Molecular Biology and Genetics of Fuchs' Endothelial Corneal Dystrophy

Abraham Anyuon Kuot

Bachelor of Science (Honours)

This thesis is submitted for the degree of Doctor of Philosophy

Department of Ophthalmology

Faculty of Medicine, Nursing, and Health Sciences

School of Medicine

Flinders University of South Australia

South Australia, AUSTRALIA

June 2015

TABLE OF CONTENTS

TABLE OF CONTENTS ..... **ii**

SUMMARY ..... **vii**

PRESENTATIONS ARISING FROM THIS THESIS ..... **x**

DECLARATION BY THE STUDENT ..... **xi**

ACKNOWLEDGEMENTS ..... **xii**

ABBREVIATIONS ..... **xvii**

CHAPTER 1 ..... **1**

Introduction **1**

1.1 STRUCTURE OF THE HUMAN EYE *1*

1.2 THE CORNEA *2*

1.3 CORNEAL DYSTROPHIES *6*

1.3.1 Definition and types 6

1.3.2 Genetic causes 6

1.4 FUCHS' ENDOTHELIAL CORNEAL DYSTROPHY *7*

1.4.1 Treatment of FECD 9

1.4.2 Epidemiological studies in FECD 9

1.4.3 Genetic studies in FECD 11

1.4.3.1 Conventional genetic studies 11

1.4.3.2 Genetic association studies 15

1.4.4 Molecular studies in FECD 19

1.4.4.1 Studies in human specimens 20

1.4.4.2 Studies in murine models 33

1.5 GAPS IN CURRENT KNOWLEDGE ABOUT FECD *35*

1.6 HYPOTHESES AND AIMS *36*

1.6.1 Hypotheses 36

1.6.2 Aims 36

CHAPTER 2 ..... **38**

Materials and Methods **38**

2.1 ETHICS APPROVALS *38*

2.2 PARTICIPANT RECRUITMENT *38*

2.2.1 Genetic studies 38
2.2.1.1 Fuchs' endothelial corneal dystrophy (FECD) cases 38
2.2.1.2 Normal South Australian controls 40
2.2.1.3 Blue Mountain Eye Study controls 40
2.2.2 Molecular studies 40

2.3 SAMPLE PREPARATION *41*
2.3.1 DNA extraction from peripheral blood 41
2.3.2 Preparation of DNA pools 42
2.3.2.1 Estimation of DNA concentration 42
2.3.2.2 Construction of DNA pools 44
2.3.3 RNA extraction 45
2.3.4 Protein extraction 46
2.3.4.1 For Western blotting 46
2.3.4.2 For Label-free mass spectrometry 48

2.4 EXPERIMENTAL PROTOCOLS *49*
2.4.1 DNA sequencing 49
2.4.1.1 Primer design 49
2.4.1.2 Polymerase chain reaction 50
2.4.1.3 Agarose Gel Electrophoresis 51
2.4.1.4 Clean-up of PCR products 52
2.4.1.5 DNA Sequencing 52
2.4.1.6 Sequence Analysis 52
2.4.2 Sequenom MassARRAY® SNP genotyping 53
2.4.3 RNA analysis 53
2.4.3.1 Primers 53
2.4.3.2 Complementary DNA synthesis 54
2.4.3.3 Quantitative reverse transcription-polymerase chain reaction 55
2.4.3.4 Data analysis 56
2.4.3.5 Statistical analysis 56
2.4.4 Protein analysis 56
2.4.4.1 Western blotting 56
2.4.4.2 Immunohistochemistry 58

CHAPTER 3 60

Mutation screening of FECD-causing genes in Caucasian Australian cases 60

3.1 INTRODUCTION *60*

3.2 MATERIALS AND METHODS *61*
3.2.1 Ion Torrent AmpliSeq™ sequencing 61
3.2.1.1 Quantitation and pooling of FECD DNA samples 62
3.2.1.2 Designing Ion AmpliSeq™ Custom Primer Pools 63
3.2.1.3 Multiplex PCR amplification of target genomic regions 63

3.2.1.4 Preparation of amplicon libraries 64
3.2.1.5 Template preparation and Emulsion PCR 67
3.2.1.6 Sequencing 69
3.2.1.7 Data Analysis 69
3.2.2 Validation of identified novel variants by Sanger sequencing 70
3.2.3 Screening of identified potential novel mutations using Custom TaqMan SNP Genotyping Assay 70
3.2.4 Prediction of effects of identified novel mutations 71

3.3 RESULTS *72*
3.3.1 Mutation screening 72
3.3.2 Validation of prioritised novel variants 78
3.3.3 Screening of the confirmed variants in NSA controls 81

3.4 DISCUSSION *85*

CHAPTER 4 89

Identification of novel genes associated with late-onset FECD in Caucasian Australian cases 89

4.1 INTRODUCTION *89*

4.2 METHODS *90*
4.2.1 Genotyping of DNA pools 90
4.2.2 Individual genotyping of candidate SNPs 92
4.2.3 Data analysis 92

4.3 RESULTS *92*
4.3.1 Pooled GWAS and replication findings 92
4.3.2 Replication study of FECD-associated SNPs identified through GWAS in a Caucasian American case-control study 100
4.3.3 Expression analyses of *ATP1B1* and *LAMC1* 106

4.4 DISCUSSION *116*

CHAPTER 5 122

Association of TGC repeat polymorphism in *TCF4* with FECD in Australian cases 122

5.1 INTRODUCTION *122*

5.2 METHODS *124*
5.2.1 Polymerase chain reaction 124
5.2.2 Short tandem repeat (STR) Assay 124
5.2.3. Statistical analyses 125

5.3 RESULTS *125*

5.4 DISCUSSION *132*

CHAPTER 6 135

Identification of differentially abundant proteins between FECD-affected and unaffected Descemet's membrane 135

6.1 INTRODUCTION 135

6.2 MATERIALS AND METHODS 136

6.2.1 Comparative mass spectrometry analysis 136

6.3 RESULTS 139

6.3.1 Identification of differentially abundant proteins between FECD-affected and unaffected Descemet's membrane 139

6.3.2 Distribution of APOE in FECD-affected and normal corneas 144

6.3.3 Relative expression of *APOE* transcript between FECD-affected and normal corneal endothelium 146

6.4 DISCUSSION 148

CHAPTER 7 154

Identification of differentially expressed genes between FECD-affected and unaffected corneal endothelium 154

7.1 INTRODUCTION 154

7.2 MATERIALS AND METHODS 155

7.2.1 Microarray analysis 155

7.2.2 Quantitative RT-PCR analysis 158

7.2.3 Bioinformatic analysis 158

7.3 RESULTS 159

7.3.1 Characteristics of participants for expression analysis 159

7.3.2 Quality assessment for RNA samples 159

7.3.3 Microarray analysis 160

7.3.4 Prioritisation of differentially expressed genes for validation 168

7.3.5 Demonstration of specificity of primers used for quantitative RT-PCR. 169

7.3.6 Determination of amplification efficiencies of primers used for quantitative RT-PCR 170

7.3.7 Validation of prioritised differentially expressed genes 172

7.3.9 Summary of results 178

7.4 DISCUSSION 178

CHAPTER 8 189

FINAL DISCUSSION 189

APPENDIX 210

REFERENCES 218

SUMMARY

Fuchs' endothelial corneal dystrophy (FECD) is a common disease of the corneal endothelium (CE) that is characterised by the formation of guttae, abnormal thickening of the Descemet's membrane (DM), and gradual loss of corneal endothelial cells (CECs). These features lead to corneal oedema, progressive vision loss, and ultimately blindness, if left untreated. Corneal grafting is the only effective treatment for FECD. In Australia, ~6% of corneal grafts are performed in FECD patients annually.

Currently, the mechanism of FECD is poorly understood. It occurs as a rare familial early-onset or a common late-onset familial or sporadic disease. This project focused on the late-onset disease that shows strong genetic heterogeneity, much of which remains undetermined.

Mutations in the *AGBL1*, *LOXHD1*, *SLC4A11*, and *ZEB1* genes contibute to the disease in Caucasian American and Asian populations. Four chromosomal loci are linked to familial disease but the underlying genes remain unknown. Strong genetic association between FECD and nucleotide variants in *TCF4* has been identified in Caucasian American, Chinese and Indian cases. Molecular studies suggest involvement of oxidative stress, apoptosis, protein misfolding, corneal endothelial ion transporters loss, and accelerated senescence in the disease pathophysiology.

This project aimed to advance the understanding of the molecular basis of FECD. The specific aims were to identify genetic causes of FECD in Caucasian Australian cases, determine relative abundance of proteins between FECD-affected and unaffected DM, and identify dysregulated genes between FECD-affected and unaffected CE.

Mutation screening of *LOXHD1, SLC4A11* and *ZEB1* genes in 128 cases by Ion Torrent Next-Generation Sequencing identified novel mutations in *LOXHD1* and *SLC4A11* as the possible causes of FECD in three cases. Genotyping of 208 cases for TGC repeat expansion in *TCF4* by short tandem repeat assay confirmed strong association with FECD.

Genome-wide association study (GWAS) using pooled DNA identified *TCF4*(*transcription factor 4)* as a major susceptibility gene for FECD in Australian cases. Genotyping of 190 cases for FECD associated potential novel genes by Sequenom MassArray Spectrometry replicated association of *ATP1B1, KANK4* and *LAMC1* genes with FECD. These genes were initially identified through GWAS performed in American cases by our collaborators in the USA. Quantitative reverse transcription-polymerase chain reaction (quantitative RT-PCR) indicated down-regulation of *ATP1B1* in FECD-affected CE compared to normal CE, and immunohistochemistry showed differential distribution of the encoded protein in diseased cornea compared to normal cornea.

Comparison of relative abundance of proteins between affected and unaffected DM by label-free mass spectrometry revealed reduced abundance of APOE in FECD-affected DM. Immunohistochemistry validated APOE as a component of the DM, with differential distribution between affected and unaffected corneas. Analysis of expression of APOE mRNA levels between FECD-affected and unaffected CE by quantitative RT-PCR revealed down-regulation of the transcript in the disease, suggesting a correlation between reduced abundance of the protein and under-expression of the gene in FECD.

Investigation of dysregulated genes between FECD-affected and unaffected CE by microarray analysis revealed 135 differentially expressed genes. Of these, 123 genes were up-regulated in the disease. Validation of 12 prioritised genes by quantitative RT-PCR confirmed significant up-regulation of *ALPK2*, *BGN*, *CLIC6*, *CPAMD8*, *CST1*, *CX3CR1*, *EDN1*, *HLA-DRA*, *NOX4* and *PPP1R1B* genes in FECD.

Ingenuity Pathway Analysis of all genes/proteins identified to be involved in FECD in this project and previous independent studies suggests involvement of activated transcription factor JUN (Jun Proto-Oncogene) and β-catenin/TCF4 complex, tumour suppressor p53, and immune response pathways in the mechanism of FECD.

Overall, this project revealed novel FECD-causing mutations in three cases, confirmed association of *TCF4* with FECD in Australian cases, and identified several new genes involved in FECD. Physiologically, the novel genes are involved in protein synthesis (*ALPK2* and *KANK4*) and assembly/structural integrity in the DM (*BGN* and *LAMC1*), electrochemical gradient homeostasis (*ATP1B1* and *CLIC6*), premature senescence (*CST1*), and oxidant-antioxidant imbalance (*APOE* and *NOX4*) pathways, known to be involved in FECD. The *CX3CR1*, *HLA-DRA* and *EDN1* genes are involved in immune response. Involvement of the immune response pathway in FECD is a novel finding.

PRESENTATIONS ARISING FROM THIS THESIS

Kuot A., Mills R., Snibson G., Burdon KP, Craig JE, Sharma S. Molecular biology and genetics of Fuchs' endothelial corneal dystrophy (Platform presentation). Australian Society for Medical Research SA Division: Annual Scientific Meeting 2013, Adelaide, June 05-06.

Kuot A, Mills R, Snibson G, Burdon KP, Craig JE, Sharma S. Molecular biology and genetics of Fuchs' endothelial corneal dystrophy (Platform presentation). 30th Australian & New Zealand Cornea Society Meeting 2013, Adelaide, March 07-08.

Kuot A, Mills R, Snibson G, Burdon KP, Craig JE, Sharma S. Differential gene expression analysis of human corneal endothelium in Fuchs' endothelial corneal dystrophy (Poster presentation). Australian Society for Medical Research SA Division: Annual Scientific Meeting 2012, Adelaide, June 04-06.

DECLARATION BY THE STUDENT

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed: Abraham K.A. Kuot **Date: 5th of June, 2015**

Abraham Kuot

ACKNOWLEDGEMENTS

This thesis is dedicated to three very special people who have immensely shaped my life, and made this academic endeavour a milestone to reach. These amazing and most humble individuals are my late inspirational father, Anyuon-titjok Kuot, Mr. Daniel Callery – a philanthropist who sacrificed every single cent of his own to ensure that I got learned, and thirdly, my lovely wife Rabecca Achol Malang.

In honour of your wisdom, vision, sacrifice, and unconditional love unto me and my brothers, I offer you, daddy, my sincere gratitude for all that I have achieved. I know that you are probably looking over my shoulder with a huge smile of approval. I wish you lived longer to witness your own investment in me.

To Daniel Callery, I have no a single word to describe your generosity, kindness, humility, and love towards me. You are a total package of humanity at its very best. How many people would offer their money to provide education to a strange, refugee kid seeking for better future? I am proud of you today, for I have made your generosity an investment that worths its best intention. Thank you so much!

Dear Rabecca, in your very humble and loving personality, I sincerely thank you so abundantly for your great understanding, unconditional love, and unshaken support throughout our eight years of marriage, and during my academic journey. I am truly blessed to marry you as my wife and a true companion; a gift of my life. I will always appreciate the sufferings you and our two lovely daughters, Ayuen and Abany, endured throughout this PhD journey. I promise to be there for you all the time.

To you my lovely Ayuen Kuot and Abany Kuot, Daddy worked so hard to ensure that you have a better future; and never the kind of life I endured during my childhood. The academic fulfilment that I acknowledge in this book is a testimony to what a human being can achieve regardless of the past misfortunes. It sets a bar for you, Ayuen and Abany, to be better citizens of the society; and as people the world could depend on. I will always be there to guide you through life challenges, as your late grandpa had done to me.

Regardless of the best building materials that can be available to any building, the foundation always remains the key determinant of the duration and structural integrity of that building. To you dear Dr. Shiwani Sharma and A/prof. Kathryn P Burdon, you have been and are the foundation to this academic milestone in my life. I am very lucky to have you as my supervisors; and I sincerely thank you from the bottom of my heart.

This journey started from the very time you offered me my first professional job in your laboratory. You tremendously helped and guided me through some of the most challenging times of my life. Whether in work-related challenges, sickness, and/or in personal life difficulties, you always steadfastly stand by my side. Your care and assistance towards my needs was wonderful and immeasurable. I am so much indebted to you for everything. Your kindness, tolerance, guidance, and passing of knowledge, which have brought me to this end, will often remain a pillar of reference to my personal life.

To Prof. Jamie Craig, I am very thankful for all who you are! I recall your beaming optimism and encouragement to carry out this project during our initial discussion on it with Dr. Sharma, A/Prof. Burdon, you and myself. You were right, for this project has become a success. You have been, and will always be a great boss, friend, and human being. You have been a source of inspiration to me; and a model of humility, great personality and a loving father. I sincerely thank you for the opportunity to be able to work with you, and for the friendship.

A/Prof. Richard Mills is another great person whose input and skills must be acknowledged. Thank you very much for your coordination and commitment in securing and providing me with surgical specimens. Without these samples, this project and its findings would have not been possible. Additionally, I sincerely thank A/Prof. Mills for financial assistance towards the extension of my scholarship. For without this valuabe assistance, I may not have excelled as much as I currently have.

I must equally acknowledge and appreciate the effort put into this project by Mrs. Margaret Philpott; the former manager of the Eye Bank of South Australia. Thank you so much for all you have done towards this project for its success. I also appreciate and sincerely acknowledge the contribution of Mr. Stephen Pulbrook, the current manager of the Eye Bank of South Australia, during this project.

Also, my sincere thanks and appreciations go to A/Prof. Sonja Klebe, Dr. Raymond Loh, Dr. Tiger Zhou, Dr. Alex Hewitt, Dr. Graham pollock, A/Prof. Grant Snibson, Ms Lisa Buckland and Dr. Mark Chehade for their great assistance in coordinating, retrieving and providing me with corneal surgical specimens that were integral to the

experimental work performed in this project. Thank you so much! Furthermore, my special thanks and appreciations go to Dr. Maurizio Ronci and Dr. Mark Corbett for their expertise and assistance in analysis of data from mass spectrometry and microarray analyses, respectively.

I am also greatly aware of the contribution of Prof. Keryn Williams towards this project. Dear Prof. Williams, thank you very much for your counselling, show of concern and understanding towards my personal misfortunes, which were linked to the family back home, and of your effort in helping to organise and secure financial assistance for me and my family, following the end of my scholarship funding. Your kind effort enabled me to successfully complete this project.

To all the past and present administrative and lab staff, and students at the Department of Ophthalmology, I offer you my sincere thanks and appreciations for your assistance, support, words of advice, and friendship. You always made my life easier and cheerful at work because of your great personalities and professionalism. In your very own names, I would say thank you Deb Sullivan, Lyn Harding, Cheri Gray, Anne Cazneaux, Miriam Keane, Emmanuelle Souzeau, Bronwyn Ridge, Sarah Martin, Mona Awadalla, Tiger Zhou, Alpana Dave, Lefta Leonardos, Shari Javadiyan, Georgia Kaidonis, David Dimasi, Kate Laurie, Lauren Mortimer, Madi Helm, Ash Hocking, Yazad Irani, Paul Badenoch, Kanchana Bandara, Binoy Appukuttan, Justine Smith, Liam Ashander, Shervi Lie, and Andrew Stempel.

Finally, I would like to thank all my extended family members of Pan-Anyuon in Australia. Without your dear support, I may not have achieved this academic

milestone. Thank you so much and maintain that unity and love among yourselves. To my special friend, Bol Ayach, and all the friends across the world, thank you so much for your support, friendship and well wishes.

ABBREVIATIONS

ABL	Anterior banded layer
ACTB	Beta actin
ACTBL2	Beta actin like protein 2
AGBL1	ATP/GTP binding protein-like 1
AGRF	Australian Genome Research Facility
AGRN	Agrin
ALB	Putative uncharacterised protein ALB
ALPK2	Alpha-kinase 2
ANOVA	Analysis of variance
APOD	Apolipoprotein D
APOE	Apolipoprotein E
ATP1B1	Na+/K+ ATPase, beta 1 subunit
BCA	Bicinchoninic acid assay
BGN	Biglycan
BL	Bowman's layer
bHLH	Basic helix-loop-helix
BMES	Blue Mountains Eye Study
bp	Base pair
BRG1	Brahma related gene 1
Caco-2	Colorectal carcinoma cell line
CDs	Corneal dystrophies
CDKN1A	Cyclin-dependent kinase inhibitor 1A
CDKN2A	Cyclin-dependent kinase inhibitor 2A
cDNA	Complementary DNA
CDSA	Corneal dystrophy South Australia
CE	Corneal endothelium
CE+DM	Corneal endothelium + Descemet's membrane
CFH	Complement Factor H
CHED1	Congenital hereditary endothelial dystrophy type 1
CHED2	Congenital hereditary endothelial dystrophy type 2

CHR	Chromosome
CI	Confidence interval
CLEC11A	C-type lectin domain family 11, member A
CLIC6	Chloride intracellular channel 6
CLU	Clusterin
COL8A2	Alpha-2 chain of collagen VIII
COL12A1	Alpha-1 chain of collagen XII
Conc.	Concentration
Corr	Correlation factor
COX2	Cyclooxygenase 2
CPAMD8	C3 and PZP-like, alpha-2-macroglobulin domain conatining 8
cRNA	Complementary RNA
CSF1R	Colony stimulating factor 1 receptor
CST1	Cystatin SN
C_T	Threshold cycle
CUG-BP1	CUG triplet repeat RNA-binding protein 1
CV	Covariance
CX3CR1	Chemokine (C-X3-C Motif) receptor 1
C9	Complement component 9
DE	Differentially expressed
DICER1	Dicer 1, ribonuclease Type III
DM	Descemet's membrane
DMD	Duchene muscular dystrophy
DMPK	Dystrophia myotonica protein kinase
DNA	Deoxyribonuleic acid
DRAM1	DNA-damaged regulated autophagy modulator 1
DSAEK	Descemet's stripping Automated Endothelial Keratoplasty
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
E	Amplification efficiency
EDN1	Endothelin 1
EDTA	Ethylenediamine tetraacetatic acid
EMT	Epithelial-to-mesenchymal transition

EMRT	Exact mass retention time
ENMT	Endothelial-mesenchymal transition
EPI	Epithelium
ERAD	Endoplasmic reticulum-associated degradation
ERBB2	V-Erb-B2 avian erythroblasttic leukemia viral oncogene homolog
ESI	Electrospray injection
ESR1	Estrogen receptor 1
ETS1	V-Ets avian erythroblastosis virus E26 oncogene homolog
E2-2	E protein E2-2 (TCF4)
F	Frequency
FAM3C	Family with sequence similarity 3, member C
FBLN1	Fibulin 1
FCD1	Fuchs' corneal dystrophy locus 1
FCD2	Fuchs' corneal dystrophy locus 2
FCD3	Fuchs' corneal dystrophy locus 3
FCD4	Fuchs' corneal dystrophy locus 4
FECD	Fuchs' endothelial corneal dystrophy
FKBP4	Fk506 binding protein 4, 59kDa
fmol	Femtomole
FMR1	*F*ragile X mental retardation 1
FN1	Fibrillin-1
FPR	False positive rate
FRDA	Friedreich ataxia
FXS	Fragile X syndrome
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
gDNA	Genomic DNA
GWAS	Genome-wide association study
HD	Huntington's disease
HEK293	Human Embryonic Kidney 293 cell line
Hg19	Human Genome sequence version 19
HLA-DRA	Major histocompatibility complex, class II, DR alpha
HMBS	Hydroxymethylbilane synthase
HPRT1	Hypoxanthine phosphoribosyltransferase 1

HSPGs	Heparan-sulfate proteoglycans
HST1H4A	Histone H4
HTRA1	Serine protease HTRA1
HTT	Huntingtin
Hz	Hertz's
IAA	Iodoacetamide
IG	Individual genotyping
IGHG1	Ig gamma 1 chain C region
IGHG2	Ig gamma 2 chain C region
IPA	Ingenuity pathway analysis
ISPs	Ion Sphere^{TM} particles
iTRAQ	Isobaric tags for relative and absolute quantification
JUN	Jun proto-oncogene
KANK4	KN Motif And Ankyrin Repeat Domains 4
kb	Kilobase
kDa	Kilodalton
KEGG	Kyoto encyclopaedia of genomes of genes
KERA	Keratocan
KRT1	Keratin type I cytoskeleton 1
KRT9	Keratin type I cytoskeleton 9
KRT10	Keratin type I cytoskeleton 10
LAMC1	Laminin, gamma 1 subunit
LC-MSE	Liquid chromatogram- mass spectrometry/mass specrtometry
LD	Linkage disequilibrium
LDL	Low density lipoprotein
LOXHD1	Lipoxygenase homology domain 1
MAF	Minor allele frequency
MBNL1	Muscleblind-like 1
MCF7	Human breast cancer cell line
mg	Milligram
MIQE	Minimum information for publication of quantitative RT-PCR experiments
ml	Millilitre
MQH_2O	MilliQ water

mRNA	Mesenger RNA
MS/MS	Tandem mass spectrometry
MT-CO2	Mitochondrially encoded cytochrome c oxidase II
mtDNA	Mitochondrial DNA
MYC	V-Myc avain myelocytomatosis viral oncogene homolog
mz	Mass-to-charge ratio
Na-K^+ ATPase	Sodium-potassium adenosine triphosphase
NCBI	National Centre for Biotechnology Information
NE	Normalised expression
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next generation sequencing
NHMRC	National Health and Medical Research Council
nm	Nanometre
MNE	Mean normalised expression
nmol	Nanomole
NOX4	NADPH oxidase 4
NSA	Normal South Australian
NTC	No template control
nUPLC-MSE	Nano-ultra pressure liquid chromatography-mass/mass spectrometry
OR	Odds ratio
ORF	Open reading frame
PAF	Pooling allele frequency
PBL	Posterior banded layer
PCA	Principal component ananlysis
PCL	Posterior collagenous layet
PCR	Polymerase chain reaction
PDDC1	Parkinson Disease 7 Domain Conatining 1
pg	Picogram
PGE2	Eicosanoid prostaglandin E2
PGM™	Personal Genome Machine™
PLAT	Polycystin-1, lipoxygenase, alpha-toxin
PLGS	ProteinLynx Global Server
PNBL	Posterior non-banded layer

PNPLA2	Patatin-like phospholipase Domain Containing 2
PolyPhen2	Polymorphism Phenotyping
POTEI	POTE ankyrin domain family member 1
PPCD	Posterior polymorphous corneal dystrophy
PRDX2	Peroxiredoxin 2
PRDX3	Peroxiredoxin 3
PRDX5	Peroxiredoxin 5
PTPRG	Protein tyrosine phosphatase, receptor type, G
PVDF	Polyvinylidine fluoride (PVDF) membrane
PPP1R1B	Protein phosphatase 1, regulatory (inhibitor) subunit 1B
QIMR	Queensland Institute of Medical Research
QLD	Queensland
Q RT-PCR	Qutantitative reverse transcription-polymerase chain reaction
R	Ratio
REFGEN	Reference genome
REFSEQ	Reference sequence
RIPA	Radio-immunoprecipitation assay
RNA	Ribonucleic acid
RS	Replication sequencing
RT+	Standard cDNA samples with reverse transcriptase
RT-	Standard cDNA sample without reverse transcriptase
RT-PCR	Reverse transcription-polymerase chain reaction
RVEEH	Royal Victorian Eye and Ear Hospital
S	Stroma
SAGE	Serial Analysis of Gene Expression
SAP	Shrimp Alkaline Phosphatase
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	Standard error of mean
SIFT	Sorting Intolerant From Tolerant program
SLC4A11	Sodium-borate co-transporter 1 protein
SLC22A5	Solute carrier family 22 (organic action/carnitine transporter), member 5

SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily B, member 1
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4
SNP(s)	Single nucleotide polymorphism(s)
SPP1	Secreted phosphoprotein 1
STD	Standard deviation
STRs	Short tandem repeats
TBE	Tris-borate-EDTA
TBS	Tris buffered saline
TBST	Tris buffered saline containing 0.1% Tween 20
TCF4	Transcription factor 4
TE	Tris-EDTA
TGC	Thymine-Guanine-Cytosine trinucleotide
TGFBI	Transforming growth factor beta-induced
TGF-β	Transforming growth factor-beta
TNRs	Trinucleotide repeats
TP53	Tumour suppressor 53
TREDS	Trinucleotide repeat expansion diseases
UBC	Ubiquitin c
UPLC	Ultra performance liquid chromatography
UTRs	Untranslated regions
VCFs	Variant called files
XECD	X-linked endothelial corneal dystrophy
ZEB1	Zinc finger E-box binding homeobox 1
µm	Micrometre
µl	Microlitre
µg	Microgram

CHAPTER 1
Introduction

1.1 STRUCTURE OF THE HUMAN EYE

The eye is the organ of vision. It is structurally made up of three main layers: outer, middle and inner [1]. The outer layer is composed of avascular connective tissue consisting of sclera and cornea (Figure 1.1).

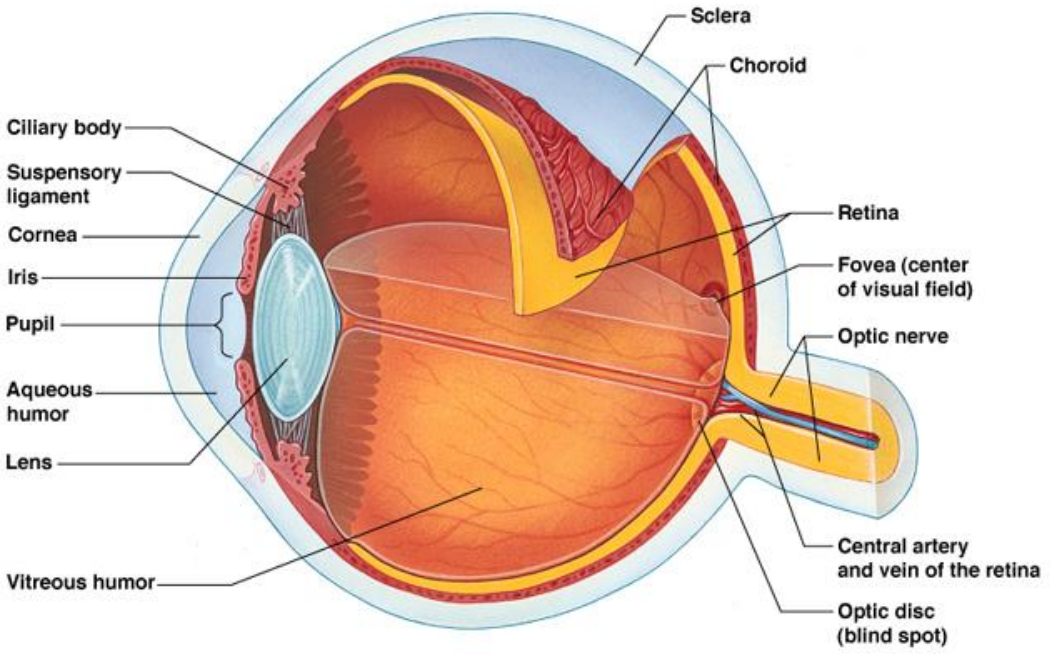


Figure 1.1 Schematic diagram of the structures of the human eye (Reece and Campbell, 2002, 6th Edition, P.1064)

The sclera is a tough white sheet of fibrous tissue that protects, and gives shape to the eye. The cornea is a modified central part of the sclera [2]. It forms a transparent front window through which light passes into the eye [3].

The middle layer consists of the lens, aqueous humour, vitreous humour, choroid, iris and pupil [1]. The pupil regulates the amount of light entering the eye.

The lens, a transparent double convex structure, focuses light rays onto the retina [1]. The aqueous and vitreous humours surround the lens. They are contained in the anterior chamber and the vitreous chamber, respectively.

The anterior chamber is the space between the cornea and the lens, whereas the vitreous chamber is the space between the lens and the retina. The aqueous humour nourishes corneal and lens cells while the vitreous humour nourishes the lens and the retina, and provides structural support to the eye. The choroid, which lies between the retina and the sclera, prevents light from scattering, and consists of blood vessels that provide nourishment to the outer layers of the retina [1].

The inner layer is the retina; located at the back of the eye [1]. Compared to a camera, the retina is the film or detector on which the images are projected. It contains millions of photoreceptor cells, which sense light that enters the eye, and send visual stimulus to the brain through the optic nerve, for interpretation. The present project relates to the cornea in the eye.

