 6.7 cases per 10 000 000 population aged below 40 years per annum may be an underestimate given that the questionnaire response rate was only 42%, which can be considered a limitation of this study. All of those clinicians who had not identified new cases but who did not respond with a questionnaire were reminded by telephone and email.

Another reason that the calculated incidence in this study might be an underestimate of the actual incidence may be a tendency to under-report mild phenotypes, especially those which may be seen more commonly in those below the age of 40 years. An example could be epithelial basement membrane dystrophy, a relatively common condition compared with many of the other corneal dystrophies reported in this chapter.

Corneal disease causes 2% of CVI in children (Rahi and Cable, 2003) and 2.6% in adults of working age (Liew et al., 2014). In this study there was one case (3.7%) of PPCD that presented with PL visual acuity and was eligible for CVI. However, two further patients exhibited moderate visual loss (Snellen 6/18 - 6/60, approximately 0.5 – 1.0 LogMAR). Furthermore, 22/27 (81.5%) experienced symptoms ranging from glare and surface irritation to reduced photopic or scotopic visual symptoms, suggesting that the burden of these conditions is considerable, even in the presence of good vision.

In the study presented in this chapter, 85.2% of the patients were Caucasian, which is comparable with the Census UK data from 2011 (11 December 2012). However, the Asian/British Asian group are a little overrepresented in

the group presented here, with 3/27 (11.1%) compared with 7.5% in the Census UK data. Two of the diagnoses in this Asian subgroup were macular and granular dystrophies but the diagnosis in the third case was unknown. Two of the three cases had a positive family history of corneal dystrophy, one of which also had a history of consanguinity.

Overall this study reports a positive family history in 50% of cases. Although hereditary factors are known to be important in the pathogenesis of these conditions, to our knowledge there are no UK studies examining this. However a study based in Saudi Arabia examining 193 corneal dystrophy cases involving the stroma indicated a positive family history of between 37.22 and 44.44% in the various dystrophies that they examined (Alzuhairy et al., 2015). This study also found that corneal histopathological examination of corneal buttons indicated that a proportion of Macular Corneal Dystrophy (MCD) cases had been misdiagnosed on clinical examination alone. In the study presented in this chapter there were only two cases that had been diagnosed with the assistance of corneal histopathology. One of these was a severe case of PPCD in which a penetrating keratoplasty was performed soon after presentation. Another was a patient with Lisch Dystrophy in whom a biopsy had been carried out on an affected first-degree relative. As observation and the provision of ocular lubricants were the most common treatments in this study, and surgical intervention was performed in only 11.1% of newly presenting cases, the fact that biopsy was not utilised to aid diagnosis more is perhaps not surprising. Despite advances in anterior segment imaging modalities (Hong et al., 2011), 74% of diagnoses were made using the slit-lamp examination alone, possibly reflecting a lack of availability or necessity of these modalities for corneal dystrophy diagnoses.

The mean age of onset was 12 years old, relatively young compared with that reported from another dataset (Musch et al., 2011) which used claims data to calculate the prevalence of corneal dystrophies in the US. Their study looked at the records of 8 million enrolees in a national managed care network. 27, 372 cases of corneal dystrophy were identified, with an overall prevalence rate 897 cases per million. The study presented in this thesis differs from that

reported by Musch et al. Even if endothelial dystrophy cases were to be excluded from their data, they identified the majority of their corneal dystrophy cases from the 45-88 age group. This BOSU study by its very nature studied examined a much younger cohort, but even taking this into account the mean age of onset of 12 years old can be considered relatively young. Additionally, their study looked at the prevalence of corneal dystrophies, whereas the results presented in this thesis looked at new cases.

Endothelial dystrophies were the most common group of dystrophies reported in this study, consistent with the US data previously described (Musch et al., 2011). Of these, PPCD was the most common dystrophy reported here. Of the three cases of PPCD that were managed by keratoplasty, two of these were lamellar endothelial keratoplasties. Endothelial keratoplasty is now performed more commonly than penetrating keratoplasty for endothelial failure in the UK (Keenan et al., 2012). As PPCD is documented to have good outcomes following keratoplasty (Nischal, 2015), this is perhaps a reason why this group of cases were more likely to undergo such surgical management than any of the other options. Interestingly, despite the known role of LASIK and surface ablation in the management of a variety of corneal dystrophies (Woreta et al., 2015), there were no cases of corneal dystrophies managed by excimer laser in this study. Corneal sensation was not documented in 78% of cases. Given that corneal sensation is known to be subnormal in some subtypes of corneal dystrophy (Rosenberg et al., 2001, Ahuja et al., 2012) and indeed might have implications for tear production and consequent dry eye, we would recommend testing corneal sensation as part of the assessment of corneal dystrophy.

Genetic testing was carried out in 4/27 cases, two of which carried a diagnosis of PPCD. In one of these PPCD cases, whole exome sequencing (WES) revealed a mutation in *ZEB1*, while in the other the result is unknown. WES was carried out in a third case, which was an undetermined stromal dystrophy. In one case of Thiel-Bhenke dystrophy, genetic screening was carried out on an affected first-degree relative. In another case of stromal dystrophy of uncertain diagnosis, a referral to the regional medical genetics

department was made for further investigations. The use of Next Generation sequencing techniques in the molecular characterisation of corneal dystrophies has been a major advance in recent years. This has been reflected in the designation of the *TGFβ1* corneal dystrophies, a category based on the molecular diagnosis of corneal dystrophies rather than the anatomical location of pathology, and it has been suggested that the existence of a new corneal dystrophy must start with identification of the clinical phenotype and culminate in characterisation of the causative gene mutation (Weiss et al., 2015). As our understanding of molecular genetics advances and pressure on traditional therapeutic options, such as corneal graft material, remains high, it may become necessary to look for alternative treatment options. The cornea is a highly accessible structure which is potentially highly advantageous when directing treatments to the target tissue. Numerous efficacious vectors, delivery techniques, and approaches have evolved in the last decade (Mohan et al., 2013), but for corneal dystrophy, these treatments have not yet found their way into clinical practice. It is not inconceivable that the identification of causative mutations might not only guide genetic counselling of prognosis and recurrence risks but in the future pave the way to the development of gene therapies.

In conclusion, corneal dystrophies are a hereditary group of corneal conditions which are extremely rare, with a minimum UK incidence of 6.7 newly-diagnosed cases per 10 000 000 population aged below 40 years per annum.

4 Genotype-Phenotype Correlations in CHED, Harboyan Syndrome and FECD

4.1 Introduction

CHED is an autosomal recessive condition, which is a cause of congenital corneal opacity. It is a rare, bilateral disease affecting the posterior aspect of the cornea, characterised by corneal oedema, thickening of the DM layer and lack of endothelial cells (Maumenee, 1960, Kirkness et al., 1987). Presentation is at birth or in the early neonatal period. FECD is a complex late-onset condition, which begins with asymptomatic corneal endothelial guttata. Its inherited forms are autosomal dominant (Krachmer et al., 2011).

Recessive mutations in *SLC4A11* (MIM*610206) on chromosome 20p13 cause CHED (formerly CHED2 see Section 1.3.6) (Vithana et al., 2006). Harboyan syndrome (MIM#217400) is characterised by CHED in association with sensorineural hearing loss. Dominant mutations in *SLC4A11* are also a rare cause of late-onset FECD (Vithana et al., 2008). The hearing deficit in Harboyan syndrome is not present at birth but is typically progressive with onset around the age of 10 to 15 years (Desir et al., 2007). Recessive mutations in *SLC4A11* also account for this phenotype. There have been over 70 different homozygous or compound heterozygous mutations reported in the *SLC4A11* gene. (Aldave et al., 2013) The mutations reported suggest that the three disorders are allelic (Desir and Abramowicz, 2008, Mehta et al., 2010) but this has not previously been demonstrated by longitudinal follow-up of the CHED patients or conclusively documented in the literature.

This chapter describes two previously reported CHED families (Vithana et al., 2006) and a newly identified family (courtesy of a collaboration with Dr. Juan Carlos Zenteno (JCZ), National Autonomous University Mexico, Mexico City) to investigate whether CHED and Harboyan syndrome are the same condition

at different stages of development, and whether the parents of CHED patients are at risk of developing late-onset FECD .

4.2 Results

4.2.1 Identification and clinical evaluation of the affected CHED cases

Three families and their relatives were identified as described below. Families A and B were ascertained for a previous study (Vithana et al., 2006) and their phenotype and mutation are reported by Vithana et al 2006. Family C was ascertained and the mutation identified in this study.

Clinical examination of the patients with endothelial dystrophy suggested a diagnosis of CHED, while family history suggested recessive inheritance. The pedigrees are shown in Figure 4.1.

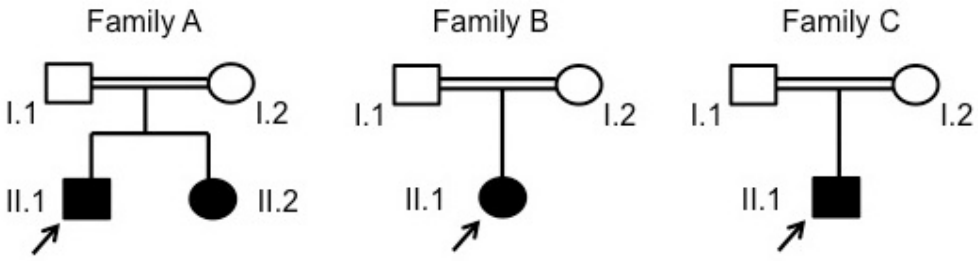


Figure 4.1. Family structure of patients with CHED. Families A, B and C are shown. The proband is indicated by an arrow. Data from Family C was provided courtesy of JCZ.

Family A was of Pakistani origin and was identified at Bradford Royal Infirmary (BRI). Clinical information from both affected individuals from Family A was derived from the patient records and outpatient follow-up appointments at the BRI Department of Ophthalmology. Patient II.1 presented initially during the immediate postnatal period with bilateral cloudy corneas. There was no evidence of buphthalmos and his mucopolysaccharide screen was negative. Examination under anaesthesia at 1 year of age revealed normal corneal diameters and signs consistent with CHED. His anterior chambers were deep

and his intraocular pressures (IOP) in both eyes were 20mm Hg. At age 5, he was registered partially sighted, with visual acuities of 0.60 and 0.70 logMAR in the right and left eye respectively and attended school with support from a visual impairment nurse. At age 24, his visual acuities had remained stable since partial sight registration. His corneal photograph was taken at this age (Figure 4.2).

His younger sister, individual II.2, presented shortly after birth due to a cloudy cornea. At this time the cause of her brother's cloudy cornea was under investigation. At 5 months of age, an examination under anaesthesia revealed bilateral cloudy corneas 11mm in diameter with an extremely poor fundal view. At age 2, her mother noted that she held objects very close to her in order to localise them and felt that her vision was significantly worse than her brothers'. Her vision was measured as 2.22 logMAR in both eyes. She underwent a right penetrating keratoplasty at age 3, but her recovery was hindered initially by a blunt trauma one month post-operatively. However after 3 months she maintained a clear graft. Her left eye underwent penetrating keratoplasty at age 4 but was later complicated by rejection and graft failure, and was subsequently re-grafted at age 9 and 19. At age 21, she had a clear graft in the right eye, but her vision in this eye has remained at 1.00 logMAR due to initial amblyopia and high astigmatism following penetrating keratoplasty. Her vision in her left eye was hand movements (HM).

Family B was also of Pakistani origin, identified at BRI. Clinical information was ascertained from the patient records and Ophthalmology outpatient follow-up appointments at the BRI. Individual II.1 was noted after birth as having a cloudy cornea. At 5 months of age, she underwent an examination under anaesthesia, which revealed generalised stromal oedema and corneal diameters of 11mm. Her IOPs were raised at 30mmHg and 40mmHg in the right and left eyes respectively. She underwent a goniotomy procedure and was commenced on Betagan BD to both eyes to reduce her intraocular pressure. Her other diagnoses included moderate learning difficulties and microcephaly. At age 16 years old, she maintained stable visual acuities of 1.00 and 1.07 logMAR in the right and left eyes respectively, and her

glaucoma was well controlled. She received continued support from a visual impairment teacher at school.

Family C were of Mexican origin, identified and recruited by collaborator JCZ. The patient presented at 5 months of age with bilateral corneal opacities that had been observed at birth. At 6 months of age, the patient received Timolol and Dorzolamide eye drops due to high IOP. At this age, a constant horizontal nystagmus was noted in both eyes. He underwent goniotomy due to elevated intraocular pressure in the left eye at the age of 3 years. At 10 years, he presented with congenital corneal clouding, nystagmus and diffuse corneal oedema in both eyes, normal intraocular pressure, no Haab striae, no buphthalmos and optic nerve head cupping of 0.7. At age 13, he had vision 1.82 and 1.60 logMAR in his right and left eye respectively. His corneal photograph was taken at this age (Figure 4.2).

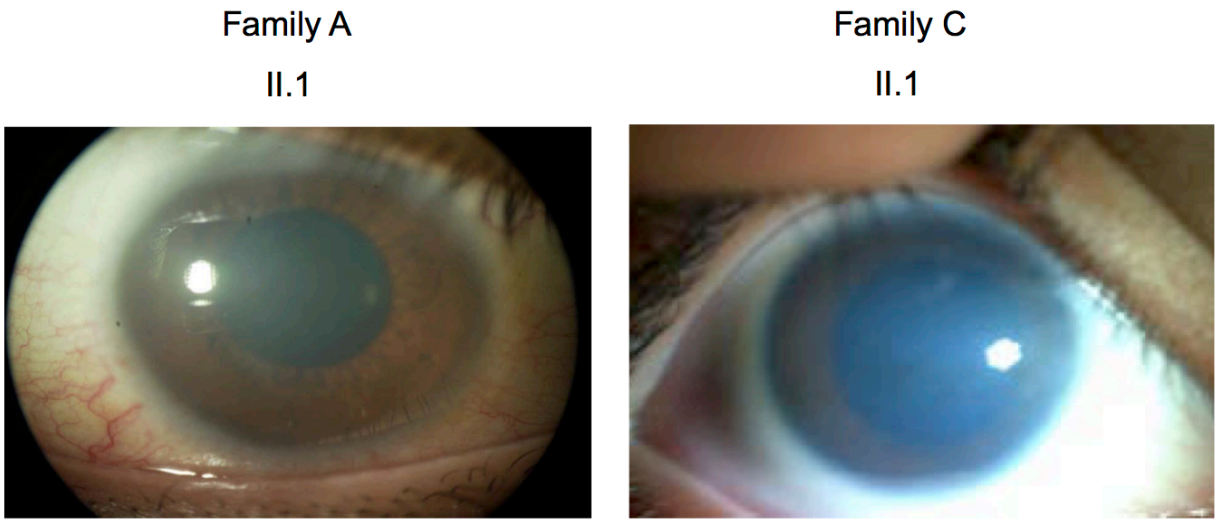


Figure 4.2 Corneal photographs of affected patients from Families A and C. Individual II.1 from Family A was 24 years old and individual II.1 from Family C was 13 years old at the time of these photographs. The typical ground-glass corneal appearance is apparent in both photographs, but more prominent in the individual from Family C.

The clinical findings are summarised in Table 4.1.

PATIENT	AGE AT EXAMINATION (years)	CLINICAL FINDINGS
Family A, II.1	Post-natal	cloudy corneas
	1	corneal diameters 10mm bilaterally, deep anterior chamber, IOP 20mmHg both eyes
	5	VA 0.60 & 0.70 logMar for OD & OS
	24	VA stable
Family A, II.2	Post-natal	cloudy corneas
	0.5	corneal diameters 11mm bilaterally, poor fundal view
	2	VA 2.22 logMar both eyes (corneal graft surgery at 3 & 4 years for OD & OS respectively)
Family B, II.1	Post-natal	cloudy corneas
	0.5	corneal diameters 11mm bilaterally, generalised oedema, IOP 30 & 20mmHg for OD & OS (underwent goniotomy)
	16	VA stable at 1.00 & 1.07 logMar for OD & OS
Family C, II.1	0.5	cloudy corneas
	0.6	bilateral nystagmus, high IOP (received eye drops & underwent goniotomy at 3 years for OS)
	10	diffuse corneal oedema, ONH cupping 0.7
	13	VA 1.82 & 1.60 logMar for OD & OS

Table 4.1 Summary of the ophthalmic findings from Families A, B and C. Age at examination and clinical findings for each CHED patient are shown. Any ophthalmic interventions are also highlighted. IOP, intraocular pressure OD, right eye; ONH, optic nerve head; OS, left eye; VA, visual acuity

4.2.2 Audiometric testing in the CHED affected cases

As CHED is associated with progressive sensorineural hearing loss (Desir and Abramowicz, 2008), hearing assessments of CHED patients from Families A-C (Figure 4.3a) were carried out. A normal audiogram is shown for comparison (Figure 4.3b).

Audiometric examination of patient II.1 from Family A at 12 years of age carried out at the BRI revealed bilateral mid to high frequency sensorineural hearing loss (Figure 4.3a). He was subsequently fitted with bilateral hearing aids. Audiometry was repeated at age 21, where some deterioration of hearing was measured. His younger affected sibling (patient II.2), whose corneal opacity was more marked at presentation, had subjective hearing problems at the age of 21. When asked about her auditory symptoms, she expressed difficulty hearing others' conversations when studying at college. Audiometry was therefore performed at the BRI and showed a mid-frequency bilateral sensorineural hearing loss consistent with these symptoms (Figure 4.3).

Audiometry in the patient from Family B, at age 11 was unremarkable. This was repeated at age 15 where bilateral high frequency hearing loss was found (Figure 4.3a).

Audiometric testing in the patient from Family C at the age of 12, disclosed bilateral sensorineural hearing loss in the range of 30 to 60 decibels, mainly affecting the higher frequencies (Figure 4.3a).

These results confirmed that the CHED patients presented in this thesis go on to develop sensorineural hearing loss. Thus their diagnosis changed from CHED to Harboyan syndrome on longitudinal follow-up. It could therefore be concluded that CHED and Harboyan Syndrome are the same condition at different stages of development.

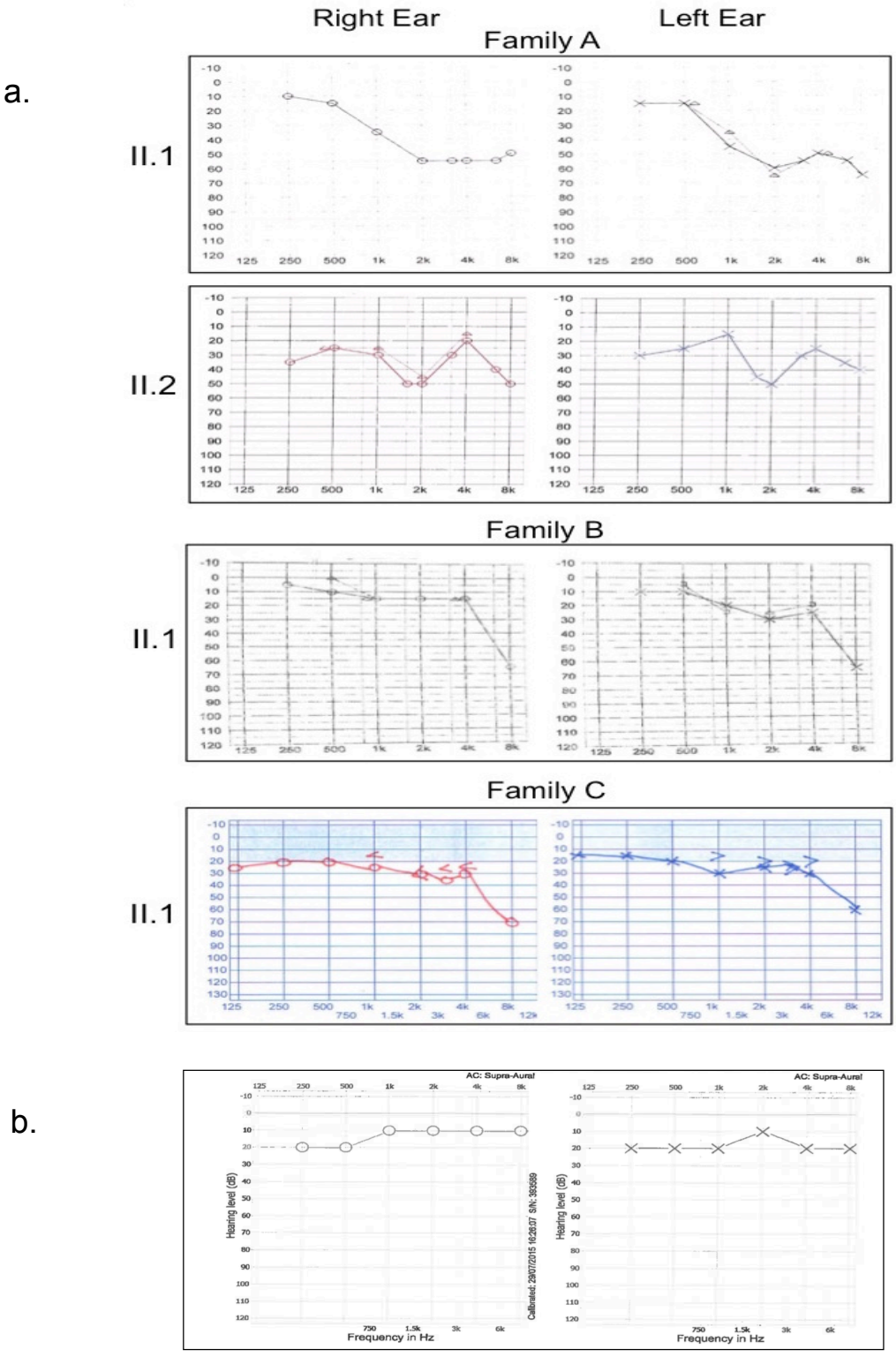


Figure 4.3 Audiometry of the CHED affected patients from Families A, B and C. a. The graphs show the frequency in hertz (x-axis) and the hearing level in decibels (y-axis). The test was performed on the patient at 12 years (Family A, II.1), 21 years (Family A, II.2), 15 years (Family B, II.1) and 12 years (Family C, II.1) respectively. Note reduced sensorineural hearing loss in the range 30 to 50 decibels at the higher frequencies. b. A normal audiogram is shown for comparison (courtesy of Mr Glen Waugh, Associate Audiologist at Bradford Royal Infirmary, Bradford.)

4.2.3 *SLC4A11* mutation screening to confirm clinical diagnosis of CHED

Peripheral blood samples were collected from the affected patient from Family C. Genomic DNA was extracted using the standard protocol described in Section 2.8. The sample from the affected patient, along with a control DNA, were screened by PCR and Sanger sequencing (Section 2.10 and 2.14). All 19 exons of *SLC4A11* transcript NM_032034 were screened, the primer pairs for which are listed in Appendix III. Sequencing of *SLC4A11* gene in the patient from Family C identified the presence of a homozygous novel mutation c.397T->C, p.F133L (Figure 4.4) contained in exon 4. To exclude this variant as a polymorphism, the EVS and ExAC databases were checked. The EVS contains the whole exome data of 6500 individuals, and variants in the *SLC4A11* gene were found in 175 Europeans and Americans. The ExAc database contains the whole exome of 60,706 individuals. This variant was not present in either database, indicating that it is a rare variant.

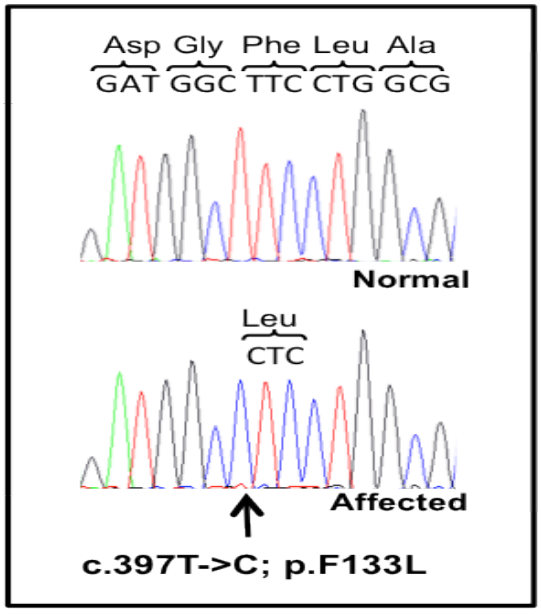


Figure 4.4. Sequence chromatogram from Family C, Individual II.1. This depicts the novel mutation in *SLC4A11* identified in the Family C Mexican CHED patient.

Where mutations have been shown to cause disease, there is often a remarkable degree of evolutionary conservation in the corresponding genes across species (Strachan and Read, 2011). In order to assess the evolutionary conservation of this phenylalanine residue, multiple sequence alignment of SLC4A11 proteins was performed (Section 2.18) and is shown in Figure 4.5. This indicates that the amino acid residue Phenylalanine in the position 133 on the SLC4A11 protein is highly conserved through evolution. The pathogenicity assessment of the variant using Polyphen2, SIFT, PROVEAN, Mutation Taster and CADD (Section 2.17) are shown in Table 4.2. All of these assessments except for SIFT indicated that this variant was likely to be pathogenic. The SLC4A11 protein BTR1 (bicarbonate transporter related protein 1) consists of cytoplasmic, transmembrane and extracellular domains (Vilas et al., 2011). This variant lies in the cytosolic domain.

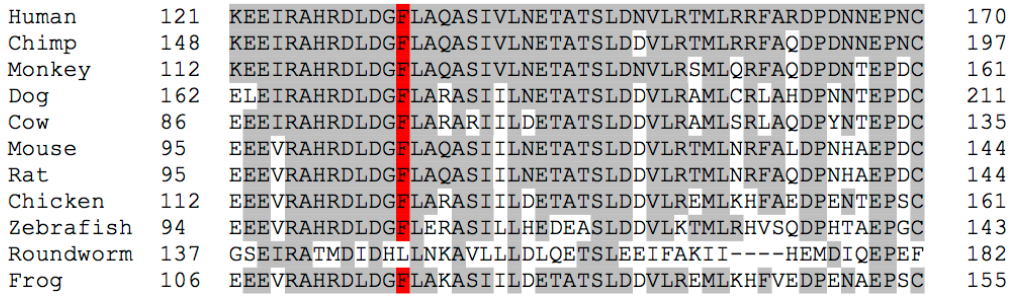


Figure 4.5 Protein sequence alignment of the human SLC4A11 protein with orthologues around the phenylalanine residue was performed using Homolgene. The variant p.F133 is highlighted in red and the surrounding amino acids that are identical to the human transcript are shaded in grey. Accession numbers for SLC4A11 sequence are NP_114423 (human), XP_001160838.1 (chimp), XP_002798266.1 (monkey), XP_005634909.1 (dog), NP_001178243.1 (cow), NP_001074631.1 (mouse), NP_001101245.1 (rat), XP_004936342.1 (chicken), NP_001153300.1 (zebrafish), NP_001033333.1 (roundworm), XP_002936409.2 (frog). Note the p.F133 residue is evolutionarily conserved suggesting an important role of this residue in the normal function or structure of SLC4A11 NP_114423.

Variant	Polyphen 2	SIFT	PROVEAN	Mutation Taster	CADD*
c.397T>C; p.F133L	Possibly damaging	Tolerated	Deleterious	Disease-causing	Scaled C-score
	(Score 0.479)	(Score 0.37)	(Score -3.328)	(Prediction probability 0.9999)	23.6

Table 4.2 Summary of bioinformatics analyses used to determine the likely pathogenicity of the c397T>C mutation in *SLC4A11.* The databases used were Polyphen2, http://genetics.bwh.harvard.edu/pph2/, (Adzhubei et al., 2010), SIFT, http://sift.jcvi.org (Ng and Henikoff, 2001), PROVEAN, http://provean.jcvi.org/index.php (Choi et al., 2012), Mutation Taster http://www.mutationtaster.org (Schwarz et al., 2010b) and CADD http://cadd.gs.washington.edu (Kircher et al., 2014). *CADD scores are reported as scaled C-scores and values ≥ 20 and ≥ 10 respectively represent the 1% and 10% most deleterious changes predicted in the human genome.

All these analyses suggest that the F133L variant identified in Family C is pathogenic and is the cause of the CHED/Harboyan phenotype.

4.2.4 Clinical examination and specular microscopy of the Parents of CHED affected cases

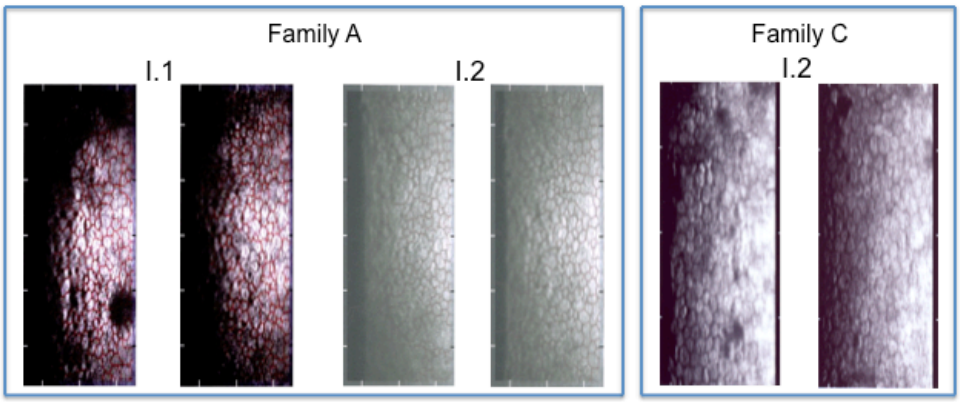
Homozygous mutations in *SLC4A11* cause CHED (Vithana et al., 2006), whereas heterozygous mutations in *SLC4A11* are a rare cause of dominantly-inherited FECD (Vithana et al., 2008). Given this, it was considered whether the parents of affected CHED patients, themselves heterozygous carriers for *SLC4A11* mutations could be at risk of developing FECD. Therefore, the parents of each CHED patient, where available, were clinically evaluated for FECD.

For each family the parents, who were related by consanguinity, did not manifest obvious visual problems. They underwent clinical examination followed by specular microscopy (SM). SM of the parents in Family A showed guttata, tiny excrescences in DM which are a hallmark of early FECD, in the 44 year old father, whereas the 46-year-old mother's scan showed moderate pleomorphism (disruption of the normal endothelial hexagonal pattern which may be seen in the initial stages of FECD before guttata are clearly visible) with a normal cell count (Figure 4.6).

SM examination of the Family B mother at age 38 showed unremarkable images with a normal endothelial cell count. The Family B father aged 40 was unavailable for examination. Specular microscopic examination of the mother in Family C, aged 37 years, revealed guttata but a normal cell count (Figures 1B and C). The father, aged 42 years, was unavailable for clinical examination.

These results suggest that the heterozygous parents of CHED and Harboyan patients go on to develop FECD.

(A)



(B)

	Family A, I.1 OD	Family A, I.1 OS	Family A, I.2 OD	Family A, I.2 OS	Family C, I.2 OD	Family C, I.2 OS
Cell Density (mm^2)	2656	2154	2730	2667	2967	2985
Coefficient of Variation	0.57	0.65	0.47	0.44	0.27	0.37
% Hexagonality	41	62	43	22	51	58

Figure 4.6 Corneal endothelium analysis. A, Specular microscopy images of the father (I.1) and mother (I.2) of family A and also the mother (I.2) of family C are shown. B, The table highlights the values for cell density (per square millimeter), coefficient of variation, and percentage hexagonality for the specular microscopic images. As a guideline, cell density of 2500 cells per square millimeter at middle age is within the normal range, and this value decreases with age. A coefficient of variation greater than 0.4 and less than 50% hexagonality are indicative of an abnormal endothelium.

4.3 Discussion

In the results reported in this chapter, two previously described local families with CHED affected cases (Vithana et al., 2006) (Families A and B) were revisited and one new family (Family C) was recruited by Dr Juan Carlos Zenteno, Mexico City. Sanger sequencing of the *SLC4A11* gene in the patients confirmed a novel homozygous missense mutation c.397T>C; p.F133L in Family C. Mutations in families A and B had been reported previously (Vithana et al., 2006).

Four patients originally diagnosed with CHED were subsequently found to have varying degrees of sensorineural hearing loss at the higher frequency range following audiometric examination, suggesting that CHED often progresses to Harboyan syndrome. The heterozygous mutation-carrying parents of these patients were also investigated for early signs of FECD. Two out of the four parents that were available to be examined had guttata in their endothelium. Interestingly this additive pattern of inheritance for a mutation in *SLC4A11* is somewhat similar to *LRP5,* in which dominant mutations and recessive mutations cause different eye phenotypes. Dominant mutations cause Familial Exudative Vitreoretinopathy (FEVR) and osteopenia or mild osteoporosis (Toomes et al., 2004) whereas recessive mutations cause OPPG (early onset retinal dysplasia and severe osteoporosis) (Gong et al., 2001). Additionally, this is not the first time that mutations in a gene causing corneal endothelial disease (Riazuddin et al., 2012) have also been shown to cause deafness (Grillet et al., 2009).

The basis for the phenotypic heterogeneity seen in homozygous *SLC4A11* mutation carriers with nonsyndromic CHED or Harboyan syndrome has been the subject of much speculation. The onset of progressive hearing loss in Harboyan syndrome has been described previously in children as young as 2 years and adults as old as 33 years (Desir et al., 2007, Mehta et al., 2010). However auditory abnormalities have not been directly tested or serially monitored in these cases of isolated CHED (Vithana et al., 2006, Jiao et al.,

2007, Kumar et al., 2007, Ramprasad et al., 2007, Aldave et al., 2007, Sultana et al., 2007, Hemadevi et al., 2008, Paliwal et al., 2010), suggesting that Harboyan syndrome may have gone undetected. As the type and location of null and missense *SLC4A11* mutations identified in both conditions is similar, with no obvious clustering (Desir and Abramowicz, 2008), and the co-existence of the conditions CHED and Harboyan syndrome within a family has been reported (Mehta et al., 2010), there is little evidence to support a genetic basis for the difference between them. Instead, the longitudinal study presented in this thesis suggests that CHED cases eventually experience some degree of sensorineural hearing loss and that the variable age of onset of these symptoms may be related to some unknown differences in the expression of genetic modifiers or exposure to environmental triggers.

The auditory phenotype seen in Harboyan syndrome is consistent with the observation that *SLC4A11* is not only expressed in the corneal endothelial cells but also in the fibrocytes of the stria vascularis in the inner ear (Lopez et al., 2009, Groger et al., 2010), cells with a common embryonic origin in the neural crest. SLC4A11 exists as a transmembrane homodimer (Park et al., 2004) and transports sodium ions coupled to hydroxide ions regulating pH in the corneal endothelium (Ogando et al., 2013, Jalimarada et al., 2013, Kao et al., 2015). It has been postulated that defective SLC4A11 expression could lead to compromised pH regulation affecting bicarbonate and lactic acid transport and depressing endothelial pump function (Jalimarada et al., 2013). Consistent with this theory is the finding that its absence in knockout mice causes accumulation of sodium chloride in the corneal stroma, collection of fluid in the normally relatively dehydrated cornea and morphological changes in fibrocytes resulting in deafness (Lopez et al., 2009, Groger et al., 2010). Examination of cells transfected with mutant SLC4A11 constructs shows that the mutant protein fails to glycosylate and is retained intracellularly, never reaching the cell surface (Vithana et al., 2006, Vithana et al., 2008).

Dominant mutations in *SLC4A11* cause FECD (Vithana et al., 2008) which in most cases are caused by missense changes. Cell based biochemical assays using SLC4A11 constructs appear to distinguish between the mutations that

cause FECD and those that cause CHED (Vilas et al., 2012a). Co-expression of mutant *SLC4A11* with the wild type construct causes partial rescue of most CHED-causing mutants but not those implicated in FECD. This is thought to be because, while most *SLC4A11* mutations do not affect cell surface processing of the wild type *SLC4A11*, presence of the FECD mutant protein reduces wild-type processing at the cell surface, suggesting a possible explanation for the dominant inheritance pattern for this disorder. However, the work presented in this thesis highlights that parents of CHED patients, carrying heterozygous missense mutations in *SLC4A11,* may go on to develop guttata as early signs of FECD onset, which might contradict the dominant negative mechanism outlined. Serial audiometric monitoring of the CHED parents would perhaps be a more helpful additional investigation in determining the genetic mechanism underlying these phenotypes. Not only would this aid the early detection of a potential hearing deficit, if heterozygous mutations do cause haploinsufficiency of the *SLC4A11* gene, then some degree of sensorineural hearing loss might be evident in the CHED parents.

Given that CHED (Vithana et al., 2008) and Harboyan syndrome (Desir et al., 2007) are caused by recessive mutations in *SLC4A11* and some FECD cases are caused by dominant mutations in *SLC4A11*, it is perhaps surprising that these conditions had never been described within the same family at the time the study described in this thesis was performed. However, the increased susceptibility of the parents of CHED patients in developing FECD has since been described in a non-consanguineous family (Kim et al., 2015b). The proband's 64-year-old father and sister were unaffected, but the asymptomatic 62-year-old mother was found to have bilateral guttata. When screened, the proband possessed a novel homozygous missense mutation c.1158C>A; p.Cys386* on exon 9 of *SLC4A11*, a mutation also harboured as a heterozygousvariant by the proband's father, mother and sister. This caused a premature stop codon, predicted to result in a truncated SLC4A11 protein. At 62-years-old the mother presented with FECD. This is consistent with the observation that presentation usually occurs in affected patients in the fifth or sixth decade (Weiss et al., 2015). It is entirely feasible that the heterozygous parents of Family B, aged 38 and 40 years-old, were not yet old enough to

present with the clinical features of FECD. This raises the question of the value of future screening for FECD in the parents of CHED affected cases. Currently pre-symptomatic screening for FECD would not alter the treatment provided for FECD, although detection of the disease prior to the advanced stages of corneal epithelial involvement (Stage 5, Table 1.4 Introduction) might avoid the need to perform a full-thickness graft. Therefore there is arguably merit in vision screening of the parents of CHED patients in the community by their optometrist or screening of the corneal endothelium for guttata by a local ophthalmologist.

Pathogenic variants tend to have markedly higher conservation than benign variants (Cooper et al., 2010) and indeed many of the variant pathogenicity prediction databases used for the assessment of missense variants incorporate conservation assessment into their algorithms (Adzhubei et al., 2010, Ng and Henikoff, 2001, Schwarz et al., 2010b, Kircher et al., 2014). However predicting the pathogenicity of a variant is still fraught with challenges. The variant was found to be “Tolerated” in SIFT, yet three of the other pathogenicity prediction sites indicated that the variant was pathogenic (Table 4.2). The variant’s CADD score was 23.6. Scores of ≥ 20 are thought to represent the 1% most deleterious changes predicted in the human genome. However, there are many non-pathogenic variants in the human genome with CADD scores of above 20. Using a combination of pathogenicity databases and comparing the outputs is therefore a reasonable way of approaching this challenge.

The 2015 International Corneal Dystrophy (IC3D) classification aimed to address much of the confusion surrounding some corneal dystrophy classification, and it has been suggested that the existence of a new corneal dystrophy must start with identification of the clinical phenotype and culminate in characterisation of the causative gene mutation (Weiss et al., 2015). This has been reflected in the designation of the *TGFβI* corneal dystrophies, a category based on the molecular diagnosis of corneal dystrophies rather than the anatomical location. In this chapter, it has been concluded that the

unaffected heterozygous parents or siblings of CHED patients are at risk of developing late-onset FECD, individuals who would otherwise be regarded as clinically normal. The work presented in this chapter highlights the impact that an accurate molecular diagnosis can have on predicting the clinical prognosis. It also indicates that the pathways involved in the pathogenesis of two different endothelial dystrophies may share some similarities.

CHED patients from Family B and C were noted to have raised IOP on initial examination under anaesthesia. In both patients this was in the absence of a previous keratoplasty, which is itself a risk factor for glaucoma (Kirkness and Moshegov, 1988). There are however no reports of the expression of *SLC4A11* in the trabecular meshwork and aqueous outflow pathways (Patel and Parker, 2015). Embryologically, congenital glaucoma is a result of abnormal neural crest cell migration, whereas CHED results from abnormal cell differentiation (Bahn et al., 1984). It is therefore feasible that the two might co-exist and indeed this has been reported (Ramamurthy et al., 2007). In the absence of molecular confirmation, if faced with the diagnosis of glaucoma in CHED, a careful examination would need to be performed in order to exclude a diagnosis of PPCD, which is more commonly diagnosed with glaucoma (Weiss et al., 2015).

