Molecular genetics and genotype-phenotype correlation of inherited corneal dystrophies

Cerys J. Evans

University College London

Doctor of Philosophy (PhD)

Supervisors: Professor Alison Hardcastle and Mr. Stephen Tuft

Declaration

I, Cerys J. Evans, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Publications

Evans, C.J., Liskova, P., Dudakova, L., Hrabcikova, P., Horinek, A., Jirsova, K., Filipec, M., Hardcastle, A., Davidson, A. and Tuft, S. (2015) 'Identification of six novel mutations in ZEB1 and description of the associated phenotypes in patients with posterior polymorphous corneal dystrophy 3.', *Annals of Human Genetics*, 79(1), pp. 1–9.

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. and 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*, 24(7), pp. 985–991.

Davidson, A., Liskova, P., **Evans, C.J**., Dudakova, L., Nosková, L., Pontikos, N., Hartmannová, H., Hodaňová, K., Stránecký, V., Kozmík, Z., Levis, H., Idigo, N., Sasai, N., Maher, G., Bellingham, J., Veli, N., Ebenezer, N., Cheetham, M., Daniels, J., Thaung, C., Jirsova, K., Plagnol, V., Filipec, M., Kmoch, S., Tuft, S. and Hardcastle, A. (2015) 'Autosomal-dominant corneal endothelial dystrophies CHED1 and PPCD1 are allelic disorders caused by non-coding mutations in the promoter of OVOL2.', *The American Journal of Human Genetics*, 98(1), pp. 75–89.

Evans, C. J., Davidson, A. E., Carnt, N., Rojas López, K. E., Veli, N., Thaung, C. M., Tuft, S. J. and Hardcastle, A. J. (2016) 'Genotype-phenotype correlation for TGFBI corneal dystrophies identifies p.(G623D) as a novel cause of epithelial basement membrane dystrophy.', *Investigative Ophthalmology & Visual Science*, 57(13), pp. 5407–5414.

Morantes, S., **Evans, C. J**., Valencia, A. V, Davidson, A. E., Hardcastle, A. J., Linares, A. R., Tuft, S. J. and Cuevas, M. (2016) 'Spectrum of clinical signs and genetic characterization of gelatinous drop-like corneal dystrophy in a Colombian family.', *Cornea*, 35(8), pp. 1141–11466.

Acknowledgements

I would like to thank my supervisors, Prof. Alison Hardcastle and Mr. Stephen Tuft, for providing me with the opportunity to undertake this project and for guiding me through the PhD process. I could not have completed this thesis without their unwavering support.

I would like to thank all my colleagues at the Institute of Ophthalmology for providing both technical and emotional support throughout the PhD, particularly, Alice Davidson, Sek-Shir Cheong, Naheed Kanuga, Wenwen Li, Nikolas Pontikos, Jessica Gardner and Michael Cheetham. I would also like to thank my collaborator Petra Liskova.

I would like to thank Fight for Sight who generously provided the funding for this project. Most importantly, I would like to thank the patients and relatives who donated their samples, without which this project could not have been completed. I hope that the results within this thesis have a positive impact and will contribute to the development of therapies in future.

Finally, I would like to thank my family and friends who encouraged me to keep going.

Abstract

Corneal dystrophies are a group of inherited, primarily monogenic, disorders that compromise the transparency of the cornea and cause visual impairment. A large cohort, consisting of 191 corneal dystrophy probands, was recruited to the study with the aim of investigating the genetic basis of phenotypically heterogeneous corneal dystrophy diagnoses. Following genetic investigation, genotype data was combined with clinical data to investigate genotype-phenotype correlations.

The most common cause of corneal dystrophy in this cohort was heterozygous dominant mutations in *TGFBI,* which were identified in 70 probands with epithelial-stromal dystrophies. The majority of *TGFBI* mutations displayed genotype-phenotype correlation, however a p.(Gly623Asp) mutation was unexpectedly associated with a broad phenotypic spectrum of disease, including epithelial basement membrane dystrophy (EBMD). There was evidence for additional locus heterogeneity for EBMD; two families were identified in which *TGFBI* coding mutations were excluded, however no compelling candidate gene(s) were identified following whole exome sequence (WES) analysis. Novel and previously reported mutations in *CHST6, UBIAD1* and *TACTSD2* were identified for other anterior corneal dystrophies; recessive macular corneal dystrophy, dominant Schnyder corneal dystrophy and recessive gelatinous drop-like dystrophy, respectively. Dominant mutations in *ZEB1* were identified in 8 patients with an endothelial dystrophy, posterior polymorphous corneal dystrophy (PPCD3), with novel heterozygous deletions encompassing the *ZEB1* gene confirming haploinsufficiency as the mechanism of pathogenicity.

Investigation of dominant endothelial dystrophy families using whole genome sequencing (WGS), including PPCD1 and congenital hereditary endothelial dystrophy (CHED1), revealed that mutations in the promoter of *OVOL2* are a novel cause of disease and therefore that PPCD1 and CHED1 are allelic disorders. *OVOL2* is a transcription factor, which is a direct repressor of *ZEB1*. *OVOL2* was not expressed in normal corneal endothelial cells, therefore it was hypothesised that aberrant expression of *OVOL2* in corneal endothelial cells causes transcriptional repression of *ZEB1* expression. A novel locus causing dominant PPCD (PPCD4) was identified in a large family from the Czech Republic and linkage analysis and WGS revealed a non-coding mutation in a novel gene as likely to be causative in this family.

Interestingly, a proportion of patients clinically diagnosed with a corneal dystrophy were found, on further genetic and clinical investigation, to have an inherited syndrome which was responsible for the presence of corneal opacities, the most common of which was Meretoja syndrome. In some cases, a phenocopy of disease was responsible, including paraproteinemic keratopathy. Genetic screening can therefore significantly improve clinical care for patients and their families.

In summary, this study provides insight into the genetic causes of corneal dystrophies including the identification of novel corneal dystrophy genes and mechanisms of disease, which is a pre-requisite for the development of targeted genetic therapies. Furthermore, it provides further understanding of genotype-phenotype correlation for specific corneal dystrophy genes and mutations that can result in improved diagnostic and prognostic accuracy for patients.

Contents

List of Figures

List of Tables

List of Abbreviations

- 1000G – 1000 Genomes Project
- AAV – adeno-associated virus
- aCGH – array comparative genomic hybridisation
- apoA-I – apoliproprotein A-I
- Cas – CRISPR associated
- CCDF – Central cloudy dystrophy of François
- CDK – climatic droplet keratopathy
- cDNA – complementary DNA
- CECM – corneal epithelial culture medium
- CHED – congenital hereditary endothelial dystrophy
- CK – cytokeratin
- CNV – copy number variation
- CRISPR - clustered regularly interspaced short palindromic repeat
- crRNA – CRISPR RNA
- CSCC – corneal stromal stem cell
- DALK – deep anterior lamellar keratoplasty
- ddNTP – dideoxynucleotide
- DMEK – Descemet's membrane endothelial keratoplasty
- DMEM – Dulbecco's Modified Eagle Medium
- dNTP – deoxynucleotide
- DoGV – The Database of Genomic Variants
- DSAEK – Descemet's stripping automated endothelial keratoplasty
- DSEK – Descemet's stripping endothelial keratoplasty
- EBMD – epithelial basement membrane dystrophy
- E-CAD – epithelial cadherin
- ECM – extracellular matrix
- EK – endothelial keratoplasty
- ELISA – enzyme-linked immunosorbent assay
- EM – electron microscopy
- EMI – N-terminal cysteine rich (domain)

- EmPCR – emulsion-PCR
- EMT – epithelial-to-mesenchymal transition
- ENCODE – Encyclopedia of DNA elements
- ERED – epithelial recurrent erosion dystrophy
- EVS – Exome Variant Server
- ExAC – Exome Aggregation Consortium
- FAS1 – fasciclin-1 (domain)
- FBS – fetal bovine serum
- FCD – fleck corneal dystrophy
- FECD – Fuchs endothelial corneal dystophy
- GAG – glycosaminoglycan
- GCD1 – granular corneal dystrophy type 1
- GCD2 – granular corneal dystrophy type 2
- GDLD – gelatinous drop-like dystrophy
- gDNA – genomic DNA
- GGPP – geranylgeranylpyrophosphate
- GoNL – Genomes of the Netherlands
- GWAS – genome-wide association study
- H&E – hematoxylin and eosin
- HCE-S – spontaneously immortalised human corneal epithelial cell line
- HDL – high density lipoproteins
- IC3D – The International Committee for Classification of Corneal Dystrophies
- ICE – iridocorneal endothelial (syndrome)
- IGV – Integrative Genomics Viewer
- iPSC – induced pluripotent stem cell
- LCD – lattice corneal dystrophy
- LDL – low density lipoprotein
- LECD – Lisch epithelial corneal dystrophy
- LESC – limbal epithelial stem cell
- LOD – logarithm of odds
- LSCD – limbal stem cell deficiency

- MAF – minor allele frequency
- MCD – macular corneal dystrophy
- MECD – Meesmann epithelial corneal dystrophy
- MEH – Moorfields Eye Hospital
- MET – mesenchymal-to-epithelial transition
- MGUS – monoclonal gammopathy of unknown significance
- MK4 – menaquinone-4
- MRI – magnetic resonance imaging
- N-CAD – neural cadherin
- NHEJ – non-homologous end joining
- OCT – optical coherence tomography
- OMIM – Online Mendelian Inheritance in Man
- P/S – penicillin/streptomycin
- PACD – posterior amorphous corneal dystrophy
- PAM – proto-spacer adjacent motif
- PAS – periodic acid-Schiff
- PCA – polymorphic corneal amyloidosis
- PCD – polymorphic corneal degeneration
- PCDC – pre-Descemet corneal dystrophy
- PCR – polymerase chain reaction
- PK – penetrating keratoplasty
- PPCD – posterior polymorphous corneal dystrophy
- PSCD – posterior stromal corneal dystrophy
- PTK – photorefractive keratectomy
- qPCR – quantitative real-time PCR
- RBCD – Reis-Bücklers corneal dystrophy
- REC – research ethics committee
- RISC – RNA-inducible silencing complex
- RNAi – RNA interference
- RNA-seq – RNA-sequencing
- RPKM – reads per kilobase million
- RT-PCR – reverse transcription PCR

- SCD – Schnyder corneal dystrophy
- siRNA – small interfering RNA
- SMCD – sub-epithelial mucinous corneal dystrophy
- SNP – small nucleotide polymorphism
- TAC – transient amplifying cell
- TAE – Tris acetate EDTA
- TBCD – Thiel-Behnke corneal dystrophy
- UTR – untranslated region
- UV – ultraviolet
- VEP – Variant Effect Predictor
- WES – whole exome sequencing
- WGS – whole genome sequencing
- WT – wild type
- XECD – X-linked endothelial corneal dystrophy

Chapter 1: Introduction

1.1 Anatomy of the eye

The outer layer of the eye consists of an opaque white sclera which protects the interior structures. Light enters the eye through a transparent window of tissue called the cornea and is focused by the lens onto the retina at the back of the eye. The ciliary muscle alters the shape of the lens via the zonule fibres to mediate the refraction and focus of light. The pupil regulates the amount of light entering the eye by altering its size in response to light stimulation; this is controlled by the muscles of the iris, a pigmented tissue located anterior to the lens. The space between the cornea and the lens and is the anterior chamber and is filled with aqueous humour, which is produced by the ciliary body and secreted into the posterior chamber. Aqueous humour drains via the trabecular meshwork, located at the junction between the peripheral iris and cornea. Between the lens and retina is a large fluid-filled region that contains the vitreous humour. The retina, located at the back of the eye, contains light-sensitive photoreceptor cells; these specialised neuronal cells transduce light signals into neuronal impulses which are then transmitted to the brain via the optic nerve. The neural retina is nourished by the retinal pigment epithelium and blood vessels which make up the choroid, located between the retina and the outer sclera (Purves *et al.*, 2001; Golb, Fernandez and Nelson, 2005). A diagram outlining the structure of the human eye is shown in Figure 1.1.

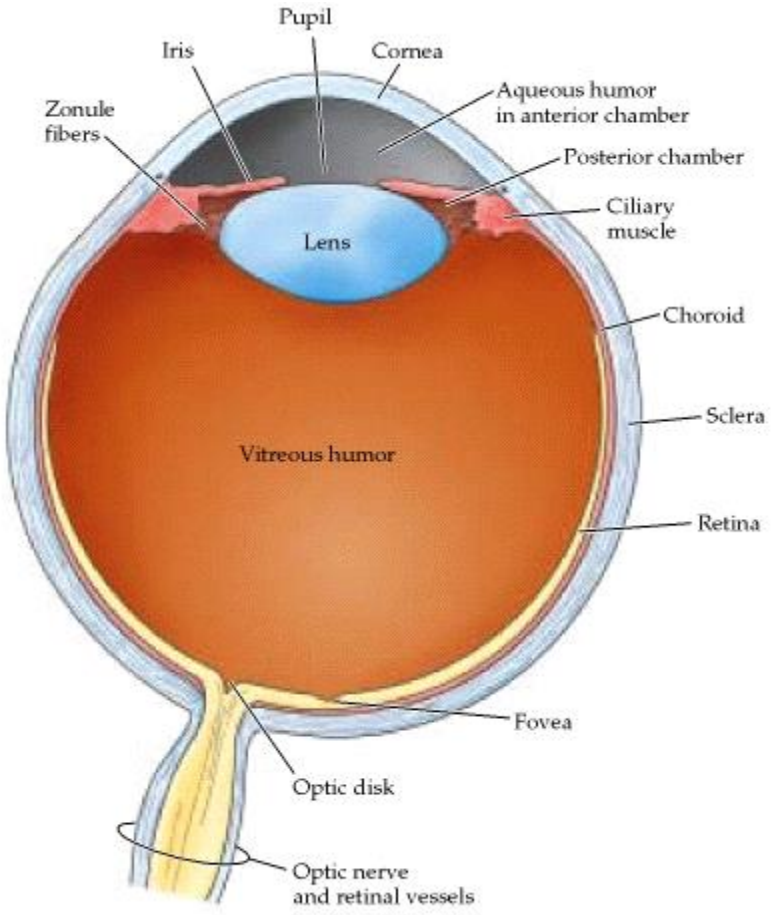


Figure 1.1 Horizontal cross section illustration of the human eye. The anterior segment contains the transparent cornea and lens, the ciliary muscle which controls lens shape, the anterior chamber containing aqueous humour, the posterior chamber, and the iris which regulates pupil size. The light-sensitive retina is found at the back of the eye, nourished by blood vessels of the choroid. The optic nerve transmits visual information to the brain. Image from (Purves *et al.*, 2001).

1.2 Structure and function of the cornea

The cornea is a transparent, avascular tissue located at the anterior surface of the eye, consisting of three cellular and two acellular layers, although a putative sixth acellular layer has been proposed (Dua *et al.*, 2013). The cornea is the major refractive surface of the eye, responsible for two-thirds of its refractive power, and also functions as a barrier to protect the internal structures of the anterior segment from trauma, infection and ultraviolet (UV) radiation (Meek, Dennis and Khan, 2003). The transparency of the cornea is essential for visual acuity.

1.2.1 Epithelium

The external layer of the cornea is a stratified non-keratinising squamous epithelium consisting of 5 – 7 cellular layers and is approximately 50 μm thick in the adult human eye (Reinstein *et al.*, 2008) (Figure 1.2). Superficial epithelial cells make up the outermost cellular layer, overlying 1-2 layers of suprabasal wing cells which are flattened in appearance. Both wing cells and superficial cells are connected by tight junctions which prevent large molecules and pathogens from entering the cornea (DelMonte and Kim, 2011). At the basal epithelial cell layer, the cells are columnar in shape and joined by gap junctions. The basal epithelial cells are anchored to the underlying epithelial basement membrane via hemidesmosomes which connect the intermediate filaments of the cellular cytoskeleton to the collagen type VII anchoring fibrils of the extracellular matrix (ECM), via a variety of structural proteins (Torricelli *et al.*, 2013).

1.2.2 Bowman's layer

Beneath the basement membrane of the epithelium is Bowman's layer, an acellular layer that is approximately 10 μm in thickness in the adult human eye (Figure 1.2), consisting of dense, randomly-orientated collagen fibrils, primarily collagen type I, III and V. The function of Bowman's layer is not known. It is only present in primates and may be absent or disrupted in a number of corneal diseases, or as a result of photorefractive keratectomy (PTK) (Wilson and Hong, 2000; Hayashi, Osawa and Tohyama, 2002).

1.2.3 Stroma

Below Bowman's layer is the stroma, which makes up ~ 90% of the thickness of the cornea and determines its shape (Figure 1.2). The bulk of the stroma consists of highly organised collagen fibrils, predominantly collagen type I and IV, which are essential for maintaining corneal transparency. Parallel arrays of collagen fibrils are packed into structures called lamellae and stacked orthogonally to reduce light scatter (Hassell and Birk, 2010). There are four major proteoglycans in the stroma; decorin, lumican, keratocan and mimecan. Decorin has glycosaminoglycan (GAG)

side chains of chondroitin/dermatan sulphate whereas lumican, keratocan and mimecan have keratan sulphate side chains. The proteoglycans bind to the collagen fibrils and are involved in maintaining regular collagen fibril spacing (Meek and Boote, 2004). Sparsely populating the stroma are keratocytes which are responsible for remodeling the highly organised ECM. In the healthy adult cornea, keratocytes are in the G0 phase of the cell cycle and therefore quiescent. Keratocytes are connected to one another via long projections called lamellipodia which form a network and allow communication via gap junctions (Hahnel *et al.*, 2000).

1.2.4 Descemet's membrane

Descemet's membrane is located at the posterior of the stroma (Figure 1.2). It is the basement membrane of the most posterior corneal cellular layer, the corneal endothelium. It is ~ 6 μm in thickness, although it increases in thickness throughout life, and consists primarily of collagen type VIII fibrils arranged in a hexagonal lattice arrangement (Sawada, Konomi and Hirosawa, 1990; Hayashi, Osawa and Tohyama, 2002).

1.2.5 Endothelium

The monolayer of corneal endothelial cells (CECs) is ~ 4 μm in thickness in a healthy adult human cornea and forms the border that separates the cornea from the aqueous humour of the anterior chamber (Figure 1.2). CECs normally have a hexagonal shape and form a regular honeycomb appearance. They are joined by tight junctions and form a barrier which prevents aqueous humour from the anterior chamber entering the stroma, but allows for the diffusion of nutrients into the cornea. Due to the 'leakiness' of the barrier there is some fluid entry, therefore CECs use an ionic pump mechanism to set up local osmotic gradients which draw fluid out of the stroma, back into the aqueous humour (Srinivas, 2012). Using this mechanism, the cornea maintains a state of relative dehydration called deturgescence, which is essential for corneal transparency.

1.2.6 Dua's layer

A putative corneal layer has been described at the posterior 6 – 13 μm of stroma, located at the interface of the stroma and Descemet's membrane, that has been designated as Dua's layer (Dua *et al.*, 2013). The layer is proposed to be an acellular membrane consisting of multiple thin lamellae that are retained to Descemet's membrane when the anterior corneal layers are removed surgically using the big bubble technique. However, the interpretation of this tissue as a separate physiological layer is disputed. It has been argued that the retained tissue occurs as a consequence of the physical properties of the stroma and the removal technique used. Additionally, stromal keratocytes have been shown to occur within the layer, therefore indicating that it is continuous with the corneal stroma and not a distinct acellular layer (Jester *et al.*, 2013).

1.2.7 Corneal nerves

The cornea is densely innervated, primarily with sensory neurons which relay information regarding injury and noxious stimuli. Nerve bundles are located at the periphery of the cornea and innervate the corneal stroma in a branching, radial manner before penetrating Bowman's layer and the corneal epithelium. Neuronal axons are unmyelinated to maintain corneal transparency (Müller *et al.*, 2003).

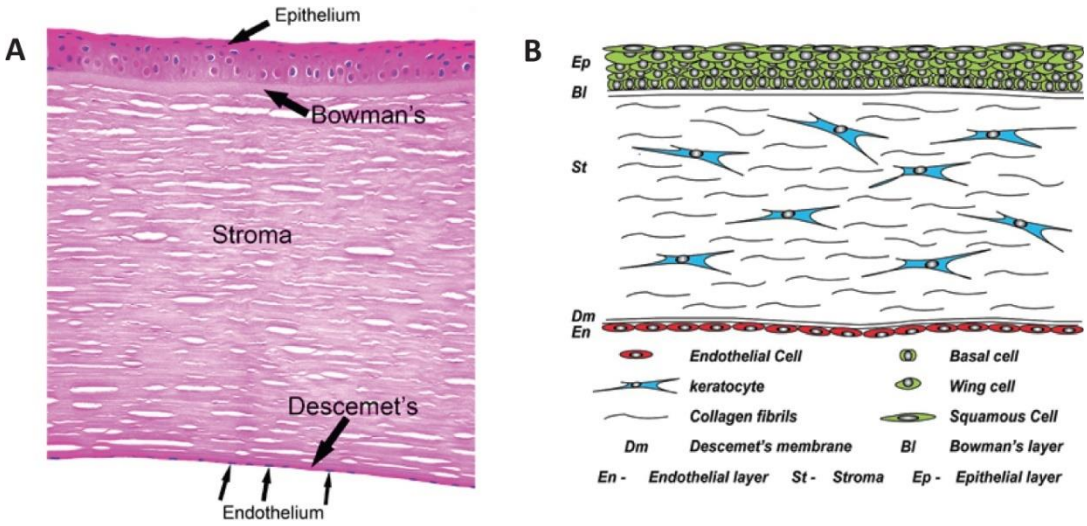


Figure 1.2 Structure of the healthy adult cornea. (A) Hematoxylin and eosin staining of a healthy human adult cornea. The five layers of the cornea are highlighted. Gaps present in the stroma are artefacts of fixation process. Image from (Meeney and Mudhar, 2013). **(B)** Illustration of cornea. The multicellular corneal epithelium (Ep) is located at the most anterior surface. Bowman's layer (Bl) is an acellular layer located at the anterior surface of the stroma. The stroma (St) is the thickest layer and is primarily an extracellular matrix with sparsely distributed keratocytes. Descemet's membrane (Dm) is the basement membrane of the most posterior cellular layer, the endothelium (En), which is a cellular monolayer. Image from (Secker and Daniels, 2009).

1.3 Corneal development

Corneal development is a dynamic process that occurs alongside development of other ocular tissues, particularly those of the anterior segment. Much of our understanding of ocular development in vertebrates originates from studies carried out primarily in chicks and mice; however, important differences have been noted between development in humans and these animals. Insights into anterior segment development have also been gained following the identification of genes underlying human anterior segment dysgenesis which disrupt this highly regulated developmental process.

1.3.1 Anterior segment development

Embryonic development of the mammalian cornea is intrinsically dependent on the development of the other structures of the anterior segment, particularly the lens (Zhang, Overbeek and Govindarajan, 2007). Ocular development begins after

gastrulation, when the embryo consists of three germ layers; ectoderm, mesoderm and endoderm. The diencephalon extends bilaterally towards the surface ectoderm to form the primary optic vesicles. Where the optic vesicles interact with the surface ectoderm, it induces the ectoderm to thicken and thereby forms a lens precursor structure called the lens placode. The lens placode invaginates and subsequently buds off from the surface to form a lens vesicle (Gilbert, 2000; Chow and Lang, 2001) (Figure 1.3). In humans, ocular development begins during the fourth week of gestation with lens placode formation occurring by the end of the fifth week (Hay, 1980). A number of genes encoding transcription factors involved in anterior segment development have been identified including *PAX6, PITX2, FOXC1* and *FOXE3* which cause inherited anterior segment dysgeneses when their function or dosage is disrupted by mutations (Sowden, 2007).

1.3.2 Epithelial development

The surface ectoderm overlying the lens vesicle will become the corneal epithelium. Initially 3 – 4 cells thick, it regresses to 2 cellular layers after eyelid fusion and is maintained in this state with limited proliferation until the 17th week of gestation. The corneal epithelial cells then undergo a period of high proliferation across the entire cornea, resulting in increased corneal diameter, which lasts until the 28th week of gestation. After this stage, cellular proliferation gradually decreases, the eyelids open, and the corneal epithelium begins to differentiate and expresses corneal epithelial specific markers. The cornea continues to grow in diameter, and the epithelium increases to 4 – 6 cell layers in thickness. By six months after birth, cell proliferation is restricted to the boundary of the cornea and the conjunctiva (the limbus) (Sevel and Isaacs, 1988; Yew *et al.*, 2001). Suppression of Wnt signaling pathways may be required for differentiation of the surface ectoderm to the corneal epithelium (Mukhopadhyay *et al.*, 2006).

1.3.3 Development of the corneal stroma and endothelium

During development, neural crest cells enter the cornea in two migratory waves; the first wave forming the corneal endothelium and the second wave differentiating into stromal keratocytes (Silla *et al.*, 2014) (Figure 1.3). Signals from the lens are crucial

for differentiation of neural crest cells into keratocytes and CECs (Zhang, Overbeek and Govindarajan, 2007). Paracrine retinoic acid signaling from the overlying surface ectoderm and the anterior lens epithelium acts on the periocular mesenchyme located between these layers, and regulates expression of transcription factors including *Pitx2* (Cvekl and Wang, 2009). The continued expression of *Pitx2* requires canonical Wnt signaling derived from the surface ectoderm (Zacharias and Gage, 2010). In mice, loss of *Pitx2* in neural crest cells specifically, results in a complete absence of the corneal endothelium and stroma (Evans and Gage, 2005). In humans, the corneal endothelium forms a monolayer by the 8th week of gestation. Once formed, the CECs secrete Descemet's membrane which matures between the second and eighth month of gestation (Wulle, 1972; Hay, 1980). Formation of the corneal endothelial monolayer is required for separation of the lens and iris from the cornea, and therefore for development of anterior chamber (Reneker *et al.*, 2000). The second wave of migrating neural crest cells infiltrate the space between the corneal epithelium and endothelium (Hay, 1980). Mesenchymal cells become more flattened in appearance as they differentiate into mature keratocytes and secrete components of the stromal matrix. As the stroma develops, lamellae are laid down in regular parallel layers (Sevel and Isaacs, 1988).

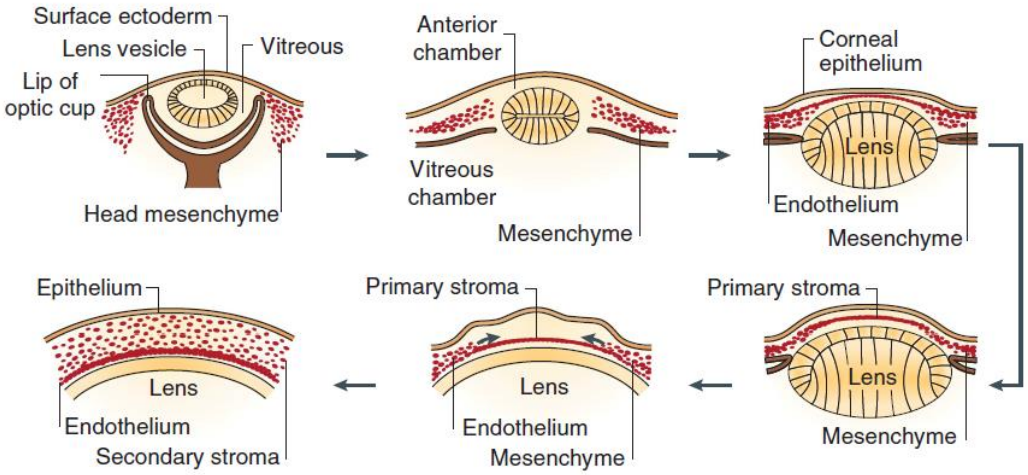


Figure 1.3 Illustration showing stages of human corneal development. The putative corneal epithelium develops due to interaction between the surface ectoderm and the lens vesicle. Waves of neural crest cells, derived from mesenchyme, infiltrate the space between the lens and the corneal epithelium and form the endothelium and stroma, respectively. Image from (Zavala *et al.*, 2013).

1.4 Corneal homeostasis and wound healing

1.4.1 Epithelial homeostasis

At the anterior surface, the superficial epithelial cells of the human cornea have a turnover of 7 – 10 days (Hanna, Bicknell and O'Brien, 1961). The current accepted model of epithelial maintenance is the X, Y, Z hypothesis which states that the loss of epithelial cells from the anterior surface (Z) is replenished by the division of basal epithelial cells (X) and the centripetal migration of cells across the corneal surface (Y), $Z = X + Y$ (Thoft and Friend, 1983). The basal epithelial cells are maintained by a population of adult stem cells in the limbus, at the perimeter between the sclera and the cornea. These limbal epithelial stem cells (LESCs) have the capacity to divide asymmetrically to produce both LESCs and transient amplifying cells (TACs). The LESCs maintain the stem cell population while the TACs migrate centripetally at the basal epithelium towards the central cornea. As the TACs divide, cells move anteriorly and become differentiated to replace the loss of the superficial cells shed from the surface into the tear film (Yoon, Ismail and Sherwin, 2014).

The limbus has several unique features that provide a suitable environment for the maintenance of stem cell characteristics. Unlike the central cornea, the limbus is vascularized allowing nutrients to be delivered. Furthermore, it contains a high proportion of melanocytes which protect the cells from UV damage. LESCs are located primarily within the Palisades of Vogt; these are undulations of the stroma which provide an increased surface area and protect the cells from shearing forces (Secker and Daniels, 2009) (Figure 1.4). Specifically, limbal crypts and focal stromal projections (Shortt *et al.*, 2007) have been proposed as potential locations of the putative stem cell niche. The importance of LESCs is exemplified by the disorder limbal stem cell deficiency (LSCD) in which the LESC population is lost, for example, by chemical burn or injury. LSCD enables conjunctival cells to migrate onto the corneal surface, often with neovascularisation of the cornea, resulting in visual impairment (Kolli *et al.*, 2010). However, there is evidence to suggest that oligopotent cells are present in the central cornea in the pig (Majo *et al.*, 2008). The authors propose that LESCs are concentrated at the limbus due to the opposing

replicative forces of the cornea and the conjunctiva, but may be present over the entire corneal surface.

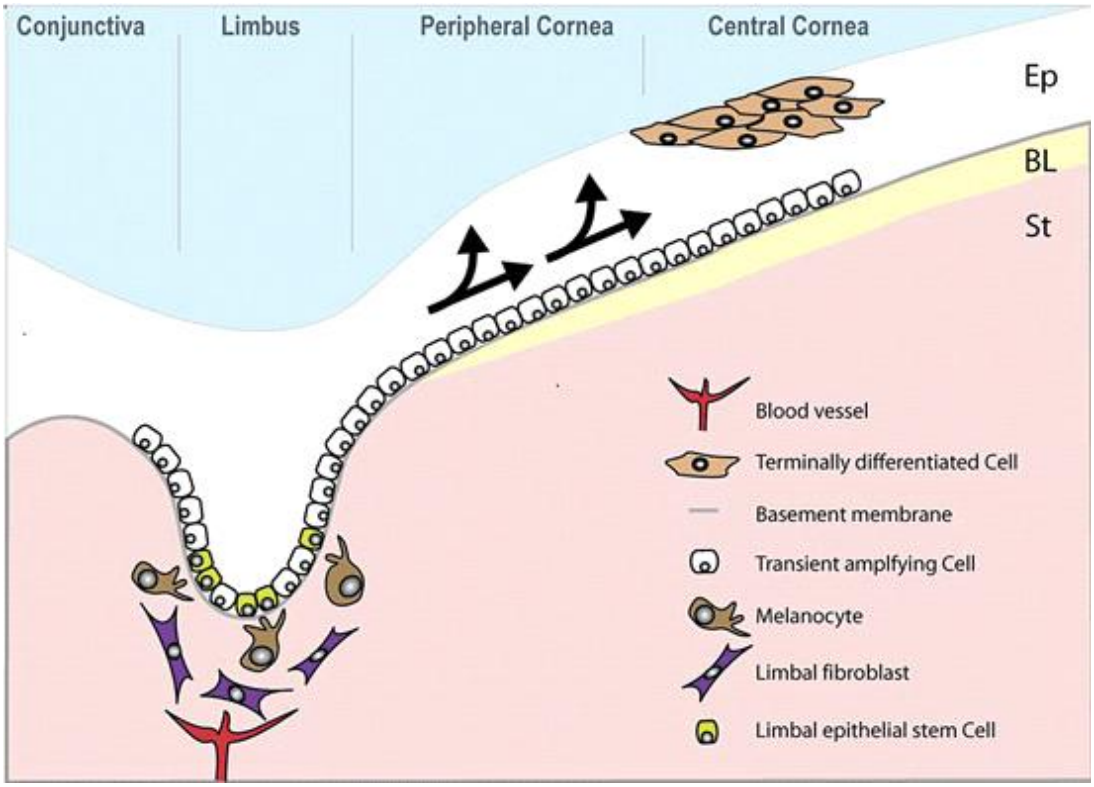


Figure 1.4 The corneal limbus and its role in epithelial maintenance. Limbal epithelial stem cells are located within undulations of the stromal surface called Palisades of Vogt, at the corneal limbus, where they are protected from shearing forces and have a source of nutrients from the blood vessels. The stem cells divide to produce transit amplifying cells which migrate across the central cornea in a centripetal pattern and become differentiated to basal epithelial cells, wing cells and superficial cells. Ep: Epithelium, BL: Bowman's layer, St: Stroma. Image from (Secker and Daniels, 2009).

1.4.2 Stromal homeostasis

In comparison to the epithelial cells which are continually replaced, stromal keratocytes are quiescent with a turnover of a few years. The existence of keratocyte precursor cells has been demonstrated in the limbus; these corneal stromal stem cells (CSSCs) are located in the anterior stroma, posteriorly to the Palisades of Vogt, and have been isolated and propagated in culture. They have stable expression of putative stem cell markers and can also be induced to differentiate into keratocytes

or chondrocytes, however their contribution to homeostasis or wound healing in the cornea has not been resolved (Du *et al.*, 2005).

1.4.3 Epithelium and stromal wound healing

Injury to the cornea, by trauma or infection, initially causes apoptosis of the keratocytes in the adjacent stroma, in a reaction which is potentially mediated by cytokines expressed by damaged corneal epithelial cells (Wilson *et al.*, 2001). The remaining keratocytes undergo mitosis to replenish the keratocyte population. Following this, some keratocytes become 'activated' and differentiate into fibroblasts. Fibroblasts differ from keratocytes in that they proliferate more rapidly, are motile and have a fusiform morphology as oppose to the stellate morphology of keratocytes. Unlike keratocytes, fibroblasts also have bundles of actin stress fibres, focal adhesions and multiple nucleoli (Fini and Stramer, 2005; West-Mays and Dwivedi, 2006). Both keratocytes and fibroblasts secrete components of the stromal ECM, however the composition differs between the two cellular phenotypes. Keratocytes secrete a highly organised extracellular matrix of collagen fibers and proteoglycans, whereas the ECM secreted by fibroblasts is disorganised and lacks expression of corneal-specific proteoglycans, which are essential for corneal transparency (Wu *et al.*, 2014). Quiescent keratocytes in the healthy cornea also express crystalline proteins which are involved in maintaining transparency by minimising light scatter from cells. Fibroblasts do not express crystalline proteins, and this can be a contributing factor to the development of corneal haze (Jester, 2008). Myofibroblasts have a distinct stromal cellular phenotype, and can be distinguished from other fibroblasts by their expression of α-small muscle actin. *In vitro* they can be differentiated from keratocytes or fibroblasts by incubation in cell culture media supplemented with TGFβ (Jester *et al.*, 1996). Myofibroblasts express hyaluronan which enhances the fibrotic wound healing response. As the stromal ECM is crucial for maintaining transparency, the fibrotic wound healing response can result in opaque scar formation (Guo *et al.*, 2010). Total levels of keratan sulphate are reduced in fibroblast and myofibroblast cultures compared to keratocytes, while chondroitin/dermatan sulphate levels are increased. In addition, there is a reduction in chain length and sulphation of keratan sulphate expressed by fibroblasts and myofibroblasts compared to keratocytes, and an increase in chain

length and sulphation of chondroitin/dermatan sulphate. Changes in levels of GAGs are thought to contribute to opaque scar formation (Funderburgh, Mann and Funderburgh, 2003).

1.4.4 Endothelial homeostasis

In an adult cornea, CECs are arrested at the G1 phase of the cell cycle, therefore do not divide. They are slowly lost throughout life at a rate of 0.3 – 0.6% per year but there is usually a sufficient quantity of CECs to maintain the leak-pump barrier function and maintain corneal deturgescence throughout life. There is some evidence to suggest that endothelial stem cells or TACs are located in the endothelial limbus, at the junction of the endothelium and the trabecular meshwork (Whikehart, Parikh and Vaughn, 2005). This region contains a higher density of cells in comparison to the central endothelium and cells located here express some putative stem cell markers; however, their role in relation to maintenance of the adult CEC population is not known (McGowan *et al.*, 2007).

1.4.5 Endothelium wound healing

When the corneal endothelium is injured, due to trauma or disease, loss of CECs is accelerated. The CECs respond to cell loss by becoming larger to maintain the barrier of the monolayer, resulting in the cells becoming pleomorphic and migrating to sites where cells have been lost. Below a certain density the cells are unable to maintain fluid regulation leading to oedema, which can significantly impair vision (Joyce, 2005).

1.5 Corneal dystrophies

1.5.1 Overview

Corneal dystrophies are a group of clinically and genetically heterogeneous inherited disorders that cause a loss of corneal transparency and result in visual impairment. With few exceptions, corneal dystrophies are monogenic disorders with autosomal dominant, autosomal recessive and X-linked modes of inheritance. The vast majority

are associated exclusively with corneal symptoms although there may be some systemic features associated with particular dystrophies (Weiss *et al.*, 2015).

The loss of corneal transparency caused by a corneal dystrophy occurs due a diverse range of pathological mechanisms. Anterior corneal dystrophies, affecting the epithelium and superficial stromal layers, are typically caused by the deposition and accumulation of protein, glycoprotein or phospholipid opacities in these corneal layers. These opacities scatter light, prevent light from entering the eye and impair vision (Oliver and Vincent, 2016). Corneal opacities can also cause painful recurrent epithelial erosions. A corneal erosion is the detachment and subsequent loss of the epithelial cell layer from the underlying epithelial basement membrane, resulting in pain, photophobia and tearing. Corneal erosions may also cause the formation of opaque scars and secondary opacities due to the corneal wound healing response (Das and Seitz, 2008). In contrast, corneal endothelial dystrophies occur due to disruption of the barrier and pump function of the corneal endothelial monolayer due to altered behavior and/or morphology of the endothelial cells. This eventually results in a loss of fluid regulation between the cornea and the anterior chamber, resulting in corneal oedema and a loss of corneal transparency (Klintworth, 2009).

1.5.2 Clinical diagnosis

The clinical diagnosis of a corneal dystrophy is typically based on the macroscopic appearance of the cornea and opacities. All layers of the cornea can be examined using slit-lamp biomicroscopy, when visualized using full light, narrow field or retroillumination. Most corneal dystrophies are characterized by a specific combination of features when examined using the slit lamp that is sufficient to determine the likely clinical diagnosis (Lisch and Seitz, 2011). In addition, specular microscopy is commonly used to visualize the endothelium at higher magnification and calculate endothelial density (Sturrock, Sherrard and Rice, 1978). *In vivo* confocal microscopy is now widely used, as it can visualize all corneal layers at a higher magnification, with the exception of Descemet's membrane (Belin *et al.*, 2010).

If the corneal dystrophy causes significant visual impairment, the entire thickness of the central cornea, or the specifically affected layer, may be surgically removed. In these cases, the tissue can be analysed histologically to identify the features of the diseased corneal tissue, and the staining pattern of key cellular marker proteins can provide support or exclude a proposed clinical diagnosis. Corneas are typically formalin-fixed paraffin embedded and stained using a variety of standard histological stains to identify the nature of the corneal opacities. Hematoxylin and eosin (H&E) stain is used routinely; this stains nuclei purple, due to negatively charged nucleic acids, while proteins which are primarily positively charged are stained pink (Figure 1.2). Masson's trichrome is another commonly used stain which stains nuclei dark brown, cytoplasm pink, collagen blue-green and keratin red. Alcian blue or Periodic acid-Schiff (PAS) can be used to identify glycoproteins, while Congo red binds to amyloid protein and exhibits yellow-green birefringence in plain polarised light. Histological characteristics, including the appearance of the corneal layers and location and composition of opacities, vary between different corneal dystrophies (Vemuganti, Rathi and Murthy, 2011).

Other relevant characteristics used to aid a clinical diagnosis include the apparent inheritance pattern in the affected family, the age of onset, rate of progression and severity of disease (Lisch and Seitz, 2011). Further complicating the clinical diagnosis, corneas that have previously undergone PTK or corneal transplantation will differ in clinical appearance from those with primary opacities. Ultimately, a definitive diagnosis can only be provided by identification of the underlying genetic cause.

1.5.3 Management options

There are no treatments available for corneal dystrophies that correct the underlying genetic lesion. Some corneal dystrophies are mild enough that they are asymptomatic and do not cause significant visual impairment, however the majority of corneal dystrophies are slowly progressive and therefore will eventually impair visual acuity to a sufficient level to require corneal transplant surgery to one or both eyes or laser ablation of opacities (Weiss *et al.*, 2015). A corneal transplant involves the surgical removal of the full thickness of the central cornea, or specifically the affected layers

of the cornea, which is then replaced with the equivalent tissue from a donor cornea. The replacement of the full thickness cornea requires a surgical procedure called penetrating keratoplasty (PK). This procedure is now less commonly performed, with the majority of surgeries removing either only the anterior epithelial and stromal layers by deep anterior lamellar keratoplasty (DALK) or the posterior endothelial layer with endothelial keratoplasty (EK), although PK may still be performed when there are opacities, damage or scarring affecting all corneal layers (Lee *et al.*, 2009; Borderie *et al.*, 2012).

1.5.3.1 Anterior corneal dystrophies

DALK is a surgical procedure which involves removing only the anterior layers of the cornea. DALK is used to treat corneal dystrophies associated with epithelial and/or stromal opacities, providing the patient's corneal endothelium is not affected. There are a range of surgical techniques to achieve this, of which the most common is the 'big bubble' technique which involves injecting air to separate the stroma from the Descemet's layer. The advantages of this surgery in comparison to PK is that it almost eliminates the risk of immune rejection, although stromal rejection can still occur, and is associated with fewer post-operative complications and higher long-term graft survival rates. Although there is a reduction in endothelial cell density following DALK surgery, this is far lower than the comparable decrease seen in the 6 months following PK surgery (Reinhart *et al.*, 2011; Borderie *et al.*, 2012). A less invasive alternative to corneal transplant for anterior corneal dystrophies is PTK which uses excimer laser generated 193 nm UV light to ablate superficial opacities and to smooth surface irregularities (Figure 1.5) (Lee and Kim, 2003).

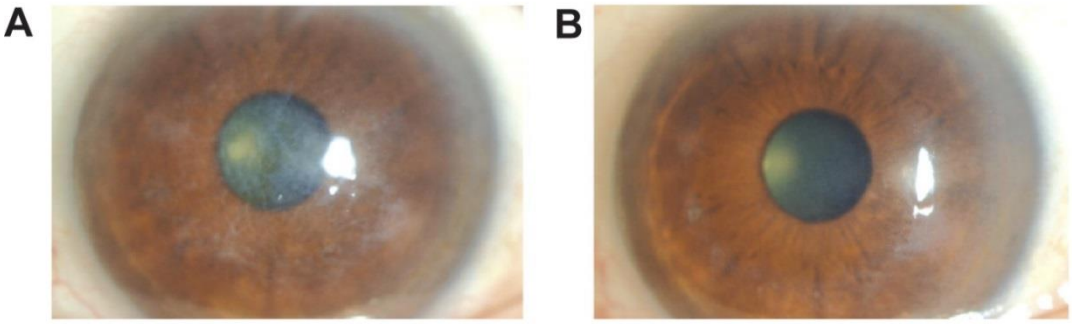


Figure 1.5 Clinical outcome of phototherapeutic keratectomy (PTK) in a patient with Thiel-Behnke corneal dystrophy. (A) Affected eye prior to PTK. **(B)** One month post-PTK, superficial opacities have been removed, restoring some visual acuity. Images from (Hieda *et al.*, 2013)**.**

1.5.3.2 Endothelial dystrophies

EK is the removal of Descemet's membrane and the attached endothelial monolayer while maintaining the patient's own stromal and epithelial layers. A number of techniques have been developed to achieve this including Descemet's stripping endothelial keratoplasty (DSEK), Descemet's stripping automated endothelial keratoplasty (DSAEK), and Descemet's membrane endothelial keratoplasty (DMEK). DSEK is a transplant of some posterior stroma in addition to Descemet's membrane and endothelium, whereas DMEK is the transplantation of a thinner graft consisting of Descemet's membrane and the endothelium only. EK surgery is associated with faster visual recovery than PK and fewer incidences of graft failure compared to PK (Price, Feng and Price, 2015).

1.5.3.3 Challenges of surgical treatments

One problem associated with these surgical treatments for corneal dystrophies is the eventual recurrence of the dystrophy in the grafted or PTK-treated tissue, which can occur for both anterior and endothelial dystrophies. Anterior dystrophies show a recurrence of opacities after surgery or PTK-treatment, although these may be phenotypically different from the primary opacities and can occur soon, or even years after surgery (Dinh *et al.*, 1999; Unal *et al.*, 2013). With endothelial dystrophies, the patient's remaining endothelial cells can overgrow the donor endothelium (Merjava *et al.*, 2011) or endothelial cell density may decrease following surgery to a density at which oedema recurs. It is not uncommon for

individuals to require repeated grafts and multiple PTK-treatments throughout their life, and therefore surgery, although initially effective for the majority of cases, is not usually a long-term cure. Surgery is also invasive and both DALK and EK surgeries have a risk of complications. DALK complications include interface irregularity, double anterior chamber, epithelial healing problems and graft rejection (Unal *et al.*, 2013). DMEK surgery can result in failure of graft attachment, endothelial graft rejection, primary graft failure and iatrogenic glaucoma (Lee *et al.*, 2009). In addition to these potential complications, there is a shortage of available suitable donor corneas for corneal transplantation; therefore, there is a need for less invasive and more effective treatments for corneal dystrophies.

As corneal dystrophies are primarily monogenic inherited disorders, there are a variety of options for the development of gene-specific therapies for corneal dystrophies, which are currently being developed for other inherited disorders. However, knowledge of the underlying genetic causes of these disorders is a pre-requisite of developing these treatments.

1.6 Genetic techniques

Corneal dystrophies are inherited monogenic disorders and, as such, obtaining a genetic diagnosis is the key to understanding these conditions. The field of genetics has undergone a revolution in the past three decades due to the development of rapid, affordable sequencing technologies.

1.6.1 Linkage analysis

Linkage analysis is a genetic technique used extensively for Mendelian disease gene discovery. Prior to the advent of next generation sequencing, linkage analysis typically represented the first step for genetic investigation of an inherited disease and is still commonly used. Linkage analysis utilises the principle that the disease-causing mutation will not be inherited in isolation, but will be inherited as a shared haplotype, alongside other benign variants which are in linkage disequilibrium with the mutation. To identify these haplotypes, linkage analysis looks for markers which are shared between affected related individuals but are absent from unaffected family

members. Historically, microsatellites were commonly used for linkage analysis but it is now more common to genotype individuals across the entire genome using dense small nucleotide polymorphism (SNP) genotyping chips. Following genotyping of polymorphic markers, analysis is performed to identify a statistically significant locus that segregates with the disease in a pedigree. The result of this analysis is a logarithm of odds (LOD) score for each marker which is a measure of the probability of linkage, calculated by log10 [probability linkage / probably of non-linkage]. An autosomal locus is generally considered to be statistically significant if the LOD score is 3.0 or higher (Strachan and Read, 2011).

The size of the mapped locus is dependent on a number of factors, including the number of informative polymorphic markers, therefore use of dense SNPs provide greater resolution. Historically, loci mapped with microsatellites were large and contained tens, if not hundreds, of genes. In these cases, once the linked region was known, higher resolution linkage analysis of the disease locus would be used to refine the region by selecting more markers within the locus. This was usually followed by directly Sanger sequencing the most likely candidate genes within the locus. Nowadays, linkage analysis is typically combined with next-generation sequencing data. Although undoubtedly useful for gene discovery, linkage analysis requires large pedigrees with DNA available from multiple affected and unaffected (clinically examined) family members, which is not always available for some inherited disorders, particularly recessive disorders in which there may be only a few affected individuals in a single generation. A common approach for recessive disorders affecting consanguineous families is autozygosity mapping. In consanguineous families, the disease-causing mutation is likely to be in the homozygous state and inherited as a haplotype that is present on both alleles, therefore this approach uses genome-wide markers to identify shared regions of homozygosity in affected individuals to identify a locus that segregates with disease (Moynihan *et al.*, 1998). The locus can then be interrogated as described previously to identify the causative mutation.

1.6.2 Sanger sequencing

The development of Sanger sequencing has revolutionized the field of genetics by providing a simple and accurate method of sequencing that ultimately resulted in the publication of the first human genome sequence (International Human Genome Sequencing Consortium, 2001). To prepare for Sanger sequencing, DNA is fragmented and cloned into a plasmid vector or a specific genomic region is amplified by polymerase chain reaction (PCR). Sanger sequencing requires a single primer which hybridizes upstream of the region of interest and a DNA polymerase which extends the oligonucleotide primer based on the sample template sequence. Deoxynucleotides (dNTPs) are incorporated into the growing DNA strand as the reaction proceeds. Dideoxynucleotides (ddNTPs) are also included in the reaction mix; these are analogues of nucleotides that lack a 3'-hydroxyl group and thus when incorporated into a strand of DNA result in termination of the growing oligonucleotide strand. This termination is due to the requirement of a free 5' phosphate group and a 3'-hydroxyl group to form a phosphodiester bond. Due to the combination of dNTPs and ddNTPs, the sequencing reaction results in multiple DNA fragments of different lengths which terminate in a labelled ddNTP. Early conceptions of the method used radiolabelled ddNTPs, with four individual reactions required for each type of ddNTP. Electrophoretic separation of the resulting fragments using slab polyacrylamide gels allowed the relative position of the nucleotide that had been incorporated into the growing oligonucleotide chain to be determined (Sanger, Nicklen and Coulson, 1977). Further refinement of the process resulted in the use of fluorescently labelled ddNTPs, allowing the use of all four ddNTPs simultaneously in one sequencing reaction, and capillary electrophoresis to separate the oligonucleotide fragments. Although initially developed in 1977, Sanger sequencing can achieve read lengths of ~ 1000 bp and remains the gold standard for sequencing. It is often used to validate variants identified by newer sequencing technologies (Shendure and Ji, 2008).

1.6.3 Next-generation sequencing

Next-generation (or second-generation) sequencing encompasses a range of sequencing technologies which use different chemistries to sequence DNA in a high-

throughput manner. Next-generation sequencing began with the release of the first commercially available next-generation sequencer in 2005 and, since then, a range of sequencing platforms have been developed. Accuracy and sequencing capacity have improved over the years while the cost of sequencing per base has decreased due to developments both in the hardware and the underlying chemistry, making next-generation sequencing more accessible for both research and clinical applications.

1.6.3.1 Library preparation and capture

Next-generation sequencing begins with shearing of the DNA template to produce fragments of sufficient size for the sequencing chemistry being used. Sequencing can proceed either by sequencing the entire human genome or by targeting of specific regions using a capture. Captures consist of arrays of oligonucleotides which contain flanking regions to the genomic regions of interest to which the fragmented sample DNA is hybridized. Liquid captures can also be used in which sample DNA is hybridized to biotinylated oligonucleotide probes and then bound to streptavidin beads (Mardis, 2013). Captures provide an enrichment of sequences of interest from the library of fragmented DNA and can be used, for example, to target all known exons in a human genome, in whole exome sequencing (WES) (Ng *et al.*, 2010). Alternatively, captures may be designed to sequence only the exons of genes known to be associated with a particular disease for sequencing gene panels (Glöckle *et al.*, 2014). The latter are increasingly being offered as genetic tests within the NHS.

Due to the relatively short read length achieved by most next-generation sequencing (typically < 250 bp) in comparison to Sanger sequencing, most sequencing technologies sequence DNA fragments in both directions, allowing the reads to be aligned with greater confidence. This can be achieved using paired-end or mate-pair methods. Paired-ends are generated by fragmenting DNA to produce fragments less than 1 kb in size and ligating adapters to each end of the fragment containing different priming sites. This allows sequencing of the fragment to be achieved from both priming sites. Alternatively, generation of a mate-pair library is achieved by ligating typically larger than 1 kb DNA fragments to a single adapter in a circularized manner, so that both ends of the fragment bind each end of the adapter. Mate pairs are generated from the circularised DNA fragments (of known size) by

digestion at specific sites from each end of the full length fragment. Specific adapters are ligated onto these mate-pair fragments to provide the template for priming the sequencing reaction. During alignment of sequencing reads, paired-end and mate-pair fragments are paired together, taking into account the average fragment size (Mardis, 2013).

Most next-generation sequencing platforms use some form of light detection system within the sequencing process, therefore multiple copies of each fragment need to be generated to produce a signal of sufficient intensity to be detected. Unlike Sanger sequencing, which uses cloning or PCR in a liquid well format, amplification of DNA fragments occurs on a solid surface and can be achieved using emulsion-PCR (emPCR) or solid-phase amplification. EmPCR uses adapter ligated, single stranded, DNA fragments and attaches them to beads. The beads containing a single template are located in a water bubble in a water-oil emulsion. The PCR reaction occurs within this bubble and results in the production of a bead containing multiple attached copies of the initial template sequence (Figure 1.6 A). EmPCR is used by the ABI SOLiD sequencing platform (Applied Biosystems) and pyrosequencing (454 Roche) (Metzker, 2010). With solid-phase amplification, fragmented DNA is immobilized by hybridizing it to adapters which act as primers for PCR amplification to produce multiple copies (~1000). The clonal DNA fragments hybridize with complementary adapters to produce a bridge (Figure 1.6 B). The solid-phase amplification method of library preparation is used by Illumina HiSeq sequencing platforms.

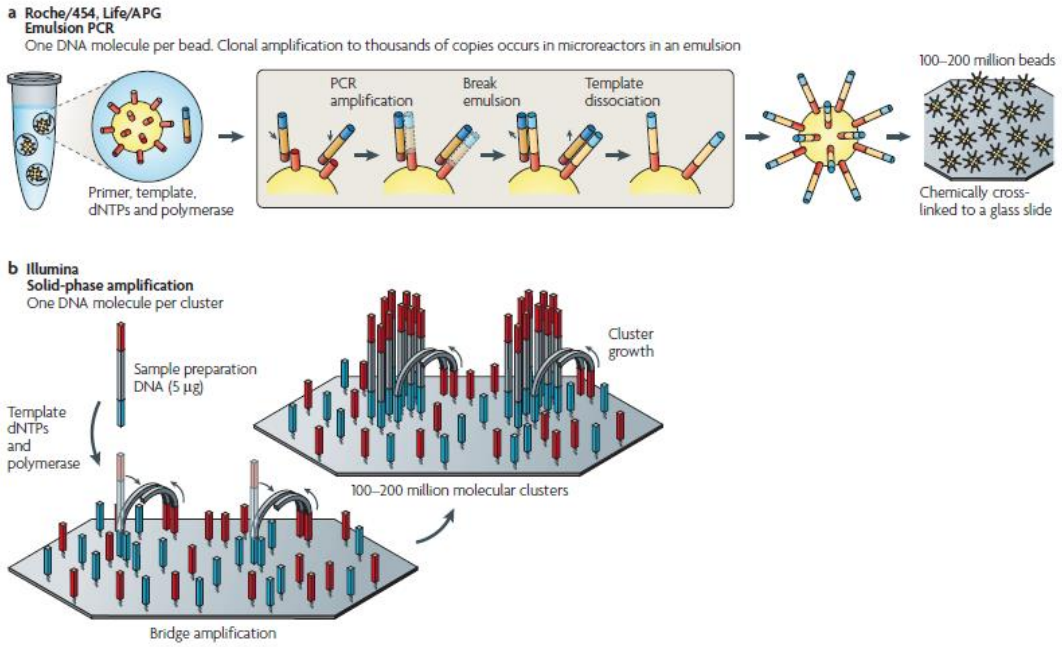


Figure 1.6 Two alternative methods of template DNA amplification prior to next-generation sequencing. (A) Emulsion PCR is PCR amplification occurring on the surface of a bead within a bubble formed by a water-oil emulsion, containing the template sequence and PCR components. After amplification, the beads (carrying multiple copies of each template) are then chemically cross-linked to a glass slide. (**B)** In **s**olid phase amplification, both ends of the template have ligated adaptors which anneal to a surface containing complementary adaptors, forming a bridge. Clusters of replicated templates are produced as amplification occurs. Image adapted from (Metzker, 2010).

1.6.3.2 Sequencing chemistries

Sequencing-by-synthesis is a sequencing method which shares similarities with Sanger sequencing in its underlying chemistry, and is used by the Illumina HiSeq platforms. Using a primer that is complementary to the adapters which were previously ligated to the DNA library fragments, and a modified DNA polymerase, it uses fluorescently labelled nucleotides with a blocked 3'OH site to determine which nucleotide is incorporated into a growing oligonucleotide chain. Termination occurs after the incorporation of a single nucleotide due to the blocked 3'OH and the nucleotide that has been incorporated is detected by a charged coupled device camera. Following the incorporation of a single nucleotide, the florescent dye located at the 3' end of the preceding base is cleaved and the 3'OH group is unblocked, allowing the cycle to be repeated and the next nucleotide to be incorporated (Figure 1.7). Following 150 cycles of nucleotide addition, the entire

process is repeated but priming also from the second adapter to achieve paired-end sequencing (Mardis, 2013).

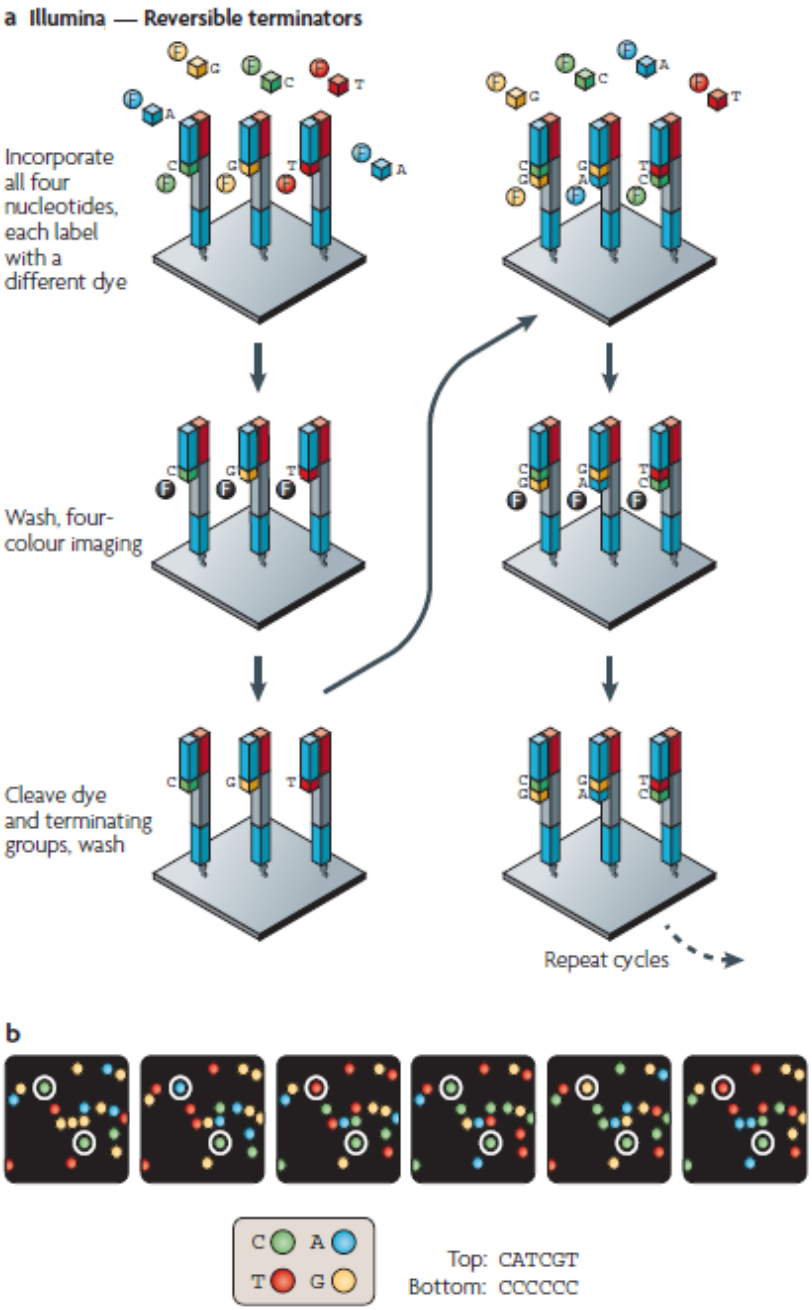


Figure 1.7 Next-generation sequencing using sequencing-by-synthesis chemistry with reversible terminators. (A) Dye labelled nucleotides containing blocked 3'OH sites are added to the sequencing reaction and are incorporated into the growing DNA strand. After imaging of the fluorescence, the blocked site is cleaved and the cycle is repeated. **(B)** Detection of the light signal is imaged by a camera to determine which nucleotide has been incorporated, using the example of two clonally amplified fragments. Image from (Metzker, 2010).

Pyrosequencing (454 Roche) is an alternative sequencing-by-synthesis approach that uses bioluminescence to determine the identity of the incorporated nucleotide in the synthesised oligonucleotide chain (Lehmann and Tost, 2015). Another technique, sequencing by ligation, is adopted by the Applied Biosystems ABI SOLiD system (Pandey, Nutter and Prediger, 2008). In addition to these most commonly utilized sequencing chemistries, there are variations and additional sequencing methods including reversible cyclic termination (Helicos Biosciences), sequencing by pH (Ion Torrent) and single molecule real-time sequencing (Pacific Biosciences). These are included in next-generation sequencing technologies or referred to as third-generation sequencing technologies.

1.6.4 Identification of copy number variation

While the previously described sequencing methods can identify small insertions and deletions, larger (sub-microscopic) regions of copy number variation (CNV) including deletions, duplications and triplications of loci containing one or multiple genes, can cause monogenic disease, and these are often undetected by Sanger sequencing.

1.6.4.1 Microarray-based methods

The most common method for detection of CNVs across the entire human genome uses DNA microarrays. Array comparative genomic hybridisation (aCGH) is a method that uses glass slides spotted with genomic clones or oligonucleotides which provide coverage across the human genome. DNA of interest, along with diploid reference DNA, is fluorescently labelled and hybridized to the microarray. The ratio of fluorescence of the test DNA in comparison to the reference sample can be used to determine copy number of each of the clones or oligonucleotides. The resolution depends on the size and number of clones or oligonucleotides used by the particular microarray (Carter, 2007).

An alternative approach is to use a SNP genotyping array, the most commonly used of which are produced by Affymetrix and Illumina. SNP arrays work by hybridizing a labelled sample DNA to a selection of probes containing SNPs across the human

genome. Signal intensities between a test sample in comparison to other samples, can result in the determination of copy number. Furthermore, the genotype information can be used to infer the presence of deletions, specifically. SNP arrays have a number of advantages over aCGH methods including less DNA sample input, and they provide additional information about genotype. The resolution achieved depends on density of SNP probes included on the SNP chip (Carter, 2007; Winchester, Yau and Ragoussis, 2009).

1.6.4.2 Quantitative real-time PCR

Quantitative real-time PCR (qPCR) can be used to determine copy number of a specific region of the genome. Primers are designed to amplify a small region (~100bp) of DNA, for example, part of a gene of interest, and two potential fluorescence-based methods allow the quantity of DNA present in the reaction to be calculated throughout the entire PCR reaction by measuring the florescent signal generated. This can be achieved either using a dye, such as SyBr green, which binds to any double stranded DNA present in the reaction, or using a fluorescent probe which hybridizes to a specific target, most commonly TaqMan probes. A value is calculated for each reaction, the cycle at which the fluorescent signal reaches a particular threshold, and this can be converted to a known amount of DNA using interpolation from a standard curve of known DNA concentrations. Comparison of the DNA levels of the gene of interest with an equivalent control DNA region which has a diploid copy number, allows the copy number of the gene of interest to be elucidated (D'haene, Vandesompele and Hellemans, 2010).

1.6.4.3 Next-generation sequencing

Next-generation sequencing technologies can also be used to detect CNVs across the genome. A number of algorithms have been designed to utilize raw sequencing data for detection of CNVs. Most commonly, programs compare sequencing read depth at a particular locus to an expected read depth at that site by comparison with flanking sites and/or other samples sequenced at the same time. An alternative approach is to look for mate-pairs which are separated by a smaller or greater distance than expected (Zhao *et al.*, 2013). Coverage is important for detection of CNVs, therefore

WGS data is more useful for CNV detection in comparison to WES data, although some programs have been developed specifically for the latter (Plagnol *et al.*, 2012).

1.7 Genetics of corneal dystrophies

1.7.1 Anterior corneal dystrophies

Advances in genetic sequencing technologies have resulted in the identification of a number of corneal dystrophy genes, which has in turn has changed the way that corneal dystrophies are diagnosed and classified. The first wave of discovery in corneal dystrophy genetics followed the adoption of Sanger sequencing. Most early genetic investigations shared a common approach of using linkage analysis in large pedigrees to identify a genetic locus containing the disease gene, and Sanger sequencing of candidate genes within the linked interval to identify the mutation underlying the disease. This approach was largely successful for the majority of anterior corneal dystrophies with the most common of these being genetically solved by 2000. One of the earliest major findings in the field of corneal dystrophy genetics was the identification that four clinically and histologically distinct autosomal dominant corneal dystrophies mapped to an overlapping locus on 5q31 (Stone and Mathers, 1994; Gregory, Evans and Bhattacharya, 1995; Small *et al.*, 1996). Following refinement of the genetic locus, mutations were identified in a shared candidate gene within the linked interval, *TGFBI* (Munier *et al.*, 1997). In the same year, heterozygous mutations in *KRT3* and *KRT12* were shown to cause Meesmann corneal dystrophy (MECD) (Irvine *et al.*, 1997; Nishida *et al.*, 1997). The autosomal recessive dystrophy, gelatinous drop-like dystrophy (GDLD), was mapped using autozygosity mapping in several affected families from Japan, where the dystrophy is most prevalent, and led to the identification of *TACSTD2* as the causative gene (Tsujikawa *et al.*, 1998, 1999). The most common autosomal recessive dystrophy, macular corneal dystrophy (MCD), was solved a year later by linkage analysis and candidate gene re-sequencing, caused by mutations in *CHST6* (Akama *et al.*, 2000). Milder and/or rarer dominant anterior corneal dystrophies including fleck corneal dystrophy (FCD), Schnyder corneal dystrophy (SCD) and posterior stromal corneal dystrophy (PSCD) were genetically solved in in the mid 2000's using a similar linkage and candidate gene screening approach and appear to be genetically

homogenous (Bredrup *et al.*, 2005; Li *et al.*, 2005; Orr *et al.*, 2007; Weiss *et al.*, 2007).

In 2014, investigation of probands with the rare unsolved dystrophy, posterior amorphous corneal dystrophy (PACD), using next-generation sequencing, failed to find a causative mutation within the mapped locus. Following this, SNP array CNV analysis identified large deletions encompassing genes *EPYC, KERA, LUM* and *DCN* in three families with this dystrophy (Kim *et al.*, 2014). Also in 2014, a study utilizing WES of a large Swedish family identified mutations in *COL17A1* as causing epithelial recurrent erosion dystrophy (ERED) (Jonsson *et al.*, 2014).

There remain some anterior corneal dystrophies for which the underlying genetic cause has not been identified, including Lisch epithelial corneal dystrophy (LECD), which has been mapped to chrXp22.3 (Lisch *et al.*, 2000). Central cloudy dystrophy of François (CCDF) has been described in multiple families, however no locus or causative gene has been identified (Strachan, 1969) and sub-epithelial mucinous corneal dystrophy (SMCD) has been described rarely but other known causes of corneal dystrophies have not been excluded. There is also some evidence for additional locus heterogeneity for some recessive anterior corneal dystrophies with genetically unsolved families following screening of the known causative genes (Ren *et al.*, 2002; Akhtar *et al.*, 2005; Alavi *et al.*, 2007; Birgani *et al.*, 2009).

1.7.2 Endothelial corneal dystrophies

In comparison to anterior corneal dystrophies, deducing the underlying cause of endothelial corneal dystrophies has been more problematic, in part due to greater genetic heterogeneity and less striking genotype-phenotype correlation. Linkage mapping in a large family with posterior polymorphous corneal dystrophy (PPCD) identified a locus on chr20p in 1995, which was refined ten years later (Héon *et al.*, 1995; Gwilliam *et al.*, 2005). In the same year, another dominant, although more severe, congenital hereditary endothelial dystrophy, (CHED), was mapped to an overlapping locus (Toma *et al.*, 1995). The causative gene had not been identified for either of these dystrophies, despite intense investigation including next-generation sequencing of the disease locus (Lai *et al.*, 2010). A second locus was mapped in a

family with a similar, but clinically distinct, endothelial dystrophy, early-onset FECD, to chr1p34.3 – p32. *COL8A2* was selected as a good candidate gene within this locus and subsequently identified to harbour a likely causative mutation in the family. Screening this gene in a PPCD cohort also revealed putative mutations in these patients (Biswas *et al.*, 2001). A third PPCD locus mapped to chr10 was identified in 2005, and soon after *ZEB1* was identified as the causative gene (Shimizu *et al.*, 2004; Krafchak *et al.*, 2005). Autosomal recessive CHED was linked to a locus on chr20p13 with candidate gene screening of *SLC4A11* identifying homozygous and compound heterozygous mutations in this gene as causative of disease (Vithana *et al.*, 2006). An X-linked endothelial dystrophy (XECD) has also been described, for which the locus has been mapped, but no gene has been identified (Schmid *et al.*, 2006). Fuchs endothelial corneal dystrophy (FECD) has been mapped to multiple loci in a wide range of families and also appears to be a complex disorder for the majority of cases. The degree of additional genetic heterogeneity for endothelial dystrophies is unknown.

1.8 The IC3D classification of corneal dystrophies

Historically, classification of corneal dystrophies was based on the morphological layer of the cornea that is predominantly affected, appearance of the corneal opacities, clinical features including age of onset, and histological staining characteristics of surgically removed corneal tissue. However, in recent decades, the increased accessibility and affordability of genetic sequencing technologies has resulted in the genetic basis being identified for a significant proportion of corneal dystrophies, as outlined in the previous section. This has resulted in the discovery that some corneal dystrophy diagnoses represent an expansion of the phenotypic spectrum of other described corneal dystrophies. Conversely, corneal dystrophies which were thought to be a single diagnostic category have been found to be genetically heterogeneous. Based on these findings, The International Committee for Classification of Corneal Dystrophies (IC3D) was created and developed a new classification system which encompassed the clinical, histological and genetic data for each corneal dystrophy. Under this classification system, all corneal dystrophies were assigned a category from 1 – 4 on the basis of the level of supporting genetic evidence for the existence of the dystrophy. Category 4 is given to descriptions of

suspected new corneal dystrophies with little supporting clinical, histological or genetic evidence. Category 3 is for clinically well described corneal dystrophies without a mapped genetic locus. Category 2 dystrophies have been clinically well described and have mapped to a specific genetic locus. Category 1 corneal dystrophies are well characterized and the specific gene which causes the dystrophy has been identified. The term degeneration is used in this classification system to refer to a non-inherited corneal loss of transparency, in contrast to an inherited corneal dystrophy (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015). A summary of corneal dystrophies is shown in Table 1.1.

Table 1.1 Summary of corneal dystrophies recognised by the International Classification of Corneal Dystrophies (IC3D) and their genetic causes. Syndromes associated with corneal opacity that are not considered to be corneal dystrophies are not included. C: category. 1 – genetically solved. 2 – locus mapped. 3 – clinically described. 4 – limited clinical description. AD: autosomal dominant. AR: autosomal recessive.

Clinical Diagnosis (IC3D)	Alternative or historical names	Anatomical classification	Abbr.	Inher.	C	Gene(s) or loci	Ref.
Meesmann corneal dystrophy	Juvenile epithelial corneal dystrophy, Stocker-Holt corneal dystrophy	Epithelial and Sub epithelial Dystrophies	MECD	AD	1	*KRT3, KRT12*	(Irvine *et al.*, 1997)
Epithelial Recurrent Erosion Dystrophy	Smolandiensis dystrophia, Helsinglandica dystrophia, Francheshetti recurrent corneal erosion dystrophy	Epithelial and Sub epithelial Dystrophies	ERED	AD	3*	*COL17A1*	(Jonsson *et al.*, 2014)
Epithelial Basement Membrane Dystrophy	Cogan's microcystic dystrophy, Map-dot fingerprint dystrophy	Epithelial and Sub epithelial Dystrophies	EBMD	AD or Degenerative	1	*TGFBI*	(Boutboul *et al.*, 2006)
Gelatinous Drop Like Dystrophy	Lattice corneal dystrophy III	Epithelial and Sub epithelial Dystrophies	GDLD	AR	1	*TACSTD2*	(Tsujikawa *et al.*, 1999)
Lisch Epithelial Corneal Dystrophy	Band-shaped and whorled microcystic dystrophy of the corneal epithelium	Epithelial and Sub epithelial Dystrophies	LECD	XD	2	chrXp22.3	
Sub epithelial mucinous corneal dystrophy		Epithelial and Sub epithelial Dystrophies	SMCD		4		
Thiel-Behnke corneal dystrophy	Dystrophy of Bowman's layer II	Epithelial-Stromal *TGFBI* Dystrophies	TBCD	AD	1	*TGFBI*	(Munier *et al.*, 1997)

Reis-Bücklers corneal dystrophy	Granular corneal dystrophy III, Dystrophy of Bowman's layer I	Epithelial-Stromal *TGFBI* Dystrophies	RBCD	AD	1	*TGFBI*	(Munier *et al.*, 1997)
Lattice corneal dystrophy	Lattice corneal dystrophy I, IIIA, I/IIIA, Biber-Haab Dimmer Dystrophy	Epithelial-Stromal *TGFBI* Dystrophies	LCD	AD	1	*TGFBI*	(Munier *et al.*, 1997)
Granular corneal dystrophy I	Groenouw corneal dystsrophy, type I	Epithelial-Stromal *TGFBI* Dystrophies	GCD1	AD	1	*TGFBI*	(Munier *et al.*, 1997)
Granular corneal dystrophy II	Avellino corneal dystrophy	Epithelial-Stromal *TGFBI* Dystrophies	GCD2	AD	1	*TGFBI*	(Munier *et al.*, 1997)
Macular corneal dystrophy	Groenouw corneal dystrophy type II	Stromal Dystrophies	MCD	AR	1	*CHST6*	(Akama *et al.*, 2000)
Schnyder corneal dystrophy	Schnyder crystalline corneal dystrophy	Stromal Dystrophies	SCD	AD	1	*UBIAD1*	(Weiss *et al.*, 2007)
Congenital stromal corneal dystrophy		Stromal Dystrophies	CSCD	AD	1	*DCN*	(Bredrup *et al.*, 2005)
Fleck corneal dystrophy	Francois-Neetens fleck corneal dystrophy	Stromal Dystrophies	FCD	AD	1	*PIKFYVE*	(Li *et al.*, 2005)
Posterior amorphous corneal dystrophy		Stromal Dystrophies	PACD	AD	1	*KERA, LUM, DCN* and *EPYC*	(Kim *et al.*, 2014)
Central cloudy dystrophy of François		Stromal Dystrophies	CCDF	AD	4		
Pre-Descemet corneal dystrophy		Stromal Dystrophies	PDCD	AD or unknown	14	*STS*	(Yen *et al.*, 1987)

Posterior polymorphous corneal dystrophy		Descemet Membrane and Endothelial Dystrophies	PPCD	AD	1 2 3	*ZEB1* (PPCD3) *COL8A2* (PPCD2) chr20p (PPCD1)	(Biswas *et al.*, 2001; Krafchak *et al.*, 2005)
Congenital hereditary endothelial dystrophy		Descemet Membrane and Endothelial Dystrophies	CHED1, CHED2	AD (type I) AR (type II)	2 1	chr20p (CHED1) *SLC4A11* (CHED2)	(Vithana *et al.*, 2006)
X-linked endothelial dystrophy		Descemet Membrane and Endothelial Dystrophies	XECD	XD	2	chrXq25	
Fuchs endothelial corneal dystrophy		Descemet Membrane and Endothelial Dystrophies	FECD	AD Complex	1, 2 or 3	*COL8A2*	(Biswas *et al.*, 2001)

* Identification of gene published after most recent IC3D release.

1.9 Aims of project

In the past two decades, the application of sequencing technologies to investigate the genetic basis of corneal dystrophies has had a significant impact on our understanding and the classification of corneal dystrophies. Despite these advances in our understanding, the genetic basis for a number of corneal dystrophies remains unsolved. Some of these clinical diagnoses may be corneal degenerations (non-genetic), or expansions of the phenotypic spectrum of existing dystrophies. However, there is evidence to suggest that novel corneal dystrophy genes have yet to be identified. Furthermore, there is also evidence for further genetic heterogeneity for some genetically solved corneal dystrophies (Ren *et al.*, 2002). In addition to the potential to discover novel corneal dystrophy genes, systematic genotyping and phenotyping of patients with these rare inherited disorders will provide more information on the onset, progression, clinical variability and response to treatments for patients carrying specific mutations, thereby improving clinical diagnosis and the accuracy of prognosis. Corneal surgery is the only available treatment to restore visual acuity for the majority of these patients, and is associated with a range of potential complications, most commonly recurrence of the dystrophy. Identification

of the spectrum of mutations causing corneal dystrophies is the first step towards the development of gene-based targeted therapies for corneal dystrophies.

Moorfields Eye Hospital (MEH) is the largest ophthalmological center in the UK and has over 400 registered corneal dystrophy families; however, less than 5% have a genetic diagnosis for their condition. This cohort therefore represents an excellent resource to address the gaps in our understanding of the genetic causes of corneal dystrophies. The increasing accessibility and affordability of next-generation sequencing technologies including WES and WGS provide the tools with which to address these problems and resolve a number of issues in this field.

The aims of this research were:

- To identify the spectrum of mutations in the MEH corneal dystrophy cohort.
- To identify novel phenotype – genotype correlations.
- To identify novel corneal dystrophy genes.

Chapter 2: Materials and Methods

2.1 Patient recruitment and sample collection

The study followed the tenets of the Declaration of Helsinki and was approved by the research ethics committees (RECs) at MEH (REC reference 13/LO/1084).

2.1.1 Clinical diagnosis

Individuals with a clinical diagnosis of a monogenic corneal dystrophy (excluding FECD) were recruited from MEH and gave informed consent to enter the study. Clinical phenotyping was carried out by predominantly by Consultant Ophthalmic Surgeon Mr. Stephen Tuft. Clinical photos were obtained with a slit lamp using direct or indirect illumination, retroillumination or specular microscopy depending on the clinical appearance of the dystrophy. Recruited individuals were anonymised and provided with a unique study number and pedigree number on entering the study. When possible, a family history was taken and a pedigree drawn. Relatives of the proband were examined at a later date, and additional samples collected, if required.

2.1.2 Clinical histology

In some cases, histological confirmation of a clinical diagnosis was obtained following corneal transplant surgery. Histopathological analysis was performed by Dr. Caroline Thaung using standard methods for diagnosis of corneal dystrophies including H&E staining, PAS staining and Congo red staining.

2.1.3 DNA extraction from blood

Blood samples were obtained from the proband, and relatives if available, for genetic analysis. DNA was extracted from peripheral leukocytes by Beverly Scott as part of the UCL Institute of Ophthalmology DNA extraction service.

2.1.4 DNA extraction from saliva

In some cases, obtaining a blood sample from the proband or their relatives was not feasible due to distance or accessibility to MEH, therefore an Oragene DNA OG-500 saliva collection kit (DNA Genotek) was posted to the individual. Upon return of the sample, DNA was extracted from 4 mL total volume of saliva and stabilisation reagent, following the recommended ethanol precipitation DNA extraction protocol.

2.1.5 Obtaining DNA samples from additional sites

DNA samples and clinical images from corneal dystrophy patients were also obtained from other collaborative centres and clinicians, including patient samples from Oxford, the Czech Republic, New Zealand and Colombia.

2.2 Sanger sequencing

2.2.1 Primer design

Primer sequences were obtained from the published literature or designed specifically for the reaction in question. Primers were designed using Primer3 (http://primer3.ut.ee/) (Untergasser *et al.*, 2012) with target DNA sequence obtained from Ensembl Genome Browser (http://www.ensembl.org/index.html) or UCSC Genome Browser (http://genome.ucsc.edu/). Primers were designed to flank the genomic region of interest by at least ~ 100 bp. Primers were designed to anneal to regions with no known SNPs, however where this was unavoidable it was ensured that any known SNPs were rare (minor allele frequency (MAF) $\leq$ 0.005) and therefore less likely to interfere with primer binding. In silico PCR analysis (http://genome.ucsc.edu/cgi-bin/hgPcr?command=start) confirmed the amplification of the correct specific DNA region and predicted product size. Unmodified oligonucleotides were purchased from Sigma-Aldrich.

2.2.2 Polymerase chain reaction

PCR amplification was carried out predominantly using GoTaq Green Master Mix (Promega). PCR reactions were carried out in a 12.5 µL volume, with reagents shown in Table 2.1. Thermal cycling parameters are shown in Table 2.2. Each PCR reaction was initially tested using an annealing temperature of 60 ^{o}C using a control DNA sample. PCR reactions which did not produce the expected product of the predicted size were further optimised using an annealing temperature gradient to identify the optimal annealing temperature for the reaction. The extension time of each PCR was adjusted depending on the predicted size of the amplimer.

Table 2.1 PCR reagents for GoTaq® Green Master Mix. Reactions were carried out in 12.5 µL total volume.

Reagent	Volume (µL)	Final concentration
2X GoTaq ® Green Master Mix	6.25	1X
Forward primer (10µM)	0.5	0.4 µM
Reverse primer (10µM)	0.5	0.4 µM
Water	4.25	
DNA template (50-100 ng/µL)	1	
Total	12.5	

Table 2.2 Thermal cycler protocol for GoTaq® Green Mastermix PCR. The extension time at 72 °C was dependent on the size of the PCR amplimer. *Annealing temperature varied with specific primers used.

Temperature (°C)	Time	
95	2:00	
95	0:30	X35 cycles
Annealing temp*	0:30	
72	0:30 per kb	
72	5:00	

PCR reactions which failed to produce the desired amplimer using GoTaqGreen Master Mix were optimised using an alternative KAPA2G Robust™ (Kapa Biosystems) enzyme. The reagents used for the PCR are shown in Table 2.3 and thermal cycler parameters in Table 2.4. The annealing temperature was optimised for each reaction and extension time adjusted depending on the size of the amplimer.

Table 2.3 PCR regents for KAPA2G Robust™ PCR. Reactions were carried out in 12.5 µL total volume.

Reagent	Volume (µL)	Final concentration
2X KapaRobust ™	6.25	1X
Forward primer (10µM)	0.625	0.5 µM
Reverse primer (10µM)	0.625	0.5 µM
DMSO	0.625	
Water	3.375	
DNA template (50-100 ng/µL)	1	
Total	12.5	

Table 2.4 Thermal cycler protocol for KAPA2G Robust™ PCR. The extension time at 72 ºC was dependent on the size of the PCR amplimer. * Annealing temperature was dependent on primers used.

Temperature (°C)	Time	
95	5:00	
95	0:15	X35 cycles
Annealing temp*	0:15	
72	0:15 per kb	
72	10:00	

The primer sequences for all PCR reactions are included in Appendix A, with optimised annealing temperature and PCR enzyme used.

2.2.3 Agarose gel electrophoresis

A 4 µL aliquot from the total 12.5 µL reaction volume of the completed PCR amplification was electrophoresed through an agarose gel in Tris acetate EDTA (TAE) buffer. Agarose gel was made by dissolving a specific amount of agarose powder, depending on the percentage gel required, in 150 mL TAE buffer, and heating until the agarose disolved. SYBR ® Safe (Life Technologies) was added to the agarose gel (2.5 uL per 50 mL of solution). The mixture was allowed to cool and poured into a gel mold with a comb to produce wells for sample loading, and left until set. The percentage of the gel was adjusted using varying amounts of agarose powder, based on the estimated size of the amplimer, with a 2% agarose gel used for most PCR amplimers up to 1 kb in size and 1% agarose gels for DNA fragments greater than 1 kb in size. Following electrophoresis at 100 mV, DNA was visualized in the presence of UV light to confirm the presence of single PCR amplimer of the predicted size, by comparison to a molecular ladder containing fragments of a known size, most commonly a 0.1 kb – 10 kb 2-log DNA ladder (New England Biolabs).

2.2.4 PCR purification

2.2.4.1 PCR product purification

PCR reactions were purified using a MultiScreen® $PCR_{\mu 96}$ plate (Millipore) which uses a vacuum pump to draw water and small molecular weight molecules including unincorporated nucleotides and primers, through a filter while larger DNA molecules remain trapped on a membrane. The PCR reaction volume (8.5 µL) was diluted to a total volume of 100 µL, loaded into one of the wells of the 96-well plate, and filtered with a vacuum pump until the membrane was dry. 40 µL of water was added to each well and filtered to further wash the membrane. Finally, 25 µL of water was added to the well and shaken lightly for 10 minutes to re-suspend the PCR product.

2.2.4.2 Gel extraction and purification

If required, amplimers were visualised using a UV illuminator and excised from the agarose gel with a scalpel. Agarose gel sections containing DNA were purified using

QIAquick PCR purification kit according to the manufacturer's protocol (Qiagen). Briefly, the agarose gel is dissolved in a buffer and applied to a spin column. The protocol uses a silica-membrane-based purification process in which the DNA anneals to the membrane, is washed using a variety of buffers to remove impurities, and is then eluted in water.

2.2.5 Sanger sequencing

Sanger sequencing of purified PCR products or plasmids was outsourced to Source Bioscience Sequencing, in Cambridge. PCR amplimers were sequenced using BigDye terminator sequencing protocol analysed on an ABI PRISM 3100 Genetic Analyser (Applied Biosystems).

2.2.5.1 Data analysis

Sanger sequencing electropherograms were aligned and compared to a reference sequence using DNAstar package software version 8.0.2 (Lasergene). Any variants were annotated in accordance with the Ensemble transcript ID shown in Table 2.5. Variant annotation was checked using Mutalyzer (https://mutalyzer.nl/).

Table 2.5 Ensembl and RefSeq transcript IDs used for annotation of variants in corneal dystrophy genes. The following genes were screened by direct sequencing within the study.

Gene	Ensembl transcript ID	RefSeq transcript ID
TGFBI	ENST00000442011	NM_000358
GSN	ENST00000373818	NM_000177
CHST6	ENST00000332272	NM_021615
ZEB1	ENST00000320985	NM_030751
COL8A2	ENST00000397799	NM_005202
UBIAD1	ENST00000376810	NM_013319
TACSTD2	ENST00000371225	NM_002353
OVOL2	ENST00000278780	NM_021220
GRHL2	ENST00000251808	NM_001330593

2.3 Assessment of variant pathogenicity

Variants identified were assessed for their likely pathogenicity using a range of factors, outlined below.

2.3.1 Frequency

The frequency of any variants identified by Sanger sequencing was determined using the Exome Aggregation Consortium (ExAC) database (http://exac.broadinstitute.org/) which includes exome data from over 60,000 individuals, and data from the 1000 Genomes project (1000G; http://www.1000genomes.org/) and the Exome Variant Server (EVS; http://evs.gs.washington.edu/EVS/). Frequency of non-coding variants was additionally checked in Kaviar (http://db.systemsbiology.net/kaviar/), which includes data from ~ 13,000 whole genome sequences. Variants with a frequency of ≤ 0.005 were candidates for further investigation for potential pathogenicity.

2.3.2 Amino acid conservation

Missense mutations were assessed to determine whether the amino acid affected was conserved in orthologous sequences. Orthologous protein sequences were obtained from Ensembl and aligned to the human protein sequence using ClustalW, in BioEdit Sequence Alignment Editor (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Variants affecting conserved amino acids were considered more likely to be pathogenic.

2.3.3 Pathogenicity prediction tools

Missense variants were assessed for likely pathogenicity using pathogenicity prediction tools SIFT (http://sift.jcvi.org/www/SIFT_enst_submit.html) (Ng and Henikoff, 2001) and Polyphen2 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei *et al.*, 2010). SIFT uses orthologous sequence comparison to provide a prediction of likely pathogenicity from 1 – 0 with 0 being 'Damaging'. Polyphen2 uses both sequence comparison and 3-D structure prediction to provide an estimate for

pathogenicity from 0 – 1 where 1 is 'Probably Damaging'. The 'HumVar' score was used, which is recommended for identification of pathogenic mutations of Mendelian diseases (Adzhubei *et al.*, 2010).

2.3.4 Segregation analysis

If DNA samples were available from relatives of the proband, targeted sequencing of the candidate variant(s) was undertaken as previously described in section 2.2, to determine whether the variants co-segregated with the corneal phenotype in the pedigree and, for recessive disease, whether two variants were present in cis or trans.

2.3.5 Splice prediction tools

To predict the functional consequences of potential splice site mutations, including intronic variants and exonic variants located near to exon-intron boundaries, three splice site prediction tools were used: Human Splicing Finder (http://www.umd.be/HSF/) (Desmet *et al.*, 2009), NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html) (Reese *et al.*, 1997) and NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) (Hebsgaard *et al.*, 1996). These programs use algorithms to provide a score that predicts whether the likelihood that a nucleotide sequence is used as a donor or acceptor splice site.

2.3.6 Transcription factor binding predictions

ENCODE (Encyclopedia of DNA elements) data was manually interrogated using Integrative Genomics Viewer (IGV) for transcription factor binding in the genomic region of interest containing a candidate variant (The ENCODE Project Consortium, 2004). The effect of a variant on transcription factor binding was predicted by Alibaba 2.1 (http://gene-regulation.com/pub/programs/alibaba2/index.html) and MatInspector (Cartharius *et al.*, 2005) (http://www.genomatix.de/matinspector.htm). Alibaba 2.1 predicts transcription factor binding sites in an input nucleotide sequence using binding sites collected from TRANSFAC. MatInspector predicts transcription factor binding sites using a library of weight matrices.

2.4 *STS* deletion screening

To determine whether the *STS* gene was deleted, a PCR based approach was used. Failure to successfully PCR exons 2 and 9 of the *STS* gene (located on chrX) in the presence of successful amplification using control DNA demonstrated that the proband had a deletion encompassing the *STS* gene. Due to the presence of two copies in females, this screening protocol was restricted to use in males. Both PCR reactions were amplified using an annealing temperature of 60 ^{o}C using GoTaq ® Green Master Mix, with primers listed in Appendix A.

2.5 *CHST6* deletion and rearrangement screening

To screen for recurrent structural mutations located upstream of *CHST6*, a PCR based assay was used, as described in Akama *et al*., (2000). The assay consists of 4 PCR reactions (F1/R1, F2/R2, F2M/R2, F1M/R1) using a combination of 6 primers, listed in appendix A.

All reactions (F1/R1, F2/R2, F2M/R2, F1/R1M) were carried out using the GoTaq ® Green Master Mix protocol with an annealing temperature of 55 ^{o}C and an extension time of 60 seconds. The production of ~ 1 kb amplimers with the F1/R1 and F2/R2 primers was taken as confirmation that at least one allele was intact. Production of a ~ 1 kb amplimer with the F2M/R2 primers indicated the presence of a ~ 40kb deletion, and the production of amplimers in both F2M/R2 and F1/R1M indicated a potential ~ 2.5kb rearrangement occurred upstream of the *CHST6* gene.

2.6 Whole exome sequencing

WES was outsourced to Otogenetics. WES was performed using an Agilent human exome V4 or V5 capture (depending on availability) on an Illumina HiSeq2000 sequencer. Sequencing data was provided in the form of FastQ files.

2.6.1 Alignment of read data and variant calling

Data were aligned by Dr. Vincent Plagnol or Dr. Niko Pontikos to the human reference genome hg19 using Novoline. ANNOVAR was used to annotate variants. Variants were annotated with their frequencies in the 1000G, EVS and a UCL internal control exome database (exomes from different mutiethnic groups across UCL). WES data in the form of FastQ files were also obtained from probands with endothelial dystrophies from Dr. Petra Liskova and this data was aligned and analysed in an identical manner to that of MEH corneal dystrophy patients. Aligned reads were provided in .bam and .bai formats, visualised using IGV and as a list of all variants provided in .xls file format.

2.6.2 Variant filtering strategies of WES data

The variant filtering strategy used varied depending on likely inheritance pattern in the proband, the availability of WES data from additional family members, and the availability of WES data from other probands with a similar corneal phenotype. Rare variants were filtered at a frequency of ≤ 0.005. In some cases, variants were filtered for changes that were absent from control databases.

2.6.2.1 Filtering by shared rare variants in multiple probands

In cases where there were multiple probands affected with the same corneal phenotype, variants were filtered for the presence of rare variants (MAF ≤ 0.005) in genes shared between multiple affected individuals. This strategy was also combined with filtering strategies in 2.6.2.3 and 2.6.2.4.

2.6.2.2 Filtering by shared familial variants

In cases where multiple family members were analysed by WES, the filtering strategy used was to filter for rare variants (MAF ≤ 0.005 or absent from control databases) that were shared between affected family members and/or absent from unaffected family members. For recessive disease, filtering was restricted to homozygous or potential compound heterozygous variants. In cases where the

inheritance pattern is unknown, no filtering restrictions were applied to the type of variant. This strategy was also combined with filtering strategies in 2.6.2.3 and 2.6.2.4.

2.6.2.3 Filtering by corneal phenotype candidate genes

A complete list of all currently known corneal dystrophy genes is shown in Table 1.1. Additionally, a list of candidate genes was compiled from literature searches and the Online Mendelian Inheritance in Man (OMIM) database, that have been associated with syndromes that include a corneal phenotype; these genes and the corresponding syndrome are listed in Appendix B. Rare variants (MAF $\leq$ 0.005) were filtered for variants present in either Table 1.1 or Appendix B. The corneal phenotype of the proband was compared to the corneal phenotype associated with the gene, and the variant was assessed for pathogenicity, as described previously in section 2.3. In cases where the corneal phenotype was consistent with published cases, and the variant was plausibly pathogenic, additional clinical assessment was carried out.

2.6.2.4 Filtering by corneal proteome candidate genes

In order to identify variants in novel genes, not previously associated with a corneal dystrophy or syndromic corneal phenotype, a list of candidate genes was compiled from the corneal proteome (Dyrlund *et al.*, 2012). The list of proteins identified in the corneal proteome was converted from Uniprot protein ID to their genes names, using Biomart (http://www.ensembl.org/biomart/martview), and is listed in Appendix C. Rare variants (MAF $\leq$ 0.005) were filtered for variants present in these genes.

2.6.3 Exome depth

Prediction of putative copy number variations (CNVs) from WES data was performed using ExomeDepth, an algorithm which utilises read depth information to predict the occurrence of CNVs (Plagnol *et al.*, 2012).

2.7 SNP array genotyping for copy number variations

Two affected individuals from F131 (III:1 and IV:1) were analysed by HumanCytoSNP-12 v2.1 BeadChip (Illumina) genotyping array containing ~ 300,000 markers, with a mean genome wide spacing of ~ 10 kb. Czech PPCD probands from families F132, F144, F145, F146 and one British proband (F133) were analysed using a HumanOmniExpress BeadChip (Illumina) genotyping array containing ~ 730,000 markers with a mean genome wide spacing of ~ 4 kb.

2.8 *ZEB1* deletion quantitative PCR assay

2.8.1 Primer design

For verification of putative *ZEB1* deletions and screening for *ZEB1* CNVs in unsolved PPCD patients, a qPCR assay was designed using genomic DNA as a template. Primers were designed to target both the 5' and 3' ends of the *ZEB1* gene. An additional primer pair was designed and used at the 3' end of the *ZEB1* gene, to verify the presence of spurious results obtained with the original 3' primers. For all qPCR reactions, *TWIST1* was used as a reference gene.

2.8.2 qPCR reactions

All qPCR reactions were carried out using labTAQ^{TM} Green Hi Rox (LabTech International) master mix on an ABI 7900HT Fast Real-Time PCR system (Applied Biosystems) following the manufacturer's protocol. The reagents used are shown in Table 2.6, using cycling parameters shown in Table 2.7. Each experimental plate was set up to create a standard curve in duplicate, with an unaffected control DNA sample for both primers used. Each reaction was carried out in triplicate for each sample, including a positive control, an unaffected negative control DNA, and a non-template control on each plate. Melt curve analysis of the qPCR product was used to ensure the amplification of a single product.

Table 2.6 PCR reagents used for qPCR. A total reaction volume of 20 µL was used. Each individual reaction was carried out in triplicate and prepared in a 96-well plate.

Reagent	Volume (µL)	Final concentration
2X LabTech Green HI ROX	10	1X
Forward primer (10µM)	0.8	0.4 µM
Reverse primer (10µM)	0.8	0.4 µM
Water	4.4	
DNA template (50-100 ng/µL)	4	
Total	20	

Table 2.7 Thermal cycler protocol used for qPCR. RT-qPCR was run on an ABI 7900HT Fast Real-Time PCR system.

Temperature (°C)	Time	
95	3:00	
95	0:05	X35 cycles
65	0:20	

+ melt curve analysis

2.8.3 qPCR data analysis

Data analysis was performed using the relative standard curve method using unaffected control DNA. Briefly, the average Ct value of the three technical replicates was converted to the amount of DNA, using interpolation from a standard curve using unaffected control DNA of known concentration. Interpolated levels of *ZEB1* were normalised to *TWIST1* levels to determine copy number: a ratio of 0.5 was considered to correspond to a copy number of 1, and a ratio of 1.0 to a copy number of 2.

2.9 *ZEB1* deletion breakpoint mapping

Predicted co-ordinates of putative *ZEB1* deletions in PPCD probands were obtained from SNP array analysis. Primers were designed to anneal to the genomic co-

ordinates flanking the predicted deletion. GoTaq ® Long Master Mix (Promega) was used to amplify across the deletion break point using the same protocol used for GoTaqGreen ® shown in Table 2.1 and thermal cycling parameters in Table 2.2. However, extension times of up to 20 minutes were used, dependent on the maximum estimated size of the amplimer. Primers to amplify across the break points in the proband of F132 were used with GoTaq ® Long Master Mix at annealing temperature of 68 ^{o}C and an extension time of 6 minutes to amplify a specific product.

Genotyping of SNPs around the breakpoint for the proband of F131 and affected relatives was carried out using amplification of potentially polymorphic SNPs. All reactions were carried out using GoTaq ® Green Master Mix at 60 ^{o}C. Subsequently, primers were designed to flank the predicted deletion in F131 and amplified with GoTaq ® Long Master Mix at annealing temperature of 66C and an extension time of 3 minutes per cycle.

2.10 Linkage analysis

Linkage analysis was performed for one corneal dystrophy pedigree (F144). Nine affected (VI:2, VI:4, VII:1, VIII:1, VIII:3, VII:7, IX:1, IX:3, IX:6) and seven (VII:2, VII:3, VIII:2, VIII:4, IX:2, IX:4, IX:5) unaffected samples were genotyped for ~2,372,784 genome-wide SNPs using an Omni2.5 Exome-8 chip (Illumina), outsourced to The Center for Applied Genomics (The Hospital for Sick Children, Toronto). Linkage analysis was performed by Dr. Niko Pontikos using MERLIN.

2.11 Whole genome sequencing

Selected DNA samples were analysed by WGS using a TruSeq Nano DNA library preparation kit on a HiSeq X Ten sequencer (Illumina), outsourced to Macrogen.

2.11.1 Alignment of read data and variant calling

Generated reads were aligned to the human reference genome hg38 by Dr. Vincent Plagnol or Dr. Nikolas Pontikos. Variant calling was performed with GATK. All coding and non-coding variants were annotated with allele frequencies from 1000G.

2.11.2 Variant filtering of WGS data

Filtering was restricted to the genetic locus previously identified by linkage analysis (see section 2.10).

In F140, WGS was performed on two distantly affected (VII:3, VI:24) and one unaffected individual (VI:22). Variants were filtered for rare (MAF $\leq$ 0.005) heterozygous variants within chr20.hg38:17,641,482 – 18,949,130 that were present in both affected individuals and absent from the unaffected individual. In F141.02, affected individual V:11 was analysed by WGS. Data was combined with previously generated targeted re-sequencing data and filtered in an identical manner to F140.

In F144, four distantly affected individuals (VI:7, VI:9, VII:8, VIII:1) from F144 were analysed by WGS. Variants were filtered for novel variants (absent from control database Kaviar and internal UCL controls) and located within chr8.hg38:100,821,039 – 119,725,923.

2.11.3 Variant effect predictor

Analysis of potential consequences of coding and non-coding variants was performed by variant effect predictor (VEP; http://www.ensembl.org/info/docs/tools/vep/index.html) which predicts effects of variants on a variety of transcripts and identifies potential regulatory regions in non-coding DNA.

2.12 Culturing corneal derived cells

2.12.1 Isolation and culturing stromal fibroblasts

Stromal fibroblasts were isolated from central graft tissue by dissection of tissue into blocks approximately 1mm x 1mm in size. Dissected tissue was aspirated in medium and seeded into a T25 flask. After a short period (~15 minutes) to allow blocks to adhere to the surface, media was removed and the flask incubated for 24 hours in the absence of medium. 2mL of Amniomax C-100 basal medium with Amniomax supplements (Invitrogen) was added after 24 hours and incubated until fibroblasts began to emerge. After passaging cells, the medium was changed to Dulbecco's Modified Eagle Medium (DMEM)-F12 medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin (P/S). Fibroblasts were cultured in DMEM-F12 medium (Life Technologies) supplemented with 10% FBS, 1% P/S.

2.12.2 Culturing limbal epithelial cells

3T3 mouse embryonic fibroblasts were cultured in DMEM + GlutaMAX^{TM} (Life Technologies) media supplemented with 10% adult bovine serum and 1X Antibiotic-Antimycotic solution (Life Technologies). 3T3 fibroblasts were growth arrested by incubating in media supplemented with 4 µg/mL mitomycin C for 2 hours.

Limbal rims were obtained from deceased donors from the MEH eye bank in Optisol media. Limbal rims were dissected into quarters and digested in 1.2 U/mL dispase for 2 hours in 37 ^{o}C, 5% CO_2. After digestion, the limbus was scraped into a small amount of media to release limbal epithelial cells and seeded into a flask containing growth arrested 3T3 cells seeded at a density of $2.4x10^4$ cells/cm^2.

Cells were cultured in corneal epithelial culture medium (CECM). CECM contained a 1:1 mix of DMEM/F-12 and DMEM + GlutaMAX^{TM}, 10% FBS, 0.4 µg/mL hydrocortisone (Sigma-Aldrich), 0.1 nM cholera toxin (Sigma-Aldrich), 0.18 nM adenine (Sigma-Aldrich), 2 nM Triiodo-L-Thyronine (Sigma-Aldrich), 5 µg/mL

transferrin (Sigma-Aldrich), 5 µg/mL Insulin (Sigma-Aldrich), 1X Antibiotic-Antimycotic solution (Life Technologies). 10 ng/mL epidermal growth factor (Life Technologies) was added as an additional supplement from the first medium change onwards.

Removal of the 3T3 feeder cells was achieved using 0.05% trypsin-EDTA, with subsequent treatment with 0.5% trypsin-EDTA required for removal of human limbal epithelial cells, allowing for the cells to be separated by differential trypsinisation.

2.12.3 Culturing immortalised human corneal epithelial cells

A spontaneously immortalised human corneal epithelial cell line (HCE-S) was obtained from Prof. Julie Daniels (Notara and Daniels, 2010). HCE-S cells were cultured in DMEM supplemented with 10% FBS and 1% P/S.

2.13 Reverse-transcription PCR

2.13.1 RNA extraction from cells

RNA was extracted from cells using the RNeasy Mini Kit (Qiagen) with QIAshredder for sample lysate homogenisation prior to extraction. This system uses a spin column based protocol for RNA extraction and includes a DNase digest step to remove contaminating DNA co-extracted with the RNA. Additionally, RNA derived from cultured corneal endothelial cells was obtained from Dr. Alice Davidson.

2.13.2 cDNA synthesis

Approximately 1 ug of RNA was reverse transcribed to generate (complementary DNA) cDNA using the Tetro cDNA Synthesis Kit (Bioline) using olido dT primers. Reaction mixture was prepared as outlined in Table 2.8. Samples were incubated at 45 ^{o}C for 2 hours, followed by 85 ^{o}C for 5 minutes to terminate the reaction.

Table 2.8 Reagents used for cDNA synthesis. Volume of RNA used varied by concentration of sample. Volume of RNA varied with concentration of sample. Water was added for a total reaction volume of 20 µL.

Reagent	Volume (µL)
Total RNA (~ 1 ug)	X
Oligo dT primers	1
5X RT buffer	4
Ribosafe RNAase inhibitor (10u /µL)	1
Tetro reverse transcriptase	1
Water	X
Total	20 µL

2.13.3 RT-PCR to investigate expression of candidate genes

Intron spanning primers were designed to determine expression of candidate genes in cDNA generated from corneal derived RNA. Reverse transcription PCR (RT-PCR) was carried out using GoTaq ® Green or KAPA2G Robust as previously described in section 2.2.2, and optimised for each specific RT-PCR reaction. Primers used are listed in appendix A.

2.14 Interrogation of RNA-sequencing data

2.14.1 Publicly available data

Publicly available data was obtained for three adult and two foetal (16- to 18-week-old) human corneal endothelial samples (accession E-GEOD-41616 from ArrayExpress) (Chen *et al.*, 2013). Furthermore, publicly available RNA-sequencing (RNA-seq) data was available for 4 corneal samples including basal limbal crypts, superficial limbal crypts, the corneal limbal stroma and the central corneal epithelium (accession E-MTAB-1498 from Array Express) (Bath *et al.*, 2013). RNA-seq data for whole corneas was obtained from an in-house experiment including 6 full thickness corneal buttons from deceased donors without a corneal dystrophy.

2.14.2 Alignment of RNA-seq data

Alignment of RNA-seq data was performed by Dr. Vincent Plagnol. Reads were aligned to the hg38 human reference genome with Star v.2.5.0. Duplicate reads were marked with Picard MarkDuplicates v.1.100. Raw read counts, excluding duplicate reads, were generated with DEXSeq python scripts (dexseq_count.py). The resulting counts were normalized according to the length of each feature (estimated with the Rsu- bread package) and a library-size factor as estimated by the DEXSeq tool. Gene annotations were based on Ensembl transcripts (GRCh38.82).

2.14.3 Interrogation of RNA-seq data

Aligned RNA-seq reads were interrogated using IGV to determine expression in corneal tissue of interest and to investigate corneal-specific transcript isoforms. Reads per kilobase million (RPKM) values were also generated.

2.15 Immunohistochemistry

Immunohistochemistry on was carried out on corneal tissue using standard methods by Dr. Caroline Thaung for an endothelial sample obtained from F148: II:3 (PPCD) and full thickness control samples (keratoconus). Details of antibodies used are shown in Table 2.9.

Table 2.9 Details of antibodies used for immunocytochemistry of corneal samples.

Targeted protein	Company and catalogue number	Concentration
GRHL2	Sigma-Aldrich HPA004820	1/100
CD56	Dako M7304	1/100
Vimentin	Dako M0725	1/3000
E-cadherin	Abcam ab11512	1/200
N-cadherin	Abcamab12221	1/500
CK7	Dako M7018	1/2000
CK8/18	Leica PA0067	Ready to use

Chapter 3: Anterior corneal dystrophies

3.1 Introduction

Anterior corneal dystrophies encompass those dystrophies which affect the corneal epithelium, Bowman's layer and/or the stroma. The majority of these corneal dystrophies are associated with a reduction of visual acuity due to an accumulation of opacities that obstruct the visual field, consisting of proteins, lipids and/or phospholipids. Some anterior dystrophies are also associated with frequent, painful recurrent erosions which occur when the epithelium becomes detached from the underlying basement membrane. Corneal scarring can occur after erosive episodes and may modify the clinical phenotype, thereby complicating the clinical diagnosis. Over ten genes have been shown to cause anterior corneal dystrophies, via a range of distinct molecular pathways. Anterior corneal dystrophies are classified by the IC3D into three groups: the epithelial-stromal *TGFBI* dystrophies, epithelial and sub-epithelial dystrophies, and stromal dystrophies, which are discussed in more detail below (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015).

3.1.1 Epithelial-Stromal TGFBI dystrophies

Epithelial-stromal *TGFBI* dystrophies is the collective name given to a range of phenotypically diverse corneal dystrophy diagnoses which are caused by dominant missense mutations in *TGFBI* (transforming growth factor β-induced). The *TGFBI* gene has 17 coding exons and two mutation hotspots at amino acid residues Arg-124 and Arg-555, located in exons 4 and 12, respectively. Six missense mutations have been described at these two amino acid residues, with the resulting clinical phenotype dependent on the specific amino acid residue substituted in place of the arginine (Munier *et al.*, 1997; Stewart *et al.*, 1999). These mutations have been described in a large number of families from different populations and have likely occurred independently on a number of occasions; however, some are founder mutations in specific populations (Korvatska *et al.*, 1998). Other rarer mutations in *TGFBI* have been described in exons 4, 11, 12, 13, 14 and 16, and cause a range of clinical phenotypes (Yamamoto *et al.*, 1998; Warren, Abbott, *et al.*, 2003; Tian *et*

al., 2005). Over 60 different pathogenic mutations in *TGFBI* have been described, to date.

3.1.1.1 Reis-Bücklers and Thiel-Behnke corneal dystrophy

Reis-Bücklers corneal dystrophy (RBCD) and Thiel-Behnke corneal dystrophy (TBCD) are dominantly inherited corneal dystrophies which primarily affect the epithelium and Bowman's layer of the cornea. They are associated with the deposition of fibrous protein deposits in the sub-epithelium and superficial stroma with an age of onset in first or second decade of life. Clinically, when observed by slit lamp, TBCD is associated with a honeycomb pattern of opacities whereas RBCD has a geographic pattern, although these cannot always be distinguished from clinical examination alone (Figure 3.1 A and D). Both of these corneal dystrophies are associated with the occurrence of recurrent erosions from childhood (Okada *et al.*, 1998). Histologically, RBCD is associated with variations in epithelial thickness and the replacement of Bowman's layer with sheet like eosinophilic opacities (Liang *et al.*, 2014). TBCD has a similar appearance with opacities located almost exclusively in Bowman's layer (Dighiero *et al.*, 2001) (Figure 3.1 B and E). RBCD and TBCD can be distinguished histologically by positive staining of RBCD deposits by Masson's trichrome stain, which is negative with TBCD opacities. Using electron microscopy (EM), RBCD protein deposits are observed as straight rod-shaped protein fibres whereas TBCD deposits consist of 'curly' fibres although this difference is rarely used in a diagnostic capacity (Küchle *et al.*, 1995) (Figure 3.1 C and F).

TBCD is caused exclusively by the c.1664G>A; p.(Arg555Gln) hot-spot mutation in exon 12 of the *TGFBI* gene, whereas RBCD is predominantly caused by the c.371G>T; p.(Arg124Leu) hot-spot mutation in exon 4 (Ridgway *et al.*, 2000). Two additional rarer mutations have been reported to cause a RBCD-like phenotype, c.1618_1620delTTT; p.(Phe540del) and c.1868G>A; p.(Gly623Asp) (Afshari *et al.*, 2001; Li *et al.*, 2008), although the latter has also been described as causative of other corneal dystrophy diagnoses (Aldave, Rayner, *et al.*, 2005; Auw-Haedrich *et al.*, 2009).

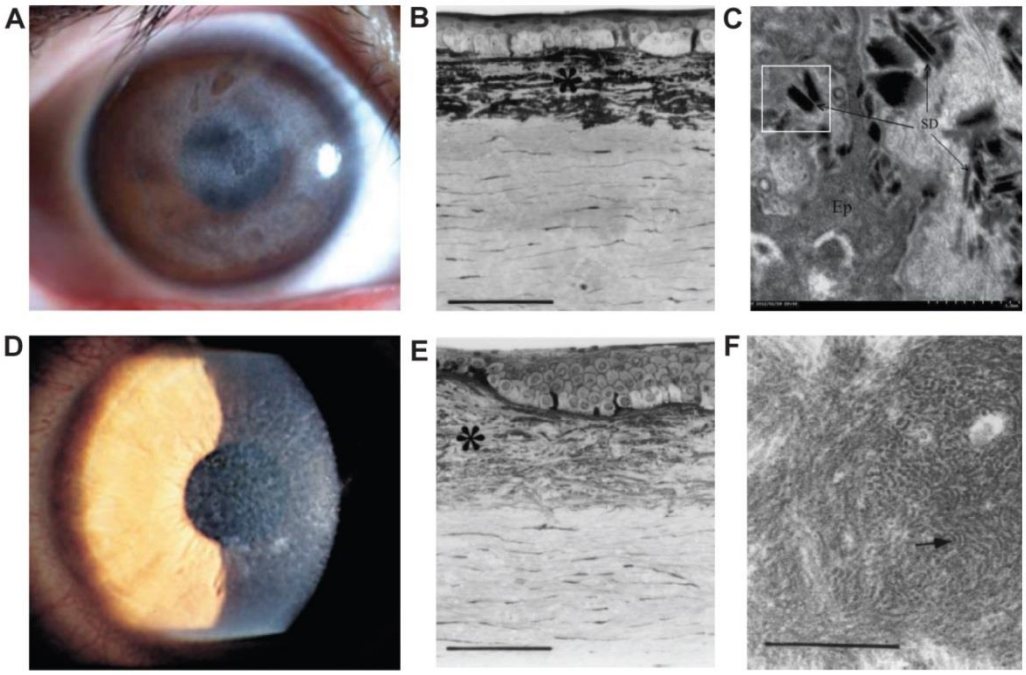


Figure 3.1 Clinical phenotype, histology and electron microscopy of Reis Bücklers corneal dystrophy (RBCD) and Thiel-Behnke corneal dystrophy (TBCD). A-C shows images of RBCD and D-F are TBCD. **(A)** Clinical appearance of RBCD with geographic-like opacities. **(B)** Histology of a RBCD corneal button showing opacities concentrated in Bowman's layer, indicated with a *, scale bar is 100 µm, X180 magnification. **(C)** Transmission electron micrograph showing rod shaped opacities, labelled with an arrow. Ep: epithelial cells, SD: sub-epithelial deposits. **(D)** Slit lamp clinical image of TBCD. **(E)** TBCD histology with opacities located in Bowman's layer, indicated with a *, scale bar is 100 µm, X180 magnification. **(F)** 'Curly' filaments characteristic of TBCD, indicated with an arrow, X40,600 magnification. Scale bar is 1µm. Images from (Dighiero *et al.*, 2001; Munier *et al.*, 2002; Liang *et al.*, 2014).

3.1.1.2 Lattice corneal dystrophy

Lattice corneal dystrophy (LCD) is characterised by a clinical appearance of lattice lines, located predominantly in the corneal stroma. Histologically, LCD is characterised by stromal amyloid deposition and therefore can be identified by Congo red staining. LCD shows considerable phenotypic variation and, historically, LCD clinical phenotypes were classified as type I, type IIIA, type IIIB, intermediate type I/IIIA or as an atypical phenotypic variant (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015). LCD type II is a not a true corneal dystrophy and is not caused by mutations in *TGFBI*. The most common clinical phenotype of LCD is called classic LCD, or LCDI. The age of onset of LCDI is typically in the first decade of life and

the clinical appearance is characterised by thin branching filaments forming a lattice pattern of opacities (Figure 3.2 A) (Lisch and Seitz, 2014). Lattice corneal dystrophy type IIIA was named due to its phenotypic similarity with the epithelial dystrophy, GDLD (section 3.1.2.3), which was historically named LCDIII. LCDIIIA is distinguished from GDLD by its autosomal dominant inheritance pattern, whereas GDLD is autosomal recessive. Lattice lines in LCDIIIA are thicker than LCDI, extend deeper into the stroma and extend to the periphery of the cornea, with a later age of onset, typically in the sixth decade or later (Yamamoto *et al.*, 1998) (Figure 3.2 D). Some LCD phenotypes are described as being intermediate I/IIIA type; this is characterised by the thicker lattice lines of IIIA, but with an age of onset in the second or third decade (Schmitt-Bernard *et al.*, 2000). Some patients with LCD cannot be categorised into any of the known lattice phenotypic subtypes due to unusual clinical features. These include amyloid deposits which penetrate deeper into the stroma than LCDI (Kawasaki *et al.*, 2011) or a unilateral or asymmetric presentation of lattice lines (Afshari *et al.*, 2008) (Figure 3.2 B). The IC3D categorises all of these as phenotypic variants of LCD (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015). In addition to variation in phenotypes associated with different mutations, there is intra- and interfamilial variation between individuals carrying the same mutation, indicating possible environmental effects and secondary changes, such as sub-epithelial scarring, as a result of recurrent corneal erosions (Tian *et al.*, 2007).

The most common cause of LCD is the c.370C>T; p.(Arg124Cys) mutation which occurs at a *TGFBI* mutation hotspot and is most commonly associated with the classic LCDI phenotype (Munier *et al.*, 1997). Mutations causing a range of phenotypic variants of the LCD phenotype have been described in exons 11, 12, 13 and 14 of the *TGFBI* gene (Schmitt-Bernard *et al.*, 2000; Munier *et al.*, 2002; Warren, Aldave, *et al.*, 2003).

3.1.1.3 Granular corneal dystrophy

Granular corneal dystrophy, type 1 (GCD1) is characterised by discrete grey-white opacities which may be oval, circular or ‘crumb-like’ in appearance when examined with the slit lamp (Figure 3.2 C). Histologically, GCD1 patients demonstrate

amorphous hyaline protein deposits located predominantly in the stroma. The opacities increase in frequency with age and merge as the disease progresses, causing a gradually increased visual impairment. The age of onset is typically in the first decade of life although diagnosis may be delayed until vision is affected. Approximately half of those affected with GCD1 experience recurrent corneal erosions (Møller and Ridgway, 1990).

Granular corneal dystrophy, type 2 (GCD2) was previously known as Avellino dystrophy. Beginning in childhood, discrete hyaline opacities are initially deposited in the sub-epithelium and anterior stroma, which increase in size as the disease progresses (Figure 3.2 D). This is followed by mild amyloid deposition in the mid to posterior stroma that can form few lattice lines. Clinically, this results in a 'snowflake'-like appearance that can eventually cause mid-stromal haze (Ferry, Benson and Weinberg, 1997; El-Ashry *et al.*, 2003). Individuals who are homozygous for the GCD2 causative mutation have also been described in certain populations (e.g. Korea) where the causative mutation is a founder mutation, resulting in an earlier disease onset, and more rapid disease progression (Moon *et al.*, 2007).