1.2 THE CORNEA

The cornea consists of five anatomical layers (epithelium, Bowman's layer, stroma, Descemet's membrane and endothelium; Figure 1.2, page 3). It is derived from the surface ectoderm during embryologic development of the eye [4]. The surface ectoderm produces the corneal epithelium. Ectodermal neural crest-derived mesenchymal cells form the corneal stroma and endothelium [5, 6]. Clarity and integrity of the cornea is essential for vision. The corneal layers are structurally formed to provide these properties.

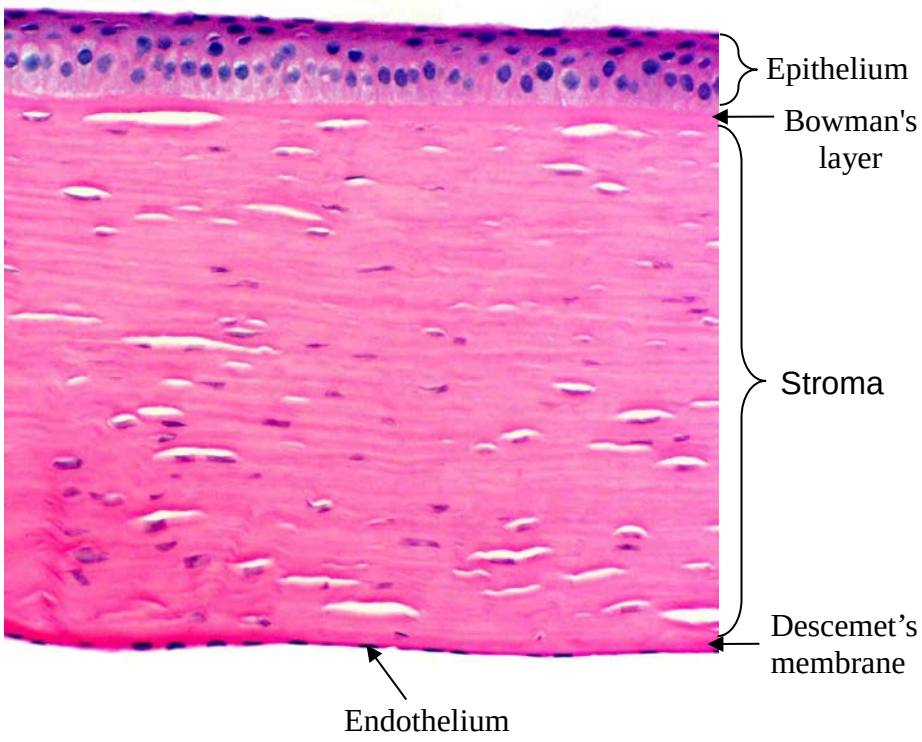


Figure 1.2 Anatomy of the human cornea. Eosin and Haemotoxylin stained section of the normal cornea (Image courtesy of A/Prof. Sonja Klebe, Anatomical Pathology, Flinders University, Adelaide, SA, Australia)

Corneal epithelium is the outermost layer and is made up of multiple layers of epithelial cells [7]. These epithelial cells are a unique type of non-keratinised cells arranged in an orderly fashion. This is essential for maintaining transparency of the cornea for penetration of light for vision. In addition, they are continuously replenished throughout an individual's lifespan by limbal stem cells present in the limbus [8, 9] to preserve corneal clarity, and contribute to general protection of the eye from injury from external foreign objects such as pathogens and dust [2].

The Bowman's layer is the epithelial basement membrane that is composed of collagen and other extracellular matrix proteins [4]. Its components protect the corneal stroma. It can be described as an acellular, condensed region of the apical stroma, composed primarily of randomly organised yet tightly woven collagen fibrils.

Stroma is the middle layer that makes up about 90% of the thickness of the cornea [10]. It consists of highly ordered collagen fibres, which keep the stroma and thus the cornea, transparent [4]; to allow optimal penetration of light to the lens. The stroma also contains keratocytes, which are specialised fibroblasts [11]. Keratocytes are interconnected and sparsely distributed to maintain the stroma, and presumably its transparency. They are also involved in wound healing and collagen formation. Activation of keratocytes in response to injury has been suggested to cause corneal haze [12].

Descemet's membrane (DM) is a specialised basement membrane of the corneal endothelium [13, 14]. It is made up of extracellular matrix proteins [14]. The DM is made up of two distinct layers, the anterior and the posterior [14]. The anterior DM is called anterior banded layer (ABL). It consists of banded collagens and is laid down *in utero* [14]. The posterior DM, normally referred to as a posterior non-banded collagen layer (PNBL), is progressively secreted by the corneal endothelium throughout life [15]. The DM is approximately, 5-20μm thick, depending on age [15].

Corneal endothelium is a monolayer of endothelial squamous cells that forms a physical barrier between the corneal stroma and aqueous humour [3]. Its main function is to preserve stromal deturgescence (relative dehydration of the cornea),

which is a prerequisite for corneal transparency [16]. The endothelium achieves stromal deturgescence through its barrier and ion pump functions [16]. It primarily uses Na-K^+ ATPase pump to pump excess fluid from the stroma across the DM into the anterior chamber [16, 17]. Simultaneously, barrier integrity of the corneal endothelium facilitates passive flow of aqueous humour into the stroma to provide nutrients [18, 19]. The barrier integrity is conferred by tight and adherens junctions of the corneal endothelial cells [20]. In a healthy cornea, fluid leak is counter-balanced by an active fluid pump mechanism, resulting in an overall mechanism that is referred to as a Pump-Leak mechanism [21, 22].

The human corneal endothelium does not contain mitotically active cells [23]; therefore the cells do not normally replicate *in vivo*. This lack of cell division results in a gradual decrease in endothelial cell density throughout life, with an average cell loss of 0.3 – 0.6% per year [24-26]. The endothelium normally has a sufficiently large reserve of cells to maintain its barrier and pump functions for a lifetime [27]. However, significant loss of endothelial cells due to increasing age [27, 28], accidental or surgical trauma [29], previous corneal transplantation [30], stress caused by diseases such as diabetes [31], treatments for glaucoma [32, 33], and endothelial corneal dystrophies [34, 35] can reduce cell density to a critical point below which the stroma becomes oedematous, and there is subsequent reduction in corneal clarity [36]. Excess stromal hydration leads to light scattering which impairs vision. Corneal dystrophies are conditions in which one or more parts of the cornea lose their normal clarity due to abnormal build-up of extraneous materials, such as lipid crystals in the cornea [37].

1.3 CORNEAL DYSTROPHIES

1.3.1 Definition and types

Corneal dystrophies (CDs) are a group of heterogeneous, bilateral, genetically determined non-inflammatory diseases of the cornea [10]. These diseases vary in their signs and symptoms, clinical severity, and genetic causes [37]. Corneal dystrophies are clinically divided into anterior, stromal, and posterior dystrophies based on the predominant layer of the cornea involved [10].

The anterior CDs affect the corneal epithelium, Bowman's layer, and superficial corneal stroma. Examples include Meesman dystrophy and Thiel-behnke dystrophy [38]. Stromal CDs affect the stroma and they include macular corneal dystrophy and Lattice corneal dystrophy type I and variants [38]. Posterior CDs are characterized by abnormalities of the corneal endothelium and DM [10]. This group includes Fuchs' endothelial corneal dystrophy, posterior polymorphous corneal dystrophy (PPCD), congenital hereditary endothelial dystrophy type 1 and 2, and X-linked endothelial corneal dystrophy [37]. In most posterior CDs, a defective active fluid transport by the corneal endothelium causes excess oedema of the corneal stroma that leads to corneal opacity, and ultimately impaired vision [37].

1.3.2 Genetic causes

Corneal dystrophies are often inherited as Mendelian traits [39]. However, these diseases have varied aetiologies. Twenty four types of CDs have been defined to date (Appendix, Table 1, page 210). Most of these CDs are caused by different mutations in up to thirteen known genes. In addition, 18 loci have been mapped but the genes responsible for the diseases have not yet been identified (Apendix, Table 1). Reported studies indicate that different mutations in the same gene can cause different CD phenotypes [10]. Alternatively, mutations in multiple genes can give

rise to the same CD phenotype [10]. A similar scenario is seen for many of the chromosomal loci onto which one or different CD phenotypes are mapped [40, 41]. For example, different mutations in the *transforming growth factor beta-induced* (*TGFBI*) gene cause two different anterior (Reis-Bücklers and Thiel-Behnke) and three different stromal (granular Type I, granular type II and Lattice type I and variants) dystrophies [41]. The *TGFBI* gene encodes for keratoepithelin – an adhesion protein found in the corneal stroma and secreted by the corneal epithelium [42]. Extracellular deposition of abnormally processed mutant TGFBI aggregates within the cornea results in corneal opacity [10, 43]. In addition, dysregulation of *TGFBI* transcript in the corneal endothelium, and of the protein in the endothelium and Descemet's membrane has been observed in Fuchs' endothelial corneal dystrophy in multiple independent studies. Fuchs' endothelial corneal dystrophy was the focus of the present project, and is thus described in more details in the next section.

1.4 FUCHS' ENDOTHELIAL CORNEAL DYSTROPHY

Fuchs' endothelial corenal dystrophy (FECD) is a common bilateral and progressive disease of the endothelial layer of the cornea. It exists in two forms: rare early-onset and late-onset form [39, 44]. The early-onset form is familial and presents from birth or early childhood and reaches advanced stages between the ages of 10 and 20 years [45]. Late-onset FECD is the most common form and its onset is typically after 40 years of age [37]. It occurs either as a familial or a sporadic disease [37].

Much of the genetic cause of FECD is yet undetermined. The earliest clinical feature of the disease is the development of guttae on the DM [35]. Other clinical features include abnormally thickened DM, gradual and progressive degeneration and loss of

endothelial cells [45], and functuation in vision, which is typically worse in the early morning and improves during the day [35]. The remaining living endothelial cells enlarge to restore severed tight cell junctions [35]. This leads to abnormal endothelial cell morphology and subsequently further degeneration or loss of the enlarged cells [45]. The loss of endothelial cells and aberrant DM thickening result in impaired corneal endothelium function [46]. These factors lead to excess corneal stromal hydration, which causes a cloudy cornea and subsequent vision impairment (Figure 1.3). Generally, FECD remains asymptomatic until such a time when the disease has reached the advanced stage [10]. Advanced FECD is characterised by severe pain and significant vision impairment, which ultimately leads to vision loss if the condition remains untreated. These conditions demand treatment to improve vision and relieve pain.

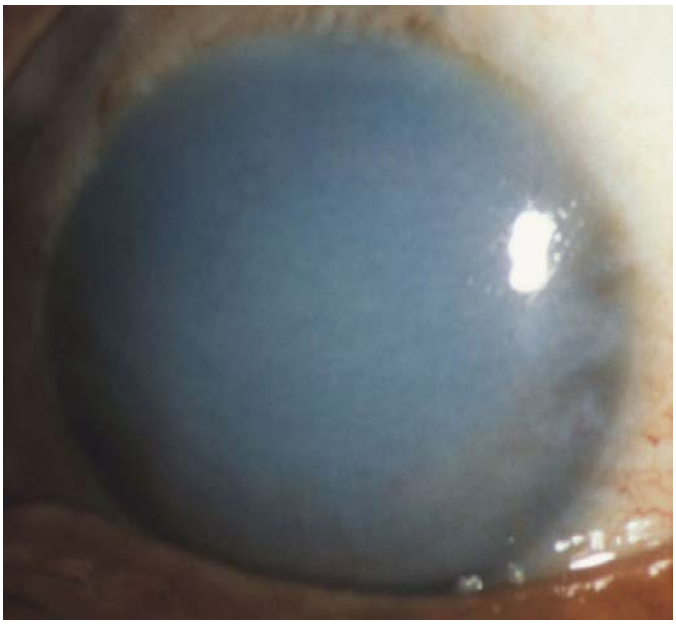


Figure 1.3 The cornea of a FECD-affected patient at an advanced stage of the disease. Adapted from Klintworth et al. [10].

1.4.1 Treatment of FECD

Corneal transplantation is currently the only effective treatment for advanced FECD [37]. Corneal transplantation in FECD patients accounts for ~ 6% of corneal grafts performed in Australia annually [47]. This disease is the second most common indication for corneal transplantation in the United States [39]. However, the treatment has some limitations. The graft can fail or lead to post-surgical complications [47, 48]. Furthermore, corneal grafts in patients with the disease last for about 15 years on an average [49]. The graft recipients also need life-long follow-up [50]. This is because visual rehabilitation following transplantation may take several years, and adjustments continue to be made for as long as 10 years [47] to manage post-graft events such as primary graft failure, infection, rejection, endothelial cell failure, and post-operative rise in intraocular pressure due to trauma [51].

Additional issues concerning the management of FECD include late diagnosis and lack of alternative forms of treatment. Collectively, these limitations call for development of alternative techniques or methods to prevent, manage and/or treat the disease. Achieving the latter depends on better understanding of the molecular basis of FECD.

1.4.2 Epidemiological studies in FECD

To date, no formal epidemiological studies have been performed to assess the prevalence or incidence of FECD. However, existing corneal transplant data indicate a disparate worldwide distribution of the disease, with Western regions such as the United States, Canada, Australia, and Europe reporting more cases than the Middle East, Asia, Oceania, and South America [10, 45, 47, 52-60]. The disparity may be due to the fact that corneal transplantation is more readily available in the developed

countries than the undeveloped or developing countries. In addition, women are disproportionately more frequently affected by FECD compared to men, at a female:male ratio of 3:1 [10, 61-64].

FECD has been shown to affect ~ 4% of the American population over the age of 40 years [62]. Based on existing corneal transplant data, similar percentages of FECD patients from America and Australia, or the Western countries, underwent corneal transplantation, suggesting approximately similar prevalence of the disease in these countries [55, 56, 61, 65]. However, the disease was observed to be prevalent in up to 11% of the inhabitants from the Tangier Island in Virginia, USA [66]. The population on this island is genetically isolated [66], therefore the residents might be more genetically homogenous than the mainland populations and hence the higher prevalence.

The reasons behind wide variations in prevalence of FECD across the world have not been established. This is partly because no formal epidemiological studies have been conducted to examine the disease prevalence on a large scale. Many investigators encountered difficulties when performing these studies due to the late onset of the disease. Due to late manifestation of FECD many patients remain undiagnosed, which potentially underestimates the prevalence of the disease [64]. Currently, only three studies have been reported on the heritability of FECD [62, 64, 67]. The first was a large heritability study of the disease in a targeted enriched sample from the United States [64], which shows that the clinical phenotype of guttae, and severe disease, are highly heritable (heritability estimates = 0.304 for severe disease; p=0.0043, and 0.368 for FECD score for the worse eye; p=0.0002) in Caucasian population. The heritability estimates in this study were determined by calculating

sibling-sibling correlations for co-inheritance of guttae and severe FECD. The heritability index, which ranges from 0 to 1, is a common estimate of heritability. The index of 0 indicates no co-inheritance and 1.0 indicates co-inheritance. The reported correlation in heritability between guttae and FECD severity is likely due to genetic predisposition, which makes Caucasians more susceptible to FECD [68].

Currently, genetic predisposition is thought to be the single most plausible factor behind the pathogenesis of FECD [37]. Furthermore, various pieces of evidence suggest that interactions between genetic factors and environmental stimuli such as oxidative stress in the corneal endothelium or geographical location, and aging related physiological processes may be contributing to the disease [55, 60, 64, 69, 70]. However there is no clear evidence for the involvement of these factors in the pathogenesis of the disease. To elucidate the molecular basis of FECD, researchers have focused their efforts on genetic and molecular studies in the disease.

1.4.3 Genetic studies in FECD

As mentioned, genetic factors are major causes of FECD [62, 67]. Therefore, identification of individual genes that contribute to the disease is paramount to the understanding of its molecular basis. This can be achieved by employing a candidate gene approach or genetic association studies.

1.4.3.1 Conventional genetic studies

In 2001, Biswas *et al* [44] successfully employed a positional candidate strategy to identify *collagen type VIII alpha 2* (*COL8A2*) as the candidate gene for FECD in a multigenerational American family with a highly penetrant, early-onset form of the disease. Mutation screening of the *COL8A2* gene for potential disease-causing mutations in members from this family revealed presence of the Gln455Lys mutation in only the affected members [44]. Subsequently, this mutation was also found to be

responsible for the disease in two additional families with early-onset FECD from Australia and Northern England [44], and also in Korean patients [71]. Further sequencing of this gene in Caucasian American cases of early-onset FECD in an independent study revealed a novel mutation, Leu450Trp, to be causative [72].

COL8A2 encodes the alpha-2 chain of collagen VIII (COL8A2) protein [44]. This protein is a major component of the DM, and is present in the abnormal posterior collagenous layer of the DM in endothelial disorders including FECD [73, 74]. Gottsch et al [73] demonstrated that abnormal secretion of COL8A2 by the endothelium in the patients with mutant *COL8A2* contributes to abnormal thickening of DM in the disease, a common clinical characteristic in the early- and late-onset FECD. Thus, *COL8A2* was considered as a candidate gene for the causation of late-onset disease.

Screening of 115 Caucasian Americans with unrelated late-onset disease for the Gln455Lys and Gln450Trp mutations revealed no association between this phenotype and the mutations [44]. Additional independent studies performed in 15 Korean probands, and 62 Caucasian American probands with familial late-onset FECD did not consistently identify any of these mutations in the screened cases [72, 75]. Based on this evidence, it was concluded that mutations in the *COL8A2* gene cause only early-onset FECD [44, 71, 72].

In a quest to identify genes responsible for late-onset FECD, various candidate gene hypotheses were investigated, and genome-wide linkage analyses were performed in large multi-generational families with late-onset disease from the United States. These pedigrees exhibited dominant inheritance with incomplete penetrance and variable expressivity. As a result, four genes, *solute carrier family 4, sodium borate*

transporter, member 11 (*SLC4A11*), *zinc finger E-box binding homeobox 1* (*ZEB1*), *lipoxygenase homology domain 1* (*LOXHD1*) and *ATP/GTP binding protein-like 1* (*AGBL1*) on chromosome 20p13, 10p11.22, 18q21.1 and 15q25.3 loci, respectively, were identified to be responsible for late-onset FECD [39, 76-79]. In addition, linkage analyses have localised the disease to four additional loci: *FCD1*, *FCD2*, *FCD3*, and *FCD4* on chromosomes 13, 18, 5, and 9, respectively [77, 80-82]. However, the genes involved at these loci are still unknown.

ZEB1 and *SLC4A11* encode for transcription factor 8 (TCF8) and sodium borate co-transporter (SLC4A11) proteins, respectively [77, 83]. TCF8 is an E-box binding transcription activator or repressor involved in proliferation, differentiation and development of various tissues [84, 85]. It acts as a master regulator of epithelial-to-mesenchymal transition (EMT) [86]. SLC4A11 is involved in ion channel transport of fluid across the corneal endothelium [76].

Disease-causing mutations in the *SLC4A11* gene were found in 4 out of 89 patients with late-onset FECD in one study from India [76]. Subsequent screening of 192 sporadic and familial cases of American origin for disease-causing mutations in *SLC4A11* revealed 7 additional mutations in 7 cases [83]. Familial data available for one of these mutations showed segregation under a dominant model in a three-generational family [83]. The most recent study identified three novel mutations in the gene in 5 out of 45 Indian-Japanese with sporadic late-onset disease [87]. Collectively, FECD-causing mutations in the *SLC4A11* gene have been shown to contribute to the pathogenesis of late-onset disease in between 2 – 4% of cases [76, 83, 87].

Similarly, FECD-causing mutations identified in *ZEB1* reportedly contribute to the disease in a similar proportion of cases. Riazuddin et al [39] identified five missense mutations in *ZEB1* in 7 of 384 unrelated patients from the United States. This finding suggests that *ZEB1* contributes ~ 2% to the genetic load of late-onset FECD. One of the seven patients bearing the identified mutation belonged to the affected large multi-generational family [39]. This patient carried the p.Q840P mutant allele. Segregation analysis showed that the mutant allele was present in only the affected family members (7/12), and was transmitted in an autosomal-dominant fashion [39]. Thus, this study revealed the first evidence for potential involvement of *ZEB1* in the development of sporadic and familial late-onset FECD.

The p.Q840P mutant allele was found to be sufficient but not necessary for pathogenesis of the disease [39]. Further investigation in the Caucasian American family with late-onset disease revealed linkage to the *FCD4* locus, and its interaction with the *ZEB1* gene [39]. This interaction was found to correlate with increased severity of the disease in those harbouring the p.Q840P allele [39].

The most recent studies reported that rare mutations in the *LOXHD1* and *AGBL1* genes have a causal role in familial and sporadic late-onset FECD [78, 79]. Whereas *LOXHD1* encodes for a membrane protein (lipoxygenase homology domain-containing protein 1; LOXHD1), the *AGBL1* gene codes a predominantly cytoplasmic glutamate decarboxylase, the ATP/GTP binding protein-like 1 protein. Both genes were considered as candidates for FECD after they were mapped to chromosomal loci *FCD2* (*LOXHD1*) and 15q25.3 (*AGBL1*) [78, 79].

Through next-generation sequencing, heterozygous missense mutations in *LOXHD1* were identified in only 1 out of 3 late-onset FECD-affected small nuclear families in

the USA [78]. Subsequent interrogation of a cohort of 207 Caucasian Americans with sporadic and familial late-onset disease by direct sequencing of all coding exons of *LOXHD1* revealed an additional 14 heterozygous missense mutations in 15 unrelated cases [78]. FECD-causing mutations in *AGBL1* were also identified for the first time in only one affected family, and in an additional 2 of the 207 individuals from the same cohort screened for mutation in the *LOXHD1* gene [79]. Further studies are required in non-American populations to confirm these findings, and ascertain the proportion of cases affected by mutations in the two genes.

In summary, all the mutations identified in the *SLC4A11* and *ZEB1* genes are collectively responsible for only a small number of FECD cases. Additionally, the involvement of *LOXHD1* and *AGBL1* in the pathogenesis of the disease has been reported for the first time only in one family each, and in approximately 7% and 1%, respectively, of the sporadic cases [78, 79]. All these findings show that FECD is a multigenic, complex disease with much of the genetic cause remaining undetermined. Therefore, further studies are required to identify novel genetic factors associated with the disease.

1.4.3.2 Genetic association studies

Conventional genetic studies often do not reveal much of the genetic determinants of complex diseases, such as FECD. In addition, the use of genome-wide linkage analysis to detect the small gene effects often seen in complex diseases is limited by rarity of large pedigrees as in the case of FECD [88, 89]. Thus, an alternative approach is needed.

Genome-wide association studies (GWAS) are an effective approach for finding the genetic determinants of complex traits or diseases [90]. Single nucleotide

polymorphism (SNP) is the genetic marker of choice since it is the commonest form of variation found within the genome and current technology permits economical genotyping of millions of SNPs in a single experiment [91]. A previous study by Baratz and colleagues identified four SNPs (rs613872, rs9954153, rs2286812, and rs17595731) in *TCF4* to be independently associated with late-onset FECD in a cohort of 130 Caucasian American cases [92]. They confirmed these findings in an independent cohort of 150 cases [92]. Individual associated SNPs and their odds ratios (ORs) are as follow: rs613872 ($p=4.25 \times 10^{-10}$, OR = 4.220), rs9954152 ($p=2.18 \times 10^{-6}$, OR = 2.79), rs2286812 ($p=4.13 \times 10^{-6}$, OR = 2.80) and rs17595731 ($p=2.97 \times 10^{-6}$, OR = 7.21) [92]. The associated SNPs increase the risk of developing the disease in the examined population by at least 2.8 times in carriers compared to non-carriers. We replicated these findings in a cohort of 103 Caucasian Australian cases: rs613872 ($p=5.25 \times 10^{-15}$, OR = 4.05), rs9954153 ($p=3.37 \times 10^{-7}$, OR = 2.58), rs2286812 ($p=4.23 \times 10^{-6}$, OR = 2.55) and rs17595731 ($p=3.57 \times 10^{-5}$, OR = 3.79) [93]. Our findings suggest that *TCF4* also contributes to FECD pathogenesis in Caucasian Australians [93].

Subsequent independent candidate gene association studies and additional GWAS confirmed association between SNPs in *TCF4* and late-onset disease in Caucasian American and Chinese populations [94-100]. All the studies have established *TCF4* as a major genetic factor for the susceptibility of developing late-onset FECD. Thus, more focus has been placed on this gene by researchers in order to understand contribution of the gene to the molecular basis of the disease.

Breschel *et al* [101] described the presence of an unstable trinucleotide repeat, TGC, within the third intron of the *TCF4* gene. They noted expanded, unstable alleles in

approximately 3% of Caucasian Americans [101]. In 2012, Wieben and colleagues [102] investigated the association of this trinucleotide repeat expansion with FECD in 66 late-onset FECD cases and 63 matched control subjects from the United States. They showed that 52 of the cases (79%) had expanded alleles [102]. Only 2 of the 63 control individuals (3%) showed evidence of an expanded allele [102]. Wieben et al's [102] findings suggest that FECD is a trinucleotide repeat expansion disorder in the majority of cases. The trinucleotide repeat expansion disorder is a disesase caused by pathogenic expansion of three-nucleotide repeat sequences, which are forms of short tandem repeats [103, 104].

Subsequently, multiple studies from the United States, China, and India have confirmed significant associations between FECD and the expanded TGC repeat in *TCF4* [99, 105, 106]. Furthermore, two of these studies have revealed a linkage of the expanded TGC repeat allele to the rs613872 SNP identified by GWAS [102, 105]. It has been reported that a repeat length longer than 150 nucleotides is highly predictive of FECD and suggested that the TGC repeat is a prime cause for pathogenesis in this degenerative disease [102, 107]. The most recent findings from *in vitro* studies by Du et al [108] have shown that expansion of the TGC repeat in the *TCF4* gene contributes to FECD through RNA toxicity and mis-splicing of the gene.

TCF4 encodes the E2-2 protein, which is a member of the ubiquitously expressed class I basic helix-loop-helix (bHLH) transcription factors that are involved in cellular growth and differentiation [109, 110]. E2-2 is expressed in developing corneal endothelium [109]. It plays an essential role in developmental processes including EMT, a normal aspect of cell migration in embryogenesis and tumour-cell invasion and metastasis [111]. The E2-2 binds to E-box promoter sequences within

target genes and either suppresses or activates tissue-specific gene transcription [109]. It also represses expression of cell-adhesion protein E-cadherin, thereby causing loss of cellular polarity and cell-to-cell contact and hence EMT.

ZEB1 protein, which also binds to E-box promoters and is involved in EMT through E-cadherin repression [112], is upregulated by E2-2. Rare pathological mutations in *ZEB1* contribute to late-onset FECD in Caucasians [39]. These observations suggest that the *TCF4* variants associated with FECD could confer the disease risk by altering the expression of *ZEB1*.

In our study of genetic analysis of candidate genes for FECD, in addition to association of SNPs in *TCF4*, we revealed significant and nominal association of tag SNPs, respectively, in the *TGFBI* and *clusterin* (*CLU*) genes with late-onset disease in Caucasian Australians [93]. The tag SNP is a representative genetic variant of certain groups of haplotypes in a region of the genome with high linkage disequilibrium. It is used to identify genetic variation in the gene/s of interest between cases and control individuals for a particular trait. The *TGFBI* and *CLU* genes were chosen because their functions were related to the pathophysiology of FECD and the encoded proteins were demonstrated to be overexpressed in the diseased corneal endothelium and DM [113, 114]. Analysis of tag SNPs in and around *CLU* revealed a significant association of rs17466684 ($p=0.003$, OR = 1.85) with advanced disease in the Caucasian Australians [93]. This was the first report of genetic association of *CLU* with FECD. Genetic analysis of five tag SNPs in *TGFBI* did not detect any association of individual SNPs with FECD, although haplotype analysis revealed that the TAAAT haplotype is significantly associated with carriers twice as likely to develop the disease ($p=0.011$, OR = 2.29) [93]. The TAAAT

haplotype possibly tags FECD-associated variants in *TGFBI* not tagged by the individual SNPs.

The most recent study by Li et al [115] has revealed nominal association of mitochondrial polymorphism A10398G (p=0.034, OR = 0.72) and Haplogroup I (1719A, 7028T, 8251A, 10398G, and 16391A) (p=0.041, OR = 0.46) with FECD in Caucasian Americans. The minor allele 10398G and Haplogroup I variant decrease the risk of disease, an effect that was strengthened when the effect of cigarette smoking was controlled [115]. This study is the first to suggest that the A10398G polymorphism and Haplogroup I decrease the risk of developing FECD; particularly in patients of European ancestry. However, additional independent studies are required to support the contribution of mitochondrial DNA (mtDNA) variants to susceptibility of the disease.

Overall, genetic association studies have identified additional genetic factors associated with the pathogenesis of late-onset FECD. However, variations in *TCF4* are associated with FECD in ~ 30% of the cases [92, 94, 97, 99]. Additionally, genetic association of *CLU*, *TGFBI*, and mitochondrial polymorphism A10398G and Haplogroup I with FECD need independent replication in other populations [93, 115].

1.4.4 Molecular studies in FECD

In order to understand how identified mutations in the *COL8A2*, *SLC4A11*, *ZEB1*, *LOXHD1* and *AGBL1* lead to the development of FECD, various functional studies were performed by independent groups of researchers. These studies were conducted using human specimens and/or murine models. *In silico* studies were also performed in some instances to investigate potential biochemical interactions between some

associated genes/proteins, and to predict how they might influence each other to cause the disease.

1.4.4.1 Studies in human specimens

According to the previous studies, patients with the L450W and Q455K mutations in the *COL8A2* gene are very distinct from non-carriers with late-onset FECD in both morphologic features of guttae and temporal progression of the disease [73, 116, 117]. Retroillumination of the cornea in representative patients with the mutations revealed a fine, patchy distribution of guttae in contrast to patients with late-onset FECD, in whom guttae were more coarse and distinct [72]. Similarly, confocal microscopy of the corneas in patients with the mutant alleles revealed small, mildly elevated guttae that were associated with individual endothelial cells [73]. The guttae were positioned near the centre of the cells [73]. In contrast, familial late-onset FECD patients had typical guttae, which were sharply raised and typically located along the borders between endothelial cells [72].

The average age of severely affected patients, grade ≥3, was observed to be roughly 40 years younger for the *COL8A2* pedigree than for the patients with familial late-onset disease [73]. However, both phenotypes progress at a similar rate that closely follows the age versus severity distribution of roughly 1 grade unit per 5 years [73, 117]. The age-versus-severity distribution of the L450W/Q455K-*COL8A2*-carrying patients is much narrower than that of the other patients with late-onset disease [73]. This is consistent with the fact that the pathogenesis of the latter involved greater heterogeneity, both in terms of the genes involved and the severity of the specific mutations [39, 76, 78, 79, 83].