The results presented in this chapter describe a novel mutation in *SLC4A11*. Discovering new mutations, especially missense ones, is of scientific interest as it informs future studies on the function of the protein. If a single amino acid change on a large protein disrupts protein function, this indicates that the amino acid performs a vital function. There is considerable allelic heterogeneity in CHED (Aldave et al., 2007) and no obvious correlation between the location of *SLC4A11* mutations, the consequent domain of the SLC4A11 protein affected (Vilas et al., 2011) and the phenotype of CHED, Harboyan syndrome or FECD have been noted previously (Figure 4.7). However, one recent study has suggested that the cytoplasmic domain may play an essential role in the transport function of SLC4A11 (Loganathan et al., 2016). Nonetheless, one implication a lack of clear mutation hotspots could be that, when ascertaining a molecular diagnosis in a newly presenting CHED

patient, it is likely that the entire coding region of the gene would need to be sequenced in order to find the causative mutation.

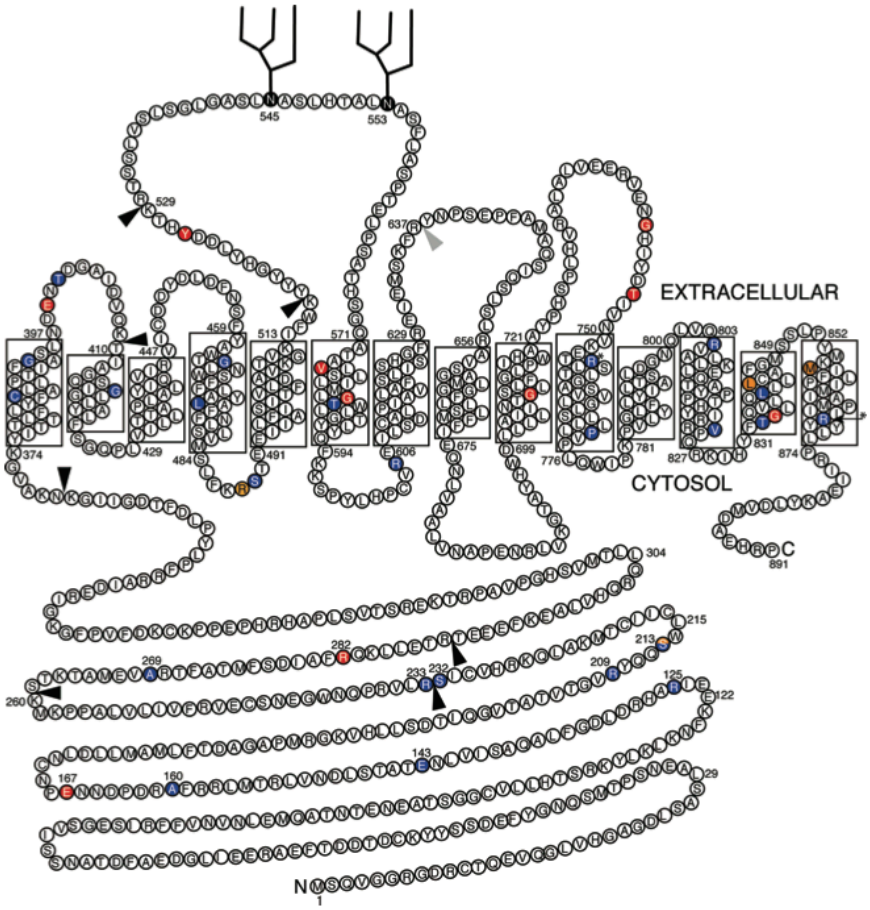


Figure 4.7 Topology model for human SLC4A11. Identified point mutations causing CHED (blue) FECD (red) and Harboyan Syndrome (orange) are shown. The numbers indicate the amino acid position. The predicted N-glycosylation sites are indicated in black. The black and grey arrowheads indicate trypsin cleavage sites identified through partial digestion of Myc-SLC4A11 and SLC4A11-Myc respectively. Reproduced with permission (Vilas et al., 2011).

To conclude, CHED progresses to Harboyan syndrome in all the cases that have been studied and presented in this thesis, such that both conditions appear to be the same disease at different stages of development. When CHED is diagnosed, patients are frequently referred for audiometric assessment. Their hearing may initially fall within normal limits but it has been shown that it may progress subsequently. Therefore, CHED patients should be monitored for progressive hearing loss. Additionally, parents of CHED patients are *SLC4A11* mutation carriers appear to be at increased risk of developing late-onset FECD and should be monitored for this.

The results presented in this chapter were published in the journal Cornea (Siddiqui et al., 2014).

5 Genetic Analysis of Fuchs Endothelial Dystrophy in Yorkshire

5.1 Introduction

FECD is a slowly progressive disease of the corneal endothelium (Krachmer et al., 2011). Its prevalence amongst Caucasians is approximately 4% over the age of 40 (Lorenzetti et al., 1967) and it accounts for about 22% of corneal transplants in the UK (Keenan et al., 2012). The condition is characterised by focal collagenous excrescences in DM called guttata seen on the posterior surface of the basement membrane and is accompanied by endothelial cell pleomorphism in the early stages. Progressive loss of endothelial cells causes loss of endothelial barrier and pump functions, leading to the influx of aqueous humour into the cornea resulting in corneal oedema, reduced corneal clarity and ultimately vision loss (Waring et al., 1978b).

Early-onset FECD may present as early as the first decade (Weiss et al., 2015). Missense changes in *COL8A2* (MIM*120252) on chromosome 1p34.3-p32 have been implicated in this rare form of disease (Biswas et al., 2001, Gottsch et al., 2005). Late-onset FECD is a common disease, typically presenting on average in the fifth decade, although guttata may be clinically apparent after the age of 40 (Elhalis et al., 2010). Many cases appear to be sporadic, although familial forms with apparent dominant inheritance have been documented (Krachmer et al., 1978). Late onset FECD is genetically heterogeneous. Rare causative mutations have been identified in *SLC4A11* (MIM*610206) (Vithana et al., 2006, Riazuddin et al., 2010b, Soumittra et al., 2014, Kim et al., 2015b) as well as *ZEB1/TCF8* (MIM*189909) (Riazuddin et al., 2010a, Lechner et al., 2013). Furthermore, familial inheritance in single pedigrees has been described for dominant loci on chromosomes 5 (FECD5, MIM%613269) (Riazuddin et al., 2009), 9 (FECD7, MIM%613271) (Riazuddin et al., 2010a) and 13 (FECD2, MIM%610158) (Sundin et al., 2006b), although the mutations involved have not yet been identified.

During the course of the studies reported in this thesis, rare dominant mutations in *LOXHD1* (MIM*613072) (Riazuddin et al., 2012), which maps close to the FECD3 locus on chromosome 18 (MIM%613267) (Sundin et al., 2006a)), and *AGBL1* (MIM*615496) (Riazuddin et al., 2013), that maps within the FECD8 locus on chromosome 15, have been reported to cause late-onset FECD (section 1.10.7). Mutations in *AGBL1* have not yet been replicated independently in another cohort. Recently, 128 FECD cases were examined and novel mutations in *LOXHD1* were detected, however as this is an unpublished thesis, full details of these variants are not yet available (Kuot, 2015).

FECD has previously been mapped to chromosome 18q and the locus designated FECD3. Subsequently, Baratz et al (Baratz et al., 2010) performed a small-scale genome wide association study (GWAS) using 130 FECD cases and 260 controls. Controls subjects were 60-years-old or over and had no observable guttata. This identified one SNP, rs613872, in intron 3 of *TCF4 (*MIM*602272) that reached genome-wide significance, with the G allele significantly enriched in FECD cases compared with controls. Close to this SNP lies an intronic trinucleotide repeat expansion, CTG18.1, which was found to be even more strongly associated with FECD and has been shown to exist in 79% of cases compared to only 3% of controls (Wieben et al., 2012). The *TCF4* gene lies within the FECD3 locus on chromosome 18q. Towards the beginning of the studies described in this thesis, Riazuddin et al published mutations in the *LOXHD1* gene, which maps just outside the critical interval for the FECD3 locus, as a cause of dominant late-onset FECD (Riazuddin et al., 2012). The positions of the rs613872 SNP and CTG18.1 repeat in *TCF4*, the *LOXHD1* gene and the FECD3 locus are shown in Figure 5.1. The physical position shows that the two genes are within 9.2Mb of each.

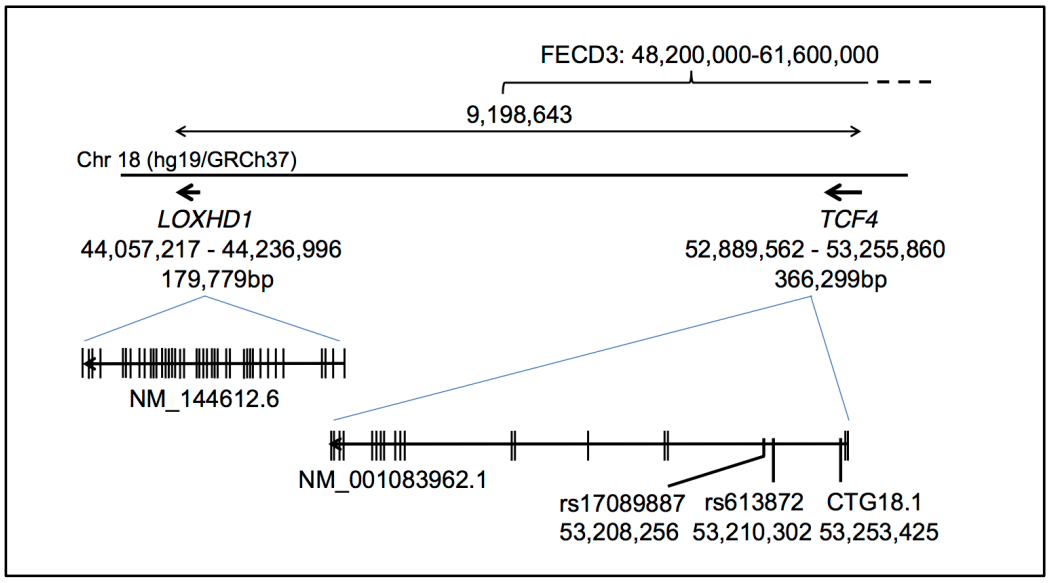


Figure 5.1 Graphical representation of part of the FECD3 locus and surrounding region, including positions of the *TCF4* and *LOXHD1* genes. All physical positions and distances are based on the hg19/GRCh37 version of the genome sequence and intervals are relative to the top of the p-arm of chromosome 18. The gene positions, their length and details of the most common transcript isoform for each are depicted. The positions of the *TCF4* variants rs17089887, rs613872 and CTG18.1, shown to be associated with FECD, are shown.

This chapter describes the genetic analysis of a Caucasian FECD cohort, recruited and sampled from Ophthalmology clinics and theatre lists in Yorkshire, that consists mostly of single cases but also includes 3 families with multiple affected members. The cohort of genomic DNA samples from FECD cases was recruited in two phases. At the start of the project, 56 genomic DNA samples from unrelated FECD cases and 1 family had been ascertained by Dr Aine Rice (University of Leeds) and SS. During the course of the genetic analysis, a further 61 DNA samples were ascertained and 2 families recruited. The total FECD cohort contained 117 unrelated cases.

The aim of this study was to analyse this cohort of Yorkshire FECD patients to identify genes involved in the pathogenesis of FECD. In the first instance this involved investigating *LOXHD1* to try and confirm the published data, especially given the inconsistencies between the FECD3 mapping data and the *LOXHD1* mutation identification data. Similarly, the *TCF4* expansion was investigated to in order to clarify the significance of the original findings. Alongside this, WES was used to try and identify the causative mutations in

the FECD families and transcriptome analysis of normal corneal endothelia was attempted to prioritise candidates in the WES data.

5.2 Results

5.2.1 Recruitment of local FECD families

Following ethical approval, the proband of a large Caucasian FECD family, FECDBRI, was identified in the BRI Ophthalmology Clinics by Mr Nigel James, Consultant Ophthalmologist at Bradford Royal Infirmary (Figure 5.2). Members of the family were subsequently examined by slit-lamp, recruited and sampled at St. James's University Hospital, Leeds following their informed consent. The proband had three affected siblings. One additional sibling had very mild pigmented endothelial changes but could not be classified as having FECD so her diagnosis remained uncertain. Specular microscopy images of the oldest unaffected individual and the proband are shown in Figure 5.3.

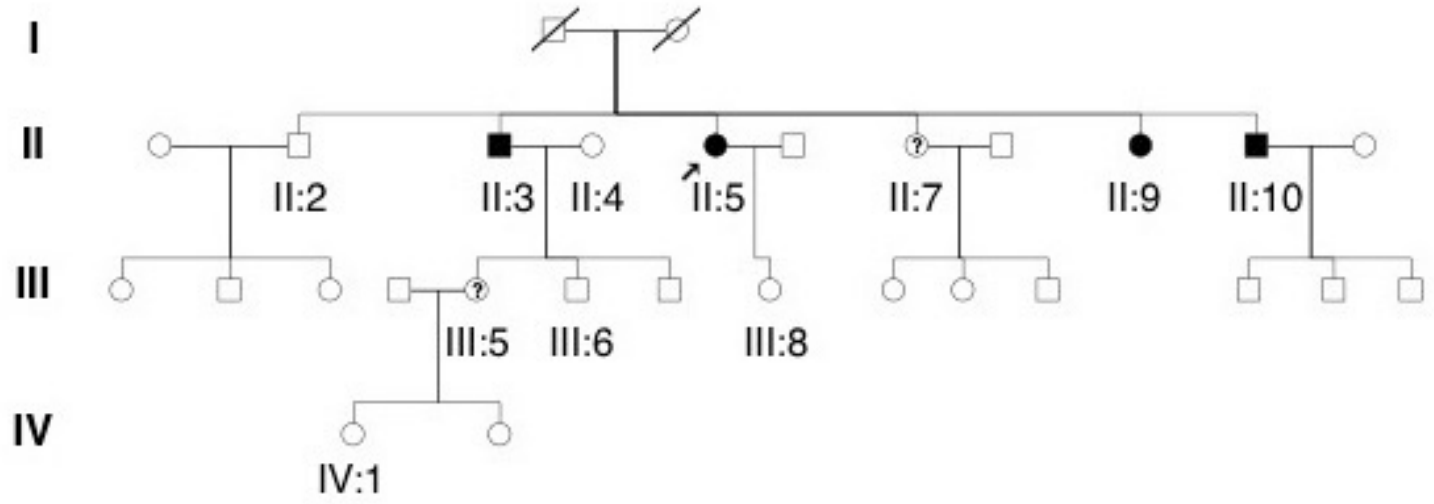
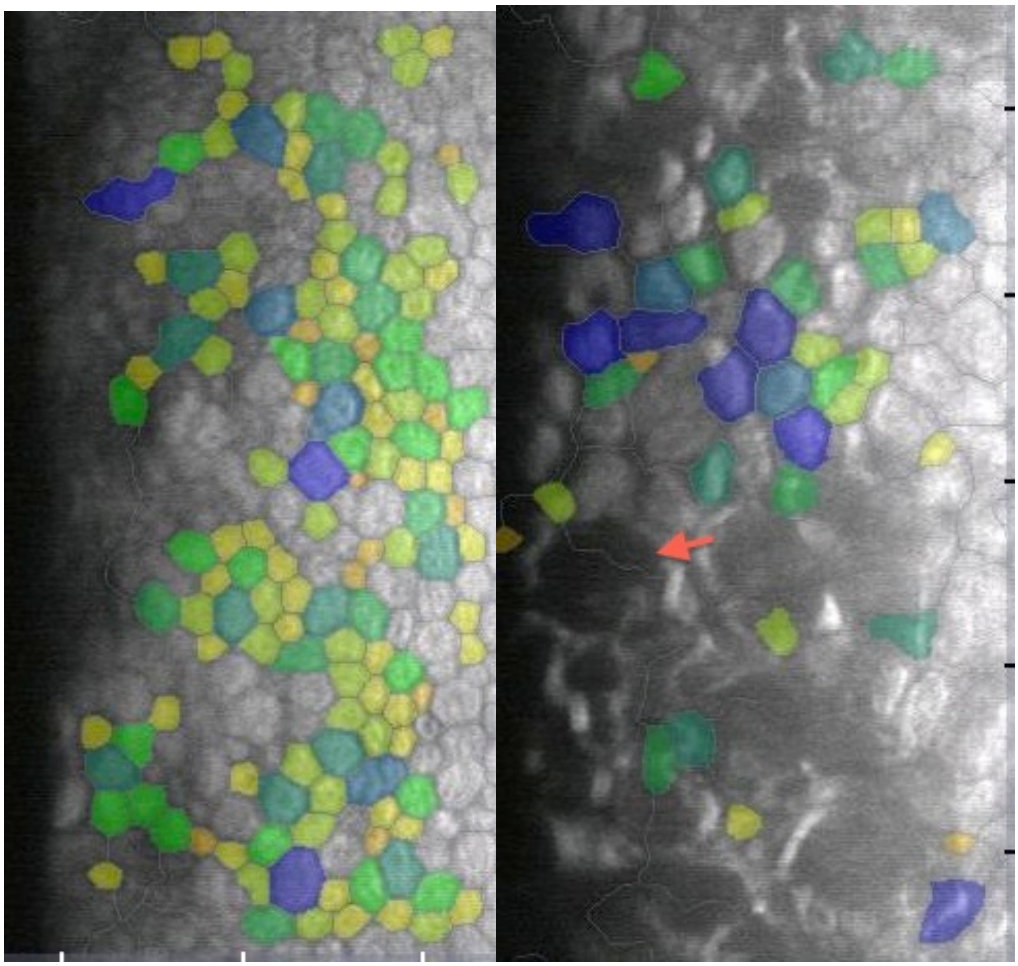


Figure 5.2 Pedigree of the FECD family FECDBRI. Family members for whom DNA was available are labelled. The arrow indicates the proband. Individuals II:9 and III:5 have an uncertain diagnosis.



A B

Figure 5.3 Specular microscopy of A. an unaffected and B. an affected member of the FECDBRI family. A shows the normal polygonal pattern of packed endothelial cells B. shows loss of the normal mosaic pattern of the endothelium with reduced endothelial cell count, polymegathism and pleomorphism. Several guttata are visible, one of which is indicated by the red arrow.

5.2.2 Screening for the previously published FECD causing *COL8A2* mutation in the FECDBRI family

Early onset FECD is a distinct form of FECD (Weiss et al., 2015) caused by mutations in the *COL8A2* gene (Biswas et al., 2001). Individuals in FECDBRI presented with their symptoms between the ages of 40 and 50 years, and were therefore thought to have late-onset FECD. However, as one of the families first reported with a *COL8A2* mutation originated from Yorkshire, the possibility that a common founder mutation existed in the FECDBRI family

was explored. Primer pairs for exon 2 of *COL8A2* (Appendix III) containing the published mutation, c.1363C>A, p.Q455K, were used to amplify and Sanger sequence the proband's DNA. For comparison, genomic DNA from an affected member of the original pedigree with a *COL8A2* mutation was also sequenced (Figure 5.4). It was found that the proband in the FECDBRI family did not have the *COL8A2* mutation, suggesting that the cause of their FECD remained unknown.

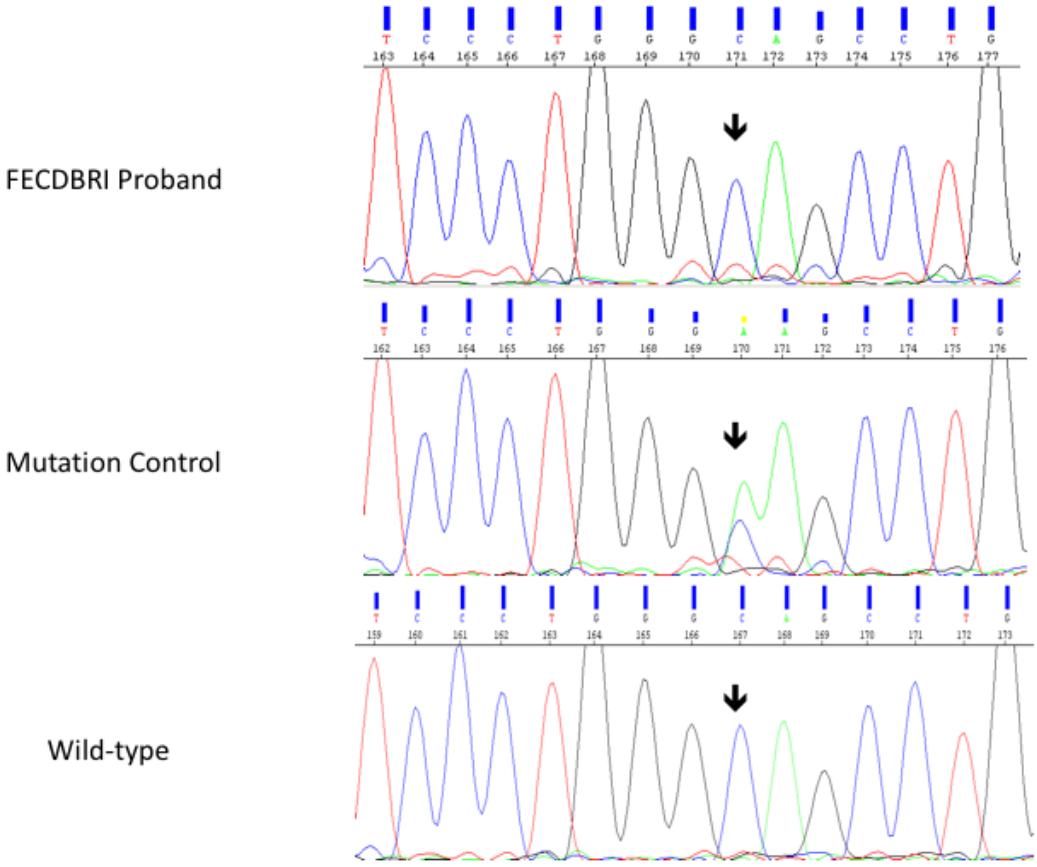


Figure 5.4 Sequence chromatograms of *COL8A2* exon 2 derived from DNA of the FECDBRI proband, a known mutation carrier and wild-type control. The heterozygous c.1363C>A, p.Q455K mutation is indicated by an arrow in the mutation control.

5.2.3 Genetic analysis of the chromosome 18 FECD locus in a local FECD family and the Yorkshire cohort

5.2.3.1 Microsatellite linkage analysis across the FECD3 locus on chromosome 18 in the FECDBRI family

At the start of this PhD, the plan had been to test this family for linkage to all of the known loci (Sundin et al., 2006b, Sundin et al., 2006a, Riazuddin et al., 2009, Riazuddin et al., 2010a). In September 2011, CFI spoke to Professor Nicholas Katsanis, Duke University Medical Center at a conference. It was discussed that his team had identified the causative mutation in the locus on chromosome 18. Therefore, it was decided that microsatellite analysis of this in the FECDBRI would commence with the locus on chromosome 18 (Sundin et al., 2006a).

Genomic DNA from family members was genotyped using microsatellite markers as described in Section 2.12. The markers used and their genetic distance from the top of the chromosome (according to the Marshfield genetic map) (Broman et al., 1998) were D18S1152 at 80.41 cM, D18S1144 at 82.25 cM, D18S1103 at 83.46 cM and D18S64 at 84.80 cM. An example of a microsatellite genotyping electropherogram is shown in Figure 5.5. The genotyped values corresponding to the size of each marker for each family member have been plotted onto the pedigree to generate haplotypes (Figure 5.6). From this analysis, it appears that all the affected cases were found to share a common haplotype 268/175/242/190, consistent with linkage to FECD3. However, this haplotype was also present in the individual of uncertain diagnosis, which might suggest that they are predisposed to developing the condition at a later stage or it may be suggest that the haplotype has nothing to do with FECD risk and has occurred by chance.

Linkage analysis was carried out using Superlink as described in Section 2.12 under a dominant model of inheritance with a 99% penetrance and zero

phenocopy rate. Disease allele frequency was taken as 0.001. Assuming the case with an uncertain diagnosis was affected, the maximum lod score that could be obtained was 2.0980. Without these assumptions about clinical status and phenocopy, a lod score of 2 would have been significant given the previous published linkage study to chromosome 18 (Sundin et al., 2006a). However, although the haplotype analysis suggested linkage to chromosome 18, this could not be confirmed as significant on the basis of statistics as the lod score was below 3.0 and suggested that association of this haplotype with FECD remained inconclusive.

A.

B.

C.

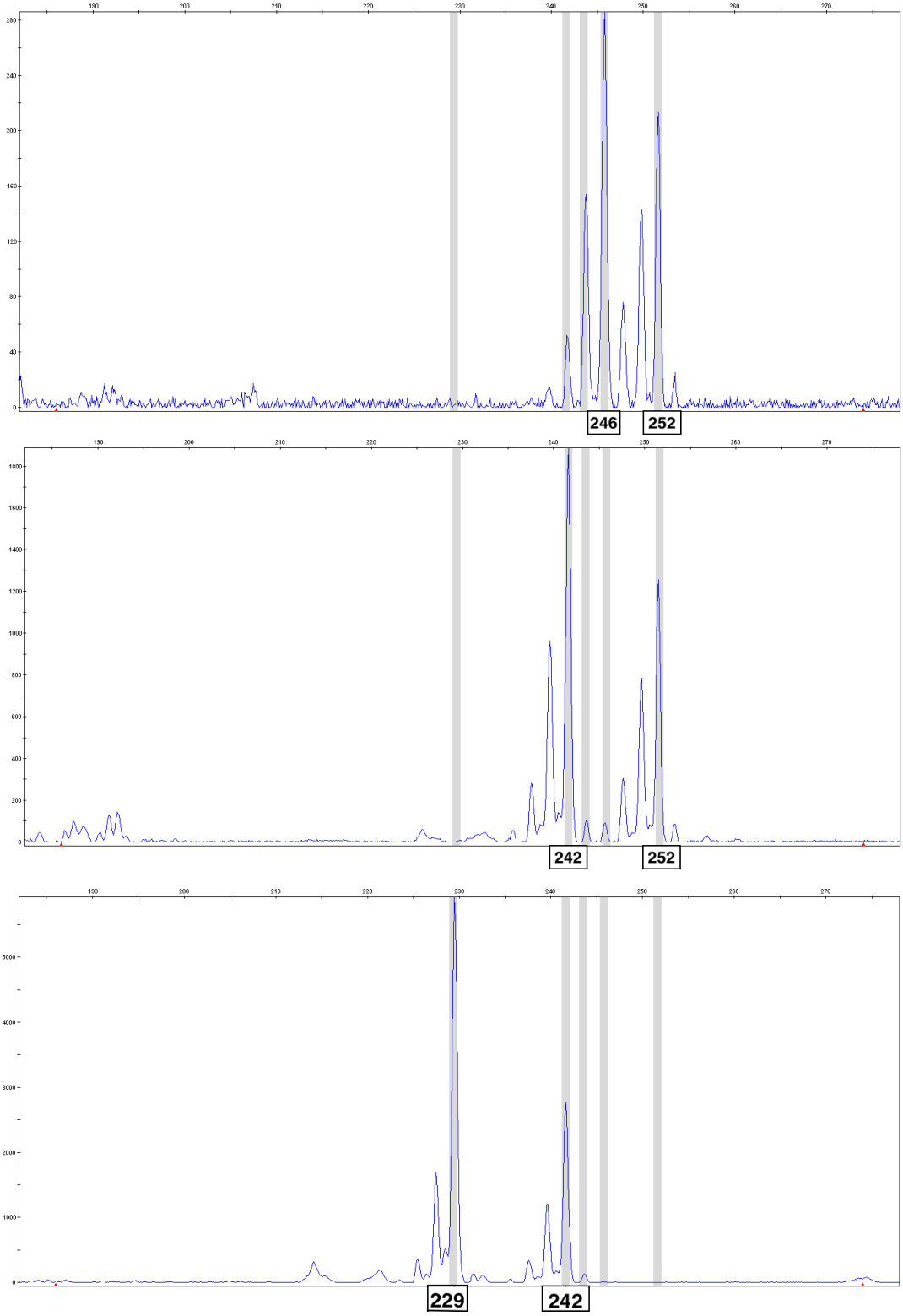


Figure 5.5 Examples of genotyping electropherograms for the chromosome 18 microsatellite marker D18S1103. Traces for FECDBRI individuals A. II:2 B. II:5 C. II:3 are shown.

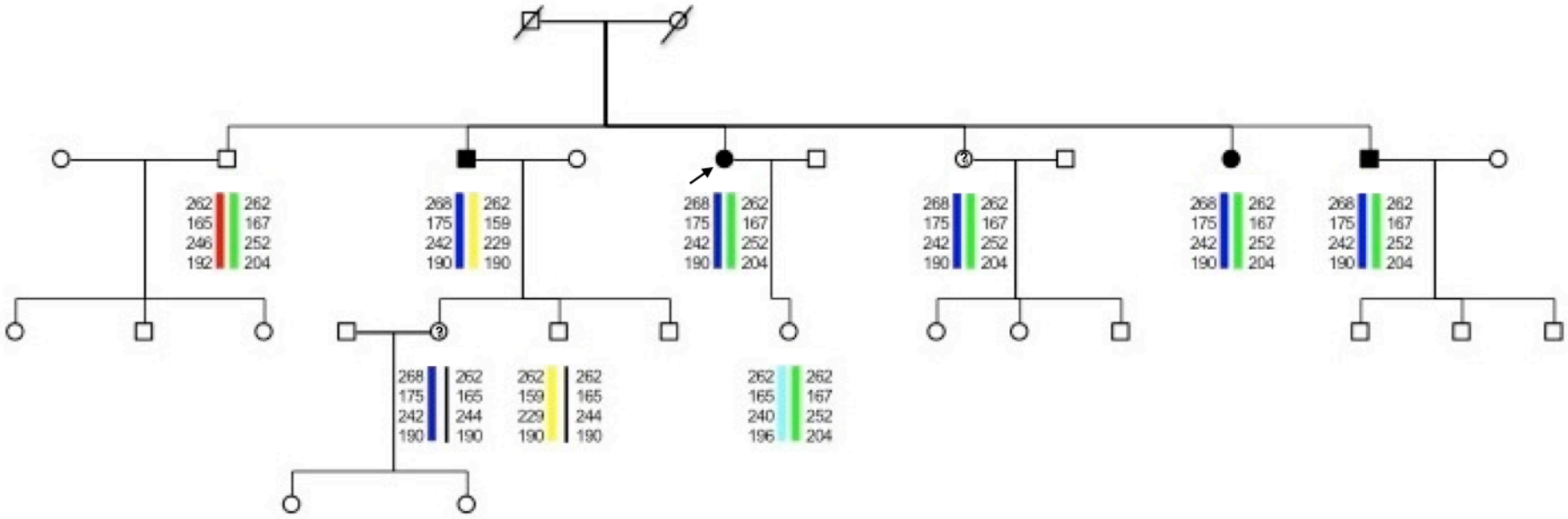


Figure 5.6 Chromosome 18 haplotypes across the FECD3 locus for the FECBRI family. The allele sizes for the markers D18S1152, D18S1144, D18S1103 and D18S64 are shown from the top to the bottom of the haplotype respectively. A common heterozygous region that is indicated in blue is shared amongst the affected individuals, as well as being carried by individuals II.7 and III.5 who are of uncertain diagnosis

5.2.3.2 The *LOXHD1* mutation spectrum in local Caucasian FECD cases

Following identification of mutations in LOXHD1 gene on chromosome 18 as causative for FECD in February 2012 (Riazuddin et al., 2012), the spectrum of *LOXHD1* variants in the Yorkshire cohort was assessed. Genomic DNA from the proband of FECDBRI and 55 unrelated FECD patients, 10 of whom had a family history, were included in this screen. The mean age of these cases at venesection was 71.7 +/- 9.5 years, ranging between 48 and 91 years old. 40 of the 56 cases (71%) were female. The cohort was screened by Sanger sequencing for variants in *LOXHD1*. To do this, primer pairs were initially designed against all the 40 exons of transcript NM_144612 and the two additional splice variant exons in transcripts NM_001145472.2 and NM_001173129 of *LOXHD1.* The primers were optimized before screening for variants in whole genome amplified DNA samples from the patient cohort. All variants found on the initial screen were confirmed in an independent PCR using unamplified DNA from stock solutions. Primer pairs and reaction conditions are shown in Appendix III.

An agarose gel showing amplification of exon 32 in samples 47 to 55 is shown in Figure 5.7.

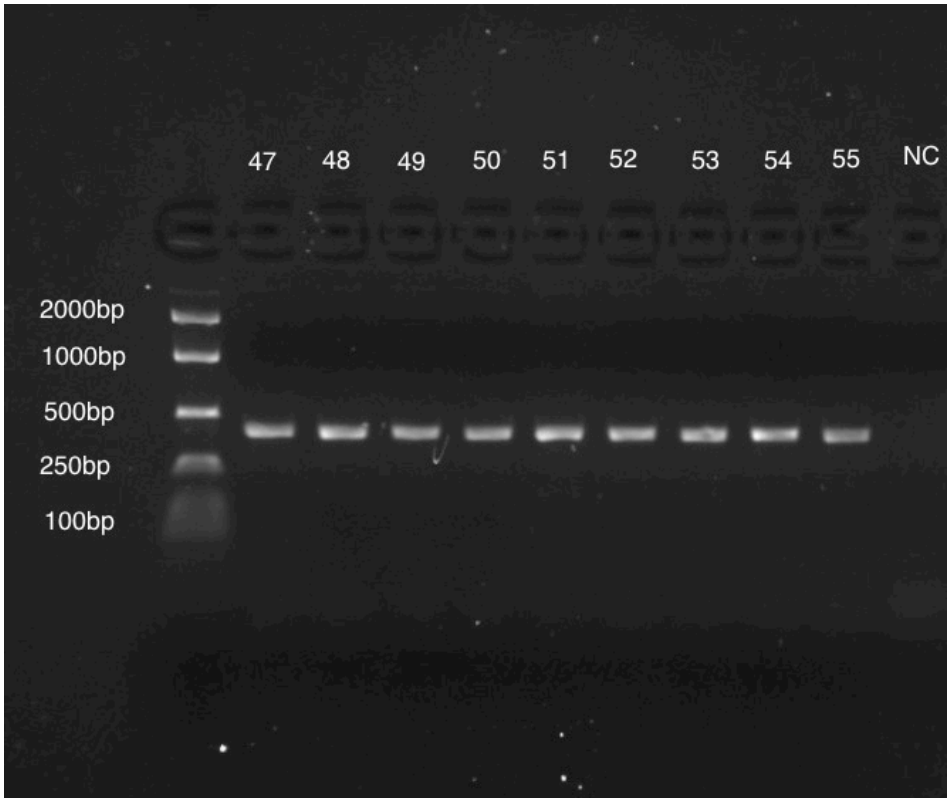


Figure 5.7 Agarose gel image after UV illumination showing amplification of exon 32 of *LOXHD1*. Samples 47-55 of the FECD cohort are shown. The expected band size is 409bp. NC indicates the negative control.

Sequence variants were considered significant if they were within the coding region of the gene or the two base pair splice recognition sites on either side of the exon, and were non-synonymous. All variants were found in *LOXHD1* transcript NM_144612. There were no variants found in transcript NM_001145472.2 or NM_001173129.

Table 5.1 summarises the 16 variants that were found. All of the sequence changes were heterozygous in the subjects. In order to identify the *LOXHD1* variants that were likely to be pathogenic, any variants that had a minor allele frequency greater than 4% in either the 1000 Genomes or the Exome Variant Server databases were excluded. The threshold of 4% was selected as this was the purported frequency of FECD (Lorenzetti et al., 1967). Variants were also assessed for pathogenicity using various mutation prediction programs. They were considered potentially pathogenic if at least one of these predictions was consistent with high pathogenicity, including scores of “Possibly” or “Probably Damaging” when assessed by Polyphen (HDIV),

“Damaging” on SIFT, “Deleterious” on PROVEAN or “Disease-causing” with Mutation Taster (Section 2.17). The *LOXHD1* sequence variants that passed these filtering criteria are highlighted in Table 5.2. The evolutionary conservation of the variants was also assessed using ClustalW (Section 2.18) and the multiple sequence alignments are shown in Figure 5.8.

Variant	No of samples containing varian	Samples containing change	Polyphen Prediction (HDIV)	SIFT prediction	PROVEAN Prediction	Mutation Taster	1000 genomes Variant Frequency (%)	EVS Variant Frequency (%)
c.2T>A, p.M1K	4	25, 26, 32, 53	Possibly Damaging	Damaging	Neutral	Polymorphism	3.5	4.0
c.889A>C ,p.T297P	1	3	Benign	Tolerated	Neutral	Polymorphism	0.3	0.1
c.1087G>A, p.V363I	11	3, 9, 11, 14, 18, 21, 23, 29, 31, 36, 53	Benign	Tolerated	Neutral	Polymorphism	7.8	9.0
c.1570C>T, p.R524C	1	31	Probably Damaging	Damaging	Deleterious	Disease Causing	0.0	0.0
c.1876G>T, p.G626C	2	32, 38	Probably Damaging	Damaging	Deleterious	Disease Causing	1.6	1.4
c.1894G>T, p.G632C	2	32, 38	Probably Damaging	Damaging	Deleterious	Disease Causing	1.6	1.4
c.2473G>A, p.V825M	4	33, 38, 41, 48	Benign	Tolerated	Neutral	Polymorphism	7.2	8.7
c.2558G>C, p.R853P	1	40	Benign	Tolerated	Deleterious	Disease Causing	0.0	0.0
c.2825_2827delAGA, pK942del	1	40	Unknown	Unknown	Neutral	Polymorphism	2.3	5.8
c.2998C>T, p.R1000W	1	16	Probably Damaging	Damaging	Deleterious	Polymorphism	0.0	0.1
c.3463A>G, p.R1155G	55	All samples except 43	Benign	Tolerated	Neutral	Polymorphism	16.8	83.5
c.4868A>G, p.E1623G	11	1, 6, 9, 11, 17, 23, 26, 36, 37, 38, 47	Probably Damaging	Tolerated	Deleterious	Polymorphism	9.2	11.5
c.5545G>A, p.G1849R	1	31	Probably Damaging	Damaging	Deleterious	Disease Causing	0.0	0.0
c.5616C>A, p.N1872K	2	29, 53	Benign	Tolerated	Neutral	Polymorphism	0.7	0.7
c.6107C>T, p.A2036V	15	4, 5, 6, 11, 18, 19, 20, 21, 25, 26, 29, 31, 35, 41, 56	Benign	Tolerated	Neutral	Polymorphism	22.3	0.0
c.6398G>A, p.R2133H	6	6, 9, 11, 17, 23, 36	Benign	Tolerated	Deleterious	Disease Causing	7.3	8.0

Table 5.1 All *LOXHD1* (NM_144616) variants detected in the first 56 patients recruited in the Yorkshire FECD cohort. An assessment of their variant frequency in known databases, pathogenicity prediction and in whom they were identified are depicted. The databases used were Polyphen2 (HDIV) http://genetics.bwh.harvard.edu/pph2/, (Adzhubei et al., 2010), SIFT, http://sift.jcvi.org (Ng and Henikoff, 2001), PROVEAN, http://provean.jcvi.org/index.php (Choi et al., 2012), Mutation Taster http://www.mutationtaster.org (Schwarz et al., 2010b), 1000 genomes (http://www.1000genomes.org) and EVS (http://evs.gs.washington.edu/EVS/databases).