The most common genetic causes of GCD1 are the c.1663C>T; p.(Arg555Trp) or c.370C>A; p.(Arg124Ser) *TGFBI* hotspot mutations (Munier *et al.*, 1997; Stewart *et al.*, 1999). Other rarer mutations have also been described which cause a GCD1 phenotype (Paliwal *et al.*, 2010). GCD2 is almost exclusively caused by a c.371G>A; p.(Arg124His) mutation in *TGFBI* and was considered to be a founder mutation originating in the Avellino region of Italy. However, this mutation has subsequently been identified in individuals with no known ancestry to this region in a variety of populations, indicating it is also a mutation hot-spot (Munier *et al.*, 1997). Another missense mutation, c.1856T>A; p.(Met619Lys), has also been described as causative of GCD2 but has only been identified in a single family, to date (Aldave *et al.*, 2008).

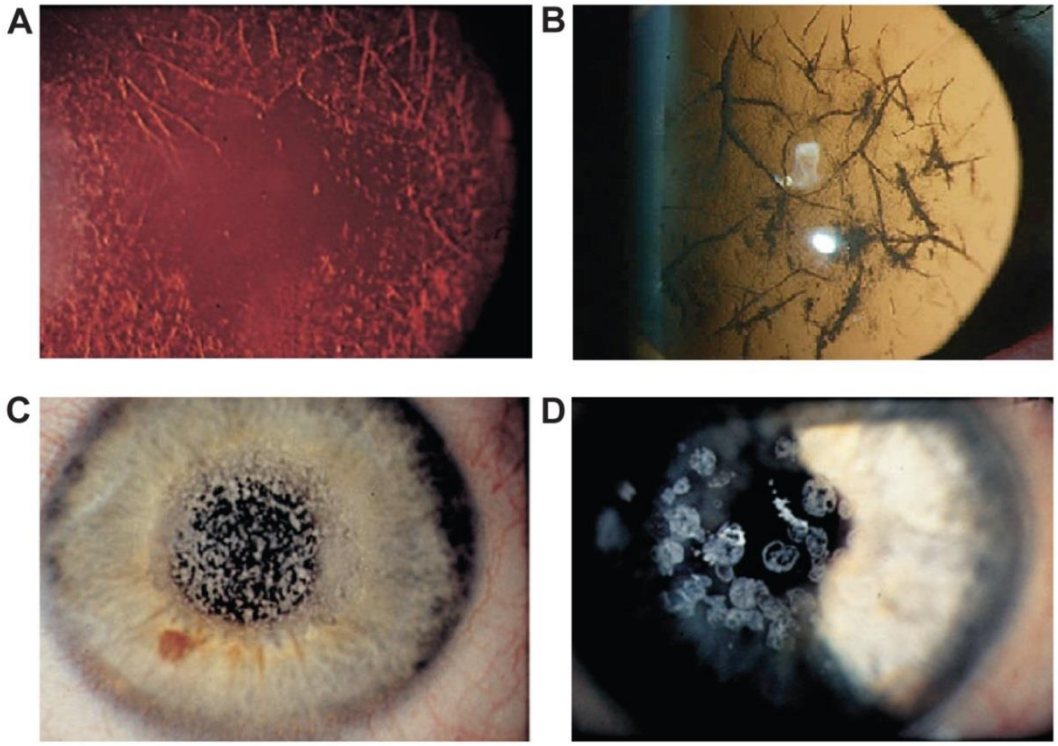


Figure 3.2 Phenotypic diversity of *TGFBI*-associated stromal dystrophies; lattice corneal dystrophy (LCD), granular type 1 (GCD1) and granular type 2 (GCD2). (A) Classic LCD appearance in retroillumination with thin branching lattice lines. **(B)** Phenotypic variant of LCD in retroillumination showing thicker, more prominent, lattice lines. **(C)** Typical clinical appearance of GCD1 showing granular 'crumb-like' opacities. **(D)** Typical appearance of GCD2. Images A, C, D from (Munier *et al.*, 2002) and B from (Zenteno *et al.*, 2009).

3.1.1.4 TGFBI pathogenesis

TGFBI encodes a 683 amino acid protein, also known as keratoepithelin. It is a secreted extracellular matrix protein of unknown function although it may play a role in cell adhesion and migration (Han *et al.*, 2015). It consists primarily of four consecutive fasciclin-1 (FAS1) domains. A secretory signal is located at the N-terminal end of the protein, followed by an N-terminal cysteine rich (EMI) domain. At the C-terminal region there are a number of potential integrin binding motifs including NKDIL, YH18, EPDIM, and RGD (Thapa, Lee and Kim, 2007). The two hot-spot mutation residues, Arg-124 and Arg-555, associated with corneal dystrophy are located in the first and forth FAS1 domain respectively. Other reported mutations cluster in exons 12, 13 and 14 of the *TGFBI* gene which encode the fourth FAS1 domain.

There is evidence to suggest that the wild type TGFBI protein has inherent amyloid forming ability; TGFBI amyloid has been found in corneas with secondary amyloidosis due to previous trauma or infection, rather than due to a primary *TGFBI* mutation (Suesskind *et al.*, 2006). The native state of the TGFBI protein is enriched with alpha helical structures, however when denatured the protein predominantly consists of beta-sheets, promoting the aggregation of the TGFBI protein into oligomeric precursors of amyloid fibrils (Grothe *et al.*, 2009). Furthermore, the fourth FAS1 domain contains an 18-residue peptide (a.a. 515-532) that has been demonstrated to readily form extensive amyloidogenic fibrils *in vitro* under physiological conditions, in addition to non-fibrillar aggregates (Yuan, Berscheit and Huang, 2007).

Mutations located in the fourth FAS1 domain affect TGFBI protein stability. The GCD1-associated p.(Arg555Trp) mutation protein has increased thermodynamic stability and is less susceptible to proteolysis compared to wild type protein, whereas, in contrast, the LCD-associated mutation p.(Ala546Thr) has reduced stability and is more susceptible to proteolysis (Runager *et al.*, 2011; Underhaug *et al.*, 2013). Less stable forms of TGFBI are susceptible to proteolytic digestion which releases a fibril-core (a.a. 571-588) with amyloid forming ability (Sorensen *et al.*, 2015).

3.1.2 Epithelial and sub-epithelial dystrophies

3.1.2.1 Epithelial basement membrane dystrophy

Although data is lacking on the prevalence of different corneal dystrophy diagnoses, epithelial basement membrane dystrophy (EBMD) is typically considered to be the most common anterior corneal dystrophy (Laibson, 2010). The age of onset can be as early as childhood and for most the onset is before the age of 30 years (Boutboul *et al.*, 2006). EBMD (previously known as map-dot-fingerprint dystrophy) is associated with a wide spectrum of clinical changes involving the epithelial basement membrane. These include lines, which may be reminiscent of fingerprints or maps, microcyst-like opacities and irregular shaped oval grey-white cysts (Figure 3.3). If cysts are the predominant feature the condition then the diagnosis is typically

called Cogan's dystrophy, although this is considered to be part of the EBMD spectrum by the IC3D. Histologically, the oval shaped opacities are accumulations of dead or dying cells (Laibson, 1976). Some patients may also have blebs, which are sub-epithelial excrescences of the epithelial basement membrane encroaching into the basal epithelial cell layer. In EBMD corneal tissue, both the epithelial basement membrane and the blebs, if present, stain positively with PAS stain, indicative of glycoprotein deposition (Dark, 1977). Most EBMD patients are asymptomatic with no significant reduction of visual acuity, however a subset experience painful recurrent epithelial erosion or reduced vision from irregular corneal astigmatism.

The majority of EBMD cases are thought to be degenerative (non-inherited), however a subset of individuals have a family history of the disorder with an apparently autosomal dominant inheritance patterns (Laibson and Krachmer, 1975). In a prior study, two potentially pathogenic heterozygous *TGFBI* mutations were identified in both familial and sporadic cases of EBMD, and the authors estimate that *TGFBI* mutations may account for ~ 10% of EBMD cases (Boutboul *et al.*, 2006). No studies have replicated these findings and, despite being one of the most common dystrophies, no additional mutations in *TGFBI* have been identified in EBMD cases. Therefore, the contribution of *TGFBI* mutations to EBMD, and the potential role of additional genetic heterogeneity, remains unresolved.

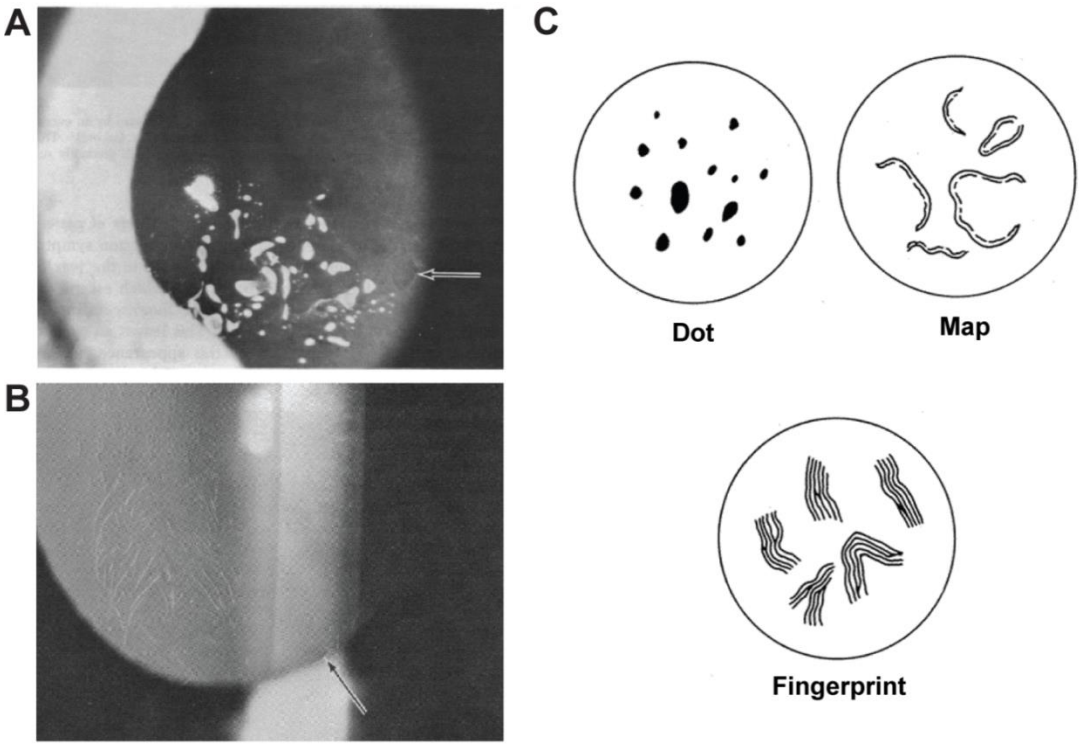


Figure 3.3 Diverse clinical appearance associated with epithelial basement membrane dystrophy (EBMD). (A) Clinical appearance of EBMD showing numerous dot-like or microcystic changes. White opacities have irregular size and shape. The arrow shows a map-like lesion. **(B)** Clinical image in retroillumination shows fingerprint lines. Images from (Laibson, 1976). **(C)** Illustration of map, dot, and fingerprint opacities. Image from (Perry and Cameron, 2006).

3.1.2.2 Epithelial recurrent erosion dystrophy

ERED is the collective name for a range of dystrophies with overlapping clinical phenotypes which have been previously described as dystrophia Smolandiensis, dystrophia Helsinglandica and Franceschetti hereditary recurrent corneal erosion dystrophy. All are associated with the occurrence of frequent corneal epithelial erosions from early childhood that become less frequent from the early twenties onwards. A proportion of individuals with ERED develop secondary opacities later in life which are thought to be a consequence of the frequent erosions (Figure 3.4). Unlike other anterior corneal dystrophies, there are no primary opacities associated with ERED, which may confound its clinical diagnosis (Hammar *et al.*, 2008, 2009; Lisch *et al.*, 2012).

Recurrent epithelial erosions can occur as a consequence of environmental factors, including a previous epithelial abrasion or infection, but are also a feature of other anterior corneal dystrophies, including RBCD, TBCD, GCD, LCD and EBMD (Küchle *et al.*, 1995; Tian *et al.*, 2005; Laibson, 2010; K. Han *et al.*, 2013). ERED is distinguished from environmental causes of recurrent epithelial erosions by an autosomal dominant pattern of inheritance. Mutations in *TGFBI* were excluded in some ERED families, either by direct sequencing or linkage analysis, indicating a different genetic cause (Hammar *et al.*, 2008, 2009; Lisch *et al.*, 2012). A heterozygous missense mutation [c.2816C>T; p.(Thr939Ile)] in *COL17A1* has been shown to cause ERED in a large Swedish pedigree. Furthermore, re-analysis of a previously published synonymous variant in *COL17A1* [c.3156C>T; p.(Gly1052Gly)] in a pedigree clinically diagnosed with TBCD, but clinically consistent with ERED, was shown to affect splicing and cause an in-frame deletion (Jonsson *et al.*, 2014). This splice mutation has been subsequently shown to be causative in multiple further pedigrees with ERED (Oliver *et al.*, 2016). Interestingly, biallelic mutations in *COL17A1* also cause the recessive skin blistering disorder epidermolysis bullosa, and corneal erosions are a common feature of this condition (Fine *et al.*, 2004). Whether the other described hereditary recurrent erosion phenotypes are also caused by mutations in *COL17A1*, is currently unresolved. They could represent a phenotypic spectrum of the same genetic disorder or alternatively, there could be additional locus heterogeneity for ERED.

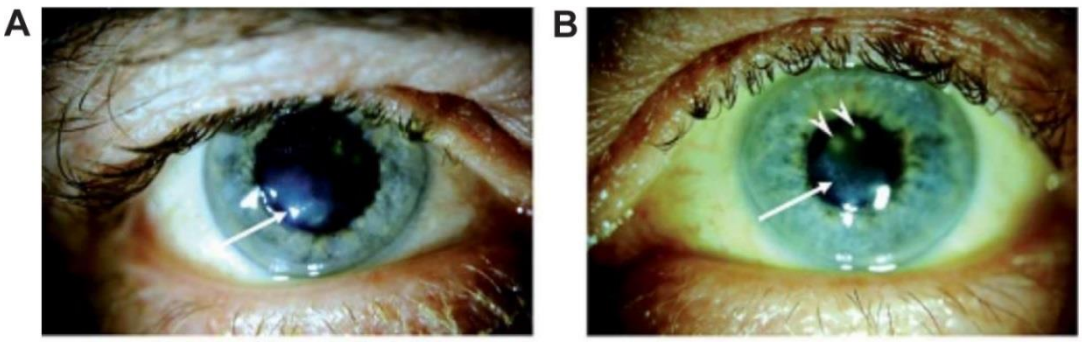


Figure 3.4 Clinical appearance and secondary opacities associated with epithelial recurrent erosion dystrophy (ERED). There are no primary opacities associated with ERED, however secondary opacities can occur due to frequent recurrent erosions. **(A)** Arrow indicates irregular surface caused by sub-epithelial opacity. **(B)** Sub-epithelial nodular opacities visible (arrow) Images from (Jonsson *et al.*, 2014).

3.1.2.3 Gelatinous drop-like dystrophy

GDLD is a rare autosomal recessive corneal dystrophy associated with the deposition of amyloid in the epithelial and sub-epithelial layers of the cornea. Symptoms of GDLD occur in the first decade of life and include photophobia, tearing, blurred vision and foreign body sensation. Typically, GDLD is associated with the appearance of nodules in the central cornea which coalesce to form a mulberry appearance (Figure 3.5 D). The advanced stages of the dystrophy are associated with neovascularisation and severely impaired vision (Kawasaki *et al.*, 2011). However, less common phenotypic manifestations have been described including band keratopathy type (Figure 3.5 A), stromal opacity type (Figure 3.5 B) and kumquat-like type (Figure 3.5 C). Clinical phenotypes have been shown to differ between individuals carrying the same pathogenic mutation and between the eyes of a single individual, indicating that environmental factors may influence the disease phenotype. It has also been proposed that the sub-types may represent different stages in disease progression (Ide *et al.*, 2004).

GDLD is caused by biallelic mutations in the *TACSTD2* (tumour associated calcium signal transducer 2) gene (Tsujikawa *et al.*, 1999). *TACSTD2* has a single coding exon and is located at chr1p32.1. *TACSTD2* encodes a single pass transmembrane glycoprotein involved in calcium signaling (Ripani *et al.*, 1998). To date, over 30 pathogenic mutations have been described, however in some cases there are suggestions of an additional disease causing locus due to the absence of mutations in *TACSTD2* (Ren *et al.*, 2002; Akhtar *et al.*, 2005; Alavi *et al.*, 2007). GDLD occurs most prevalently in Japan with an estimated incidence of 1 in 300,000 (Kawano, Fujiki and Kanai, 1992). This is due to the presence of a founder mutation [c.352C>T; p.(Gln118*)] in the population, that is responsible for an estimated 90% of GDLD cases in Japan (Tsujikawa, Tsujikawa and Maeda, 2000).

The epithelial barrier function is compromised in GDLD. Altered sub-cellular localisation of claudin-1, caludin-7, ZO-1 and occludin proteins, which are required for the formation of tight junctions, results in abnormal morphology of epithelial cell junctions and ultimately, increased epithelial permeability (Kinoshita *et al.*, 2000; Nakatsukasa *et al.*, 2010). Sub-epithelial amyloid deposits contain lactoferrin, the

major protein in tear fluid, indicating that the permeation of tear fluid through the corneal epithelium is the source of the amyloidogenesis (Klintworth *et al.*, 1997). Advanced glycation end products and D-amino acid containing proteins also co-localise with amyloid fibrils in GDLD corneas (Kaji *et al.*, 2012).

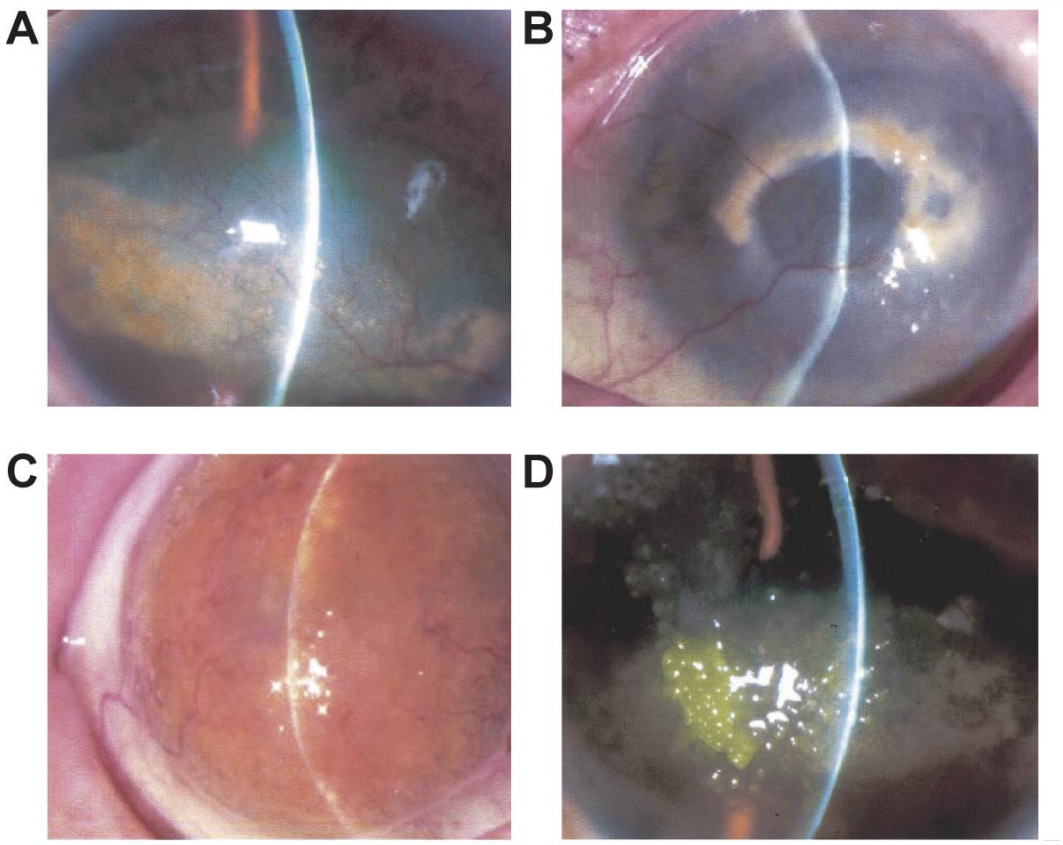


Figure 3.5 Phenotypic spectrum of gelatinous drop-like dystrophy (GDLD). Four clinical sub-types of GDLD have been identified. **(A)** Band keratopathy sub-type showing a yellow-opaque band across the corneal surface. **(B)** Stromal opacity sub-type showing yellow-white lesions in central cornea. **(C)** Kumquat-like sub-type. **(D)** Typical mulberry sub-type with cluster of raised nodules. Images from (Ide *et al.*, 2004).

3.1.2.4 Meesmann corneal dystrophy

MECD is an autosomal dominant corneal dystrophy affecting the epithelium. Large numbers of intra-epithelial cysts develop in the first few years of life (Figure 3.6 A); however, the disorder is usually asymptomatic or causes only mild symptoms of photophobia, irritation and/or foreign body sensation. Some patients experience painful episodes due to eruption of cysts on the epithelial surface. It is non-progressive and generally does not cause severe visual impairment although

recurrent erosions can result in scarring (Ehlers *et al.*, 2008). There is some phenotypic variation, as demonstrated by a recently identified c.395T>C; p.(Leu132Pro) mutation in *KRT12* which is associated with a severe phenotype including a large numbers of cysts present from an early age with amblyopia, sub-epithelial fibrosis and superficial vascularisation (Hassan *et al.*, 2013) (Figure 3.6 B). Histologically, there are variations in the thickness of the epithelial cell layer with microcysts found primarily in the superficial third of the epithelium. Histological staining shows that the microcysts contain debris which stain positive for PAS and alcian blue, indicative of glycoprotein, with a characteristic linear deposit beneath the epithelial layer (Burns, 1968).

MECD is caused by heterozygous mutations in the *KRT3* (keratin 3) or *KRT12* (keratin 12) genes; the predominant keratin pair expressed in the mature corneal epithelium (Irvine *et al.*, 1997). There is also limited evidence to suggest an additional disease causing locus for MECD (Cao *et al.*, 2013). Keratins are structural proteins that form intermediate filaments of the cytoskeleton in epithelial cells. Epithelial cytokeratins are either acidic (type I; KRT1-8) or basic (type II; KRT9-20) and pairs of keratins from each type form an obligate heterodimer. The heterodimers polymerise to form intermediate filaments, which are approximately 10 nm in diameter (Corden and McLean, 1996). Keratin proteins share a common structure: they consist of a central α-helical rod domain that is subdivided into 4 coiled-coil domains (1A, 1B, 2A, 2B), joined by 3 linker regions. The rod domain is flanked by non-helical head and tail domains. The first 20 amino acids of the 1A domain consist of the helix initiation motif, and the terminal 20 amino acids of the 2B domain make up the helix termination motif. These motifs differ between the type I and type II keratins but are highly conserved within their keratin sub-types (Irvine and McLean, 1999). Pathogenic mutations causing MECD in *KRT12* occur within the helix initiation motif and helix termination motif, encoded by exons 1 and 6 (Yoon *et al.*, 2004) whereas in *KRT3*, mutations identified are found exclusively in the helix termination motif encoded by exon 7 (Szaflik *et al.*, 2008). Disruption of the epithelial cell intermediate filament cytoskeleton causes fragility of the epithelial cells resulting in detachment, blistering or breakage of skin, hair, nails or cornea, depending on the expression pattern of the affected keratin protein (Corden and McLean, 1996).

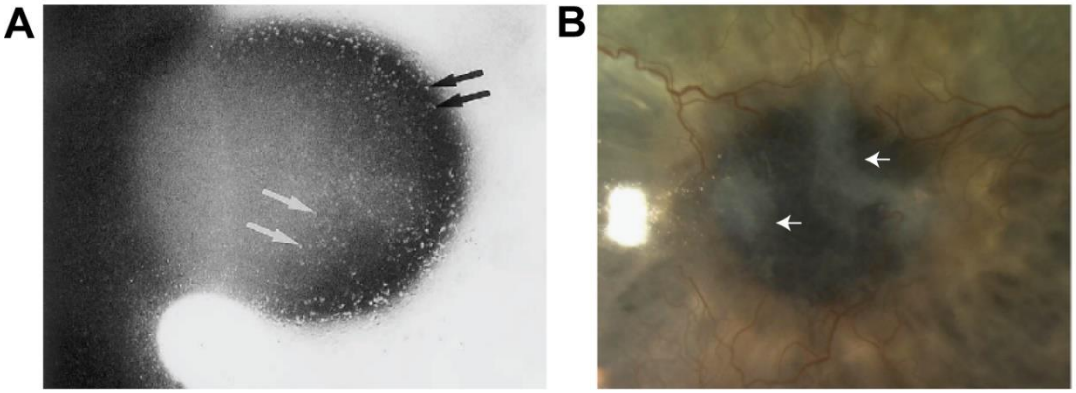


Figure 3.6 Clinical appearance of Meesmann corneal dystrophy (MECD). (A) Typical appearance of MECD showing presence of numerous microcysts, labelled with arrows. Image from (Nishida *et al.*, 1997). **(B)** Severe MECD clinical phenotype with superficial vascularisation, and a sub-epithelial scar (arrowheads). Image from (Hassan *et al.*, 2013).

3.1.3 Stromal corneal dystrophies

3.1.3.1 Macular corneal dystrophy

MCD is an autosomal recessive corneal dystrophy, predominantly affecting the stroma. The clinical appearance consists of diffuse grey-white opacities (macules) which can coalesce to produce a cloudy opaque cornea that causes increasing levels of visual impairment and with an age of onset in the first or second decade of life (Figure 3.7 A). The opacities may also cause photophobia and irritation (Jonasson *et al.*, 1989). MCD is associated with irregular corneal thinning and steepening of the anterior corneal surface (Dudakova *et al.*, 2014). Histologically, stromal cells contain numerous vacuoles containing glycosaminoglycans originating from the endoplasmic reticulum, in addition to extracellular deposits in the stroma and Bowman's layer (Klintworth and Vogel, 1964) (Figure 3.7 B). The posterior zone of Descemet's membrane is thickened with deposits occurring in a honeycomb pattern, and excrescences of Descemet's membrane called guttae may also be present. Despite its classification as a stromal dystrophy, endothelial cells have also been shown to contain vacuoles containing a fibrillogranular material (Snip, Kenyon and Green, 1973).

MCD cases are caused by biallelic mutations in the *CHST6* gene, or structural mutations upstream that disrupt its expression, located at chr16q32.1. *CHST6* consists of a single coding exon that encodes the enzyme corneal N-acetylglucosamine-6-sulfotransferase and is required for the production of keratan sulphate (Akama *et al.*, 2000). Keratan sulphate is an important GAG side chain of stromal proteoglycans including lumican, keratocan and mimecan. These are required for regular spacing of collagen lamellae and therefore for stromal transparency (section 1.2.3). Over 150 pathogenic mutations have been described in *CHST6* including nonsense mutations, frameshifts and missense mutations.

The *CHST6* gene is located ~ 30 kb upstream from the *CHST5* gene and they are both transcribed from the negative DNA strand. The genes share ~ 90% identity at the cDNA level and ~ 89% at the amino acid level, suggesting that they likely arose from a gene duplication event. The non-coding regions immediately upstream of both genes also contain two regions (regions A and B) which are flanked by highly homologous sequences and disruptions of these regions are prone to occur by non-allelic homologous recombination (Akama *et al.*, 2000). Figure 3.8 demonstrates the two mechanisms by which the two reported disruptions occur, resulting in the same functional outcome. These include a ~ 40 kb deletion encompassing *CHST5* and the upstream region of *CHST6*, or a ~ 2.5 kb replacement of the *CHST5* upstream region with the upstream region of *CHST6* (Klintworth, Smith and Bowling, 2006). These structural mutations are associated with a loss of *CHST6* transcript expression specifically in the cornea, suggesting that an important cornea specific regulatory element is located within region B. Although biallelic coding mutations in *CHST6* are most common, the relative contribution of *CHST6* coding mutations and structural mutations to the development of MCD is unknown.

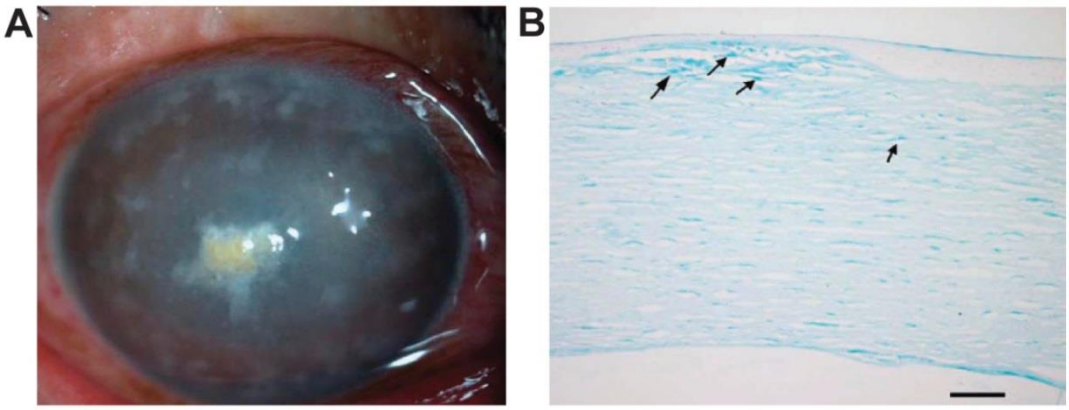


Figure 3.7 Clinical and histological appearance of macular corneal dystrophy (MCD). (A) Clinical slit lamp image of MCD showing the presence of white-yellow stromal opacities. **(B)** Histology of cornea stained with alcian blue which is indicative of glycoprotein. Arrows indicate accumulations of abnormal material in the sub-epithelium and stroma. Scale bar is 100 uM. Images from (Liu *et al.*, 2010).

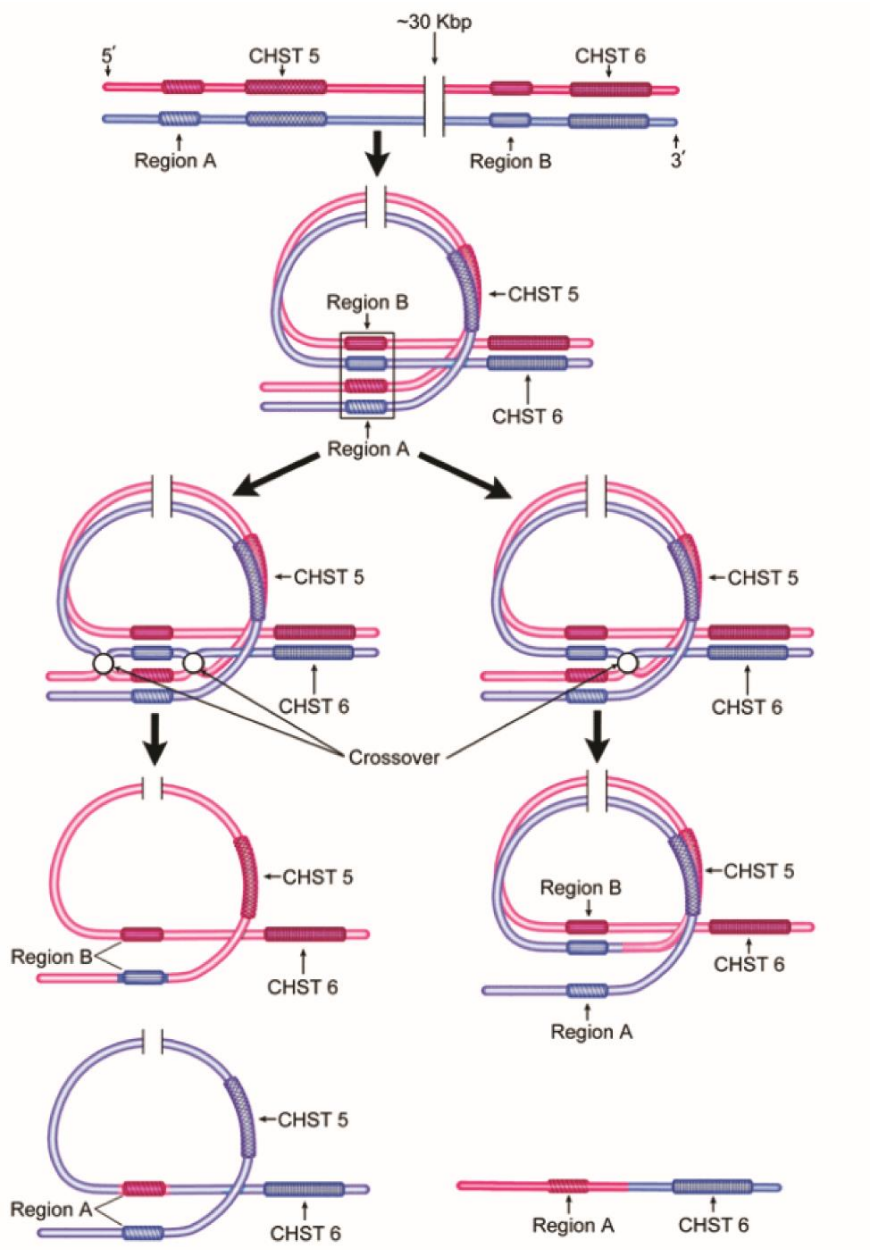


Figure 3.8 Disruption of the *CHST6* upstream putative regulatory region can occur by two possible mechanisms (deletion or re-arrangement) with the same functional outcome. Regions A and B, located upstream of *CHST5* and *CHST6* respectively, mis-pair during homologous recombination due to stretches of homology flanking the regions. One possibility is a double crossover resulting in region B upstream from *CHST6* on one allele being replaced with region A upstream of *CHST5* (left). Alternatively, a single crossover also results in the replacement of region B with region A, but with the deletion of *CHST5* and region B (right). Diagram from (Klintworth, Smith and Bowling, 2006).

3.1.3.2 Schnyder corneal dystrophy

SCD is an autosomal dominant corneal dystrophy associated with the deposition of cholesterol and phospholipids in the corneal stroma (Garner and Tripathi, 1972). Sub-epithelial cholesterol crystals become apparent as early as the first decade of life in ~ 50% of patients (Figure 3.9 A). The remaining ~ 50% do not present with crystalline deposition, meaning that diagnosis of SCD can be delayed until the second decade of life (Weiss, 1996). Corneal arcus, a narrow band of cholesterol deposited in the circumference of the cornea, is present in both groups of patients from the second decade onwards, although this is a non-specific sign that is also present in some forms of anterior segment dysgenesis, hypercholesterolaemia, and old age (Weiss and Khemichian, 2011). In the fourth decade of life, central stromal haze begins to cause significant visual impairment and progressively worsens with age, when surgical intervention may be required (Figure 3.9 B). As the disease progresses there is a loss of corneal sensation, caused by a loss of corneal nerves (Kobayashi *et al.*, 2009). Systemic features associated with SCD have also been reported. Hyperlipoproteinemia reportedly occurs in ~ 66% of SCD patients; however, it has also been diagnosed in individuals from SCD families who do not have a corneal dystrophy, therefore may not be associated with the underlying genetic cause (Du *et al.*, 2011).

SCD is caused by heterozygous mutations in the *UBIAD1* (UbiA prenyltransferase domain containing 1) gene (Orr *et al.*, 2007; Weiss *et al.*, 2007). *UBAID1* is a two exon coding gene, located at chr1p36.22. Individuals within the same family may differ in their development of SCD phenotype, including the presence or absence of discrete cholesterol crystals, indicating that other genetic or environmental factors affect the clinical phenotype (Du *et al.*, 2011). Initially proposed as a novel corneal dystrophy, central discoid dystrophy was subsequently identified to be a phenotypic variant of SCD without crystalline deposition, caused by mutations in *UBIAD1* (Weiss *et al.*, 2010).

UBIAD1 encodes an enzyme involved in the biosynthesis of menaquinone-4 (vitamin K) and coenzyme Q10 although the relevance of these functions to the pathogenesis of SCD is unknown (Nakagawa *et al.*, 2010; Mugoni *et al.*, 2013). The corneal

clouding associated with SCD is caused by the abnormal accumulation of unesterified cholesterol and phospholipids including sphingomyelin, phosphatidylserine and phosphatidylinositol (Yamada *et al.*, 1998). High density lipoproteins (HDL) including Apo-AI, AII and ApoE are specifically raised in affected corneas and the molar ratio of cholesterol to phospholipids is increased three-fold (Gaynor *et al.*, 1996). Oxidative stress has also been implicated in the pathogenesis of SCD with greater levels of immunofluorescent staining of inducible and constitutive nitric oxide synthase in SCD corneal buttons compared to controls (Gatzioufas *et al.*, 2010). In hepatocarcinoma cells, reduction of *UBIAD1* expression has been shown to result in increased levels of reactive oxidative species and impairment of mitochondrial morphology (Morales *et al.*, 2014).

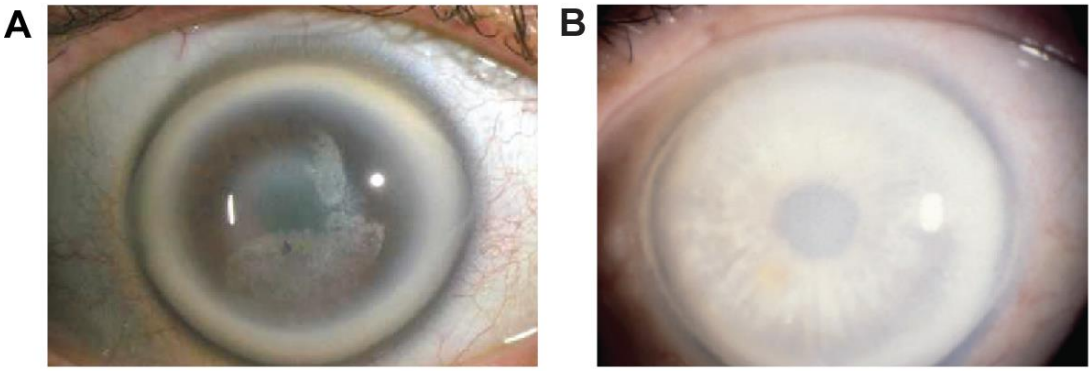


Figure 3.9 Two potential clinical phenotypes of Schnyder corneal dystrophy (SCD). (A) Classic phenotype showing discrete crystal deposition, found in ~ 50% of SCD cases. Image from (Rittenbach, 2013). **(B)** Central stromal haze without crystals occurs in ~ 50% of SCD cases. Image from (Weiss, Kruth, *et al.*, 2008). Both (A) and (B) show corneal arcus.

3.1.4 Aims of chapter

In this chapter, patients with anterior corneal dystrophies were recruited from MEH or referred from other clinical centres, and screened for known genetic causes of corneal dystrophies. The aim was to identify the most common causes of anterior corneal dystrophies within this cohort, to correlate the clinical phenotype with underlying genotype, and to investigate any potentially novel mutations and genetic causes of anterior corneal dystrophies. Patients who were initially clinically diagnosed with an anterior corneal dystrophy, but subsequently identified to have a different underlying cause of corneal opacities are discussed in chapter 5.

3.2 *TGFBI* mutations

3.2.1 TGFBI cohort and gene screening

Mutations in *TGFBI* are responsible for GCD, LCD, RBCD and TBCD (section 3.1.1). With the exception of RBCD and TBCD, these dystrophies are sufficiently clinically distinct to be distinguished by slit lamp examination. We recruited a cohort of 68 probands (F1 – F68), and their relatives, with a diagnosis of a *TGFBI* associated epithelial-stromal dystrophy. The majority of the probands were positive for a family history of corneal disease, which was consistent with an autosomal dominant pattern of inheritance. Additionally, 23 probands (F69 – F91) diagnosed with EBMD were included in this cohort, of which three were positive for a family history of corneal dystrophy. These patients were included because two mutations in *TGFBI* have previously been reported in EBMD patients, and it is estimated that these mutations may explain ~ 10% of EBMD cases; however, this has not been replicated in an independent study (Boutboul *et al.*, 2006) (section 3.1.2.1). In total, including additional family members, 113 DNA samples were obtained for this study.

For all probands, *TGFBI* exons 4, 11, 12, 13, 14 and 16 (previously reported to harbour mutations) were amplified by PCR and Sanger sequenced (section 2.2; Appendix A). The remaining coding exons in the 17 coding exon gene were directly screened in those probands who were negative for mutations. Any variants identified in *TGFBI* were assessed for potential pathogenicity using frequency, pathogenicity prediction programs and conservation of the residue in orthologous protein sequences (section 2.3). Segregation analysis was performed when DNA was available for additional family members. A mutation in *TGFBI* was identified in all 68 probands (F1 – F68) with a clinical diagnosis of an epithelial-stromal *TGFBI* dystrophy and in two probands (F69 – F70) clinically diagnosed with EBMD. The remaining 21 probands with EBMD (F71 – F91) were negative for mutations in all *TGFBI* coding exons.

3.2.2 Recurrent mutations at amino acid residues Arg-124 and Arg-555

Mutations occurring at one of two *TGFBI* mutation hot-spot amino acid residues Arg-124 or Arg-555, located in exons 4 and 12 respectively, were found to be responsible for disease in 89.7% (61/68) of cases of epithelial-stromal *TGFBI* dystrophies (Figure 3.10 A; Figure 3.11). These mutations showed complete correlation between their clinical phenotype and the underlying causative mutation (Figure 3.10).

A heterozygous c.1664G>A; p.(Arg555Gln) mutation, previously reported to be associated with TBCD, was identified in 20 unrelated probands diagnosed with RBCD/TBCD and was the most common mutation in this cohort (Figure 3.10 A; Figure 3.11 A). A RBCD associated mutation [c.371G>T; p.(Arg124Leu)] was found in only one patient, demonstrating that TBCD is much more prevalent than RBCD in our patient population (Figure 3.11 A). Clinically, all individuals in this group had anterior stromal opacity as previously described (Figure 3.10 B).

A heterozygous c.370C>T; p.(Arg124Cys) mutation was the second most common *TGFBI* mutation, identified in 19 probands (Figure 3.10 A; Figure 3.11 A). All probands with this mutation demonstrated a phenotype which was consistent with a classic LCD1 phenotype, with thin, branching refractile lattice lines observed on clinical examination (Munier *et al.*, 1997) (Figure 3.10 B).

Twenty-one probands were recruited with a clinical diagnosis of GCD. A c.1663C>T; p.(Arg555Trp) mutation, which correlates with a clinical diagnosis of GCD1, was identified in 13 probands, and the c.371G>A; p.(Arg124His) mutation (GCD2) was identified in 8 probands (Figure 3.10 A; Figure 3.11 A). Individuals with GCD1 had predominantly deep granular opacities while individuals with GCD2 had either exclusively granular opacities or anterior stromal 'snowflake-like' opacities, consistent with previous reports (Munier *et al.*, 1997; Stewart *et al.*, 1999), however mid-stromal haze was absent (Figure 3.10 B).

With the exception of c.1664G>A; p.(Arg555Gln), all mutations identified were predicted to be pathogenic by SIFT and Polyphen2 (Table 3.2). All were absent from

ExAC, except for c.370C>T; p.(Arg124Cys) and c.371G>A; p.(Arg124His), which were present at low frequencies (Table 3.2). When DNA from other members was available, the identified mutations segregated with the corneal dystrophy phenotype. An alignment of orthologous TGFBI protein sequences revealed that the Arg-124 and Arg-555 residues are conserved in mammals, but not in birds, fish or reptiles (Figure 3.11 B and C).

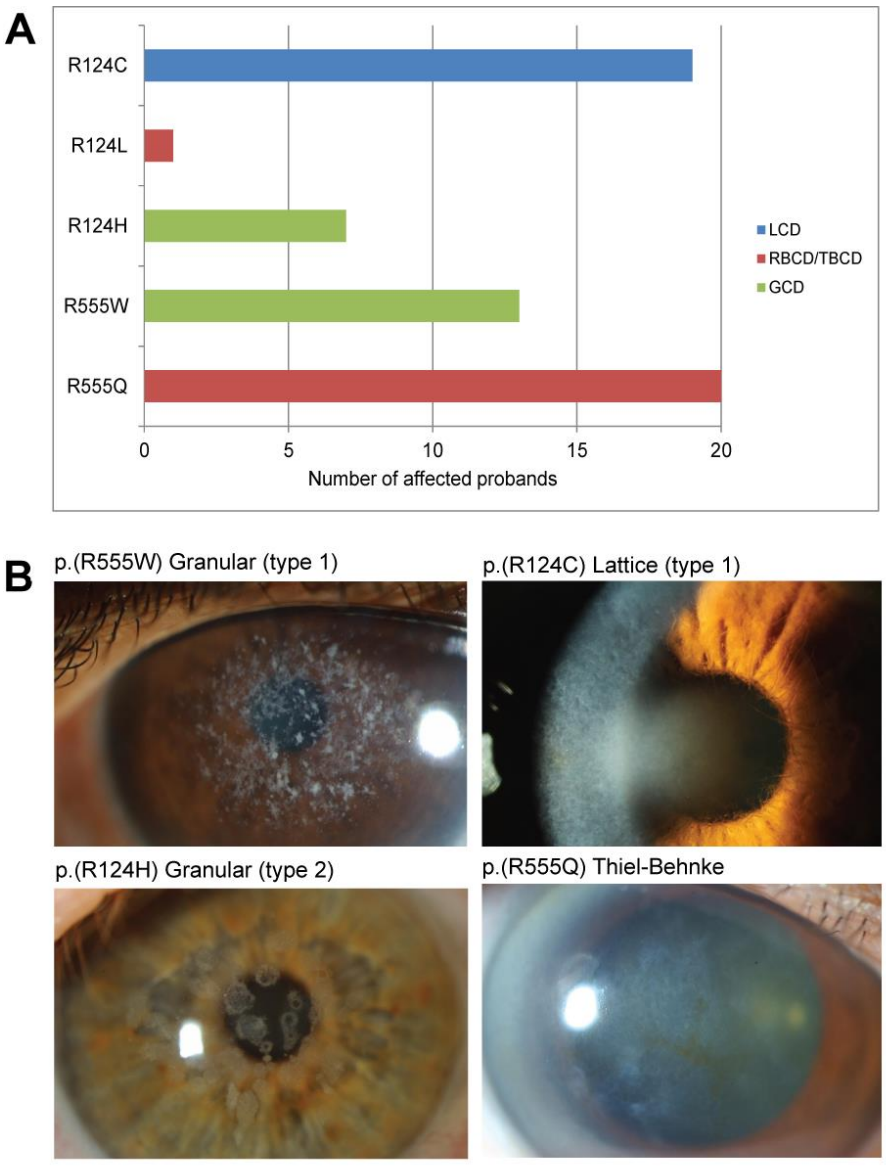


Figure 3.10 *TGFBI* hotspot mutations affecting residues Arg-124 and Arg-555 show complete genotype-phenotype correlation. (A) Graphical representation of the number of probands carrying a mutation affecting *TGFBI* amino acid residue Arg-124 and Arg-555 and associated clinical diagnosis. There is complete genotype-phenotype correlation for mutations affecting these residues. **(B)** Representative clinical phenotype associated with each mutation.

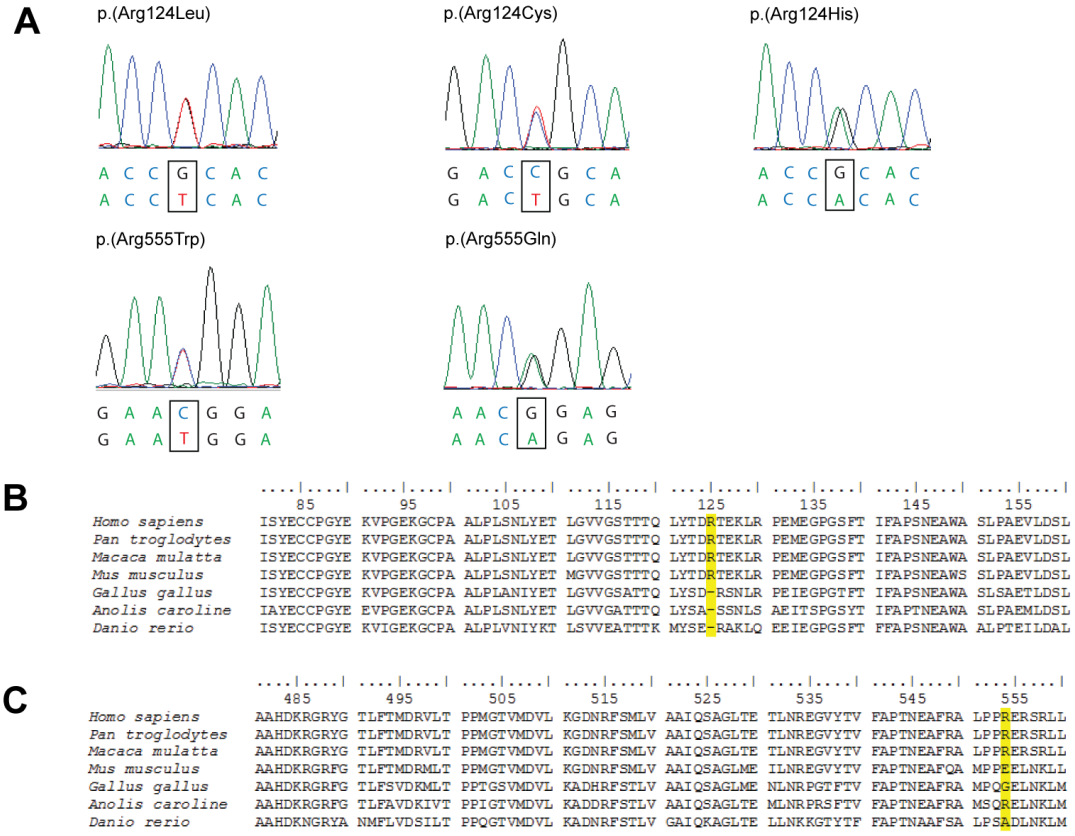


Figure 3.11 Sequence electropherograms of *TGFBI* hot-spot mutations and protein alignment of residues Arg-124 and Arg-555. (A) Sequence electropherograms showing mutations affecting hot-spot residues Arg-124 and Arg-555. **(B)** Alignment showing Arg-124 residue (highlighted) in orthologous *TGFBI* protein sequences. The Arg-124 residue is conserved in mammals but absent from birds, fish and reptiles. **(C)** Alignment showing Arg-555 residue (highlighted) in orthologous *TGFBI* protein sequences. The residue is conserved in mammals only.

3.2.3 Variant-LCD

Four probands (F62-F65) diagnosed with classic or phenotypic variants of LCD were identified to harbour three mutations in exon 14 of *TGFBI*.

The proband of family #62 (F62: III:2; Figure 3.12 A) was initially clinically diagnosed with RBCD with diffuse anterior stromal haze and small granular focal opacities (Figure 3.12 C). However, histology was negative for Masson's trichrome staining and did not demonstrate the sub-epithelial fibrous sheet-like deposits which are characteristic of RBCD or TBCD, therefore these diagnoses were excluded. Histology was positive for Congo red staining, indicating the presence of amyloid in the subepithelium and few sub-epithelial deposits in the superficial stroma,

consistent with a diagnosis of LCD (Figure 3.12 D). Screening revealed the presence of a heterozygous c.1874T>A; p.(Val625Asp) mutation (Figure 3.12 B). This variant is absent from ExAC, affects a conserved residue and is predicted to be pathogenic by both SIFT and Polyphen2 (Table 3.2; Figure 3.12 E).

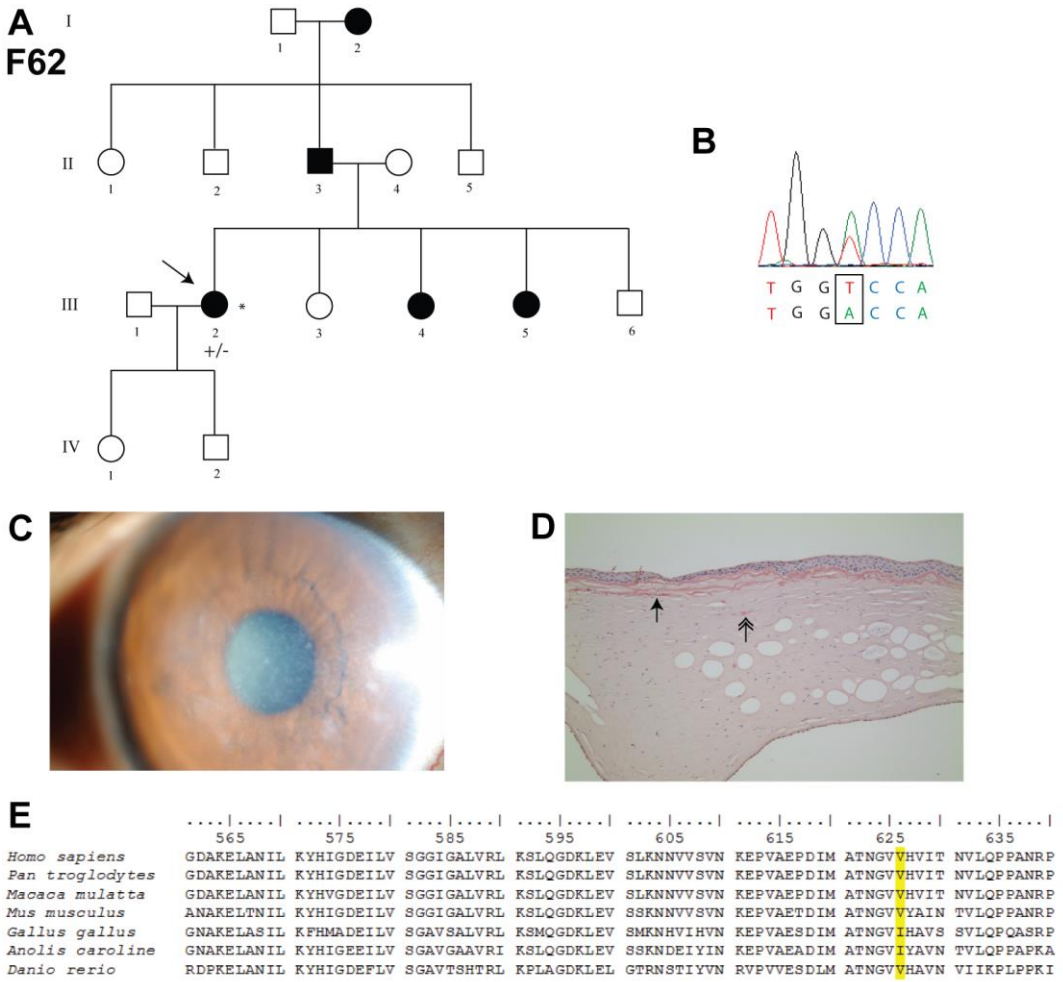


Figure 3.12 *TGFBI* c.1874T>A; p.(Val625Asp) mutation identified in the proband of family #62 with lattice corneal dystrophy. (A) Pedigree structure showing dominant pattern of inheritance. * indicates DNA sample was available for testing. +/- indicates that the mutation was identified in the heterozygous state. **(B)** Sequence electropherogram showing a c.1874T>A; p.(Val625Asp) mutation. **(C)** Clinical phenotype of proband F62: III:2 showing anterior stromal haze combined with small granular focal opacities. **(D)** Histology showing amyloid predominantly located near Bowman's layer (arrow) with few stromal deposits (doubled-headed arrow), resulting in a diagnosis of LCD. **(E)** Alignment of orthologous TGFBI protein sequences with Val-625 highlighted in yellow.

The proband of family #63 (F63: IV:3; Figure 3.13 A) was diagnosed with LCD due to mild symptoms of bilateral thin, branching lattice lines. Her older brother (F63:

IV:2) was also affected with similar symptoms but with only minor changes in one eye and an asymmetric progression of disease (Figure 3.13 C). The proband of family #64 (F64: II:1; Figure 3.13 B) was also mildly affected with LCD, with lattice lines predominantly below the visual axis. Both probands harboured the c.1877A>G; p.(His626Arg) mutation in exon 14 of *TGFBI* (Figure 3.13 C). This mutation was predicted to be pathogenic by SIFT and Polyphen2 and was absent from the control database ExAC (Table 3.2).

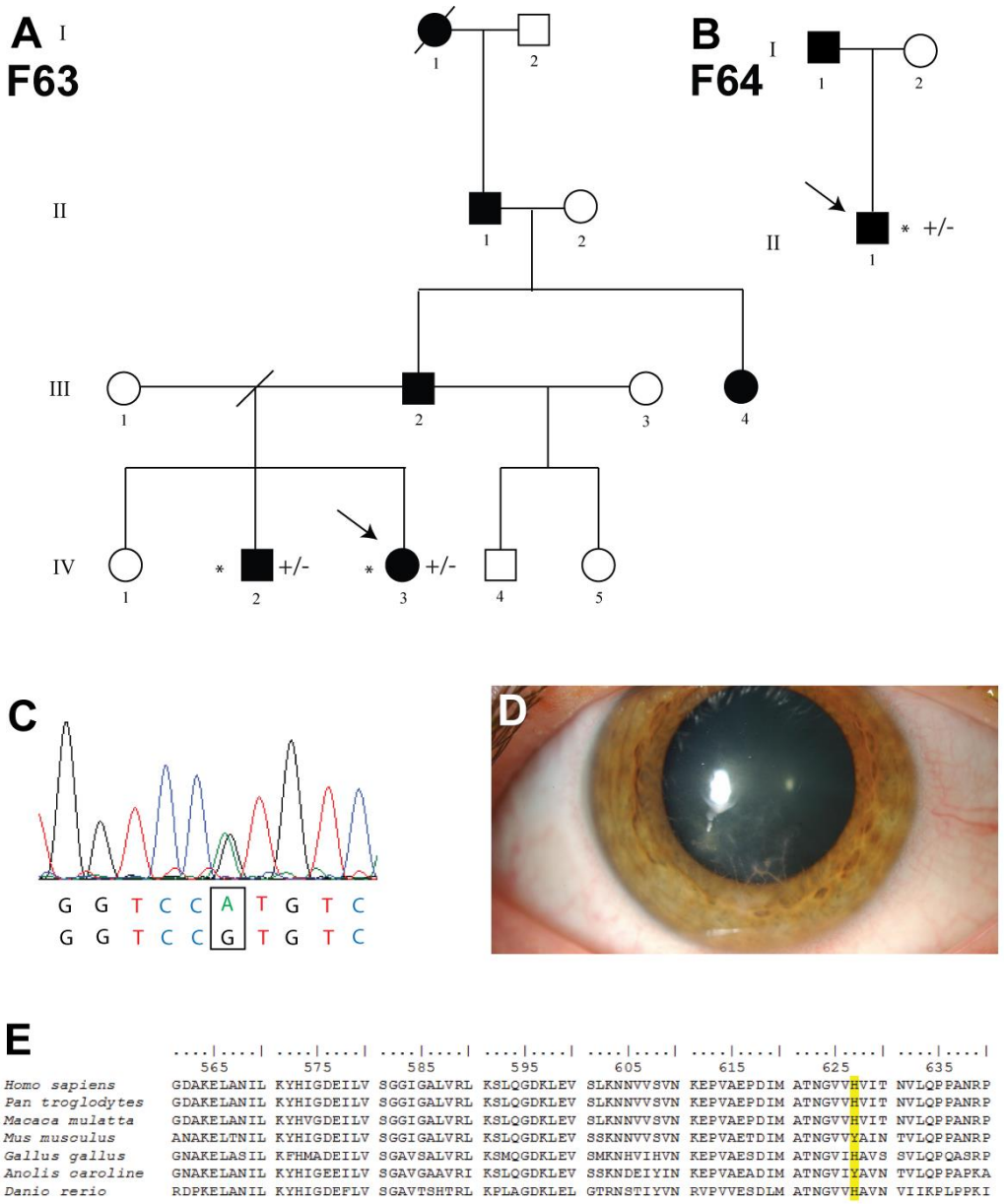


Figure 3.13 *TGFBI* c.1877A>G; p.(His626Arg) mutation in F63 and F64 with lattice corneal dystrophy. (A) Pedigree structure of F63 showing dominant inheritance of the corneal dystrophy. **(B)** Pedigree structure of F64 is consistent with autosomal dominant inheritance. * indicates DNA sample is available for testing. +/- indicates that the mutation was identified in the heterozygous state. **(C)** Sequence electropherogram showing the c.1877A>G; p.(His626Arg) mutation. **(D)** Clinical appearance of lattice lines in F63: IV:2. **(E)** Alignment of His-626 residue in orthologous protein sequences shows that it is conserved for the majority of sequences.

The proband of family #65 (F65: II:10; Figure 3.14 A), of East Asian origin, was diagnosed clinically with LCD with thin, branching lattice lines consistent with the classic phenotypic subtype. Sequencing of *TGFBI* identified a c.1859C>A; p.(Ala620Asp) mutation in exon 14 (Figure 3.14 B). This missense change was

predicted to be pathogenic by SIFT and Polyphen2, absent from ExAC (Table 3.2), and affects a conserved residue (Figure 3.14 C).

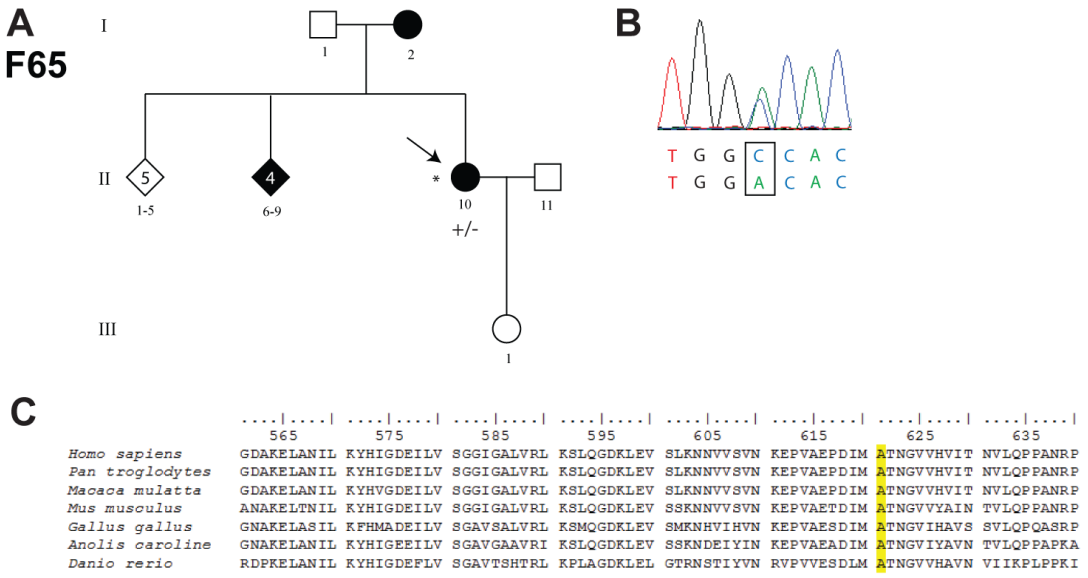


Figure 3.14 *TGFBI* c.1859C>A; p.(Ala620Asp) mutation in F65 with lattice corneal dystrophy. **(A)** Pedigree structure of proband II:10 showing dominant inheritance of the dystrophy. * indicates DNA sample is available for testing. +/- indicates that the mutation was identified in the heterozygous state. **(B)** Sequence electropherogram showing the c.1859C>A; p.(Ala620Asp) mutation. **(C)** Alignment of orthologous TGFBI protein sequences with Ala-620 residue highlighted in yellow.

3.2.4 Phenotypic spectrum associated with p.(Gly623Asp)

Probands from five families (F66 – F70; Figure 3.15 A) were identified to harbour a c.1868G>A; p.(Gly623Asp) mutation in exon 14 (Figure 3.15 F), associated with a range of clinical diagnoses, including two probands from the EBMD cohort. This mutation is absent from ExAC, is predicted to be pathogenic by SIFT and Polyphen2 and affects a conserved residue (Table 3.2; Figure 3.15 G).

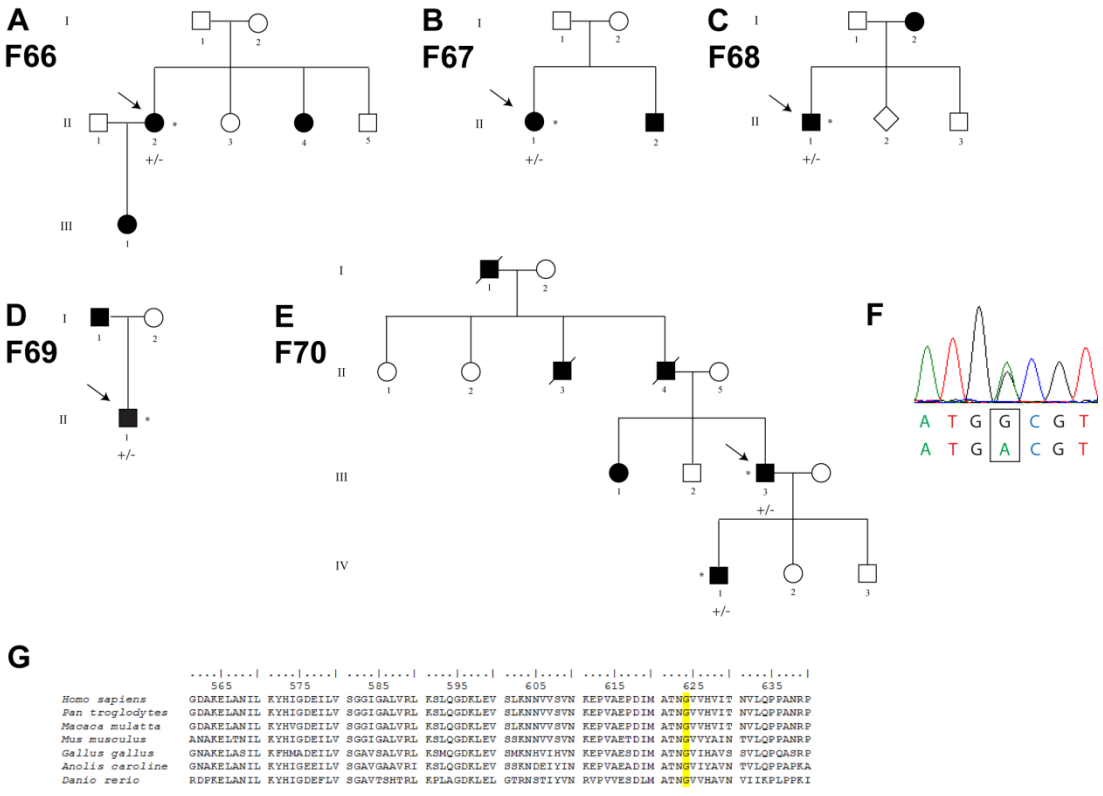


Figure 3.15 Five families (F66 – F70) harbouring a *TGFBI* c.1868G>A; p.(Gly623Asp) mutation. Pedigrees showing inheritance pattern of corneal dystrophy in families harbouring the c.1868G>A; p.(Gly623Asp) mutation. * indicates DNA sample was available for testing. +/- indicates presence of heterozygous mutation**. (A)** F66 **(B)** F67 **(C)** F68 **(D)** F69 **(E)** F70. **(F)** Sequence electropherogram showing presence of heterozygous c.1868G>A; p.(Gly623Asp) mutation **(G)** Alignment of orthologous protein residues with Gly-623 highlighted.

The proband of family #68 (F68: II:1, Figure 3.15 C) was the only proband in the p.(Gly623Asp) cohort with a clinical diagnosis of LCD. He presented in his sixties with symptoms of severe recurrent erosion which were resolved with PTK. Clinical examination revealed small lattice structures and anterior diffuse stromal haze (Figure 3.16 A).

Probands from families #66 (F66: II:2) and #67 (F67: II:1) were both diagnosed with RBCD/TBCD. F66: II:2 had previously been managed clinically for several years with a diagnosis of bilateral herpes simplex keratitis prior to being diagnosed with a corneal dystrophy. Clinical examination revealed a geographic pattern of sub-epithelial haze and tiny granular structures (Figure 3.16 B). F67: II:1 also had sub-

epithelial haze and tiny granular structures, and had been diagnosed previously with both EBMD and RBCD/TBCD on different occasions (Figure 3.16 C).

Two probands of families #69 (F69: II:1) and #70 (F70: III:3) from the EBMD cohort also harboured the c.1868G>A; p.(Gly623Asp) mutation. In family F69, II:1 had elevated serpiginous superficial opacities of both corneas with no clinical signs of an epithelial-stromal *TGFBI* dystrophy (Figure 3.16 F). He was initially considered to be a sporadic case, without a family history of corneal disease; however, following identification of this mutation it was discovered that his father had been previously diagnosed with dry eye disease. In family F70, III:3 and his son (IV:1) were both diagnosed with EBMD due to superficial stromal map-like opacities (Figure 3.16 D and E). It was unknown whether these were primary opacities or secondary to recurrent erosive episodes.

Collectively, the c.1868G>A; p.(Gly623Asp) mutation positive cohort demonstrated a consistently late-onset (4^{th} decade or later) of severe recurrent erosions. Clinically, a diffuse or geographic pattern of superficial stromal opacity, with small refractile granules or lines were observed in some probands, although this may be modified by secondary scar formation. PTK treatment was beneficial to relieving symptoms for those treated in this cohort.

A summary of all *TGFBI* mutations identified in this cohort is shown in Table 3.1 and diagrammatically shown in Figure 3.17.

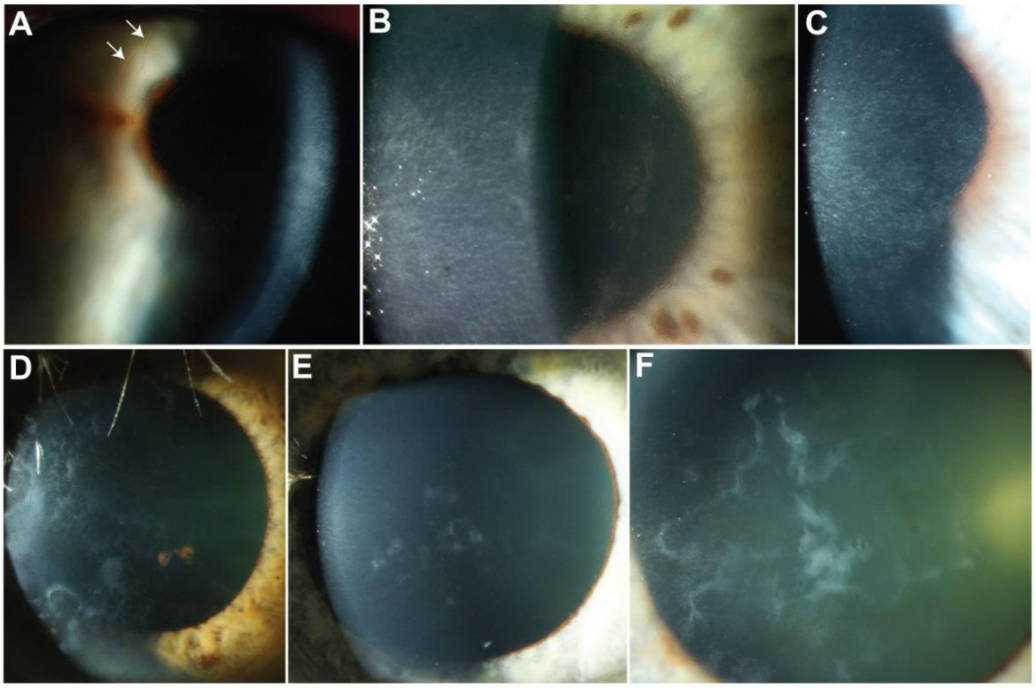


Figure 3.16 Spectrum of phenotypes associated with a c.1868G>A; p.(Gly623Asp) mutation in families F66 – F70. (A) F68: II:1 with small lattice structures (arrows) resulting in a clinical diagnosis of lattice corneal dystrophy. **(B)** Bowman's layer opacities in F66: II:2. **(C)** Bowman's layer opacities in F67: II:1. Both B and C were diagnosed with Reis-Bücklers corneal dystrophy or Thiel-Behnke corneal dystrophy. (**D** and **E**) Map-like opacities reminiscent of epithelial basement membrane dystrophy (EBMD) were observed in the father and son, respectively, of family F70: III:3 and IV:1. **(F)** Map-like opacities in F69: II:1, diagnosed with EBMD.

Table 3.1 Summary of *TGFBI* mutations identified in 91 unrelated probands diagnosed with an epithelial-stromal *TGFBI* dystrophy or epithelial basement membrane dystrophy. A summary of the causative mutation is shown for each of the clinical diagnoses. Thiel-Behnke and Reis-Bücklers corneal dystrophy are grouped because histology (Masson's trichrome staining) is required to reliably distinguish between the two conditions. N = no of probands.

Clinical Diagnosis	N (%)	*TGFBI* Mutation (N)
Lattice corneal dystrophy	24 (26%)	p.(R124C) (19) p.(V625D) (1) p.(H626R) (2) p.(A620D) (1)
Granular corneal dystrophy	21 (23%)	p.(G623D) (1) p.(R555W) (13) p.(R124H) (8)
Thiel-Behnke/Reis-Bücklers corneal dystrophy	23 (25%)	p.(R555Q) (20) p.(R124L) (1) p.(G623D) (2)
Epithelial basement membrane dystrophy	23 (25%)	p.(G623D) (2) Negative (21)

Table 3.2 Assessment of pathogenicity of *TGFBI* mutations. All *TGFBI* mutations identified in the cohort were assessed for potential pathogenicity using prediction programs SIFT and Polyphen2 which use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. The frequency of the variant was taken from the ExAC database. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Nucleotide	Protein	SIFT (1-0)	Polyphen2 (0-1)	ExAC	Ref.
c.370C>T	p.(Arg124Cys)	0 (D)	0.958 (PrD)	0.0000083 (1/120340)	(Munier *et al.*, 1997)
c.371G>T	p.(Arg124Leu)	0.02 (D)	0.854 (PoD)	NP	(Mashima *et al.*, 1999)
c.371G>A	p.(Arg124His)	0.02 (D)	0.958 (PrD)	0.000058 (7/120220)	(Munier *et al.*, 1997)
c.1663C>T	p.(Arg555Trp)	0.02 (D)	0.846 (PoD)	NP	(Munier *et al.*, 1997)
c.1664G>A	p.(Arg555Gln)	0.27 (T)	0.252(B)	NP	(Munier *et al.*, 1997)
c.1859C>A	p.(Ala620Asp)	0 (D)	1 (PrD)	NP	(Lakshminarayanan *et al.*, 2011)
c.1868 G>A	p.(Gly623Asp)	0 (D)	1 (PrD)	NP	(Afshari *et al.*, 2001)
c.1874T>A	p.(Val625Asp)	0 (D)	0.956 (PrD)	NP	(Tian *et al.*, 2007)
c.1877 A>G	p.(His626Arg)	0.01 (D)	0.583 (PoD)	NP	(Chau *et al.*, 2003)

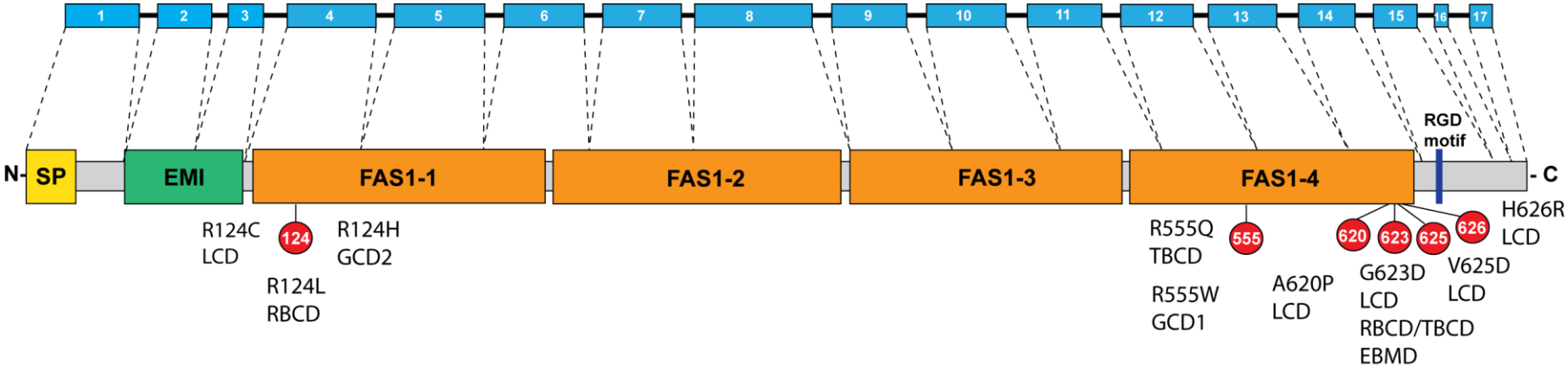


Figure 3.17 Summary of pathogenic mutations in the study cohort mapped onto *TGFBI* gene and protein structure, showing functional domains. Mutations cluster at hot-spots Arg-124 in the first fasciclin-1 (FAS-1) and Arg-555 in the fourth FAS-1 domain. Rarer non hot-spot mutations cluster towards the end of the fourth FAS1 domain. SP: Signal petide. EMI: N-terminal cysteine rich (domain), R-G-D (amino acid) motif. LCD: Lattice corneal dystrophy, RBCD: Reis-Bücklers corneal dystrophy, TBCD: Thiel-Behnke corneal dystrophy, EBMD: Epithelial basement membrane dystrophy.

3.3 Next-generation sequence analysis of epithelial basement membrane dystrophy

Corneal erosions are the detachment of the epithelial cell layer from the underlying epithelial basement membrane, which causes episodes of severe pain and photophobia. They can be caused by environmental factors, primarily trauma, or occur as a consequence of an anterior corneal dystrophy, including *TGFBI*-associated corneal dystrophies RBCD/TBCD, GCD and LCD, and other corneal dystrophies including ERED and MECD (Reidy, Paulus and Gona, 2000). EBMD, a dystrophy associated with changes of the epithelial basement membrane and recurrent corneal erosions, is not included in the IC3D classification as a *TGFBI*-associated corneal dystrophy, but in a study of 30 EBMD patients it was estimated to be caused by mutations in *TGFBI* in ~ 10% of cases (Boutboul *et al.*, 2006) (section 3.1.2.1). The cause of the dystrophy in the remaining EBMD cases is unknown. In this section, genetic investigation was performed for the cohort of 21 *TGFBI*-mutation negative EBMD probands (F71 – 91) who were screened in section 3.2.

Notably in our *TGFBI*-negative EBMD cohort (F71 – 91), two of the probands reported a family history of corneal disease. The proband of family #71 (F71: II:2; Figure 3.18 A), diagnosed with EBMD, reported a sibling (II:3) who had also been diagnosed with EBMD, who was subsequently recruited to the study, and an additional sibling (II:1) who had experienced similar symptoms, but had not been clinically examined and did not have a clinical diagnosis. Neither parent was clinically examined but there was no history of corneal disease, which could indicate a potentially recessive or partially penetrant dominant inheritance pattern. Providing both parents are unaffected then *de novo* dominant is another possibility. The proband of family #72 (F72: II:1; Figure 3.18 B) was diagnosed with EBMD and indicated that his late father (I:2) had also been diagnosed and treated for the same condition. DNA was obtained by saliva sample from his half-brother (II:2) who was also diagnosed with EBMD, however this individual was not clinically examined at MEH. This pedigree therefore indicated a potentially dominant pattern of inheritance. The remaining 19 *TGFBI*-negative EBMD probands (F73 – 91) were negative for a family history of corneal disease.

Three families from this cohort (F71 – F73) were analysed by WES (section 2.6). For families F71 and F72, two affected individuals from each pedigree were sequenced. A sporadically affected proband (F73: II:1) was also analysed by WES (Figure 3.18 C and D). Given the shared clinical diagnosis of EBMD in all three probands, it was hypothesised that they may have a shared genetic cause. Variants in three probands were filtered for rare non-synonymous or splice site variants (MAF ≤ 0.005) in the 1000G, EVS and UCL internal control databases, irrespective of homozygous or heterozygous; however, no variants in shared genes were identified (section 2.6.2.1). This analysis was repeated excluding the proband F73: II:1 due to the lack of family history of corneal disease, but this also failed to identify a shared candidate gene(s). Therefore, a family specific based filtering strategy was applied to each individual family (section 2.6.2.2).

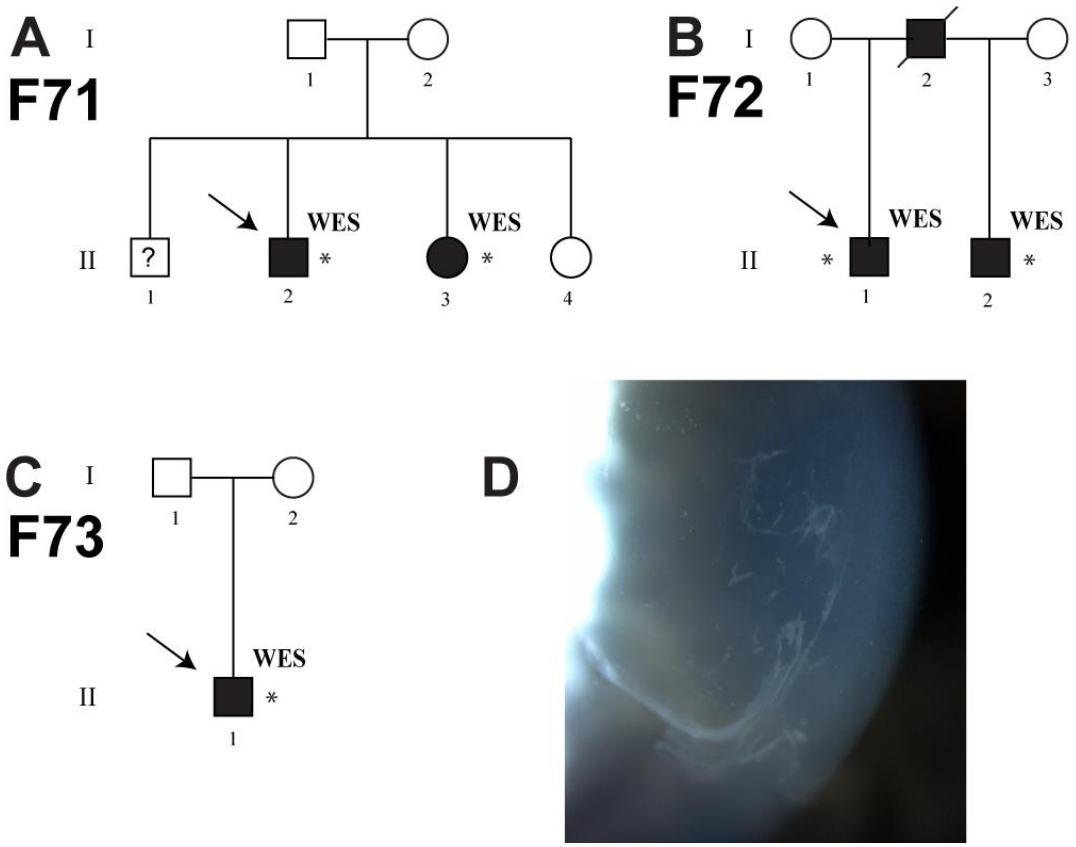


Figure 3.18 Three families with *TGFBI*-negative epithelial basement membrane dystrophy (EBMD) analysed by whole exome sequencing (WES). * indicates DNA was available for those individuals. **(A)** Pedigree of family #71 shows potentially recessive or dominant disease with reduced penetrance. **(B)** Pedigree of family #72 is consistent with a dominant pattern of inheritance. **(C)** Pedigree of family #73 shows a single affected individual in the family which is consistent with recessive, *de novo* or partially penetrant dominant inheritance or may indicate the condition is of non-genetic aetiology. **(D)** Clinical image of F73: II:1 showing 'fingerprint' lines which are characteristic of EBMD.

3.3.1 Analysis of WES data for family #71

The proband of F71: II:2 was a male proband in his sixties diagnosed with EBMD. He had a history of several years of recurrent erosions. He underwent laser PTK in his right eye in 2013 which was successful and is currently asymptomatic. His sister (F71: II:3), also in her sixties, was also seen in clinic at MEH and had bilateral changes consistent with EBMD, however she had been asymptomatic for years, with the most recent corneal erosion reported as five years prior. The proband also reported a brother with similar symptoms; however, he had not been clinically

examined at MEH and therefore lacked a clear diagnosis. There was no knowledge of corneal disease in the parents (Figure 3.18 A).

Following negative screening of *TGFBI* in the proband, DNA from both affected siblings was analysed by WES (section 2.6). Given the pedigree, the possibility of a recessive pattern of inheritance was initially investigated. Variants were filtered for rare (MAF< 0.005) homozygous variants shared between the two affected siblings in the pedigree, resulting in the identification of one candidate gene (*KANK2*) containing a homozygous non-synonymous variant (Table 3.3; section 2.6.2.2). Biallelic mutations in *KANK2* cause keratoderma and wooly hair syndrome (Ramot *et al.*, 2014).

Two candidate gene lists were compiled for filtering WES data. The first is a manually compiled list of 127 genes associated with inherited corneal disorders including both corneal dystrophies (Table 1.1) and syndromes with corneal symptoms (Appendix B; section 2.6.2.3). The second list includes ~3000 genes which encode proteins that are present in the corneal proteome which encompasses proteins produced by corneal cell types and additionally a significant proportion (1050) of plasma proteins (Dyrlund *et al.*, 2012) (Appendix C; section 2.6.2.4). *KANK2* was absent from both these candidate gene lists.

To investigate the possibility of partially penetrant dominant disease, heterozygous variants were filtered for shared non-synonymous or splice site variants with a frequency ≤ 0.005. Sixteen variants were shared between both affected siblings (Table 3.4). None of the genes contained potentially compound heterozygous variants that could be causative of recessive disease. Furthermore, no variants were present in genes known to cause corneal disease when filtered against the inherited corneal disorders gene lists (section 2.6.2.3; Table 1.1; Appendix B). To further refine this list of candidate variants, they were filtered against a corneal proteome candidate gene list (section 2.6.2.4; Appendix C). Two of the 16 variants were present in genes that encode proteins which are present in the cornea (*PDIA4* and *HSP90B1)*. Both variants have a higher frequency in ExAC than would be expected for dominant disease, however as EBMD is one of the most common anterior corneal dystrophies, this does not exclude either variant as causative. Both variants are

predicted to be tolerated by SIFT, although the *PDIA4* variant is predicted to be 'probably pathogenic' by Polyphen2.

Table 3.3 Rare (≤ 0.005) shared homozygous non-synonymous or splice site variants in two affected individuals with epithelial basement membrane dystrophy from family #71. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
KANK2 NM_001136191.2	c.10G>A; p.(Val4Ile)	0.15 (T)	0.956 (PrD)	0.0031 (368/120166)

Table 3.4 Rare (≤ 0.005) shared heterozygous non-synonymous or splice site variants in two affected individuals with epithelial basement membrane dystrophy from family #71. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Those present in the corneal proteome are highlighted in bold.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
KIAA1324 NM_001267048.1	c.368G>A; p.(Arg123Gln)	0.02 (D)	0.968 (PrD)	0.00067 (81/121338)
SYT6 NM_001270805.1	c.353G>A; p.(Ser118Asn)	0.52 (T)	0.04 (B)	0.000033 (4/121390)
RASGRP3 NM_170672.2	c.314T>C; p.(Val105Ala)	0.76 (T)	0.001 (B)	0.0031 (370/120764)
DNAH1 NM_015512.4	c.10438A>G; p.(Ile3480Val)	0.16 (T)	0.459 (PoD)	0.00080 (96/120684)
ILDR1 NM_001199799.1	c.1305C>A; p.(Ser435Arg)	0.68 (T)	0.003 (B)	0.0015 (180/118898)
TAAR8 NM_053278.2	c.575T>C; p.(Ile192Thr)	0.56 (T)	0.014 (B)	0.00085 (103/121396)
TAAR6 NM_175067.1	c.557G>A; p.(Cys186Tyr)	0 (D)	0.985 (PrD)	0.00088 (107/121372)
PTPRZ1 NM_001206838.1	c.1022C>T; p.(Thr341Ile)	0.002 (D)	0.628 (PoD)	0.0000082 (1/121398)
***PDIA4* NM_004911.4**	**c.1645G>A; p.(Glu549Lys)**	**0.008 (T)**	**0.988 (PrD)**	**0.002 (240/121408)**
WDR60 NM_018051.4	c.2398G>A; p.(Ala800Thr)	0.003 (D)	0.992 (PrD)	0.00030 (33/110560)
MADD NM_001135943.1	c.2435G>A; p.(Arg812Gln)	0.15 (T)	0.245 (B)	0.00023 (28/121358)
CTSW NM_001335.3	c.317G>A; p.(Arg106Gln)	0.6 (T)	0.004 (B)	0.00078 (95/121272)
***HSP90B1* NM_003299.2**	**c.877A>G; p.(Met293Val)**	**0.017 (T)**	**0 (B)**	**0.0009 (26/28408)**
SAMD15 NM_001010860.1	c.1618T>A; p.(Phe540Ile)	0.26 (T)	0.428 (B)	0.0029 (353/121042)
PRKD2 NM_001079880.1	c.1457C>T; p.(Pro486Leu)	0.02 (D)	0.029 (B)	0.00085 (102/119736)
NLRP11 NM_145007.3	c.1163A>G; p.(Gln388Arg)	0.71 (T)	0.929 (D)	0.0017 (211/121324)

3.3.2 WES analysis of family #72

The proband of family #72 (F72: II:1) was recruited with a diagnosis of bilateral EBMD and meibomian gland dysfunction. Corneal changes were prominent in the right eye with milder changes in the left eye. He underwent PTK in his right eye following recurrent erosions and responded well to this treatment. He reported that his father had also been diagnosed with EBMD but he was deceased. His (paternal) half-brother (F72: II:2) had also been diagnosed with EBMD and provided a DNA sample via a remotely collected saliva sample which precluded clinical evaluation at MEH (Figure 3.18 B). On the basis of this pedigree, a dominant pattern of inheritance was presumed for this family. The filtering strategy applied was to filter variants for rare (MAF $\leq$ 0.005) variants shared between the two half-siblings, resulting in the identification of four variants (Table 3.5; section 2.6.2.2). The four candidate genes were cross-referenced with the corneal dystophy (Table 1.1) and corneal syndrome candidate gene lists (Appendix B); however, all four candidate genes were absent from both. All were also absent from the corneal proteome (Appendix C).

Table 3.5 Rare (≤ 0.005) shared non-synonymous or splice site variants in two affected half siblings with epithelial basement membrane dystrophy in F72. Variants were filtered for rare shared non-synonymous heterozygous variants (≤ 0.005 MAF). SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
OBSCN NM_001098623.2	c.3806G>A; p.(Arg1269Gln)	0.55 (T)	0.01 (B)	0.00042 (51/120748)
SLC41A3 NM_001164475.1	c.283A>C; p.(Lys95Gln)	0.14 (T)	0.828 (PoD)	0.000058 (7/121390)
PCDHA13 NM_018904.2	c.742A>G; p.(Lys248Glu)	0.26 (T)	0.062 (B)	0.00050 (61/121206)
KLHDC4 NM_017566.3	c.205C>G; p.(Leu69Val)	0.17 (T)	0.062 (B)	0.0031 (370/120406)

3.3.3 Analysis of WES data for proband of family #73

The proband of family #73 was a sporadically affected patient recruited with a diagnosis of EBMD (Figure 3.18 C and D). He also had surface irregularity in this left eye crossing the visual axis and causing blurred vision. This was treated with epithelial debridement and left eye PTK. He subsequently underwent right eye PTK and is currently asymptomatic.

Due to the lack of known family history of corneal disease, recessive disease was initially investigated in the proband. Filtering for rare (MAF ≤ 0.005) homozygous, non-synonymous or splice site variants resulted in the identification of no candidate variants. To investigate potentially partially penetrant dominant disease, WES variants were filtered for rare (MAF ≤ 0.005) variants, resulting in the identification of 446 variants. This was refined by filtering for variants which were also present in the corneal proteome candidate gene list, resulting in the identification of 49 variants (section 2.6.2.4; Appendix C). All were heterozygous with the exception of one hemizygous variant in *JADE3*, located on the X chromosome. No further refinement of these variants could be achieved due to the absence of additional familial samples. If parental samples could be obtained in future, *de novo* dominant inheritance could be explored by identifying variants which are absent from either parent.

To investigate whether the corneal symptoms in the proband could be caused by an alternative inherited corneal disorder, WES variants were filtered for rare (MAF ≤ 0.005) variants in genes from the corneal dystrophy gene list (Table 1.1) and the inherited corneal syndromes gene list (Appendix B), which resulted in the identification of six variants in five candidate genes (Table 3.6; section 2.6.2.3). None of the associated inherited corneal disorders were consistent with the clinical phenotype of the proband and/or inheritance pattern; however, an interesting heterozygous variant was identified in *KRT12*. Mutations in *KRT12* cause the dominantly inherited epithelial corneal dystrophy, MECD, which is associated with corneal epithelial microcysts and acute episodes of corneal pain (Section 3.1.2.4). The c.728C>T; p.(Thr243Ile) variant in exon 3 of *KRT12*, which was confirmed by direct sequencing (section 2.2; Appendix A), was predicted to be 'Damaging' (0.01) and 'Probably Damaging' (0.997) by SIFT and Polyphen2, respectively. The KRT12

protein is also present in the corneal proteome. However, the frequency of this variant in ExAC was 0.0013 (160/121412) and is higher than expected for dominant disease. Manual interrogation of *KRT12* sequencing data using IGV revealed that all *KRT12* coding exons had adequate coverage and did not reveal the presence of a second rare *KRT12* variant that could be causative of a recessive disease.

Table 3.6 Rare (≤ 0.005) non-synonymous or splice site variants in genes causing corneal dystrophies or inherited corneal disorders in F73: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Candidate genes are highlighted in bold.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
LAMC2 NM_005562.2	c.2419A>G; p.(Ser807Gly)	0.01 (D)	0.997 (PrD)	0.000033 (4/121178)	Epidermolysis bullosa (AR)
PLEC NM_201380.3	c.10421C>T; p.(Thr3474Met)	0 (D)	0.306 (B)	0.00010 (12/116700)	Epidermolysis bullosa (AR)
ERCC6 NM_000124.3	c.3122A>C; p.(Gln1041Pro)	0.29 (T)	0 (B)	0.0016 (197/121200)	Cockayne syndrome
	c.1996C>T; p.(Arg666Cys)	0 (D)	0.987 (PrD)	0.0016 (198/121264)	
B3GLCT NM_194318.3	c.1207G>A; p.(Val403Ile)	0.44 (T)	0.038 (B)	0.0012 (141/121410)	Peters-plus syndrome
***KRT12* NM_000223.3**	**c.728C>T; p.(Thr243Ile)**	**0.01 (D)**	**0.997 (PrD)**	**0.003 (160/121412)**	**Meesmann corneal dystrophy**

3.4 Epithelial recurrent erosion dystrophy

ERED is a corneal dystrophy with no primary corneal opacities, but is characterised by frequent recurrent erosions, recently shown to be caused by heterozygous *COL17A1* mutations (Jonsson *et al.*, 2014) (section 3.1.2.2). It is unknown whether ERED is associated with additional genetic heterogeneity. ERED is rarely used as a clinical diagnosis due to the difficulty in differential diagnosis with other causes of corneal erosions; however, two probands were recruited with clinical signs consistent with this diagnosis.

3.4.1 WES analysis of proband of family #92

The proband of family #92 (F92: II:1) first presented at the age of 18 months with episodes of acute corneal pain and photophobia. The presence of corneal microcysts on clinical examination during one episode prompted a diagnosis of MECD. MECD is typically a relatively asymptomatic epithelial microcystic dystrophy; however, individuals may experience acute painful episodes due to the eruption of cysts (section 3.1.2.4). She was treated with alcohol delamination to remove the epithelium and allow the re-epithelialisation of the cornea, however this did not resolve the symptoms. A DNA sample had been previously sent to a non-UCL research laboratory for genetic screening of *KRT3* and *KRT12* genes, and was negative, therefore excluding this clinical diagnosis. She was examined at MEH in 2013 during an asymptomatic period and no microcysts were present at this time. The episodes of acute corneal pain were attributed to the occurrence of recurrent bilateral corneal erosions with unknown aetiology. Subsequently, a corneal erosive episode in the left eye required hospitalisation and may have caused scarring. There is no family history of corneal erosions however previous examinations have revealed a few microcysts in the father of the proband, although these were not observed in the most recent examination at MEH. A diagnosis of ERED was consistent with her symptoms and clinical examination. When questioned about possible skin symptoms, the proband had recently been diagnosed with Keratosis Pilaris and both her father and younger sibling had similar skin symptoms.

Due to the lack of a clear clinical diagnosis in the proband, DNA was sent for WES, to allow all corneal dystrophy genes to be analysed. All corneal dystrophy genes, including the ERED gene *COL17A1*, were adequately covered by WES analysis except from exon 16 of the 55 coding exon *COL17A1* gene, which was subsequently covered by direct Sanger sequencing (section 2.2; Appendix A). WES data was filtered for rare variants (MAF $\leq$ 0.005) in corneal dystrophy genes (Table 1.1) and in genes causing inherited corneal syndromes (Appendix B); resulting in the identification of three variants (Table 3.7; section 2.6.2.3). Only one was associated with clinical diagnosis phenotypically relevant to the proband; epidermolysis bullosa, which is a skin blistering disorder which is characterised by fragile skin, dystrophic nails, and amelogenesis imperfecta. Epidermolysis bullosa is caused by

mutations in multiple genes including *LAMC2*. Interestingly, ERED is caused by heterozygous mutations in *COL17A1*; however, biallelic loss-of-function mutations in *COL17A1* also cause epidermolysis bullosa (Jonsson *et al.*, 2014).

Table 3.7 Rare (≤ 0.005) non-synonymous or splice site variants in genes causing corneal dystrophies or inherited corneal disorders in F92: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Candidates are highlighted in bold.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
***LAMC2* NM_005562.2**	**c.2153_2160del; p.(Asp718Glufs*14)**	**N/A**	**N/A**	**N/A**	**Epidermolysis bullosa (AR)**
TRPV3 NM_145068.3	c.2185C>T; p.(Arg729*)	N/A	N/A	0.00039 (47/121400)	Olmstead syndrome (AD)
XPC NM_004628.4	c.872C>G; p.(Ser291Cys)	0.03 (D)	0.968 (PrD)	0.0028 (340/120064)	Xeroderma pigmentosum (AR)

The variant identified in F92: II:1 was a novel heterozygous frameshift mutation c.2153_2160del; p.(Asp718Glufs*14) in the *LAMC2* gene. Although WES filtering revealed no other potentially pathogenic mutations in the coding region or splice sites of the *LAMC2* gene, a rare variant was identified in intron 15 by manual visualisation of the WES data using IGV which is located outside of the usual cut-off for the WES data annotation. The variant c.2301+25T>G (rs751561812) is also absent from both EVS and 1000 Genomes Project, and is present in the ExAC database at a frequency of 0.0000255 (3/117416).

Segregation analysis revealed that the proband inherited the c.2153_2160del; p.(Asp718Glufs*14) variant from her father, and the c.2301+25T>G variant from her mother, therefore demonstrating that the variants are present on alternate alleles (Figure 3.19; section 2.2; Appendix A). Corneal erosions occur in a quarter of patients with the clinical subtype of epidermolysis bullosa caused by *LAMC2* (Fine *et al.*, 2004). Due to the presence of two rare variants in the *LAMC2* gene, it was

therefore hypothesised that the intronic variant may be affecting splicing of the *LAMC2* transcript, or influence transcript levels. To test this hypothesis in silico, splice prediction tools NetGene, Human Splice Finder and NNSplice were used (section 2.3.5). The presence of the variant did not appear to affect the predicted strength of the canonical splice donor site (Table 3.8). This does not exclude a possible role for the c.2153_2160del; p.(Asp718Glufs*14) variant that was inherited from the father.

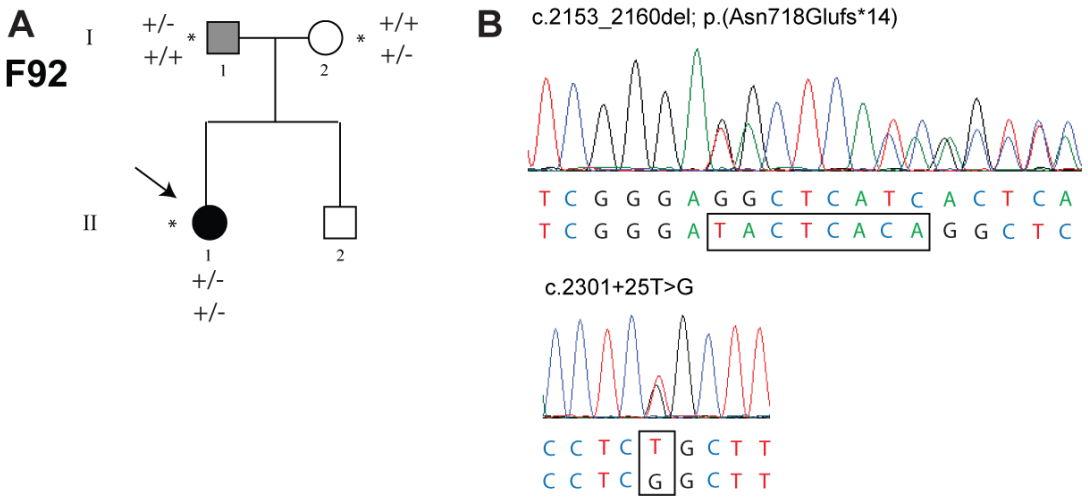


Figure 3.19 Biallelic *LAMC2* variants in a proband of family #92 with recurrent epithelial erosions. (A) Pedigree structure, with proband F92: II:1, shown with an arrow. The proband had experienced recurrent erosions from age of 18 months, occasionally with the presence of microcysts. The proband's father occasionally presented with microcysts but experienced no erosions. * indicates individual has been examined and a blood sample available. +/- indicates heterozygosity. **(B)** Whole exome sequencing analysis revealed the presence of a novel c.2153_2160del; p.(Asp718Glufs*14) variant in *LAMC2*. A second rare variant c.2301+25T>G was identified in an intron, and is of unknown significance. Variants are present on alternate alleles.

Table 3.8 In silico prediction of a *LAMC2* c.2301+25T>G variant in F92:II:1 on splice donor preference. The c.2301+25T>G variant was not predicted to influence splice donor preference by three independent splice prediction tools. All three provide a score (0-100) on the likely strength of the canonical splice donor site. Netgene predicted a small reduction in strength whereas NNSplice and Human Splice Finder predicted no change. No other potential splice site was introduced by the presence of the variant.

	WT canonical splice donor site score	**Canonical donor site score in presence of c.2301+25T>G variant**
NNSplice	0.99	0.99
Netgene	0.60	0.54
Human Spice Finder	70.88	70.88

3.4.2 WES analysis for proband of family #93

The proband of family #93 (F93: II:1) was recruited to the study with previous diagnoses of bilateral recurrent erosion syndrome and bilateral dry eye disease. She was in her fifties, of Indian ethnicity, and had no known family history of corneal disease. She experienced a gritty sensation in both eyes with photophobia and episodes of severe pain. Subsequent follow up appointments revealed lid margin disease and nocturnal lagophthalmos. She was referred for dermatological assessment for skin legions. She underwent laser PTK in her left eye for recurrent erosions with no improvement in symptoms one year post-intervention.

DNA from the proband was pre-screened for mutations in *TGFBI* to exclude a *TGFBI*-associated corneal dystrophy, and was analysed by WES following negative pre-screening (section 2.2; 2.6). Variants were filtered for rare non-synonymous or splice site variants (MAF $\leq$ 0.005) in the corneal dystrophy (Table 1.1) and corneal syndrome candidate gene lists (section 2.6.2.3; Appendix B), resulting in the identification eight rare variants in seven genes (Table 3.9). Three of these genes (*LUM, CTNS, LOXHD1*) were excluded due to lack of associated clinical symptoms in the proband. Four of the genes have been previously associated with corneal

erosions including the ERED associated gene *COL17A1*. The presence of the *COL17A1* c.3190G>A; p.(Val1064Ile) variant was confirmed by direct sequencing (section 2.2; Appendix A). Interestingly, a rare variant in *LAMC2* was also identified in the splice donor site of exon 16. All three *in silico* splice prediction programs used predicted that the variant would abolish the splice acceptor/donor site and an alternative +192_193 GT site within intron 16 would be used preferentially (section 2.3.5; Table 3.10). Biallelic variants in *PLEC* and a heterozygous variant in *SPINT2* were also considered to be potential candidates. Biallelic mutations in *PLEC* also cause epidermolysis bullosa while biallelic mutations in *SPINT2* cause congenital sodium diarrhoea, which is also associated with episodes of corneal erosions (Varki *et al.*, 2006; Heinz-Erian *et al.*, 2008). No other family members were available and therefore segregation analysis could not be performed in this family to eliminate any of these four candidate genes.

Table 3.9 Rare (≤ 0.005 MAF) non-synonymous or splice site variants in F93: II:1 in genes causing corneal dystrophies and inherited corneal disorders. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Candidate genes are highlighted in bold.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
LAMC2 **NM_005562.2**	**c.2456+1G>T**	**N/A**	**N/A**	**N/A**	**Epidermolysis bullosa (AR)**
PLEC **NM_201380.3**	**c.7631G>A; p.(Arg2544Gln)**	**0.35 (T)**	**0.084 (B)**	**20/109558**	**Epidermolysis bullosa (AD or AR)**
	c.6010C>T; p.(Arg2004Trp)	**0.01 (D)**	**0.996 (PrD)**	**15/95302**	
COL17A1 **NM_000494.3**	**c.3190G>A; p.(Val1064Ile)**	**0.21 (T)**	**0.979 (PrD)**	**44/121356**	**Epidermolysis bullosa (AR)** **Epithelial recurrent erosion dystrophy (AD)**
LUM NM_002345.3	c.968C>T; p.(Pro323Leu)	0.59 (T)	0.158 (B)	163/121058	Posterior amorphous corneal dystrophy
CTNS NM_001031681.2	c.356G>A; p.(Arg119His)	0.48 (T)	0.001 (B)	415/121406	Cystinosis (AR)
LOXHD1 NM_144612.6	c.6523G>A; p.(Val2175Met)	0.03 (D)	0.243 (B)	N/A	Fuch's endothelial corneal dystrophy
SPINT2 **NM_021102.3**	**c.689G>A; p.(Arg230His)**	**0.12 (T)**	**0.425 (B)**	**163/119144**	**Congenital sodium diarrhoea (AR)**

Table 3.10 In silico prediction of the effect of a *LAMC2* c.2456+1G>T variant identified in F93: II:1 on splice donor preference. Introduction of a c.2456+1G>T variant was predicted to abolish the canonical donor site by two independent splice prediction tools but was not recognised by Human Splice Finder. All three programs predict the same strongest alternative donor site to be located at the +192-193 GT site within intron 16.

	WT canonical splice donor site score	**Canonical donor site score in presence of c.2456+1G>T**	**Highest scoring alternative donor site (+192/193)**
NNSplice (0-1)	0.93	N/A	0.95
Netgene (0-1)	0.88	N/A	0.46
Human Spice Finder (0-100)	N/A	N/A	88.24

It is interesting to note that both probands with an ERED-like clinical phenotype recruited to the study and analysed by WES were found to harbour novel heterozygous loss-of-function mutations in *LAMC2*. Interrogation of the ExAC database reveals that loss-of-function mutations in *LAMC2* are rare, with the most prevalent loss-of-function mutation having a frequency of 0.00004 (5/120768). While other probands in the cohort analysed by WES were identified to harbour rare variants in *LAMC2*, these were rare missense changes as opposed to putative loss-of-function variants. It is therefore possible that loss of one copy of *LAMC2* confers a susceptibility to developing corneal erosions, potentially interacting with other genetic (including variants in other genes involved in the hemidesmosome complex) and environmental factors such as trauma. However, further research is necessary to prove a causative role for *LAMC2* in ERED and similar dystrophies.

3.5 Gelatinous drop-like dystrophy

3.5.1 Large consanguineous family from Colombia

GDLD is a rare autosomal recessive corneal dystrophy associated with deposition of amyloid in sub-epithelial layers of the cornea (section 3.1.2.3). A large consanguineous family (F94) consisting of 5 affected and 5 unaffected siblings was

recruited from Universidad de Antioquia, Colombia (Figure 3.20 A). Neither parent was affected. The proband (F94: II:10) was the youngest of the affected siblings and had experienced symptoms of photophobia from the first year of life with corneal opacity developing aged 3. When recruited to the study at ten years of age, central corneal opacity and peripheral superficial neovascularization was causing a significant reduction in visual acuity, prompting superficial keratotomy. All affected siblings had symptoms of photophobia, foreign body sensation and blurred vision. Two affected siblings (II:3 and II:5), examined in their early twenties, also shared the same characteristic mulberry appearance and peripheral superficial neovascularization as the proband (Figure 3.20 D). However, siblings II:8 and II:9, aged 14 and 11 years, were less severely affected than their siblings with polymorphic grey sub-epithelial band-shaped opacities without neovascularization, with late epithelial stain following instillation of fluorescein (Figure 3.20 D). Additionally, two of the siblings (II:3 and II:8) had peri-limbal deposits suggestive of amyloid (Figure 3.20 D).

Mutations in the *TACSTD2* gene, a single coding exon gene on chr1p32.1, cause GDLD. Screening the *TACSTD2* gene by PCR amplification and direct sequencing (section 2.2; Appendix A) revealed the presence of a homozygous c.551A>G; p.(Tyr184Cys) mutation in the proband II:10 (Figure 3.20 B). The mutation segregated fully in the eleven individuals for whom DNA samples were available from the pedigree, with all affected siblings (II:3, II:8, II:5, II:9) homozygous for the mutation and both parents (I:1 and I:2) identified as heterozygous carriers. Unaffected siblings (II:6, II:7, II:11) and the unaffected son of II:3 (III:1) were either heterozygous for the mutation or wild type (Figure 13.20 A, B). No other variants were identified in the *TACSTD2* gene.

Alignment of orthologous TACSTD2 protein sequences demonstrates that the non-synonymous variant, c.551A>G; p.(Tyr184Cys) alters a highly conserved residue (Figure 3.20 C). Using SIFT and Polyphen2 prediction tools, the c.551A>G; p.(Tyr184Cys) substitution is predicted to be Damaging (0) and Probably Damaging (0.994), respectively. This mutation has been identified previously in a GDLD patient of Chinese ethnicity; the mutation was identified in the heterozygous state in combination with the Japanese founder mutation c.352C.T; p.(Gln118*) (Tian *et al.*,

2004). The c.551A>G; p.(Tyr184Cys) mutation is present in the ExAC database at an overall frequency of 0.0000093 (1/107098). The variant was identified in the heterozygous state once in the Colombian sub-population within the 1000 Genomes Project. A total of 93 people from Medellin, Antioquia were included in this cohort, therefore the frequency of the mutation in this sub-population is 0.53%.

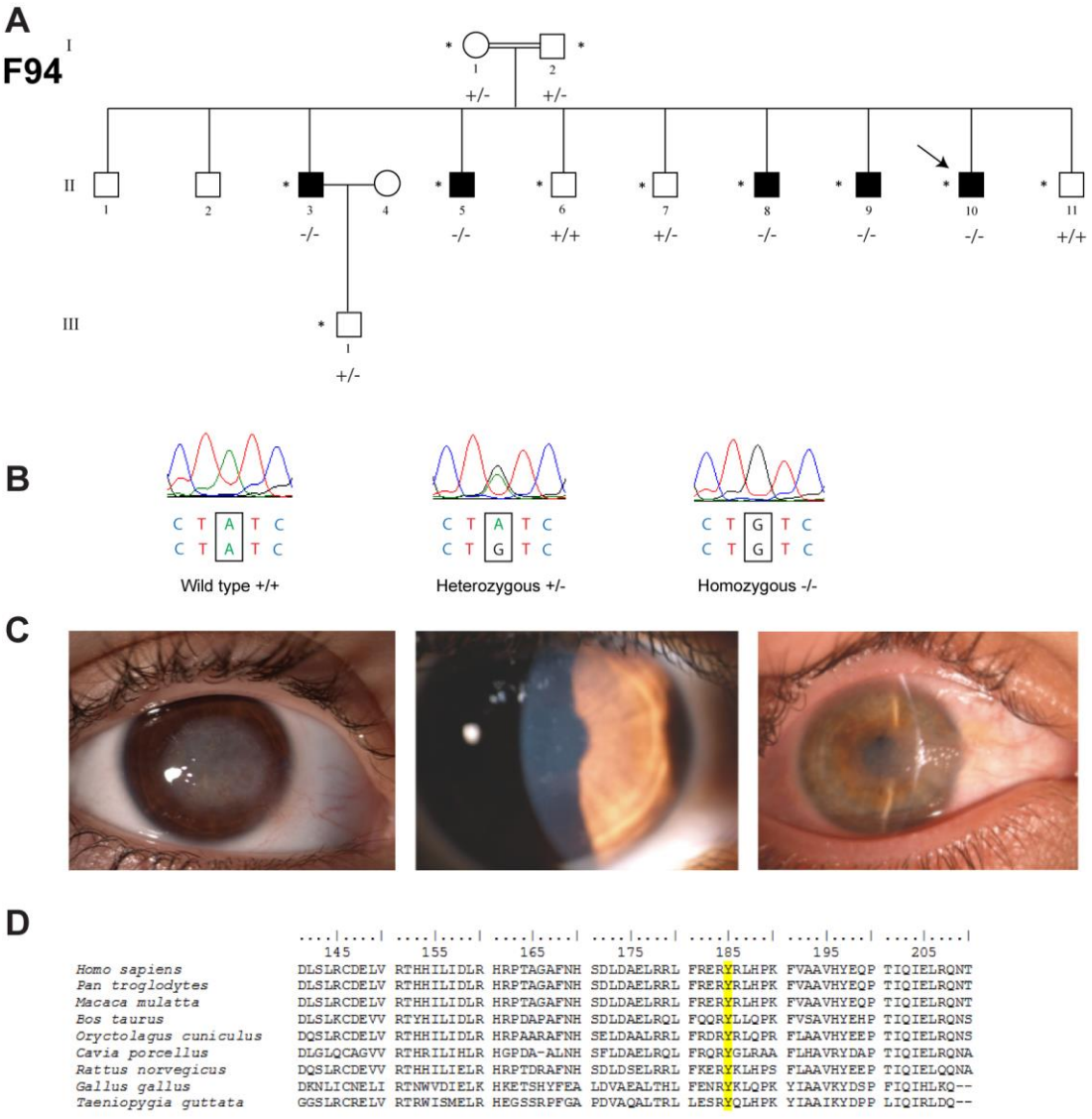


Figure 3.20 Identification of a homozygous *TACSTD2* c.551A>G; p.(Tyr184Cys) mutation in a Colombian family F94 with gelatinous drop like dystrophy. (A) Pedigree showing segregation of the mutation in each individual tested. * indicates DNA sample available for testing. Proband is shown with an arrow. +/- indicates heterozygous and -/- homozygous for the mutation. **(B)** Sequence electropherogram for each genotype. Mutation is highlighted with an arrow. **(C)** Representative clinical phenotypes of affected family members. (Left) classic severe mulberry phenotype in the proband II:10, (middle) milder band keratopathy phenotype in II:8 and (right) peri-limbal amyloid in individual II:3. **(D)** ClustalW alignment of mammalian and avian orthologous *TACSTD2* sequences showing conservation of the Tyrosine (Y) residue, highlighted in yellow.

3.5.2 Proband with gelatinous drop-like dystrophy from the UK

The proband (F95) was recruited in her mid-fifties from MEH. She had a previous medical history of a childhood ocular infection and had previously undergone

multiple keratoplasties in both eyes due to corneal opacities of unknown aetiology. When examined at MEH, she demonstrated vascularisation with nodular lesions and secondary lipid keratopathy. The proband's parents were not affected but she reported a sibling with similar symptoms who had not been clinically examined at MEH. A diagnosis of GDLD was considered most consistent with the clinical symptoms and likely recessive inheritance pattern.

Screening *TACSTD2* in the proband F95: II:1 identified the presence of compound heterozygous mutations; c.355T>A; p.(Cys119Ser) and c.661delG; p.(Asp221Metfs*50) (section 2.2; Appendix A). The c.355T>A; p.(Cys119Ser) mutation has previously been reported to be disease causing and is predicted to be Damaging (0) by SIFT and 'Probably Damaging' (1) by Polyphen2. It is absent from ExAC and is conserved in orthologous protein sequences (Figure 3.21 D). The c.661delG; p.(Asp221Metfs*50) is also absent from ExAC and is a novel frameshift mutation. As a frameshift mutation that introduces a premature stop codon, this is likely to be a null allele. A summary diagram showing the location of the *TACSTD2* mutations identified is shown in Figure 3.22.