Both affected individuals and total members of the *COL8A2* pedigree approximated a 1:1 female:male ratio, whereas affected individuals in the late-onset pedigrees were approximately 2.5:1 [73]. This observation confirmed the earlier findings in the general population [62, 118]. Thus, it is clear that the Mendelian *COL8A2* mutation follows conventional expectations of a 1:1 female:male ratio, and does not conform to the 2.5:1 ratio of common FECD.

The mechanism by which *COL8A2* mutations cause the clinical and histopathological features of FECD remains unknown. It is suggested that the mechanism of formation and positioning of guttae in patients with *COL8A2* mutations slightly differs from that in patients with the late-onset disease, in whom the guttae generally appear first at the cell-cell junctions, near the basolateral face of the endothelial cells [72]. COL8A2 is a short-chain, nonfibrillar, extracellular matrix component with a composition that varies between tissues [119, 120] and there is evidence that it may serve a structural role [121] or be involved in cellular differentiation [120].

The discovery of two FECD-related dominant mutations affecting a small, conserved interval of the COL8A2 collagen helix domain suggests that this region may be of selective importance to the function of this protein in the cornea, as these mutations have no obvious effects elsewhere in the body. Gottsch et al [73] suggest that mutant COL8A2 improperly interacts with other molecules in the extracellular matrix, resulting in sites of structural weakness in DM [122, 123] that allows the extrusion of material that forms the guttae. The irregular basal lamina topography appears to indent physically and stretch the endothelial cells [124, 125], which may compromise their ability to transport electrolytes and maintain corneal clarity.

Normally, COL8A2 interacts with collagenous and non-collagenous components of DM, such as TGFBI, to achieve the specialised arrangement of extracellular matrix proteins for optimal clarity of the cornea [3]. It is unknown whether the mutant COL8A2 is unable to achieve this specialised arrangment of extracllular matrix proteins, thus, possibly, leads to TGFBI accumulation in FECD-affected cornea [114]. Further studies may provide an insight to the molecular interactions of mutant proteins in the cornea in FECD.

Pathogenic mutations in *SLC4A11* were shown to lead to abnormal accumulation of mutant SLC4A11 in the endoplasmic reticulum in FECD-affected CE, and reduced trafficking of the accumulated protein to the cell membrane [76, 87, 126]. In their recent *in vitro* study for characterisation of the functional effects of *SLC4A11* mutations, Soumittra and colleagues [87] reported that the SLC4A11 protein processed to the cell surface was functionally compromised; with ion transport functional activity of the mutant protein in the CE being 5-42% that of the wild-type protein. Consistently, a serial analysis of gene expression (SAGE) study by independent researchers revealed a 5-fold decrease in the expression of *SLC4A11* transcript in FECD versus non-FECD CE [127]; SAGE is a technique used to compare global gene expression profile changes between two conditions. Collectively, the functional studies imply that less SLC4A11 protein may be available at the corneal endothelial cell membrane in FECD corneas, therefore compromising the ion transport. This could result in impaired corneal endothelial pump function.

To determine the impact of missense mutations in *ZEB1* on protein production and cellular localisation, Chung and colleagues [128] ectopically expressed in human

corneal endothelial cells (HCEnC-21T) tagged wild-type and mutant ZEB1 proteins comprising each of the six *ZEB1* missense mutations previously associated with FECD [39]. The data revealed no difference in the production and localisation between the wild-type and the mutant proteins [83, 128]. However, *in silico* analyses indicated that the mutations may alter putative ZEB1 regulatory or modification sites [128]. This necessitates further functional studies.

Different functional studies on characterisation of expression and effects of FECD-associated rare alleles in *LOXHD1* were performed by Riazuddin et al [78]. Expression analysis of the gene transcript in normal cultured human corneal endothelial cells showed expression of *LOXHD1* in human CE [78]. Further investigation for expression of the protein in corneal sections between FECD-affected proband with c.1639C>T (p.Arg547Cys) *LOXHD1* mutation, and from controls, FECD patient without the mutation and keratoconus patient, showed distinct protein aggregation in the CE and DM of the proband when compared to the controls [78]. Additional *in silico* studies using a structural model of LOXHD1 biochemically predicted that most of the reported FECD-associated mutations in *LOXHD1* reside on the surface of the protein [78], suggesting that they are more likely to affect the way in which the protein interacts with other proteins. To further investigate the importance of the finding in a biological system, Riazuddin and colleagues [78] conducted expression analysis of wild-type LOXHD1 and 3 mutant proteins in cells from human retinal pigment epithelium cell line (ARPE-19). Consistently, cells transfected with the mutant proteins showed distinct cytoplasmic puncta, which were rarely presence in cells expressing the wild-type [78]. LOXHD1 is an evolutionarily conserved protein predicted to consist of 15 polycystin-1, lipoxygenase, alpha-toxin (PLAT) domains. The biological function of PLAT

domains is not well established, but it is postulated to target proteins to the plasma membrane [129]. Based on this, and the findings from the studies by Riazuddin and colleagues [78], it is reasonable to hypothesise that some fraction of FECD might be caused by harmful LOXHD1 protein aggregation in the CE in the affected individuals.

Investigations into the impact of FECD-causing mutations in *AGBL1* were conducted by Riazuddin and colleagues [79] using similar approach as in the molecular studies of *LOXHD1* [78]. Examination of the ectopical expression of tagged wild-type and two AGBL1 mutants in NIH 3T3 mouse embryonic fibroblast cells revealed predominant cytoplasmic localisation of the wild type protein compared to distinct nuclear localisation of the mutant [79]. To interrogate the biological significance of this finding, Riazuddin et al [79] conducted additional expression analysis on AGBL1, TCF4 and TCF8 to examine whether the three might participate in the same biological processes. Cells from human embryonic kidney cell line (HEK293) were transfected with each tagged wild-type protein and analysed for any biochemical interaction. The data revealed specific interaction between AGBL1 and TCF4, but not TCF8 [79]. Further transfection of the cells with AGBL1 mutants and wild-type TCF4 showed significantly reduced binding affinity of the mutants to TCF4, thus suggesting that ablation of the AGBL1-TCF4 interaction might contribute to the pathogenesis of FECD.

Protein and gene expression studies have led to some understanding of FECD. Through comparative 2-D SDS-PAGE (two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis) differential abundance of PRDX2, PRDX3, CLU, and TGFBI proteins between FECD-affected and unaffected CE+DM (corneal

endothelium + Descemet's membrane) complexes were revealed [113, 114, 130]. The expression of the PRDX2 and PRDX3 were significantly lower in FECD-affected than unaffected tissue [130]. In contrast, CLU and TGFBI were expressed at significantly higher levels in diseased CE+DM than normal [114, 130]. Both, the CLU and TGFBI proteins were also shown to be overexpressed in corneas from normal older individuals compared to younger individuals [114], which could be relevant for late-onset FECD. Quantitative RT-PCR and western blotting confirmed that the *PRDX2* and *PRDX3* genes and encoded proteins are downregulated, and the *CLU* and *TGFBI* genes and encoded proteins upregulated in FECD-affected compared to the unaffected CE [113, 114, 130].

In order to compare the localisation of CLU and TGFBI in the CE of affected and normal individuals, and further investigate the relationship between these proteins, indirect immunofluorescence analysis was performed by Jurkunas and colleagues [114, 130]. They found unique pattern of colocalisation of both proteins in the centre of guttae in the affected compared to the unaffected CE [114]. TGFBI was prominent throughout the DM and showed a marked increase in intensity at the centre of the guttae [114]. CLU was also observed in the centre of guttae but closer to the apical side of the CE [114].

In our previous study, we further investigated expression profiles of the CLU and TGFBI proteins between FECD-affected and unaffected corneas using immunohistochemistry. Our immunolabelling analysis revealed differential distribution of CLU and TGFBI in the affected cornea compared to unaffected normal cornea [93]. These findings support the earlier findings. It is not known whether involvement of these proteins in the disease is a cause or an effect.

Two recent molecular studies between FECD-affected and unaffected CE and DM revealed substantial dysregulation of the genes, and their encoded proteins in the disease [131, 132]. Specifically, Weller and colleagues confirmed FECD-specific upregulation, production, and deposition of CLU and TGFBI, as well as revealing for the first time similar findings for agrin (AGRN), and collagen types III (COL3A1) and XVI (COL16A1) proteins in corneal endothelial cells [131]. Immunolocalisation of the proteins in CE and DM of corneal buttons from FECD patients and normal donor eyes support the earlier findings [131]. Consistently, Poulsen et al [132] revealed higher relative abundance of CLU, TGFBI, and AGRN in FECD-affected DM compared to normal DM. They also identified up-regulation of apolipoprotein D (APOD), C-type lectin domain family 11, member A (CLEC11A), collagen, type VI, alpha 3 (COL6A3), elastin microfibril interfacer 1 (EMILIN-1) and Keratocan (KERA), and down-regulation of COL1A2, COL2A1, COL4A4, COL5A2, Fibrillin-1, and Keratin type I cytoskeletal 9 and 10 and type II cytoskeletal 1 [132]. Collectively, the findings support the involvement of extracellular matrix alterations in the pathogenesis of FECD.

CLU is a molecular chaperone and is implicated in protecting cells from effects of physiological stress caused by aging, oxidative stress and apoptosis [69, 133, 134]. TGFBI is a secreted extracellular matrix protein that mediates cell-adhesion by interacting with major protein components of the DM, including collagens [114]. FECD is characterised by breakdown of tight junctions between corneal endothelial cells, aberrant cell adhesion, and migration of the surviving cells in a process resembling that of the EMT. TGFBI is induced by the *TGF-β* (*transforming growth factor-beta*) gene [135]. Interestingly, *TGF-β* regulates both *TCF4* and *ZEB1*, which have been implicated in FECD [109]. Therefore, these findings suggest that

dysregulation of cell adhesion and mesenchymal transformation of corneal endothelial cells underlie FECD pathogenesis.

Differential expression of nuclear ferritin, glutathione S-transferase-π, heat shock 70-kDa protein and APOD between FECD-affected and unaffected CE suggest that the basis for the disease is related to dysregulation of genes involved in cellular energy activity, antioxidant, and apoptosis induced by oxidative stress related damage of the corneal endothelial cells [127]. This is consistent with other findings indicating that oxidative stress and apoptosis are potential mechanisms underlying FECD pathophysiology [69, 134].

Accumulating evidence suggests that oxidative stress may contribute to the pathogenesis of FECD through mainly targeting of mitochondria in corneal endothelial cells [69, 136, 137]. Changes in the ultrastructure of mitochondria and large numbers of degenerated mitochondria found in CE of FECD patients suggest that these organelles may mediate oxidative stress-related pathogenesis of the disease [69, 117, 138, 139]. To quantify DNA damage and repair in FECD, Czarny and colleagues evaluated the extent of endogenous DNA damage induced by hydrogen peroxide as well as the kinetics of DNA repair in peripheral blood mononuclear cells of 50 patients with FECD and 43 age-matched controls without visual disturbances [140]. They found a lower efficacy of DNA repair in FECD patients as compared with control individuals [140], leading to the conclusion that lowering of the DNA repair capacity may be one of the mechanisms underlying the role of oxidative stress in FECD pathology [140]. Recently, this group reported another study that investigated the effect of oxidative stress in mtDNA damage and repair, and its copy number and 4977-base pairs common deletion in corneal endothelial cells and

peripheral blood lymphocytes isolated from FECD patients and normal participants [141]. The data demonstrated increased mtDNA damage in corneal endothelial cells of FECD patients compared with the controls [141], thus implying decreased mtDNA repair in the disease. Furthermore, the investigators observed higher copy number, and a higher ratio of the common 4977-base pair deletion of mtDNA in corneal endothelial cells of FECD patients compared with the normal indvidiuals [141]. They thus concluded that mutagenesis of mtDNA may be involved in FECD pathogenesis and disturbance in mtDNA sensitivity to damaging agents and changes in mtDNA damage repair along with alterations in mtDNA copy number may underlie this involvement [141].

Since aqueous humor supports CE [142, 143], it has been speculated its composition may play a role in FECD [144, 145]. To investigate this possibility, Richardson and colleagues [146] investigated differentially abundant proteins in aqueous humor proteome of individuals with and without the disease. The study identified several protein alterations in aqueous humour with FECD which may have important implications in the disease process. FECD patients exhibited downregulation in aqueous humor of afamin, histidine-rich glycoprotein, complement component 3, immunoglobulin heavy chain, and FAM3C proteins [146]. Additionally, common serum proteins (haemoglobin fragment 1, immunoglobulin kappa chain, Immunoglobulin light chain, and putative uncharacterised protein ALB), and CLU were found to be upregulated in aqueous humor of patients with FECD compared to non-FECD individuals [146]. These findings have not been confirmed by an independent technique or study.

The possible role of complement activation in the pathogenesis of FECD was raised in a study published over 30 years ago [147]. In another study from 1990 the C3a complement activation product in the aqueous humour was significantly increased in patients with FECD compared to non-FECD controls [148]. Fust et al [149] have recently investigated the role of the complement system in FECD by measuring the levels of the activation products in the aqueous humour of patients with and without the disease. The data revealed that C1rs-C1inh and the C3bBbP complexes were significantly increased in the aqueous humour of FECD patients compared to the control patients [149]. These results suggest that complement system is activated in FECD, and confirmed the previous findings implicating complement activation in the pathophysiology of the disease [147, 148]. However, these findings have not been validated by independent methods.

Few systemic disorders such as cardiovascular diseases, and progressive hearing loss have been reported to be associated with FECD. In his study, Olsen [150] has observed an increased rate of cardiovascular diseases in a series of 27 FECD patients, but this has not since been confirmed. In addition, mutations in the *LOXHD1* and *SLC4A11* genes have been found to cause FECD and hearing loss in some cases by two independent groups [78, 151]. However, to date reports of the association of FECD with systemic disorders have been generally limited by small sample size. Thus, definitive conclusions will require future studies with adequate number of the patients to determine causal or correlation reltionships between FECD and systemic diseases.

In summary, functional studies have identified various mechanisms that lead to the pathogenesis of FECD due to mutations in known disease-causing genes. Mutant COL8A2 is sufficient to cause early-onset FECD in the patients through its improper interaction and disassembly with constituents of DM, causing potential compromise in ability of the corneal endothelium to transport electrolytes and maintain corneal clarity. Pathogenic SLC4A11 contributes to the disease through its aberrant accumulation in the CE of some FECD patients, leading to reduced ion transporter in the cell membrane and corneal endothelial cell density. This culminates in compromised osmoregulation, and subsequent corneal oedema. The AGBL1 and LOXHD1 mutants are shown to cause FECD through accumulation of cytotoxic proteins in CE of relevant FECD patients, thus loss of corneal endothelial cells, leading to reduced cell density. Other functional studies have demonstrated potential roles for oxidative stress-induced mtDNA damage, decreased capacity for DNA or mtDNA repair, complement activation, for dysregulation of genes and proteins in the CE and DM in FECD patients. A summary of key studies described in this section is given in Table 1.1 (page 31 – 32).

Table 1.1 A summary table of key studies on specimens from patients with Fuchs' endothelial corneal dystrophy (FECD).

Specimen used (FECD vs non-FECD)	Type of study	Finding	Reference
	Gene expression analysis by:		
Corneal endothelium (CE)	Serial Analysis of Gene Expression	Differential expression of several corneal endothelial genes in FECD	[127]
Human CE culture	Quantitative RT-PCR	Over-expression of apoptotic-related genes in CE in FECD	[134]
Corneal endothelial-Descemet's membrane (CE-DM) complexes	Quantitative RT-PCR	Under-expression of *PRDX*-2, -3 and -5 genes in CE in FECD Over-expression of *CLU* and *TGFB*I genes in CE in FECD	[113, 114, 130]
CE-DM complexes	Quantitative RT-PCR array	Over-expression of extracellular matrix-related genes in CE FECD	[131]
	Protein expression analysis by:		
Human CE culture		Increased corneal endothelial cell apoptosis in FECD	
CE-DM complexes	Two-Dimension Gel Electrophoresis	PRDX-2, -3, -5 and -6 proteins identified in CE-DM tissue	[130]
CE-DM complexes	MALDI-TOF mass spectrometry (MS)	Expression of PRDX-2, -3 and -5 identified in human CE	[130]
CE-DM complexes	iTRAQ and Label-free MS	Differential relative abundances of extracellular matrix proteins in DM/endothelial layer in FECD	[130, 132]
CE-DM complexes	Western blotting	Under-expression of PRDX-2 and -3 proteins in CE in FECD	[130]
CE-DM complexes	Western blotting	Over-expression of CLU and TGFBI in CE in FECD	[113, 114]
Corneal buttons/corneas	Immunocytochemistry	Co-localisation of CLU and TGFBI in CE-DM layer and in guttae	[114]
Corneal buttons/corneas	Immunocytochemistry	Increased evidence for oxidative stress in CE in FECD	[69]
Corneal buttons/corneas	TUNEL assay	Colocalisation of DNA damage and apoptotic CE cells in FECD	
Corneal buttons/corneas	Immunohistochemistry, TEM	Activation of unfolded protein response and its relationship	[126]

		with endothelial cell apoptosis in FECD	
	Aqueous humor analysis by:		
Aqueous humor	ELISA	Activation of complement system in aqueous humor in FECD	[149]
	Quantification of mtDNA damage and repair in FECD by:		
Peripheral whole blood cells	Comet assay	Oxidative stress-induced mtDNA damage in CE in FECD	[140]
Peripheral whole blood cells	Comet assay	Decreased capacity for DNA or mtDNA repair	[141]

PRDX, = peroxiredoxin; CLU = Clusterin; TGFBI = Transforming growth factor beta-induced; RT-PCR = reverse transcription-polymerase chain reaction; MALDI-TOF MS = matrix-assisted desorption ionization-time of flight mass spectrometry; iTRAQ = isobaric tag for relative and absolute quantitation; TUNEL = terminal deoxynucleotidyl transferase dUTP nick-end labeling; TEM = Transmission electron microscopy; Enzyme-linked immunosorbent assay; mtDNA = mitochondrial DNA.

1.4.4.2 Studies in murine models

Murine models of FECD generated by manipulation of the *COL8A2*, *ZEB1* and *SLC4A11* genes have been investigated to understand the mechanism/s of the disease. *COL8A2 L450W* and/or *Q455K* mutant transgenic knock-in mouse models showed characteristic features of the early-onset FECD, early endothelial cell unfolded protein response and apoptosis, and distinct phenotypes and evidence for altered autophagy [152, 153]. These mice are thought to be good models for the early-onset form of the disease [152, 153].

Quantitative RT-PCR analysis of corneal endothelial cells from *L450W* and *Q455K* *Col8a2* knock-in mouse models respectively revealed upregulation of autophagy marker *Dram1* [153]. Analysis of human late-onset FECD CE by the same technique also showed upregulation of *DRAM1* compared to normal CE [153]. Consistent upregulation of the gene in mouse and human corneal endothelial cells suggested a role for altered autophagy in FECD. Based on this finding, Matthaei et al [154] performed an endothelial cell whole genome expression microarray analysis in a *Col8a2 Q455K* mutant knock-in mouse model to identify potential targets which could be correlated to human late-onset disease. This study revealed significantly increased mRNA levels of *cyclooxygenase 2* (*Cox2*) and *Jun proto-oncogene* (*Jun*) in $Col8a2^{Q455K/Q455K}$ mutant compared to wild-type mice [154]. Investigations of differential expression of the *COX2* and *JUN* genes between late-onset FECD and normal human CE confirmed a statistically significant increase in levels of *COX2* and *JUN* transcripts, and COX2 and JUN proteins [154]. Generally, cyclooxygenases, such as COX2, influence changes in morphology, mitosis, and migration of CECs through their most important product, the eicosanoid prostaglandin E2 (PGE2) [155-157]. Jun is a protein of the Ap-1 complex with a

complex spectrum of functional properties such as proliferation, cellular survival or death, and differentiation [158, 159]. Upregulation of these genes in FECD is likely to lead to dysregulation of cellular morphology, survival and death, and abnormal cell migration of the corneal endothelium.

COL8A2 knockout mice manifest a dysgenesis of anterior segment of the eye with a globloid, keratoglobus-like protrusion of the anterior chamber [160]. Some of the characteristic features include an increased depth of anterior chamber, thin corneal stroma, and thin DM that separates corneal stroma from enlarged and fewer corneal endothelial cells [160]. DM thinning in *COL8A2* knockout mice is different from what is seen in patients with FECD or PPCD [10, 44]. Therefore, the *COL8A2* knockout mice do not serve as a model for FECD.

ZEB1 knockout mice exhibit characteristics similar to patients with PPCD, rather than FECD [161]. Reactivation of epithelial specification genes in mesenchymal, neuroectodermal, and endothelial cells is seen in these mice [161]. These observations are consistent with the role of *ZEB1* and suggest that it has a more general role in suppression of an endothelial phenotype, which is involved in both PPCD and FECD [77, 162, 163].

Corneas in the *SLC4A11* knockout mice appeared generally normal with subtle changes in the epithelial cell architecture and a normal endothelium, which remained unchanged as the animals aged [164]. Besides, corneal phenotype in the *SLC4A11* knockout mice differs significantly from the severe corneal phenotype described in FECD patients with mutations in the *SLC4A11* gene [164]. Thus, this mouse is not considered a robust model of FECD.

The differences in corneal phenotypes between humans and *COL8A2*, *ZEB1* and *SLC4A11* knockout mice may be explained by physiological differences between human and mouse cornea. For instance, mouse has a substantially thinner cornea lacking a Bowman's layer, and a corneal endothelium that possesses a replicative potential not observed in human [165]. In addition, these knockout mice lack either the *COL8A2*, *ZEB1*, or *SLC4A11* gene, which is a more severe effect than the presence of a mutant gene in humans. This difference may be the reason why only *COL8A2* transgenic knock-in mouse shows characteristic features of early-onset FECD. Presently, no good animal models are available for late-onset FECD.

1.5 GAPS IN CURRENT KNOWLEDGE ABOUT FECD

Many of the findings about FECD described in this chapter were only made in the first one to two years, prior to the beginning of this project, and within the last last two years before its ending. These findings have been discussed in this thesis. Although many facts have been revealed about the genetic and molecular biology of FECD, much remains to be known about this genetically heterogenous and complex disease.

Reported mutations in known disease-causing genes are responsible for only a small proportion of FECD cases in non-Australian populations. In addition, genes within the mapped disease loci, or genetic cause of the disease in Australian population are yet undetermined. Genetic association of *TCF4* with FECD has been observed in Caucasian American, Chinese, Indian and Australian populations. However, it does not account for the entire disease burden in these populations, thus suggesting that more genes are involved including in the Australian population.

To date, molecular studies have provided information about a few genes or proteins implicated in FECD pathophysiology. However it is not yet known whether they are the cause or effect or how are they functionally related and contribute to the disease. Therefore, further studies are needed for understanding the pathopathology of the disease. Better understanding of the pathophysiology of the disease will improve the knowledge about the mechanism of FECD, which will in time facilitate development of strategies for early intervention/prevention or better disease management and improved visual outcomes for patients.

1.6 HYPOTHESES AND AIMS

1.6.1 Hypotheses

1. The majority of genetic factors involved in the pathogenesis of late-onset FECD, particularly in Australian cases, remain unknown and can be identified through genetic approaches.

2. Pathological changes in the Descemet's membrane in FECD are accompanied by changes in the corneal endothelium. These cellular changes can influence gene expression and thereby protein composition in the Descemet's membrane, therefore, identification of proteins with altered abundance in Descemet's membrane and genes with altered expression in corneal endothelium in FECD can provide insights into the disease mechanism.

1.6.2 Aims

The main aim of this project was to improve the understanding of the molecular basis of FECD. The specific aims were:

Aim 1: To identify genetic factors involved in the pathogenesis of late-onset FECD in Australian cases.

Aim 2: To identify differentially abundant proteins between FECD-affected and unaffected Descemet's membrane.

Aim 3: To identify differentially expressed genes between FECD-affected and unaffected corneal endothelium.

All the aims were successfully achieved and have led to novel findings. These findings are described in Chapters 3 to 8.

CHAPTER 2
Materials and Methods

The materials and methods used in more than one study in the present project are presented in this chapter. The materials and methods specific to individual studies have been included in the relevant chapters.

2.1 ETHICS APPROVALS

Ethics approvals for the research reported in this thesis were obtained from the Human Research Ethics Committees of the Flinders Medical Centre/Flinders University of South Australia (Adelaide, South Australia, Australia), the Royal Victorian Eye and Ear Hospital (Melbourne, Victoria, Australia), and the University of Sydney (Sydney, New South Wales, Australia). The research was conducted in accordance with the requirements of the National Health and Medical Research Council (NHMRC) Statement on ethical Conduct in Human Research (2007, Updated March 2014) (http://www.nhmrc.gov.au). Written informed consent was obtained from all the individuals who participated in the genetic studies. Consent was also obtained from participants from Victoria and Western Australia (Australia) who provided their relevant surgical specimens for molecular studies. The research adhered to the tenets of the Declaration of Helsinki.

2.2 PARTICIPANT RECRUITMENT

2.2.1 Genetic studies

2.2.1.1 Fuchs' endothelial corneal dystrophy (FECD) cases

FECD patients were recruited by independent ophthalmologists at the Flinders Eye Clinic (Adelaide, SA, Australia), and the Royal Victorian Eye and Ear Hospital

(RVEEH) (Melbourne, VIC, Australia). All the patients consulted the ophthalmologists for their patient-doctor appointment due to impaired vision. Each participant underwent detailed ophthalmic examination, including slit lamp biomicroscopy to investigate the presence and nature of guttae, corneal oedema and haze, and abnormal thickening of the cornea to determine severity of the disease. The majority of the recruited participants received corneal grafts, and histopathological examination of their surgical specimens confirmed the diagnosis of FECD. All the recruited patients were Caucasian Australians and were diagnosed with advanced disease (Grades 3 – 6). The disease was graded on a semi-quantitative scale from 0 – 6, modified from previous severity scale [62, 64]. The grading scale is given in Table 2.1. A total of 190 cases were recruited for this project.

Table 2.1 Fuchs' endothelial corneal dystrophy grading scale used for determining disease severity.

Disease Status	Grade	Criteria (central/paracentral corneal guttae)
Unaffected	0	No guttae
Intermediate	1	1-12 non-confluent
	2	More than 12 non-confluent
	3	1-2 mm confluent
Severe	4	2-5 mm confluent
	5	Greater than 5 mm confluent
	6	Greater than 5 mm confluent with clinically apparent stromal/epithelial oedema

2.2.1.2 Normal South Australian controls

The Normal South Australian (NSA) controls consisted of healthy elderly Caucasian Australians, aged over 50, who were ascertained from the Flinders Eye Clinic and residential retirement villages and nursing homes within Adelaide, South Australia, Australia. These individuals were previously recruited for use as controls in a variety of ocular genetic studies in our laboratory [166-168]. A total of 282 controls were recruited for the present study.

2.2.1.3 Blue Mountain Eye Study controls

The Blue Mountains Eye Study (BMES) controls were from a previous population-based survey of vision and common eye diseases in Australians over the age of 49 years living in the Blue Mountains region, west of Sydney, New South Wales, Australia. Participants recruited for this study were predominantly Caucasian; the population and full recruitment methodology has been described in detail previously [169, 170]. Briefly, all participants, who were permanent, non-institutionalised residents of the defined geographical region, were identified by a door-to-door census of all dwellings and by closely matched findings from the national census. At the baseline visit, the participants received a detailed eye examination. They also underwent two follow-up visits at 5 and 10 years, respectively. A total of 2761 participants with normal threshold or suprathreshold field tests and no sign of any ocular disease were recruited for study in this project.

2.2.2 Molecular studies

Surgical specimens of corneal endothelium and Descemet's membrane (CE+DM) complex were obtained from Caucasian Australian patients undergoing Descemet's stripping Automated Endothelial Keratoplasty (DSAEK)

procedure due to clinically diagnosed end-stage FECD, at the Flinders Eye Clinic, and through collaboration with Dr. Grant Snibson at the RVEEH (Melbourne, Victoria, Australia), and Dr. Steven Wiffen at the Lions Eye Bank (Western Australia, Australia). Control CE+DM complex specimens were from deceased donors whose corneas were deemed unsuitable for corneal transplantation and were obtained through the Eye Bank of South Australia (Flinders Medical Centre, Bedford Park, South Australia, Australia) following consent from donor families. Each specimen was placed in 1mL RNA*later*® Solution (Life Technologies Australia Pty Ltd., Mulgrave, VIC, Australia) immediately after surgery or dissection, and stored at 4°C for at least two days. Then the RNA*Later*® solution was removed and the specimen stored at -80°C for later protein or RNA analysis. For immunohistochemical analyses, formalin-fixed, paraffin-embedded sections of FECD-affected corneas, and normal corneas were used. The affected corneas were initially used for histopathological diagnosis following corneal transplantation and left over tissue was available for this study. The normal corneas were provided by the Department of Anatomical Pathology (Flinders Medical Centre, Bedford Park, South Australia, Australia), and/or the Eye Bank of South Australia (Flinders Medical Centre). Normal corneas were from deceased donors whose corneas were deemed unsuitable for corneal transplantation and consented for use in research by the donor families.

2.3 SAMPLE PREPARATION

2.3.1 DNA extraction from peripheral blood

From cases and controls, venous blood samples were collected in 9-mL EDTA tubes. Genomic DNA was extracted from peripheral blood using QIAmp DNA Blood Maxi

Kit (Qiagen Pty Ltd, Doncaster, VIC, Australia) according to the manufacturer's protocol. Extracted DNA samples were quantified using mass spectrophotometer and stored in -20°C until required.

2.3.2 Preparation of DNA pools

DNA pools were prepared for use in genotyping of genomic DNA samples of cases and controls in Chapters 3 and 4. Pooling of DNA samples significantly reduces the amount of genotyping required to perform large scale studies [171, 172], thus making the technique available to laboratories with limited resources. Construction of DNA pools from cases or control samples required the addition of equal amounts of DNA from each individual. Therefore an accurate quantitation of each individual DNA sample was necessary, which was achieved by Picogreen assay using the Quant-iT™ PicoGreen® dsDNA reagent (Invitrogen, Carlsbad, USA). The PicoGreen® fluorochrome selectively binds double-stranded DNA and has the advantage of producing little background due to the unbound dye having virtually no fluorescence. The Fluoroskan® Ascent microplate fluorometer (ThermoFisher Scientific, Waltham, USA) was used to measure the fluorescence readings from the PicoGreen® assay.