Variants	Variant frequency: 1000 genomes (%)	Variant frequency: Exome variant server (%)	Polyphen (HDIV)	SIFT	PROVEAN	Mutation Taster	Sample ID containing the variant
c.1570C>T, p.R524C	0.0	0	Probably Damaging	Damaging	Deleterious	Disease Causing	31
c.1876G>T, p.G626C	1.6	1.4	Probably Damaging	Damaging	Deleterious	Disease Causing	32, 38
c.1894G>T, p.G632C	1.6	1.4	Probably Damaging	Damaging	Deleterious	Disease Causing	32, 38
c.2558G>C, p.R853P	0.0	0.0	Benign	Tolerated	Deleterious	Disease Causing	40
c.2998C>T, p.R1000W	0.0	0.1	Probably Damaging	Damaging	Deleterious	Polymorphism	16
c.5545G>A, p.G1849R	0.0	0.0	Probably Damaging	Damaging	Deleterious	Disease Causing	31

Table 5.2 List of putative pathogenic *LOXHD1* variants that were identified in Caucasian FECD cases. The *LOXHD1* variants remaining after filtering for variant frequency in known databases and pathogenicity prediction. ID for the sample in whom they were identified is also given.

a. p.R524C

```
Human     490 RVWHDKRSSGSGWHLERMTLMNTLNKDKYNFNCNRWLDANEDDNEIVREM 539
Chimp     490 RVWHDKRSSGSGWHLERMTLMNTLNKDKYNFNCNRWLDANEDDNEIVREM 539
Rat       490 RAWHDRQNPGSGWHLEKMTLMNTLNKDKYNFNCNRWLDANEDDNEIVREM 539
Dog       490 RAWHDRRSPGSGWHLEKMTLMNTLTKDKYNFNCNRWLDANEDDNEIVREM 539
Cow       490 RVWHDKRNPGSGWHLEKMTLLNTLTKDKYNFNCNRWLDANEDDNEIVREM 539
Chicken   360 RIWHEKRNPFAGWHLDKVTLLKTLTKDKYSFNCGRWLDINEDDNEIVREL 409
Turtle    501 RIWHEKRSPFAGWHLDKITLLKTLTKEKYTFNCGRWLDINEDDNEIIREL 550
```

b. p.G626C, p.G632C

```
Human     590 YNCRNNTDLFEKGNADEFTIESVTMRNVRRVRIRHDGKGSGSGWYLDRVL 639
Chimp     590 YNCRNNTDLFEKGNADEFTIESVTMRNVRRVRIRHDGKGSGSGWYLDRVL 639
Rat       590 YNCRNNTDLFEKGNADEFTIESVTMRKVRRVRVRHDGKGSGSGWYLDRVL 639
Dog       590 YNCRNNTDLFEKGNADEFTIESVTMRKVRRVRIRHDGKGGGSGWFLERVL 639
Cow       590 YNCRNNTDLFEKGNADEFTIESVTMRKVRRVRIRHDGKGSGSGWFLERVL 639
Chicken   460 YKCINNVNKFEKGNADEFFVEAVTLKQVRRVRIGHDGKGGSSGWYLAKVI 509
Turtle    601 YNSINTMNKFEKGSADEFYIEAVTLKQVRRVRIGHDGKGGSSGWYLDKVI 650
```

c. p.R853P

```
Human     840 GEKGKTEVLFLSSRSKVFERASKDTFQLEAADVGEVYKLRLGHTGEGFGP 889
Chimp     840 GEKGKTEVLFLSSRSKVFERASKDTFQLEAADVGEVYKLRLGHTGEGFGP 889
Rat       840 GEEGKTEVLFLSSRSKVFDRGSKDIFQLEAADVGEIYKIRLGHTGEGFGP 889
Dog       840 GEDGKTEVLFLSSRSKVFDRASKDIFQLEAQDVGEVFKIRLGHTGEGFGP 889
Cow       840 GEEGKTEVLFLSSRSKVFERASKDTFQLEAADVGQVFKIRLGHTGEGFSP 889
Chicken   710 GETGKTELIILENRSNNFERGATDIFKVEAADVGKIYKIRIGHDGKGIGD 759
Turtle    851 GETGKSELITLKSRSNNFERGATDIFKVEAADVGKIYKLRIGHDGKGIGD 900
```

d. p.R1000W

```
Human     985 EVIEQHKFEAHRWLARGKEDNELVVELVPAGKPGPERNTYEVQVVTGNVP 1034
Chimp     986 EVIEQHKFEALRWLARG--------------------NTYEVQVVTGNVP 1015
Rat       989 EVIVQYKFVANRWLARGKEDNELVVELVPAGQPGPEPNTYEVQVITGNVP 1038
Dog       989 EVIEEYKFEAHRWLARGKEDNELVVELVPAGRPGPEPNTYEVQVITGNVP 1038
Cow       990 EVIQEYKFDAHRWLAQGKEDNELVVELVPTGQEGPEPNTYEIQVITGNVP 1039
Chicken   802 EVMDVYTFVAHRWLAKDEGDKELVVELVPDGESELEENTYEVHVLTGSVW 851
Turtle    915 EVMEVCKFVAHRWLATDEGDKELVVELIPEDGSGLEENTYEVHVITGNVW 964
```

e. p.G1894R

```
Human    1821 EWTSYTVAVKTSDILGAGTDANVFIIIFGENGDSGTLALKQSANWNKFER 1870
Chimp    1802 EWTSYTVAVKTSNILGAGTDANVFIIIFGENGDSGTLALKQSANWNKFER 1851
Rat      1826 EWTSYTVSVKTSDILGAGTDANVFIIIFGENGDSGTLALKQSANWNKFER 1875
Dog      1826 EWTSYTVSVKTSDILGAGTDANVFIIIFGENGDSGTLALKQSANWNKFER 1875
Cow      1827 EWTSYTVSVKTSDVLGAGTDANVFIIIFGENGDSGTLALKQSANWNKFER 1876
Chicken  1691 EDTTYTIQVKTSDIGGAGTDANVSLILFGENGDSGTLALKESNKSNKFER 1740
Turtle   1743 ENTTYTVQVKTSDIGGAGTDANVSLIMFGENGDSGTLALKESNKSNKFER 1792
```

Figure 5.8 Protein sequence alignment of LOXHD1 orthologues around the predicted pathogenic variants that were identified in the FECD cases. Each variant, a. p.R524C, b. p.G626C, p.G632C, c. p.R853P, d. p.R1000W and e. p.G1894R, is highlighted in red and the surrounding amino acids that are identical to the human transcript are shaded in grey. Accession numbers for LOXHD1 sequence are NP_653213.6 (human), XP_009432217.1 (chimp), NP_001099602.3 (rat), XP_547589.3 (dog), XP_003587840.2 (cow), XP_425221.4 (chicken) and XP_007055242.1 (turtle).

5.2.3.3 The *LOXHD1* mutation spectrum in 1000 Genomes database

The published LOXHD1 study (Riazuddin et al., 2012) used a Mendelian model to implicate *LOXHD1* in a dominant family but used a case-control approach to assess the mutational load. A similar approach was used in this thesis except that at this point in the PhD there were no ethnically-matched controls that had been examined to exclude endothelial disease. Therefore, the 1000 genomes database was utilised for the ascertainment of controls.

For the *LOXHD1* variant dataset from population controls, the 1000 Genomes database was accessed in July 2014. The database was interrogated for any cases of Caucasian origin that had a read depth of 5 for at least 85% of the *LOXHD1* gene. A dataset of 467 individuals (934 alleles) who satisfied these criteria was subsequently mined for sequence variation in *LOXHD1*. Any sequence variants were included into the study if they were within the coding region of the gene as well as the two base pair splice recognition sites either side of the exon. However synonymous variants were excluded. A complete list of the 21 *LOXHD1* variants found in controls from the 1000 Genomes database is given in Table 5.3.

In order to establish which *LOXHD1* variants were more likely to be pathogenic, any variants that had a minor allele frequency greater than 4% in either the 1000 Genomes or the Exome Variant Server databases were excluded. Variants were also assessed for pathogenicity according to the same mutation prediction programs used to filter variants found in the FECD cohort. As for the FECD cases, they were retained if at least one of these predictions was consistent with high pathogenicity such as “Possibly” or “Probably Damaging” when assessed by Polyphen (HDIV), “Damaging” on SIFT, “Deleterious” on PROVEAN or “Disease-causing” by Mutation Taster (Section 2.17). The *LOXHD1* sequence variants that passed these filtering criteria are highlighted in Table 5.4.

Variants	1000 Genome Variant Frequency	1000 Genome Variant Frequency (%)	SNP ID	Polyphen Prediction (HDIV)	SIFT Prediction	Provean prediction	Mutation Taster Prediction	Variant Frequency in EVS (%)
c.2T>A, p.M1K	33/934	3.5	rs36024592	Possibly Damaging	Damaging	Neutral	Polymorphism	4.0
c.410G>A, p.R137H	3/934	0.3	rs151268914	Probably Damaging	Tolerated	Neutral	Disease causing	0.0
c.722A>G, p.N241S	2/934	0.2	rs191697915	Probably Damaging	Tolerated	Deleterious	Polymorphism	0.0
c.889A>C, p.T297P	3/934	0.3	rs117747744	Benign	Tolerated	Neutral	Polymorphism	0.1
c.1087G>A, p.V363I	73/934	7.8	rs10163657	Benign	Tolerated	Neutral	Polymorphism	9.0
c.1876G>T, p.G626C	15/934	1.6	rs34589386	Probably Damaging	Damaging	Deleterious	Disease causing	1.4
c.1894G>T, p.G632C	15/934	1.6	rs35088381	Probably Damaging	Damaging	Deleterious	Disease causing	1.4
c.2027A>G, p.D676G	2/934	0.2	rs16978578	Benign	Tolerated	Neutral	Polymorphism	0.0
c.2080G>T, p.D694Y	1/934	0.1	rs35727744	Probably Damaging	Damaging	Deleterious	Disease causing	0.0
c.2473G>A, p.V825M	68/934	7.2	rs36086089	Benign	Tolerated	Neutral	Polymorphism	8.7
c.2554A>G, p.S852G	2/934	0.2	rs183848033	Probably Damaging	Tolerated	Deleterious	Disease causing	0.0
c.2771G>A, p.R924Q	1/934	0.1	rs140904207	Benign	Tolerated	Neutral	Polymorphism	0.0
c.2825_2827del, p.K942del	22/934	2.3	rs142960762	Unknown	Unknown	Neutral	Polymorphism	5.8
c.3269G>A, p.R1090Q	23/934	2.5	rs118174674	Probably Damaging	Tolerated	Deleterious	Disease causing	1.9
c.3463A>G, p.R1155G	157/934	16.8	rs1893566	Benign	Tolerated	Neutral	Polymorphism	83.5
c.4148C>T, p.T1383M	4/934	0.4	rs7244681	Possibly Damaging	Tolerated	Neutral	Disease causing	0.1
c.4217C>T, p.A1406V	3/934	0.3	rs146739496	Possibly Damaging	Damaging	Neutral	Disease causing	0.2
c.4868A>G, p.E1623G	86/934	9.2	rs12606417	Possibly Damaging	Tolerated	Deleterious	Polymorphism	11.5
c.5616C>A, p.N1872K	7/934	0.7	rs61733519	Benign	Tolerated	Neutral	Polymorphism	0.7
c.6107C>T, p.A2036V	208/934	22.3	rs1377016	Benign	Tolerated	Neutral	Polymorphism	0.0
c.6398G>A, p.R2133H	68/934	7.3	rs74316327	Benign	Tolerated	Deleterious	Disease causing	8.0

Table 5.3 All *LOXHD1* variants found in Caucasian controls derived from the 1000 Genomes, together with assessment of likely pathogenicity. The cohort of Caucasian samples that were screened for variants consisted of 467 controls. The databases used were Polyphen2 (HDIV) http://genetics.bwh.harvard.edu/pph2/, (Adzhubei et al., 2010), SIFT, http://sift.jcvi.org (Ng and Henikoff, 2001), PROVEAN, http://provean.jcvi.org/index.php (Choi et al., 2012) and Mutation Taster http://www.mutationtaster.org (Schwarz et al., 2010b). Additionally their frequency in the 1000 genomes (http://www.1000genomes.org) and EVS (http://evs.gs.washington.edu/EVS/databases) are shown.

Variants	Variant frequency: 1000 genomes (%)	Variant frequency: Exome variant server (%)	Polyphen (HDIV)	SIFT	PROVEAN	Mutation Taster	Sample ID containing the variant
c.410G>A, p.R137H	0.2	0.0	Probably Damaging	Tolerated	Neutral	Disease Causing	19752
c.722A>G, p.N241S	0.2	0.0	Probably Damaging	Tolerated	Deleterious	Polymorphism	1492
c.1876G>T, p.G626C	2.3	1.4	Probably Damaging	Damaging	Deleterious	Disease Causing	1149, 1247, 1624, 1779, 7347, 11840, 20519, 20586, 20758, 20765, 20792, 20806; homo alt 1241
c.1894G>T, p.G632C	2.3	1.4	Probably Damaging	Damaging	Deleterious	Disease Causing	1149, 1247, 1624, 1779, 7347, 11840, 20519, 20586, 20758, 20765, 20792, 20806; homo alt 1241
c.2080G>T, p.D694Y	0.2	0.0	Probably Damaging	Damaging	Deleterious	Disease Causing	19734
c.3269G>A, p.R1090Q	2.8	0.0	Probably Damaging	Tolerated	Deleterious	Disease Causing	128, 149, 156, 1167, 1271, 1624, 1747, 1775, 12154, 20507, 20509, 20517, 20756, 20761, 20803; homo alt 6985
c.4148C>T, p.T1383M	0.3	0.1	Possibly Damaging	Tolerated	Neutral	Disease Causing	1704, 19734
c.4217C>T, p.A1406V	0.3	0.2	Possibly Damaging	Damaging	Neutral	Disease Causing	135, 1679

Table 5.4 List of putative pathogenic *LOXHD1* variants identified in 1000 Genomes database controls. The *LOXHD1* variants remaining after filtering for pathogenicity, their allele frequency in known databases, pathogenicity predictions and in whom they were identified are depicted. All the sequence changes identified were heterozygous in the subjects apart from three cases designated “homo alt”. The databases used were Polyphen2 (HDIV) http://genetics.bwh.harvard.edu/pph2/, (Adzhubei et al., 2010), SIFT, http://sift.jcvi.org (Ng and Henikoff, 2001), PROVEAN, http://provean.jcvi.org/index.php (Choi et al., 2012) and Mutation Taster http://www.mutationtaster.org (Schwarz et al., 2010b). Additionally their frequency in the 1000 genomes (http://www.1000genomes.org) and EVS (http://evs.gs.washington.edu/EVS/databases) are shown

5.2.3.4 Comparison of the predicted pathogenic *LOXHD1* variant findings in the Caucasian FECD cohort and 1000 Genomes database controls

From Table 5.1 and Table 5.3 some of the variants were present in individuals who carried only a single heterozygous *LOXHD1* coding variant, while other individuals had two or more variants. Although it was not easily possible to establish the phase of multiple variants found in a single individual, it was observed that *LOXHD1* variants p.G626C and p.G632C were always found to co-exist in the same individuals and were present at exactly the same frequency in the 1000 Genomes and Exome Variant Server databases, suggesting that these variants are almost certainly in phase and represent a single complex allele. These variants are therefore considered to be a single allele in subsequent analyses.

In terms of comparison of the results, there were two ways the analysis could be interpreted, either based on the number of predicted pathogenic alleles identified in each cohort or the number of cases with predicted pathogenic alleles. Table 5.5 summarises these comparisons that were derived from Table 5.2 and Table 5.4 for FECD and controls respectively. The results highlight that 6 out of 112 (5.4%) potentially pathogenic *LOXHD1* alleles were identified in the FECD cohort, whereas 38 out of 934 (4.1%) pathogenic alleles were identified in the controls, suggesting a modest 1.33 fold increased prevalence of alleles amongst the FECD cohort compared to controls (Table 5.5). The pathogenic *LOXHD1* alleles exist in 5 out of 56 (8.9%) FECD cases compared to 34 out of 467 (7.3%) controls, suggesting a 1.24 fold increased enrichment amongst the FECD cases (Table 5.5).

In summary, the Yorkshire FECD cohort was screened for sequence variations in the *LOXHD1* gene, and the results were compared with controls from the 1000 Genomes database. This analysis revealed only a modest enrichment of *LOXHD1* putative pathogenic alleles in FECD cases compared with controls, suggesting that

mutations in *LOXHD1* are unlikely to be the major contributor to FECD onset that was previously identified on chromosome 18 (Sundin et al., 2006a).

		Cases	Controls	*P*	OR
		n = 56	n = 467		
A	Number of alleles	112	934	0.45	1.33 (0.55 - 3.23)
	Predicted pathogenic alleles	6	38		
	Frequency	0.054	0.041		
B	Number of subjects	56	467	0.65	1.24 (0.46 - 3.33)
	Subjects with pathogenic variant	5	34		
	Frequency	0.089	0.073		

Table 5.5 Comparison of *LOXHD1* variant findings in FECD cases and controls. The frequency of *LOXHD1* alleles predicted to be pathogenic (A) and subjects containing a putative pathogenic *LOXHD1* variant (B) are shown. n= the number of cases, P= the p value, OR= odds ratio.

5.2.3.5 Genetic analysis of the *TCF4* intronic polymorphism, rs613872, in Caucasian FECD cases, ECACC controls and 1000 Genomes database population controls

In order to identify whether the GWAS SNP that was associated with FECD in Caucasians (Baratz et al., 2010) was associated in the Yorkshire FECD cohort, genomic DNA from the 56 unrelated FECD patients and 192 normal control DNAs obtained from the ECACC, were genotyped for the *TCF4* intronic polymorphism, rs613872 by Sanger sequencing. The ECACC control cohort were all Caucasians of UK origin. The gender balance was 48% (92/192) females in the ECACC human random control group compared to 71% (40/56) in the cases and mean age at venesection was 38.7 +/- 8.4 and 71.6 +/- 9.5 years respectively. Sequence chromatograms of the three representative genotypes are shown in Figure 5.9 and the results summarised in Table 5.6. As a comparison for population control dataset, the rs613872 genotypes were also downloaded from the 1000 Genomes database for 467 unrelated Caucasian individuals. The results showed that the G allele for the

intronic polymorphism, rs613872, which had previously been shown to be associated with FECD (Baratz et al., 2010, Li et al., 2011, Riazuddin et al., 2011, Kuot et al., 2012, Eghrari et al., 2012, Igo et al., 2012, Stamler et al., 2013), was identified in 73% (41/56) of FECD cases compared to 32% (62/192) and 26% (123/467) in the different control subject groups. These differences were statistically significant (P = 2.77 x 10^{-7} and 9.77 x 10^{-12}) with an odds ratio of 3.81 and 4.47 respectively.

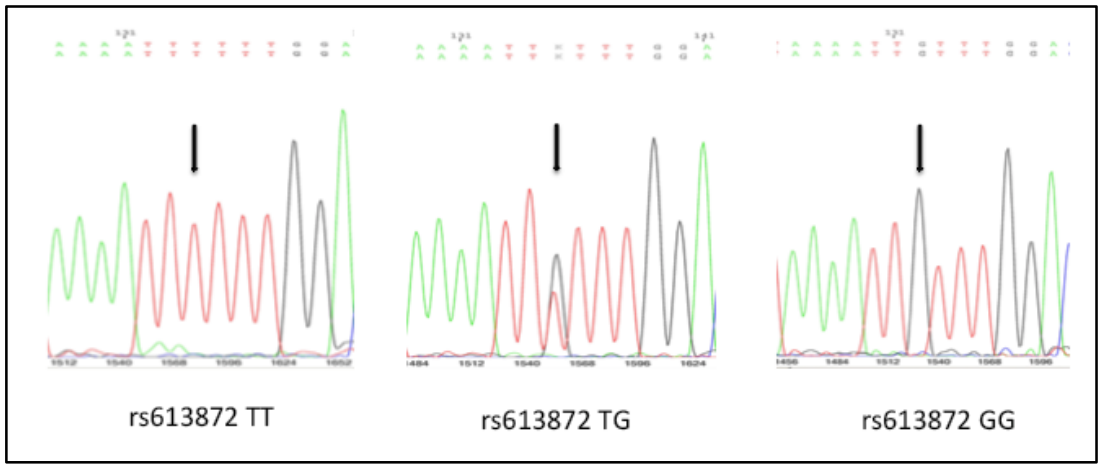


Figure 5.9 Chromatograms from three FECD samples that had been sequenced for the polymorphic SNP, rs613872. The TT, TG and GG genotypes are shown.

		Cases n = 56	Controls* n = 192	*P*	OR (95% CI)	Controls** n = 467	*P*	OR (95% CI)
rs613872	GG	8	7	2.77×10^{-7}	3.81 (2.29 - 6.34)	15	9.77×10^{-12}	4.47 (2.95 - 6.76)
	GT	33	55			108		
	TT	15	130			344		

Table 5.6 Summary of SNP rs613872 genotyping on Caucasian FECD cases and population controls. The subjects were categorised for rs613872 according to their G/T genotyped allele status. Caucasian controls were from ECACC* or the 1000 Genomes database (July 2014)**. The rs613872 genotyped alleles were in Hardy-Weinberg Equilibrium.

5.2.3.6 Genetic analysis of the trinucleotide expansion, CTG18.1, in Caucasian FECD cases and ECACC controls

The FECD and ECACC control cohorts were also genotyped for the previously associated *TCF4* trinucleotide repeat polymorphism, CTG18.1. The existence of the CTG18.1 trinucleotide repeat was established using the short tandem repeat (STR) and Triplet-Primed PCR (TP-PCR) previously described (Warner et al., 1996). Primer sequences for both assays are detailed in Appendix IV.

The STR genotyping assay uses primer pairs flanking the CTG repeat sequence and measures unexpanded CTG18.1 alleles up to 43 trinucleotide repeats in length. This threshold is determined by the assay. However, this assay can fail to amplify expanded alleles, and hence the requirement for parallel analysis using the TP-PCR assay. This detects expanded CTG18.1 alleles by using a trinucleotide repeat specific 3' primer that binds at numerous sites within the CTG repeat paired with a fixed locus-specific 5' primer, resulting in a mixture of products. The TP-PCR assay allows discrimination between the cases with a homozygous unexpanded allele and heterozygous cases with one unexpanded and one expanded allele. Characteristic traces for the TP-PCR assay are shown in Figure 5.10.

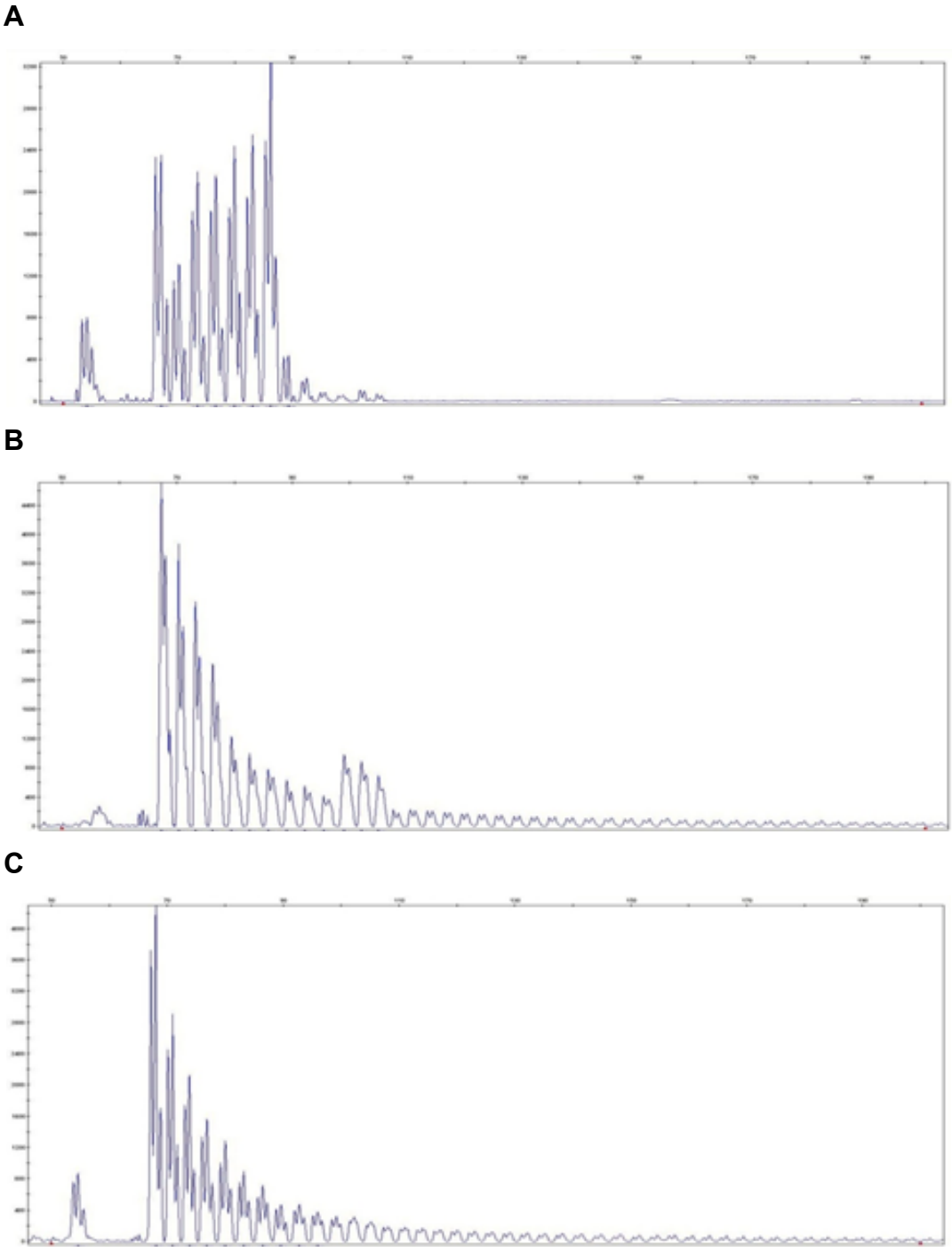


Figure 5.10 TP-PCR electropherograms for assessing CTG18.1 trinucleotide repeat. A. Pattern for a subject with 2 unexpanded alleles at CTG18.1 (the sizes determined by STR assay), B. An individual with an expansion and one unexpanded allele (again the size of the unexpanded allele determined by STR) and C. A case with two expanded alleles (since no product on STR assay).

A summary of the results of the STR and TP-PCR analysis to determine CTG18.1 trinucleotide repeat in FECD cases and controls is given in Table 5.7. The STR assay failed to detect any alleles in 7% (4/56) of FECD cases compared to 0% (0/192) in controls. The 4 FECD cases that failed to detect any alleles following STR assay gave the characteristic pattern for an expanded allele after the TP-PCR assay, suggesting homozygosity for the expansion. It was noted that 68% (38/56) of the FECD cases in the study described here had at least one expanded CTG18.1 allele compared with only 7% (14/192) in the control subjects. The differences between the cases and controls were analysed using Fischer's exact test. This was statistically significant ($P = 3.43 \times 10^{-14}$) with an odds ratio of 15.88.

		Cases n = 56	Controls* n = 192	*P*	OR (95% CI)
CTG18.1	XX	4	0	3.43×10^{-14}	15.88 (7.77 - 32.45)
	SX	34	21		
	SS	18	171		

Table 5.7 Summary of the CTG18.1 trinucleotide repeat analysis on the FECD cases and ECACC controls. The subjects were categorised for CTG18.1 according to the presence (X) or absence (S) of the trinucleotide expanded allele. The CTG18.1 genotyped alleles, which consist of two category groups (denoted X and S) based on trinucleotide repeat number being greater than or less than 43 repeats in length respectively were out of Hardy-Weinberg Equilibrium in cases (*P* = 0.045)

5.2.3.7 Relationship between the *TCF4* trinucleotide expansion and the *LOXHD1* variants found in the FECD cohort

The 5 cases with *LOXHD1* variant alleles described in Section 5.2.3.4 were investigated for their trinucleotide expansion status. Table 5.8 summarises the number of cases with (*LOXHD1*/X) and without (*LOXHD1*/S) an expansion. The observations show that more than 50% of the *LOXHD1* variant cases have an expansion polymorphism at CTG18.1, suggesting that the TCF4 expansion rather than the *LOXHD1* variant may be causal in these cases. This provides further evidence that *LOXHD1* variants are not a significant cause of FECD in this cohort.

Co-existence	Cases n = 5
LOXHD1/X	3
LOXHD1/S	2

Table 5.8 Relationship between the *LOXHD1* variants and the CTG18.1 trinucleotide repeat in FECD cases.

5.2.3.8 Recruitment of endothelium-checked controls and additional FECD cases

The analysis of *LOXHD1* variants and *TCF4* expansion in 56 patients (Sections 5.2.3.2 to 5.2.3.7) was written up in a manuscript and submitted for publication. The reviewers were positive about many aspects of the study, but felt that the number of cases in the *LOXHD1* part of the study was inadequate to confidently refute the findings of the original study. Additionally, they were critical of the lack of controls that had been examined for endothelial disease. The ECACC controls used for the *TCF4* analysis were commercially available, but were not age-or sex-matched and had not had their corneas examined. Therefore, it was feasible that a proportion of the control individuals who carried the *TCF4* expansion would go on to develop FECD. Cognisant of this

potential weakness in study design, 83 age-matched controls with healthy endothelium were identified from Leeds Teaching Hospitals Trust NHS Cataract Clinics by Consultant Ophthalmologist, Mr John Buchan and recruited at St. James's University Hospital, Leeds. This control group had a mean age at venesection of 76.7 +/- 7.7 years (range 56 – 93 years old) and 52% (43/83) were female.

During this time, further recruitment of 61 additional FECD cases had also taken place onto the NIHR portfolio project UKCRN 11297. In total, the FECD cohort now included 117 patients, of whom 65% (71/117) were female and the mean age at venesection was 70.8 +/- 9.2 years (range 48-94 years).

The publication of the *TCF4* RNA foci experiments (Du et al., 2015, Mootha et al., 2015) (Section 5.3.2) provided a clear mechanism of disease for mutations in this gene and cemented the decision to further analyse the *TCF4* mutation spectrum in the full cohort of cases and newly-recruited controls. This was therefore prioritised over further experiments to assess *LOXHD1* mutation spectrum in the case/control cohort.

5.2.3.9 Genotyping of the TCF4 intronic polymorphism, rs613872, and CTG18.1 trinucleotide repeat in the full cohort of 117 FECD cases and 83 age-matched, endothelium-checked controls

The additional 61 FECD samples and 83 endothelium-checked controls were genotyped for the *TCF4* intronic polymorphism, rs613872, and CTG18.1 trinucleotide repeat as described before (Sections 5.2.3.2 to 5.2.3.7) and the data combined with the previous study to increase its power. A summary of the results is shown in Table 5.9. TCF4 rs613872 and CTG18.1 genotypes for all patients in the case/control cohort are listed in Appendix VII.

It was observed that 77.2% (85/117) of the FECD cases in the study have at least one expanded CTG18.1 allele compared with only 6.0% (5/83) and

10.9% (21/192) in the endothelium-checked subjects and the ECACC controls respectively. This was statistically significant (P = 1.95 x 10^{-18} and 8.43 x 10^{-28} with odds ratios of 23.20 and 12.47 respectively. The significantly greater values between the cases and the different controls used to perform the analysis highlights the value in selecting controls that have had their endothelium checked. These findings also show that the CTG18.1 expanded allele of *TCF4* appears to be a major risk allele for FECD in Caucasians. The controls were in HWE, but the cases were not. Linkage disequilibrium between the rs613872 SNP and the CTG18.1 was 0.79 as calculated by r^2 suggesting that these alleles were inherited together.

A.

		Cases n = 117	Controls n = 83	*P*	OR (95% CI)
rs613872	GG	21	0	5.45×10^{-14}	6.113 (3.646 – 10.25)
	GT	71	22		
	TT	25	61		
CTG18.1	XX	13	0	1.95×10^{-18}	23.2 (9.181 – 58.64)
	SX	72	5		
	SS	32	78		

B.

		Cases n = 117	Controls n = 192	*P*	OR (95% CI)
rs613872	GG	21	7	3.384×10^{-15}	4.263 (2.958 – 6.145)
	GT	71	55		
	TT	25	130		
CTG18.1	XX	13	0	8.43×10^{-28}	12.46 (7.473 – 20.76)
	SX	72	21		
	SS	32	171		

Table 5.9 Summary of the intronic SNP rs613872 and CTG18.1 genotyping on the full cohort of FECD cases and controls. The subjects were categorised for rs613872, according to their G/T genotyped allele status, and for CTG18.1, according to the presence (X) or absence (S) of the trinucleotide expanded allele. Caucasian controls were recruited in St James's University Hospital Cataract Clinics (A). As a comparison the genotyping data from the ECACC human random control group was included (B).

5.2.4 Recruitment of two further local FECD families

Two further families, FECDWAK and FECDBAR, were identified at Pinderfields Hospital, Wakefield (Figure 5.11 and Figure 5.12) by Mr Andrew Chung, Consultant Ophthalmologist at Pinderfields Hospital. The recruitment of the family members took place between September 2013 and September 2015.

FECDWAK had five affected individuals in generation II. The brother of the proband was managed clinically by Mr James Ball (JLB) in SJUH and had undergone bilateral DSEKs in 2008. Unaffected individuals from generation III were recruited at a later point in the study.

The son of the FECDBAR proband, individual III:28, was referred from Pinderfields Hospital to St James's University Hospital for an endothelial graft in 2015. DMEKs were carried out by JLB in April and July 2015 to the right and left eyes respectively. His daughter IV:1, aged 28, accompanied him to his follow-up appointments and expressed some concerns about her vision. She was examined and noted to have FECD. Both her and her asymptomatic half-brother IV:II, who was phenotypically normal when examined at the slit-lamp, were recruited, sampled and underwent specular microscopy. The specular microscopy images of both grandchildren of the proband are shown in Figure 5.13.

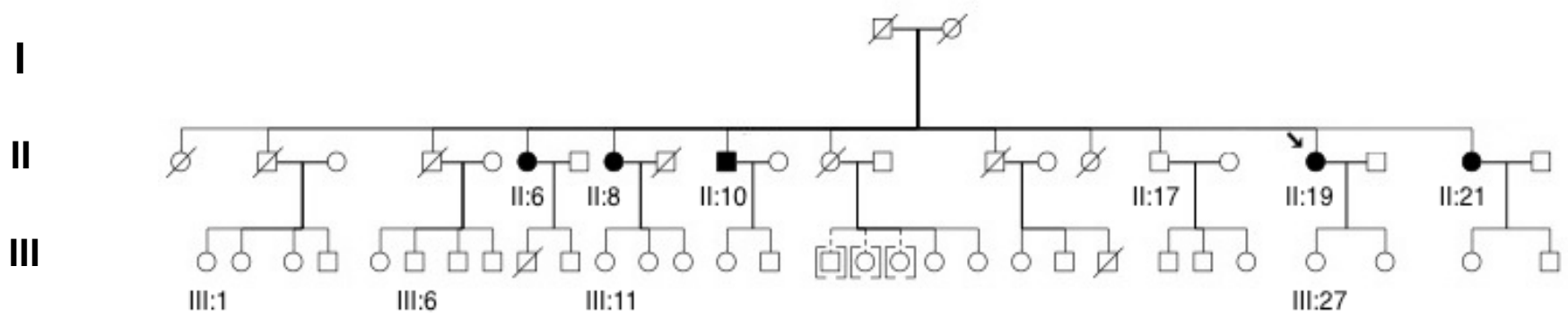


Figure 5.11 Pedigree of the FECD family FECDWAK. Family members for whom DNA was available are labelled. Individuals from generation II were recruited initially, followed by individuals in generation III. The arrow indicates the proband. Individuals in [ ] brackets indicated that the offspring were adopted. The dashed line indicates that the parents were the adoptive parents.

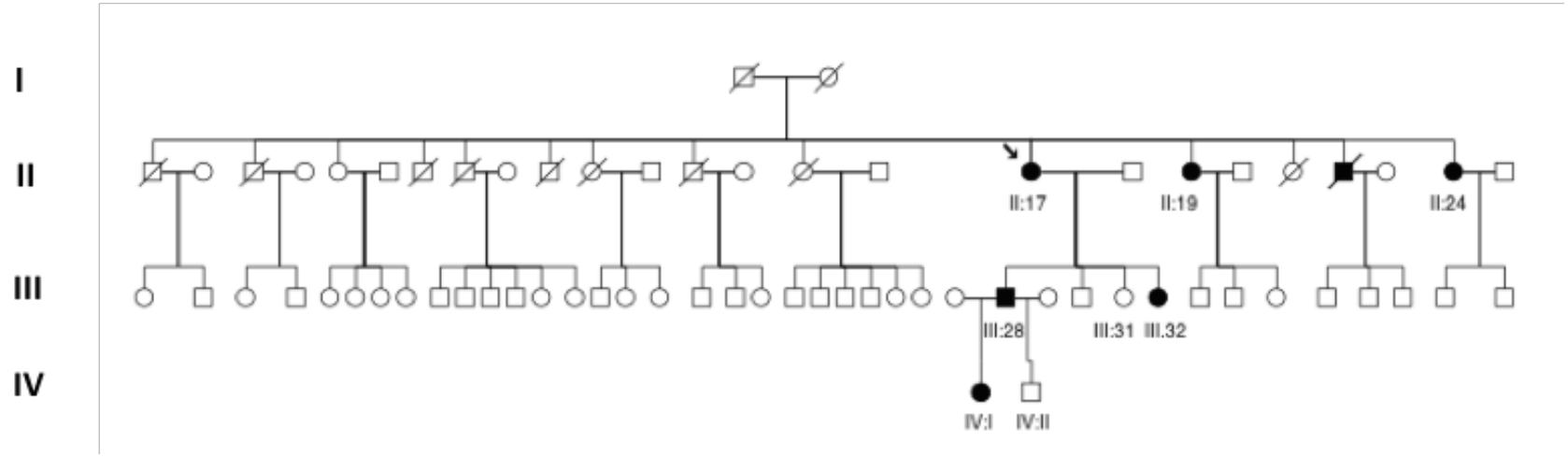


Figure 5.12 Pedigree of the FECD family FECDBAR. Family members for whom DNA was available are labelled. Individuals from generation II and III were recruited 2013-2014. The two individuals from generation IV were recruited in September 2015. The arrow indicates the proband.

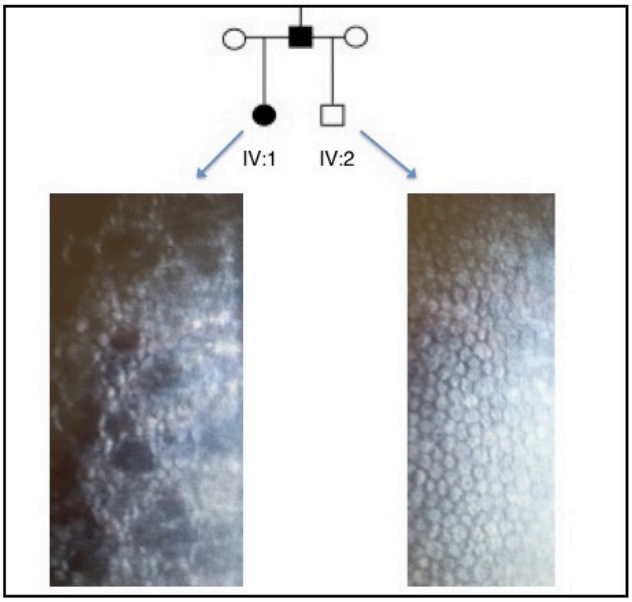


Figure 5.13 Specular microscopy scans of individuals IV:1 and IV:2 from FECDBAR family. The SM scan of IV:1 shows multiple gutatta with associated loss of endothelial cells consistent with FECD. The SM scan of IV:2 exhibits the normal arrangement of densely packed polygonal endothelial cells (Photos courtesy of Mr James Ball, Consultant Ophthalmic Surgeon, St James's University Hospital, Leeds).

5.2.5 Genotyping the multiplex FECD families for *TCF4* risk alleles of SNP rs613872 and the CTG18.1 trinucleotide repeat

Given that the *TCF4* CTG18.1 expansion is a common variant in FECD patients (Wieben et al., 2012) and has been shown to segregate with FECD in families (Mootha et al., 2014), all members of the three families for whom DNA was available were tested for the *TCF4* rs613872 SNP and the CTG18.1 trinucleotide repeat as described before (Sections 5.2.3.5 to 5.2.3.9).