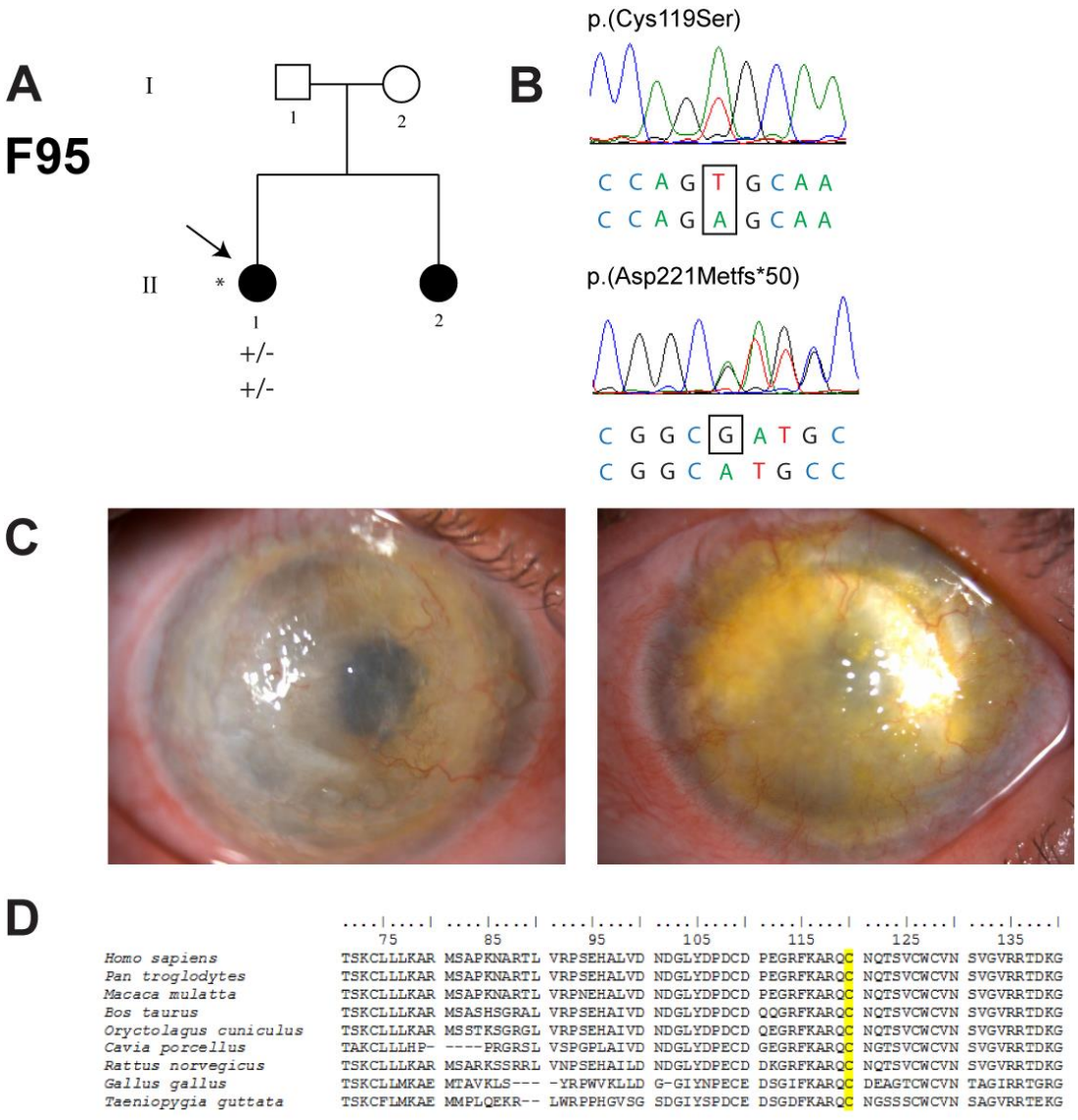


Figure 3.21 Compound heterozygous *TACSTD2* mutations in a British family (F95) diagnosed with gelatinous drop like dystrophy. (A) Pedigree structure of family showing affected siblings and unaffected parents. Proband indicated with an arrow. +/- indicates person is heterozygous for the mutation. **(B)** Compound heterozygous mutations c.355T>A; p.(Cys119Ser) and c.661delG; p.(Asp221Metfs*50) identified by Sanger sequencing of *TACSTD2* in the proband. **(C)** Clinical appearance of corneal dystrophy in the proband. **(D)** Alignment of orthologous residues showing conservation of the Cys-119 residue.

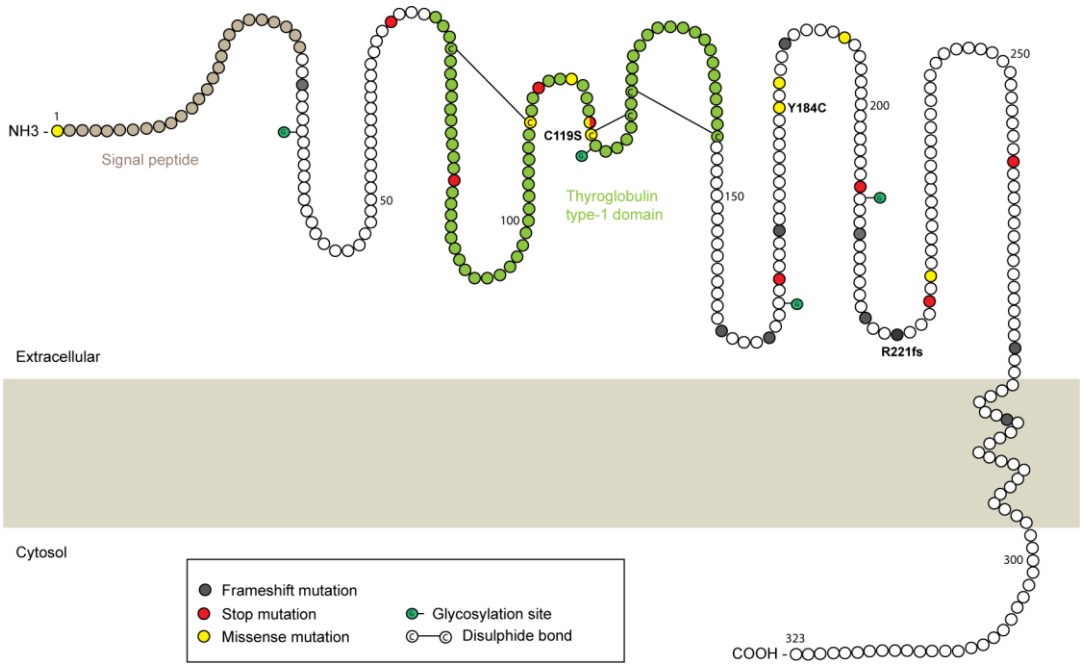


Figure 3.22 Topology diagram of the TACSTD2 protein showing location of novel and previously identified mutations. The protein has 323 amino acids and includes an extracellular, transmembrane and cytosolic domain. The extracellular domain contains a Thyroglobulin type-1 domain with three disulfide bridges. Many of the missense mutations involve the destruction or creation of cysteine residues. Missense mutations p.(Cys119Ser) and p.(Tyr184Cys), and a novel frameshift p.(Asp221Metfs*50), found in this study, are labelled.

3.6 Macular corneal dystrophy

3.6.1 MCD cohort

MCD is an autosomal recessive dystrophy caused by mutations in the single exon *CHST6* gene on chr16q23.1 (section 3.1.3.1). Probands from 20 families (F96 – 115) were recruited from MEH with a diagnosis of MCD, and a total of 27 DNA samples from extended pedigrees were obtained for analysis. For all probands recruited, only a single generation of affected individuals was documented or no known family history of corneal disease was described, which is consistent with an autosomal recessive pattern of inheritance. All probands had a characteristic MCD clinical phenotype with diffuse amorphous white stromal opacities (Figure 3.23).

PCR amplification of three overlapping fragments encompassing the single coding exon of *CHST6* and Sanger sequencing was performed for all 20 recruited MCD probands (F95 – F115) (section 2.2; Appendix A). Individuals without biallelic mutations in the coding exon of *CHST6* were subsequently analysed for upstream deletions and rearrangements using a PCR-based assay that had previously been designed to detect the two previously reported structural mutations (section 2.5). This uses a combination of six primers in four PCR reactions to determine whether a recurrent ~ 40 kb deletion has occurred upstream of *CHST6*, or whether a ~ 2.5 kb region upstream of *CHST6* is replaced with a homologous region upstream of neighbouring gene, *CHST5*.

Figure 3.23 Clinical appearance of macular corneal dystrophy (MCD) in the study cohort. The majority of patients diagnosed with MCD had white, cloudy amorphous stromal deposits. Images shown are probands of F108, F103 and F109, respectively.

3.6.2 CHST6 coding mutations

Screening the single coding exon of *CHST6* resulted in the identification of previously reported biallelic missense or nonsense mutations in 11 out of 20 families (F96 – 106; Tables 3.11 and 3.12). Families harbouring novel mutations are discussed in detail below.

3.6.2.1 Compound heterozygous mutations in family #107

CHST6 sequence analysis of the proband of family #107 (F107: II:3) resulted in the identification of a heterozygous mutation in *CHST6*: c.827T>C; p.(Leu276Pro) that has been previously identified as causative of MCD (El-Ashry *et al.*, 2005) (Figure 3.24 B). Screening for recurrent upstream *CHST6* disruptions in this proband was

negative, removing the possibility of a previously reported structural mutation on the alternate allele. Two other heterozygous missense variants were present in the *CHST6* gene: c.484C>G, p.(Arg162Gly) and c.656C>T; p.(Ala219Val). The former has a MAF of 0.03572 (4253/ 119070) in the ExAC database and includes 99 homozygotes who would be expected to be affected if this variant was disease-causing. This variant has previously been described in MCD patients; however due to its high frequency it was considered to be a non-pathogenic missense variant (Gruenauer-Kloevekorn *et al.*, 2008). In contrast, the c.656C>T, p.(Ala219Val) variant has a frequency of 0.0002774 (32/115358) and is therefore considerably rarer. DNA was available from two other family members: an affected sibling II:2 and an unaffected niece III:2, therefore segregation of these variants were tested. The c.827T>C; p.(Leu276Pro) mutation and c.656C>T; p.(Ala219Val) variant were both present in the affected sister, II:2, in the heterozygous state. The proband's niece III:2 was a carrier of the c.656C>T; p.(Ala219Val) variant but not the c.827T>C; p.(Leu276Pro) mutation, indicating that these variants are present on different alleles (Figure 3.24 A). Based on segregation in this family, it is likely that the c.656C>T; p.(Ala219Val) variant is the second pathogenic mutation. Although it is present in ExAC, the occurrence of the mutation in the control population at a low frequency is not unexpected for alleles associated with recessive conditions. The SIFT and Polyphen2 scores for the p.(Ala219Val) mutation are 'Tolerated' (0.52) and 'Benign' (0.107), respectively although interestingly, this is similar to the prediction scores for pathogenicity of other known pathogenic *CHST6* missense mutations (Table 3.10). Furthermore, the affected amino acid residue Ala-219 is conserved in orthologous primate *CHST6* sequences (Figure 3.24 C).

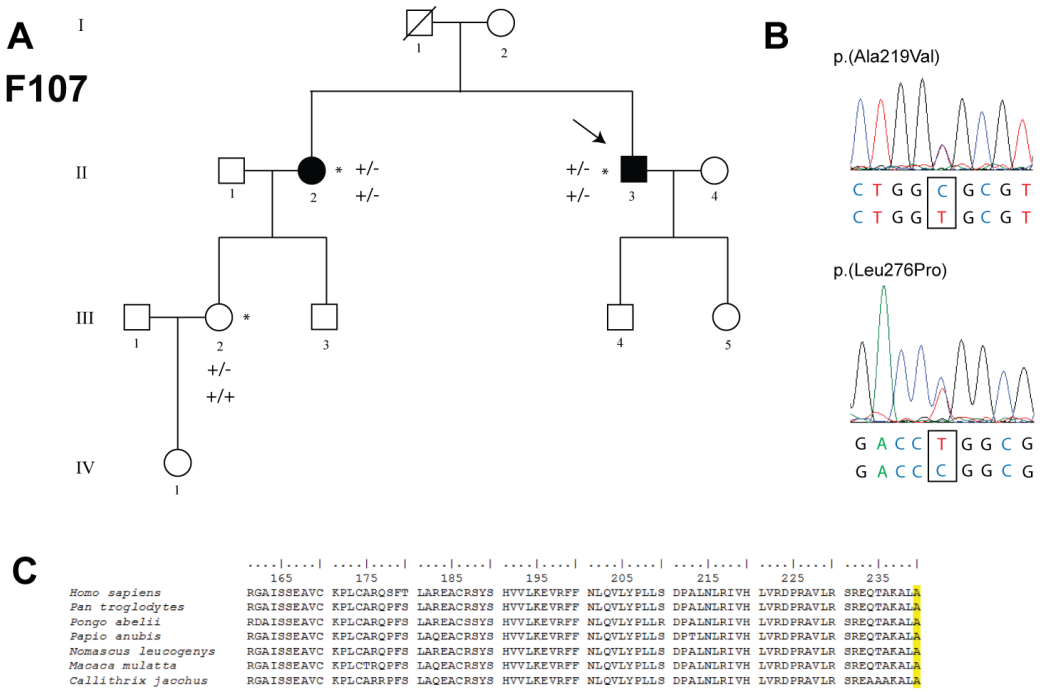


Figure 3.24 Segregation of *CHST6* compound heterozygous mutations in family #107 with macular corneal dystrophy (MCD). (A) Pedigree structure showing inheritance of MCD in the family and segregation of two heterozygous variants. +/+ indicates that individual is homozygous for wild type and +/- is heterozygous. * indicates individual has been examined and a blood sample taken. Proband II:3 is indicated with an arrow. **(B)** Sequence electrophoretogram of two heterozygous variants identified by PCR amplification and Sanger sequencing of the proband F107: II:3. The c.827T>C; p.(Leu276Pro) variant is a previously identified pathogenic mutation whereas c.656C>T; p.(Ala219Val) is a rare (MAF ≤ 0.005) missense variant of unknown pathological significance. **(C)** Alignment of the CHST6 protein sequence with ClustalW shows conservation of the amino acid residue 219 in primates.

3.6.2.2 Novel homozygous missense mutation in families #108 and #109

The proband of family #108 of white British ethnicity (F108: II:1) was recruited with MCD with no family history of disease. The proband of family #109 (F109: II:1) was also recruited with MCD and no known family history of disease (Figure 3.25 A). There was no evidence that the families were related. Screening of *CHST6* identified the presence of a homozygous c.997T>C; p.(Trp333Arg) mutation in both probands (Figure 3.25 B). The mutation is present at very low frequency in the ExAC database (1/119996) and is predicted to be Damaging (0) and Probably Damaging (1) by SIFT and Polyphen2, respectively. The mutation affects a conserved amino acid residue (Figure 3.25 C), which has been previously associated

with a different MCD causative mutation; the c.997T>G; p.(Trp333Gly) identified in a compound heterozygous state with another pathogenic mutation (Liu *et al.*, 2010).

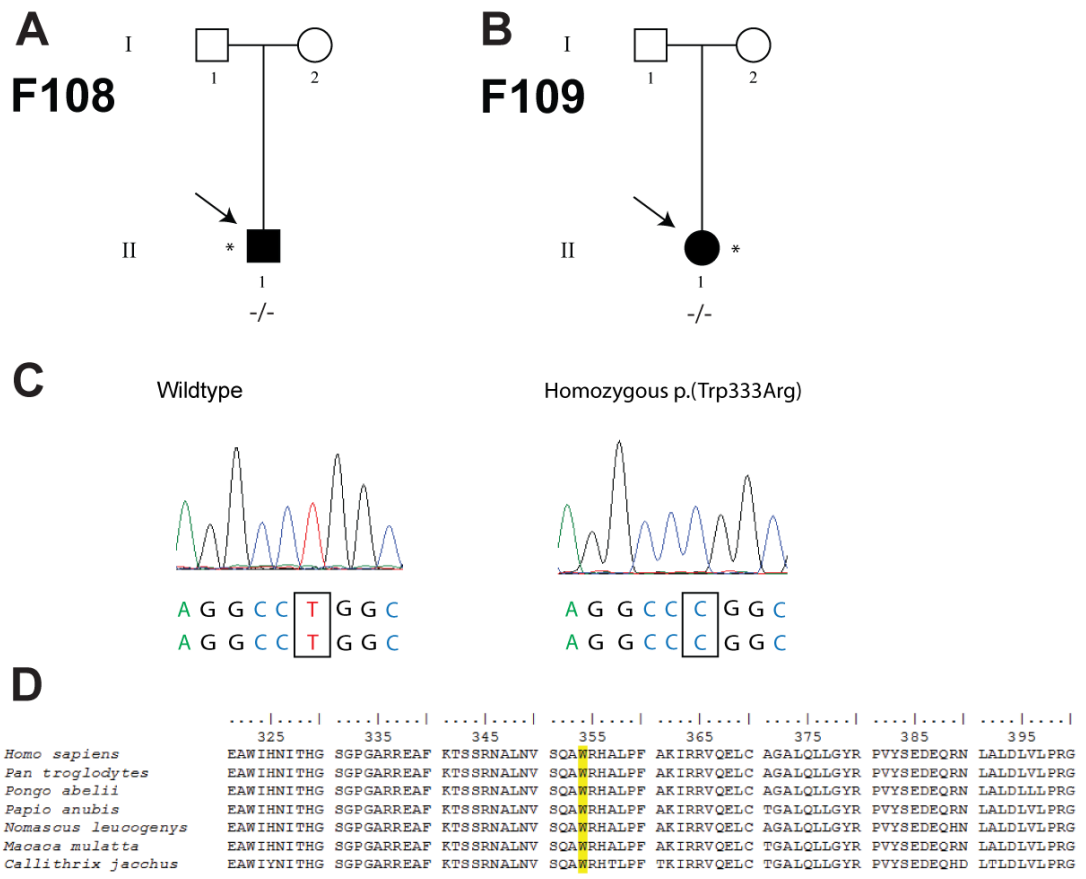


Figure 3.25 Identification of a novel homozygous *CHST6* mutation c.997T>C; p.(Trp333Arg) in two unrelated families (F108, F109). Pedigrees are shown that demonstrate a lack of family history for both probands. **(A)** F108. **(B)** S109. There is no evidence that the two probands were related. **(C)** Sequence electropherogram showing wild-type sequence and a homozygous c.997T>C; p.(Trp333Arg) mutation found in both probands. **(D)** Alignment of orthologous primate protein sequences shows the amino acid residue of interest is conserved.

3.6.2.3 Novel homozygous frameshift mutation in family #110

The proband of family #110 (F110: II:2; Figure 3.26 A) was diagnosed with MCD, and recruited with no family history of corneal disease; however, subsequently the proband's younger brother (II:3) was found to have corneal changes consistent with MCD. Screening *CHST6* identified the presence of a novel homozygous frameshift mutation due to a small duplication c.250dupA: p.(Ser84Lysfs*25) (Figure 3.26 B). This variant is absent from ExAC and as a frameshift that incorporates a premature

stop codon, the mutation is likely to be a loss of function mutation responsible for disease in the proband.

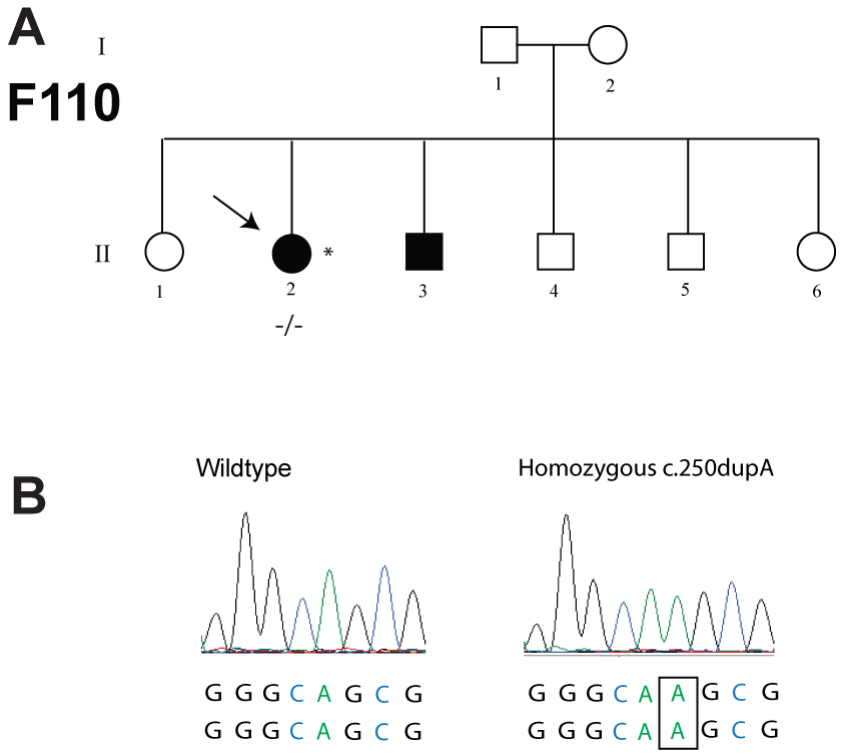


Figure 3.26 Identification of a novel duplication in *CHST6* causing MCD in family #110. (A) Pedigree structure of F110. **(B)** Sequence electropherogram showing presence of a single adenine residue in the wild-type (WT) sequence and a homozygous duplication of this residue in the proband, leading to a frameshift c.250dupA; p.(Ser84Lysfs*25).

3.6.3 CHST6 upstream deletions

In cases where screening the single coding exon of *CHST6* did not result in the identification of biallelic mutations in *CHST6*, a PCR based assay was used to screen for recurrent structural mutations upstream of the gene (section 3.1.3.1). Briefly, the PCR assay consists of six primers in four PCR reactions that determine the integrity of a critical region upstream of *CHST6* (section 2.5). When no mutation is present the PCR will amplify a product for F1/R1 and F2/R2 primers only. When a ~ 40 kb deletion is present a product will be amplified by FM/R2M and when the ~2.5 kb region upstream of *CHST5* replaces the equivalent region of *CHST6*, products will be present in the FM/R2M and F1M/R1 reactions. In the presence of a structural

mutation the F1/R1 and F2/R2 products will fail to amplify, provided both alleles are disrupted (Figure 3.30).

3.6.3.1 MCD family #111

The proband of family #111 (F111: II:1) was diagnosed with MCD, however there was no known information on inheritance of the dystrophy in the family. Screening the *CHST6* single coding exon revealed a single novel heterozygous *CHST6* nonsense mutation, c.1089dupT; p.(Glu364*), which is highly likely to be pathogenic. It is not present in the ExAC database and is likely to produce a transcript which is degraded by nonsense-mediated decay or produce a truncated protein if translated. No other rare heterozygous variants were identified by direct sequencing that could explain the patient's MCD therefore screening for structural mutations was undertaken. Using the strategy described two PCR products of the expected sizes were observed for the F1/R1 and F2/R2 reactions, indicating that at least one allele was not disrupted. No products were amplified for FM2/R2 and F1/R1M reactions, excluding the known upstream deletion or rearrangement as a cause of disease (Figure 3.27 A). However, in addition to amplification of the expected F2/R2 product, an additional smaller amplicon was detected. This product was gel extracted and directly sequenced, which indicated that a 444 bp deletion had occurred between the F2 and R2 primers. The breakpoints of the deletion occur at identical 8 bp sequences (GAAAACC) flanking the deletion (Figure 3.27 B). Due to lack of familial samples, it could not be identified whether the upstream deletion occurred on the same, or different allele to the c.1089dupT; p.(Glu364*) mutation, however given that the individual has MCD it is likely to be the second causative allele.

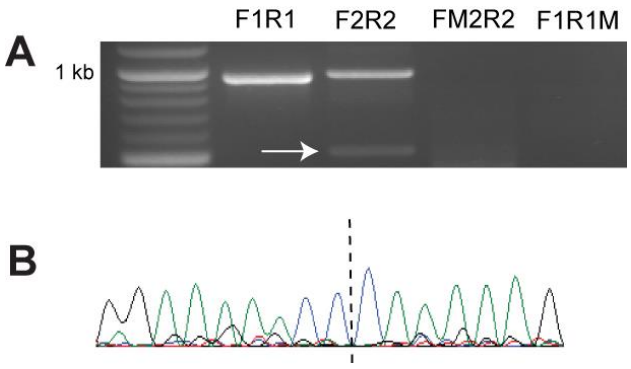


Figure 3.27 Identification of a novel deletion upstream of the *CHST6* gene in family #111. (A) PCR amplification of 4 amplimers in an assay to detect a recurrent ~40 kb deletion or ~2.5 kb replacement mutation upstream of the *CHST6* gene identified a second smaller amplimer in using F2/R2 primers, shown with an arrow. **(B)** Sequence electropherogram define the breakpoints of the deletion as.chr16.hg38: 75,499,098 –75,499,541del.

3.6.3.2 MCD family #112

The proband of family #112 was recruited with no known family history of disease. Screening the *CHST6* coding exon by direct screening was negative for any potentially pathogenic mutations. Analysis of the upstream region in proband F112: II:1 revealed PCR products with primer combinations F1/R1, F2/R2 and F2M/R2 (Figure 3.28 A). The F1/R1 and F2/R2 amplimers indicate that at least one allele is not disrupted for this region upstream of the *CHST6* gene. The F2M/R2 amplicon is indicative that a previously described ~40 kb deletion was present (Akama *et al.*, 2000). Gel extraction and sequencing of the product confirmed a deletion as chr16.hg38:75,499,407 – 75,539,615del (section 2.2; Figure 3.28 B). In the original paper describing this ~ 40 kb recurrent deletion the PCR product was not sequenced therefore the deletion breakpoints cannot be compared to those published, although it is assumed to be similar in size and span due to its amplification with the F2M/R2 primers (Akama *et al.*, 2000). Due to the lack of a second *CHST6* mutation, this patient remains unsolved with no heterozygous coding variants present in the coding region of the *CHST6* gene, however it is likely that a second mutation occurs outside the regions tested in this assay.

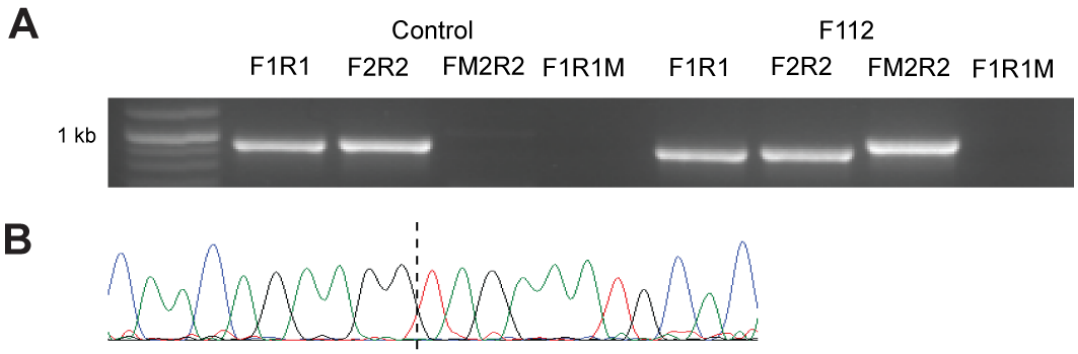


Figure 3.28 Identification of a ~ 40 kb deletion upstream of *CHST6* in F112: II:1. (A) Presence of a band with the F2M/R2 primers indicates the occurrence of a ~ 40kb deletion of the *CHST5* gene and the *CHST6* critical upstream region. The alternate allele is not affected by other rearrangements as demonstrated by the presence of the F1/R1 and F2/R2 bands. Sequencing across the breakpoint confirmed the presence of a deletion with co-ordinates chr16.hg38:75,499,407 – 75,539,615del.

3.6.3.3 MCD family #113

The proband of family #113 (F113: II:2; Figure 3.29 A) was not included in the original MCD cohort. He presented at MEH with an unusual clinical phenotype; he was, on different occasions, diagnosed with both keratoconus and PPCD due to steep corneas and changes of the corneal endothelium and Descemet's membrane. Stromal clouding was also noted and considered to be a consequence of an endothelial dystrophy. His daughter (III:1) was not clinically examined but previously was noted to have changes of Descemet's membrane and was diagnosed with the endothelial dystrophy CHED. Clinical photos were not available for either the proband or his affected daughter therefore these diagnoses could not be reviewed.

Known PPCD genes *ZEB1* and *COL8A2* were screened for potentially pathogenic mutations, but no rare variants were identified (section 2.2; Appendix A). DNA from the proband was subsequently analysed by WES (section 2.6). Aligned WES data was filtered for rare (MAF ≤ 0.005) heterozygous variants, assuming an autosomal dominant inheritance on the basis of the affected daughter. Variants were filtered for rare (≤ 0.005) non-synonymous or splice site variants in genes associated with corneal dystrophies (Table 1.1) or inherited corneal disease (section 2.6.2.3; Appendix B). This filtering strategy revealed a novel heterozygous variant in the *CHST6* gene, c.505A>C; p.(Ser169Arg), which was confirmed by Sanger

sequencing (Figure 3.29 B). The variant was predicted to be 'Probably Damaging' (0.991) by Polyphen2 but 'Tolerated' (0.48) by SIFT. The variant occurs at a conserved residue (Figure 3.28 C), and is absent from the ExAC database. As a single heterozygous mutation in *CHST6* is not sufficient to develop MCD, the proband was then also screened for *CHST6* upstream rearrangements that would not be detected by WES. A second smaller amplimer was identified using the F2/R2 primer combination (Figure 3.29 C), similar in size to the amplimer previously been identified in F111: II:1. Sanger sequencing of the gel extracted product confirmed that the proband had inherited the same novel recurrent 444 bp deletion. It was possible that both *CHST6* changes were inherited on the same allele, therefore segregation analysis was performed using DNA samples from the unaffected parents (I:1 and I:2) of the proband. The novel c.505A>C; p.(Ser169Arg) mutation was inherited from I:2, whereas the *CHST6* upstream deletion was absent from both the proband's mother and father indicating it is a *de novo* change (Figure 3.29 E). The proband's daughter could not be contacted for clinical examination and DNA analysis. Although the functional consequences of this deletion have not been fully resolved, its proximity to other known deleterious rearrangements and occurrence in more than one unrelated MCD patient indicate that it is highly likely to be pathogenic as a compound heterozygous allele or in the homozygous state. However, it is currently not clear if this upstream deletion would prevent any functional protein being expressed in the cornea.

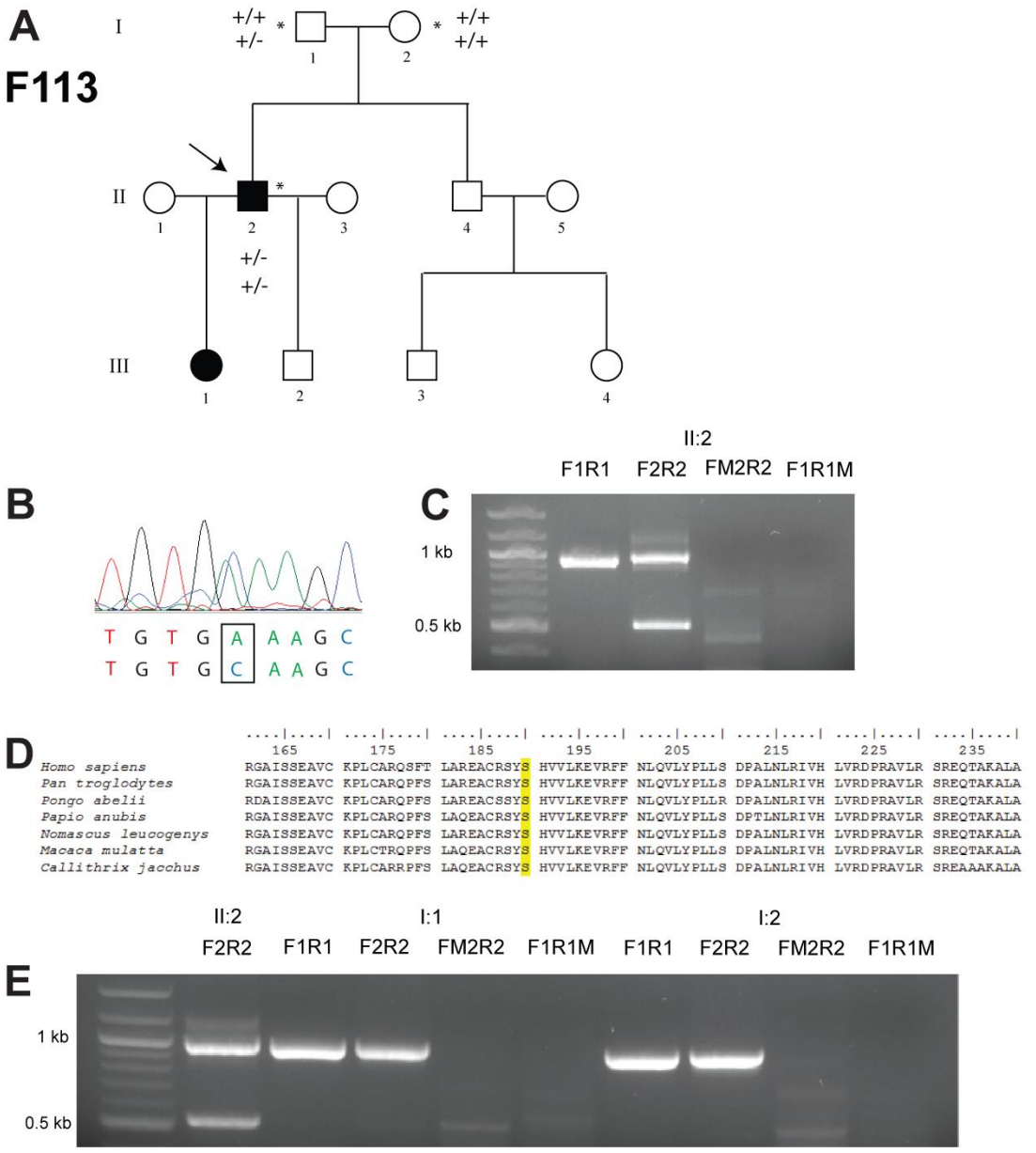


Figure 3.29 Diagnosis of macular corneal dystrophy following whole exome sequencing in family #113. (A) Pedigree showing presumed autosomal dominant inheritance of an irregular corneal dystrophy phenotype with features of both macular corneal dystrophy and an endothelial dystrophy. Segregation of two *CHST6* variants is shown; a novel missense change and a novel upstream deletion. -/- indicates homozygous for the mutation, +/+ homozygous wild type and +/- heterozygous. * indicates individual has been examined and a blood sample taken. The proband II:2 is shown with an arrow. **(B)** A novel heterozygous *CHST6* mutation, c.505A>C; p.(Ser169Arg) was identified by whole exome sequencing of the proband, and confirmed by Sanger sequencing, shown with an arrow. **(C)** Screening of the upstream region of *CHST6* revealed the presence of a deletion of unknown pathogenic significance in the proband. **(D)** ClustalW alignment shows conservation of the affected residue in orthologous primate *CHST6* sequences. **(E)** The *CHST6* upstream deletion was not inherited from either parent and must have occurred *de novo*.

Probands from families #114 and #115 were both negative for upstream *CHST6* disruptions (data not shown) and remain unsolved, with a single coding *CHST6* mutation identified in each individual (Table 3.11).

3.6.4 Summary of CHST6 mutations

In total, fourteen *CHST6* coding mutations were identified that have been previously shown to be disease-causing, and five novel coding mutations that were deemed highly likely to be pathogenic, were identified (Table 3.11). Predictions of pathogenicity for all *CHST6* mutations identified and their MAF in the ExAC database, are shown in Table 3.12, although in some cases known mutations are predicted to be benign/tolerated. A novel deletion upstream of the *CHST6* gene which overlaps with a known pathogenic deletion was also identified in two probands. For the majority of MCD probands, *CHST6* sequence analysis and screening for upstream deletions and rearrangements identified two likely pathogenic *CHST6* mutations, either in the homozygous (10 probands) or compound heterozygous (8 probands) state, however two probands were not fully genetically solved (Table 3.11). Proband F112: II:1 has a heterozygous upstream ~ 40 kb deletion, but the *CHST6* coding exon was negative for any coding mutations. A single heterozygous coding *CHST6* mutation was identified in F115: II:1, and the upstream region was negative for known disruptions. Diagrammatic summaries of the coding and structural mutations identified are shown in Figure 3.30 and Figure 3.31. Retrospectively, it was determined that families F96, F97, F99, F100 and F104 have been previously been described (El-Ashry *et al.*, 2002).

Table 3.11 Summary of *CHST6* mutations identified in the macular corneal dystrophy cohort. Biallelic mutations were identified for the majority of probands (18/20). The majority of mutations were private to the affected family within the cohort.

Family ID	*CHST6* allele 1	*CHST6* allele 2
F96	c.213_214delGCinsTT; p.(Glu71_Pro72delinsAspSer)	c.213_214delGCinsTT; p.(Glu71_Pro72delinsAspSer)
F97	c.217C>T; p.(Ala206Val)	c.217C>T; p.(Ala206Val)
F98	c.392C>T; p.(Ser131Leu)	c.629C>T; p.(Ser210Phe)
F99	c.320T>C; p.(Phe107Ser)	c.320T>C; p.(Phe107Ser)
F100	c.91C>T; p.(Pro31Ser)	c.599T>G; p.(Leu200Arg)
F101	c.379C>T; p.(Arg127Cys)	c.379C>T; p.(Arg127Cys)
F102	c.611C>A; p.(Pro204Gln)	c.629C>T; p.(Ser210Phe)
F103	c.599T>G; p.(Leu200Arg)	c.599T>G; p.(Leu200Arg)
F104	c.91C>T; p.(Pro31Ser)	c.599T>G; p.(Leu200Arg)
F105	c.545delA; p.(Gln182Argfs*199)	c.545delA; p.(Gln182Argfs*199)
F106	c.494_495delGCinsCT; p.(Cys165Ser)	c.494_495delGCinsCT; p.(Cys165Ser)
F107	c.827T>C; p.(Leu276Pro)	c.484C>G; p.(Ala219Val)
F108	c. 997T>C; p.(Trp333Arg)	c. 997T>C; p.(Trp333Arg)
F109	c. 997T>C; p.(Trp333Arg)	c. 997T>C; p.(Trp333Arg)
F110	c.250dupA; p.(Ser84Lysfs*25)	c.250dupA; p.(Ser84Lysfs*25)
F111	c.1089dupT; p.(Glu364*)	chr16.hg38: 75,499,098 - 75,499,541del
F112	chr16.hg38: 75,499,407 - 75,539,615del	unknown
F113	c.505A>C; p.(Ser169Arg)	chr16.hg38: 75,499,098 - 75,499,541del
F114	c.6_7delinsAA; p.(Trp2*)	unknown
F115	c.614G>A; p.(Arg205Gln)	unknown

Table 3.12 Assessment of pathogenicity of *CHST6* nonsense, frameshift and missense mutations in the macular corneal dystrophy cohort. All missense variants were checked for potential pathogenicity using pathogenicity prediction tools SIFT (1-0) and Polyphen2 (0-1), and frequency in ExAC database. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Novel mutations are highlighted in bold.

Nucleotide	Protein	SIFT (1-0)	Polyphen2 (1-0)	ExAC	Ref.
c.6_7delinsAA	p.(Trp2*)	N/A	N/A	3/19222	(Sultana *et al.*, 2005)
c.91C>T	p.(Pro31Ser)	0.76 (T)	0.054 (B)	NP	(Klintworth, Smith and Bowling, 2006)
c.213_214delGCinsTT	p.(Glu71_Pro72delinsAspSer)	N/A	N/A	NP	(El-Ashry *et al.*, 2002)
c.250dupA	**p.(Ser84Lysfs*25)**	**N/A**	**N/A**	**NP**	
c.320T>C	p.(Phe107Ser)	0.41 (T)	0.998 (PrD)	6/119826	(Sultana *et al.*, 2003)
c.379C>T	p.(Arg127Cys)	0.01 (D)	1 (PrD)	NP	(Warren, Aldave, *et al.*, 2003)
c.392C>T	p.(Ser131Leu)	0.19 (T)	0.986 (PrD)	5/119530	(Klintworth, Smith and Bowling, 2006)
c.494_495delGCinsCT	p.(Cys165Ser)	0 (D)	1 (PrD)	NP	(Sultana *et al.*, 2005)
c.505A>C	**p.(Ser169Arg)**	**0.48 (T)**	**0.991 (PrD)**	**NP**	
c.545delA	p.(Gln182Argfs*199)	N/A	N/A	NP	(Sultana *et al.*, 2005)
c.599T>G	p.(Leu200Arg)	0 (D)	1 (PrD)	31/117292	(Niel *et al.*, 2003)
c.611C>A	p.(Pro204Gln)	0 (D)	1 (PrD)	2/116362	(Sultana *et al.*, 2005)
c.614G>A	p.(Arg205Gln)	0 (D)	1 (PrD)	2/115940	(Warren, Aldave, *et al.*, 2003)
c.617C>T	p.(Ala206Val)	0 (D)	0.999 (PrD)	NP	(El-Ashry *et al.*, 2005)

c.629C>T	p.(Ser210Phe)	0 (D)	0.999 (PrD)	2/115952	(Sultana *et al.*, 2005)
c.656C>T	**p.(Ala219Val)**	**0.52 (T)**	**0.107 (B)**	**32/115358**	
c.827T>C	p.(Leu276Pro)	0.01 (D)	0.992 (PrD)	6/118222	(Sultana *et al.*, 2005)
c.997T>C	**p.(Trp333Arg)**	**0 (D)**	**1 (PrD)**	**1/119996**	
c.1089dupT	**p.(Glu364*)**	**N/A**	**N/A**	**1/121656**	

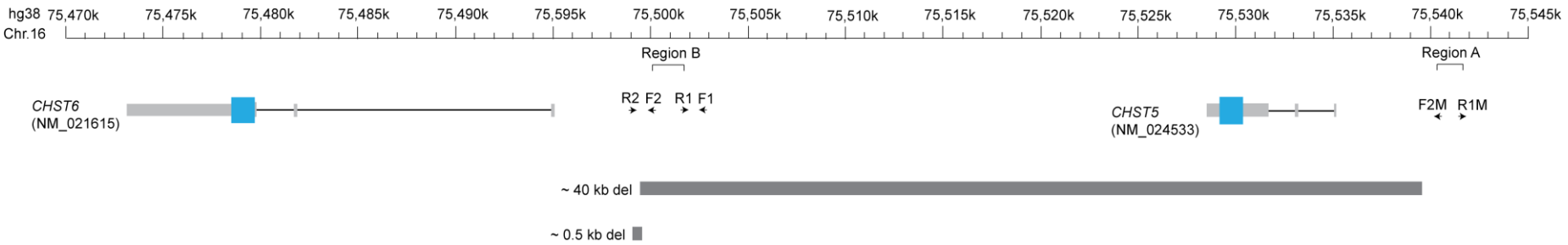


Figure 3.30 Schematic showing relative locations of *CHST6* and *CHST5* genes (hg38) and location of two deletions identified in three probands with macular corneal dystrophy. The location of primers used to screen for upstream deletions are also shown.

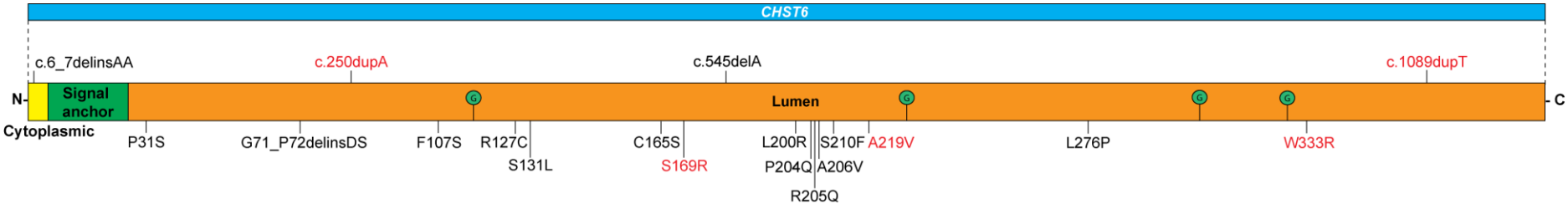


Figure 3.31 Schematic showing structure of *CHST6* gene and the encoded single pass transmembrane protein. Mutations identified in the MEH cohort are labelled; mutations in red are novel.

3.7 Schnyder corneal dystrophy

SCD is an autosomal dominant corneal dystrophy caused by heterozygous mutations in the *UBIAD1* gene located on chr1p36.22. Approximately half of patients have discrete crystalline deposits and diagnosis of SCD may be delayed in individuals presenting without crystals (Weiss, 1996) (section 3.2.3.2). A total of six probands were recruited from MEH with a clinical diagnosis, or suspected diagnosis, of SCD and one additional individual in which the diagnosis of SCD was made following WES analysis. Two probands with SCD were recruited from the Czech Republic. All probands recruited with a diagnosis of SCD were screened for mutations in the two coding exons of *UBIAD1* by PCR amplification and Sanger sequencing (section 2.2; Appendix A).

3.7.1 SCD families #116 and #117

Two unrelated probands diagnosed with SCD (F116: II:1 and F117: II:2; Figure 3.32 A and B) were found to carry the same recurrent causative mutation, a heterozygous c.361C>T; p.(Leu121Phe) mutation in exon 1 of *UBIAD1* (Figure 3.32 C). This mutation has previously been identified as a pathogenic mutation causing SCD in one British, one American and one Saudi Arabian family (Weiss, Kruth, *et al.*, 2008; Al-Ghadeer, Mohamed and Khan, 2011). The mutation is absent from ExAC, predicted to be 'Damaging' (0.01) and 'Probably Damaging' (0.950) by SIFT and Polyphen2, respectively (Table 3.13), and the affected leucine residue is fully conserved in orthologous sequences (Figure 3.32 F). Proband F115: II:1 reported no known family history of corneal disease but was unable to provide information regarding maternal relatives. The parents of the proband were not clinically examined and their DNA samples were not available for testing. The parents of proband F116: II:2 had not been diagnosed with a corneal dystrophy but she reported an affected sibling. Both probands presented with the crystalline phenotypic variant of SCD, but F116: II:2 had significant corneal arcus, possibly due to the age difference between the two probands (F116: II:1, aged 12, and F117: II:2, aged 59) (Figure 3.32 D and E).

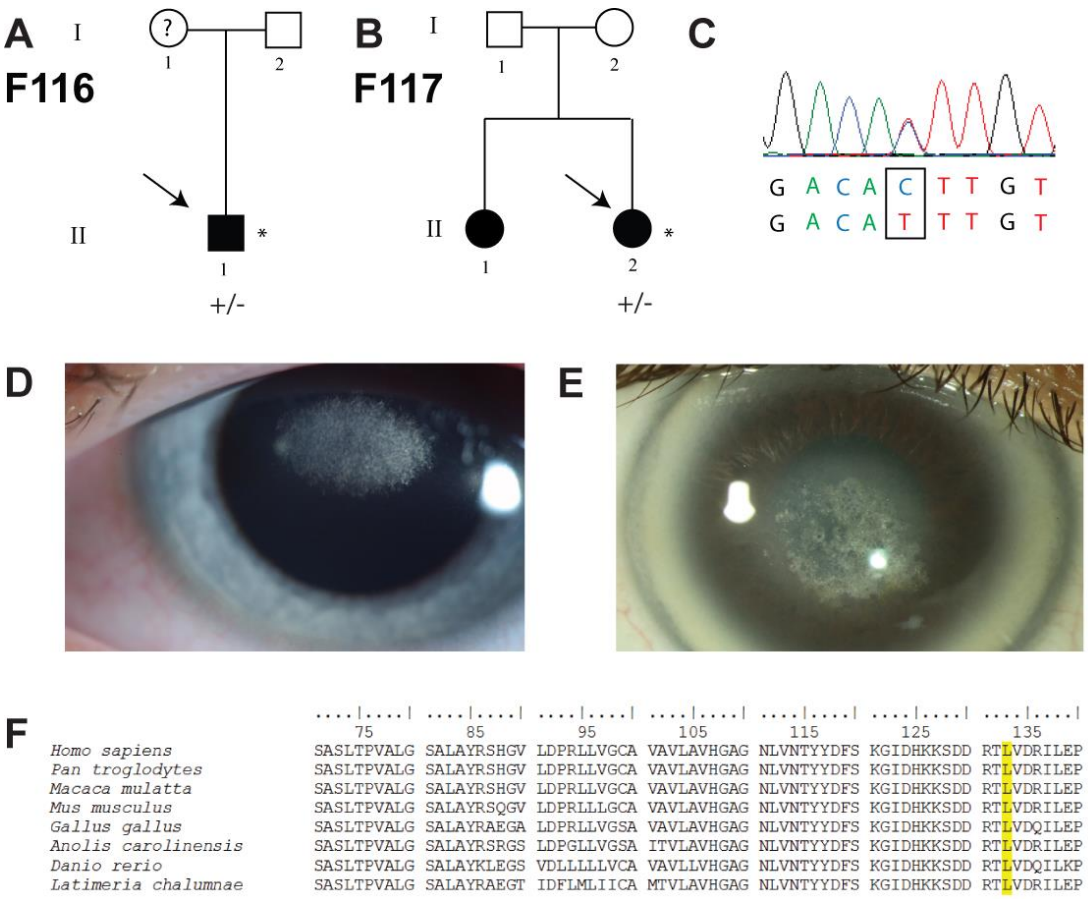


Figure 3.32 Identification of a recurrent *UBIAD1* c.361C>T; p.(Leu121Phe) mutation in unrelated probands F116: II:1 and F117: II:2 with Schnyder corneal dystrophy. * indicates patient has been examined and a blood sample taken. +/- indicates individual is heterozygous for the mutation. **(A)** Pedigree structure of family F116. No family history was indicated however the proband's parents were not available for clinical examination. **(B)** Pedigree structure of F117. **(C)** Sequence electropherogram showing identification of c.361C>T; p.(Leu121Phe) mutation in both probands **(D)** Clinical image of F116: II:1 showing crystalline deposition **(E)** Clinical image of F117:II:2 showing crystalline deposition and corneal arcus **(F)** ClustalW alignment of *UBIAD1* orthologous sequences shows the affected Leu-121 residue is fully conserved.

3.7.2 SCD family #118

The proband of family #118 (F118: III:2) presented with bilateral corneal clouding affecting the entire thickness of the stroma, prominent corneal arcus and peripheral crystalline deposits with a disease onset in early twenties (Figure 3.33 C and D). The proband had a family history of corneal clouding on the paternal side (Figure 3.33 B). She had originally been diagnosed with familial crocodile shagreen; however, on clinical examination at MEH the phenotype was consistent with SCD. Sequencing

two *UBIAD1* coding exons revealed a novel heterozygous missense variant c.308C>T; p.(Thr103Ile) (Figure 3.33 B). The variant was absent from the ExAC database and is predicted to be 'Damaging' (0) and 'Probably Damaging' (0.993) by SIFT and Polyphen2, respectively (Table 3.13). Alignment of orthologous protein sequences revealed the residue is highly conserved (Figure 3.33 E). Subsequently, the same mutation has been associated with SCD patient in which it occurred *de novo* (Lin *et al.*, 2016).

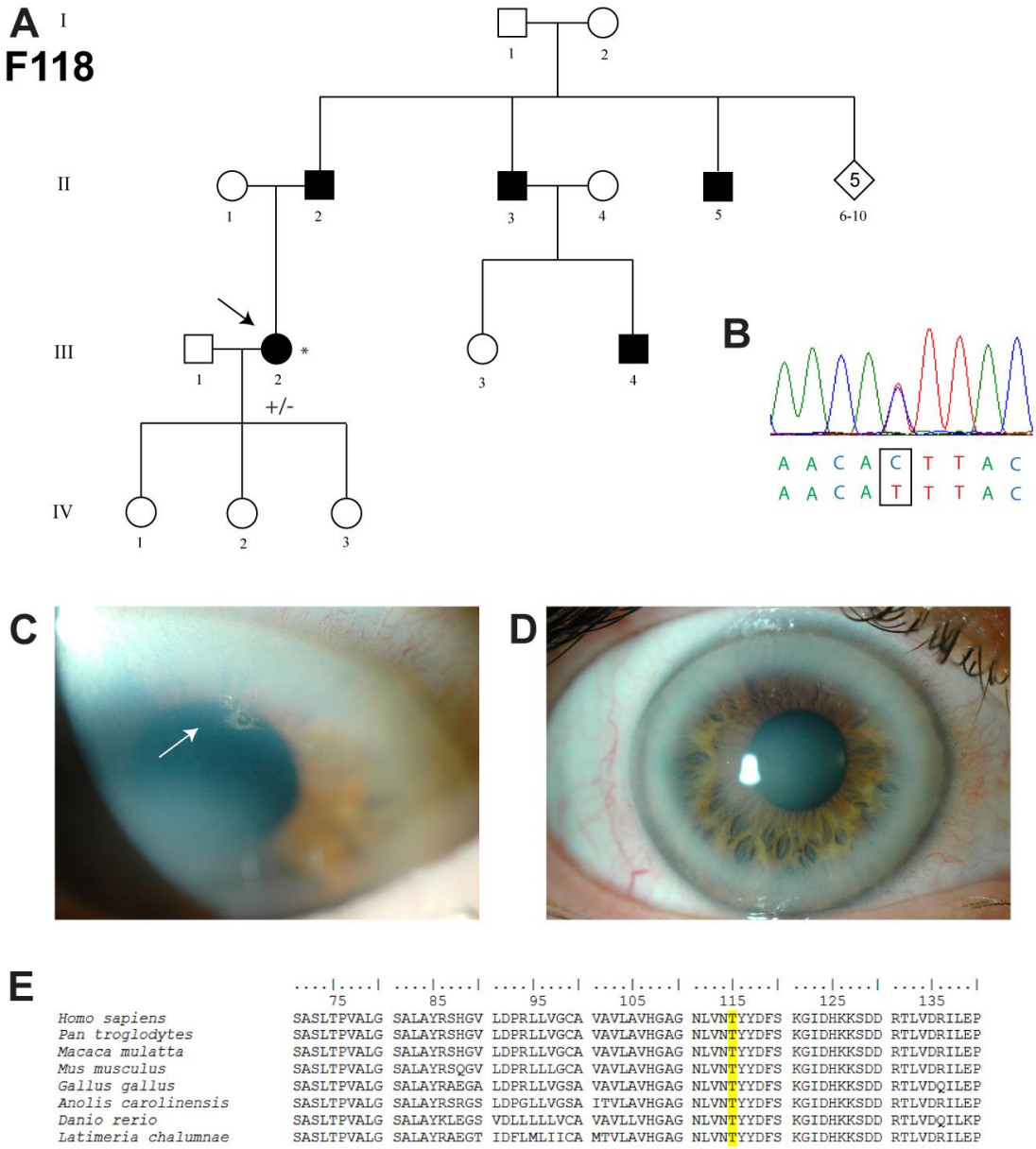


Figure 3.33 Identification of a c.308C>T; p.(Thr103Ile) mutation in *UBIAD1* in family #118 with Schnyder corneal dystrophy. * indicates patient has been examined and a blood sample taken. +/- indicates individual is heterozygous for the mutation. **(A)** Pedigree structure shows at least two generations of individuals affected with Schnyder corneal dystrophy. **(B)** Sequence electropherogram showing presence of a heterozygous *UBIAD1* mutation: c.308C>T; p.(Thr103Ile) in the proband. **(C)** Clinical images of proband showing small peripheral crystalline deposition and **(D)** prominent corneal arcus. **(E)** Alignment of orthologous protein sequences shows that Thr-103 is a conserved residue.

3.7.3 SCD families #119 – #121

Proband F119: III:1 was recruited with suspected SCD due to corneal clouding (Figure 3.34 A). An inherited syndrome was also considered possible due to

additional symptoms in the proband including knee deformities and learning difficulties. Family history was positive for corneal clouding and knee deformities in the proband's affected sister (Figure 3.34 B). The proband's mother and grandmother were deceased however there was photographic evidence of corneal clouding in the mother. Screening *UBIAD1* revealed the presence of a c.305G>A; p.(Asn102Ser) mutation in exon 1 (Figure 3.34 E); this is the most frequently reported *UBIAD1* mutation. There is no evidence for a founder effect with this mutation, and it has been described in families originating from different populations, indicating that it arose on multiple occasions and is therefore a mutation hot-spot. The mutation affects a conserved residue (Figure 3.34 E), is absent from ExAC and is predicted to be 'Damaging' (0) and 'Probably Damaging' (0.999) by SIFT and Polyphen2, respectively (Table 3.13).

Proband F120: III:1 was clinically diagnosed with SCD due to the presence of central crystalline deposits, however she had been previously diagnosed with RBCD. She had a family history of SCD for at least three generations (Figure 3.34 C). The proband harboured the same mutation c.305G>A; p.(Asn102Ser) mutation in *UBIAD1*.

The proband of F121: III:3 was recruited to the study with a clinical diagnosis of LCD and had been treated with bilateral penetrating keratoplasties prior to being recruited to the study. His family history was positive for affected individuals in at least three generations (Figure 3.34 D), consistent with an autosomal dominant pattern of inheritance, however sequencing *TGFBI* and *GSN* exon 4 was negative for any potentially pathogenic mutations (section 2.2; Appendix A). Given the autosomal dominant family history, a genetic cause of the corneal opacities was considered likely therefore the proband's DNA was analysed by WES (section 2.5). Filtering for rare non-synonymous or splice site variants (MAF $\leq$ 0.005) in genes associated with corneal dystrophies (Table 1.1) and inherited corneal disorders (Appendix B) revealed the presence of a c.305G>A; p.(Asn102Ser) mutation in *UBIAD1*, providing a genetic diagnosis of SCD.

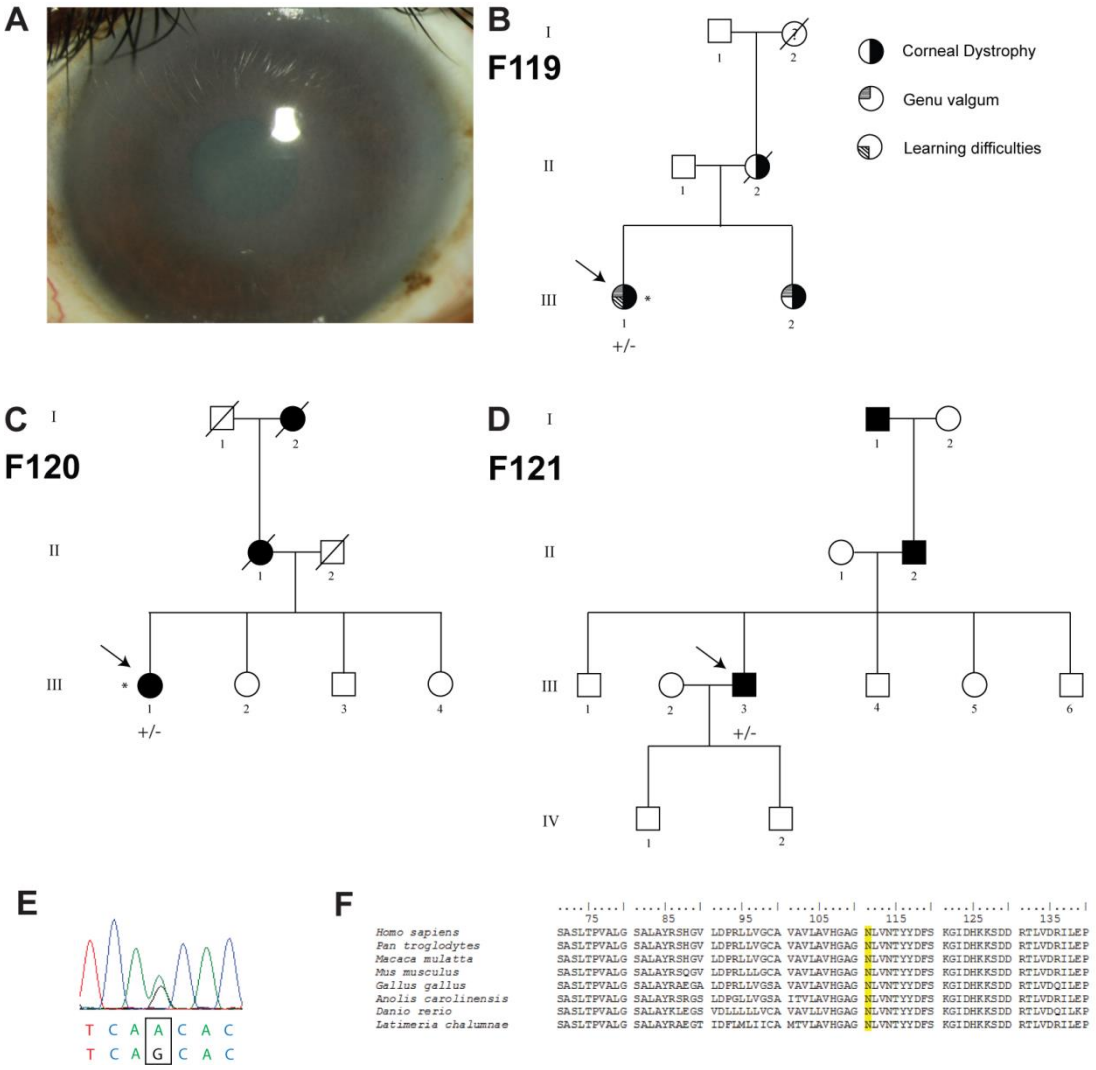


Figure 3.34 Identification of a *UBIAD1* mutation c.305G>A; p.(Asn102Ser) in three unrelated families with Schnyder corneal dystrophy (SCD).* indicates patient has been examined and a blood sample taken. +/- indicates individual is heterozygous for the mutation. **(A)** Clinical image showing central clouding in proband F119: III:1. **(B)** Pedigree structure of family #119 with suspected SCD and additional symptoms **(C)** Pedigree structure of #120 showing autosomal dominant inhertiance of SCD. **(D)** Pedigree structure of #121 showing autosomal dominant inhertiance of SCD. **(E)** Sequence electropherogram showing presence of a c.305G>A; p.(Asn102Ser) *UBIAD1* mutation. **(F)** Conservation of affected Asn-102 residue in orthologous protein sequences.

3.7.4 SCD family #122

Proband F122: IV:2 was recruited from the Czech Republic with SCD, aged in his thirties. Clinical examination demonstrated sub-epithelial central and mid peripheral crystals in a circular pattern. There was minimal corneal arcus (Figure 3.35 C). Family history was positive for corneal dystrophy for at least three generations,

consistent with an autosomal dominant pattern of inheritance; however, DNA samples were not obtained from other family members (Figure 3.35 A). Serum lipid blood screening revealed borderline high total cholesterol (4.89 mmol/l) and increased triglycerides (2.05 mmol/l). Screening *UBIAD1* revealed the presence of a novel mutation: c. 527G>A; p.(Gly176Glu) (Figure 3.35 B). The mutation was absent from the ExAC database and was predicted to be 'Damaging' (0) and 'Probably Damaging' (0.992) by SIFT and Polyphen2, respectively (Table 3.13). Alignment shows complete conservation of the residue in orthologous protein sequences (Figure 3.35 D).

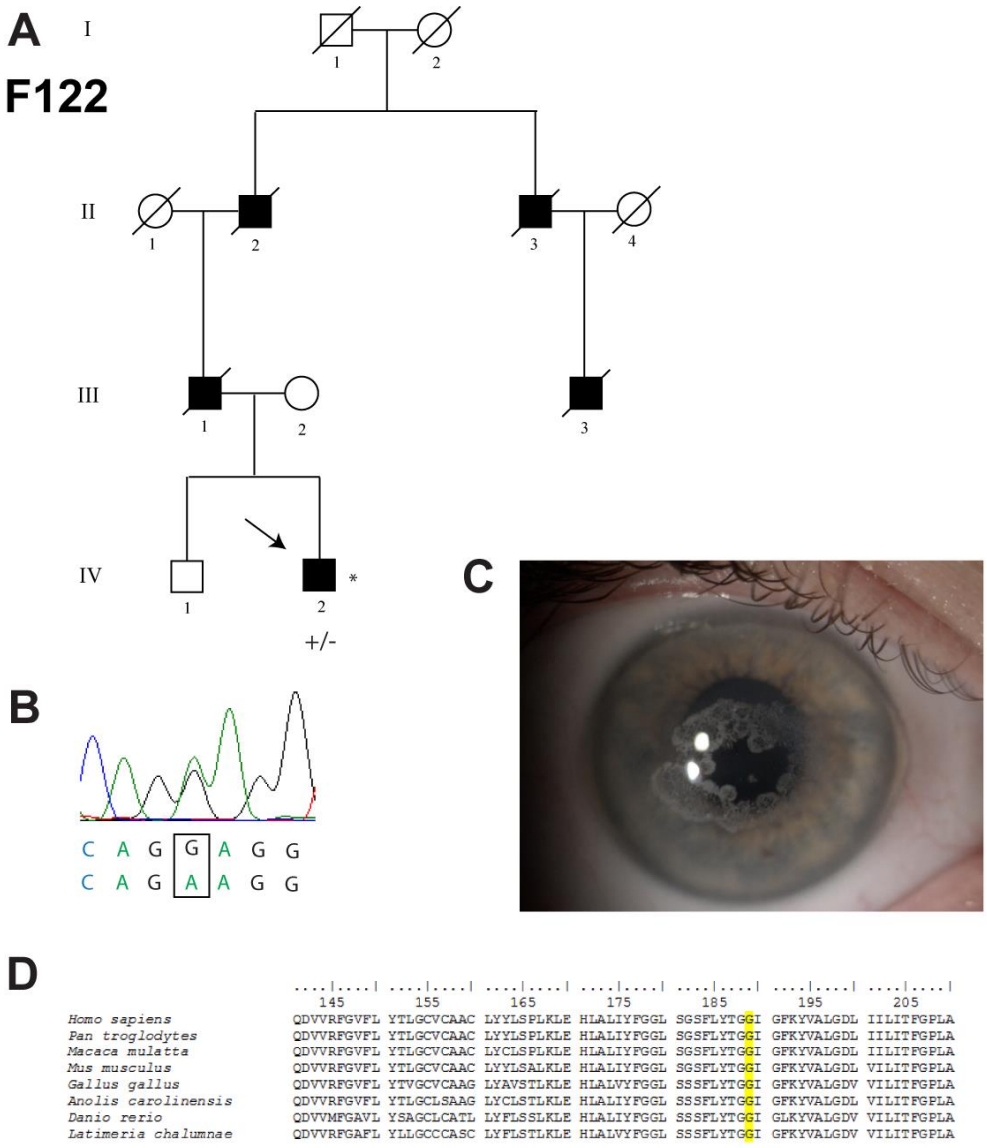


Figure 3.35 Identification of a novel *UBIAD1* mutation in a Czech family F122 with Schnyder corneal dystrophy. (A) Pedigree showing autosomal dominant inheritance of the disease in the family. * indicates patient has been examined and a blood sample taken. +/- indicates individual is heterozygous for the mutation. **(B)** Sequence electropherogram showing presence of a novel *UBIAD1* mutation c. 527G>A; p.(Gly176Glu). **(C)** Clinical phenotype of the proband F120: IV:2 showing extensive crystalline deposition. **(D)** Alignment of orthologous sequences showing complete conservation of the Gly-176 amino acid residue.

3.7.5 SCD family #123

DNA was obtained from a proband F123: II:1 of Czech origin diagnosed with SCD, who was recruited at six years of age. She exhibited very few sub-epithelial and mid peripheral crystals in a ring-like pattern (Figure 3.36 A). There was no history of

corneal disease in the family history (Figure 3.36 B). Screening of *UBIAD1* exons 1 and 2 revealed the presence of a *UBIAD1* mutation in exon 1: c.289G>A; p.(Ala97Thr) in the proband (Figure 3.36 C). The mutation is predicted to be pathogenic by SIFT and Polyphen2 and is absent from ExAC (Table 3.13). The mutation affects a residue that is conserved in orthologous sequences (Figure 3.36 D). DNA was obtained from her unaffected mother, father and sibling, and segregation analysis revealed the mutation was absent from the three other family members tested. Paternity testing performed in the Czech Republic confirmed the paternity of the father, therefore the mutation has occurred *de novo* in this proband.

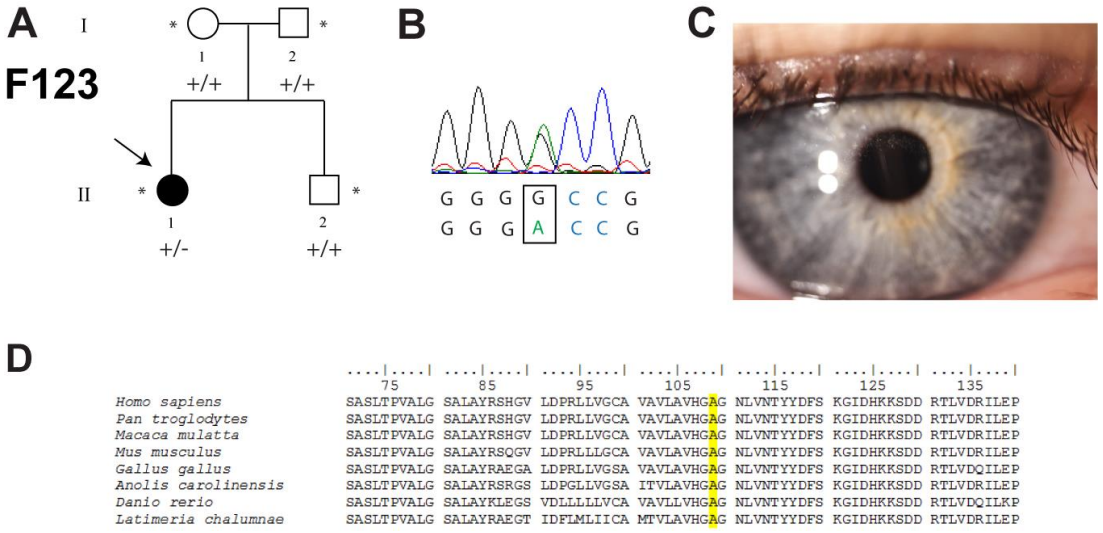


Figure 3.36 Identification of a *de novo UBIAD1* mutation in a Czech proband F123: II:1 diagnosed with Schnyder corneal dystrophy. (A) Pedigree showing lack of family history of corneal disease in the family. * indicates patient has been examined and a blood sample taken. +/- indicates individual is heterozygous for the mutation and +/+ indicates wild-type. Segregation of the *UBIAD1* c.289G>A; p.(Ala97Thr) mutation shows it is present in the proband but absent from both parent samples and an unaffected sibling**. (B)** Sequence electropherogram showing the *UBIAD1* c.289G>A; p.(Ala97Thr) mutation. **(C)** Small patches of crystals observed in the proband. **(D)** Alignment of orthologous protein sequences with the Ala-97 amino acid residue highlighted, which is conserved in all species.

3.7.6 Summary of UBIAD1 mutations

In total, 5 mutations were identified in a cohort of eight families with SCD from the MEH and Czech cohorts. Analyses of potential pathogenicity for these mutations are

shown in Table 3.13. Four mutations had been previously reported and one was novel. One was confirmed to have occurred *de novo*.

Table 3.13 Summary of *UBIAD1* mutations identified in the MEH and Czech cohort with Schnyder corneal dystrophy and assessment of pathogenicity. SIFT and Polyphen2 provide predictions of likelihood of pathogenicity for a particular variant. ExAC shows reported frequency of variant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Novel mutation is highlighted in bold.

Family ID	Nucleotide	Protein	SIFT (1-0)	Polyphen2 (0-1)	ExAC	Ref.
F116, F117	c.361C>T	p.(Leu121Phe)	0.01 (D)	0.950 (PrD)	NP	(Weiss, Kruth, *et al.*, 2008)
F118	c.308C>T	p.(Thr103Ile)	0 (D)	0.993 (PrD)	NP	(Lin *et al.*, 2016)
F119, F120, F121	c.305G>A	p.(Asn102Ser)	0 (D)	0.999 (PrD)	NP	(Weiss *et al.*, 2007)
F122	**c.527G>A**	**p.(Gly176Glu)**	**0 (D)**	**0.992 (PrD)**	**NP**	
F123	c.289G>A	p.(Ala97Thr)	0.05 (D)	0.996 (PrD)	NP	(Nickerson *et al.*, 2010)

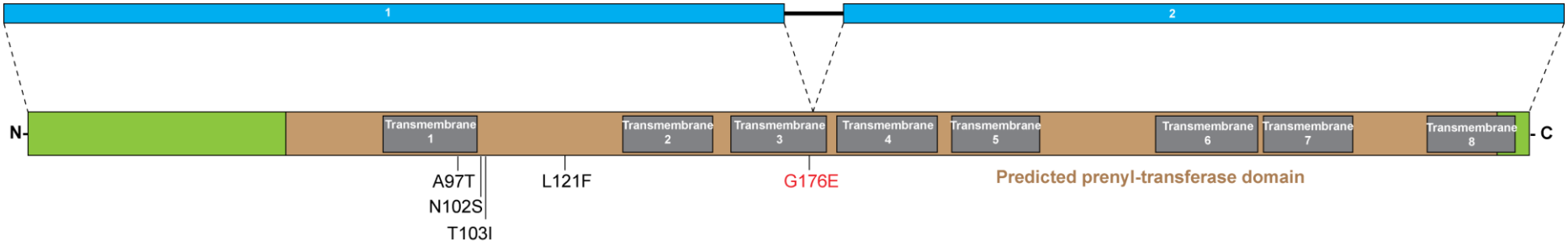


Figure 3.37 Summary of the gene and protein structure of UBIAD1. The mutations identified in this study are shown in their respective locations with novel mutations highlighted in red. All mutations occur within the predicted prenyl-transferase domain, and the majority cluster within or on the border of the first transmembrane domain (TM1). Structure adapted from (Nickerson *et al.*, 2010).

Table 3.14 Summary of families analysed in chapter 3 (F1 – F123), their associated clinical diagnosis and the causative mutation identified. Standard acronyms are used for corneal dystrophy diagnoses. ? prior to a diagnosis indicates diagnostic uncertainty.

Family ID	Clinical Diagnosis	Mutation
F1	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F2	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F3	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F4	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F5	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F6	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F7	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F8	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F9	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F10	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F11	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F12	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F13	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F14	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F15	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F16	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F17	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F18	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F19	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F20	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F21	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F22	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F23	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F24	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F25	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F26	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F27	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F28	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F29	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F30	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F31	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F32	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F33	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F34	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)

F35	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F36	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F37	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F38	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F39	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F40	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F41	RBCD/TBCD	Het *TGFBI:* c.371G>T; p.(Arg124Leu)
F42	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F43	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F44	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F45	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F46	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F47	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F48	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F49	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F50	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F51	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F52	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F53	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F54	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F55	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F56	RBCD/TBCD	Het *TGFBI:* c.1664G>A; p.(Arg555Gln)
F57	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F58	GCD	Het *TGFBI:* c.371G>A; p.(Arg124His)
F59	GCD	Het *TGFBI:* c.1663C>T; p.(Arg555Trp)
F60	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F61	LCD	Het *TGFBI:* c.370C>T; p.(Arg124Cys)
F62	RBCD/TBCD	Het *TGFBI:* c.1874T>A; p.(Val625Asp)
F63	LCD	Het *TGFBI:* c.1877A>G; p.(His626Arg)
F64	LCD	Het *TGFBI:* c.1877A>G; p.(His626Arg)
F65	LCD	Het *TGFBI:* c.1859C>A; p.(Ala620Asp)
F66	RBCD/TBCD	Het *TGFBI:* c.1868G>A; p.(Gly623Asp)
F67	RBCD/TBCD	Het *TGFBI:* c.1868G>A; p.(Gly623Asp)
F68	LCD	Het *TGFBI:* c.1868G>A; p.(Gly623Asp)
F69	EBMD	Het *TGFBI:* c.1868G>A; p.(Gly623Asp)
F70	EBMD	Het *TGFBI:* c.1868G>A; p.(Gly623Asp)
F71	EBMD	Unsolved
F72	EBMD	Unsolved
F73	EBMD	Unsolved

F74	EBMD	Unsolved
F75	EBMD	Unsolved
F76	EBMD	Unsolved
F77	EBMD	Unsolved
F78	EBMD	Unsolved
F79	EBMD	Unsolved
F80	EBMD	Unsolved
F81	EBMD	Unsolved
F82	EBMD	Unsolved
F83	EBMD	Unsolved
F84	EBMD	Unsolved
F85	EBMD	Unsolved
F86	EBMD	Unsolved
F87	EBMD	Unsolved
F88	EBMD	Unsolved
F89	EBMD	Unsolved
F90	EBMD	Unsolved
F91	EBMD	Unsolved
F92	ERED	Unsolved
F93	ERED	Unsolved
F94	GDLD	Hom *TACSTD2:* c.551A>G; p.(Tyr184Cys)
F95	GDLD	Het *TACSTD2:* c.661delG; p.(Asp221Metfs*50) Het *TACSTD2*: c.355T>A; p.(Cys119Ser)
F96	MCD	Hom *CHST6*: c.213_214delGCinsTT; p.(Glu71_Pro72delinsAspSer)
F97	MCD	Hom *CHST6*: c.617C>T ; p.(Ala206Val)
F98	MCD	Het *CHST6*: c.392C>T; p.(Ser131Leu) Het *CHST6*: c.629C>T; p.(Ser210Phe)
F99	MCD	Hom *CHST6*: c.320T>C; p.(Phe107Ser)
F100	MCD	Het *CHST6*: c.91C>T; p.(Pro31Ser) Het *CHST6*: c.599T>G; p.(Leu200Arg)
F101	MCD	Hom *CHST6*: c.379C>T; p.(Arg127Cys)
F102	MCD	Het *CHST6*: c.611C>A; p.(Pro204Gln) Het *CHST6*: c.629C>T; p.(Ser210Phe)
F103	MCD	Hom *CHST6:* c.599T>G; p.(Leu200Arg)
F104	MCD	Het *CHST6*: c.91C>T; p.(Pro31Ser) Het *CHST6*: c.599T>G; p.(Leu200Arg)
F105	MCD	Hom *CHST6:* c.545delA; p.(Gln182Argfs*199)
F106	MCD	Hom *CHST6:* c.494_495delGCinsCT; p.(Cys165Ser)

F107	MCD	Het *CHST6:* c.656C>T; p.(Ala219Val) Het *CHST6*: c.827T>C; p.(Leu276Pro)
F108	MCD	Hom *CHST6:* c. 997T>C; p.(Trp333Arg)
F109	MCD	Hom *CHST6:* c. 997T>C; p.(Trp333Arg)
F110	MCD	Hom *CHST6:* c.250dupA; p.(Ser84Lysfs*25)
F111	MCD	Het *CHST6:* c.1089dupT; p.(Glu364*) Het chr16.hg38: 75,499,098_75,499,541del
F112	MCD	Het chr16.hg38: 75,499,407_75,539,615del
F113	?PPCD	Het *CHST6:* c.505C>A; p.(Ser169Arg) Het chr16.hg38: 75,499,098_75,499,541del
F114	MCD	Het *CHST6*: c.6_7delinsaa; p.(Trp2*) Het chr16.hg38: 75,499,407_75,539,615del
F115	MCD	Het *CHST6*: c.614G>A; p.(Arg205Gln)
F116	SCD	Het *UBIAD1*: c.361C>T; p.(Leu121Phe)
F117	SCD	Het *UBIAD1*: c.361C>T; p.(Leu121Phe)
F118	SCD	Het *UBIAD1*: c.308C>T; p.(Thr103Ile)
F119	?SCD ?Syndromic	Het *UBIAD1*: c.305G>A; p.(Asn102Ser)
F120	SCD	Het *UBIAD1*: c.305G>A; p.(Asn102Ser)
F121	LCD	Het *UBIAD1*: c.305G>A; p.(Asn102Ser)
F122	SCD	Het *UBIAD1*: c.527G>A; p.(Gly176Glu)
F123	SCD	Het *UBIAD1*: c.289G>A; p.(Ala97Thr)

3.8 Discussion

3.8.1 Mutational spectrum of TGFBI mutations

TGFBI mutations are responsible for causing the epithelial-stromal *TGFBI* dystrophies, which consist of a range of clinically distinct corneal dystrophy diagnoses. Over 60 *TGFBI* mutations have been described, however studies have shown that ~ 50% of epithelial-stromal *TGFBI* dystrophies are caused by one of six mutations located at amino acid residues Arg-124 and Arg-555 (Munier *et al.*, 2002). In our large multi-ethnic cohort recruited from MEH, we found that five of the six reported potential hotspot mutations were responsible for ~ 90% (61/68) of all *TGFBI*-associated cases, with nine different *TGFBI* mutations identified in total. The p.(Gly623Asp) mutation was found to be more prevalent in our cohort than previously reported, and was identified in three patients clinically diagnosed with an epithelial-stromal *TGFBI* dystrophy and a further two probands with EBMD. This data demonstrates that despite a large number of reported *TGFBI* mutations and a diverse multi-ethnic population, a limited number of mutations are responsible for the majority of the *TGFBI*-associated dystrophy cases in the MEH cohort.

3.8.2 Genotype – phenotype correlation of TGFBI mutations

Correlation between the specific *TGFBI* mutation that has been inherited and the clinical appearance is a consistent feature of the epithelial-stromal *TGFBI* dystrophies. This has been demonstrated in cohorts from a range of diverse populations, particularly for the hot-spot mutations affecting Arg-124 and Arg-555. This was replicated in the MEH cohort with mutations affecting these residues showing complete genotype-phenotype correlation. For rarer *TGFBI* mutations, determining genotype-phenotype correlation was more difficult due to the smaller numbers of probands carrying the same mutation for comparison, however the majority show similar clinical phenotypes to those previously reported.

The *TGFBI* mutation p.(Val625Asp) was identified in the proband of family F62. This mutation has been reported only once previously in the literature (Tian *et al.*, 2007). The clinical phenotype caused by the mutation in the previously reported

family was variable between the two affected individuals. The detection of lattice lines in the proband led the authors to conclude the phenotype was a variant of LCD; however, no lattice lines were identified in the proband's affected mother. She was described as having grey-white opacities and retractile dots. The clinical phenotype of the MEH proband has been modified as she has previously undergone bilateral PTK surgery, preventing any meaningful phenotypic assessment. Although she had been clinically diagnosed with RBCD, histology excluded this as a clinical diagnosis and revealed the presence of amyloid, confirming that the p.(Val625Asp) mutation is associated with the development of a LCD phenotype.

Two white British families (F63 and F64) were found to harbour the *TGFBI* mutation p.(His626Arg). This mutation has been previously reported to cause a phenotypic variant of LCD in both white and Asian populations. It is particularly prevalent in Vietnam, where it is more common than the classic LCD p.(Arg124Cys) mutation (Chau *et al.*, 2003). Clinically, this mutation has been reported to be associated with fewer, thicker lattice lines, compared to those of classic LCD that occur deeper in the stroma (Schmitt-Bernard *et al.*, 2000; Schmitt-Bernard, Schneider and Argiles, 2002). These features were not noted for the MEH families and other reports have also described a LCD phenotype that is indistinguishable from classic LCD (Nowińska *et al.*, 2011). It has been previously reported that the age of onset is mid- to late-thirties (Chau *et al.*, 2003), however F63: IV:3 presented with lattice lines in her mid-twenties. Late-onset of corneal erosions (aged 50+) has been described in patients with this mutation (Chau *et al.*, 2003), however in F63 both the proband and her brother experienced recurrent corneal erosions previously in affected eyes at the ages of 27 and 35. Asymmetric progression is a commonly reported feature of this mutation (Schmitt-Bernard *et al.*, 2000; Schmitt-Bernard, Schneider and Argiles, 2002; Chau *et al.*, 2003; Lai *et al.*, 2014) and one patient (F63: IV:2) demonstrated a severely affected right eye requiring surgery but with only minor changes in the left eye. Collectively, these data indicate that the p.(His626Arg) mutation is associated with a clinical diagnosis of LCD with variable appearance of lattice lines, age of onset and onset of corneal erosions, although asymmetric progression appears to be a common but not consistent feature of disease.

Proband F65: III:10 with LCD harboured the *TGFBI* p.(Ala620Asp) mutation. This mutation has previously been described only once, in a patient of Chinese ethnicity. This patient was clinically described as having lattice lines, and sub-epithelial and stromal scarring due to frequent recurrent erosions, with an onset of early twenties (Lakshminarayanan *et al.*, 2011). Based on limited clinical information for the MEH patient and the previously reported patient, the p.(Ala620Asp) mutation appears to have no distinctive clinical features that would distinguish the phenotype from that of classic LCD. Both families were of East Asian origin, indicating that this mutation may be more common in this population.

Due to this striking *TGFBI* genotype-phenotype correlation, it was therefore unexpected to discover that the *TGFBI* p.(Gly623Asp) mutation was associated with a wide range of clinical diagnoses including LCD, RBCD/TBCD and EBMD in five families. Six families carrying the p.(G623D) mutation have been described previously (Afshari *et al.*, 2001; Munier *et al.*, 2002; Aldave, Rayner, *et al.*, 2005; Auw-Haedrich *et al.*, 2009; Gruenauer-Kloevekorn *et al.*, 2009). The most common phenotype is a geographic pattern of opacities at Bowman's layer, consistent with RBCD, although some individuals also have the stromal lattice lines of LCD. The majority of families described have similar clinical features within the affected family, although some have intrafamilial variation. The phenotypic variability associated with p.(G623D) may be the result of environmental or genetic modifiers of this mutation, secondary corneal scarring (Salzmann degeneration), or the accumulation of amyloid with age to form lattice lines in older individuals.

Although all of the probands carrying the p.(Gly623Asp) mutation were eventually diagnosed with a corneal dystrophy, some of them were treated for chronic ocular surface disorders prior to this molecular diagnosis. This indicates that there may be other patients harboring the same mutation which were not recruited to our study or other studies, because they have been incorrectly diagnosed. Diagnosis of a corneal dystrophy was required for inclusion in the study, therefore those individuals diagnosed with dry eye or bilateral herpes simplex were excluded unless there was diagnostic uncertainty. Therefore, the identification of 5 probands in our cohort carrying the p.(Gly623Asp) mutation is likely to be an underestimate.

3.8.3 Epithelial basement membrane dystrophy

EBMD is typically considered to be a sporadic dystrophy with a degenerative (non-genetic) aetiology. However, two previous mutations in *TGFBI* have been associated with the characteristic map-dot-fingerprint changes occurring at the epithelial basement membrane [p.(Leu509Arg) and p.(Arg666Ser)] (Boutboul *et al.*, 2006). Following the original identification of these two mutations in three probands with EBMD, no follow up cohorts have been reported and therefore these findings have not been replicated. In large cohorts of corneal dystrophy patients, probands with EBMD are typically excluded from genetic screens as they are assumed to have a degenerative aetiology. Within our cohort, we have now replicated the finding that ~ 10% cases of EBMD can be explained by mutations in *TGFBI* and expanded the mutation spectrum associated with this clinical phenotype. As EBMD is usually considered to be degenerative, identification of the p.(G623D) mutation in the EBMD cohort also has important implications for clinical management and counseling for patients and their families.

This study provides the first evidence that additional genes may contribute to the EBMD phenotype for some patients. Two families, one with an apparent dominant inheritance pattern and another with a potentially recessive or dominant with reduced penetrance, were recruited. WES analysis did not identify a mutation in any corneal dystrophy or syndrome-associated gene which could be responsible for the corneal phenotype. Filtering WES data for rare variants in shared genes also failed to identify the causative gene for these families. One possibility is the causative variant is a synonymous variant which effects splicing, but is not located at canonical splice site. This has been demonstrated for the *COL17A1* mutation c.3156C>T; p.(Gly1052Gly). Another possibility is that the variant occurs in a region which did not have sufficient coverage by WES, or occurs in a non-coding region such as regulatory or promoter sequences. Further recruitment of familial cases of EBMD will be essential to further investigate the possibility of additional genetic heterogeneity in EBMD.

3.8.4 Epithelial recurrent erosion dystrophy

ERED is not a commonly used clinical diagnosis as the condition is often indistinguishable from recurrent erosions of sporadic or environmental (traumatic) aetiology. However, two probands were recruited with symptoms consistent with this diagnosis, one of which also demonstrated some skin symptoms, although further clinical detail could not be obtained. For one proband multiple potential genetic causes of their corneal dystrophy were identified.

A heterozygous c.3190G>A; p.(Val1064Ile) missense change was identified in the ERED gene *COL17A1* which is a rare variant with a MAF of 0.0004 (44/121356), however it is more common in South Asians with a population specific MAF of 0.002 (41/16510) indicating that is may be a rare polymorphism is that is more common in the South Asian population, and not causative. Three variants were identified in two genes, *LAMC2* and *PLEC*, which also cause (autosomal recessive) epidermolysis bullosa. A large majority of previously reported pathogenic *PLEC* mutations are nonsense mutations; however, a heterozygous *PLEC* missense mutation [p.(Arg2000Trp)] has been reported as causative in one family demonstrating autosomal dominant transmission, which involves the same amino acid substitution and is located nearby to the heterozygous c.6010C>T; p.(Arg2004Trp) change identified in our proband (Koss-Harnes *et al.*, 2002). Our proband was referred to a dermatologist, so epidermolysis bullosa cannot be excluded as a potential diagnosis although further clinical details on skin symptoms were not available. The heterozygous *LAMC2* splice site variant identified, c.2456+1G>T, is a novel variant located in the canonical splice donor site for exon 16. The presence of this mutation is predicted to abolish the canonical splice donor site of this exon (Table 3.8). This could result in exon skipping or the use of an alternative donor site, such as the one located within intron 16 that is predicted to be used by three splice prediction programs. Either circumstance would be predicted to result in a frameshift and incorporation of a premature stop codon. Furthermore, biallelic mutations in *SPINT2* cause congenital diarrhoea syndrome associated with corneal erosions (Heinz-Erian *et al.*, 2008). A heterozygous *SPINT2* variant was identified in the proband, however this has a higher MAF than the other candidates 0.001 (163/119144) and is predicted to be benign and tolerated by pathogenicity

prediction programs. No other relevant symptoms of this condition were noted in the proband.

Mutations in *COL17A1* cause both the skin blistering disorder epidermolysis bullosa (autosomal recessive) and the corneal dystrophy ERED (autosomal dominant) (Jonsson *et al.*, 2014). In epidermolysis bullosa, loss of function mutations affecting both copies of *COL17A1* result in the epidermis being prone to separation from the underlying dermis, leading to severe skin blistering. Corneal erosions are common in individuals with epidermolysis bullosa, ranging from 47.5 – 74.1% experiencing at least one erosion, depending on the clinical sub-type (Fine *et al.*, 2004). The mechanism of pathogenicity of mutations causing ERED is unknown, although gain-of-function has been proposed given the lack of ERED in the parental carriers of epidermolysis bullosa patients (Jonsson *et al.*, 2014). WES analysis of two probands with ERED revealed possible loss of function variants in the *LAMC2* gene in both affected individuals [F92 and F93]. This gene is an excellent candidate as, in combination with two other proteins, it forms a complex called laminin-332 which directly interacts with collagen XVII (encoded by *COL17A1*) in the hemidesmosome complex; junctions which adhere epithelial cells to the underlying basement membrane in both the skin and in the cornea (Kiritsi, Has and Bruckner-Tuderman, 2013; Torricelli *et al.*, 2013). Given the lack of family history and samples from family F93 and only two families interrogated in this cohort, it is not possible to prove a causative role for the loss-of-function mutations in *LAMC2* although this remains a compelling candidate. Further families should be recruited to the study to provide strong genetic evidence for this candidate gene.

3.8.5 Gelatinous drop like dystrophy

GDLD is a recessive corneal dystrophy affecting the corneal epithelium that is extremely rare in the United Kingdom. GDLD is most prevalent in Japan due to the presence of a founder mutation (Kawano, Fujiki and Kanai, 1992). Two families were recruited with GDLD; a family recruited from MEH (F95) and a consanguineous family (F94) recruited from Antioquia, Colombia. Both families were positive for biallelic mutations in *TACSTD2*; we were therefore unable to investigate additional genetic heterogeneity for GDLD.

The molecular diagnosis of a family with GDLD from Antioquia, Colombia was the first identification of a pathogenic *TACSTD2* mutation in the Latin American population. There have been a number of clinical descriptions of GDLD patients from Mexico in the literature indicating either one or multiple mutations present in the population (Pineda-Cárdenas and González-Almaráz, 2002). Recently, the underlying mutation has been reported for one of these families, who were found to harbour a novel homozygous nonsense mutation in *TACSTD2* (Cabral-Macias *et al.*, 2016). The mutation identified in our Colombian family was a homozygous missense mutation p.(Tyr184Cys) which has previously been observed once in a Chinese proband as a heterozygous mutation in combination with the Japanese founder mutation p.(Gln188*) (Tian *et al.*, 2004). It has not previously been identified in the homozygous state. The p.(Tyr184Cys) mutation has an estimated frequency of 0.53% in the Antioquia population. This frequency indicates that GDLD may be more prevalent in Colombia than currently recognised and may be underdiagnosed; however, this estimate is based on a limited number of Colombian samples.

There are limited reports of patients in the UK with GDLD; however, none of these are genetically solved. This includes a 7 year old white proband, diagnosed with GDLD and global developmental delay, who was negative for pathogenic mutations in *TACSTD2* (Akhtar *et al.*, 2005). It is presumed therefore that *TACSTD2* mutation carriers are extremely rare in the UK, however in our white British proband compound heterozygous *TACSTD2* mutations were identified. One was a novel frameshift, however the p.(Cys119Ser) mutation identified has previously been identified in two previous families, both of Tunisian ethnicity (Ren *et al.*, 2002). No parental samples could be obtained to determine which mutation was inherited from each parent, or if either change has a *de novo* origin.

Four clinical sub-types of GDLD have been described; band keratopathy, stromal opacity, kumquat and typical mulberry, however their relationship to the underlying genotype or to environmental factors has not been elucidated. It has previously been proposed that the clinical sub-types may be associated with an age dependent progression of the disease (Ide *et al.*, 2004). F94 consisted of a large consanguineous family with 5 affected individuals. Of the affected individuals in the pedigree, two

showed a mild clinical phenotype of the band keratopathy sub-type that caused minimal visual disturbance, whereas the remaining three had a characteristic severe mulberry appearance with neovascularisation. The milder phenotype observed in this family may still progress to more severe subtypes, however, interestingly the mulberry phenotype appeared as early as age 3 in the proband, whereas other siblings, who have reached the age of 14 still retain a mild band-keratopathy phenotype. This does not rule out an age-related progression but it does indicate that other, likely environmental, factors influence either the development of a particular clinical sub-type or the rate of disease progression from milder to more severe sub-types. The clinical appearance of the British proband had been modified by multiple previous grafts in both eyes but was most consistent with the mulberry sub-type.

TACSTD2 encodes a single pass transmembrane glycoprotein involved in calcium signaling which localises to cell-to-cell contacts (Ripani *et al.*, 1998). It is likely that *TACSTD2* mutations are loss-of-function as many are nonsense or frameshift mutations. The p.(Tyr184Cys) and p.(Cys119Ser) mutations are two missense mutations in *TACSTD2* that involve cysteine residues. The TACSTD2 protein contains 3 disulphide bridges within its thyroglobulin type-1 domain which are important for the secondary structure of the protein. The p.(Tyr184Cys) mutation, which involves the creation of a new cysteine residue, may affect the secondary structure of the protein by disrupting the formation of one of these disulphide bridges or by inducing the formation of aberrant disulphide bridges. The p.(Cys119Ser) missense mutation disrupts a known disulphide bridge formed by the Cys-119 and the Cys-125 residues.

3.8.6 Macular corneal dystrophy

MCD was the most common recessive corneal dystrophy in our cohort, and *CHST6* the second most common cause of anterior corneal dystrophies after *TGFBI*. In total 20 probands were recruited from MEH with MCD. Most *CHST6* mutations identified were private to the affected family within the cohort, although the majority have been previously reported in the literature. This indicates that *CHST6* carriers are present in the control population at low frequencies, which is common for recessive disease.

In our cohort, 10 probands were found to harbour a homozygous mutation and 10 had compound heterozygous changes. Missense mutations were most frequent with 15 identified plus an additional mutation which involved the substitution of two amino acids. In comparison, there were only two nonsense mutations and one frameshift, which were predicted to result in the incorporation of a premature stop codon. Two different deletions upstream of *CHST6* were identified, of which one was novel, and are predicted to disrupt a transcriptional regulatory region. Five novel *CHST6* mutations were identified in the *CHST6* coding sequence. The nonsense mutation [p.(Glu364*)] and frameshift mutation [p.(Ser84Lysfs*25)] are highly likely to be pathogenic as they introduce a premature stop codon in the *CHST6* transcript and therefore would be predicted to be degraded by nonsense mediated decay or, if translated, produce a truncated protein. The remaining three are missense mutations, with varying levels of evidence for a pathogenic role. Pathogenicity prediction programs SIFT and Polyphen2 predicted a number of *CHST6* missense changes as benign and tolerated, respectively. This occurred even in cases where the mutation had been previously identified as pathogenic. Both SIFT and Polyphen2 utilise data from the conservation of the amino acid residue in orthologous sequences (Ng and Henikoff, 2003; Adzhubei *et al.*, 2010). As there is only a direct orthologue in primates for *CHST6*, the lack of accurate prediction may be due to lack of orthologous sequences in other species.

The *CHST6* gene is neighboured by a highly homologous gene, *CHST5*, and the two genes likely arose from a gene duplication event. Upstream from the *CHST6* transcript is a ~ 2.5 kb region (region B) flanked by two sequences which have significant homology with the equivalent sequences flanking the ~ 2.5 kb region (region A) upstream of *CHST5*. Mis-pairing of regions A and B during homologous recombination produce one of two outcomes: a ~ 40 kb deletion encompassing *CHST5* and region B, or a rearrangement in which region B is replaced by region A. Both these events result in region A, normally upstream of *CHST5* being relocated to the 5' region upstream of *CHST6*, in place of region B (Akama *et al.*, 2000). A screening protocol for recurrent *CHST5* and *CHST6* deletions and replacements was used to test any MCD patients negative for one or two mutations in *CHST6*. In the MEH cohort, the ~ 40 kb deletion was identified, in a heterozygous state, in one proband with MCD. No upstream rearrangements were detected. However, no

positive control DNA for the F1/R1M PCR reaction was available prior to screening, therefore there is no way to distinguish between a failed PCR reaction and the absence of this amplimer. The F1/R1M amplimer distinguishes between the *CHST6* upstream ~ 40 kb deletion and the ~ 2.5 kb rearrangement, with both types of mutation showing an amplified product with the F2M/R2 primers, therefore this should not change the outcome of the screening in terms of whether a mutation is present in the upstream region or not. A novel ~ 0.5 kb deletion was identified which spanned the genomic interval between the F2 and R2 primers. When directly sequenced, the upstream deletion was identified to be 444 bp in size and occurs at a region with 8 bp region of microhomology. The 444 bp deletion overlaps with the known pathogenic ~ 40 kb upstream deletion. This deletion was detected in two probands, one with MCD, and one with an unclassified corneal dystrophy and represents a new recurrent mechanism, likely due to the homology between the regions.

CHST6 encodes an enzyme responsible for catalysing the sulphation of keratan sulphate. Keratan sulphate is a common glycosaminoglycan side chain of corneal proteoglycans lumican, keratocan and mimecan; these are present at high levels in the cornea and are thought to maintain regular collagen fibril spacing, which is essential for maintenance of transparency (section 1.2.3). An enzyme-linked immunosorbent assay (ELISA) using a keratan sulphate antibody has defined three immunophenotypes within MCD. Type I have no detectable levels of keratan sulphate in corneal buttons or in serum, type IA have no detectable keratan sulphate in corneal buttons or serum but keratan sulphate positive accumulations are detectable in keratocytes. The type II immunophenotype has no detectable keratan sulphate levels in corneal buttons but keratan sulphate is present in serum (Hassell *et al.*, 1980). The correlation between the genetic basis of the disease and the resulting immunophenotype is not fully resolved. This was beyond the scope of this study but could provide additional insights into pathogenicity if investigated in future.

The occurrence of multiple mutations across the entirety of the coding exon, and the lack of specific hot-spots are consistent with a loss of function mechanism. *CHST6* upstream disruptions including the ~ 40 kb deletion and the replacement of a ~ 2.5 kb *CHST6* upstream region with the equivalent region of *CHST5* are hypothesised to

disrupt a regulatory region driving corneal specific expression. *CHST5* has a different expression pattern to *CHST6* and explains why replacement of the *CHST6* upstream region with that of *CHST5* causes a loss of corneal specific expression. The underlying genetic architecture driving corneal expression upstream of *CHST6* is undefined. Presumably the regulatory elements are present within the ~ 2.5 kb region that is flanked by two regions of homology with *CHST5*. In our cohort, two MCD patients were identified with a novel 444 bp upstream deletion which overlapped with this critical region. This could therefore define and refine the region containing the regulatory elements which drive corneal specific expression. Identification of these elements and the transcription factors that bind them could be important for elucidating the molecular pathways that maintain transparency in the cornea.

Potentially due to the loss of function mechanism, mutations in *CHST6* do not show genotype-phenotype correlation. The clinical phenotype is consistent in the majority of probands and there are few reported incidences of misdiagnosis. However, in one proband clinically diagnosed with PPCD (or another endothelial dystrophy) potentially pathogenic mutations in *CHST6* were identified; one mutation was detected by WES and subsequently confirmed by Sanger sequencing, and the other, a novel upstream deletion was identified using the PCR-based screening assay. There are some similarities between the phenotype of MCD and endothelial dystrophies. MCD causes stromal clouding that can be similar to stromal oedema, which occurs in late stages of endothelial dystrophies and MCD can also affect Descemet's membrane and the endothelium, therefore the previous diagnoses of PPCD or CHED does not rule out a diagnosis of MCD in this family. Stromal thinning and localised elevation of anterior corneal curvature is also a feature of MCD (Dudakova *et al.*, 2014), which is reminiscent of the corneal steepening reported in many PPCD3 patients (Aldave *et al.*, 2013) and may also explain why the proband was previously diagnosed with keratoconus. Despite the identification of two *CHST6* variants in this proband, a diagnosis of MCD is also not consistent with the inheritance of the disease in the family which is presumed to be autosomal dominant. Although a DNA sample was not available for testing, the affected daughter of the proband would have inherited only one of these *CHST6* mutations. For the recessive disorder to occur in the second generation the affected daughter's mother would also have to be a carrier of a *CHST6* mutation, which would result in pseudo-dominance. Although

this seems unlikely, two generations of affected individuals have previously been reported for MCD in two unrelated Vietnamese families (Ha *et al.*, 2003). Without a DNA sample for the daughter and further clinical examination, evaluating the pathogenicity of the upstream deletion and novel *CHST6* mutation is not possible in this case. It is possible that these variants are not responsible for the corneal dystrophy in the proband and that he has an autosomal dominant endothelial dystrophy of unknown genetic cause. Although known corneal dystrophy genes were excluded in pre-screening and confirmed negative in the WES data analysis, there are still endothelial corneal dystrophy genes yet to be discovered and mutations that lie outside of regions covered by WES will have been missed by the analysis.

Biallelic mutations were identified for the majority of the MEH MCD cohort using a combination of screening the coding region of the *CHST6* gene for mutations and using a PCR-based assay to detect deletions and rearrangements upstream of the *CHST6* gene, however a complete genetic diagnosis was not achieved for all MCD probands. Locus heterogeneity is possible for MCD as mutation negative families have been described (Birgani *et al.*, 2009). However, in our cohort both unsolved probands have a single pathogenic *CHST6* mutation, and therefore it is more likely that there are additional non-coding mutations which affect the expression of *CHST6* in the cornea. These would not be detected by current screening assay protocols. Using a candidate screening approach, the ~ 2.5 kb region upstream of *CHST6* would be the best candidate to sequence to identify novel variation which causes MCD. Alternatively, WGS may provide a more comprehensive analysis of the locus. WGS has the advantage that novel gross deletions and rearrangements would also be detectable using this method in addition to smaller mutations in important regulatory regions affecting *CHST6* expression.

Deletion of the entire coding exon has previously been described as causative of MCD (Akama *et al.*, 2000). In the homozygous state only, this would be identifiable as a failure to amplify a product with *CHST6* PCR primers or amplification of a smaller product providing the deletion was within the boundaries of the PCR primers used. However, due to their similarity, co-amplification of the *CHST5* coding exon is also possible and would be identifiable from sequencing as it would appear to contain a greater than expected number of SNPs when aligned to the *CHST6*

reference sequence (Liskova, personal communication). As deletion of a heterozygous deletion of the coding region of *CHST6* was not included in the screening protocol, this may account for proband F112 who has only a single upstream mutation and no mutations in the coding region of the gene. Deletions of the *CHST6* coding exon could be detectable using a custom designed qPCR assay or WGS.

3.8.7 Schnyder corneal dystrophy

Eight probands, six from MEH and two from the Czech Republic, had genetically confirmed SCD with five mutations identified in total in *UBIAD1*. The majority of pedigrees were small with only one or two generations of affected individuals. Clinically, all eight probands examined had corneal clouding with the presence of at least some crystalline deposition. In some this was extensive whereas others had only small peripheral crystals, but this did not appear to correlate with their underlying genotype. It is estimated that ~ 50% of SCD cases present without crystalline deposition (Weiss, 1996). Although we have examined a small cohort, the lack of equivalent number of patients with the non-crystalline phenotype indicates that our cohort is likely to be an underestimate of the true number of patients with SCD. Whereas crystalline corneal deposits are highly indicative of a clinical diagnosis of SCD, both corneal clouding and corneal arcus are observed commonly in other syndromes and phenocopies of SCD. This is exemplified in the proband of F119 who was initially diagnosed with suspected SCD, but other syndromic diagnoses were considered possible due to the presence of additional symptoms including learning difficulties and knee deformities. In this case the proband was positive for a mutation in *UBIAD1* however the difficulties in diagnosing corneal clouding after exclusion of *UBIAD1* mutations are discussed further in chapter 5.

In addition to corneal symptoms, there have been reports of systemic features associated with SCD. Early reports indicated an increased incidence of hyperlipidemia in individuals with SCD. Genu valgum, a knee deformity which affect the angle of the knee, has been reported in an estimated ~ 4% of SCD cases (Weiss, 2009). In our cohort, genu valgum was present in two cases within the same family (1/8 families; 12.5%). This could indicate an association with specific

UBIAD1 mutations or a contribution of other genetic variation to the development of this symptom.

Over 25 mutations have been described in *UBIAD1*. Four of the five mutations identified in this cohort have been previously identified in SCD patients. Two of our probands harboured the p.(Leu121Phe) mutation, with no evidence that they are related. This mutation has previously been identified in one British, one American, and one Saudi-Arabian family (Weiss, Kruth, *et al.*, 2008; Al-Ghadeer, Mohamed and Khan, 2011). Neither of these probands reported an affected parent, although one had an affected sibling. In the absence of clinical examination of the parents of either proband, it is not possible to determine whether any of the parents were mildly affected but not diagnosed due to variable expressivity, whether the mutation demonstrates incomplete penetrance, or whether the mutation is *de novo* in these families. In total, there are five families identified with the p.(Leu121Phe) mutation. The p.(Thr103Ile) change is the second time this mutation has been identified (Lin *et al.*, 2016). In the previously reported family, the mutation was shown to occur *de novo*, therefore this mutation has arisen independently on at least two occasions. Furthermore, the p.(Asn102Ser) mutation, found in two South Asian families in the MEH cohort, has now been reported in 22 families from a variety of populations (Orr *et al.*, 2007; Weiss *et al.*, 2007; Yellore, Khan, *et al.*, 2007; Weiss, Kruth, *et al.*, 2008; Nickerson *et al.*, 2010; Du *et al.*, 2011; Nowinska, Wylegala and Teper, 2014). The *UBIAD1* p.(Ala97Thr) mutation, found to occur *de novo* in a proband (F122) of Czech origin, has been previously reported in a male proband in his thirties with no reported family history of corneal disease. The proband was from the USA and presented with central and paracentral crystalline deposition (Nickerson *et al.*, 2010). The occurrence of multiple families harbouring the same mutation indicates that these amino acid residues are mutational hot-spots. For at least two of these mutations, *de novo* occurrence has been demonstrated and a founder mutation is unlikely for the remaining cases given that they occur in a range of populations.

UBIAD1 is a two coding exon gene that encodes a 338 amino acid residue protein which is predicted to contain between seven and ten transmembrane domains (Nickerson *et al.*, 2013). The largest functional domain of the protein is the prenyl-transferase domain which makes up the majority of the protein and encompasses the

transmembrane domains; to date, all mutations have been located in this domain. UBIAD1 is an enzyme which catalyzes the biosynthesis of menaquinone-4 (MK-4; vitamin K_2) from phylloquinone (vitamin K_1) derivatives. UBIAD1 cleaves the phytyl side chain from ingested plant-derived phylloquinone and catalyses the transfer of a geranylgeranyl group from geranylgeranylpyrophosphate (GGPP) to the substrate to form MK-4 (Nakagawa *et al.*, 2010).

Currently, there is no crystal structure of the human UBIAD1 protein, or even the isolated prenyl-transferase domain, therefore computational modelling has been used extensively using either the *Escherichia coli* ubiA or menA, or argebacteria Ubia. However, these computer models differ in both the number and location of transmembrane helices present in the human UBIAD1 protein. All SCD associated mutations identified to date cluster in specific regions of the protein, exclusively affecting either the transmembrane helices or amino acid residues in the aqueous portion of the protein. Mutation hot-spots include the region between the first transmembrane helices (loop 1), the third and fourth transmembrane helices (loop 2), and the region between helices five and six (loop 3). Modelling has also identified that loop 1 contains the putative binding site of GGPP, therefore mutations within these region may affect substrate binding (Weiss, Kruth, *et al.*, 2008). All the mutations identified in this study are located within these loops including the novel p.(Glu176Gly) mutation which is located within the third transmembrane helix. Other mutations cluster nearby, including a p.(Thr175Ile) mutation which occurs at the adjacent residue and p.(Glu177Gly) which affects an adjacent residue and also includes the same amino acid substitution (Orr *et al.*, 2007; Nickerson *et al.*, 2013).

It has been argued that *UBIAD1* mutations have a loss-of-function effect either by affecting substrate binding or disrupting the enzymatic activity of the protein (Nickerson *et al.*, 2010). It has been demonstrated experimentally that 3 out of the 4 SCD mutations tested have a statistically significant reduction of MK-4 biosynthetic activity. Computational modelling of the protein structure for one mutation revealed that this is likely due to the destabilisation of transmembrane helices affecting the structure of the protein, resulting in the loss of close contact between menadione (intermediate in the MK-4 biosynthetic pathway) and GGPP (Nickerson *et al.*, 2013). Despite this, clustering of mutations in specific regions of the protein and

recurrence of the same changes indicate that the mechanism of pathogenicity is gain-of-function or dominant negative as opposed to haploinsufficiency. Furthermore, none of the *UBIAD1* mutations identified to date have been nonsense or frameshift mutations. Although some loss of enzymatic activity has been demonstrated, this does not preclude a dominant negative role of disease for example, if the mutant protein binds to the substrate but does not catalyse the reaction, but prevents interaction with the wild type enzyme.

The mechanism by which mutations in *UBIAD1* lead to corneal cholesterol deposition is unresolved. Computational modelling predicted that cholesterol binds to UBIAD1 at an overlapping site with GGPP and may compete for binding (Nickerson et al., 2013). The UBIAD1 protein contains a cholesterol recognition/interaction motif [(L/V)-X_{1-5}-(Y)-X_{1-5}-(K/R)] which may be disrupted by mutations nearby and influence cholesterol binding (Lin *et al.,* 2016). UBIAD1 also binds to 3-hydroxy-3-methylglutaryl-CoA reductase and sterol o-acyltransferase 1; two enzymes involved in cholesterol metabolism (Nickerson et al., 2013). Cholesterol deposition in the cornea therefore may occur as a downstream effect of reduced MK-4 levels, due to disruption of cholesterol binding to UBIAD1, and/or due to disruption of interactions of UBIAD1 with enzymes involved in cholesterol synthesis, storage or degradation.

Chapter 4: Endothelial corneal dystrophies

4.1 Introduction

Endothelial corneal dystrophies affect the posterior-most layer of the cornea; the corneal endothelium. Unlike anterior corneal dystrophies, which are associated with an accumulation of opacities and/or recurrent erosions, endothelial corneal dystrophies occur due to dysfunction of corneal endothelial cells. Clinical signs observed at the corneal endothelium and Descemet's membrane can be used to distinguish the various endothelial corneal dystrophies. The end stage of all endothelial dystrophies occurs due to loss of fluid regulation between the anterior chamber and the cornea, resulting in oedema that affects the stroma and the epithelium, called bullous keratopathy. The corneal cloudiness caused by oedema leads to a reduction of visual acuity. Autosomal dominant, autosomal recessive and X-linked endothelial dystrophies have been described (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015).

4.1.1 Posterior polymorphous corneal dystrophy

PPCD is an autosomal dominant endothelial corneal dystrophy, associated with vesicles, band-like lesions and Descemet's membrane thickening on slit lamp examination (Figure 4.1). Specular microscopy reveals loss of the regular size and shape of endothelial cells. Although PPCD is typically described as a mild or asymptomatic condition, disease progression in some individuals lead to corneal oedema and significant visual impairment (Laganowski *et al.*, 1991). PPCD patients also have an increased risk of developing glaucoma (Héon *et al.*, 1995). Histologically, the endothelium typically shows proliferative, migratory endothelial cells with focal multi-layering. Immunohistochemical staining reveals that affected corneal endothelial cells express several keratins that are typically expressed in epithelial tissues, including cytokeratin (CK) 7 and CK19, and are not usually expressed in a healthy endothelium (Jirsova *et al.*, 2007).

PPCD is genetically heterogeneous, with clinically similar phenotypes mapping to three separate genetic loci; PPCD1, PPCD2 and PPCD3. The disease genes and causative mutations have been identified for PPCD2 and PPCD3. PPCD3 is caused by heterozygous mutations in the *ZEB1* gene, located on chr10p11.22. The proportion of PPCD cases associated with *ZEB1* mutations differs between populations and is estimated to be 25-45% in the USA, 5% in the Czech Republic, 9% in New Zealand and 60% in the UK (Liskova *et al.*, 2013). All *ZEB1* mutations described to date are nonsense mutations or frameshifts resulting the incorporation of a premature stop codon, suggesting that haploinsufficiency of the ZEB1 protein is responsible for pathogenesis of the disease (Lechner *et al.*, 2013; Liskova *et al.*, 2013). *ZEB1* encodes a transcription factor which is a driver of epithelial-to-mesenchymal transition (EMT) via transcriptional repression of several genes, and has a role in cancer metastasis (Spaderna *et al.*, 2008). One of the genes repressed by *ZEB1* is *COL4A3,* via binding of ZEB1 to E2-box motifs in the *COL4A3* promoter. *ZEB1* haploinsufficiency results in ectopic *COL4A3* expression in primary endothelial tissue (Yellore *et al.*, 2012).

Mutations in *COL8A2* have been reported to cause PPCD2. The locus was originally mapped to chr1p34.3 – p32 in a pedigree with a different endothelial dystrophy: familial early-onset FECD (section 4.1.4). *COL8A2* was selected as the most promising candidate gene within this locus and screening identified a p.(Gly455Lys) mutation which segregated with disease in the pedigree. Screening two additional families, one with early-onset FECD and another with PPCD, identified the same mutation in *COL8A2* (Biswas *et al.*, 2001). Subsequently, over 5 additional *COL8A2* mutations have been identified in families affected with early-onset FECD (Gottsch *et al.*, 2005; Mok, Kim and Joo, 2009), however there have been no further cases of PPCD caused by *COL8A2* mutations, therefore the full extent of the contribution of this gene to PPCD has not been resolved.

The PPCD1 locus was originally mapped to chr20p11 – q11 in a large 5-generation family diagnosed with PPCD (Héon *et al.*, 1995). Screening candidate genes from the linked region revealed a potentially pathogenic mutation in *VSX1* that segregated with disease. Subsequent screening in additional PPCD probands identified a second variant in *VSX1* which was thought to be pathogenic (Héon *et al.*, 2002). However,

these variants have since been identified at higher than expected frequencies in control populations, raising doubts over their pathogenicity. Furthermore, no additional *VSX1* mutations have been identified in PPCD patients (Aldave, Yellore, *et al.*, 2005). In 2005, two Czech pedigrees with PPCD also mapped to an overlapping locus on chr20p11.2, but crucially, this locus excluded the *VSX1* gene (Gwilliam *et al.*, 2005). Genotyping of microsatellites around the linked region revealed that an additional 12 PPCD families from the Czech Republic were genetically linked to this locus and shared a common haplotype. This indicates that a founder mutation located within this locus is likely to be responsible for the high prevalence of PPCD in the Czech population (Liskova *et al.*, 2012). Linkage to the PPCD1 locus has also been demonstrated in American PPCD families of European origin (Yellore, Papp, *et al.*, 2007). In the UK, a large family has been described attending MEH which was genetically linked to the PPCD1 locus (Toma *et al.*, 1995). The family were initially diagnosed as an autosomal dominant form of CHED (CHED1) (section 4.1.2); however, subsequent re-analysis of the clinical phenotype, and the likelihood that this is allelic with the PPCD1 locus, has resulted in autosomal dominant CHED1 being re-categorised as PPCD1 in the most recent IC3D classification (Weiss *et al.*, 2015). Determining the underlying genetic cause(s) of PPCD1 and CHED1 will determine whether they are allelic disorders or are caused by mutations in different genes located within an overlapping genetic interval.

An additional PPCD phenotype has been described. PPCD bands have previously been described in the literature and are also known as posterior corneal vesicles. They are usually detected during routine ophthalmological examination but do not typically interfere with vision or require treatment (Harada *et al.*, 1990). PPCD bands are distinct in appearance from Haab's striae, which are splits of Descemet membrane secondary to congenital glaucoma (Figure 4.1 B). PPCD bands are thought to result from focal/linear thickening of Descemet membrane, whereas Haab's straiae are areas of membrane thinning associated with splits, although the borders of the striae may be scrolled and thickened (Cibis and Tripathi, 1982). To date, no cases of band-type PPCD have been associated with a mutation in any of the PPCD-associated genes, however it is still considered to be part of the PPCD clinical spectrum.

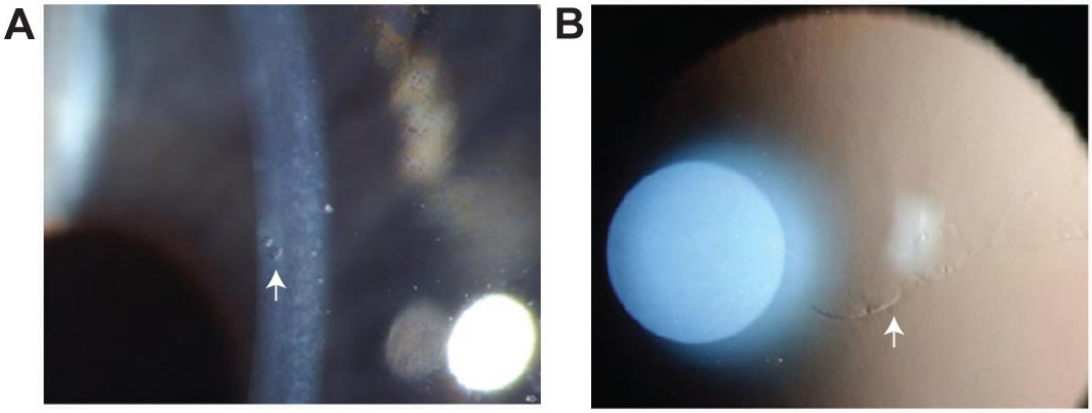


Figure 4.1 Typical clinical appearance of posterior polymorphous corneal dystrophy. (A) Vesicles are present when viewed with slit lamp (example indicated by arrow head). **(B)** Affected cornea in retroillumination shows the presence of a band like lesion with scalloped edge (arrow head indicates the edge of the band). Images from (Vincent *et al.*, 2009).

4.1.2 Congenital hereditary corneal dystrophy

CHED has historically been divided into autosomal recessive CHED2 or autosomal dominant CHED1 sub-types (Callaghan *et al.*, 1999). As described above, in 4.1.1, CHED1 was recently reclassified as PPCD1 in the most recent release of the IC3D classification due to the overlap of clinical features and shared genetic locus of the index CHED1 family with PPCD1. In this IC3D release, CHED2 was updated to CHED (Weiss *et al.*, 2015); however, until the genetic basis of CHED1 and PPCD1 is resolved, the terms are useful for distinguishing the dominant from the recessive forms of CHED and are used throughout this thesis.

A single large family with early-onset corneal oedema has been described as autosomal dominant CHED1. Clinically, CHED1 is associated with a diffuse corneal opacification which occurs from birth or in the first year of life (Figure 4.2 B). Nystagmus is also present in some individuals, and older affected individuals may have secondary scarring and band shaped keratopathy (calcification). Histology shows significant endothelial cell loss and thickening of Descemet's membrane (Pearce, Tripathi and Morgan, 1969). The genetic basis of CHED1 has not been resolved although linkage has identified a locus on chr20p11.2, which overlaps with the PPCD1 locus (section 4.1.1). It is not known whether CHED1 and PPCD1 share a common causative gene or whether there is a different underlying genetic cause within the overlapping loci (Toma *et al.*, 1995).