2.3.2.1 Estimation of DNA concentration

Individual genomic DNA samples were initially quantitated using a spectrophotometer. The DNA was then diluted into a 96-well plate according to its concentration. The initial dilution was performed a day prior to performing PicoGreen® assay. The samples were diluted in TE buffer (10mM Tris, 1mM EDTA; pH 8.0) as follows.

i. <175 ng/µl = neat (80 µl DNA)

ii. 175-300 ng/µl = 1:1 (60 µl DNA + 60 µl TE buffer)

iii. 300-600 ng/µl = 1:2 (40 µl DNA + 80 µl TE buffer)

iv. >600 ng/µl = 1:3 (25 µl DNA + 75 µl TE buffer)

The diluted samples were further diluted 100× in TE buffer in a new 96-well plate. On this same day, preparation of a DNA standard for the PicoGreen® assay was also undertaken using a stock of 100ng/µl Lambda DNA (Invitrogen, Carlsbad, USA). The stock DNA was diluted to 1ng/µl in TE buffer, which was then serially diluted to the following concentrations: 0.5 ng/µl; 0.25 ng/µl; 0.125 ng/µl; 0.1 ng/µl; 0.05 ng/µl; 0.025 ng/µl; 0.01 ng/µl; 0.005 ng/µl; 0.0025 ng/µl. The DNA samples dilution plate and the standard dilutions were left at room temperature overnight to achieve a homogenous solution.

On the day of quantitation, sufficient PicoGreen® reagent was diluted 200× in TE buffer. The DNA standard solutions were then measured, with 5 µl of each standard and 45µl of TE buffer added to a black OptiPlate (PerkinElmer, Waltham, USA), with blanks also included. Each standard had four replicates. The PicoGreen® reagent reservoir and DNA standards plate were then placed in the Fluoroskan® Ascent microplate fluorometer. PicoGreen® reagent was added into each DNA standard well, and the plate was agitated to mix the samples with the PicoGreen® reagent. The mixture was incubated for 4 minutes. The fluorescence was measured, with the excitation wavelength set between 475 to 505nm and the emission wavelength at 520 to 550nm. The software generated the standard curve from fluorescence measurements.

To prepare DNA sample dilutions for PicoGreen® quantitation, 25µl of each DNA sample was loaded in duplicate onto a black OptiPlate (PerkinElmer), to which a further 25µl of TE buffer was added. The DNA samples plate was placed in the Fluoroskan® Ascent microplate fluorometer. As described above, the PicoGreen® reagent was added, and the mixture was agitated and incubated for 4 minutes, and fluorescence was measured. Using the generated standard curve, concentrations of DNA samples were calculated. Based on the calculated concentration each sample was diluted to 75ng/µl, whilst those with a concentration under 85ng/µl were used neat. As with the first round of quantitation, all samples were then diluted 1:100 in TE buffer in a 96-well plate and quantitation measurements repeated as described above. The DNA concentrations obtained following the second PicoGreen® assay were then used to construct the case and controls pools.

2.3.2.2 Construction of DNA pools

To create the pools, the same amount of genomic DNA from each FECD/control sample was included in the pool. A minimum of 5µl of each DNA sample was added to the final pool. This was to minimise the impact of pipetting error from using smaller volumes. To achieve this, 5µl of DNA from the sample with the highest concentration was added to the pool. This determined the amount (nmol) of each sample in the pool. All samples with lower concentrations required the addition of more than 5µl to achieve the identical amount of DNA. The prepared case and control DNA pools were used for genotyping studies reported in Chapters 3 and 4 in this thesis.

2.3.3 RNA extraction

Total RNA was isolated from diseased and normal CE+DM specimens using Trizol reagent (Invitrogen, Carlsbad, CA, USA), and an RNeasy mini-kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. A stainless bead was thoroughly cleaned using RNAzap and dried with Kim wipe, and then placed into a sterile 2-mL eppendorf tube with the CE+DM specimen and 1ml of Trizol lysis reagent. The specimen was homogenised on a TissueLyser (Retsch GmbH & CO. KG, Haan, Germany) at 30 Hertz's (Hz) four times for 2 minutes each; the homogenate was placed on ice between bursts for up to 2 minutes each. The homogenate was carefully pipetted into a clean 1.5 mL microfuge tube and incubated at room temperature for 5 minutes. Two hundred microlitres of chloroform was added to the homogenate. The tube was capped and vigorously vortexed for 15 seconds, and left for 3 minutes at room temperature. The homogenate-chloroform mixture was then spun at 12,000 x g for 15 minutes at 4°C to separate organic phase from aqueous phase. The aqueous phase was carefully removed avoiding cellular debris and placed into 1.5ml sterile tube. One-to-one volume of 70% ethanol was added into the aqueous phase and thoroughly mixed by pipetting five times. 700µl of the lysate was transferred to an RNeasy Mini Spin column placed in a 2ml collection tube (Qiagen). The column was capped and spun at 10,000 × g for 15 seconds at room temperature. The flow-through was discarded. The remainder of the sample mixture was transferred to the spin column and spun at 10,000 × g for 15 seconds at room temperature. The flow-through was again discarded. 350µl of buffer RW1 was added into the RNeasy Spin column, and the column was spun at 10,000 × g for 15 seconds to wash the membrane. The flow-through was discarded. 80µl of DNase I incubation mix, prepared from 10µl of DNase 1 Stock solution (Cat #:79254, DNA-

free, Ambion, Austin, TX, USA) and 70µl buffer RDD, was directly added to the RNeasy spin column membrane. This was to treat total RNA sample on-column with Dnase I (DNA-free, Ambion, Austin, TX, USA) to remove any contaminating genomic DNA. The column was incubated for 15 minutes at room temperature, before a 350µl of buffer RW1 was added to it. The mixture was spun 10,000 × g for 15 seconds at room temperature, and flow-through was discarded. A 500-µl volume of buffer RPE was added into the column, and spun at 10,000 × g for 2 minutes. The column was carefully removed from the collection tube and placed into new 2-mL collection tube (Qiagen). The column was capped and centrifuged at full speed for 1 minute to remove any residual buffer RPE. It was then placed in a new 1.5 mL collection tube (Qiagen). A 20µl of RNase-free water was directly added to the column membrane, and the collection tube was centrifuged at 10,000 × g for 1 minute to elute the RNA. The RNA eluent was returned to the column, and the collection tube was spun again at 10,000 × g for 1 minute. Isolated RNA samples were quantified on the NanoDrop 8000 (Thermo Scientific, Wilmington, DE). The absorbance ratio at 260 nm and 280 nm was used as a guide to determine RNA quality. Total RNA with a ratio of ≥1.8 was considered acceptable for use.

2.3.4 Protein extraction

2.3.4.1 For Western blotting

For Western blotting, protein extracts from human serum, and human breast cancer (MCF7) and colorectal carcinoma (Caco2) cell lines were kindly provided by Ms Alpana Dave (Department of Ophthalmology, Flinders University, Adelaide, Australia). The serum protein extract was used to determine the specificity of mouse monoclonal anti-human APOE antibody (Calbiochem, Merck Pty, VIC, Australia)

and MCF7 and Caco2 lysates were used for determining the specificity of the mouse anti-human ATP1B1 (Thermo Scientific, Rockford, IL, USA) antibody.

Western blotting was also performed on human normal colon lysate to determine the specificity of the rat anti-human LAMC1 antibody (Calbiochem, Merck Pty, VIC, Australia). The colon specimen was a kind gift from Associate Professor Michael Michael (Department of Gastroenterology and Hepatology, Flinders University, Adelaide, Australia). The protein was extracted from the tissue in 200µl of radio-immunoprecipitation assay (RIPA) buffer [10mM HEPES pH 7.5, 150mM sodium chloride, 2mM ethylenediaminetetraacetatic acid (EDTA), 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 25x protease inhibitor cocktail (Roche Diagnostics Australia Pty Ltd., Castle Hill, Australia), 0.057mM phenylmethylsulfonyl fluoride, 2mM sodium orthovanadate, 10mM sodium pyrophosphate and 20mM sodium fluoride]. Homogenisation was performed in a glass homogeniser. The homogenate was chilled on ice for 20 minutes, then pipetted into a 1.7ml maximum recovery tube (Brand: MCT-175-L-C; Axygen Inc., Union City, CA, USA) and spun at 18000 × g for 30 minutes at 4°C to pellet the debris. Supernatant was transferred to another maximum recovery tube and protein concentration was estimated by bicinchoninic acid assay (BCA) protein assay (Thermo Fisher Scientific, VIC, Australia) according to the manufacturer's protocol and using the VESRAmax tunable microplate reader (Molecular Devices, LL, Sunnyvale, CA, USA). Ovalbumin (2mg/ml) protein was used as standard. The standard sample was serially diluted into 7 two-fold dilution series. Each standard dilution and the protein concentrate were quantified in triplicate. A standard curve was generated from net signal intensity and corresponding mass (in mg) from each of

the seven standard sample dilutions. The standard curve was used to calculate protein concentrate.

2.3.4.2 For Label-free mass spectrometry

CE+DM specimen was washed thrice with 100 µL of Ultrapure Water (CascadaAN water; PALL Corporation). Then the specimen was incubated in 100µl of Ultrapure Water at room temperature for 20 minutes to lyse the corneal endothelial cells. The specimen was removed from water and washed three times in 100µl of 10mM Tris-HCl buffer (pH 7.5) each, to remove any cellular proteins present on the specimen. It was then transferred into a 2-mL low-binding tube. Proteins were extracted by chemical cleavage followed by hydroxylamine and guanidine-hydrochloride extraction. One hundred and forty microlitres of 98% formic acid and 1 mg of cyanogen bromide were added to the sample and the sample incubated at 30°C overnight to cleave peptide bonds at the C-terminus of the methionine residues. The sample was dried in a SpeedVacuum (Labcono Corporation, Kansas City, MO, USA) for at least 45 minutes. It was then homogenised in 75µl of 2M hydroxylamine and 6M Guanidine-HCl extraction buffer (pH 9). Homogenisation was performed, using a cleaned stainless steel ball, on a TissueLyser (Retsch GmbH & CO. KG, Haan, Germany) at 30 Hertz's (Hz) four times for 30 seconds each; the homogenate was placed on ice between bursts for up to 2 minutes each. The lysate was incubated at 45°C for 4 hours before it was spun at 18,000 x g in a microfuge for 10 minutes. Supernatant was transferred into 1.5ml fresh low-binding tube. The protein samples were buffer exchanged into a nUPLC-MSE compatible buffer (1.6M urea, 100mM Tris-HCl, pH 7.5) by volume-exchange using Vivaspin 500 columns (3,000 kDa, PES membrane; Sartorius, Melbourne, Australia). In brief, the lysate was made

up to 500µl with the urea extraction buffer and applied onto the Vivaspin 500 PES column. The column was then spun at 15,000 x g for 30 minutes at room temperature. The filtrate was discarded. This process was repeated at least five times. The final 30µl of protein concentrate was collected in a low-binding tube, and quantified using the EZQ Protein Quantitation method (Molecular Probes, Eugene, Oregon, USA) according to the manufacturer's protocol. Ovalbumin (2 mg/ml) protein was used as standard. The standard sample was serially diluted into 5 two-fold dilution series. Each standard dilution and the protein concentrate were quantified in triplicate. A standard curve was generated from net signal intensity and corresponding mass (in mg) from each of the five standard sample dilutions. The standard curve was used to calculate the mass of the protein concentrate.

2.4 EXPERIMENTAL PROTOCOLS

2.4.1 DNA sequencing

2.4.1.1 Primer design

The gene-specific polymerase chain reaction (PCR) primers used for DNA sequencing in the studies described in Chapters 3 and 5 are tabulated in Table 2.2 (page 50). Reference sequences for the *LOXHD1* (NM_144612.6) and *SLC4A11* (NM_032034.3) genes were retrieved from GenBank through the UCSC genome browser (http://genome.ucsc.edu.au/). Primers were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/) to amplify the coding exons and splice sites of exons of interest in the *LOXHD1* (exons 2, 32, 37, 40) and *SLC4A11* (exon 3) genes. Primers used for amplification of the TGC trinucleotide repeat region in the third intron in the *TCF4* gene (in Chapter 5) were the same as those used by Wieben et al [102]. All primers were synthesised by GeneWorks (GeneWorks Pty Ltd, Thebarton,

SA, Australia). The primers were rehydrated to 200μM solution with sterilised milliQ water (MQH_2O). A 10μM working dilution of each primer was prepared from the 200μM stock by adding 5μl of the stock solution into 95μl of MQH_2O. Both the stock solution and working dilutions were stored at -20°C for later use.

Table 2.2 Primers used for mutation screening and *TCF4* repeat expansion studies. Gene names, amplified regions, primer sequences, annealing temperature for PCR and expected product sizes in base pairs (bp) are given.

Gene	Exon	Forward primer (5' > 3')	Reverse primer (5' > 3')	Annealing temperature (°C)	Product size (bp)
LOXHD1	2	TACAGGATGCA GCAGGGATT	TGTGGTCAGAT GTGCCTTCT	57	356
	32	GCAGTGCCTTG TCTATCTGG	AGGTAGGCTGT TCTTCCCAC	57	370
	37	CCTAACCCAAC CCCTCACTT	CGGTGTCTTCTT ATTCCAGCA	57	396
	40	ACCAAGACGAC AGAGAGCTT	CAAGGTGGAGG GCAGAAATG	57	474
SLC4A11	3	GGGAATGCTGG AGACTCACT	GGCCTGCATCT CAAGGTTG	57	356
TCF4	3	CAGATGAGTTT GGTGTAAGATG	ACAAGCAGAAA GGGGGCTGCAA	68	264-481

2.4.1.2 Polymerase chain reaction

PCR mixtures were prepared in a laboratory that was free of any PCR product contamination. The reactions were set up in 96-well PCR plates (Axygen Scientific, Inc., Union City, CA, USA). Each reaction contained 40 ng of template DNA, 0.05 μM each of forward and reverse primer, 400 μM dNTP mix (Rochie Diagnostics GmbH, Mannheim, Germany) and 0.5U HotStar Plus Taq® DNA Polymerase with 1× final concentration of supplied PCR buffer (Qiagen GmbH, Hilden, Germany). The final volume per reaction was made up to 20μl for Sanger sequencing (Chapter 3) or 10μl for short tandem repeat (STR) assay (Chapter 5) with sterilised MQH_2O.

The set-up of a standard PCR mix is shown in Appendix, Table 2 (page 211). The template DNA was replaced with sterilised MQH_2O in the negative control well/s. The plates were heat sealed with MF-111 heat sealing film (Axygen Scientific, Inc.) using Minisealer (Axygen Scientific, Inc.) at 150°C for 4 seconds. They were then briefly centrifuged in a mini-plate spinner (Axygen Scientific, Inc.). DNA amplification was performed in a Bio-Rad thermal cycler (Bio-Rad Laboratories Pty Ltd., Gladesville, NSW, Australia) under gene-specific conditions. The PCR conditions are given in the relevant chapters.

2.4.1.3 Agarose Gel Electrophoresis

PCR products were analysed by agarose gel electrophoresis on a 1% gel. The gel was prepared by dissolving 0.5g agarose (Promega, Madison, WI, USA) in 50mL of 1× Tris-borate-EDTA (TBE) buffer (0.045M Tris, 0.045M boric acid, and 0.001M EDTA; pH 8.0) in a microwave for about 1-1.5 minutes. Two microlitres of 5000× GelRed stain (Biotium Inc, Hayward, CA, USA; stock concentration:10,000× in water) were added to the gel before setting. Five microlitres of each product was mixed with 2µl of gel loading buffer (containing 80% glycerol, 250µl of sterilised milliQ water, 0.02mM Bromophenol blue, and 0.03mM Xylene cyanole) and loaded onto the gel submerged in 1× TBE buffer. Half microgram of 100 base pair DNA ladder (New England Biolabs, Beverly, USA) was loaded as a size standard. Samples were subjected to electrophoresis at 130V for 40 minutes using an external power supply (Bio-Rad Laboratories Pty. Ltd., Gladesville, NSW, Australia). Gels were imaged on a UV transilluminator using the GeneGenius imaging system (Synoptics, Cambridge, England).

2.4.1.4 Clean-up of PCR products

To prepare PCR products for sequencing, 5µl of each PCR product was treated with 10U Exonuclease 1 (New England Biolabs, Ipswich, MA, USA) and 2U Shrimp Alkaline Phosphatase (SAP) (USB Corporation, Cleveland, Ohio USA) to digest single stranded primer DNA and dephosphorylate dNTPs, respectively. The reaction mix was incubated in a Bio-Rad thermal cycle (Bio-Rad Laboratories Pty Ltd, Gladesville, NSW, Australia) at 37°C for 1 hour followed by 80°C for 20 minutes to terminate the reaction.

2.4.1.5 DNA Sequencing

Cleaned PCR products were sequenced using BigDye® Terminator v3.1 (Applied Biosystems, Foster City, CA) and 5 µM of forward primer, and electrophoresed on a 3130xL Genetic Analyzer (Applied Biosystems) at the SA Pathology Sequencing Facility (Flinders Medical Centre, Adelaide, Australia) according to standard protocols. Data was received in the form of chromatogram files.

2.4.1.6 Sequence Analysis

Using the Sequencher® software 5.0 (GeneCodes Corporation, Ann Arbor, MI), DNA sequences from cases and controls were aligned to relevant reference sequences of *LOXHD1* (NM_144612.6) and *SLC4A11* (NM_032034.3), and *TCF4* (NM_001083962.1). The reference sequences were retrieved from the Reference Human Genome sequence version 19 (hg19_RefGen). Reported or novel variants of interest were noted by visual inspection of chromatograms of individual sequences in Sequencher®.

2.4.2 Sequenom MassARRAY® SNP genotyping

SNP genotyping in individual FECD, BMES and NSA samples was performed using iPLEX Gold chemistry (Sequenom Inc., Herston, QLD, Australia) on a MassARRAY Spectrophotometer (Sequenom Inc.) at the Australian Genome Research Facility (AGRF) (Brisbane, QLD, Australia). Genotypes were analysed using PLINK [173]. Chi-square (χ^2) and/or Fisher's exact tests were conducted to assess allelic and genotypic associations. Haplotype analysis was also performed using PLINK to investigate the effect of haplotypes of the associated SNPs at each locus, on FECD.

2.4.3 RNA analysis

2.4.3.1 Primers

With the exception of primers for *APOE*, all primers used in quantitative reverse transcription-polymerase chain reaction (quantitative RT-PCR) in this project were commercially acquired from SABiosciences (SABiosciences, Doncaster, VIC, Australia). Details of these primers are given in Table 2.3 (Page 54).

Table 2.3 Details of quantitative RT-PCR primer assays. Gene symbols, reference sequence (RefSeq), GeneBank accession number, primer assay ID, expected product sizes in base pairs (bp), and positions of the amplicon in the RefSeq sequence (reference position) of each gene are tabulated.

Gene symbol	RefSeq accession number	Assay ID	Band size (bp)	Reference position
ACTB	NM_001101	PPH00073E	174	730
ALPK2	NM_052947.3	PPH18889A	64	6547
ATP1B1	NM_001677.3	PPH01367A	136	871
BGN	NM_001711.4	PPH01899A	91	1232
CLIC6	NM_053277.1	PPH17669A	110	1958
CPAMD8	NM_015692.2	PPH58080A	125	2219
CSF1R	NM_005211.3	PPH00191F	95	1941
CST1	NM_001898.2	PPH11620A	115	173
CX3CR1	NM_001337.3	PPH00620A	155	580
EDN1	NM_001955.4	PPH00653A	96	888
HLA-DRA	NM_019111.4	PPH00857F	148	310
LAMC1	NM_002293.3	PPH00184F	127	4941
NOX4	NM_016931.3	PPH06078A	88	1643
PPP1R1B	NM_181505.3	PPH05948A	84	519
SPP1	NM_000582.2	PPH00582E	88	892

Specific primer sequences for the *APOE* gene (Forward: 5'-TTGCTGGTCACATTCCTGG-3'; Reverse: 5'-CAGGTAATCCCAAAAGCGAC-3') were retrieved from Primer Depot, the online database (http://primerdepot.nci.nih.gov/cgi-bin/testdb.pl). These primers were synthesised by GeneWorks (GeneWorks Pty Ltd, Thebarton, SA, Australia).

2.4.3.2 Complementary DNA synthesis

Complementary DNA (cDNA) was synthesised by reverse-transcription from 0.471µg of total RNA from FECD-affected and unaffected corneal endothelium using first-strand cDNA synthesis kit (SuperScript III First-Strand Synthesis System, Invitrogen) and random hexamer primers. Standard cDNA samples with and without reverse transcriptase (RT+ and RT-) were synthesised from RNA of all the analysed

samples. To create the standard cDNA sample, the same amount of total RNA from each sample was pooled. The pooled RNA sample was then divided into two equal parts and used as template for reverse-transcription with the reverse transcripatse (RT+ sample) and without the enzyme (RT- sample). The amount of RNA template in the RT+ and RT- cDNA samples was the same as that in each FECD or control cDNA sample.

2.4.3.3 Quantitative reverse transcription-polymerase chain reaction

Quantitative RT-PCR was performed on a StepOne Plus real-time PCR system (Applied Biosystems, Foster City, CA, USA) using RT^2 $SYBR^R$ Green Rox^{TM} master mix (SABiosciences). A PCR master mix was prepared in the MicroAmp® Fast Optical 96-well reaction plate (Applied Biosystems) for each gene of interest and the *Beta actin* (*ACTB)* reference gene. Each reaction contained 5µl of RT^2 $SYBR^R$ Green Rox^{TM} master mix, 0.4µl of gene-specific forward and reverse primer mix (0.5µM final concentrations each), 2.2µl of autoclaved milliQ water, and 2.4µl of cDNA template. Each sample was analysed in duplicate.

For each gene of interest and the reference gene, the RT+ cDNA standard dilution series was included on each analysed plate; five-fold serial dilutions beginning at 1:20 were made to generate a standard curve, which was used for calculating amplification efficiency (*E*) of the experiment. The RT- cDNA standard individual FECD and normal control cDNA samples were used at 1:20 dilution. A no template control (autoclaved milliQ water) was included to monitor specificity of amplification. Amplification was performed as follows: DNA polymerase was activated at 95°C for 10 minutes. This was followed by 40 cycles of denaturation at

95°C for 15 seconds and annealing and extension at either 60°C, 62°C or 64°C for 1 minute. Data were analysed using the StepOne Plus software.

2.4.3.4 Data analysis

For data analysis, standard curves for each gene of interest and the reference gene were generated from quantitative RT-PCR threshold cycle (C_t) values obtained from the dilution series of the standard sample. The average C_t values of duplicate wells for each dilution was plotted against an arbitrarily estimated corresponding log cDNA concentration. The *E* was calculated using the equation, ($10^{-1/slope}$ of the standard curve) -1), and then converted into percentage.

Normalised expression (NE) of each gene of interest in FECD-affected and unaffected corneal endothelium was calculated using the equation $NE = (E_{target})^{Ct\_target} \div (E_{ref})^{Ct\_ref}$ [174, 175] in the Q-gene software (http://www.gene-quantification.de/download.html). The normalised mean expression (NME) was calculated from the arithmetic mean of the duplicate normalised expression values. Data was expressed as NME ± standard error of mean (SEM).

2.4.3.5 Statistical analysis

Student's *t-test* was performed to investigate any significant difference in NME of the gene of interest between diseased and normal corneal endothelium. The threshold for significance was set at $p = 0.05$. The statistical analyses were performed using the Statistical Package for Social Sciences (IBM SPSS 22, Chicago, IL).

2.4.4 Protein analysis

2.4.4.1 Western blotting

For western blotting, 15µg of human serum, 40µg of MCF7, and 46µg of normal colon protein were used. The human serum and MCF7 samples were were resolved

on 12% gels by SDS-PAGE using the Lamelli method [176]. The colon sample was resolved on 7% gels using the same method as the previous two proteins. Separated proteins were transferred to polyvinylidine fluoride (PVDF) membranes (Bio-Rad). The transfer was conducted in Hoeffer (Bio-Rad) at 4°C overnight at 30 volt. Ponceau staining test was used to indicate successful transfer of the protein. The blot was blocked with 5% skimmed milk in 1× Tris Buffered Saline (0.02M Tris, 0.15M sodium chloride) containing 0.1% Tween 20 (TBST) at room temperature for one hour. It was then hybridised with mouse anti-human ATP1B1 (1:1000) or APOE (1:1000), or rat anti-human LAMC1 primary antibody (1:250) diluted in appropriate volume of 5% skimmed milk/TBST, at room temperature for one hour. The blot was washed washed three times in 1× TBST for 10 minutes each, and hybridised with rabbit anti-mouse anti-ATP1B1 or anti-APOE conjugated to horseradish peroxidase (HRP), or with goat anti-rat anti-LAMC1 secondary antibody that was conjugated to biotin. The secondary antibody hyrbidisation was performed at room temperature for one hour. The blot for LAMC1 immunobloting was washed three times in 1× TBST for 10 minutes each, and hybridised with 1:1000 streptavidin-HRP (SA-RHP; catalog number: SNN4004, Invitrogen) diluted in 1× TBST, at room temperature for 20 minutes. The SA-HRP hybridised blot and the blots hybridised with the rabbit anti-mouse secondary antibodies were washed three times in 1× TBST for 10 minutes each, before detecting the antibody binding. Antibody binding was detected using the chemiluminescence pierce reagents (Thermo Fisher Scientific) in case of LAMC1 or Amersham ECL Prime Western blotting reagents (catalog number: RPN2232, GE Healthcare, Buckinghamshire, UK) for APOE and ATP1B1.The images were taken using the ImageQuant LAS4000 (GE Healthcare).

2.4.4.2 Immunohistochemistry

Paraffin-embedded sections, 4-5µm thick, of FECD and normal corneas were immunolabelled for the APOE, ATP1B1 and LAMC1 proteins. The sections were provided by the Department of Anatomical Pathology (Flinders University of South Australia). Sections were incubated at 60°C overnight to dewax and to promote adherence to the slide. The sections were then washed sequentially, for 3 minutes each, twice in Xylene and in ethanol. Sections were quickly rinsed three times in deionised water, and then in 1× Tris Buffered Saline (TBS; pH 7.6) before being placed in 1% hydrogen peroxide in 50% ethanol for 10 minutes, on a stirrer. This was to quench endogenous hydrogen peroxidase contained in fixed tissues. Slides were removed and washed twice, for 5 minutes each, in 1× TBS (pH 7.6).

Antigen retrieval under alkaline conditions was performed for APOE immunolabelling. For this, slides were placed in retrieval solution (1×Dako Target Retrieval Solution, pH 9; Dako Australia Pty Ltd, Scoresby, VIC, Australia) heated to boiling point in a closed coplin jar, and the jar left for about one hour in a water bath (Model No: NBCT2, Laboratory Equipment Pty Ltd., Marrickville, NSW, Australia) at 100°C. Sections were cooled for one hour at room temperature, and washed in 1× TBS (pH 7.6) before incubation with mouse monoclonal anti-human APOE antibody (1:2000; cat#NE1004; Calbiochem, Merck Pty, VIC, Australia) at 4°C overnight.

Labelling of ATP1B1 and LAMC1 on corneal sections did not require antigen retrieval. Following quenching of endogenous hydrogen peroxidase, the sections were incubated with mouse monoclonal anti-human ATP1B1 (1:1000; clone:M17-P5-F11; cat#PIEMA3-930; Thermo Scientific Pierce) or rat monoclonal anti-human

LAMC1 antibody (1:50; cat# AB80580; Sapphire Bioscience Pty Ltd., Waterloo, NSW, Australia) at 4°C overnight.

The sections were then washed twice in 1× TBS for 5 minutes each, and incubated with the NovoLink Polymer complex reagent (Leica Microsystms, Bannockburn, IL, USA) at room temperature for one hour. Two additional washes were performed for 5 minutes each before primary antibody binding was detected with Liquid DAB+ substrate Chromogen System (K3468; Dako Australia Pty Ltd). Sections were counterstained with haematoxylin and mounted in DePex (Merck KGaA, Darmstadt, Germany). Labelling was imaged on an Olympus BX50 microscope fitted with QImaging Micropublisher RTV 5 Megapixel Digital Camera using QCapture Imaging software (Olympus Corporation, Tokyo, Japan).

CHAPTER 3
Mutation screening of FECD-causing genes in Caucasian Australian cases

3.1 INTRODUCTION

Findings from previous genetic studies indicate that FECD is a multigenic, complex disease [177]. Thus, identification of various mutations in different disease-causing genes is critical to the understanding of the molecular basis of the disease. Currently, mutations in *LOXHD1*, *SLC4A11*, *ZEB1*, and the recently identified *AGBL1* gene are known to cause both familial and sporadic late-onset FECD [39, 76, 78, 79, 83, 178]. However these mutations are responsible for a very small number of cases from the USA and India. The involvement of *LOXHD1* and *AGBL1* in FECD pathogenesis has been recently reported only in one family each [78, 79]. In addition, only one study so far has reported FECD-causing mutations in *ZEB1* [77]. Different rare mutations in *SLC4A11* are reported to cause late-onset FECD in both American and Asian populations [39, 76]. Further studies are required to identify additional novel mutations, and ascertain the proportion of cases affected by mutations in the *LOXHD1*, *SLC4A11*, *ZEB1* and *AGBL1* genes in other populations.

In the present study, I aimed to identify novel mutations in *LOXHD1*, *SLC4A11*, and *ZEB1* in Caucasian Australian cases with advanced late-onset FECD (grade ≥3) [62]. The *AGBL1* gene was not included because its involvement in the disease was reported in October, 2013 [79] after initiation of the study. We employed Ion Torrent Next generation sequencing (NGS) to sequence the targeted genomic regions in the three genes. This technology uses ion semiconductor chip to sequence amplicons of interest. Sanger sequencing and Custom TaqMan SNP genotyping assay were used to validate the results.

3.2 MATERIALS AND METHODS

3.2.1 Ion Torrent AmpliSeq™ sequencing

Ion AmpliSeq™ Library Kit 2.0 (Life Technologies) and Ion AmpliSeq™ Custom Primer Pools were used to prepare barcoded amplicon libraries for sequencing on the Ion Personal Genome Machine® (PGM™). All reagents used in the study were from Life Technologies (Life Technologies, Carlsbad, CA, USA) unless otherwise specified. The study was conducted according to the Ion Torrent AmpliSeq™ workflow shown in Figure 3.1.