The results for the FECDBRI family are summarised in Figure 5.14. All affected individuals in the family possessed the rs613872 TT genotype. The CTG18.1 expansion was also absent in all individuals that were tested suggesting that the cause of FECD in this family is unknown.

The results for the FECDWAK family are summarised in Figure 5.15. All the affected individuals possessed the rs613872 G allele and the CTG18.1

expansion that predisposes to FECD onset. Two of the affected individuals were found to be homozygous for the expansion. There was no clear phenotypic difference in terms of disease severity or age of onset of the FECD in those who had a single expanded allele when compared with those who had two expanded alleles. There were unaffected cases in generation III of FECDWAK who possessed the CTG18.1 expansion but were not old enough to manifest signs or symptoms of FECD.

The results for the FECDBAR family are summarised in Figure 5.16. Most of the affected individuals were found to possess the rs613872 G allele and a CTG18.1 expansion that could account for their FECD phenotype, with the exception of one subject who has an rs613872 TT genotype and no expanded alleles. It was concluded that this individual with no expanded alleles was likely to be a phenocopy for which the cause of FECD is unknown. Following the clinical analysis of FECD affected female IV.1 and her unaffected cousin IV.2, both were found to be heterozygous for the rs613872 G allele and the CTG18.1 expansion, suggesting that the unaffected cousin may be at risk of developing FECD at a later stage in life. It was also noted in this pedigree that affected family members with a *TCF4* CTG18.1 expanded allele existed in three successive generations and that the FECD phenotype appeared to present at an earlier age at each generation.

On the basis of this segregation analysis, it was concluded that the *TCF4* CTG18.1 expanded allele accounted for the FECD in FECDWAK and FECDBAR. In FECDBRI no CTG18.1 expansion was detected, therefore the causative mutation remained unknown and warranted further work.

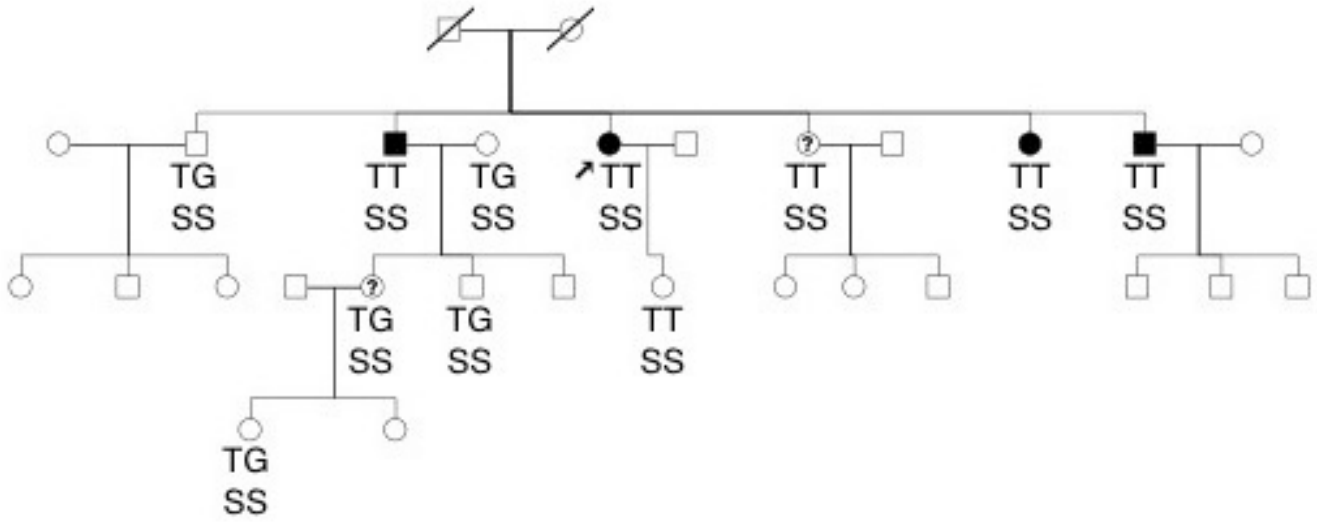


Figure 5.14 Segregation of *TCF4* rs613872 and the CTG18.1 trinucleotide repeat in FECDBRI. rs613872 is represented by the genotypes TT, TG or GG. SS = no expanded allele, SX = one expanded allele and XX = two expanded alleles. The arrow represents the proband.

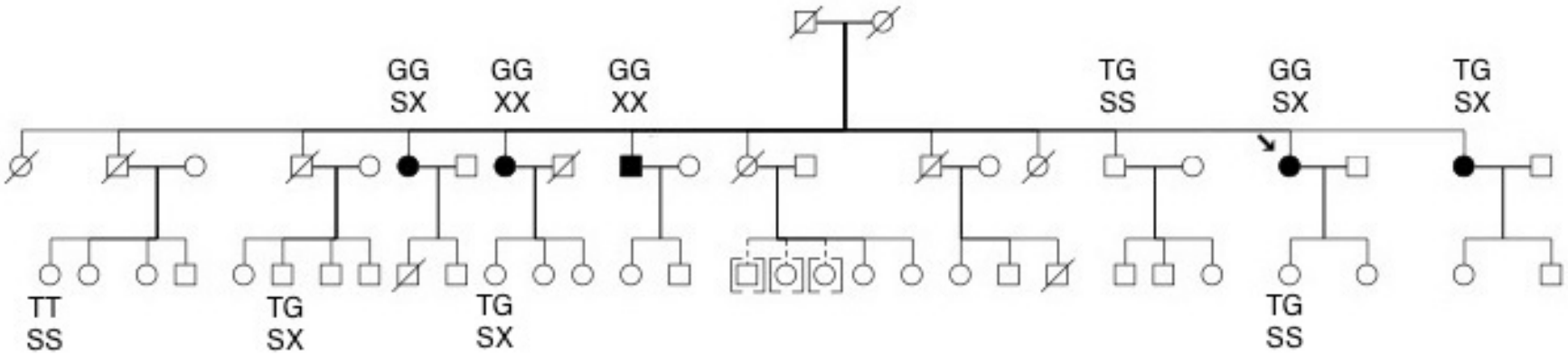


Figure 5.15 Segregation of *TCF4* rs613872 and the CTG18.1 trinucleotide repeat in FECDWAK. rs613872 is represented by the genotypes TT, TG or GG. SS = no expanded allele, SX = one expanded allele and XX = two expanded alleles. The individual in brackets denote that offspring were adopted. The arrow represents the proband.

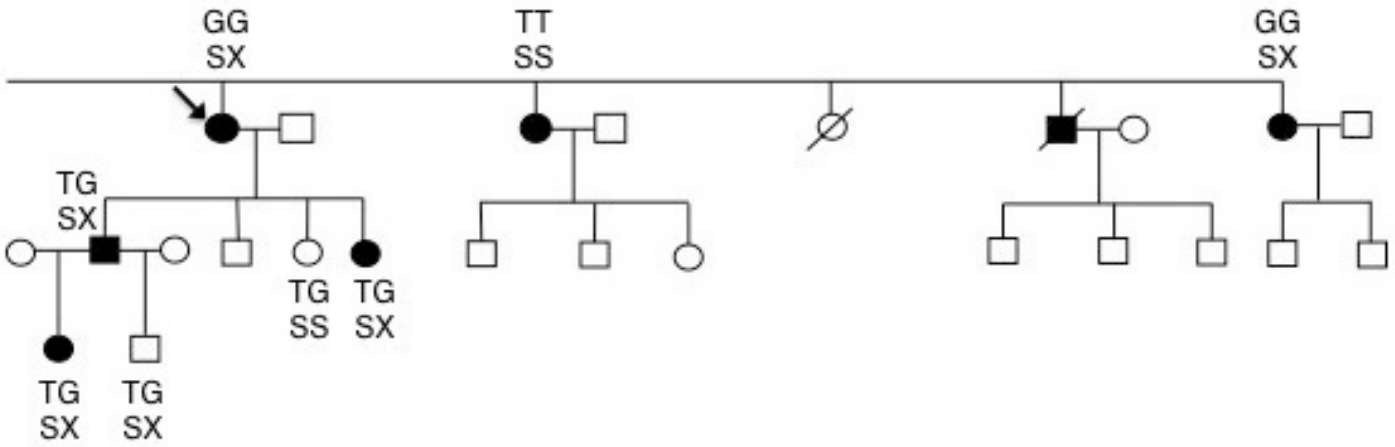


Figure 5.16 Segregation of *TCF4* rs613872 and the CTG18.1 trinucleotide repeat in FECDBAR. rs613872 is represented by the genotypes TT, TG or GG. SS = no expanded allele, SX = one expanded allele and XX = two expanded alleles. The part of the pedigree containing recruited patients is shown. The arrow represents the proband.

5.2.6 Next Generation Sequencing

To assist in finding mutations in genes causing FECD in the recruited families and cohort of FECD cases, RNA sequencing of normal corneal tissue was attempted to determine the normal corneal expression.

5.2.6.1 RNA-seq analysis of the normal corneal endothelium

Determining which genes are expressed in the normal corneal endothelium would provide a useful data set with which to prioritise candidate genes for analysis in inherited diseases of the corneal endothelium. A study by Chng and co-workers took this approach to identify the genes expressed in corneal endothelial cells from young and old research-grade donors tissue (tissue unsuitable for corneal grafting) as well as from cell culture, by RNA extraction using a Qiagen RNeasy column, then performing RNA sequencing (Chng et al., 2013).

This study attempted a similar analysis except that the tissue used here was corneal endothelial and epithelial/stromal tissue remaining from lamellar graft surgery (endothelial tissue remaining when the epithelium and anterior stroma was utilised for corneal grafting). This was collected from corneal operating theatres at SJUH and immediately stored in RNAlater. RNA was extracted simultaneously using 2 methods. These were either using the Trizol method according to manufacturer's instructions or the RNAeasy Plus Universal Minikit (Qiagen) (section 2.14).

Following Trizol extraction of 6 tissue samples, an aliquot of the RNA was subjected to reverse transcription PCR (RT-PCR) using primers against a ubiquitous housekeeping gene, p53, as described in section 2.14. A sample of the PCR product was analysed by agarose gel electrophoresis (Figure 5.17). The gel shows that no band could be seen for the endothelial samples, indicating the absence of a PCR product, which in turn shows that the RNA

extraction was unsuccessful. However, the presence of a band for PCR derived from the epithelial/stromal tissue confirms that RNA had been extracted from those samples.

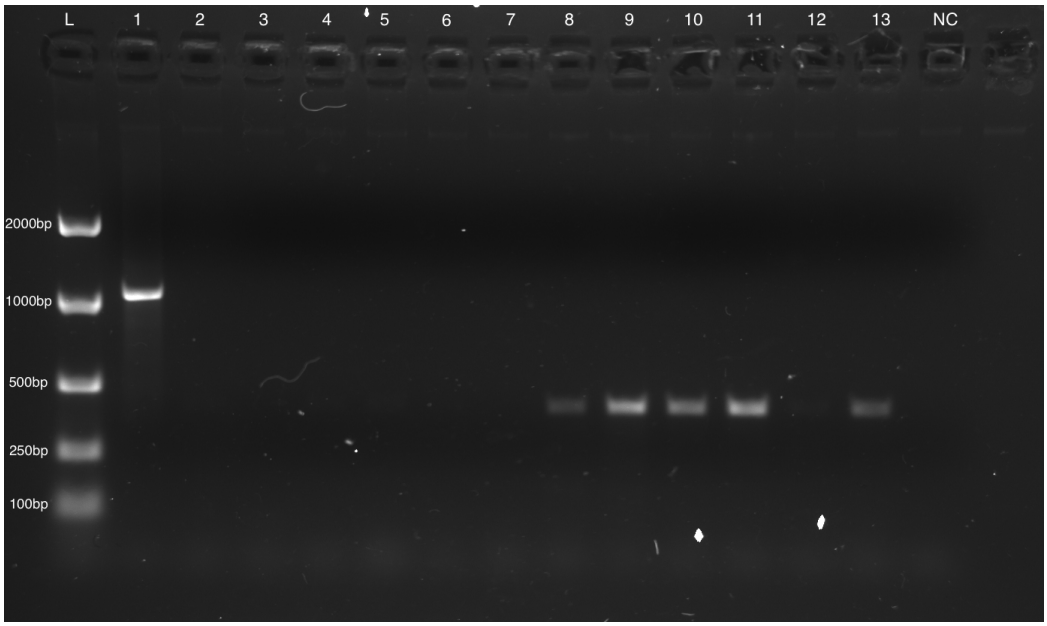


Figure 5.17 Agarose gel showing PCR amplification of cDNA generated by reverse transcription RNA from endothelial and epithelial/stromal tissue. L = Ladder, Lane 1 = Genomic DNA expected band size 1056bp, Lanes 2-7 = Endothelial samples, Lanes 8-13 Epithelial/Stromal Samples, RNA expected band size 407bp. NC = No DNA control. No band is visible in Lanes 2-7. There is a visible band in lanes 8-11 and 13.

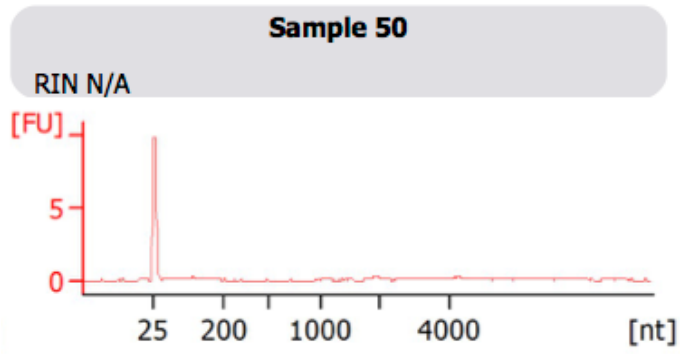
As a further verification, the RNA samples were run on an Agilent 2100 Bioanalyser, as this has the capacity to detect very small amounts of RNA (Figure 5.18). The electropherogram traces indicated that there was insufficient RNA extracted from the corneal endothelial samples to proceed with RNA sequencing. The RNA Integrity number (RIN) helps to estimate the integrity of the total RNA and is determined by the whole electrophoretic trace rather than just the ratio of the ribosomal bands alone. It can be used to directly compare the RNA integrity of different RNA samples.

Following RNA extraction using the RNAeasy Plus Universal Minikit (Qiagen) on 6 further samples and Bioanalyser measurements, the total RNA concentration was 2ng/µl, which again was inadequate to proceed with RNA

sequencing. Therefore, RNA sequencing of corneal endothelial samples was not possible.

One possible explanation for this was that the RNA had degraded during the time it was stored in RNAlater. Therefore RNA extractions were carried out on 6 freshly dissected bovine endothelial samples (section 2.14) using the RNAeasy Plus Universal Minikit (Qiagen). Although the RNA concentrations were greater than detected in human endothelium, average 8ng/µl, this was the total RNA, a proportion of which would be ribosomal RNA.

A.



B.

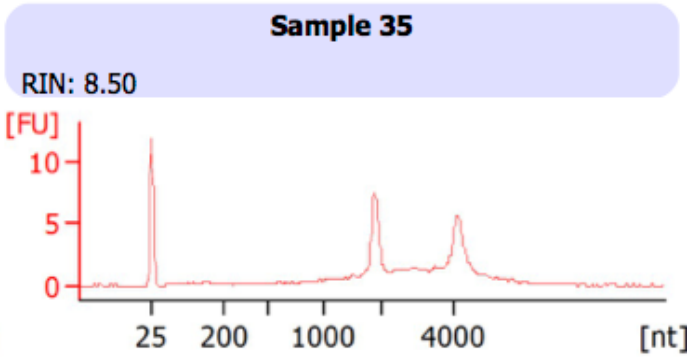


Figure 5.18 Bioanalyser electropherogram traces for the detection of RNA concentration and quality. The two samples are (A) an endothelial corneal tissue sample at an RNA concentration of 3µg/l and RNA Integrity Number (RIN) not detectable, and (B) an epithelial/stromal corneal tissue sample at 35ng/µl with RIN 8.5. FU = Fluoresence units, nt = nucleotide.

5.2.6.2 Whole Exome sequencing of FECD Families

The whole exome sequencing of FECD family members was carried out over a number of years as samples became available and in small batches to reduce the cost. The WES analysis of FECDBAR and FECDWAK was performed alongside the *TCF4* CTG18.1 expansion analysis described in section 5.2.5. As it was concluded from this that the *TCF4* trinucleotide repeat expansion was causative in these two families, an assessment of the genes known to cause FECD was carried out. No further analysis of WES data was performed in these two families.

5.2.6.2.1 Whole Exome Sequencing Library Preparation and Evaluation of the known FECD genes in the three FECD families

FECDBRI samples II:3, II.5 and II:10 were processed commercially by Otogenetics Corporation whereas samples II:9 and II:2 were processed by the author and run on the in-house sequencer (section 2.15). Additionally, FECDWAK II:6, II:8, II:10, II:17, II:19, II:21 and FECDBAR II:17, II:19, II:24 III:28 and III:32 were sequenced in-house.

A representative Bioanalyser trace after the different stages of library preparation is shown in Figure 5.19.

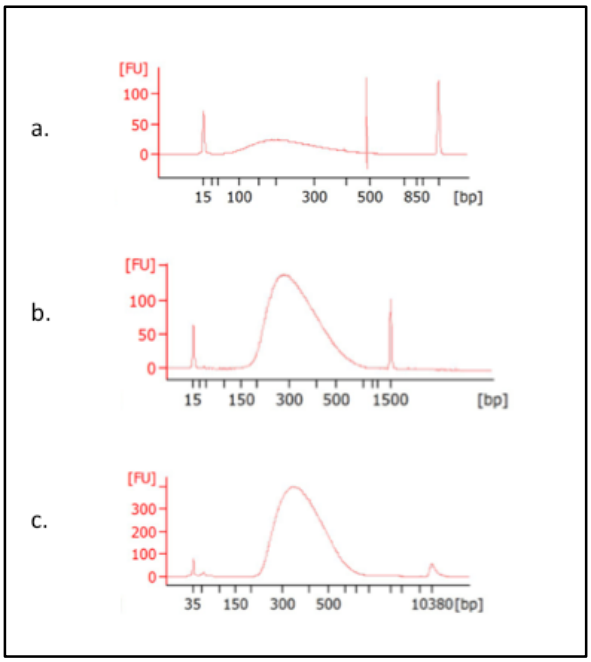


Figure 5.19 Bioanalyser traces of sheared genomic DNA from the different stages of library preparation prior to whole exome sequencing for FECDBAR II:17. The stages of library preparation are (a) after shearing, (b) after amplification and (c) after hybridisation of the DNA during library preparation for exome sequencing. FU = Fluorescence units, bp = base pairs

The resulting FASTQ next generation sequencing output files were analysed as detailed in Section 2.22, looking for variants common to all the affected cases in the family but absent from the unaffected case if there was one sequenced from the family.

Analysis of shared variants was carried out using two methods. Firstly, following annotation of the variants using Annovar (http://annovar.openbioinformatics.org/en/latest/), the individual lists of annotated variants from affected individuals from the same family were merged so that variants shared by affected individuals could be analysed using the Agile Variant Selector (software designed by Dr Ian Carr, LIMM Bioinformatician). This was carried out on all three families.

Pipeline 2 was a more advanced pipeline developed in 2015 after refinements were made to Pipeline 1 (Section 2.21) and also incorporated CADD scores at the annotation stage, therefore the variant list was further prioritised based on this. Pipeline 1 utilised Bowtie2 for the alignment whereas in pipeline 2

Novoalign was used. Additionally, Pipeline 2 incorporated the use of the VCF on VCF perl script (https://github.com/gantzgraf/vcfhacks) developed by Dr David Parry. This script merged the individual VCF files into a single file of shared variants, which was then subsequently annotated using Annovar. Both pipelines allowed the subtraction of variants from an unaffected individual from the total list of shared variants, which was carried out in FECDBRI. Read depth was not used for filtering purposes but was taken into account for prioritisation of variants. This was carried out on FECDBRI.

Prior to filtering the refined list, variants in the known genes that cause FECD *COL8A2, SLC4A11, TCF4, TCF8, LOXHD1 and AGBL1* were selected in all three families

For FECDBRI, there were 39 variants present, two of which were in the coding regions or splice site recognition signal in *SLC4A11* (Table 5.10A). One was a synonymous variant c.G639A: p.S213S and the other was a non-synonymous variant c.C77G: p.P26R with a frequency of 0.64 in 1000G, graded as non-pathogenic by the pathogenicity prediction tools Polyphen, SIFT and Mutation Taster. These variants were therefore excluded. Given the tentative linkage of FECDBRI to chromosome 18 (Section 5.2.3.1), an intronic *TCF4* A>G change at position 53,177,742 from the top of chromosome 18 (according to the human Genome Browser version, hg19) was also highlighted. This variant, termed rs192075715 in the dbSNP database, has a frequency of 0.32% in the 1000 Genomes database.

Primers were designed across rs192075715 that were used in the PCR to confirm the segregation of this variant in all available FECDBRI family members (Appendix IV). This variant was found to segregate with FECD in FECDBRI and was also present in the individual of unknown disease status II:7 (Table 5.12).

As this intronic variant rs192075715 was only 68 base pairs away from an exon, the possibility that it might affect splicing was considered. The wild type and variant sequences entered into the BGBP splicing prediction website

(http://www.fruitfly.org/seq_tools/splice.html) (Section 2.17.7) predicted that there was no effect on splicing.

In FECDWAK, there were 96 variants present in the known genes, all of which were intronic except for one synonymous coding variant in *TCF4* rs143944746 (0.09% frequency) (Table 5.10B). All variants were present in the 1000 genomes and/or EVS databases at a frequency greater than 2% except for two variants in *TCF4*; rs143944746 and a *TCF4* intronic SNP at position 53173173 T>A rs17512480 (frequency 1%).

In FECDBAR, there were 199 variants in the genes known to be causative for FECD, 20 of which were coding variants (Table 5.10C). Of these 20, 7 were common non-synonymous SNPs graded as non-pathogenic by the pathogenicity prediction tools Polyphen, SIFT and Mutation Taster. One intronic SNP in TCF4 G>A at position 52942827 (rs1788027), which had a 1000Genome frequency of 50% was common to both FECDWAK and FECDBAR.

As the CTG18.1 expanded allele was found to segregate with the FECD phenotype in FECDBAR and FECDWAK (section 5.2.5), and as no clear pathogenic variants were detected, no further analysis of the WES data on these families was carried out in these families. FECBRI was selected for further analysis of variants from WES.

A.

Chr	Start	End	Ref	Alt	Gene.refGene	ExonicFunc.refGene	Otherinfo	AAChange.refGene	EVS	1000g	snp137	SIFT	Polyphen2	MutationTaster
chr20	3214581	3214581	C	T	*SLC4A11*	synonymous SNV	het	NM_032034:exon5:c.G639A:p.S213S,SLC4A11	0.186222	0.16	rs3803956	Unknown	Unknown	Unknown
chr20	3218634	3218634	G	C	*SLC4A11*	nonsynonymous SNV	het	NM_001174090:exon1:c.C77G:p.P26R	Unknown	0.64	rs3810562	0.66	Benign	Polymorphism (Automatic)

B.

Chr	Start	End	Ref	Alt	Gene.refGene	ExonicFunc.refGene	Otherinfo	AAChange.refGene	EVS	1000g	snp137	SIFT	PolyPhen2	MutationTaster
chr18	52901846	52901846	C	G	*TCF4*	synonymous SNV	het	NM_001243226:exon17:c.G1725C:p.P575P	0.001922	0.0009	rs143944746	Unknown	Unknown	Unknown

C.

Chr	Start	End	Ref	Alt	Gene.refGene	ExonicFunc.refGene	Otherinfo	AAChange.refGene	EVS	1000g	snp137	SIFT	PolyPhen2	MutationTaster
chr15	86800209	86800209	C	T	*AGBL1*	synonymous SNV	hom	NM_152336:exon7:c.C723T:p.P241P	0.645368	0.52	rs1353578	Unknown	Unknown	Unknown
chr15	86806029	86806029	C	T	*AGBL1*	synonymous SNV	hom	NM_152336:exon9:c.C852T:p.D284D	0.691817	0.56	rs1566088	Unknown	Benign	Unknown
chr15	86807542	86807542	C	T	*AGBL1*	synonymous SNV	hom	NM_152336:exon10:c.C1002T:p.A334A	0.160255	0.12	rs10520617	Unknown	Unknown	Unknown
chr15	86807843	86807843	T	C	*AGBL1*	nonsynonymous SNV	hom	NM_152336:exon10:c.T1303C:p.S435P	0.169	0.12	rs11857527	Tolerated	Unknown	Polymorphism
chr15	86807884	86807884	A	C	*AGBL1*	synonymous SNV	hom	NM_152336:exon10:c.A1344C:p.V448V	0.174007	0.12	rs11856833	Unknown	Unknown	Unknown
chr15	86814866	86814866	A	G	*AGBL1*	synonymous SNV	hom	NM_152336:exon14:c.A1866G:p.L622L	0.731923	0.64	rs11858435	Unknown	Benign	Unknown
chr15	86940622	86940622	T	C	*AGBL1*	synonymous SNV	het	NM_152336:exon17:c.T2262C:p.T754T	0.607208	0.7	rs4362360	Unknown	Benign	Unknown
chr15	86940673	86940673	C	T	*AGBL1*	synonymous SNV	het	NM_152336:exon17:c.C2313T:p.T771T	0.109367	0.09	rs1367647	Unknown	Unknown	Unknown
chr15	87217613	87217613	A	G	*AGBL1*	nonsynonymous SNV	hom	NM_152336:exon22:c.A3029G:p.Q1010R	0.8897	0.87	rs8028043	Tolerated	Unknown	Polymorphism
chr15	87531281	87531281	A	C	*AGBL1*	synonymous SNV	hom	NM_152336:exon23:c.A3147C:p.T1049T	0.806037	0.82	rs1006030	Unknown	Unknown	Unknown
chr18	44063598	44063598	G	A	*LOXHD1*	nonsynonymous SNV	het	NM_144612:exon39:c.C6107T:p.A2036V	NA	0.41	rs1377016	Tolerated	Benign	Polymorphism
chr18	44126909	44126909	T	C	*LOXHD1*	nonsynonymous SNV	het	NM_144612:exon22:c.A3463G:p.R1155G	0.771572	0.7	rs1893566	Tolerated	Unknown	Polymorphism
chr18	52895531	52895531	T	C	*TCF4*	synonymous SNV	het	NM_001243226:exon20:c.A2247G:p.S749S	0.364678	0.38	rs8766	Unknown	Unknown	Unknown
chr18	53303101	53303101	C	G	*TCF4*	nonsynonymous SNV	hom	NM_001243226:exon1:c.G28C:p.A10P	0.998774	1	rs611326	Tolerated	Unknown	Polymorphism
chr20	3210301	3210301	G	A	*SLC4A11*	synonymous SNV	het	NM_032034:exon13:c.C1659T:p.N553N,SLC4A11	0.093803	0.06	rs41281860	Unknown	Unknown	Unknown
chr20	3211235	3211235	C	T	*SLC4A11*	synonymous SNV	het	NM_032034:exon11:c.G1389A:p.T463T,SLC4A11	0.090804	0.06	rs6084312	Unknown	Benign	Unknown
chr20	3214819	3214819	T	G	*SLC4A11*	synonymous SNV	het	NM_032034:exon4:c.A481C:p.R161R,SLC4A11	0.480932	0.49	rs3827075	Unknown	Unknown	Unknown
chr20	3218563	3218563	T	C	*SLC4A11*	nonsynonymous SNV	het	NM_001174090:exon1:c.A148G:p.R50G	NA	0.06	rs79057061	Tolerated	Benign	Polymorphism (Automatic)
chr20	3218634	3218634	G	C	*SLC4A11*	nonsynonymous SNV	het	NM_001174090:exon1:c.C77G:p.P26R	NA	0.64	rs3810562	Tolerated	Benign	Polymorphism (Automatic)

Table 5.10 The list of coding variants in the FECDBRI, FECDWAK and FECDBAR affected cases in genes known to cause FECD. The databases used were Polyphen2 (HDIV), SIFT, PROVEAN, and Mutation Taster. Additionally their frequency in the 1000 genomes and EVS databases are shown (Genome Browser version hg19).

5.2.6.2.2 Analysis of WES in FECDBRI to identify the pathogenic mutation causing FECD

The refined list containing variants shared by all the affected cases in family FECDBRI was filtered by removing the homozygous variants, followed by the synonymous and intronic variants. Spice site variants at positions +/-1-5, along with exons and exon/splicing variants (variants within the exon but close to the intron/exon boundary) were included in the analysis. Variants that had a frequency of more than 2% within EVS or 1000 Genomes databases were also excluded. The reason for choosing this threshold was that as TCF4 accounted for the majority of FECD, it was felt likely that the remainder of variants would be rare. This left a list of 14 variants, which are shown in Table 5.11A. The analysis was also repeated using the more advanced Pipeline 2 in 2015, and after applying CADD scores to the list of variants, the ones with a CADD score greater than or equal to 10 were retained. 5 variants were highlighted (Table 5.11B). These were all variants that had been found using Pipeline 1. This included variants in Oxysterol-binding Protein-like Protein 1A (*OSBPL1A)* c.115_116insAATT: p.C39_K40delinsX, Tandem C2 domains, nuclear *(TC2N)* c.T962C: p.I321T, Tetratricopeptide repeat domain 40 *(TTC40)* c.T1814C: p.L605P, URB1 Ribosome biogenesis 1 Homolog (*URB1)* c.C2126T: p.A709V and Crystallin, Zeta-like 1 (*CRYZL1)* c.C461G: p.A154G.

A.

Gene.refGene	ExonicFunc.refGene	Otherinfo	AAChange.refGene	EVS	1000g	snp137	SIFT	Polyphen2	MutationTaster
TTC40	nonsynonymous SNV	het	NM_001200049:exon15:c.T1814C: p.L605P	Not present	0.0005	rs146983053	Damaging	Probably Damaging	Deleterious
TDRD3	nonsynonymous SNV	het	NM_030794:exon13:c.G1850A: p.G617D	0.000846	0.0009	rs140870765	Tolerated	Benign	Deleterious
SLAIN1	nonsynonymous SNV	het	NM_001242868:exon4:c.C1069G: p.P357A	0.001768	0.0023	rs144139933	Tolerated	Damaging	Deleterious
TC2N	nonsynonymous SNV	het	NM_001128596:exon9:c.T962C: p.I321T	0.000231	0.0005	rs148609061	Damaging	Benign	Deleterious
URB1	nonsynonymous SNV	het	NM_014825:exon17:c.C2126T: p.A709V	Not present	0.0005	rs190315797	Tolerated	Benign	Polymorphism (Automatic)
RPUSD3	nonsynonymous SNV	het	NM_173659:exon8:c.C799T: p.R267C	0.002691	0.0023	rs146799821	Tolerated	Benign	Deleterious
IL17RC	nonsynonymous SNV	het	NM_153460:exon11:c.G970A: p.A324T	0.013314	0.01	rs115461448	Tolerated	Damaging	Polymorphism (Automatic)
HOMEZ	nonframeshift deletion	het	NM_020834:exon2:c.1634_1636del: p.545_546del	Not present	Not present	rs148005528	NA	NA	Unknown
KCNJ12,KCNJ18	nonsynonymous SNV	het	NM_021012:exon3:c.C1206A: p.D402E	Not present	Not present	rs2917720	Tolerated	Benign	Polymorphism (Automatic)
KCNJ12,KCNJ18	nonsynonymous SNV	het	NM_021012:exon3:c.G1214T: p.S405I	Not present	Not present	rs73979902	Tolerated	Benign	Deleterious
KCNJ12,KCNJ18	nonsynonymous SNV	het	NM_021012:exon3:c.A1289G: p.E430G	Not present	Not present	rs5021699	Damaging	Damaging	Deleterious
OSBPL1A	stopgain SNV	het	NM_080597:exon2:c.115_116insAATT: p.C39_K40delinsX	0.002796	Not present	NA	Unknown	NA	Unknown
CRYZL1	nonsynonymous SNV	het	NM_145858:exon7:c.C461G: p.A154G	Not present	Not present	NA	Tolerated	Possibly Damaging	Deleterious
RP1L1	nonsynonymous SNV	het	NM_178857:exon4:c.C3956G: p.A1319G	Not present	Not present	rs4840501	Tolerated	Benign	Polymorphism (Automatic)

B.

symbol	variant	gene	feature	allele	consequence	cds_position	protein_position	amino_acids	codons	existing_variation	exon	CaddPhredScore
OSBPL1A	c.115_116insAATT: p.C39_K40delinsX	ENSG00000141447	ENST00000319481	AATT	stop_gained&frameshift_variant	115-116	39	C/*LX	tgc/tAATTgc	~rs74793804	2/28	35
TC2N	c.T962C: p.I321T	ENSG00000165929	ENST00000435962	G	missense_variant	962	321	I/T	aTt/aCt	rs148609061	9/12	25.2
TTC40	c.T1814C: p.L605P	ENSG00000171811	ENST00000368586	G	missense_variant	1814	605	L/P	cTc/cCc	rs146983053	15/58	25
URB1	c.C2126T: p.A709V	ENSG00000142207	ENST00000382751	A	missense_variant	2126	709	A/V	gCg/gTg	rs190315797	17/39	23
CRYZL1	c.C461G: p.A154G	ENSG00000205758	ENST00000381554	C	missense_variant	461	154	A/G	gCa/gGa		7/13	21

Table 5.11 Filtered variants in FECDBRI family common to the affected cases but absent from the unaffected member after whole exome sequencing using A. Pipeline 1 (annotated by Annovar) and B. Pipeline 2 (annotated by VEP) in FECDBRI. Pipeline 2 was a more advanced pipeline developed in 2015 and incorporated the use of CADD scores (http://cadd.gs.washington.edu). The 5 variants in B were all seen in A. and all had CADD scores of above 10 (Genome Browser version hg19).

The variants were Sanger sequenced in the FECDBRI family to confirm their presence in the family members that were exome sequenced and also to assess segregation in other family members (Table 5.12). The findings showed that all of the variants segregated perfectly with those known to have FECD, some segregated with the individual with an unknown disease status from generation II, individual II:7. Additionally, family members from generation III and IV might have been asymptomatic mutation carriers who were not yet old enough to have developed the condition, and so caution ought to be considered in interpreting the analysis.

FECDBRI	rs192075715	OSBPL1A c.115_116insAATT:p.C39_K40 delinsX	TC2N c.T962C:p.I321T	TTC40 c.T1814C:p.L605P	URB1 c.C2126T:p.A709V	CRYZL1 c.C461G:p.A154G
II:2	AA	WT	TT	TT	CC	CC
II.3	AG	INS	TC	TC	CT	CG
II:4	AA	WT	TT	TT	CC	CC
II:5	AG	INS	TC	TC	CT	CG
II:7	AG	INS	TC	TC	CC	CC
II:9	AG	INS	TC	TC	CT	CG
II:10	AG	INS	TC	TC	CT	CG
III:5	AG	WT	TT	TT	CT	CG
III:6	AA	INS	TT	TT	CT	CG
III:8	AA	WT	TC	TT	CC	CC
IV:1	AA	WT	TT	TT	CT	CG

Table 5.12 Segregation analysis of the 5 coding variants identified in the FECDBRI family after exome sequencing. Affected individuals are highlighted in green, unaffected in grey and those of uncertain diagnosis in pink. The variant residues are highlighted in blue. Segregation of the intronic SNP, rs192075715, is also shown as a comparison. All variants shown are heterozygous.

The *OSBPL1A* variant c.115_116insAATT: p.C39_K40delsinX represents a null variant whereas the other 4 variants are missense changes. Protein sequence alignments of the residues affected by these missense changes in orthologues highlight that the normal residue in *CRYZL1* is fully conserved through evolution while the residues altered by other variants are not (Figure 5.20). It is worth highlighting that *CRYZL1, URB1* and *OSBPL1A* were found to be expressed at significant levels in corneal endothelial cells whereas *TC2N* and *TTC40* were not (Table 5.14). *OSBPL1A* is on chromosome 18, but not within or near the published locus (Sundin et al., 2006a). The other variants were not within the other published loci (Sundin et al., 2006b, Riazuddin et al., 2009, Riazuddin et al., 2010a). Based on this evidence, it seemed likely that *CRYZL1* and *OSBPL1A* were the best gene candidates for further screening. However due to time constraints, Sanger sequencing in an

FECD case/control cohort was not possible. A summary of these variants is provided below.

5.2.6.2.3 Oxysterol-binding Protein-like Protein 1A

Oxysterol-binding Protein-like Protein 1A (*OSBPL1A)* (MIM*606730) has 28 exons and encodes a member of the oxysterol-binding protein (OSBP) family, a group of intracellular lipid receptors (Jaworski et al., 2001). It is differentially regulated by *FOXC1*, which has a role in cell viability and resistance to oxidative stress in the eye, and has been implicated in Axenfeld-Rieger syndrome, a condition characterised by anterior segment malformations in the eye and glaucoma (Berry et al., 2008). As detailed above, it has been shown to be expressed in the corneal endothelium (Chng et al., 2013) (Table 5.14). The insertion/frameshift variant identified has an allele frequency of 0.18% in ExAc (http://exac.broadinstitute.org). It was originally assigned a reference of rs74793804 but was removed from dbSNP in 2013 due to mapping errors (http://www.ncbi.nlm.nih.gov/SNP/).

5.2.6.2.4 Crystallin, Zeta-like 1

Crystallin, zeta(quinone reductase)-like 1 (*CRYZL1)* (MIM*603920) is a 13 exon gene on chromosome 21 which encodes a protein that has sequence similarity to zeta crystallin, also known as quinone oxidoreductase (Kim et al., 1999). The missense variant identified is present in ExAc (http://exac.broadinstitute.org) at a frequency of 0.0001%, and the transcript is expressed in the corneal endothelium (Chng et al., 2013) (Table 5.14). A study assessing the differential response of lens and corneal crystallins in degenerative corneas suggested a potential role of crystallins in the maintenance of corneal clarity, although the study concluded that lens crystallins might play a greater role in this than corneal crystallins. (Gong et al., 2012).