Individuals with autosomal recessive CHED2 demonstrate diffuse corneal oedema resulting in clouding of corneas noted soon after birth, which is often described as having a 'ground-glass' appearance (Figure 4.2 A). The cornea is thickened, and the stroma is swollen due to corneal oedema, but the epithelium is not usually affected. Due to the congenital nature of CHED2, affected individuals can have nystagmus (Mehta *et al.*, 2010). In the second decade, some CHED2 patients also develop sensorineural hearing loss, in which case the condition is defined as Harboyan syndrome (Desir and Abramowicz, 2008). However, examination of 4 patients previously diagnosed with CHED2 found evidence of hearing loss at higher frequencies, therefore hearing loss may be a consistent feature of CHED2, with variable severity and age of onset (Siddiqui *et al.*, 2013). Both CHED2 and Harboyan syndrome are caused by biallelic mutations in *SLC4A11* located on chr20p13 (Vithana *et al.*, 2006). There is no correlation between hearing loss severity and specific *SLC4A11* mutations (Mehta *et al.*, 2010). The SLC4A11 protein is a transmembrane protein with Na^+ coupled OH^- co-transport activity (Jalimarada *et al.*, 2013). SLC4A11 proteins harbouring a mutation fail to localise to the plasma membrane and are retained in the endoplasmic reticulum. Some mutant SLC4A11 proteins retain a degree of water flux function when they are driven to localise to the plasma membrane (Loganathan and Casey, 2014).

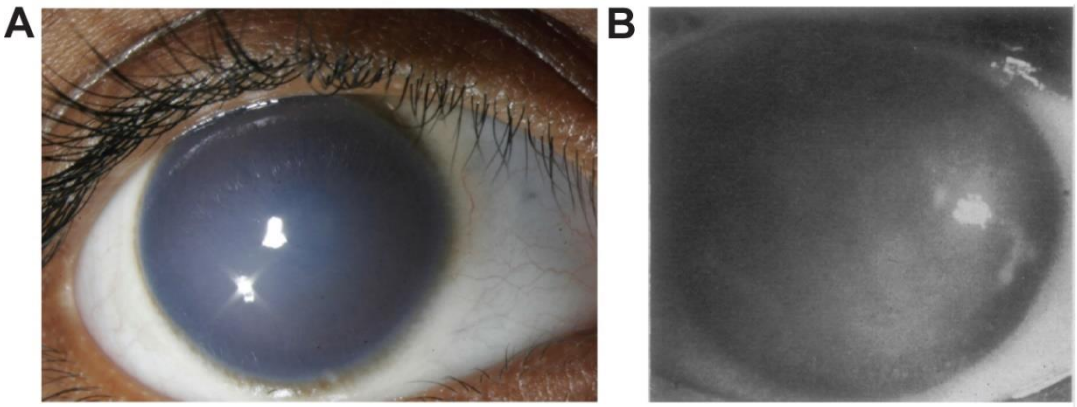


Figure 4.2 Clinical appearance of autosomal dominant and autosomal recessive forms of congenital hereditary endothelial dystrophy (CHED). (A) Clinical image of a cornea with autosomal recessive CHED2 showing opaque cornea due to corneal oedema. Image from (Kodaganur *et al.*, 2013). **(B)** Opaque cloudy cornea of a child with autosomal recessive CHED1 (aged 14 years). Image from (Pearce, Tripathi and Morgan, 1969).

4.1.3 X-linked endothelial corneal dystrophy

XECD has been described in a single large pedigree consisting of 35 affected individuals of both males and females. Affected individuals have corneal clouding and a 'ground-glass' appearance similar to CHED2 (Schmid *et al.*, 2006) (Figure 4.3 A). 'Moon crater'- like lesions of the endothelium were also observed on slit lamp examination that are similar in appearance to vesicles seen in PPCD (Figure 4.3 B). Visual impairment was more severe in males than females and secondary band keratopathy was present in 7 affected males (aged 23 – 75) (Figure 4.3 C). Histology revealed changes within all corneal layers. The epithelium was thinned with irregular collagen lamellae at the anterior stroma. Amorphous material was present in superficial layers consistent with band keratopathy. Descemet's membrane was thickened and the endothelial monolayer was disrupted with regions of multilayered cells and regions with cell loss (Schmid *et al.*, 2006). The pedigree demonstrated affected family members in each generation, consistent with an autosomal dominant or X-linked inheritance pattern. The absence of male-to-male transmission and the greater severity of visual impairment in males, led the authors to conclude that the dystrophy was likely X-linked, with symptomatic females. Linkage analysis with markers on the X chromosome revealed a significant LOD score at chrXq25, however the gene underlying XECD is yet to be identified (Schmid *et al.*, 2006).

No additional reports of XECD have been described in the literature, however due to the overlapping phenotypes of endothelial dystrophies and the difficultly in distinguishing between autosomal dominant and X-linked inheritance with symptomatic females in small pedigrees, it is possible that some cases which have been diagnosed as PPCD or CHED1 may have mutations in the XECD causative gene.

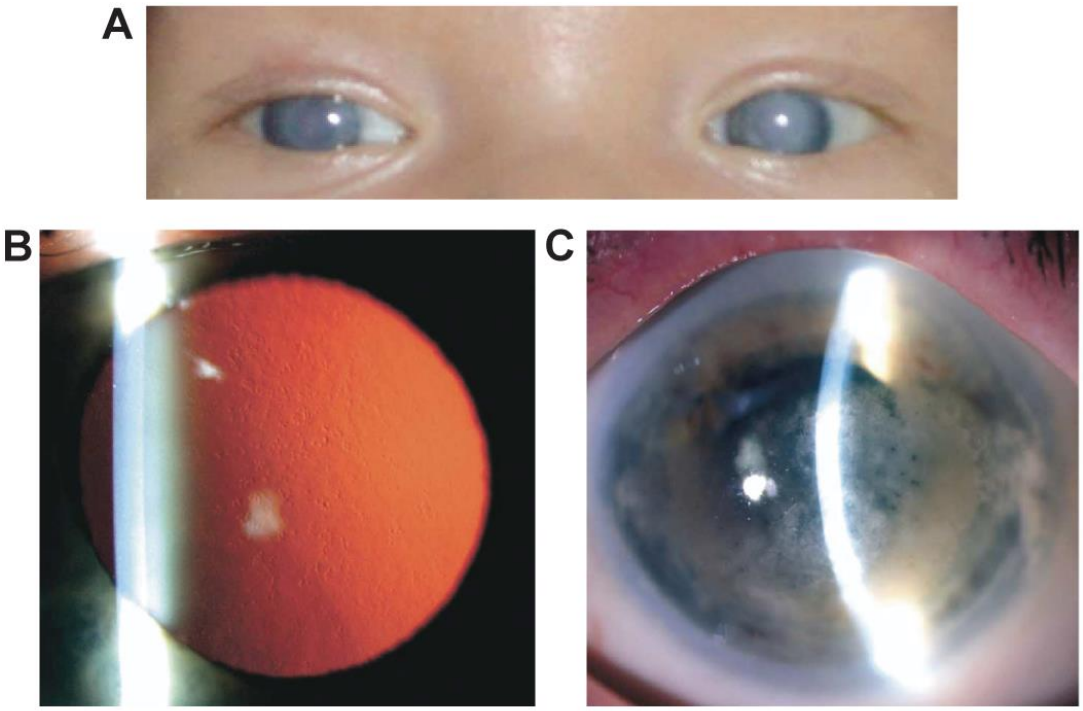


Figure 4.3 Clinical appearance of X-linked endothelial dystrophy. (A) Soon after birth corneas appear cloudy with a ground glass appearance. **(B)** Slit lamp appearance in retroillumination shows presence of 'moon-craters'. **(C)** Clinical appearance showing presence of band keratopathy in affected individual (aged 75) which occurs only in males. Images from (Schmid *et al.*, 2006).

4.1.4 Fuchs endothelial corneal dystrophy

FECD is the most common corneal dystrophy. The prevalence is unknown in the UK, however it is predicted to affect ~ 4% of individuals over the age of 40 in the USA (Lorenzetti *et al.*, 1967). It is characterised by the presence of excrescences of Descemet's membrane called guttae (Figure 4.4). These occur predominantly in the central cornea during disease onset, but become more frequent, spread to the periphery and coalesce as the disease progresses. The increase in guttae is correlated with a decrease in endothelial cell density. As endothelial cells are lost, the remaining cells become larger and lose their regular hexagonal shape. Although patients are initially asymptomatic, as the disease progresses the endothelium becomes compromised due to cell loss, causing blur from corneal oedema and painful epithelial bullae (Elhalis, Azizi and Jurkunas, 2010).

There are several subtypes of FECD, which differ in their underlying genetic basis. The most common is sporadic late-onset FECD which has a complex age-related

aetiology with both genetic and environmental components. A genome-wide association study (GWAS), with a combined discovery and replication cohort of 280 FECD cases and 410 controls, identified a genome-wide significant association between FECD and a SNP (rs613872) located in the third intron of *TCF4* (Baratz *et al.*, 2010). This association was subsequently replicated in other independent cohorts and a meta-analysis that analysed thirteen of these studies, including 2468 FECD cases and 2902 controls, found a statistically significant association between FECD and four SNPs spanning *TCF4* (Li, Peng and Sun, 2015). An intronic trinucleotide repeat within the *TCF4* gene was subsequently found to be expanded in a substantial proportion of FECD cases, with 79% of cases having greater than 50 repeats compared to only 3% of age-matched controls. This trinucleotide expansion is thought to be the functional genetic variation underlying the GWAS association, and is the most significant genetic factor underlying sporadic late-onset FECD (Wieben *et al.*, 2012). Monogenic endothelial dystrophy genes causing PPCD and CHED have been implicated in FECD and may have an overlapping pathogenesis. The identification of mutations in genes associated with other endothelial dystrophies including *COL8A2*, *ZEB1* and *SLC4A11* have been reported from screening cohorts of FECD patients (Gottsch *et al.*, 2005; Vithana *et al.*, 2008; Mok, Kim and Joo, 2009; Riazuddin *et al.*, 2010).

There are a number of familial FECD cases that have an apparent autosomal dominant inheritance pattern. The FCD2 locus was mapped to a common overlapping locus (chr18q21.2 – q21.32) in three unrelated large pedigrees with late-onset FECD (Riazuddin *et al.*, 2012). Inheritance of the dystrophy was consistent with an autosomal dominant pattern of inheritance although with reduced penetrance and phenocopies. Targeted re-sequencing in these families revealed the presence of a mutation in *LOXHD1* in one family that was hypothesised to be causative. In the two remaining linked families there were no variants within the locus that could be potentially causative. Screening *LOXHD1* directly in 207 FECD probands identified a further 15 rare missense changes. The pathogenicity of these variants is not fully resolved as screening a control panel of 288 individuals identified 8 novel missense changes that were also predicted to be pathogenic by pathogenicity prediction programs. Furthermore, the FCD1 locus was mapped to chr13pTel – 13q12.13 in a four generation white family with late-onset FECD (Sundin *et al.*, 2006). The FCD3

locus was mapped to chr5q33.1 – q35.2 in a three generation family with a mild phenotypic variant of late-onset FECD (Riazuddin *et al.*, 2009). The genetic variation underlying these associations has not yet been identified.

Digenic inheritance has been described in one FECD pedigree. This family consisted of a large autosomal dominant pedigree with late-onset FECD, for which the other known FECD loci were excluded. Linkage analysis failed to identify a single significant locus, therefore the analysis was repeated using a multi-locus model; this identified two suggestive loci on chr3p and chr15q, however both of these failed to reach statistical significance. WES analysis of two affected and one unaffected individual from the pedigree identified no potentially pathogenic variants on chr3p, however a c.3082C>T; p.(Arg1028*) variant in *AGBL1* was considered to be causative in the chr15q linkage interval. Screening *AGBL1* in an unsolved FECD cohort identified the same variant in a further two cases and an additional *AGBL1* variant in a third. The authors argue that the occurrence of both these variants in control databases at low frequencies does not exclude a causative role given that ~ 4% of the population will develop FECD (Riazuddin *et al.*, 2013)**.** The majority of these genetic findings for FECD, with the exception of *TCF4* repeat expansion, require further replication and validation.

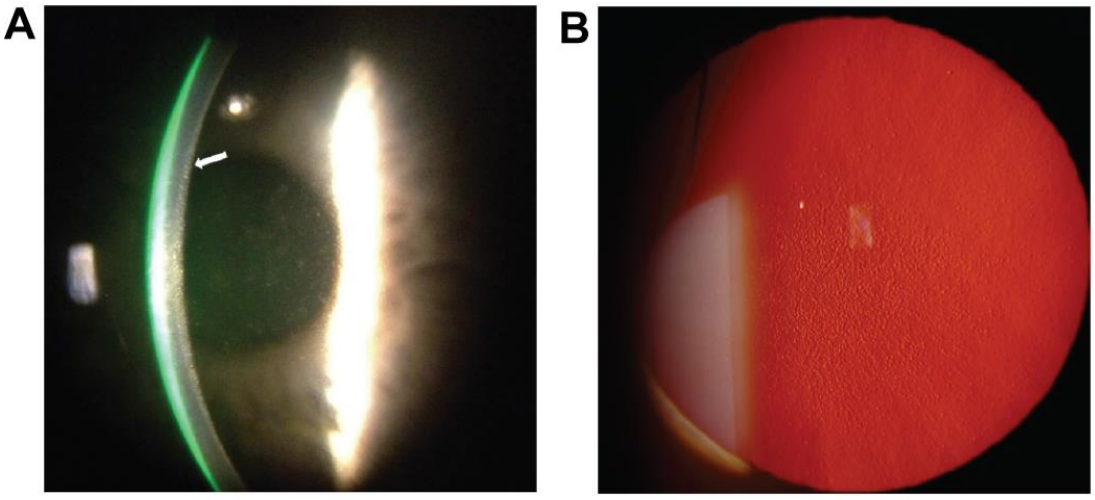


Figure 4.4 Clinical images of Fuchs endothelial corneal dystrophy. (A) Direct illumination and **(B)** retroillumination showing appearance of guttae (arrow). Images from (Eghrari, Riazuddin and Gottsch, 2015).

4.1.5 Aims of chapter

There is evidence that several genes causing corneal endothelial dystrophies remain to be identified, including PPCD1 and CHED1 on chr20p, XECD and FECD. The aim of this study was to identify the genetic cause and mutational spectrum of monogenic endothelial dystrophies in patients attending Moorfields Eye Hospital, and to investigate any potentially novel endothelial disease genes. Recruitment and screening of FECD cases was beyond the scope of this study, however given the overlap between genetic causes of endothelial dystrophies, genes implicated in FECD were considered potential candidates when analysing probands diagnosed with a corneal endothelial dystrophy.

4.2 Clinical features of the PPCD cohort

PPCD was the most common endothelial corneal dystrophy recruited to the study with 28 probands recruited from MEH. Analysis of the clinical features of this cohort revealed two clinical sub-types. The majority displayed typical signs of PPCD with vesicles; however, an asymptomatic phenotypic variant of PPCD that consisted of one or multiple band-like lesions with a scalloped edge, often with markedly asymmetric corneal involvement in the absence of other clinical signs of PPCD, was also observed. Henceforth these cases are referred to as band-type PPCD. Although typically asymptomatic, one proband with band-type PPCD (F168: II:2) required DSAEK surgery due to corneal oedema in her left eye, indicating that disease progression does occur in some cases. The proband of F165 had bleb-like endothelial lesions in addition to band-type PPCD, which represents an intermediate clinical state between band-type and typical PPCD.

The endothelial dystrophy cohort was expanded through collaboration with Dr. Petra Liskova (First Faculty of Medicine, Charles University of Prague) and Dr. Andrea Vincent (University of Auckland, New Zealand). An additional 14 probands with PPCD were included in the study from the Czech Republic and 6 probands from New Zealand.

4.3 *ZEB1* coding mutations

All nine coding exons of *ZEB1* were amplified and sequenced for 28 PPCD patients recruited from MEH, regardless of clinical sub-type (section 2.2; Appendix A) Czech and New Zealand probands had previously been screened by the referring centres for *ZEB1* mutations and were negative. Screening resulted in the identification of *ZEB1* mutations in probands from seven families (F124 – F130; Figure 4.5; 4.7; Table 4.2), all of which had typical PPCD symptoms of endothelial vesicles when examined with a slit lamp and abnormalities of endothelial cell shape and morphology when examined with specular microscopy (Figure 4.6).

The proband II:1 from family F124:II:1 (Figure 4.5 A) had poor vision since childhood and had undergone surgery for strabismus at the age of 7 years. Apart from changes in the posterior corneal layers there were no other documented ocular abnormalities. The father (I:1) of the proband was diagnosed with PPCD at the age of 4 years, and although his parents were not examined, both were known to have good visual acuity. No systemic abnormalities were documented in I:1 and II:1, but specific enquiry about the extraocular features associated with PPCD3 was not possible as the family was lost to follow-up. Screening identified a novel *ZEB1* c.1100C>A; p.(Ser367*) mutation in exon 7 in the proband which was inherited from his affected father.

The proband from family F125: II:1 (Figure 4.5 B) was first seen at the age of 32 years when she was referred for a PK in her right eye. She subsequently had a further three grafts in the same eye. There was no history of ocular disease in her parents. Details of systemic features were unavailable as the patient was lost to follow up. Screening identified a previously reported *ZEB1* c.1576dupG; p.(Val526Glyfs*3) mutation in exon 7.

The proband F126: II:1 (Figure 4.5 C) was clinically examined at 13 years old, when there were diffuse endothelial changes. She was negative for iris or other anterior segment abnormalities and intraocular pressures were normal. Her mother was examined and found to be unaffected. Screening identified a novel *ZEB1* mutation c.1669C>T; p.(Gln557*) in exon 7.

The proband F127: II:1 (Figure 4.5 D) was recruited in her fifties with bilateral PPCD. She underwent right eye DMEK, which improved her visual acuity, and the same operation is planned for her left eye in the near future. Screening revealed the presence of a novel *ZEB1* c.1918C>T; p.(Gln640*) mutation in exon 7.

The proband F128: II:1 (Figure 4.5 E) of Sri Lankan origin (Tamil) was negative for a family history of corneal disease although no other family members were available for examination. He was noted to have poor vision at 7 years of age. At the age of 29 years he was suspected to have glaucoma, with IOP measurements of 24 mmHg in his left eye, but with normal optic discs and open anterior chamber angles. Vesicular endothelial lesions were noted on clinical examination (Figure 4.6 C) and specular microscopy showed enlarged endothelial cells and an uneven posterior corneal surface (Figure 4.6 D). He had a right posterior lamellar keratoplasty with a subsequent best corrected visual acuity of only 0.18 despite a clear graft, suggesting the presence of amblyopia. The patient did not report any general health problems; however, examination for the extraocular features associated with PPCD3 was not possible. The *ZEB1* mutation identified in F128: II:1 occurs in the canonical splice acceptor site flanking exon 6: c.685-2A>G. Splice acceptor (AG) and donor sites (GT) are highly conserved motifs that flank exon-intron boundaries and are essential for splicing of the pre-mRNA molecule. The c.685-2A>G mutation is therefore expected to result in aberrant splicing of the *ZEB1* transcript. Ideally, to test this hypothesis, RNA would be derived from patient leukocytes and splicing of the transcript would be observed using RT-PCR to amplify across exons 5, 6 and 7. However in this case, the proband had returned to Sri Lanka and it was not possible to obtain a fresh blood sample. Evidence for this hypothesis was therefore sought using in silico splice prediction tools. Three individual algorithms, Human Splice Finder, NNSPLICE and NetGene2, identified the presence of canonical splice acceptor site in the wild type sequence as the highest scoring site and all predicted that it would be abolished by the presence of the c.685-2A>G mutation (Table 4.1). One possible consequence is that exon 6 would be skipped in the transcript, which would result in exon 5 being fused to exon 7. If this were to occur, the functional consequences for the resulting protein would be a frameshift and incorporation of a premature stop codon: p.(Arg229Glufs*24). Alternatively, all three splice prediction tools predicted that an alternative splice acceptor site 440 bp upstream from the

canonical splice acceptor site would be the highest scoring acceptor site in the mutant transcript (Table 4.1). If this splice acceptor site was used then the resulting mRNA transcript would incorporate an additional 440 bp of intronic sequence into the mRNA transcript and result in a frameshift and incorporation of a premature stop codon on translation: p.(Arg229Aspfs*9) (Figure 4.7 B).

Table 4.1 In silico prediction of potential effect of the *ZEB1* c.685-2A>G mutation proband F128: II:1 on the canonical splice acceptor site. Human Splicing Finder, NNSPLICE and NetGene2 each identify potential splice sites in a given sequence and score them from 0 to 100, 0-1 and 0-1, respectively. In each instance, a higher number corresponds to an increased likelihood of the motif functioning as a splice site. All three programs predict that the presence of the mutation abolishes the canonical splice acceptor site.

	Canonical splice acceptor score	**Canonical splice acceptor score in presence of c.685-2A>G**	**Alternative splice acceptor score**
Human Splice Finder (0 – 100)	93.2	0	92.89
NNSplice (0 – 1)	0.98	0	0.90
NetGene2 (0 – 1)	0.85	0	0.16

Individual II:1 from family F129 (Figure 4.5 F) reported poor vision since childhood with gradual further deterioration. He was first examined at MEH at the age of 38 years. At the age of 39 years he underwent a left PK followed by a right PK when he was 41 years old. Except for corneal changes consistent with the diagnosis of PPCD, no other abnormalities of the anterior segment were noted. The patient reported that his father also had poor vision; however, no further details could be obtained. His three siblings had no eyesight problems but were not available for examination. The proband had a history of an enlarged heart since childhood and was recommended for pericardiectomy when he was 37 years old. Screening identified a *ZEB1* c.627delT; p.(Phe209Leufs*11) mutation in exon 5.

The proband F130: I:2 (Figure 4.5 G) was recruited in her fifties, and had been diagnosed with PPCD over twenty years previously. She was severely affected with corneal oedema and areas of Decemet's membrane detachment. She also had

extensive anterior synechiae (Figure 4.6 B). Her 25 year old son (II:1) was also affected, although less severely and no iris abnormalities were present (Figure 4.6 A). Screening revealed the presence of a *ZEB1* c.2577delA; p.(Val860*) mutation in exon 7 that was present in both the proband and her affected son.

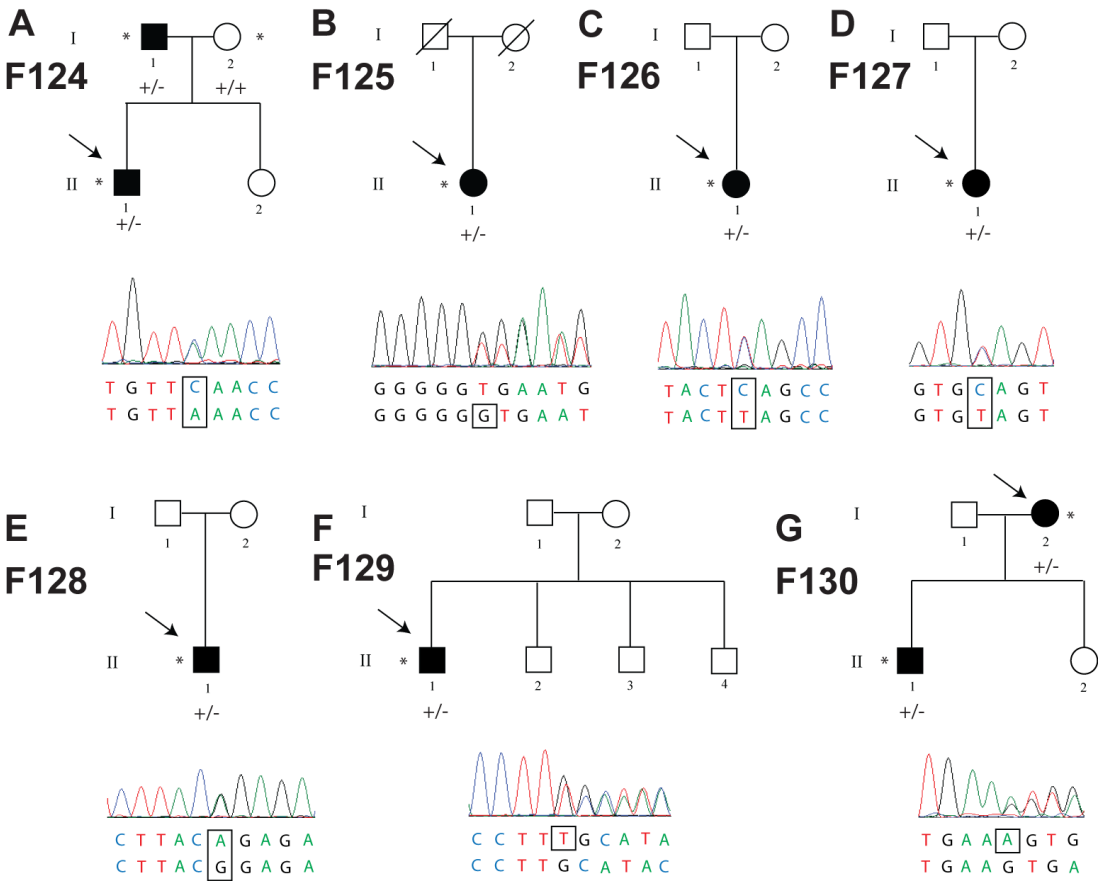


Figure 4.5 Summary of PPCD families F124 – F130 with coding mutations in *ZEB1*. Seven families from MEH are shown with corresponding heterozygous mutations in the coding region of *ZEB1*. +/+ indicates individual is wild type and +/- heterozygous. * indicates patient has been examined and a blood sample taken. **(A)** F124 and c.1100C>A; p.(Ser367*) mutation. **(B)** F125 and c.1576dupG; p.(Val526Glyfs*3). **(C)** F126 and c.1918C>T; p.(Gln640*). **(D)** F127 and c.1669C>T; p.(Gln557*). **(E)** F128 and c.685-2A>G. **(F)** F129 and c.627delT; p.(Phe209Leufs*11). **(G)** F130 and c.2577delA; p.(Val860*).

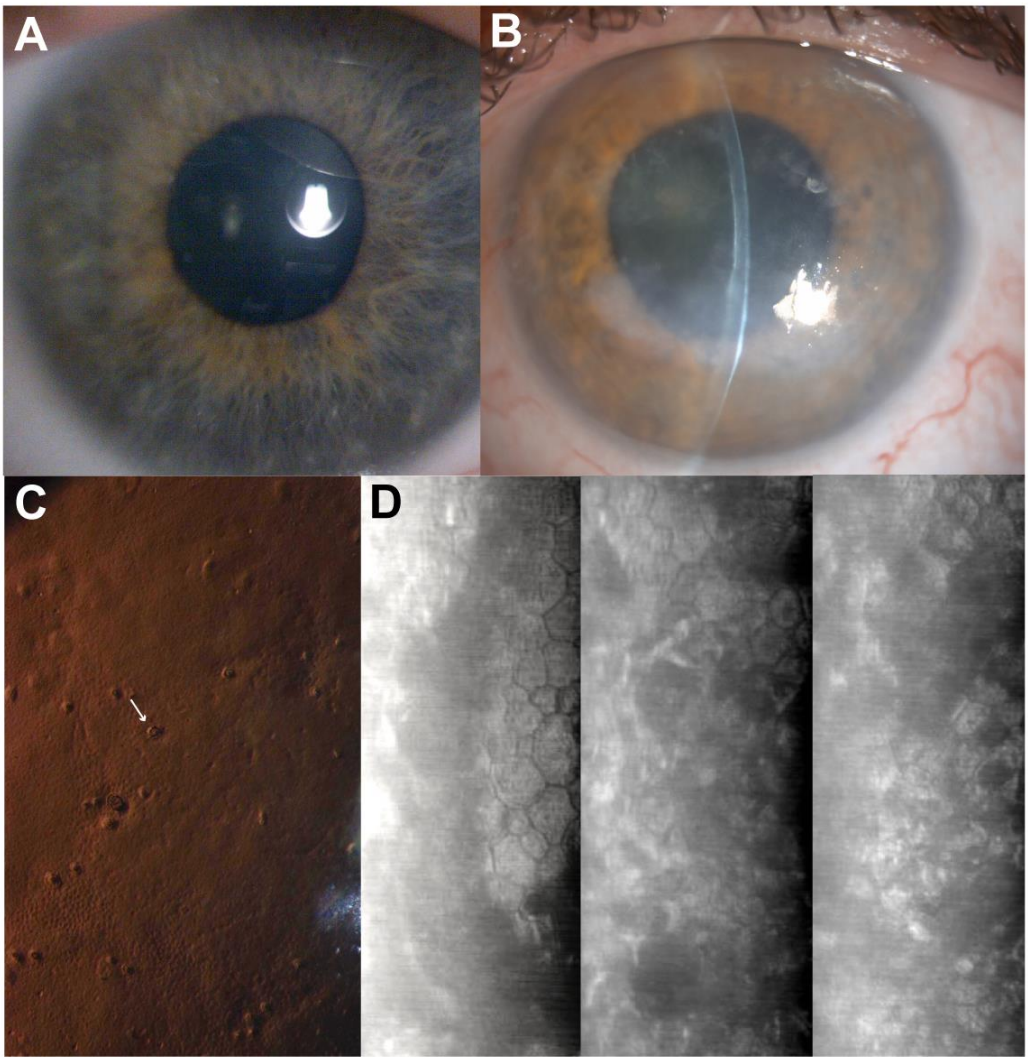


Figure 4.6 Summary of clinical features of *ZEB1* mutation positive patients with posterior polymorphous corneal dystrophy. (A) F130:II:1 **(B)** F130:I:2 **(C)** Slit lamp in retroillumination with presence of vesicles in proband F128: II:1, an example of which is indicated with a white arrow. **(D)** Specular microscopy showing overgrowth of endothelial cells and irregular endothelial cell shape and size in proband F128: II:1. Endothelial cells should be regular in hexagonal shape and size.

Six of the mutations were novel whereas one had been previously described as pathogenic. All seven were absent from the ExAC database (Table 4.2). All *ZEB1* mutations identified were private to the affected family, except for the c.1576dup; p.(Val526Glyfs*3) mutation which is the fourth reported occurrence of this mutation, indicating that this may be a mutation hotspot (Krafchak *et al.*, 2005; Bakhtiari *et al.*, 2013; Lechner *et al.*, 2013). Consistent with previously published reports, all seven of the mutations identified are predicted to cause *ZEB1* haploinsufficiency due to the incorporation of a premature stop codon into the transcript. The resulting mRNA transcripts would be expected to be degraded by nonsense mediated decay or be translated into non-functional truncated proteins. Mutations are annotated on the ZEB1 protein structure in Figure 4.7 showing their distribution across the protein.

Table 4.2 Summary of coding mutations in *ZEB1* in the posterior polymorphous corneal dystrophy cohort. One mutation affected a canonical splice acceptor, four are nonsense mutations and two are frameshifts. All mutations identified were absent from the ExAC database. Six were novel and one had been previously reported as a cause of PPCD. NP: not present.

Family	Nucleotide	Protein	ExAC	Ref.
F124	c.1100C>A	p.(Ser367*)	NP	(Krafchak *et al.*, 2005)
F125	c.1576dupG	p.(Val526Glyfs*3)	NP	
F126	c.1669C>T	p.(Gln557*)	NP	
F127	c.1918C>T	p.(Gln640*)	NP	
F128	c.685-2A>G	Uncharacterised	NP	
F129	c.627delT	p.(Phe209Leufs*11)	NP	
F130	c.2577delA	p.(Val860*)	NP	

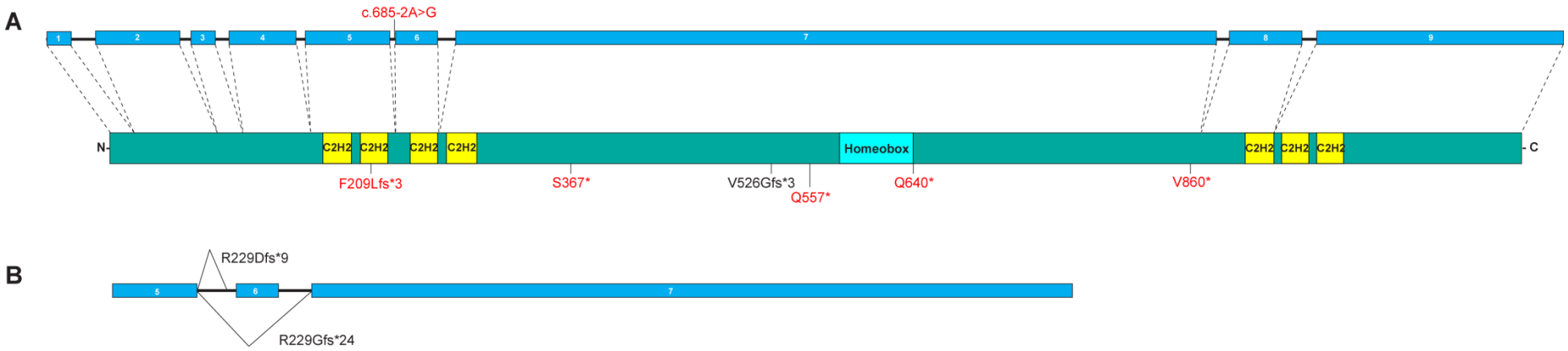


Figure 4.7 Gene and protein structure of ZEB1 and summary of coding mutations identified in the study cohort. (A) Seven *ZEB1* coding mutations were identified in probands with posterior polymorphous corneal dystrophy, of which 6 were located in exons and one was located in intron 5/6. The majority of mutations occurred in exon 7, the largest exon of the 9 coding exons. Exons are drawn to scale, introns are included for exon/intron boundaries but are not to scale. The ZEB1 protein includes 7 C2H2 domains and an atypical homeobox domain. **(B)** Prediction of splicing consequences of c.685-2A>G mutation. One possibility is use of an alternative splice acceptor site 440 bp upstream from the canonical splice acceptor site, resulting in p.(Arg229Aspfs*9). Alternatively, exon 6 may be skipped resulting in p.(Arg229Glufs*24).

4.4 *ZEB1* deletions

4.4.1 Identification of a ZEB1 deletion in a Czech PPCD pedigree

A white Czech family (F131), with at least three generations of individuals diagnosed with PPCD, was recruited to the study (Figure 4.8 A). Following negative pre-screening for mutations in *ZEB1* coding exons, DNA from the proband F131: III:1 and affected relatives II:2 and IV:1 was analysed by WES (section 2.2; 2.6; Appendix A). Manual interrogation of WES data revealed inconsistencies in the segregation of variants flanking the *ZEB1* gene between related individuals; for example, the SNP rs2994626 was homozygous T in II:2 but present as a homozygous C allele in her son (III:1). The WES data also revealed a run of putative homozygous SNPs spanning the *ZEB1* locus and flanking genes, indicating that a deletion may be present at this locus. On the basis of this evidence, two affected individuals from the family (III:1 and IV:1) were analysed by HumanCytoSNP-12 v2.1 BeadChip (Illumina) genotyping array containing ~ 300,000 markers, with a mean genome wide spacing of ~ 10 kb (section 2.7). This resulted in the identification of a large deletion spanning *ZEB1* in the proband (Figure 4.8 B).

Data from the SNP array genotyping for proband F131: III:1 predicted that the distal break point would be located between chr10.hg19: 29,272,115 – 29,290,203. This was refined by ~ 3 kb using WES data by identification of heterogeneous SNPs. A homozygous SNP was identified at chr10.hg19:29,287,696 (rs788051) which did not correctly segregate in the family, indicating that this region is deleted (data not shown). The proximal breakpoint was predicted to occur between chr10.hg19: 32,538,176 – 32,563,922, however this could not be refined using WES data. Due to the large interval where the breakpoints may lie, the region containing the proximal break point was manually refined by genotyping potentially polymorphic SNPs across the interval (MAF ~ 0.3 – 0.7) (section 2.2; Appendix A). Manual genotyping of SNPs surrounding the predicted breakpoint was undertaken to refine the interval, and the data are summarised in Table 4.3. Using this strategy, the size of the region containing the proximal breakpoint was refined from ~ 26 kb to ~ 12 kb.

Table 4.3 Refinement of deletion breakpoints in F131 by manual genotyping of small nucleotide polymorphisms (SNPs). Sanger sequencing of putative polymorphic SNPs was undertaken to identify inconsistencies in segregation of alleles that would be indicative of a deletion at that locus. Heterozygosity was taken as evidence of the presence of two alleles.

rs number	Genomic co-ordinates (hg19)	Reference allele	Alternative allele	Conclusion
rs12268212	chr10: 32,538,176	Homozygous. III:1	Homozygous. II:2, III:2, IV:1	Deleted
rs2806139	chr10: 32,542,456	Homozygous. III:1	Homozygous. II:2, III:2, IV:1	Deleted
rs2806144	chr10: 32,554,586	Heterozygous: II:2, III:2		Not Deleted

Once refined, long range PCR was used to amplify across the co-ordinates of the deletion in three affected family members from the pedigree (II:2, III:1, IV:1). This produced an amplimer approximately 3 kb in size for all affected individuals, but was unable to amplify a wild-type larger product in an unaffected control DNA sample (Figure 4.8 C). Gel extraction and sequencing across the deletion break points with an internal primer confirmed the presence of a ~ 3.2 Mb deletion in affected family members (section 2.2.4.2). The deletion spanned chr10.hg19: 29,277,299 – 32,553,577 and occurred at a 36 bp sequence that was identical between the proximal and distal breakpoints, one copy of which was encompassed by the deletion (Figure 4.8 D). The deletion encompassed genes *LYZ1, PTCHD3P1, SVIL, KIAA1462, MTPAP, MAP3K8, LYZL2, ZNF438, ZEB1, ARHGAP12* and *KIF5B* in their entirety; however, no other non-ocular phenotype was identified in the family, indicating that haploinsufficiency of these genes, except for *ZEB1*, does not cause any overt clinical phenotype.

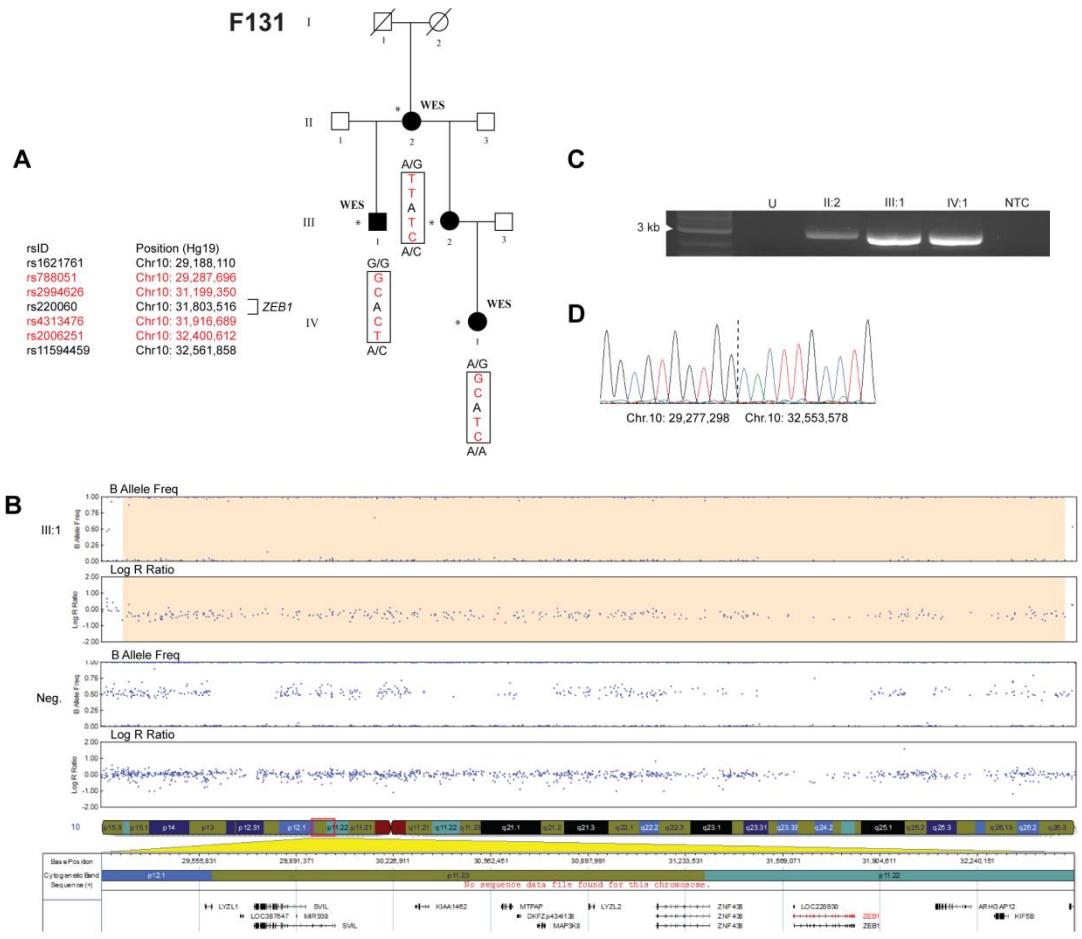


Figure 4.8 Large heterozygous deletion encompassing *ZEB1* in Czech PPCD family #131. (A) Pedigree structure of affected family showing inconsistent segregation of WES variants between affected family members (red) spanning the *ZEB1* locus. **(B)** SNP array copy number analysis revealed the presence of a large deletion encompassing multiple genes including *ZEB1*. Highlighted regions show a heterozygous ~ 3.3 Mb deletion was detected in individual F131: III:1 (top two traces) compared to a patient sample who was negative for a *ZEB1* deletion (Neg). Blue dots represent individual array probes. The horizontal axis indicates genomic position of probes and the vertical axis shows allele frequency (top trace for each individual) or log R ratio (bottom trace). **(C)** Amplification product in three affected individuals from F131 (II:2, III:1, IV:1) using primers flanking the predicted deletion break points. No PCR amplification was achieved using DNA from an unaffected (U) control or in non-template control (NTC). Size marker shows 3 kb marker band highlighted with white arrowhead. **(D)** Direct sequencing of PCR product identifies the deletion breakpoints annotated as chr10.hg19: 29,277,299_32,553,577del.

4.4.2 SNP array analysis of Czech PPCD probands

Probands from four additional unsolved Czech PPCD families (F132, F144, F145, F146), previously identified as *ZEB1* negative by direct Sanger sequencing, were analysed using a HumanOmniExpress BeadChip (Illumina) genotyping array containing ~ 730,000 markers with a mean genome wide spacing of ~ 4 kb (section 2.7). This analysis identified a partial *ZEB1* deletion in the proband of F132: II:1 (Figure 4.9 A) who was negative for a family history of PPCD (Figure 4.9 B).

Due to the use of a denser genotyping SNP array, manual refinement of deletion break points was not required for F132: II:1. Primers were designed to flanking regions surrounding runs of putative homozygosity and long range PCR was used to amplify across the putative break points. A PCR product ~ 6 kb in size was observed, which was not present in the unaffected control DNA (Figure 4.9 C). The band was excised from the agarose gel, and the DNA extracted for direct Sanger sequencing (section 2.2.4.2). Sequencing across the break point revealed the deletion has coordinates chr10.hg19: 31,784,795 – 31,855,747 (70,953 bp in size) and extends from *ZEB1* intron 3-4 to beyond the gene although no other genes are encompassed. Within the deletion, a short dinucleotide AC repeat is inserted (x4), which matches the terminal dinucleotides of the distal breakpoint (Figure 4.9 D).

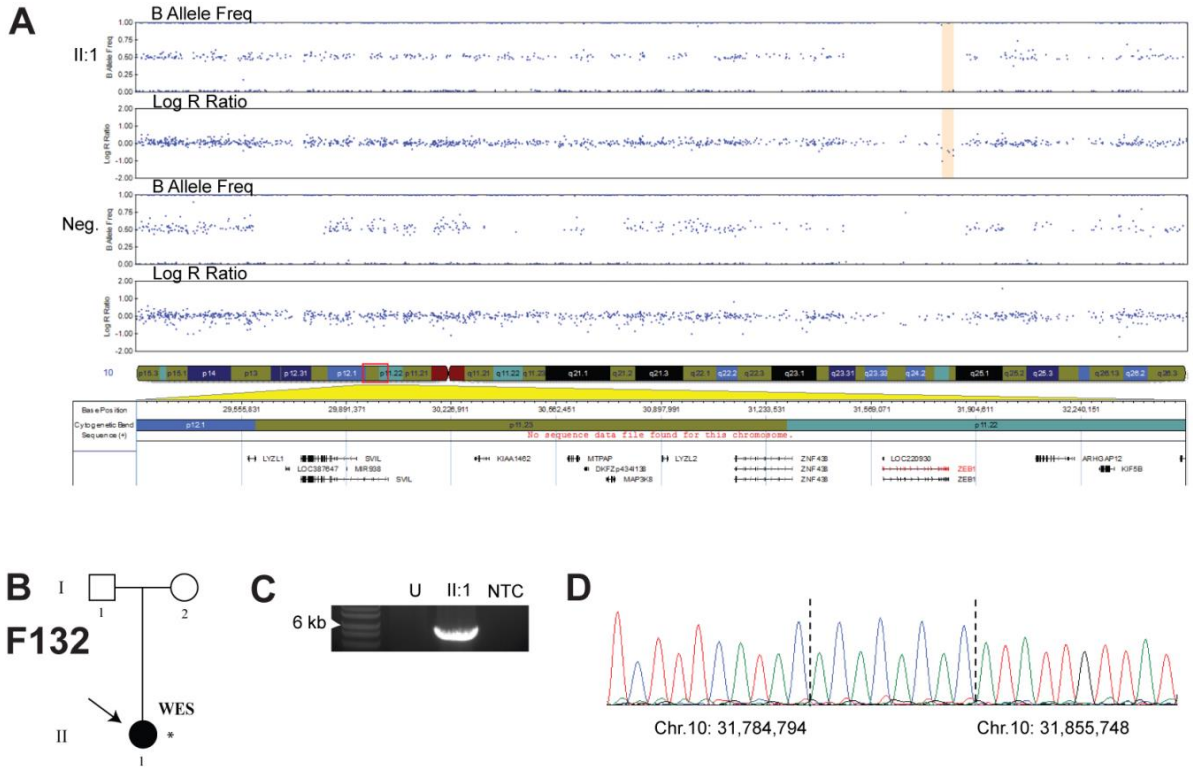


Figure 4.9 Identification of a partial *ZEB1* deletion in family #132. (A) Highlighted region shows a putative deletion spanning *ZEB1* in F132: II:1 (first and second traces) compared to a patient who was negative for a *ZEB1* deletion (Neg). Blue dots represent individual array probes. The horizontal axis indicates genomic position of probes and the vertical axis shows allele frequency (top trace for each individual) or log R ratio (bottom trace). **(B)** Pedigree of F132: II:1 showing absence of family history. **(C)** PCR amplification using primers flanking the predicted deletion for F132: II:1 results in amplification in the proband but not in unaffected (U) or non-template control (NTC). Size marker shows 6 kb marker band highlighted with white arrowhead. **(D)** Direct sequencing reveals a partial *ZEB1* spanning intron 3-4 to the terminal exon and is annotated as chr10.hg19:g.31,784,795_31,855,747delinsACACACAC.

4.4.3 Screening an unsolved PPCD cohort using a qPCR assay

Following the identification of two *ZEB1* deletions in the genetically unsolved Czech PPCD cohort, it was hypothesised that the proportion of individuals with PPCD3 was likely an underestimate, as large deletions that could be identified by the genotyping array would not be detected by Sanger sequencing of *ZEB1*. A qPCR assay was therefore designed as a method to screen the unsolved British (F133, F140) and Czech (F141, F144, F145, F146, F155, F157) PPCD cohort (section 2.8). In addition, DNA from six *ZEB1*-negative PPCD probands (F134 – 139) was obtained from New Zealand for inclusion in the screen. Czech probands (F131, F132, F145 and F146), previously analysed for copy number variations by SNP array, were included as controls.

Primers were designed to target the 5' (intron 1) and 3' (intron 8) end of the *ZEB1* gene, allowing the assay to detect partial *ZEB1* deletions, as previously identified in F131. Following amplification, fluorescence levels of amplified product for each proband were compared to fluorescence levels from a standard curve using control DNA. A reference gene, *TWIST1*, with an assumed copy number of 2 was also amplified. The Database of Genomic Variants (DoGV) revealed that there are no copy number variations in *TWIST1* in the control population and deletion of *TWIST1* causes Saethre–Chotzen syndrome which has a distinctive faciocranial phenotype (Howard *et al.*, 1997), therefore PPCD patients were assumed to have 2 copies of this gene. Amplification for each sample was carried out in triplicate (n=1).

Average interpolated genomic DNA (gDNA) levels of *ZEB1* (5' or 3') in the patient samples were normalised to *TWIST1* gDNA levels to obtain a *ZEB1/TWIST1* ratio (Figure 4.10). A ratio of ~ 1.0 indicates a *ZEB1* copy number of 2, with a ratio of ~0.5 corresponding to a copy number of 1. The proband of F131 (III:1) had a copy number of 1 for both the 5' and 3' *ZEB1* qPCR assays, which acted as a positive control for this assay, as the SNP array, WES and breakpoint sequencing data confirmed heterozygous deletion at these loci. Consistent with the SNP array data, the 3' *ZEB1* qPCR assay identified F132: II:1 as having 1 copy, whereas the 5' *ZEB1* qPCR assay demonstrated the presence of 2 copies. These data indicate a partial deletion of the *ZEB1* gene encompassing the 3' end of the gene which is supported

by SNP array data and sequencing of the breakpoints. The unsolved PPCD proband from MEH, F133, was also found to have a deletion spanning the 5' end of the *ZEB1* gene (Figure 4.10).

The remaining probands all had a copy number of 2 for the *ZEB1* 5' end. The majority of samples also demonstrated a copy number of 2 for the *ZEB1* 3' end, however this assay appeared less robust and a sub-set of samples consistently resulted in a *ZEB1*/*TWIST* ratio of 1.5 or greater, corresponding to copy numbers of 3 or greater (data not shown). This is likely to be an artefact of the assay, as the proband of F146 showed no evidence of copy number variations in *ZEB1* with SNP array analysis to support a true duplication or triplication event. In addition, several attempts of this assay with different control DNA samples confirmed that the 3' assay is less robust, as several resulted in a *ZEB1*/*TWIST* ratio of 1.5 or greater (data not shown). The effect appeared to be sample specific, as it occurred on each re-run of the assay in specific samples only, perhaps due to DNA quality. Melt curve analysis indicated a single product was being amplified by during the qPCR reaction. Primer pairs were re-designed for the 3' end of *ZEB1* targeting the untranslated region (UTR), however these also resulted in the same pattern of results with an estimated copy number of 2 for an unaffected control, a copy number of 1 for F131 and a spurious result for sample F146 (Appendix A). In conclusion, although a robust assay was developed for 5' *ZEB1* deletions, the 3' assay was not consistent and therefore could not be used for conclusive screening of the unsolved cohort.

Results from the *ZEB1* qPCR assay at the 5' end of the *ZEB1* gene consistently showed a *ZEB1/TWIST1* ratio of 1.0 for the unaffected control and 0.5 for the F131 positive control, corresponding to a *ZEB1* copy number of 2 and 1, respectively. The unsolved PPCD samples screened all displayed a copy number of 2 except for F133: II:1, who was a MEH patient previously designated as negative for *ZEB1* coding mutations with no family history of PPCD (Figure 4.10 A). It was concluded that F133: II:1 had a deletion that encompassed at least the 5' end of the *ZEB1* gene and likely extended to the 3' end of the gene (Figure 4.10 B). Confirmation of the deletion in this proband was obtained using the HumanCytoSNP-12 v2.1 BeadChip (Illumina) genotyping array, which identified the presence of a deletion encompassing the entire *ZEB1* gene with predicted co-ordinates

chr10.hg19:g.(31,298,325_31,350,494)_(32,214,122_32,285,574) del (section 2.7; Figure 4.11). A summary of all deletions spanning the *ZEB1* locus is shown in Figure 4.12.

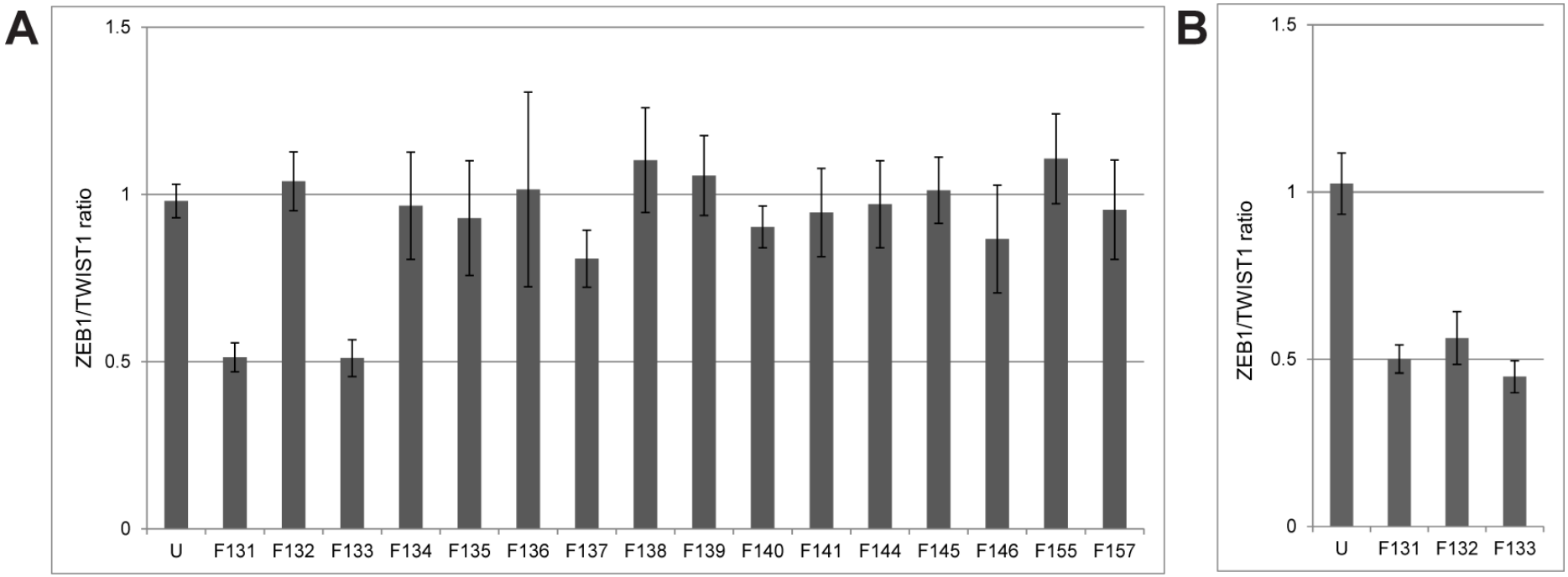


Figure 4.10 Summary of qPCR screening for copy number variations in *ZEB1*. Positive and negative controls include samples that were also analysed by other methods and an unaffected (U) control sample. **(A)** 5' *ZEB1* qPCR assay and **(B)** 3' *ZEB1* qPCR assay. Deletions at the 5' end were identified in F131 and F133 (ratio of 0.5), and at the 3' end of the gene in F131, F132 and F133 (ration of 0.5). Each reaction was carried out in triplicate, with n=1. Results show average ZEB1/TWIST1 ratio +/- standard deviation (n=1).

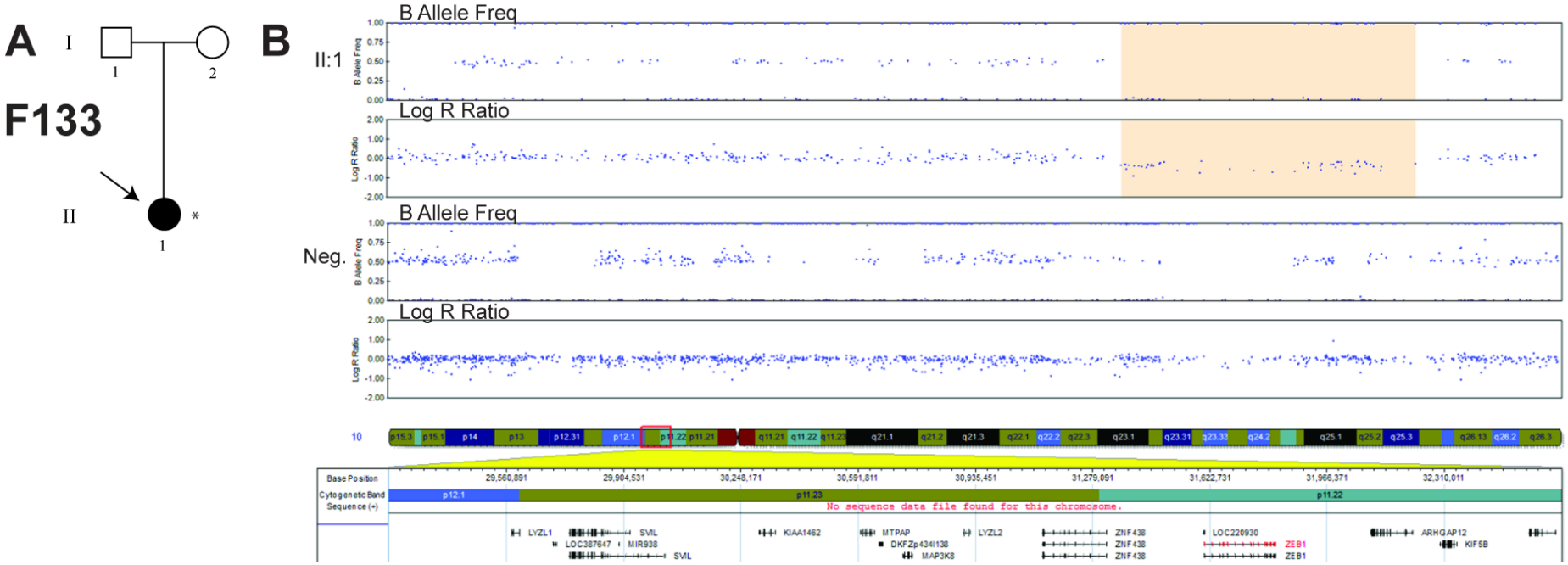


Figure 4.11 SNP array copy number variation analysis of the *ZEB1* region in family #133 confirms the presence of a deletion identified using qPCR. (A) Pedigree structure of F133: II:1 harbouring a deletion encompassing *ZEB1*. **(B)** The results of proband F133:II:1 are shown with a proband who was negative for a *ZEB1* deletion, for comparison. Blue dots represent individual array probes. The horizontal axis indicates genomic position of array probes and the vertical axis indicates the allele frequency (top trace for each individual) or log R ratio (bottom trace for each individual).

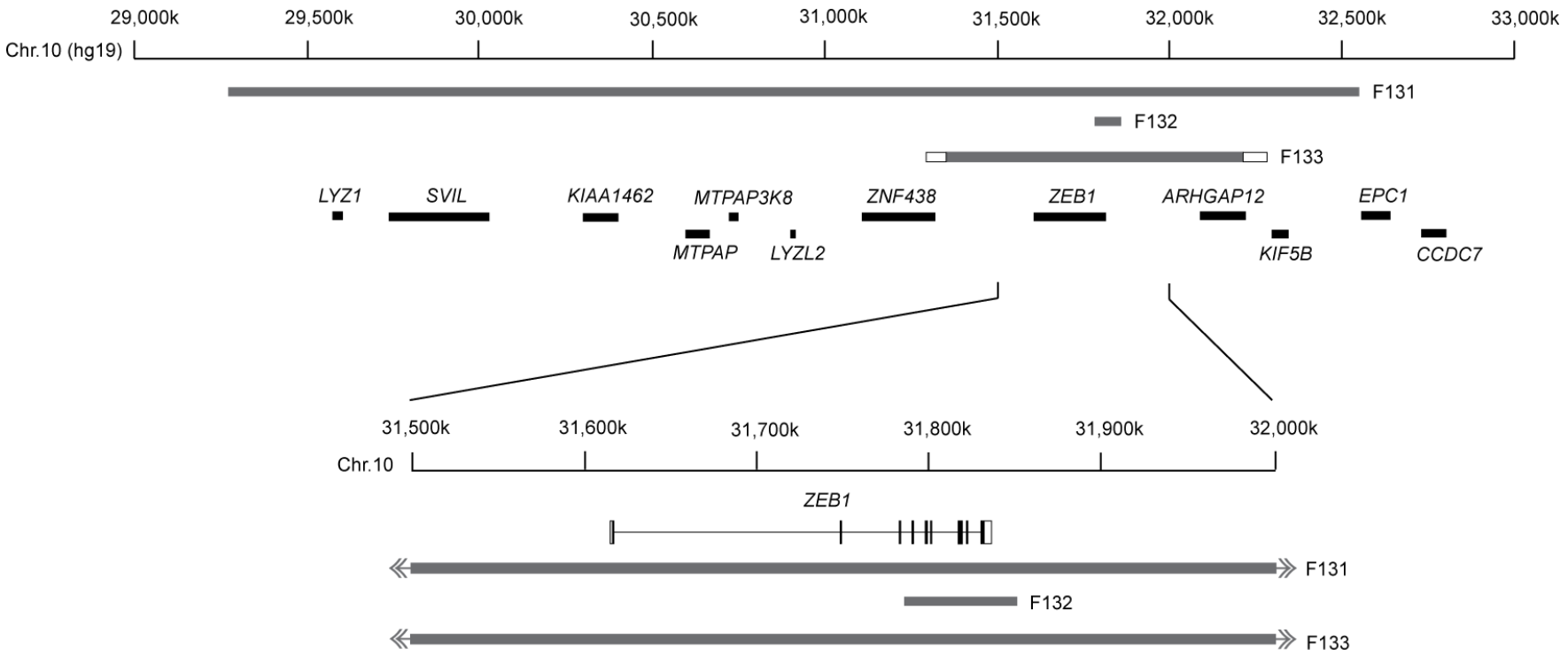


Figure 4.12 Summary of *ZEB1* deletions identified in probands with posterior polymorphous corneal dystrophy (PPCD). Locations of genes and PPCD causing deletions are shown on hg19 scale. The deletions identified in F130, F131 and F132 are different sizes and do not share any common breakpoints. The deletion in F131 spans ~ 3.3 Mb and encompasses the entire *ZEB1* gene and nine other protein coding genes. F132 had a ~ 71 kb deletion which spans the 3' end of the *ZEB1* gene from intron 3 onwards. F133 had a 0.8 – 1 Mb deletion encompassing the entire *ZEB1* gene and at least partially encompasses the flanking gene *ARHGAP12*. Grey boxes indicate boundaries of deletions identified in the PPCD cohort. White boxes are unknown regions in probands where the exact boundaries of the deletion have not been mapped (F133).

4.5 PPCD1/CHED1 locus

4.5.1 Linkage of two endothelial dystrophy families to chr20p

Two families with autosomal dominant endothelial dystrophies that genetically link to an overlapping locus on chr20p had been identified previously. Family F140 from MEH was diagnosed with CHED1; a congenital or early-onset disorder with symptoms of epiphora, photophobia and corneal haze (section 4.1.2; Figure 4.16). All affected individuals have received at least one corneal graft with many undergoing multiple procedures following graft failure and secondary glaucoma. Keratoprosthesis have been implanted in three individuals and three individuals have had an eye enucleated. Previous linkage analysis has identified the genetic locus to be located between markers D20S48 and D20S471 on chr20 (Toma *et al.*, 1995).

The second pedigree included 16 families from the Czech Republic diagnosed with PPCD who were initially presumed to be unrelated, but were subsequently shown to share a common haplotype on chr20p and therefore likely inherited the mutation from a common ancestor (Liskova *et al.*, 2012). These families are collectively referred to as F141 as they can be considered to be part of larger pedigree. Disease onset was primarily in the second decade although in rare cases it was as early as 5 years old. Clinical signs include irregularities of the posterior corneal surface, focal opacities and geographic regions of endothelial cells with an abnormal appearance (Figure 4.16). Histology demonstrated that the corneal endothelium was occasionally multilayered. Approximately 1/3rd of affected individuals have undergone keratoplasty in at least one eye.

Given the similar phenotypes and overlapping genetic locus, it was hypothesised that CHED1 and PPCD1 are allelic conditions. Mutations in the *ZEB1* coding exons had been previously excluded in both families. The *ZEB1* qPCR 5' copy number assay was used to eliminate the possibility of a large *ZEB1* deletion using an affected individual from both F140 and F141 (Figure 4.10).

4.5.2 Genetic analysis of the CHED1 pedigree

The locus in the extended pedigree in F140 (Figure 4.13 A) had been previously refined from 2.7 Mb to 1.3 Mb by manually genotyping microsatellite markers and SNPs in eight affected individuals from different branches of the pedigree (individuals V:20, VI:2, VI:7, VI:17, VI:24, VII:3, VII:7, and VII:13). The disease interval spanned chr20.hg38: 17,641,482 – 18,949,130, which encompasses 46 annotated transcripts. WES analysis was undertaken for two distantly related affected individuals (VI:5, VII:13) and filtered for rare (MAF ≤ 0.005) heterozygous genome-wide variants that were shared between these two and absent from the unaffected individuals (section 2.6.2.2). This analysis revealed a single non-synonymous variant [c.1540A>C; (p.Ile514Leu)] in the *DZANK1* gene. The filtering was not restricted to the linked locus; however, this variant was present within the genetically linked interval on chr20p. The *DZANK1* c.1540A>C; (p.Ile514Leu) variant segregated with the disease in the family, but was present in ExAC with a frequency of 5/120650 alleles. Both SIFT and Polyphen2 predicted that the variant was not pathogenic with scores of 0.33 (Tolerated) and 0.008 (Benign), respectively, therefore the variant was considered likely to be a rare polymorphism and not disease causing in this family but present on the same causative haplotype as the unidentified causative variant. Additionally, a *C20ORF78*: c.181C>T variant was identified within a non-coding transcript with unknown functional consequences. ExomeDepth, a bioinformatic tool which infers copy number variations from WES data was unable to identify any potential disruptions within the genetic locus on chr20p. Furthermore, a dense chromosome 20 specific aCGH performed on individual VII:13 has previously confirmed the lack of copy number variations within the disease locus.

Given the lack of causative mutation in the WES data analysis, it was hypothesised that the causative variant may occur in a non-coding region within the genetic locus. To investigate this possibility, WGS was performed on two distantly related affected (VII:3, VI:24) and one unaffected individual (VI:22) (section 2.11). All variants within the refined genetic locus (chr20.hg38: 17,641,482 – 18,949,130) were filtered for rare (MAF ≤ 0.005) changes in a variety of external (1000G, ExAC, UK10K and Genomes of the Netherlands (GoNL)) and internal WGS (UCL-WGS) datasets

(section 2.11.2). Variants present in the unaffected individual and a further eight unrelated samples analysed by WGS at the same time as the CHED1 samples were removed. This approach resulted in the identification of 19 heterozygous variants, including the previously identified *DZANK1* missense variant (Table 4.4). Only two of the 19 variants were novel, one which was intergenic and one which was in the promoter of the *OVOL2* gene. Sanger sequencing confirmed the presence of an *OVOL2* c.-339_361dup (hg38.chr20:18,057,947 – 18,057,995dup) and segregation with disease (Figure 4.13).

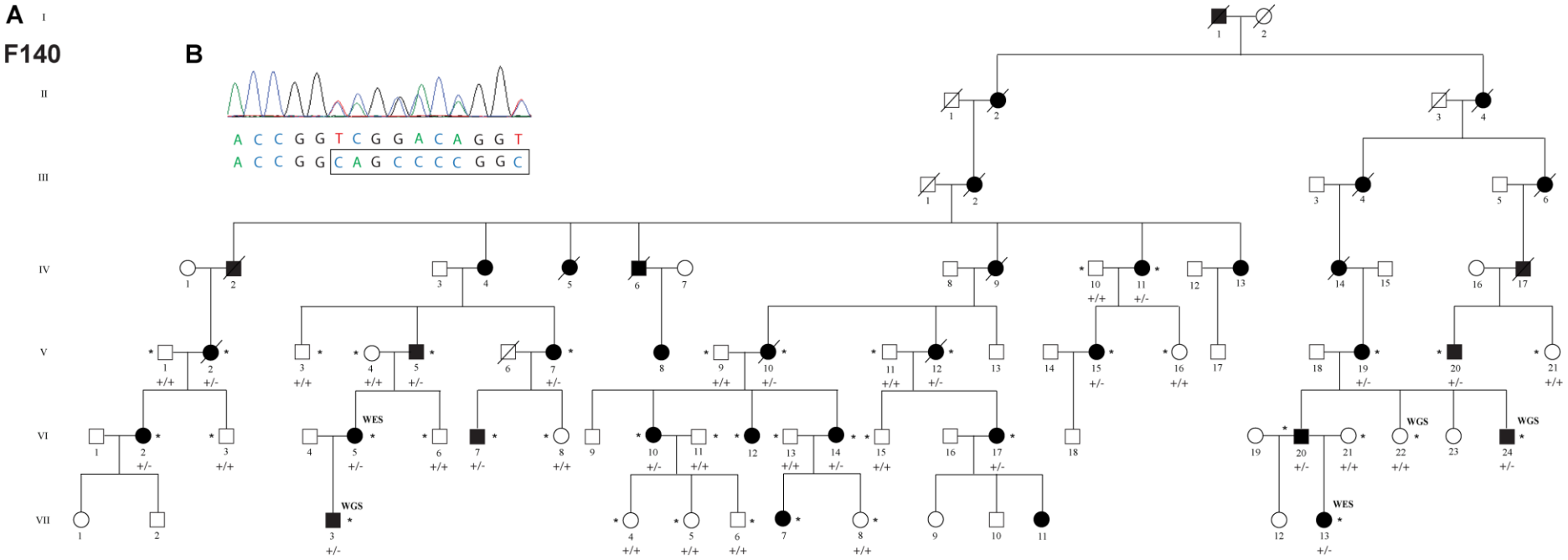


Figure 4.13 Pedigree of family #140 diagnosed with autosomal dominant congenital hereditary endothelial dystrophy (CHED1). (A) DNA was available for individuals highlighted with a *. Individuals analysed by whole exome sequencing are labelled with WES and whole genome sequencing with WGS. **(B)** Sanger sequencing trace showing presence of *OVOL2* mutation c.-339_361dup which segregates with disease in the family in individuals for which DNA was available.

Table 4.4 Rare (≤ 0.005) heterozygous variants within chr20.hg38: 17,641,482 – 18,949,130 in F140 family detected by whole genome sequencing (WGS). WGS was performed on two distantly related affected (VII:3, VI:24) and one unaffected individual (VI:22). Variants were filtered for rare (≤ 0.005) shared variants within the locus in affected individuals. Variants present in the unaffected individual were removed.

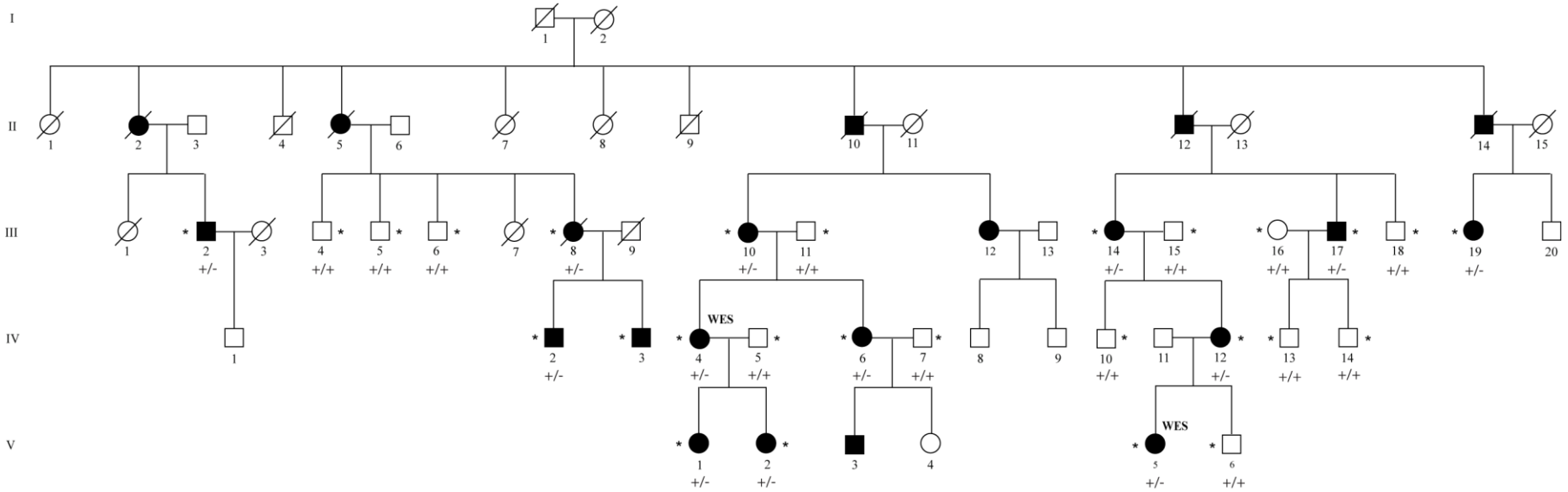
Co-ordinates (hg38)	Variant	Location	Closest transcript	dbSNP	Control Frequency				
					1000G	UK10K	GoNL	ExAC	UCL-WGS
18,057,973	A>ACCGGTTCCGGCGGCCGGGGCTG	Promoter	*OVOL2*	-	-	-	-	-	-
18,112,700	T>A	Intergenic	*PET117*	rs556855465	2/5008	4/7562	-	-	-
18,114,422	A>G	Intergenic	*PET117*	rs560139714	-	-	1/998	-	-
18,119,439	A>G	Intergenic	*PET117*	rs563340932	-	-	1/998	-	-
18,124,502	TAGA>T	Intergenic	*PET117*	-	-	-	-	-	-
18,189,275	CT>C	Intergenic	*CSRP2BP*	-	-	1/7562	-	-	-
18,189,278	C>T	Intergenic	*CSRP2BP*	rs772649261	-	-	-	-	-
18,240,797	G>A	Intergenic	*CSRP2BP*	rs552441504	5/5008	-	7/998	-	-
18,273,229	G>A	Intergenic	*ZNF133*	rs184537558	5/5008	-	-	-	-
18,292,765	C>T	Deep Intronic	*ZNF133*	rs542530373	-	-	1/998	-	-
18,319,195	A>G	Intergenic	*ZNF133*	rs530751423	-	-	1/998	-	-
18,379,703	G>A	Deep Intronic	*LINC00851*	rs558852368	2/5008	-	-	-	-
18,396,543	T>G	Missense	*DZANK1*	rs560809093	-	-	1/998	5/120,650	-
18,476,283	G>A	Deep Intronic	*POLR3F*	rs537806334	-	-	1/998	-	-
18,646,498	G>A	Deep Intronic	*DTD1*	rs532426738	-	-	1/998	-	-
18,784,373	C>A	Intergenic	*DTD1*	rs570664591	-	-	1/998	-	-
18,810,051	G>A	Non coding transcript	*C20ORF78*	rs543631581	-	-	1/998	-	-
18,818,189	T>C	Deep Intronic	*C20ORF78*	rs566681725	-	-	1/998	-	-
18,851,016	C>T	Intergenic	*C20ORF78*	rs550023958	-	-	1/998	-	-

4.5.3 Genetic analysis of Czech founder families

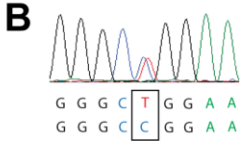
DNA samples were collected from sixteen Czech families which shared a common haplotype on chr20. In total, 75 DNA samples were available from affected individuals, 21 from unaffected first degree relatives and 15 from unaffected spouses. aCGH using probes spanning chr20 was negative in individuals III:19 (affected) and III:16 (unaffected spouse) from family F141.01. Furthermore, individuals IV:16, IV:17, V:9, V10, V:11 from F141.02 were analysed for copy number variations by SNP array, and were negative (section 2.7). In the Czech founder families (F141.01 – 16), data was provided from a pooled WES experiment performed using DNA from 10 individuals (F141.01: IV:4 and V:5, F141.02: IV:8 and V:9, F141.03: III:5, F141.06: IV:6, F141.09: IV:1, F141.11: III:1, F141.13: IV:6, F141.15: II:3) with PPCD spanning eight of the sixteen pedigrees (Figure 4.13 A). No rare (MAF $\leq$ 0.005) non-synonymous variants were shared among the affected individuals (section 2.6.2.2). In addition, data from targeted re-sequencing of the linked interval was performed on one affected (V:11) and one unaffected individual (V:10) from F141.02 (Figure 4.14 A). Due to relatively low coverage, V:11 was also analysed by WGS in an identical manner to the British CHED1 family F140. The combined targeted re-sequencing and WGS data were aligned and filtered using the same method described previously, resulting in the identification of 18 heterozygous variants in the linkage interval (section 2.11.2; Table 4.5).

The filtered variants identified by WGS in the CHED1 family (F140) and by targeted re-sequencing and WGS in the Czech founder families (F141) were compared for variants in the same functional regions. Ensembl VEP was used to predict disruption of potential regulatory regions by any of the candidate variants (section 2.11.3). In the CHED1 family, a heterozygous duplication was identified in the predicted promoter of *OVOL2* c.-339_361dup (chr20.hg38:18,057,947 – 18,057,995dup), which was absent from all WGS control databases interrogated. This mutation fully segregated with the disease in the pedigree. In the Czech combined re-sequencing and WGS data, a novel variant was also identified in the predicted promoter of *OVOL2* c.-370C>T (chr20.hg38:18,058,004A>G), which segregates with disease in all sixteen families (Figure 4.14 A and B).

A F141.01

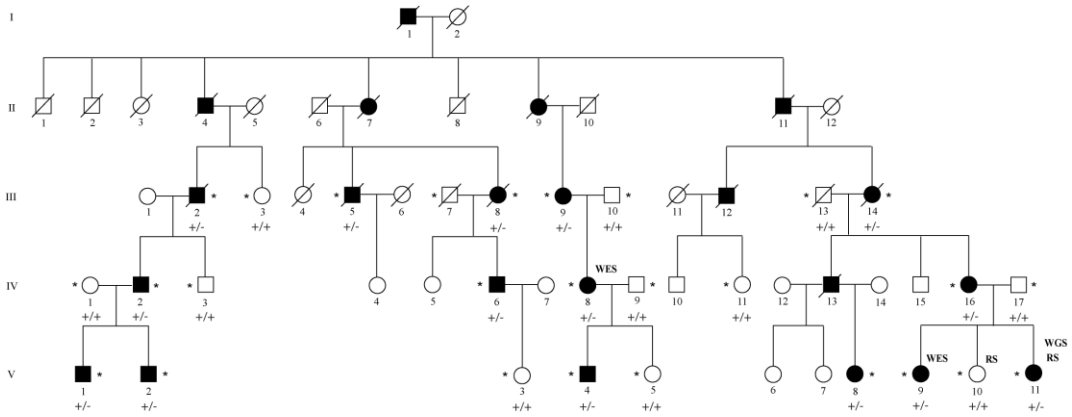


B

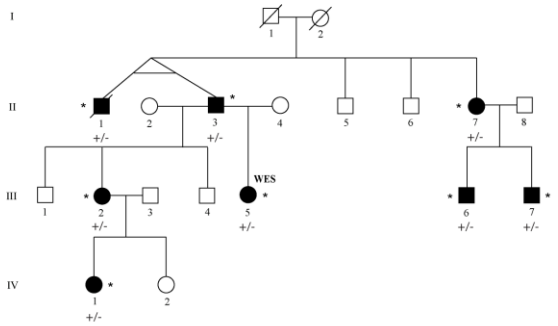


A (cont)

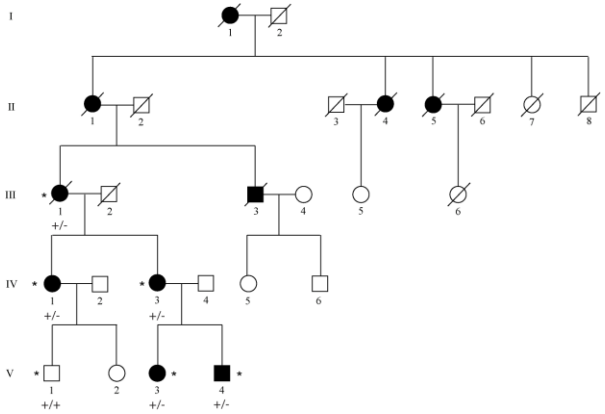
F141.02



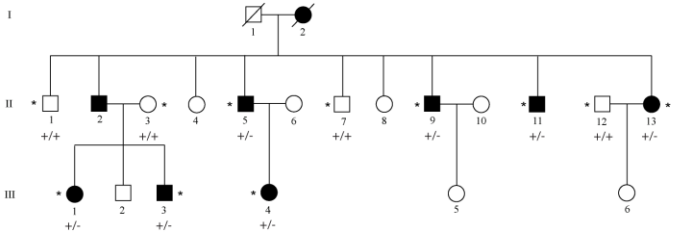
F141.03



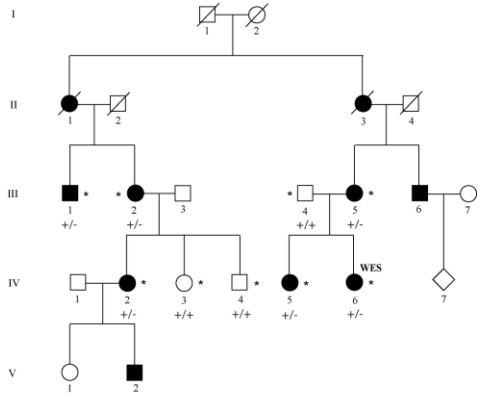
F141.04



F141.05



F141.06



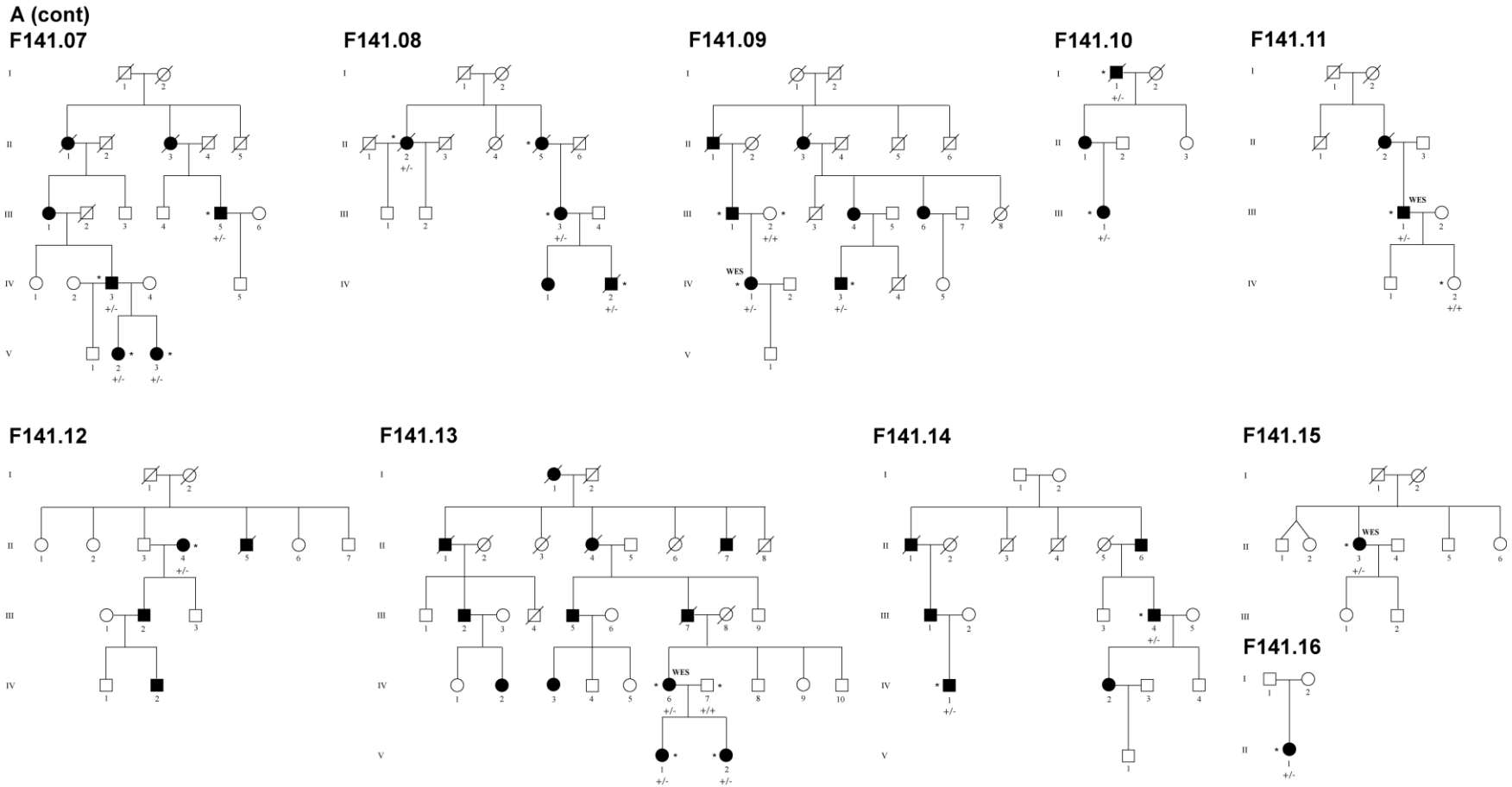


Figure 4.14 Sixteen Czech PPCD families (F141.01 – 16) share a common disease associated haplotype on chr20. (A) Pedigree structure of sixteen families showing autosomal dominant inheritance. DNA was available for individuals highlighted with a *. Individuals analysed by targeted resequencing, whole exome sequencing and whole genome sequencing are labelled with RS, WES and WGS, respectively. **(B)** Sanger sequence electropherogram showing *OVOL2* c.-370C>T mutation.

Table 4.5 Rare (≤ 0.005) heterozygous variants within chr20.hg38: 17,335,789 – 19,665,902 in a Czech PPCD1 individual analysed by targeted re-sequencing and whole genome sequencing (WGS). Targeted re-sequencing and WGS was performed on one affected (V:11), and re-sequencing on one unaffected individual (V:10) from F141.02. Variants were filtered for rare (≤ 0.005) shared variants within the locus. Variants present in the unaffected individual were removed.

Co-ordinates (hg38)	Variant	Location	Closest transcript	dbSNP	Control Frequency				
					1000G	UK10K	GoNL	ExAC	UCL-WGS
17,779,178	T>C	Intergenic	*BANF2*	-	-	-	-	-	-
17,962,621	T>TTCGGGGGAGGGGGG	Deep Intronic	*SNX5*	-	-	-	-	-	-
18,010,425	T>C	Intergenic	*OVOL2*	rs62206463	-	-	-	-	-
18,058,004	A>G	Promoter	*OVOL2*	-	-	-	-	-	-
18,373,822	G>A	Intergenic	*LINC00851*	-	-	-	-	-	-
18,504,053	T>G	Intergenic	*SEC23B*	-	-	-	-	-	-
18,800,653	T>C	Intergenic	*LOC100270804*	-	-	-	-	-	-
18,831,605	T>C	Intergenic	*C20ORF78*	-	-	-	-	-	-
18,836,917	A>G	Intergenic	*C20ORF78*	-	-	-	-	-	-
18,863,686	G>C	Intergenic	*C20ORF78*	-	-	-	-	-	-
18,870,484	C>G	Intergenic	*C20ORF78*	rs148906570	10/5008	-	-	-	-
18,937,206	C>G	Intergenic	*C20ORF78*	rs150426313	10/5008	60/7562	9/998	-	-
18,950,428	C>T	Intergenic	*C20ORF78*	-	-	-	-	-	-
19,140,530	G>A	Intergenic	*SLC24A3*	-	-	-	-	-	-
19,325,002	A>G	Deep Intronic	*SLC24A3*	-	-	-	-	-	-
19,331,987	C>T	Deep Intronic	*SLC24A3*	-	-	-	-	-	-
19,601,355	C>A	Deep Intronic	*SLC24A3*	rs537549121	9/5008	2/7562	-	-	-
19,637,683	A>G	Deep Intronic	*SLC24A3*	-	-	-	-	-	-

4.5.4 OVOL2 screening in unsolved PPCD cohort

A ~ 1.8 kb genomic region containing the proximal promoter of *OVOL2* was sequenced in an unsolved PPCD cohort, in which *ZEB1* mutations had been previously excluded including probands of British, Czech and New Zealand origin. Two further novel variants were identified in the *OVOL2* proximal promoter in probands from MEH.

The proband of family #142 had corneal opacity from childhood. When clinically examined at MEH her phenotype was not reportable due to previous bilateral multiple corneal grafts, glaucoma drainage surgery and enucleation of her left eye. She had, throughout her medical history, been diagnosed with both PPCD and anterior segment dysgenesis. The proband described a family history of corneal disease consistent with autosomal dominant pattern of inheritance; however, no relatives were available for clinical examination or able to provide a DNA sample (Figure 4.15 A). Mutations in *ZEB1* were excluded by direct sequencing (section 2.2; Appendix A). Direct screening of a ~1.8 kb fragment of the *OVOL2* promoter in the proband revealed a heterozygous c.-274T>G mutation (chr20.hg38:18,057,908A>C), absent from 209 ethnically matched controls that were also screened by targeted Sanger sequencing, 1000 Genomes and UCL WGS datasets (Figure 4.15 B) (section 2.2; Appendix A).

Family #143 consisted of a family with at least 5 generations of affected individuals diagnosed with PPCD. Mutations in *ZEB1* were excluded in the proband (III:3) (section 2.2; Appendix A). The clinical phenotype of examined individuals III:3 and IV:2 was mild with no requirement for surgical grafting in either and no signs of glaucoma. Both had bilateral opacities at Descemet's membrane and peripheral adhesions between the iris and cornea causing pupillary distortion (Figure 4.16 E). Screening the ~ 1.8 kb promoter of *OVOL2* revealed the presence of a c.307T>C variant (chr20.hg38: 18,057,941A>G) in the proband (Figure 4.15 D) (section 2.2; Appendix A). Segregation analysis revealed the presence of the mutation in her affected daughter (IV:2) but not her unaffected sister (III:1). Three grandchildren of the proband provided DNA by saliva sample but were not seen in clinic; one of the proband's three undiagnosed grandchildren (IV:3) was found to harbour the *OVOL2*

promoter mutation; therefore this individual was referred to MEH for clinical assessment (Figure 4.15 C).

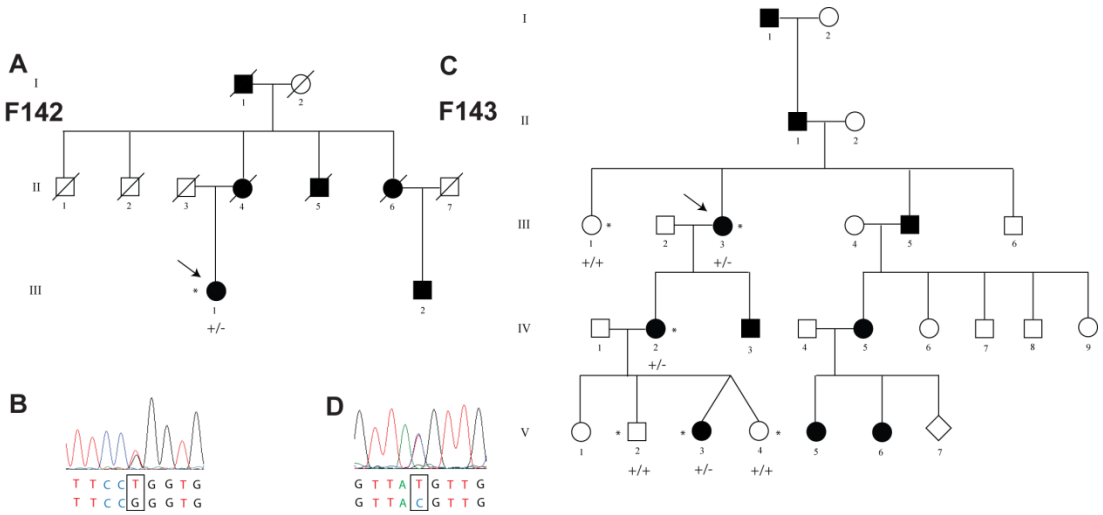


Figure 4.15 *OVOL2* mutations identified by direct screening of a genetically unsolved posterior polymorphous corneal dystrophy cohort. (A) Family #142 show autosomal dominant inheritance in at least three generations. No other familial DNA samples could be obtained for segregation analysis. **(B)** Sequence electropherogram of *OVOL2* c.-274T>G mutation identified in F142. **(C)** Family 143 show autosomal dominant inheritance in at least five generations. The *OVOL2* mutation segregated with disease in the family but was also present in asymptomatic V:3 who is a child and has not been clinically examined. **(D)** Sequence eletropherogram of *OVOL2* c.-274T>G mutation identified in family #143.

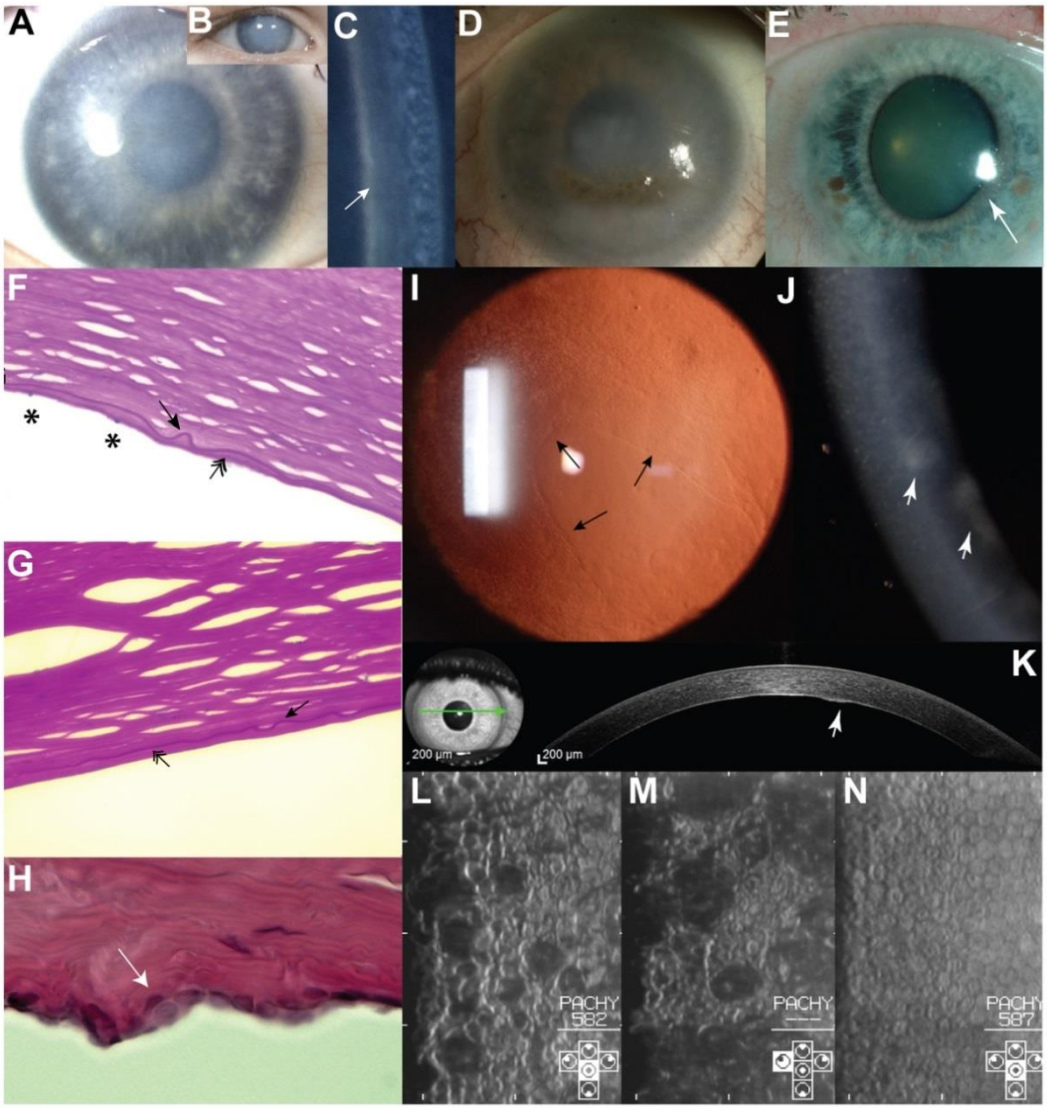


Figure 4.16 Clinical spectrum of disease associated with *OVOL2* mutations. (A-C) Individual VII: 13 from F140 **(D)** Unoperated eye in VI:2 from F140 aged 52 years old showing secondary lipoidal degeneration. **(E)** Individual III:3 from F143 showing distortion of iris (arrow) due to iridocorneal adhesion. **(F)** Histology from individual F140: VII:13 at 6 years old. **(G)** Histology from F140: VII:7 aged eleven years. Thin, irregular Descemet membrane shown with arrow, reduced endothelial cell count indicated with asterisks. Abnormal material accumulation which may be retrocorneal fibrosis located posterior to Descemet membrane is shown with double headed arrow. **(H)** Histology from a F141.11: III:1 aged 42 years demonstrating focal multi-layering (arrow) and undulation of the posterior corneal surface. **(I)** Retroillumination of cornea in Czech F141.16: II:1 aged 29 years old. Arrows indicate boundary of normal endothelial cells surrounded by abnormal cells. **(J)** Narrow beam section from F141.05: III:5 (aged 33) showing corneal haze and focal areas (arrowhead) **(K)** Ocular coherence tomography image from the same individual showing raised lesion on poster corneal surface (arrowhead). **(L, M)** Specular microscopy from F141.02: V:11 (aged 13 years) showing irregular shape and size of endothelial cells. **(N)** Specular microscopy of an unaffected sibling F141.02: V10 (aged 25), for comparison.

Table 4.6 Summary of pathogenic mutations in *OVOL2* in the posterior polymorphous corneal dystrophy cohort. Mutations were checked for their frequency in Kaviar, which includes data from ~13,000 whole genomes. NP: not present.

Family ID	Mutation	Kaviar frequency
F140	c.-339_361dup	NP
F141	c.-370C>T	NP
F142	c.-274T>G	NP
F143	c.-307T>C	NP

4.5.5 Expression of OVOL2 in the cornea

To determine the expression pattern of *OVOL2* in the cornea, RNA was extracted from both tissue and cultured cells (section 2.13.1). RNA was extracted from whole corneal buttons donated after enucleation surgery for posterior segment melanoma. Corneal endothelial tissue was obtained from individuals undergoing DMEK for FECD. Primary corneal epithelial cells were expanded from limbal epithelial cells extracted from deceased donor corneoscleral rims and the corneal limbal epithelial cell line HCE-S was obtained from Professor Julie Daniels (section 2.12.2; 2.12.3). Stromal fibroblasts were cultured from corneal buttons obtained from donors without a corneal dystrophy and cultured in serum-containing media (section 2.12.1). It is important to note that the presence of serum induces quiescent stromal keratocytes into proliferative fibroblasts (section 1.4.3). RNA from cultured primary corneal endothelial cells was obtained from Dr. Alice Davidson.

OVOL2 encodes a transcription factor that induces mesenchymal-to-epithelial transition (MET). Importantly, it is a direct transcriptional repressor of established PPCD3 gene, *ZEB1*, and is itself negatively regulated by *ZEB1* in a negative feedback loop (Hong *et al.*, 2015). RT-PCR demonstrated expression of *OVOL2* in the full thickness cornea, however we were unable to detect expression in adult human endothelial tissue or expanded adult human corneal cells (Figure 4.17 A) (section 2.13). The absence of *OVOL2* expression in the corneal endothelium was confirmed by interrogation of publicly available RNA-seq data obtained from three adult human corneal endothelial tissue samples and two foetal (16 – 18 weeks of

gestation) samples (Chen *et al.*, 2013) (Figure 4.17 B; section 2.14.1). In the remaining corneal layers, *OVOL2* was absent from cultured stromal fibroblasts, but was present in cultured epithelial cells expanded from adult human limbal tissue and the HCE-S cell line (section 2.13). This data indicates that the source of the *OVOL2* expression in the full thickness cornea is the epithelial layer. This is consistent with expression of *OVOL2* in other epithelial tissues including the skin, kidney and germline epithelium of the testis.

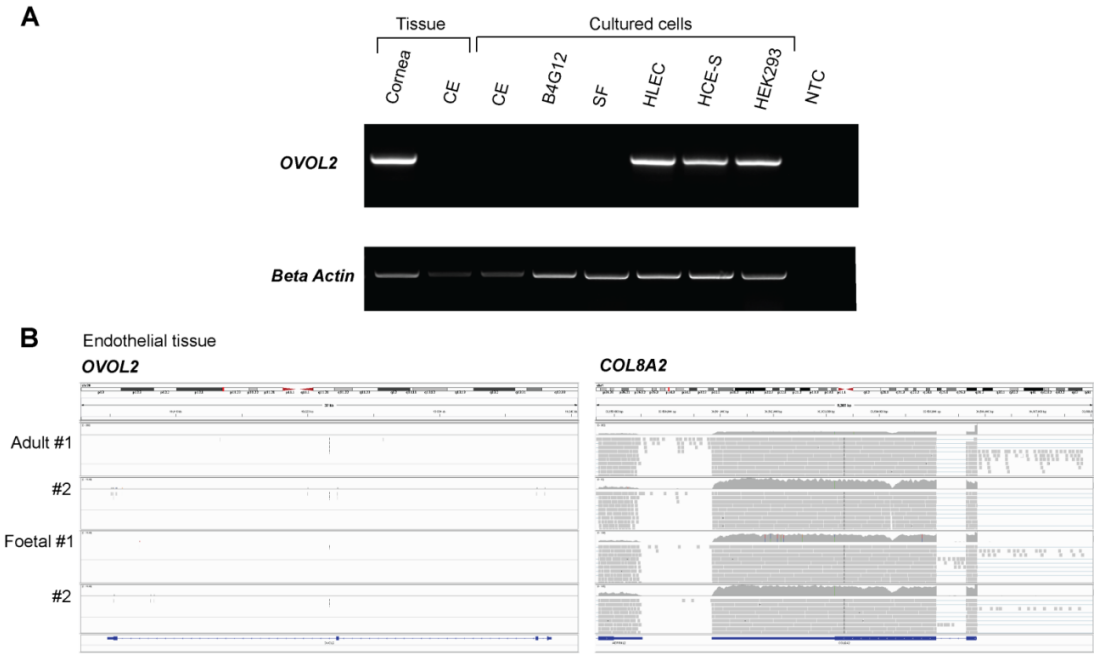


Figure 4.17 *OVOL2* is not expressed in normal corneal endothelial tissue or cultured cells. (A) RT-PCR expression of *OVOL2* in corneal tissue and cultured corneal cells. **(B)** Interrogation of corneal endothelial RNA-seq data showing lack of transcripts mapping to *OVOL2*. For comparison, multiple reads map to *COL8A2*, which is highly expressed in the corneal endothelium and is the genetic cause of early-onset Fuchs endothelial corneal dystrophy.

4.5.6 Transcriptional regulation of OVOL2 expression

The ENCODE project aims to catalogue the function of regulatory DNA elements in the human genome including cis-acting elements such as promoters and long acting regulatory elements such as enhancers, silencers and insulators (The ENCODE Project Consortium, 2011). The project uses a range of techniques including chromatin immunoprecipitation and DNA footprinting to identify active regulatory

elements in various cell types. ENCODE transcription factor binding data was interrogated for the *OVOL2* promoter region containing all four identified mutations. All were found to occur within a known binding site of multiple transcription factors including FOXA1, Nrf1, p300, CtBP2, and SP1 (Figure 4.18 A). Alignment of orthologous nucleotide sequences revealed that all four *OVOL2* mutations are conserved in mammals, including chimpanzees and mice, but not all are conserved in chicken, lizards or amphibians. However, for mutations that were single base pair changes, the substituted base pair was not observed in orthologous sequences (Figure 4.18 B).

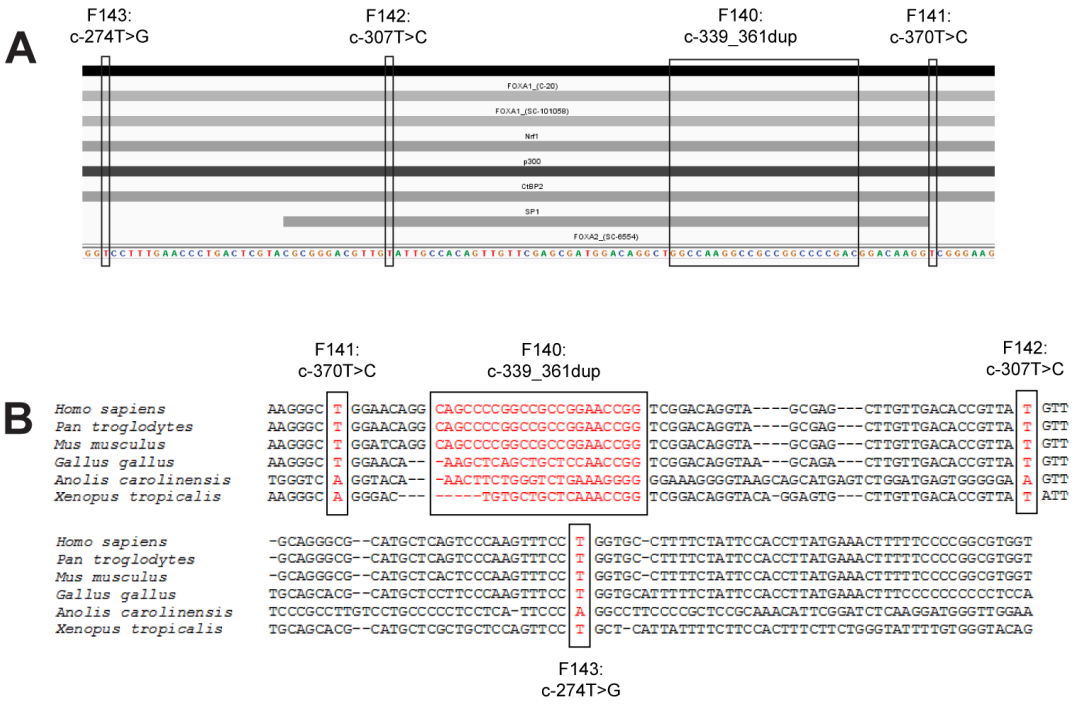


Figure 4.18 Predicted sites of transcription factor binding and location of *OVOL2* promoter mutations. (A) Mutations occur upstream of *OVOL2* and are located in regions that are targeted by binding of multiple transcription factors. **(B)** Conservation of orthologous nucleotide sequences with mutations highlighted in red.

Transcription factor prediction programs AliBaba2.1 and MatInspector were used to predict the consequences of the four mutations identified (section 2.3.6). Both c.-339_361dup and c.-370T>C were predicted to introduce additional transcription factor binding sites, whereas c.-274T>G and c.-307T>C were predicted to both gain additional transcription factor binding sites and loss of some transcription factor

binding sites. Some of these are expressed in either adult or foetal endothelial tissue (Table 4.7).

The hypothesis that the *OVOL2* mutations cause inappropriate ectopic expression was supported additionally by *in vitro* luciferase assay data performed by Dr. Alice Davidson which demonstrated increased expression of a reporter gene when driven by the *OVOL2* promoter containing any of the identified mutations, in comparison to the wild type promoter activity.

Table 4.7 In silico prediction of transcription factor protein binding altered by the presence of *OVOL2* mutations. AliBaba 2.1 is a program for predicting transcription factor binding sites in an unknown DNA sequence utilising binding sites collected in TRANSFAC public database. Mat Inspector is a software tool that utilizes a large library of matrix descriptions for transcription factor binding sites to locate matches in DNA sequences. *RKPM values also encompass flanking transcripts. All four mutations were predicted to alter transcription factor binding.

Transcription factor	F140 c.-339_361dup	F141 c.-370T>C	F142 c.-274T>G	F143 c.-307T>C	Expression in endothelium (RKPM)	
					Adult	Foetal
ELK1	Site gained[2]	Site gained[1,2]	Site gained[2]		14.01	17.42
GRHL1	Site gained[2]				0.98	0.04
SLC2A4RG	Site gained[2]	Site gained[2]			11.36*	20.46*
FLI1	Site gained[2]				7.85	1.79
ZNF239		Site gained[2]			2.18	4.30
REL		Site gained[2]			2.17	1.00
ZNF143			Site lost[2]		12.06	7.54
DMRTA2				Site gained[2]	0.65	0.02
RABL6				Site gained[2]	14.98*	10.10*
RFX3				Site gained[2]	2.39	5.07
T				Site lost[2]	0.00	0.01
SP1	Site gained[1]	Site gained[1]			12.52	17.53
EGR1	Site gained[1]				134.09	84.08
ATF2				Site gained[1]	11.78	13.26
SRF			Site gained[1]		21.42	16.48
ETS1			Site lost[1]		10.92	9.01

[1]AliBaba 2.1 [2]Mat Inspector

4.6 PPCD4 locus

4.6.1 Exclusion of known PPCD genes in a large unsolved Czech PPCD pedigree

The majority of unsolved families in the MEH cohort were sporadic cases or small nuclear families, for which linkage analysis would not be informative. However, a large pedigree (F144) was identified by Dr. Petra Liskova (Figure 4.19 A). The majority of affected individuals clinically examined had typical signs of PPCD including an irregular posterior surface with vesicular lesions and geographical opacities (Figure 4.22 C – E; H – K). 25.9 % of (7/27) individuals required corneal grafting, of these 3 were grafted bilaterally. Glaucoma was also diagnosed in 25.9 % individuals (7/27), of which two developed glaucoma after corneal grafting. Corectopia was noted in 4 eyes of 3 individuals and was associated with secondary glaucoma in all cases (Figure 4.22 F). Screening the *ZEB1* coding exons revealed no mutations, and inclusion of the proband (F144: VII:8) in the *ZEB1* qPCR copy number assay had previously indicated that there was no deletion, at least at the 5' end of the gene (section 4.4.3). Analysis of copy number variation using DNA from VII:10 on an Illumina HumanOmniExpress BeadChip SNP array was also negative (section 2.7).

Eight affected (V:18, VI:7, VI:13, VII:6, VII:8, VII:10, VIII:1, VIII:5) and three unaffected individuals (VI:14, VI:15, VII:9) from the large Czech PPCD pedigree were analysed by WES (section 2.6). No rare variants (MAF $\leq$ 0.005) were identified which were shared between all affected individuals that were absent in unaffected individuals. In a single branch of the pedigree, a heterozygous nonsense variant in *MYOC* [c.1102C>T; p.(Gln368*)] was identified in 5/8 affected individuals (section 2.6.2.2). This variant has been reported as a heterozygous pathogenic mutation that causes juvenile onset primary open angle glaucoma (Stone *et al.*, 1997). However, this variant has a frequency of 0.001 (132/121392) in ExAC, arguing against a pathogenic role. No clinical differences were apparent between individuals harbouring the *MYOC* variant in comparison to those affected individuals without the *MYOC* variant.

4.6.2 Linkage analysis

For the linkage study, a large branch of the family for which samples were available was analysed. Nine affected (VI:2, VI:4, VII:1, VIII:1, VIII:3, VII:7, IX:1, IX:3, IX:6) and seven (VII:2, VII:3, VIII:2, VIII:4, IX:2, IX:4, IX:5) unaffected samples were genotyped using an Illumina 2.5-Exome8 array (section 2.10). Linkage analysis was performed by Dr. Nikolas Pontikos using MERLIN. Only one locus resulted in a significant LOD score, spanning chr8.hg38:100,821,039 – 119,725,923 had a LOD score of greater than 3, with a maximum LOD of 3.17 (Figure 4.20 A).

4.6.3 Whole genome sequencing

Four distantly affected individuals (VI:7, VI:9, VII:8, VIII:1) from F144 were analysed by WGS (Figure 4.19; section 2.11). Variants located between chr8.hg38:100,821,039 – 119,725,923 were extracted from the data and filtered for novel variants which were shared between all four affected individuals absent from control database Kaviar and internal UCL controls. Variant effect predictor (VEP) identified a single variant c.20+544G>T (chr8.hg38: 101,493,333G>T) located in intron one of *GRHL2*, which is predicted to be a putative promoter region for this gene and segregated with disease in the family (Figure 4.20 B; section 2.11.3). Interestingly, this variant also lies upstream of antisense transcript KB-1562D12.1 on the opposite strand to *GRHL2*. A second variant was identified in *GRHL2* intron 11, while the remaining variants were intergenic (Table 4.8).

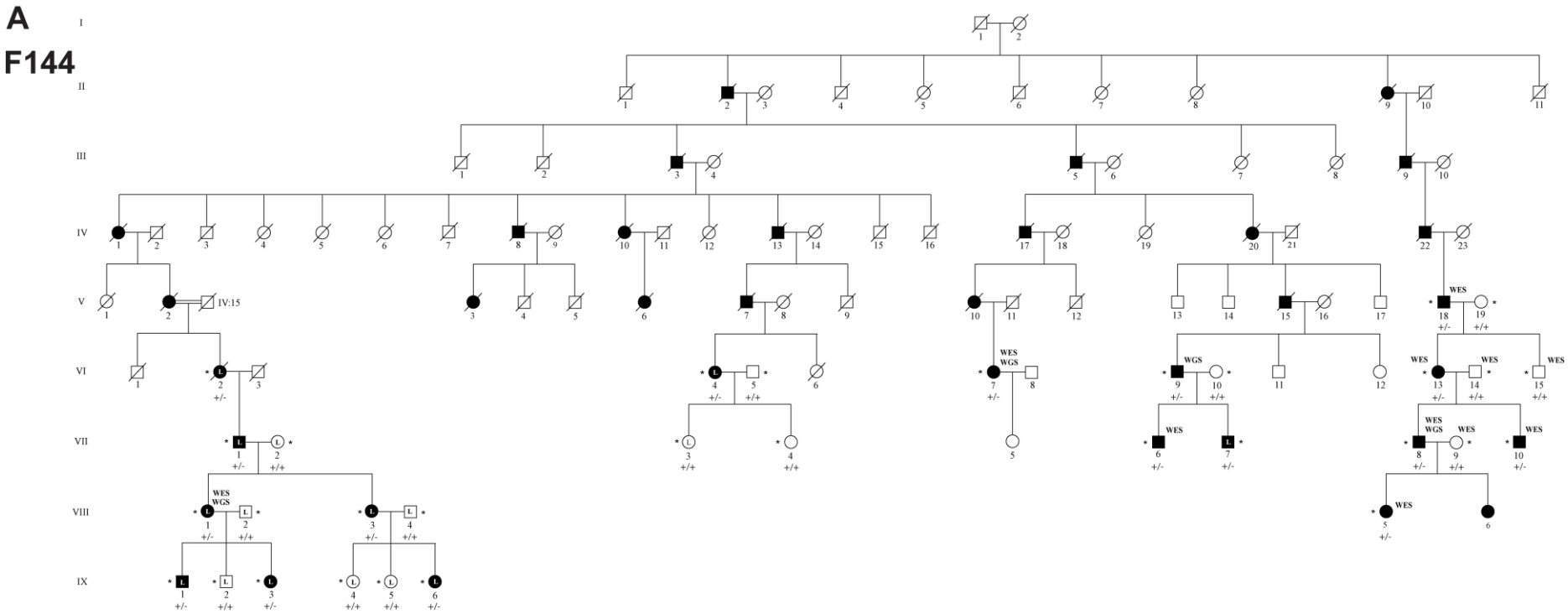


Figure 4.19 Pedigree of Czech family #144 with posterior polymorphous corneal dystrophy. DNA was available for individuals highlighted with a *. Individuals analysed by whole exome sequencing are labelled with WES and for whole genome sequencing, WGS. SNP genotyping and linkage analysis was performed on individuals labelled with 'L'.

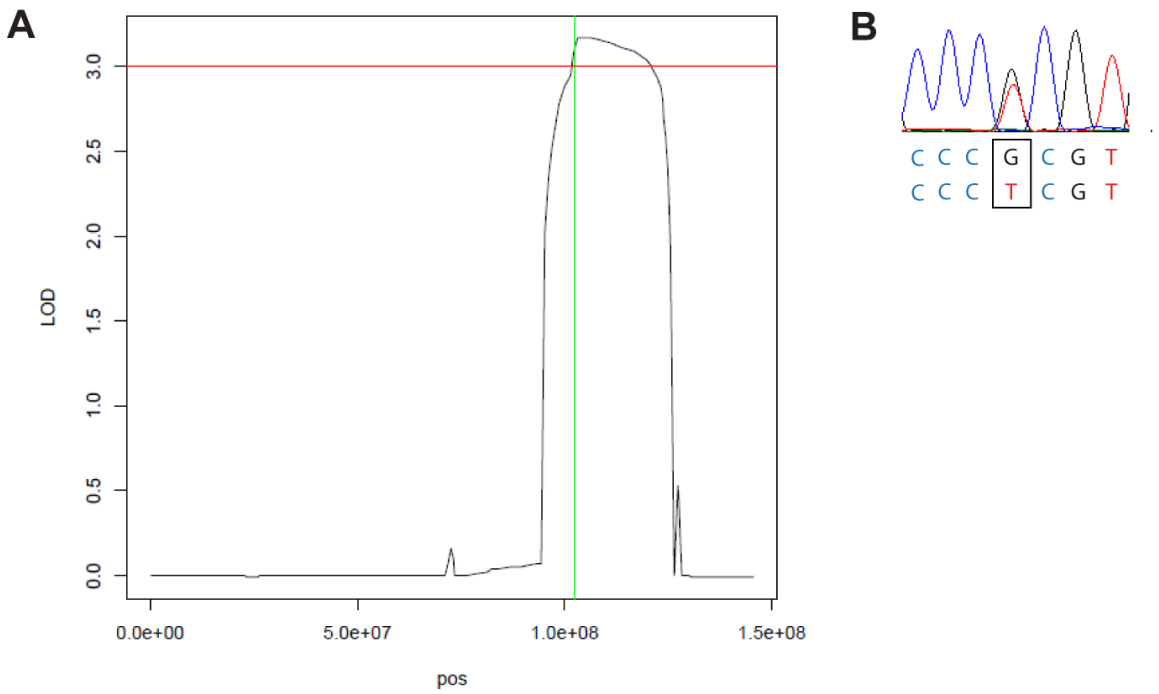


Figure 4.20 Linkage analysis in F144 reveals a significant LOD score on chr8q. (A) Plot of LOD scores shows a LOD score of > 3.0 spanning chr8.hg38:100,821,039 – 119,725,923. Line indicates location of a novel non-coding variant *GRHL2:* c.20+544G>T, identified by whole genome sequencing. **(B)** Sequence electropherogram confirming presence of *GRHL2:* c.20+544G>T variant.