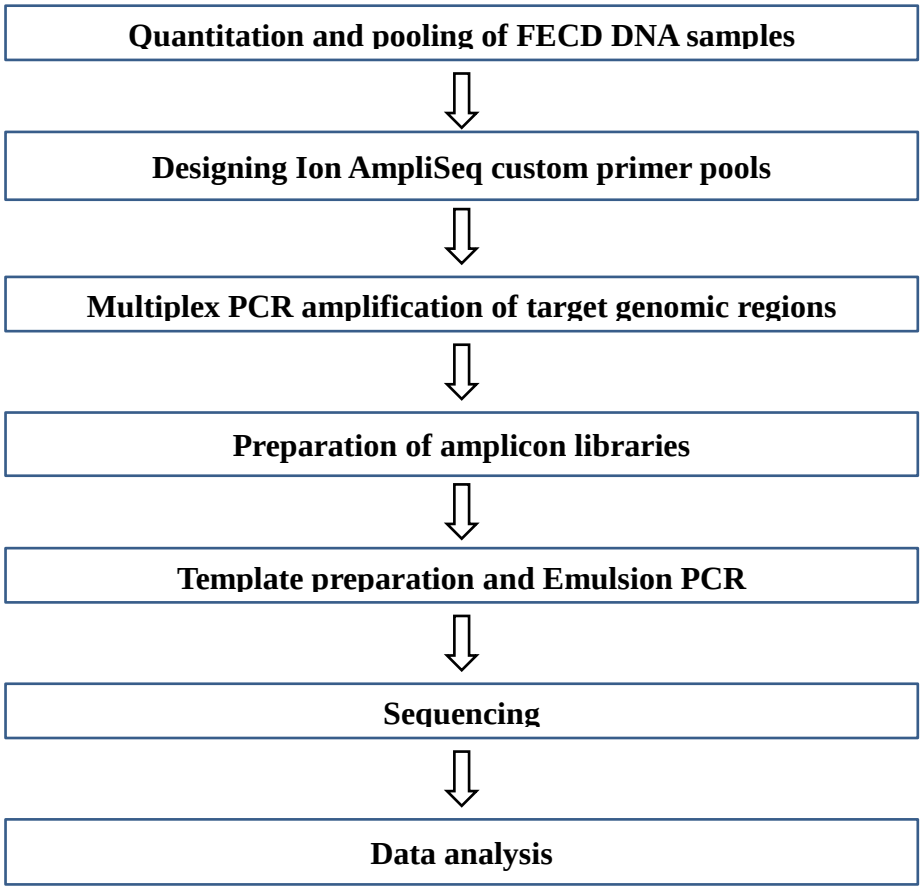


Figure 3.1: Ion Torrent AmpliSeq™ workflow for mutation screening of FECD cases in known FECD-causing genes.

3.2.1.1 Quantitation and pooling of FECD DNA samples

Sixteen initial DNA pools, with a concentration range between 67.4 – 81.5ng/µl in each pool, were prepared as described in Chapter 2 (Section 2.3.1). However, 10ng is the optimal amount for amplification of gDNA template with custom primers in a multiplex PCR for Ion Torrent AmpliSeq^{TM} sequencing. Thus, the pools were diluted to 1:10, and a total of 10ng gDNA was used for library preparation. Table 3.1 lists the details of pooled DNA samples, and concentrations of the stock pools.

Table 3.1 FECD DNA samples pooled for sequencing on the Ion Torrent PGM^{TM}. DNA pool number, FECD samples in the pool and the concentration of each DNA pool stock in ng/µl are given. CDSA, corneal dystrophy South Australia; conc., concentration.

Pool #	CDSA IDs in each pool	DNA conc.in each Stock pool (ng/ul)
3	7, 10, 15, 17, 80, 87, 161, 179	80.4
4	51, 55, 84, 106, 108, 144, 181, 199	81.5
5	34, 49, 63, 64, 110, 176, 184, 185	79.6
6	18, 19, 29, 39, 58, 109, 148, 182	80.8
7	37, 57, 126, 173, 178, 190, 193, 212	78.9
8	13, 14, 24.1, 45, 79, 111, 124, 157	75.1
9	12, 23, 36, 66, 76, 113, 147, 201	76.9
10	8, 20, 30, 74, 112, 136, 137, 172	77.8
11	6, 48, 71, 100, 102, 103, 153, 166	73.4
12	9, 31, 46, 47, 50, 59, 61, 65	76.8
13	67, 69, 72, 75, 85, 86, 88, 89	67.7
14	91, 96, 97, 104, 107, 114, 120, 127	73.9
15	150, 154, 159, 164, 171, 180, 183, 187	67.4
16	192, 194, 195, 198, 200, 202, 213, 214	78.4
17	5, 32, 33, 62, 73, 94, 98, 115	69.1
18	129, 130, 132, 133, 134, 138, 139, 141	67.8

3.2.1.2 Designing Ion AmpliSeq™ Custom Primer Pools

Ion AmpliSeq™ custom primers for amplification of targeted genomic regions in the *LOXHD1*, *SLC4A11* and *ZEB1* genes were designed against the human genome 19 (hg19) reference sequence using the Ion AmpliSeq™ Designer web-based software (pipeline version 1.2) available at www.ampliseq.com. The target region was 18.43 kilobases (kbs), and was comprised of sequences of the exons, exon/intron junctions, and the 5' and 3' untranslated regions (UTRs) of the three genes. The expected average size of amplicons was 159 base pairs. Two custom primer pools (pool 1 = 80 primer pairs, pool 2 = 79 primer pairs), covering 95.54% of the target sequence, were synthesised by Applied Biosystems (Applied Biosystems) and supplied at the standard 2× concentration for direct use.

3.2.1.3 Multiplex PCR amplification of target genomic regions

In order to amplify each pooled DNA template using each primer pool in a multiplex PCR, a 20µl reaction was prepared using 1µl of pooled gDNA (10ng), 4µl of 5× Ion AmpliSeq™ HiFi Master Mix, 10µl of 2× Ion AmpliSeq custom primer pool 1 or 2, and 5µl of Nuclease-free water. The multiplex PCR was performed in the Bio-Rad thermal cycler (Bio-Rad Laboratories Pty. Ltd., Gladesville, NSW, Australia) as follows. The enzyme was activated at 99°C for 2 minutes. This was followed by 18 cycles of denaturing at 99°C for 15 seconds and annealing and extension at 60°C for 4 minutes. Samples were held at 10°C overnight before partial digestion of primer sequences with the FuPa enzyme.

3.2.1.4 Preparation of amplicon libraries

3.2.1.4.1 Partial digestion of primer sequences

Two microlitres of the FuPA reagent was mixed with 20µl of each amplified pooled DNA sample. The mixture was incubated at 50°C for 10 minutes, followed by 55°C for another 10 minutes in the Bio-Rad thermal cycler (Bio-Rad Laboratories Pty. Ltd) to, respectively, activate the enzyme, and partially digest primer sequences. The enzyme was then inactivated at 65°C for 20 minutes.

3.2.1.4.2 Adapter and barcode ligation

A unique individual Ion Xpress™ barcode, (numbers 1 – 16 designated by Life Technologies), was assigned to each library to facilitate combining multiple barcoded libraries on each sequencing chip. Eight microlitres mixture of Ion P1 adapter and Ion Xpress™ barcode was prepared at a final dilution of 1:4 for each adapter by adding 2µl of Ion P1 adapter, 2µl of Ion Xpress barcode, and 4µl of Nuclease-free water. Ligation reaction was prepared by adding 22µl of partially digested amplicon, 4µl of Switch solution, 2µl of diluted barcode-adapter mix, and 2µl of DNA ligase. The barcode and adapter were ligated to the ends of the digested amplicons with ligase in the Bio-Rad thermal cycler (Bio-Rad Laboratories Pty. Ltd) for 30 minutes at 22°C. The ligase was inactivated at 72°C for 10 minutes, and the reaction held at 10°C for 5 minutes.

3.2.1.4.3 Purification of barcoded amplicon libraries

Barcoded libraries were purified with the AgenCourt® AMPure® XP system (Beckman Coulter Genomics, Danvers, MA) according to the manufacturers' protocol. In brief, the AMPure® reagent was equilibrated at room temperature for at

least 30 minutes, and then vortexed thoroughly to disperse beads before use. A 45µl (1.5× sample volume) of the reagent was added to each library and thoroughly mixed by pipetting. The mixture was incubated at room temperature for five minutes to allow the library to bind to the beads before it was transferred to a 1.5ml microfuge tube and placed in a DynaMag® Magnet for 2 minutes, or until the solution cleared. Supernatant was carefully removed and discarded without disturbing the pellet. One hundred and fifty microlitres of freshly prepared 85% molecular biology grade ethanol (catalogue number: E7023-500mL; Sigma-Aldrich, St. Louis, MO, USA) was added to each tube. The tube was then rotated 90° at a time in the DynaMag® Magnet to wash the beads and the solution removed without disturbing the pellet. A second wash using 150µl of 85% ethanol was similarly performed. All ethanol droplets were removed from the tubes, and pellets were briefly air-dried, without over-drying, at room temperature for 5 minutes while still in the DynaMag® Magnet.

3.2.1.4.4 Amplification of purified barcoded amplicon libraries

In order to enrich amplifiable amplicon material and obtain sufficient material for accurate quantification of the libraries, amplicon libraries were first amplified. Platinum® PCR SuperMix High Fidelity solution (50µl) and library amplification primer mix (2µl) were added to each library-bead pellet. The mixture was thoroughly mixed to elute amplicons. The tube was returned to DynaMag® magnet and left for at least 2 minutes before 50µl of the supernatant was carefully transferred to a PCR tube, without disturbing the pellet. To amplify the library, the supernant was incubated at 98°C for 2 minutes in the Bio-Rad thermal cycler to activate the enzyme, and at 98°C for 15 seconds and 60°C for 1 minutes, for five cycles, for

annealing and extension, respectively. The reaction was held at 10°C for 2 minutes to inactivate the enzyme.

3.2.1.4.5 Purification of amplified barcoded amplicon libraries

The libraries were purified with the AMPure® reagent in a two-round purification process according to the manufacturers' protocol. First, 25µl (0.5× bead-to-sample-volume ratio) of the reagent was added to the PCR tube containing the library. The mixture was thoroughly mixed and left at room temperature for 5 minutes to equilibrate. The tube was placed in the DynaMag® magnet for at least 5 minutes, or until solution was completely clear. The supernatant was carefully transferred into a new microfuge tube without disturbing the pellet.

Then, 60µl (1.2× bead-to-original-sample-volume ratio) of the reagent was added to the supernatant. The mixture was repeatedly pipetted to thoroughly mix the bead suspension with amplicons, and incubated at room temperature for 5 minutes to equilibrate. The tube was then placed in the DynaMag® magnet for 3 minutes, or until the solution was clear. This step allowed binding of amplicons to the beads and retaining of the primers in solution. The solution was carefully removed and discarded without disturbing the pellet.

The pellet was washed and air-dried as described in section 3.2.1.4.3. It was then dissolved in 50µl of Ion Torrent Low TE (Tris-EDTA (Ethylenediamine Tetraacetic Acid) buffered solution (Catalogue number: 602-1297-01; Life Technologies) to disperse the beads and elute the amplicons. The solution was thoroughly mixed and incubated at room temperature in the DynaMag® magnet for at least 2 minutes to

pellet beads and elute amplicons. The supernatant was carefully pipetted into a cleaned 1.5ml microfuge tube.

3.2.1.4.6 Quantification of amplified barcoded amplicon libraries

The libraries were quantified using the Qubit 2.0 Fluorometer (Invitrogen) and the Qubit™ dsDNA (double-stranded DNA) HS (High Sensitivity) Assay Kit (Invitrogen) according to the manufacturers' protocol. This kit provided concentrated assay reagent (200× concentrate in DMSO), dilution buffer, and two pre-diluted DNA standards 1 (0ng/µl) and 2 (10ng/ul) in TE buffer. In summary, appropriate number of thin-wall, clear 0.5 mL Axygen PCR-05-tubes (VWR, part no. 10011-830; Axygen) were set-up for quantifying the libraries. A Qubit® working dilution of 1:200 was made with 1µl of dsDNA HS reagent (Invitrogen) and 199µl of dsDNA HS buffer (Invitrogen). Each standard solution was prepared by adding 10µl of the Qubit™ standard into 190µl of Qubit™ working solution in the tube. Additionally, each DNA library assay was prepared by adding 2µl of the library into 198 µl of the working solution. The standard and assay tubes were mixed by vortexing for 2 – 3 seconds, and incubated at room temperature for 2 minutes to equilibrate. The Qubit® 2.0 Fluorometer was calibrated with prepared standard 1 and standard 2 solutions according to the manufacturers' protocol before analysing the samples. The range of concentration expected to be achieved was 10ng/ml to 100,000ng/ml. The sample concentrations are given in Appendix, Table 3.

3.2.1.5 Template preparation and Emulsion PCR

Each amplicon library was diluted to 10pM and 8 libraries were pooled by adding equal amounts from each library to form a sequence reaction or multiplex. Two multiplex samples were thus prepared from 16 different barcoded libraries. The

multiplexed samples underwent emulsion PCR and sequencing at the Flinders Genomic Facility (Flinders Medical Centre, Adelaide, Australia) according to standard protocols.

Briefly, library fragments in multiplex samples were clonally amplified onto the proprietary Ion Sphere™ particles (ISPs) by emulsion PCR with the Ion OneTouch™ 2 System and the Ion OneTouch™ 200 Template Kit v2 according to the manufacturers' procedure (Publication Part Number 4478371 Rev. B Revision Date 13 June 2012). Template-positive ISPs were then enriched with the Dynabeads® MyOne™ Streptavidin C1 Magnetic Beads and washed with Ion OneTouch™ Wash Solution in the kit.

Quality of the template-positive ISPs was assessed by Dr. Renee Smith (Flinders Genomic Facility, Flinders Medical Centre) using a Qubit® 2.0 Fluorometer (Invitrogen) according to the manufacturer's protocol. In summary, Ion probes tube, annealing buffer, and quality control wash buffer were thawed. Enriched ISPs sample volume was adjusted to 100µl in Ion OneTouch™ wash solution, and 10µl of the sample transferred to a 0.2-mL PCR tube. A volume of 19µl of annealing buffer and of 1µl of Ion Probes were added to the tube containing the ISPs, and mixed well by pipetting. The mixture was loaded into a thermal cycler, and incubated at 95°C for 2 minutes, followed by 37°C for another 2 minutes, to anneal the Ion Probes.

The sample was then washed three times with 200µl of quality control wash buffer to remove unbound probes. ISPs were spun down at 15,000 × g for 90 seconds, and all but 10µl of the supernatant was carefully removed so as not to disturb the pelleted ISPs. A 190µl of the quality control wash buffer was added to the tube, and thoroughly mixed by pipetting. The entire volume was transferred to a Qubit® assay

tube for quality assessemnt. A negative control was set up by adding 200µl of quality control wash buffer to a fresh Qubit® assay tube. The test sample and negative control were measured using the Qubit® 2.0 Fluorometer as described in the manufacturer's protocol.

3.2.1.6 Sequencing

Emulsion PCR products were sequenced on the Ion Torrent PGM™ system using the Ion PGM™ 200 Sequencing Kit and Ion 318™ Chip (Life Technologies) according to the established procedures (Publication Part Number 4474596 Rev. B Revision Date 14 July 2012). Data generated were automatically transferred from the Ion PGM™ sequencer to the PGM™ Torrent Server for analysis.

3.2.1.7 Data Analysis

Base calling was performed by Ion Torrent Suite™ software (version 3.4.2). Reads were mapped to human genome sequence (Build GRCh37/hg19) with Torrent Mapping Alignment Program. DNA variants were called with the Ion Torrent Variant Caller plug-in software (version 3.6.6335). Since 8 FECD cases were pooled and sequenced in this study, a heterozygous allele in 1 FECD patient would be present in one sixteenth of the reads, that is ~6% of reads. Somatic workflows configuration was set at low stringency (i.e, the threshold was set to 4% frequency for SNPs and 20% for indels) and used for variant analysis. The threshold percentage for detecting the indels was set at 20 because their occurance rate is ~8-fold lower than that of the SNPs [179]. The 4% frequency filter threshold was measured in unfiltered reads, and should detect at least 90% of SNPs that occur at a frequency of 10% or more in a sample. Variant data were exported as variant call files (VCFs). Variants identified were annotated to dbSNP 137 using the Ion Reporter™ software (version 1.6.2) (https://ionreporter.lifetechnologies.com/ir/), as well as by using an

independent method, SeattleSeq (http://snp.gs.washington.edu/SeattleSeqAnnotation137/), to increase confidence in the annotations. Novel variants identified were prioritised for validation if they were not present in the dbSNP 137 and 1000 Genomes public databases, and predicted with the Sorting Intolerant from Tolerant (SIFT; http://sift.jcvi.org/), and Polymorphism Phenotyping (PolyPhen2; http://genetics.bwh.harvard.edu/pph2/) functional prediction programs to be protein changing.

3.2.2 Validation of identified novel variants by Sanger sequencing

Novel prioritised variants for *LOXHD1* (NM_144612.6) and *SLC4A11* (NM_032034.3) were confirmed by Sanger sequencing according to the standard protocols. Gene-specific primers used for sequencing are listed in Table 2.2 (page 50). PCR amplification was performed in a Bio-Rad thermal cycler (Bio-Rad Laboratories Pty Ltd) as follows. Hotstar Plus Taq® DNA polymerase (Qiagen) was activated at 95°C for 15 minutes followed by 30 cycles of denaturation at 95°C, annealing at 57°C, and elongation at 72°C for 30 seconds each. The final elongation was at 72°C for 5 minutes. DNA of each FECD case from the pool in which the variant was identified was individually sequenced.

3.2.3 Screening of identified potential novel mutations using Custom TaqMan SNP Genotyping Assay

Sequence flanking the variant of interest was submitted in the TaqMan Genotyping SNP assay design tool (http://www.appliedbiosystems.com.au/) to design custom primers and allele specific labelled probes for the allelic discrimination assays. The primers and probes (Table 3.2, page 71) were synthesized and supplied by Applied Biosystems. The assay supplied at 40× concentration was diluted to 20× concentration by adding 188µl of sterilised MQH_2O, and then divided into smaller

aliquots for storage at -20°C for later use. A 10µl final volume for each reaction was prepared with 5µl of 2× TaqMan Genotyping Mix, 0.5µl of 20× SNP Assay Mix, 2µl of 20ng/µl genomic DNA and 2.5µl of MQH_20. A total of 282 NSA controls were screened for each of the potential novel mutations identified, on a StepOne Plus real-time PCR (Applied Biosystems) according to the manufacturers' standard ('fast') protocol. The specific PCR conditions were as follows. The reaction mix was held at 60°C for seconds, then at 95°C for 10 minutes activation. This step was followed by 40 cycles of denaturation at 95°C for 15 seconds, and annealing and elongation at 60°C for 1 minute each. The final elongation was at 60°C for 30 seconds.

Table 3.2 Sequence of probes used in custom TaqMan SNP Genotyping Assays. Gene symbol, name of the assay and the heterozygous novel variant identified in each of the genes are given. The wild-type and the variant alleles are respectively named Reporter 1 base and reporter 2 base; and labeled with VIC or FAM fluorescent dyes respectively. In the DNA sequences, the wild-type allele preceeds the variant allele (in red).

Gene	**Assay Name**	**Reporter 1 base (VIC)/Reporter 2 base (FAM)**
LOXHD1	LOXHD132-GC assay_AHVJJHG	CATCCTCCC[C/G]GATGAGA
	LOXHD1_R2133H assay_AHRSO4Q	AAAATGC[G/A]CAACCTC
SLC4A11	SLC4A11_S-3	CACACTTGTAGTAGCC[T/C]AGAGA

3.2.4 Prediction of effects of identified novel mutations

Effects of identified novel missense variants were predicted with the Sorting Intolerant from Tolerant (http://sift.jcvi.org/), and Polymorphism Phenotyping (http://genetics.bwh.harvard.edu/pph2/) functional prediction programs. Additionally, evolutionary conservation of mutant amino acid residues among

orthologous genes was assessed using CLUSTALW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

3.3 RESULTS

3.3.1 Mutation screening

Screening of 128 cases for FECD-causing mutations in the *LOXHD1*, *SLC4A11* and *ZEB1* genes was sucessfully conducted using the Ion Torrent NGS. Coding exons, exon-intron junctions and untranslated regions of the three genes were sequenced from genomic DNA samples. Analysis of sequencing data revealed credible sequencing data as demonstrated by the coverage statistics presented in Table 3.3 (Page 73).

The first column in the table indicates the number of each sequenced FECD DNA pool, and an average ± standard deviation (SD) of each parameter corresponding to the pools. The mapped reads represent total number of reads in each DNA pool that have been correctly aligned to the reference genome. Percentages of the ratio of number of reads within a target region to the total number of mapped reads for each DNA pool are given in the third column. Similarly, the ratio of number of on-target bases to the total aligned bases for each DNA pool has been expressed as percantages, and presented in the fourth column. Read depth in the fifth column shows the total number of times a particular base was read during the sequencing process of each DNA pool. Thus, a 1×, 20×, and 100× coverage columns given in the sixth, seventh and eighth column, respectively, correspond to the number of times, on an average, each base has been respectively read by 1, 20 and 100 sequences for each pool. The last column presents the number of variants identified in each sequenced DNA pool.

Table 3.3 Sequencing results of pooled FECD DNA samples using Ion Torrent PGM™ sequencer. SD = standard deviation.

Description	Mapped Reads	Reads On-Target	Bases On-Target	Read Depth	Coverage 1x	Coverage 20x	Coverage 100x	Variants Detected
Pool 3	626,968	79.79%	81.78%	2,350.52	99.87%	98.76%	98.12%	49
Pool 4	830,547	81.51%	83.33%	3,180.61	99.89%	98.91%	98.54%	26
Pool 5	549,262	84.91%	86.19%	2,230.74	99.88%	98.88%	97.72%	27
Pool 6	686,908	85.43%	86.67%	2,814.78	99.87%	98.72%	97.73%	27
Pool 7	663,406	85.44%	86.90%	2,668.23	99.88%	98.91%	98.17%	52
Pool 8	731,758	82.36%	84.00%	2,899.71	99.86%	99.09%	98.53%	49
Pool 9	843,485	82.32%	83.88%	3,257.20	99.84%	99.09%	98.51%	59
Pool 10	512,513	84.32%	85.67%	2,030.21	99.85%	98.76%	97.96%	42
Pool 11	675,718	80.24%	81.79%	2,728.54	99.53%	98.81%	98.38%	38
Pool 12	499,889	83.99%	85.33%	2,094.49	99.93%	98.65%	98.36%	36
Pool 13	473,298	86.04%	87.25%	2,032.22	99.82%	98.64%	98.35%	35
Pool 14	525,860	84.89%	86.89%	2,245.95	99.83%	98.80%	98.20%	32
Pool 15	478,322	86.67%	87.80%	2,089.32	99.75%	98.73%	97.73%	27
Pool 16	474,167	86.02%	87.14%	2,039.02	99.83%	98.68%	97.74%	39
Pool 17	453,517	83.79%	85.36%	1,887.80	99.48%	98.73%	97.71%	32
Pool 18	490,823	77.58%	80.10%	1,875.11	99.59%	98.71%	98.23%	25
Average ± SD	594,778 ± 129,926	83.46% ± 0.026	85.01% ± 0.023	2,401.53 ± 456.84	99.79% ± 0.001	98.80% ± 0.001	98.12% ± 0.003	37.19 ± 10.48

In summary, the sequencing results indicated that on an average 594,778 ± 129,926 mapped reads were on target genomic regions sequenced. The variation in the number of mapped reads among the pools was 129,926, and reflected a high variability in concentrations of the DNA libraries. Importantly, similar high average percentages of reads or bases on-target were achieved for each DNA pool thus leading to good read depth, with an average of 2,401 ± 456.84, in each DNA pool (Table 3.3, page 73). This is an average read depth of 300 for each person in the pool. Similarly, 98.12% ± 0.003 of the target region is covered at least 100 fold. have an average read depth of 2,401 ± 456.84. This coverage of the target region indicates the reliablity and sensivitiy of the sequencing assay. Across the 16 FECD DNA pools sequenced, an average of ~37 ± 10 variants were detected per pool. This is an average read depth of 300 for each person in the pool. Similarly, 98.12% ± 0.003 of the target region is covered at least 100 fold. have an average read depth of 2,401 ± 456.84. This coverage of the target region indicates the reliablity and sensivitiy of the sequencing assay. Across the 16 FECD DNA pools sequenced, an average of ~37 ± 10 variants were detected per pool.

Annotation of the identified variants to dbSNP 137 and 1000 Genomes revealed the presence of SNPs and novel variants. A total of 51 SNPs were revealed, and found to make up the majority of identified variants. Twenty novel variants were identified and are given in Table 3.4, (page 76 – 77). Four variants were heterozygous missense changes, and all located in the *LOXHD1* gene. The remaining comprised one heterozygous splice variant in the 3'end of the *SLC4A11* gene, 2 coding-synonymous variants each in the *LOXHD1* and *ZEB1* genes, seven intronic variants 3 each in *LOXHD1* and *SLC4A11*, and one in *ZEB1*, one 5'UTR and two 3'UTR variants in *ZEB1*, and three 3'UTR variants in *LOXHD1*.

All the missense variants, and the 3’splice variant were prioritised for validation assuming that they would likely be pathogenic.

Table 3.4 Details of the 20 novel variants identified by next generation sequencing in the *LOXHD1* (accession no: NM_144612.6), *SLC4A11* (accesion no: NM_032034.3) and *ZEB1* (accession no: NM_030751.5) genes. The annotation was performed using the SeatleSeq program and Ion Reporter. Each gene symbol and its chromosomal position, reference allele and genotypes of cases harbouring the novel variant/s, the type of variant and/or encoded amino acid, DNA pool number in which variant identified, and the pool and variant allele coverages in reads are given. *LOXHD1 = lipoxygnase homology domains*; *SLC4A11 = solute carrier family 4, sodium borate transporter, member 11*; *ZEB1 = zinc finger E-box binding homeobox 1*; UTR = untranslated region; Chr = chromosome.

Gene	Chr:position	Reference Allele	Genotypes in cases	Variant identified and encoded amino acid position	Encoded amino acid position/exon	Pool # in which identified	Pool coverage in reads	Variant Allele coverage in reads
LOXHD1	18:44057339	C	C/T	3'UTR		12	1418	T = 71
	18:44057362	T	A/T	3'UTR		12	1523	A = 86
	18:44057416	G	A/G	3'UTR		12	1720	A = 92
	18:44057643	C	A/C	c.6428G>T	p.R2143L	13	1459	A = 82
	18:44068985	C	C/T	c.5813G>A	p.R1938H	6	1999	T = 98
	18:44087453	C	C/T	intron		11	1962	T = 100
	18:44102192	C	C/G	c.4957G>C	p.G1653R	18	1331	G = 98
	18:44137737	T	C/T	intron		8	1786	C = 117
	18:44152027	C	C/T	intron		9	2000	T = 105
	18:44157732	G	A/G	coding-synonymous	636	9	1997	A = 115
	18:44229159	A	A/C	c.204T>G	p.F68L	16	726	C = 482
SLC4A11	20:3209942	C	C/T	intron		16	530	T = 24
	20:3215279	G	A/G	intron		14	2000	A = 140
	20:3215375	G	A/G	intron		14	1985	A = 140
	20:3215542	T	C/T	3'splice site; c.4752A>G	(IVS2-2A>G)	6	2000	C = 106
	10:31608152	A	A/T	5'UTR		4	39	T = 8

ZEB1	10:31812857	C	C/G	intron		6	1999	G = 117
	10:31815823	G	A/G	coding-synonymous	1002	14	2000	A = 118
	10:31816722	A	A/G	3’UTR		16	1727	G = 83
	10:31817929	T	T/G	3’UTR		5	2000	G = 130

3.3.2 Validation of prioritised novel variants

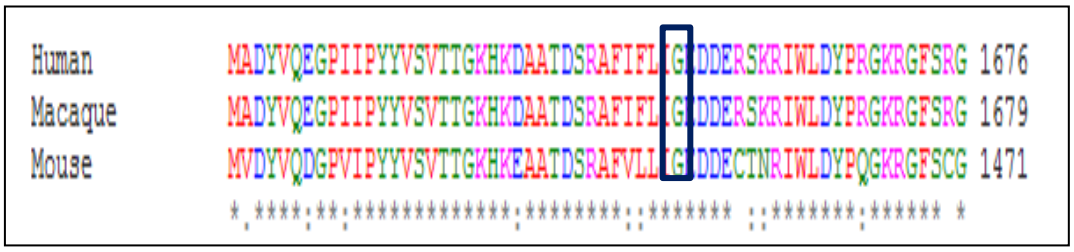
Validation of c.6428G>T (p.R2143L), c.5813G>A (p.R1938H), c.4957G>C (p.G1653R), c.204T>G (p.F68L) and c.475-2A>G (IVS2-2A>G) variants was performed in individual FECD cases from the DNA pools in which the variants were identified. This was to confirm whether the variants were positive findings, and to identify the cases carrying them. The variants were identified in one pool each (Table 3.4, page 76 – 77). Gene specific primer pairs listed in Table 2.2 (In Chapter 2, section 2.4.1.1) were used to partially sequence exons 2, 32, 37, and 40 in *LOXHD1*, and the 3'splice site in the *SLC4A11* gene. Sequencing results are summarised in Table 3.5.

Table 3.5 Results of individual validation of novel variants identified in individual FECD cases by Sanger sequencing. The table indicates symbols and accession numbers of the genes containing the variants, which underwent validation. The identified variants, the case DNA pool number in which the variants were identified, and the validation outcome are given. *LOXHD1 = lipoxygnase homology domains*; *SLC4A11 = solute carrier family 4, sodium borate transporter, member 11*

Gene symbol	Accession number	Variant identified	Pool # in which variant identified	Validation outcome
LOXHD1	NM_144612.6	c.204T>G (p.F68L)	16	SNP rs34763877
		c.4957G>C (p.G1653R)	18	confirmed
		c.5813G>A (p.R1938H)	6	confirmed
		c.6428G>T (p.R2143L)	13	Not confirmed
SLC4A11	NM_032034.3	c.475-2A>G (IVS2-2A>G)	6	confirmed

One of the potential missense variants was not found to be present in individual FECD cases from the relevant DNA pools. This suggests that the variant (p.R2143L) was a false positive. The c.204T>G (p.F68L) variant was confirmed to be an SNP rs34763877. The other two heterozygous missense variants (c.4957G>C (p.G1653R) and c.5813G>A (p.R1938H)) and a heterozygous 3'splice site variant (c.475-2A>G (IVS2-2A>G)) were confirmed to be present in patient CDSA-141, CDSA-39 and CDSA-29, respectively. Sequence chromatograms of the mutations identified in individual cases, and the illustrations of evolutianry conservation of the residues are given in Figure 3.2 (page 80), Figure 3.3 (page 80), and Figure 3.4 (page 81). Figure 3.2 shows the c.4957G>C base change, which would result in amino acid change from glycine (G) to arginine (R) at position 1653. The c.5813G>A transition in CDSA-39 individual (Figure 3.3) resulted in amino acid change from arginine to histidine (H) at position 1938. Alignment of LOXHD1 sequences from different species by ClustalW2 program have shown both mutated residues, p.G1653 and p.R1938, to be evolutionarily conserved in human, macaque and mouse (Figure 3.2 and 3.3). A novel variant, c.475-2A>G (IVS2-2A>G; 2 bases upstream of the exon 3 splice site), was confirmed in the 3' acceptor splice site of *SLC4A11* exon 3 in patient CDSA-29 (Figure 3.4). The G nucleotide leads to substitution of the adenosine nucleotide in the invariant AG splice acceptor consensus sequence; leading to GG dinucleotide. The change is likely to result in inactivation of the 3' splice acceptor site, potentially causing skipping of the exon 3. This might lead to production of an abnormal SLC4A11 protein.