A. URB1 c.C2126T:p.A709V

```
Human       688   TMEEDKETVIQFLERILLTLVANPYSYTDKASDFVQEASMLQATMTKQEA   737
Chimp       688   TVEEDKETVIQFLERILLTLVANPYSYTDKASDFVQEASTLQATMTKQEA   737
Dog         688   TVEEKQETVIQFLERVLLTLVANPYLYTDKASDFVQEASTLQATGTKQDA   737
Cow         688   MAEENKEAVIQFLERVLLTLVANPYSYTDKASDFVQEASALQAPGTRQDA   737
Mouse       691   TAEERKEAVIQFLERILLTLVMNPYSYTDKASEFVQEASTLQASMGKQDA   740
Rat         691   TAEKHKETVIQFLERILLTLVVNPYSYTDKASEFVQEASTLQGAVGKQDA   740
Chicken     695   TKEDKKEAVIQFLERILLKLVTNPYPYTDKAADLVQEASMLQVNVFKQDS   744
Zebrafish   678   LGLSQQETVIQFLDHVLTRVVCNSHVYTDKAASMVQEAANLQANLSGQDG   727
Frog        695   VTPDSKEAVIQFLEQVLVKLVSNPYPYTDKAAESVQEASMLQSSLSKQDS   744
```

B. CRYZL1 c.C461G:p.A154G

```
Human       143   SPGKSVLIMDGASAFGTIAIQLAHHRGAKVISTACSLEDKQCLERFRP--   190
Chimp       143   SPGKSVLIMDGASAFGTIAIQLAHHRGAKVISTACSLEDKQCLERFRP--   190
Monkey      143   SPGKSVLIMDGASAFGTIAIQLAHHRGAKVISTACSLEDKQCLERFRP--   190
Dog         143   SPGKSVLIMDGASALGMIAIQLAHHRGAKVISTAGSLEDKQCLERLRP--   190
Cow         143   SPGKSVLIMDGASALGTIAIQLAHHRGAKVISTACSLEDKQFLERFRP--   190
Mouse       143   SPGKSVLIMDGASAFGTIAIQLAHHRGAKVISTAHSLEDKQHLERLRP--   190
Rat         143   SPGKSVLIMDGASAFGTIAIQLAHHRGAKVISTAHSLEDKQHLERLRP--   190
Chicken     151   SPGKTVLVMDGASPFGTIAIQLAQHRGAKVISTAYSLEDKQYLERLRPAV   200
Zebrafish   142   AAGQTVLVLDGASPFGVLAIQLAHYHGVKVLATALSPEDQKFLQELRP--   189
Frog        143   GPGKSVLVMDGASPLGALAVQLSLHRGATVITTADSEESKLYFEKLRP--   190
```

C. TC2N c.T962C:p.I321T

```
Human       298   KLQNLQTVRLVFKIQTQTPRKKTIGECSMSLRTLSTQEMDYSLDITPPSK   347
Chimp       298   KLQSLQTVRLVFKIQTQTPRKKTIGECSLSLRTLSTQEMDYSLDITPPSK   347
Monkey      298   KLQSLQTVRLIFKIQTQTPRKKTIGKCSLSLRTLSTQEMDYSLDITPPSK   347
Dog         298   KLQSLQTVRLVFKIQTQTPRKKTIGECSLSLRTLSTQEMDYSLDIMPPSK   347
Cow         298   KLQSLQTVRLVFKIQTQTPRKKTIGECSLSLRTLSTQEMDYSLDISPPSK   347
Mouse       297   KLQNLQAVRLAFKIQTQTPKKKTIGECSLSLRTLSTQEMEYSLEIIAPSK   346
Rat         297   KLQNLQAVRLAFKIQTQTPKKKTIGECSLSLRTLSTQEMDYSLEIIPPSK   346
Chicken     297   KLQSVQAVRLVFKIQMQTPRKRTIGECSLSLRELSSEESNHWLDISPPSK   346
Zebrafish   283   NLEQIRRSALVLRLQAHTPRKRTLGECVLSLRSLGPQETKHWLEFKHPSK   332
Frog        295   KHQSLQASRLVFKVLSLTPRKRTISHCSISLRDLSEEETDHWLELVPSSK   344
```

D. TTC40 c.T1814C:p.L605P

```
Human       603   FCLLYDNVK--VKK---LRLRRGKKK---RGRDGSVQDTWSQPEVVLQRQ   644
Chimp       548   FCLLYDNVX--VKN---LKLRRGKKK---QGRDGSVQDTWSQPEVVLQRQ   589
Dog         545   FCLLYDNVK--VKKPTWLKKGTGRRK---KGGDSSAQDSWGPSEVPLQKQ   589
Cow               ----------------------------------------------------
Rat         603   FCLLYDNAK--SKKS--TKPKRGKKK---KGMEGSVTE--GPSEVVLNKQ   643
Zebrafish   735   FCLLYDDGRWRNT----------------------------GEQESAVSPE   757
Frog        308   FCLLYDDGRWNIQKKDTLRKNASTANSMDEGKGLGLEGELSNAKAAFTNE   357
```

Figure 5.20 Protein sequence alignment of WES variants segregating in FECDBRI with orthologues. A. URB1 c.C2126T:p.A709V, B. CRYZL1 c.C461G:p.A154G, C. TC2N c.T962C:p.I321T and D TTC40 c.T1814C:p.L605P. Analysis was performed using Homologene. Each variant is highlighted in red and the surrounding amino acids that are identical to the human transcript are shaded in grey. Accession numbers for URB1 are NP_055640.2 (human), XP_531425.3 (chimp), XP_005638888.1 (dog), NP_001192909.1(cow), NP_083773.1 (mouse), NP_001178590.1 (rat), XP_003640527.1 (chicken), XP_002664599.3 (zebrafish), NP_001131087.1 (frog) for CRYZL1 are NP_114423.1 (human), XP_531441.3 (chimp), XP_001090445.1 (monkey), XP_535585.1 (dog), NP_001030209.2 (cow), NP_598440.1 (mouse), NP_001013062.1 (rat), XP_004934579.1 (chicken), NP_001002633.1 (zebrafish), NP_001016084.1 (frog) for TC2N are NP_689545.1(human),XP_001145068.1 (chimp),XP_001091431.1(monkey), XP_547711.1(dog), NP_001180133.1 (cow), NP_001273293.1 (mouse), NP_001020323.1 (rat), XP_004941928.1 chicken, XP_005157017.1 zebrafish), XP_002933250.2(frog) and for TTC40 are NP_001186978.2 (human), XP_003951835.1 (chimp), XP_005637957.1 (dog), XP_005225941.1 (cow), XP_006230566.1 (rat), XP_005157133.1 (zebrafish), XP_002936855.2 (frog).

5.2.6.2.5 ExomeDepth analysis of whole exome sequencing data from case II.9 in the FECDBRI family

In order to investigate whether the causative mutation in affected cases of the FECDBRI family was a large copy number variation (CNV) or other structural variation that affects the coding region captured by exome sequencing, ExomeDepth analysis was carried out on the WES from affected case II:9. The FASTQ files were processed as described in Section 2 using Pipeline 1, up to and including the removal of duplicate reads by Picard (Table 2.2). The programme requires the exome samples to have good depth of coverage. The sample that met these requirements was sequenced in the Next Generation facility in Leeds. The other affected samples were sequenced at Otogenetics, and therefore had reduced depth of coverage compared to the data from individual II:9. For ExomeDepth, control samples were also required, preferably those run on the same Illumina sequencing lane as the test sample. The other 4 samples run on the same lane as individual II.9 were FECD samples from FECDBAR and FECDWAK.

A list of Bayes Factor (BF) scores are shown in Table 5.13. The higher the BF score, the more confidence of the presence of a CNV. The BF score calculates the log10 of the likelihood of the testing data having a CNV relative to the normal copy number at that position using the control exomes. The BF does not however have a clearly defined threshold.

The highest BF score in case II.9 was a heterozygous duplication at chromosome 10, positions 46,965,003-47,087,911 (hg19) that includes the start of *SYT15*, the entire *GPRIN2* and end of *NPY4R* genes and encompasses 8 exons (Figure 5.21). It was noted that only *SYT15* and *GPRIN2* are both expressed in the corneal endothelium (Table 5.14) so any aberrant transcripts as a result of the duplication event could be considered candidates for FECD.

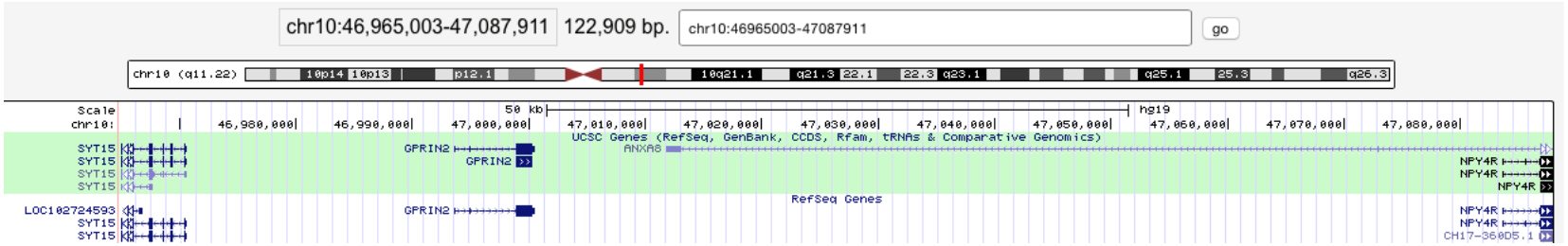


Figure 5.21 The genomic region identified to have a putative CNV in case II.9 from FECDBRI family.
The interval with a putative CNV is chr10:46965003-47087911 as indicated on the Genome browser version hg19. This 122,909bp duplication encompasses the start of the *SYT15*, the entire *GPRIN2* and the end of *NPY4R* genes.

start.p	end.p	Type	nexons	start	end	chromosome	id	BF	reads.expected	reads.observed	reads.ratio	Conrad.hg19	exons.hg19
21864	21871	duplication	8	46965003	47087911	chr10	chrchr10:46965003-47087911	44.1	2912	4112	1.41	NA	NA
154672	154692	deletion	21	32485517	32632844	chr6	chrchr6:32485517-32632844	29.6	134	1	0.00746	NA	NA
167146	167151	deletion	6	100331793	100336236	chr7	chrchr7:100331793-100336236	29.4	394	187	0.475	NA	NA

Table 5.13 ExomeDepth results of individual II:9 from the FECDBRI family. The variant with the highest BF score is shown in the top row. The read ratio of 1.41 indicates that this is a heterozygous duplication event encompassing 8 exons. None of the structural variants listed were present in the Conrad database (Conrad et al., 2010).

			Old endothelial cells			Cultured endothelial cells			Young endothelial cells			Stroma		
RefSeq ID	**Gene ID**	**Total Length**	**Total Count**	**RPKM**	**RPM**	**Total Count**	**RPKM**	**RPM**	**Total Count**	**RPKM**	**RPM**	**Total Count**	**RPKM**	**RPM**
NM_018030	OSBPL1A	3301	0	0.00	0.00	22	1.44	4.75	2	0.22	0.72	30	5.04	16.65
NM_145858	CRYZL1	1697	0	0.00	0.00	43	5.47	9.28	98	20.73	35.17	2	0.65	1.11
NM_014825	URB1	10769	0	0.00	0.00	50	1.00	10.79	1	0.03	0.36	4	0.21	2.22
NM_031912	SYT15	5500	3	0.29	1.57	104	4.08	22.44	5	0.33	1.79	0	0.00	0.00
NM_014696	GPRIN2	1820	0	0.00	0.00	1	0.12	0.22	10	1.97	3.59	0	0.00	0.00

Table 5.14 RNA expression levels in old, cultured and young endothelia cells and stroma (Chng et al., 2013). Levels of expression in the corneal endothelium are shown for transcripts of genes containing candidate FECD causal variants identified in exome sequencing; *OSPBL1A, CRYZL1* and *URB1*; or in Exomedepth analysis of WES; *STY15* and *GPRIN2.* RPKM = Reads Per Kilobase of transcript per Million, RPM = reads per million

5.3 Discussion

FECD is a genetically heterogeneous disease, with mutations in *COL8A2*, *SLC4A11*, *LOXHD1*, *TCF4*, *ZEB1* and *AGBL1* having been identified as causative (Biswas et al., 2001, Vithana et al., 2008, Riazuddin et al., 2012, Wieben et al., 2012, Riazuddin et al., 2010a, Riazuddin et al., 2013). In this chapter, families and cases were examined to look for the genetic contributors to FECD. An assessment of the mutational load of *LOXHD1* variants was carried out in cases and controls, and the results suggested that *LOXHD1* was not significant contributor to the FECD in the Yorkshire cohort. Whole exome sequencing (WES) was carried out on the affected members of FECDBRI, FECDBAR and FECDWAK. Additionally, unaffected family members of FECDBRI and FECDWAK were also selected for WES. Samples from FECDBAR and FECDWAK were analysed for variants in the known genes, however as both of these families were found to harbour a CTG18.1 expansion which fully either (FECDWAK) or partially (FECDBAR) segregated with the disease, further analysis of WES variants focussed on FECDBRI in which a CTG18.1 expansion was not seen in any family member.

5.3.1 Genetic contribution of *LOXHD1* variant alleles causing FECD

The FECD3 locus on chromosome 18q21.2-q21.32 was identified in 2006 following genome- wide microsatellite linkage analysis on 3 families with multiple FECD affected members. The disease allele frequency was set very low at 0.0001. Significant linkage was detected using a dominant model of inheritance, taking into account non-penetrance and allowing for an phenocopy rate of 5% (Sundin et al., 2006a). However only one of the families individually reported in this study independently reached a significant lod score and is therefore the only one that can be used to define the locus boundaries. *LOXHD1*, which maps close to (but not within the locus originally reported) the FECD3 locus, appeared to be a strong candidate for involvement in FECD. Riazuddin et al went on to report an enrichment of

mutations in their DNA cohort of 208 cases compared to 838 ethically-matched controls (Riazuddin et al., 2012), implying that heterozygous *LOXHD1* missense mutations were implicated in 7.2% of FECD cases.

Work presented in this thesis included Sanger sequencing of the *LOXHD1* gene to screen for mutations in a cohort of 56 unrelated FECD patients. The results show that 6/112 (5.4%) potentially pathogenic *LOXHD1* alleles were identified in FECD cases, whereas 38/934 (4.1%) pathogenic alleles were identified in the controls. The abundance of rare pathogenic *LOXHD1* alleles identified in this study was therefore only modestly elevated at 1.33-fold in the FECD cohort compared to controls, considerably less than the 3.53-fold enrichment that was described before (Riazuddin et al., 2012) (Table 5.5). Riazuddin and colleagues began their analysis assuming Mendelian inheritance at this locus, as illustrated by use of linkage data under a dominant inheritance model and variant filtering criteria that exclude all changes present in the 1000 genomes database. Their conclusion is that *LOXHD1* alleles are enriched in FECD cases but are nevertheless also present in controls, implying complex etiology for this relatively common condition. The identification of a rare mutation in familial cases, followed by the screening for the mutation in a complex cohort is a strategy that has been successfully employed in a previous study, which found a link between apolipoprotein E and Alzheimer's Disease (Strittmatter et al., 1993).

The study by Riazuddin et al (2012) Sanger sequenced a large gene in a relatively large patient cohort and also in controls, carried out cell-based functional assessments of key mutations, and provided evidence of *LOXHD1* corneal expression as well as altered expression in cornea from a patient, all of which are notable strengths. However, their rationale for screening *LOXHD1* as a candidate gene was based on a study of dominant FECD families which showed linkage to the region 18q21.2-q21.32 (the FECD3 locus (Sundin et al., 2006a)). The fact that they then only identified a *LOXHD1* coding sequence mutation in one of the three families poses questions as to whether *LOXHD1* mutations account for the published FECD locus. It is possible that the other two families may have had a non-coding mutation in

LOXHD1, which would have been missed by their screening strategy. Though plausible, this could also imply that the two families without a *LOXHD1* mutation may have had a mutation in another gene at the FECD3 locus. Although rare, microheterogeneity has been reported in other genetic eye conditions (Toomes et al., 2004). However, in the more recent study (Riazuddin et al., 2012) it would appear that different microsatellite markers were used and the locus boundaries are broader compared with the original linkage study (Sundin et al., 2006a) and encompass the region 18q21.1 where *LOXHD1* is located.

It must be acknowledged that the analysis described herein also has significant weaknesses. The cohort of 56 FECD patients is modest in size by comparison with the 207 cases tested by Riazuddin et al, so may be too under-powered to attain significant enrichment for a small effect size. Although both studies utilised database controls, the ethnically matched control samples from the 1000 Genomes database will undoubtedly include a small number of individuals who later go on to develop FECD. The Yorkshire endothelial-checked controls were not recruited and sampled until later in the study and were therefore not available at the time that the screen was performed. Even if these controls had been available, it would be costly both financially and in terms of time to screen a control cohort for *LOXHD1* variants. By contrast, the 192 control individuals used by Riazuddin et al were all aged over 60 and were phenotyped by slit lamp examination so it can be stated with certainty that they did not have FECD. The lower abundance of pathogenic *LOXHD1* alleles found in the controls used in their study could therefore possibly reflect the exclusion of presymptomatic cases from their control population. It is, however, unclear whether these same controls were among the 288 control individuals sequenced for *LOXHD1* at a later stage of their study, about whom no information on ethnicity, age or ophthalmic status was given. It was not explicitly stated that these were the same controls whose variants were listed in their report. Additionally, the prevalence of FECD is possibly higher in Europe than the US (Eghrari and Gottsch, 2010) and the control individuals analysed from the 1000 Genomes database, though all Caucasian, have been drawn from a variety of different populations,

so they may not be the best available match for FECD cases from the north of England.

Another significant difference between the study by Riazuddin and colleagues and the study presented in this thesis is the assessment criteria for variant pathogenicity. Riazuddin et al used only a single bioinformatic software program, Polyphen2 (Adzhubei et al., 2010) to predict mutation pathogenicity, Of the two available versions of the Polyphen database (http://genetics.bwh.harvard.edu/pph2/), the HDIV model is the most suitable for looking for rare alleles in complex disease, and it was felt to be the most appropriate model to use in this study. Riazuddin et al (Riazuddin et al., 2012) did not state which model they used, and given the variation in results that the two versions produce, and the fact that this was the only pathogenicity prediction software used in the study, this could potentially produce variation in the results. The analysis presented in this chapter used four prediction programs, Polyphen2, SIFT (Ng and Henikoff, 2001), PROVEAN (Choi et al., 2012) and Mutation Taster (Schwarz et al., 2010a) and included *LOXHD1* variants if they were predicted to be pathogenic by any one.

Unlike Polyphen and SIFT, PROVEAN and Mutation Taster are able to process and score small indels as well as missense changes. However, since this analysis was performed, more comprehensive algorithms have been introduced and are now widely utilised. CADD scores can be used to grade nonsense, missense and splice site variants as well as those in intronic regions (Kircher et al., 2014). The Eigan score is the most recently published approach to the analysis of variant pathogenicity (Ionita-Laza et al., 2016). The authors state that this scoring system, which used a large set of variants in its development but no labelled training data set, performs better than CADD in discriminating between disease-causing and benign variants. Given the fact that different pathogenicity prediction tools utilised can give different pathogenicity gradings for the same variant, it seems prudent to utilise all of the available tools in deciding if a variant is potentially pathogenic.

All of the rare *LOXHD1* variants described in the previous study (Riazuddin et al., 2012) were absent from the 1000 Genomes database whereas the ones described in this study were considered if their allele frequency was less than or equal to 4% in the databases. This less stringent approach was felt to better reflect the approximately 4% FECD frequency in Caucasians over the age of 40 years old in the general population (Lorenzetti et al., 1967). Furthermore, for the assessment of evolutionary conservation of the amino acid residues that were mutated in *LOXHD1,* numerous orthologues were listed in the report by Riazuddin et al (2012). The way that this data was presented suggested that the mutations were highly conserved. However, as Chimp, Orangutan, Rhesus and Baboon are all higher-order mammals from an evolutionary perspective, there seemed limited value in listing these in their evaluation. In the assessment of conservation presented in this thesis (Figure 5.8), the putative *LOXHD1* mutations had fewer organisms listed with a greater variation in species order, which made the analysis of evolutionary conservation more meaningful.

Recessive mutations in *LOXHD1* cause non-syndromic sensorineural hearing loss (ARNSHL) with defects of mild-moderate and high frequency range typically presenting during childhood and adolescence. (Grillet et al., 2009). The involvement of mutations in *LOXHD1* in endothelial corneal dystrophy would therefore have some similarity to the finding that dominant mutations in *SLC4A11* cause FECD and recessive mutations cause CHED2 and sensorineural hearing loss (Harboyan syndrome) (Desir et al., 2007, Siddiqui et al., 2014). Late-onset FECD has also been associated hearing loss (Stehouwer et al., 2011) but given that these symptoms affect 20-30% of patients by the age of 70 and 45-55% by age 80 years old (Roth et al., 2011) and that presbyacusis is also a common complex condition with significant environmental contributory factors, any link between heterozygous *LOXHD1* mutations and hearing loss should be interpreted with caution.

The failure of the study presented in this thesis to replicate significant enrichment of *LOXHD1* mutations in FECD cases does not disprove the involvement of *LOXHD1* variants in FECD. However, considered alongside

the lack of *LOXHD1* mutations in two out of three linked families that were studied in the original report, the presence of the nearby trinucleotide expansion in *TCF4* which is very strongly associated with FECD (Wieben et al., 2012) and the data presented herein, it does cast some doubt on the involvement of *LOXHD1* variants in FECD. When inspecting the five cases with predicted pathogenic *LOXHD1* variants for their *TCF4* CTG18.1 status, it was noted that they were inherited independently of one another, but nevertheless 3/5 cases with *LOXHD1* predicted pathogenic variants co-existed with an expanded CTG18.1 trinucleotide variant. This suggests that the *LOXHD1* variants may contribute to disease symptoms in the absence of the major *TCF4* FECD-predisposing expanded allele only in a small number of cases.

5.3.2 Genetic contribution of *TCF4* variant alleles to FECD

In 2010, Baratz et al published the results of a modestly-sized GWAS in which a significant association was found in Caucasians with the G allele of an intronic SNP rs613872 in the *TCF4* gene on Chromosome 18 (Baratz et al., 2010). In contrast to the *LOXHD1* analysis, work presented in this thesis replicated this association in the Yorkshire FECD cohort compared with controls. The association of rs613872 in *TCF4* has also been replicated independently in Caucasian populations in a number of other studies (Li et al., 2011, Riazuddin et al., 2011, Kuot et al., 2012, Igo et al., 2012, Stamler et al., 2013). In Chinese cases (Thalamuthu et al., 2011) and in one study of Indian cases (Nanda et al., 2014) there was an association with another *TCF4* intronic SNP, rs17089887. However, since then one published study of North Indian FECD cases has replicated the association with rs613872 (Gupta et al., 2015). Igo et al (Igo et al., 2012) found a highly significant association between FECD and SNP, rs613872 (P = 2.0 x 10^{-19}) and additionally performed a meta-analysis of other studies (Baratz et al., 2010, Li et al., 2011, Riazuddin et al., 2011, Kuot et al., 2012). The combined odds ratio for each G allele in this meta-analysis was 4.96. Subsequent replication studies have

also supported the association of this SNP with FECD (Mootha et al., 2014, Stamler et al., 2013).

Researchers have gone on to show that a trinucleotide repeat expansion within the *TCF4* gene gives an odds ratio of more than 10 for FECD risk in Caucasian (Wieben et al., 2012, Mootha et al., 2014), Chinese (Xing et al., 2014), Indian (Nanda et al., 2014) and Japanese (Nakano et al., 2015) populations, and cosegregates with the disease in approximately 50% of families with complete and 10% with incomplete penetrance (Mootha et al., 2014). Mootha et al examined 120 Caucasian subjects and 100 controls and found that the two polymorphisms, rs613872 and CTG18.1 were in linkage disequilibrium (Mootha et al., 2014). Both rs613872 and the expanded CTG18.1 were found to be significantly associated with FECD in the FECD cohort presented in this thesis as well as in two of the three FECD families recruited, and it can be concluded that *TCF4* mutations are a major contributor to the pathogenesis of FECD, greater than any other gene identified to date.

The major association of the trinucleotide expansion within the *TCF4* gene in FECD cases previously suggested a number of possible disease mechanisms. One possibility, that the expansion could be in linkage disequilibrium with a functional variant in *TCF4,* has already been investigated but comprehensive sequencing of the coding regions as well as the splice recognition signals have previously failed to identify any potential pathogenic variants in cases to account for disease (Riazuddin et al., 2011, Wieben et al., 2014). This may not be too surprising since haploinsufficiency caused by heterozygous gene deletions, nonsense, frameshift, splice-site and missense mutations in *TCF4* (http://www.LOVD.nl/TCF4/) that impact on the helix-loop-helix, Rep repressor or second activation protein domains, give rise to the neurological condition Pitt-Hopkins Syndrome (MIM#610954). This condition is characterised by intellectual disability, developmental delay, intermittent hyperventilation and distinctive facial features, as well as problems with the eyes, testes and skin (Amiel et al., 2007, Brockschmidt et al., 2007, Zweier et al., 2007).

Alternatively the expansion could be directly causative, affecting *TCF4* gene transcription either by inactivating some of the transcript isoforms that are generated through different 5'-exon usage (Sepp et al., 2011) or it may stabilise transcripts containing the expansion. Recent studies confirm a direct causative effect since transcripts containing expanded repeats have been shown to form RNA foci in the nuclei of cultured fibroblasts and cornea endothelium (Du et al., 2015, Mootha et al., 2015).

In Du et al's study FECD patients' endothelia were collected at the time of corneal transplant, whereas control tissue was obtained as corneoscleral buttons from an eye bank or fresh from enucleation specimens that did not involve the anterior segment. Endothelia from four FECD patients with a repeat expansion, one FECD patient without an expansion and three unaffected individuals were collected. RNAseq data suggested sequences from the intron containing the repeat preferentially accumulated in patients with FECD. In order to explore the mechanism by which the expanded CTG18.1 causes FECD, fibroblast lines from FECD patients were examined. They determined that the CTG-CAG trinucleotide repeat expansion was transcribed into stable RNA, which causes the formation of CUG RNA foci in the affected tissue. There was selective abundance of poly(CUG) RNA foci in FECD corneal endothelial cells compared to fibroblasts, suggesting that TCF4 poly(CUG) transcripts predominantly accumulate in the corneal endothelium, leading to FECD. When present, the foci were shown to sequester RNA binding proteins such as splicing factor, muscleblind-like 1 (MBNL1). RNAseq analysis of the relative abundance of alternatively spliced isoforms for each gene demonstrated widespread changes in splicing in many genes including those involved in EMT, when four corneal endothelial samples harbouring a trinucleotide repeat were compared with three controls samples. This suggests that the presence of a trinucleotide repeat leads to aberrant splicing in FECD. As *TCF4* is known to have a regulatory role in EMT and *MBNL1* splicing alterations have been implicated as an important factor in EMT, the RNAseq results that the authors describe indicate a possible disease mechanism in patients with a *TCF4* trinucleotide repeat (Du et al., 2015). This

is the first time a trinucleotide repeat expansion has been identified as being causative for FECD.

Another group (Mootha et al., 2015) published similar findings using fluorescence-in-situ hybridisation. They examined 8 FECD expansion-positive endothelial samples and noted abundant discrete, spheroidal RNA foci, which were absent in controls samples that did not harbour a *TCF4* expanded CTG18.1. There were no correlations between the CTG18.1 repeat length and the percentage of cells that contained RNA foci. The authors postulated an RNA gain-of-function model in which mutant expanded CUG transcripts were stabilized through their interaction with RNA binding proteins to form nuclear inclusions triggering corneal endothelium-specific aberrant splicing an possibly also apoptosis. Like Du at al, they concluded that RNA nuclear foci were pathognomonic for CTG18.1-mediated FECD and that the disease mechanism was similar to other rare neurodegenerative trinucleotide disorders, such as Myotonic Dystrophy Type I.

On statistical analysis of the *LOXHD1* and *TCF4* variants, both control cohorts were found to be in HWE, but the cases were not in HWE (there were insufficient homozygote expansions to be in HWE). This has also been found in previous FECD case control studies (Mootha et al., 2014). Departure from the HWE have been documented in other studies and in some cases it has been assumed that there was a biological reason for this, whereas others assume a genotyping error has occurred (Wittke-Thompson et al., 2005). In the study presented in this thesis there could be a number of reasons behind the departure from HWE. It is possible, although unlikely, that those homozygous for the CTG18.1 expanded allele are disadvantaged in some way from an evolutionary perspective, and therefore fewer case with this genotype exist in the case population. Another could be that the cases selection is not random as FECD is largely a dominant disease, and therefore the case population will contain a greater number of heterozygotes than homozygotes. Other reasons relate to the TP-PCR assay. Perhaps the threshold of “normal” and “disease” is different clinically from that determined by the laboratory TP-PCR assay), and that if the threshold were actually lower

then there would be a greater number of homozygotes. A Southern Blot would be useful in defining this threshold by accurately sizing the CTG18.1 trinucleotide repeat lengths in cases and controls (this is explored fully in Section 5.3.3).

5.3.3 The role of pre-symptomatic testing in FECD diagnosis

The CTG18.1 *TCF4* trinucleotide expansion appears to be the major presymptomatic factor in causing FECD in Caucasians. Of the 3 families presented in this thesis, FECDBAR and FECDWAK had unaffected members who were identified as having a CTG18.1 expanded allele, some of whom were too young to exhibit clinical signs of the disease. This raises the question of whether pre-symptomatic testing would be of value in individuals with a family history of FECD, by looking for either a trinucleotide expansion or another mutation before the disease had manifested. The role of pre-symptomatic diagnostic testing in diseases caused by trinucleotide expansion mutations is clearer in conditions which present with rare and more severe phenotypes than FECD, such as Huntingtons Disease (Katsanis and Katsanis, 2013), and where the disease is completely penetrant (Mielcarek, 2015). The existence of the CTG.18.1 expansion in controls might be due to the fact that the threshold of a disease-causing repeats is yet to be determined or that the mutation is incompletely penetrant. Therefore pre-symptomatic testing and genetic counselling in FECD patients is potentially fraught with challenges and caveats. Perhaps more information needs to be ascertained about the full genetic architecture of FECD, along with more comprehensive knowledge of contribution of disease modifiers and environmental influences before pre-symptomatic testing can be widely available. Certainly, in the cases of FECDWAK and FECDBAR families who possess an expanded CTG18.1, it would be of benefit to analyse the exact size of the expansion and follow these patients longitudinally.

Of the 83 controls tested using the TP-PCR assay, 5 of the controls, who were aged between 64 and 88, were found to have a CTG18.1 expansion.

There were many advantages of the PCR based assay used in this study (Warner et al., 1996). It allowed the analysis of a large number of cases and controls to be performed accurately and relatively cheaply. The limit of this assay is that, while it can determine the boundary of a “normal” allele and “expanded” allele, clinically this is an arbitrary cut-off point. In Mootha et al’s study the authors distinguish between two groups based on an expansion greater than or less than 43 repeats (Mootha et al., 2014). However, the repeat length is a linear variable, and the exact threshold at which the expansion becomes unstable or pathogenic remains unknown. Therefore, whether an individual subject with a repeat in the 40 to 60 range has an increased risk of disease is also unknown.

There is evidence that repeat length can influence the severity of the disease (Mootha et al., 2015), but the Krachmer grading (Table 1.4) on which disease severity is based in this study, is not a simple objective measure of disease severity as it varies with time over the course of the disease. It is therefore entirely possible that repeat length can influence penetrance or severity of FECD, but this remains largely speculative. The definition of the threshold for a repeat expansion may directly influence the odds ratio. For the association between CTG expansion and FECD, the odds ratio may increase with increasing repeat length. Hypothetically speaking, it is entirely possible that a repeat length of 55 might cause disease, perhaps with a lower odds ratio than that estimated in this or other studies, whereas a repeat length of 2000 will certainly cause disease, perhaps with a much higher odds ratio. Unlike Myotonic Dystrophy Type, the *TCF4* CTG18.1 in FECD does not yet have the benefit of a long and detailed history of validation. The challenge of better defining these boundaries relates to the technical challenges of determining large trinucleotide repeat lengths in the laboratory, and Southern blotting will almost invariably play a role in defining this. It would be interesting to know the exact size of the expansion in the 5 controls in whom an expanded CTG18.1 was found. Southern blotting could be a means of ascertaining whether the trinucleotide repeat length in these controls exists in the intermediate range. This could be an avenue of further work. Additionally, long-term follow-up of the controls that possess the CTG18.1 expanded allele

to assess for the development of FECD could be useful in determining whether the mutation is incompletely penetrant.

5.3.4 Anticipation and trinucleotide instability in FECD patients with the CTG18.1 expansion

Trinucleotide expansions are the causative mutation for many genetic diseases, including Myotonic Dystrophy Type I and 2 (DM1 and DM2 respectively), Fragile X and Huntington's Disease (HD). The mechanism of RNA toxicity seen in FECD patients who possess a CTG18.1 expansion is similar to that seen in DM1 (McMurray, 2010). Both myotonic dystrophy and HD exhibit anticipation, worsening of disease severity and earlier age of onset in successive generations (Orr and Zoghbi, 2007). The mechanism by which anticipation occurs may vary between trinucleotide repeat diseases. In Fragile X and Myotonic Dystrophy, CGG repeats in the non-coding region expand almost exclusively through maternal transmission, whereas in HD CAG repeats expand more commonly through paternal transmission (McMurray, 2010) However, the sex preponderance of these diseases remains unaffected by this mechanism.

Instability of the repeat sequence is the hallmark of trinucleotide disease and can be characterized by anticipation, worsening of the clinical phenotype, of progressively younger age of onset in successive generations. There are 32,448 trinucleotide repeat sequences in the human genome and 878 trinucleotide repeat sequence-containing genes (Kozlowski et al., 2010). While the mechanisms of triplet repeat instability remains unclear, possible contributions include polymerase slippage, DNA secondary structure, chromatin structure and triplex DNA formations (Longshore and Tarleton, 1996). When the CTG18.1 was originally identified, some evidence of instability was indicated on Southern blotting (Breschel et al., 1997). While anticipation is not considered a typical feature of FECD, the presence of the CTG18.1 in the FECDBAR family and the earlier age of onset of affected individuals in the younger generation suggests that instability of the CTG18.1

mechanism might underlie this process. This is particularly true in the case of individual IV:1 (Figure 5.12), who was 28 -years-old when she was found to have guttata. A Southern blot analysis of individuals II:17, III:28 and IV:I would help to determine whether the repeat expansion length was increasing in successive generations in this family.

5.3.5 The use of endothelial-checked controls or ECACC population controls in FECD studies

For FECD, whether the controls have had their endothelium checked could have a profound effect on the findings of studies. From the work presented in this thesis, the mean age of the ECACC controls, who had not had their endothelium checked, was 38.7 +/- 8.4 years, which was considerably less than the mean age of the 56 FECD cases initially recruited. Had these controls been examined for endothelial abnormalities, a small proportion may had been found to have guttata, indicating that they would go on to develop FECD, and this proportion would have increased if they were checked when they reached an age equivalent to the patient cohort. This issue was addressed with the recruitment of age-matched endothelial checked controls from NHS cataract clinics. Unsurprisingly, a greater number of CTG18.1 expansions had been noted in the ECACC controls compared with the endothelial-checked controls. This reduction in the frequency of such pre-symptomatic signs in controls increased the odds ratio and significance of findings when using endothelial-checked as opposed to unchecked population controls.

The presence of the expanded allele in a small proportion of the general population could have implications for the use of donor corneas to treat FECD, as some donor corneas may in fact be taken from pre-symptomatic carriers. One might expect that if the cornea belonging to a patient with a *TCF4* trinucleotide expansion were used for transplant, that the recipient might be at risk of developing FECD. However, FECD recurs extremely rarely following corneal transplantation. Furthermore, corneas that have been

donated for transplantation are checked under light microscopy for abnormal corneal endothelial cell morphology. The data presented in this thesis suggests that examination of the corneal endothelium appears to be adequate in screening out patients with a repeat expansion. This would suggest that, until such a time as corneal gene therapies reach the clinic and theatre, there is no current need for genotyping corneal graft tissue, as the existing quality control appears to exclude abnormal morphology and therefore largely excludes those patients that might develop FECD. This data therefore highlights the importance of these checks, though one exception might be the use of tissue from a very young donor, who might be too young even to show these presymptomatic abnormalities. Around 65% of patients who donate their corneas are above the age of 60-years-old (http://www.nerc-charity.org.uk/cornea-donation), leaving a significant minority who might be too young to manifest signs of endothelial disease. A further study examining the genotypes of corneal tissue assessed and not ultimately utilised for corneal grafting could add weight to this argument. An alternative explanation for this lack of recurrence of FECD in grafted individuals is that perhaps on occasion a corneal graft harbouring a *TCF4* CTG18.1 expanded allele is transplanted, but the disease does not manifest itself during the lifetime of the patient or the individual experiences graft failure for a different reason prior to this. There might therefore still be value in tissue genotyping in younger FECD patients being grafted. As genetic mechanisms in FECD continue to be fully elucidated, and as patient management becomes increasingly individualised, the role of genotyping tissue may become increasingly prominent in the future.

5.3.6 The corneal endothelium transcriptome

From the work presented in this thesis, RNA extractions performed on endothelial tissue using two different methods were unsuccessful (Section 2.15). The use of either method led to unmeasurable levels of RNA, and therefore RNA sequencing of this tissue could not be carried out. The advantages of using graft tissue left over from lamellar graft procedures was

that it was tissue that was deemed suitable for corneal transplant and accurately dissected into the appropriate layers under sterile operating conditions. There is a lengthy exclusion criteria for the use of graft tissue in patients (Zuberbuhler et al., 2013). Therefore, the tissue used for grafting has come from healthy patients without significant co-pathology.

This strategy worked well for the RNA sequencing of epithelial and anterior stromal tissue, where RNA sequencing was successfully carried out for another project. One reason for the lack of RNA extracted from endothelial corneal tissue was most likely degradation of the RNA from the single endothelial monolayer from the time the eyes were enucleated until the time they were placed in RNA later. This could have been overcome by using tissue from freshly enucleated eyes from donors with other general pathology not deemed suitable for corneal transplant. This would have had the advantage of being freshly extracted, although overall this tissue might have been less suitable as control tissue due to other general pathology in the patient, which potentially could affect corneal expression. However, had degradation of tissue been the sole reason for failure of extracting sufficient quantities of RNA, then one would expect the bovine tissue to yield sufficient amounts of RNA, which it did not. It may be that the volumes of endothelial tissue used were too small for the RNA kits to extract useful levels of RNA.

A published database of RNA sequencing in control tissue can be used to aid the prioritisation of gene candidates following next generation sequencing studies (Chng et al., 2013, Du et al., 2015). The study by Du et al (2015) used tissue from an eye bank, or eyes enucleated for pathology that did not involve the anterior segment. The corneal endothelial tissue was bluntly stripped away from the corneal button and the RNA extracted using two different Qiagen RNA extraction kits. The authors do not mention whether the patients whose eyes were used as controls had any other general pathology.

5.3.7 Whole exome sequencing of members of the FECD Families

Two different methods of NGS were employed in FECDBRI. Three affected family member's DNA samples were sent to Otogenetics. However, as the read depth and data quality were significantly better at the Leeds NGS facility, a fourth affected sample and an unaffected sample were prepared and sequenced in Leeds. When the shared variants were analysed, the same 5 variants were present after analysis with both Pipeline 1 and Pipeline 2 (Section 2.21).The reason for the use of two NGS pipelines in this study was broadly the development and refinement of existing NGS pipelines over time with updated guidance on best practice of variant discovery (https://www.broadinstitute.org/gatk/guide/topic?name=methods) and experience within the VRG. Specifically, in pipeline 1 Bowtie 2 was used for the alignment of the FASTQ files, whereas in pipeline 2 Novoalign was the utilised. Experience from the members of the VRG and from other research groups in the Section of Ophthalmology and Neuroscience indicated that Novoalign outperformed Bowtie2 in terms of alignment accuracy, including where a large number of indels were present and overall in the identification of common and novel variants. Additionally, a published article (Yu et al., 2012) compared four different aligners including Bowtie2 and Novoalign and concluded that all aligners perform similarly, with a slight advantage of using Novoalign in terms of accuracy and sensitivity to improved data quality. Therefore when Novoalign was ultimately used by most members of the VRG and included on the standard NGS pipelines developed by local bioinformaticians.

The use of an unaffected individual in the analysis of data significantly reduced the number of variants when subtracted from the list of shared variants in affected individuals. Individual II:2 of FECDBRI was identified as clinically unaffected with FECD, presymptomatic endothelial cell changes were also excluded and this individual did not develop the disease during the course of the study. However, the issue of non-penetrance in FECD could be considered a potential problem in the analysis of exome data that utilises a sample from an unaffected individual. If the individual possesses the mutation

but does not manifest the disease then this could erroneously exclude a causative mutation from the analysis.

In the analysis of exome data, the frequency of each variant in the 1000 Genomes or ExAc databases is provided. It is therefore possible to filter variants on frequency. For the study of rare autosomal recessive disease a threshold of <1% is usually employed. While FECD is a relatively common disease, mutations in published genes have largely been rare (Biswas et al., 2001, Vithana et al., 2006, Riazuddin et al., 2010a, Riazuddin et al., 2012, Riazuddin et al., 2013). Therefore rather than selecting a threshold of 4%, the cited prevalence of FECD (Lorenzetti et al., 1967), a threshold of ≤2% was selected.

In relation to this, one of the major challenges of studying the segregation of putative mutations in a late-onset disease in a pedigree such as FECDBRI was the uncertainty of whether members of generation III will develop the disease. While individual III:5 did not have definitive signs of FECD, there was some evidence of pleomorphism, and the consultant examining her felt that she might well be at risk of developing FECD. This has implications for the variant segregation of the potential gene candidates in FECDBRI. On Sanger sequencing she was found to possess the wild-type allele for *OSBPL1A* but the c.C461G variant of *CRYZL1A.* As the disease status of the younger individuals of generation III in FECDBRI was not ultimately confirmed as they may not have been old enough to exhibit clinical signs of FECD, it was difficult to make a definite decision on screening the cohort based on this segregation analysis.