Table 4.8 Novel shared variants within chr8.hg38:100,821,039 – 119,725,923 in affected individuals VI:7, VI:9, VII:8, VIII:1 from F144, analysed by whole genome sequencing. Six shared novel variants were identified within the mapped locus. Two variants occurred within the *GRHL2* gene (intronic) whereas the remaining three were intergenic.

Co-ordinates (hg38)	Variant	Location	Gene	dbSNP	Control Frequency				
					1000G	UK10K	GoNL	ExAC	UCL-WGS
chr8: 101,109,636	C>T	Intergenic	-	-	-	-	-	-	-
chr8: 101,493,333	G>T	Intronic	*GRHL2*	-	-	-	-	-	-
chr8: 101,636,724	T>TACACAC	Intronic	*GRHL2*	-	-	-	-	-	-
chr8: 109,214,441	CTT>C	Intergenic	-	-	-	-	-	-	-
chr8: 109,214,443	T>TAAAACA	Intergenic	-	-	-	-	-	-	-
chr8: 114,635,792	A>T	Intergenic	-	-	-	-	-	-	-

4.6.4 Screening the unsolved PPCD cohort for GRHL2

Given the identification of a putative causative change in the first intron of *GRHL2*, a ~ 1.2 kb region encompassing the 5' UTR, exon 1 and the beginning of intron 1 of *GRHL2* was amplified and screened in unsolved PPCD patients from MEH, the Czech Republic and New Zealand (section 2.2; Appendix A).

Probands from families F145 and F146 (of Czech origin) were previously identified as negative for *ZEB1* coding mutations and *ZEB1* deletions by Sanger sequencing, SNP genotyping array and qPCR. WES analysis was performed, but filtering for rare (MAF $\leq$ 0.005) non-synonymous or splice site variants; however, no potentially pathogenic variants in any known endothelial dystrophy gene was identified (section 2.6.2.3). Direct Sanger sequencing of a ~ 1.2 kb region of the *GRHL2* in both F145 and F146 identified the same variant (*GRHL2*: c.20+544G>T) that was originally detected in the large pedigree F144 (Figure 4.19; Figure 4.20 B) (section 2.2; Appendix A). Neither proband was knowingly related to the original pedigree, or to each other. A fourth Czech family (F147) was also found to harbour the *GRHL2*: c.20+544G>T variant. Interestingly, in this family segregation analysis demonstrated that the variant was absent from both parents. Paternity testing (performed by Dr. Petra Liskova) confirmed the identity of the proband's biological father, therefore the variant appears to have occurred *de novo* in the proband (Figure 4.21 C).

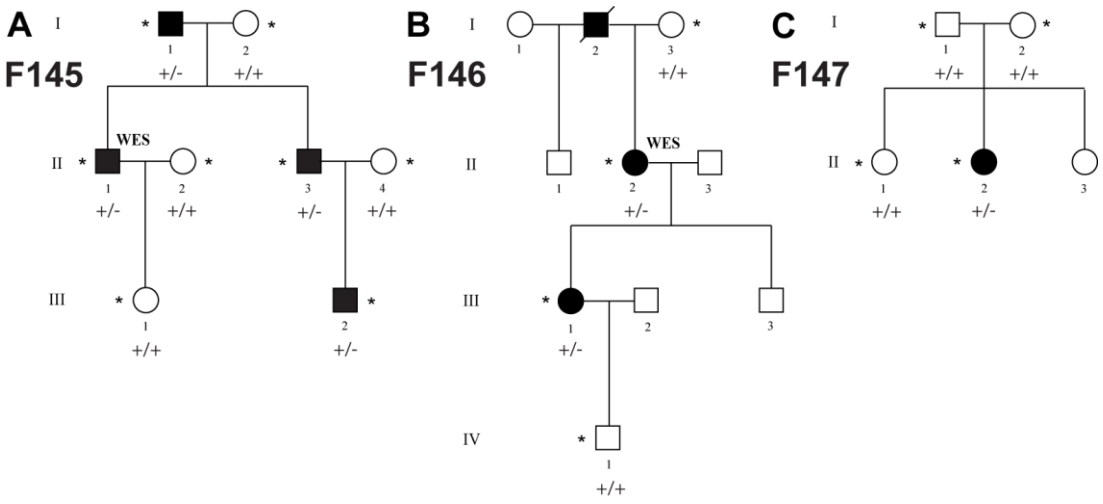


Figure 4.21 *GRHL2* c.20+544G>T variant segregates in three additional Czech families diagnosed with posterior polymorphous corneal dystrophy. +/- indicates that the individual is heterozygous for the variant and +/+ is wild type. Segregation analysis was performed for **(A)** F145 **(B)** F146 **(C)** F147 with *de novo* occurrence of the c.20+544G>T variant.

A white British proband, F148: II:3, was recruited from MEH. He had typical bilateral signs of PPCD which were more severe in his left eye and he also had amblyopia of his right eye. He underwent DSAEK in his left eye in 2014, resulting in visual improvement. Screening for mutations in *ZEB1*, *COL8A2* and the *OVOL2* promoter was negative (section 2.2; Appendix A). Screening a ~ 1.2 kb region of the first intron of *GRHL2* identified a deletion, c.20+32_20+51del (chr8.hg38: 101,492,821_101,492,840del) located 31 bp from the first exon (Figure 4.22 B). This variant was absent in the Kaviar database. There was no known family history of PPCD or other corneal disease; however, family members were clinically examined following identification of a possible causative mutation (Figure 4.22 A). The proband's nephew, III:2, was found to have changes of the corneal endothelium consistent with band-type PPCD. DNA was obtained from saliva kits from the proband's niece (III:3) and her two sons (IV:3, IV:4). The *GRHL2* variant was only identified in III:3 and was absent in III:2. She had not been diagnosed with PPCD and had no corneal abnormalities when examined by an optometrist (aged 81). In conclusion, if the *GRHL2* intronic deletion is contributing to the phenotype in this family, then it is partially penetrant.

F149 was a Caucasian proband from MEH with an atypical PPCD phenotype (Figure 4.23 A). There was no family history of PPCD; however, the proband's brother had

previously had his right eye enucleated due to a tumour (Figure 4.22 C). Mutations in the *ZEB1* coding exon and *OVOL2* promoter were excluded (section 2.2; Appendix A). Screening the ~ 1.8 kb *GRHL2* region identified a single nucleotide deletion c.20+257delT (chr8.hg38: 101,493,046delT; Figure 4.22 D) that was absent from the Kaviar database.

F150 was a female proband of Bangladeshi origin with no known family history of PPCD (Figure 4.22 E). The endothelium was abnormal with the presence of endothelial vesicles restricted to her left eye (Figure 4.23 B). She also had left amblyopia and a decompensated left exotropia. Intraocular pressures were normal in both eyes and there were no iris abnormalities. The endothelial cell density was reduced in her affected eye with 871 cells/mm^2 compared to 3165 cells/mm^2 in her right eye. Mutations in the *ZEB1* coding exon and *OVOL2* promoter were excluded (section 2.2; Appendix A). A single nucleotide deletion within the first intron of *GRHL2* was identified; c.20+133delA (chr8.hg38: 101,492,922delA; Figure 4.22 F) that was absent from the Kaviar database.

All identified *GRHL2* variants in PPCD patients were also absent from 1000G Project, the UCL-WGS bipolar WGS cohort (~100 individuals, 200 alleles) and Kaviar variant database (Table 4.9)

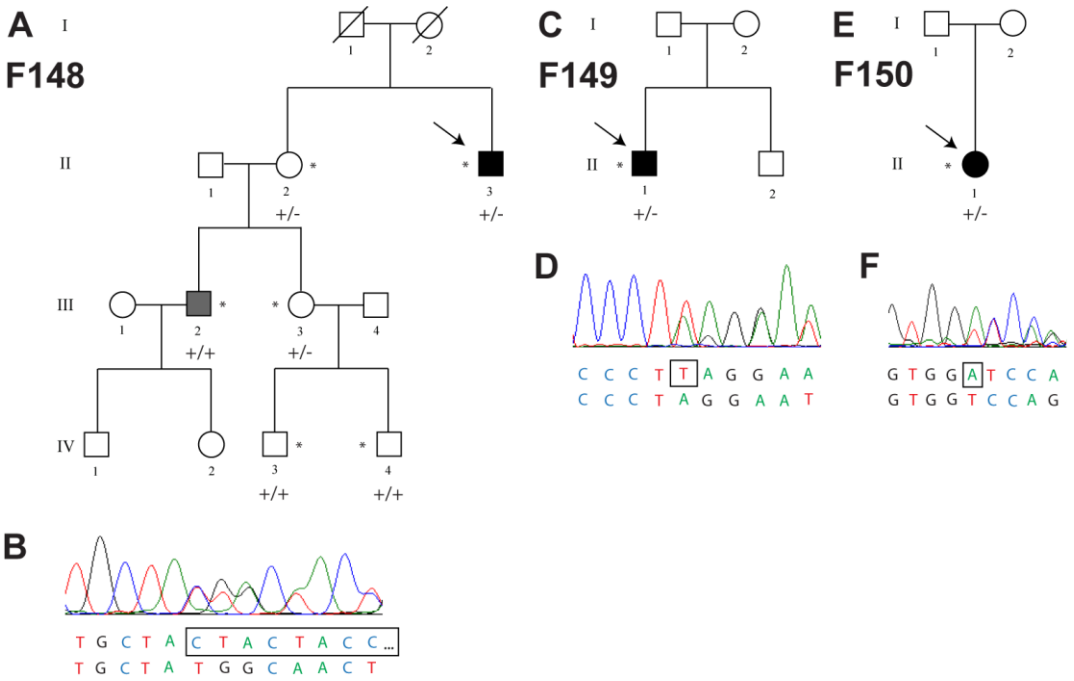


Figure 4.22 Three MEH families with posterior polymorphous corneal dystrophy harbouring variants in the first intron of *GRHL2*. DNA samples were obtained from those with a *, +/- indicates that the individual is heterozygous for the variant, +/+ indicates wild type. **(A)** Pedigree of family #148 shows lack of segregation of variant with corneal changes. Black filled shape indicates affected with typical PPCD, dark grey indicates band-type PPCD. **(B)** Sequence electropherogram showing *GRHL2:* c.20+32_20+51del variant. **(C)** Pedigree of family #149. **(D)** Sequence electropherogram showing *GRHL2:* c.20+257delT variant. **(E)** Pedigree of family #150. **(F)** Sequence electropherogram showing *GRHL2*: c.20+133delA variant.

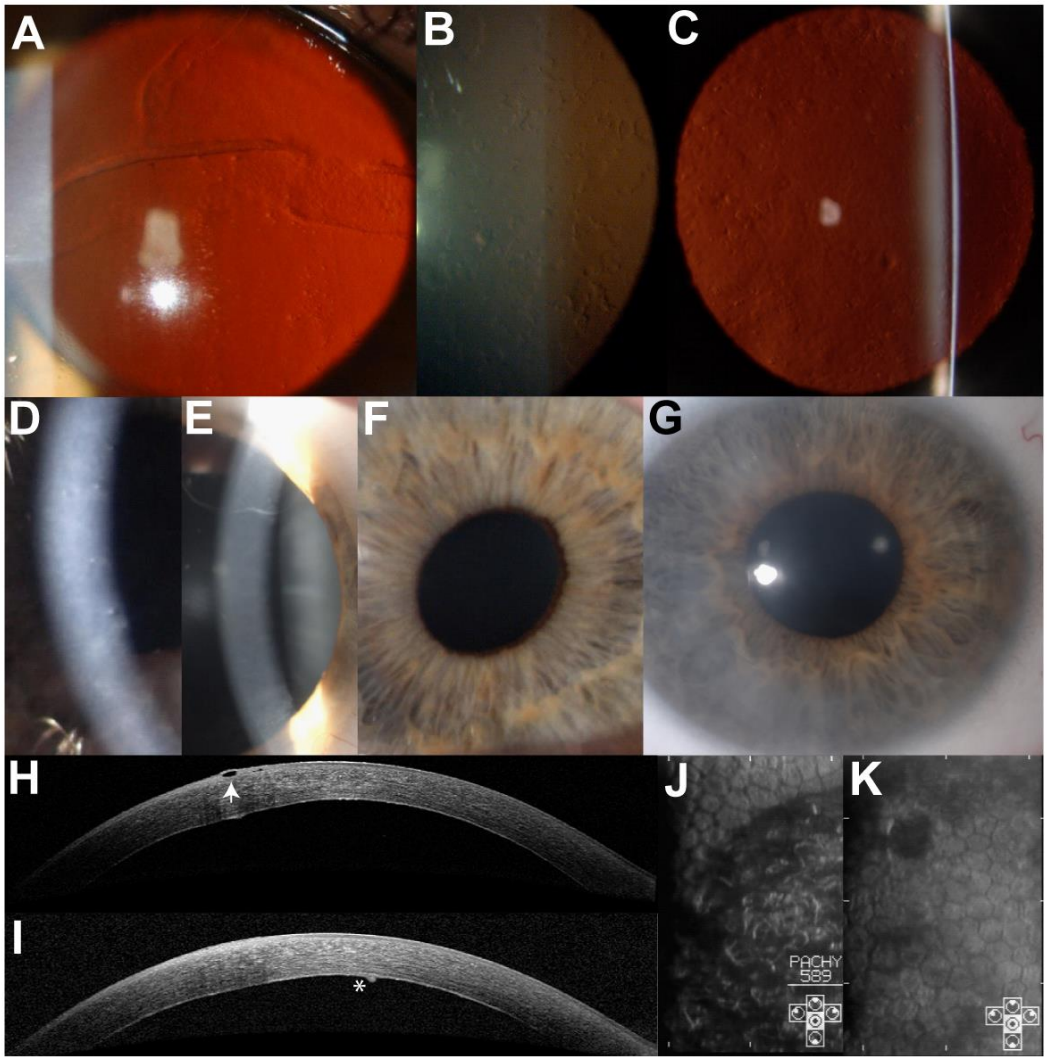


Figure 4.23 Spectrum of clinical phenotypes in individuals harbouring *GRHL2* variants. (A-C) Affected corneas in retroillumination: **(A)** Proband of F149 (55 years old). **(B)** Proband of F150 (30 years old) showing presence of vesicles. **(C)** F144: VI:9 (aged 69 years old) showing presence of vesicles. **(D)** Vesicles in direct illumination in F144: VII:7 (23 years old). **(E)** A prominent peripheral line of opacification on direct illumination with broader slit beam in F146: II:3 (46 years old). **(F)** Corectopia in the right eye of F144: VII:6 (29 years old). **(G)** Diffuse corneal opacification causing a hazy appearance in individual III:2 from F146. **(H - I)** Ocular coherence tomography cross section of the cornea of individual F144: VI:9 (58 years old) showing paracentral corneal oedema manifesting as epithelial bullae and increased stroma thickness (arrow), increased reflexivity of the posterior corneal layers with focal undulation and protrusion into the anterior chamber (asterisk). **(J - K)** Corneal endothelial specular images of left and right eye in individual F144: VIII:3 (37 years old) showing differences in endothelial cell morphology and size. Dark areas correspond to regions with cell multilayering and abberant formation of irregular posterior layer.

Table 4.9 Summary of variants in *GRHL2* in the posterior polymorphous corneal dystrophy cohort. Mutations were checked for their frequency in Kaviar, which includes data from ~13,000 whole genomes. NP: not present.

Family ID	*GRHL2* variant	Kaviar frequency
F144, F145, F146, F147	c.20+544G>T	N/A
F148	c.20+32_20+51del	N/A
F149	c.20+257delT	N/A
F150	c.20+133delA	N/A

4.6.5 Screening a Czech control cohort

Due to the occurrence of multiple families from the Czech Republic carrying the *GRHL2*: c.20+544G>T variant, one possible explanation was that the variant was a rare benign polymorphism present in the Czech population. Given the lack of representation of the Czech population in online control databases, we examined the frequency of this variant in a control Czech cohort by direct sequencing of a ~ 700 bp amplimer encompassing all previously identified *GRHL2* variants (section 2.2; Appendix A). In total 210 Czech control samples (420 alleles) were directly screened encompassing the position of the Czech c.20+544G>T variant and other variants identified in PPCD patients. No variant alleles found in the PPCD cohort were detected in controls. Interestingly, in contrast, this genomic region is conserved, as only a single heterozygous variant (rs548346355) was identified in a single control in the Czech cohort, supporting the hypothesis that this region is an important regulatory/functional region.

4.6.6 Expression of GRHL2 in the cornea

GRHL2 encodes a transcription factor that is a direct transcriptional repressor of the PPCD3 gene, *ZEB1* (Cieply *et al.*, 2012). RNA-seq data generated in-house demonstrated that *GRHL2* is expressed in control corneal buttons (section 2.14). However, publicly available RNA-seq data from both adult and foetal human endothelial tissue showed no evidence of *GRHL2* expression (Figure 4.24) (Chen *et*

al., 2013). Interrogation of a different publicly available RNA-seq data demonstrated that *GRHL2* was expressed in the central corneal epithelium and in putative stem cell niches basal limbal crypts and superficial limbal crypts, but absent from the corneal stroma (data not shown) (Bath *et al.*, 2013).

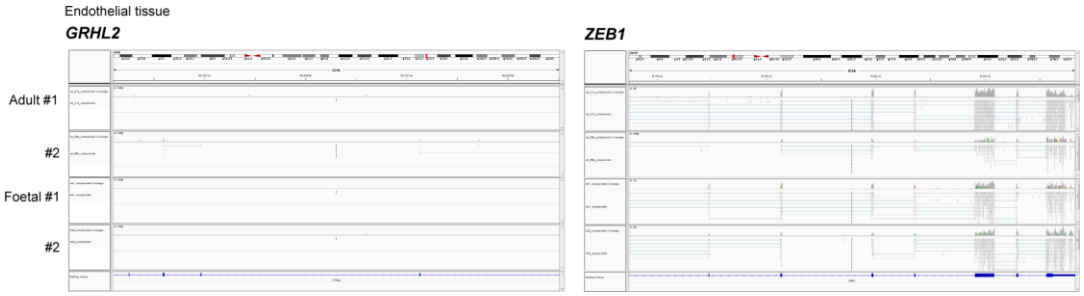


Figure 4.24 *GRHL2* is not expressed in healthy corneal endothelium. Interrogation of endothelial tissue RNA-seq data reveals absence of *GRHL2* transcripts. Multiple reads are shown for *ZEB1* expression for comparison.

4.6.7 Transcriptional regulation of GRHL2

The four unique variants identified in PPCD patients were all located within the first intron of *GRHL2*. To determine whether this is likely to be a regulatory region, we interrogated publicly available ENCODE data (section 2.3.6). The data revealed two distinct clusters of transcription factor binding; the first occurring upstream of the *GRHL2* gene and spanning the 5' UTR and first exon. This cluster encompasses variants c.20+32_20+51del and c.20+133delA. The second is a cluster deeper into the first intron, and overlaps with the c.20+544G>T variant, with binding sites for transcription factors TAF1, YY1, Pol2 and c-Myc. The c.20+257delT variant occurs between these two clusters of transcription factor binding (Figure 4.25 A). Alignment of orthologous nucleotide sequences revealed that the c.20+32__20+51del variant is conserved in chimpanzee only and the c.20+133delA variant is conserved in both chimpanzee and mice. The c.20+257delT variant is conserved in all orthologous sequences aligned, except for mice. The c.20+544G>T variant is conserved in chimpanzee, mice, and African clawed frog, but not chickens or Carolina anole lizard, where a 'C' is present (Figure 4.25 B). The bioinformatics transcription factor binding prediction tools MatInspector and AliBaba 2.1 were used to determine whether the variants identified in PPCD patients are predicted to disrupt

the binding of a transcription factor, or introduce a binding site for a new transcription factor (section 2.3.6). All four variants altered multiple transcription factor binding sites with the introduction and loss of at least one binding site for each variant tested. The majority of transcription factors show expression in adult or foetal endothelial tissue (Table 4.10).

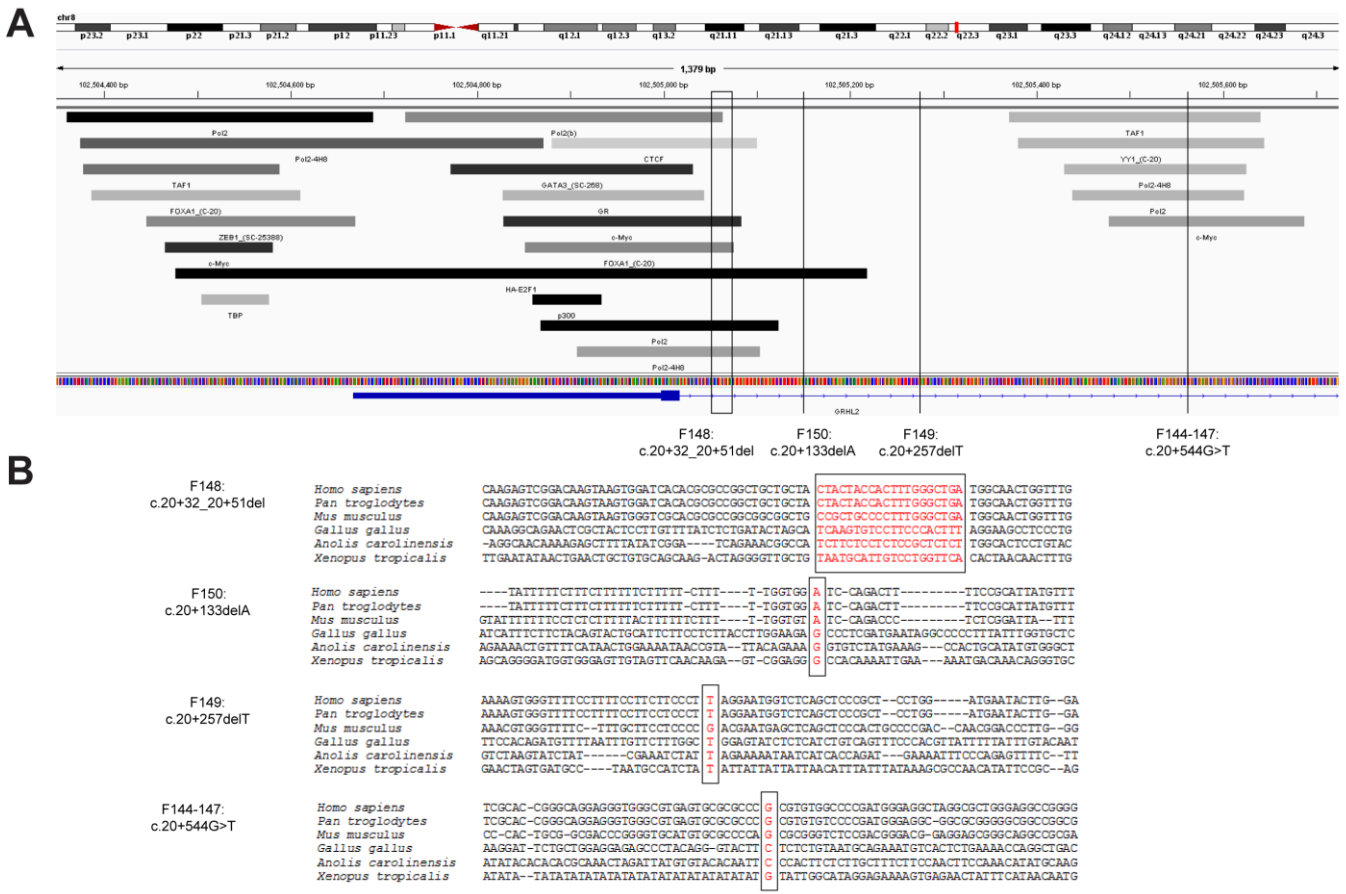


Figure 4.25 Predicted sites of transcription factor binding and location of *GRHL2* non-coding variants. (A) Clusters of transcription factor binding within *GRHL2* exon 1 and intron 1 region spanning the location of the variants identified. Three of the four variants occur at known transcription factor binding sites. **(B)** Alignment of orthologous nucleotide sequences with *GRHL2* variants highlighted in red.

Table 4.10 In silico predictions of the consequences of *GRHL2* variants on transcription factor binding. AliBaba 2.1 is a program for predicting transcription factor binding sites in an unknown DNA sequence utilising binding sites collected in TRANSFAC public database. Mat Inspector is a software tool that utilizes a large library of matrix descriptions for transcription factor binding sites to locate matches in DNA sequences. *RKPM values also encompass flanking transcripts. All four mutations were predicted to alter transcription factor binding. All four variants were predicted to cause both the loss of transcription factor binding and gain of new binding sites.

Transcription Factor	F144-F147 c.20+544G>T	F148 c.20+32_20+51del	F149 c.20+257delT	F150 c.20+133delA	Expression in endothelium (RKPM)	
					Adult	Foetal
SP1	Site lost[1]				12.52	17.53
NRF1*	4 sites lost[2]				9.31	12.78
ESRRA*	Site gained[2]				41.52	8.74
GLIS1	Site gained[2]				42.75	3.19
AHR	Site lost[2]				32.84	28.87
MYC	Site lost[2]				25.00	5.98
E2F2	Site lost[2]				0.07	1.44
E2F3	Site gained[2]				0.17	0.02
MYOD1		Site lost[1]			0.00	0.00
NF1		Site lost[1]			5.41	6.01
MYF6		Site gained[1]			0.00	0.48
TFAP4		Site lost[2]			1.73	3.18
MAF		Site gained[2]			16.70	15.49
RFX1		Site gained[2]			8.38	10.28
HMX3		Site lost[2]			0.01	0.02
NEUROD1*		Site lost[2]			0.04	0.62
RFX2		Site lost[2]			1.40	2.26
LTSM binding proteins		Site lost[2]			N/A	N/A
EBF1			Site lost[2]		0.05	5.98
POZ/zinc finger proteins	Site lost[2]		Site gained[2]		N/A	N/A
STAT6			Site gained[2]		27.53	21.94
ZNF354C				Site lost[2]	3.53	5.61
GLI3				Site gained[2]	20.15	14.68
ZBTB7A				Site gained[2]	17.91	7.01

[1] Alibab 2.1 [2] MatInspector

4.6.8 Histological staining of patient endothelial tissue

Proband F148: II:3 had previously undergone EK surgery in 2014, with the removed endothelium submitted for histological analysis to provide confirmation of the diagnosis of PPCD. Histological analysis was carried out by Dr. Caroline Thaung. Notably, the sample showed a greater endothelial cell density than would be typically expected from a patient of this age, although no multi-layering was observed. Following the identification of a c.20+32_20+51del variant in the first intron of *GRHL2* in the proband of F148, ethics approval was sought and granted for use of the histological specimen as a research sample for immunohistochemistry (Figure 4.26; section 2.15).

A number of proteins define corneal cell types, so antibodies for these proteins were selected for immunohistochemistry. In addition, a GRHL2 antibody was used to test for ectopic expression in primary diseased tissue. CD56 is a cell adhesion protein which is expressed on membranes of corneal endothelial cells (He *et al.*, 2016). This antibody stained positive in both the patient and the control samples, excluding the cell nuclei (Figure 4.26 A). Vimentin is a mesenchymal cell marker that is a component of the cytoskeleton intermediate filaments (Kidd, Shumaker and Ridge, 2014) which was positive in both case and control tissue, excluding the nuclei. In both cases, the entire tissue was stained positive (Figure 4.26 B).

Staining of proteins involved in EMT/MET revealed differences between the endothelium of the patient sample when compared to a control corneal endothelium. Neural cadherin (N-CAD) is a cadherin expressed on the membranes of mesenchymal cells, including corneal endothelial cells (He *et al.*, 2016). N-CAD stained all cells of the control endothelium, but showed clear patches of both positive and negative staining in the patient sample, indicating that the cells in the diseased tissue were abnormal, and in some places cells were diverging from their mesenchymal phenotype (Figure 4.26 C). Membrane protein epithelial cadherin (E-CAD), a component of adherens junctions, is not expressed in healthy corneal endothelium (Lamouille, Xu and Derynck, 2014; He *et al.*, 2016), however clear regions of positive staining for E-CAD were detected in the patient sample, which was completely negative in the control endothelium (Figure 4.26 D). This

downregulation of N-CAD and corresponding upregulation of E-CAD is consistent with cells undergoing MET.

Previous immunohistochemistry of PPCD samples has indicated that cytokeratins are ectopically expressed in the diseased tissue (Jirsova *et al.*, 2007). Antibodies for two cytokeratin markers (CK8/18 and CK7) were positive in the patient endothelium and were negative in the control sample (Figure 4.26 E and F). Taken together these data indicate that the patient endothelial cells were in transition from a mesenchymal to an epithelial cell type, or had already diverged.

Given that *GRHL2* is not usually expressed in the corneal endothelium, and that GRHL2 protein is a repressor of *ZEB1* gene expression, we hypothesized that GRHL2 would be ectopically expressed in the patient sample, but absent from control endothelium. Interestingly, some patient 'endothelial cells' were positive for GRHL2 protein expression localised to the nuclei, consistent with its role as a transcription factor (Figure 4.27). Notably, there was no GRHL2 expression in control samples (without corneal endothelial disease), as predicted by the transcript analysis in corneal cells and tissue (section 4.6.6).

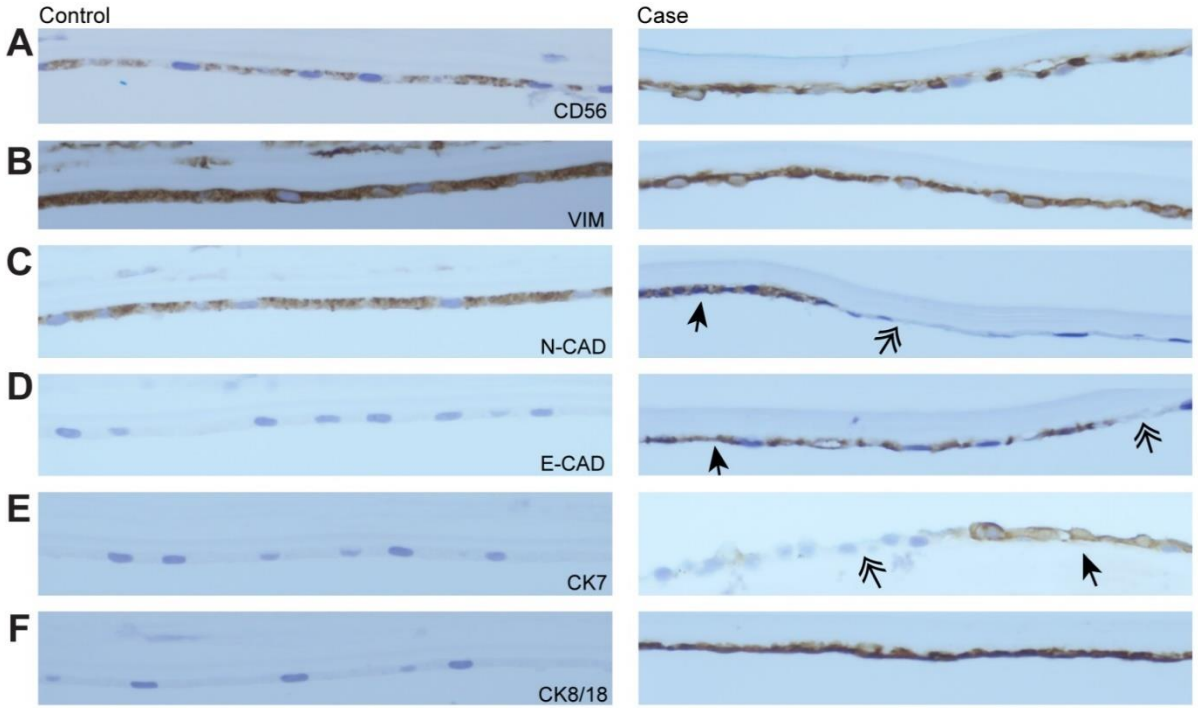


Figure 4.26 Immunohistochemistry of a patient with posterior polymorphous corneal dystrophy reveals diseased endothelial cells undergo mesenchymal-to-epithelial transition. Control cornea is shown in left panel for each antibody and patient's corneal endothelium in the right panel. **(A)** Corneal endothelial cell marker CD56 is positive for both control and patient endothelial tissue. **(B)** Mesenchymal marker vimentin (VIM) was positive for both control and patient endothelial tissue. **(C)** N-cadherin (N-CAD) was positive for control endothelial tissue but demonstrated patches of positive staining (arrowhead) and areas of negative staining (double arrowhead). **(D)** E-cadherin is negative for control tissue but shows positive staining in patient tissue (arrowhead) and some areas of negative staining (double arrowhead). **(E)** Cytokeratin 7 (CK7 was negative for control but areas of both positive (arrowhead) and negative (double arrowhead) were observed in patient tissue. **(F)** Cytokeratin-8 and -18 (CK8/18) are negative in control tissue but stain positive in patient tissue.

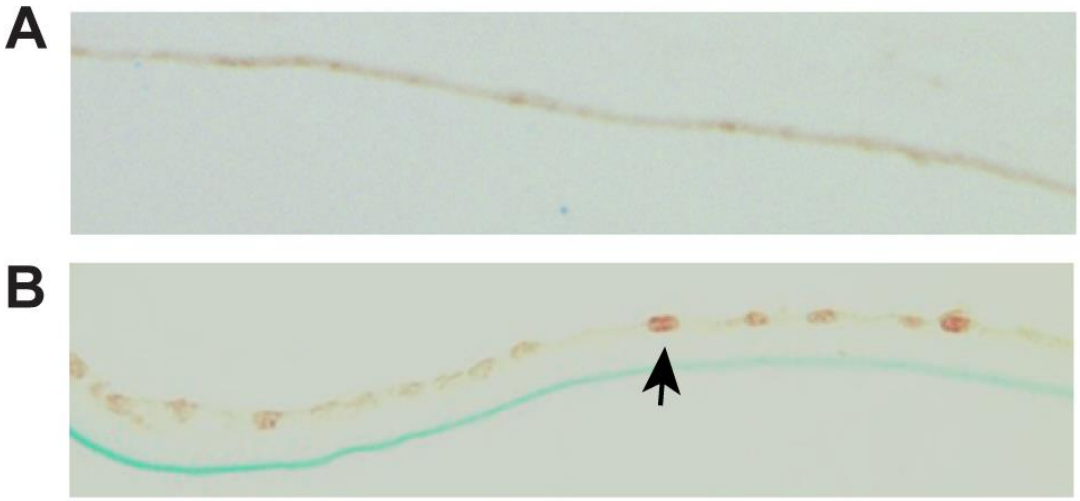


Figure 4.27 Immunocytochemistry reveals ectopic expression of GRHL2 in posterior polymorphous corneal dystrophy (PPCD) endothelial cells. (A) No evidence for positive GRHL2 staining in control corneal endothelial cells, although there is some background staining. **(B)** GRHL2 staining in PPCD case, without counterstain, shows positive staining of nuclei (arrowhead).

4.7 Additional genetic heterogeneity for PPCD

Following screening of *ZEB1, OVOL* and *GRHL2*, 28 probands diagnosed with PPCD remained without a genetic diagnosis from the MEH (15), Czech (7) and New Zealand (6) cohorts. No further analysis could be performed on the New Zealand probands due to lack of DNA, however the others were further investigated. Other endothelial dystrophy genes have been implicated as a potential cause of PPCD (*COL8A2*), CHED (*SLC4A11*) and FECD (*LOXHD1*, *AGBL1*) which were not been previously excluded by direct screening. WES was therefore performed on six unsolved families (section 2.6). Initially, each family was analysed individually for rare (MAF $\leq$ 0.005) variants in genes associated with endothelial corneal dystrophies, followed by a combined analysis of WES data of unsolved PPCD cases.

4.7.1 Known endothelial dystrophy genes

The proband F151: II:1 was recruited from the Czech Republic but was originally from Greece. He was diagnosed with PPCD and had an affected brother (II:2). Clinical information was minimal as the patient was lost to follow up. However, he had previously undergone DSAEK in his right eye (aged 64). Pre-screening was

negative for *ZEB1* coding mutations, the *OVOL2* promoter and the *GRHL2* region of interest (section 2.2; Appendix A). DNA from the proband was analysed by WES (section 2.6). Filtering for rare (MAF $\leq$ 0.005) non-synonymous or splice site variants (heterozygous or homozygous) in genes associated with corneal dystrophies (Table 1.1) and inherited corneal disorders (Appendix B) identified six potential variants, of which two associated with endothelial dystrophies were relevant (Table 4.11; section 2.6.2.3). A heterozygous variant was identified in the second coding exon of *OVOL2*: c.167_184del; p.(Gly56_Ser61del). This is a rare variant with a frequency of 1/95233 (rs767318236) in ExAC. The second was a novel heterozygous missense change in *SLC4A11:* c.2246A>T; p.(Tyr749Phe) that was absent from control database ExAC. This was an interesting candidate as it could be indicative of autosomal recessive CHED, with a second variant not detected by WES due to lack of coverage of the all *SLC4A11* exons or alternatively, could be indicative of early-onset FECD which is associated with heterozygous *SLC4A11* missense mutations. However, segregation analysis of both candidate variants in DNA from his affected brother revealed that both were absent and therefore, unlikely to be causative of disease (Figure 4.28).

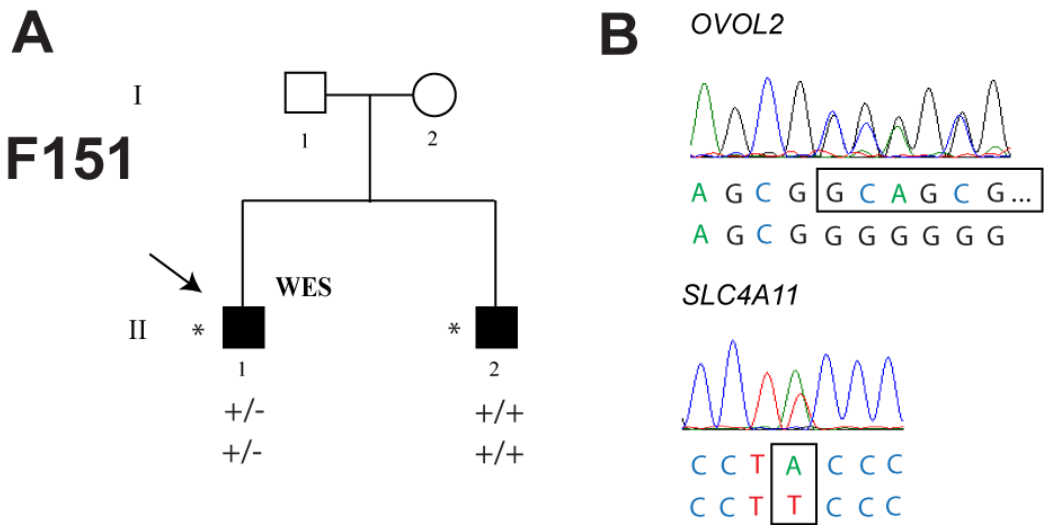


Figure 4.28 Variants identified in *OVOL2* and *SLC4A11* in a family #151 with posterior polymorphous corneal dystrophy. (A) Pedigree showing two affected brothers diagnosed with PPCD. DNA samples were obtained from those with a *. +/- indicates heterozygous for the mutation and +/+ indicates wild type. **(B)** Two variants were identified in endothelial dystrophy genes, *OVOL2*: c.167_184del; p.(Gly56_Ser61del), and *SLC4A11*: c.2246A>T; p.(Tyr749Phe), following whole exome sequencing analysis. Both were excluded following segregation analysis in the pedigree, as both were absent in his affected brother.

Table 4.11 Rare (≤ 0.005) non-synonymous or splice site variants in genes associated with endothelial dystrophies in F151: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
SLC4A11 NM_032034.3	c.2246A>T; p.(Tyr722Phe)	1 (T)	0.021 (B)	NP	Congenital hereditary endothelial dystrophy (AR)
OVOL2 NM_021220.3	c.167_184del; p.(Gly56_Ser61del)	N/A	N/A	1/ 95233	Posterior Polymorphous Corneal Dystrophy (AD)

The proband of F152 (II:1) was a white individual recruited from the Czech Republic with a diagnosis with PPCD, although there was diagnostic uncertainty. She was severely affected, although the age of onset was 11 years, which is more

consistent with a diagnosis of PPCD than CHED2. Given the severity, *SLC4A11* was directly sequenced by Sanger sequencing, in addition to known PPCD genes (*ZEB1*, *OVOL2* and *GRHL2*) prior to WES analysis (section 2.2; Appendix A). Her mother had been diagnosed with early-onset FECD and was only mildly affected (Figure 4.29 B). Due to this diagnosis both the proband and her mother were analysed for *TCF4* intronic repeat expansion, performed by Dr. Alice Davidson. The proband was negative for a *TCF4* repeat expansion, with repeat sizes 18 and 31, whereas her mother had an expansion of one allele (87, 31). WES analysis was performed on DNA from the proband, and a heterozygous splice site variant in *ABGL1*, a gene implicated in FECD, was identified (section 2.6.2.3; Table 4.12). The variant (c.3185+1_2insT) is an insertion that occurs at a canonical splice site between the +1 and the +2 sites with the canonical splice donor 'GT' becoming 'GTT'. As the crucial 'GT' sequence is not affected by this change, is it possible that this variant does not affect splicing. To investigate this further, in silico splice prediction tools were used to predict the effect of this variant. One program predicted no significant change in splice donor usage whereas the two other programs predicted that the site would be abolished (Table 4.13). If the site was abolished, this could lead to exon skipping or use of an alternative splice donor site. Segregation analysis was not possible in this family due to lack of DNA from the proband's affected mother.

Table 4.12 Rare (≤ 0.005) non-synonymous or splice site variants in genes associated with endothelial dystrophies in F152: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
ABGL1 NM_152336	c.3185+1_2insT	N/A	N/A	NP	Fuchs endothelial corneal dystrophy

Table 4.13 In silico predicted effect of the *ABGL1* c.3185+1_2insT variant identified in F152: II:1 on the canonical splice acceptor site. Human Splicing Finder, NNSPLICE and NetGene2 each identify potential splice sites in a given sequence and score them from 0 to 100, 0-1 and 0-1, respectively. In each instance, a higher number corresponds to an increased likelihood of the motif functioning as a splice site. All three programs predict that the presence of the mutation abolishes the canonical splice acceptor site.

	Canonical splice donor score	**Canonical splice donor score in presence of c.3185+1_2insT**	**Alternative splice acceptor score (+97_98)**
Human Splice Finder (0 – 100)	77.32	77.35	N/A
NNSplice (0 – 1)	0.94	0	0.64
NetGene2 (0 – 1)	0.35	0	N/A

The remaining four families were negative for any potentially pathogenic changes in known endothelial dystrophy genes (section 2.6.2.3).

The proband F153: II:1 recruited from MEH was black African, and diagnosed with PPCD aged 11. He also had high myopia, astigmatism and consecutive exotropia. Both parents and two younger siblings were examined following the diagnosis, but none were affected. He also had an older sibling who could not be examined (Figure 4.28 C). Two subsequent annual appointments revealed stable vision and no progression of PPCD symptoms.

The proband F154: II:1 was first clinically examined age 8 and was of black African ethnicity, recruited from MEH. He was bilaterally affected with focal and linear changes at the level of Descemet's membrane and the corneal endothelium (Figure 4.29 A). Corneal tomography (Pentacam) showed mild irregularity of corneal shape. Corneal endothelial density was reduced in the right eye in comparison to the left eye. The changes had not progressed when he was clinically examined a year following diagnosis. Although asymptomatic, his mother (I:2) was also found to have milder endothelial changes including band-type PPCD in her left eye (Figure 4.29 D).

The proband of family F155 was of white Czech origin (Figure 4.29 E). She had mild band-type PPCD which was non-progressive. The proband of family F156 was of white Czech origin with a typical PPCD clinical phenotype (Figure 4.29 F). There was no history of PPCD or corneal disease in either family.

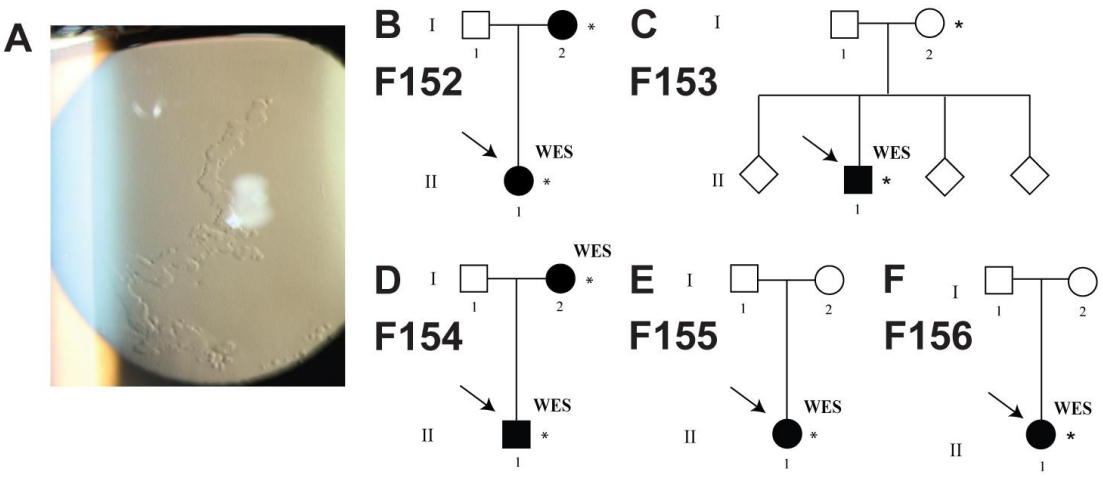


Figure 4.29 Unsolved families with posterior polymorphous corneal dystrophy analysed by whole exome sequencing (WES) and direct screening. DNA samples were obtained from those with a *. WES indicates sample was analysed by whole exome sequencing. **(A)** Clinical appearance of band-type PPCD in F154: II:1 viewed in slit lamp retroillumination. **(B)** F152. **(C)** F153. **(D)** F154. **(E)** F155. **(F)** F156.

4.7.2 Combined WES analysis of unsolved PPCD families

As Sanger sequencing and WES analyses revealed that no known endothelial dystrophy genes were causative of PPCD in these six probands, we hypothesised that a novel PPCD gene could be responsible for the dystrophy in some or all the families. Variants from all six probands were filtered for non-synonymous or splice site changes that were absent from internal controls, 1000G and EVS (section 2.6.2.1). When the analyses from each proband were combined, no gene was found to harbour rare heterozygous or homozygous coding or splice site variants from four or more unsolved PPCD probands; however, one gene, *PCDH1*, was found to have rare variants in three probands (Table 4.14).

Table 4.14 Three genetically unsolved probands with posterior polymorphous corneal dystrophy harbour a rare variant in *PCDH1*. Only one gene contained variants absent from 1000G, EVS and UCL internal controls, in more than 2 unsolved PPCD probands. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Family	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
PCDH1 NM_032420.3	F154	c.3488G>A p.(Gly1163Glu)	0.29 (T)	0.012 (B)	2/108008
	F151	c.2908C>G p.(Arg970Gly)	0 (D)	0.0962 (PrD)	NP
	F156	c.1562C>G p.(Pro521Arg)	0.22 (T)	0.457 (PoD)	NP

The *PCDH1* variant in the proband of F154 was also present in his affected mother, who was also analysed by WES. The proband of F156 did not have a family history of disease and no familial samples were available. Interrogation of RNA-seq data revealed low levels of expression of *PCDH1* in adult, but not foetal, corneal endothelium (section 2.14) (Chen *et al.*, 2013).

In addition to these six unsolved PPCD probands, the unsolved PPCD cohort also contains a further 5 probands with typical PPCD (3 Czech, 2 British; F156 – 160) and 11 probands with the band-type PPCD subtype recruited from MEH (F161 – 171), for which mutations in *ZEB1*, *OVOL2* and *GRHL2* have been excluded. Further WES analysis of this cohort should confirm or exclude a role for this gene in PPCD and identify further candidate genes for investigation.

4.8 CHED2

Five probands were recruited with autosomal recessive CHED2. Three were excluded because they had already been genetically solved (Ramprasad *et al.*, 2007).

Two were screened for mutations in *SLC4A11* by direct sequencing of coding exons (section 2.2; Appendix A).

The proband from F173 was from the United Arab Emirates and was diagnosed with CHED in the first year after birth due to cloudy corneas. The proband was from a consanguineous family, with no family history of corneal disease, consistent with recessive disease (Figure 4.30 A). Screening the coding exons of the *SLC4A11* gene revealed a novel homozygous in-frame deletion, resulting in the deletion of a single amino acid residue in exon 12: c.1376_1378del; p.(Phe459del) (Figure 4.30 B).

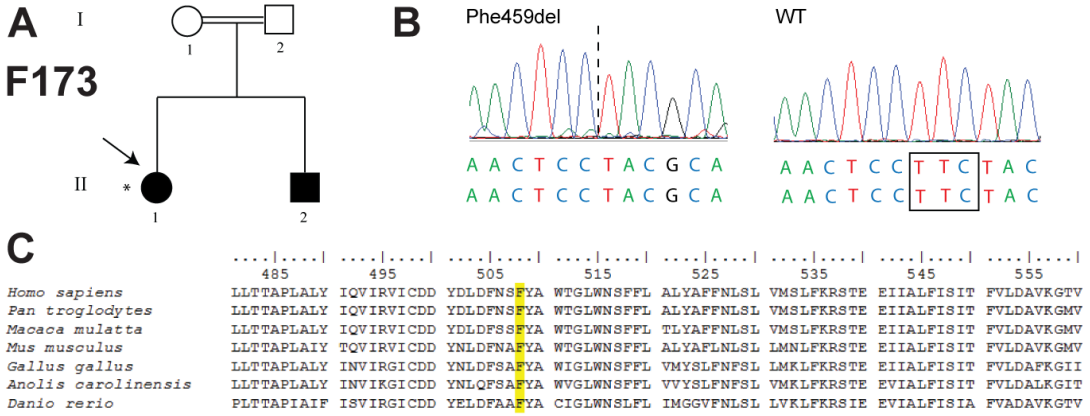


Figure 4.30 Novel homozygous in-frame deletion in *SLC4A11* in a consanguineous family F173 with congenital hereditary endothelial dystrophy (CHED). (A) Pedigree showing two affected siblings affected with autosomal recessive CHED2. **(B)** Sequence electropherogram demonstrating identification of a homozygous in-frame deletion c.1376_1378del; p.(Phe459del). **(C)** Alignment showing conservation of the Phe-459 residue deleted by the mutation.

The proband of family #174 was recruited with a clinical diagnosis of CHED. He was treated previously with a Boston keratoprosthesis following the failure of corneal grafting in the left eye. Three corneal grafts had previously failed in the right eye. He was first seen at MEH in his forties. There was no evidence of consanguinity or a family history of corneal disease (Figure 4.31 A). Screening *SLC4A11* revealed a novel heterozygous variant c.2003T>C; p.(Leu668Pro) (Figure 4.31 B). This variant is predicted to be 'Damaging' (0.01) and 'Probably Damaging' (1.000) by SIFT and Polyphen2, respectively and is conserved in homologous protein sequences (Figure 4.31 C). No other potentially pathogenic mutations were identified in the remaining coding exons of the gene.

SLC4A11 has multiple annotated transcripts listed on the Ensembl database. These transcripts are similar in size (875, 891 and 918 amino acids), two (SLC-001 and SLC-003) with 20 exons and SLC-002 with 19 exons in which the first exon used encompasses the second exon of the other transcripts. Most exons are common to all potential transcripts, except for first exon usage. In the literature, the majority of mutations are annotated against the SLC-002 transcript, which is also used here. The alternative first exons used by SLC-001 (NM_001174090) and the SLC-003 (NM_001174089) were amplified and directly screened. One missense variant was identified in the alternative exon one in the SLC-001 transcript: c.148A>G; p.(Arg50Gly); however, this has an average MAF of 0.06, and is as high as 0.12 in some populations, arguing against a pathogenic role. In the canonical transcript SLC-002 this variant is located upstream of the first exon, and is in the first intron in SLC-003.

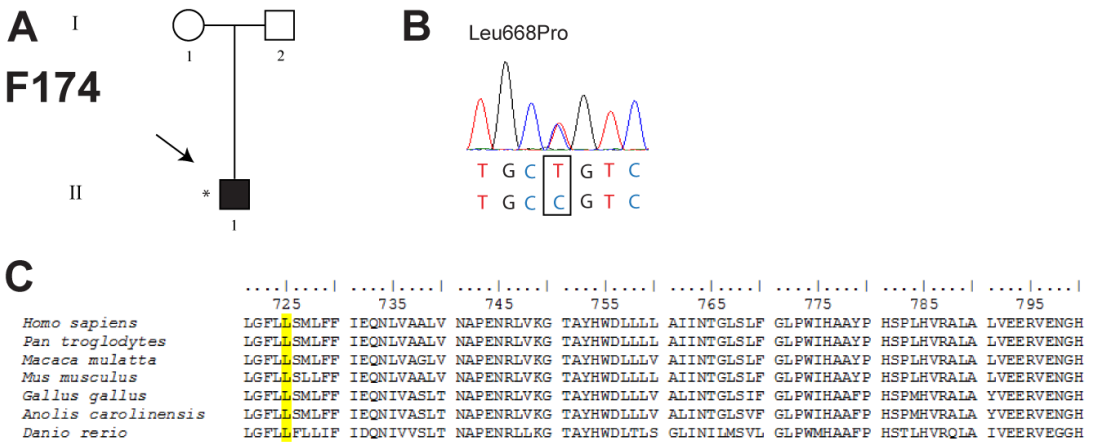


Figure 4.31 A single novel heterozygous *SLC4A11* mutation in recessive congenital hereditary endothelial dystrophy in family #174. (A) Pedigree showing no family history of corneal disease. **(B)** Sequence electropherogram showing presence of novel heterozygous c.2003T>C; p.(Leu668Pro) variant. No other potentially pathogenic variants were found by direct sequencing of *SLC4A11*. **(C)** Alignment of orthologous SLC4A11 protein sequences shows that Leu-668 is a conserved residue.

4.9 Screening *OVOL2* and *GRHL2* in a FECD cohort

Typical late-onset FECD is caused by an intronic trinucleotide repeat expansion in the *TCF4* gene in approximately 80% of patients, however a minority of cases are associated with missense mutations in endothelial dystrophy genes including *ZEB1*

(as opposed to frameshift, stop and deletion mutations which cause PPCD3), *COL8A2*, and *SLC4A11*. We therefore hypothesised that novel PPCD genes *OVOL2* and *GRHL2* were good candidates for FECD.

An enriched FECD cohort of 25 probands was selected for screening. We included individuals in which the disease was early-onset and they were negative for the intronic *TCF4* repeat expansion (Dr. Alice Davidson), although the possibility of a repeat expansion beyond the limits of detection could not be fully excluded. All samples were negative for the previously identified *OVOL2* and *GRHL2* mutations (section 2.2; Appendix A).

Table 4.15 Summary of families (F124 – F174) analysed in chapter 4, their respective clinical diagnoses and the causative mutation identified. Standard acronyms are use d for corneal dystrophies. (T) indicates the standard phenotype of PPCD whereas (B) is band-type PPCD. ? prior to a diagnosis indicates diagnostic uncertainty.

Family ID	Clinical Diagnosis	Mutation
F124	PPCD (T)	Het *ZEB1:* c.1100C>A; p.(Ser367*)
F125	PPCD (T)	Het *ZEB1:* c.1576dupG; p.(Val526Glyfs*3)
F126	PPCD (T)	Het *ZEB1:* c.1669C>T; p.(Gln557*)
F127	PPCD (T)	Het *ZEB1:* c.1918C>T; p.(Gln640*)
F128	PPCD (T)	Het *ZEB1:* c.685-2A>G
F129	PPCD (T)	Het *ZEB1:* c.627delT; p.(Phe209Leufs*11)
F130	PPCD (T)	Het *ZEB1:* c.2577delA; p.(Val860*)
F131	PPCD (T)	Het chr10.hg19: 29,277,299_32,553,577del
F132	PPCD (T)	Het chr10.hg19:g.31,784,795_31,855,747delinsACACACAC
F133	PPCD (T)	Het chr10.hg19: g.(31,298,325_31,350,494)_(32,214,122_32,285,574) del
F134	PPCD (T)	Unsolved
F135	PPCD (T)	Unsolved
F136	PPCD (T)	Unsolved
F137	PPCD (T)	Unsolved
F138	PPCD (T)	Unsolved
F139	PPCD (T)	Unsolved
F140	CHED	*OVOL2:* c.-340_339ins22
F141	PPCD (T)	*OVOL2:* c.-370T>C
F142	PPCD (T)	*OVOL2:* c.-274T>G
F143	PPCD (T)	*OVOL2:* c.-307T>C
F144	PPCD (T)	*GRHL2:* c.20+544G>T
F145	PPCD (T)	*GRHL2:* c.20+544G>T
F146	PPCD (T)	*GRHL2:* c.20+544G>T
F147	PPCD (T)	*GRHL2:* c.20+544G>T
F148	PPCD (T)	*GRHL2:* c.20+32_20+51del
F149	PPCD (T)	*GRHL2:* c.20+257delT
F150	PPCD (T)	*GRHL2:* c.20+133delA
F151	PPCD (T)	Unsolved
F152	?PPCD ?FECD	Unsolved
F153	PPCD (T)	Unsolved
F154	PPCD (B)	Unsolved

F155	PPCD (B)	Unsolved
F156	PPCD (T)	Unsolved
F157	PPCD (T)	Unsolved
F158	PPCD (T)	Unsolved
F159	PPCD (T)	Unsolved
F160	PPCD (T)	Unsolved
F161	PPCD (T)	Unsolved
F162	PPCD (B)	Unsolved
F163	PPCD (B)	Unsolved
F164	PPCD (B)	Unsolved
F165	PPCD (B)	Unsolved
F166	PPCD (B)	Unsolved
F167	PPCD (B)	Unsolved
F168	PPCD (B)	Unsolved
F169	PPCD (B)	Unsolved
F170	PPCD (B)	Unsolved
F171	PPCD (B)	Unsolved
F172	PPCD (B)	Unsolved
F173	CHED	Hom *SLC4A11:* c.1376_1378del; p.(Phe459del)
F174	CHED	Het *SLC4A11:* c.2003T>C; p.(Leu668Pro)

4.10 Discussion

4.10.1 Spectrum of ZEB1 mutations causing PPCD

ZEB1 mutations were the most common cause of PPCD in the MEH cohort. Seven mutations were identified in the coding region or splice sites of the *ZEB1* gene and a deletion encompassing *ZEB1* was identified in a further PPCD proband from MEH. There are therefore eight *ZEB*1 mutation-positive families from a total cohort of 28 PPCD families (29%) recruited from MEH. *ZEB1* mutations frequently occur *de novo* resulting in a lack of family history or small pedigrees with few affected individuals. In this study, of the eight MEH families who were positive for *ZEB1* mutations, only two reported a family history of disease and this was restricted to two generations. Non-penetrance has also been reported in a small number of cases but was not observed in the families in our cohort (Liskova *et al*., 2013). Six of the seven coding or splice mutations identified were private to the affected family. The remaining *ZEB1* mutation [c.1576dup; p.(Val526Glyfs*2)], identified in F125, has been previously reported in three presumed unrelated families (Krafchak *et al.*, 2005; Bakhtiari *et al.*, 2013; Lechner *et al.*, 2013). To date, this is the only recurrent *ZEB1* mutation identified. The parents of the proband were deceased therefore testing for *de novo* occurrence of this mutation was not possible, but neither parent had been diagnosed with a corneal dystrophy during their lifetimes.

4.10.2 Identification of ZEB1 deletions confirms haploinsufficiency as the mechanism of disease

Sanger sequencing of *ZEB1* is unable to detect the presence of large heterozygous deletions as the remaining allele is amplified during PCR, which explains why some probands who were negative for *ZEB1* coding mutations by direct screening were later found to be caused by *ZEB1* deletions. Sanger sequence data of a single allele is indistinguishable to that obtained from sequencing two alleles, with the exception that all variants appear homozygous. As *ZEB1* is not a highly polymorphic gene, it is typical for the coding region of the gene to contain no heterozygous variants, even when both alleles are present; therefore, the absence of heterozygous changes cannot be relied upon as an indication of a potential deletion. Conversely, given the

identification of a partial *ZEB1* deletion in one proband, the occurrence of heterozygous variants does not exclude the possibility of a deletion encompassing part of the *ZEB1* gene. Familial WES data, however, is useful for assessing the likelihood of a *ZEB1* deletion, particularly those encompassing flanking genes. In this study, a *ZEB1* deletion was detected due to the inconsistencies in segregation of homozygous variants located within, and flanking, the *ZEB1* locus. However, this is not always possible due to the large number of probands with no apparent family history of disease, either because of *de-novo* occurrence or non-penetrance. In these cases, bioinformatics programs (e.g. ExomeDepth) are able to detect deletions from WES data using read count information (Plagnol *et al.*, 2012). This does not require familial WES samples but it does require WES to be performed in batches in order to compare expected read depth at a particular region across different DNA samples.

A qPCR assay was designed to confirm and detect *ZEB1* deletions, using primers designed to the 5' and 3' end of the gene. The 5' qPCR assay was consistent and reliable, providing ZEB1/TWIST1 ratios of either 1.0 or 0.5 for all of the DNA samples screened. However, the 3' assay, while consistently detecting 1 or 2 alleles for the majority of samples, including positive PPCD controls, produced spurious results for a sub-set of the samples screened. One possibility was the amplification of more than one PCR product, however the melt curve revealed a single peak and only a single product was observed on an agarose gel. This explanation is also not consistent with spurious results occurring for only a sub-set, rather than all, of the samples. The presence of a SNP interfering with primer annealing was considered, and different primers were designed which targeted a different region of the 3' end of the gene however the spurious results occurred again with the same samples, even though the majority of samples worked well. It is possible that DNA quality of some of the samples is inadequate for this assay, although if this is the case it does not appear to affect the results of the 5' assay, which is more robust. The use of qPCR as a mechanism for detecting *ZEB1* deletions is therefore currently only useful as an accurate screening test for the 5' end of the gene. This assay would be expected to detect most deletions that encompass the entire gene but will miss smaller 3' deletions. In our cohort, the qPCR assay detected a deletion in one proband which was subsequently confirmed using other methods. However, given that the three deletions identified differ in size and in deletion break points, a negative result using

the qPCR assay cannot exclude the possibility of a deletion. The most useful and reliable method to detect *ZEB1* deletions is the use of a SNP genotyping array, which can detect deletions and provides information regarding the chromosomal location of the deletion breakpoints.

Using a variety of methods, a total of three deletions encompassing *ZEB1* were identified in two probands of Czech origin (F131 and F132) and one British proband (F133). The deletion breakpoints were mapped for both F131 and F132 using long range PCR from estimates of breakpoints derived from the SNP array data. An estimate of the boundaries for the proximal and distal deletion breakpoints was obtained for F133 from a confirmatory SNP genotype array. Due to the difference in the deletion breakpoints between the three cases, and that the deletions are unique to each family, it is likely that non-homologous end joining and/or micro-homologous end joining is responsible for the deletions. This is in contrast to MCD in which a deletion or rearrangement upstream of *CHST6* occurs in multiple families due to homologous recombination (Akama *et al.*, 2000). Interestingly, a 36 bp region of microhomology was identified that flanks the large deletion identified in F131, which may have originated from the proximal or distal break point, indicating that this may have influenced the extent of the deletion (McVey and Lee, 2008).

To date, all coding mutations identified in *ZEB1* which cause PPCD3 are predicted to be truncating; no missense mutations are reported. The seven *ZEB1* coding mutations identified by Sanger sequencing in the MEH cohort were either nonsense, frameshift or splice site mutations, all of which are predicted to cause the inclusion of a premature stop codon in the *ZEB1* transcript. It is likely that these transcripts will be degraded by nonsense mediated decay (NMD) or, if translated, will produce truncated proteins; therefore, haploinsufficiency of *ZEB1* has long been assumed to be the mechanism of disease for PPCD3. Supporting this hypothesis, heterologous expression studies demonstrate that some of these truncated proteins have cytosolic rather than nuclear localisation or decreased protein abundance (Chung *et al.*, 2014). Although unlikely, an alternative possibility is a dominant-negative function of the translated truncated *ZEB1* proteins. With the identification of deletions described here, which encompass the entire *ZEB1* gene and cause PPCD3, this now conclusively demonstrates that haploinsufficiency is the mechanism of disease.

4.10.3 Identification of OVOL2 as the CHED1/PPCD1 gene

Despite the identification of an overlapping genetic locus on chr20p11 for both CHED1 and PPCD1 pedigrees in 1995, the genetic cause of disease has remained elusive for both disorders (Héon *et al.*, 1995; Toma *et al.*, 1995). Using parallel strategies of linkage analysis, followed by WGS, for the British CHED1 family and the Czech founder families, without the *a priori* assumption that they are both caused by the same gene, we have identified the common cause of these conditions as mutations located in the promoter of *OVOL2*. Further targeted screening in the unsolved PPCD cohort resulted in a total of four novel mutations located in the promoter of *OVOL2* in two British PPCD probands, one British CHED1 family and one mutation which is causative in 16 PPCD families in the Czech Republic comprising over 100 affected individuals. As we have demonstrated that CHED1 and PPCD1 are allelic conditions, both conditions can now be considered to be a clinical spectrum of PPCD1. Since the assignment of the nomenclature CHED1 is now redundant, CHED2 (autosomal recessive) is the only form of CHED, in agreement with recommendation by the recent update of the IC3D classification of corneal dystrophies (Weiss *et al.*, 2015).

OVOL2 encodes a transcription factor which is a direct transcriptional repressor of *ZEB1* (Hong *et al.*, 2015). Unlike *ZEB1*, which is normally expressed in corneal endothelium, *OVOL2* expression was not detected in the corneal endothelium, in contrast *OVOL2* expression was detected in the corneal epithelium, which is not relevant to endothelial disease. Given the absence of *OVOL2* expression in the endothelium and the location of the mutations in the promoter of *OVOL2*, we hypothesised that these mutations may cause aberrant expression of *OVOL2*, during development and/or in adult corneal endothelial tissue. It is possible that *OVOL2* is expressed in the corneal endothelium at a specific developmental stage that was not tested in this study, and that disruption of this expression pattern may be the disease mechanism. However, ectopic expression could be caused by the introduction of new transcription factor binding sites which drive expression in the corneal endothelium, or via disruption of transcriptional repressors. The hypothesis that these mutations affect transcriptional regulation of *OVOL2* was supported by *in vitro* luciferase assay data performed by Dr. Alice Davidson which demonstrated increased expression of a

reporter gene when driven by the *OVOL2* promoter containing any of the identified mutations, in comparison to the wild type promoter activity.

There are differences in the clinical appearance and severity of PPCD in the four families harbouring *OVOL2* promoter mutations. These differences may be in part due to the different mechanisms by which *OVOL2* is aberrantly expressed in corneal endothelium cells by each mutation. Specifically, different transcription factor binding sites are either lost or gained that would influence when *OVOL2* expression is switched on, and/or the level of expression. For example, ectopic expression earlier in development may result in early-onset disease, more features of anterior segment dysgenesis and a more severe clinical phenotype.

4.10.4 Identification of GRHL2 as the cause of a novel PPCD locus

Linkage analysis in the large Czech family F144 represents the first evidence for an additional locus for PPCD, with a statistically significant new locus, PPCD4, mapping to chr8.hg38:100,821,039 – 119,725,923 and exclusion of known PPCD loci. Extensive analysis by WES of eleven individuals from the pedigree failed to identify a causative change in a coding region within the locus, leading us to hypothesise that the causative mutation, like *OVOL2* mutations causing PPCD1, could lie in a non-coding regulatory region. Using a genetic strategy similar to that previously used to identify *OVOL2* promoter mutations; we performed WGS of four distantly related affected individuals. Shared novel variants within the genetic locus were assessed for potential function consequences, leading to the identification of a putative mutation in the first intron of *GRHL2* within a transcriptionally active conserved genomic domain. Further screening in an unsolved PPCD cohort identified three British families with novel variants in the first intron of *GRHL2*, although one was partially penetrant if causative of PPCD in the proband. These are also compelling candidate variants given the regions are highly conserved, only one polymorphism was found in 420 alleles by direct sequencing, and they are unique variants not found on publicly available databases. However, they require functional validation.

Biallelic coding loss of function mutations in *GRHL2* have been identified in two families with the inherited recessive disorder ectodermal dysplasia (OMIM: 616029)(Petrof *et al.*, 2014). Affected individuals in the families experienced dystrophic nails and/or nail loss, pigmentation of the oral mucosa and/or tongue, tooth abnormalities, and keroderma and hyperkeratosis of the hands and feet. In addition, sensorineural hearing loss and bronchial asthma were noted in some affected individuals. Keratinocytes from affected individuals revealed a reduction in *GRHL*2 expression and a loss of GRHL2 localisation at the cell membrane. Mutant cells failed to form cell junctions and had reduced expression of E-cadherin and zona occludens-2. Since loss of function is the mechanism of disease for this condition cause by coding mutations in *GRHL2*, loss of function is unlikely to be the mechanism of pathogenicity for PPCD.

Similar to *OVOL2*, *GRHL2* is not expressed in the healthy corneal endothelium as determined by both RNA-seq data and RT-PCR, and encodes a transcription factor which is a negatively regulates *ZEB1* expression (Cieply *et al.*, 2012) (section 4.10.6). *GRHL2* is expressed in epithelial tissues, including the corneal epithelium and therefore is present in whole corneal buttons but absent from both the stroma and the endothelium. We therefore hypothesised that the mechanism of disease caused by promoter mutations is likely very similar to PPCD1; ectopic expression of *GRHL2* in the corneal endothelium results in repression of *ZEB1*, thereby causing PPCD.

All putative *GRHL2* mutations are located in the first intron of the *GRHL2* gene, with the furthest variant being 544 bp away from the first exon, and this genomic region is predicted to be a promoter by VEP. It is possible that this region may contain enhancer, repressor or insulator sequences which influence expression of *GRHL2*, and that mutations in this region disrupt these regulatory regions, leading to ectopic expression in the corneal endothelium. Interrogation of publicly available ENCODE data confirms that this genomic region binds to multiple transcription factors, and in silico prediction programs predict that addition and loss of several transcription factor binding sites occurs when the putative mutations are introduced. Further evidence for an important regulatory role of this region is the highly conserved nature of the sequence, with only few reported rare variants occurring in this region.

Whether all the *GRHL2* variants identified in our PPCD cohort are causative of disease has not been fully resolved. For family F144, in which linkage analysis and WGS has been performed, the data are compelling as there is a lack of any other potential locus and causative changes within the locus. The LOD score is statistically significant for this locus and the *GRHL2* variant fully segregates with disease in the large extended pedigree. In contrast, we were unable to show segregation for the three British families. In two families, this was due to the lack of available samples from additional family members and the aim is to recruit more individuals to the study. Given the lack of known family history, convincing evidence would be the appearance of clinical signs of PPCD in other family members who also harbour the *GRHL2* variant, or alternatively, demonstration that the variant has occurred *de novo* in the proband. However, for family F148, other carriers of the *GRHL2* variant, had no clinical signs of PPCD. While this does not fully exclude *GRHL2* as the cause in this family, it does suggest that if it is causative, then it has reduced penetrance, or that there is a different underlying cause of PPCD in this family.

One experimental way to test the pathogenicity of the variants identified would be to conduct a luciferase assay in which the *GRHL2* promoter is cloned upstream of a luciferase open reading frame. An increase in *GRHL2* expression levels in the promoter carrying a variant when compared to the wild type promoter would be evidence for a pathogenic role.

4.10.5 Pathogenicity of PPCD

ZEB1, *OVOL2* and *GRHL2* encode transcription factors that form a functional network with a crucial role in the regulation of EMT and MET (Cieply *et al.*, 2012; Hong *et al.*, 2015). EMT/MET are cellular processes that occur during embryonic development, wound healing and cancer. During EMT, epithelial cells lose their apical-basal polarity, change their morphology and undergo re-organisation of the cytoskeleton. Both tight junctions and adherens junctions are disassembled. The gene expression profile changes from those characteristic of epithelial cells to those of mesenchymal cells. During MET, the process is reversed (Lamouille, Xu and Derynck, 2014). A network of transcription factors regulate complete and partial EMT/MET, however the role of ZEB1 is crucial, acting as a molecular switch

between epithelial and mesenchymal states. ZEB1 is a driver of EMT and is a direct transcriptional repressor of E-cadherin, a component of adherens junctions.

Both GRHL2 and OVOL2 are negative regulators of *ZEB1* transcription, and are themselves negatively transcriptionally regulated by ZEB1 (Cieply *et al.*, 2012; Hong *et al.*, 2015). Haploinsufficiency of *ZEB1* is the mechanism causing PPCD3, as demonstrated by the identification of entire *ZEB1* deletions in this study, eliminating the possibility of a toxic gain-of-function effect. Therefore, given this mechanism, combined with the lack of expression of *OVOL2* and *GRHL2* in healthy corneal endothelium, it can be predicted that ectopic expression of either of these genes in the corneal endothelium would result in transcriptional suppression of *ZEB1*.

Ectopic expression of *OVOL2* due to a promoter mutation would induce its expression in the corneal endothelium in patients, resulting in a corresponding decrease in *ZEB1* expression. A reduction in the levels of ZEB1, a known repressor of *GRHL2*, would therefore then result in lack of repression of *GRHL2* expression. Similarly, ectopic expression of *GRHL2* due to a promoter mutation could induce its expression in the corneal endothelium in patients, resulting in a corresponding decrease in *ZEB1* expression. A reduction in the levels of ZEB1, a known repressor of *OVOL2* and *GRHL2*, would therefore then result in lack of repression of both *OVOL2* and *GRHL2* expression, in a negative feedback loop continually suppressing expression of ZEB1, which is required for normal corneal endothelial cell function (Figure 4.32).

Dysregulation of this regulatory network by reduction of ZEB1 levels (directly or indirectly) in the corneal endothelium, is likely to lead to repression of EMT and/or induction of MET. The healthy adult corneal endothelium consists of mesenchymal cells (section 1.3.3) as demonstrated by positive histological staining for the mesenchymal markers vimentin and N-CAD, and lack of E-CAD expression (epithelial specific) in a histological specimen of control corneal endothelium (He *et al.*, 2016). Corneal endothelial cells in individuals with PPCD demonstrate a number of abnormal characteristics including proliferation and overgrowth resulting in focal multi-layering. Cellular overgrowth onto the posterior surface can occur post-EK

surgery and this has shown to occur due to host cells overgrowing onto the donor tissue (Merjava *et al.*, 2011). It has been proposed that the diseased endothelial cells are 'epithelial-like' with the appearance of microvilli, keratofibrils and desmosomes. Additionally, they express keratins which are usually absent from the corneal endothelium (Jirsova *et al.*, 2007). We also observed expression of cytokeratins in our *GRHL2*-positive patient endothelial tissue. We therefore hypothesise that diseased corneal endothelial cells have undergone a partial MET process, due to repression of *ZEB1*, resulting in epithelial-like characteristics and morphology, which compromise the barrier function and integrity of the corneal endothelium. Consistent with this hypothesis, the *GRHL2*-positive patient endothelial tissue had positive regions of E-CAD expression, and N-CAD was clearly absent from certain areas of the endothelium, however positive staining for N-CAD was observed in some cells and endothelial markers were expressed in the entire tissue.

The congenital or early-onset nature of the disease in some probands indicates that PPCD is, at least in part, a developmental disorder. Supporting this hypothesis, the phenotypic spectrum of PPCD encompasses defects of the anterior chamber in some probands including iris flocculi and iridocorneal adhesions. Developmentally, the corneal endothelium originates from neural crest cells which migrate into the cornea during foetal development to form the endothelial monolayer (Hay, 1980) (section 1.3.3). These corneal endothelial cells share a common embryonic origin with several structures which make up the anterior segment, including the iris. The important role of OVOL2 and GRHL2 in neural development is exemplified by knockout mice generated for these genes. *OVOL2* knockout mice are embryonic lethal at E10.5, and display a range of defects including failure of neural tube closure, abnormal gut morphology and severe heart defects. At the neural tube, *Ovol2* -/- mice had a lateral shift in the border between the neuroectoderm and the surface ectoderm which appeared to be due to neuroectodermal expansion. Furthermore, the border between these two epithelial tissues was less morphologically distinct than observed in wild type (WT) mice, with knockout mice having a gradual change from neuroectoderm to surface ectoderm rather than a distinct border in wild type mice. *Ovol2* -/- mice also failed to develop an optic eminence or optic pits; an early stage of eye development (Mackay *et al.*, 2006). *Grhl2* knockout mice are embryonic lethal at E11.5 and, like *Ovol2* -/- mice, display

a range of developmental defects in multiple tissues including split face malformation and incomplete neural tube closure (Rifat *et al.*, 2010).

In summary, there is an evolving convergent pathological mechanism for dominant PPCD that involves the regulatory loop of transcription factors involved in EMT/MET (Figure 4.32). This supports the candidacy of *GRHL2*, and in the large Czech family there is enough genetic evidence to support the conclusion that the *GRHL2* mutation is causative. However, the *GRHL2* variants in the other families require further experimental validation. This could be achieved using a luciferase assay to determine the effect of the *GRHL2* non-coding variants on the transcript levels of *GRHL2* with the hypothesis that pathogenic variants would cause an increase in *GRHL2* transcription.

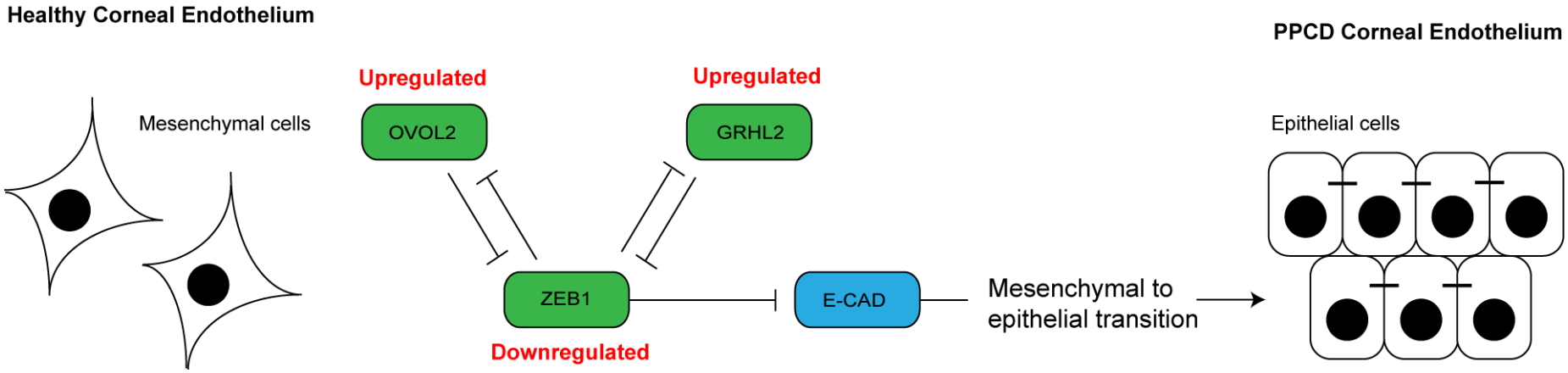


Figure 4.32 Summary of transcription factor network disrupted in posterior polymorphous corneal dystrophy leading to mesenchymal-to-epithelial transition. Dysregulation of this network is achieved by mutations associated with the upregulation of *OVOL2* or *GRHL2* expression or haploinsufficiency of *ZEB1*. OVOL2 and ZEB1 are transcriptional repressors of *ZEB1*, and are themselves transcriptionally repressed by ZEB1. As ZEB1 is also a negative transcriptional repressor of E-cadherin, downregulation of ZEB1 results in the activation of mesenchymal-to-epithelial transition and leading to epithelialisation of the corneal endothelium.

4.10.6 Unsolved PPCD

Following genetic investigation of the PPCD cohort with a combination of targeted sequencing, several individuals from both the typical PPCD phenotype and the band-type PPCD subtype, remained genetically unsolved. Unlike anterior corneal dystrophies, which show a distinct clinical appearance, endothelial dystrophies have a range of overlapping clinical characteristics, therefore any proband who was negative for mutations in PPCD genes was also assessed for mutations in genes causing other endothelial dystrophies. As the age of onset for PPCD is the first or second decade, early-onset FECD associated genes were considered candidates. To investigate this possibility, a sub-set of the unsolved PPCD cohort was analysed by WES. All known genes associated with FECD, regardless of age of onset, were included in the corneal gene WES filtering strategy. Although some candidate variants were identified in some probands, there was no compelling evidence for a role of previously identified genes. In one family with PPCD (F151) filtering of WES data revealed variants associated with two endothelial dystrophies – a heterozygous in-frame deletion in *OVOL2*, and a heterozygous missense mutation in *SLC4A11*, however, both variants were excluded by segregation analysis. It is possible that some cases with XECD may be misdiagnosed with PPCD, particularly as the endothelial changes are similar, but as the genetic cause of this dystrophy has not been solved however as the genetic cause of this dystrophy has not been resolved, it is not currently possible to exclude this hypothesis.

The WES data of all six unsolved PPCD probands (4 with typical PPCD, 2 with band-type PPCD) was then combined to facilitate identification of common shared gene(s). Therefore, within this analysis, the possibility of a shared gene causing both typical PPCD and band-type PPCD was not excluded. The filtering strategy focussed on non-synonymous changes in shared genes between probands within the cohort. Only one gene was found to harbour rare variants in more than 2 probands – *PCDH1*; three missense variants were identified in three unrelated PPCD families. Two changes were novel, whereas one was extremely rare. However, assessing the potential pathogenicity of these variants is more difficult in cases where only small pedigrees are available. The putative *PCDH1* mutation in F154 segregated in this small pedigree consisting of an affected mother and son, however F156 had no

known family history and no additional samples were available. If there is another PPCD gene, identification is currently hindered by the lack of potential for linkage analysis, which requires a large pedigree.

Future research should focus on WES or WGS analysis of the additional 5 unsolved PPCD probands with typical PPCD and 11 unsolved PPCD probands with band-type PPCD who were negative for targeted sequencing of mutations in *ZEB1*, *OVOL2* and *GRHL2*. An alternative hypothesis is that the mutations in the unsolved cohort are located in a region outside of the pre-screened region, which is particularly relevant for *OVOL2* and *GRHL2*. Given that the mutations are predicted to affect a promoter and/or intronic sequences, it is plausible that mutations in different regulatory regions could affect the expression of either *OVOL2* or *GRHL2*. These could be detected by WGS analysis. Unsolved probands should also be analysed for *ZEB1* deletions by qPCR or SNP array copy number analysis, as this mechanism remains a possibility for a sub-set of the unsolved cohort.

A non-inherited disorder is also possible, particularly for those probands of the PPCD-band subtype without a family history and/or who are unilaterally affected. The most likely alternative diagnosis is iridocorneal endothelial (ICE) syndrome; a corneal degeneration of unknown aetiology which affects the corneal endothelium (Sacchetti *et al.*, 2015). The clinical phenotype associated with ICE syndrome overlaps with that of endothelial corneal dystrophies although it is almost always unilateral. The appearance of the endothelium has a 'beaten silver' appearance; however, corneal endothelial cells associated with ICE syndrome are similar in behaviour and morphology to those of PPCD with endothelial metaplasia, abnormal morphology and epithelialisation. Loss of the corneal endothelial function eventually results in oedema in individuals with ICE syndrome. Furthermore, ICE is associated with other abnormalities of the anterior segment including the anterior chamber, iris, and pupil, with an increased risk of developing secondary glaucoma. The biggest differences between typical PPCD and ICE syndrome is that typical PPCD is predominantly bilateral whereas ICE is unilateral, and typical PPCD is familial whereas ICE is sporadic. However, as we have shown in our cohort, band-type PPCD is predominantly sporadic and unilateral and therefore cannot be distinguished from ICE in this manner. This is consistent with a previous study that demonstrated

that 19/24 cases of band-type PPCD were unilateral in comparison to only 1/19 with typical PPCD (Laganowski *et al.*, 1991). Furthermore, even in cases where a diagnosis of PPCD is genetically confirmed, there may not be a positive family history due to *de novo* occurrence or non-penetrance. Therefore, ICE should be considered as a possible clinical diagnosis for probands who are unilaterally affected and with an absence of family history, although these features alone do not exclude PPCD as a diagnosis.

4.10.7 Summary of the genetic causes of PPCD

The relative proportion of PPCD cases that are caused by mutations in *ZEB1* vary by population. Prior to the identification of *ZEB1* deletions, we proposed that PPCD3 accounted for ~ 71% of British cases, ~ 30% of cases in the USA, and ~ 20% in the Czech Republic. The differences are partially explained by the high prevalence of PPCD1 in the Czech Republic due the presence of a founder mutation in this population (Liskova *et al.*, 2012). However, following the identification of *ZEB1* deletions as a cause of PPCD3, it is likely that these figures are underestimates.

Based on currently published data and this study, within the British cohort, 8 probands harbour mutations in *ZEB1*, 3 have mutations in the *OVOL2* promoter and 3 have potential mutations in *GRHL2*. In comparison, 10 PPCD cases in the Czech Republic were caused by mutations in in *ZEB1*, 16 PPCD families caused by a shared *OVOL2* founder mutation and four families harbouring a shared mutation in *GRHL2* (Liskova, Personal Communication). None of the MEH probands carried a mutation in *COL8A2*, and it appears that this gene, although causative of early-onset FECD, is not responsible for a significant proportion of PPCD cases.

It has previously been reported that mutations in *VSX1* cause PPCD and keratoconus (Héon *et al.*, 2002). However, the *VSX1* p.(Gly160Asp) variant, which was proposed to cause PPCD is present in ExAC in a frequency of 0.002 (224/106,272), which although rare, is more frequent than would be expected for a dominantly inherited disease (Héon *et al.*, 2002). Although *VSX1* was not included in the PPCD pre-screening, *VSX1* was included as a candidate gene in the list of genes associated with corneal dystrophies or corneal syndromes for which all WES data was filtered. Since

the original identification of this gene as causative of PPCD, no follow up papers have confirmed this association and identified further *VSX1* mutations causing PPCD. Given that the PPCD1/CHED1 locus is now explained by mutations in *OVOL2*, it is highly likely that any variant identified in *VSX1* in PPCD cases was a rare benign polymorphism and *VSX1* should not now be considered a causative gene of PPCD.

4.10.8 Genotype-phenotype correlation in PPCD

Prior to the genetic analysis in this study, two clinical sub-types of PPCD were described; a classic sub-type with typical endothelial opacities and vesicles characteristic of PPCD, and a band-type PPCD subtype which appeared to be less severe and for which the genetic basis had not been resolved. With the identification of the genetic basis of PPCD in 21 families in this study, and comparison of their clinical features, insights can be made pertaining to the characteristic features of each underlying genetic cause.

The commonly described and consistent feature of PPCD3, caused by coding mutations or deletions of *ZEB1*, is the occurrence of abnormal corneal steepening, and this is likely to be the most useful clinical feature to differentiate between PPCD3 and other genetic causes (Liskova *et al.*, 2010, 2013). All mutation-positive probands in the MEH cohort presented with classic PPCD, with no probands of the band-type PPCD positive for mutations in *ZEB1*. Inguinal or abdominal hernia has been reported in 17 individuals harbouring *ZEB1* mutations, including the proband of F128 in our cohort, and hydrocele in five cases (Krafchak *et al.*, 2005; Aldave *et al.*, 2007; Lai *et al.*, 2010; Liskova *et al.*, 2013). There are few reports of bone spurs, kneecap dislocations and extra vertebrae, but the association is not compelling and none were identified in our cohort (Krafchak *et al.*, 2005). Agenesis of the corpus callosum has been reported in two probands with PPCD3, however as brain imaging is required to detect this feature there may be additional undetected cases (Jang *et al.*, 2014).

The four probands who were positive for *OVOL2* mutations differed in their clinical features and severity of disease, and there was also evidence of inter-familial

variability. The British CHED1 family (F140) were the most severely affected, with the earliest onset, either at birth or within the first year of life, which explains why historically they were given a diagnosis of CHED over PPCD, which is typically milder. The proband of F142 was also severely affected, eventually resulting in enucleation of her left eye following multiple failed grafts. The prior diagnosis of anterior segment dysgenesis in this proband is indicative of other abnormalities of the anterior chamber in addition to corneal opacity. In comparison, the Czech founder families (F141) are affected in the first or second decade of life and only a third of patients required surgery. The examined individuals from F143 were also mildly affected and have not had corneal transplants. In this family, there were some abnormalities of the anterior segment such as pupillary distortion due to peripheral adhesions.

The band-type clinical subtype of PPCD poses a number of unanswered questions. Clinically, these probands rarely overlapped with the clinical symptoms of a proband with typical PPCD and only one proband was severely affected enough to require surgery. The remaining cohort was asymptomatic and had only mild clinical signs. Some individuals only had unilateral corneal changes. One proband was positive for a family history of disease; an affected son and mother who were recruited and subsequently analysed by WES, with the remainder of the cohort sporadically affected. None of the probands with band-type PPCD were found to harbour mutations in *ZEB1*, *OVOL2* or *GRHL2*. This poses the question of whether these individuals have an inherited form of mild band-type PPCD or an environmental phenocopy of the disease, such as ICE. PPCD is typically described as a mild, non-progressive disease; however, this is likely to be a misconception based primarily on band-type PPCD, which may not be necessarily related to other forms of PPCD.

4.10.9 CHED2

Recruitment of autosomal recessive CHED was not prioritised for this study; however, two individuals were recruited with a clinical diagnosis of CHED2, now referred to as CHED. Clinically, the phenotype was consistent with that previously described, however nystagmus was absent from our patients.

Causative mutations in *SLC4A11* are loss of function and include missense, nonsense and frameshift mutations as well as small deletions and indels. There is no evidence of additional locus heterogeneity for CHED, however, in one case in our cohort only a single mutation was identified on screening *SLC4A11* coding exons, despite screening alternatively spliced exons 1 and 2 from the three annotated transcripts. Given the single heterozygous variant, it is likely that there is a disruption of the second allele that has not been identified by direct sequencing. Detection of only a single heterozygous *SLC4A11* mutation or absence of any mutations has also been reported by other studies (Sultana *et al.*, 2007). This could be due to the second heterozygous mutation being present in an unannotated exon, or a partial gene deletion which spans beyond the primer boundaries.

Chapter 5: Syndromes and phenocopies of corneal dystrophies

5.1 Introduction

The clinical appearance of some corneal dystrophy diagnoses overlap, complicating the process of making a diagnosis based only on clinical signs; however, this can be resolved by identification of the underlying genetic cause. There are also several inherited syndromes that are associated with a corneal phenotype that has a similar clinical presentation to a corneal dystrophy. These differential diagnoses need to be considered in cases where a patient with corneal opacity is negative for a mutation in the associated corneal dystrophy gene(s). Alternatively, environmental and age-related factors can cause acquired corneal opacities (corneal degeneration) which may mimic the appearance of a corneal dystrophy and may explain the corneal symptoms in the absence of a mutation. While some misdiagnoses of corneal dystrophies have been reported in the literature, there are currently no estimates on the number of cases initially diagnosed with corneal dystrophy that have an alternative genetic or environmental cause. This chapter describes the genetic investigation of individuals who were recruited to the study with a potential corneal dystrophy, who were subsequently identified to have an alternative cause for their corneal opacity, following genetic and clinical investigation. Additionally, corneal phenotypes that have previously been clinically described, but have an unknown aetiology, were investigated. For these cases, the aim was to confirm or exclude whether these corneal phenotypes are expansions of the clinical phenotype of a known corneal dystrophy or syndrome, or whether they have a novel genetic cause of disease.

5.1.1 Meretoja syndrome

Meretoja syndrome has previously been described as LCD, type 2 and was included in the first edition of the IC3D classification of corneal dystrophies, although this diagnosis has been excluded from the most recent edition given that it does not meet the criteria to be considered a corneal dystrophy (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015). Unlike most corneal dystrophies, which exclusively affect the cornea,

Meretoja syndrome is characterised by systemic amyloid deposition, primarily in the corneal stroma, skin, cranial nerves, kidneys and the heart, causing a range of clinical symptoms. The corneal amyloid deposition causes the appearance of lattice lines which are phenotypically similar to those of *TGFBI*-associated dystrophy LCD, and cause visual impairment (Gorevic *et al.*, 1991) (Figure 5.1 A).The ocular symptoms typically develop first, in the third or fourth decade of life, often resulting in clinical misdiagnosis. In later stages of the disease affected individuals develop a mask-like facial appearance due to deposition of amyloid in cranial nerves (Figure 5.1 B), loose skin, and disrupted renal and cardiac function due to amyloid deposition in other tissues (Kiuru, 1992). It is therefore important that patients are referred by their ophthalmologists for monitoring and treatment of systemic disease.

Meretoja syndrome is caused by mutations in *GSN* (gelsolin), a 17 exon gene located on chr9q33.2 (Levy *et al.*, 1990; Maury *et al.*, 1990). To date, only two mutations have been identified to cause Meretoja syndrome, both of which affect the same amino acid residue; Asp-214 in exon 4. Meretoja syndrome is common in Finland with all reported cases from this region associated with a *GSN* c.640G>A; p.(Asp214Asn) mutation, likely due to a founder effect (Levy *et al.*, 1990). However, the same mutation has been described in an American patient of Irish descent and a Japanese family, indicating that the mutation may also occur *de novo* (Gorevic *et al.*, 1991; Sunada *et al.*, 1993). An alternative pathogenic mutation, c.640G>T; p.(Asp214Tyr), was described initially in one affected family of Danish origin and another of Czech origin. It was concluded that this is likely to have occurred independently in both families as they do not share a common haplotype (De La Chapelle *et al.*, 1992). Recently, 2 additional *GSN* pathogenic mutations [c.580G>A; p.(Gly194Arg) and c.633C>A; p.(Asn211Lys)] have been described that are associated specifically with renal dysfunction caused by amyloid deposition exclusively in the kidneys. There was no evidence of amyloid deposited in other tissues and no corneal phenotype (Sethi, Theis and Quint, 2013; Efebera *et al.*, 2014).

There are at least two isoforms of the GSN protein; a cytoplasmic form which is an actin binding protein involved in cytoskeleton remodeling, and a longer, secreted isoform. Mutations of amino acid residue Asp-214 result in abnormal cleavage of the

secreted isoform, potentially via disruption of a disulphide bond, which results in the secretion of the 83 kDa full length protein and a 68 kDa C-terminal fragment (Paunio *et al.*, 1994, 1998). A second cleavage event of the C-terminal fragment produces a 60 kDa protein fragment containing the amyloidogenic residues 173-243.

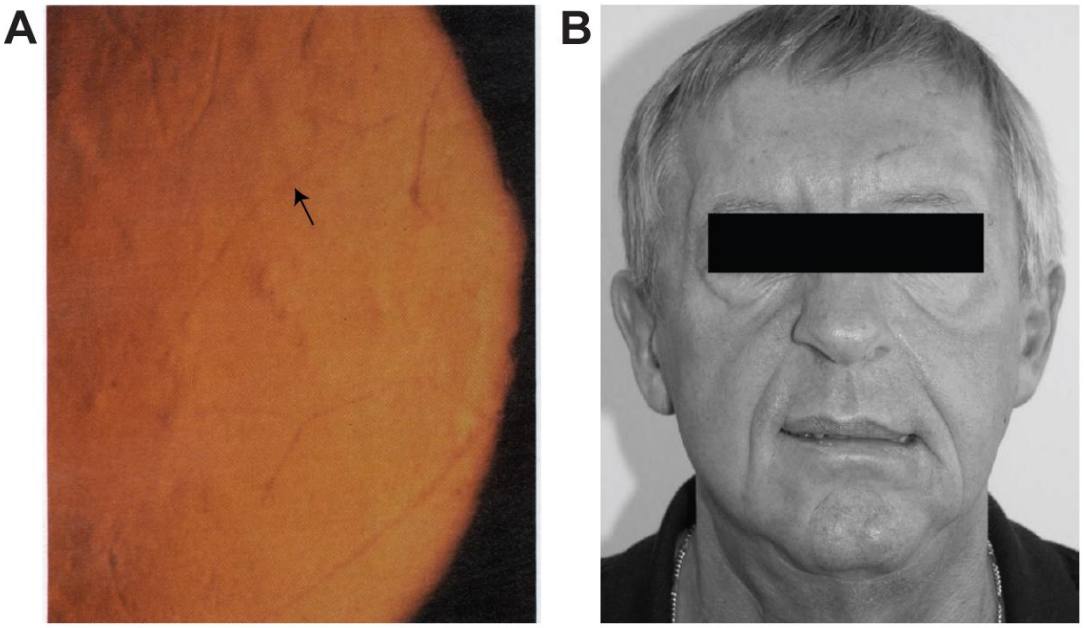


Figure 5.1 Typical clinical appearance of Meretoja syndrome. (A) Cornea in retroillumination shows thin, branching lattice lines (arrow), similar to those observed for *TGFBI*-associated corneal dystrophies. **(B)** Mask-like appearance due to facial nerve palsy from the deposition of amyloid in facial nerves and skin. Image A from (Lüttmann *et al.*, 2010) and B from (Gorevic *et al.*, 1991).

5.1.2 Polymorphic corneal degeneration /amyloidosis

Polymorphic corneal degeneration (PCD) is a corneal phenotype that usually occurs sporadically in individuals aged fifty years or over. It is characterised by a combination of bilateral deep stromal punctate opacities which appear grey-white in direct illumination and transparent in retroillumination, and filamentous lines due to an accumulation of amyloid (Figure 5.2). These are do not usually affect vision and are asymptomatic (Thomsitt and Bron, 1975). Although it was initially considered to be a corneal degeneration rather than an inherited corneal dystrophy, an indistinguishable clinical and histological phenotype was subsequently noted in individuals who were positive for a family history of corneal opacities. In these

families, the disorder was classified as polymorphic corneal amyloidosis (PCA) (Eifrig *et al.*, 2004).

Four families have been reported with PCA, all of which are have the same c.1637; p.(Ala546Asp) mutation in exon 12 of *TGFBI* (Eifrig *et al.*, 2004; Correa-Gomez, Villalvazo-Cordero and Zenteno, 2007; Zenteno *et al.*, 2009). The *TGFBI* mutation, c.1580T>G; p.(Leu527Arg), a founder mutation in Japan, also causes a similar clinical phenotype. This data suggest that the PCD/PCA phenotype is an atypical form of *TGFBI*-associated LCD, however screening *TGFBI* in a cohort of 8 patients with clinical appearance of PCD/PCA, which were described as clinically and histopathologically indistinguishable to those cases with *TGFBI* mutations, did not identify any pathogenic *TGFBI* mutations (Aldave *et al.*, 2006). It is not known whether there are additional genetic aetiologies for this corneal phenotype, or whether the *TGFBI*-negative PCD patients represent those with an environmental age-related aetiology (degeneration).

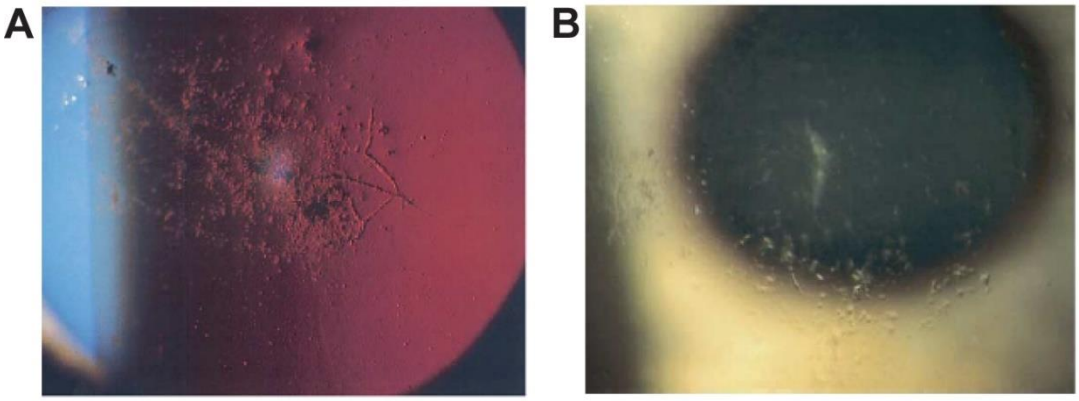


Figure 5.2 Characteristic clinical appearance of polymorphic corneal amyloidosis. (A) Cornea in retroillumination shows a combination of line and dot-like opacities. **(B)** Cornea in direct illumination showing the same combination of opacities. Images from (Eifrig *et al.*, 2004).

5.1.3 Paraproteinemic keratopathy

Paraproteinemic keratopathy refers to the deposition of immunoglobulin proteins in the cornea. Clinically, this can cause corneal clouding, corneal verticillata, small reticular or granular opacities, crystalline deposits and/or lattice lines and therefore can be misdiagnosed as LCD, GCD, SCD or GDLD (Milman *et al.*, 2015) (Figure

5.3). The most common cause of paraproteinemic keratopathy is monoclonal gammopathy of unknown significance (MGUS). MGUS is a diagnosis given to individuals with serum monoclonal protein at a concentration greater than 3 g/dL. It is asymptomatic and does not require medical treatment, but individuals are regularly monitored as transformation may occur to multiple myeloma. MGUS affects an estimated ~ 2% of individuals aged over 50 and ~ 3% over 70 years of age, with ~ 1% of sufferers per year developing multiple myeloma (Kyle *et al.*, 2002). Other causes of paraproteinemic keratopathy include lymphoma and cryoglobulinemia.

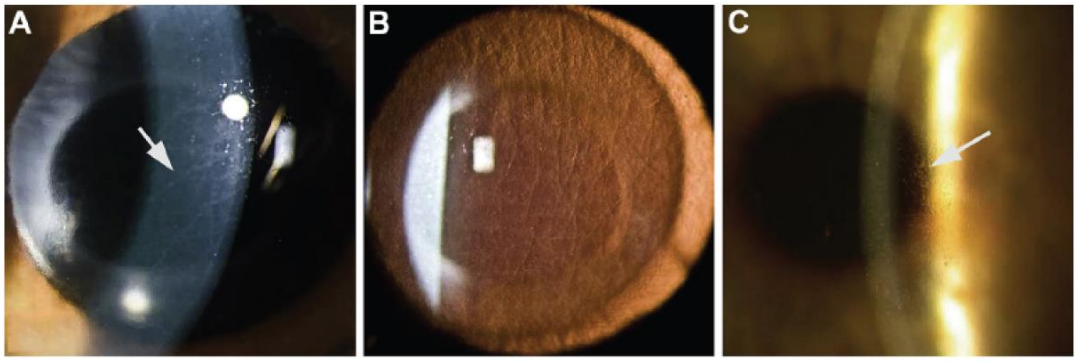


Figure 5.3 Clinical appearance of corneas with paraproteinemic keratopathy. (A) Direct illumination showing appearance of straight overlapping lattice lines, indicated with an arrow. **(B)** Retroillumination showing presence of lattice lines. **(C)** Corneal opacities in a different affected individual indicated with an arrow, appear to be crystalline in appearance. Images from (Milman *et al.*, 2015).

5.1.4 Climatic droplet keratopathy

Climatic droplet keratopathy (CDK) is a corneal degeneration characterised by corneal haze caused by an accumulation of translucent droplet opacities in the superficial corneal stroma. Opacities occur initially in the periphery and are deposited centrally as the disease progresses (Figure 5.4). It was initially described in Somalia and the Persian Gulf, but was subsequently described in a range of populations under different names including Bietti's band-shaped nodular dystrophy, Labrador keratopathy, spheroidal degeneration, chronic actinic keratopathy, oil droplet degeneration, elastoid degeneration and keratinoid corneal degeneration. CDK predominantly affects males over 40 years with a history of working outdoors. A range of causative factors have been proposed including microtrauma from dust

particles, evaporation due to dry heat, and poor nutrition. Although these may have an influence, UV radiation is considered the most likely primary cause (Gray, Johnson and Freedman, 1992; Serra *et al.*, 2015).

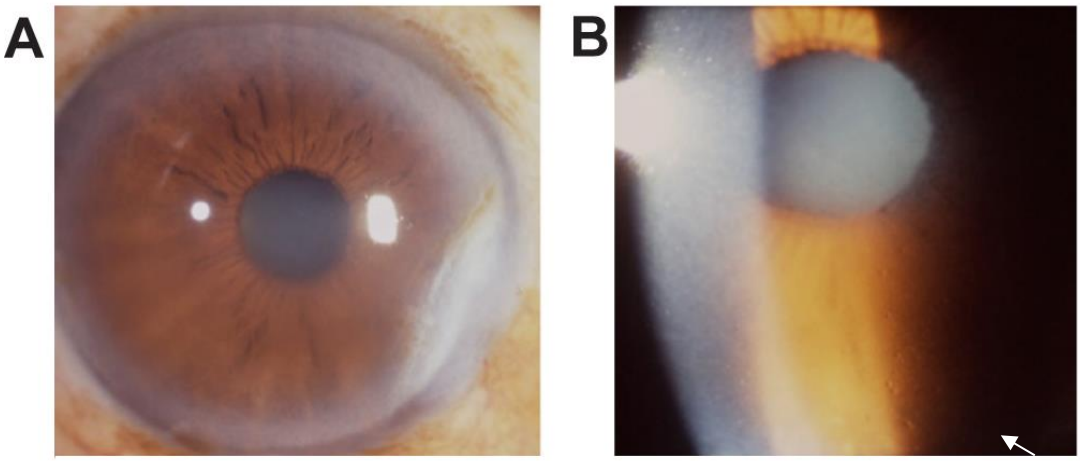


Figure 5.4 Clinical appearance of climatic droplet keratopathy. (A) Haziness occurs predominantly in the periphery of cornea. **(B)** Central cornea of same eye begins to show evidence of microdroplet deposition (arrow). Images taken from (Serra *et al.*, 2015).

5.1.5 Pre-Descemet corneal dystrophy and X-linked ichthyosis

Pre-Descemet corneal dystrophy (PDCD) is a rare corneal dystrophy with an unknown genetic basis. It is asymptomatic and characterised by fine grey-white pleomorphic opacities in the posterior stroma, located anteriorly to Descemet's membrane (Ye *et al.*, 2006). These white dot-like opacities are visible on slit lamp examination (Figure 5.5 A). Confocal *in vivo* microscopy reveals multiple highly reflective dot-like opacities localised to the posterior stroma, both intracellular and extracellular. These hyper reflective dots are larger and more posterior than the intracellular opacities observed in a similar corneal dystrophy, FCD (Kontadakis *et al.*, 2014). FCD can also be distinguished from PDCD by the appearance of opacities in the full thickness of the stroma in the former as opposed to the posterior third of the stroma in PDCD (Holopainen, Moilanen and Tervo, 2003). Corneal opacities indistinguishable from that of PDCD also occur in 10 - 50% of cases of the inherited syndrome, X-linked ichthyosis, which is characterised by skin flaking and scaling (Figure 5.5 B) (Fernandes, Janniger and Schwartz, 2010).

Non-syndromic PDCD cases are of unknown aetiology with both sporadic and familial cases previously described. Familial cases of PDCD are rarer, although this may be related to the asymptomatic nature of the dystrophy. X-linked ichthyosis is caused by a deletion of the *STS* (steroid sulphatase) gene on chrXp22.3 in ~ 90% of cases, or rarely by missense mutations in this gene (Hung *et al.*, 2013).

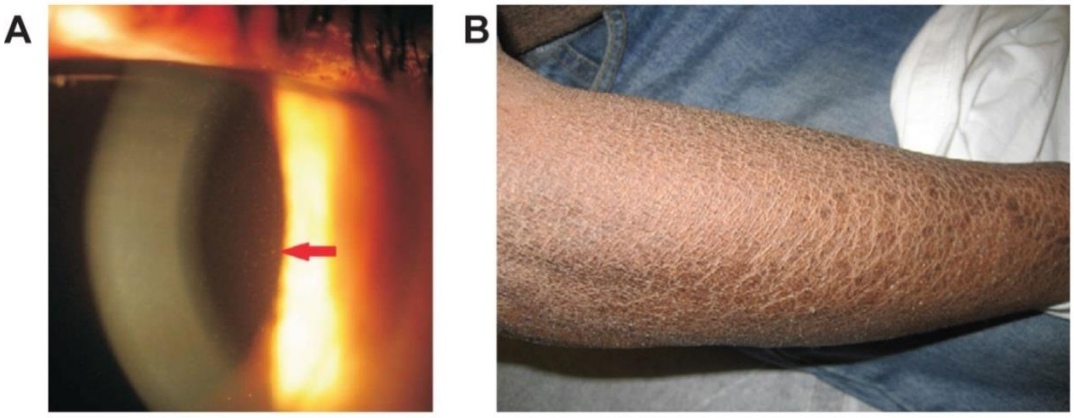


Figure 5.5 Typical clinical image of Pre-Descemet corneal dystrophy and X-linked ichthyosis. (A) Slit lamp image with indirect illumination shows the presence of fine white opacities, labelled with an arrow. **(B)** Skin scaling symptoms typical of X-linked ichthyosis. Image A from (Kontadakis *et al.*, 2014) and B from (Fernandes, Janniger and Schwartz, 2010).

5.1.6 Corneal clouding

Corneal clouding is a non-specific corneal sign, observed in corneal dystrophies, inherited corneal syndromes and corneal degenerations, many of which are caused by dysregulation of lipid metabolism. SCD is the only known corneal dystrophy to cause lipid and cholesterol deposition in the cornea. SCD corneas contain ten-fold more cholesterol and 5-fold more phospholipids than unaffected corneas. There is also a higher proportion of unesterified cholesterol in SCD corneas. Specifically, apolipoprotein components of HDL, including apoA-I, apoA-II and apoE, but not those of low density lipoprotein (LDL), are present in affected corneal buttons (Gaynor *et al.*, 1996).

Interestingly, there are a range of inherited syndromes involving dysregulation of HDL-cholesterol metabolism that are associated with corneal clouding and several

other symptoms. An example is Tangier disease, a recessive condition which is associated with low plasma levels of cholesterol, apoA-I and apoA-II and deposition of cholesterol esters in a variety of tissues. Clinically, it results in corneal clouding, yellow tonsils, splenomegaly, lymphadenopathy, hepatomegaly and peripheral neuropathy (Pressly *et al.*, 1987). Tangier disease is caused by biallelic mutations in *ABCA1* (Brooks-Wilson *et al.*, 1999). Mutations in *LCAT* cause lecithin-cholesterol acetyltransferase deficiency characterised by corneal opacity and reduced HDL cholesterol. Some individuals demonstrate other systemic symptoms including anaemia and renal failure. Where only corneal opacity is present, the disorder is also known as fish eye disease (Funke *et al.*, 1991). These syndromes highlight that disruption of lipid metabolism; particularly HDL-cholesterol, can result in corneal opacity and clouding.

ABCA1 and LCAT are essential to the biosynthesis of mature HDL-cholesterol, of which ApoA-I is a major protein component. Unlipified apoA-I is secreted by the liver and intestines and lipidated by ATP binding cassette transporter A1 (encoded by *ABCA1*), which is responsible for the efflux of cholesterol from cells to lipid-poor apoA-I. As unlipidated apoA-I is catabolised, the loss of *ABCA1* activity in Tangier disease is responsible for low apo-A-I levels and therefore low HDL-cholesterol levels. Lecithin-cholesterol acetyltransferase (encoded by *LCAT*) catalyses the formation of cholesterol esters, which are incorporated into the hydrophobic core of mature HDL particles (Lewis and Rader, 2005).

5.2 Revised diagnoses of suspected lattice corneal dystrophy

5.2.1 Meretoja Syndrome

Proband F175: II:1 was recruited to the study at 71 years of age with a clinical diagnosis of LCD. He had been treated previously with PK in his left eye. In his un-operated right cornea, thin branching lattice lines were observed with elevated sub-epithelial focal lesions. The intervening corneal stroma was clear (Figure 5.5 A). He had an affected daughter, who was not clinically examined, and a potentially affected father who was deceased, consistent with autosomal dominant inheritance (Figure

5.5 C). Sequence analysis of all *TGFBI* coding exons and splice sites revealed no potentially pathogenic mutations (section 2.2; Appendix A).

The proband of family #176 (II:1) was a 64 year old female of Arabic ethnicity, recruited with symptoms of dry eyes and poor vision, and had recently undergone bilateral upper and lower lid blepharoplasties. Clinical examination revealed bilateral nuclear cataracts and linear corneal opacities. Both LCD and Meretoja syndrome were considered as potential causes for the corneal changes. There was no known family history (Figure 5.6 D). Screening *TGFBI* (exons 4, 11, 12, 13, 14 and 16) were negative for potentially causative changes (section 2.2; Appendix A).

The proband of family #177 (I:2) was recruited with a diagnosis of LCD, having previously had bilateral penetrating keratoplasties. Thin lattice lines were visible in both corneas located peripheral to the grafts. He had facial weakness and had bilateral partial tarsorrhaphies for corneal exposure. The pedigree was consistent with an autosomal dominant pattern of inheritance with two children and one grandchild affected (Figure 5.6 E). Screening *TGFBI* (exons 4, 11, 12, 13, 14 and 16) were negative for potentially causative changes (section 2.2; Appendix A).

Given the lack of mutations in *TGFBI*, exon 4 of the *GSN* gene was screened in all three probands; this exon contains the Asp-214 residue which is the site of both previously described Meretoja syndrome mutations, although exon 4 does not contain the two *GSN* mutations associated with kidney specific amyloidosis (section 2.2; Appendix A). Sequence analysis resulted in the identification of a heterozygous *GSN* c.640G>A; p.(Asp214Asn) mutation in all three probands (Figure 5.6 B). This mutation has previously been reported in multiple families with Meretoja syndrome (De La Chapelle *et al.*, 1992). SIFT and Polyphen2 scores were 'Damaging' (0.02) and 'Probably Damaging' (1), respectively and alignment of orthologous sequences shows that the Asp-214 residue is conserved in different species (Figure 5.6 F). All three patients and their relatives were offered referral to the UK National Amyloid Centre (Royal Free Hospital) for assessment.

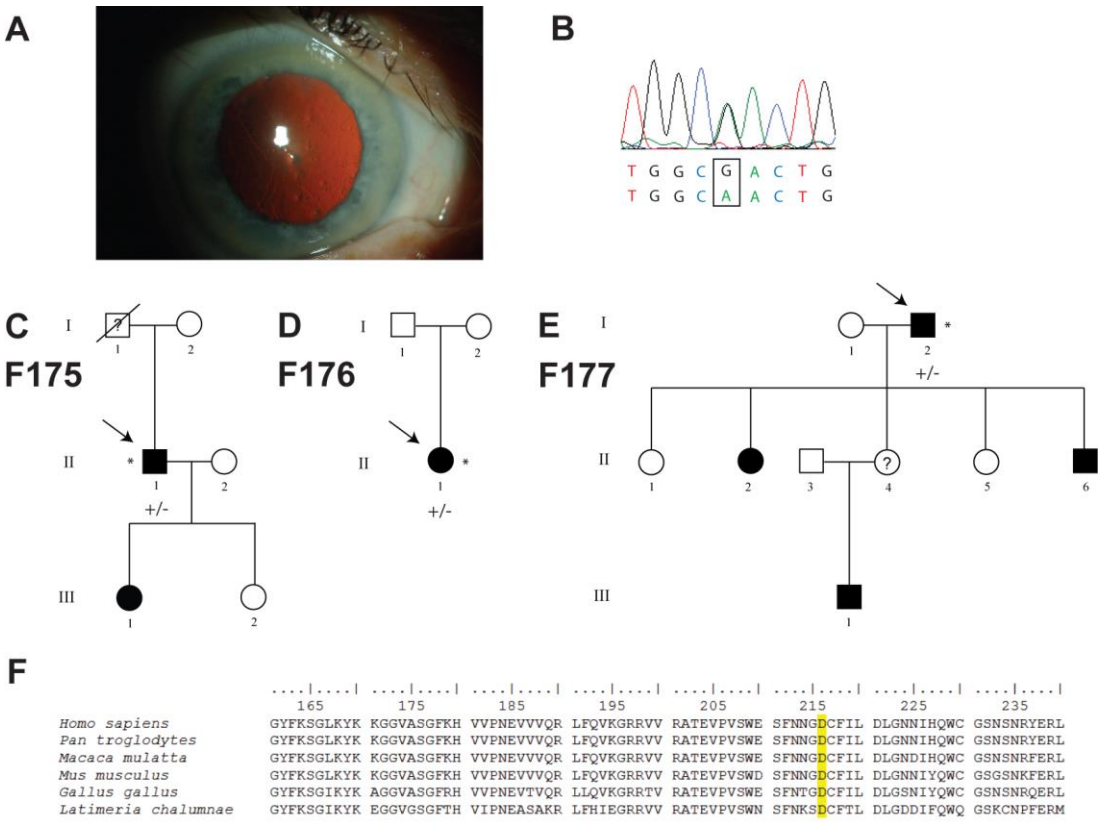


Figure 5.6 Meretoja syndrome due to a *GSN*: c.640G>A; p.(Asp214Asn) mutation in three families clinically diagnosed with lattice corneal dystrophy. +/- indicates heterozygous for the mutation. * indicates individual has been examined and a blood sample taken. The proband is shown with an arrow. **(A)** Corneal phenotype of proband F175: II:1 in retroillumination, showing thin lattice lines. **(B)** Sequence electropherogram showing presence of a c.640G>A; p.(Asp214Asn) mutation. **(C)** Pedigree structure of F175. **(D)** F176. **(E)** F177. **(F)** Alignment with ClustalW, showing conservation of the Asp-214 residue.

5.2.2 Paraproteinemic keratopathy

The proband of family #178 was referred to MEH due to a late-onset deterioration of vision. Clinically, there was an unusual reticular pattern of predominantly anterior stromal haze. Lattice structures were observed in the corneal mid-periphery which could indicate a diagnosis of LCD or Meretoja syndrome (Figure 5.7 A and B). Screening revealed that the proband was negative for mutations in both *TGFBI* and *GSN* exon 4 (section 2.2; Appendix A). He was negative for a family history of corneal disease (Figure 5.7 C). Due to the lack of family history and negative result of genetic screening, paraproteinemic keratopathy was suspected as a potential cause of the corneal opacities. Serum electrophoresis was performed that showed a

monoclonal spike and resulted in a diagnosis of paraproteinemic keratopathy caused by MGUS.

The proband of #179 was referred for ophthalmological assessment following the identification of corneal changes by his optometrist in 2002 and again in 2007. When clinically examined he demonstrated an unusual reticular pattern of small granular structures predominantly in the anterior stroma, similar to the opacities observed in proband F178: II:1. Screening *TGFBI* and exon 4 of *GSN* was negative for potentially pathogenic mutations, therefore excluding the most likely genetic causes of his condition (section 2.2; Appendix A). There was no family history of corneal disease (Figure 5.7 D). The proband had been diagnosed with MGUS 3 years prior to the examination, which indicated that paraproteinemic keratopathy was the likely cause of the corneal opacities, therefore serum electrophoresis was not repeated.