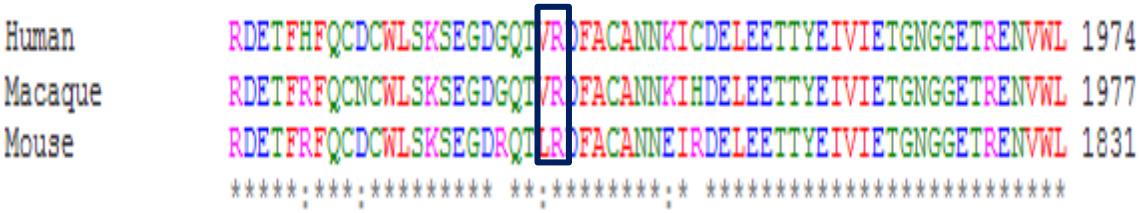
A. **B.**



C.

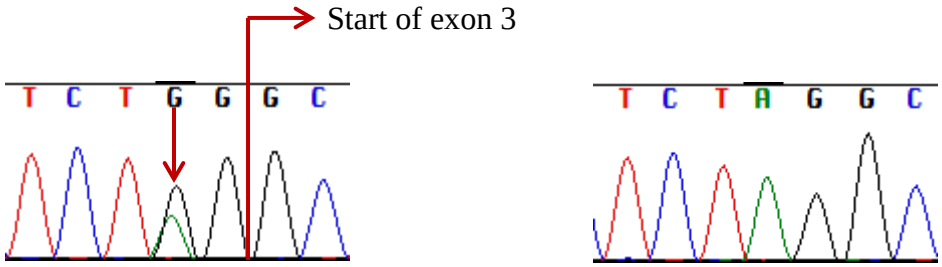
Figure 3.2 Chromatograms of sequences of the genomic region containing c.4957G>C (p.G1653R) variant in exon 32 of *LOXHD1* in patient CDSA-141 (A) when compared with patient without the variant (B). **C** shows that the p.G1653 residue is conserved in human, macaque and mouse LOXHD1 orthologs (protein accession number: F5GZB4).

A. **B.**



C.

Figure 3.3 Chromatograms of sequences of the genomic region containing c.5813G>A (p.R1938H) variant in exon 37 of *LOXHD1* in patient CDSA-39 (A) when compared with patient without the variant (**B**). **C** shows that the p.R1938 residue is conserved in human, macaque and mouse (protein accession number: F5GZB4).



A. **B.**

Figure 3.4 A chromatogram of sequences of the genomic region containing c.475-2A>G (IVS2–2A>G) variant observed in the splice site of *SLC4A11* exon3 in patient CDSA-29 (A) when compared with a patient without the variant (B).

PolyPhen2 and Sift programs were used to predict the functional significance of the allele replacement for the p.G1653R and p.R1938H missense mutations in the *LOXHD1* gene. PolyPhen2 predicted both mutations as deleterious, each with a total score of 1.000 (specificity: 1.000; sensitivity: 0.00). Consistently, Sift predicted both mutations as deleterious, with a total score of 0.67 and 0.74 for the p.G1653R and p.R1938H, respectively. To the best of our knowledge, these potential mutations have not been reported before in any other study in FECD. In addition, this study is the second to report that mutations in *LOXHD1* contribute to late-onset FECD.

3.3.3 Screening of the confirmed variants in NSA controls

To investigate whether the three confirmed novel variants are not common polymorphisms, but mutations responsible for late-onset FECD, we screened the variants in 282 NSA controls. None of the controls showed the new variants which

were clearly detectable in the CDSA-141, CDSA-39 and CDSA-29 patients, as shown in Figure 3.5, 3.6 (page 83), and 3.7 (page 84), respectively.

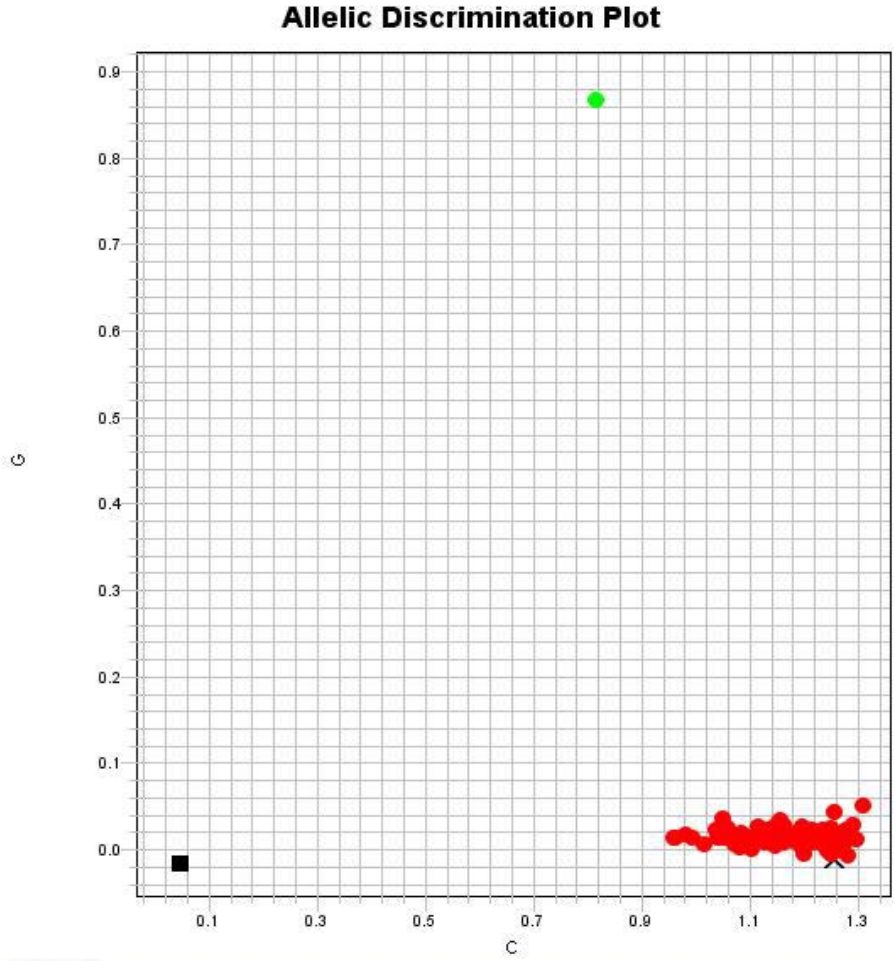


Figure 3.5 Allelic discrimination plot of p.G1653R mutation in 282 NSA controls. X-axis represents C allele (reference allele) and Y-axis represents G allele (mutated allele). As shown at the extreme ends of the respective axes, the NSA controls have the C allele (red dots) and the CDSA-141 FECD patient (CDSA-141) has the heterozygous G allele (green dot). The black square represents no template control. The black x indicates incomplete selection of an NSA control that is supposed to have the C allele (red dots).

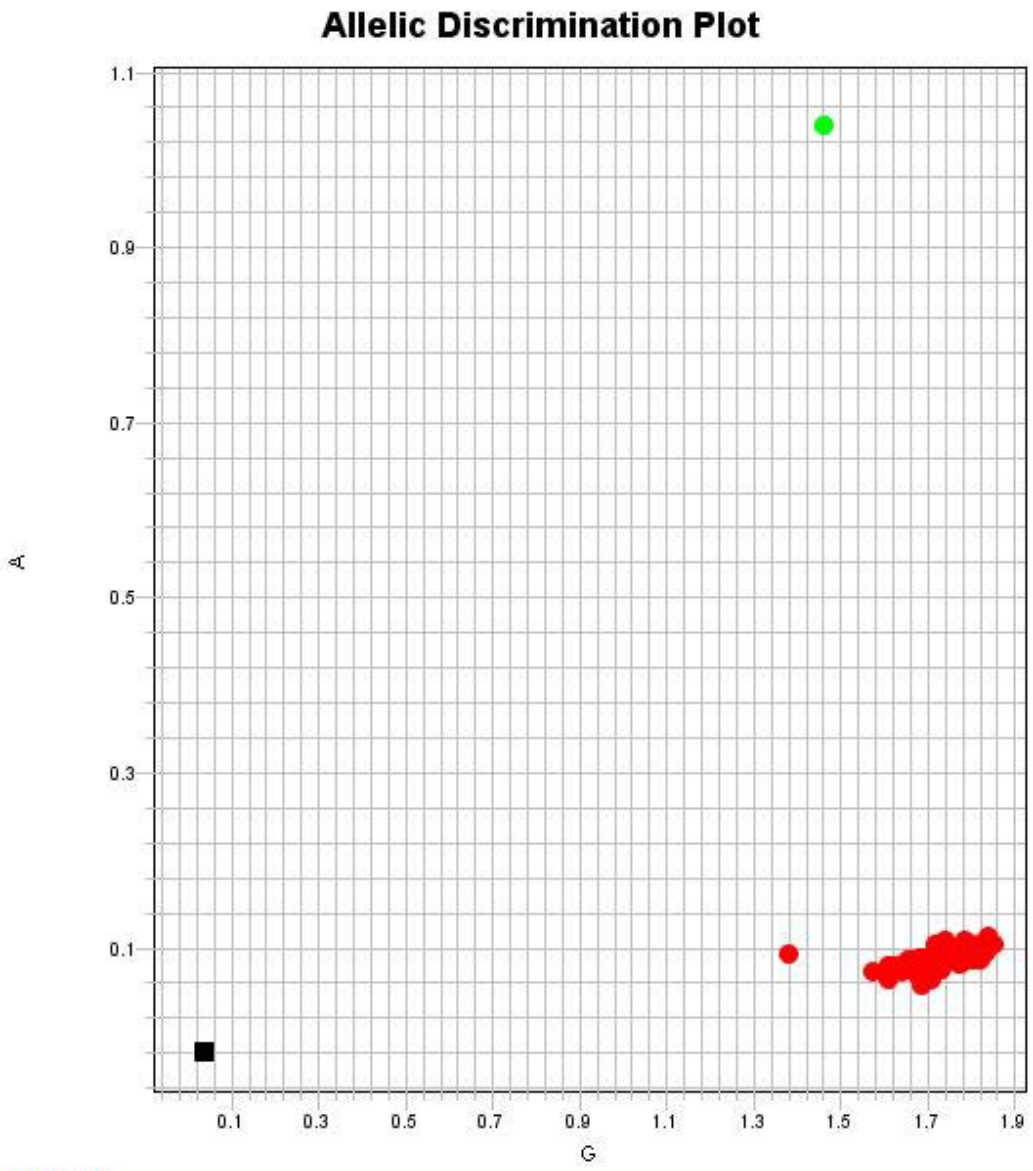


Figure 3.6 Allelic discrimination plot of p.R1938H mutation in 282 NSAcontrols. X-axis represents G allele (reference allele) and Y-axis represents A allele (mutated allele). The NSA controls have the G allele (red dots) and the CDSA-39 FECD patient (CDSA-39) has the heterozygous A allele (green dot) are shown at the extreme ends of the respective axes. The black square represents no template control.

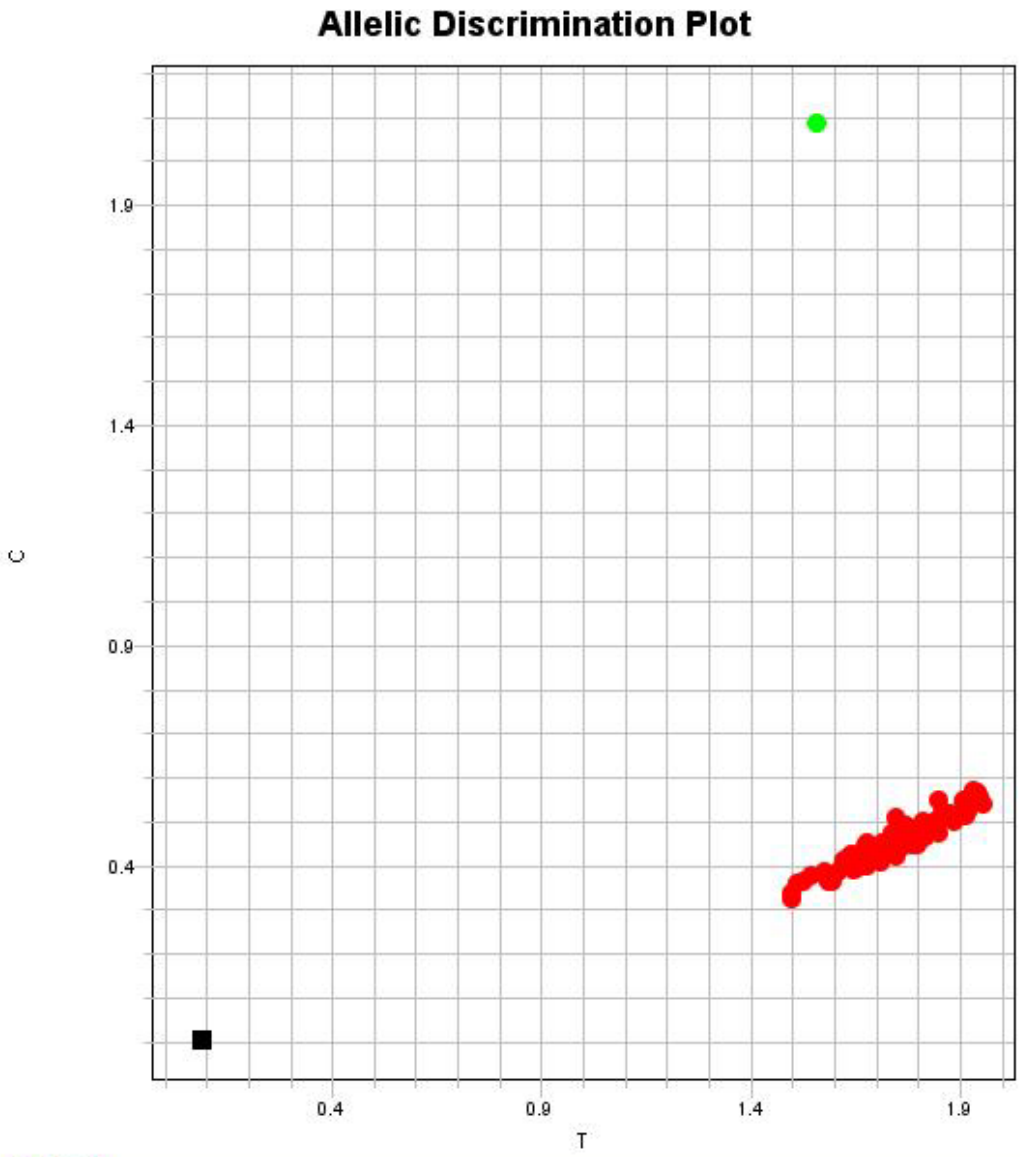


Figure 3.7 Allelic discrimination plot of IVS2–2T>C mutation in 282 NSA controls. X-axis represents T allele (reference allele) and Y-axis represents C allele (mutated allele). As shown at the extreme ends of the respective axes, the NSA controls have the T allele (red dots) and the CDSA-29 FECD patient (CDSA-29) has the heterozygous C allele (green dot). The black square represents no template control.

3.4 DISCUSSION

In previous studies in non-Australian populations, rare mutations in the *LOXHD1*, *SLC4A11* and *ZEB1* genes have been identified to be responsible for FECD in a small proprotion of cases [39, 76, 78, 83]. This study aimed to identify novel mutations in these genes in Caucasian Australians with advanced late-onset FECD. The study revealed 20 novel variants in the three genes. These variants consisted of 4 missense, 2 coding-synonymous, a 3'splice site, 7 intronic, a 5'UTR, and 5 3'UTR changes. With the exception of 4 heterozygous missense variants in the *LOXHD1* gene and the splice site variant in the *SLC4A11* gene, the remaining variants were not prioritised for individual validation since they were less likely to be pathogenic and contribute to the disease.

Individual validation confirmed c.4957G>C (p.G1653R) and c.5813G>A (p.R1938H) in *LOXHD1*, and c.475-2A>G (IVS2-2A>G) in *SLC4A11* as potential pathogenic mutations in patients CDSA-141, CDSA-39 and CDSA-29, respectively. These mutations were absent in 282 normal controls, thus strengthening the case that they contribute to FECD in the three individuals. The c.4957G>C (p.G1653R) and c.5813G>A (p.R1938H) mutations affect conserved residues in orthologous genes, and were predicted to be deleterious. A functional effect of the c.475-2A>G (IVS2-2A>G) mutation remains to be determined.

LOXHD1 encodes an evolutionarily, highly conserved protein that is predicted to consist of 15 PLAT (polycystin-1, lipoxygenase, alpha-toxin) domains [129]. The biological function of PLAT domains is not well established. They are predicted to target proteins to the plasma membrane, and in some cases mediate protein-protein

interactions [129]. The p.G1653R and p.R1938H residues lie in PLAT domains 11 and 14, respectively.

Riazuddin et al. demonstrated that a heterozygous missense c.1639C>T (p.R547C) mutation in *LOXHD1* produces mutant LOXHD1 in the corneal endothelium and Descemet's membrane of the FECD patients with the mutation [78]. Their examination of corneal sections from a FECD patient with the mutation and another without the mutation, and an individual with keratoconus revealed a distinct increase in labelling for LOXHD1 in the endothelium and Descemet's membrane of the patient with the mutation when compared to the endothelium and Descemet's membrane from the other patients [78]. In addition, they observed multiple aggregate stainings in the endothelium, and prominently increased protein abundance in the Descemet's membrane only in the corneal sections from FECD patient harbouring the *LOXHD1* mutation [78]. These findings suggest that missense mutations in *LOXHD1* lead to aberrant accumulation of the protein in the corneal endothelium and Desecemet's membrane in FECD cases.

Based on their mutation screening and immunohistochemical data, Riazuddin and colleagues proposed that missense mutations in the *LOXHD1* gene contribute to FECD, most likely through a mechanism of protein aggregation in the corneal endothelium and Descemet's membrane of the affected individual [78]. In this context, immunohistochemical analysis of corneal sections from our FECD cases with p.G1653R and p.R1938H mutations to investigate presence of aggregation defects might reveal the mechanism of the disease related to the mutations.

The IVS2-2A>G mutation in *SLC4A11* affects conserved AG dinucleotide at the 3' splice site [180]. The AG acts as the splice acceptor in pre-mRNA splicing, and is

required for the second step of splicing [180]. For some introns, referred to as 'AG dependent', the AG dinucleotide is also essential for the earliest step/s in spliceosome assembly and the first step of splicing [181, 182]. It is reported that an existence of AG-dependent introns strongly suggests that the splicing factor $U2AF^{35}$ must recognise the AG dinucleotide during initiation of spliceosome assembly [180]. Changing AG to either GG or UG reportedly abolishes crosslinking of $U2AF^{35}$ [77, 180, 183], leading to aberrant splicing. Thus, aberant splicing may occur in CD-29 FECD patient harbouring the identified IVS2-2A>G mutation in the *SLC4A11* gene.

We postulate that the IVS2-2A>G mutation may lead to direct splicing of exon 2 to exon 4, thus skipping of exon 3 in mature mRNA transcript of S*LC4A11*, and generation of mutant SLC4A11 protein. Investigation for the presence of inappropriately spliced mRNA in the CDSA-29 patient by RT-PCR can be used to prove the functional effect of this mutation on splicing.

Each of the novel mutations identified in this study was found in only 1 of the 128 cases. This implies that only very rare mutations in *LOXHD1* and *SLC4A11* contribute to FECD pathogenesis in affected Australian Caucasians. The present study revealed no novel variant in the coding/splicing regions in the *ZEB1* gene in any of the screened cases. Therefore, mutations in this gene do not contribute to FECD in Caucasian Australian cases, at least in the screened cases. My findings are consistent with prevoius studies reporting contribution of rare pathological mutations in *LOXHD1* and *SLC4A11* to late-onset FECD in a small proportion of affected American Caucasians [39, 78, 83]. Collectively, this and previous studies strongly suggest that additional FECD-causing genes remained to be identified.

In conclusion, this study led to identification of three novel mutations in *LOXHD1* (p.G1653R and p.R1938H) and *SLC4A11* (IVS2-2A>G) in 3 of 128 unrelated Caucasian Australians with late-onset FECD. Both p.G1653R and p.R1938H were predicted to be pathogenic, and could cause FECD. The proportion of our cases in whom these mutations were identified is <1%, suggesting that they are rare mutations. Further studies might be required to functionally characterise biological effects of the novel mutations, and the mechanisms of their involvement in the disease.

CHAPTER 4
Identification of novel genes associated with late-onset FECD in Caucasian Australian cases

4.1 INTRODUCTION

Fuchs' endothelial corneal dystrophy is a complex disease, and heterogeneity underlies its genetic determinants [79]. Studies performed using the candidate gene approaches have identified mutations in the *AGBL1*, *LOXHD1*, *SLC4A11* and *ZEB1* genes as causes of FECD in non-Australian populations [39, 76, 78, 79]. However, these genetic mutations are responsible for a small proportion of cases in the examined populations. Thus, it is likely that additional genes are responsible for the disease. The research being reported in this chapter was based on the hypothesis that further novel genes are associated with FECD in Caucasian Australian cases and could be identified through a genome-wide association study (GWAS).

GWAS is an effective and an unbiased approach for finding genetic determinants of complex traits or diseases such as FECD [90]. They can effectively identify genetic loci associated with a disease or trait by interrogating the entire genome for association/s of genetic markers with the phenotype of interest rather than just focusing on selected candidate genes [90].

The SNP is the genetic marker used for GWAS; since it is the commonest form of variation found within the genome and current technology permits economical genotyping of millions of SNPs in a single experiment [91]. Previous studies on ocular and non-ocular diseases have successfully applied GWAS to identify SNPs associated with pathogenesis of those diseases [184-186]. In FECD, SNPs in *TCF4*

were found to be significantly associated with development of the disease in American and Chinese populations [92, 95].

The concept of linkage disequilibrium (LD) is fundamental to a GWAS. It enables the majority of variations within the genome to be assessed by genotyping only a subset of the SNPs. Thus, the ability to identify the casual variant depends on the degree of LD between the marker and the causal SNPs, genome coverage by genotyped SNPs, and more importantly on the power of the GWAS [187]. The power relies on sample size, effect size and frequency of the risk variant, and the correction for multiple testing [187].

In order to maximise detection of SNPs in LD, a large-scale genotyping is required. The GWAS employ 'chip'-based technologies, in which arrays of hybridisation assays are embedded. Due to the large number of hypotheses being tested , a *p* value of $< 5 \times 10^{-8}$ is generally accepted as genome-wide significant following Bonferroni correction for 1 million SNPs [88]. Achieving such a threshold requires a well-powered study.

The aim of this study was to identify novel genes associated with FECD in Caucasian Australians with advanced disease. To achieve this aim, a GWAS was performed on pooled DNA of cases and of NSA controls to identify top candidate SNPs associated with the disease.

4.2 METHODS

4.2.1 Genotyping of DNA pools

DNA pools were prepared from 94 cases and 216 NSA controls as described in Chapter 2 (Section 2.3.1). Genotyping of the case and control DNA pools and

analysis of the data were undertaken at the Department of Genetics and Population Health, Queensland Institute of Medical Research (QIMR), Berghofer Medical Research Institute, Brisbane, QLD, Australia. The methods have been described in the study by Lu *et al* [188]. Briefly, each DNA pool was genotyped in triplicate on Illumina Human 1M-Duo V3 arrays. The output of the raw red and green bead scores from the genotyping stage was available for data analysis. Same data processing protocol was applied to both pools, using the method described by Brown *et al.* [189].

Before calibrating the raw scores, SNPs with more than 10% negative scores on each array were excluded, as well as the SNPs with the sum of mean red and green scores across each array less than 1200. A normalisation/correction factor (corr) was calculated by forcing the mean value of the pooling allele frequency to be 0.5 over all SNPs on each stripe per array. The pooling allele frequency (PAF) was then estimated based on the raw red intensities and the corrected green intensities for all the SNPs [PAF = red/(red + green/corr)]. A final set of autosomal SNPs met the following criteria: more than 5 probes in each pool; with a minor allele frequency (MAF) greater than 1%; without a significant variance difference between case and control pools, was taken forward to a linear regression model [190]. The underlying concept was to regress the pooling allele frequency on the case/control status for each SNP and estimate the pooling error across all the SNPs, as described in previous studies [190, 191]. The *p* value from comparing the test statistic described by MacGregor *et al* was computed to assess the significance of allele frequency difference between the two pools [190].

4.2.2 Individual genotyping of candidate SNPs

All genotyping in individual FECD, BMES and NSA samples was performed using iPLEX Gold chemistry (Sequenom Inc., Herston, QLD, Australia) on a Sequenom MassArray® Spectrophotometer (Sequenom Inc.) at the Australian Genome Research Facility (Brisbane, QLD, Australia). Genotypes were analysed using PLINK program [173]. Chi-square (χ^2) and/or Fisher's exact tests were conducted to assess inheritance models.

4.2.3 Data analysis

For association of SNPs with FECD in the studied cohorts, $p<5 \times 10^{-08}$ was accepted as significant for the genome-wide genotyping, whilst $p<0.05$ was considered significance in the replication studies. Analysis of haplotypes with >1% frequency was also performed. P-values and odds ratios were calculated for each haplotype compared with all other haplotypes.

4.3 RESULTS

4.3.1 Pooled GWAS and replication findings

Following analysis of data from the GWAS on pooled DNA of cases and NSA controls, data for 10,000 SNPs with genome-wide association values at $p\leq0.01$ was received from the Australian Genome Research Facility. All these SNPs were used to construct the Manhattan plot for the GWAS of the discovery cohort (Figure 4.1, page 93). Of the top 10,000 SNPs, 1,541 failed quality control test, and 4,223 passed it. The remaining 4,236 SNPs were reported as NA (not available), because the quality control on their calls were not reported due to missing or low quality data. Thus, it was not possible to know whether they would pass the stringent quality control test.

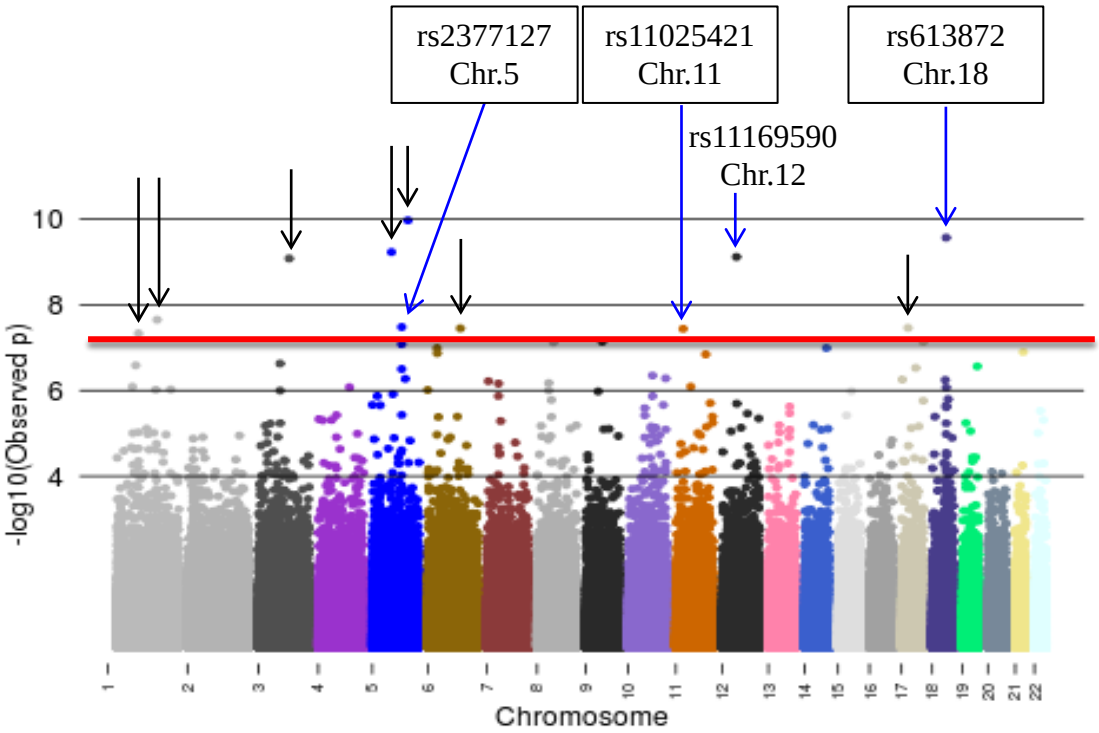


Figure 4.1: Manhattan plot from pooled GWAS of FECD-NSA cohort. The -log10 of the p value of association (y-axis) has been plotted against position in the genome (represented as chromosome on the x-axis). Each chromosome is depicted in a different colour. Four genomic regions on chromosomes 5, 11, 12 and 18 include SNPs that exceeded the genome-wide significance value of $p < 5 \times 10^{-8}$ (indicated by a horizontal red line). Of these, only four (indicated with blue arrows, and named) passed the quality control test. The rest, pointed out with black arrows, have either failed or have NA (not available) report for quality control.

Of the SNPs that passed quality control, only 4 (rs613872, rs11169590, rs2377127, and rs11025421 on chromososmes 18, 12, 5 and 11, respectively) met the genome-wide significance of $p<5 \times 10^{-8}$ (Table 4.1). The first three SNPs were intronic, and respectively found in the *TCF4*, *TMPRSS12*, and *DTWD2* genes. The SNP rs11025421 was found at the 5' flanking region in the *DBX1* gene. The Manhattan

plot (Figure 4.1, page 93) illustrates these four genomic regions on chromosomes 18, 12, 5 and 11, which contain these SNPs.

Seven additional SNPs, rs3776319, rs168752, rs705224, rs11576306, rs4646794, rs9480703 and rs1333971, exceeded the genome-wide significance of $p<5 \times 10^{-8}$, and are also shown above the horizontal red line illustrated in the Manhattan plot (Figure 4.1). The first two SNPs reportedly failed the quality control test and were not considered any further. The remaining five SNPs were reported to have NA report for quality control, and are given in Table 4.1 (page 95 – 97).

To validate the findings from the GWAS, the four SNPs that passed the quality control test, and met the genome-wide significance threshold were considered for individual genotyping in the FECD-NSA discovery cohort. Since these SNPs were few, the five SNPs with reported NA quality control and genome-wide significance of $p<5 \times 10^{-8}$ were included for validation. The number of the prioritised SNPs for validation was still small, therefore additional SNPs were considered for inclusion.

Given that the number of pooled cases in the discovery cohort was small, the study was somewhat under-powered and SNPs that may be significantly associated with FECD could fall short of reaching the genome-wide significance threshold.

Table 4.1. Results of allelic association between FECD and the top 66 candidate SNPs identified through GWAS and genotyped in Caucasian Australians. Results of individual genotyping in the discovery cohort, replication in an independent cohort, and combined analysis of results from the individual genotyping and the replication study are shown. Significant associations are in bold font. SNP = single nucleotide polymorphism; CHR = chromosome; Allele 1 = minor allele; GWAS = genome-wide association study; IG = individual genotyping; RS = replication study; OR = Odds ratio; and Combined analysis = IG plus RS.