With all the caveats in this analysis listed above, *CRYZL1* and *OSPBL1A* were considered the best candidate genes highlighted by WES of FECDBRI. *CRYZL1* encodes a corneal crystalline, is expressed in the corneal endothelium (Chng et al., 2013), and there is a potential mechanism by which a mutation in this gene might compromise corneal clarity (Gong et al., 2012). The variant c.C461G:p.A154G is not only rare or absent from publicly available databases but is also highly conserved throughout evolution (Figure

5.20). *OSBPL1A* is also expressed in the corneal endothelium (Chng et al., 2013) and additionally it interacts with *FOXC1*, a gene known to be involved in the development of the anterior segment (Berry et al., 2008). Furthermore, the *OSBPL1A* variant c.115_116insAATT:p.C39_K40delinsX is a stopgain mutation which leads to a frameshift and premature stop codon, so if pathogenic would have a clear effect on the protein translation and function. However, the relatively frequent occurrence of other *OSBPL1A* homozygous variants in ExAc make a loss-of-function mutation less likely to be disease-causing.

Had the study been carried out on an early-onset disease and therefore led to a clearer result from the Sanger sequencing of variants in FECDBRI, these data might have led to the screening of the FECD cohort for mutations in these genes. However, doubts over diagnostic accuracy in the younger generations, together with restricted PhD lengths, meant that this follow-up work was not completed. Although an accurate method of sequencing a gene, Sanger sequencing (Sanger and Coulson, 1975) is a time-consuming and expensive method of sequencing large genes. Alternative high throughput methods avoiding the full costs of exome sequencing include long-range PCR followed by sequencing of products on a MiSeq machine, as utilized clinically in the screening of BRCA1 mutations (Morgan et al., 2010) or using molecular inversion probes (MIPs), a strategy that has been employed in the sequencing of genes that cause autism spectrum disorders (O'Roak et al., 2012). Both of these methods require meticulous optimisation of primers, which itself can be extremely time-consuming, and which can ultimately prove costly. Additionally, they can only be used in the screening of a selection of genes, unlike WES which provides the coding sequence for all genes in an individual. It may be that, as the cost of exome sequencing diminishes further, WES of either the entire cohort of FECD patients, or those who did not possess a CTG18.1 expansion, might become feasible. This is a potential future avenue of this project (Chapter 6).

The rare *TCF4* SNP rs192075715 was found to segregate with FECD in the FECDBRI family, and when combined with the initial linkage to chromosome

18 (Section 5.2.6.2.1), pointed towards the same region of chromosome 18 to be linked in the FECDBRI family. However, no expansion was found on TP-PCR in this family, nor any potential pathogenic variant in *LOXHD1*. Given that the CADD score of SNP rs192075715 is 2.5 and the fact that, using the BDBG splicing prediction modeler (Section 2.17.7), it has no discernable effect on splicing, it is unlikely that this SNP is itself increasing susceptibility to FECD in this family. It remains possible that this SNP is in linkage disequilibrium with an FECD causing variant in this region. There are well-documented examples of causative intronic mutations, such as that seen in Lebers Congenital Amaurosis (den Hollander et al., 2006) or mutations within a promoter region in PPCD (Davidson et al., 2016). Such mutations would have been missed by the strategy presented in this thesis but could be found using WGS. The *TCF4* CTG18.1 is the first example of an association of a trinucleotide repeat with a common, complex disease. Therefore it is possible that the causative mutation in FECDBRI is another trinucleotide expansion on chromosome 18. A disadvantage of NGS is that it cannot directly sequence trinucleotide repeat sequences in the genome. Therefore, in the absence of other linkage data, it is not currently possible to ascertain whether other trunucleotide repeats cause or are associated with FECD.

Another potential drawback of exome sequencing is, while it is more robust when it comes to determining structural variants such as deletions and insertions than Sanger sequencing, this cannot be done using the standard WES pipelines (Section 2.22). ExomeDepth (Plagnol et al., 2012) addresses this by its capacity to determine the presence of indels and larger structural variants. The need for controls run on the same next generation sequencing lane was a potential disadvantage in this study as the controls were FECD patients from different FECD families. The other FECD patients were from families which had been identified as having an expanded CTG18.1, therefore perhaps this was not a significant limitation. Nonetheless the ExomeDepth output presented for the FECDBRI family gave a list of pathogenic structural variants, with the highest BF score of 44.1 being a heterozygous duplication of 122,909 base pairs on chromosome 10. This was not identified as a common duplication in the Conrad database (Conrad et al., 2010) (Section

2.23). A large duplication would be challenging to verify on Sanger sequencing, however whole genome sequencing (WGS) could be performed in another affected individual to assess a limited form of segregation and examine the potential breakpoints. A pipeline is being developed locally that could perform this analysis on WGS data, and can manage the large FASTQ file sizes more efficiently that the standard aligners used in WES.

In the FECD cohort presented in this thesis, 33/117 (28%) of the patients did not harbour a CTG18.1 expanded allele. A further study could also evaluate these cases by WES to determine their genetic basis.

5.3.8 The contribution of phenocopies in FECD families

An individual affected by a condition despite not having the disease genotype is called a phenocopy (Strachan and Read, 2011). Phenocopies have been reported previously in large FECD pedigrees where putative causative variants have been identified (Sundin et al., 2006a, Riazuddin et al., 2012). Additionally they have been reported in rare severe conditions caused by trinucleotide repeat expansion (Wild and Tabrizi, 2007, Abbruzzese et al., 1996). Individual II:19 in FECDBAR exhibited signs of FECD but possessed the rs613872 genotype TT and was negative for the CTG18.1 expansion. This result was verified with stock DNAs in this individual. This supports the concept that FECD is a complex disease, implying that perhaps there are significant environmental factors or variants in other genes which played a part in the pathogenesis of the disease in this individual. Other explanations for the lack of CTG18.1 expanding allele in FECDBAR individual II:19 include the possibility of a sample mix up during the DNA extraction process or non-paternity of the individual.

6 General Summary, Concluding Remarks and Future Directions

This thesis aimed to examine the genetic basis of corneal endothelial dystrophies. Chapter 3 looked broadly at corneal dystrophies as a whole, examining their incidence in patients below the age of 40 years. This was carried out in conjunction with the British Ophthalmic Surveillance Unit. In Chapters 4 and 5, the focus turned to corneal endothelial dystrophies. In Chapter 4, a rare Mendelian congenital form of endothelial dystrophy, CHED, was explored. A newly recruited CHED patient from Mexico was screened by Sanger sequencing for mutations in *SLC4A11*. This patient's hearing was also tested by audiometry along with two other CHED families previously confirmed as harboring a mutation in *SLC4A11.* Additionally, parents of CHED patients that were available for testing were examined for evidence of adult onset corneal endothelial disease. Chapter 5 focused on the genetic analysis of a Caucasian cohort of 117 patients with the more common but potentially more genetically complex endothelial dystrophy FECD, recruited and sampled for genomic DNA from Ophthalmology clinics and theatre lists in Yorkshire. This cohort consisted mostly of single cases but also one member of each of 3 families with multiple affected members. The genetic analysis evaluated the relative contributions of *LOXHD1* variants and the *TCF4* expansion to endothelial disease susceptibility in the FECD cohort compared with controls, and to identify the genetic cause of FECD in the families using whole exome sequencing and segregation analysis. The major findings of these studies are discussed below, together with the implications for diagnostic testing, current treatment regimes and future therapies for endothelial dystrophies.

6.1 Summary of key findings and future directions.

6.1.1 A study of young-onset corneal dystrophies in conjunction with the British Ophthalmic Surveillance Unit

The aim of the study described in Chapter 3 was to ascertain the incidence of new-onset corneal dystrophies in patients presenting below the age of 40. This was carried out in conjunction with BOSU, which runs a nation-wide surveillance system for the epidemiological study of incidence, clinical features and management of rare ophthalmic diseases. For any study adopted by the BOSU, reporting cards naming the conditions of interest are sent to UK ophthalmologists monthly by the BOSU team. This reduces the burden of work for doctors identifying patients and avoids the reporting of different conditions to multiple sources.

Corneal dystrophies are an ideal group of conditions to be studied in conjunction with BOSU. There are no existing national incidence studies of corneal dystrophy, and there is likely to be significant variation in practice. As young-onset corneal dystrophy is a relatively rare but important cause of severe visual impairment, BOSU provided an ideal way of collaboratively identifying rare cases. The study found that the majority of referrals originated either from an ophthalmologist or an optometrist. In terms of patient demographics, nearly 80% of the patients studied were White-British and females accounted for 48% of the patients identified. The mean age at presentation was 12 years old and the most commonly-reported type of dystrophy was endothelial. The incidence of corneal dystrophy in patients younger than 40 years was compared with UK census data and it was concluded that young-onset corneal dystrophies are extremely rare and this study would suggest a minimum UK incidence of 6.7 newly-diagnosed cases per 10 000 000 population aged below 40 years per annum.

The knowledge of epidemiology and presentation gained could be used to inform clinicians about the expected presentation, common visual symptoms, clinical course and commonly utilised management strategies. In turn

clinicians could use this information to help manage and counsel their patients. Additionally, knowledge of the epidemiology of corneal dystrophies could be utilized for service planning in the NHS. Half of the patients identified in the study had a positive family history of a corneal dystrophy. However, a molecular diagnosis was sought in only four patients. It would be interesting to further explore the results of whole exome sequencing in those patients who underwent genetic testing and indeed to perform genetic testing in all of those with a family history of corneal dystrophy. The corneal dystrophy classification has been recently amended to incorporate those dystrophies caused by mutations in the *TGFβI* gene. It is therefore entirely probable that in the future, next generation sequencing will become a routine part of corneal dystrophy work-up and management.

6.1.2 Genotype-Phenotype Correlations in CHED, Harboyan Syndrome and FECD

CHED is a rare, bilateral autosomal recessive condition affecting the corneal endothelium. It is characterized by corneal oedema and opacity. Harboyan syndrome is a considered a distinct syndrome consisting of the symptoms of CHED combined with later onset sensorineural hearing loss. Both conditions are caused by mutations in the *SLC4A11* gene. In Chapter 4 of this thesis, a new mutation in the *SLC4A11* gene, c397T>C F133L, was identified in a newly-recruited CHED patient from Mexico. Additionally, sensorineural hearing loss was found to be present in this patient and in three other patients who had been diagnosed with CHED when they were younger. This suggested that, rather than Harboyan syndrome being a distinct condition, it is more likely to be the same condition as CHED but at a later stage of progression. The study examines just a small number of cases and their families. A further study could longitudinally monitor a larger cohort of CHED patients to verify this finding of eventual high-frequency sensorineural hearing loss determined by audiometry. The results of this chapter suggest that in clinical practice, serial monitoring by audiometry should be carried out in CHED patients, even if the initial audiometry findings were unremarkable.

When individual II.2 from Family A (Figure 4.1) attended her routine Ophthalmology appointment at the BRI in 2012, she was coping with low vision (acuities of 6/60 in the right eye and hand movements in the left eye), and therefore the additional impact of hearing loss to her day-to-day activities could not be understated. Repeating the audiometry at this stage meant that her hearing impairment was correctly diagnosed and she was fitted with bilateral hearing aids. It also enabled her to receive additional social support.

The later onset endothelial dystrophy FECD also results from mutations in *SLC4A11*, but has onset in late adulthood and is the result of heterozygous rather than homozygous mutations. This implies that parents of CHED patients, who are almost certainly carriers of *SLC4A11* mutations, may also be at risk of developing endothelial dystrophy in later life. The heterozygous parents of the CHED patients examined herein were found to have early signs of corneal endothelial disease including guttata and reduced endothelial cell counts, implying that they are indeed at increased risk of developing FECD. Future studies monitoring the parents for the development of FECD would be useful. Additionally, studying more parents of CHED patients for corneal disease outcome would further validate the link presented in this thesis between CHED patients and FECD in their heterozygous parents, and would mean that prognostic information including the likelihood of requiring corneal graft surgery could be offered to other parents of CHED patients.

6.1.2.1 Genetic Analysis of Fuchs Endothelial Dystrophy in Yorkshire

FECD is a progressive disease of the corneal endothelium. In Chapter 5, the genetic basis of FECD was explored. The relative contributions of the *TCF4* SNP rs613872, the intronic CTG18.1 trinucleotide expansion and *LOXHD1* variants in a UK Caucasian FECD cohort were compared in order to clarify the significance of the original findings at the FECD3 locus on chromosome 18. Mutations in *LOXHD1* were not found to be significantly enriched in cases compared with public databases. The *TCF4* CTG18.1 expanded allele, in

contrast, was present in 72.6% of cases compared with only 6.0% of endothelial-checked controls (odds ratio 23.2). Furthermore, the results of the segregation analysis of the *TCF4* CTG18.1 expansion in three local FECD families showed that the CTG18.1 was also likely to be involved as a causative or susceptibility allele in two of the three FECD families. The mutation in the third family, FECDBRI, was not explained by the *TCF4* CTG18.1 expanded allele, and therefore further analysis of the WES in this family was carried out. Segregation analysis of five of the variants obtained from WES after filtering was performed, two of which, *OSBPL1A* and *CRYZL1* were considered to be good candidates. However, due to the unknown disease status of members of generation III of FECDBRI, segregation analysis of the entire family was not sufficiently powerful to provide compelling evidence of the involvement of one or both of these variants. Screening of the FECD cohort for these variants did not take place, in part for this reason and also, due to time constraints of the PhD, which meant that other experiments were prioritized.

The main conclusions drawn from this chapter were that, in our cohort, *LOXHD1* mutations did not appear to be significantly enriched in the FECD cases compared with controls, suggesting that the role of *LOXHD1* mutations is minor. By contrast, in the Leeds FECD cohort there is a highly significant association of *TCF4* mutations in FECD cases compared with controls, suggesting that mutations in *TCF4* contribute significantly to the pathogenesis and more than any other gene found to date. Nevertheless, there remain a proportion of FECD cases which do not carry *TCF4* expansions, and some unaffected individuals do carry the expanded *TCF4* allele yet do not develop the condition, which confirms the contribution of other genetic, and potentially environmental, factors in protection from or susceptibility to FECD.

In our cohort there were 6% of control individuals who were of an age at which FECD might be expected to develop, had normal corneas on examination, and yet who possessed an expanded CTG18.1 allele. The TP-PCR assay was only able to determine whether the allele exceeded 43 repeats, which was a laboratory rather than clinical threshold. It was

considered possible that the threshold of “normal” and “disease” could be slightly above or slightly below 43 repeats, and that the control individuals with an expanded allele might in fact carry a CTG18.1 allele of this intermediate length (around 43 repeats). Future avenues of work include Southern blotting to determine the trinucleotide length in controls who possessed the expansion, to determine whether their expansion is of an intermediate length. This technique would also be of use in individuals from generation III of the FECD families, FECDBAR and FECDWAK, who may still be too young to manifest signs of the condition and therefore whose clinical status remains unknown. If very large repeats were found in these individuals, they could be considered at risk of developing FECD and followed longitudinally. Studying FECD in families could also give valuable information about how repeat lengths are inherited by an individual’s offspring, and would allow clarification of whether anticipation had occurred as the expanded allele passed from one generation to the next. In this thesis, the three generations of the FECDBAR family with affected individuals could be examined for this purpose.

Southern Blotting allows the detection of specific DNA sequences or small repeats separated by gel electrophoresis which are then transferred from the gel to a porous membrane by capillary action using absorbent paper to soak solution through the gel and the membrane. These specific sequences are detected in the membrane by hybridization with labelled nucleic acid probes (Southern, 2006). It is an accurate means of assessing variable repeat lengths, however does have disadvantages. It is expensive, time-consuming to perform when large number of samples are processed, and traditionally requires the use of radiation.

A major limitation of NGS is its inability to accurately size trinucleotide repeats. Newer techniques are evolving, utilising high throughput technology, which may make the characterization of trinucleotide repeat lengths a possibility without the need for Southern Blotting. Single-molecule real-time (SMRT) sequencing generates longer reads than traditional WES (Guo et al., 2015). This technology utilises a sequencing-by-synthesis approach, in which a circular DNA molecule is used as a template for a single DNA polymerase.

Linear amplicons are converted to a circular form, and the resulting SMRT library is then sequenced using a SMRT cell containing 150,000 DNA polymerases. The main advantage of this technique is the creation of single-molecule reads that exceed 10kb in length. This technology has been used to sequence large trinucleotide repeat sequences that are causative for Fragile X (Loomis et al., 2013), and there is great potential to use this to accurately ascertain the CTG18.1 repeat length in both FECD patients and controls.

As tentative linkage to chromosome 18 was found in FECDBRI in chapter 5, whole genome sequencing (WGS) would be of use in order to fully examine the FECD3 locus on chromosome 18, as well as rest of the genome, in this family. This would screen for large structural variants, especially those in genes known to cause FECD. Although this analysis could be carried out using WES data, for example using the program ExomeDepth (Section 2.23.1), WGS provides a much better dataset for such analysis, and includes the sequence for intronic and intragenic regions. Therefore, if a large structural variant were found, the breakpoints of this could be accurately determined. This would also screen for mutations in promoter regions and introns, and would specifically assess the presence of the heterozygous duplication described in Section 5.2.6.2.5.

As *TCF4* is genetically a major contributor to the pathogenesis of FECD, WES of the patient samples from the Yorkshire FECD cohort without a *TCF4* CTG18.1 expansion could be carried out in order to screen for mutations in the best gene candidates from WES in FECDBRI, *CRYZL1* and *OSBPL1A.* This strategy could highlight novel variants in other genes as potential causative mutations in FECD, and if performed collaboratively with different research groups, could result in sufficient patients recruited to perform a GWAS to assess for genetic associations in expansion-negative patients.

6.2 Implications of research for corneal dystrophy patients

6.2.1 Strategies in the screening of candidate genes

Despite advances in Next Generation capabilities over recent years, if the next step in this project were to involve screening a large cohort of FECD patients, the total cost of WES is still prohibitive at this time. It therefore becomes necessary to consider the available target enrichment strategies. These include those based on hybridisation strategies such as the Agilent Sureselect^{QXT} kit, the use of molecular inversion probes (MIPS) and PCR based enrichment strategies. The method utilised might depend on the ease of experimental design, cost, scalability and the uniformity of coverage (Kozarewa et al., 2015). The employment of such methods could have been useful for example in carrying out the screen of a large gene such as *LOXHD1* for mutations in a larger cohort of FECD cases and controls, as well as for screening for mutations in the other genes known to cause FECD. However, at the time that the *LOXHD1* screen was carried out, these technologies were not freely available or fully developed, and may not have proved cost-effective in a cohort of the size studied in this project. Assessing the FECD case/control cohort for mutations in *OSBPL1A* and *CRYZL1* might have been possible using one of these high throughput methods. However, the ideal way of analyzing the FECD cohort if further time allowed would be to exclude FECD with expansions by means of Southern Blot, then analyse smaller numbers of patients negative for the CTG18.1 expanded allele using WES.

6.2.2 Challenges of NGS

One of the current main disadvantages of WES is the associated cost of performing it. However, this has diminished in recent years and it is feasible that it could be reduced further to allow large cohorts of patients to be exome sequenced. The ability to perform WES in a cohort of ethnically-matched

examined controls would be invaluable for use in studies of other corneal diseases. The major drawback in the use of WES data drawn from databases of control individuals in the study of common, complex, late onset disease is that inevitably some of these patients sequenced will have, or go on to develop, corneal disease, and therefore disease alleles will be present in the resulting datasets. The major advantage of an exome sequenced cohort of examined patients would be that the number of corneal disease alleles present would be reduced, allowing greater power in the detection of mutations in case cohorts.

The advent of WGS would increase the ability to assess such cohorts for mutations in non-coding or promoter regions. WGS has allowed researchers to make a more detailed assessment of intronic and intragenic variants and a more comprehensive analysis of structural variants in genomic DNA. There are notable examples in ophthalmology (Davidson et al., 2016, Small et al., 2016) where this technology has led to the detection of a causative mutation where previously WES alone had failed to reveal this. However, such advances are not without their challenges. Firstly, WGS generates millions of genetic variants, and the task of finding the causative mutation amongst these is an extremely difficult task. This can be addressed in part by developing effective filtering strategies but these strategies are still evolving and this is therefore an ongoing issue. Secondly, there are challenges relating to the storage of the large files generated from WGS, which can be several hundred gigabytes in size and may contain sensitive information about an individual's genetic make-up, which raises additional issues regarding the confidentiality of data. Locally in Leeds, this has been addressed with the installation of an advanced computational infrastructure with over 2.5 petabytes of secure data storage capacity, through the Leeds Institute for Data Analytics (LIDA - http://www.lida.leeds.ac.uk/).

Next generation sequencing as a whole is usually requested for a specific clinical or research reason, and may yield a primary finding specific to the indication for which the test was carried out. However, this approach also generates data on other genetic variants, some of which may have medical

significance and can be challenging both in their interpretation and in the process of reporting back to patients and their clinicians. NGS variants which are not related to the diagnostic indication, but which are likely to have medical consequences, and for which preventative treatment may be beneficial, are referred to as secondary findings. Variants of unknown significance, or those variants of medical significance but for which no preventative measures are available, are referred to as incidental findings. On the whole, genetic professionals feel that secondary results, but not incidental findings, should be offered to adult patients, healthy adults and parents of a child with a medical condition (Yu et al., 2014). Among patients themselves there is considerable diversity of opinion in terms of what information should or should not be returned and most patients feel that they should have a choice and participate in this decision. It is therefore crucial that patients are allowed to opt out of receiving certain results (Clift et al., 2015). Fortunately there are guidelines available to assist those counselling patients in making these difficult decisions (Green et al., 2013). The working group involved in the development of these guidelines acknowledged that there was insufficient evidence about the benefits, risks and costs of disclosing incidental findings to make evidence-based recommendations, but felt that secondary findings were likely to have a medical benefit for patients and their families undergoing sequencing. The current Leeds corneal dystrophy ethics approval covers the use of Next Generation Sequencing for WES of patients' DNA samples, and the WES of the CHED and FECD samples used in this study so far have not raised any such issues. However, the question of whether the patient does, or does not, want to be made aware of secondary findings is not specifically addressed in the consent form used. A new ethics approval is currently being drafted to address this issue.

6.2.3 Clinical genetic testing of the *TCF4* CTG18.1 allele in FECD patients

The results presented in Chapter 5 indicate a highly significant association between FECD cases and the CTG18.1 expansion, with carriers 23.2 times

more likely to develop FECD than non-carriers. It can therefore be concluded that *TCF4* mutations contribute more to the pathogenesis of FECD than any other gene identified to date. This may lead to demand for a clinical genetic test to look for an expanded CTG18.1 allele in the families of FECD patients. For the TP-PCR assay to become an accredited method of clinical diagnostic screening for the CTG18.1 expansion in FECD patients, further validation would be required. Analysis of the CTG18.1 allele indicated that the FECD cases were not in Hardy Weinberg Equilibrium (HWE). Departure from HWE can occur for biological results or as a result of genotyping error. As FECD is a dominant disease, the population from which these alleles are drawn are largely heterozygous, which could account for the relative lack of homozygous CTG18.1 variants. Alternatively, homozygous variants might be evolutionarily disadvantaged, resulting fewer than expected homozygotes, although given that the disease occurs in otherwise healthy adults, this is unlikely. Errors in genotyping as a result of laboratory assay may also be responsible. If this were the case, one possible explanation for this could be that the TP-PCR was failing to detect those homozygous for the expanded allele. Designing primers enabling the TP-PCR assay to be performed in both directions would be an additional confirmatory step to ensure that the genotyping calls were correct. Southern Blotting of several FECD patients and examined control individuals would also provide further validation by establishing the threshold of repeat expansion associated with “normal” and “disease” status. This knowledge would be invaluable for genetic counselling of FECD patients with an expanded CTG18.1. In both FECDBAR and FECDWAK there are apparently normal individuals who carry the CTG18.1 expansion but are potentially too young to manifest the symptoms and signs of FECD. A detailed understanding of the link between trinucleotide repeat number and FECD causation, phenocopy rates and penetrance are essential for correct counselling of these individuals.

6.2.4 Gene therapy for corneal disease

Much of the focus of gene therapies in eye disease has been on the treatment of retinal disorders. One example of this is the gene therapy trials for Leber's Congenital Amaurosis caused by mutations in the RPE65 gene. Long-term follow-up of treated patients has indicated improved vision for three years, although in the longer term eventual progressive photoreceptor degeneration hindered sustained improvements in vision (Cideciyan et al., 2013).

A number of approaches to treating corneal disease using gene therapies have also been considered. These include the use of adeno-associated viral vectors to facilitate gene transfer, siRNA-based approaches and therapy for corneal scarring by ameliorating *TGFβI* or matrix metalloprotease expression in the cornea. However these have yet to be developed to the point of reaching clinical trials (Williams and Klebe, 2012). Treatment of CHED by gene replacement seems feasible in some respects as, compared with the retina, the corneal endothelium is relatively easily accessible. Additionally, as the mechanism is thought to be deficiency of the SLC4A11 protein, this makes CHED an ideal target for gene replacement. Approaches to gene therapy for trinucleotide expansion-type diseases, of which FECD due the *TCF4* CTG18.1 expansion is thought to be one, are also under development. These approaches include the use of Zinc finger nucleases, transcription-activator–like effector nucleases and CRISPR-Cas nucleases to shorten trinucleotide repeats. However the development of these methods is very much in its infancy and it is not known whether they will ultimately lead to effective therapies (Richard, 2015).

6.2.5 Strategies for cellular replacement in corneal disease

Traditionally, it was thought that endothelial cells do not divide (Bourne, 2003) and that therefore, as they die with age, the cells that remain have a limited regenerative capacity (Waring et al., 1982). However, a more recent study suggests that the corneal periphery contains a reservoir of stem-like cells that

replace damaged or dead endothelium (He et al., 2012). As human donor corneal tissue for use in corneal transplants is a limited resource and demand for it remains high, the possibility of endothelial regeneration is an exciting avenue for research. Novel treatments based on this approach might overcome the problems with existing graft operations, with the challenges of limited availability of graft tissue, graft rejection and the surgical complications of infection and glaucoma being avoided. Currently, research is aimed at identifying the optimal conditions for the isolation and culture of corneal endothelial cells and the results of animal studies in this area are promising (Zavala et al., 2013). If the progenitor cells in the corneal periphery could be used in this way, corneal endothelium prepared from the patient's corneal endothelial cells, bioengineered to correct the endothelial dystrophy-causing defect and grown in culture, could present an exciting therapeutic prospect for corneal endothelial disease. Full knowledge of the genetic architecture of endothelial dystrophies and their mechanisms of disease are an essential step in progressing towards these therapeutic avenues.

References

ABBRUZZESE, C., KRAHE, R., LIGUORI, M., TESSAROLO, D., SICILIANO, M. J., ASHIZAWA, T. & GIACANELLI, M. 1996. Myotonic dystrophy phenotype without expansion of (CTG)n repeat: an entity distinct from proximal myotonic myopathy (PROMM)? *Journal of Neurology,* 243**,** 715-21.

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nature methods,* 7**,** 248-9.

AHUJA, Y., BARATZ, K. H., MCLAREN, J. W., BOURNE, W. M. & PATEL, S. V. 2012. Decreased corneal sensitivity and abnormal corneal nerves in Fuchs endothelial dystrophy. *Cornea,* 31**,** 1257-63.

AKAMA, T. O., NISHIDA, K., NAKAYAMA, J., WATANABE, H., OZAKI, K., NAKAMURA, T., DOTA, A., KAWASAKI, S., INOUE, Y., MAEDA, N., YAMAMOTO, S., FUJIWARA, T., THONAR, E. J., SHIMOMURA, Y., KINOSHITA, S., TANIGAMI, A. & FUKUDA, M. N. 2000. Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulphotransferase gene. *Nature Genetics,* 26**,** 237-41.

ALBIN, R. L. 1998. Fuch's corneal dystrophy in a patient with mitochondrial DNA mutations. *Journal of Medical Genetics,* 35**,** 258-9.

ALDAVE, A. J., HAN, J. & FRAUSTO, R. F. 2013. Genetics of the corneal endothelial dystrophies: an evidence‐based review. *Clinical Genetics*

ALDAVE, A. J., RAYNER, S. A., SALEM, A. K., YOO, G. L., KIM, B. T., SAEEDIAN, M., SONMEZ, B. & YELLORE, V. S. 2006. No pathogenic mutations identified in the COL8A1 and COL8A2 genes in familial Fuchs corneal dystrophy. *Investigative Ophthalmology and Visual Science,* 47**,** 3787-90.

ALDAVE, A. J., YELLORE, V. S., BOURLA, N., MOMI, R. S., KHAN, M. A., SALEM, A. K., RAYNER, S. A., GLASGOW, B. J. & KURTZ, I. 2007. Autosomal recessive CHED associated with novel compound heterozygous mutations in SLC4A11. *Cornea,* 26**,** 896-900.

ALI, R. R. & SOWDEN, J. C. 2011. Regenerative medicine: DIY eye. *Nature,* 472**,** 42-3.

ALZUHAIRY, S., ALKATAN, H. M. & AL-RAJHI, A. A. 2015. Prevalence and histopathological characteristics of corneal stromal dystrophies in Saudi Arabia. *Middle East African Journal Ophthalmology,* 22**,** 179-85.

AMIEL, J., RIO, M., DE PONTUAL, L., REDON, R., MALAN, V., BODDAERT, N., PLOUIN, P., CARTER, N. P., LYONNET, S., MUNNICH, A. & COLLEAUX, L. 2007. Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *American Journal of Human Genetics,* 80**,** 988-93.

ASHAR, J. N., RAMAPPA, M. & VADDAVALLI, P. K. 2013. Paired-eye comparison of Descemet's stripping endothelial keratoplasty and penetrating keratoplasty in children with congenital hereditary endothelial dystrophy. *British Journal of Ophthalmology,* 97**,** 1247-9.

AZIZI, B., ZIAEI, A., FUCHSLUGER, T., SCHMEDT, T., CHEN, Y. & JURKUNAS, U. V. 2011. p53-regulated increase in oxidative-stress-induced apoptosis in Fuchs endothelial corneal dystrophy: A native tissue model. *Investigative Ophthalmology and Visual Science, 52,* 2278-89.

BAHN, C. F., FALLS, H. F., VARLEY, G. A., MEYER, R. F., EDELHAUSER, H. F. & BOURNE, W. M. 1984. Classification of corneal endothelial disorders based on neural crest origin. *Ophthalmology,* 91**,** 558-63.

BARATZ, K. H., TOSAKULWONG, N., RYU, E., BROWN, W. L., BRANHAM, K., CHEN, W., TRAN, K. D., SCHMID-KUBISTA, K. E., HECKENLIVELY, J. R. & SWAROOP, A. 2010. E2-2 protein and Fuchs's corneal dystrophy. *New England Journal of Medicine,* 363**,** 1016-1024.

BARRETT, J. C., FRY, B., MALLER, J. & DALY, M. J. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics,* 21**,** 263-5.

BERRY, F. B., SKARIE, J. M., MIRZAYANS, F., FORTIN, Y., HUDSON, T. J., RAYMOND, V., LINK, B. A. & WALTER, M. A. 2008. FOXC1 is required for cell viability and resistance to oxidative stress in the eye through the transcriptional regulation of FOXO1A. *Human Molecular Genetics,* 17**,** 490-505.

BISWAS, S., MUNIER, F. L., YARDLEY, J., HART-HOLDEN, N., PERVEEN, R., COUSIN, P., SUTPHIN, J. E., NOBLE, B., BATTERBURY, M. & KIELTY, C. 2001. Missense mutations in COL8A2, the gene encoding the 2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy. *Human Molecular Genetics,* 10**,** 2415-2423.

BOURNE, W. M. 2003. Biology of the corneal endothelium in health and disease. *Eye,* 17**,** 912-8.

BOURNE, W. M. & MCLAREN, J. W. 2004. Clinical responses of the corneal endothelium. *Experimental Eye Research,* 78**,** 561-572.

BOUTBOUL, S., BLACK, G. C., MOORE, J. E., SINTON, J., MENASCHE, M., MUNIER, F. L., LAROCHE, L., ABITBOL, M. & SCHORDERET, D. F. 2006. A subset of patients with epithelial basement membrane corneal dystrophy have mutations in TGFBI/BIGH3. *Human Mutation,* 27**,** 553-7.

BRESCHEL, T. S., MCINNIS, M. G., MARGOLIS, R. L., SIRUGO, G., CORNELIUSSEN, B., SIMPSON, S. G., MCMAHON, F. J., MACKINNON, D. F., XU, J. F., PLEASANT, N., HUO, Y., ASHWORTH, R. G., GRUNDSTROM, C., GRUNDSTROM, T., KIDD, K. K., DEPAULO, J. R. & ROSS, C. A. 1997. A novel, heritable, expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1. *Human Molecular Genetics,* 6**,** 1855-63.

BROCKSCHMIDT, A., TODT, U., RYU, S., HOISCHEN, A., LANDWEHR, C., BIRNBAUM, S., FRENCK, W., RADLWIMMER, B., LICHTER, P., ENGELS, H., DRIEVER, W., KUBISCH, C. & WEBER, R. G. 2007. Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Human Molecular Genetics,* 16**,** 1488-94.

BROMAN, K. W., MURRAY, J. C., SHEFFIELD, V. C., WHITE, R. L. & WEBER, J. L. 1998. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *American Journal Human Genetics,* 63**,** 861-9.

BRUINSMA, M., TONG, C. M. & MELLES, G. R. 2013. What does the future hold for the treatment of Fuchs endothelial dystrophy; will 'keratoplasty' still be a valid procedure? *Eye,* 27**,** 1115-22.

BUNCE, C., XING, W. & WORMALD, R. 2010. Causes of blind and partial sight certifications in England and Wales: April 2007-March 2008. *Eye,* 24**,** 1692-9.

CALLAGHAN, M., HAND, C. K., KENNEDY, S. M., FITZSIMON, J. S., COLLUM, L. M. T. & PARFREY, N. A. 1999. Homozygosity mapping and linkage analysis demonstrate that autosomal recessive congenital hereditary endothelial dystrophy (CHED) and autosomal dominant CHED are genetically distinct. *British Journal of Ophthalmology,* 83**,** 115-119.

CHIOU, A. G., KAUFMAN, S. C., BEUERMAN, R. W., OHTA, T., SOLIMAN, H. & KAUFMAN, H. E. 1999. Confocal microscopy in cornea guttata and Fuchs' endothelial dystrophy. *British Journal of Ophthalmology,* 83**,** 185-189.

CHNG, Z., PEH, G. S., HERATH, W. B., CHENG, T. Y., ANG, H. P., TOH, K. P., ROBSON, P., MEHTA, J. S. & COLMAN, A. 2013. High throughput gene expression analysis identifies reliable expression markers of human corneal endothelial cells. *PLoS One.,* 8**,** e67546. doi: 10.1371/journal.pone.0067546. Print 2013.

CHOI, Y., SIMS, G. E., MURPHY, S., MILLER, J. R. & CHAN, A. P. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One,* 7**,** e46688.

CHUNG, D. W., FRAUSTO, R. F., ANN, L. B., JANG, M. S. & ALDAVE, A. J. 2014. Functional impact of ZEB1 mutations associated with posterior polymorphous and Fuchs' endothelial corneal dystrophies. *Investigative Ophthalmology and Visual Science,* 55**,** 6159-66.

CIDECIYAN, A. V., JACOBSON, S. G., BELTRAN, W. A., SUMAROKA, A., SWIDER, M., IWABE, S., ROMAN, A. J., OLIVARES, M. B., SCHWARTZ, S. B., KOMAROMY, A. M., HAUSWIRTH, W. W. & AGUIRRE, G. D. 2013. Human retinal gene therapy for Leber congenital amaurosis shows advancing retinal degeneration despite enduring visual improvement. Proceedings of the National Academy of Sciences of the United States of America*,* 110**,** E517-25.

CLIFT, K. E., HALVERSON, C. M., FIKSDAL, A. S., KUMBAMU, A., SHARP, R. R. & MCCORMICK, J. B. 2015. Patients' views on incidental findings from clinical exome sequencing. *Applied and Translational Genomics,* 4:38-43.

COGAN, D. G., DONALDSON, D. D., KUWABARA, T. & MARSHALL, D. 1964. Microcystic Dystrophy of the Corneal Epithelium. *Transactions of the American Ophthalmology Society,* 62**,** 213-25.

CONRAD, D. F., PINTO, D., REDON, R., FEUK, L., GOKCUMEN, O., ZHANG, Y., AERTS, J., ANDREWS, T. D., BARNES, C., CAMPBELL, P., FITZGERALD, T., HU, M., IHM, C. H., KRISTIANSSON, K., MACARTHUR, D. G., MACDONALD, J. R., ONYIAH, I., PANG, A. W., ROBSON, S., STIRRUPS, K., VALSESIA, A., WALTER, K., WEI, J., TYLER-SMITH, C., CARTER, N. P., LEE, C., SCHERER, S. W. & HURLES, M. E. 2010. Origins and functional impact of copy number variation in the human genome. *Nature,* 464**,** 704-12.

COOPER, G. M., GOODE, D. L., NG, S. B., SIDOW, A., BAMSHAD, M. J., SHENDURE, J. & NICKERSON, D. A. 2010. Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nature methods,* 7**,** 250-1.

CZARNY, P., SEDA, A., WIELGORSKI, M., BINCZYK, E., MARKIEWICZ, B., KASPRZAK, E., JIMENEZ-GARCIA, M. P., GRABSKA-LIBEREK, I., PAWLOWSKA, E., BLASIAK, J., SZAFLIK, J. & SZAFLIK, J. P. 2014. Mutagenesis of mitochondrial DNA in Fuchs endothelial corneal dystrophy. *Mutation Research,* 760:42-7.

DAVIDSON, A. E., LISKOVA, P., EVANS, C. J., DUDAKOVA, L., NOSKOVA, L., PONTIKOS, N., HARTMANNOVA, H., HODANOVA, K., STRANECKY, V., KOZMIK, Z., LEVIS, H. J., IDIGO, N., SASAI, N., MAHER, G. J., BELLINGHAM, J., VELI, N., EBENEZER, N. D., CHEETHAM, M. E., DANIELS, J. T., THAUNG, C. M., JIRSOVA, K., PLAGNOL, V., FILIPEC, M., KMOCH, S., TUFT, S. J. & HARDCASTLE, A. J. 2016. Autosomal-Dominant Corneal Endothelial Dystrophies CHED1 and PPCD1 Are Allelic Disorders Caused by Non-coding Mutations in the Promoter of OVOL2. *American Journal of Human Genetics,* 98**,** 75-89.

DEN HOLLANDER, A. I., KOENEKOOP, R. K., YZER, S., LOPEZ, I., ARENDS, M. L., VOESENEK, K. E., ZONNEVELD, M. N., STROM, T. M., MEITINGER, T., BRUNNER, H. G., HOYNG, C. B., VAN DEN BORN, L. I., ROHRSCHNEIDER, K. & CREMERS, F. P. 2006. Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. *American Journal of Human Genetics,* 79**,** 556-61.

DESIR, J. & ABRAMOWICZ, M. 2008. Congenital hereditary endothelial dystrophy with progressive sensorineural deafness (Harboyan syndrome). *Orphanet Journal of Rare Diseases,* 3**,** 28.

DESIR, J., MOYA, G., REISH, O., VAN REGEMORTER, N., DECONINCK, H., DAVID, K. L., MEIRE, F. M. & ABRAMOWICZ, M. J. 2007. Borate transporter SLC4A11 mutations cause both Harboyan syndrome and non-syndromic corneal endothelial dystrophy. *Journal of Medical Genetics,* 44**,** 322-326.

DU, J., ALEFF, R. A., SORAGNI, E., KALARI, K., NIE, J., TANG, X., DAVILA, J., KOCHER, J. P., PATEL, S. V., GOTTESFELD, J. M., BARATZ, K. H. & WIEBEN, E. D. 2015. RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. *Journal of Biological Chemistry,* 290**,** 5979-90.

EDELHAUSER, H. F. 2006. The balance between corneal transparency and edema: the Proctor Lecture. *Investigative Ophthalmology and Visual Science,* 47**,** 1754-67.

EDVARDSON, S., JALAS, C., SHAAG, A., ZENVIRT, S., LANDAU, C., LERER, I. & ELPELEG, O. 2011. A deleterious mutation in the LOXHD1 gene causes autosomal recessive hearing loss in Ashkenazi Jews. *American Journal of Medical Genetics Part A,* 155A**,** 1170-2.