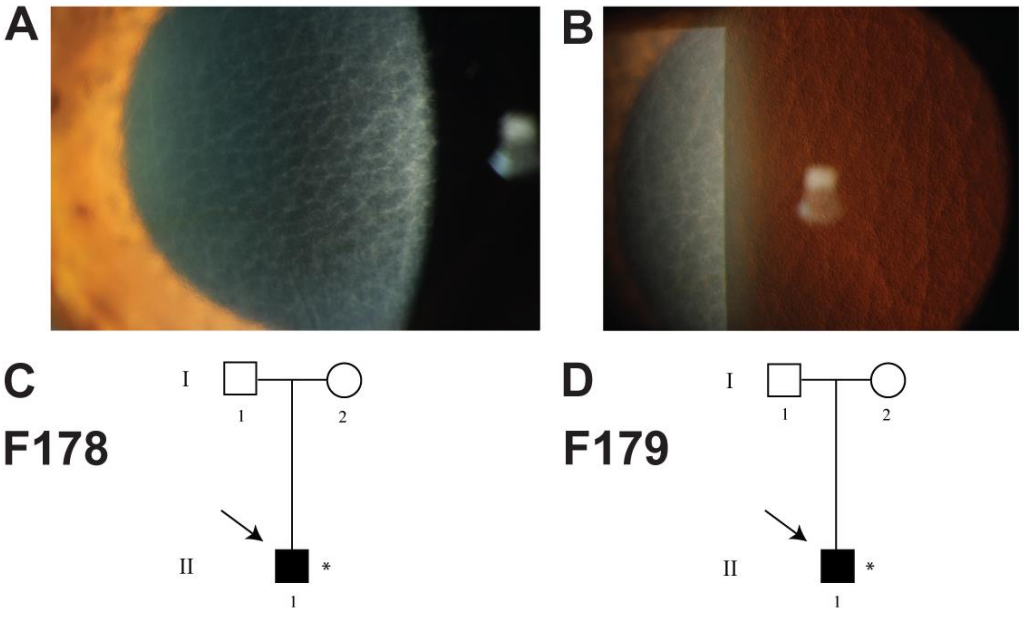


Figure 5.7 Two patients with paraproteinemic keratopathy, previously diagnosed with lattice corneal dystrophy. (A) Corneal phenotype of F178:II:1 in direct illumination. **(B)** Corneal phenotype of F178:II:1 in retroillumination. **(C)** Lack of family history of corneal opacity in family #178. **(D)** Lack of family history in family #179.

5.2.3 Climatic droplet keratopathy

The proband of family #180 (F180: II:1) was suspected to have LCD, predominantly in his left eye, identified during a routine follow up examination following cataract surgery in his seventies (Figure 5.8 A). His family history was negative for corneal disease, however no additional family members were available for clinical examination (Figure 5.8 B). CKD was also considered to be a potential diagnosis due to his age and prior history of working outdoors in Somalia. Screening the *TGFBI* gene revealed a homozygous missense mutation in exon 12: c.1645G>A; p.(Ala549Thr) (Figure 5.8 C) (section 2.2; Appendix A). This variant is rare, with a frequency of 0.00038 (45/118842) in the ExAC database and is predicted to be 'Damaging' (0.04) and 'Probably damaging' (0.938) by SIFT and Polyphen2, respectively. However, there is a lack of control genomic and exome data for individuals of African ethnicity therefore the frequency specifically in Somalia is not known. An alignment of orthologous sequences revealed that the variant affects a conserved residue, and is only 6 amino acid residues from the hot-spot residue Arg-555 (Figure 5.8 D). Mutations affecting residues 550 and 547 have also been reported to cause GCD and LCD, respectively. The contribution of the homozygous *TGFBI* variant to the corneal phenotype is unknown.

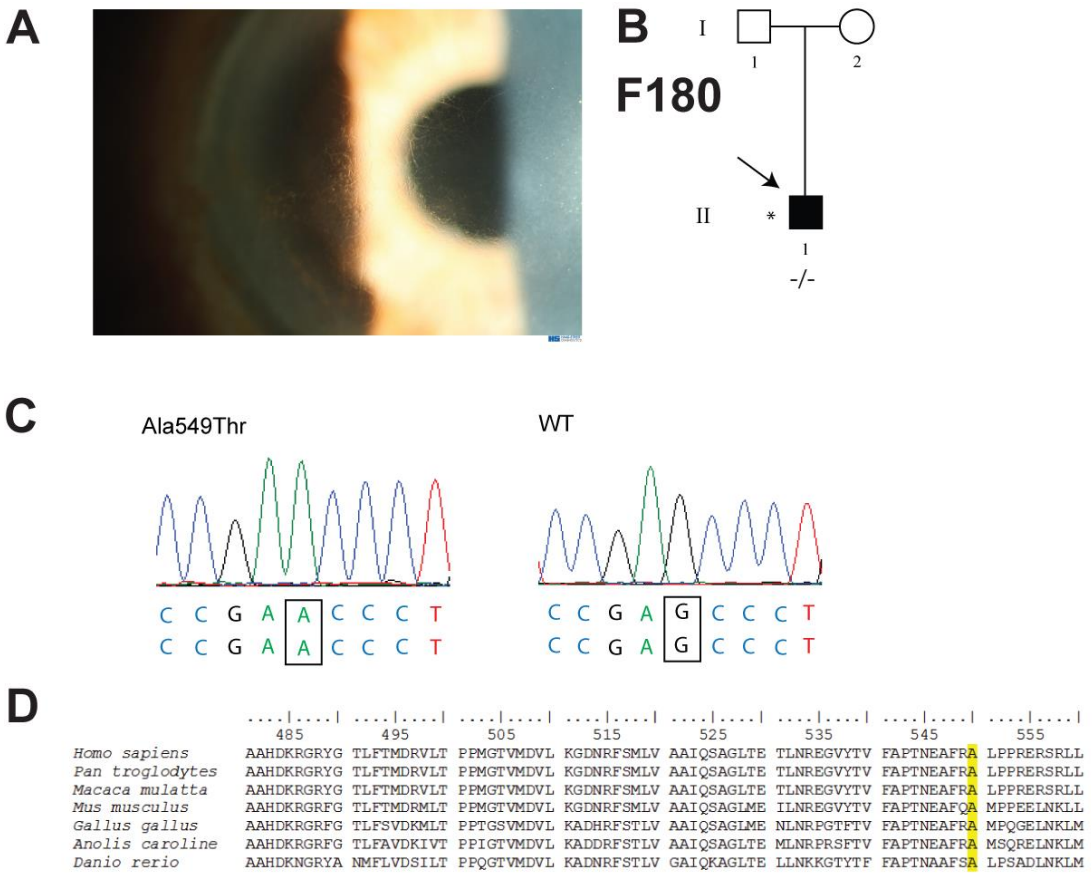


Figure 5.8 Identification of a homozygous variant in *TGFBI* of unknown significance in a proband with lattice corneal dystrophy or climatic droplet keratopathy. (A) Clinical phenotype of F180: II:1. **(B)** Pedigree structure of F180. * indicates DNA sample was available for testing. -/- indicates presence of mutation in homozygous state. There was no family history of corneal disease. **(C)** Sequence electropherograms showing presence of homozygous mutation and reference sequence in a control sample. **(D)** Alignment of orthologous TGFBI protein sequences with the conserved Ala-549 residue highlighted in yellow.

5.3 Polymorphic corneal amyloidosis

5.3.1 Polymorphic corneal amyloidosis and myelopathy

Family #181 was referred from Oxford Neurology for genetic analysis with an unknown inherited syndrome. The proband (I:1) presented in his early twenties in 1985 with bilateral asymmetric muscle wasting in his hands, more severe in the left, after noticing a decrease in finger grip of his left hand. At the initial examination, corneal opacity was also noted and clinical images were taken. These were clinically consistent with the appearance of PCD/PCA (Figure 5.9 A). Interestingly, the

proband had noticed a painless visual disturbance 14 years before the onset of other symptoms. PK was performed in one eye, a year prior to his first neurological examination. On referral, the proband was in his sixties and the disease has been slowly progressive since its onset. Recent magnetic resonance imaging (MRI) of his cervical spine revealed markedly reduced anteroposterior diameter, cord atrophy and superimposed degenerative changes with disc bulges at C5/6 and C6/7. The proband subsequently underwent lumbar spine decompression. The proband's daughter (II:4) also presented in her twenties with asymmetric bilateral hand muscle wasting. At the most recent examination, the disease had progressed and her symptoms included lower limb spasticity, dorsal column involvement, extensor plantar responses, ankle clonus and spinothalamic tract dysfunction. Two additional affected individuals are known in the pedigree. Corneal opacities were present in all individuals of the pedigree who were also affected with myelopathy. The co-inheritance of both traits is consistent with an autosomal dominant condition (Figure 5.9 B). This is the first description of a disorder combining the ophthalmic phenotype of PCA with a myelopathy.

Affected individuals I:1 and II:4 from the pedigree were analysed by WES (section 2.6). Variants were filtered for rare, non-synonymous or splice site changes (≤ 0.005 in EVS, 1000G and UCL controls), shared between the two affected individuals. Initial filtering revealed that 121 rare (MAF ≤ 0.005) variants were shared between the two affected individuals. Rare variants were further filtered against a candidate list of genes reported to cause a corneal dystrophy (Table 1.1) or syndromes with the involvement of corneal symptoms (Appendix B) resulting in three shared rare variants in candidate genes (Table 5.1; section 2.6.2.3). Biallelic mutations in *PLEC* and *POMT1* are associated with inherited syndromes that have been reported to include corneal erosions and corneal clouding, respectively and therefore were excluded on the basis of inheritance and lack of association with the described corneal phenotype. In addition, both the variants identified in these genes were predicted to be benign/tolerated by prediction programs SIFT and Polyphen2.

The remaining variant identified was a heterozygous mutation in *GSN*. Mutations in this gene cause a form of systemic amyloidosis associated with amyloid deposition

in the cornea and therefore is consistent with the corneal phenotype of the proband (section 5.1.1). The presence of the *GSN* c.1738G>A; p.(Glu580Lys) variant was confirmed in both the proband and his affected daughter using Sanger sequencing (section 2.2; Appendix A). The variant is absent from the ExAC database and alters a conserved residue (Figure 5.9 D). In addition, it was predicted to be pathogenic by in silico pathogenicity prediction tools. As the *GSN* mutation affected the terminal codon of exon 12, in silico splice prediction tools were used to assess whether the mutation was likely to affect splicing. Loss of a splice site is associated with a lower splice prediction score, however in this case a slight increase in the canonical splice donor site score was predicted by all three programs used indicating that it is unlikely to disrupt the splice site (Table 5.2). No other alternative splice donor sites were identified by these programs in the presence of the p.(Glu580Lys) variant. The *GSN* mutation is likely a novel mutation causing gelsolin amyloidosis in this family, and represents a new clinical association as a corneal phenotype of PCA has not been previously described with this myelopathy.

Table 5.15 Rare (≤ 0.005) shared non-synonymous or splice site shared variants in individuals I:1 and II:4 from F181 in genes causing inherited corneal disorders. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Most promising candidate highlighted in bold.

Gene (Accession)	Heterozygous Variants	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
PLEC NM_201380.3	c.8462G>A; p.(Arg2821Gln)	1 (T)	0 (B)	0.0015	Epidermolysis bullosa (AR)
POMT1 NM_001077366.1	c.592G>C ; p.(Val198Leu)	0.58 (T)	0.012 (B)	NP	Muscular Dystrophy-Dystroglycanopathy (AR)
***GSN* NM_000177.4**	**c.1738G>A; p.(Glu580Lys)**	**0 (D)**	**0.990 (PrD)**	**NP**	**Meretoja Syndrome (AD)**

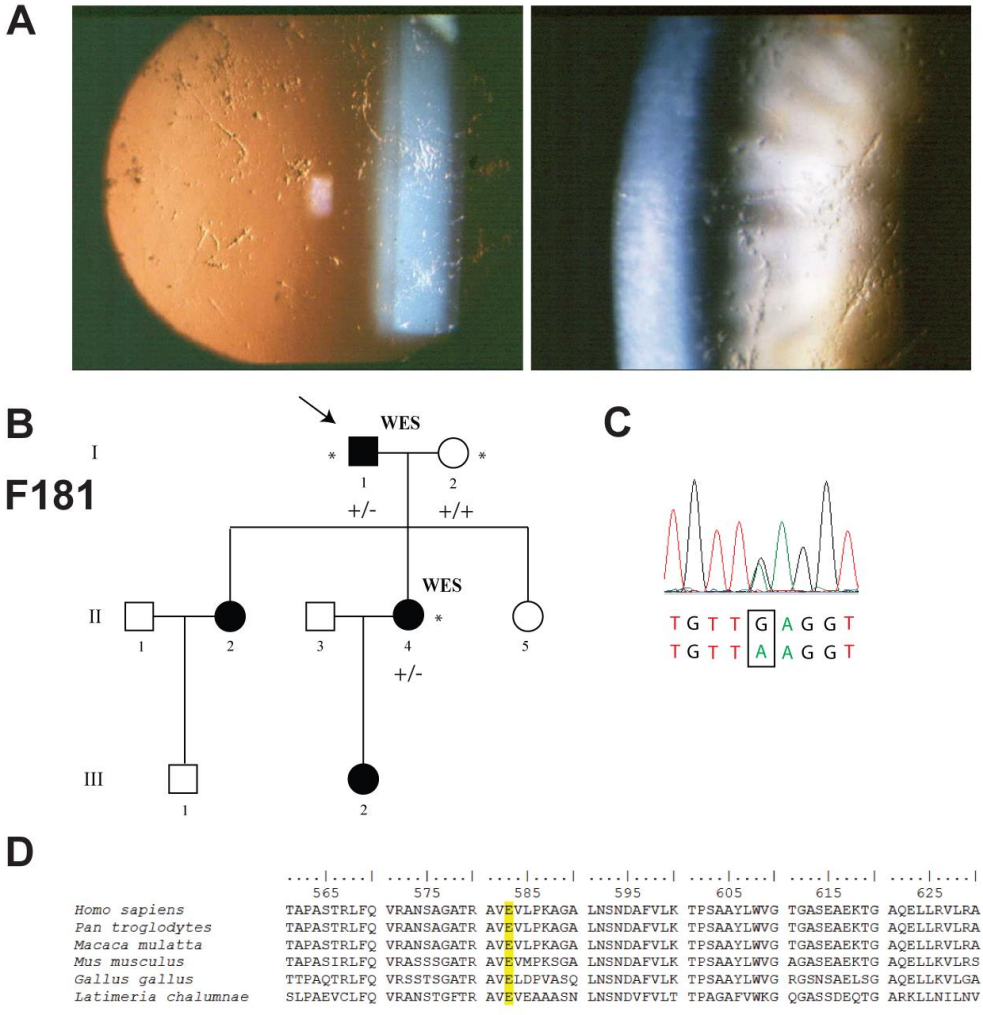


Figure 5.9 Identification of a novel *GSN* mutation c1738G>A; p.(Glu580Lys) in family #181 with polymorphic corneal amyloidosis and myelopathy. (A) Corneal phenotype showing presence of lattice lines, reminiscent of polymorphic corneal amyloidosis. **(B)** Pedigree demonstrating dominant segregation of *GSN* variant identified by whole exome sequencing in the family. **(C)** Sequence electropherogram of *GSN* variant. **(D)** Alignment of orthologous protein residues showing conservation of the affected amino acid residue in the GSN protein (highlighted).

Table 5.2 In silico prediction of potential effect of a *GSN* c.1738G>A mutation on the canonical splice donor site. Human Splicing Finder, NNSPLICE and NetGene2 each identify potential splice sites in a given sequence and score them from 0 to 100, 0-1 and 0-1, respectively. In each instance, a higher number corresponds to an increased likelihood of the motif functioning as a splice site. No change in score was observed for any other splice donor site.

	Canonical splice donor score	**Canonical splice donor score in presence of c.1738G>A**
Human Splice Finder (0 – 100)	74.05	74.96
NNSplice (0 – 1)	0.93	0.96
NetGene2 (0 – 1)	0.91	0.93

5.3.2 Sporadic polymorphic corneal amyloidosis

A single proband was recruited from MEH with sporadic PCA. The proband of family #182 was recruited with a typical clinical appearance of PCA/PCD (Figure 5.10 A) and no known family history of corneal disease (Figure 5.10 B). Screening *TGFBI* exons 4, 11, 12, 13, 14 and 16 revealed no potentially pathogenic mutations (section 2.2; Appendix A). An alternative diagnosis of Meretoja syndrome was also excluded by sequencing exon 4 of the *GSN* gene (section 2.2; Appendix A). Subsequently, DNA from the proband was sequenced by WES (section 2.6).

The initially filtering strategy was to determine whether the corneal symptoms could be caused by an inherited syndrome or was an expansion of the clinical phenotype of a corneal dystrophy (section 2.6.2.3). WES variants were filtered for rare (MAF ≤ 0.005) variants present in genes associated with corneal dystrophy (Table 1.1) or inherited corneal syndromes (Appendix B), resulting in the identification of variants in two genes, both of which cause autosomal recessive epidermolysis bullosa (Table 5.3; section 2.6). The c.8371C>T; p.(Arg2791Trp) variant in *COL7A1* is annotated in mutation databases as likely pathogenic. Further investigation of the WES data for this gene identified another variant c.3830C>T; p.(Pro1277Leu) which had been excluded during the filtering stage due to a frequency of >0.05 in internal controls, although this was still relatively rare (MAF: 0.052). Interestingly, this variant has

also previously been annotated as pathogenic, however frequency data in ExAC showed 5075/121356 including 146 homozygous individuals, which argues against a role for this variant in disease. Furthermore, the amyloidogenic phenotype is not consistent with a clinical diagnosis of epidermolysis bullosa. WES data excluded mutations in the *GSN* gene, which was previously considered one of the most promising candidates, due to the phenotypic similarity with the corneal phenotype of family F181.

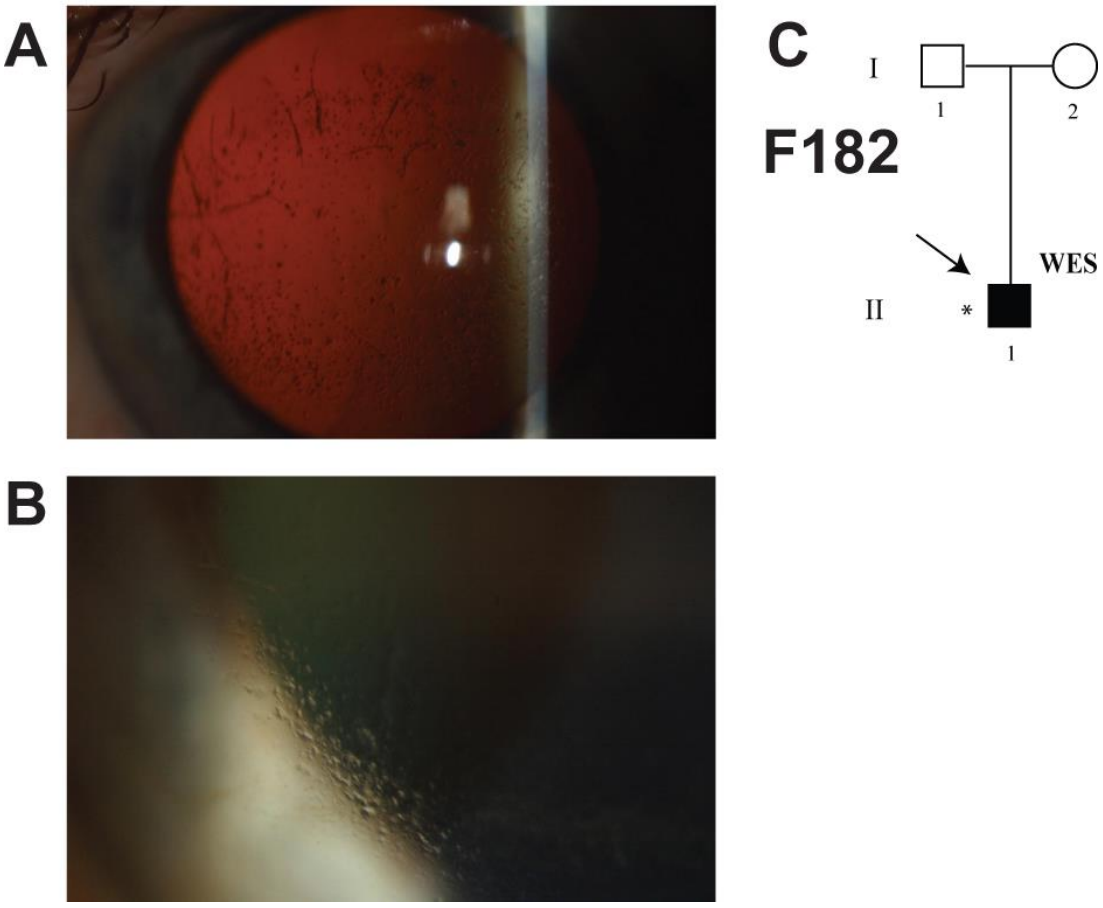


Figure 5.10 Summary of the proband of family #182 with sporadic polymorphic corneal amyloidosis. (A) Corneal images of proband F182: II:1 in retroillumination showing lattice and punctate opacities. **(B)** Punctate opacities are observable. **(C)** Pedigree of F182 demonstrating lack of family history of corneal disease.

Table 5.3 Rare (≤ 0.005) non-synonymous or splice site variants in genes causing inherited corneal disorders in individual F182: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Heterozygous Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
COL7A1 NM_000094.3	c.8371C>T; p.(Arg2791Trp)	0.16 (T)	0.999 (PrD)	0.0020 (140/69660)	Epidermolysis bullosa (AR)

Following exclusion of an inherited syndrome, potentially novel genetic causes were considered. Due to the lack of family history, the potential inheritance modes to consider were recessive, partially penetrant dominant and *de novo*. Alternatively, the phenotype may be degenerative. To explore possible recessive disease, WES variants were filtered for homozygous or (potential) compound heterozygous rare (MAF ≤ 0.005) non-synonymous or splice site variants resulting in the identification of 5 candidate genes (Table 5.4). One of the candidate genes, *NDUFB6,* was present in the corneal proteome (section 2.6.2.4; Appendix C).

Table 5.4 Rare (≤ 0.005) non-synonymous or splice site homozygous or compound heterozygous variants in F182: II:1. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Those present in the corneal proteome are highlighted in bold.

Gene (Accession)	Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
NDUFB6 **NM_002493.4**	**Hom c.76C>G; p.(Gln26Glu)**	**0 (D)**	**0.987 (PrD)**	**0.00021 (25/119416)**
SPTA1 NM_003126.2	c.7104G>T; p.(Lys2368Asn)	0 (D)	0.322 (B)	0.0015 (181/120604)
	c.33G>C; p.(Glu11Asp)	0.72 (T)	0.001 (B)	0.0038 (453/120662)
PTPRJ NM_002843.3	c.1274G>A; p.(Arg425His)	0.26 (T)	0.734 (PrD)	0.0010 (125/121372)
	c.3683A>G; p.(Gln1228Arg)	0.24 (T)	0.996 (PrD)	0.000016 (25/121276)
PCSK6 NM_002570.4	c.536G>A; p.(Arg179His)	N/A	0 (D)	NP
	c.2891C>T; p.(Thr964Met)	N/A	0 (D)	0.0037 (400/107326)
ZNF460 NM_006635.3	c.351A>T; p.(Lys117Asn)	0.47 (T)	0 (B)	0.0017 (208/121394)
	c.1512C>A; p.(His504Gln)	0 (D)	0.077 (B)	0.000074 (9/121408)

The possibility of dominant disease was also investigated. Variants were filtered for non-synonymous or splice site variants in the corneal proteome candidate gene list that were absent from 1000G, EVS, UCL internal controls and ExAC (Dyrlund *et al.*, 2012) (section 2.6.2.4; Appendix C). Ten variants were prioritised, but could not be further refined due to a lack of additional DNA samples (Table 5.5). The most compelling candidate was a non-synonymous mutation in *PAPLN*. In 2014, a conference abstract was published outlining a family with a corneal dystrophy of Bowman's layer without a mutation in *TGFBI*. WES analysis of two affected and one unaffected individual from the family identified ten candidate genes containing shared novel variants, in genes expressed in the cornea. Two of these variants were predicted to be deleterious: *BAZ2A* and *PAPLN* (Le, Frausto and Aldave, 2014).

Although no details are provided about the *PAPLN* variant identified, or the corneal phenotype except that is a Bowman's layer dystrophy, the presence of a novel *PAPLN* variant in a second family with a corneal dystrophy is intriguing. Due to the absence of familial samples, segregation analysis could not be performed. Little is known about PAPLN function in humans, however in *Drosophila melanogaster* and *Caenorhabditis elegans* it a secreted extracellular matrix protein that is primarily found in basement membranes (Fessler *et al.*, 2004).

Table 5.5 Novel non-synonymous or splice site variants in F182: II:1 in genes encoding proteins present in the corneal proteome. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database. Best candidate highlighted in bold.

Gene (Accession)	Heterozygous Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC
IARS2 NM_018060.3	c.2342A>G; p.(Lys781Arg)	0.63 (T)	0.019 (B)	NP
VARS NM_006295.2	c.3352G>C; p.(Val1118Leu)	0.11 (T)	0.542 (PrD)	NP
ITPR3 NM_002224.3	c.859G>C; p.(Val287Leu)	0.02 (D)	0.997 (PrD)	NP
RMDN1 NM_016033.2	c.58G>T; p.(Gly20Trp)	0.04 (D)	0.521 (PoD)	NP
	c.49G>T; p.(Ala17Ser)	0.86 (T)	0.003 (B)	NP
WNK1 NM_018979.3	c.3623G>T; p.(Ser1208Ile)	0 (D)	0.022 (B)	NP
***PAPLN* NM_173462.3**	**c.2962G>A; p.(Asp988Asn)**	**0.69 (T)**	0.004 (B)	**NP**
EMP2 NM_001424.5	c.54G>C; p.(Leu18Phe)	0.09 (T)	1 (PrD)	NP
ARL6IP1 NM_015161.2	c.577_580del p.(Lys193Phefsdel*37)	N/A	N/A	NP
CAPS NM_080590.3	c.387+4C>A	N/A	N/A	NP
RAD23A NM_001270362.1	c.892G>C; p.(Gly298Arg)	0.47 (T)	0.912 (PrD)	NP

5.4 Pre-Descemet's corneal dystrophies

5.4.1 STS deficiency

The proband of family #183 (III:1), of Italian origin, was referred to MEH following identification of opacities at a routine optometry examination and was subsequently diagnosed with PDCD. Spectral-domain optical coherence tomography (OCT) identified the presence of bilateral, central pre-Descemet amorphous deposits in the cornea which was consistent with this diagnosis (Figure 5.11 A). No other symptoms were noted. Previously, PDCD has been identified both as an isolated condition, and in combination with the presence of the skin condition X-linked ichthyosis (section 5.1.5). The proband's family were located in Sicily therefore were not available for clinical examination however his brother had also previously been diagnosed with an undisclosed ocular condition (Figure 5.11 B).

The proband of family #184 (II:1) was initially seen at MEH due to adenoviral keratoconjunctivitis in his right eye causing scarring and resulting in blurred vision. Clinical examination subsequently revealed bilateral posterior corneal haze and the appearance of granular opacities in the deep stroma, anterior to Descemet membrane. He was recruited to the study with a diagnosis of PDCD (Figure 5.11 D).

Both missense mutations and deletions encompassing *STS* have been shown to cause X-linked ichthyosis, with ~ 90% of cases caused by deletions encompassing the entire gene. The gene causing isolated PDCD without skin symptoms has not yet been identified. Despite the lack of major skin symptoms on initial examination in F183, the X-linked ichthyosis gene *STS* was considered to be a possible candidate gene for PDCD to be excluded in the first instance. As both probands were male, a deletion of the *STS* gene on the X chromosome would result in a failure to amplify by PCR. Primers were designed to amplify exon 2 and exon 9 of the ten exon *STS* gene, to detect a deletion encompassing either the 5' or 3' end of the *STS* gene (section 2.4; Appendix A). The PCR reaction successfully amplified *STS* exons 2 and 9 using two control DNA samples as a template, however no amplimer was observed using either proband's DNA sample for two independent PCRs (Figure

5.11 D). Amplification of *TGFBI* exons using their DNA as a template demonstrated that the DNA was of sufficient quantity and quality to produce an amplimer when the complementary sequence was present in the genome (data not shown). Subsequent clinical follow up revealed that F183: III:1 suffered from dry skin, which was asymptomatic when treated, and he also indicated that both his grandfather and brother were affected with an undiagnosed dry skin disorder. F183: III:1 had been previously examined by a dermatologist, however no diagnosis was provided. Furthermore, it became apparent that F184: II:1 had been diagnosed with X-linked ichthyosis in childhood and had an affected brother (Figure 5.11 C). The symptoms of both probands are therefore consistent with a clinical and molecular diagnosis of *STS* deficiency resulting in X-linked ichthyosis with PDCD.

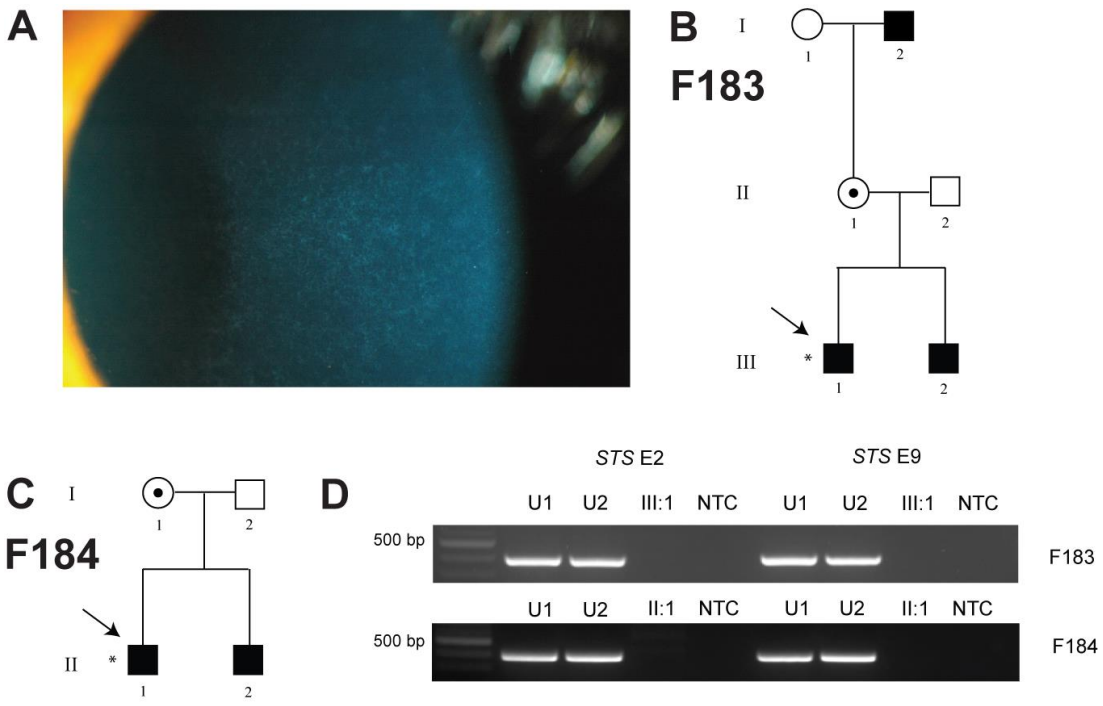


Figure 5.11 X-linked ichthyosis and pre-Descemet corneal dystrophy in families #183 and #184. * indicates the patient was examined and a DNA sample was available for genetic testing. **(A)** Corneal phenotype showing small, fine opacities which did not affect vision in proband of F183: III:1 consistent with a diagnosis of PDCD. Skin symptoms were revealed post-genetic testing. **(B)** Pedigree structure of F183 with the proband indicated with an arrow. Inheritance of the dystrophy in the family is consistent with X-linked inheritance where the mother is presumed to be a carrier. The proband's brother and grandfather were affected with dry skin but were not available for ophthalmological examination. **(C)** Pedigree structure of F184. **(D)** PCR amplification of the *STS* gene (exon 2 and exon 9), located on the X chromosome, in unaffected male control DNA (U1: unaffected DNA sample one, U2: unaffected DNA sample two), the probands of F183 and F184 and a non-template control (NTC). No amplification was observed in the proband indicating a deletion of the *STS* gene that spans at least from exons 2 to 9 of the ten exon gene.

5.4.2 Posterior stromal opacities of unknown aetiology

Two probands from families #185 (II:1) and #186 (II:1) were recruited with posterior stromal comma-shaped opacities of unknown aetiology that were not consistent with a previously described corneal dystrophy. Proband F185: II:1 was male in his sixties with bilateral oval and comma-shaped opacities in the posterior stroma, located adjacent to Descemet's membrane (Figure 5.12 A). The endothelium was unaffected and the proband was asymptomatic, with no reduction in visual acuity. There was no

known family history, although other family members were not examined (Figure 5.12 B). Proband F186: II:1 was a female in her sixties who was referred to MEH due to opacities being identified during a routine optometry examination. On clinical examination, bilateral peripheral opacities were noted in the poster stroma. The endothelium was not affected and the endothelial cell count was normal. She was also asymptomatic with no known family history (Figure 5.12 C).

To determine whether these opacities were caused by X-linked ichthyosis gene, *STS* exons 2 and 9 were amplified using DNA from the proband F185: II:1, which indicated that there was no large deletion of the gene (section 2.4; Appendix A; Figure 5.12 D). It was not possible to use this assay for F186: II:1, who was female. While a coding mutation in *STS* remains a possibility, the opacities differ from those typically observed in PDCD. The most common cause of stromal corneal dystrophies in the MEH cohort, *TGFBI* (exons 4, 11, 12, 13, 14 and 16) was screened in both probands but was negative (section 2.2; Appendix A).

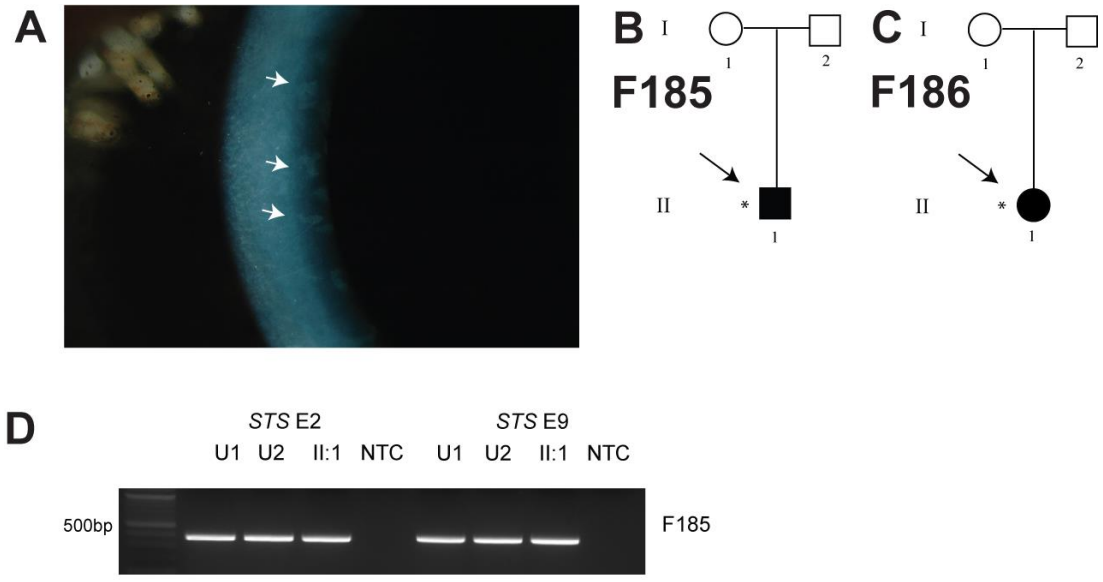


Figure 5.12 Two cases of genetically unsolved Pre-Descemet corneal dystrophy in families #185 and #186. (A) Corneal image of the proband from F185 with opacities shown with arrows. **(B)** Pedigree structure of F185 showing no known family of corneal opacities. **(C)** Pedigree structure of F186. **(D)** Amplification of exon 2 and 9 of *STS* gene in F185: II:1 eliminates a whole gene deletion as a potential cause of opacities. Deletion of *STS* could not be tested in F186 as the proband was female.

5.5 Corneal clouding of unknown aetiology

5.5.1 Familial corneal clouding

The proband of family F187 (II:2) was a 56 year old female referred with bilateral corneal opacity including a floccular haze involving the whole corneal stroma and mild corneal arcus (Figure 5.13 A). The proband was initially clinically diagnosed with MCD, although paraproteinemic keratopathy was also considered as a possible diagnosis. The documented family history of the disorder was consistent with a dominant pattern of inheritance, as the proband had an affected mother and sister, although neither the clinical diagnosis of MCD nor paraproteinemic keratopathy is consistent with an autosomal dominant inheritance pattern (Figure 5.13 B). A diagnosis of MCD was eliminated by Sanger sequencing of *CHST6*, which was negative for both coding mutations and upstream rearrangements or deletions (data not shown) (section 2.2; 2.5; Appendix A). Both coding exons of *UBIAD1* were screened and this eliminated a diagnosis of SCD given the potential for a phenotypic variant without crystalline deposition. Subsequently, other corneal dystrophies associated with corneal haze including *TGFBI* and exon 4 of *GSN* were screened and excluded.

DNA from F187: II:2 was analysed by WES (section 2.6). Variants were filtered for non-synonymous or splice site variants with a MAF of ≤ 0.005 in genes associated with corneal dystrophy (Table 1.1) or inherited corneal disease (section 2.6.2.3; Appendix B). The only candidate was a variant in a gene, *APOA1*, which is associated with both autosomal dominant and recessive conditions (Table 5.6). Biallelic mutations in *APOA1* cause Apo-A1 deficiency which is associated with corneal clouding in a sub-set of patients (Miccoli *et al.*, 1996; Pisciotta *et al.*, 2003). Autosomal dominant mutations in *APOA1* are associated with systemic amyloidosis; however, no cases of corneal deposition have been associated with ApoA1 amyloidosis, to date (Rowczenio *et al.*, 2011). Despite manual interrogation of the aligned WES reads, no second variant was identified in the *APOA1* gene.

Table 5.6 Rare (≤ 0.005) non-synonymous or splice site variants in genes causing inherited corneal disorders in F187: II:2. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Heterozygous Variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
APOA1 NM_000039.2	c.517C>T; p.(Arg173Cys)	0 (D)	1 (PrD)	NP	Amyloidosis (AD) ApoA-1 deficiency (AR)

The *APOA1*: c.517C>T; p.(Arg173Cys) variant is absent from control databases, and predicted to be pathogenic by both SIFT and Polyphen2. Alignment of orthologous protein sequences using ClustalW revealed that the residue is conserved in the majority of organisms, with the Arg-173 residue replaced with a tyrosine in fish (*Gadus morhua;* Figure 5.13 D). The variant is absent from the unaffected father of the proband and was inherited from the proband's affected mother. RNA sequencing data revealed that the *APOA1* gene is not expressed in the corneal stroma, basal or superficial limbal crypts (suspected corneal epithelial stem cells niches) or in total cornea. However, in contrast the protein is present in all corneal layers in the corneal proteome (Dyrlund *et al.*, 2012).

The proband underwent DALK surgery and the sample sent for histological analysis. The corneal button was fixed using standard methods which does not preserve lipid material. The histology revealed the presence of vacuoles in the stroma, which were absent of material. There was no amyloid or other protein accumulation in the stroma that could account for the cloudy corneal opacities observed clinically. It was hypothesised that the vacuoles may have once been associated with lipid deposits that were abolished by the fixation process (Figure 5.13 B).

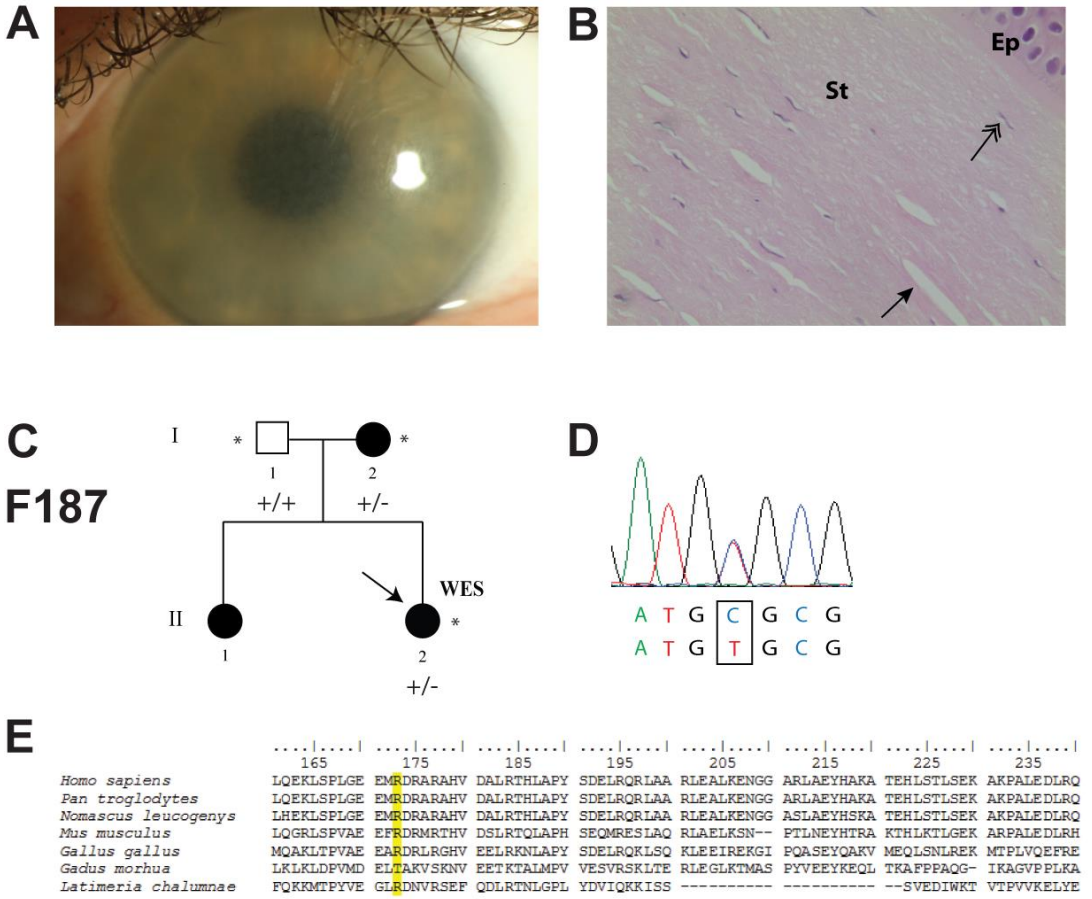


Figure 5.13 Novel *APOA1* c.517C>T; p.(Arg173Cys) variant of unknown significance identified in family #187 with corneal clouding. (A) Corneal images showing cloudy corneas in the proband. **(B)** Histology from deep anterior lamellar keratoplasty shows gaps in the stroma, which are devoid of material (arrow). The fixing process dissolves lipids which may have been present. Keratocytes are shown with double arrow. Ep: epithelium, St: stroma. There is no evidence of amyloid deposition. **(C)** Pedigree showing autosomal dominant pattern of inheritance in the family. +/- indicates individual is heterozygous for the mutation. The mutation was inherited from the proband's affected mother. **(D)** Sequence electropherogram showing presence of the *APOA1* c.517C>T; p.(Arg173Cys) mutation in the proband. **(E)** Alignment of orthologous APOA1 protein sequences shows affected residue Arg-173 is conserved in the majority of organisms.

Following identification of the *APOA1* variant, a serum lipid profile was obtained from the proband. This demonstrated high total levels of cholesterol, with low HDL-cholesterol and high LDL-cholesterol although triglyceride levels were normal (Table 5.7).

Table 5.7 Lipid profile of proband F187: II:2 carrying a heterozygous *APOA1* c.517C>T; p.(Arg173Cys) variant. Lipid profile is shown in mmol/L and the normal expected range for each variable. The proband has total cholesterol above normal range, low HDL cholesterol and high LDL cholesterol. Triglyceride levels are within the standard range.

	F187: II:2 results (mmol/L)	**Normal range (mmol/L)**
Cholesterol	5.2	2.3 – 4.9
HDL Cholesterol	0.8	1.2 – 1.7
HDL % of total	15%	20-99%
LDL cholesterol	3.5	0 – 3.0
Triglycerides	1.9	0 – 2.2

5.5.2 Sporadic corneal clouding

Three probands were recruited with corneal clouding of unknown aetiology. The proband F188: II:1 presented in her sixties with bilateral corneal opacity, and central mid-stromal haze (Figure 5.14 A). She was originally from Zambia and there was no family history of corneal disease. There had been a gradual deterioration of vision over several years prior to her examination; however, her visual acuity was not considered severe enough to require surgery. Paraproteinemic keratopathy was eliminated by plasma electrophoresis. Proband F189: II:1 also presented in his sixties with visual deterioration over several years. Diagnoses of SCD and CCDF were both considered due to the presence of symmetric, round, central stromal opacity (Figure 5.14 B). He was one of the youngest of fifteen siblings, with no other known relatives affected with corneal disease (Figure 5.14 E). His medical history was positive for type II diabetes and he had previously had a myocardial infarction in 2003. The proband of F190: II:1 was referred to MEH following identification of mild central corneal opacity during a routine optometry appointment. She was asymptomatic with no family history of corneal disease (Figure 5.14 F). During clinical examination at MEH, bilateral central corneal haze and peripheral corneal arcus was noted and a diagnosis of SCD was considered most likely (Figure 5.14 C).

Screening *UBIAD1* in all three probands was negative for potentially pathogenic mutations (section 2.2; Appendix A). Due to late age of onset of corneal symptoms and the absence of family history WES analysis was not undertaken.

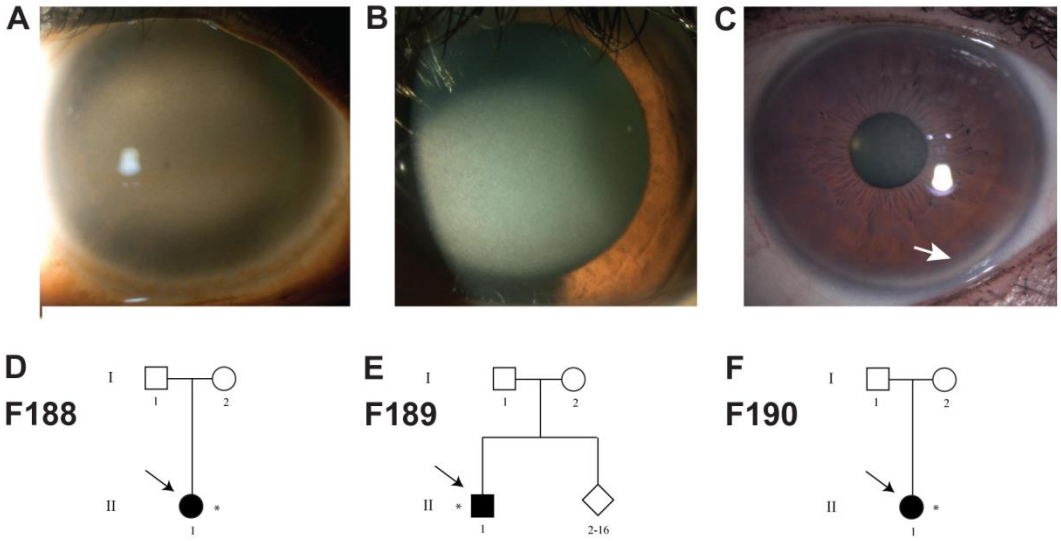


Figure 5.14 Three probands presenting with corneal clouding of unknown aetiology. Screening *UBIAD1* eliminated a diagnosis of Schnyder corneal dystrophy in all cases. Corneal images are shown of probands **(A)** F188: II:1 and **(B)** F189: II:1 showing extensive corneal clouding. **(C)** Clinical image of F190: II:1 with mild corneal clouding and corneal arcus (arrow). **(D)** Pedigree structure of F188 **(E)** F189 and **(F)** F190. *indicates DNA sample was available. None of the probands had a family history of corneal disease.

5.6 Corneal epithelial microcysts

Proband F191: II:3 was a 35 year old male initially diagnosed with MECD at 2 years old (section 3.1.2.4). He had myriads of intraepithelial cysts extending to the corneal limbus (Figure 5.15 A). He had previously undergone right laser PTK with rapid recurrence of the corneal changes. There was no medical history of cornel erosions, which are a typical feature of this dystrophy. The coding regions of both *KRT3* and *KRT12* genes were screened in their entirety in the proband (section 2.2; Appendix A). A heterozygous variant [c.43C>T; p.(Pro15Ser)] was identified in exon 1 of *KRT12*, however this is a common variant with an allele frequency of 0.3556 in the ExAC database. Three heterozygous variants were present in the coding region of the *KRT3* gene; c.1089C>T; p.(Arg363Arg), c.1123C>G; p. (Arg375Gly), c.1297G>A; p.(Glu433Lys), which have allele frequencies of 0.7937, 0.5240 and

0.1249, respectively, and therefore were not considered pathogenic. MECD was therefore excluded as a diagnosis, given the absence of potentially pathogenic changes in either *KRT3* or *KRT12* genes.

To investigate the genetic basis of the corneal epithelial cysts, DNA from the proband was analysed by WES (section 2.6). Given that he had an affected brother and unaffected parents, recessive and partially penetrant dominant inheritance patterns were considered. To exclude the possibility of a misdiagnosis of another corneal dystrophy, WES variants were filtered for non-synonymous exonic or flanking intronic variants with MAF $\leq$ 0.005 in the corneal dystrophy (Table 1.1) and inherited corneal disorders candidate gene lists (section 2.6.2.3; Appendix B). Six variants were identified in five genes associated with corneal syndromes, however only one gene, *MCOLN1,* which causes mucolipidosis IV was consistent with the corneal symptoms reported (Bach, 2001) (Table 5.8). The presence of the two *MCOLN1* variants was confirmed by Sanger sequencing in the proband, and segregation analysis was undertaken to determine whether the variants were present in cis or trans. Segregation analysis demonstrated that the c.572-2A>C variant was inherited on the maternal allele and the c.1517T>G; p.(Ile506Ser) variant on the paternal allele. Importantly, both variants were also inherited by the proband's affected brother (Figure 5.15 B). The proband declined referral to a metabolic unit for additional clinical testing.

Functionally, the *MCOLN1* c.1517T>G; p.(Ile506Ser) missense change in exon 12 is predicted to be damaging by SIFT and Polyphen2 and affects an amino acid residue conserved from mammals to the coelacanth (Figure 5.15 D). On the alternate allele, the c.572-2A>C variant is predicted to abolish the canonical splice acceptor site of exon 5 (Table 5.9). This would be predicted to result in skipping of exon 5 in the transcript, or use of an alternative splice acceptor site. The highest scoring predicted splice acceptor site would result in the loss of the first 11 nucleotides of exon 5, causing a frameshift and resulting in the incorporation of a premature termination codon.

Table 5.8 Rare (≤ 0.005) non-synonymous or splice site variants in genes causing inherited corneal disorders in individual F191: II:3. SIFT and Polyphen2 use conservation and structural data to predict the likelihood of a variant being pathogenic using a scoring system of 1-0 and 0-1, respectively. Disease information from OMIM. AR: autosomal recessive, AD: autosomal dominant. T: Tolerated, D: Damaging, B: Benign, PoD: Possibly Damaging, PrD: Probably damaging. NP: Not present in database.

Gene (Accession)	Heterozygous variant	SIFT (1-0)	Polyphen2 (0-1)	ExAC	OMIM (Inheritance)
TNXB NM_019105	c.8012C>T; p.(Ala6271Val)	0.65 (T)	0.007 (B)	76/116318	Ehlers Danlos syndrome (AR)
PLEC NM_201380	c.8164G>C; p.(Glu2722Gln)	0.5 (T)	0.0001 (B)	62/51838	Epidermolysis bullosa (AR)
LOXHD1 NM_001145472	c.1287G>C; p.(Trp429Cys)	0 (D)	0.999 (PrD)	NP	Fuch's endothelial corneal dystrophy (AD)
***MCOLN1* NM_020533**	**c.572-2A>C**	**NA**	**NA**	**NP**	**Mucolipidosis IV (AR)**
	c.1517T>G; p.(Ile506Ser)	**0 (D)**	**0.955 (PrD)**	**NP**	
ERCC2 NM_000400	c.1277C>T; p.(Pro426Leu)	0.02 (D)	0.241 (B)	22/120900	Trichothiodystrophy (AR)

Table 5.9 In silico prediction of potential effect of a *MCOLN1* c.572-2A>C mutation on the canonical splice acceptor site. Human Splicing Finder, NNSPLICE and NetGene2 each identify potential splice sites in a given sequence and score them from 0 to 100, 0-1 and 0-1, respectively. In each instance, a higher number corresponds to an increased likelihood of the motif functioning as a splice site. All three programs predict that the presence of the mutation abolishes the canonical splice acceptor site.

	Canonical splice acceptor score	Canonical splice acceptor in presence of c.572-2A>C	Alternative splice acceptor score
Human Splice Finder (0 – 100)	80.02	0	86.26
NNSplice (0 – 1)	0.93	0	0.82
NetGene2 (0 – 1)	0.95	0	0.95

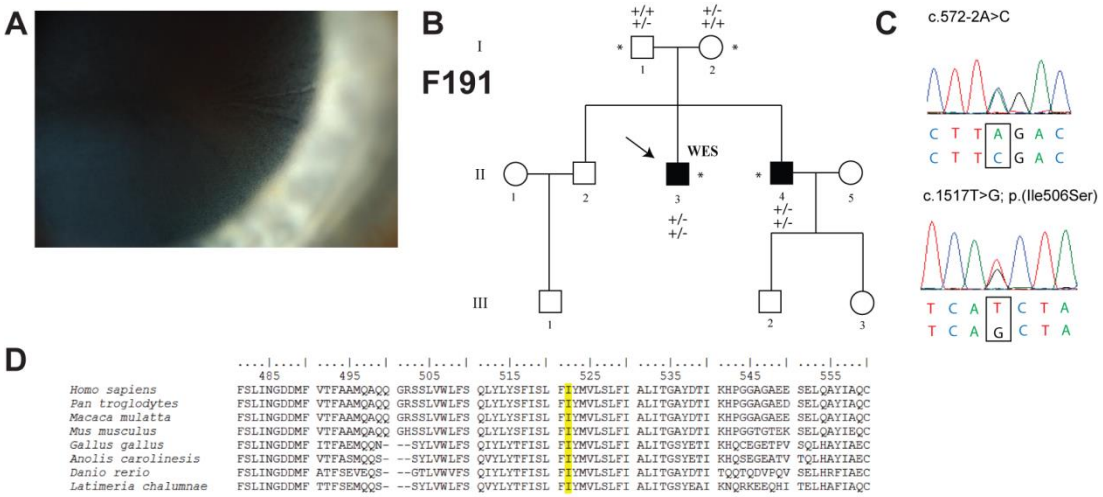


Figure 5.15 Family #191 clinically diagnosed with Meesmann corneal dystrophy suspected to have mucolipidosis IV following identification of biallelic *MCOLN1* variants by whole exome sequencing. (A) Corneal image demonstrating myriads of microcysts across entire surface of cornea in F191: II:3. **(B)** Pedigree showing recessive inheritance of the dystrophy in the affected family with the proband highlighted with an arrow. The c.572-2A>C variant was maternally inherited and the c.1517T>G; p.(Ile506Ser) variant paternally inherited. Both variants were present in the proband and his affected brother. **(C)** Sequence electropherogram of both *MCOLN1* variants identified in this family. **(D)** An alignment of orthologous protein sequences reveals that the Ile506 residue is a conserved residue.

Table 5.10 Summary of families (F175 – F191) analysed in chapter 5, their respective clinical diagnosis, the causative mutation or most compelling candidate identified and the updated syndromic or phenocopy diagnosis, if relevant. Standard corneal dystrophy acroyms are used. ? prior to a diagnosis indicates diagnostic uncertainty.

Family ID	**Clinical Diagnosis**	**Mutation (*or most compelling variant)**	**Updated Diagnosis**
F175	LCD	*GSN:* c.640G>A; P.(Asp214Asn)	Meretoja Syndrome
F176	?LCD ?Meretoja	*GSN:* c.640G>A; P.(Asp214Asn)	Meretoja Syndrome
F177	LCD	*GSN:* c.640G>A; P.(Asp214Asn)	Meretoja Syndrome
F178	LCD	*TGFBI* and *GSN* mutation negative	Paraproteinemic keratopathy
F179	LCD	*TGFBI* and *GSN* mutation negative	Paraproteinemic keratopathy
F180	?LCD ?CDK	*Hom *TGFBI:* c.1645G>A; p.(Ala549Thr)	
F181	?Syndrome	*GSN:* c.1738G>A; p.(Glu580Lys)	Meretoja Syndrome
F182	PCA	Unsolved	
F183	PDCD	*STS* deletion	X-linked ichthyosis
F184	PDCD	*STS* deletion	X-linked ichthyosis
F185	PDCD	Unsolved	
F186	PDCD	Unsolved	
F187	Corneal clouding	* Het *APOA1*: c.517C>T; p.(Arg173Cys)	?ApoA1 deficiency
F188	Corneal clouding	Unsolved	
F189	Corneal clouding	Unsolved	
F190	Corneal clouding	Unsolved	
F191	MECD	*Het *MCOLN1:* c.572-2A>C *Het *MCOLN1:* .1517T>G; p.(Ile506Ser)	?Mucolipidosis IV

5.7 Discussion

5.7.1 Gelsolin amyloidosis

Three probands who were initially clinically diagnosed with *TGFBI*-associated LCD were subsequently re-classified as Meretoja syndrome due to the presence of a p.(Asp214Asn) mutation in exon 4 of the *GSN* gene (section 5.1.1). The p.(Asp214Asn) mutation that was identified is a founder mutation in Finland, but has been shown to occur *de novo* in other populations (De La Chapelle *et al.*, 1992). In one case, both LCD and Meretoja were considered as potential clinical diagnoses with genetic screening requested to distinguish between these two possibilities, whereas in the remaining two cases there were no systemic symptoms identified prior to genetic screening. Meretoja syndrome is associated with systemic amyloid deposition and therefore with additional non-ocular symptoms; however, amyloid in the cornea resulting in a clinical phenotype reminiscent of *TGFBI*-associated LCD is typically the first presenting sign (Carrwik and Stenevi, 2009), therefore at this stage an individual may be mistaken for *TGFBI*-associated LCD.

In addition to the three probands with Meretoja syndrome, a family referred from Oxford was found to harbour a novel mutation [p.(Glu580Lys)] in *GSN* (section 5.3.1). Four *GSN* mutations have been described previously. The two mutations which cause Meretoja syndrome are both substitutions of amino acid residue Asp-214 [p.(Asp214Asn) and p.(Asp214Tyr)] (De La Chapelle *et al.*, 1992). Clinically, Meretoja syndrome is associated with symptoms of LCD, cranial neuropathy, peripheral neuropathy, and cutis laxa (Kiuru, 1992). However, two alternative mutations in *GSN* [p.(Gly194Arg) and p.(Asn211Lys)] have been described which cause gelsolin amyloidosis that is restricted specifically to the kidneys, specifically resulting in renal dysfunction. Affected individuals carrying these *GSN* mutations did not have corneal opacities or the other systemic features of Meretoja syndrome (Sethi, Theis and Quint, 2013; Efebera *et al.*, 2014). In the family harbouring the novel p.(Glu580Lys) *GSN* mutation, affected individuals appear to have a corneal phenotype which is similar to PCD/PCA, a subtype of LCD, and ataxia due to a myelopathy caused by neuronal amyloid deposition. Late-onset ataxia is a rarely reported symptom of Meretoja syndrome. In one published case, late-onset ataxia

and peripheral sensory polyneuropathy was described in a 78 year old patient with Meretoja syndrome, with the resulting immobility contributing to his death. Autopsy revealed severe loss of peripheral nerves, atrophy of lower limb muscles and moderate degeneration of posterior columns of the spinal cord (Tanskanen *et al.*, 2009). Based on these five mutations, there appears to be a relatively stable genotype-phenotype correlation between the clinical syndromes of the syndrome and the specific *GSN* mutation that has been inherited.

The GSN protein has 6 domains (G1-G6), each containing a type 2 Ca^{2+} binding site, with additional type 1 Ca^{2+} binding sites (requiring co-binding of actin) in domains G1 and G4 . Mutations affecting the Asp-214 residue, located in G2, compromise the binding of Ca^{2+} to the GSN protein, required for its activation. Although the mutant protein can be activated by use of the alternative Ca^{2} binding sites, it spends longer in an intermediate state that contains a furin cleavage site, compared to the wild type protein. The mutant protein is cleaved by furin to produce a fragment (C68) that is further cleaved by additional proteases to produce two amyloidogenic 8 kDa and 5 kDa fragments (Solomon *et al.*, 2009, 2012). The other known mutations (causing an isolated amyloidosis affecting only the kidneys) are located nearby at residues 194 and 211 and therefore may also act via a similar mechanism, whereas the affected amino acid affected by the novel mutation is located at residue 580 and predicted to be located in the G5 domain. The mechanism of pathogenicity may contribute to the development of different symptoms in patients harbouring different *GSN* mutations.

5.7.2 Polymorphic corneal amyloidosis

PCA is a distinctive corneal amyloidogenic phenotype that has been described both in sporadically affected individuals and in some families. The majority of cases are considered to be degenerative, however a sub-set of cases are associated with specific mutations in *TGFBI* (Eifrig *et al.*, 2004). One sporadic proband was recruited to the study with a diagnosis of PCA and a family referred from Oxford with a corneal phenotype consistent with the PCA. Both were negative for mutations in *TGFBI*. WES analysis of the sporadically affected proband and two affected individuals from the familial case revealed different underlying causes for each proband. The familial case was subsequently diagnosed with gelsolin amyloidosis

following WES analysis due to the identification of a novel mutation in *GSN*, discussed previously in section 5.7.1, whereas WES data from our proband with an isolated PCA phenotype lacked a mutation in *GSN*.

A causative mutation was not conclusively identified for the sporadically affected proband. Analysis of WES data was limited by the recruitment of a single proband with this phenotype. Using corneal proteome data to filter variants, a list of potential variants was identified, but this could not be further refined due to lack of additional DNA samples. The most promising candidate was a missense variant in *PAPLN*, a gene which has previously been identified in another published family with a Bowman's layer corneal dystrophy and encoding a protein which is a component of basement membranes (Le, Frausto and Aldave, 2014). Until more is known about the clinical phenotype of this proband, or more families are recruited with the PCA clinical phenotype, the contribution of this gene to corneal dystrophies remains unresolved.

Amyloid deposited in the cornea can also occur as a result of age, trauma or environmental degenerative processes (Suesskind *et al.*, 2006). Given these findings, it is plausible that PCA is a non-specific amyloidogenic corneal phenotype that can have either a degenerative or genetic aetiology but does not seem to be specific for a particular genetic diagnosis. This is comparable to other causes of corneal lattice lines such as those seen in classic LCD, which can be caused by mutations in *TGFBI, GSN* or occur as a result of environmental, traumatic or age-related aetiologies. If there is a specific novel genetic cause of familial PCA cases, further families will need to be recruited to investigate this possibility.

5.7.3 Paraproteinemic keratopathy

Two probands recruited to the study were subsequently identified to have paraproteinemic keratopathy; one of which was previously diagnosed with LCD, and a second in which paraproteinemic keratopathy was the most likely diagnosis, but LCD was considered a possibility. In both cases, the corneal phenotype was most reminiscent of LCD. The lack of family history of corneal disease was an important indication of a diagnosis of paraproteinemic keratopathy. However other reported

corneal phenotypes mimic the cloudy corneas or crystalline deposits of SCD (Milman *et al.*, 2015); this makes diagnosis more complicated as small families and *de novo* occurrence of causative mutations are more common with this dystrophy. Clinically, paraproteinemia can sometimes be distinguished from LCD by the unusual reticular pattern of opacities, however diverse clinical presentations have been described in the literature, therefore paraproteinemia should be excluded as a potential cause in all cases of LCD or other corneal dystrophies without a positive family history of corneal disease and a definite genetic diagnosis.

Distinguishing paraproteinemic keratopathy from an inherited corneal dystrophy has implications for both the patient and their families. Diagnosis of a corneal dystrophy may result in the proband's children or other relatives erroneously believing that they are also at risk of developing corneal opacity. It also fails to address the additional clinical risk of those with paraproteinemia who may have an underlying malignancy or may go on to develop a malignancy in the future.

5.7.4 Climatic droplet keratopathy

CDK is a corneal degeneration which occurs predominantly in individuals over 40 who work outdoors in rural environments (section 5.1.4). It was suspected in a single proband who presented at MEH with an unusual pattern of lattice lines and a history of working outdoors, in which both LCD and CDK were considered potential clinical diagnoses. Screening *TGFBI* revealed the presence of a homozygous *TGFBI* p.(Ala549Thr) variant of uncertain significance. This variant has been previously reported to occur in a heterozygous state on the alternate allele of a patient with GCD1 due to a p.(Arg555Trp) mutation. The reported patient had an unusually severe disease progression for this mutation, which the authors hypothesised may be due to the contribution of the p.(Ala549Thr) variant, although this could not be proven. The reported patient's son was a carrier of the heterozygous p.(Ala549Thr) variant only, and showed no sign of corneal disease at age 32 (Frising *et al.*, 2006). Our proband had no family history of corneal disease, which could indicate that in the heterozygous state the variant is not pathogenic, presuming both parents were carriers of the mutation. It is possible that the *TGFBI* variant is only penetrant when it occurs in the homozygous state, in which case this would be the first recessive

TGFBI mutation. An alternative explanation is that the variant may pre-dispose an individual to developing corneal opacity when exposed to a specific environment, for example working outdoors, and therefore may cause an increased susceptibility to acquiring corneal opacity. It is also possible that the *TGFBI* variant is benign and has no contribution to the corneal phenotype.

5.7.5 Pre-Descemet corneal dystrophies

PDCD is described as "neither a well-defined entity nor is it a clearly hereditary or degenerative disorder" in the IC3D classification of corneal dystrophies (Weiss *et al.*, 2015). A number of cases have been described in which the opacities vary in the size, colour (white, grey, polychromatic), shape (punctate, fusiform threads, comma-shaped, circular or dot-like) and distribution, however all share a common localisation to the posterior stroma, just anterior to Descemet's membrane (Klintworth, 2003). The crystalline sub-type (also called punctiform polychromatic) has the most evidence for being a distinct corneal dystrophy, having been described in at least two families with an apparent autosomal dominant pattern of inheritance. In this sub-type, pre-Descemet opacities are equal size, are distributed across the entire cornea to the limbus and have polychromatic reflections (Fernandez-Sasso, Acosta and Malbran, 1979; Dolz-Marco *et al.*, 2014). Other described PDCD sub-types are cornea farinata, deep filiform and deep punctiform varieties (Fernandez-Sasso, Acosta and Malbran, 1979).

In total, four probands were recruited with an initial diagnosis of PDCD or had corneal opacities consistent with this diagnosis. Two of these probands, in which opacities were small and punctate, were subsequently identified to have a deletion of *STS* and therefore a diagnosis of X-linked ichthyosis. Although initially thought to be sporadic, after diagnosis proband F183 revealed that both his brother and maternal grandfather had an undiagnosed skin condition, although as they lived in Italy it was not possible to examine either of these affected relatives to determine whether they also had corneal opacities. The deletion was likely inherited on the maternal allele; therefore, investigation of potential corneal opacities in the female carrier would be interesting to determine whether *STS* mutations could also explain descriptions of females in the literature diagnosed with PDCD. The proband of F184

was also found to harbour a deletion of *STS* following a diagnosis of PDCD, however it was subsequently discovered that he had been diagnosed with X-linked ichthyosis in childhood. These findings demonstrate the importance of a full clinical workup in suspected instances of PDCD for other systemic, particularly dermatological, symptoms.

An additional two probands were recruited with a diagnosis consistent with PDCD. Both F185 and F186 had deep stromal comma-shaped opacities which differed in appearance to those observed in F183 and F184, and a lack of family history of corneal disease. A large deletion of *STS* was eliminated by PCR amplification in F185, however F186 was female, and indicating that the disorder is unlikely to be X-linked in this case and therefore detection of a deletion of *STS* was not possible in this case. Although *STS* deletions are the most common mechanism, missense mutations in *STS* have also been described, therefore this remains a potential cause of disease in these families as the gene was not screened for these mutations. Additionally, smaller deletions which do not encompass *STS* exons 2 or 9 have not been excluded. WES analysis of these probands may be useful to elucidate whether this corneal phenotype is caused by a recessive or *de novo* dominant mutation, or whether this is a corneal degeneration.

5.7.6 Corneal clouding associated with apoA-I

Corneal clouding is most often an indication of a clinical diagnosis of SCD. This diagnosis is typically accurate in cases where crystalline deposition is observed as these are pathognomonic for this condition. Approximately 50 % of probands with SCD present with corneal clouding without crystals; however, in these cases there are a variety of other possible diagnoses, some of which are inherited disorders and others which have an age or environmental aetiology (section 5.1.6). Four probands were recruited with corneal clouding in the absence of a *UBAID1* mutation. One these probands had an autosomal dominant pattern of inheritance, whereas three were sporadically affected.

WES analysis of the proband (F187: II:2) of autosomal dominantly inherited corneal clouding revealed the presence of a heterozygous variant in *APOA1* which

segregated with disease. Given the association between mutations in *APOA1* and corneal clouding, this was considered a good candidate. Biallelic loss of function mutations cause apoA-I deficiency/familial HDL deficiency. Affected individuals have very low or absent levels of apoA-I and HDL-cholesterol and in some cases have myocardial infarctions at a relatively young age compared to the general population (Römling *et al.*, 1994). Diffuse corneal clouding, corneal opacity and corneal arcus are often reported in individuals with biallelic *APOA1* null alleles (Miccoli *et al.*, 1996; Pisciotta *et al.*, 2003). Other clinical features include the presence of planar xanthomas. Heterozygous carriers of null *APOA1* alleles typically have intermediate levels of apoA-I and HDL-cholesterol but are usually asymptomatic (Pisciotta *et al.*, 2003). Furthermore, a number of dominantly inherited *APOA1* missense mutations have been described which cause familial systemic amyloidosis, predominantly in the liver, kidneys and spleen with no evidence of amyloid deposition in the cornea (Rowczenio *et al.*, 2011). Amyloidosis was excluded in our family F187 by histological analysis following DALK surgery.

This identified mutation [p.(Arg173Cys)] was absent from control databases, but has been described in one large Italian family. In 1980, a proband was identified with high triglyceride levels, low plasma HDL and low HDL-cholesterol, which was resistant both to diet and drug treatment. Two of three offspring of the proband were found to share the same lipid profile as the proband. There was no evidence of atherosclerosis in the proband and his family history was negative for atherosclerotic disease. LPL and LCAT enzyme activities were normal. In both the proband and two affected children, there was a reduction in apoliproprotein A-I (apoA-I) levels and isoelectric focusing revealed an abnormal banding pattern in the range of apoA-I with additional bands observed in patients in comparison to controls. This combination of factors had not previously been described (Franceschini, Sirtori and Capurso, 1980). Further characterisation revealed the identity of the additional bands as a variant apoA-I protein sequence monomer (A-I $_{\text{MILANO}}$), a A-I $_{\text{MILANO}}$ disulphide-linked dimer, and a heterodimer formed of A-I $_{\text{MILANO}}$ and ApoA-II (Weisgraber, Bersot and Mahley, 1980). Isoeletric focusing was used to determine that the A-I $_{\text{MILANO}}$ protein variant differed from that of the wild-type protein with the substitution of the Arg-173 residue to cysteine (Weisgraber, Rall and Bersot, 1983). Sampling the entire ~ 1000 population of a small village, Limone sul Garda in

Northern Italy, resulted in the identification of 33 carriers of the A-I $_{MILANO}$ variant and demonstrated inheritance of the variant as an autosomal dominant trait and inherited from a shared ancestor (Gualandri *et al.*, 1985).

The unique clinical features of carriers of the A-I $_{MILANO}$ protein variant lead to the hypothesis that the protein variant was protective against coronary heart disease, given that there was no increased incidence of myocardial infarction despite the low levels of HDL-cholesterol, therefore the recombinant protein has been developed a potential therapeutic. A randomised control trial in patients with acute coronary syndromes showed a significant regression of coronary atherosclerosis following application of intravenous recombinant A-I $_{MILANO}$ protein/phospholipid complex (ETC-216) (Schoenhagen *et al.*, 2003). Recently a phase 1 clinical trial was published which administered a single dose of ETC-216 to patients with coronary artery disease (Kallend *et al.*, 2016).

The lipid profile obtained by our proband F187: II:2 is not typical for a carrier of the apoA-I variant. She did not exhibit hypertriglyceridemia and although her HDL-cholesterol levels are reduced in comparison to the normal range, this is not as severe as previously reported for carriers of this variant (Gualandri *et al.*, 1985). Affected members of the originally described Italian family were clinically examined due to the presence of corneal clouding in other inherited disorders associated with low HDL-cholesterol but all were negative for corneal opacity, unlike our family in which corneal clouding was present in at least two individuals harbouring the *APOA1* variant (Franceschini, Sirtori and Capurso, 1980; Franceschini *et al.*, 1981).

The *APOA1* p.(Arg173Cys) variant is of uncertain significance with regards to the corneal clouding observed in F187: II:2 and her affected mother. Given the important role of apoA-I in regulating HDL-cholesterol levels and the presence of corneal clouding in HDL-deficiency syndromes, including apoA-I deficiency, it represents the most likely current candidate. Given the development of ETC-216 as a therapeutic agent, the possible association of this variant in our family with corneal clouding has implications beyond the affected pedigree.

5.7.7 Corneal clouding of unknown aetiology

The association between lipid metabolism and corneal clouding may also explain the three sporadically affected probands with corneal clouding but no family history of corneal disease. Corneal arcus is the deposition of lipid material in the limbus, primarily originating from LDL-cholesterol, and occurs in older individuals without inherited disease and in younger individuals with SCD or inherited disorders of lipid metabolism. Corneal arcus has been associated with both hypercholesterolemia and hypertriglyceridemia in individuals aged 50 or older and is a risk factor for developing coronary heart disease (Chua *et al.*, 2004; Fernández, Sorokin and Thompson, 2007). The proband of F189 previously had a myocardial infarction, indicating that lipid dysregulation may also cause corneal clouding. Lipid profiling would be a useful investigation in these three probands, given that SCD has been excluded as a potential diagnosis.

5.7.8 Corneal epithelial microcysts

The proband of F191 was originally clinically diagnosed with MECD, an autosomal dominant dystrophy affecting the corneal epithelium (section 3.1.2.4). DNA from the proband was analysed by WES following negative pre-screening of *KRT3* and *KRT12*, resulting in the identification of two variants in *MCOLN1* and a putative diagnosis of mucolipidosis IV; typically, a severe lysosomal storage disease associated with psychomotor retardation and variant ocular abnormalities including corneal clouding and verticillata, retinal degeneration and optic nerve pallor. The majority of affected individuals have no motor function progression past 12 – 15 months of age and no language; therefore it seems unlikely that this diagnosis could go undetected. The age of onset is typically in the first or second year of life, and psychomotor symptoms are relatively stable for two or three decades, although ophthalmological changes may be slowly progressive (Chitayat *et al.*, 1991; Wakabayashi *et al.*, 2011). Our proband has been monitored at MEH for corneal clouding since he was two years old. He is described as having corneal microcysts similar to verticillata resulting in corneal haze, which is consistent with the corneal phenotype of mucolipidosis IV. However, he has no recorded developmental delays or other known ophthalmological symptoms.

Although severe in the majority of cases, there have been case reports of milder, atypical cases of mucolipidosis IV. In one case study, the authors reported a patient aged twenty who had ophthalmological symptoms in the absence of any psychomotor retardation (Dobrovolny *et al.*, 2007). She presented at two years old with corneal cloudiness. Corneal opacity was described as a diffuse fine dot-like epithelial opacity in a whorl-like pattern reminiscent of cornea verticillata. This did not progress during regular examinations up to age 20. This is the same age of onset and similar corneal opacities presenting in our proband. At 12 years old, she was first noted to have signs of retinal degeneration and her visual acuity was reduced to light perception only by 20 years old. Cultured fibroblasts from patients with mucolipidosis IV show autoflourescence and is considered an important sign for diagnosis, however this was absent in this individual. Interestingly, as with our proband, genetic analysis of *MCOLN1* revealed the presence of a heterozygous missense mutation and a splice site mutation. The authors propose that this mild presentation of disease may be due to the 'leakiness' of the splicing mutation resulting in a proportion of normal transcript being produced alongside the abnormal transcript. The splicing mutation is predicted to produce a transcript with a frameshift resulting in a premature stop codon, which if translated would produce a partially functional protein. It is possible that a similar mechanism is operating in our patient with a degree of full length *MCOLN1* transcript being produced from the allele harbouring the splice mutation. Alternatively, the missense change on the alternate allele may not be a functional null allele, but have some residual MCOLN1 protein activity. It has been shown previously that missense changes cause disease by retention of the MCOLN1 protein in the endoplasmic reticulum, whereas others correctly localise to late endosomes/lysosomes but have a lower pH in comparison to WT (Raychowdhury *et al.*, 2004). Due to a refusal of clinical follow up, we were unable to assess whether there was any evidence for retinal degeneration or mild motor deficits in the MEH proband.

As there have been limited reports of other cases of MECD that are not associated with mutations in *KRT12* or *KRT3*, one other possibility is that there is additional locus heterogeneity for MECD. This should be considered following further clinical follow up in the proband and a determination is made as to whether he has additional

clinical symptoms. If these are not present, then WES data should be re-interrogated for a novel genetic cause of MECD in this family.

5.7.9 Investigating alternative diagnoses for corneal dystrophies

Although there are isolated case reports in the literature of misdiagnoses of a proband with a corneal dystrophy, there have been no systematic reports on how often this occurs and the spectrum of underlying causes, in a large cohort study. Here, 18 probands were identified who were initially clinically diagnosed with a corneal dystrophy, or a suspected corneal dystrophy, within our total recruited cohort of 191 families. To aid future studies and to provide a diagnostic pathway, a workflow was developed which combined genetic, clinical and biochemical testing to investigate the underlying cause of the corneal opacity. This resulted in plausible candidate genes for the majority of probands investigated.

The method established is to initially pre-screen for the most likely genetic cause using Sanger sequencing of candidate genes depending on the given clinical diagnosis. For example, corneal clouding required screening for *UBIAD1* mutations to exclude a diagnosis of SCD. In cases of suspected LCD it is also recommended to screen *GSN* following negative genetic pre-screening of corneal dystrophy genes. Following negative pre-screening, paraproteinemic keratopathy should be investigated as a diagnosis using serum electrophoresis in probands lacking a positive family history of disease. At this stage, it is also recommended to obtain a lipid profile. Probands with a clear recessive or dominant inheritance pattern, and sporadic probands for which a diagnosis of paraproteinemic keratopathy was excluded, should then be analysed by WES. Using a combination of literature searches and the OMIM database, a list of genes was compiled which have been previously associated with corneal symptoms, predominantly corneal opacity, clouding, haze or oedema (Appendix B). WES variants should be filtered for rare variants in corneal dystrophies (Table 1.1) and genes causing syndromes with corneal involvement (Appendix B) and assessed for their likely role in pathogenicity. This is summarised in Figure 5.16 A. In the MEH corneal dystrophy cohort, this system this identified a plausible candidate gene which was at least consistent with the corneal appearance and symptoms of the proband in most cases (Figure 5.16 B).

When compiling candidate genes for the WES candidate gene list, no attention was given to symptoms in other tissues, either in their type or severity. The reason for this was that additional associated symptoms may display variable expressivity, be mild, or require further specific clinical tests in order to identify non-ocular symptoms. Furthermore, specific mutations in the same gene may cause different constellations of symptoms or may be specific for an eye disease and not associated with other syndromic features. For example, there are two mutations in *GSN* reported to cause Meretoja syndrome (systemic amyloidosis including cornea) and a further two mutations in the same gene which cause amyloidosis specifically affecting the kidneys. Furthermore, the same mutation inherited by different probands can cause different combinations and severity of symptoms due to environmental effects and genetic background; a common mutation may cause severe disease in one individual and mild disease in another. Given that the individuals in our cohort were initially diagnosed with corneal dystrophies, they were unlikely to also have severe symptoms affecting other tissues if they were not present at presentation, however in cases where the most likely diagnoses have been excluded, other rarer possibilities need to be assessed to determine the likely diagnosis.

Although, in some cases, additional symptoms were present on clinical examination following genetic diagnosis, there were some cases which were indistinguishable from a corneal dystrophy by clinical examination alone. These cases demonstrate the utility of genetic confirmation of a clinical diagnosis of a corneal dystrophy, even in cases where the clinical appearance and features are consistent with the diagnosis.

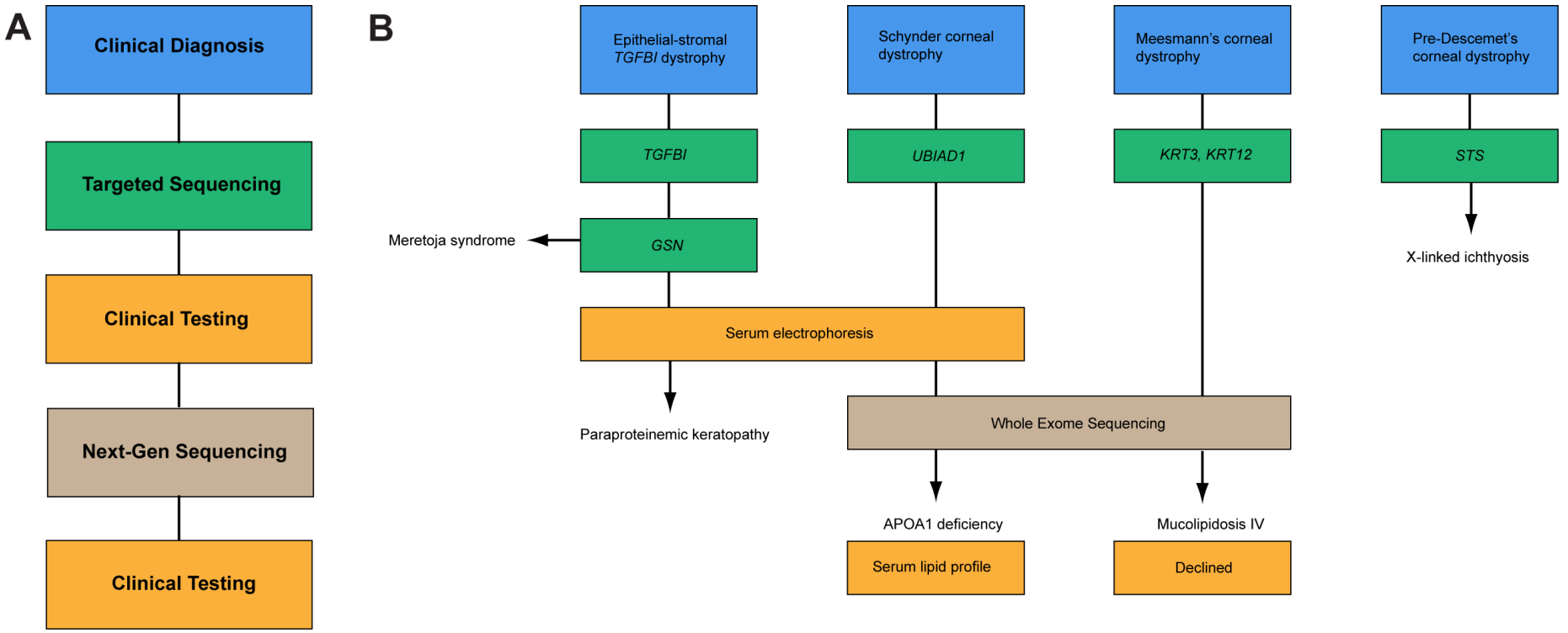


Figure 5.16 Method to achieve a genetic or clinical diagnosis of inherited syndrome or phenocopy of a corneal dystrophy. (A) Workflow designed for identification of syndromes or phenocopies includes initial clinical examination leading to clinical diagnosis of suspected corneal dystrophy, followed by targeted sequencing of the associated corneal dystrophy gene and likely syndromic causes. In the absence of a mutation, clinical testing for paraproteinemic keratopathy should be performed. If negative, this is followed by next-generation sequencing for identification of potential genetic causes of disease. If a candidate variant is identified, further clinical investigation should be sought to confirm the diagnosis. **(B)** Results of this workflow in the study cohort, leading to diagnoses of X-linked ichthyosis, Meretoja syndrome and paraproteinemic keratopathy. In two cases, whole exome sequencing identified potential diagnoses of ApoA1 deficiency and Mucolipidosis IV.

5.7.10 Classification of corneal dystrophies and syndromes

The classification of corneal dystrophies is largely historical. The current definition of a corneal dystrophy as defined by the IC3D classification is an inherited, monogenic disorder that is non-progressive, bilateral and not associated with any systemic symptoms; however, there are exceptions to each of these criteria (Weiss, Møller, *et al.*, 2008; Weiss *et al.*, 2015). For example, the majority of cases of EBMD and FECD are not monogenic, and there are cases of LCD that are unilateral or highly asymmetric (section 3.2.3). While some are stable, many corneal dystrophies show a slow progression in the severity of opacities, eventually resulting in corneal transplant when visual acuity is severely impaired. Most importantly, additional non-ocular symptoms have been associated with corneal dystrophies including inguinal hernia in PPCD3 and hearing loss in CHED2, therefore the distinction between corneal dystrophies and syndromes with corneal involvement is ambiguous. We have shown in this chapter that syndromes with corneal involvement may be mis-diagnosed as corneal dystrophies due to lack of other presenting symptoms or mild presenation of non-ocular symptoms that may be missed during routine ophthalmological examination. Furthermore, it is unresolved whether the corneal dystrophy PDCD is a distinct genetic entity, or whether all reported cases are misdiagnosed cases of X-linked ichythosis. Given this, it could be argued that some syndromes, such as Meretoja syndrome and X-linked ichythosis, should be included in the IC3D classification of corneal dystrophies as syndromic corneal dystrophies. It is interesting to note that Meretoja syndrome was included in the first edition of the IC3D but removed from the second as it was not considered to be a 'true' corneal dystrophy; however, PDCD caused by *STS* deficiency is still included. If the aim of the classification is to aid in diagnosis of these conditions then their inclusion will alert clinicians to the possibility of syndromic causes, even in cases where the patient's non-ocular symptoms are mild, or not present.

5.7.11 Summary of syndromes and phenocopies

The findings presented in this study demonstrate that misdiagnosis of corneal dystrophies is an uncommon occurrence, but that it can have significant implications for patients, given the true diagnosis is typically more severe and can involve other

organs which may need to be monitored or treated. Misdiagnosis was most common for syndromes or phenocopies that have been previously reported in the literature, confirming the need for these conditions to be considered or excluded during diagnosis. The most common diagnosis given to those with a syndrome or phenocopy was LCD – three probands were subsequently identified to have Meretoja syndrome (although it was already suspected in one) and 2 with paraproteinemic keratopathy. This is not unexpected given that both LCD and Meretoja syndrome are both associated with corneal amyloidosis, whereas other corneal dystrophy diagnoses (e.g. GCD and RBCD/TBCD) are associated with unique corneal opacities that are characteristic of a single corneal dystrophy diagnosis. Paraproteinemic keratopathy is clinically variable and can masquerade as several corneal dystrophies, however lack of positive family history can provide the best indication that the diagnosis is not autosomal dominant LCD. Although not observed here, paraproteinemic keratopathy may also appear as corneal clouding or haze which can look like SCD. Lack of family history is less informative in these cases as our cohort shows that many SCD families are small with few affected individuals and *de novo* occurrence of *UBIAD1* mutations has also been demonstrated (Lin *et al.*, 2016). In some cases, the true diagnosis was not fully resolved as further clinical investigation was required; however, WES was used to exclude mutations in known corneal dystrophy genes and to identify the most likely candidates causing the corneal opacity.

While it seems unlikely that a severe inherited syndrome can go undiagnosed, our identification of a patient with suspected mucolipidosis IV, and recent reports of mild cases in the literature, indicate that these should at least be considered when the genetic cause of corneal opacity is unknown. In some cases, additional symptoms that were unknown at the time of initial ocular examination, or were not considered to be relevant, were later shown to be part of a syndromic diagnosis following genetic investigation. Therefore, when making a clinical diagnosis of a corneal dystrophy all additional symptoms and medical history should be noted alongside ocular signs and symptoms, prior to genetic confirmation.

Chapter 6: Discussion

In this thesis, genetic investigation of a large cohort of patients with a suspected corneal dystrophy attending MEH is described, supplemented with referrals from additional sites from the UK and internationally. The aim of this study was to identify the spectrum of mutations causing corneal dystrophies as a pre-requisite for the development of targeted therapies, to identify genotype-phenotype correlations, and to identify novel corneal dystrophy genes.

6.1 Mutational spectrum and clinical findings in the cohort

Patients with a diagnosis of a corneal dystrophy attending MEH were recruited to the study. A minority (~ 5%) of the total MEH corneal dystrophy patient cohort had previously been genetically diagnosed, therefore these individuals were excluded. Not all corneal dystrophy diagnoses were included in this study, either due to their specific exclusion or lack of available patients with specific diagnoses. Patients with FECD were excluded from the current study due to its complex aetiology and the large number of cases at MEH (~ 450 patients); however, a small enriched cohort of FECD patients was screened for potential variants in novel endothelial corneal dystrophy genes. No patients were recruited with a diagnosis of FCD, CSCD, PACD, CCDF, SMCD, LECD or XECD.

A spectrum of mutations was identified in a large, ethnically diverse cohort of corneal dystrophy patients. When combined with clinical and histological data, genetic diagnosis is a powerful approach to improve clinical diagnostic and prognostic accuracy.

6.1.1 Anterior corneal dystrophies

Of the dominantly inherited dystrophies, epithelial-stromal *TGFBI* dystrophies were the most common, with 70 families harbouring a heterozygous mutation in *TGFBI* (section 3.2). Most mutations in *TGFBI* correlated with a specific corneal dystrophy diagnosis including LCD, TBCD/RBCD, and GCD. The only exception was the

p.(Gly623Asp) mutation that was identified in 5 probands, therefore a major cause of epithelial-stromal *TGFBI* dystrophies in the cohort, but was also identified in patients with EBMD and previous diagnoses of bilateral chronic ocular surface disorders (herpes simplex or dry eye disease). Despite their different clinical appearance, all patients carrying the p.(Gly623Asp) mutation responded well to PTK.

A single family with recessive GDLD was recruited from MEH and is the first description of this dystrophy in a patient of British origin. An additional family was referred from Colombia. Both families had biallelic mutations in *TACSTD2* (section 3.5). GDLD has previously been classified into four phenotypic sub-types, which vary in their severity. An age-dependent progression of disease from the mild to the severe clinical sub-types has been proposed, but this was not consistent with the clinical presentation observed in our study.

MCD was the most common recessive dystrophy with 20 families recruited in total; of which 18/20 had biallelic *CHST6* mutations and 2/20 harboured a single mutation in *CHST6* (section 3.6). Diagnosis of MCD was highly accurate, with the exception of a single proband that was initially diagnosed with an endothelial dystrophy, however, following genetic investigation two *CHST6* mutations were identified and a differential diagnosis of MCD was made.

Six families were recruited with SCD and the cohort was supplemented with two additional families from the Czech Republic; all were positive for heterozygous mutations in *UBIAD1* (section 3.7). It has previously been reported that ~50% of patients with SCD do not have crystalline deposition, instead presenting with cloudy corneas and corneal arcus (Weiss, 2009), however in our cohort, all patients had at least some minimal crystalline deposition. One possible explanation is that those without crystalline deposition were not diagnosed with SCD and therefore not recruited to the study, in which case this is likely to be an underestimate of the number of probands with SCD. The mutations identified in *UBIAD1* include those previously reported with both the crystalline and the non-crystalline phenotype, so it appears unlikely that there is any genotype-phenotype correlation, although an environmental effect cannot be excluded.

6.1.2 Endothelial corneal dystrophies

Of the 28 families with a diagnosis of PPCD recruited from MEH, heterozygous coding mutations in *ZEB1* were identified in 7 families (section 4.3), one had a heterozygous *ZEB1* encompassing deletion (section 4.4), 2 were found to have heterozygous promoter mutations in *OVOL2* (section 4.5), and three were potentially caused by heterozygous mutations in the putative *GRHL2* promoter region (section 4.6). Of the 14 Czech families with PPCD (excluded for *ZEB1* coding mutations prior to this study), 2 families had heterozygous *ZEB1* deletions, one large founder family (consisting of 16 individual pedigrees) was caused by an ancestral *OVOL2* promoter mutation and 4 families had a putative *GRHL2* mutation. A single large pedigree was recruited from MEH with a clinical diagnosis of CHED1 which was also caused by a heterozygous *OVOL2* promoter mutation (section 4.5). Only two families with CHED2 were included in this study, one with biallelic mutations in *SLC4A11* and another in which a single *SLC4A11* mutation was identified (section 4.8). However, this diagnosis was not prioritised for recruitment and therefore does not represent prevalence.

There appeared to be no robust genotype-phenotype correlation for patients with PPCD. Individuals with mutations in *ZEB1* occasionally have a characteristic appearance of corneal steepening and rarely, associated systemic symptoms, most commonly inguinal hernia (Aldave *et al.*, 2013). This was consistent with our *ZEB1* mutation positive cohort (section 4.3). However, of the four PPCD families with *OVOL2* mutations, two were mildly affected, with one family not requiring corneal surgery, whereas at the other end of the spectrum, the MEH family F140 originally diagnosed with CHED1 were so severely affected that the majority required surgery (section 4.5.2). In addition, two of the four families had symptoms of anterior segment dysgenesis, although further research is required to identify whether this is a common complication associated with *OVOL2* mutations (section 4.5.2: Figure 4.16). Different severities and clinical signs were also observed in probands harbouring *GRHL2* variants (section 4.6.4: Figure 4.22). The lack of genotype-phenotype correlation associated with non-coding mutations affecting *OVOL2* and *GRHL2* may be due to their mechanism of pathogenicity. It is hypothesised that mutations in the promoter or regulatory regions of these genes disrupt the binding of

transcription factors and/or produce novel transcription factor binding sites that ultimately result in the ectopic expression of *OVOL2* or *GRHL2* in the corneal endothelium (section 4.5.6, 4.6.7). However, as the variants differ in type and location, with some single nucleotide substitutions and some larger duplications or deletions, it is likely that the variants have different consequences and produce a unique pattern of transcription factor binding. This is supported by transcription factor prediction tools that predict a loss and gain of different transcription factors for different variants. The binding of different transcription factors would be expected to cause different effects on the expression of *OVOL2* or *GRHL2*, with the affected gene expressed at different levels and/or at different developmental timepoints, or persistent expression into adulthood. These differences in expression may explain the differences in clinical phenotype between individuals harbouring different variants.

No patients were recruited with a clinical diagnosis of XECD (section 4.1.3). However, given the overlap in the phenotypic spectrums of XECD and PPCD, it is possible that a proportion of the unsolved PPCD cases in the study cohort are XECD (Schmid *et al.*, 2006). XECD is an X-linked dystrophy, therefore an affected male cannot inherit the disease from his father and father to son transmission of the dystrophy can be used to exclude XECD as a potential diagnosis. Typically for X-linked disorders, females are unaffected or affected with milder symptoms due to random X inactivation of one of the two X chromosomes. One genetically unsolved family (F154) with the PPCD-band type phenotypic variant of PPCD included a proband that was initially thought to be sporadically affected, however, on clinical examination the proband's mother was found to have minor clinical changes of her endothelium, despite being asymptomatic (section 4.7). This family is therefore a potential candidate for a diagnosis of XECD.

6.1.3 Syndromes

Seventeen families were recruited to the study with a corneal dystrophy or a corneal opacity of unknown aetiology and mutations in corneal dystrophy genes were excluded in these individuals. However, rather than indicative of additional locus heterogeneity the evidence presented in this study suggests that the majority of these

probands have an inherited syndrome or environmental phenocopy of a true corneal dystrophy (chapter 5).

Misdiagnosis of corneal dystrophies occurred exclusively for anterior corneal dystrophies with some syndromes and phenocopies found in multiple probands in the corneal dystrophy cohort. In comparison, no endothelial dystrophies were found to be syndromic or phenocopies. However, genetically unsolved PPCD cases were predominantly of the PPCD-band clinical sub-type, some unilaterally affected and the majority without a family history of corneal disease, therefore it is possible that the PPCD-band sub-type may be a non-genetic phenocopy of PPCD. Identification of novel corneal dystrophy genes.

6.1.4 Locus heterogeneity for anterior corneal dystrophies

Previously, locus heterogeneity has been suggested for MCD and GDLD due to negative genetic screening of probands for mutations in *CHST6* and *TACSTD2*, respectively (Alavi *et al.*, 2007; Birgani *et al.*, 2009). In the study cohort, both GDLD families harboured biallelic mutations in *TACSTD2;* therefore there was no evidence for additional locus heterogeneity (section 3.5). Similarly, there was no evidence for additional locus heterogeneity for MCD. Although a definitive genetic diagnosis was not obtained for all 20 probands, all were found to harbour at least one mutation in the MCD gene, *CHST6*, therefore it is likely that disruption of *CHST6* is responsible for all MCD cases in the study cohort, with no additional contributing locus (section 3.6). As the coding region of the *CHST6* gene was covered by direct sequencing this missing variation is likely to be in a non-coding region. Alternatively, although a PCR-based assay was used to identify recurrent deletions occurring upstream of the *CHST6* gene, it is possible that another structural mutation (deletion or rearrangement) that is not encompassed by these primers has occurred in some unsolved patients. In future research, WGS of the unsolved MCD patients, with analysis restricted to the *CHST6* locus should be used to identify the missing variation.

With the exception of rare anterior corneal dystrophy diagnoses including SMCD and CCDF, none of which were present in the MEH study cohort, most anterior

corneal dystrophies are genetically solved. However, two probands were identified with clinical signs of EBMD and a family history of corneal disease, who were negative for mutations in both *TGFBI* and *COL17A1*; this is the first evidence to suggest additional locus heterogeneity for patients with EBMD (section 3.3). Following WES analysis of these families, no variants were identified in any other corneal dystrophy gene and no candidates were identified for genes known to cause syndromes associated with a corneal phenotype, in either proband. However, it should be noted that the causative mutation may have been missed if it occurred within a region of low coverage, or if it occurs within a gene that was not included in the corneal syndromes candidate list. There was also limited evidence for an additional locus for ERED (section 3.4), however due to a lack of primary opacities associated with this diagnosis and the variety of other potential traumatic and environmental causes, clinical diagnosis is difficult for this condition. Only one proband with genetically unsolved ERED had a family history of corneal disease (an affected father) but it was not certain whether his corneal symptoms were related to the disease in the proband. For the majority of cases of anterior corneal dystrophy where locus heterogeneity was considered to be a potential option, further investigation revealed a likely syndrome or phenocopy that was consistent with the symptoms of the proband (section 5).

6.1.5 Identification of novel corneal endothelial dystrophy genes

Prior to this study, there was at least one additional locus for PPCD (PPCD1), which overlapped with the locus of another endothelial dystrophy, CHED1. Using WGS analysis, and combining data from two families with these disorders, we have now shown that mutations in the promoter of *OVOL2* are causative of both these endothelial dystrophies (section 4.5). We have also found evidence for another new locus for PPCD with linkage to chr8q (PPCD4; section 4.6). Following WGS analysis and targeted sequencing, the most likely cause is variants within a potential promoter sequence of *GRHL2*, although further work is required to genetically and functionally validate this discovery.

There is evidence to suggest that there may be additional endothelial dystrophy genes, including XECD, and possibly additional locus heterogeneity for PPCD with

14 genetically unsolved PPCD probands from MEH and 7 from the Czech Republic (section 4.7). However, the majority of these probands are sporadically affected, although for most probands additional family members have not been clinically examined. Recruitment of pedigrees with multiple affected individuals, and clinical assessment of relatives to determine whether they are affected or unaffected, are required to explore the possibility of additional genetic heterogeneity for PPCD. Some of the genetically unsolved cases may be due to deletions that did not encompass the 5' end of the *ZEB1* gene targeted by qPCR primers or non-coding variants occurring in promoter or intronic sequences associated with *OVOL2* or *GRHL2* that were not targeted by direct sequencing that affect transcriptional regulation or splicing of the gene. These cases will likely require SNP genotyping and CNV analysis and/or WGS to identify the causative change.

6.1.6 Molecular pathways governing the maintenance of corneal transparency

The identification of novel corneal dystrophy genes can provide important insight into the understanding of the molecular pathways underlying corneal transparency. This has been demonstrated previously for the majority of genetically solved anterior corneal dystrophies. We are now beginning to understand the crucial role that stromal proteoglycans play in maintenance of corneal transparency. These bind to the collagen fibrils and are involved in maintaining regular collagen fibril spacing (Meek and Boote, 2004). Deletion of the stromal proteoglycans decorin, lumican, keratocan and mimecan cause PACD while mutations of decorin (*DCN*) cause PSCD. Furthermore, mutations in keratocan (*KERA*) cause another inherited corneal disorder; corneal plana (Pellegata *et al.*, 2000; Bredrup *et al.*, 2005; Kim *et al.*, 2014). These stromal proteoglycans have GAG side chains of chondroitin/dermatan sulphate or keratan sulphate. The generation of keratan sulphate requires the enzyme encoded by *CHST6*, and biallelic mutations in this gene are associated with MCD. Loss of CHST6 enzyme in the cornea results in production of unsulfated keratan sulphate, and disorganisation of the stromal matrix.

For other solved corneal dystrophies, the contribution of the protein to normal corneal transparency, if any, remains unknown. TGFBI protein accumulates in the cornea due to specific gain-of-function mutations that affect the protein's

thermodynamic and proteolytic stability (Runager *et al.*, 2011; Underhaug *et al.*, 2013). Increases in stability result in formation of amorphous granular opacities, while decreases in stability release a fibril-core that forms the amyloid deposits observed in LCD (Sorensen *et al.*, 2015). Despite ubiquitous expression of *TGFBI*, protein deposits caused by *TGFBI* mutations are restricted to the cornea (El Kochairi *et al.*, 2006). This may be due to the abundance of TGFBI in the cornea, which makes up ~1% of corneal epithelial protein, ~17.6% of stromal protein and is the most abundant protein in the endothelium, making up ~36.8% of the total protein (Dyrlund *et al.*, 2012). Evidence suggests that the origin of TGFBI protein opacities for epithelial-stromal *TGFBI* dystrophies is the corneal epithelium, which is then deposited in Bowman's layer and the stroma (Ridgway *et al.*, 2000). Alternatively, the composition of the corneal extracellular matrix or the lack of vascularisation in cornea may explain why the *TGFBI* protein accumulates only in the cornea.

In this thesis, two endothelial corneal dystrophy genes encoding transcription factors were identified (*OVOL2* and *GRHL2*) that play crucial roles in the process of regulation of EMT/MET, and directly interact with another corneal dystrophy gene, *ZEB1*. These three transcription factors form a network in which ZEB1 negatively regulates the expression of both *OVOL2* and *GRHL2*, and is, in turn, negatively regulated by both OVOL2 and GRHL2 (section 4.9.5: Figure 4.31). While EMT has known roles in development, wound healing and cancer metastasis, it has not previously been associated with maintenance of corneal transparency, and this provides an important area of future research to determine how this functional network of transcription factors, and the EMT/MET pathway in general, contribute to normal corneal function. The mechanism by which this network influences corneal transparency is by maintenance of the regular, hexagonal monolayer of corneal endothelial cells, which forms a barrier that maintains corneal deturgescence. When this network is disrupted, corneal endothelial cells change their morphology, express proteins typically seen in epithelial cells, and show focal multilayering potentially via a partial induction of MET (section 4.6.8). This change in behaviour can lead to disruption of water homeostasis and loss of corneal transparency due to corneal oedema.

6.2 Strategies to obtain a genetic diagnosis for corneal dystrophies

6.2.1 Sanger sequencing of corneal dystrophy genes

In this study, a large proportion of the corneal dystrophy cohort was genetically solved by targeted Sanger sequencing of the gene(s) associated with their clinical diagnosis. Sanger sequencing is a quick and simple method that can provide confirmation or exclusion of clinical diagnosis (section 1.6.2) (Sanger, Nicklen and Coulson, 1977). This was demonstrated predominantly for anterior corneal dystrophies, for which the majority of clinical diagnoses have been genetically solved, with the associated causative gene identified and with little evidence for additional genetic heterogeneity (chapter 3). There is also genotype-phenotype correlation such that most clinical diagnoses are associated with a specific identifiable phenotype. However, while the majority of the clinical diagnoses were genetically confirmed, a sub-set of individuals were identified who were negative for mutations in the associated corneal dystrophy gene and subsequently found to have an inherited syndrome or phenocopy of disease (chapter 5). On the basis of these findings, genetic confirmation of an anterior corneal dystrophy diagnosis should be sought, even in cases where the clinical features appear to be completely consistent with the diagnosis. In comparison, a significant proportion of probands with endothelial dystrophies remained unsolved following Sanger sequencing of known genes, most likely due to the less clinically distinct phenotypes caused by mutations in endothelial dystrophy genes and greater genetic heterogeneity (chapter 4).

6.2.2 Whole exome sequencing

In recent years, WES has become a routinely used genetic technique for a variety of purposes both in research and clinical practice (section 1.6.3). It has been used to identify rare variants which may contribute to the pathogenesis of complex, multifactorial disorders (Cirulli and Goldstein, 2010), and can be used clinically for genetic diagnosis of Mendelian disorders. WES is especially important for conditions with high genetic heterogeneity, where Sanger sequencing of all potential causative genes would be time consuming and expensive (Rabbani, Tekin and Mahdieh, 2014). WES has been extensively used to provide a genetic diagnosis for

cases where the clinical diagnosis is not known, or where there are potential differential clinical diagnoses (Yang *et al.*, 2013). WES has also replaced traditional linkage and candidate gene re-sequencing approaches for novel Mendelian disease gene discovery (Bamshad *et al.*, 2011). Despite covering only ~ 1% of the entire genome, ~ 85% of Mendelian disease gene mutations are estimated to occur in protein coding regions, therefore it is an excellent genetic technique for investigating monogenic diseases, including corneal dystrophies (Rabbani, Tekin and Mahdieh, 2014). WES was used in this study to investigate individuals diagnosed with corneal dystrophy who were negative when pre-screened for mutations in known corneal dystrophy genes (sections 3.3 – 3.4, 4.5 – 4.7, 5.3 – 5.6). Analysis was carried out by first prioritising rare variants in any known corneal dystrophy gene or gene which causes an inherited syndrome with corneal involvement, with the aim of providing a genetic diagnosis, and therefore a clinical diagnosis for the proband. If this approach was negative, the WES data was interrogated to look for novel genetic causes of corneal disease.

One challenge associated with the use of WES analysis to identify causative mutations is prioritisation of the vast number of variants identified in each proband. On average, WES identifies ~ 24,000 variants for African Americans and ~ 20,000 variants in European Americans (Bamshad *et al.*, 2011). Filtering by frequency is a standard approach for identification of mutations causing rare Mendelian disease as any variant with a high frequency in the control population could not be responsible for the disorder in the proband. Dominant disease-causing mutations would be expected to be absent from control populations or present at extremely low frequencies, whereas recessive disease alleles may be present in the heterozygous state in the control population at a low frequency, which differs depending on the rarity of the disease and potentially ethnicity. With the decreasing cost of WES, large scale re-sequencing projects have been undertaken to catalogue human genetic variation and to investigate genetic causes of disease. Some of these projects have released their data for public use including the 1000G project and EVS. ExAC has collated data from these and a range of other studies and released the data from ~ 60,000 samples (Lek *et al.*, 2016). For this corneal dystrophy study, the bioinformatics pipeline used annotated variants with their frequency the 1000G and EVS datasets, along with their frequency in an internal UCL exome dataset. Any

variant considered to be a possible candidate for disease was manually interrogated in the ExAC database for their allele frequency. Furthermore, several bioinformatics tools have been designed to predict the functional consequences of mutations occurring in coding or splice sites (Ng and Henikoff, 2003; Adzhubei *et al.*, 2010). Two of these, SIFT and Polyphen2, were incorporated into the bioinformatics pipeline used in this study, although due to their limitations they were used to support the assessment of variants rather than as a specific filtering criteria.

Filtering for rare variants is common to all filtering strategies used in the study; however additional filtering criteria were used depending on the likely pattern of inheritance, phenotype, availability of other unrelated probands with the same condition, and availability of DNA from additional family members. One previously reported successful strategy is to compare exomes of multiple unrelated probands with the same disorder for rare variants in shared genes (Ng, Bigham, *et al.*, 2010; Ng, Buckingham, *et al.*, 2010). This is particularly useful when the disease locus has not been mapped. This approach was used to initially analyse variants for the two *TGFBI*-negative EBMD families; however, no shared rare variants were identified which overlapped between the two probands (section 3.3). One possibility for the lack of shared rare variants is locus heterogeneity for EBMD with different genetic causes in each family. This could also explain why one EBMD family had an autosomal dominant pattern of inheritance whereas the other family appeared to be autosomal recessive. An alternative explanation is lack of coverage in some exons within the unknown disease gene so that the variant was not called in one or two of the families. Comparison of multiple unrelated unsolved probands was also used for PPCD and identified that three probands with PPCD harboured a rare variant in a gene, *PCDH1,* although the relevance of this to disease in unresolved (section 4.7.2).

In cases where the pedigree had multiple affected individuals who had been analysed by WES, a family based filtering strategy was applied. This facilitated filtering for rare variants that were shared between affected family members (sections 3.3). This approach typically resulted in a candidate list of multiple variants that was further refined by filtering for those genes that encoded proteins that were present in the corneal proteome (Dyrlund *et al.*, 2012). One advantage of using the proteomics data is that it included a significant proportion of proteins that appear to be deposited in

the cornea from the blood plasma, despite the avascularity of cornea, however; proteins that are expressed at low levels may not be detected.

Filtering for rare variants in overlapping genes shared by unrelated individuals was not a suitable filtering strategy for other anterior corneal dystrophy probands due to the lack of patients recruited to the study with a similar clinical phenotype; therefore, an alternative candidate gene approach was used. Variants occurring within genes which have been previously identified to cause a corneal dystrophy or genes which cause a syndrome associated with a corneal phenotype were filtered using a manually curated list of candidate genes and prioritised (sections 5.3 – 5.6). This strategy typically identified fewer than 10 variants for further assessment, per proband. Both inheritance pattern and clinical phenotype of the proband was compared to that associated with the candidate gene. In cases where the variant was a good candidate, further clinical investigation or segregation analysis was undertaken to confirm or exclude the diagnosis. For the majority of anterior corneal dystrophy cases analysed using this filtering strategy, at least one good candidate variant was identified which was considered potentially likely to contribute to their corneal phenotype.

There are limitations of using WES to identify genetic causes of corneal dystrophies, and indeed for other conditions. One is the lack of coverage of some coding exons which may lead to false negative results. To overcome this in this study, where the gene was consistent with the clinical phenotype of the proband, missing exons were covered by Sanger sequencing; for example, manual sequencing of *COL17A1* exons not covered by WES in probands with ERED (section 3.4). However, these cannot be negated when using WES for novel gene discovery and not all missing exons in corneal dystrophy gene were directly sequenced, therefore this could result in a false negative result in some cases. Another limitation is lack of ability to interrogation of non-coding regions and the difficulties of identifying CNVs using WES data (Bamshad *et al.*, 2011).

6.2.3 Whole genome sequencing

As discussed in the previous section, one of the limitations associated with WES is the lack of coverage of some exons which means that some mutations, in known corneal dystrophy genes or novel genes, may be missed. Although more expensive, WGS is superior to WES for identification of variants in protein-coding regions of the genome due to greater coverage of coding exons (Belkadi *et al.*, 2015). It is now becoming more common and routine to select WGS of samples over WES, with data analysis initially restricted to the coding regions of the genome. Analysis of non-coding regions may then be undertaken in cases which are negative for mutations in coding regions.

In this study, WGS was undertaken for three large families (one British, two Czech) with PPCD. Due to prior knowledge of the locus obtained by linkage analyses (chr20p and chr8q), variants could be filtered by frequency and restricted to those variants located within the disease locus, vastly reducing the number of potential disease-causing variants (section 4.5.1 – 4.5.2, 4.6.3). Although there are fewer bioinformatic prediction tools available for non-coding variants, their number is growing. In this study, VEP was used to identify variants which occurred within predicted regulatory regions, and included promoter regions and non-coding transcripts which may have a regulatory function. For both the PPCD3 and PPCD4 locus, novel variants were identified which were predicted to occur in the promoter of *OVOL2* and *GRHL2*, respectively, and likely causative when investigated with additional bioinformatics tools and experimental data. The successful use of WGS to identify these non-coding variants was dependent on prior knowledge of the locus, however linkage analysis can only be performed on large pedigrees, and therefore this strategy cannot be used for all probands and small families (section 1.6.1) (Strachan and Read, 2011).

Several PPCD probands remained genetically unsolved following both targeted Sanger sequencing of associated PPCD genes and WES. WGS could be performed in these cases; however, analysis would be restricted to known endothelial corneal dystrophy genes. Additional candidate gene approaches could be adopted such as rare variants in shared genes or more extensive WGS of other family members such

as parents (trios) and unaffected siblings. Clinical examination of asymptomatic individuals would be essential, and age of onset and clinical variability would also need to be considered.

Although WGS is becoming more commonly used for genetic studies, there are still fewer studies which choose to use WGS in comparison to the cheaper WES approach with more manageable data files, and therefore there is less publicly available data available to use as control data. However, this is likely to increase in the future as there are vast amounts of both WES and WGS data being generated each year as the cost of next-generation sequencing is expected to fall, and the sharing of anonymised genetic data for research purposes will likely increase (Belkadi *et al.*, 2015). The majority of publicly available WGS data is from white European cohorts; therefore, it is important that WGS data is available for control purposes from individuals from multiple ethnicities in order to distinguish benign population specific variants from novel pathogenic variation.

When mutations in protein coding genes have been eliminated, WGS offers the potential to investigate non-coding variation, however currently our lack of functional understanding of non-coding regions hinders their interpretation. The sheer number of variants generated in a single proband by WGS means that filtering based on frequency (novel variants) alone would still provide an overwhelming number of variants to manually assess. While monogenic deep intronic mutations have been described previously, the majority of these affect splicing (Liquori *et al.*, 2015; Stepensky *et al.*, 2016). There are a variety of algorithms which predict whether a variant is likely to affect splicing, and there may be some merit for these to be included in a standard bioinformatics pipeline of variant annotations (Hebsgaard *et al.*, 1996; Reese *et al.*, 1997; Desmet *et al.*, 2009). However, in this study, non-coding mutations in both novel corneal dystrophy genes do not appear to affect splicing and instead appear to be influencing regulation of gene expression.

6.2.4 Challenges of interpreting genetic data

Several online databases aim to catalogue human genetic variation, and the majority of these distinguish benign variants from disease-causing pathogenic mutations

including The Human Gene Mutation database (www.hgmd.cf.ac.uk) and dbSNP (https://www.ncbi.nlm.nih.gov/SNP/). However, as more sequencing data is being generated, it is becoming apparent that many of the mutations reported historically in the literature are rare benign variants, or population specific variants, which are not causative of disease (Richards *et al.*, 2015). In some cases, even where the scientific record is corrected, the databases continue to report the variant as a disease-causing change. These false positives can mistakenly result in the wrong diagnosis being given to the clinician while the true genetic cause goes undiscovered. A number of false positives were identified within this study. For example, a variant in *MYOC* [c.1102C>T; p.(Gln368*)] was identified in a subset of affected individuals in F144 with PPCD, that is reported as causing glaucoma, which if true, could plausibly be responsible for some or all their ocular symptoms, however ExAC revealed a minor allele frequency of frequency of 0.001, which is high for a dominantly inherited disease (section 4.6.1). In addition, many databases continue to report *VSX1* mutations as causative of PPCD including the The Human Gene Mutation Database although there is currently little evidence to suggest that this is the case.

Another challenge of working with large genetic datasets is the identification of variants of uncertain significance (Richards *et al.*, 2015). These are variants which cannot be reasonably excluded, but there is not enough supporting evidence to plausibly deem the variant causative. One example of this is the homozygous *TGFBI* variant identified in F180 (section 5.2.3). The proband was referred for genetic screening to distinguish between two differential diagnoses – a *TGFBI*-associated LCD or climatic keratopathy. The identification of the variant of uncertain significance meant that neither of these two clinical diagnoses could be excluded. Further confirmation of a causative role for the *TGFBI* variant would require segregation analysis in the extended family or functional validation. A further variant of uncertain significance was identified in proband F187, in which a heterozygous variant was identified in *APOA1* that may or may not contribute to the proband's autosomal dominantly inherited cloudy corneas (section 5.5.1).

Incidental findings are genetic findings which are not causative of the condition which is being studied, but may contribute to the patient's health either now or in the future (Richards *et al.*, 2015). There are a number of ethical issues regarding the

reporting of incidental genetic findings, as the consequences can involve the entire family, not only the person who has consented to the study. In this study, only the most likely contributing variant(s) were reported to the referring clinician as a research finding, allowing further confirmation to be sought by an accredited clinical genetics laboratory if the patient wished to pursue this and access further genetic services. Incidental findings were not reported, however, in the future if a patient's genetic data is interrogated by their clinician directly, ethically they may be compelled to inform a patient of a finding which is relevant to their health and is actionable.

6.3 Gene-based therapies for corneal dystrophies

Genetic diagnosis is a pre-requisite for development of gene therapies for corneal dystrophies. The cornea has several features which make it amenable to the development of gene directed therapeutics (Klausner *et al.,* 2007). It is an easily accessible tissue, allowing any therapeutic to be delivered in a non-invasive manner in the form of eye drops, contact lens mediated delivery or injection. As a bilateral organ, one eye can be treated while the untreated eye acts as an internal control for clinical trials, and there are a wide variety of imaging techniques for measuring various aspects of the corneal transparency and function; therefore, the clearance of opacities from the cornea, or lack of accumulation of amyloid and non-amyloid deposits, can be easily monitored and tracked.

There is a lack of available mouse models for anterior corneal dystrophies, which may hinder efforts to develop gene or mutation specific therapeutics. Attempts to generate *TGFBI* mouse models using the mouse *TGFBI* sequence have not been successful in recapitulating the clinical phenotype (Bustamante *et al.,* 2008). A single mouse model has been described for epithelial-stromal *TGFBI* dystrophies; generated using a knock-in method that replaced the mouse *TGFBI* gene with the human cDNA sequence containing the GCD2 associated mutation p.(Arg124His) (Yamazoe *et al.,* 2015). In this model, 42.1% of heterozygotes and 75.0% of homozygotes older than 80 weeks had corneal opacity, and were absent from WT mice. Consistent with the corneal phenotype in humans, the opacities were predominantly granular with some lattice lines, and were confirmed as hyaline and

amyloid in nature by Masson's trichrome and Congo red staining. Overexpression of human WT *TGFBI* cDNA in mice has also been shown to result in reduced corneal transparency (Liao, Cui and Wang, 2013). Another humanized knock-in corneal dystrophy mouse model has been generated harbouring a mutation causing MECD (Allen *et al.*, 2015). This produced a corneal phenotype in homozygotes when compared to WT mice by histology and electron microscopy, including thickening of the corneal epithelium and prominent intracellular spaces, but did not result in any observable corneal surface defects in living mice. In comparison, mouse models have been developed for several corneal endothelial dystrophies that recapitulate the human phenotype using knock-out and knock-in strategies including PPCD1 (*ZEB1*), CHED (*SLC4A11*), and early-onset FECD (*COL8A2*) (Liu *et al.*, 2008; Meng *et al.*, 2013; S. Han *et al.*, 2013).