SNP	CHR	Gene	Allele 1		p-value (OR)			
			Pooled GWAS	IG/RS	Pooled GWAS	Individual genotyping	Replication study	Combined analysis
rs613872	**18**	***TCF4***	**A**	**G**	**2.69×10^{-10}**	**2.47×10^{-11} (3.62)**	**3.42×10^{-12} (3.75)**	**4.90×10^{-29}**
rs11169590	12	*TMPRSS12*	A	G	7.62×10^{-10}	0.004 (0.2)	0.2279 (1.5)	0.127
rs705224	3	*IGSF11*	T	T	8.25×10^{-10}	0.0897 (1.35)	0.2143 (0.72)	0.433
rs11576306	1	*FMO4*	A	G	2.23×10^{-8}	0.0007 (0)	0.3064 (0.6)	0.0007
rs2377127	5	*DTWD2*	A	C	3.30×10^{-8}	0.0012 (0.26)	0.3481 (1.34)	0.086
rs4646794	17	*ALDH3A2*	A	A	3.45×10^{-8}	0.265 (0.32)	0.3774 (0)	0.386
rs9480703	6	*QRSL1*	C	C	3.54×10^{-8}	0.0076 (0.17)	0.1364 (1.68)	0.243
rs11025421	11	*DBX1*	A	A	3.64×10^{-8}	0.002 (0.14)	0.2998 (1.44)	0.08
rs1333971	1	*IFI44*	A	A	4.67×10^{-8}	0.0039 (0.3)	0.6975 (0.86)	0.029
rs906181	17	*FLJ44861*	A	A	7.37×10^{-8}	0.0002 (0.29)	0.617 (1.21)	0.006
rs16918853	8	*OPRK1*	A	T	7.54×10^{-8}	0.0002 (0.14)	0.9592 (0.98)	0.004
rs6870923	5	*DTWD2*	A	C	8.23×10^{-8}	0.0013 (0.29)	0.4488 (0.78)	0.001
rs2145976	14	*GLRX5*	A	T	1.02×10^{-7}	0.0016 (0.24)	0.1191 (1.62)	0.262
rs760172	21	*DSCAM*	A	C	1.26×10^{-7}	0.0236 (0.27)	0.3679 (0.59)	0.03
rs12638019	3	*PPP4R2*	A	G	2.34×10^{-7}	0.0003 (0.27)	0.7838 (0.91)	0.015
rs1276295	1	*GNG12*	A	A	2.59×10^{-7}	0.0019 (0.18)	0.8533 (1.14)	0.0056
rs7251341	19	*ZNF415*	A	G	2.73×10^{-7}	0.0075 (0.22)	0.6936 (1.16)	0.05

rs9896237	17	*TMEM100*	A	C	2.97×10^{-7}	0.0137 (0.25)	0.9597 (0.98)	0.022
rs17144526	5	*DTWD2*	A	C	3.15×10^{-7}	0.0002 (0.26)	0.9398 (0.98)	0.007
rs12783237	10	*KIAA1128*	A	T	4.46×10^{-7}	0.0036 (0.15)	0.8541 (0.93)	0.0156
rs10234867	7	*RBAK*	A	C	6.06×10^{-7}	0.002 (0.25)	0.2746 (1.39)	0.068
rs17385833	8	*CCDC25*	A	C	6.70×10^{-7}	0.0013 (0.28)	0.8003 (1.09)	0.03
rs7925158	11	*OR4A15*	A	G	8.12×10^{-7}	0.0064 (0.42)	0.1783 (1.46)	0.483
rs13133783	4	*LOC391698*	A	C	8.35×10^{-7}	0.0002 (0.2)	0.8898 (0.94)	0.0011
rs618869	**18**	***TCF4***	**A**	**T**	**8.54×10^{-7}**	**3.37×10^{-8} (3.14)**	**6.73×10^{-8} (3.03)**	**3.93×10^{-19}**
rs865152	1	*CREG1*	A	G	9.60×10^{-7}	0.0181 (0.3)	0.1805 (1.51)	0.145
rs10216643	8	*CCDC25*	A	C	9.97×10^{-7}	0.002 (0.3)	0.4379 (1.44)	0.027
rs11128302	3	*PPP4R2*	A	G	1.01×10^{-6}	0.00120 (0.32)	0.7373 (0.89)	0.033
rs1877422	15	*LOC646574*	A	C	1.04×10^{-6}	0.0144 (0.29)	0.9145 (0.96)	0.019
rs4723373	7	*AAA1*	A	G	1.35×10^{-6}	0.0007 (0.12)	0.553 (1.25)	0.039
rs4737151	8	*ZMAT4*	A	A	1.66×10^{-6}	0.0004 (0.15)	0.969 (1.01)	0.007
rs2659046	17	*TFLJ44861*	A	T	1.73×10^{-6}	3.35×10^{-5} (0.21)	0.756 (1.09)	0.002
rs977628	5	*CDH9*	A	C	2.22×10^{-6}	0.0013 (0.35)	0.1749 (0.61)	0.003
rs2029036	5	*DTWD2*	A	G	3.71×10^{-6}	0.00230 (0.28)	0.3622 (1.33)	0.095
rs2393261	10	*IPMK*	A	C	3.81×10^{-6}	0.0004 (0.26)	0.321 (0.68)	0.005
rs920139	11	*HNT*	A	T	4.07×10^{-6}	0.0023 (0.32)	0.3328 (1.3)	0.043
rs1386274	6	*EPHA7*	A	C	4.08×10^{-6}	0.0007 (0.49)	0.832 (0.95)	0.01
rs523927	18	*PTPRM*	A	A	4.08×10^{-6}	0.0019 (0.18)	0.764 (1.13)	0.052
rs4736872	8	*ZMAT4*	A	G	4.13×10^{-6}	0.0003 (0.15)	0.996 (1)	0.005
rs4864829	4	*LOC402176*	A	T	4.86×10^{-6}	0.0021 (0.35)	0.9056 (0.96)	0.022
rs10865671	3	*PPP4R2*	A	G	5.83×10^{-6}	0.0007 (0.29)	0.714 (0.88)	0.021
rs12427294	12	*DYRK2*	A	T	7.39×10^{-6}	0.0426 (0.52)	0.1889 (0.63)	0.007
rs8080584	17	*LOC284100*	A	A	7.55×10^{-6}	0.0044 (2.49)	0.7994 (0.9)	0.055
rs438931	14	*POMT2*	A	A	8.28×10^{-6}	0.0014 (0.387)	0.3523 (0.74)	0.0097

rs9527458	13	*FLJ40296*	A	C	9.45×10^{-6}	0.002 (0.47)	0.7263 (0.92)	0.002
rs7511972	1	*DAB1*	A	C	9.91×10^{-6}	0.958 (0.97)	0.734 (1.16)	0.551
rs12644376	4	*TLL1*	A	T	1.02×10^{-5}	0.0024 (0.14)	0.646 (1.2)	0.05
rs12189640	6	*LOC645950*	A	C	1.08×10^{-5}	0.0023 (0.22)	0.864 (0.94)	0.021
rs10887261	10	*RGR*	A	T	1.13×10^{-5}	0.0235 (0.37)	0.7002 (1.15)	0.104
rs9937968	16	*LOC729979*	A	G	1.44×10^{-5}	0.0017 (0.43)	0.203 (0.67)	0.015
rs2256700	10	*DNMBP*	A	G	1.52×10^{-5}	0.0088 (0.3)	0.212 (0.53)	0.008
rs7744039	6	*ENPP3*	A	T	1.92×10^{-5}	0.0003 (1.92)	0.595 (0.89)	0.015
rs2193965	5	*DTWD2*	A	C	2.47×10^{-5}	0.0006 (0.36)	0.831 (0.94)	0.01
rs2302319	**17**	***MINK1***	**A**	**T**	**4.34×10^{-5}**	**3.64×10^{-5} (3.32)**	0.853 (0.93)	0.002
rs11064694	12	*HSPB8*	A	G	4.60×10^{-5}	0.0066 (0.22)	0.594 (0.78)	0.016
rs1190892	14	*C14ORF70*	A	A	6.59×10^{-5}	0.0008 (0.27)	0.496 (1.23)	0.039
rs1127473	6	*C6ORF85*	A	A	7.08×10^{-5}	0.0113 (0.36)	0.2887 (1.43)	0.397
rs16927158	9	*PTPRD*	A	G	7.27×10^{-5}	0.003 (0.35)	0.768 (1.1)	0.049
rs6134174	20	*hCG_2045828*	A	A	7.91×10^{-5}	0.0073 (0.52)	0.0742 (1.51)	0.769
rs10497711	2	*STAT4*	A	G	8.10×10^{-5}	0.012 (0.36)	0.7085 (1.15)	0.204
rs7018431	8	*LOC643042*	A	G	9.33×10^{-5}	0.0055 (0.28)	0.906 (0.96)	0.024
rs168371	3	*PTPRG*	A	C	6.93×10^{-4}	0.0363 (0.5)	0.5248 (0.81)	0.029
rs17767976	3	*PTPRG*	A	C	7.97×10^{-4}	0.0307 (0.48)	0.9662 (1.01)	0.158
rs6453815	6	*COL12A1*	A	A	0.003	0.0357 (1.65)	0.26 (0.71)	0.753
rs6424883	**1**	***LAMC1***	**A**	**C**	**0.009**	**0.0046 (0.6)**	**0.012 (0.58)**	**0.004**
rs2300431	10	*HTRA1*	A	A	0.0130	0.0092 (1.64)	0.991 (1)	0.034

To avoid leaving out such associated SNPs, an additional threshold of $p \leq 10^{-5}$ for the SNPs that had passed the quality control test or had the NA report was considered suggestive of association. A total of 52 additional SNPs met this cut-off (Table 4.1, page 95 - 97).

Furthermore, we considered 5 additional SNPs for individual genotyping from the *PTPRG* (rs168371, p = 6.93 × 10^{-4}; rs17767976, p = 7.97 × 10^{-4}), *COL12A1* (rs6453815, p = 0.003), *LAMC1* (rs6424883, p = 0.009), and *HTRA1* (rs2300431, p = 0.0130) genes that were previously reported or thought as good candidates for association with FECD [73, 92], and were identified in the present study (Table 4.1). The *PTPRG* (*Protein Tyrosine Phosphatase, Receptor Type, G)* gene was initially identified to be associated with FECD in Caucasian American cases in a GWAS performed by Baratz et al [92]; but the finding did not replicate in independent studies in Caucasian American and Chinese populations [92, 94]. SNPs in *COL12A1* (*Collagen, Type XII, Alpha 1*) were associated with FECD in Caucasian Australian cases, in a candidate gene study in my Honours thesis [192]. Proteins encoded by *LAMC1* (*Laminin, Gamma 1*) and *HTRA1* (*HtrA Serine peptidase 1*) were among the most frequently identified proteins in Descemet's membrane by a mass spectrometry analysis performed by myself during my Honours project on FECD [192]. Overall, a total of 66 SNPs were prioritised for individual genotyping in the case-control discovery cohort.

Results of individual genotyping of the prioritised SNPs showed similar or more significant p-values of associations for the SNPs rs613872 (p = 2.47 x 10^{-11}, OR = 3.62) and rs6188892 (p = 3.37 x 10^{-8}, OR = 3.14) in *TCF4*, rs2302319 (p = 3.64 x 10^{-5}, OR = 3.32) in *MINK1*, rs6424883 (p = 0.0046, OR = 0.60) in *LAMC1*, and

rs2300431 (p = 0.0092, OR = 1.64) in *HTRA1*, when compared to the results from GWAS on pooled DNA (Table 4.1). Thus, these SNPs were considered to be validated. The rest of the individually genotyped SNPs (n = 61) had larger p-values of association in the individual genotyping compared to the GWAS on pooled DNA (Table 4.1). They were thus considered not validated. The SNPs that did not validate were considered to be more likely false-positives in the GWAS on pooled DNA.

Along with individual genotyping for validation, the prioritised SNPs were genotyped in a an independent replication cohort of 50 advanced FECD cases and 2760 normal controls of Caucasian Australians. The replication study confirmed three SNPs, rs613872 (p = p = 3.42 x 10^{-12}, OR = 3.75) and rs618869 (p = 6.73 x 10^{-8}, OR = 3.03) in *TCF4* and rs6424883 in *LAMC1* (p = 0.01, OR = 0.58), to be significantly associated with FECD at the p-value of ≤0.05 (Table 4.1). SNPs rs613872 and rs618869 in *TCF4* have been associated with FECD in previous studies [92-95]. Association of rs6424883 is a novel finding; although it is weaker than the genome-wide significance of $p<5 \times 10^{-8}$ (Table 4.1) The replication study did not confirm any association between FECD and the remaining 63 SNPs (Table 4.1).

Combined analysis of the findings from individual genotyping study and replication study cohorts showed similar or better association of only 13 SNPs with the disease. The strength of association of the rs613872 and rs618869 SNPs in *TCF4* with FECD more than doubled as the power to identify significant association with the disease increased with sample size (Table 4.1). Consistently, SNPs rs12427294 (*DYRK2)* and rs6424883 (*LAMC1)* showed improved p-value of association with the disease (Table 4.1). In addition, strength of association of SNPs rs11576306 (*FMO4*), rs760172 (*DSCAM*), rs1276295 (*GNG12*), rs9896237 (*TMEM100*), rs977628

(*CDH9*), rs438931 (*POMT2*), rs9527458 (*FLJ40296*), rs2256700 (*DNMBP*) and rs168371 (*PTPRG*) with FECD in the replication study was similar to that in the individual genotyping study. This observation suggested no genome-wide association of these SNPs with the disease.

In spite of increased sample size and consequently power to identify associations in the combined analysis, the p-values of association for 53 of the 66 SNPs showed reduced or non-significant association with FECD (Table 4.1, last column) and likely indicate false positive associations. SNPs with reduced association with FECD in the individual genotyping and/or combined analysis may indicate a weaker association with the disease, probably due to the small size of the replication cohort. These observations call for genotyping of more cases to find the truly associated SNPs. This would also be the case for *LAMC1* SNP rs6424883.

To summarise, I identified strong association between the rs613872 and rs6188892 SNPs in *TCF4* and late-onset FECD in Caucasian Australian cases. In addition, this study revealed an association of *LAMC1* SNP rs6424883 with the disease. Genotyping of the SNPs with similar or better p-values of association with FECD in the combined analysis compared with the individual genotyping study in a larger case-control cohort may reveal genome-wide significant associations with the disease.

4.3.2 Replication study of FECD-associated SNPs identified through GWAS in a Caucasian American case-control study

During this project, researchers from Case Western Reserve University (Cleveland, Ohio, USA) identified significant association of SNPs in 7 novel loci and the *TCF4* gene with late-onset FECD following GWAS in 2500 Caucasian American cases and

1100 normal controls. As a collaboration, we performed a replication study for 19 candidate SNPs at the 8 loci (Table 4.2, page 102) in Caucasian Australian case-control cohorts. The SNPs were genotyped in 190 FECD cases and 285 unaffected controls. For association of the genotyped SNPs with FECD in the present study, $p<0.05$ was considered significant. Analysis of haplotypes with >1% frequency was also performed, and the P-values and odds ratios were calculated for each haplotype compared with all other haplotypes. Allele frequencies and association results for genotyped SNPs are given in Table 4.2. We confirmed allelic association between FECD and four SNPs at the *TCF4* locus. The G allele of the *TCF4* intronic SNP rs613872, which was previously identified as the most significantly associated with the disease in the GWAS on pooled DNA (Table 4.1), showed the most significant association in this replication study ($p=2.14\times10^{-22}$; OR = 4.23 (95% CI:3.13-5.71)).

Significant SNPs (rs72932713, rs11659764 and rs784257) in/near *TCF4* (Table 4.2) were not analysed in the pooled GWAS. The rs72932713 is also intronic, and its minor allele C was observed to confer a risk for the development of FECD by as much as 4.6 times ($p=1.45\times10^{-15}$; OR = 4.62 (95% CI:3.10-6.88)) in Caucasian Australian carriers compared with non-carriers. Both, the rs11659764 and rs784257 are located near the *TCF4*, and were observed to have the A allele that confers the risk for developing the disease in Australians (rs11659764, $p=5.67\times10^{-18}$; OR = 5.64 (95% CI:3.69-8.64); rs784257, $p=2.80\times10^{-20}$; OR = 3.86 (95% CI:2.87-5.17)).

Table 4.2 Results of replication study of FECD-associated SNPs identified through GWAS in a Caucasian American case-control cohort. Significant associations are in bold font. SNP = single nucleotide polymorphism; 1/2 = minor allele/major allele; Freq = frequency; OR = Odds ratio for minor allele; and CI = Confidence Interval.

SNP	Chromosome: Position	Allele 1/2	Allele 1 Freq		p-value	OR (95%CI)	Gene
			cases	controls			
rs12082238	**1: 62,760,625**	**C/T**	**0.23**	**0.30**	**0.034**	**0.73 (0.54-0.98)**	***KANK4***
rs12058486	1: 62,767,265	T/C	0.24	0.28	0.179	0.82 (0.61-1.10)	*KANK4*
rs79742895	**1: 62,782,860**	**C/T**	**0.07**	**0.03**	**0.003**	**2.44 (1.33-4.48)**	***KANK4***
rs1200114	**1: 169,060,489**	**G/A**	**0.43**	**0.34**	**0.006**	**1.45 (1.11-1.89)**	***ATP1B1***
rs1200115	**1: 169,061,558**	**T/C**	**0.41**	**0.33**	**0.015**	**1.40 (1.07-1.83)**	***ATP1B1***
rs6424883	**1: 183,037,695**	**C/T**	**0.34**	**0.46**	**3.15×10^{-4}**	**0.61 (0.47-0.80)**	***LAMC1***
rs2296292	**1: 183,086,757**	**A/C**	**0.50**	**0.42**	**0.011**	**1.41 (1.08-1.82)**	***LAMC1***
rs3768617	**1: 183,092,500**	**A/G**	**0.34**	**0.46**	**2.39×10^{-4}**	**0.61 (0.46-0.79)**	***LAMC1***
rs1413386	**1: 183,104,811**	**C/G**	**0.38**	**0.48**	**0.002**	**0.66 (0.51-0.86)**	***LAMC1***
rs20561	**1: 183,105,705**	**T/C**	**0.50**	**0.41**	**0.008**	**1.42(1.09-1.85)**	***LAMC1***
rs2274700	1: 196,682,947	T/C	0.43	0.39	0.240	1.17 (0.90-1.53)	*CFH*
rs1329428	1: 196,702,810	A/G	0.42	0.38	0.224	1.18 (0.90-1.54)	*CFH*
rs12223324	11:772,701	G/A	0.49	0.48	0.633	1.07 (0.82-1.38)	*PDDC1*
rs4963153	11:791,462	T/C	0.48	0.46	0.524	1.09 (0.84-1.41)	*SLC25A22*
rs1138714	11:825,110	A/G	0.46	0.46	0.937	0.99 (0.76-1.28)	*PNPLA2*
rs72932713	**18:53,186,911**	**C/T**	**0.25**	**0.07**	**1.45×10^{-15}**	**4.62 (3.10-6.88)**	***TCF4***
rs11659764	**18:53,335,512**	**A/T**	**0.25**	**0.06**	**5.67×10^{-18}**	**5.64 (3.69-8.64)**	***TCF4***
rs784257	**18: 53,397,199**	**A/G**	**0.46**	**0.18**	**2.80×10^{-20}**	**3.86 (2.87-5.17)**	***TCF4***
rs613872	**18: 55,543,071**	**G/T**	**0.46**	**0.16**	**2.14×10^{-22}**	**4.23 (3.13-5.71)**	***TCF4***

All the five genotyped SNPs in *LAMC1* showed significant allelic association with FECD (Table 4.2). The rs6424883 ($p=3.15 \times 10^{-4}$; OR = 0.61 (95% CI:0.47-0.80)) and rs3768617 ($p=2.39 \times 10^{-4}$; OR = 0.61 (95% CI:0.46-0.79)) showed the most significant association with the disease. The former SNP was also identified in the GWAS on pooled DNA to associate with protection against the risk of developing FECD; but the association was not genome-wide significant (Table 4.1). The present study confirmed the GWAS finding. It also showed an improved p-value of association of the SNP rs6424883 with the disease compared to the GWAS, more likely due to a larger sample size (n = 190) in the present study. The rs6424883, rs3768617, and the rs1413386 are intronic, and their minor alleles confer protection against FECD development in carriers compared with non-carrires by about 40% (Table 4.2). Significantly associated SNPs rs2296292 (p=0.011; OR = 1.41 (95% CI:1.08-1.82)) and rs20561 (p=0.008; OR = 1.42 (95% CI:1.09-1.85)) are exonic and carry minor alleles that confer the risk of developing the disease by at least 40% in carriers compared with non-carriers. These findings are consistent with those of our collaborators.

The genotyped SNPs rs1200114 and rs1200115 are located upstream of the *ATP1B1* gene. Both showed significant association with FECD (Table 4.2). Consistent with our collaborator's findings, the G allele in the SNP rs1200114 (p=0.006; OR = 1.45 (95% CI:1.11-1.89)) or T allele in the rs1200115 (p=0.015; OR = 1.40 (95% CI:1.07-1.83)) confer protection against susceptiblity to FECD by 40-45% in carriers compared with non-carriers. These findings confirm *ATP1B1* as a novel gene in the pathogenesis of FECD.

Of the three genotyped SNPs from *KANK4*, two (rs79742895 and rs12082238) showed significant association with the disease (Table 4.2). The carriers of the C allele in the rs7974289 have at least 2.4 times the risk of developing FECD compared with non-carriers (p=0.003; OR = 2.44 (95% CI:1.33-4.48)). In contrast, the minor allele in the rs12082238 of the carriers is protective, with carriers being about 27% less susceptible to the disease than non-carriers (p=0.034; OR = 0.73 (95% CI:0.54-0.98)). Both, the rs7974289 and rs12082238 SNPs are found in the introns of the *KANK4* gene.

None of the SNPs in the *CFH*, *PDDC1*, *PNPLA2* and *SLC5A22* genes revealed significant association with FECD, at p-value of ≤0.05, in the Australian cohort (Table 4.2). Consistently, haplotype analysis of the associated SNP/s in or near the *CFH*, *PDDC1*, *PNPLA2* and *SLC5A22* genes did not confirm any association with the disease. This finding suggested lack of association between these loci and FECD in the Australian population. Given the smaller number of cases involved in the present study compared to that of the initial GWAS, it may be possible that the former was under-powered to detect association of these loci.

Haplotype analyses of the associated SNPs in *LAMC1*, *ATP1B1* and *KANK4* carry the associated minor alleles at the SNPs with significant associations with FECD, with overall *P*-values of 0.013 (Table 4.3, page 105), 0.017 (Table 4.4, page 106), and 0.003 (Table 4.5, page 106), respectively. Two specific haplotypes of the associated SNPs in *LAMC1* harboured protective alleles (haplotype: CCACC (*P*=5.19 × 10^{-4}; OR = 0.63 (95% CI:0.48 – 0.82)) and susceptible alleles (haplotype: TAGGT, *P*=0.011; OR =1.41 (95% CI:1.08 - 1.85)). Similarly, results of haplotype analysis of the associated SNPs upstream of *ATP1B1* revealed that the specific GT

haplotype (*P*=0.037; OR = 1.32 (95% CI:1.02 – 1.72)) habours the G and T risk alleles at the rs1200114 and rs1200115, respectively. Individuals without the GT, but AC, haploytpe are less susceptible to developing FECD than the GT carriers by 31% (p=0.005; OR = 0.69 (95% CI:0.54 – 0.90)). For haplotype analysis of the *KANK4* SNPs, two specific haplotypes were significantly associated with the protection or susceptibility of FECD. The TCC haplotype habour the C risk allele at the rs7974289 (*P*=0.001; OR = 3.17 (95% CI:1.53 - 6.56)). Carriers of this haplotype are more than three times likely to develop FECD than non-carriers. Meanwhile, the CCT haplotype contains the C protective allele at the SNP rs12082238 (*P*=0.033; OR = 0.24 (95% CI:0.05 - 1.09)). The carriers of the CCT haplotype are less susceptible to developing the disease by 76% than those without. In summary, carriers of the risk haplotypes observed in the three genes analysed in this study are more likely to develop the disease compared with non-carriers.

Table 4.3 Association between common haplotypes of genotyped SNPs (**rs6424883, rs2296292, rs3768617, rs1413386, rs20561**) in the *LAMC1* gene and FECD. Significant associations are in bold font. f = frequency; OR = Odds ratio; and CI = Confidence Interval.

Over-all p-value = 0.013				
Haplotype	**f cases**	**f controls**	**p-value**	**OR (95% CI)**
TAGGT	**0.50**	**0.41**	**0.011**	**1.41 (1.08 - 1.85)**
CCACC	**0.34**	**0.45**	**5.19×10^{-4}**	**0.63 (0.48 - 0.82)**
TCACC	0.01	0.01	0.921	0.94 (0.31 - 2.93)
TCGCC	0.02	0.01	0.322	1.80 (0.56 - 5.72)
TCGGC	0.13	0.11	0.200	1.32 (0.87 - 2.00)

Table 4.4 Association between common haplotypes of genotyped SNPs (**rs1200114, rs1200115**) in the *ATP1B1* gene and FECD. Bold font indicates significant association. f = frequency; OR = Odds ratio; and CI = Confidence Interval.

	Over-all p-value = 0.017			
Haplotype	**f cases**	**f controls**	**P-value**	**OR (95% CI)**
GT	**0.39**	**0.32**	**0.037**	**1.32 (1.02 – 1.72)**
AT	0.02	0.01	0.195	2.30 (0.64 – 8.27)
GC	0.04	0.02	0.091	1.96 (0.90 – 4.29)
AC	**0.55**	**0.65**	**0.005**	**0.69 (0.54 –0.90)**

Table 4.5 Association between common haplotypes of genotyped SNPs (**rs12082238; rs12058486; rs79742895**) in the *KANK4* gene and FECD. Bold font indicates significant association. f = frequency; OR = Odds ratio; and CI = Confidence Interval.

	Over-all p-value = 0.003			
Haplotype	**f cases**	**f controls**	**p-value**	**OR (95% CI)**
TCC	**0.06**	**0.02**	**0.001**	**3.17 (1.53 - 6.56)**
CTT	0.22	0.27	0.097	0.78 (0.58 - 1.05)
TTT	0.01	0.01	0.383	1.73 (0.51 - 5.94)
CCT	**0.01**	**0.02**	**0.033**	**0.24 (0.05 - 1.09)**
TCT	0.69	0.68	0.606	1.08 (0.82 - 1.42)

4.3.3 Expression analyses of *ATP1B1* and *LAMC1*

In order to find independent supporting evidence of involvement of *ATP1B1* and *LAMC1* in the pathogenesis of FECD, expression analysis of each gene at the mRNA and protein level was performed. These genes were prioritised over the *KANK4* gene because the association result indicated that they have stronger genetic association

with the disease than *KANK4*. For gene expression analysis, 3 pools of 3 DSAEK specimens from FECD patients and 3 individual DSAEK specimens from controls were analysed using quantitative RT-PCR. This was to investigate relative expression of the genes between affected and unaffected individuals. *ATP1B1* mRNA level was found to be statistically significantly lower in FECD-affected corneal endothelium (Figure 4.2; p = 0.035) compared to unaffected corneal endothelium. In contrast, LAMC1 mRNA level did not differ between affected and unaffected corneal endothelium (Figure 4.3, page 108).

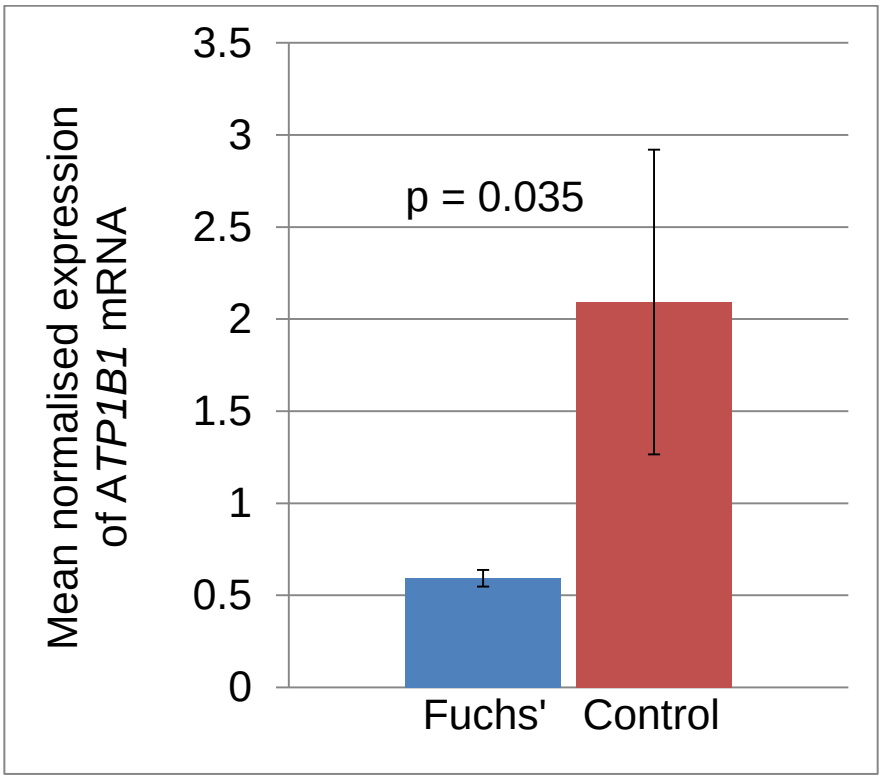


Figure 4.2 Comparison of expression levels of *ATP1B1* mRNA in FECD-affected (n = 3 DSAEK pools of 3 specimens per pool) and unaffected corneal endothelium (n = 3 DSAEK specimens). Bar graph shows the mean normalised expression of *ATP1B1* mRNA in corneal endothelium of FECD-affected (Fuchs') and unaffected controls (Control). *ATP1B1* was significantly down-regulated (p = 0.035) in the affected compared to unaffected corneal endothelium. mRNA, messenger RNA; ATP1B1, Na+/K+ ATPase subunit, beta 1; FECD, Fuchs' endothelial corneal

dystrophy; DSAEK, Descemet's membrane stripping automated endothelial keratoplasty; FECD, Fuchs' endothelial corneal dystrophy.