EGHRARI, A. O. & GOTTSCH, J. D. 2010. Fuchs' corneal dystrophy. *Expert Review of Ophthalmology,* 5**,** 147-159.

EGHRARI, A. O., MCGLUMPHY, E. J., ILIFF, B. W., WANG, J., EMMERT, D., RIAZUDDIN, S. A., KATSANIS, N. & GOTTSCH, J. D. 2012. Prevalence and severity of fuchs corneal dystrophy in Tangier Island. *American Journal of Ophthalmology,* 153**,** 1067-72.

EGHRARI, A. O., RIAZUDDIN, S. A. & GOTTSCH, J. D. 2015a. Fuchs Corneal Dystrophy. *Progress in Molecular Biology and Translational Science,* 134:79-97.

EGHRARI, A. O., RIAZUDDIN, S. A. & GOTTSCH, J. D. 2015b. Overview of the Cornea: Structure, Function, and Development. *Progress in Molecular Biology and Translational Science,* 134:7-23.

ELHALIS, H., AZIZI, B. & JURKUNAS, U. V. 2010. Fuchs endothelial corneal dystrophy. *Ocular Surface,* 8**,** 173-84.

ELKINGTON, A. R., FRANK, H. & GREANEY, M. 1999. Clinical optics.

ETHNICITY AND NATIONAL IDENTITY IN ENGLAND AND WALES 2011 [Online]. [Accessed 29 December 2015]. Available from: http://www.ons.gov.uk/ons/rel/census/2011-census/key-statistics-for-local-authorities-in-england-and-wales/rpt-ethnicity.html - tab-Ethnicity-in-England-and-Wales.

FISHELSON, M. & GEIGER, D. 2002. Exact genetic linkage computations for general pedigrees. *Bioinformatics,* 18 Suppl 1**,** S189-98.

FOSTER, A., GILBERT, C. & JOHNSON, G. 2008. Changing patterns in global blindness: 1988–2008. *Community Eye Health,* 21**,** 37.

FOSTER, A. & RESNIKOFF, S. 2005. The impact of Vision 2020 on global blindness. *Eye,* 19**,** 1133-1135.

FREEGARD, T. J. 1997. The physical basis of transparency of the normal cornea. *Eye,* 11**,** 465-71.

FRICK, K. D. & FOSTER, A. 2003. The magnitude and cost of global blindness: an increasing problem that can be alleviated. *American Journal of Ophthalmology,* 135**,** 471-6.

FUCHS, E. 1910. Dystrophia epithelialis corneae. *Albrecht von Graefes Archiv für Ophthalmologie,* 76**,** 478-508.

GONG, H., WANG, Y., QI, X., WANG, C., LIU, T., REN, S. & WANG, Y. 2012. Differential response of lens crystallins and corneal crystallins in degenerative corneas. *Experimental Eye Research,* 96**,** 55-64.

GONG, Y., SLEE, R. B., FUKAI, N., RAWADI, G., ROMAN-ROMAN, S., REGINATO, A. M., WANG, H., CUNDY, T., GLORIEUX, F. H., LEV, D., ZACHARIN, M., OEXLE, K., MARCELINO, J., SUWAIRI, W., HEEGER, S., SABATAKOS, G., APTE, S., ADKINS, W. N., ALLGROVE, J., ARSLAN-KIRCHNER, M., BATCH, J. A., BEIGHTON, P., BLACK, G. C., BOLES, R. G., BOON, L. M., BORRONE, C., BRUNNER, H. G., CARLE, G. F., DALLAPICCOLA, B., DE PAEPE, A., FLOEGE, B., HALFHIDE, M. L., HALL, B., HENNEKAM, R. C., HIROSE, T., JANS, A., JUPPNER, H., KIM, C. A., KEPPLER-NOREUIL, K., KOHLSCHUETTER, A., LACOMBE, D., LAMBERT, M., LEMYRE, E., LETTEBOER, T., PELTONEN, L., RAMESAR, R. S., ROMANENGO, M., SOMER, H., STEICHEN-GERSDORF, E., STEINMANN, B., SULLIVAN, B., SUPERTI-FURGA, A., SWOBODA, W., VAN DEN BOOGAARD, M. J., VAN HUL, W., VIKKULA, M., VOTRUBA, M., ZABEL, B., GARCIA, T., BARON, R., OLSEN, B. R. & WARMAN, M. L. 2001. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. *Cell,* 107**,** 513-23.

GOTTSCH, J. D., BOWERS, A. L., MARGULIES, E. H., SEITZMAN, G. D., KIM, S. W., SAHA, S., JUN, A. S., STARK, W. J. & LIU, S. H. 2003. Serial analysis of gene expression in the corneal endothelium of Fuchs dystrophy. *Investigative Ophthalmology Visual Science,* 44**,** 594.

GOTTSCH, J. D., SUNDIN, O. H., RENCS, E. V., EMMERT, D. G., STARK, W. J., CHENG, C. J. & SCHMIDT, G. W. 2006. Analysis and documentation of progression of Fuchs corneal dystrophy with retroillumination photography. *Cornea,* 25**,** 485-9.

GOTTSCH, J. D., ZHANG, C., SUNDIN, O. H., BELL, W. R., STARK, W. J. & GREEN, W. R. 2005. Fuchs corneal dystrophy: aberrant collagen distribution in an L450W mutant of the COL8A2 gene. *Investigative Ophthalmology and Visual Science,* 46**,** 4504-11.

GREEN, R. C., BERG, J. S., GRODY, W. W., KALIA, S. S., KORF, B. R., MARTIN, C. L., MCGUIRE, A. L., NUSSBAUM, R. L., O'DANIEL, J. M., ORMOND, K. E., REHM, H. L., WATSON, M. S., WILLIAMS, M. S. & BIESECKER, L. G. 2013. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genetics in Medicine,* 15**,** 565-74.

GRILLET, N., SCHWANDER, M., HILDEBRAND, M. S., SCZANIECKA, A., KOLATKAR, A., VELASCO, J., WEBSTER, J. A., KAHRIZI, K., NAJMABADI, H. & KIMBERLING, W. J. 2009. Mutations in LOXHD1, an evolutionarily conserved stereociliary protein, disrupt hair cell function in mice and cause progressive hearing loss in humans. *American Journal of Human Genetics,* 85**,** 328-337.

GROGER, N., FROHLICH, H., MAIER, H., OLBRICH, A., KOSTIN, S., BRAUN, T. & BOETTGER, T. 2010. SLC4A11 prevents osmotic imbalance leading to corneal endothelial dystrophy, deafness, and polyuria. *Journal of Biological Chemistry,* 285**,** 14467-14474.

GUELL, J. L., EL HUSSEINY, M. A., MANERO, F., GRIS, O. & ELIES, D. 2014. Historical Review and Update of Surgical Treatment for Corneal Endothelial Diseases. *Ophthalmology and Therapy***,** 3, 1-15.

GUPTA, R., KUMAWAT, B. L., PALIWAL, P., TANDON, R., SHARMA, N., SEN, S., KASHYAP, S., NAG, T. C., VAJPAYEE, R. B. & SHARMA, A. 2015. Association of ZEB1 and TCF4 rs613872 changes with late onset Fuchs endothelial corneal dystrophy in patients from northern India. *Molecular Vision,* 21**,** 1252-60.

GUO, X., LEHNER, K., O'CONNELL, K., ZHANG, J., DAVE, S. S. & JINKS-ROBERTSON, S. 2015. SMRT Sequencing for Parallel Analysis of Multiple Targets and Accurate SNP Phasing. *G3 (Bethesda).* 5**,** 2801-8.

HAN, K. E., CHOI, S. I., KIM, T. I., MAENG, Y. S., STULTING, R. D., JI, Y. W. & KIM, E. K. 2016. Pathogenesis and treatments of TGFBI corneal dystrophies. *Progress in Retinal and Eye Research,* 50:67-88.

HAN, S. B., ANG, H. P., POH, R., CHAURASIA, S. S., PEH, G., LIU, J., TAN, D. T., VITHANA, E. N. & MEHTA, J. S. 2013. Mice with a targeted disruption of Slc4a11 model the progressive corneal changes of congenital hereditary endothelial dystrophy. *Investigative Ophthalmology and Visual Science,* 54**,** 6179-89.

HARRISON, D. A., JOOS, C. & AMBROSIO JR, R. 2003. Morphology of corneal Basal epithelial cells by in vivo slit-scanning confocal microscopy. *Cornea,* 22**,** 246-8.

HE, Z., CAMPOLMI, N., GAIN, P., HA THI, B. M., DUMOLLARD, J. M., DUBAND, S., PEOC'H, M., PISELLI, S., GARRAUD, O. & THURET, G. 2012. Revisited microanatomy of the corneal endothelial periphery: new evidence for continuous centripetal migration of endothelial cells in humans. *Stem Cells,* 30**,** 2523-34.

HEJTMANCIK, J. F. & NICKERSON, J. M. 2015. Overview of the Visual System. *Progress in Molecular Biology and Translational Science,* 134:1-4.

HEMADEVI, B., VEITIA, R. A., SRINIVASAN, M., ARUNKUMAR, J., PRAJNA, N. V., LESAFFRE, C. & SUNDARESAN, P. 2008. Identification of mutations in the SLC4A11 gene in patients with recessive congenital hereditary endothelial dystrophy. *Archives of Ophthalmology,* 126**,** 700-708.

HENIKOFF, S. & HENIKOFF, J. G. 1992. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 10915-9.

HÉON, E., GREENBERG, A., KOPP, K. K., ROOTMAN, D., VINCENT, A. L., BILLINGSLEY, G., PRISTON, M., DORVAL, K. M., CHOW, R. L. & MCINNES, R. R. 2002. VSX1: a gene for posterior polymorphous dystrophy and keratoconus. *Human Molecular Genetics,* 11**,** 1029.

HONG, J. P., KIM, T. I., CHUNG, J. L., HUANG, D., CHO, H. S. & KIM, E. K. 2011. Analysis of deposit depth and morphology in granular corneal dystrophy type 2 using fourier domain optical coherence tomography. *Cornea,* 30**,** 729-38.

IGO, R. P., JR., KOPPLIN, L. J., JOSEPH, P., TRUITT, B., FONDRAN, J., BARDENSTEIN, D., ALDAVE, A. J., CROASDALE, C. R., PRICE, M. O., ROSENWASSER, M., LASS, J. H. & IYENGAR, S. K. 2012. Differing roles for TCF4 and COL8A2 in central corneal thickness and fuchs endothelial corneal dystrophy. *PLoS One,* 7**,** e46742.

IONITA-LAZA, I., MCCALLUM, K., XU, B. & BUXBAUM, J. D. 2016. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nature Genetics,* 48**,** 214-20.

IRVINE, A. D., CORDEN, L. D., SWENSSON, O., SWENSSON, B., MOORE, J. E., FRAZER, D. G., SMITH, F. J., KNOWLTON, R. G., CHRISTOPHERS, E., ROCHELS, R., UITTO, J. & MCLEAN, W. H. 1997. Mutations in cornea-specific keratin K3 or K12 genes cause Meesmann's corneal dystrophy. *Nature Genetics,* 16**,** 184-7.

JACKSON, A. J., ROBINSON, F. O., FRAZER, D. G. & ARCHER, D. B. 1999. Corneal guttata: a comparative clinical and specular micrographic study. *Eye,* 13**,** 737-43.

JALIMARADA, S. S., OGANDO, D. G., VITHANA, E. N. & BONANNO, J. A. 2013. Ion transport function of SLC4A11 in corneal endothelium. *Investigative Ophthalmology and Visual Science,* 54**,** 4330-40.

JAWORSKI, C. J., MOREIRA, E., LI, A., LEE, R. & RODRIGUEZ, I. R. 2001. A family of 12 human genes containing oxysterol-binding domains. *Genomics,* 78**,** 185-96.

JIAO, X., SULTANA, A., GARG, P., RAMAMURTHY, B., VEMUGANTI, G. K., GANGOPADHYAY, N., HEJTMANCIK, J. F. & KANNABIRAN, C. 2007. Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11. *Journal of Medical Genetics,* 44**,** 64-68.

JOHNSON, D. H., BOURNE, W. M. & CAMPBELL, R. J. 1982. The ultrastructure of Descemet's membrane. I. Changes with age in normal corneas. *Archives in Ophthalmology,* 100**,** 1942-7.

JONSSON, F., BYSTROM, B., DAVIDSON, A. E., BACKMAN, L. J., KELLGREN, T. G., TUFT, S. J., KOSKELA, T., RYDEN, P., SANDGREN, O., DANIELSON, P., HARDCASTLE, A. J. & GOLOVLEVA, I. 2015. Mutations in collagen, type XVII, alpha 1 (COL17A1) cause epithelial recurrent erosion dystrophy (ERED). *Human Mutation,* 36**,** 463-73

JUDISCH, G. F. & MAUMENEE, I. H. 1978. Clinical differentiation of recessive congenital hereditary endothelial dystrophy and dominant hereditary endothelial dystrophy. *American Journal of Ophthalmology,* 85**,** 606-12.

JURKUNAS, U. V., BITAR, M. S., FUNAKI, T. & AZIZI, B. 2010. Evidence of oxidative stress in the pathogenesis of fuchs endothelial corneal dystrophy. *American Journal of Pathology,* 177**,** 2278-89.

KAO, L., AZIMOV, R., ABULADZE, N., NEWMAN, D. & KURTZ, I. 2015. Human SLC4A11-C functions as a DIDS-stimulatable H(+)(OH(-)) permeation pathway: partial correction of R109H mutant transport. *American Journal of Physiology - Cell Physiology* 308**,** C176-88.

KATSANIS, S. H. & KATSANIS, N. 2013. Molecular genetic testing and the future of clinical genomics. *Nature Reviews Genetics,* 14**,** 415-26.

KAUFMAN, S. C., MUSCH, D. C., BELIN, M. W., COHEN, E. J., MEISLER, D. M., REINHART, W. J., UDELL, I. J. & VAN METER, W. S. 2004. Confocal microscopy: a report by the American Academy of Ophthalmology. *Ophthalmology,* 111**,** 396-406.

KEATES, R. H. & CVINTAL, T. 1965. Congenital hereditary corneal dystrophy. *American Journal of Ophthalmology,* 60**,** 892-4.

KEENAN, T. D., JONES, M. N., RUSHTON, S. & CARLEY, F. M. 2012. Trends in the indications for corneal graft surgery in the United Kingdom: 1999 Through 2009. *Archives of Ophthalmology,* 130**,** 621-628.

KIM, E. K., LEE, H. & CHOI, S. I. 2015a. Molecular Pathogenesis of Corneal Dystrophies: Schnyder Dystrophy and Granular Corneal Dystrophy type 2. *Progress in Molecular Biology and Translational Science,* 134:99-115.

KIM, J. H., KO, J. M. & TCHAH, H. 2015b. Fuchs Endothelial Corneal Dystrophy in a Heterozygous Carrier of Congenital Hereditary Endothelial Dystrophy Type 2 with a Novel Mutation in SLC4A11. *Ophthalmic Genetics,* 36**,** 284-6.

KIM, M. Y., LEE, H. K., PARK, J. S., PARK, S. H., KWON, H. B. & SOH, J. 1999. Identification of a zeta-crystallin (quinone reductase)-like 1 gene (CRYZL1) mapped to human chromosome 21q22.1. *Genomics,* 57**,** 156-9.

KIRCHER, M., WITTEN, D. M., JAIN, P., O'ROAK, B. J., COOPER, G. M. & SHENDURE, J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics,* 46**,** 310-5.

KIRKNESS, C. M., MCCARTNEY, A., RICE, N., GARNER, A. & STEELE, A. 1987. Congenital hereditary corneal oedema of Maumenee: its clinical features, management, and pathology. *British Journal Ophthalmology,* 71**,** 130-144.

KIRKNESS, C. M. & MOSHEGOV, C. 1988. Post-keratoplasty glaucoma. *Eye,* 2**,** S19-26.

KLINTWORTH, G. K. 2009. Corneal dystrophies. *Orphanet Journal of Rare Diseases,* 23, 4, 7.

KO, F., PAPADOPOULOS, M. & KHAW, P. T. 2015. Primary congenital glaucoma. *Progress in Brain Research,* 221:177-89.

KOBAYASHI, A., FUJIKI, K., MURAKAMI, A., KATO, T., CHEN, L. Z., ONOE, H., NAKAYASU, K., SAKURAI, M., TAKAHASHI, M. & SUGIYAMA, K. 2004. Analysis of COL8A2 gene mutation in Japanese patients with Fuchs endothelial dystrophy and posterior polymorphous dystrophy. *Japanese Journal of Ophthalmology,* 48**,** 195-198.

KOZAREWA, I., ARMISEN, J., GARDNER, A. F., SLATKO, B. E. & HENDRICKSON, C. L. 2015. Overview of Target Enrichment Strategies. *Current Protocols in Molecular Biology,* 112:7, 1-23.

KOZLOWSKI, P., DE MEZER, M. & KRZYZOSIAK, W. J. 2010. Trinucleotide repeats in human genome and exome. *Nucleic Acids Research,* 38**,** 4027-39.

KRACHMER, J. H., MANNIS, M. J. & HOLLAND, E. J. 2011. *Cornea*, Moseby.

KRACHMER, J. H., PURCELL JR, J. J., YOUNG, C. W. & BUCHER, K. D. 1978. Corneal endothelial dystrophy: a study of 64 families. *Archives of Ophthalmology,* 96**,** 2036-9.

KRAFCHAK, C. M., PAWAR, H., MOROI, S. E., SUGAR, A., LICHTER, P. R., MACKEY, D. A., MIAN, S., NAIRUS, T., ELNER, V. & SCHTEINGART, M. T. 2005. Mutations in TCF8 cause posterior polymorphous corneal dystrophy and ectopic expression of COL4A3 by corneal endothelial cells. *The American Journal of Human Genetics,* 77**,** 694-708.

KUMAR, A., BHATTACHARJEE, S., PRAKASH, D. R. & SADANAND, C. S. 2007. Genetic analysis of two Indian families affected with congenital hereditary endothelial dystrophy: two novel mutations in SLC4A11. *Molecular Vision,* 13**,** 39-46.

KUOT, A. 2015. Molecular Biology and Genetics of Fuchs' Endothelial Corneal Dystrophy. Ph.D Thesis. Flinders University, School of Medicine.

KUOT, A., HEWITT, A. W., GRIGGS, K., KLEBE, S., MILLS, R., JHANJI, V., CRAIG, J. E., SHARMA, S. & BURDON, K. P. 2012. Association of TCF4 and CLU polymorphisms with Fuchs' endothelial dystrophy and implication of CLU and TGFBI proteins in the disease process. *European Journal of Human Genetics,* 20**,** 632-8.

KUOT, A., MILLS, R., CRAIG, J. E., SHARMA, S. & BURDON, K. P. 2013. Screening of the COL8A2 gene in an Australian family with early-onset Fuchs' endothelial corneal dystrophy. *Clinical and Experimental Ophthalmology*, 42, 198-200.

LAGANOWSKI, H. C., SHERRARD, E. S., MUIR, M. G. & BUCKLEY, R. J. 1991. Distinguishing features of the iridocorneal endothelial syndrome and posterior polymorphous dystrophy: value of endothelial specular microscopy. *British Journal of Ophthalmology,* 75**,** 212-6.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9**,** 357-9.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LARSEN, W. J., SCHOENWOLF, G. C., BLEYL, S. B., BRAUER, P. R. & FRANCIS-WEST, P. H. 2009. *Larsen's Human Embryology*, Churchill Livingstone.

LECHNER, J., DASH, D. P., MUSZYNSKA, D., HOSSEINI, M., SEGEV, F., GEORGE, S., FRAZER, D. G., MOORE, J. E., KAYE, S. B., YOUNG, T., SIMPSON, D. A., CHURCHILL, A. J., HEON, E. & WILLOUGHBY, C. E. 2013. Mutational spectrum of the ZEB1 gene in corneal dystrophies supports a genotype-phenotype correlation. *Investigative Ophthalmology and Visual Science,* 54**,** 3215-23.

LI, Y. J., MINEAR, M. A., QIN, X., RIMMLER, J., HAUSER, M. A., ALLINGHAM, R. R., IGO, R. P., LASS, J. H., IYENGAR, S. K., KLINTWORTH, G. K., AFSHARI, N. A. & GREGORY, S. G. 2014. Mitochondrial polymorphism A10398G and Haplogroup I are associated with Fuchs' endothelial corneal dystrophy. *Investigative Ophthalmology and Visual Science,* 55**,** 4577-84.

LI, Y. J., MINEAR, M. A., RIMMLER, J., ZHAO, B., BALAJONDA, E., HAUSER, M. A., ALLINGHAM, R. R., EGHRARI, A. O., RIAZUDDIN, S. A., KATSANIS, N., GOTTSCH, J. D., GREGORY, S. G., KLINTWORTH, G. K. & AFSHARI, N. A. 2011. Replication of TCF4 through association and linkage studies in late-onset Fuchs endothelial corneal dystrophy. *PLoS One,* 6**,** e18044.

LIEW, G., MICHAELIDES, M. & BUNCE, C. 2014. A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. *BMJ Open,* 4**,** e004015. doi: 10.1136/bmjopen-2013-004015.

LISCH, W., BUTTNER, A., OEFFNER, F., BODDEKER, I., ENGEL, H., LISCH, C., ZIEGLER, A. & GRZESCHIK, K. 2000. Lisch corneal dystrophy is genetically distinct from Meesmann corneal dystrophy and maps to xp22.3. *American Journal of Ophthalmology,* 130**,** 461-8.

LISCH, W., STEUHL, K. P., LISCH, C., WEIDLE, E. G., EMMIG, C. T., COHEN, K. L. & PERRY, H. D. 1992. A new, band-shaped and whorled microcystic dystrophy of the corneal epithelium. *American Journal of Ophthalmology,* 114**,** 35-44.

LISKOVA, P., EVANS, C. J., DAVIDSON, A. E., ZALIOVA, M., DUDAKOVA, L., TRKOVA, M., STRANECKY, V., CARNT, N., PLAGNOL, V., VINCENT, A. L., TUFT, S. J. & HARDCASTLE, A. J. 2015. Heterozygous deletions at the ZEB1 locus verify haploinsufficiency as the mechanism of disease for posterior polymorphous corneal dystrophy type 3. *European Journal of Human Genetics***,** doi: 10.1038/ejhg.2015.232.

LISKOVA, P., PRESCOTT, Q., BHATTACHARYA, S. S. & TUFT, S. J. 2007. British family with early-onset Fuchs' endothelial corneal dystrophy associated with p.L450W mutation in the COL8A2 gene. *British Journal of Ophthalmology,* 91**,** 1717-8.

LOEWENSTEIN, A., GEYER, O., HOURVITZ, D. & LAZAR, M. 1991. The association of Fuch's corneal endothelial dystrophy with angle closure glaucoma. *British Journal of Ophthalmology,* 75**,** 510.

LOGANATHAN, S. K., LUKOWSKI, C. M. & CASEY, J. R. 2016. The cytoplasmic domain is essential for transport function of the integral membrane transport protein SLC4A11. *American Journal of Physiology - Cell Physiology,* 310**,** C161-74.

LONGSHORE, J. W. & TARLETON, J. 1996. Dynamic mutations in human genes: a review of trinucleotide repeat diseases. *Journal of Genetics,* 75**,** 193-217.

LOOMIS, E. W., EID, J. S., PELUSO, P., YIN, J., HICKEY, L., RANK, D., MCCALMON, S., HAGERMAN, R. J., TASSONE, F. & HAGERMAN, P. J. 2013. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Research,* 23**,** 121-8.

LOPEZ, I. A., ROSENBLATT, M. I., KIM, C., GALBRAITH, G. C., JONES, S. M., KAO, L., NEWMAN, D., LIU, W., YEH, S. & PUSHKIN, A. 2009. Slc4a11 Gene disruption in mice: cellular targets of sensorineural abnormalities. *Journal of Biological Chemistry,* 284**,** 26882-26896.

LORENZETTI, D. W., UOTILA, M. H., PARIKH, N. & KAUFMAN, H. E. 1967. Central cornea guttata. Incidence in the general population. *American Journal of Ophthalmology,* 64**,** 1155-8.

LOUTTIT, M. D., KOPPLIN, L. J., IGO, R. P., JR., FONDRAN, J. R., TAGLIAFERRI, A., BARDENSTEIN, D., ALDAVE, A. J., CROASDALE, C. R., PRICE, M. O., ROSENWASSER, G. O., LASS, J. H. & IYENGAR, S. K. 2012. A multicenter study to map genes for Fuchs endothelial corneal dystrophy: baseline characteristics and heritability. *Cornea,* 31**,** 26-35.

LWIGALE, P. Y. 2015. Corneal Development: Different Cells from a Common Progenitor. *Progress in Molecular and Translational Science,* 134:43-59.

MAUMENEE, A. E. 1960. Congenital hereditary corneal dystrophy. *American Journal Ophthalmology,* 50**,** 1114-24.

MAURICE, D. M. 1974. A scanning slit optical microscope. *Investigative Ophthalmology,* 13**,** 1033-7.

MCCAREY, B. E., EDELHAUSER, H. F. & LYNN, M. J. 2008. Review of corneal endothelial specular microscopy for FDA clinical trials of refractive procedures, surgical devices and new intraocular drugs and solutions. *Cornea,* 27**,** 1.

MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., GARIMELLA, K., ALTSHULER, D., GABRIEL, S., DALY, M. & DEPRISTO, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research,* 20**,** 1297-303.

MCMURRAY, C. T. 2010. Mechanisms of trinucleotide repeat instability during human development. *Nature Reviews Genetics,* 11**,** 786-99.

MEHTA, J. S., HEMADEVI, B., VITHANA, E. N., ARUNKUMAR, J., SRINIVASAN, M., PRAJNA, V., TAN, D. T., AUNG, T. & SUNDARESAN, P. 2010. Absence of phenotype-genotype correlation of patients expressing mutations in the SLC4A11 gene. *Cornea,* 29**,** 302-306.

MELLES, G., EGGINK, F., LANDER, F., PELS, E., RIETVELD, F., BEEKHUIS, W. H. & BINDER, P. S. 1998. A surgical technique for posterior lamellar keratoplasty. *Cornea,* 17**,** 618-626.

MELLES, G. R., SAN ONG, T., VERVERS, B. & VAN DER WEES, J. 2006. Descemet membrane endothelial keratoplasty (DMEK). *Cornea,* 25**,** 987-990.

MIELCAREK, M. 2015. Huntington's disease is a multi-system disorder. *Rare Diseases,* 3**,** e1058464. doi: 10.1080/21675511.2015.1058464. eCollection 2015.

MITTAL, V., MITTAL, R. & SANGWAN, V. S. 2011. Successful Descemet stripping endothelial keratoplasty in congenital hereditary endothelial dystrophy. *Cornea,* 30**,** 354-356.

MOHAN, R. R., RODIER, J. T. & SHARMA, A. 2013. Corneal gene therapy: basic science and translational perspective. *Ocular Surface,* 11**,** 150-64.

MOK, J. W., KIM, H. S. & JOO, C. K. 2009. Q455V mutation in COL8A2 is associated with Fuchs' corneal dystrophy in Korean patients. *Eye,* 23**,** 895-903.

MOOTHA, V. V., GONG, X., KU, H. C. & XING, C. 2014. Association and Familial Segregation of CTG18.1 Trinucleotide Repeat Expansion of TCF4 Gene in Fuchs' Endothelial Corneal Dystrophy. *Investigative Ophthalmology and Visual Science,* 55**,** 33-42.

MOOTHA, V. V., HUSSAIN, I., CUNNUSAMY, K., GRAHAM, E., GONG, X., NEELAM, S., XING, C., KITTLER, R. & PETROLL, W. M. 2015. TCF4 Triplet Repeat Expansion and Nuclear RNA Foci in Fuchs' Endothelial Corneal Dystrophy. *Investigative Ophthalmology and Visual Science,* 56**,** 2003-11.

MORGAN, J. E., CARR, I. M., SHERIDAN, E., CHU, C. E., HAYWARD, B., CAMM, N., LINDSAY, H. A., MATTOCKS, C. J., MARKHAM, A. F., BONTHRON, D. T. & TAYLOR, G. R. 2010. Genetic diagnosis of familial breast cancer using clonal sequencing. *Human Mutation,* 31**,** 484-91.

MUNIER, F. L., FRUEH, B. E., OTHENIN-GIRARD, P., UFFER, S., COUSIN, P., WANG, M. X., HEON, E., BLACK, G. C., BLASI, M. A., BALESTRAZZI, E., LORENZ, B., ESCOTO, R., BARRAQUER, R., HOELTZENBEIN, M., GLOOR, B., FOSSARELLO, M., SINGH, A. D., ARSENIJEVIC, Y., ZOGRAFOS, L. & SCHORDERET, D. F. 2002. BIGH3 mutation spectrum in corneal dystrophies. *Investigative Ophthalmology and Visual Science,* 43**,** 949-54.

MURAGAKI, Y., JACENKO, O., APTE, S., MATTEI, M. G., NINOMIYA, Y. & OLSEN, B. R. 1991. The alpha 2(VIII) collagen gene. A novel member of the short chain collagen family located on the human chromosome 1. *Journal of Biological Chemistry,* 266**,** 7721-7.

MUSCH, D. C., NIZIOL, L. M., STEIN, J. D., KAMYAR, R. M. & SUGAR, A. 2011. Prevalence of corneal dystrophies in the United States: estimates from claims data. *Investigative Ophthalmology Visual Science,* 52**,** 6959-63.

NAGAKI, Y., HAYASAKA, S., KITAGAWA, K. & YAMAMOTO, S. 1996. Primary cornea guttata in Japanese patients with cataract: specular microscopic observations. *Japanese Journal Ophthalmology,* 40**,** 520-5.

NAGARSHETH, M., SINGH, A., SCHMOTZER, B., BABINEAU, D. C., SUGAR, J., LEE, W. B., IYENGAR, S. K. & LASS, J. H. 2012. Relationship Between Fuchs Endothelial Corneal Dystrophy Severity and Glaucoma and/or Ocular Hypertension. *Archives Ophthalmology,* 130**,** 1384-8.

NAKANO, M., OKUMURA, N., NAKAGAWA, H., KOIZUMI, N., IKEDA, Y., UENO, M., YOSHII, K., ADACHI, H., ALEFF, R. A., BUTZ, M. L., HIGHSMITH, W. E., TASHIRO, K., WIEBEN, E. D., KINOSHITA, S. & BARATZ, K. H. 2015. Trinucleotide Repeat Expansion in the TCF4 Gene in Fuchs' Endothelial Corneal Dystrophy in Japanese. *Investigative Ophthalmology and Visual Science,* 56**,** 4865-9.

NANDA, G. G., PADHY, B., SAMAL, S., DAS, S. & ALONE, D. P. 2014. Genetic association of TCF4 intronic polymorphisms, CTG18.1 and rs17089887, with Fuchs' endothelial corneal dystrophy in an Indian population. *Investigative Ophthalmology and Visual Science,* 55**,** 7674-80.

NG, P. C. & HENIKOFF, S. 2001. Predicting deleterious amino acid substitutions. *Genome Research,* 11**,** 863-74.

NISCHAL, K. K. 2015. Genetics of Congenital Corneal Opacification--Impact on Diagnosis and Treatment. *Cornea,* 34**,** S24-34.

O'ROAK, B. J., VIVES, L., FU, W., EGERTSON, J. D., STANAWAY, I. B., PHELPS, I. G., CARVILL, G., KUMAR, A., LEE, C., ANKENMAN, K., MUNSON, J., HIATT, J. B., TURNER, E. H., LEVY, R., O'DAY, D. R., KRUMM, N., COE, B. P., MARTIN, B. K., BORENSTEIN, E., NICKERSON, D. A., MEFFORD, H. C., DOHERTY, D., AKEY, J. M., BERNIER, R., EICHLER, E. E. & SHENDURE, J. 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science.,* 338**,** 1619-22.

OGANDO, D. G., JALIMARADA, S. S., ZHANG, W., VITHANA, E. N. & BONANNO, J. A. 2013. SLC4A11 is an EIPA-sensitive Na(+) permeable pHi regulator. *American Journal of Physioliology - Cell Physiology,* 305**,** C716-27.

ORR, H. T. & ZOGHBI, H. Y. 2007. Trinucleotide repeat disorders. *Annual Review of Neuroscience,* 30**,** 575-621.

OZDEMIR, B., KUBALOGLU, A., KOYTAK, A., COSKUN, E., CINAR, Y., SARI, E. S. & OZERTURK, Y. 2012. Penetrating keratoplasty in congenital hereditary endothelial dystrophy. *Cornea,* 31**,** 359-65.

PALIWAL, P., SHARMA, A., TANDON, R., SHARMA, N., TITIYAL, J. S., SEN, S., NAG, T. C. & VAJPAYEE, R. B. 2010. Congenital hereditary endothelial dystrophy-mutation analysis of SLC4A11 and genotype-phenotype correlation in a North Indian patient cohort. *Molecular Vision,* 16**,** 2955-2963.

PAPADOPOULOS, M., CABLE, N., RAHI, J. & KHAW, P. T. 2007. The British Infantile and Childhood Glaucoma (BIG) Eye Study. *Investigative Ophthalmology and Visual Science,* 48**,** 4100-6.

PARK, C. Y., LEE, J. K., GORE, P. K., LIM, C. Y. & CHUCK, R. S. 2015. Keratoplasty in the United States: A 10-Year Review from 2005 through 2014. *Ophthalmology,* 122**,** 2432-42.

PARK, M., LI, Q., SHCHEYNIKOV, N., ZENG, W. & MUALLEM, S. 2004. NaBC1 is a ubiquitous electrogenic Na+ coupled borate transporter essential for cellular boron homeostasis and cell growth and proliferation. *Molecular Cell,* 16**,** 331-341.

PATEL, S. P. & PARKER, M. D. 2015. SLC4A11 and the Pathophysiology of Congenital Hereditary Endothelial Dystrophy. *Biomed Research International,* 2015:475392.**,** 10.1155/2015/475392. Epub 2015 Sep 16.

PEDERSEN, O. O., RUSHOOD, A. & OLSEN, E. G. 1989. Anterior mesenchymal dysgenesis of the eye. Congenital hereditary endothelial dystrophy and congenital glaucoma. *Acta Ophthalmologica* 67**,** 470-6.

PITTS, J. F. & JAY, J. L. 1990. The association of Fuchs's corneal endothelial dystrophy with axial hypermetropia, shallow anterior chamber, and angle closure glaucoma. *British Journal of Ophthalmology,* 74**,** 601-4.

PLAGNOL, V., CURTIS, J., EPSTEIN, M., MOK, K. Y., STEBBINGS, E., GRIGORIADOU, S., WOOD, N. W., HAMBLETON, S., BURNS, S. O., THRASHER, A. J., KUMARARATNE, D., DOFFINGER, R. & NEJENTSEV, S. 2012. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics,* 28**,** 2747-54.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I., DALY, M. J. & SHAM, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics,* 81**,** 559-75.

RAHI, J. S. & CABLE, N. 2003. Severe visual impairment and blindness in children in the UK. *Lancet,* 362**,** 1359-1365.

RAMAMURTHY, B., SACHDEVA, V., MANDAL, A. K., VEMUGANTI, G. K., GARG, P. & SANGWAN, V. S. 2007. Coexistent congenital hereditary endothelial dystrophy and congenital glaucoma. *Cornea,* 26**,** 647-649.

RAMPRASAD, V. L., EBENEZER, N. D., AUNG, T., RAJAGOPAL, R., YONG, V. H., TUFT, S. J., VISWANATHAN, D., EL‐ASHRY, M. F., LISKOVA, P. & TAN, D. T. 2007. Novel SLC4A11 mutations in patients with recessive congenital hereditary endothelial dystrophy (CHED2). *Human Mutation,* 28**,** 522-523.

RESNIKOFF, S., PASCOLINI, D., ETYA'ALE, D., KOCUR, I., PARARAJASEGARAM, R., POKHAREL, G. P. & MARIOTTI, S. P. 2004. Global data on visual impairment in the year 2002. *Bulletin of the World Health Organisation,* 82**,** 844-51.

RIAZUDDIN, S., ZAGHLOUL, N. A., AL-SAIF, A., DAVEY, L., DIPLAS, B. H., MEADOWS, D. N., EGHRARI, A. O., MINEAR, M. A., LI, Y. J. & KLINTWORTH, G. K. 2010a. Missense mutations in TCF8 cause late-onset Fuchs corneal dystrophy and interact with FCD4 on chromosome 9p. *American Journal of Human Genetics,* 86**,** 45-53.

RIAZUDDIN, S. A., EGHRARI, A. O., AL-SAIF, A., DAVEY, L., MEADOWS, D. N., KATSANIS, N. & GOTTSCH, J. D. 2009. Linkage of a mild late-onset phenotype of Fuchs corneal dystrophy to a novel locus at 5q33.1-q35.2. *Investigative Ophthalmology and Visual Science,* 50**,** 5667-71.

RIAZUDDIN, S. A., MCGLUMPHY, E. J., YEO, W. S., WANG, J., KATSANIS, N. & GOTTSCH, J. D. 2011. Replication of the TCF4 intronic variant in late-onset Fuchs corneal dystrophy and evidence of independence from the FCD2 locus. *Investigative Ophthalmology and Visual Science,* 52**,** 2825-9.

RIAZUDDIN, S. A., PARKER, D. S., MCGLUMPHY, E. J., OH, E. C., ILIFF, B. W., SCHMEDT, T., JURKUNAS, U., SCHLEIF, R., KATSANIS, N. & GOTTSCH, J. D. 2012. Mutations in LOXHD1, a recessive-deafness locus, cause dominant late-onset Fuchs corneal dystrophy. *American Journal of Human Genetics,* 90**,** 533-539.

RIAZUDDIN, S. A., VASANTH, S., KATSANIS, N. & GOTTSCH, J. D. 2013. Mutations in AGBL1 cause dominant late-onset Fuchs corneal dystrophy and alter protein-protein interaction with TCF4. *American Journal of Human Genetics,* 93**,** 758-64.

RIAZUDDIN, S. A., VITHANA, E. N., SEET, L. F., LIU, Y., AL-SAIF, A., KOH, L. W., HENG, Y. M., AUNG, T., MEADOWS, D. N., EGHRARI, A. O., GOTTSCH, J. D. & KATSANIS, N. 2010b. Missense mutations in the sodium borate cotransporter SLC4A11 cause late-onset Fuchs corneal dystrophy. *Human Mutation,* 31**,** 1261-8.

RIAZUDDIN, S. A., ZAGHLOUL, N. A., AL-SAIF, A., DAVEY, L., DIPLAS, B. H., MEADOWS, D. N., EGHRARI, A. O., MINEAR, M. A., LI, Y. J., KLINTWORTH, G. K., AFSHARI, N., GREGORY, S. G., GOTTSCH, J. D. & KATSANIS, N. 2010c. Missense mutations in TCF8 cause late-onset Fuchs corneal dystrophy and interact with FCD4 on chromosome 9p. *American Journal of Human Genetics,* 86**,** 45-53.

RICE, G. D., WRIGHT, K. & SILVERSTEIN, S. M. 2014. A retrospective study of the association between Fuchs' endothelial dystrophy and glaucoma. *Clinical Ophthalmology,* 8:2155-9.**,**

RICHARD, G. F. 2015. Shortening trinucleotide repeats using highly specific endonucleases: a possible approach to gene therapy? *Trends in Genetics,* 31**,** 177-86.

ROGOWSKI, K., VAN DIJK, J., MAGIERA, M. M., BOSC, C., DELOULME, J. C., BOSSON, A., PERIS, L., GOLD, N. D., LACROIX, B., BOSCH GRAU, M., BEC, N., LARROQUE, C., DESAGHER, S., HOLZER, M., ANDRIEUX, A., MOUTIN, M. J. & JANKE, C. 2010. A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell,* 143**,** 564-78.

ROSENBERG, M. E., TERVO, T. M., GALLAR, J., ACOSTA, M. C., MULLER, L. J., MOILANEN, J. A., TARKKANEN, A. H. & VESALUOMA, M. H. 2001. Corneal morphology and sensitivity in lattice dystrophy type II (familial amyloidosis, Finnish type). *Investigative Ophthalmology and Visual Science,* 42**,** 634-41.