6.3.1 Traditional gene therapy approaches

The traditional gene therapy approach aims to restore functional WT protein to the affected tissue for recessive loss-of-function diseases. This can be achieved by introduction of the open reading frame encoding the protein, driven by a suitable promoter and delivered by a viral or non-viral vector. For viral mediated delivery, viral genes are removed to prevent the virus from replicating, and allow the transgene to be inserted in the viral genome. The most common viral vectors which have been developed for gene therapy include both DNA (adenoviruses, adeno-associated viruses) and RNA viruses (retroviruses), and each are associated with their own advantages and disadvantages (Vannucci *et al.*, 2013).

Adenoviruses are non-enveloped, non-integrating double stranded DNA viruses which are delivered to the nucleus of cells and produce high levels of transient transgene expression. They allow up to 36 kb of DNA to be delivered and can transduce non-dividing or dividing cells. However, there are some problems including dose-limiting effects due to pre-existing neutralising immunity (Appaiahgari and Vrati, 2015). Furthermore, they have a high immunogenic potential; one gene therapy clinical trial using an adenoviral vector resulted in the death of a patient due to acute respiratory distress syndrome following a systemic inflammatory response and multi-system organ failure due to an excessive dose

(Raper *et al.*, 2003). Adeno-associated viruses (AAV) are non-integrating single stranded DNA viruses that require helper viruses to complete their replication cycle. They have not been associated with disease, have low immunogenicity and can infect dividing and non-dividing cells, however they can only can deliver up to 5 kb of transgene and pre-existing immunity may pose a problem (Badalà, Nouri-mahdavi and Raoof, 2008).

Retroviruses are single stranded RNA viruses that can deliver transgenes up to 8 kb in size. They integrate in the cellular genome to provide long term transgene expression as daughter cells produced by replication will also contain the replicated transgene (Vannucci *et al.*, 2013). However, they can only infect dividing cells which excludes their use in stromal keratocytes or the corneal endothelium but theoretically should be an option for the corneal epithelium (section 1.2). They have relatively low immunogenicity, and little problems with pre-existing immunity but integration of the transgene introduces the additional risk of insertional mutagenesis, which occurs when the integrating transgene disrupts a gene present in the host's genome. In a previous clinical trial, insertional mutagenesis resulted in cellular transformation and leukemia in 4/9 patients, resulting in one death (Hacein-Bey-Abina *et al.*, 2010). Lentiviruses are a sub-type of retroviruses which show promise for gene therapy for a variety of tissues. They can infect both dividing and non-dividing cells and can deliver up to 9 kb of DNA. They have a lower risk of insertional mutagenesis than other retroviral vectors as non-integration-defective vectors have been developed, but still provide long term stable transgene expression (Vannucci *et al.*, 2013).

Viral vectors can achieve high efficiency transgene delivery and the use of specific promoters (e.g. *KRT12* promoter for the corneal epithelium) can be used to target specific cells in which the transgene is to be expressed; however, there remain safety concerns with their use. Therefore, several non-viral vectors have also been developed as alternatives to viral delivery, including naked DNA, plasmid DNA, cationic liposomes, minimalistic immunologically defined gene expression vectors, and nanoparticles (Klausner *et al.*, 2007; Mohan *et al.*, 2012). Although less toxic and safer, most non-viral vectors tend to have lower transfection efficiencies in comparison to viral-mediated delivery. Administration into the cornea can be

achieved by topical application, injection (subconjunctival, intrastromal, intracameral, intravitreal), electroporation, gene gun, sonoporation or *ex vivo* transduction.

To date, traditional gene therapy pre-clinical research in the cornea *in vivo* in animal models has focused on the treatment of acquired corneal disorders including corneal graft rejection, herpetic stromal keratitis, vascularization and post PTK-haze (Klausner *et al.*, 2007; Mohan *et al.*, 2012). Despite a wide range of pre-clinical studies in a range of animal models, using a variety of different transgene delivery methods, no clinical trial has been published which aims to introduce a transgene into the cornea. However, other ocular disorders have demonstrated that the gene therapy approach can be successful and gene therapy for retinal disorders has made significantly more progress, in comparison, largely due to the availability of suitable animal models. For example, an AAV based gene therapy is currently in phase III clinical trials for the treatment of Leber's congenital amaurosis (Sengillo *et al.*, 2016).

The most suitable target for the development of a traditional gene therapy approach of the corneal dystrophies is MCD (section 3.6). This was the most common cause of recessive corneal dystrophy in the MEH cohort, due to biallelic mutations in *CHST6*. As stromal keratocytes are affected, the most suitable approach would be injection and delivery using a viral vector which is able to infect slowly dividing and non-dividing cells such as AAV or a lentivirus. In the MEH cohort, the majority of mutations were private to the affected family; however the benefit of traditional gene therapy is that the same therapeutic protocol can be used to treat individuals with different underlying causative mutations in the same gene as it involves the restoration of functional protein, transcribed from the transgene.

6.3.2 Oligonucleotide gene silencing

The basis of oligonucleotide-based therapies is the process of RNA interference (RNAi). RNAi was first discovered in *c.elegans* as a mechanism of defense against viruses which does not activate the innate immune system (Wilkins *et al.*, 2005). RNAi is triggered by the presence of long double stranded RNA or short hairpin

double stranded RNA within the cell which is targeted by an enzyme called Dicer. Dicer generates a small interfering RNA (siRNA) which is an oligonucleotide, typically 21 nt in length, consisting of 19 nt of double stranded RNA with 2 nt 3' overhangs. siRNAs associate with RISC (RNA-inducible silencing complex) and Argonaute-2 protein to form activated RISC. The antisense siRNA strand provides the template for the active RISC which cleaves complementary mRNA transcripts and thereby prevents their translation. Once discovered, it was quickly realized that this system could be utilised as both a powerful experimental tool and a potential clinical therapeutic (Kim and Rossi, 2007). The specificity of RNAi allows a mutant allele to be silenced while the WT allele remains intact; therefore it is potentially treatment for dominant gain-of-function or dominant negative disorders. However, oligonucleotide-based therapeutics suffers from some of the same problems as traditional gene therapy with regards to the delivery of the oligonucleotide to the tissue of interest, and preventing its degradation once delivered. An additional potential problem associated with RNAi is off-target effects; which is the silencing of other mRNA transcripts which share a degree of complementarity with the transcript of interest (Jackson and Linsley, 2010).

The most common cause of corneal dystrophies in patients attending MEH was mutations in *TGFBI*. Despite over 60 reported mutations, and an ethnically diverse patient population, a limited spectrum of 9 mutations was responsible for all cases of *TGFBI*-associated corneal dystrophies with mutations affecting residues Arg-124 and Arg-555 responsible for 89.7% of disease (section 3.2.2). Due to the large patient cohort at MEH and limited spectrum of *TGFBI* mutations, the *TGFBI* associated epithelial-stromal dystrophies are an excellent candidate for the development of a gene mutation targeted therapy. As *TGFBI* mutations are considered to have a toxic gain-of-function effect and are heterozygous, one potential therapeutic strategy could be to use an oligonucleotide therapy to silence the mutant allele, while allowing the *TGFBI* transcript to be expressed from the WT allele. For this therapeutic strategy to be suitable, haploinsufficiency of the gene must not be associated with any clinical disorder. Twenty eight likely loss of function alleles in *TGFBI* have been reported in ExAC, however clinical examination of individuals harbouring one of these alleles would be preferable to confirm the absence of a clinical phenotype. The selective silencing of the *TGFBI*

p.(Arg124Cys) allele using an oligonucleotide has been demonstrated in cells transiently transfected with *TGFBI* constructs (Courtney *et al.*, 2014).

Although no patients were recruited with MECD from our cohort, others studies have targeted the MECD associated genes for oligonucleotide therapy. A founder *KRT12* mutation [p.(Arg135Thr)] common in Europeans was selected, and all possible siRNAs were screened for suppression of the mutant allele with minimal repression of expression of the wild type allele using transiently transfected *KRT12* constructs (Allen *et al.*, 2013).

UBIAD1 is also a potential target, albeit with a smaller potential pool of families than *TGFBI*. Of the 5 *UBIAD1* mutations identified in 8 SCD patients, 4 had been previously reported, with one mutation associated with over twenty families in the literature. The *de novo* occurrence of some of these mutations argues for a mechanism by which recurrent mutations occur at the same residue and therefore are likely gain-of-function; however, this has not been proven for *UBIAD1* mutations. Furthermore, the corneal proteome revealed a number of corneal proteins that have a potential origin from blood plasma, therefore although *UBIAD1* is expressed by corneal cells, mutant UBIAD1 protein from other cellular sources may influence lipid deposition in the cornea which may limit the effectiveness of any locally applied therapy (Dyrlund *et al.*, 2012).

Dominantly inherited endothelial dystrophies are less suitable than anterior corneal dystrophies for gene therapy approaches due to the underlying pathogenic mechanism. Heterozygous mutations in *ZEB1* cause PPCD by haploinsufficiency therefore are not suitable for a knock-down strategy (section 4.4). Given that *OVOL2* and *GRHL2* promoter mutations appear to cause an upregulation of the respective genes expression (section 4.5 – 4.6), it is likely that a partial knockdown of expression would be of some benefit. However as the feedback system in which *ZEB1*, *OVOL2* and *GRHL2* operate is a complex, highly regulated network and sensitive to changes in dosage, we do not currently know the consequences of altering gene dosage in this system. Altering this network is likely to influence other proteins, and given the known role of this network in cancer, it could result in severe consequences (Cieply *et al.*, 2013).

Although development is in early pre-clinical stages for the use of RNAi for corneal dystrophies, other fields have demonstrated the safety of this approach in early human clinical trials for other inherited disorders. In 2010, the first clinical trial using siRNA was published. This was a phase Ib proof of concept study using siRNA to treat the inherited skin disorder plantar keratoderma in a single patient. No adverse effects were reported and there was a regression on the treated foot versus the vehicle treated (Leachman *et al.*, 2010). A number of siRNA based therapeutics are in various clinical trial stages, including ocular disorders (Guzman-Aranguez, Loma and Pintor, 2013). Based on these studies, it appears that this approach could be feasible for dominant negative and gain-of-function corneal dystrophies.

6.3.3 Gene editing

Gene editing using CRISPR-Cas9 is a relatively new technique that is revolutionising the field of molecular biology and in the future is likely to be a powerful tool for curing inherited disease; with the potential to treat both dominant and recessive disorders. Clustered regularly interspaced short palindromic repeat (CRISPR) sequences are components of an adaptive immune system against phage infection that is widespread in bacteria and archaea. CRISPR sequences have a distinct structure; stretches of tandem repeats with intervening non-repeat spacer sequences that are highly homologous to phage DNA, and nearby CRISPR associated (Cas) genes (Mojica *et al.*, 2005; Sorek, Kunin and Hugenholtz, 2008). Immunity to phage infection is acquired by incorporating additional spacer sequences of phage DNA into the CRISPR locus. For type II CRISPR systems, Cas9 nuclease confers phage resistance by cleaving the DNA of the target phage, guided by two RNA molecules. CRISPR RNA (crRNA) is a 61 nucleotide transcript transcribed from the CRISPR locus containing the target DNA sequence, and tracrRNA is a small molecule transcribed adjacently from the CRISPR locus, required for processing of crRNA and double stranded DNA cleavage. At the target DNA, double stranded breaks are generated 3 bp upstream from a proto-spacer adjacent motif (PAM; NGG) immediately following the spacer sequence (Gasiunas *et al.*, 2012; Jinek *et al.*, 2012; Lander, 2016).

The CRISPR-Cas9 components were subsequently adapted for use in eukaryotic cells. Using the Cas9 protein and by fusing the two RNAs into a single guide RNA, the complex could be used to induce double-stranded breaks at a target site (Ran *et al.*, 2013; Cong *et al.*, 2014). In eukaryotic cells, there are two possible repair mechanisms. The first is NHEJ which occurs when the double stranded break is re-joined without a template. This is an error prone process that typically results in the deletion or insertion of nucleotides. If the target DNA is a coding region then this indel can produce a frameshift mutation that introduces a premature stop codon and therefore produces a null allele. Alternatively, if a complementary sequence is available then the sequence may be used as a template for homologous directed repair. However, by co-transfecting a repair template with the CRISPR-Cas9 components, the target DNA sequence can be edited to change a specific nucleotide. This process occurs at a much lower efficiency than NHEJ, but as each cell repairs the break independently, can be achieved by screening multiple cell lines obtained from clonal expansion of single cells obtained from a transfection. CRISPR-Cas9 was rapidly adopted for gene editing in cell lines and various animal models. Recently, the first successful use of CRISPR-Cas9 for gene editing of human embryos was published (Liang *et al.*, 2015); however, the use of gene editing in human embryos has been restricted in many countries due to ethical considerations. While the human germline is unlikely to be edited in the near future, gene editing has a wide range of potential uses both for inherited and complex disorders by directly targeting the diseased tissue.

6.3.4 Cell based therapies

In addition to targeting the cornea *in situ*, the cornea is also suitable for *ex vivo* applications. For corneal dystrophies affecting the epithelium, cells could be removed from the patient's cornea in the form of a limbal biopsy and expanded *in vitro*. Following treatment with gene therapy, oligonucleotide, or CRISPR-Cas9 gene editing, the autologous cells could be transplanted back into the patient's eye in the form of a limbal stem cell transplant. Limbal stem cell transplants are currently used to treat limbal stem cell deficiency and anirida (Shortt, Tuft and Daniels, 2011). This would also be suitable for *TGFBI*-associated corneal dystrophies as the source

of protein expression is predominantly the epithelium as opposed to the stromal keratocytes.

It is more difficult to determine how this would be achieved for endothelial dystrophies, given their greater genetic heterogeneity and lack of corneal endothelial cell division. However, one way would be to obtain a skin biopsy and reprogram the keratinocytes into induced pluripotent stem cells (iPSC) and administer the gene based treatment. Protocols to induce the differentiation of corneal endothelial-like cells from human embryonic stem cells have been published, although further work is needed to determine whether they can function as a corneal endothelium when transplanted *in vivo* (Mccabe *et al.*, 2015).

6.4 Future research in corneal dystrophy genetics

There remain several unanswered questions arising from this thesis. Firstly, a small number of corneal dystrophies remain genetically unsolved, particularly for rare corneal dystrophies which we were unable to recruit at MEH. There are two explanations for the remaining genetically unsolved corneal dystrophies. It is likely that there are additional corneal dystrophy genes yet to be identified, but due to the lack of coding variants identified, non-coding mutations are likely to be the source of the missing genetic variation. In other cases, the identification of the causative gene may be hindered by lack of available families. Recruitment of additional unsolved families would be aided by collaboration with other research groups, as some dystrophies are so rare that few or no affected families may exist in the UK. LECD, for example, is unsolved despite a known locus, however most reported families live in Germany. Once recruited, genetic linkage followed by WGS analysis should be used to identify the likely causative change, once all known corneal dystrophy genes and potential syndromic causes or phenocopies have been excluded. This process will also be aided by the increase in publicly available control WGS genomes and exomes in future years, as the release of the ~ 60,000 WES exome data by ExAC has aided the interpretation of patient WES data. The possibility of additional genetic variation contributing to the EBMD phenotype, suggested by the research presented here, should be further investigated.

Given the limitations of currently available surgical treatments for corneal dystrophies and the many advantages of the cornea as a target for a gene or mutation-based therapy, future work should focus on the pre-clinical development of a potential therapeutic for the most common genetic causes of corneal dystrophies; the epithelial-stromal *TGFBI* corneal dystrophies.

6.5 Conclusions

This thesis presents the first systematic genotyping of a large cohort of corneal dystrophy patients attending MEH. The most common mutations have been described for the predominantly genetically solved anterior corneal dystrophies, and combined with clinical data, novel insights into genotype-phenotype correlation have been identified. Significant progress has been made investigating the genetic basis of endothelial dystrophies including the identification of two potential novel genes causing PPCD. Furthermore, unexpectedly several patients were identified within the corneal dystrophy cohort who did not have a corneal dystrophy, and were subsequently diagnosed with an inherited syndrome or a phenocopy of disease. Given the clinical implications of some of these diagnoses, it is essential that these patients are distinguished from those with true corneal dystrophies.

In past two decades, advances in genetic technologies have resulted in the identification of the causative gene for the majority of corneal dystrophies. The initial discovery of most cornel dystrophy genes correlated with the onset of Sanger sequencing, followed by a second wave of discovery following the advent of WES. It now appears that the lower cost and increased availability of WGS is resulting in another wave of genetic discovery for unsolved corneal dystrophies, and the important contribution of non-coding variation to corneal dystrophies is becoming apparent. These novel corneal dystrophy genes are elucidating new molecular pathways which play an important role in the maintenance of corneal transparency and have implications for the field of genetics in general, providing insights into how gene transcription is altered by non-coding variation. These and future discoveries, in combination with novel methods for manipulating and editing the human genome, is likely to result in less invasive, more effective therapeutics for corneal dystrophy patients in the future.

References

Aberdam, D., Galliano, M., Vailley, J., Pulkkinen, L., Bonifas, J., Christiano, A., Tryggvason, K., Uitto, J., Epstein, E., Ortonne, J. and Meneguzzi, G. (1994) 'Herlitz's junctional epidermolysis bullosa is linked to mutations in the gene (LAMC2) for the gamma-2 subunit of nicein/kalinin (LAMININ-5)', *Nature Genetics*, 6, pp. 299–304.

Abu, A., Frydman, M., Marek, D., Pras, E., Nir, U., Reznik-wolf, H. and Pras, E. (2008) 'Deleterious mutations in the zinc-finger 469 gene cause brittle cornea syndrome.', *American Journal of Human Genetics*, 82, pp. 1217–1222.

Adzhubei, I., Schmidt, S., Peshkin, L., Ramensky, V., Gerasimova, A., Bork, P., Kondrashov, A. and Sunyaev, S. (2010) 'A method and server for predicting damaging missense mutations.', *Nature Methods*, 7(4), pp. 248–249.

Afshari, N., Bahadur, R., Eifrig, D., Thogersen, I., Enghild, J. and Klintworth, G. (2008) 'Atypical asymmetric lattice corneal dystrophy associated with a novel homozygous mutation (Val624Met) in the TGFBI gene.', *Molecular Vision*, 14, pp. 495–499.

Afshari, N., Mullally, J., Afshari, M., Steinert, R., Adamis, A., Azar, D., Talamo, J., Dohlman, C. and Djyra, T. (2001) 'Survey of patients with granular, lattice, avellino, and Reis-Bucklers corneal dystrophies for mutations in the BIGH3 and gelsolin genes.', *Archives of Ophthalmology*, 119(1), pp. 16–22.

Ahram, D., Sato, T. S., Kohilan, A., Tayeh, M., Chen, S., Leal, S., Al-salem, M. and El-shanti, H. (2009) 'A homozygous mutation in ADAMTSL4 causes autosomal recessive isolated ectopia lentis.', *The American Journal of Human Genetics*. The American Society of Human Genetics, 84(2), pp. 274–278.

Akama, T., 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. and 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(2), pp. 237–241.

Akhtar, S., Bron, A., Qin, X., Creer, R., Guggenheim, J. and Meek, K. (2005) 'Gelatinous drop-like corneal dystrophy in a child with developmental delay: clinicopathological features and exclusion of the M1S1 gene.', *Eye*, 19(2), pp. 198–204.

Al-Ghadeer, H., Mohamed, J. and Khan, K. (2011) 'Schnyder corneal dystrophy in a Saudi Arabian family with heterozygous UBIAD1 mutation (p. L121F).', *Middle East African Journal of Ophthalmology*, 18(1), pp. 61–64.

Alavi, A., Elahi, E., Amoli, F. and Tehrani, M. (2007) 'Exclusion of TACSTD2 in an Iranian GDLD pedigree.', *Molecular Vision*, 13, pp. 1441–1445.

Alazami, A., Shaheen, R., Alzahrani, F., Snape, K., Saggar, A., Brinkmann, B., Bavi, P., Al-gazali, L. and Alkuraya, F. (2009) 'FREM1 Mutations Cause bifid nose, renal agenesis, and anorectal malformations syndrome.', *American Journal of Human Genetics*, 85, pp. 414–418.

Aldave, A., Ann, L., Frausto, R., Nguyen, C., Yu, F. and Raber, I. (2013)

‘Classification of posterior polymorphous corneal dystrophy as a corneal ectatic disorder following confirmation of associated significant corneal steepening.’, *JAMA Ophthalmology*, 131(12), pp. 1583–1590.

Aldave, A., Rayner, S., King, J., Affeldt, J. and Yellore, V. (2005) ‘A unique corneal dystrophy of Bowman’s layer and stroma associated with the Gly623Asp mutation in the transforming growth factor beta-induced (TGFBI) gene.’, *Ophthalmology*, 112(6), pp. 1017–1022.

Aldave, A., Rayner, S., King, J., Salem, A., Prechanond, A., Hashida, S., Affeldt, J., Meallet, M., Glasgow, B., Small, K. and Yellore, V. (2006) ‘No pathogenic mutations identified in the TGFBI gene in polymorphic corneal amyloid deposition.’, *Cornea*, 25(4), pp. 413–415.

Aldave, A., Yellore, V., Principe, A., Abedi, G., Merrill, K., Chalukya, M., Small, K. and Udar, N. (2005) ‘Candidate gene screening for posterior polymorphous dystrophy.’, *Cornea*, 24(2), pp. 151–155.

Aldave, A., Yellore, V., Sonmez, B., Bourla, N., Salem, A., Khan, M., Rayner, S. and Glasgow, B. (2008) ‘A novel variant of combined granular-lattice corneal dystrophy associated with the Met619Lys mutation in the TGFBI gene.’, *Archives of Ophthalmology*, 126(3), pp. 371–377.

Aldave, A., Yellore, V., Yu, F., Bourla, N., Sonmez, B., Salem, A., Rayner, S., Sampat, K., Krafchak, C. and Richards, J. (2007) ‘Posterior polymorphous corneal dystrophy is associated with TCF8 gene mutations and abdominal hernia.’, *American Journal of Medical Genetics*, 143(21), pp. 2549–2556.

Aligianis, I., Johnson, C., Gissen, P., Chen, D., Hampshire, D., Hoffmann, K., Maina, E., Morgan, N., Tee, L., Morton, J., Ainsworth, J., Horn, D., Rosser, E., Cole, T., Stolte-dijkstra, I., Fieggen, K., Kjaer, K., Warburg, M., Bond, J., Trembath, R., Harris, L. and Takai, Y. (2005) ‘Mutations of the catalytic subunit of RAB3GAP cause Warburg Micro syndrome.’, *Nature Genetics*, 37(3), pp. 221–223.

Allen, E., Atkinson, S., Liao, H., Moore, J., Leslie Pedrioli, D., Smith, F. J., McLean, W. and Moore, C. (2013) ‘Allele-specific siRNA silencing for the common keratin 12 founder mutation in Meesmann epithelial corneal dystrophy.’, *Investigative Ophthalmology & Visual Science*, 54(1), pp. 494–502.

Allen, E., Courtney, D., Atkinson, S., Moore, J., Mairs, L., Poulsen, E., Schiroli, D., Maurizi, E., Cole, C., Hickerson, R., James, J., Murgatroyd, H., Smith, F. J., MacEwen, C., Enghild, J., Nesbit, M., Leslie Pedrioli, D., McLean, W. and Moore, C. (2015) ‘Keratin 12 missense mutation induces the unfolded protein response and apoptosis in meesmann epithelial corneal dystrophy’, *Human Molecular Genetics*, 25(6), pp. 1176–1191.

Appaiahgari, M. and Vrati, S. (2015) ‘Adenoviruses as gene/vaccine delivery vectors: promises and pitfalls.’, *Expert Opinion on Biological Therapy*, 15(3), pp. 337–351.

Auw-Haedrich, C., Agostini, H., Clausen, I., Reinhard, T., Eberwein, P., Schorderet, D. and Gruenauer-Kloevekorn, C. (2009) ‘A corneal dystrophy associated with transforming growth factor beta-induced Gly623Asp mutation: an amyloidogenic phenotype.’, *Ophthalmology*, 116(1), pp. 46–51.

Bach, G. (2001) ‘Mucolipidosis type IV.’, *Molecular Genetics and Metabolism*,

73(3), pp. 197–203.

Badalà, F., Nouri-mahdavi, K. and Raoof, D. (2008) 'Adeno-associated virus: fit to serve', *Current Opinion in Virology*, 144(5), pp. 724–732.

Bakhtiari, P., Frausto, R., Roldan, A., Wang, C., Yu, F. and Aldave, A. (2013) 'Exclusion of pathogenic promoter region variants and identification of novel nonsense mutations in the zinc finger E-box binding homeobox 1 gene in posterior polymorphous corneal dystrophy.', *Molecular Vision*, 19, pp. 575–580.

Bamshad, M., Ng, S., Bigham, A., Tabor, H., Emond, M., Nickerson, D. and Shendure, J. (2011) 'Exome sequencing as a tool for Mendelian disease gene discovery.', *Nature Reviews Genetics*, 12(11), pp. 745–755.

Baratz, K., Tosakulwong, N., Ryu, E., Brown, W., Branham, K., Chen, W., Tran, K., Schmid-Kubista, K., Heckenlively, J., Swaroop, A., Abecasis, G., Bailey, K. and Edwards, A. (2010) 'E2-2 protein and Fuchs's corneal dystrophy.', *The New England Journal of Medicine*, 363(11), pp. 1016–1024.

Bargal, R., Avidan, N., Ben-asher, E., Olender, Z., Zeigler, M., Frumkin, A., Glusman, G., Lancet, D. and Bach, G. (2000) 'Identification of the gene causing mucolipidosis type IV.', *Nature Genetics*, 26, pp. 120–123.

Barker, D. F., Hostikka, S. L., Zhou, J., Chow, L. T., Oliphant, A. R., Gerken, S. C., Gregory, M. C., Skolnick, M. H., Atkin, C. L. and Tryggvason, K. (1990) 'Identification of mutations in the COL4A5 collagen gene in Alport syndrome.', *Science*, 248(4960), pp. 1224–1227.

Basel-Vanagaite, L., Attia, R., Ishida-Yamamoto, A., Rainshtein, L., Ben Amitai, D., Lurie, R., Pasmanik-Chor, M., Indelman, M., Zvulunov, A., Saban, S., Magal, N., Sprecher, E. and Shohat, M. (2007) 'Autosomal recessive ichthyosis with hypotrichosis caused by a mutation in ST14, encoding type II transmembrane serine protease matriptase.', *American Journal of Human Genetics*, 80(3), pp. 467–477.

Basler, E., Grompe, M., Parenti, G., Yates, J. and Ballabio, A. (1992) 'Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis.', *American Journal Human Genetics*, 50, pp. 483–491.

Bath, C., Muttuvelu, D., Emmersen, J., Vorum, H., Hjortdal, J. and Zachar, V. (2013) 'Transcriptional dissection of human limbal niche compartments by massive parallel sequencing.', *PLoS one*, 8(5), p. e64244.

Belin, M., Khachikian, S., McGhee, C. and Patel, D. (2010) 'New technology in corneal imaging.', *International Ophthalmology Clinics*, 50(3), pp. 177–189.

Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q., Antipenko, A., Shang, L., Boisson, B., Casanova, J. and Abel, L. (2015) 'Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(17), pp. 5473–5478.

Beltra, D., Currier, S., Steinbrecher, A., Celli, J., Beusekom, E., Zwaag, B., Merlini, L., Chitayat, D., Dobyns, W., Cormand, B., Lehesjoki, A., Voit, T., Walsh, C., Bokhoven, H. and Brunner, H. (2002) 'Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe reuronal migration disorder Walker-Warburg syndrome.', *American Journal Human Genetics*, 71, pp. 1033–1043.

Berger, W., Meindl, A., van de Pol, T., Cremers, F., Ropers, H., Doerner, C., Monaco, A., Bergen, A., Lebo, R., Warburg, M., Zorgollern, L., Lorenz, B., Gal, A., Bleeker-Wagemakers, E. and Meitinger, T. (1992) 'Isolation of a candidate gene or Norrie disease by positional cloning.', *Nature Genetics*, 1, pp. 199–203.

Bernstein, H. S., Bishop, D. F., Astrin, K. H., Komreich, R., Eng, C. M., Sakuraba, H., Desnick, R. J., Kornreich, R., Eng, C. M., Sakuraba, H. and Desnick, R. J. (1989) 'Fabry disease: six gene rearrangements and an exonic point mutation in the alpha-galactosidase gene.', *The Journal of Clinical Investigation*, 83(21), pp. 1390–1399.

Beutler, E., Kuhl, W. and Comings, D. (1975) 'Hexosaminidase isozyme in type 0 GM2 gangliosidosis (Sandhoff-Jatzkewitz disease).', *American Journal of Human Genetics*, 27, pp. 628–638.

Birgani, S., Salehi, Z., Houshmand, M., Mohamadi, M., Promehr, L. and Mozafarzadeh, Z. (2009) 'Novel mutations of CHST6 in Iranian patients with macular corneal dystrophy.', *Molecular Vision*, 15, pp. 373–377.

Biswas, S., Munier, F., Yardley, J., Hart-Holden, N., Perveen, R., Cousin, P., Sutphin, J., Noble, B., Batterbury, M., Kielty, C., Hackett, A., Bonshek, R., Ridgway, A., McLeod, D., Sheffield, V., Stone, E., Schorderet, D. and Black, G. (2001) 'Missense mutations in COL8A2, the gene encoding the α2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy.', *Human Molecular Genetics*, 10(21), pp. 2415–2423.

Bodzioch, M., Orsó, E., Klucken, J., Langmann, T., Böttcher, A., Diederich, W., Drobnik, W., Barlage, S., Büchler, C., Porsch-özcürümez, M., Kaminski, W., Hahmann, H., Oette, K., Rothe, G., Aslanidis, C., Lackner, K. and Schmitz, G. (1999) 'The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease.', *Nature Genetics*, 22, pp. 347–351.

Boerkoel, C., Takashima, H., John, J., Yan, J., Stankiewicz, P., Rosenbarker, L., André, J., Bogdanovic, R., Burguet, A., Cockfield, S., Cordeiro, I., Illies, F., Joseph, M., Kaitila, I., Lama, G., Loirat, C., Ross, D., Milford, D., Petty, E., Rodrigo, F., Saraiva, J., Schmidt, B., Smith, G., Spranger, J., Stein, A., Thiele, H., Tizard, J., Weksberg, R., Lupski, J. and Stockton, D. (2002) 'Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia.', *Nature Genetics*, 30, pp. 215–220.

Bonifas, J., Rothman, A. and Epstein, E. (1991) 'Epidermolysis bullosa simplex : evidence in two families for keratin gene abnormalities', *Science*, 254, pp. 1202–1205.

Borderie, V., Sandali, O., Bullet, J., Gaujoux, T., Touzeau, O. and Laroche, L. (2012) 'Long-term results of deep anterior lamellar versus penetrating keratoplasty.', *Ophthalmology*, 119(2), pp. 249–255.

Boutboul, S., Black, G., Moore, J., Sinton, J., Menasche, M., Munier, F., Laroche, L., Abitbol, M. and Schorderet, D. (2006) 'A subset of patients with epithelial basement membrane corneal dystrophy have mutations in TGFBI/BIGH3.', *Human mutation*, 27(6), pp. 553–557.

Bowden, P., Haley, J., Kansky, A., Rothnagel, J., Jones, D. and Turner, R. (1995) 'Mutation of a type II keratin gene (K6a) in pachyonychia congentia.', *Nature Genetics*, 10, pp. 363–365.

Bredrup, C., Knappskog, P., Majewski, J., Rødahl, E. and Boman, H. (2005) 'Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene.', *Investigative Opthalmology & Visual Science*, 46(2), pp. 420–426.

Brockington, M., Blake, D., Prandini, P., Brown, S., Torelli, S., Benson, M., Ponting, C., Estournet, B., Romero, N., Mercuri, E., Voit, T., Sewry, C., Guicheney, P. and Muntoni, F. (2001) 'Mutations in the fukutin-related protein gene (FKRP) cause a form of congenital muscular dystrophy with secondary laminin alpha-2 deficiency and abnormal glycosylation of alpha-dystroglycan.', *American Journal Human Genetics*, 69, pp. 1198–1209.

Brooks-Wilson, A., Marcil, M., Clee, S., Zhang, L., Roomp, K., van Dam, M., Yu, L., Brewer, C., Collins, J., Molhuizen, H., Loubser, O., Ouelette, B., Fichter, K., Ashbourne-Excoffon, K., Sensen, C., Scherer, S., Mott, S., Denis, M., Martindale, D., Frohlich, J., Morgan, K., Koop, B., Pimstone, S., Kastelein, J., Genest, J. and Hayden, M. (1999) 'Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency.', *Nature Genetics*, 22(4), pp. 336–345.

Bull, P., Thomas, G., Rommens, J., Forbes, J. and Cox, D. (1993) 'The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene.', *Nature Genetics*, 5, pp. 327–337.

Burch, G., Gong, Y., Liu, W., Dettman, R., Curry, C., Smith, L., Miller, W. and Bristow, J. (1997) 'Tenascin-X deficiency is associated with Ehlers-Danlos syndrome.', *Nature Genetics*, 17, pp. 104–108.

Burns, R. (1968) 'Meesman's corneal dystrophy.', *Transactions of the American Ophthalmological Society*, 66, pp. 530–635.

Burrows, N., Nicholls, A., Yates, J., Gatward, G., Sarathachandra, P., Richards, A. and Pope, F. (1996) 'The gene encoding collagen α1(V) (COL5A1) is linked to mixed Ehlers-Danlos syndrome type I/II.', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 106, pp. 1273–1276.

Bustamante, M., Tasinato, A., Maurer, F., Elkochairi, I., Lepore, M., Arsenijevic, Y., Pedrazzini, T., Munier, F. and Schorderet, D. (2008) 'Overexpression of a mutant form of TGFBI/BIGH3 induces retinal degeneration in transgenic mice.', *Molecular Vision*, 14, pp. 1129–1137.

Cabral-Macias, J., Zenteno, J., Ramirez-Miranda, A., Navas, A., Bermudez-Magner, J., Boullosa-Graña, V., Graue-Hernandez, E. and Buentello-Volante, B. (2016) 'Familial gelatinous drop-like corneal dystrophy caused by a novel nonsense TACSTD2 mutation.', *Cornea*, 35(7), pp. 987–990.

Callaghan, M., Hand, C., Kennedy, S., FitzSimon, J., Collum, L. and Parfrey, N. (1999) 'Homozygosity mapping and linkage analysis demonstrate that autosomal recessive congenital hereditary endothelial dystrophy (CHED) and autosomal dominant CHED are genetically distinct.', *The British Journal of Ophthalmology*, 83(1), pp. 115–119.

Cao, W., Yan, M., Hao, Q., Wang, S., Wu, L., Liu, Q., Li, M., Biddle, F. and Wu, W. (2013) 'Autosomal-dominant Meesmann epithelial corneal dystrophy without an exon mutation in the keratin-3 or keratin-12 gene in a Chinese family.', *The Journal of International Medical Research*, 41(2), pp. 511–518.

Carrwik, C. and Stenevi, U. (2009) 'Lattice corneal dystrophy, gelsolin type

(Meretoja's syndrome).', *Acta Ophthalmologica*, 87(8), pp. 813–819.

Carter, N. (2007) 'Methods and strategies for analyzing copy number variation using DNA microarrays.', *Nature Genetics*, 39, pp. S16-21.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005) 'MatInspector and beyond: Promoter analysis based on transcription factor binding sites.', *Bioinformatics*, 21(13), pp. 2933–2942.

Chau, H., Ha, N., Cung, L., Thanh, T., Fujiki, K., Murakami, A. and Kanai, A. (2003) 'H626R and R124C mutations of the TGFBI (BIGH3) gene caused lattice corneal dystrophy in Vietnamese people.', *The British Journal of Ophthalmology*, 87(6), pp. 686–9.

Chen, Y., Huang, K., Nakatsu, M., Xue, Z., Deng, S. and Fan, G. (2013) 'Identification of novel molecular markers through transcriptomic analysis in human fetal and adult corneal endothelial cells.', *Human Molecular Genetics*, 22(7), pp. 1271–1279.

Chitayat, D., Meunier, C., Hodgkinson, K., Silver, K., Flanders, M., Anderson, I. J., Little, J. M., Whiteman, D. and Carpenter, S. (1991) 'Mucolipidosis type IV: clinical manifestations and natural history.', *American Journal of Medical Genetics*, 41, pp. 313–318.

Chow, R. and Lang, R. (2001) 'Early eye development in vertebrates.', *Annual Review of Cell and Developmental Biology*, 17, pp. 255–296.

Christiano, A., Greenspan, D., Hoffman, G., Zhang, X., Tamai, Y., Lin, A., Dietz, H., Hovnanian, A. and Uitto, J. (1993) 'A missense mutation in type VII collage in two affected siblings with recessive dystrophic epidermolysis bullosa.', *Nature Genetics*, 4, pp. 62–66.

Chua, B. E., Mitchell, P., Wang, J. J. and Rochtchina, E. (2004) 'Corneal arcus and hyperlipidemia: findings from an older population.', *American Journal of Ophthalmology*, 137(2), pp. 363–365.

Chung, D., Frausto, R., Ann, L., Jang, M. and Aldave, A. (2014) 'Functional impact of ZEB1 mutations associated with posterior polymorphous and fuchs' endothelial corneal dystrophies.', *Investigative Ophthalmology and Visual Science*, 55(10), pp. 6159–6166.

Cibis, G. and Tripathi, R. (1982) 'The differential diagnosis of Descemet's tears (Haab's striae) and posterior polymorpous dystrophy bands. A clinicopathologic study.', *Ophthalmology*, 89(6), pp. 614–620.

Cieply, B., Farris, J., Denvir, J., Ford, H. L. and Frisch, S. M. (2013) 'Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2', *Cancer Research*, 73(20), pp. 6299–6309.

Cieply, B., Riley IV, P., Pifer, P., Widmeyer, J., Addison, J., Ivanov, A., Denvir, J., Frisch, S., Riley, P., Pifer, P., Widmeyer, J., Addison, J. B., Ivanov, A. V., Denvir, J. and Frisch, S. M. (2012) 'Suppression of the epithelial-mesenchymal transition by grainyhead-like-2.', *Cancer Research*, 72(9), pp. 2440–2453.

Cirulli, E. and Goldstein, D. (2010) 'Uncovering the roles of rare variants in

common disease through whole-genome sequencing.', *Nature Reviews Genetics*, 11(6), pp. 415–425.

Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P., Wu, X., Jiang, W., Marraffini, L. and Zhang, F. (2014) 'Multiplex Genome Engineering Using CRISPR/Cas Systems', *Science*, 344(6121), pp. 299–304.

Corden, L. and McLean, W. (1996) 'Human keratin diseases: hereditary fragility of specific epithelial tissues.', *Experimental Dermatology*, 5(6), pp. 297–307.

Correa-Gomez, V., Villalvazo-Cordero, L. and Zenteno, J. (2007) 'The TGFBI A546D mutation causes an atypical type of lattice corneal dystrophy.', *Molecular Vision*, 13, pp. 1695–1700.

Courtney, D., Atkinson, S., Moore, J., Maurizi, E., Serafini, C., Pellegrini, G., Black, G., Manson, F., Yam, G., Macewen, C., Allen, E., McLean, W. and Moore, C. (2014) 'Development of allele-specific gene-silencing siRNAs for TGFBI Arg124Cys in lattice corneal dystrophy type I.', *Investigative Ophthalmology & Visual Science*, 55(2), pp. 977–985.

Cvekl, A. and Wang, W.-L. (2009) 'Retinoic acid signaling in mammalian eye development.', *Experimental Eye Research*, 89, pp. 280–291.

D'haene, B., Vandesompele, J. and Hellemans, J. (2010) 'Accurate and objective copy number profiling using real-time quantitative PCR.', *Methods*, 50, pp. 262–270.

Dagoneau, N., Scheffer, D., Al-gazali, L., Rocco, M., Godard, A., Martinovic, J., Raas-rothschild, A., Sigaudy, S., Unger, S., Nicole, S., Fontaine, B., Taupin, J., Superti-furga, A., Merrer, M., Bonaventure, J., Munnich, A. and Legeai-mallet, L. (2004) 'Null leukemia inhibitory factor receptor (LIFR) mutations in Stuve-Wiedemann/Schwartz-Jampel type 2 syndrome.', *American Journal of Human Genetics*, 74, pp. 298–305.

Dark, A. (1977) 'Bleb dystrophy of the cornea: histochemistry and ultrastructure.', *British Journal of Ophthalmology*, 61(1), pp. 65–69.

Das, S. and Seitz, B. (2008) 'Recurrent corneal erosion syndrome.', *Survey of Ophthalmology*, 53(1), pp. 3–15.

DelMonte, D. and Kim, T. (2011) 'Anatomy and physiology of the cornea.', *Journal of Cataract and Refractive Surgery*, 37, pp. 588–598.

Desir, J. and Abramowicz, M. (2008) 'Congenital hereditary endothelial dystrophy with progressive sensorineural deafness (Harboyan syndrome).', *Orphanet Journal of Rare Diseases*, 3(28).

Desmet, F., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M. and Béroud, C. (2009) 'Human Splicing Finder: an online bioinformatics tool to predict splicing signals.', *Nucleic Acids Research*, 37(9), p. e67.

Dierks, T., Schmidt, B., Borissenko, L., Peng, J., Preusser, A., Mariappan, M. and Figura, K. (2003) 'Multiple sulfatase deficiency is caused by mutations in the gene encoding the human Calpha-formylglycine generating enzyme.', *Cell*, 113, pp. 435–444.

Dighiero, P., Niel, F., Ellies, P. and D'Hermies, F. (2001) 'Histologic phenotype–genotype correlation of corneal dystrophies associated with eight distinct mutations

in the TGFBI gene.', *Ophthalmology*, 6420, pp. 818–823.

Dinh, R., Rapuano, C. J., Cohen, E. J. and Laibson, P. R. (1999) 'Recurrence of corneal dystrophy after excimer laser phototherapeutic keratectomy.', *Ophthalmology*, 106(8), pp. 1490–1497.

Dobrovolny, R., Liskova, P., Ledvinova, J., Poupetova, H., Asfaw, B., Filipec, M., Jirsova, K., Kraus, J. and Elleder, M. (2007) 'Mucolipidosis IV: report of a case with ocular restricted phenotype caused by leaky splice mutation.', *American Journal of Ophthalmology*, 143(4), pp. 663–671.

Dodt, G., Braverman, N., Wong, C., Moser, A., Moser, H., Watkins, P., Valle, D. and Gould, S. (1995) 'Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the peroxisome biogenesis disorders.', *Nature Genetics*, 9, pp. 115–125.

Dolz-Marco, R., Gallego-Pinazo, R., Pinazo-Duran, M. and Diaz-Llopis, M. (2014) 'Crystalline subtype of pre-descemetic corneal dystrophy.', *Journal of Ophthalmic and Vision Research*, 9(2), pp. 269–271.

Du, C., Li, Y., Dai, L., Gong, L. and Han, C. (2011) 'A mutation in the UBIAD1 gene in a Han Chinese family with Schnyder corneal dystrophy.', *Molecular Vision*, 17, pp. 2685–2692.

Du, Y., Funderburgh, M., Mann, M., SundarRaj, N. and Funderburgh, J. (2005) 'Multipotent stem cells in human corneal stroma.', *Stem Cells*, 23(9), pp. 1266–1275.

Dua, H., Faraj, L., Said, D., Gray, T. and Lowe, J. (2013) 'Human corneal anatomy redefined: a novel pre-Descemet's layer (Dua's layer).', *Ophthalmology*, 120(9), pp. 1778–1785.

Dudakova, L., Palos, M., Svobodova, M., Bydzovsky, J., Huna, L., Jirsova, K., Hardcastle, A., Tuft, S. and Liskova, P. (2014) 'Macular corneal dystrophy and associated corneal thinning.', *Eye*, 28(10), pp. 1201–1205.

Dyrlund, T., Poulsen, E., Scavenius, C., Nikolajsen, C., Thøgersen, I., Vorum, H. and Enghild, J. (2012) 'Human cornea proteome: identification and quantitation of the proteins of the three main layers including epithelium, stroma, and endothelium.', *Journal of Proteome Research*, 11(8), pp. 4231–4239.

Efebera, Y., Sturm, A., Baack, E., Hofmeister, C., Satoskar, A., Nadasdy, T., Nadasdy, G., Benson, D., Gillmore, J., Hawkins, P. and Rowczenio, D. (2014) 'Novel gelsolin variant as the cause of nephrotic syndrome and renal amyloidosis in a large kindred.', *Amyloid*, 21(2), pp. 110–112.

Eghrari, A., Riazuddin, S. and Gottsch, J. (2015) 'Fuchs corneal dystrophy.', in *Progress in Molecular Biology and Translational Science*, pp. 79–97.

Ehlers, N., Hjortdal, J., Nielsen, K., Thiel, H.-J. and Ørntoft, T. (2008) 'Phenotypic variability in Meesmann's dystrophy: clinical review of the literature and presentation of a family genetically identical to the original family.', *Acta Ophthalmologica*, 86(1), pp. 40–44.

Eifrig, D., Afshari, N., Buchanan, H., Bowling, B. and Klintworth, G. (2004) 'Polymorphic corneal amyloidosis: a disorder due to a novel mutation in the transforming growth factor beta-induced (BIGH3) gene.', *Ophthalmology*, 111(6),

pp. 1108–1114.

El-Ashry, M., Abd El-Aziz, M., Larkin, D., Clarke, B., Cree, I., Hardcastle, A., Bhattacharya, S. and Ebenezer, N. (2003) 'A clinical, histopathological, and genetic study of Avellino corneal dystrophy in British families.', *The British Journal of Ophthalmology*, 87(7), pp. 839–842.

El-Ashry, M., El-Aziz, M. A., Shalaby, O., Wilkins, S., Poopalasundaram, S., Cheetham, M. E., Tuft, S. J., Hardcastle, A. J., Bhattacharya, S. S. and Ebenezer, N. D. (2005) 'Novel CHST6 nonsense and missense mutations responsible for macular corneal dystrophy', *American Journal of Ophthalmology*, 139, pp. 192–193.

El-Ashry, M., El-Aziz, M., Wilkins, S., Cheetham, M. and Wilkie, S. (2002) 'Identification of novel mutations in the carbohydrate sulfotransferase gene (CHST6) causing macular corneal dystrophy.', *Investigative Ophthalmology & Visual Science*, 43(2), pp. 377–382.

Elhalis, H., Azizi, B. and Jurkunas, U. (2010) 'Fuchs endothelial corneal dystrophy.', *The Ocular Surface*, 8(4), pp. 173–184.

Escudier, E., Chapelin, C., Bridoux, A., Roger, G., Cle, A., Goossens, M. and Amselem, S. (1999) 'Loss-of-function mutations in a human gene related to Chlamydomonas reinhardtii dynein IC78 result in primary ciliary dyskinesia.', *American Journal of Human Genetics*, 65, pp. 1508–1519.

Evans, A. and Gage, P. (2005) 'Expression of the homeobox gene Pitx2 in neural crest is required for optic stalk and ocular anterior segment development.', *Human Molecular Genetics*, 14(22), pp. 3347–3359.

Fang, Dagenais, S., Erickson, R., Arlt, M., Glynn, M., Gorski, J., Seaver, L. and Glover, T. (2000) 'Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome.', *American Journal of Human Genetics*, 67, pp. 1382–1388.

Fernandes, N., Janniger, C. and Schwartz, R. (2010) 'X-linked ichthyosis: an oculocutaneous genodermatosis.', *Journal of the American Academy of Dermatology*, 62(3), pp. 480–485.

Fernandez-Sasso, D., Acosta, J. and Malbran, E. (1979) 'Punctiform and polychromatic pre-Descemet's dominant corneal dystrophy.', *British Journal of Ophthalmology*, 63(5), pp. 336–338.

Fernández, A., Sorokin, A. and Thompson, P. (2007) 'Corneal arcus as coronary artery disease risk factor.', *Atherosclerosis*, 193(2), pp. 235–240.

Ferry, A., Benson, W. and Weinberg, R. (1997) 'Combined granular-lattice ('Avellino') corneal dystrophy.', *Transactions of the Ophthalmic Society*, 95, pp. 61–77.

Fessler, J., Kramerova, I., Kramerov, A., Chen, Y. and Fessler, L. (2004) 'Papilin, a novel component of basement membranes, in relation to ADAMTS metalloproteases and ECM development', *International Journal of Biochemistry and Cell Biology*, 36(6), pp. 1079–1084.

Fine, J., Johnson, L., Weiner, M., Stein, A., Cash, S., Deleoz, J., Devries, D. and Suchindran, C. (2004) 'Eye involvement in inherited epidermolysis bullosa: Experience of the national epidermolysis bullosa registry.', *American Journal of*

Ophthalmology, 138(2), pp. 254–262.

Fini, E. and Stramer, B. (2005) 'How the cornea heals: cornea-specific repair mechanisms affecting surgical outcomes.', *Cornea*, 24, pp. S2–S11.

Franceschini, G., Sirtori, C. and Capurso, A. (1980) 'A-I(Milano) apoprotein. Decreased high density lipoprotein cholesterol levels with significant lipoprotein modifications and without clinical atherosclerosis in an Italian family.', *Journal of Clinical Investigation*, 66(5), pp. 892–900.

Franceschini, G., Sirtori, M., Gianfranceschi, G. and Sirtori, C. (1981) 'Relation between the HDL apoproteins and AI isoproteins in subjects with the AIMilano abnormality.', *Metabolism*, 30(5), pp. 502–509.

Frederick, G., Amirkhan, R., Schultz, R. and Friedberg, E. (1994) 'Structural and mutational analysis of the xeroderma pigmentosum group D (XPD) gene.', *Human Molecular Genetics*, 3(10), pp. 1783–1788.

Freeman, A., Datta, S., Torabi-parizi, P., Subramanian, N., Bunney, T., Baxendale, R., Martins, M., Romberg, N., Komarow, H., Aksentijevich, I., Kim, H., Ho, J., Cruse, G., Jung, M., Gilfillan, A., Metcalfe, D., Nelson, C., Brien, M., Wisch, L., Shianna, K., Cirulli, E., Goldstein, D., Long, E. O., Moir, S., Meffre, E., Holland, S., Kastner, D., Katan, M., Hoffman, H. and Milner, J. (2012) 'Cold urticaria, immunodeficiency, and autoimmunity related to PLCG2 deletions.', *The New England Journal of Medicine*, 366(4), pp. 330–338.

Frising, M., Wildhardt, G., Frisch, L. and Pitz, S. (2006) 'Recurrent granular dystrophy of the cornea: an unusual case.', *Cornea*, 25(5), pp. 614–617.

Funderburgh, J., Mann, M. and Funderburgh, M. (2003) 'Keratocyte phenotype mediates proteoglycan structure: a role for fibroblasts in corneal fibrosis.', *The Journal of Biological Chemistry*, 278, pp. 45629–45637.

Funke, H., von Eckardstein, A., Pritchard, P., Albers, J., Kastelein, J., Droste, C. and Assmann, G. (1991) 'A molecular defect causing fish eye disease: an amino acid exchange in lecithin-cholesterol acyltransferase (LCAT) leads to the selective loss of alpha-LCAT activity.', *Proceedings of the National Academy of Sciences of the United States of America*, 88, pp. 4855–4859.

Garner, A. and Tripathi, R. (1972) 'Hereditary crystalline stromal dystrophy of Schnyder. II. Histopathology and ultrastructure.', *The British Journal of Ophthalmology*, 56(5), pp. 400–408.

Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012) 'Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), pp. E2579–E2586.

Gatzioufas, Z., Charalambous, P., Loew, U., Kozobolis, V., Schirra, F., Krause, M. and Seitz, B. (2010) 'Evidence of oxidative stress in Schnyder corneal dystrophy.', *The British Journal of Ophthalmology*, 94(9), pp. 1262–1264.

Gaynor, P., Zhang, W., Weiss, J., Skarlatos, S., Rodrigues, M. and Kruth, H. (1996) 'Accumulation of HDL apolipoproteins accompanies abnormal cholesterol accumulation in Schnyder's corneal dystrophy.', *Arteriosclerosis, Thrombosis and Vascular Biology*, 16(8), pp. 992–999.

Giglia-Mari, G., Coin, F., Ranish, J., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N., Raams, A., Argentini, M., Spek, P., Botta, E., Stefanini, M., Egly, J., Aebersold, R., Hoeijmakers, J. and Vermeulen, W. (2004) 'A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A.', *Nature Genetics*, 36(7), pp. 714–719.

Gilbert, S. (2000) *Developmental biology*. 6th ed. Sunderland (MA): Sinauer Associates.

Gimelli, G., Giglio, S., Zuffardi, O., Alhonen, L., Suppola, S. and Cusano, R. (2002) 'Gene dosage of the spermidine/ spermine N1-acetyltransferase (SSAT) gene with putrescine accumulation in a patient with a Xp21.1p22. 12 duplication and keratosis follicularis spinulosa decalvans (KFSD).', *Human Genetics*, 111, pp. 235–241.

Giunta, C., Elc, N., Albrecht, B., Eich, G., Janecke, A., Yeowell, H., Weis, M., Eyre, D., Kraenzlin, M. and Steinmann, B. (2008) 'Spondylocheiro dysplastic form of the Ehlers-Danlos syndrome — an autosomal-recessive entity caused by mutations in the zinc transporter gene SLC39A13.', *American Journal of Human Genetics*, 82, pp. 1290–1305.

Glöckle, N., Kohl, S., Mohr, J., Scheurenbrand, T., Sprecher, A., Weisschuh, N., Bernd, A., Rudolph, G., Schubach, M., Poloschek, C., Zrenner, E., Biskup, S., Berger, W., Wissinger, B. and Neidhardt, J. (2014) 'Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies.', *European Journal of Human Genetics*, 22, pp. 99–104.

Golb, H., Fernandez, E. and Nelson, R. (2005) *Gross anatomy of the eye., Webvision: The Organization of the Retina and Visual System*. Salt Lake City.

Gorevic, P., Munoz, P., Gorgone, G., Purcell Jr, J., Rodrigues, M., Ghiso, J., Levy, E., Haltia, M. and Frangione, B. (1991) 'Amyloidosis due to a mutation of the gelsolin gene in an American family with lattice corneal dystrophy type II.', *The New England Journal of Medicine*, 325(25), pp. 1780–1785.

Gottsch, J., Sundin, O., Liu, S., Jun, A., Broman, K., Stark, W., Vito, E., Narang, A., Thompson, J. and Magovern, M. (2005) 'Inheritance of a novel COL8A2 mutation defines a distinct early-onset subtype of fuchs corneal dystrophy.', *Investigative Ophthalmology & Visual Science*, 46(6), pp. 1934–1939.

Gould, D., Phalan, F., Breedveld, G., van Mil, S., Smith, R., Schimenti, J., Aguglia, U., van der Knaap, M., Heutink, P. and John, S. (2005) 'Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly.', *Science*, 308, pp. 1167–1172.

Gray, R., Johnson, G. and Freedman, A. (1992) 'Climatic droplet keratopathy', *Survey of Ophthalmology*, 36(4), pp. 241–253.

Gregory, C. Y., Evans, K. and Bhattacharya, S. S. (1995) 'Genetic refinement of the chromosome 5q lattice corneal dystrophy type I locus to within 2 cM interval.', *Journal of Medical Genetics*, 32, pp. 224–226.

Groesser, L., Herschberger, E., Ruetten, A., Ruivenkamp, C., Lopriore, E., Zutt, M., Langmann, T., Singer, S., Klingseisen, L., Schneider-Brachert, W., Toll, A., Real, F., Landthaler, M. and Hafner, C. (2012) 'Postzygotic HRAS and KRAS mutations cause nevus sebaceous and Schimmelpenning syndrome.', *Nature Genetics*. Nature Publishing Group, 44(7), pp. 783–787.

Grothe, H., Little, M., Cho, A., Huang, A. and Yuan, C. (2009) 'Denaturation and solvent effect on the conformation and fibril formation of TGFBIp.', *Molecular Vision*, 15, pp. 2617–2626.

Groves, R., Liu, L., Dopping-Hepenstal, P., Markus, H., Lovell, P., Ozoemena, L., Lai-cheong, J., Gawler, J., Owaribe, K., Hashimoto, T., Mellerio, J., Mee, J. and Mcgrath, J. (2010) 'A homozygous nonsense mutation within the dystonin gene coding for the coiled-coil domain of the epithelial isoform of BPAG1 underlies a new subtype of autosomal recessive epidermolysis bullosa simplex.', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 130(6), pp. 1551–1557.

Gruenauer-Kloevekorn, C., Braeutigam, S., Heinritz, W., Froster, U. and Duncker, G. (2008) 'Macular corneal dystrophy: mutational spectrum in German patients, novel mutations and therapeutic options.', *Graefe's Archive for Clinical and Experimental Ophthalmology*, 246(10), pp. 1441–1447.

Gruenauer-Kloevekorn, C., Clausen, I., Weidle, E., Wolter-Roessler, M., Tost, F., Völcker, H., Schulze, D. P., Heinritz, W., Reinhard, T., Froster, U., Duncker, G., Schorderet, D. and Auw-Haedrich, C. (2009) 'TGFBI (BIGH3) gene mutations in German families: two novel mutations associated with unique clinical and histopathological findings.', *The British Journal of Ophthalmology*, 93(7), pp. 932–937.

Gualandri, V., Franceschini, G., Sirtori, C., Gianfranceschi, G., Orsini, G., Cerrone, A. and Menotti, A. (1985) 'AI-Milano apoprotein identification of the complete kindred and evidence of a dominant genetic transmission.', *American Journal of Human Genetics*, 37(6), pp. 1083–1097.

Guo, N., Li, X., Mann, M., Funderburgh, M., Du, Y. and Funderburgh, J. (2010) 'Hyaluronan synthesis mediates the fibrotic response of keratocytes to transforming growth factor beta.', *The Journal of Biological Chemistry*, 285(42), pp. 32012–32019.

Guzman-Aranguez, A., Loma, P. and Pintor, J. (2013) 'Small-interfering RNAs (siRNAs) as a promising tool for ocular therapy.', *British Journal of Pharmacology*, 170(4), pp. 730–747.

Gwilliam, R., Liskova, P., Filipec, M., Kmoch, S., Jirsova, K., Huckle, E., Stables, C., Bhattacharya, S., Hardcastle, A., Deloukas, P. and Ebenezer, N. (2005) 'Posterior polymorphous corneal dystrophy in Czech families maps to chromosome 20 and excludes the VSX1 gene.', *Investigative Ophthalmology & Visual Science*, 46(12), pp. 4480–4484.

Ha, N., Chau, H., Cung, L., Thanh, T., Fujiki, K., Murakami, A., Hiratsuka, Y., Hasegawa, N. and Kanai, A. (2003) 'Identification of novel mutations of the CHST6 gene in Vietnamese families affected with macular corneal dystrophy in two generations.', *Cornea*, 22(6), pp. 508–511.

Hacein-Bey-Abina, S., Hauer, J., Lim, A., Picard, C., Wang, G., Berry, C., Martinache, C., Rieux-Laucat, F., Latour, S., Belohradsky, B., Leiva, L., Sorensen, R., Debré, M., Casanova, J., Blanche, S., Durandy, A., Bushman, F., Fischer, A. and Cavazzana-Calvo, M. (2010) 'Efficacy of gene therapy for X-linked severe combined immunodeficiency.', *The New England Journal of Medicine*, 363(4), pp. 355–364.

Hadj-rabia, S., Has, C., Jobard, F., Bouadjar, B., Matsuda, F., Weissenbach, J.,

Lathrop, M. and Fischer, J. (2003) ‘Identification of mutations in a new gene encoding a FERM family protein with a pleckstrin homology domain in Kindler syndrome.’, *Human Molecular Genetics*, 12(8), pp. 925–935.

Hahn, S., Krasnewich, D., Brantly, M., Kvittingen, E. and Gahl, W. (1995) ‘Heterozygosity for an exon 12 splicing mutation and a W234G missense mutation in an American child With chronic tyrosinemia type 1.’, *Human Mutation*, 73, pp. 66–73.

Hahnel, C., Somodi, S., Weiss, D. and Guthoff, R. (2000) ‘The keratocyte network of human cornea: a three-dimensional study using confocal laser scanning fluorescence microscopy.’, *Cornea*, 19(2), pp. 185–193.

Hammar, B., Björck, E., Lagerstedt, K., Dellby, A. and Fagerholm, P. (2008) ‘A new corneal disease with recurrent erosive episodes and autosomal-dominant inheritance.’, *Acta Ophthalmologica*, 86(7), pp. 758–763.

Hammar, B., Björck, E., Lind, H., Lagerstedt, K., Dellby, A. and Fagerholm, P. (2009) ‘Dystrophia Helsinglandica: a new type of hereditary corneal recurrent erosions with late subepithelial fibrosis.’, *Acta Ophthalmologica*, 87(6), pp. 659–665.

Han, K., Choi, S., Kim, T., Maeng, Y., Stulting, R., Ji, Y. and Kim, E. (2015) ‘Pathogenesis and treatments of TGFBI corneal dystrophies.’, *Progress in Retinal and Eye Research*, 50, pp. 67–88.

Han, K., Chung, W., Kim, T., Kim, K., Kim, T. and Kim, E. (2013) ‘Changes of clinical manifestation of granular corneal deposits because of recurrent corneal erosion in granular corneal dystrophy types 1 and 2.’, *Cornea*, 32(5), pp. e113–e120.

Han, S., Ang, H., Poh, R., Chaurasia, S., Peh, G., Liu, J., Tan, D., Vithana, E. and Mehta, J. (2013) ‘Mice with a targeted disruption of Slc4a11 model the progressive corneal changes of congenital hereditary endothelial dystrophy.’, *Investigative Ophthalmology and Visual Science*, 54(9), pp. 6179–6189.

Hanna, C., Bicknell, D. and O’Brien, J. (1961) ‘Cell turnover in the adult human eye.’, *Archives of Ophthalmology*, 65(5), pp. 695–698.

Harada, T., Tanaka, H., Ikema, T., Asakura, K., Miura, M. and Ozeki, Y. (1990) ‘Specular microscopic observation of posterior corneal vesicles.’, *Ophthalmologica*, 201(3), pp. 122–127.

Hassan, H., Thaung, C., Ebenezer, N., Larkin, G., Hardcastle, A. and Tuft, S. (2013) ‘Severe Meesmann’s epithelial corneal dystrophy phenotype due to a missense mutation in the helix-initiation motif of keratin 12.’, *Eye*, 27(3), pp. 367–373.

Hassell, J. and Birk, D. (2010) ‘The molecular basis of corneal transparency.’, *Experimental Eye Research*, 91, pp. 326–335.

Hassell, J., Newsome, D., Krachmer, J. and Rodrigues, M. (1980) ‘Macular corneal dystrophy: failure to synthesize a mature keratan sulfate proteoglycan.’, *Proceedings of the National Academy of Sciences of the United States of America*, 77(6), pp. 3705–3709.

Hautala, T., Heikkinen, J., Kivirikko, K. and Myllyla, R. (1993) ‘A large duplication in the gene for lysyl hydroxylase accounts for the type VI variant of Ehlers-Danlos syndrome in two siblings.’, *Genomics*, 15, pp. 399–404.

Hay, E. D. (1980) ‘Development of the vertebrate cornea.’, *International Review of Cytology*, 63, pp. 263–322.

Hayashi, S., Osawa, T. and Tohyama, K. (2002) ‘Comparative observations on corneas, with special reference to Bowman’s layer and Descemet’s membrane in mammals and amphibians.’, *Journal of Morphology*, 254, pp. 247–258.

He, Z., Forest, F., Gain, P., Rageade, D., Bernard, A., Acquart, S., Peoc ’h, M., Defoe, D. M. and Thuret, G. (2016) ‘3D map of the human corneal endothelial cell.’, *Scientific Reports*, 6, pp. 1–14.

Hebsgaard, S., Korning, P., Tolstrup, N., Engelbrecht, J., Rouze, P. and Brunak, S. (1996) ‘Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information.’, *Nucleic Acids Research*, 24(17), pp. 3439–3452.

Heinz-Erian, P., Müller, T., Krabichler, B., Schranz, M., Becker, C., Rüschendorf, F., Nürnberg, P., Rossier, B., Booth, I., Holmberg, C., Wijmenga, C., Grigelioniene, G., Kneepkens, C., Rosipal, S., Mistrik, M., Kappler, M., Michaud, L., Dóczy, L., Siu, V., Krantz, M., Zoller, H., Uterman, G. and Janecke, A. R. (2008) ‘Mutations in SPINT2 cause a syndromic form of congenital sodium diarrhea.’, *American Journal of Human Genetics*, 84(2), pp. 188–196.

Héon, E., Greenberg, A., Kopp, K., Rootman, D., Vincent, A., Billingsley, G., Priston, M., Dorval, K., Chow, R., McInnes, R., Heathcote, G., Westall, C., Sutphin, J., Semina, E., Bremner, R. and Stone, E. (2002) ‘VSX1: a gene for posterior polymorphous dystrophy and keratoconus.’, *Human Molecular Genetics*, 11(9), pp. 1029–1036.

Héon, E., Mathers, W., Alward, W., Weisenthal, R., Sunden, S., Fishbaugh, J., Taylor, C., Krachmer, J., Sheffield, V. and Stone, E. (1995) ‘Linkage of posterior polymorphous corneal dystrophy to 20q11.’, *Human Molecular Genetics*, 4(3), pp. 485–488.

Hieda, O., Kawasaki, S., Wakimasu, K., Yamasaki, K., Inatomi, T. and Kinoshita, S. (2013) ‘Clinical outcomes of phototherapeutic keratectomy in eyes with thiel-behnke corneal dystrophy.’, *American Journal of Ophthalmology*, 155(1), pp. 66–72.

Holopainen, J., Moilanen, J. and Tervo, T. (2003) ‘In vivo confocal microscopy of fleck dystrophy and pre-Descemet’s membrane corneal dystrophy.’, *Cornea*, 22(2), pp. 160–163.

Hong, T., Watanabe, K., Ta, C., Villarreal-Ponce, A., Nie, Q. and Dai, X. (2015) ‘An Ovol2-Zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states.’, *PLoS Computational Biology*, 11(11), p. e1004569.

Howard, T., Paznekas, W., Green, E., Chiang, L., Ma, N., Ortiz de Luna, R., Garcia Delgado, C., Gonzalez-Ramos, M., Kline, A. and Jabs, E. (1997) ‘Mutations in TWIST, a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome.’, *Nature Genetics*, 15(1), pp. 36–41.

Hung, C., Ayabe, R., Wang, C., Frausto, R. and Aldave, A. (2013) ‘Pre-Descemet corneal dystrophy and X-linked ichthyosis associated with deletion of Xp22. 31 containing the STS gene.’, *Cornea*, 32(9), pp. 1283–1287.

Ide, T., Nishida, K., Maeda, N., Tsujikawa, M., Yamamoto, S., Watanabe, H. and Tano, Y. (2004) ‘A spectrum of clinical manifestations of gelatinous drop-like

corneal dystrophy in Japan.', *American Journal of Ophthalmology*, 137(6), pp. 1081–1084.

Iliff, B., Riazuddin, S. and Gottsch, J. (2012) 'A single-base substitution in the seed region of miR-184 causes EDICT syndrome.', *Investigative Ophthalmology & Visual Science*, 53, pp. 348–533.

Irvine, A., Corden, L., Swensson, O., Swensson, B., Moore, J., Frazer, D., Smith, F., Knowlton, R., Christophers, E., Rochels, R., Uitto, J. and McLean, I. (1997) 'Mutations in cornea-specific keratin K3 or K12 genes cause Meesmann's corneal dystrophy.', *Nature Genetics*, 16, pp. 184–187.

Irvine, A. and McLean, W. (1999) 'Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype-genotype correlation.', *The British Journal of Dermatology*, 140(5), pp. 815–828.

Itoh, T., Mori, T., Ohkubo, H. and Yamaizumi, M. (1999) 'A newly identified patient with clinical xeroderma pigmentosum phenotype has a non-sense mutation in the DDB2 gene and incomplete repair in (6-4) photoproducts.', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 113, pp. 251–257.

Jackson, A. and Linsley, P. (2010) 'Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application.', *Nature Reviews Drug Discovery*, 9(1), pp. 57–67.

Jalimarada, S., Ogando, D., Vithana, E. and Bonanno, J. (2013) 'Ion transport function of SLC4A11 in corneal endothelium.', *Investigative Ophthalmology & Visual Science*, 54(6), pp. 4330–4340.

Jang, M., Roldan, A., Frausto, R. and Aldave, A. (2014) 'Posterior polymorphous corneal dystrophy 3 is associated with agenesis and hypoplasia of the corpus callosum.', *Vision Research*, 100, pp. 88–92.

Jaspers, N., Raams, A., Silengo, M., Wijgers, N., Niedernhofer, L., Robinson, A., Giglia-Mari, G., Hoogstraten, D., Kleijer, W., Hoeijmakers, J. and Vermeulen, W. (2007) 'First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure.', *American Journal of Human Genetics*, 80, pp. 457–466.

Jenkins, D., Seelow, D., Jehee, F., Perlyn, C., Alonso, G., Bueno, D., Donnai, D., Josifiova, D., Mathijssen, I., Morton, J., Ørstavik, K., Sweeney, E., Wall, S., Marsh, J., Nu, P., Passos-bueno, M. and Wilkie, A. (2007) 'RAB23 mutations in Carpenter syndrome imply an unexpected role for hedgehog signaling in cranial-suture development and obesity.', 80, pp. 1162–1170.

Jester, J. (2008) 'Corneal crystallins and the development of cellular transparency.', *Seminars in Cell & Developmental Biology*, 19, pp. 82–93.

Jester, J., Barry-Lane, P., Cavanagh, H. and Petroll, W. (1996) 'Induction of α-smooth muscle actin expression and myofibroblast transformation in cultured corneal keratocytes.', *Cornea*, 15(5), pp. 505–516.

Jester, J., Murphy, C., Winkler, M., Bergmanson, J., Brown, D., Steinert, R. and Mannis, M. (2013) 'Lessons in corneal structure and mechanics to guide the corneal surgeon.', *Ophthalmology*, 120(9), pp. 1715–1717.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. and Charpentier, E.

(2012) ‘A programmable dual-RNA – guided DNA endonuclease in adaptive bacterial immunity’, *Science*, 337, pp. 816–822.

Jirsova, K., Merjava, S., Martincova, R., Gwilliam, R., Ebenezer, N., Liskova, P. and Filipec, M. (2007) ‘Immunohistochemical characterization of cytokeratins in the abnormal corneal endothelium of posterior polymorphous corneal dystrophy patients.’, *Experimental Eye Research*, 84(4), pp. 680–686.

Jm, M., Zhao, G., Ta, M., Colby, R. and Ladd, E. (2006) ‘LADD syndrome is caused by FGF10 mutations.’, *Clinical Genetics*, 69, pp. 349–354.

Jonasson, F., Johannsson, J. H., Garner, A. and Rice, N. S. (1989) ‘Macular corneal dystrophy in Iceland.’, *Eye*, 3, pp. 446–454.

Jonsson, F., Byström, B., Davidson, A., Backman, L. and Therese, G. (2014) ‘Mutations in Collagen type XVII alpha 1 (COL17A) cause Epithelial Recurrent Erosion Dystrophy (ERED)’, *Human Mutation*, 36(4), pp. 463–473.

Jordan, T., Hanson, I., Zaletayev, D., Hodgson, S., Prosser, J., Seawright, A., Hastie, N. and van Heyningen, V. (1992) ‘The human PAX6 gene is mutated in two patients with aniridia.’, *Nature Genetics*, 1, pp. 328–332.

Joyce, N. (2005) ‘Cell cycle status in human corneal endothelium.’, *Experimental Eye Research*, 81(6), pp. 629–638.

Kaji, Y., Oshika, T., Takazawa, Y., Fukayama, M. and Fujii, N. (2012) ‘Co-localisation of advanced glycation end products and D-β-aspartic acid-containing proteins in gelatinous drop-like corneal dystrophy.’, *The British Journal of Ophthalmology*, 96(8), pp. 1127–1131.

Kallend, D., Reijers, J., Bellibas, S., Bobillier, A., Kempen, H., Burggraaf, J., Moerland, M. and Wijngaard, P. (2016) ‘A single infusion of MDCO-216 (ApoA-1 Milano/POPC) increases ABCA1-mediated cholesterol efflux and pre-beta 1 HDL in healthy volunteers and patients with stable coronary artery disease.’, *European Heart Journal - Cardiovascular Pharmacotherapy*, 2(1), pp. 23–29.

Kawano, H., Fujiki, K. and Kanai, A. (1992) ‘Prevalence of gelatinous drop-like corneal dystrophy in Japan.’, *Atarashii Ganka*, 9(11), pp. 1879–1882.

Kawasaki, S., Yagi, H., Yamasaki, K., Matsuda, A., Takeda, K. and Kinoshita, S. (2011) ‘A novel mutation of the TGFBI gene causing a lattice corneal dystrophy with deep stromal involvement.’, *The British Journal of Ophthalmology*, 95(1), pp. 150–151.

Khan, K., Rudkin, A., Parry, D., Burdon, K., Mckibbin, M., Logan, C., Abdelhamed, Z., Muecke, J., Fernandez-fuentes, N., Laurie, K., Shires, M., Fogarty, R., Carr, I., Poulter, J., Morgan, J., Mohamed, M., Jafri, H., Raashid, Y., Meng, N., Piseth, H., Toomes, C., Casson, R., Taylor, G., Hammerton, M., Sheridan, E., Johnson, C., Inglehearn, C., Craig, J. and Ali, M. (2011) ‘Homozygous mutations in PXDN cause congenital cataract, corneal opacity, and developmental glaucoma.’, *The American Journal of Human Genetics*. The American Society of Human Genetics, 89(3), pp. 464–473.

Kidd, M. E., Shumaker, D. K. and Ridge, K. M. (2014) ‘The role of vimentin intermediate filaments in the progression of lung cancer.’, *American Journal of Respiratory Cell and Molecular Biology*, 50(1), pp. 1–6.

Kim, D. and Rossi, J. (2007) ‘Strategies for silencing human disease using RNA interference.’, *Nature Reviews Genetics*, 8(3), pp. 173–184.

Kim, M., Frausto, R., Rosenwasser, G., Bui, T., Le, D., Stone, E. and Aldave, A. (2014) ‘Posterior amorphous corneal dystrophy is associated with a deletion of small leucine-rich proteoglycans on chromosome 12.’, *PLoS One*, 9(4), p. e95037.

Kinoshita, S., Nishida, K., Dota, A., Inatomi, T., Koizumi, N., Elliott, A., Lewis, D., Quantock, A. and Fullwood, N. (2000) ‘Epithelial barrier function and ultrastructure of gelatinous drop-like corneal dystrophy.’, *Cornea*, 19(4), pp. 551–555.

Kiritsi, D., Has, C. and Bruckner-Tuderman, L. (2013) ‘Laminin 332 in junctional epidermolysis bullosa’, *Cell Adhesion and Migration*, 7(1), pp. 135–141.

Kiuru, S. (1992) ‘Familial amyloidosis of the Finnish type (FAF): a clinical study of 30 patients.’, *Acta Neurologica Scandinavica*, 86(4), pp. 346–353.

Klausner, E., Peer, D., Chapman, R., Multack, R. and Andurkar, S. (2007) ‘Corneal gene therapy.’, *Journal of Controlled Release*, 124(3), pp. 107–133.

Klintworth, G. (2003) ‘The molecular genetics of corneal dystrophies - current status.’, *Frontiers in Bioscience*, 8, pp. 687–713.

Klintworth, G. K. (2009) ‘Corneal dystrophies.’, *Orphanet Journal of Rare Diseases*, 4(7).

Klintworth, G. K., Valnickova, Z., Kielar, R. a, Baratz, K. H., Campbell, R. J. and Enghild, J. J. (1997) ‘Familial subepithelial corneal amyloidosis: a lactoferrin-related amyloidosis.’, *Investigative Ophthalmology & Visual Science*, 38(13), pp. 2756–2763.

Klintworth, G., Smith, C. and Bowling, B. (2006) ‘CHST6 mutations in North American subjects with macular corneal dystrophy: a comprehensive molecular genetic review.’, *Molecular Vision*, 12, pp. 159–176.

Klintworth, G. and Vogel, F. (1964) ‘Macular corneal dystrophy: an inherited acid mucopolysaccharide storage disease of the corneal fibroblast’, *The American Journal of Pathology*, 45(4), pp. 565–586.

Kobayashi, A., Fujiki, K., Murakami, A. and Sugiyama, K. (2009) ‘In vivo laser confocal microscopy findings and mutational analysis for Schnyder’s crystalline corneal dystrophy.’, *Ophthalmology*, 116(6), pp. 1029–1037.

El Kochairi, I., Letovanec, I., Uffer, S., Munier, F., Chaubert, P. and Schorderet, D. (2006) ‘Systemic investigation of keratoepithelin deposits in TGFBI/BIGH3-related corneal dystrophy.’, *Molecular Vision*, 12, pp. 461–466.

Kodaganur, S., Kapoor, S., Veerappa, A., Tontanahal, S., Sarda, A., Yathish, S., Prakash, D. and Kumar, A. (2013) ‘Mutation analysis of the SLC4A11 gene in Indian families with congenital hereditary endothelial dystrophy 2 and a review of the literature.’, *Molecular Vision*, 19, pp. 1694–1706.

Kolli, S., Ahmad, S., Lako, M. and Figueiredo, F. (2010) ‘Successful clinical implementation of corneal epithelial stem cell therapy for treatment of unilateral limbal stem cell deficiency.’, *Stem Cells*, 28(3), pp. 597–610.

Kontadakis, G., Kymionid, G., Kankariya, V., Papadiamantis, A. and Pallikaris, A. (2014) ‘Corneal confocal microscopy findings in sporadic cases of Pre-Decemet corneal dystrophy.’, *Eye & Contact Lens*, 40(2), pp. e8–e12.

Korvatska, E., Munier, F., Djemai, A., Wang, M., Frueh, B., Choiu, A., Uffer, S., Ballestrazzi, E., Braunstein, R., Forster, R., Culberston, W. and Boman, H. (1998) 'Mutation hot spots in 5q31-linked corneal dystrophies.', *American Journal of Human Genetics*, 62, pp. 320–324.

Koss-Harnes, D., Høyheim, B., Anton-Lamprecht, I., Gjesti, A., Jørgensen, R., Jahnsen, F., Olaisen, B., Wiche, G. and Gedde-Dahl, T. (2002) 'A site-specific plectin mutation causes dominant epidermolysis bullosa simplex Ogna: Two identical de novo mutations.', *Journal of Investigative Dermatology*, 118(1), pp. 87–93.

Krafchak, C., Pawar, H., Moroi, S., Sugar, A., Lichter, P., Mackey, D., Mian, S., Nairus, T., Elner, V., Schteingart, M., Downs, C., Kijek, T., Johnson, J., Trager, E., Rozsa, F., Mandal, M., Epstein, M., Vollrath, D., Ayyagari, R., Boehnke, M. and Richards, J. (2005) 'Mutations in TCF8 cause posterior polymorphous corneal dystrophy and ectopic expression of COL4A3 by corneal endothelial cells.', *American Journal of Human Genetics*, 77(5), pp. 694–708.

Küchle, M., Green, W., Völcker, H. and Barraquer, J. (1995) 'Reevaluation of corneal dystrophies of Bowman's layer and the anterior stroma (Reis-Bucklers and Thiel-Behnke Types): a light and electron microscopic study of eight corneas and a review of the literature.', *Cornea*, 14(4), pp. 333–354.

Kyle, R., Therneau, T., Rajkumar, S., Offord, J., Larson, D., Plevak, M. and Melton, L. (2002) 'A long-term study of prognosis in monoclonal gammopathy of undetermined significance.', *The New England Journal of Medicine*, 346(8), pp. 564–569.

De La Chapelle, A., Tolvanen, R., Boysen, G., Santavy, J., Bleeker-Wagemakers, L., Maury, C. and Kere, J. (1992) 'Gelsolin–derived familial amyloidosis caused by asparagine or tyrosine substitution for aspartic acid at residue 187.', *Nature Genetics*, 2, pp. 157–160.

Laganowski, H., Sherrard, E., Muir, M. and Buckley, R. (1991) 'Distinguishing features of the iridocorneal endothelial syndrome and posterior polymorphous dystrophy: value of endothelial specular microscopy.', *British Journal of Ophthalmology*, 75(4), pp. 212–216.

Lai, I., Yellore, V., Rayner, S., D'Silva, N., Nguyen, C. and Aldave, A. (2010) 'The utility of next-generation sequencing in the evaluation of the posterior polymorphous corneal dystrophy 1 locus.', *Molecular Vision*, 16, pp. 2829–2838.

Lai, K., Reidy, J., Bert, B. and Milman, T. (2014) 'Spheroidal degeneration in H626R TGFBI variant lattice dystrophy: a multimodality analysis.', *Cornea*, 33(7), pp. 726–732.

Laibson, P. (1976) 'Microcystic corneal dystrophy.', *Transactions of the American Ophthalmological Society*, 74, pp. 488–531.

Laibson, P. (2010) 'Recurrent corneal erosions and epithelial basement membrane dystrophy.', *Eye & Contact Lens*, 36(5), pp. 315–317.

Laibson, P. and Krachmer, J. (1975) 'Familial occurrence of dot (microcystic), map, fingerprint dystrophy of the cornea.', *Investigative Ophthalmology & Visual Science*, 14(5), pp. 397–399.

Lakshminarayanan, R., Vithana, E., Chai, S., Chaurasia, S., Saraswathi, P.,

Venkatraman, A., Rojare, C., Venkataraman, D., Tan, D., Aung, T., Beuerman, R. and Mehta, J. (2011) 'A novel mutation in transforming growth factor-beta induced protein (TGFβIp) reveals secondary structure perturbation in lattice corneal dystrophy.', *The British Journal of Ophthalmology*, 95(10), pp. 1457–1462.

Lamouille, S., Xu, J. and Derynck, R. (2014) 'Molecular mechanisms of epithelial-mesenchymal transition.', *National Review Molecular Cell Biology*, 15(3), pp. 178–196.

Lander, E. (2016) 'The Heroes of CRISPR.', *Cell*, 164(1), pp. 18–28.

Lane, E., Rugg, E., Navsaria, H., Leigh, I., Heagerty, E., Ishida-Yamamoto, A. and Eady, R. (1992) 'A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering.', *Nature*, 356, pp. 244–246.

Le, D., Frausto, R. and Aldave, A. (2014) 'Identification of candidate genes for a corneal dystrophy of Bowman layer not associated with a TGFBI mutation.', *Investigative Ophthalmology & Visual Science*, 55(13), p. 1012.

Leachman, S., Hickerson, R., Schwartz, M., Bullough, E., Hutcherson, S., Boucher, K., Hansen, C., Eliason, M., Srivatsa, G., Kornbrust, D., Smith, F., McLean, W., Milstone, L. and Kaspar, R. (2010) 'First-in-human mutation-targeted siRNA phase Ib trial of an inherited skin disorder.', *Molecular Therapy*, 18(2), pp. 442–446.

Lechner, J., Dash, D., Muszynska, D., Hosseini, M., Segev, F., George, S., Frazer, D., Moore, J., Kaye, S., Young, T., Simpson, D., Churchill, A., Heon, E. and Willoughby, C. (2013) 'Mutational spectrum of the ZEB1 gene in corneal dystrophies supports a genotype-phenotype correlation', *Investigative Ophthalmology and Visual Science*, 54(5), pp. 3215–3223.

Lee, B., Jacobs, D., Musch, D., Kaufman, S., Reinhart, W. and Shtein, R. (2009) 'Descemet's stripping endothelial keratoplasty: safety and outcomes.', *Ophthalmology*, 116(9), pp. 1818–1830.

Lee, E. and Kim, E. (2003) 'Surgical do's and don'ts of corneal dystrophies.', *Current Opinion in Ophthalmology*, 14(4), pp. 186–191.

Lehmann, U. and Tost, J. (eds) (2015) *Pyrosequencing: methods and protocols.* 2nd ed. Humana Press.

Lek, M., Karczewski, K., Samocha, K., Banks, E., Fennell, T., O, A., Ware, J. S., Hill, A. J., Cummings, B. B., Birnbaum, D. P., Kosmicki, J. A., Duncan, L., Zhao, F., Zou, J., Pierce-Hoffman, E., Cooper, D. N., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Stenson, D., Stevens, C., Tiao, G., Tusie-Luna, M. T., Weisburd, B., Wilson, G., Daly, M. J. and MacArthur, D. G. (2016) 'Analysis of protein-coding genetic variation in 60,706 humans.', *Nature*, 536(7616), pp. 285–291.

Levy, E., Haltia, M., Fernandez-Madrid, I., Koivunen, O., Ghiso, J., Prelli, F. and Frangione, B. (1990) 'Mutation in gelsolin gene in Finnish hereditary amyloidosis.', *The Journal of Experimental Medicine*, 172(6), pp. 1865–1867.

Lewis, G. and Rader, D. (2005) 'New insights into the regulation of HDL metabolism and reverse cholesterol transport.', *Circulation Research*, 96(12), pp. 1221–1232.

Li, A., Jiao, X., Munier, F., Schorderet, D., Yao, W., Iwata, F., Hayakawa, M., Kanai, A., Chen, M., Lewis, R., Heckenlively, J., Weleber, R., Traboulsi, E., Zhang,

Q., Xiao, X., Kaiser-kupfer, M., Sergeev, Y. and Hejtmancik, J. (2004) 'Bietti crystalline corneoretinal dystrophy is caused by mutations in the novel gene CYP4V2.', *American Journal of Human Genetics*, 74, pp. 817–826.

Li, D., Peng, X. and Sun, H. (2015) 'Association of TCF4 polymorphisms and fuchs' endothelial dystrophy: a meta-analysis.', *BMC Ophthalmology*, 15(1), p. 61.

Li, D., Qi, Y., Wang, L., Lin, H., Zhou, N. and Zhao, L. (2008) 'An atypical phenotype of Reis-Bücklers corneal dystrophy caused by the G623D mutation in TGFBI.', *Molecular Vision*, 14, pp. 1298–1302.

Li, L., Bales, E., Peterson, C. and Legerski, R. (1993) 'Characterization of molecular defects in xeroderma pigmentosum group C.', *Nature Genetics*, 5, pp. 413–417.

Li, S., Tiab, L., Jiao, X., Munier, F., Zografos, L., Frueh, B., Sergeev, Y., Smith, J., Rubin, B., Meallet, M., Forster, R., Hejtmancik, J. and Schorderet, D. (2005) 'Mutations in PIP5K3 are associated with François-Neetens mouchetée fleck corneal dystrophy', *The American Journal of Human Genetics*, 77(1), pp. 54–63.

Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., Lv, J., Xie, X., Chen, Y., Li, Y., Sun, Y., Bai, Y., Songyang, Z., Ma, W., Zhou, C. and Huang, J. (2015) 'CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes.', *Protein & Cell*, 6(5), pp. 363–372.

Liang, Q., Pan, Z., Sun, X., Baudouin, C. and Labbé, A. (2014) 'Reis-Bücklers Corneal Dystrophy: A Reappraisal Using in vivo and ex vivo imaging techniques.', *Ophthalmic Research*, 51(4), pp. 187–95.

Liao, X., Cui, H. and Wang, F. (2013) 'Establishment of a transgenic mouse model of corneal dystrophy overexpressing human BIGH3.', *International Journal of Molecular Medicine*, 32(5), pp. 1110–1114.

Lim, Y., Ovejero, D., Sugarman, J., Deklotz, C., Maruri, A., Eichenfield, L., Kelley, P., Ju, H., Gottschalk, M., Tifft, C., Gafni, R., Boyce, A., Cowen, E., Bhattacharyya, N., Guthrie, L., Gahl, W., Golas, G., Loring, E., Overton, J. and Mane, S. (2014) 'Multilineage somatic activating mutations in HRAS and NRAS cause mosaic cutaneous and skeletal lesions, elevated FGF23 and hypophosphatemia.', *Human Molecular Genetics*, 23(2), pp. 397–407.

Lin, B. R., Frausto, R. F., Vo, R. C., Chiu, S. Y., Chen, J. L. and Aldave, A. J. (2016) 'Identification of the first de novo UBIAD1 gene mutation associated with Schnyder corneal dystrophy.', *Journal of Ophthalmology*.

Lin, Z., Chen, Q., Lee, M., Cao, X., Zhang, J., Ma, D., Chen, L., Hu, X., Wang, H., Wang, X., Zhang, P., Liu, X., Guan, L., Tang, Y., Yang, H., Tu, P., Bu, D., Zhu, X., Wang, K. and Li, R. (2012) 'Exome Sequencing Reveals Mutations in TRPV3 as a Cause of Olmsted Syndrome', *The American Journal of Human Genetics*. The American Society of Human Genetics, 90(3), pp. 558–564.

Liquori, A., Vaché, C., Baux, D., Blanchet, C., Hamel, C., Malcolm, S., Koenig, M., Claustres, M. and Roux, A. (2015) 'Whole USH2A gene sequencing identifies several new deep intronic mutations.', *Human Mutation*, 37(2), pp. 185–193.

Lisch, W., Bron, A., Munier, F., Schorderet, D., Tiab, L., Lange, C., Saikia, P., Reinhard, T., Weiss, J., Gundlach, E., Pleyer, U., Lisch, C. and Auw-Haedrich, C. (2012) 'Franceschetti hereditary recurrent corneal erosion.', *American Journal of Ophthalmology*, 153(6), pp. 1073–1081.

Lisch, W., Büttner, A., Oeffner, F., Böddeker, I., Engel, H., Lisch, C., Ziegler, A. and Grzeschik, K. (2000) 'Lisch corneal dystrophy is genetically distinct from Meesmann corneal dystrophy and maps to xp22.3.', *American Journal of Ophthalmology*, 130(4), pp. 461–468.

Lisch, W. and Seitz, B. (2011) 'The clinical landmarks of corneal dystrophies.', in *Corneal Dystrophies*, pp. 9–23.

Lisch, W. and Seitz, B. (2014) 'Lattice corneal dystrophy type 1: an epithelial or stromal entity?', *Cornea*, 33(10), pp. 1109–1112.

Liskova, P., Filipec, M., Merjava, S., Jirsova, K. and Tuft, S. (2010) 'Variable ocular phenotypes of posterior polymorphous corneal dystrophy caused by mutations in the ZEB1 gene.', *Ophthalmic Genetics*, 31(4), pp. 230–234.

Liskova, P., Gwilliam, R., Filipec, M., Jirsova, K., Reinstein Merjava, S., Deloukas, Webb, T., Bhattacharya, S., Ebenezer, N., Morris, A. and Hardcastle, A. (2012) 'High prevalence of posterior polymorphous corneal dystrophy in the Czech Republic; linkage disequilibrium mapping and dating an ancestral mutation.', *PLoS One*, 7(9), p. e45495.

Liskova, P., Palos, M., Hardcastle, A. and Vincent, A. (2013) 'Further genetic and clinical insights of posterior polymorphous corneal dystrophy 3.', *JAMA Ophthalmology*, 131(10), pp. 1296–1303.

Liu, Y., Peng, X., Tan, J., Darling, D., Kaplan, H. and Dean, D. (2008) 'Zeb1 mutant mice as a model of posterior corneal dystrophy.', *Investigative Ophthalmology and Visual Science*, 49(5), pp. 1843–1849.

Liu, Z., Tian, X., Iida, N., Fujiki, K., Xie, P. and Wang, W. (2010) 'Mutation analysis of CHST6 gene in Chinese patients with macular corneal dystrophy', *Cornea*, 29(8), pp. 883–888.

Loganathan, S. and Casey, J. (2014) 'Corneal dystrophy-causing SLC4A11 mutants: suitability for folding-correction therapy.', *Human Mutation*, 35(9), pp. 1082–1091.

Lorenzetti, D., Uotila, M., Parikh, N. and Kaufman, H. (1967) 'Central cornea guttata: incidence in the general population.', *American Journal of Ophthalmology*, 64(6), pp. 1155–1158.

Lüttmann, R., Teismann, I., Husstedt, I., Ringelstein, E. and Kuhlenbäumer, G. (2010) 'Hereditary amyloidosis of the Finnish type in a German family: clinical and electrophysiological presentation.', *Muscle & Nerve*, 41(5), pp. 679–684.

Mackay, D., Hu, M., Li, B., Rhéaume, C. and Dai, X. (2006) 'The mouse Ovol2 gene is required for cranial neural tube development', *Developmental Biology*, 291(1), pp. 38–52.

Majo, F., Rochat, A., Nicolas, M., Jaoudé, G. and Barrandon, Y. (2008) 'Oligopotent stem cells are distributed throughout the mammalian ocular surface.', *Nature*, 456(7219), pp. 250–254.

Mallery, D., Tanganelli, B., Colella, S., Steingrimsdottir, H., Gool, A., Troelstra, C., Stefanini, M. and Lehmann, A. (1998) 'Molecular analysis of mutations in the CSB (ERCC6) gene in patients with Cockayne syndrome.', *American Journal Human Genetics*, 62, pp. 77–85.

Mardis, E. (2013) 'Next-generation sequencing platforms.', *Annual Review of*

Analytical Chemistry, 6, pp. 287–303.

Mardy, S., Miura, Y., Endo, F., Matsuda, I., Frossard, P., Moosa, A., Ismail, E., Macaya, A., Andria, G., Toscano, E., Gibson, W., Graham, G. and Indo, Y. (1999) 'Congenital insensitivity to pain with anhidrosis: novel mutations in the TRKA (NTRK1) gene encoding a high-affinity receptor for nerve growth factor.', *American Journal of Human Genetics*, 64, pp. 1570–1579.

Mashima, Y., Nakamura, Y., Noda, K., Konishi, M., Yamada, M., Kudoh, J. and Shimizu, N. (1999) 'A novel mutation at codon 124 (R124L) in the BIGH3 gene is associated with a superficial variant of Granular corneal dystrophy.', *Opthalmic Molecular Genetics*, 117, pp. 90–93.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takoi, K. and Hanaoka, F. (1999) 'The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase.', *Nature*, 399, pp. 700–705.

Maury, C., Kere, J., Tolvanen, R. and de la Chapelle, A. (1990) 'Finnish hereditary amyloidosis is caused by a single nucleotide substitution in the gelsolin gene.', *FEBS Letters*, 276(1–2), pp. 75–77.

Mccabe, K., Kunzevitzky, N., Chiswell, B., Xia, X., Goldberg, J. and Lanza, R. (2015) 'Efficient generation of human embryonic stem cell-derived corneal endothelial cells by directed differentiation.', *PLoS One.*, 10(12), p. e0145266.

McGowan, S., Edelhauser, H., Pfister, R. and Whikehart, D. (2007) 'Stem cell markers in the human posterior limbus and corneal endothelium of unwounded and wounded corneas.', *Molecular Vision*, 13, pp. 1984–2000.

McGrath, J., Gatalica, B., Christiano, A., Li, K., Owaribe, K., McMillan, J., Eady, R. and Uitto, J. (1995) 'Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa.', *Nature Genetics*, 11, pp. 83–86.

McGrath, J., Kivirikko, S., Ciatti, S., Moss, C., Dunnill, G., Eady, R., Rodeck, C., Christiano, A. and Uitto, J. (1995) 'A homozygous nonsense mutation in the alpha-3 chain gene of laminin 5 (LAMA3) in Herlitz junctional epidermolysis bullosa: prenatal exclusion in a fetus at risk.', *Genomics*, 29, p. 2820284.

McLean, W., Rugg, E., Lunny, D., Morley, S., Lane, E., Swensson, O., Dopping-Hepenstal, P., Griffiths, W., Eady, R., Higgins, C., Navsaria, H., Leigh, I., Strachan, T., Kunkeler, L. and Munro, C. (1995) 'Keratin 16 and keratin 17 mutations cause pachyonychia congenita.', *Nature Genetics*, 9, pp. 273–278.

McVey, M. and Lee, S. (2008) 'MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings', *Trends in Genetics*, 24(11), pp. 529–538.

Meek, K. and Boote, C. (2004) 'The organization of collagen in the corneal stroma.', *Experimental Eye Research*, 78(3), pp. 503–512.

Meek, K., Dennis, S. and Khan, S. (2003) 'Changes in the refractive index of the stroma and its extrafibrillar matrix when the cornea swells.', *Biophysical Journal*, 85, pp. 2205–2212.

Meeney, A. and Mudhar, H. (2013) 'Histopathological reporting of corneal

pathology by a biomedical scientist: the Sheffield experience.', *Eye*, 27(2), pp. 272–276.

Mehta, J. S., Hemadevi, B., Vithana, E. N., Arunkumar, J., Srinivasan, M., Prajna, V., Tan, D. T., Aung, T. and Sundaresan, P. (2010) 'Absence of phenotype-genotype correlation of patients expressing mutations in the SLC4A11 gene.', *Cornea*, 29(3), pp. 302–6.

Meng, H., Matthaei, M., Ramanan, N., Grebe, R., Chakravarti, S., Speck, C., Kimos, M., Vij, N., Eberhart, C. G. and Jun, A. S. (2013) 'L450W and Q455K Col8a2 knock-in mouse models of fuchs endothelial corneal dystrophy show distinct phenotypes and evidence for altered autophagy.', *Investigative Ophthalmology and Visual Science*, 54(3), pp. 1887–1897.

Merjava, S., Malinova, E., Liskova, P., Filipec, M., Zemanova, Z., Michalova, K. and Jirsova, K. (2011) 'Recurrence of posterior polymorphous corneal dystrophy is caused by the overgrowth of the original diseased host endothelium.', *Histochemistry and Cell Biology*, 136(1), pp. 93–101.

Metzker, M. (2010) 'Sequencing technologies - the next generation.', *Nature reviews. Genetics*, 11(1), pp. 31–46.

Miccoli, R., Bertolotto, A., Odoguardi, L. and Boni, A. (1996) 'Compound heterozygosity for a structural apolipoprotein A-I variant, Apo A-I (L141R) Pisa and an apolipoprotein A-I null allele in patients with absence of HDL cholesterol, corneal opacifications, and coronary heart disease.', *Circulation*, 94(7), pp. 1622–1628.

Michalickova, K., Susic, M., Willing, M., Wenstrup, R. and Cole, W. (1998) 'Mutations of the α2(V) chain of type V collagen impair matrix assembly and produce Ehlers – Danlos syndrome type I.', *Human Molecular Genetics*, 7(2), pp. 249–255.

Milman, T., Kao, A., Chu, D., Gorski, M., Steiner, A., Simon, C., Shih, C., Aldave, A., Eagle, R., Jakobiec, F. and Udell, I. (2015) 'Paraproteinemic keratopathy: the expanding diversity of clinical and pathologic manifestations.', *Ophthalmology*, 122(9), pp. 1748–1756.

Mochizuki, T., Lemmink, H., Mariyama, M., Antignac, C., Gubler, M., Pirson, Y., Verellen-Dumoulin, C., Chan, B., Schroder, C., Smeets, H. and Reeders, S. (1994) 'Identification of mutations in the alpha-3(IV) and alpha-4(IV) collagen genes in autosomal recessive Alport syndrome.', *Nature Genetics*, 8, pp. 77–82.

Mohan, R., Tovey, J., Sharma, A. and Tandon, A. (2012) 'Gene therapy in the cornea: 2005-present.', *Progress in Retinal and Eye Research*, 31(1), pp. 43–64.

Mojica, F., Díez-Villaseñor, C., García-Martínez, J. and Soria, E. (2005) 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements.', *Journal of Molecular Evolution*, 60(2), pp. 174–182.

Mok, J., Kim, H. and Joo, C. (2009) 'Q455V mutation in COL8A2 is associated with Fuchs' corneal dystrophy in Korean patients.', *Eye*, 23(4), pp. 895–903.

Møller, H. U. and Ridgway, A. E. a. (1990) 'Granular corneal dystrophy Groenouw type I.', *Acta Ophthalmologica*, 68(1), pp. 384–389.

Moon, J., Kim, S., Kim, T., Cristol, S., Chung, E. and Kim, E. (2007) 'Homozygous

granular corneal dystrophy type II (Avellino corneal dystrophy): natural history and progression after treatment.', *Cornea*, 26(9), pp. 1095–1100.

Morales, C., Grigoryeva, L., Pan, X., Bruno, L., Hickson, G., Ngo, M., McMaster, C., Samuels, M. and Pshezhetsky, A. (2014) 'Mitochondrial damage and cholesterol storage in human hepatocellular carcinoma cells with silencing of UBIAD1 gene expression.', *Molecular Genetics and Metabolism Reports*, 1, pp. 407–411.

Moynihan, L., Bundey, S., Heath, D., Jones, E., McHale, D., Mueller, R., Markham, A. and Lench, N. (1998) 'Autozygosity mapping, to chromosome 11q25, of a rare autosomal recessive syndrome causing histiocytosis, joint contractures, and sensorineural deafness.', *American Journal of Human Genetics*, 62(5), pp. 1123–1128.

Mu, T., Zhang, Q., Pan, J., Steinmann, B., Vodopiutz, J., Gruber, R., Sonoda, T., Krabichler, B., Utermann, G., Baenziger, J. U., Zhang, L. and Janecke, A. R. (2009) 'Loss of dermatan-4-sulfotransferase 1 function results in adducted thumb-clubfoot syndrome.', *American Journal of Human Genetics*, 85, pp. 873–882.

Mugoni, V., Postel, R., Catanzaro, V., De Luca, E., Turco, E., Digilio, G., Silengo, L., Murphy, M., Medana, C., Stainier, D., Bakkers, J. and Santoro, M. (2013) 'Ubiad1 is an antioxidant enzyme that regulates eNOS activity by CoQ10 synthesis.', *Cell*, 152(3), pp. 504–518.

Mukhopadhyay, M., Gorivodsky, M., Shtrom, S., Grinberg, A., Niehrs, C., Morasso, M. and Westphal, H. (2006) 'Dkk2 plays an essential role in the corneal fate of the ocular surface epithelium.', *Development*, 133(11), pp. 2149–2154.

Müller, L., Marfurt, C., Kruse, F. and Tervo, T. (2003) 'Corneal nerves: structure, contents and function.', *Experimental Eye Research*, 76(5), pp. 521–542.

Munier, F., Frueh, B., Othenin-Girard, P., Uffer, S., Cousin, P., Wang, M., Héon, E., Black, G., Blasi, M., Balestrazzi, E., Lorenz, B., Escoto, R., Barraquer, R., Hoeltzenbein, M., Gloor, B., Fossarello, M., Singh, A., Arsenijevic, Y., Zografos, L. and Schorderet, D. (2002) 'BIGH3 mutation spectrum in corneal dystrophies.', *Investigative Ophthalmology & Visual Science*, 43(4), pp. 949–954.

Munier, F., Korvatska, E., Djemai, A., Paslier, D., Zografos, L., Pescia, G. and Schorderet, D. (1997) 'Kerato-epithelin mutations in four 5q31-linked corneal dystrophies.', *Nature Genetics*, 15(3), pp. 247–251.

Nakagawa, K., Hirota, Y., Sawada, N., Yuge, N., Watanabe, M., Uchino, Y., Okuda, N., Shimomura, Y., Suhara, Y. and Okano, T. (2010) 'Identification of UBIAD1 as a novel human menaquinone-4 biosynthetic enzyme.', *Nature*, 468(7320), pp. 117–121.

Nakajima, M., Mizumoto, S., Miyake, N., Kogawa, R., Iida, A., Ito, H., Kitoh, H., Hirayama, A., Mitsubuchi, H., Miyazaki, O., Kosaki, R., Horikawa, R., Lai, A., Mendoza-londono, R., Dupuis, L., Chitayat, D., Howard, A., Leal, G., Cavalcanti, D., Tsurusaki, Y., Saitsu, H., Superti-furga, A., Matsumoto, N., Sugahara, K., Nishimura, G. and Ikegawa, S. (2013) 'Mutations in B3GALT6, which encodes a glycosaminoglycan linker region enzyme, cause a spectrum of skeletal and connective tissue disorders.', *American Journal of Human Genetics*, 92, pp. 927–934.

Nakatsukasa, M., Kawasaki, S., Yamasaki, K., Fukuoka, H., Matsuda, A.,

Tsujikawa, M., Tanioka, H., Nagata-Takaoka, M., Hamuro, J. and Kinoshita, S. (2010) 'Tumor-associated calcium signal transducer 2 is required for the proper subcellular localization of claudin 1 and 7: implications in the pathogenesis of gelatinous drop-like corneal dystrophy.', *The American Journal of Pathology*, 177(3), pp. 1344–1355.

Nakazawa, Y., Sasaki, K., Mitsutake, N., Matsuse, M., Shimada, M., Nardo, T., Takahashi, Y., Ohyama, K., Ito, K., Mishima, H., Nomura, M., Kinoshita, A., Ono, S., Takenaka, K., Masuyama, R., Kudo, T., Slor, H., Utani, A., Tateishi, S., Yamashita, S., Stefanini, M., Lehmann, A., Yoshiura, K. and Ogi, T. (2012) 'Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription- coupled nucleotide-excision repair.', *Nature Genetics*. Nature Publishing Group, 44(5), pp. 586–592.

Natt, E., Kida, K., Odievre, M., Rocco, M. and Scherer, G. (1992) 'Point mutations in the tyrosine aminotransferase gene in tyrosinemia type II.', *Proceedings of the National Academy of Sciences of the United States of America*, 89, pp. 9297–9301.

Ng, P. C. and Henikoff, S. (2001) 'Predicting deleterious amino acid substitutions.', *Genome Research*, 11(5), pp. 863–874.

Ng, P. and Henikoff, S. (2003) 'SIFT: Predicting amino acid changes that affect protein function', *Nucleic Acids Research*, 31(13), pp. 3812–3814.

Ng, S., Bigham, A., Buckingham, K., Hannibal, M., McMillin, M., Gildersleeve, H., Beck, A., Tabor, H., Cooper, G., Mefford, H., Lee, C., Turner, E., Smith, J., Rieder, M., Yoshiura, K., Matsumoto, N., Ohta, T., Niikawa, N., Nickerson, D., Bamshad, M. and Shendure, J. (2010) 'Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome.', *Nature Genetics*, 42(9), pp. 790–793.

Ng, S., Buckingham, K., Lee, C., Bigham, A., Tabor, H., Dent, K., Huff, C., Shannon, P., Jabs, E., Nickerson, D., Shendure, J. and Bamshad, M. (2010) 'Exome sequencing identifies the cause of a mendelian disorder.', *Nature Genetics*, 42(1), pp. 30–35.

Ng, S., Turner, E., Robertson, P., Flygare, S., Abigail, W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Evan, E., Bamshad, M., Nickerson, D. and Shendure, J. (2010) 'Targeted capture and massicely parallel Sequencing of twelve human exomes.', *Nature*, 461(7261), pp. 272–276.

Nickerson, M., Bosley, A., Weiss, J., Kostiha, B., Hirota, Y., Brandt, W., Esposito, D., Kinoshita, S., Wessjohann, L., Morham, S., Andresson, T., Kruth, H., Okano, T. and Dean, M. (2013) 'The UBIAD1 prenyltransferase links menaquinone-4 synthesis to cholesterol metabolic enzymes.', *Human Mutation*, 34(2), pp. 317–329.

Nickerson, M., Kostiha, B., Brandt, W., Fredericks, W., Xu, K., Yu, F., Gold, B., Chodosh, J., Goldberg, M., Lu, D., Yamada, M., Tervo, T., Grutzmacher, R., Croasdale, C., Hoeltzenbein, M., Sutphin, J., Malkowicz, S., Wessjohann, L., Kruth, H., Dean, M. and Weiss, J. (2010) 'UBIAD1 mutation alters a mitochondrial prenyltransferase to cause Schnyder corneal dystrophy.', *PLoS One*, 5(5), p. e10760.

Niel, F., Ellies, P., Dighiero, P., Soria, J., Sabbagh, C., San, C., Renard, G., Delpech, M. and Valleix, S. (2003) 'Truncating Mutations in the Carbohydrate Sulfotransferase 6 Gene (CHST6) Result in Macular Corneal Dystrophy', *Investigative Ophthalmology & Visual Science*, 44(7), pp. 2949–2953.

Nishida, K., Honma, Y., Dota, A., Kawasaki, S., Adachi, W., Nakamura, T., Quantock, A., Hosotani, H., Yamamoto, S., Okada, M., Shimomura, Y. and Kinoshita, S. (1997) 'Isolation and chromosomal localization of a cornea-specific human keratin 12 gene and detection of four mutations in Meesmann corneal epithelial dystrophy.', *American Journal of Human Genetics*, 61(6), pp. 1268–1275.

Notara, M. and Daniels, J. (2010) 'Characterisation and functional features of a spontaneously immortalised human corneal epithelial cell line with progenitor-like characteristics.', *Brain Research Bulletin*, 81(2), pp. 279–286.

Nouspikel, T. and Ciarkson, S. (1994) 'Mutations that disable the DNA repair gene XPG in a xeroderma pigmentosum group G patient.', *Human Molecular Genetics*, 3(6), pp. 963–967.

Nowińska, A., Wylegala, E., Janiszewska, D., Dobrowolski, D., Aragona, P., Roszkowska, A. and Puzzolo, D. (2011) 'Genotype-phenotype correlation of TGFBI corneal dystrophies in Polish patients.', *Molecular Vision*, 17, pp. 2333–2342.

Nowinska, A., Wylegala, E. and Teper, S. (2014) 'Phenotype–genotype correlation in patients With Schnyder corneal dystrophy.', *Cornea*, 33(5), pp. 497–503.

Oberstein, S., Kriek, M., White, S., Kalf, M., Szuhai, K., Dunnen, J., Breuning, M. and Hennekam, R. (2006) 'Peters plus syndrome is caused by mutations in B3GALTL, a putative glycosyltransferase.', *American Journal Human Genetics*, 79, pp. 562–566.

Oeffner, F., Fischer, G., Happle, R., Konig, A., Betz, R., Bornholdt, D., Neidel, U., Boente, M., Redler, S., Romero-Gomez, J., Salhi, A., Vera-Casano, A., Weirich, C. and Grzeschik, K.-H. (2009) 'IFAP Syndrome is caused by deficiency in MBTPS2, an intramembrane zinc metalloprotease essential for cholesterol homeostasis and ER stress response.', *American Journal Human Genetics*, 84, pp. 459–467.

Okada, M., Yamamoto, S., Watanabe, H., Inoue, Y., Tsujikawa, M., Maeda, N., Shimomura, Y., Nishida, K., Kinoshita, S. and Tano, Y. (1998) 'Granular corneal dystrophy with homozygous mutations in the kerato-epithelin gene', *American Journal of Ophthalmology*, 126(2), pp. 169–176.

Oliver, V., van Bysterveldt, K., Cadzow, M., Steger, B., Romano, V., Markie, D., Hewitt, A., Mackey, D., Willoughby, C., Sherwin, T., Crosier, P., McGhee, C. and Vincent, A. (2016) 'A COL17A1 Splice-Altering Mutation Is Prevalent in Inherited Recurrent Corneal Erosions', *Ophthalmology*, 123(4), pp. 709–722.

Oliver, V. and Vincent, A. (2016) 'The genetics and pathophysiology of IC3D category 1 corneal dystrophies.', *Asia-Pacific Journal of Ophthalmology*, 5(4), pp. 272–281.

Orr, A., Dubé, M., Marcadier, J., Jiang, H., Federico, A., George, S., Seamone, C., Andrews, D., Dubord, P., Holland, S., Provost, S., Mongrain, V., Evans, S., Higgins, B., Bowman, S., Guernsey, D. and Samuels, M. (2007) 'Mutations in the UBIAD1 gene, encoding a potential prenyltransferase, are causal for Schnyder crystalline corneal dystrophy.', *PLoS One*, 2(8), p. e685.

Paliwal, P., Sharma, A., Tandon, R., Sharma, N., Titiyal, J., Sen, S., Kaur, P., Dube, D. and Vajpayee, R. (2010) 'TGFBI mutation screening and genotype-phenotype correlation in north Indian patients with corneal dystrophies.', *Molecular Vision*, 16, pp. 1429–1438.

Pandey, V., Nutter, R. and Prediger, E. (2008) 'Applied Biosystems SOLiD system: ligation-based sequencing.', in *Next Generation Genome Sequencing: Towards Personalized Medicine*, pp. 29–42.

Paunio, T., Kangas, H., Heinonen, O., Buc-Caron, M., Robert, J., Kaasinen, S., Julkunen, I., Mallet, J. and Peltonen, L. (1998) 'Cells of the neuronal lineage play a major role in the generation of amyloid precursor fragments in gelsolin-related amyloidosis', *Journal of Biological Chemistry*, 273(26), pp. 16319–16324.

Paunio, T., Kangas, H., Kalkklnen, N., Haltia, M., Palo, J. and Peltonen, L. (1994) 'Toward understanding the pathogenic mechanisms in gelsolin-related amyloidosis: in vitro expression reveals an abnormal gelsolin fragment.', *Human Molecular Genetics*, 3(12), pp. 2223–2229.

Pearce, W., Tripathi, R. and Morgan, G. (1969) 'Congenital endothelial corneal dystrophy. Clinical, pathological, and genetic study.', *British Journal of Ophthalmology*, 53(9), pp. 577–591.

Pellegata, N., Dieguez-Lucena, J., Joensuu, T., Lau, S., Montgomery, K., Krahe, R., Kivelä, T., Kucherlapati, R., Forsius, H. and de la Chapelle, A. (2000) 'Mutations in KERA, encoding keratocan, cause cornea plana.', *Nature Genetics*, 25, pp. 91–95.

Perry, H. and Cameron, J. (2006) 'Pathology of the cornea-sclera.', *Duane's Foundations of Clinical Ophthalmology*. Edited by W. Tasman and E. Jaeger, 3(9), pp. 33–37.

Petrof, G., Nanda, A., Howden, J., Takeichi, T., McMillan, J., Aristodemou, S., Ozoemena, L., Liu, L., South, A., Pourreyron, C., Dafou, D., Proudfoot, L., Al-Ajmi, H., Akiyama, M., Irwin McLean, W., Simpson, M., Parsons, M. and McGrath, J. (2014) 'Mutations in GRHL2 result in an autosomal-recessive ectodermal dysplasia syndrome.', *American Journal of Human Genetics*, 95(3), pp. 308–314.

Pineda-Cárdenas, D. and González-Almaráz, G. (2002) 'Manifestaciones ultraestructurales , histoquímicas e inmunopatológicas de la betafibrilosis corneal primaria.', *Revista Mexicana de Oftalmología*, 76(4), pp. 128–135.

Pisciotta, L., Miccoli, R., Cantafora, A., Calabresi, L., Tarugi, P., Alessandrini, P., Bittolo Bon, G., Franceschini, G., Cortese, C., Calandra, S. and Bertolini, S. (2003) 'Recurrent mutations of the apolipoprotein A-I gene in three kindreds with severe HDL deficiency', *Atherosclerosis*, 167(2), pp. 335–345.

Plagnol, V., Curtis, J., Epstein, M., Mok, K., Stebbings, E., Grigoriadou, S., Wood, N., Hambleton, S., Burns, S., Thrasher, A., Kumararatne, D., Doffinger, R. and Nejentsev, S. (2012) 'A robust model for read count data in exome sequencing experiments and implications for copy number variant calling.', *Bioinformatics*, 28(21), pp. 2747–2754.

Pressly, T. A., Scott, W. J., Ide, C. H., Winkler, A. and Reams, G. P. (1987) 'Ocular complications of tangier disease', *The American Journal of Medicine*, 83(5), pp. 991–994.

Price, F., Feng, M. and Price, M. (2015) 'Evolution of endothelial keratoplasty : Where Are We headed? Lowering the visual impairment threshold.', *Cornea*, 34(10S), pp. S41–S47.

Pulkkinen, L., Christiano, A., Gerecke, D., Wagman, D., Burgeson, R., Pittelkow, M. and Uitto, J. (1994) 'A homozygous nonsense mutation in the beta-3 chain gene

of Laminin-5 (LAMB3) in Herltiz junctional epidermolysis bullosa.', *Genomics*, 24, pp. 357–360.

Pullinger, C., Hennessy, L., Chatterton, J., Liu, I., Love, J., Mendel, C., Frost, P., Malloy, M., Schumaker, V. and Kanet, J. (1995) 'Familial ligand-defective apolipoprotein B.', *Journal of Clinical Investigation*, 95, pp. 1225–1234.

Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J. and Williams, S. (2001) *Neuroscience.* 2nd edn.

Rabbani, B., Tekin, M. and Mahdieh, N. (2014) 'The promise of whole-exome sequencing in medical genetics.', *Journal of Human Genetics*, 59(1), pp. 5–15.

Ramot, Y., Molho-Pessach, V., Meir, T., Alper-Pinus, R., Siam, I., Tams, S., Babay, S. and Zlotogorski, A. (2014) 'Mutation in KANK2, encoding a sequestering protein for steroid receptor coactivators, causes keratoderma and woolly hair.', *Journal of Medical Genetics*, 51, pp. 388–394.

Ramprasad, V., Ebenezer, N., Aung, T., Rajagopal, R., Yong, V., Tuft, S., Viswanathan, D., El-Ashry, M., Liskova, P., Tan, D., Bhattacharya, S., Kumaramanickavel, G. and Vithana, E. (2007) 'Novel SLC4A11 mutations in patients with recessive congenital hereditary endothelial dystrophy (CHED2).', *Human Mutation*, 28(5), pp. 522–523.

Ran, F., Hsu, P., Wright, J., Agarwala, V., Scott, D. and Zhang, F. (2013) 'Genome engineering using the CRISPR-Cas9 system.', *Nature Protocols*, 8(11), pp. 2281–2308.

Raper, S. E., Chirmule, N., Lee, F. S., Wivel, N. A., Bagg, A., Gao, G. P., Wilson, J. M. and Batshaw, M. L. (2003) 'Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer.', *Molecular Genetics and Metabolism*, 80, pp. 148–158.

Raychowdhury, M., Gonzalez-Perrett, S., Montalbetti, N., Timpanaro, G., Chasan, B., Goldmann, W., Stahl, S., Cooney, A., Goldin, E. and Cantiello, H. (2004) 'Molecular pathophysiology of mucolipidosis type IV: pH dysregulation of the mucolipin-1 cation channel.', *Human Molecular Genetics*, 13(6), pp. 617–627.

Reese, M., Eeckman, F., Kulp, D. and Haussler, D. (1997) 'Improved splice site detection in Genie.', *Journal of Computational Biology*, 4(3), pp. 311–323.

Reidy, J., Paulus, M. and Gona, S. (2000) 'Recurrent erosions of the cornea: epidemiology and treatment.', *Cornea*, 19(6), pp. 767–771.

Reinhart, W., Musch, D., Jacobs, D., Lee, W., Kaufman, S. and Shtein, R. (2011) 'Deep anterior lamellar keratoplasty as an alternative to penetrating keratoplasty.', *Ophthalmology*, 118(1), pp. 209–218.

Reinstein, D., Archer, T., Gobbe, M., Silverman, R. and Coleman, D. (2008) 'Epithelial thickness in the normal cornea: three-dimentional display with very high frequency ultrasound.', *Journal of Refractive Surgery*, 24(6), pp. 571–581.

Ren, Z., Lin, P., Klintworth, G., Iwata, F., Munier, F., Schorderet, D., El Matri, L., Theendakara, V., Basti, S., Reddy, M. and Hejtmancik, J. (2002) 'Allelic and locus heterogeneity in autosomal recessive gelatinous drop-like corneal dystrophy.', *Human Genetics*, 110(6), pp. 568–577.

Reneker, L., Silversides, D., Xu, L. and Overbeek, P. (2000) 'Formation of corneal

endothelium is essential for anterior segment development - a transgenic mouse model of anterior segment dysgenesis.', *Development*, 127(3), pp. 533–542.

Riazuddin, S., Eghrari, A., Al-Saif, A., Davey, L., Meadows, D., Katsanis, N. and Gottsch, J. (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(12), pp. 5667–5671.

Riazuddin, S., Parker, D., McGlumphy, E., Oh, E., Iliff, B., Schmedt, T., Jurkunas, U., Schleif, R., Katsanis, N. and Gottsch, J. (2012) 'Mutations in LOXHD1, a recessive-deafness locus, cause dominant late-onset Fuchs corneal dystrophy.', *American Journal of Human Genetics*, 90(3), pp. 533–539.

Riazuddin, S., Vasanth, S., Katsanis, N. and Gottsch, J. (2013) 'Mutations in AGBL1 cause dominant late-onset Fuchs corneal dystrophy and alter protein-protein interaction with TCF4.', *American Journal of Human Genetics*, 93(4), pp. 758–764.

Riazuddin, S., Zaghloul, N., Al-Saif, A., Davey, L., Diplas, B., Meadows, D., Eghrari, A., Minear, M., Li, Y., Klintworth, G., Afshari, N., Gregory, S., Gottsch, J. and Katsanis, N. (2010) 'Missense mutations in TCF8 cause late-onset Fuchs corneal dystrophy and interact with FCD4 on chromosome 9p.', *American Journal of Human Genetics*, 86(1), pp. 45–53.

Richard, Rouan, F., Willoughby, C., Brown, N., Chung, P., Jabs, E., Bale, S., Digiovanna, J., Uitto, J. and Russell, L. (2002) 'Missense mutations in GJB2 encoding connexin-26 cause the ectodermal dysplasia keratitis-ichthyosis-deafness syndrome.', *American Journal of Human Genetics*, 70, pp. 1341–1348.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K. and Rehm, H. L. (2015) 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology', *Genetics in Medicine*, 17(5), pp. 405–423.

Ridgway, A., Akhtar, S., Munier, F., Schorderet, D., Stewart, H., Perveen, R., Bonshek, R., Odenthal, M., Dixon, M., Barraquer, R., Escoto, R. and Black, G. (2000) 'Ultrastructural and molecular analysis of Bowman's layer corneal dystrophies: an epithelial origin?', *Investigative Ophthalmology and Visual Science*, 41(11), pp. 3286–3292.

Rifat, Y., Parekh, V., Wilanowski, T., Hislop, N., Auden, A., Ting, S., Cunningham, J. and Jane, S. (2010) 'Regional neural tube closure defined by the grainy head-like transcription factors', *Developmental Biology*, 345(2), pp. 237–245.

Ripani, E., Sacchetti, A., Corda, D. and Alberti, S. (1998) 'Human Trop-2 is a tumor-associated calcium signal transducer.', *International Journal of Cancer.*, 76(5), pp. 671–676.

Rittenbach, T. (2013) 'A case of Schnyder corneal dystrophy with crystals.', *Optometry & Vision Science*, 90(12), pp. 301–304.

Rohmann, E., Brunner, H., Kayserili, H., Uyguner, O., Nurnberg, G., Lew, E., Dobbie, A., Eswarakumar, V., Uzumcu, A., Ulubil-Emeroglu, M., Leroy, J., Li, Y., Becker, C., Legnerdt, K., Cremers, C., Memnune, Y.-A., Schlessinger, J., van Bokhoven, H. and Wollnik, B. (2006) 'Mutations in different components of FGF

signaling in LADD syndrome.', *Nature Genetics*, 38(4), pp. 414–417.

Römling, R., von Eckardstein, A., Funke, H., Motti, C., Fragiacomo, G., Noseda, G. and Assmann, G. (1994) 'A nonsense mutation in the apolipoprotein A-I gene is associated with high-density lipoprotein deficiency and periorbital xanthelasmas.', *Arteriosclerosis,Thrombosis and Vascular Biology*, 14(12), pp. 1915–1922.

Rowczenio, D., Dogan, A., Theis, J. D., Vrana, J. a., Lachmann, H. J., Wechalekar, A. D., Gilbertson, J. a., Hunt, T., Gibbs, S. D. J., Sattianayagam, P. T., Pinney, J. H., Hawkins, P. N. and Gillmore, J. D. (2011) 'Amyloidogenicity and clinical phenotype associated with five novel mutations in apolipoprotein A-I.', *The American Journal of Pathology*, 179(4), pp. 1978–1987.

Rüetschi, U., Cerone, R., Cristina, M., Sue, S., Ugarte, M. and Holme, E. (2000) 'Mutations in the 4-hydroxyphenylpyruvate dioxygenase gene (HPD) in patients with tyrosinemia type III.', *Human Genetics*, 106, pp. 654–662.

Runager, K., Basaiawmoit, R., Deva, T., Andreasen, M., Valnickova, Z., Sørensen, C., Karring, H., Thøgersen, I., Christiansen, G., Underhaug, J., Kristensen, T., Nielsen, N., Klintworth, G., Otzen, D. and Enghild, J. (2011) 'Human phenotypically distinct TGFBI corneal dystrophies are linked to the stability of the fourth FAS1 domain of TGFBIp.', *The Journal of Biological Chemistry*, 286(7), pp. 4951–4958.

Ruzzi, L., Gagnoux-palacios, L., Pinola, M., Belli, S., Meneguzzi, G., Alessio, M. and Zambruno, G. (1997) 'A homozygous mutation in the integrin alpha-6 gene in junctional epidermolysis bullosa with pyloric atresia.', *Journal of Clinical Investigation*, 99, pp. 2826–2831.

Sacchetti, M., Mantelli, F., Marenco, M., Macchi, I., Ambrosio, O. and Rama, P. (2015) 'Diagnosis and management of iridocorneal endothelial syndrome.', *BioMed Research International*.

Sanger, F., Nicklen, S. and Coulson, A. (1977) 'DNA sequencing with chain-terminating inhibitors.', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463–5467.

Sawada, H., Konomi, H. and Hirosawa, K. (1990) 'Characterization of the collagen in the hexagonal lattice of Descemet's membrane: its relation to type VIII collagen.', *The Journal of Cell Biology*, 110, pp. 219–227.

Schmid, E., Lisch, W., Philipp, W., Lechner, S., Göttinger, W., Schlötzer-Schrehardt, U., Müller, T., Utermann, G. and Janecke, A. (2006) 'A new, X-linked endothelial corneal dystrophy.', *American Journal of Ophthalmology*, 141(3), pp. 478–487.

Schmitt-Bernard, C., Guittard, C., Arnaud, B., Demaille, J., Argiles, A., Claustres, M. and Tuffery-Giraud, S. (2000) 'BIGH3 exon 14 mutations lead to intermediate type I/IIIA of lattice corneal dystrophies.', *Investigative Ophthalmology & Visual Science*, 41(6), pp. 1302–1308.

Schmitt-Bernard, C., Schneider, C. and Argiles, A. (2002) 'Clinical, histopathologic, and ultrastructural characteristics of βIGH3 (TGFBI) amyloid corneal dystrophies are supportive of the existence of a new type of LCD: The LCDi', *Cornea*, 21(5), pp. 463–468.

Schoenhagen, P., Cooper, C., Yasin, M., Eaton, G., Lauer, M., Sheldon, W., Grines, C., Halpern, S., Crowe, T., Blankenship, J., Nissen, S., Tsunoda, T., Tuzcu, E. and

Kerensky, R. (2003) ‘Effect of recombinant ApoA-I Milano on coronary atherosclerosis in patients with acute coronary syndromes.’, *JAMA*, 290(17), pp. 2292–2300.

Scott, H., Litjens, T., Hopwood, J. and Morris, C. (1992) ‘A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype.’, *Human Mutation*, 108, pp. 103–108.

Secker, G. and Daniels, J. (2009) *Limbal epithelial stem cells of the cornea., StemBook*.

Semina, E., Brownell, I., Mintz-hittner, H., Murray, J. and Jamrich, M. (2001) ‘Mutations in the human forkhead transcription factor FOXE3 associated with anterior segment ocular dysgenesis and cataracts.’, *Human Molecular Genetics*, 10(3), pp. 231–236.

Semina, E., Ferrell, R., Mintz-hittner, H., Bitoun, P., Alward, W., Reiter, R., Funkhauser, C., Daack-hirsch, S. and Murray, J. (1998) ‘A novel homeobox gene PITX3 is mutated in families with autosomal-dominant cataracts and ASMD.’, *Nature Genetics*, 19, pp. 167–170.

Sengillo, J., Justus, S., Tsai, Y., Cabral, T. and Tsang, S. (2016) ‘Gene and cell-based therapies for inherited retinal disorders: an update.’, *American Journal of Medical Genetics*, 172(4), pp. 349–366.

Serra, H., Holopainen, J., Beuerman, R., Kaarniranta, K., Suarez, M. and Urrets-Zavalia, J. (2015) ‘Climatic droplet keratopathy: an old disease in new clothes.’, *Acta Ophthalmologica*, 93(6), pp. 496–504.

Sethi, S., Theis, J. and Quint, P. (2013) ‘Renal amyloidosis associated with a novel sequence variant of gelsolin.’, *American Journal of Kidney Disease*, 61(1), pp. 161–166.

Sevel, D. and Isaacs, R. (1988) ‘A re-evaluation of corneal development.’, *Transactions of the American Ophthalmological Society*, 86, pp. 178–207.

Shendure, J. and Ji, H. (2008) ‘Next-generation DNA sequencing.’, *Nature Biotechnology*, 26(10), pp. 1135–1145.

Shimizu, S., Krafchak, C., Fuse, N., Epstein, M., Schteingart, M., Sugar, A., Eibschitz-Tsimhoni, M., Downs, C., Rozsa, F., Trager, E., Reed, D., Boehnke, M., Moroi, S. and Richards, J. (2004) ‘A locus for posterior polymorphous corneal dystrophy (PPCD3) maps to chromosome 10.’, *American Journal of Medical Genetics*, 130A(4), pp. 372–377.

Shortt, A., Secker, G., Munro, P., Khaw, P., Tuft, S. and Daniels, J. (2007) ‘Characterization of the limbal epithelial stem cell niche: novel imaging techniques permit in vivo observation and targeted biopsy of limbal epithelial stem cells.’, *Stem Cells*, 25(6), pp. 1402–1409.

Shortt, A., Tuft, S. and Daniels, J. (2011) ‘Corneal stem cells in the eye clinic.’, *British Medical Bulletin*, 100(1), pp. 209–225.

Siddiqui, S., Zenteno, J., Rice, A., Chacón-Camacho, O., Naylor, S., Rivera-de la Parra, D., Spokes, D., James, N., Toomes, C., Inglehearn, C. and Ali, M. (2013) ‘Congenital Hereditary Endothelial Dystrophy Caused by SLC4A11 Mutations Progresses to Harboyan Syndrome.’, *Cornea*, 33(3), pp. 247–251.

Silla, Z., Naidoo, J., Kidson, S. and Sommer, P. (2014) ‘Signals from the lens and Foxc1 regulate the expression of key genes during the onset of corneal endothelial development.’, *Experimental Cell Research*, 322(2), pp. 381–388.

Small, K., Mullen, L., Barletta, J., Graham, K., Glasgow, B., Stern, G. and Yee, R. (1996) ‘Mapping of Reis-Bucklers’ corneal dystrophy to chromosome 5q.’, *American Journal of Ophthalmology*, 121(4), pp. 384–390.

Smith, F., Eady, R., Leigh, I., McMillan, J., Rugg, E., Kelsell, D., Bryant, S., Spurr, N., Kirtschig, G., Milana, G., de Bono, A., Owaribe, K., Wiche, G., Pulkkinen, L., Uitto, J., McLean, W. and Lane, E. (1996) ‘Plectin deficiency results in muscular dystrophy with epidermolysis bullosa’, *Nature Genetics*, 13, pp. 450–457.

Smith, F., Jonkman, M., Goor, H., Coleman, C., Covello, S., Uitto, J. and Mclean, W. (1998) ‘A mutation in human keratin K6b produces a phenocopy of the K17 disorder pachyonychia congenita type 2.’, *Human Molecular Genetics*, 7(7), pp. 1143–1148.

Smith, L., Schwarze, U., Goldstein, J. and Byerstt, P. (1997) ‘Mutations in the COL3A1 gene result in the Ehlers–Danlos syndrome type IV and alterations in the size and distribution of the major collagen fibrils of the dermis.’, *Journal of Investigative Dermatology*. Elsevier Masson SAS, 108, pp. 241–247.

Snip, R., Kenyon, K. and Green, W. (1973) ‘Macular corneal dystrophy: ultrastructural pathology of the corneal endothelium and Descemet’s membrane.’, *Investigative Ophthalmology and Visual Science*, 12(2), pp. 88–97.

Solomon, J., Page, L., Balch, W. and Kelly, J. (2012) ‘Gelsolin amyloidosis: genetics, biochemistry, pathology and possible strategies for therapeutic intervention.’, *Critical Reviews in Biochemistry and Molecular Biology*, 47(3), pp. 282–296.

Solomon, J., Yonemoto, I., Murray, A., Price, J., Powers, E., Balch, W. and Kelly, J. (2009) ‘The 8 and 5 kDa fragments of plasma gelsolin form amyloid fibrils by a nucleated polymerization mechanism, while the 68 kDa fragment is not amyloidogenic.’, *Biochemistry*, 48(48), pp. 11370–11380.

Sorek, R., Kunin, V. and Hugenholtz, P. (2008) ‘CRISPR- a widespread system that provides acquired resistance against phages in bacteria and archaea.’, *Nature Reviews. Microbiology*, 6(3), pp. 181–186.

Sorensen, C., Runager, K., Scavenius, C., Jensen, M., Nielsen, N., Christiansen, G., Petersen, S., Karring, H., Sanggaard, K. and Enghild, J. (2015) ‘Fibril core of transforming growth factor beta-induced protein (TGFBIp) facilitates aggregation of corneal TGFBIp.’, *Biochemistry*, 54(19), pp. 2943–2956.

Sowden, J. (2007) ‘Molecular and developmental mechanisms of anterior segment dysgenesis.’, *Eye*, 21(10), pp. 1310–1318.

Spaderna, S., Schmalhofer, O., Wahlbuhl, M., Dimmler, A., Bauer, K., Sultan, A., Hlubek, F., Jung, A., Strand, D., Eger, A., Kirchner, T., Behrens, J. and Brabletz, T. (2008) ‘The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer.’, *Cancer Research*, 68(2), pp. 537–544.

Srinivas, S. (2012) ‘Cell signaling in regulation of the barrier integrity of the corneal endothelium.’, *Experimental Eye Research*, 95(1), pp. 8–15.

Stepensky, P., Keller, B., Shamriz, O., NaserEddin, A., Rumman, N., Weintraub, M., Warnatz, K., Elpeleg, O. and Barak, Y. (2016) 'Deep intronic mis-splicing mutation in JAK3 gene underlies T-B+NK- severe combined immunodeficiency phenotype.', *Clinical Immunology*, 163, pp. 91–95.

Stewart, H., Ridgway, A., Dixon, M., Bonshek, R., Parveen, R. and Black, G. (1999) 'Heterogeneity in granular corneal dystrophy: identification of three causative mutations in the TGFBI (BIGH3) gene-lessons for corneal amyloidogenesis.', *Human Mutation*, 14(2), pp. 126–132.

Stone, E., Fingert, J., Alward, W., Nguyen, T., Polansky, J., Sunden, S., Nishimura, D., Clark, A., Nystuen, A., Nichols, B., Mackey, D., Ritch, R., Kalenak, J., Craven, E. and Sheffield, V. (1997) 'Identification of a Gene That Causes Primary Open Angle Glaucoma', *Science*, 275(5300), pp. 668–670.

Stone, E. and Mathers, W. (1994) 'Three autosomal dominant corneal dystrophies map to chromosome 5q.', *Nature*, 6, pp. 47–51.

Strachan, I. (1969) 'Cloudy central corneal dystrophy of François. Five cases in the same family.', *The British Journal of Ophthalmology*, 53(3), pp. 192–194.

Strachan, T. and Read, A. P. (2011) *Human molecular genetics.* New York: Garland Science.

Sturrock, G., Sherrard, E. and Rice, N. (1978) 'Specular microscopy of the corneal endothelium.', *The British Journal of Ophthalmology*, 62(12), pp. 809–814.

Suesskind, D., Auw-Haedrich, C., Schorderet, D., Munier, F. and Loeffler, K. (2006) 'Keratoepithelin in secondary corneal amyloidosis.', *Graefe's Archive for Clinical and Experimental Ophthalmology*, 244(6), pp. 725–731.

Sultana, A., Garg, P., Ramamurthy, B., Vemuganti, G. and Kannabiran, C. (2007) 'Mutational spectrum of the SLC4A11 gene in autosomal recessive congenital hereditary endothelial dystrophy.', *Molecular Vision*, 13, pp. 1327–1332.

Sultana, A., Sridhar, M., Klintworth, G., Balasubramanian, D. and Kannabiran, C. (2005) 'Allelic heterogeneity of the carbohydrate sulfotransferase-6 gene in patients with macular corneal dystrophy.', *Clinical Genetics*, 68(5), pp. 454–460.

Sultana, A., Sridhar, M. S., Jagannathan, A., Balasubramanian, D., Kannabiran, C. and Klintworth, G. K. (2003) 'Novel mutations of the carbohydrate sulfotransferase-6 (CHST6) gene causing macular corneal dystrophy in India.', *Molecular vision*, 9(December), pp. 730–4.

Sunada, Y., Shimizu, T., Nakase, H., Ohta, S., Asaoka, T., Amano, S., Sawa, M., Kagawa, Y., Kanazawa, I. and Mannen, T. (1993) 'Inherited amyloid polyneuropathy type IV (gelsolin variant) in a Japanese family.', *Annals of Neurology*, 33(1), pp. 57–62.

Sundin, O., Jun, A., Broman, K., Liu, S., Sheehan, S., Vito, E., Stark, W. and Gottsch, J. (2006) 'Linkage of late-onset Fuchs corneal dystrophy to a novel locus at 13pTel-13q12.13.', *Investigative Ophthalmology & Visual Science*, 47(1), pp. 140–145.

Szaflik, J., Ołdak, M., Maksym, R., Kamińska, A., Pollak, A., Udziela, M., Płoski, R. and Szaflik, J. (2008) 'Genetics of Meesmann corneal dystrophy: a novel mutation in the keratin 3 gene in an asymptomatic family suggests genotype-

phenotype correlation.', *Molecular Vision*, 14, pp. 1713–1718.

Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M., Satoh, Y., Kondo, S., Yasui, A., Okayama, H. and Okada, Y. (1990) 'Analysis of a human DNA excision repair gene involved in group A xeroderma pimentosum and containing a zinc-finger domain.', *Nature*, 348, pp. 73–76.

Tanskanen, M., Paetau, A., Salonen, O., Salmi, T., Lamminen, A., Lindsberg, P., Somer, H. and Kiuru-Enari, S. (2009) 'Severe ataxia with neuropathy in hereditary gelsolin amyloidosis.', *Amyloid*, 14(1), pp. 89–95.

Thapa, N., Lee, B. and Kim, I. (2007) 'TGFBIp/βig-h3 protein: a versatile matrix molecule induced by TGF-β.', *The International Journal of Biochemistry & Cell Biology*, 39(12), pp. 2183–2194.

The ENCODE Project Consortium. (2004) 'The ENCODE (ENCyclopedia Of DNA Elements) Project.', *Science*, 306(5696), pp. 636–640.

The ENCODE Project Consortium. (2011) 'A user's guide to the Encyclopedia of DNA elements (ENCODE).', *PLoS Biology*, 9(4), p. e1001046.

The International Human Genome Sequencing Consortium. (2001) 'Initial sequencing and analysis of the human genome.', *Nature*, 409(6822), pp. 860–921.

Thoft, R. and Friend, J. (1983) 'The X, Y, Z hypothesis of corneal epithelial maintenance.', *Investigative Opthalmology & Visual Science*, 24(10), pp. 1442–1443.

Thomsitt, J. and Bron, A. (1975) 'Polymorphic stromal dystrophy.', *The British Journal of Ophthalmology*, 59(3), pp. 125–132.

Tian, X., Fujiki, K., Li, Q., Murakami, A., Xie, P., Kanai, A., Wang, W. and Liu, Z. (2004) 'Compound heterozygous mutations of M1S1 gene in gelatinous droplike corneal dystrophy.', *American Journal of Ophthalmology*, 137(3), pp. 567–569.

Tian, X., Fujiki, K., Wang, W., Murakami, A., Xie, P., Kanai, A. and Liu, Z. (2005) 'Novel mutation (V505D) of the TGFBI gene found in a Chinese family with lattice corneal dystrophy, type I.', *Japanese Journal of Ophthalmology*, 49(2), pp. 84–88.

Tian, X., Fujiki, K., Zhang, Y., Murakami, A., Li, Q., Kanai, A., Wang, W., Hao, Y. and Ma, Z. (2007) 'A novel variant lattice corneal dystrophy caused by association of mutation (V625D) in TGFBI gene.', *American Journal of Ophthalmology*, 144(3), pp. 473–475.

Tiede, S., Storch, S., Lubke, T., Henrissat, B., Bargal, R., Raas-Rothschild, A. and Braulke, T. (2005) 'Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase.', *Nature Medicine*, 11(10), pp. 1109–1112.

Toma, N., Ebenezer, N., Inglehearn, C., Plant, C., Ficker, L. and Bhattacharya, S. (1995) 'Linkage of congenital hereditary endothelial dystrophy to chromosome 20.', *Human Molecular Genetics*, 4(12), pp. 2395–2398.

Torricelli, A., Singh, V., Santhiago, M. and Wilson, S. (2013) 'The corneal epithelial basement membrane: structure, function, and disease.', *Investigative Ophthalmology and Visual Science*, 54(9), pp. 6390–6400.

Town, M., Jean, G., Cherqui, S., Attard, M., Forestier, L., Whitmore, S., Callen, D., Gribouval, O., Broyer, M., Bates, G., Hoff, W. and Antignac, C. (1998) 'A novel

gene encoding an integral membrane protein is mutated in nephropathic cystinosis.', *Nature Genetics*, 18, pp. 319–324.

Tsuji, S., Choudary, P., Martin, B., Stubblefield, B., Mayor, J., Barranger, J. and Ginns, E. (1987) 'A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease.', *The New England Journal of Medicine*, 316(10), pp. 570–575.

Tsujikawa, M., Kurahashi, H., Tanaka, T., Nishida, K., Shimomura, Y., Tano, Y. and Nakamura, Y. (1999) 'Identification of the gene responsible for gelatinous drop-like corneal dystrophy.', *Nature Genetics*, 21(4), pp. 420–423.

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. and Tano, Y. (1998) 'Homozygosity mapping of a gene responsible for gelatinous drop-like corneal dystrophy to chromosome 1p.', *American Journal of Human Genetics*, 63(4), pp. 1073–1077.

Tsujikawa, M., Tsujikawa, K. and Maeda, N. (2000) 'Rapid detection of M1S1 mutations by the protein truncation test.', *Investigative Ophthalmology & Visual Science*, 41, pp. 2466–2468.

Unal, M., Arslan, O., Atalay, E., Mangan, M. and Bilgin, A. (2013) 'Deep anterior lamellar keratoplasty for the treatment of stromal corneal dystrophies.', *Cornea*, 32(3), pp. 301–305.

Underhaug, J., Koldsø, H., Runager, K., Nielsen, J., Sørensen, C., Kristensen, T., Otzen, D., Karring, H., Malmendal, A., Schiøtt, B., Enghild, J. and Nielsen, N. (2013) 'Mutation in transforming growth factor beta induced protein associated with granular corneal dystrophy type 1 reduces the proteolytic susceptibility through local structural stabilization.', *Biochimica et Biophysica Acta*, 1834(12), pp. 2812–2822.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B., Remm, M. and Rozen, S. (2012) 'Primer3-new capabilities and interfaces', *Nucleic Acids Research*, 40(15), p. e115.

Vannucci, L., Lai, M., Chiuppesi, F., Ceccherini-nelli, L. and Pistello, M. (2013) 'Viral vectors: a look back and ahead on gene transfer technology.', *New Microbial*, 36(1), pp. 1–22.

Varki, R., Sadowski, S., Pfendner, E. and Uitto, J. (2006) 'Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants.', *Journal of Medical Genetics*, 43, pp. 641–652.

Vemuganti, G., Rathi, V. and Murthy, S. (2011) 'Histological landmarks in corneal dystrophy: pathology of corneal dystrophies.', *Developments in Ophthalmology*, 48, pp. 24–50.

Vidal, F., Aberdam, D., Miquel, C., Christiano, A., Pulkkinen, L., Uitto, J., Ortonne, J. and Meneguzzi, G. (1995) 'Integrin beta-4 mutations associated with junctional epidermolysis bullosa with pyrloric atresia', *Nature Genetics*, 10, pp. 229–234.

Vincent, A., Niederer, R., Richards, A., Karolyi, B., Patel, D. and McGhee, C. (2009) 'Phenotypic characterisation and ZEB1 mutational analysis in posterior polymorphous corneal dystrophy in a New Zealand population.', *Molecular Vision*, 15, pp. 2544–2553.

Vithana, E., Morgan, P., Ramprasad, V., Tan, D., Yong, V., Venkataraman, D., Venkatraman, A., Yam, G., Nagasamy, S., Law, R., Rajagopal, R., Pang, C., Kumaramanickevel, G., Casey, J. and Aung, T. (2008) 'SLC4A11 mutations in Fuchs endothelial corneal dystrophy.', *Human Molecular Genetics*, 17(5), pp. 656–666.

Vithana, E., Morgan, P., Sundaresan, P., Ebenezer, N., Tan, D., Mohamed, M., Anand, S., Khine, K., Venkataraman, D., Yong, V., Salto-Tellez, M., Venkatraman, A., Guo, K., Hemadevi, B., Srinivasan, M., Prajna, V., Khine, M., Casey, J., Inglehearn, C. and Aung, T. (2006) 'Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2).', *Nature Genetics*, 38(7), pp. 755–757.

Wakabayashi, K., Gustafson, A., Sidransky, E. and Goldin, E. (2011) 'Mucolipidosis type IV: an update', *Molecular Genetics and Metabolism*, 104(3), pp. 206–213.

Warren, J., Abbott, R., Yoon, M., Crawford, J., Spencer, W. and Margolis, T. (2003) 'A new mutation (Leu569Arg) within Exon 13 of the TGFBI (BIGH3) gene causes lattice corneal dystrophy type I.', *American Journal of Ophthalmology*, 136(5), pp. 872–878.

Warren, J., Aldave, A., Srinivasan, M., Eugene, J., Kumar, A., Cevallos, V., Whitcher, J. and Margolis, T. (2003) 'Novel mutations in the CHST6 gene associated with macular corneal dystrophy in southern India.', *Archives of Ophthalmology*, 121, pp. 1608–1612.

Webb, T. R., Matarin, M., Gardner, J. C., Kelberman, D., Hassan, H., Ang, W., Michaelides, M., Ruddle, J. B., Pennell, C. E., Yazar, S., Khor, C. C., Aung, T., Yogarajah, M., Robson, A. G., Holder, G. E., Cheetham, M. E., Traboulsi, E. I., Moore, A. T., Sowden, J. C., Sisodiya, S. M., MacKey, D. a., Tuft, S. J. and Hardcastle, A. J. (2012) 'X-linked megalocornea caused by mutations in CHRDL1 identifies an essential role for ventroptin in anterior segment development.', *American Journal of Human Genetics*. The American Society of Human Genetics, 90(2), pp. 247–259.

Weedaft, G., Ham, R., Bootsma, D., van der Eb, A. and Hoeijmakerst, J. (1990) 'A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome.', *Cell*, 62, pp. 777–791.

Weisgraber, K., Bersot, T. and Mahley, R. (1980) 'A-I(Milano) apoprotein. Isolation and characterization of a cysteine-containing variant of the A-I apoprotein from human high density lipoproteins.', *Journal of Clinical Investigation*, pp. 901–907.

Weisgraber, K., Rall, S. and Bersot, T. (1983) 'Apolipoprotein A-I (Milano). Detection of normal A-I in affected subjects and evidence for a cysteine for arginine substitution in the variant A-I.', *Journal of Biological Chemistry*, pp. 2508–2513.

Weiss, J. (1996) 'Schnyder crystalline dystrophy sine crystals.', *Ophthalmology*, 103(3), pp. 465–473.

Weiss, J. (2009) 'Schnyder corneal dystrophy.', *Current Opinion in Opthalmology*, 20(4), pp. 292–298.

Weiss, J. and Khemichian, A. (2011) 'Differential diagnosis of Schnyder corneal dystrophy.', in Lisch, W. and Seitz, B. (eds) *Corneal Dystrophies*, pp. 67–96.

Weiss, J., Kruth, H., Kuivaniemi, H., Tromp, G., Karkera, J., Mahurkar, S., Lisch, W., Dupps Jr, W., White, P., Winters, R., Kim, C., Rapuano, C., Sutphin, J., Reidy, J., Hum, F., Lu, D., Ebenezer, N. and Nickerson, M. (2008) 'Genetic analysis of 14 families with Schnyder crystalline corneal dystrophy reveals clues to UBIAD1 protein function.', *American Journal of Medical Genetics*, 283, pp. 271–283.

Weiss, J., Kruth, H., Kuivaniemi, H., Tromp, G., White, P., Winters, R., Lisch, W., Henn, W., Denninger, E., Krause, M., Wasson, P., Ebenezer, N., Mahurkar, S. and Nickerson, M. (2007) 'Mutations in the UBIAD1 gene on chromosome short arm 1, region 36, cause Schnyder crystalline corneal dystrophy.', *Investigative Ophthalmology and Visual Science*, 48(11), pp. 5007–5012.

Weiss, J., Møller, H., Aldave, A., Seitz, B., Bredrup, C., Kivelä, T., Munier, F., Rapuano, C., Nischal, K., Kim, E., Sutphin, J., Busin, M., Labbé, A., Kenyon, K., Kinoshita, S. and Lisch, W. (2015) 'IC3D classification of corneal dystrophies-edition 2.', *Cornea*, 34(2), pp. 117–159.

Weiss, J., Møller, H., Lisch, W., Kinoshita, S., Aldave, A., Belin, M., Kivela, T., Busin, M., Munier, F., Seitz, B., Sutphin, J., Bredrup, C., Mannis, M., Rapuano, C., Rij, G., Kim, E. and Klintworth, G. (2008) 'IC3D classification of corneal dystrophies.', *Developments in Ophthalmology*, 27(2).

Weiss, J., Wiaux, C., Yellore, V., Raber, I., Eagle, R., Mequio, M. and Aldave, A. (2010) 'Newly reported p.Asp240Asn mutation in UBIAD1 suggests central discoid corneal dystrophy is a variant of Schnyder corneal dystrophy.', *Cornea*, 29(7), pp. 777–780.

Wenstrup, R. J., Langland, G. T., Willing, M. C., Souza, V. N. D. and Cole, W. G. (1996) 'A splice-junction mutation in the region of COL5A1 that codes for the carboxyl propeptide of pro-alpha1-(V) chains results in the gravis form of the Ehlers – Danlos syndrome (type I).', *Human Molecular Genetics*, 5(11), pp. 1733–1736.

West-Mays, J. and Dwivedi, D. (2006) 'The keratocyte: corneal stromal cell with variable repair phenotypes.', *The International Journal of Biochemistry & Cell Biology*, 38(10), pp. 1625–1631.

Whikehart, D., Parikh, C. and Vaughn, A. (2005) 'Evidence suggesting the existence of stem cells for the human corneal endothelium.', *Molecular Vision*, 11, pp. 816–824.

Wieben, E., Aleff, R., Tosakulwong, N., Butz, M., Highsmith, W., Edwards, A. and Baratz, K. (2012) 'A common trinucleotide repeat expansion within the transcription factor 4 (TCF4, E2-2) gene predicts Fuchs corneal dystrophy.', *PLoS One*, 7(11), p. e49083.

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. a, Chow, M. and Machaca, K. (2005) 'RNA interference is an antiviral defence mechanism in Caenorhabditis elegans.', *Nature*, 436(7053), pp. 1044–1047.

Wilson, N., Messenger, A., Leachman, S., Toole, E., Lane, E., Mclean, W. and Smith, F. (2010) 'Keratin K6c mutations cause focal palmoplantar keratoderma.', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 130(2), pp. 425–429.

Wilson, S. and Hong, J. (2000) 'Bowman's layer structure and function: critical or dispensable to corneal function? A hypothesis', *Cornea*, 19(4), pp. 417–420.

Wilson, S., Mohan, R., Ambrósio, R., Hong, J. and Lee, J. (2001) 'The corneal

wound healing response: cytokine-mediated interaction of the epithelium, stroma, and inflammatory cells.', *Progress in Retinal and Eye Research*, 20(5), pp. 625–637.

Winchester, L., Yau, C. and Ragoussis, J. (2009) 'Comparing CNV detection methods for SNP arrays.', *Briefings in Functional Genomics and Proteomics*, 8(5), pp. 353–366.

Wright, E., Spencer, H., Daly, S., Manson, F., Zeef, L., Urquhart, J., Zoppi, N., Bonshek, R., Tosounidis, I., Mohan, M., Madden, C., Dodds, A., Chandler, K., Banka, S., Au, L., Clayton-smith, J., Khan, N., Biesecker, L., Wilson, M., Rohrbach, M., Colombi, M., Giunta, C. and Black, G. (2011) 'Mutations in PRDM5 in Brittle Cornea syndrome identify a pathway regulating extracellular matrix development and maintenance.', *The American Journal of Human Genetics*. The American Society of Human Genetics, 88, pp. 767–777.

Wu, J., Du, Y., Mann, M., Funderburgh, J. L. and Wagner, W. (2014) 'Corneal stromal stem cells versus corneal fibroblasts in generating structurally appropriate corneal stromal tissue.', *Experimental Eye Research*, 120, pp. 71–81.

Wulle, K. (1972) 'Electron microscopy of the fetal development of the corneal endothelium and Descemet's membrane of the human eye.', *Investigative Ophthalmology & Visual Science*, 11(11), pp. 897–904.

Xia, K., Wu, L., Xi, X., Liang, D., Zheng, D., Cai, F., Pan, Q., Long, Z., Dai, H., Tang, B., Zhang, Z., Xia, J., Liu, X., Xi, X., Liang, D., Zheng, D., Cai, F., Pan, Q., Long, Z., Dai, H., Hu, Z., Tang, B., Zhang, Z. and Xia, J. (2004) 'Mutation in PITX2 is associated with ring dermoid of the cornea.', *Journal of Medical Genetics*, 41, p. e129.

Yamada, M., Mochizuki, H., Kamata, Y., Nakamura, Y. and Mashima, Y. (1998) 'Quantitative analysis of lipid deposits from Schnyder's corneal dystrophy.', *British Journal of Ophthalmology*, 82(4), pp. 444–447.

Yamamoto, S., Okada, M., Tsujikawa, M., Shimomura, Y., Nishida, K., Inoue, Y., Watanabe, H., Maeda, N., Kurahashi, H., Kinoshita, S., Nakamura, Y. and Tano, Y. (1998) 'A kerato-epithelin (Big-h3) mutation in lattice corneal dystrophy type IIIA.', *Ophthalmology*, 62(3), pp. 719–722.

Yamazoe, K., Yoshida, S., Yasuda, M., Hatou, S., Inagaki, E., Ogawa, Y., Tsubota, K. and Shimmura, S. (2015) 'Development of a transgenic mouse with R124H human TGFBI mutation associated with granular corneal dystrophy type 2.', *PLoS One*, 10(7), p. e0133397.

Yang, Y., Muzny, D., Reid, J., Bainbridge, M., Willis, A., Ward, P., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M., Leduc, M., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., Plon, S., Lupski, J., Beaudet, A., Gibbs, R. and Eng, C. (2013) 'Clinical whole-exome sequencing for the diagnosis of mendelian disorders.', *The New England Journal of Medicine*, 369(16), pp. 1502–1511.

Ye, Y., Yao, Y., Zhou, P. and Pan, F. (2006) 'In vivo confocal microscopy of pre-Descemet's membrane corneal dystrophy.', *Clinical & Experimental Ophthalmology*, 34, pp. 614–616.

Yellore, V., Khan, M., Bourla, N., Rayner, S., Chen, M., Sonmez, B., Momi, R., Sampat, K., Gorin, M. and Aldave, A. (2007) 'Identification of mutations in

UBIAD1 following exclusion of coding mutations in the chromosome 1p36 locus for Schnyder crystalline corneal dystrophy.', *Molecular Vision*, 13, pp. 1777–1782.

Yellore, V., Papp, J., Sobel, E., Khan, M., Rayner, S., Farber, D. and Aldave, A. (2007) 'Replication and refinement of linkage of posterior polymorphous corneal dystrophy to the posterior polymorphous corneal dystrophy 1 locus on chromosome 20.', *Genetics in Medicine*, 9(4), pp. 228–234.

Yellore, V., Rayner, S., Nguyen, C., Gangalum, R., Jing, Z., Bhat, S. and Aldave, A. (2012) 'Analysis of the role of ZEB1 in the pathogenesis of posterior polymorphous corneal dystrophy.', *Investigative Ophthalmology & Visual Science*, 53(1), pp. 273–278.

Yen, P., Allen, E., Marsh, B., Mohandas, T., Wang, N., Taggart, R. and Shapiro, L. (1987) 'Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: Implications for X-Y interchange.', *Cell*, 49(4), pp. 443–454.

Yew, D., Sha, O., Li, W., Lam, T. and Lorke, D. (2001) 'Proliferation and apoptosis in the epithelium of the developing human cornea and conjunctiva', *Life Sciences*, 68(26), pp. 2987–3003.

Yoon, J., Ismail, S. and Sherwin, T. (2014) 'Limbal stem cells: central concepts of corneal epithelial homeostasis.', *World Journal of Stem Cells*, 6(4), pp. 391–403.

Yoon, M., Warren, J., Holsclaw, D., Gritz, D. and Margolis, T. (2004) 'A novel arginine substitution mutation in 1A domain and a novel 27 bp insertion mutation in 2B domain of keratin 12 gene associated with Meesmann's corneal dystrophy.', *British Journal of Ophthalmology*, 88(6), pp. 752–756.

Yoshida, K., Oshima, A., Shimmoto, M., Fukuhara, Y., Sakuraba, H., Yanagisawa, N. and Suzuki, Y. (1991) 'Human beta-galactosidase gene mutations in GM1-gangliosidosis: a common mutation among Japanese adult/chronic cases.', *American Journal of Human Genetics*, 49, pp. 435–442.

Young, S., Bertics, S., Curtiss, L., Dubois, B. and Witztum, J. (1987) 'Genetic analysis of a kindred with familial hypobetalipoproteinemia.', *Journal of Clinical Investigation*, 79, pp. 1842–1851.

Yuan, C., Berscheit, H. and Huang, A. (2007) 'Identification of an amyloidogenic region on keratoepithelin via synthetic peptides.', *FEBS Letters*, 581(2), pp. 241–247.

Zacharias, A. and Gage, P. (2010) 'Canonical Wnt/β-catenin signaling is required for maintenance but not activation of Pitx2 expression in neural crest during eye development.', *Developmental Dynamics*, 239(12), pp. 3215–3225.

Zavala, J., López Jaime, G., Rodríguez Barrientos, C. and Valdez-Garcia, J. (2013) 'Corneal endothelium: developmental strategies for regeneration.', *Eye*, 27(5), pp. 579–588.

Zenteno, J., Correa-Gomez, V., Santacruz-Valdez, C., Suarez-Sanchez, R. and Villanueva-Mendoza, C. (2009) 'Clinical and genetic features of TGFBI-linked corneal dystrophies in Mexican population: description of novel mutations and novel genotype-phenotype correlations.', *Experimental Eye Research*, 89(2), pp. 172–177.

Zhang, Y., Overbeek, P. and Govindarajan, V. (2007) 'Perinatal ablation of the

mouse lens causes multiple anterior chamber defects.', *Molecular Vision*, 13, pp. 2289–2300.

Zhao, H., Li, H., Bacht, G., Schmidtchen, A. and Neufeld, E. (1996) 'The molecular basis of Sanfilippo syndrome type B.', *Proceedings of the National Academy of Sciences of the United States of America*, 93, pp. 6101–6105.

Zhao, M., Wang, Q., Wang, Q., Jia, P. and Zhao, Z. (2013) 'Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives.', *BMC Bioinformatics*, 14(11), p. S1.

Appendix A – List of primers used in the study for PCR, RT-PCR and qPCR.

Target	Forward primer	Reverse Primer	Size (bp)	Function	Enzyme	Annealing Temp
TGFBI exon 1	TTGGTTTGAGGAAGACTGTGG	CACTTCCCCACTCCCTTCTA	689	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 2	AACAATTTTTGCAGGGGAAG	TGGCCAGCTTCCTAAAAATG	489	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 3	GCTCAGGAAAGGCAGACCTA	CTCAGCCCCTCGCATAGA	399	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 4&5	GTCAGAGAAGGGAGGGTGTG	AGCTTAAACCCCAGAAACCA	843	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 6	GCTTGTGGAACCCACATTTT	AAGGGCATTCAGGGGAAC	400	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 7	TGGGTTTGGCTTCTGTTTTC	AGCAACAGGACAGGATGACC	434	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 8	CCCAGCTGGTTTTCTAGGG	GTCACAACCCACACATTTGC	800	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 9	AGAGGGGTTGTTGACTCACG	AAGTCACATCTGCTTTTGGTTG	400	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 10	CTCCATAGAAGATACCAGATGTTAAGG	TCCTTCTGCAGGATCTCATTT	487	PCR and Sanger sequencing	GoTaqGreen	55
TGFBI exon 11	CCATCCCAGTGTATACTCCTTCA	TCTGGAAGGTCGCAGCAGT	400	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 12	AGCCTGGAATCACTCCCTCT	AGGGTTGCTAGTCCCTGGTT	499	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 13	TGCTTTGTGTCCTCTGACCA	TGATTTCCCTGAAGACCCTCT	500	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 14	GGCGACAAGATTGAAACTCC	GCAATCAGTCACATAGGCACA	392	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 15	CCATGGGCCAAGTTCTACC	AGGATGCCTCAGTGGGAGT	386	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 16	AGGAACTCCCTGGTGCCTAT	AGAAGGAAATGGGGTGCTCT	363	PCR and Sanger sequencing	GoTaqGreen	60
TGFBI exon 17	TCACAAACCACAAAGCACCT	ACATCTCTCTCCCTCCTCCC	357	PCR and Sanger sequencing	GoTaqGreen	60
COL17A1 exon 16	CGTAGTCACTCAGGCCAGAA	TGGGCCCATTCACAGATTCC	390	PCR and Sanger sequencing	GoTaqGreen	60
LAMC2 exon 14	ACCATAAGCCAGTCAACCCT	AGTAGGTCTCCCAACACAGC	349	PCR and Sanger sequencing	GoTaqGreen	60

LAMC2 exon 15	TGGAACCCTGTGTGGCTAAT	GCCCGAGGAATCACATCTGA	485	PCR and Sanger sequencing	GoTaqGreen	60
TACSTD2	CCTGCAGACCATCCCAGAC	CAGGAAGCGTGACTCACTTG	1140	PCR and Sanger sequencing	KAPA Robust	60
CHST6 exon 1A	GCCCCTAACCGCTGCGCTCTC	GGCTTGCACACGGCCTCGCT	489	PCR and Sanger sequencing	GoTaqGreen	60
CHST6 exon 1B	GACATGGACGTGTTTGATGC	GAGACTGAGCCCAGTGAAGG	579	PCR and Sanger sequencing	GoTaqGreen	60
CHST6 exon 1C	CTCCCGGGAGCAGACAGCCAA	CTCCCGGGCCTAGCGCCT	599	PCR and Sanger sequencing	GoTaqGreen	60
CHST6 F1/R1	CCACAGAAGGAAGGACAGAGTAAATGAA	TTCCCTTTACTATTATAAAAATGCTGCTAATG	956	PCR based assay	GoTaqGreen	55
CHST6 F2/R2	CATATCCTGTCTGGCCTAAACCTTAGTTTAC	GGGCACAGACAGAGGGAAAAACC	966	PCR based assay	GoTaqGreen	55
CHST6 F2M/R2	GGCCAAGTTCAGGTCAGCTTCCA	GGGCACAGACAGAGGGAAAAACC	N/A	PCR based assay	GoTaqGreen	55
CHST6 F1/R1M	CCACAGAAGGAAGGACAGAGTAAATGAA	TGCTGAATGGCTAACTGAAGGAATACTATAC	N/A	PCR based assay	GoTaqGreen	55
UBIAD1 exon 1	CCGTCCTTCCTCCTTCCC	AAGCCACCTTTGACATCCCT	700	PCR and Sanger sequencing	GoTaqGreen	65
UBIAD1 exon 2	CCACCTGCACAGTCTAAGGA	CTGCCAAATCACATTCCTTCCT	689	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 1	GTAAAGCCGGGAGTGTCGTA	GCGGAGAGAGACCAGGTAAG	422	PCR and Sanger sequencing	KAPA Robust	57
ZEB1 exon 2	CATTGAATTACAATCTGTTTTAAGCAT	TTGATTTCAAACTTTTCATCCA	405	PCR and Sanger sequencing	GoTaqGreen	57
ZEB1 exon 3	CCTTTTCAGATTTCGGGAAG	TGTAGTGCACCATATGCCTAAGT	452	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 4	TTCTGCAGATTCAAGAACAATCA	TGCATGGTCATCATAGTGTTCC	456	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 5	GTGGGTAGCACAATATCTGG	AGGCTGCAGATATAGCACTG	435	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 6	CAACCATCAGGCTCACAAAA	TGATTAGGAAATCTGGAGTATGTCA	433	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7A	CCGCTTGTTTTAGGGAAATG	CCAGTGAAAACCCCATTTTG	423	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7B	CCCCTTCAAGAACAACTTTCTG	AGTTGGCTAGGCTGCTCAAG	449	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7C	AGCCATCAGTCTTCCTTTGG	TCTTCTGCACTTGGTTGTGC	456	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7D	GGAGATGGCAATTTGTCTCC	CTCTTCTTGTGCACCCTCAG	456	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7E	GATCAACCACCAATGGTTCC	ATTGGCAGCTAGCGCTCTTA	422	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 7F	CAAGTGCCAACCCCATAAAT	CCCTGGATTAGCAAACAACC	408	PCR and Sanger sequencing	GoTaqGreen	60

ZEB1 exon 8	TCAGTGTGCTTGCTTTGGTC	GAGCCAGACCTTGTCTCAAAA	450	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 9A	GAGTTTGGGACCTGGAAATG	TCATCCTCTTCCCTTGTCAAA	457	PCR and Sanger sequencing	GoTaqGreen	60
ZEB1 exon 9B	GCCTGAAATCCTCTCGAATG	TGTTTCCATGAAAAGCAAGG	447	PCR and Sanger sequencing	GoTaqGreen	60
rs10826649	AGTCCCTCGTCACTGCATAC	CACTTATTCCAGCATCGTTTGT	913	PCR and Sanger sequencing	GoTaqGreen	60
rs10826650	TGCAATTTCCAGTAACCAGAGG	TGGGATCACAGGTGCACAC	215	PCR and Sanger sequencing	GoTaqGreen	60
rs11007625	CATCCCTGGACGAACAACATC	CGGCCTCTGTGCTAATGAAC	399	PCR and Sanger sequencing	GoTaqGreen	60
rs12268212	CTAGGGGCTGAAAGGTCTCC	CTGTTCAGGACCAGGCATG	600	PCR and Sanger sequencing	GoTaqGreen	60
rs16930343	AGAGTTCACCGTCTGCAGAA	GAAATGATGCACCGGTGACA	992	PCR and Sanger sequencing	GoTaqGreen	60
rs2368394	ATTCCTCTGCACCCTACGTC	CTCACCCAGCAAAACCATGT	667	PCR and Sanger sequencing	GoTaqGreen	60
rs2479356	GTGGAATCCCTAGGTCATATGG	GGAGGCAGAGGTAGGTTGT	903	PCR and Sanger sequencing	GoTaqGreen	60
rs2806139	TATCCAGCCCCTACTCCAGA	TAAAGGGGCAGGACAGACAC	840	PCR and Sanger sequencing	GoTaqGreen	60
rs2806144	CAGCGAGCCAAGATTGCAC	TCAGATAGTGATTGGCGGTCT	404	PCR and Sanger sequencing	GoTaqGreen	60
rs6481606	TGCAGCTGTGAGTTCTGAGG	GTTCGTCCAGGGATGCATTG	594	PCR and Sanger sequencing	GoTaqGreen	60
Deletion in F131	ACGACACAAAGCCTCCTGTA	ACTTCAGGGTATTTGGCCTGA	N/A	PCR and Sanger sequencing	GoTaqLong	66
Deletion in F132	AAGAAATCCTGGGGCCTGAA	TGGGTGACGGATTTTGCAAA	N/A	PCR and Sanger sequencing	GoTaqLong	68
ZEB1 5' (intron 1)	ACACGTACATTTCGGACCGA	GGGGAGCACGAGAGAGTTTT	157	qPCR	labTAQ™ Green Hi Rox	65
ZEB1 3' (intron 8)	AAGGTCTGGCTCTAGTCACC	CAGGAGGCTGAGATGAAAGGA	112	qPCR	labTAQ™ Green Hi Rox	65
ZEB1 3' re-designed	GCTCTCCCTGAACCGTTATG	AAAAGTGCGGAAAGAAGCAA	190	qPCR	labTAQ™ Green Hi Rox	65
TWIST 3' UTR	ACTGGCCTGCAAAACCATAG	TGCATTTTACCATGGGTCCT	128	qPCR	labTAQ™ Green Hi Rox	65

OVOL2 P1	CTTCTGACCGCCCAGATT	AGGAAGACTTTGGGCATGGT	689	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2 P2	CGCAACGTGAGCACAACT	CCGGTGGATTTGTCTTCGAC	676	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2 P3	TCCATTCACCCAAGCTCCTT	TCCTATCAACCTTTATGCCTGGA	769	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2 exon 1	CGAGCTTGTTGACACCGTTA	CCTAACCCGCCTAGAAGACC	575	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2 exon 2	CTAGCTCTGTGGGCCGATT	CAGTGATGCCTGTGCCTTT	689	PCR and Sanger sequencing	KAPA Robust	65
OVOL2 exon 3	AGCCAAAACCTGCTTGAAAA	ACCTCGTGATCCGCATCTT	470	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2 exon 4	GTCCTCATGCTGCCATGAAT	CACAGGTTACAGGGCCTGAC	630	PCR and Sanger sequencing	GoTaqGreen	60
OVOL2	CTCGCGATTTAAGGCATAGG	ACAGCTGTGAACCACCGAGT	587	RT-PCR	GoTaqGreen	60
GRHL2 5'UTR & exon 1	AACCAGTAGCCTCCACCTTC	AGTATTCATCCAGGAGCGGG	930	PCR and Sanger sequencing	GoTaqGreen	63
GRHL2 intron 1	TGGATCAAACATGTCACAAGAGT	GGAGGAAGGCGGGATGTC	669	PCR and Sanger sequencing	GoTaqGreen	60
GSN exon 4	TGAGAATCACTGTGTTGCTGG	TGCATTTGTTCACGTGGATGT	384	PCR and Sanger sequencing	GoTaqGreen	60
GSN exon 12	ACAGACACAAGCATGATGGC	TACCCCGTCATTCACCATCC	392	PCR and Sanger sequencing	GoTaqGreen	60
STS exon 2	CCAGGACTTCAAGGTTGCAG	GGCATGGTTTGGAGGGTTTT	387	PCR based assay	GoTaqGreen	60
STS exon 9	CAGGTGGGGTCAGTCTTCTT	AGGACAAAGCAAAGCATGGG	390	PCR based assay	GoTaqGreen	60
APOA1 exon 4	AGGGGTGTTGGTTGAGAGTG	GGCAGACTCTACTCCCTCAC	998	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 1	GGATCCAATTTTGAGTGGAGA	AGTACAGCTAAATTGGAAAAT	804	PCR and Sanger sequencing	GoTaqGreen	57
KRT12 exon 2	TGCAGCAATTTTCTCATCCCA	GCACACTACAAATATCGTGATCC	590	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 3	AGGTGCACAGATTTAAGCAAGA	GGCCATTTTAACAGGGAGACA	385	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 4&5	ACAAAAGTATCCAGGGCCCA	ACCGAAAAGAGGAGGGTAGC	549	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 6	CGCAAACAGACGTAGCATCC	GAGCATTTAGTAAATCCCAGGCA	498	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 7	CAGGCCTACAAATTGGGAAATCT	AGCTATGAGGTTACAGGCATGA	400	PCR and Sanger sequencing	GoTaqGreen	60
KRT12 exon 8	TCATTGTCATTGGCTAAAACACA	AGACCAACATGACCAAAATGACT	473	PCR and Sanger sequencing	GoTaqGreen	60

KRT3 exon 1	GTGTTTGGTAGCCTGGAACA	CCTAACTCCTCACCCTCCAG	975	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 2	GCTGCAAGTGTAGTGTCCTG	TAAGGAGAGGGAAATGGGCC	538	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 3&4	TGGAGGGAGGGAAGAGATGA	GGGGCCTATATACCAAAATGCA	698	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 5&6	ACATGGGGTTCTCTGAGTGG	CCAGACGTCTATTCCAGGGG	958	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 7	AATCCATTGCATGTCAGGAAGGGC	TATCTGGCCCTTGGCCTATGACTT	345	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 8	CTCTGTTCAGCTCTGGTTGC	CTCCCAGAAAACCCCTCCAA	251	PCR and Sanger sequencing	GoTaqGreen	60
KRT3 exon 9	TCAGCTCTGGATAGTCGTGG	CTGGAAGGAGGAGCAAGAGG	563	PCR and Sanger sequencing	GoTaqGreen	60
MCOLN1 exons 5&6	CACAGGGAAGAAGGCCACT	CCAAGACACTCCCTGACCAG	598	PCR and Sanger sequencing	KAPA Robust	60
MCOLN1 exon 12	CTAGCCTGGAACCCGACC	TGGGTCTTCTCTGTTCACCC	485	PCR and Sanger sequencing	GoTaqGreen	60

Appendix B – Manually curated list of syndromes with corneal involvement including corneal opacity, developmental corneal abnormalities and loss of transparency, derived from OMIM database and literature searches.

Gene	Syndrome	Inher.	Ref.
ABCA1	Tangier disease	AR	(Bodzioch *et al.*, 1999)
ADAMTSL4	Ectopia lentis	AR	(Ahram *et al.*, 2009)
APOB	Hypercholesterolemia Hypobetalipoproteinemia	AD AR	(Pullinger *et al.*, 1995) (Young *et al.*, 1987)
ATP7B	Wilson disease	AR	(Bull *et al.*, 1993)
B3GALT6	Ehlers-Danlos syndrome	AR	(Nakajima *et al.*, 2013)
B3GALTL	Peter's plus syndrome	AR	(Oberstein *et al.*, 2006)
CHRDL1	Megalocornea	XLR	(Webb *et al.*, 2012)
CHST14	Ehlers-Danlos syndrome	AR	(Mu *et al.*, 2009)
COL17A1	Epidermolysis bullosa	AR	(McGrath, Gatalica, *et al.*, 1995)
COL1A1	Ehlers-Danlos syndrome	AD	(Wenstrup *et al.*, 1996)
COL3A1	Ehlers-Danlos syndrome	AD	(Smith *et al.*, 1997)
COL4A1	Brain small vessel disease with or without ocular anomalies	AD	(Gould *et al.*, 2005)
COL4A3	Alport syndrome	AR	(Mochizuki *et al.*, 1994)
COL4A4	Alport syndrome	AR	(Mochizuki *et al.*, 1994)
COL4A5	Alport syndrome	XLD	(Barker *et al.*, 1990)
COL5A1	Ehlers-Danlos syndrome	AD	(Burrows *et al.*, 1996)
COL5A2	Ehlers-Danlos syndrome	AD	(Michalickova *et al.*, 1998)
COL7A1	Epidermolysis bullosa	AD/AR	(Christiano *et al.*, 1993)
CTNS	Cystinosis	AR	(Town *et al.*, 1998)
CYP4V2	Bietti crystalline corneoretinal dystrophy	AR	(Li *et al.*, 2004)
DDB2	Xeroderma pigmentosum	AR	(Itoh *et al.*, 1999)
DNAI1	Ciliary dyskinesia	AR	(Escudier *et al.*, 1999)
DST	Epidermolysis bullosa simplex	AR	(Groves *et al.*, 2010)
ERCC1	Cerebrooculofacioskeletal syndrome	AR	(Jaspers *et al.*, 2007)
ERCC2	Xeroderma pigmentosum	AR	(Frederick *et al.*, 1994)
ERCC3	Xeroderma pigmentosum	AR	(Weedaft *et al.*, 1990)
ERCC5	Xeroderma pigmentosum	AR	(Nouspikel and Ciarkson, 1994)
ERCC6	Xeroderma pigmentosum	AR	(Mallery *et al.*, 1998)
FAH	Tyrosinemia I	AR	(Hahn *et al.*, 1995)
FERMT1	Kindler syndrome	AR	(Hadj-rabia *et al.*, 2003)

FGF10	Lacrimoauriculodentodigital syndrome	AD	(Jm *et al.*, 2006)
FGFR2	Lacrimoauriculodentodigital syndrome	AD	(Rohmann *et al.*, 2006)
FGFR3	Lacrimoauriculodentodigital syndrome	AD	(Rohmann *et al.*, 2006)
FKRP	Muscular dystrophy-dystroglycanopathy	AR	(Brockington *et al.*, 2001)
FOXC2	Lymphedema-distichiasis syndrome	AD	(Fang *et al.*, 2000)
FOXE3	Anterior segment dysgenesis	AR	(Semina *et al.*, 2001)
FREM1	Manitoba oculotrichoanal syndrome	AR	(Alazami *et al.*, 2009)
GBA	Gaucher's disease	AR	(Tsuji *et al.*, 1987)
GJB2	Keratitis-ichthyosis-deafness syndrome	AD	(Richard *et al.*, 2002)
GLA	Fabry disease	XL	(Bernstein *et al.*, 1989)
GLB1	GM1-gangliosidosis Mucopolysaccharidosis type IVB	AR	(Yoshida *et al.*, 1991)
GNPTAB	Mucolipidosis II/III	AR	(Tiede *et al.*, 2005)
GSN	Meretoja syndrome	AD	(Maury *et al.*, 1990)
GTF2H5	Xeroderma pigmentosum	AR	(Giglia-Mari *et al.*, 2004)
HEXA	Tay-Sachs disease	AR	(Beutler, Kuhl and Comings, 1975)
HPD	Tyrosinemia III	AR	(Rüetschi *et al.*, 2000)
HRAS	Schimmelpenning-Feuerstein-Mims syndrome	Somatic mosiac	(Groesser *et al.*, 2012)
IDUA	Mucopolysaccharidosis Ih/s	AR	(Scott *et al.*, 1992)
ITGA6	Epidermolysis bullosa	AR	(Ruzzi *et al.*, 1997)
ITGB4	Epidermolysis bullosa	AD/AR	(Vidal *et al.*, 1995)
K6A	Pachyonychia congenita	AD	(Bowden *et al.*, 1995)
K6B	Pachyonychia congenita	AD	(Smith *et al.*, 1998)
K6C	Pachyonychia congenita	AD	(Wilson *et al.*, 2010)
KERA	Cornea plana	AR	(Pellegata *et al.*, 2000)
KRAS	Schimmelpenning-Feuerstein-Mims syndrome	Somatic mosiac	(Groesser *et al.*, 2012)
KRT14	Epidermolysis bullosa	AD	(Bonifas, Rothman and Epstein, 1991)
KRT16	Pachyonychia congenita	AD	(McLean *et al.*, 1995)
KRT17	Pachyonychia congenita	AD	(McLean *et al.*, 1995)
KRT5	Epidermolysis bullosa	AD	(Lane *et al.*, 1992)
LAMA3	Epidermolysis bullosa	AR	(McGrath, Kivirikko, *et al.*, 1995)
LAMB3	Epidermolysis bullosa	AR	(Pulkkinen *et al.*, 1994)
LAMC2	Epidermolysis bullosa	AR	(Aberdam *et al.*, 1994)

LCAT	Fish-eye disease	AR	(Funke *et al.*, 1991)
LIFR	Stuve-Wiedemann syndrome	AR	(Dagoneau *et al.*, 2004)
MBTPS2	IFAP (ichthyosis follicularis with atrichia and photophobia) syndrome	XLR	(Oeffner *et al.*, 2009)
MCOLN1	Mucolipidosis IV	AR	(Bargal *et al.*, 2000)
MIR184	EDICT syndrome	AD	(Iliff, Riazuddin and Gottsch, 2012)
NAGLU	Mucopolysaccharidosis IIIB	AR	(Zhao *et al.*, 1996)
NDP	Norrie disease	XLR	(Berger *et al.*, 1992)
NRAS	Schimmelpenning-Feuerstein-Mims syndrome	Somatic mosiac	(Lim *et al.*, 2014)
NTRK1	Congenital insensitivity to pain with anhidrosis	AR	(Mardy *et al.*, 1999)
PAX6	Aniridia	AD	(Jordan *et al.*, 1992)
PEX5	Peroxisome biogenesis disorder 2A	AR	(Dodt *et al.*, 1995)
PITX2	Ring dermoid of cornea	AD	(Xia *et al.*, 2004)
PITX3	Anterior segment dysgenesis	AD	(Semina *et al.*, 1998)
PLCG2	Autoinflammation, antibody deficiency, and immune dysregulation syndrome	AD	(Freeman *et al.*, 2012)
PLEC	Epidermolysis bullosa	AD/AR	(Smith *et al.*, 1996)
PLOD1	Ehlers-Danlos syndrome	AR	(Hautala *et al.*, 1993)
POLH	Xeroderma pigmentosum	AR	(Masutani *et al.*, 1999)
POMT1	Muscular dystrophy-dystroglycanopathy	AR	(Beltra *et al.*, 2002)
PRDM5	Brittle cornea syndrome	AR	(Wright *et al.*, 2011)
PXDN	Anterior segment dysgenesis with sclerocornea	AR	(Khan *et al.*, 2011)
RAB23	Carpenter syndrome	AR	(Jenkins *et al.*, 2007)
RAB3GAP1	Warburg micro syndrome	AR	(Aligianis *et al.*, 2005)
SAT1	Keratosis follicularis spinulosa decalvans	XL	(Gimelli *et al.*, 2002)
SLC39A13	Ehlers-Danlos syndrome	AR	(Giunta *et al.*, 2008)
SMARCAL1	Schimke immune-osseous dysplasia	AR	(Boerkoel *et al.*, 2002)
SPINT2	Syndromic congenital sodium diarrhoea	AR	(Heinz-Erian *et al.*, 2008)
ST14	Ichthyosis with hypertrichosis	AR	(Basel-Vanagaite *et al.*, 2007)
STS	X-linked ichthyosis	XLR	(Basler *et al.*, 1992)
SUMF1	Multiple sulfatase deficiency	AR	(Dierks *et al.*, 2003)
TAT	Tyrosinemia II	AR	(Natt *et al.*, 1992)
TNXB	Ehlers-Danlos syndrome	AR	(Burch *et al.*, 1997)

TRPV3	Olmsted syndrome	AD	(Lin *et al.*, 2012)
UVSSA	Xeroderma pigmentosum	AR	(Nakazawa *et al.*, 2012)
XPA	Xeroderma pigmentosum	AR	(Tanaka *et al.*, 1990)
XPC	Xeroderma pigmentosum	AR	(Li *et al.*, 1993)
ZNF469	Brittle cornea syndrome	AR	(Abu *et al.*, 2008)

Appendix C – List of genes encoding proteins present in the human corneal proteome. Proteins were downloaded from Dyrlund *et al.*, (2012) in UNIPROT format and converted to gene name using Biomart (Ensembl). Several genes responsible for anterior corneal dystrophies are found in the human corneal proteome; however, some known corneal dystrophy genes (primarily those causing endothelial dystrophies) are absent (e.g. *ZEB1*).

FER1L5	*BTBD19*	*DNM1L*	*GOSR2*	*GIPC1*	*SLC25A12*	*CEP152*	*CGB5*	*PCBD2*
UBA6	*C7orf73*	*RTCA*	*IRS4*	*LIN7A*	*EIF3G*	*YIF1A*	*CGB8*	*HDHD2*
ESYT2	*TK2*	*AGRN*	*ADAM10*	*TFEC*	*EIF3J*	*UBR5*	*CGB7*	*MAGT1*
KIAA1598	*DDX39A*	*EXOC5*	*PTGES*	*UQCRQ*	*CBR3*	*NDUFB6*	*CGB*	*RAB1B*
DAPL1	*PDLIM1*	*BTN3A1*	*KMT2D*	*LECT2*	*PSMD10*	*DUSP14*	*IGJ*	*MRPL18*
ILVBL	*MYO1C*	*NDUFA4*	*CDIPT*	*RAB29*	*UPK1B*	*NDUFA3*	*PIGR*	*C11orf54*
XIRP2	*SNAP23*	*PSMD14*	*PDCD5*	*DSCR3*	*AP1G2*	*NDUFB4*	*IGHG1*	*SDCBP2*
CNOT1	*AIP*	*KPNA3*	*SLC9A3R1*	*HNRNPDL*	*IDH1*	*NDUFB8*	*IGHG2*	*UNC93B1*
CCDC88B	*GTPBP1*	*LAD1*	*CHEK1*	*XPO1*	*ATRN*	*SCEL*	*IGHG4*	*NUCKS1*
GRXCR2	*RAB27B*	*VWA5A*	*ZNF213*	*SPTBN2*	*RBBP9*	*NDUFA7*	*IGHA1*	*TMX4*
TCAF2	*AP3B1*	*NOP56*	*TPP1*	*U2SURP*	*ALDH1L1*	*RTN3*	*HBZ*	*KIF13A*
C2orf71	*NDUFS8*	*DDX3X*	*APOL1*	*ANKRD28*	*ARL6IP5*	*LYPD3*	*MB*	*C2orf40*
DDTL	*PSMD11*	*RNASET2*	*PSMA7*	*FYB*	*BCAS2*	*VAPB*	*COL1A1*	*CDT1*
TUBAL3	*PSMD12*	*PIR*	*MGST3*	*NPC1*	*DCTN3*	*RASAL1*	*COL2A1*	*VPS16*
SMCHD1	*PSMD9*	*KPNA4*	*TAX1BP3*	*AGPAT2*	*BBOX1*	*SNAPIN*	*COL3A1*	*SH3BGRL3*
WTIP	*ATOX1*	*PPP6C*	*GMDS*	*SCAMP2*	*ATP5H*	*NDUFC2*	*COL4A1*	*UBA3*
ZNF839	*PGRMC1*	*MAN2B1*	*TLR5*	*ARPC1B*	*FLOT1*	*NDUFA10*	*CRYAA*	*BAHD1*
CEACAM18	*DFFA*	*ACACB*	*DIAPH1*	*FLNB*	*TRIO*	*FKBP9*	*CRYAB*	*FAM81A*
FAM171A2	*HIP1*	*PDXK*	*EXOC3*	*NDUFS6*	*ATP5L*	*PGLS*	*KRT14*	*TMEM167A*
COL6A5	*CLIC1*	*CLDN4*	*PLIN3*	*CS*	*GLRX3*	*ATG7*	*KRT6A*	*MDM1*
UPK3BL	*EIF3F*	*PPAP2B*	*KPNA6*	*MRPL33*	*KRT36*	*TRIM16*	*LMNA*	*TBC1D15*
FHAD1	*WWP2*	*ISLR*	*SRPX2*	*SEC22B*	*NEBL*	*LYPLA2*	*APOA1*	*RDH11*
IGLL5	*MATN2*	*COX7A2L*	*UGDH*	*PRPF40A*	*SNCG*	*IPO7*	*APOE*	*NT5C*
ARPC2	*IPO5*	*GAPDHS*	*CTNND1*	*VPS26A*	*CIAO1*	*PGM3*	*APOA2*	*FAM45A*
ARPC3	*EEF2K*	*NDUFAB1*	*EIF1B*	*NDUFB1*	*SRP72*	*PAPLN*	*APOC1*	*HM13*
POLR1C	*PSMD3*	*COPE*	*SNX2*	*CRLF1*	*DDAH1*	*AHSA1*	*APOC2*	*PREX1*

PGRMC2	*PAPSS1*	*AP3D1*	*DPM1*	*ERLIN1*	*ALDH1A2*	*CLDN7*	*APOC3*	*BICD2*
PFDN6	*SPINT1*	*CCS*	*USO1*	*NDUFS3*	*TOMM70A*	*ABCA1*	*FGA*	*ZNF398*
LAMA5	*CCP110*	*NARS*	*TOM1*	*SRSF10*	*MYO1D*	*SEC24B*	*FGB*	*DNAH3*
ZNF185	*EEF1E1*	*MYO1B*	*HNRNPCL1*	*RBMXL2*	*UFL1*	*MPC2*	*FGG*	*MGARP*
ACOX3	*HNRNPR*	*NUDT21*	*CCDC22*	*BANF1*	*PROSC*	*ETHE1*	*CRP*	*CDK5R2*
SURF4	*TXNL1*	*LANCL1*	*PRAF2*	*SF3B1*	*ERLIN2*	*ACSL3*	*APCS*	*EIF3I*
MAPK13	*TPD52L2*	*STRN*	*DKC1*	*CSDE1*	*PRPF6*	*NTN1*	*C1QA*	*UBE2V1*
SPTLC1	*PPIH*	*ANGPTL7*	*EIF5B*	*SCGB2A1*	*PCF11*	*ASMTL*	*C1QB*	*DYNC1I2*
PMM2	*AKR7A2*	*CALU*	*EDF1*	*PRKRA*	*FRYL*	*KRT75*	*C1QC*	*MSLN*
CHAD	*TGOLN2*	*AHCYL1*	*DNAJA2*	*MYCBP2*	*AP2A2*	*CXCL14*	*C9*	*NNT*
HMGB3	*DENR*	*CD5L*	*CUTA*	*LYPLA1*	*SEC31A*	*RAB3D*	*APOH*	*PPIG*
PPM1G	*XPOT*	*KIF1C*	*CTSV*	*SNRNP200*	*LSM8*	*SNAP29*	*LRG1*	*TCOF1*
EIF3D	*DNPH1*	*NDUFS5*	*PFDN1*	*TIPRL*	*AP2A1*	*OXSR1*	*FN1*	*SF3B2*
EIF3H	*TIMM44*	*PDE6D*	*NOL3*	*UTP20*	*STAU1*	*HSPA4L*	*RBP4*	*PDAP1*
STX7	*TRAPPC3*	*PEX1*	*KERA*	*NUP155*	*CER1*	*BCL10*	*AMBP*	*FKBP5*
CAPN5	*CHMP2A*	*AKR1B10*	*KIF21B*	*DHFR*	*BAG3*	*NDUFB10*	*ORM1*	*PICALM*
YKT6	*PSCA*	*TIMM8A*	*PLXNA2*	*CYB5R3*	*AIFM1*	*MOCS2*	*AHSG*	*TUBB3*
ARPC5	*ZNF207*	*TMPRSS11D*	*WDR1*	*GSR*	*CLDN1*	*TOMM40*	*TTR*	*ASAH1*
POLR2D	*NDUFB5*	*SOX15*	*CPNE3*	*MT-CO1*	*CLIC3*	*PEX11B*	*ALB*	*PIN1*
FGF10	*NDUFB3*	*PRPSAP2*	*CLUH*	*MT-CO2*	*EML2*	*ACTL6A*	*GC*	*VAMP7*
FABP7	*NDUFA2*	*SMARCA5*	*DNAJC13*	*SOD1*	*NUDT14*	*MOCS2*	*TFRC*	*RPS6KA3*
CLDN3	*ASNA1*	*MAST3*	*DNAJB6*	*CP*	*BPNT1*	*MT-CYB*	*TF*	*PRKX*
DHX15	*BUB3*	*ATP10A*	*GGCT*	*F8*	*DDAH2*	*CYB5A*	*LTF*	*HDGF*
PHGDH	*ACTN4*	*OPA1*	*NDUFS7*	*F13A1*	*TXNDC12*	*ADH1B*	*HPX*	*LUM*
NDUFS4	*GSTZ1*	*ECE2*	*NDUFS2*	*PNP*	*ITGBL1*	*LDHA*	*FTL*	*PRELP*
DYNC1LI2	*AP1G1*	*PPL*	*GBAS*	*HPRT1*	*EFEMP2*	*ALDH1A1*	*FTH1*	*NDUFA8*
HP	*SYNGR1*	*NRP2*	*PDCD6*	*GOT2*	*SCGB1D1*	*GLUD1*	*MT2A*	*HNRNPA3*
HPR	*SYNGR2*	*MPZL2*	*TBCA*	*EGFR*	*NUDT3*	*ENO1*	*MT-ND2*	*PGD*
F10	*SGTA*	*ACSL4*	*ATP6V1G1*	*MOS*	*AGR2*	*PYGL*	*MT-ND6*	*HNRNPM*

CFD	*ENSA*	*SNX3*	*VPS4B*	*PGK1*	*COX5B*	*GPI*	*MT-ATP8*	*KPNA1*
PLG	*REG1A*	*MGEA5*	*H2AFY*	*AK1*	*CTSA*	*NPM1*	*ANG*	*NCBP2*
F12	*CLEC3B*	*SYNCRIP*	*SH3BGRL*	*F2*	*MGST1*	*TPM3*	*KLKB1*	*POLR2H*
CFB	*SSB*	*CD14*	*DBT*	*C1R*	*C7*	*ITGAV*	*AMH*	*ARHGDIA*
MT-ATP6	*SERPINA7*	*COL4A2*	*PYGB*	*CAPN1*	*KIT*	*CRH*	*SLPI*	*ARHGDIB*
CA2	*SERPIND1*	*CYC1*	*PYGM*	*TUBB*	*RBP3*	*EPHX1*	*VTN*	*SEC23IP*
ASS1	*ITGB1*	*SNRPB2*	*RALA*	*CA3*	*ESD*	*DBI*	*FUCA1*	*SYNE2*
OAS1	*KRT18*	*CFH*	*RALB*	*GRP*	*HSPD1*	*LDHB*	*ALDOA*	*CTNNBL1*
SERPINC1	*KRT8*	*SNRNP70*	*SPTB*	*DCN*	*CLU*	*NEFL*	*CSTB*	*TTN*
SERPINA1	*CRYBA1*	*VIM*	*LAMP1*	*PSAP*	*HSPA5*	*GPX1*	*ANXA1*	*ARAP2*
SERPINA3	*COL5A2*	*SERPINF2*	*ACADM*	*HEXB*	*LAMC1*	*PGK2*	*SOD2*	*CTTNBP2*
AGT	*IGKV4-1*	*KRT19*	*TOP1*	*PFN1*	*MAP2*	*P4HB*	*HRG*	*DNASE2B*
A2M	*GSN*	*KRT7*	*G6PD*	*APRT*	*HSPA8*	*H1F0*	*A1BG*	*PSME3*
C3	*PTMA*	*GNAI3*	*UBL4A*	*EPRS*	*SLC2A1*	*ACYP1*	*KRT6B*	*EXOSC4*
C5	*ATP5B*	*ANXA5*	*DMD*	*CTSB*	*UMPS*	*CRYGD*	*KRT1*	*NOP10*
TIMP1	*C2*	*KRT16*	*MTHFD1*	*HSP90AA1*	*PDHB*	*ADH1A*	*S100B*	*IL1RAP*
CST3	*S100A9*	*RPSA*	*IGF2R*	*HNRNPC*	*CD44*	*CTSD*	*SHBG*	*ACOT13*
CST4	*S100A6*	*CD63*	*ADH5*	*UQCRH*	*NQO2*	*ANXA2*	*TUBB4A*	*SPINK5*
CSTA	*APOA4*	*SNRPA*	*PRPS2*	*LAMB1*	*ITGB4*	*C8A*	*GAPDH*	*MRPL40*
KNG1	*EIF4E*	*FGF2*	*PABPC1*	*FH*	*CBR1*	*C8B*	*ASL*	*CRTAC1*
NRAS	*CKM*	*ENO2*	*PCNA*	*THBS1*	*NCK1*	*C8G*	*IGKV3-11*	*TRIM36*
HRAS	*FABP4*	*ACAA1*	*KRT3*	*RNASE1*	*HIST1H1B*	*MYL12A*	*CAPNS1*	*RTN4*
PDGFB	*GLUL*	*GSTP1*	*NEFH*	*HSPA1B*	*HIST1H1D*	*NCL*	*MT1A*	*CYLD*
MIF	*AKR1B1*	*SNRPC*	*HARS*	*HSPA1A*	*HIST1H1C*	*HK1*	*MT1E*	*MANBAL*
CD99	*ANPEP*	*LGALS1*	*COL11A1*	*COL1A2*	*POR*	*POLR2E*	*HSPB1*	*RPRD1B*
PRKCSH	*GSPT1*	*QDPR*	*COL6A1*	*ANXA6*	*MGMT*	*NDUFV2*	*RPN1*	*PFDN4*
FDPS	*EZR*	*HMGB1*	*COL6A2*	*RHOC*	*ATP2A2*	*EIF2AK2*	*RPN2*	*NIT2*
NID1	*ATP6V1B1*	*RBP1*	*COL6A3*	*CD55*	*FAH*	*SRM*	*GNAI2*	*KIF13B*
AKR1A1	*UCHL3*	*FBP1*	*SLC25A4*	*SERPINA6*	*STMN1*	*ORM2*	*HIST1H2AE*	*XPNPEP1*

PLA2G2A	*CD46*	*TPM1*	*SLC25A6*	*SLC3A2*	*YBX3*	*ITIH2*	*HIST1H2AB*	*GPHN*
PKM	*NME1*	*CLTA*	*IMPDH2*	*PFKM*	*HSPA6*	*ITIH1*	*ATP1A1*	*MYO5C*
HSP90B1	*NQO1*	*CLTB*	*TPR*	*HSP90AB1*	*HMGA1*	*ANXA7*	*ATP1B1*	*DIABLO*
MYL6B	*IGLL1*	*ANXA4*	*PIP*	*SRPR*	*TMEM11*	*BTF3*	*APP*	*DDX21*
SNRPB	*RPS2*	*CNP*	*CKB*	*ELANE*	*GOT1*	*RAB3A*	*ARG1*	*SIAE*
IDE	*DSP*	*DLD*	*ANXA3*	*MMP2*	*ITGA2*	*RAB4A*	*APOD*	*XPO5*
COX6B1	*RPA2*	*HNRNPA1*	*CKMT1A*	*MMP3*	*RPL35A*	*RAB5A*	*ALDH2*	*GRPEL1*
HNRNPL	*COX7C*	*SNRPA1*	*CKMT1B*	*SOD3*	*ARF4*	*RAB6A*	*S100A8*	*C12orf10*
DARS	*TIMP2*	*COX6C*	*BMP3*	*MME*	*RPL7*	*PVALB*	*HMGN1*	*CYP3A43*
JUP	*ACO1*	*TACSTD2*	*ACTN1*	*MGP*	*VCL*	*NPR2*	*SLC25A5*	*CACYBP*
UQCRB	*VDAC1*	*C1S*	*CDH1*	*LPA*	*GPX2*	*PSMB1*	*SERPINA5*	*NMUR1*
LIF	*BGN*	*PARP1*	*PEPD*	*PDHA1*	*LBP*	*M6PR*	*SERPING1*	*PLGRKT*
OGN	*SDHB*	*UCHL1*	*XRCC6*	*GJA1*	*PTPRA*	*HIST1H2AD*	*CFI*	*EML4*
CAST	*CD9*	*LTA4H*	*XRCC5*	*UBTF*	*IL1RN*	*COX5A*	*EIF2S1*	*GLOD4*
COL9A1	*BCKDHB*	*ALDOC*	*COX4I1*	*NDUFB7*	*RPL17*	*LMNB1*	*HMGN2*	*NEK6*
COL5A1	*COMT*	*FREM3*	*IFITM1*	*PRKACA*	*PGAM1*	*PZP*	*RPLP1*	*SPON1*
NEB	*TGM2*	*C4A*	*LAMP2*	*CAPN2*	*SDC1*	*RPS12*	*RPLP2*	*MOV10*
FLG	*KCNA3*	*C4B*	*RNH1*	*DES*	*ATP5J*	*DNAJB1*	*RPLP0*	*ARHGEF10L*
PTMS	*MUT*	*C4B_2*	*BMP1*	*CTPS1*	*KRT15*	*DNAJB2*	*FABP3*	*RPL15*
MAL	*OSBP*	*FBXO45*	*NCAM1*	*DDX5*	*KRT4*	*ATP5A1*	*RPL13*	*MAGOH*
GSTM3	*PCMT1*	*CLEC2L*	*EEF2*	*PFKL*	*FKBP2*	*CTSS*	*IVD*	*RPL27*
ATP6V1B2	*FBL*	*GSTT2*	*MT1G*	*GM2A*	*CRYBB3*	*PSMA1*	*S100A4*	*TTLL12*
ATP6V1C1	*GART*	*RPS17*	*KRT10*	*LGALS3*	*STOM*	*PSMA2*	*MGAT1*	*FHL2*
CSRP1	*TNXB*	*ZBTB48*	*KRT13*	*PSMC3*	*AK4*	*PSMA3*	*HMGB2*	*DPYSL3*
FLNA	*PAICS*	*FDX1*	*KRT5*	*TCP1*	*PON1*	*PSMA4*	*PTBP1*	*ICT1*
MAOA	*ITGA6*	*RAP2A*	*PDIA4*	*PTPN1*	*YWHAQ*	*S100P*	*TARS*	*DCTN1*
IDS	*SFPQ*	*TROVE2*	*C6*	*IGFBP2*	*MAPK3*	*COL5A3*	*VARS*	*DYNC1H1*
SCP2	*PPIB*	*GAA*	*TPT1*	*RPA1*	*ATP6V0C*	*ITGA3*	*EEF1G*	*EIF2B1*
UBA1	*WARS*	*RRAS*	*ALAD*	*APEX1*	*CALML3*	*MSN*	*CRABP1*	*EIF4A2*

NME2	*RPS3*	*ACR*	*HLA-E*	*CAD*	*RPL10*	*DDX6*	*MARCKS*	*CTTN*
ENPP1	*AHCY*	*HIST1H1E*	*LCP1*	*CALR*	*COL8A1*	*CTNNA2*	*AQP1*	*FLOT2*
HIST1H1T	*CFL1*	*SPP1*	*PLS3*	*MAP4*	*X*	*DNMT1*	*GNA11*	*RCN2*
SPRR2D	*EIF4B*	*DLAT*	*APEH*	*CANX*	*LOX*	*PAX6*	*ERP29*	*TRIM25*
FDXR	*ATP2B4*	*TXN*	*ETFA*	*CFP*	*NDUFS1*	*U2AF2*	*PRDX6*	*GNA13*
HNRNPA2B1	*DGKA*	*IGFBP6*	*GYS1*	*PSMB8*	*MAPK1*	*CRABP2*	*C21orf33*	*GALE*
UQCRC2	*VAMP1*	*IGFBP5*	*PRKAR2A*	*PSMB9*	*GCA*	*COL4A5*	*BLVRB*	*GALNT3*
TGM1	*CPT2*	*ACP1*	*ANXA8*	*PSMA5*	*GRN*	*TKT*	*PRDX5*	*CAPRIN1*
CA4	*DTYMK*	*ACAT1*	*ENO3*	*PSMB4*	*LAP3*	*SERPINB3*	*DDT*	*RBM39*
CRYGS	*NR2F2*	*COL8A2*	*COL11A2*	*PSMB6*	*PTPN4*	*LMOD1*	*PRDX3*	*SPARCL1*
FBLN1	*EEF1B2*	*AZGP1*	*CD59*	*PSMB5*	*TPP2*	*SERPINA4*	*ATP5D*	*KRT81*
TCEA1	*ATP5F1*	*MPST*	*CD34*	*S100A2*	*IMPA1*	*EEF1D*	*RPL12*	*PRPSAP1*
ECHS1	*HIST2H4B*	*PTGES3*	*NOMO1*	*PLEC*	*PMVK*	*THOC3*	*CFL2*	*PSMD6*
CMPK1	*HIST2H4A*	*PLCH1*	*TXNRD1*	*GMIP*	*SAR1A*	*ABCC11*	*FARSA*	*SEPT2*
PEBP1	*HIST4H4*	*SVEP1*	*TPD52L1*	*ACTR3B*	*NANS*	*CDK5RAP3*	*ITM2B*	*SART3*
PDIA3	*HIST1H4F*	*ARID4B*	*IMMT*	*KIAA1522*	*SH3GLB2*	*VPS39*	*DRG1*	*EFTUD2*
PPP2R1A	*HIST1H4I*	*GRIPAP1*	*FAM98C*	*LRFN1*	*ASH1L*	*SEC16B*	*TRAPPC4*	*ARL6IP1*
PPP2R1B	*HIST1H4H*	*FNDC1*	*COL24A1*	*RCC2*	*ARHGEF3*	*VCPIP1*	*NCKAP1*	*KARS*
NMT1	*HIST1H4E*	*FAM98B*	*HNRNPUL2*	*KIAA1468*	*OLFML3*	*CDHR1*	*CNPY2*	*EIF4H*
HLA-A	*HIST1H2BC*	*PDCD4*	*KIAA1109*	*DIP2B*	*OSTC*	*BTF3L4*	*EXOC6B*	*ACAP2*
ADSS	*HIST1H2BG*	*TP53I3*	*KIF7*	*STK26*	*PLSCR4*	*RNF170*	*SLC27A6*	*EEA1*
LRPAP1	*HIST1H2BE*	*C11orf73*	*KRT24*	*PTGFRN*	*CYTL1*	*EXOC2*	*GSTK1*	*PDIA6*
TSPO	*HIST1H2BF*	*HSD17B12*	*KIAA1033*	*RRBP1*	*HEBP1*	*CNDP2*	*LAMTOR2*	*PAFAH1B3*
ADSL	*HIST1H2BI*	*LACTB2*	*SGSM1*	*LARS*	*RAB6B*	*EGLN2*	*MRPS28*	*PCOLCE*
CLIP1	*RAB1A*	*FBLN7*	*IAH1*	*ZNF624*	*APOBEC3C*	*RPL10L*	*COA3*	*PEA15*
SRI	*RAN*	*CYBRD1*	*COL28A1*	*KANSL3*	*VPS45*	*CAPNS2*	*MRPS17*	*PGM5*
GSTT1	*RPL23*	*ACTBL2*	*EIF2S3L*	*WRAP73*	*NDUFA4L2*	*SNX27*	*CRYL1*	*EBP*
SERPINB1	*RAP1A*	*PRR16*	*PRTG*	*GMPR2*	*PHPT1*	*IFT74*	*TMA7*	*PCBD1*
ALDH1B1	*UBE2D2*	*ZNF385B*	*TMEM249*	*DPM3*	*TOMM22*	*DHRS1*	*AP3M1*	*RPL37A*

ALDH3A1	*RPS15*	*SPIN4*	*TRIM16L*	*MYO1A*	*CHST7*	*C5orf34*	*GDA*	*DRG2*
LCN1	*RPS24*	*ESCO2*	*LRRFIP1*	*SAE1*	*LANCL2*	*MBOAT7*	*CARHSP1*	*PLTP*
SDHA	*RPS25*	*OCIAD2*	*HNRNPA1L2*	*COPG2*	*KRT84*	*FRMD6*	*THRAP3*	*CSE1L*
GDI1	*RPS26*	*ZNF326*	*LGALS9B*	*MRC2*	*KRT82*	*PVRL4*	*NOP58*	*VCP*
TIA1	*RPS28*	*FSIP2*	*PTCHD3*	*GNG12*	*HOMER3*	*COL21A1*	*ARL2BP*	*HADHB*
DNAJA1	*FAU*	*FLG2*	*DAK*	*OPTC*	*FARSB*	*SERPINB12*	*SUGT1*	*MANF*
UQCRC1	*GNB1*	*TNFAIP8L3*	*LSM12*	*DHCR7*	*IARS2*	*IPO9*	*ACOT9*	*MLLT4*
HIBADH	*POLR2L*	*HERC4*	*KRT71*	*HDAC6*	*ATG3*	*PGLYRP2*	*DERA*	*NAP1L1*
ATIC	*RBX1*	*PPP6R3*	*TBRG1*	*DKK3*	*SMC4*	*MCEE*	*MEMO1*	*CASP7*
APOBEC3A	*GNB2*	*DDX60L*	*VMA21*	*VPS29*	*OLA1*	*RBM14*	*LSM2*	*CASP6*
HNRNPH3	*RPL30*	*RABL3*	*LGALSL*	*EIF3K*	*ECHDC1*	*FANCD2OS*	*STARD10*	*GP2*
HNRNPH1	*RPL31*	*EMC4*	*C2CD3*	*GRHPR*	*ABHD10*	*SMG1*	*SH3GLB1*	*ADK*
YWHAB	*RPL10A*	*C3orf38*	*VPS26B*	*CTSZ*	*LIN7C*	*PPP1R10*	*CAB39*	*ADAR*
SFN	*RPL32*	*PLAC9*	*LARP7*	*DNAJB11*	*ABCF3*	*VPS35*	*LUC7L2*	*LAMB2*
STIP1	*RPL11*	*GNAS*	*PPA1*	*UBA2*	*FAM49B*	*PURB*	*DHRS7*	*CDH11*
S100A11	*RPL8*	*SAMD9*	*LZIC*	*DKK2*	*DDX19A*	*MAP2K1*	*FA2H*	*CDH13*
PRDX2	*PPIA*	*COLEC12*	*IRGQ*	*NXF1*	*EXOC1*	*FKBP4*	*GET4*	*TPD52*
DCTD	*FKBP1A*	*SPDYA*	*DDB2*	*SEL1L*	*SEPT11*	*PLOD1*	*GDPD3*	*SEC13*
DSG3	*RPS27A*	*TCHHL1*	*AQP3*	*PEF1*	*PANK4*	*NUCB1*	*COPS6*	*NHP2L1*
RPL9	*GRB2*	*ELFN2*	*DDX1*	*COPS7A*	*FANCI*	*RPL6*	*FASTKD5*	*NPEPPS*
ACSL1	*TRA2B*	*EXOSC6*	*HSD17B8*	*FBLN5*	*ATAD3A*	*SLC25A11*	*MOB1B*	*EPS8L2*
KIF5B	*RAC1*	*PHYHD1*	*FAM3C*	*PPP1R1B*	*ARL8B*	*DST*	*CHMP1B*	*ESRP2*
DUT	*AP2B1*	*MUC21*	*H1FX*	*CLIP2*	*DNAJC17*	*CAV1*	*RRM2B*	*DCTPP1*
HIST1H2BB	*VAMP2*	*HP1BP3*	*CCT6B*	*UQCR10*	*PNPO*	*ACY1*	*NLRC3*	*AAMDC*
RNASE4	*GNAI1*	*LDLRAP1*	*PSMF1*	*TJP2*	*NECAP2*	*KMT2A*	*KRT74*	*UBE2Z*
SHMT1	*PPP3R1*	*FRMPD1*	*AP3S1*	*MALT1*	*GID8*	*TNFAIP2*	*PHF5A*	*METTL7A*
SHMT2	*YWHAZ*	*C6orf132*	*UBXN4*	*VTI1B*	*C1orf123*	*MVK*	*SLC16A9*	*MANSC1*
EPHX2	*PPP2R2A*	*ARMC4*	*NDRG1*	*STK39*	*C2orf42*	*TAP1*	*USP17L1*	*PLEKHF2*
HSPA1L	*DYNLL1*	*HECTD3*	*HSPH1*	*DNAH17*	*NDUFB11*	*CFHR1*	*ERMP1*	*HN1L*

HSPA4	*DYNLT1*	*UBR4*	*SEPT8*	*NIPSNAP3A*	*NHP2*	*COL10A1*	*MRPL21*	*SP140L*
MPI	*RPL38*	*KIAA1217*	*RQCD1*	*TES*	*OCIAD1*	*CEBPZ*	*NRK*	*MED20*
PFN2	*SKP1*	*MRPL2*	*JADE3*	*DMBT1*	*ADPRHL2*	*UBXN1*	*EFTUD1*	*SFXN1*
CTNNA1	*GNG5*	*CFAP58*	*MYO18A*	*SEC63*	*CHCHD3*	*GBE1*	*MAMDC2*	*AKIRIN1*
CTNNB1	*RPS21*	*L1TD1*	*GCN1L1*	*SUSD2*	*MIOS*	*EIF4G1*	*ATG9A*	*CAB39L*
PHB	*EIF5A*	*CPSF3L*	*NUP205*	*SWAP70*	*MTMR10*	*KRT17*	*CCDC186*	*FN3KRP*
SERPINB6	*GNB2L1*	*MLK4*	*NSMAF*	*SUN2*	*TEX10*	*GLO1*	*C16orf62*	*TBC1D17*
RDX	*UBE2I*	*SH3PXD2A*	*SORL1*	*LAMTOR3*	*THUMPD1*	*AKR1C1*	*MON2*	*PPCS*
RPA3	*PPP2CA*	*DDI2*	*ANP32B*	*SRP68*	*BRE*	*SSBP1*	*MYH14*	*SRPRB*
RPL22	*YBX1*	*NT5DC1*	*TFG*	*NPC1L1*	*CDKN2AIP*	*YWHAH*	*SZRD1*	*GMPPB*
SPR	*SEC11A*	*MINOS1*	*HTRA1*	*SEPT9*	*GAR1*	*PLP2*	*DHX29*	*TRAPPC1*
THBS4	*CSNK2B*	*RASIP1*	*ARPC1A*	*UBQLN2*	*DPP3*	*CLC*	*COMMD6*	*CDC42BPB*
KRT9	*TPM4*	*ATP5EP2*	*FRZB*	*PCYOX1*	*DACT1*	*PRKCD*	*HDDC2*	*RBM8A*
SAA4	*ACTC1*	*ZNF318*	*TAF15*	*SHPK*	*BCLAF1*	*COL14A1*	*KDELC2*	*MPC1*
FBN1	*UBE2L3*	*SDHAF4*	*GOLGA1*	*CYB5R1*	*TMOD3*	*RNASEL*	*KIF21A*	*IER3IP1*
AGL	*EEF1A1*	*ARHGEF16*	*EVPL*	*PFDN2*	*NKIRAS2*	*KDSR*	*DCXR*	*SNX5*
MYH9	*TUBA1B*	*BROX*	*GGH*	*PUF60*	*NKIRAS1*	*REG3A*	*SPATA21*	*HEBP2*
MYH10	*TUBA4A*	*LYPLAL1*	*DDX17*	*ENOPH1*	*UGGT1*	*GFPT1*	*GOLGA7*	*CSAD*
COPB2	*TUBB4B*	*KIAA0368*	*RAD50*	*NDUFA12*	*TECR*	*PVRL1*	*PDGFRL*	*PUS1*
ACTN2	*CSNK2A1*	*C9orf131*	*OSTF1*	*EIF2B4*	*ERAP1*	*RPS7*	*HSPE1*	*LRRFIP2*
BSG	*PAFAH1B2*	*KIAA1614*	*UFD1L*	*ATP6V1H*	*ACTR10*	*PPP1CA*	*SEC61A1*	*PSAT1*
TIMP3	*HIST1H3A*	*KRT79*	*GLG1*	*RABAC1*	*CISD1*	*PPP1CB*	*LYZ*	*F11R*
FUS	*HIST1H3C*	*C19orf70*	*UPF1*	*TRMT112*	*CLIC5*	*CALM1*	*VBP1*	*NPTN*
DEK	*HIST1H3H*	*TRIQK*	*RPL3L*	*CTNNA3*	*FAM120A*	*CALM2*	*B2M*	*CPQ*
MYH11	*HIST1H3B*	*TNS2*	*COPS5*	*XPO7*	*GLTP*	*CALM3*	*DAD1*	*COPG1*
GLRX	*HIST1H3I*	*METRNL*	*SMARCD2*	*VPS51*	*UBP1*	*PSMC1*	*NPC2*	*FKBP7*
PPM1A	*HIST1H3J*	*TBC1D9B*	*RAB8B*	*BAZ2A*	*MTCH1*	*PSMC5*	*COPZ1*	*CLIC4*
IGFALS	*HIST1H3D*	*MAP1S*	*BAD*	*ATPIF1*	*MAT2B*	*RPS8*	*PKIA*	*CYP46A1*
KRT2	*HIST1H3E*	*STEAP4*	*KHSRP*	*AK3*	*MYOF*	*RPS15A*	*SUMO2*	*CFAP20*

HMGCL	*HIST1H3F*	*CYB5R2*	*TNPO1*	*KIF25*	*EHD3*	*RPS16*	*UFM1*	*SAR1B*
PSMC2	*HIST1H3G*	*ATL3*	*DVL3*	*LNPEP*	*C1RL*	*YWHAE*	*WDR5*	*MTCH2*
CHI3L1	*HBB*	*RPL7L1*	*USP9X*	*MSRA*	*CALML5*	*RPS14*	*AP1S1*	*ARFGEF1*
GGT5	*HBA2*	*NXN*	*USP7*	*NAGK*	*CHMP5*	*RPS23*	*NUTF2*	*BZW2*
ARL2	*HBA1*	*VASN*	*CUL5*	*SH3BGRL2*	*COMMD9*	*RPS18*	*HNRNPK*	*WNT6*
ARL3	*CXADR*	*HIST2H2AA4*	*ATP6V0A1*	*SLC25A13*	*CWC15*	*WIBG*	*YWHAG*	*COMMD10*
MAP2K2	*SIRPA*	*HIST2H2AA3*	*HIST1H2BH*	*DBNL*	*THYN1*	*MIEN1*	*RRAS2*	*CHCHD2*
ATP5C1	*CYP4F2*	*CIAPIN1*	*TCEAL3*	*ANAPC5*	*CCDC59*	*ERP44*	*CDC42*	*OAS3*
ATP6V1E1	*PSPH*	*PTRHD1*	*MYDGF*	*ANAPC4*	*HACD3*	*LXN*	*DSTN*	*NDUFB9*
RPL4	*EIF4G2*	*NADSYN1*	*LOH12CR1*	*STOML2*	*NDUFA13*	*RAB10*	*GMFB*	*SQRDL*
PGM1	*GTF2I*	*LAMTOR1*	*MAL2*	*FBXO2*	*RAI14*	*UBE2D3*	*RAB8A*	*AP1M2*
GNL1	*CCT2*	*DHRS7B*	*RPL36AL*	*ZNF580*	*VAPA*	*UBE2M*	*SPCS3*	*SLC4A4*
SERPINB5	*KRT83*	*TWF2*	*RAB24*	*VPS28*	*H2AFY2*	*UBE2K*	*NONO*	*SCIN*
POLR2I	*PRAME*	*ZNF770*	*SLC35B4*	*LSM7*	*MRPL27*	*UBE2N*	*PPP2R4*	*WASF2*
SERPINF1	*RAE1*	*RAB12*	*WBP2*	*USP21*	*GSKIP*	*RAB14*	*RAB35*	*FAM169A*
DLST	*GSTO1*	*SLC25A25*	*TMBIM1*	*APPL1*	*ORMDL1*	*ACTR3*	*RCN1*	*IVNS1ABP*
GPX4	*PRKDC*	*PDE12*	*ERGIC1*	*PARP4*	*TMEM14C*	*ACTR2*	*KRT31*	*CPNE1*
CFHR2	*BLOC1S1*	*CAPN13*	*ATP6V1E2*	*NUDT5*	*SEP10*	*ACTR1A*	*TMED2*	*CCT7*
ACVRL1	*IL13RA1*	*SLC25A24*	*HIST1H2BA*	*RALY*	*UNC5B*	*SRP54*	*PCBP1*	*PRMT1*
HSD17B2	*CRADD*	*RETSAT*	*FAM162A*	*C14orf1*	*ABI1*	*RAB4B*	*PCBP2*	*BAG1*
SRP14	*ADARB1*	*TCTN3*	*MYLPF*	*PACSIN3*	*C11orf31*	*RAB2A*	*TCEB1*	*PKP2*
NUP62	*LCN2*	*UGT3A1*	*TMEM230*	*ACIN1*	*ASPM*	*RAB5B*	*TCEB2*	*RARRES2*
HPCAL1	*MT1X*	*HIBCH*	*EXOC4*	*AGO2*	*APCDD1*	*NTPCR*	*RHEB*	*MYOC*
TAGLN2	*DLK1*	*RWDD4*	*ISOC2*	*MYH2*	*C1orf87*	*HDHD3*	*TOMM20*	*SEMA3C*
TALDO1	*ABAT*	*ESRP1*	*FUBP1*	*CDV3*	*VKORC1L1*	*ESYT1*	*SF3B3*	*NIPSNAP1*
ETFB	*BASP1*	*H3F3C*	*LRRC59*	*BAG5*	*FUK*	*COPS4*	*RSU1*	*ARPC5L*
RBMX	*DCD*	*KANK3*	*CCDC74A*	*AGO1*	*SPICE1*	*TMEM121*	*CNN3*	*FAM110A*
GNAL	*MRPS36*	*PTRF*	*CLDN23*	*RAB21*	*ABCA6*	*DCTN5*	*RPS6KA1*	*WDR77*
ATP6V1A	*MRPS9*	*TATDN1*	*C10orf32*	*RAB22A*	*CCAR2*	*ANP32E*	*SAFB*	*TUBA1C*

HSPA9	*SARNP*	*EDC4*	*TOMM6*	*PSME2*	*FAM58A*	*TMEM43*	*SF3B4*	*FERMT1*
EIF4A3	*RPL24*	*PRPF8*	*CREB3L1*	*MKLN1*	*NUP93*	*FSD1*	*SF3A2*	*CORO1B*
RPS19	*TXNL4A*	*FBXO42*	*ARL8A*	*DNPEP*	*LRRC47*	*DHRS4*	*PPP1R7*	*TXNDC17*
RPL3	*RPL36A*	*SCYL2*	*PTER*	*MCTS1*	*KRT78*	*NDUFAF3*	*SEC23A*	*CPPED1*
COL18A1	*CBX1*	*PLBD1*	*ARMC5*	*OGDHL*	*MYO1H*	*TMEM70*	*SEC23B*	*VPS25*
DDOST	*HPCA*	*FAHD1*	*EFHD2*	*FAM184B*	*PROM2*	*TUBB6*	*SF3A1*	*NUDT16L1*
ANP32A	*ARF1*	*ERICH5*	*GALM*	*NWD2*	*ARMC10*	*C9orf142*	*TGFBI*	*PAAF1*
FEN1	*ARF5*	*PGM2L1*	*DCPS*	*PRR12*	*ABHD12*	*HNRNPUL1*	*TSN*	*API5*
CAPG	*ERH*	*TTC37*	*PPP1R14B*	*MYO5B*	*GPD1L*	*PDCD10*	*SF1*	*DPY30*
ADCY8	*RHOG*	*FIGNL1*	*ISOC1*	*CORO1C*	*ASCC3*	*DERL1*	*TRIP10*	*UNK*
CCT6A	*RPL19*	*POTEE*	*AP2M1*	*PYCARD*	*MPP5*	*EFHD1*	*HMGN3*	*TNKS1BP1*
NNMT	*SRSF3*	*CRTC3*	*KCTD12*	*TMCO1*	*CPXM2*	*HTATIP2*	*JMJD1C*	*CCSER1*
PSMB10	*MXRA7*	*MOXD1*	*RCN3*	*MYO6*	*MINK1*	*APOO*	*TRIP6*	*C20orf27*
ADH7	*H3F3B*	*TMEM205*	*REPS1*	*COL17A1*	*CBR4*	*BDH2*	*MAPRE1*	*PITHD1*
RPL13A	*H3F3A*	*SBSN*	*ZG16B*	*NFU1*	*OXR1*	*MRI1*	*TSC22D1*	*WDR61*
ARL1	*TNFAIP6*	*C6orf89*	*RMDN1*	*PRPF19*	*SMIM20*	*VAMP8*	*ELAVL1*	*SLIRP*
STAT3	*HSPG2*	*OLFML1*	*NUDT16*	*UBQLN1*	*CISD2*	*ADI1*	*NSDHL*	*SRR*
MDH1	*RBM3*	*PLXDC2*	*CMBL*	*NENF*	*JAGN1*	*EMC6*	*MYLK*	*NIF3L1*
MDH2	*CYCS*	*LYPD2*	*MSI2*	*SNX12*	*CCDC33*	*TMEM109*	*SLC1A5*	*CHST6*
IL15	*TFAM*	*VIT*	*SNRNP40*	*NDRG2*	*EIF4E3*	*PBDC1*	*MLF2*	*NAA50*
HADHA	*PITPNA*	*NPNT*	*DNAH11*	*PCDHA4*	*COMMD1*	*DUSP23*	*SMAD2*	*LACRT*
EIF2S3	*MAT1A*	*WDR82*	*MTBP*	*G3BP2*	*KRTCAP2*	*TMED9*	*UBE2V2*	*XRN2*
CETN2	*SLC25A3*	*IGSF10*	*MRRF*	*STUB1*	*ARL6IP6*	*GGACT*	*CST6*	*TOLLIP*
PTGDS	*HDLBP*	*NAPRT*	*RBMXL1*	*PACSIN2*	*LRRC17*	*C1orf116*	*STK11*	*WWP1*
KDM5C	*CDK16*	*CD109*	*SEH1L*	*SNX6*	*CSNK1A1L*	*KIFC1*	*STXBP2*	*ATP6V0D1*
CSK	*PURA*	*HSDL2*	*GNPNAT1*	*STX8*	*AMER2*	*IFT27*	*RPS29*	*NMNAT1*
GARS	*CLTC*	*CCBL2*	*MRPS24*	*SSR3*	*ANKRD44*	*ACAT2*	*RPS13*	*RHOA*
IARS	*FKBP3*	*KDM7A*	*DAZAP1*	*PSMD13*	*KIAA1958*	*FUNDC2*	*RPS11*	*FAM114A1*
ACTR1B	*PLCB2*	*ZNF813*	*DTNBP1*	*COPS3*	*PTGR2*	*SFXN3*	*SNRPE*	*MGA*

ECI1	*REEP5*	*NBEAL2*	*MVB12A*	*RPL26L1*	*MRPL43*	*CHID1*	*SNRPF*	*UBR1*
ASIP	*SORD*	*CATSPERG*	*MMAB*	*NSFL1C*	*LRRC57*	*SORBS1*	*LSM3*	*CCAR1*
TMPO	*HNRNPU*	*SRCAP*	*DYNLL2*	*SLC12A4*	*MARCH10*	*JAM3*	*LSM6*	*DNAJC10*
TMPO	*U2AF1*	*FAM83H*	*CPNE2*	*BSN*	*KDF1*	*HINT2*	*SNRPD1*	*RHOT2*
STAT1	*SPTBN1*	*FAM179A*	*S100A16*	*MACF1*	*PDDC1*	*NAA15*	*SNRPD2*	*MSRB3*
SKIV2L2	*TIAL1*	*NYAP1*	*SCRN2*	*SCAF8*	*NHLRC2*	*BCL2L13*	*SNRPD3*	*NRM*
AKR1C3	*SET*	*TOM1L2*	*OTUB1*	*PLCL2*	*COLGALT1*	*PRX*	*TMSB4X*	*SP8*
MTOR	*SRSF2*	*NCCRP1*	*CHMP6*	*EXOC7*	*SUMF2*	*SRRT*	*ARF6*	*DDX60*
CASP3	*AMPD3*	*KCTD8*	*PGM2*	*MAPK8IP3*	*RDH13*	*TLR10*	*PSMC6*	*SRRM1*
RPS27	*PMP22*	*NEK10*	*DCUN1D1*	*ZC3H4*	*HSD17B11*	*CACNG6*	*RPL7A*	*EXOC8*
LRPPRC	*FABP5*	*TMTC3*	*SMARCD1*	*SAMD4A*	*TP53I13*	*ANKRD30B*	*POLR2G*	*SYTL1*
ACAA2	*ANK2*	*VPS13C*	*BCAN*	*PA2G4*	*TXNDC5*	*ITPA*	*RAB11A*	*CCDC105*
RPL35	*CAP1*	*ACAD11*	*SNF8*	*BAIAP2*	*ATAD1*	*SEC11C*	*ETF1*	*FAM186B*
PAFAH1B1	*KRT76*	*MOB2*	*ZNF764*	*SMC3*	*UXS1*	*CADM1*	*CNBP*	*KATNAL2*
MATR3	*DR1*	*TUBA1A*	*PDLIM5*	*CHMP2B*	*SERBP1*	*KIAA1671*	*YPEL5*	*ELMOD2*
GPD2	*MC1R*	*H2AFV*	*ERO1L*	*TNN*	*CHST14*	*MRPL13*	*RPS4X*	*ALDH16A1*
SSR1	*PFKP*	*RPS27L*	*TMEM19*	*GNE*	*C2CD4A*	*SRXN1*	*ACTA2*	*H1FOO*
CRYBB2	*XPC*	*CBLL1*	*C7orf55*	*C14orf166*	*APOA1BP*	*NLN*	*RPL23A*	*COL27A1*
SYK	*EWSR1*	*PDZD4*	*DOCK6*	*RUVBL2*	*SMG8*	*TBC1D2*	*RPS6*	*FBXO22*
RANBP1	*COL4A3*	*BTN2A1*	*REEP6*	*PIN4*	*MROH1*	*UACA*	*VSNL1*	*SYNE1*
NAMPT	*TAGLN*	*SND1*	*DDRGK1*	*CLEC11A*	*RBM15B*	*WDR11*	*HIST1H4A*	*ANKK1*
AFM	*OGDH*	*MARK2*	*FUBP3*	*TCF19*	*PDZD8*	*CRNKL1*	*HIST1H4B*	*TSTD1*
PSMC4	*ALDH6A1*	*DDX46*	*COX14*	*EIF3L*	*LFNG*	*UBL5*	*HIST1H4C*	*FDCSP*
PPIC	*COL7A1*	*BZW1*	*SUCLG2*	*PLAA*	*APPL2*	*FAM129A*	*HIST1H4K*	*ABHD11*
VDAC2	*DSG1*	*CYB5R4*	*GMPPA*	*RUVBL1*	*XDH*	*RAP2B*	*HIST1H4J*	*HELB*
ACADSB	*DSC2*	*EIF3M*	*ABHD14B*	*NUDC*	*ATP5O*	*RPS3A*	*HIST1H4L*	*ATL2*
USP5	*RPL18A*	*RRAGA*	*USMG5*	*VDAC3*	*PREP*	*RPL26*	*HIST1H4D*	*KDM2B*
MAPK8	*NHLH1*	*CYFIP1*	*CPB2*	*VSIG4*	*ARCN1*	*RAP1B*	*COPS2*	*RFWD2*
LSS	*PSME1*	*RAI1*	*HIST1H2AA*	*MOB4*	*HNRNPF*	*HDAC1*	*NEDD8*	*IL31RA*

GCLC	*FMR1*	*TMC3*	*VPS13A*	*SBDS*	*GTF2A2*	*CAMK2G*	*ADIRF*	*MCFD2*
PCP4	*FMOD*	*ARPIN*	*GFM1*	*TMED5*	*MSH6*	*DCTN2*	*RAB11B*	*WDR36*
PSMD8	*PRDX1*	*HUWE1*	*C16orf13*	*TMED7*	*VAV2*	*NAE1*	*ZYX*	*PSMA8*
GSS	*RPL18*	*KRT77*	*GPX7*	*SF3B6*	*HRSP12*	*IRF5*	*SEPT7*	*SMARCC2*
CCT5	*C1QBP*	*CNTRL*	*CDK5RAP1*	*MRPL11*	*MRPL12*	*STIM1*	*ADRM1*	*SUN3*
KRT6C	*CKAP4*	*ABI3BP*	*RSF1*	*REXO2*	*STC1*	*TRA2A*	*PRR4*	*NPLOC4*
CSNK1A1	*COL16A1*	*TMED4*	*RBM15*	*CCDC53*	*DGKQ*	*SNX1*	*PSMD5*	*EFEMP1*
IDH2	*TJP1*	*RALYL*	*MMS19*	*TPRKB*	*THOP1*	*CUL1*	*GRIK5*	*SF3A3*
PITPNB	*DPT*	*S100A7A*	*FAM129B*	*PPIL1*	*AKR1C2*	*CUL2*	*CSRP2*	*AIMP1*
ALDH9A1	*KHDRBS1*	*SCD5*	*MYCBP*	*UFC1*	*CAPZA1*	*CUL3*	*DDB1*	*ILF2*
RPL34	*BAX*	*TRAPPC6B*	*TBCB*	*FAM96B*	*CRIP2*	*CUL4B*	*CDC37*	*ILF3*
LMAN1	*BCL2L1*	*SERPINA11*	*PSMB7*	*FIS1*	*BLVRA*	*TSTA3*	*DPYSL2*	*LMAN2*
NASP	*KLC1*	*PABPN1*	*CNN2*	*MRPS23*	*SLC25A1*	*FHL1*	*SYPL1*	*LRMP*
FASN	*SOS1*	*METTL3*	*SEC62*	*HDGFRP3*	*ARFIP2*	*MOGS*	*ECM1*	*TRAP1*
FNTA	*LRP1*	*MTDH*	*PCYT2*	*PTRH2*	*ARFIP1*	*ALCAM*	*SRSF7*	*DLG1*
CCT3	*SRSF1*	*KTN1*	*KRT12*	*CHMP3*	*ACLY*	*LAMB3*	*CPSF6*	*MYO1E*
TUFM	*ARHGAP1*	*FAM83A*	*CDC5L*	*STRAP*	*COL4A4*	*LAMC2*	*FSCN1*	*CSTF3*
ALDH7A1	*SRSF4*	*MDP1*	*PSMD1*	*RTCB*	*METAP1*	*NACA*	*IFI16*	*ECH1*
INPP1	*PCDH1*	*CAND1*	*PFDN5*	*SH3BP1*	*SUCLG1*	*SPTAN1*	*HIST3H3*	*ARHGAP5*
SRP9	*DHX9*	*COMMD7*	*AGAP2*	*RAP2C*	*COPB1*	*PKP1*	*DECR1*	*FLII*
PCYT1A	*CRYZ*	*TPRA1*	*PARK7*	*STARD13*	*COPA*	*DDX39B*	*NDUFA5*	*COASY*
AARS	*DDR1*	*CCDC25*	*VAT1*	*TMED3*	*CTSC*	*BLMH*	*TST*	*GPS1*
CARS	*GOLGA2*	*RALGAPB*	*POP1*	*HCN4*	*CRYBA4*	*TUBB2A*	*QPCT*	*USP4*
SARS	*LGALS3BP*	*NIT1*	*S100A13*	*RPL36*	*CRYBB1*	*IDI1*	*GUK1*	*MTAP*
PSMB3	*MFGE8*	*COMMD2*	*TSNAX*	*CHTOP*	*AP2S1*	*CAPS*	*HAGH*	*PRKAA1*
PSMB2	*DSC1*	*C19orf68*	*EIF3C*	*PKP3*	*IST1*	*CCIN*	*LAMA3*	*TARDBP*
COMP	*HDHD1*	*COL23A1*	*C12orf57*	*TLN1*	*SEC24C*	*CKAP5*	*NDUFA9*	*HNRNPA0*
ACADVL	*PPID*	*NDUFA11*	*PHB2*	*AIM1*	*SUB1*	*CIRBP*	*HADH*	*AIMP2*
VPS41	*SSRP1*	*KRT73*	*COPS8*	*HYOU1*	*CLNS1A*	*COTL1*	*UGP2*	*PRDX4*

TMED10	*CAMSAP2*	*STX12*	*CHP1*	*AFG3L2*	*RARS*	*COL4A6*	*ATP6V1F*	*PAK2*
RBM25	*SLFN5*	*ERO1LB*	*HSD17B10*	*LSM4*	*YARS*	*COL9A2*	*ABCE1*	*CBX3*
HINT1	*ACSM2A*	*GPHB5*	*COL12A1*	*PRRC2C*	*USP14*	*COX17*	*SDF2L1*	*PSMD2*
FHIT	*NSUN2*	*HRNR*	*SIGMAR1*	*OARD1*	*PRKAG1*	*HNRNPD*	*CELSR2*	*SEMA3B*
RANBP2	*RBBP4*	*KRTAP11-1*	*BARD1*	*HES2*	*HSPA2*	*SCARB2*	*S100A14*	*SELENBP1*
NDUFV1	*NCBP1*	*HIST2H2AB*	*HNRNPAB*	*PPME1*	*ATP1B3*	*IL18*	*CHMP1A*	*NME3*
GSK3B	*AHNAK*	*TRAPPC5*	*NAP1L4*	*AP4S1*	*RAD23A*	*DSG2*	*TM9SF3*	*SRSF9*
NT5C2	*AP1B1*	*AEBP1*	*MGST2*	*SUPT16H*	*RAD23B*	*TRIM29*	*APMAP*	*SRSF5*
SEPP1	*CPSF1*	*AHNAK2*	*CDC6*	*PCDHAC2*	*AK2*	*UBE4A*	*ABCB6*	*SRSF6*
MTHFS	*KIF1A*	*NAV2*	*NAPG*	*PHLDA3*	*EMP2*	*SEPT6*	*RAB18*	*NOG*
GMPS	*KIAA0196*	*CARKD*	*ATP5S*	*UCHL5*	*NAPA*	*EIF3A*	*VTA1*	*TRIM28*
LIG3	*NUP160*	*SLC44A2*	*OLFM1*	*ATP6V1D*	*EIF5*	*UBAP2L*	*PNPLA8*	*STX3*
MRE11A	*TWF1*	*WDFY1*	*ACO2*	*TNPO3*	*SLC12A2*	*SCRIB*	*DYNLRB1*	*G3BP1*
GNAQ	*ASPH*	*MFSD6L*	*TSG101*	*TIMM13*	*PSMD4*	*MLEC*	*ENY2*	*PABPC4*
IDH3A	*ST13*	*ATP1A2*	*GALK1*	*SULT1A1*	*VASP*	*ANXA11*	*SSR4*	*CRIP1*
DNM2	*FXR1*	*BCAP31*	*MMP14*	*NUDT2*	*RAB5C*	*TPMT*	*GDI2*	*LRBA*
RAB7A	*ALDH3A2*	*EMD*	*BCAM*	*RAB9A*	*HSD17B4*	*SLC26A2*	*PPT1*	*RAB13*
PSMD7	*SERPINB8*	*RPL14*	*PLCD1*	*SERPINB9*	*CCT8*	*BLK*	*SERPINH1*	*CCT4*
DUSP3								