ROSENBLUM, P., STARK, W., MAUMENEE, I., HIRST, L. & MAUMENEE, A. 1980. Hereditary Fuchs' Dystrophy. *American Journal of Ophthalmology,* 90**,** 455.

ROTH, T. N., HANEBUTH, D. & PROBST, R. 2011. Prevalence of age-related hearing loss in Europe: a review. *European Archives of Oto-Rhino-Laryngology,* 268**,** 1101-7.

SANGER, F. & COULSON, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology,* 94**,** 441-8.

SCHMID, E., LISCH, W., PHILIPP, W., LECHNER, S., GˆTTINGER, W., SCHLˆTZER-SCHREHARDT, U., M¸LLER, T., UTERMANN, G. & JANECKE, A. R. 2006. A new, X-linked endothelial corneal dystrophy. *American Journal of Ophthamology,* 141**,** 478-478.

SCHORDERET, D. 2015. Corneal Dystrophies: Overview and Summary. *Progress in Molecular Biology Translational Science,* 134:73-8.

SCHWARZ, J. M., RODELSPERGER, C., SCHUELKE, M. & SEELOW, D. 2010a. MutationTaster evaluates disease-causing potential of sequence alterations. *Nature methods,* 7**,** 575-6.

SCHWARZ, J. M., RODELSPERGER, C., SCHUELKE, M. & SEELOW, D. 2010b. MutationTaster evaluates disease-causing potential of sequence alterations. *Nature methods,* 7**,** 575-6.

SEPP, M., KANNIKE, K., EESMAA, A., URB, M. & TIMMUSK, T. 2011. Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS One,* 6**,** e22138.

SIDDIQUI, S., ZENTENO, J. C., RICE, A., CHACON-CAMACHO, O., NAYLOR, S. G., RIVERA-DE LA PARRA, D., SPOKES, D. M., JAMES, N., TOOMES, C., INGLEHEARN, C. F. & ALI, M. 2014. Congenital Hereditary Endothelial Dystrophy Caused by SLC4A11 Mutations Progresses to Harboyan Syndrome. *Cornea*, 33, 247-51.

SMALL, K. W., DELUCA, A. P., WHITMORE, S. S., ROSENBERG, T., SILVA-GARCIA, R., UDAR, N., PUECH, B., GARCIA, C. A., RICE, T. A., FISHMAN, G. A., HEON, E., FOLK, J. C., STREB, L. M., HAAS, C. M., WILEY, L. A., SCHEETZ, T. E., FINGERT, J. H., MULLINS, R. F., TUCKER, B. A. & STONE, E. M. 2016. North Carolina Macular Dystrophy Is Caused by Dysregulation of the Retinal Transcription Factor PRDM13. *Ophthalmology,* 123**,** 9-18.

SNELL, R. S. & LEMP, M. A. 1997. *Clinical anatomy of the eye*, Wiley-Blackwell.

SOUMITTRA, N., LOGANATHAN, S. K., MADHAVAN, D., RAMPRASAD, V. L., AROKIASAMY, T., SUMATHI, S., KARTHIYAYINI, T., RACHAPALLI, S. R., KUMARAMANICKAVEL, G., CASEY, J. R. & RAJAGOPAL, R. 2014. Biosynthetic and functional defects in newly identified SLC4A11 mutants and absence of COL8A2 mutations in Fuchs endothelial corneal dystrophy. *Journal of Human Genetics,* 59**,** 444-53.

SOUTHERN, E. 2006. Southern blotting. *Nature Protocols,* 1**,** 518-25.

STAMLER, J. F., ROOS, B. R., WAGONER, M. D., GOINS, K. M., KITZMANN, A. S., RILEY, J. B., STONE, E. M. & FINGERT, J. H. 2013. Confirmation of the association between the TCF4 risk allele and Fuchs endothelial corneal dystrophy in patients from the Midwestern United States. *Ophthalmic Genetics,* 34**,** 32-4.

STEHOUWER, M., BIJLSMA, W. R. & VAN DER LELIJ, A. 2011. Hearing disability in patients with Fuchs' endothelial corneal dystrophy: unrecognized co-pathology? *Clinical Ophthalmology,* 5:1297-301.

STRACHAN, T. & READ, A. 2011. *Human Molecular Genetics,* Oxford, Taylor and Francis Group.

STRITTMATTER, W. J., SAUNDERS, A. M., SCHMECHEL, D., PERICAK-VANCE, M., ENGHILD, J., SALVESEN, G. S. & ROSES, A. D. 1993. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 1977-81.

SULTANA, A., GARG, P., RAMAMURTHY, B., VEMUGANTI, G. K. & KANNABIRAN, C. 2007. Mutational spectrum of the SLC4A11 gene in autosomal recessive congenital hereditary endothelial dystrophy. *Molecular Vision,* 13**,** 1327-1332.

SUNDIN, O. H., BROMAN, K. W., CHANG, H. H., VITO, E. C. L., STARK, W. J. & GOTTSCH, J. D. 2006a. A common locus for late-onset Fuchs corneal dystrophy maps to 18q21. 2-q21. 32. *Investigative Ophthalmology and Visual Science,* 47**,** 3919-3926.

SUNDIN, O. H., JUN, A. S., BROMAN, K. W., LIU, S. H., SHEEHAN, S. E., VITO, E. C. L., STARK, W. J. & GOTTSCH, J. D. 2006b. Linkage of late-onset Fuchs corneal dystrophy to a novel locus at 13pTel-13q12. 13. *Investigative Ophthalmology and Visual Science,* 47**,** 140-145.

SWEATT, J. D. 2013. Pitt-Hopkins Syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. *Experimental and Molecular Medicine,* 45:e21.**,** 10.1038/emm.2013.32.

THALAMUTHU, A., KHOR, C. C., VENKATARAMAN, D., KOH, L. W., TAN, D. T., AUNG, T., MEHTA, J. S. & VITHANA, E. N. 2011. Association of TCF4 gene polymorphisms with Fuchs' corneal dystrophy in the Chinese. *Investigative Ophthalmology and Visual Science,* 52**,** 5573-8.

TOMA, N. M. G., EBENEZER, N. D., INGLEHEARN, C. F., PLANT, C., FICKER, L. A. & BHATTACHARYA, S. S. 1995. Linkage of congenital hereditary endothelial dystrophy to chromosome 20. *Human Molecular Genetics,* 4**,** 2395.

TOOMES, C., BOTTOMLEY, H. M., JACKSON, R. M., TOWNS, K. V., SCOTT, S., MACKEY, D. A., CRAIG, J. E., JIANG, L., YANG, Z., TREMBATH, R., WOODRUFF, G., GREGORY-EVANS, C. Y., GREGORY-EVANS, K., PARKER, M. J., BLACK, G. C., DOWNEY, L. M., ZHANG, K. & INGLEHEARN, C. F. 2004a. Mutations in LRP5 or FZD4 underlie the common familial exudative vitreoretinopathy locus on chromosome 11q. *American Journal of Human Genetics.,* 74**,** 721-30.

TORRICELLI, A. A., SINGH, V., SANTHIAGO, M. R. & WILSON, S. E. 2013. The corneal epithelial basement membrane: structure, function, and disease. *Investigative Ophthalmology and Visual Science,* 54**,** 6390-400.

TOURTAS, T., LAASER, K., BACHMANN, B. O., CURSIEFEN, C. & KRUSE, F. E. 2012. Descemet membrane endothelial keratoplasty versus descemet stripping automated endothelial keratoplasty. *American Journal of Ophthalmology,* 153**,** 1082-90.

TSUJIKAWA, M., KURAHASHI, H., TANAKA, T., OKADA, M., YAMAMOTO, S., MAEDA, N., WATANABE, H., INOUE, Y., KIRIDOSHI, A., MATSUMOTO, K., OHASHI, Y., KINOSHITA, S., SHIMOMURA, Y., NAKAMURA, Y. & TANO, Y. 1998. Homozygosity mapping of a gene responsible for gelatinous drop-like corneal dystrophy to chromosome 1p. *American Journal of Human Genetics.,* 63**,** 1073-7.

VANDEWALLE, C., VAN ROY, F. & BERX, G. 2009. The role of the ZEB family of transcription factors in development and disease. *Cell and Molecular Life Sciences,* 66**,** 773-87.

VILAS, G. L., LOGANATHAN, S. K., QUON, A., SUNDARESAN, P., VITHANA, E. N. & CASEY, J. 2012a. Oligomerization of SLC4A11 protein and the severity of FECD and CHED2 corneal dystrophies caused by SLC4A11 mutations. *Human Mutation,* 33**,** 419-428.

VILAS, G. L., LOGANATHAN, S. K., QUON, A., SUNDARESAN, P., VITHANA, E. N. & CASEY, J. 2012b. Oligomerization of SLC4A11 protein and the severity of FECD and CHED2 corneal dystrophies caused by SLC4A11 mutations. *Human Mutation,* 33**,** 419-428.

VILAS, G. L., MORGAN, P. E., LOGANATHAN, S. K., QUON, A. & CASEY, J. R. 2011. A biochemical framework for SLC4A11, the plasma membrane protein defective in corneal dystrophies. *Biochemistry (Mosc).* 50**,** 2157-69.

VITHANA, E. N., MORGAN, P., SUNDARESAN, P., EBENEZER, N. D., TAN, D. T. H., MOHAMED, M. D., ANAND, S., KHINE, K. O., VENKATARAMAN, D. & YONG, V. H. K. 2006. Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2). *Nature Genetics,* 38**,** 755-757.

VITHANA, E. N., MORGAN, P. E., RAMPRASAD, V., TAN, D. T. H., YONG, V. H. K., VENKATARAMAN, D., VENKATRAMAN, A., YAM, G. H. F., NAGASAMY, S. & LAW, R. W. K. 2008. SLC4A11 mutations in Fuchs endothelial corneal dystrophy. *Human Molecular Genetics,* 17**,** 656-666.

WANG, K., LI, M. & HAKONARSON, H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research,* 38**,** e164. doi: 10.1093/nar/gkq603.

WANG, K. J., JHANJI, V., CHEN, J., LAW, R. W., LEUNG, A. T., ZHANG, M., WANG, N., PANG, C. P. & YAM, G. H. 2013. Association of Transcription Factor 4 (TCF4) and Protein Tyrosine Phosphatase, Receptor Type G (PTPRG) with Corneal Dystrophies in Southern Chinese. *Ophthalmic Genetics, 35, 138-41.*

WARING, G. O., BOURNE, W. M., EDELHAUSER, H. F. & KENYON, K. R. 1982. The corneal endothelium. Normal and pathologic structure and function. *Ophthalmology.,* 89**,** 531-90.

WARING, G. O., RODRIGUES, M. M. & LAIBSON, P. R. 1978a. Corneal dystrophies. I. Dystrophies of the epithelium, Bowman's layer and stroma. *Survey in Ophthalmology,* 23**,** 71-122.

WARING, G. O., RODRIGUES, M. M. & LAIBSON, P. R. 1978b. Corneal dystrophies. II. Endothelial dystrophies. *Survey in Ophthalmology,* 23**,** 147-68.

WARNER, J. P., BARRON, L. H., GOUDIE, D., KELLY, K., DOW, D., FITZPATRICK, D. R. & BROCK, D. J. 1996. A general method for the detection of large CAG repeat expansions by fluorescent PCR. *Journal of Medical Genetics.,* 33**,** 1022-6.

WEISS, J. S., MOLLER, H., LISCH, W., KINOSHITA, S., ALDAVE, A. J., BELIN, M. W., KIVEL‰, T., BUSIN, M., MUNIER, F. L. & SEITZ, B. 2008. The IC3D classification of the corneal dystrophies. *Cornea,* 27**,** S1.

WEISS, J. S., MOLLER, H. U., ALDAVE, A. J., SEITZ, B., BREDRUP, C., KIVELA, T., MUNIER, F. L., RAPUANO, C. J., NISCHAL, K. K., KIM, E. K., SUTPHIN, J., BUSIN, M., LABBE, A., KENYON, K. R., KINOSHITA, S. & LISCH, W. 2015. IC3D classification of corneal dystrophies--edition 2. *Cornea,* 34**,** 117-59.

WIEBEN, E. D., ALEFF, R. A., ECKLOFF, B. W., ATKINSON, E. J., BAHETI, S., MIDDHA, S., BROWN, W. L., PATEL, S. V., KOCHER, J. P. & BARATZ, K. H. 2014. Comprehensive assessment of genetic variants within TCF4 in Fuchs' endothelial corneal dystrophy. *Investigative Ophthalmology and Visual Science,* 55**,** 6101-7.

WIEBEN, E. D., ALEFF, R. A., TOSAKULWONG, N., BUTZ, M. L., HIGHSMITH, W. E., EDWARDS, A. O. & BARATZ, K. H. 2012. A common trinucleotide repeat expansion within the transcription factor 4 (TCF4, E2-2) gene predicts Fuchs corneal dystrophy. *PLoS One,* 7**,** e49083.

WILD, E. J. & TABRIZI, S. J. 2007. Huntington's disease phenocopy syndromes. *Currrent Opinion in Neurology,* 20**,** 681-7.

WILLIAMS, K. A. & KLEBE, S. 2012. Gene therapy for corneal dystrophies and disease, where are we? *Current Opinion in Ophthalmology,* 23**,** 276-9.

WITTKE-THOMPSON, J. K., PLUZHNIKOV, A. & COX, N. J. 2005. Rational inferences about departures from Hardy-Weinberg equilibrium. *American Journal of Human Genetics,* 76**,** 967-86.

WOJCIK, K. A., SYNOWIEC, E., POLAKOWSKI, P., BLASIAK, J., SZAFLIK, J. & SZAFLIK, J. P. 2015. Variation in DNA Base Excision Repair Genes in Fuchs Endothelial Corneal Dystrophy. *Medical Science Monitor,* 21:2809-27.

WORETA, F. A., DAVIS, G. W. & BOWER, K. S. 2015. LASIK and surface ablation in corneal dystrophies. *Survey Ophthalmology,* 60**,** 115-22.

WRIGHT, A. F. & DHILLON, B. 2010. Major progress in Fuchs's corneal dystrophy. *New England Journal of Medicine,* 363**,** 1072-5.

XING, C., GONG, X., HUSSAIN, I., KHOR, C. C., TAN, D. T., AUNG, T., MEHTA, J. S., VITHANA, E. N. & MOOTHA, V. V. 2014. Transethnic replication of association of CTG18.1 repeat expansion of TCF4 gene with Fuchs' corneal dystrophy in Chinese implies common causal variant. *Investigative Ophthalmology and Visual Science,* 55**,** 7073-8.

YANOFF, M. & DUKER, J. S., J. 2009. *Ophthalmology*. 3rd ed. Edinburgh, UK: Mosby Elsevier.

YEE, R. W., SULLIVAN, L. S., LAI, H. T., STOCK, E. L., LU, Y., KHAN, M. N., BLANTON, S. H. & DAIGER, S. P. 1997. Linkage mapping of Thiel-Behnke corneal dystrophy (CDB2) to chromosome 10q23-q24. *Genomics,* 46**,** 152-4.

YELLORE, V. S., RAYNER, S. A., EMMERT-BUCK, L., TABIN, G. C., RABER, I., HANNUSH, S. B., STULTING, R. D., SAMPAT, K., MOMI, R., PRINCIPE, A. H. & ALDAVE, A. J. 2005. No pathogenic mutations identified in the COL8A2 gene or four positional candidate genes in patients with posterior polymorphous corneal dystrophy. *Investigative Ophthalmology and Visual Science,* 46**,** 1599-603.

YORSTON, D. 1999. The global initiative vison 2020: the right to sight childhood blindness. *Community Eye Health,* 12**,** 44-5.

YU, J. H., HARRELL, T. M., JAMAL, S. M., TABOR, H. K. & BAMSHAD, M. J. 2014. Attitudes of genetics professionals toward the return of incidental results from exome and whole-genome sequencing. *American Journal of Human Genetics,* 95**,** 77-84.

YU, X., GUDA, K., WILLIS, J., VEIGL, M., WANG, Z., MARKOWITZ, S., ADAMS, M. D. & SUN, S. 2012. How do alignment programs perform on sequencing data with varying qualities and from repetitive regions? *BioData Mining,* 5**,** 6.

YUEN, H. K. L., RASSIER, C. E., JARDELEZA, M. S. R., GREEN, W. R., DE LA CRUZ, Z., STARK, W. J. & GOTTSCH, J. D. 2005. A morphologic study of Fuchs dystrophy and bullous keratopathy. *Cornea,* 24**,** 319.

ZAVALA, J., LOPEZ JAIME, G. R., RODRIGUEZ BARRIENTOS, C. A. & VALDEZ-GARCIA, J. 2013. Corneal endothelium: developmental strategies for regeneration. *Eye,* 27**,** 579-88.

ZHANG, X., IGO, R. P., JR., FONDRAN, J., MOOTHA, V. V., OLIVA, M., HAMMERSMITH, K., SUGAR, A., LASS, J. H. & IYENGAR, S. K. 2013. Association of smoking and other risk factors with Fuchs' endothelial corneal dystrophy severity and corneal thickness. *Investigative Ophthalmology and Visual Science,* 54**,** 5829-35.

ZOEGA, G. M., FUJISAWA, A., SASAKI, H., KUBOTA, A., SASAKI, K., KITAGAWA, K. & JONASSON, F. 2006. Prevalence and risk factors for cornea guttata in the Reykjavik Eye Study. *Ophthalmology,* 113**,** 565-9.

ZUBERBUHLER, B., TUFT, S., GARTRY, D. & SPOKES, D. 2013. *Corneal Surgery: Essential Techniques,* New York: Springer.

ZWEIER, C., PEIPPO, M. M., HOYER, J., SOUSA, S., BOTTANI, A., CLAYTON-SMITH, J., REARDON, W., SARAIVA, J., CABRAL, A., GOHRING, I., DEVRIENDT, K., DE RAVEL, T., BIJLSMA, E. K., HENNEKAM, R. C., ORRICO, A., COHEN, M., DREWEKE, A., REIS, A., NURNBERG, P. & RAUCH, A. 2007. Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *American Journal Human Genetics,* 80**,** 994-1001.

Appendix I

A sample of the BOSU Corneal Dystrophy Questionnaire is shown below

Study number 59

Corneal Dystrophy Questionnaire

In association with the British Ophthalmological Surveillance Unit (BOSU)

Patient Details

Hospital No. _____

Month and Year of Birth (mm/yyyy) _____

Full postcode _____

Gender ❑Male ❑ Female

Ethnicity

White	Asian or Asian British	Black or Black British	Chinese	Mixed Race	Other ethnic group
❑ British ❑ Irish ❑ Other (please specify)	❑ Indian ❑ Pakistani ❑ Bangladeshi ❑ Other (please specify)	❑ Caribbean ❑ African ❑ Other (please specify)	❑ Chinese ❑ Other (please specify)	❑ White & Black Caribbean ❑ White & Black African ❑ White & Asian ❑ Other (please specify)	❑ Other (please specify)

Patient Profile

Source of Referral

❑ Patient ❑ Parent/Relative ❑ GP ❑ Optometrist ❑ unknown

❑ Ophthalmologist (please specify unit_____)

❑ Other (please specify_____)

Family History of Corneal Dystrophy ❑ None ❑Siblings ❑ Parents

❑ Other (please specify……..) ❑ Not Known

Are the parents related ❑ Yes ❑ No ❑ Not Known

Study number 59

Clinical Details

Date of initial presentation _____

Diagnosis

❑ Epithelial basement membrane dystrophy
❑ Thiel-Bhenke corneal dystrophy
❑ Meesman corneal dystrophy
❑ Reis-Bucklers corneal dystrophy
❑ Lattice corneal dystrophy
❑ Fleck corneal dystrophy
❑ Granular corneal dystrophy
❑ Macular corneal dystrophy
❑ Fuchs endothelial dystrophy
❑ Congenital hereditary endothelial dystrophy
❑ Posterior polymorphous corneal dystrophy
❑ Currently unknown/yet to be determined
❑ Other (please specify_____)

❑ Isolated corneal dystrophy
❑ Other associated features (please specify_____)

Symptoms

❑ None ❑ Reduced photopic/scotopic VA ❑ Glare ❑ Surface irritation

Best Corrected Visual Acuity Right Eye ..... Left eye .....

Method (please specify_____)

Symmetrical disease ❑ yes ❑ no (if yes only fill below for one eye)

	Right Eye	Left Eye
Epithelial Signs		
Fluorescein stain	❑ Yes ❑ No	❑ Yes ❑ No
Opacity		
level	❑ epithelial ❑ sub-epithelial ❑ ant half of stroma ❑ post half stroma ❑ endothelial	
position	❑ central pupillary zone ❑ peripheral	❑ central pupillary zone ❑ peripheral
oedema	❑ focal ❑ diffuse ❑ absent	❑ focal ❑ diffuse ❑ absent
areas of normal cornea	❑ yes ❑ no	❑ yes ❑ no

If Schnyders are crystals present? ❑ yes ❑ no

Study number 59

Endothelial Signs

guttate changes ❑ Yes ❑No ❑ Yes ❑No

View of anterior segment

❑ full iris details ❑ hazy iris view ❑ no view ❑ full iris details ❑ hazy iris view ❑ no view

Conjunctiva

❑ normal ❑ injected ❑ normal ❑ injected

Corneal vascularisation

❑ yes ❑ no ❑ yes ❑ no
❑ superficial ❑ deep ❑ superficial ❑ deep

Corneal sensation

❑ not checked ❑ normal ❑ abnormal

Investigations undertaken for diagnosis (please specify)

❑ slit lamp exam alone

❑ imaging necessary _____

❑ biopsy/histology _____

❑ tests or referral for known associated morbidity _____

Genetics

testing performed ❑yes ❑no result if known_____

Intended initial Treatment

❑ Observation
❑ Lubricants
❑ Chemical surface ablation
❑ Laser surface ablation
❑ Lamellar surgery – anterior/ deep stromal
❑ Endothelial
❑ Penetrating keratoplasty
❑ Other (.....)

Are they eligible for CVI?

❑no ❑sight impaired ❑severely sight impaired

Intended Follow up

❑ Discharged ❑ appointment within 6months
❑ Next appointment after 6months
❑ Referral to other ophthalmologist

Study number 59

Other comments

Name of person completing questionnaire _____

Contact Telephone Number _____

Many thanks for taking time to complete this questionnaire, your help is genuinely appreciated. Please return in the paid envelope to:

Kamron Khan, Level 8 Ophthalmology and Neuroscience, Wellcome Trust Brenner Building, St.James' Hospital, Beckett Street, Leeds LS7 9TF

Appendix II

Corneal tissue processing

Donor tissue was tested for the presence of Hepatitis B Surface Antigen, Hepatitis C Antibody, Human Immunodeficiency virus (HIV) 1 and 2, Antibody, Human T-lymphotropic Virus (HTLV) and 2, Syphilis serology, Hepatitis B core Antibody (HBcAb). Following enucleation the donor eye was cleaned in 3% Povidone-iodine and 0.3% Sodium Thiosulphate prior to corneal excision. The corneascleral disc was then stored at 34°C in Eagles Essential Medium containing 2% FBS and the following antibiotics:

Penicillin 100 units/ml
Streptomycin 0.1mg/ml
Amphotericin B 0.25 µl/ml

Manchester Eye Bank Standards

The corneal endothelium was examined by light microscopy a few days before use to ensure its suitability for transplantation in patients with corneal endothelial disease/deficiency. Organ-cultured corneas were delivered to hospitals in medium containing 5% dextran to reverse the stromal oedema that occurs during storage. Corneas with an inadequate endothelium could still be suitable for anterior lamellar grafts. The corneas may also be transferred to 70% ethanol and stored at room temperature for up to 12 months or used in glaucoma surgery. Sclera, which was also stored in 70% ethanol for up to 12 months, can also be used for glaucoma or other reconstructive surgery. Ocular surface stem cells may be isolated from the limbus and expanded in *ex vivo* culture fortreating limbal stem cell deficiency.

Appendix III

Primer Pair Sequences

Standard PCR Conditions and Mastermix were used unless otherwise stated

SLC4A11

Exon	Forward Sequence 5'-3'	Reverse Sequence 5'-3'	Product Size (bp)	Annealing Temperature (°C)/Programme
1	CCTAGCAGATGGGCTAAGCA	GAGCAAAGCCACAGGACTCT	374	59
2 -3	CGAGAGTGGGACAGTCCAG	CTCCCTGTTGAGCTGCTCCT	554	61
4 -5	TCCAGGAGCAGCTCAACAG	CAGCCCTCTTCTCCCAAGTT	686	59
6	CCAACCAACTTGGGAGAAGA	CCTTCAGAGGCCAGGACAT	391	57
7 - 8	AAAACCTGCTGCCAGTTCAT	AAAACCTGCTGCCAGTTCAT	589	55
9 -10	ACTGATGGTACGTGGCCTCT	CGTCCATGCGTAGAAGGAGT	567	59
11 -12	TCTACATCCAGGGTGCAGTG	ACTCAGCTTGAGCCAGTCCT	660	59
13-14	GAGCCCTTTCTCCCTGAGAT	GGTTGTAGCGGAACTTGCTC	623	59
15 -16	CGGGAAATCGAGAGTGAGTT	CGTCTCCTTCACGTTCACAA	673	57
17-18	CTGGCCACATGGGACATAG	CTAGGCAGGACCCCTCCTC	678	59
19	CAGGAGGGGCTCCAGTCTA	CTGTCCCTTGCATTCCACTT	692	57

COL8A2

Exon	Forward 5'-3'	Reverse 5'-3'	Product Size (bp)	Annealing Temperature (°C)/Programme
2	AGGTGAGAGGGGACTTCCTG	GATGCCAGTCTCATCGAAGG	400	57

LOXHD1

Exon	Forward Sequence 5'-3'	Reverse Sequence 5'-3'	Product Size (bp)	Annealing Temperature (°C)/Programme
1	CAGAGCTCAGGGAGGAGG	AATCAGTGAGGAAGGGCTTG	336	69
2	AAGAGTCCTTTGTGCTTGGG	CTTCTCCCCAGAGAAGCAGG	319	Touchdown
3	GGGATGATGGAGGAAGAAGC	TGGGAAGTAATTCATACCCAGAA	273	57
4	ATGATGGAGGAAGAAGCGG	TGGGAAGTAATTCATACCCAGA	377	57
5	GCTCTGTTATTGACAGTAATGGTGAC	GCTCAATAGAGGAGCCCAAA	355	*
6	TGGAAGAGCATCTTTCAGTGG	GAGTGGATGCAGATGGACCT	149	57
7	AAAGTAGACTCAAGTGATTGGAAGA	GTTTGATCACAGGCCTCCAG	314	Touchdown
8	GTGGAGGAGGAGGGCTTT	TTCAGAGAAGTAGCATTCAGGTT	455	57
9	GTCTCTAAATGTGGGGCTGG	GAGTGGACTGCCCTCATAGC	329	65
10	GTATACCCCGCCCTTCAGTC	TTGGTCCAAAACCTGGCTTA	371	57
11	GCTGAAGAAAGAGCCCAAAG	CAGAGGCAAATTTATGTGACAG	287	57
12	TGGCTTCTCTGCAATGAGAT	CTAAGGGGCCTGAAGATGC	346	57
13	TCAGCCCAGATGAGAACTAGA	CAGCTCAACTTTAACAGGGCA	364	57
14	TTGCCTAACCCATCAGCTCT	TTGCTTGCTGGTCATGGTAG	368	57
15	CTCTGGGCCTCCATAGTGAC	GCCACCTCTGTGAAACATGA	283	57
16	CTGGATTTTCAATTCCCAGC	TCTTCAAAATGTGTTTACCCTTATGT	470	57
17	CTGGAGACCTGGGTTGTGTT	TGTCAGCAAGACCTGCTTTG	401	57
18	GGAACAGGCTCAGGAAAGG	CGGGTGAGTATTGACTGAGGA	349	61
19	GCCTTGTGCCTGGCTCT	TGCCCACCTATTTGGCCT	673	69
20	AGGATCTGGCTGGATCTGAG	TGAACAAGTCACACTGCCAA	673	57
21	CTGCCCTGGTTCTTGGG	CCTCACCCTCCACCGTC	338	*
22	GACAGGGGAGAGTTTGGGAC	ACAGGGGAGGGAAGGAAGAT	368	67
23	GGGAGTGAATCAAGGAAGCC	CATAATTAGGATTCCCCTTGGA	304	57
24	TGGAAGTTCAGAAATTTGGTCA	CAGATGGCATTCAAATTTCC	339	57
25	CAGCCACCTAAGGAAGGAAG	GCCCACAGTCAATCCTGAAG	361	57
26	CTGGTTGTGCTGGTGAAGAA	CATCAGGATGAAGGGCATGT	372	57
27	GAAATGCAAAGGGACTCAGG	GCTGATCTAGCCAGTAGGTCC	322	57
28	CCATGATCCTTGTCCTTGGT	tccTGGGTGAAGAGGCTTAG	367	57
29	CCTAGGCCAGAAGCTTAGCA	GATGTCCCCAGGAACCAAG	351	57
30-31	AGGTCTGTTCAGTGCAGCAA	GATGGTGGGGCTCAAGAAT	658	69
32	GCCCTCAGAGGTCACTTCC	CAGGTAGGCTGTTCTTCCCA	409	69
33	GCTGTGGAACTTGGACAGTG	GGCATGTGAGAATCAGCCT	329	57
34	ATCTTCTCCCTCACCCCAAA	TTTGTGCTTTAACAAGGTCCA	311	65
35	GATCTCCAGGGTTGGGATTC	TGGAAGGCCTTATGAAGAAAA	385	57
36	TGTCCACCTGTACCCCTGAC	ATCAGAGTCAATGTGCTGCC	389	57
37	GATAACTTGTCCAGGGCCAC	CTGTATCTGGCACCTGACCC	377	*
38	GATCAAATGAAGGACCGGAG	CTGGTTAGGCCCATTTGTGT	341	61
39	CATCCCTGTTCCCTGGC	GAGACCTCATCATACCCTGC	366	57
40	CACCTTGGGAAGGGATCATT	GCCAATGCTAGAGGCTTTGA	689	61
32B	GAGCACTTCCTTCCCAGTTG	GATGCCCCAGTGATGAGTCT	356	57
40B	GAGATGCCCCAATCTCACAT	TGCCAATGCGTTCTCTGTAA	568	57
5II*	GATGGTCATACATCCAATGGC	ATTGCAATCAACCCACACAC	622	61
5III*	TTTGGAACCAATGTGGACTG	AGGACAGGTCACTCCAAACC	**	**
21*	GTGGGGAGGGGTAGGTCTTA	CCCCAGTCTTCTTCCAGGAC	271	57
37*	ATGGGTTGTGGGGATGTAGA	CCTCCTGAGCCAATGATCTC	281	57

* Redesigned primer

**Primer Pair 5III was a nested primer used for DNA sequencing only

Hotshot Mastermix was used for Exons 1 - 4, 6, 7, 10, 11,15, 18, 19, 22, 23, 30-32, 34, 38 and 40, For all of these Hotshot Cycling Conditions were used except for 2 and 7 where Touchdown was utilised.

TCF4

SNP	Forward 5'-3'	Reverse 5'-3'	Product Size (bp)	Annealing Temperature (°C)/Programme
rs613872	GTTGGGAACACCCATTTGTC	ACCCCAGTAGGGTTGTGATG	275	65
rs192075715	TTCTCATTTATATGTGTCCAACCTG	CACCAGATATATTGGGGGAA	253	60

Hotshot was used for rs613872

FECDBRI Exome Sequencing Variants

Gene	Forward 5'-3'	Reverse 5'-3'	Product Size (bp)	Annealing Temperature (°C)/Programme
OSBPL1A	TCTGTGGGGTTCTTCCTAGC	CACCCTGCAACGGATTTATT	398	59
CRYZL1	GGAGAGAAATAGATCCAGGAGG*	CAAATGGATGGCTCATTGCT*	286	55
TC2N	TGAGAGCCTTCCAGATCCTC	CAAGTCATTGCCATTCTTATTTTC*	466	55
TTC40	GAACCACAACCCTTCCCAC	CCACACTTGTCTACTCACCTTTCA	250	59
URB1	TCACTTGCCTTGATTTTGACC*	TTCTCAATCTCTTAACACGTCACT*	307	63

Appendix IV

Standard PCR Mastermix was used for all PCRS using these primers

Chromosome 18 Linkage Markers

Marker	Physical Location (hg19)	Forward 5'-3'	Reverse 5'-3'	Size (bp)	Dye
D18S1152	54716520 - 54716915	GTTTGGAGACAGGGCG	TTATAGTTCAGGCTCTTGTGTATTT	222-274	HEX
D18S1144	55568578 - 55568922	CTGGATTAGCCAGGCCC	TGACTTGTGGACACATCACTC	159-181	TET
D18S1103	56930137 - 56930437	GAATCTCTTGAACCAGGGA	AACCAGTAGGCATTTGGAA	206-254	FAM
D18S64	57426031 - 57426378	ATACTGGTGGTGGTTATACAACAT	AAATCAGGAAATCGGCA	188-208	FAM

TCF4 STR and TP-PCR Primers (to measure CTG18.1 expansion)

Primer	Sequence
P1	AATCCAAACCGCCTTCCAAGT
P2	CAAAACTTCCGAAAGCCATTTCT
P3	TACGCATCCCAGTTTGAGACG
P4	TACGCATCCCACTTTGAGACGCAGCAGCAGCAGCAG

Primer Ratios:

STR P1:P2 1:1

TP-PCR P1:P3:P4 1:1:0.5

TCF4 CTG18.1 Genotyping	Annealing Temperature (°C)/Programme
STR	61
TP PCR	60

Appendix V

Rs613872 and CTG18.1 Genotype and Demographic Data

Cases and Controls

ID	rs613872	CTG18.1	Age at Venesection	Gender
P1	TT	SS	68	F
P2	TG	SX	60	F
P3	TT	SS	57	F
P4	TG	SX	60	M
P5	TT	SS	65	F
P6	TG	SS	48	M
P7	TT	SS	49	F
P8	TG	SX	60	F
P9	TG	SS	91	M
P10	GG	XX	89	F
P11	TT	SS	80	F
P12	GG	SX	78	F
P13	TG	SS	79	F
P14	TG	SX	76	M
P15	TT	SS	69	M
P16	GG	SX	81	M
P17	TG	SS	65	F
P18	TG	SX	70	F
P19	TT	SX	69	F
P20	GG	XX	74	M
P21	TG	SX	66	F
P22	TG	SX	76	F
P23	TT	SS	69	M
P24	TG	SX	67	F
P25	TT	SS	59	M
P26	TT	SX	56	F
P27	GG	XX	77	F
P28	TT	SS	72	F
P29	TG	SX	69	F
P30	TG	SX	66	F
P31	TG	SX	76	F
P32	TT	SS	65	M
P33	TG	SX	80	F
P34	TG	SX	72	F
P35	TG	SX	67	M

P36	TG	SX	86	F
P37	TT	SS	88	M
P38	TG	SX	80	F
P39	TG	SX	62	F
P40	TT	SS	79	F
P41	TT	SS	70	M
P42	TG	SX	85	F
P43	TG	SX	82	F
P44	TG	SX	64	M
P45	GG	SX	63	F
P46	TG	SX	68	F
P47	TG	SX	75	F
P48	TG	SS	73	F
P49	TG	SX	83	F
P50	TG	SX	70	M
P51	GG	XX	68	F
P52	TG	SX	77	F
P53	TG	SX	77	F
P54	TG	SX	88	F
P55	TG	SX	75	M
P56	GG	SX	75	F
P57	TT	SS	94	M
P58	GG	SX	66	M
P59	TG	SX	68	M
P60	TT	SS	48	F
P61	TT	SS	82	F
P62	TG	SX	67	F
P63	TG	SX	74	M
P64	TG	SX	81	M
P65	TG	SX	63	M
P66	TT	SS	85	F
P67	TG	SX	78	M
P70	TG	SS	64	F
P71	TT	SS	89	F
P72	GG	SX	63	F
P73	GG	XX	76	M
P74	TG	SX	76	F
P75	TG	SX	84	F
P76	TT	SS	69	F
P77	TG	SX	76	M
P78	TG	SX	82	F
P79	GG	SX	65	M
P80	TG	SX	65	F

P81	TG	SX	68	F
P82	GG	XX	64	M
P83	TG	SX	73	F
P84	TG	SX	68	F
P85	TG	SX	68	F
P86	TG	SX	69	M
P87	TG	SX	71	F
P88	TG	SX	69	F
P89	TT	SS	76	F
P90	TG	SS	85	F
P91	TG	SX	56	M
P92	TG	SX	73	F
P93	TG	SS	69	F
P94	TG	SX	74	F
P95	TG	SX	79	M
P96	GG	XX	69	F
P97	TG	SX	59	F
P98	GG	SX	68	M
P100	TG	SX	60	F
P101	TG	SX	73	F
P103	GG	XX	54	M
P104	TT	SS	52	F
P105	TG	SX	67	M
P106	TG	SX	80	F
P107	TG	SX	72	M
P108	TG	XX	66	M
P109	TG	SX	66	M
P110	GG	SS	71	F
P111	TG	SX	68	M
P112	TG	SX	67	F
P115	GG	XX	72	M
P116	GG	XX	73	M
P117	TG	SX	73	F
P118	TG	SX	62	M
P119	TT	SS	58	F
P120	TT	SS	63	F
P125	GG	XX	74	F
P126	GG	XX	68	F
P127	TG	SX	58	M

ID	rs613872	CTG18.1	Age at Venesection	Gender
C1	TT	SS	74	M
C2	TT	SS	80	F
C3	TT	SS	67	F
C4	TG	SS	79	M
C5	TT	SS	78	M
C6	TT	SS	85	M
C7	TT	SS	76	M
C8	TT	SS	80	F
C9	TT	SS	81	F
C10	TT	SS	80	F
C11	TT	SS	82	M
C12	TT	SS	82	F
C13	TG	SS	77	M
C14	TT	SS	78	F
C15	TG	SS	80	F
C16	TT	SS	67	M
C17	TT	SS	75	M
C18	TT	SS	67	F
C19	TT	SS	84	F
C20	TT	SS	77	M
C21	TT	SS	78	F
C22	TT	SS	77	M
C23	TG	SS	76	F
C24	TT	SS	62	F
C25	TT	SS	78	M
C26	TT	SS	56	M
C27	TT	SS	74	F
C28	TT	SS	81	M
C29	TT	SS	69	M
C30	TG	SX	88	F
C31	TG	SS	81	F
C32	TT	SS	70	F
C33	TT	SS	74	F
C34	TG	SX	83	M
C35	TT	SS	87	F
C36	TG	SX	73	F
C37	TG	SS	86	F
C38	TT	SS	83	M
C39	TT	SS	66	M
C40	TT	SS	68	M

C41	TT	SS	81	M
C42	TG	SS	93	M
C43	TG	SX	64	M
C44	TG	SS	88	M
C45	TT	SS	76	F
C46	TG	SX	66	M
C47	TT	SS	77	M
C48	TG	SS	88	F
C49	TT	SS	79	F
C50	TT	SS	80	M
C51	TG	SS	73	F
C52	TG	SS	83	M
C53	TG	SS	68	F
C54	TT	SS	71	M
C55	TT	SS	70	F
C56	TG	SS	81	F
C57	TT	SS	73	M
C58	TT	SS	87	M
C59	TT	SS	83	F
C60	TT	SS	73	F
C61	TT	SS	77	F
C62	TT	SS	77	M
C63	TT	SS	92	F
C64	TT	SS	87	M
C65	TT	SS	86	F
C66	TT	SS	69	F
C67	TT	SS	78	M
C68	TG	SS	73	F
C69	TT	SS	79	F
C70	TT	SS	73	M
C71	TG	SS	67	M
C72	TT	SS	71	F
C73	TT	SS	86	F
C74	TT	SS	90	M
C75	TT	SS	78	F
C76	TT	SS	72	F
C77	TG	SS	61	M
C78	TT	SS	82	M
C79	TT	SS	86	F
C80	TT	SS	75	F
C85	TT	SS	68	M
C86	TG	SS	63	F
C87	TT	SS	67	M