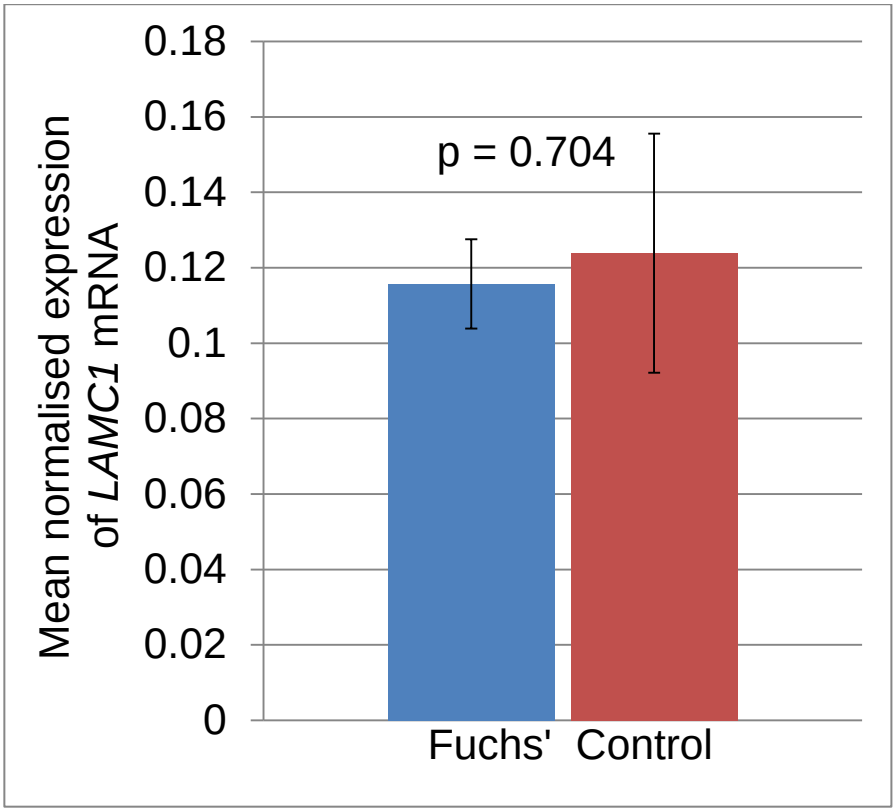


Figure 4.3 Comparison of expression levels of *LAMC1* mRNA in FECD-affected (n =3 DSAEK pools of 3 specimens per pool) and unaffected corneal endothelium (n = 3 DSAEK specimens). Bar graph shows the mean normalised expression of *LAMC1* mRNA in corneal endothelium of FECD-affected (Fuchs') and unaffected controls (Control). *LAMC1* expression did not vary between the affected and unaffected corneal endothelium (p = 0.704). DSAEK, Descemet's membrane stripping automated endothelial keratoplasty; FECD, Fuchs' endothelial corneal dystrophy; LAMC1, laminin subunit, gamma 1; mRNA, messenger RNA.

Expression analysis of ATP1B1 and LAMC1 proteins was performed in corneal sections from three pairs of corneas from FECD patients and healthy deceased donors using immunohistochemistry. The aim was to investigate expression profiles of the proteins. Before immunohistochemical analysis of each protein, commercially acquired mouse anti-ATP1B1 human monoclonal antibody and rat anti-LAMC1 human monoclonal antibody were probed for specificity for each protein of interest by Western blotting. Western blotting could not be attempted on corneal proteins because human corneal tissue was not available for protein extraction. Thus, MCF-7 human cell line and human normal colon tissue that are known to respectively express ATP1B1 and LAMC1 were used for determining anbtibody specificities.

ATP1B1 exists as a tetramer Na^+/K^+-ATPase protein, which is composed of two dimers. Each dimer consists of an alpha subunit of ~112 kDa, and a beta subunit of ~55 kDa [193, 194]. The beta 1 subunit undergoes heavy post-translation glycosylation. Mouse anti-ATP1B1 human antibody reportedly identifies a monomer beta subunit protein, and a dimer of ~55 kDa and ~270 kDa, respectively, in human ciliary body tissue [193]. In this study, western blot analysis identified two protein bands of mass ~55 kDa and more than 250 kDa in MCF-7 breast cancer cell line (Figure 4.4A, page 111). These protein bands correspond to the beta subunit and alpha and beta subunit dimer, respectively. The greater than expected mass of the dimer may be due to post-translational modification of the beta subunit. Thus, the mouse anti-ATP1B1 was inferred to be specific.

In corneal sections immunolabelled for ATP1B1, similar ATP1B1-positive labelling was observed in the epithelium in both disease-affected and normal corneas (Figure 4.4B, page 111). The protein was seen localised in the membrane/between the

epithelial cells, consistent with its known expression as a membrane transporter [193]. Positive antibody staining for the protein was also observed in the endothelium in diseased and normal corneas, but partial loss of endothelium made it difficult to ascertain similar labelling in the affected cornea. Similar faint ATP1B1 labelling was seen in the stroma in both FECD-affected and unaffected corneas. No labelling for ATP1B1 was observed in the Descemet's membrane and Bowman's layer in either disease-affected or unaffected cornea. Descemet's membrane and Bowman's are composed of collagens and extracellular matrix proteins. Thus, lack of immunolabelling for ATP1B1 in these corneal layers is consistent with it being a membranous protein.

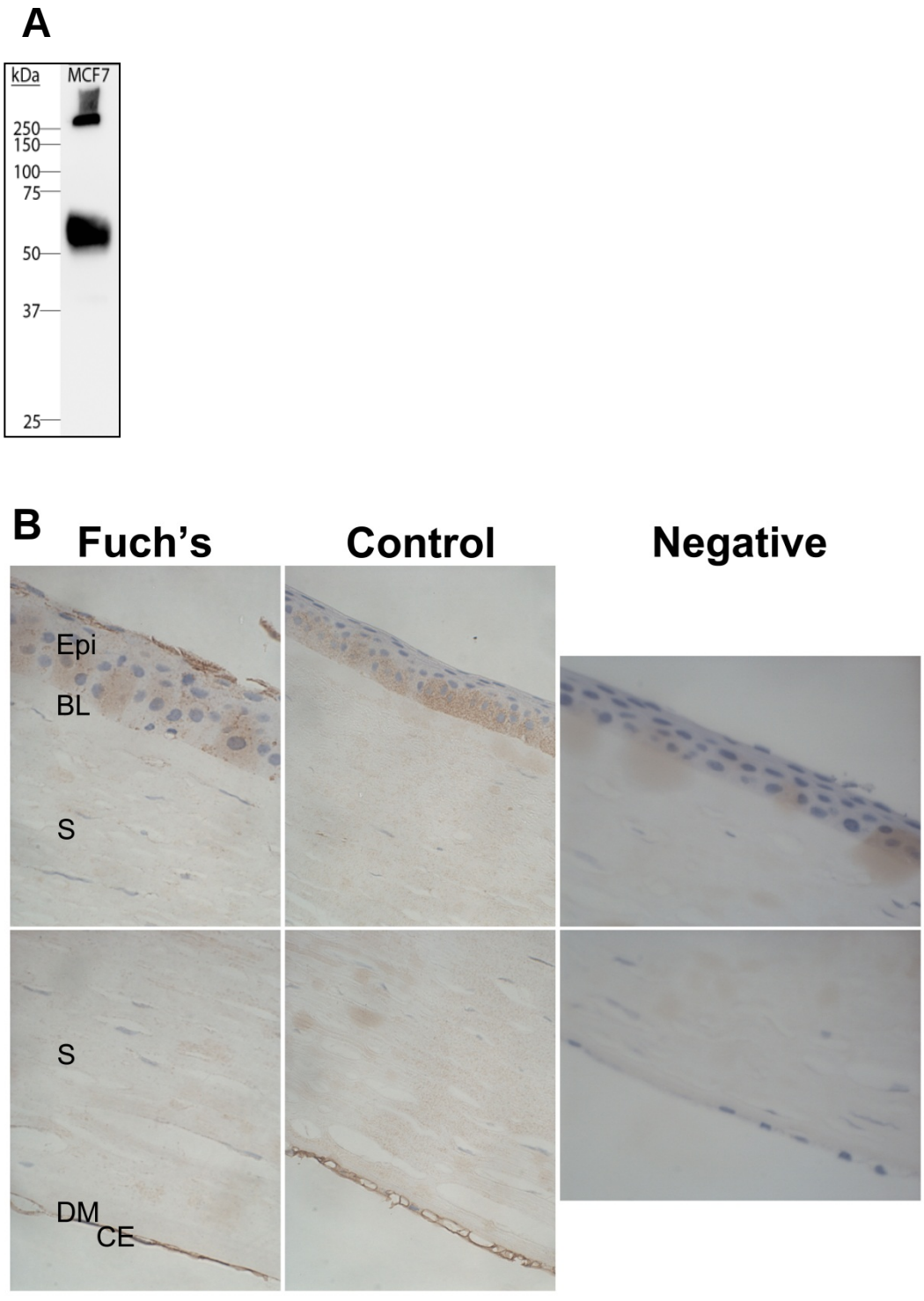


Figure 4.4 (**A**) Western blot showing specificity of the mouse anti-human ATP1B1 primary antibody. ATP1B1 was deteced in protein lysate of MCF-7 human breast cancer cells. The blot shows two protein bands of mass ~55 kDa and >250 kDa. The bands correspond to predicted sizes of monomer beta subunit protein, and a tetramer

of ~55 kDa and ~270 kDa, respectively [193]. (**B**) Immunohistochemical labelling for the ATP1B1 protein in sections of corneas from FECD (Left panels), and normal (middle panels) individuals. A negative control (no primary antibody) is shown in the right panels. Top and bottom panels respectively depict anterior and posterior halves of the corneas. The sections were immunolabelled with mouse anti-human ATP1B1 primary antibody, and rabbit anti-mouse IgG secondary antibody. Similar ATP1B1-positive labelling was observed in the epithelium in both disease-affected and normal corneas. Similar faint ATP1B1 labelling was seen in the stroma in both FECD-affected and unaffected corneas. No labelling for ATP1B1 was observed in the Descemet's membrane and Bowman's layer in either disease-affected or unaffected cornea. Positive labelling was observed in the corneal endothelium in unaffected cornea. However, partial loss of endothelium made it difficult to ascertain similar labelling in the affected corneal endothelium. No labelling in the negative sections indicates the specificity of the mouse anti-human ATP1B1 primary antibody. Representative images from experiments in three independent pairs of corneas from FECD patients and controls are presented. Images are at 400× magnification. Epi, Epithelium; BL, Bowman layer; S, Stroma; DM, Descemet's membrane; CE, corneal endothelium; kDa, kilodalton

The rat anti-human LAMC1 antibody has been shown to identify a LAMC1 protein band of ~177 kDa in human colon polyps and adenocarcinomas [195, 196]. In addition, LAMC1 protein undergoes glycosylation and forms disulphide bonds at several residues, possibly resulting in protein mass greater than the expected size of ~177 kDa [197]. Western blot analysis showed a protein band greater than 250 kDa in human colon (Figure 4.5A, page 114). This band possibly represents LAMC1 protein arising from post-translational modification, such as glycosylation. Thus the rat anti-human LAMC1 antibody was considered to be specific and suitable for use in immunohistochemical analysis.

Immunolabelling for LAMC1 showed detectable protein expression in endothelium, stroma and epithelium in both disease-affected and unaffected corneas. No detectable labelling was observed for the protein in the Descemet's membrane and Bowman's layer either in diseased or normal corneas (Figure 4.5B, page 114). The distribution pattern of LAMC1 was observed to be similar between the diseased and normal cornea. However, corneal endothelium in all normal corneas was more intact than in FECD corneas, and may thus account for widely distributed immunolabelling for LAMC1 in normal corneas. LAMC1 is an extracellular matrix protein, and a major component of the Descemet's membrane. Lack of detectable staining for LAMC1 expression in the Descemet's membrane may be accounted for by less sensitivity of the immunohistochemistry method to identify the protein. In conclusion, investigation of LAMC1 expression and distribution between diseased and normal corneas revealed similar immunolabelling of the protein.

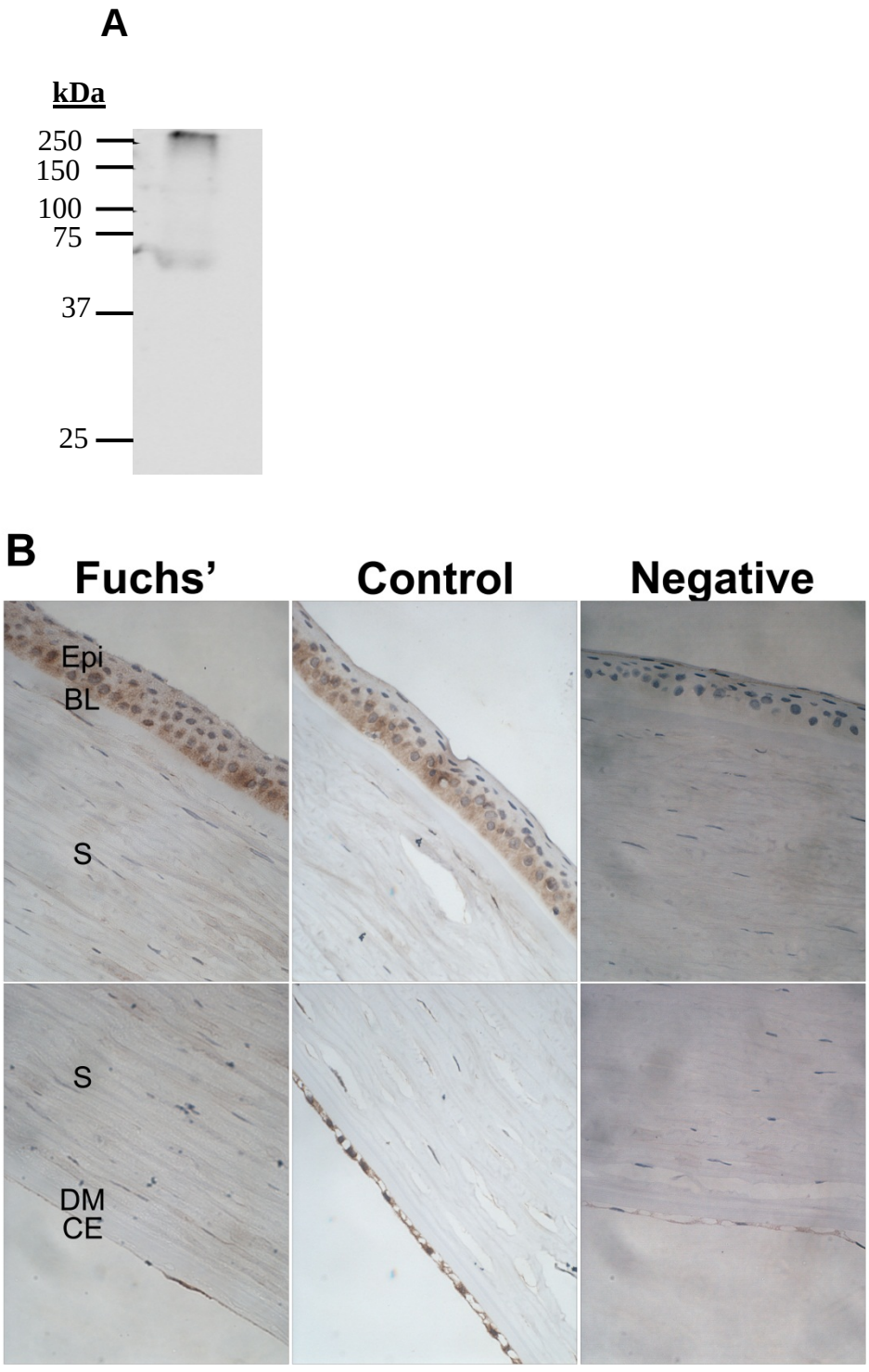


Figure 4.5 (**A**) Western blot showing specificity of the rat anti-human LAMC1 primary antibody. LAMC1 was detected in protein lysate of human normal colon. The blot shows two protein bands of mass ~175 kDa and >250 kDa. The ~175 kDa

band corresponds to predicted size of ~177 kDa [195, 196]. The >250 kDa band is likely a LAMC1 protein arising from post-translational modification. (**B**) Immunohistochemical labelling for the LAMC1 protein in sections of corneas from FECD (Left panels), and normal (middle panels) individuals. A negative control (no primary antibody) is shown in the right panels. Top and bottom panels respectively depict anterior and posterior halves of the corneas. The sections were immunolabelled with rat anti-human LAMC1 primary antibody, and goat anti-rat IgG secondary antibody. Similar LAMC1-positive labelling was observed in the epithelium (Epi) in both disease-affected and normal corneas. Similar very faint LAMC1 labelling was seen in the stroma (S) in both FECD-affected and unaffected corneas. No labelling for LAMC1 was observed in the Descemet's membrane (DM) and Bowman's layer (BL) in either disease-affected or unaffected cornea. Positive labelling was observed in the corneal endothelium (CE). However, partial loss of the endothelium made it difficult to ascertain similar labelling in the affected corneal endothelium. No labelling in the negative sections indicates that the rat anti-human LAMC1 primary antibody was specific. Representative images from experiments in three independent pairs of corneas from FECD patients and controls are presented. Images are at 400× magnification. kDa, kilodalton

4.4 DISCUSSION

The goal of this study was to identify novel genes associated with late-onset FECD in Caucasian Australian cases. The results have revealed that SNPs in the *TCF4*, *LAMC1, ATP1B1* and *KANK4* genes are significantly associated with FECD, and suggest that these genes contribute to pathogenesis of the disease in Caucasian Australian cases. I observed that individuals with risk alleles at the associated SNPs in *TCF4* are at least three times more likely to develop FECD than the non-carriers, in the pooled GWAS and replication study of FECD-associated SNPs performed in this Chapter. These observations are consistent with previous findings of association of *TCF4* with FECD in Caucasian, Chinese and Indian populations [92, 94, 95, 97]. However, this study is the first to report *TCF4* as the major genetic factor associated with the risk of developing the disease in Australia.

TCF4 maps to the *FCD2* locus that has been linked with late-onset FECD in multiple families from the United States [80]. However, variations in the *TCF4* gene do not appear to account for this linkage signal [94]. *TCF4* encodes E2-2 protein, a member of the ubiquitously expressed class I basic helix-loop-helix (bHLH) transcription factors that are involved in cellular growth and differentiation [109]. E2-2 is expressed in developing corneal endothelium [109]. It plays an essential role in developmental processes including epithelial-mesenchymal transition, a normal aspect of cell migration in embryogenesis and tumour-cell invasion and metastasis [111].

E2-2 binds to E-box promoter sequences within target genes and either suppresses or activates their tissue specific transcription [109]. Repression of cell-adhesion protein, E-cadherin expression by E2-2 causes loss of cellular polarity and cell-to-cell contact

and hence EMT. *ZEB1* is a downstream target of *TCF4*, and induces EMT by repressing E-cadherin and recruiting the SWI/SNF chromatin-remodelling protein BRG1 (Brahma Related Gene 1) [198].

Rare pathological mutations in *ZEB1* reportedly contribute to late-onset FECD in Caucasians [39]. In addition, *ZEB1* has been shown to regulate type I collagen expression and is thus critical for maintenance of an endothelial phenotype [199, 200]. These findings suggest that *TCF4* variants associated with FECD could confer the disease risk by altering regulation of *ZEB1* expression or any of other target genes.

In keeping with this hypothesis, *TCF4* rs613872 has been shown by chromatin immunoprecipitation sequence (ChIP-seq) to be present within the binding site for the SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily B, Member 1 (SMARCB1), and subfamily A, Member 4 (SMARCA4) transcription factors [201]. These two factors are components of the SWI/SNF chromatin remodelling complex involved in transcriptional regulation [202, 203]. It has been hypothesised that variation at the rs613872 SNP would affect the spatiotemporal expression of *TCF4* through SMARCB1 and SMARCA4 and hence its targets [96]. However, experimental evidence is needed to support this hypothesis.

FECD-associated SNPs in *LAMC1*, which encodes the laminin subunit, gamma 1 (LAMC1) protein, predict both protective and risk associated role for the gene in development of the disease in carriers of the protective or risk alleles. Carriers of the A or T risk alleles at the rs2296292 or rs20560, respectively, are at least 1.4 times more likely to develop the disease than non-carriers. In contrast, carriers of the C or

A protective alleles at the rs6424883 and rs1413386, or rs3768617 are about 40% less likely to develop the disease than the non-carriers.

Expression analyses of *LAMC1* further supports the involvement of the gene in FECD. The gene was shown to be expressed in the corneal endothelium of both affected and unaffected individuals. However, no statistically significant difference was observed in the expression of *LAMC1* mRNA levels between FECD-affected and unaffected corneal endothelium, suggesting that the gene is likely not dysregulated in corneal endothelium in the disease.

Immunohistochemical analysis for LAMC1 expression in corneal sections from diseased and normal corneas showed expression of the protein in the endothelium, stroma and epithlium in the affected and unaffected corneas. This implies that LAMC1 plays an important biological role in the physiology of these corneal layers. Additionally, LAMC1 expression does not distribute differentially between diseased and normal corneas, suggesting that distribution of the protein is not altered in FECD. This observation correlates with the gene expression analysis between the affected and unaffected corneal endothelium in this study.

A recent study by Matthaei and colleagues has revealed expression of LAMC1 protein in the corneal endothelium but not Descemet's membrane in FECD-affected and unaffected corneas [204]. Their findings correlate with ours. Nevertheless, other independent studies have identified the protein and other laminin subunuits as major component of the Descemet's membrane in both normal and diseased condition [72, 74, 205, 206]. Furthermore, LAMC1 is reported to be the most widely expressed subunit, and a main component of laminin isoforms [207, 208], which form basement membranes, including the Descemet's membrane [209-211]. LAMC1 acts

as a regulator of extracellular matrix proteins in the basement membranes [208]. The failure to detect any immunolabelling for LAMC1 in the Descemet's membrane in the present study and that by Matthaei et al [204] might be due to the presence of undetectable amount of the protein in this layer. Comparative expression analysis of protein components in Descemet's membrane from FECD-affected and unaffected individuals using a more sensitive and quantitative techniques, such as label-free mass spectrophotometry, may reveal relative abundance of LAMC1 in Descemet's membrane.

FECD-associated SNPs, rs1200114 and rs1200115, in *ATP1B1* increase the risk of developing the disease by at least 1.4 times in carriers of these alleles compared with non-carriers. Down-regulation of *ATP1B1* in diseased corneal endothelium compared to normal corneal endothelium, and immunolabelling of ATP1B1 in corneal endothelium and epithelium in FECD and normal corneas further support involvement of the gene in the pathogenesis of the disease.

ATP1B1 encodes the Na+/K+-ATPase, beta 1 polypeptide (ATP1B1) protein [193]. ATP1B1 and the alpha-subunit of the Na+/K+-ATPase form a sodium pump that is responsible for establishing and maintaining electrochemical gradients in most eukaryotic cells; including corneal endothelium [193]. In addition, several *in vitro* studies have demonstrated that ATP1B1, in synergy with E-cadherin, facilitates cell-cell adhesion, epithelial polarisation, and suppression of invasiveness and motility of carcinoma cells [212-215].

In FECD, electrochemical gradient homeostasis is compromised, leading to excessive accumulation of fluid in the stroma and subsequent opacification of the cornea [216]. Also, progressive death of corneal endothelial cells in the disease leads

to gaps between cells, causing the remaining live cells to enlarge to restore severed tight cell junctions [35]. This leads to abnormal endothelial cell morphology and further degeneration or loss of the enlarged cells [45]. Consistently, observed down-regulation of the *ATP1B1* transcript in FECD-affected corneal endothelium may disturb the electrochemical gradient homeostasis. It may also result in compromised cell-cell adhesion between the endothelial cells, and may contribute to severed tight junctions and enlargement of the remaining live cells.

Pathological mutations in the *SLC4A11* gene lead to FECD in a small number of cases [83]. Independent functional studies reported abnormal accumulation of mutant SLC4A11 protein in the endoplasmic reticulum in FECD-affected corneal endothelium, and reduced trafficking of the protein to the cell membrane [76, 87, 126]. Furthermore, characterisation of functional effects of *SLC4A11* mutations recently revealed that the mutant protein present at the cell membrane was functionally compromised [87]. Both of these effects could result in impaired corneal endothelial pump function in the disease [216]. It would be interesting to investigate *ATP1B1* for FECD-causing mutations.

The *KANK4* gene codes KN motif and Ankyrin repeat domains 4 (KANK4) protein [217]. Little is known about the biological function of KANK4. [218]. It is reported to regulate polymerisation of actin stress fibres through inhibition of RhoA activity, and thus inhibit cell migration [218, 219]. Actin stress fibres reportedly play an important role in cell motility and contractility, providing a force for a number of cell functions including cell adhesion and morphogenesis [220]. KANK4 is expressed in the corneal endothelium, as shown by microarray analysis of FECD-affected and unaffected corneal endothelium presented in Chapter 7 in this thesis. There was no

differential expression of *KANK4* observed in corneal endothelium between diseased and normal corneal endothelium. However, the microarray result requires validation by an independent method, such as quantitative RT-PCR, to reach a conclusion of lack of dysregulation of the gene in FECD. Immunohistochemical analysis of KANK4 in diseased and normal corneas is advisable, as it might provide further support for involvement of the gene in the pathogenesis of the disease.

In conclusion, I identified *TCF4* as the major susceptibility gene for development of FECD in Caucasian Australians. In addition, *LAMC1*, *ATP1B1* and *KANK4* were identified as novel genetic factors associated with the risk of developing the disease in Australia, similar to that found by our collaborators in the United States. The *ATP1B1* gene is down-regulated in FECD. Expression of *KANK4* and *LAMC1* in corneal endothelium, and likely dysregulation of ATP1B1 protein in the corneal endothelium in FECD provide further support for involvement of the associated genes in the disease.

CHAPTER 5
Association of TGC repeat polymorphism in *TCF4* with FECD in Australian cases

5.1 INTRODUCTION

Common repetitive DNA elements comprise approximately half of the entire sequence of the human genome. They include large segmental duplications, long and short interspersed transposon-derived elements and tandem repeats [221, 222]. The tandem repeats encompass satellites, mini-satellites and microsatellites or short tandem repeats (STRs) [103]. The STRs account for about 2% of the genome [103]. Most STR tracts occur in the inter-genic regions and introns. A fraction of them, predominantly trinucleotide repeats, also reside in exons and may be beneficial, neutral or deleterious [104].

Among the beneficial roles of trinucleotide repeats is their potential to modulate cellular processes, including transcriptional splicing and translation [223]. These trinucleotide repeats include repeats of CGG, CAG and AGG, which are over-represented in human exons [224, 225]. On the other hand, AAT, AAC and AAG are probably disadvantageous as they are negatively selected in exons [224, 225].

Trinucleotide repeat sequences undergo mutations at a very high frequency due to their highly polymorphic nature [226]. This may increase disease risk or trigger disease in specific conditions [227, 228]. At least 20 neurological and muscular diseases have been shown to derive from pathogenic expansion of trinucleotide repeats, located either in coding or non-coding gene sequences [229]. These diseases are known as trinucleotide repeat expansion diseases (TREDs). The best examples

include fragile X syndrome (FXS), myotonic dystrophy type 1 and type 2, Huntington's disease (HD), and Friedreich ataxia (FRDA) [104, 230, 231].

FXS is caused by an expanded CGG repeat located in the 5'UTR of the *fragile X mental retardation 1* gene (*FMR1*). Whereas myotonic dystrophy type 1 is triggered by an expanded CUG repeat in the 3'UTR of the *dystrophia myotonica protein kinase* (*DMPK*) gene [104], FRDA and HD respectively arise from pathogenic expansion of GAA repeat in intron 1 of the *FRDA* gene, and from an abnormally elongated CAG repeat located in the open reading frame (ORF) of the *Huntingtin* (*HTT*) gene [104].

The repeat type and localisation determine the mechanism of pathogenesis. The current models of patho-mechanisms of TREDs postulate two toxic agents, RNA and protein [103]. In addition, recent reports also speculate DNA toxicity. RNA gain-of-function toxicity has been reported to underlie molecular pathogenesis of myotonic dystrophy type 1. Interestingly, Gattey et al. [232] have recently reported a series of four cases of FECD in patients with an established diagnosis of myotonic dystrophy, thus showing for the first time an association between the two diseases.

Similar to myotonic dystrophy type 1 disease, the unstable TGC trinucleotide repeat in the *TCF4* gene has been identified to be associated with FECD in cases from non-Australian populations [99, 102, 105, 106]. In this project, *TCF4* was identified as the major susceptibility gene for development of FECD in Australia (See Chapter 4). In the present study, I aimed to determine whether the reported expanded TGC repeat polymorphism in *TCF4* is associated with FECD in Australian cases.

5.2 METHODS

5.2.1 Polymerase chain reaction

Information about primers used for the amplification of the TGC trinucleotide repeat region in the *TCF4* gene (Table 2.2), and the polymerase chain reaction are given in Chapter 2, Sections 2.4.1.1 and 2.4.1.2, respectively. PCR amplification was performed in a Bio-Rad thermal cycler (Bio-Rad Laboratories Pty Ltd.) as follows. HotStar Plus Taq® DNA Polymerase (Qiagen GmbH, Hilden, Germany) was activated at 95°C for 6 minutes followed by 35 cycles of denaturation at 95°C for 1 minute, annealing at 64°C for 1 minute, and elongation at 68°C for 3 minutes cycles. The final elongation was at 68°C for 7 minutes.

5.2.2 Short tandem repeat (STR) Assay

Each PCR product was diluted to 1:10 with autoclaved MQH_2O. One microlitre of diluted PCR product, 0.15µl of GeneScan™ 1200LIZ® Dye Internal Size Standard (Applied Biosystems, Foster City, CA.) and 8.85µl of Hidi formamide (Applied Biosystems) were mixed and the mixture was electrophoresed on a 3130xL Genetic Analyzer (Applied Biosystems), at the Southpath and Flinders Sequencing Facility (Flinders Medical Centre, Adelaide, Australia) according to standard protocols. The TGC repeat alleles were manually called using Peak Scanner™ Software v1.0 (Applied Biosystems) and recorded. In order to know whether the STR assay accurately detected the number of TGC repeats, Sanger sequencing was performed on selected cases and controls homozygous for the shortest and the longest TGC repeat length alleles.

5.2.3. Statistical analyses

Baseline characteristics of cases and controls were compared using a Student's t-test for age, and a chi-square test for gender. Mann-Whitney U test was employed to assess any difference in the distribution of the expanded alleles between cases and controls. The distribution of the alleles in cases and controls was visually displayed by Boxplot generated in the IBM SPSS (version 22). Hardy-Weinberg equilibrium for the alleles was examined in cases and controls using a chi-square test, in PLINK [173]. Logistic regression analysis was performed in PLINK with age and gender as covariates to explore the effects of age and gender on TGC repeat polymorphism. Conditional analysis was performed on the *TCF4* SNP rs613872 and the expanded TGC allele to know if the SNP rs613872 was independently associated with FECD.

5.3 RESULTS

A total of 189 unrelated cases with advanced late-onset FECD and 183 controls were screened for TGC repeat polymorphism in the *TCF4* gene by short tandem repeat (STR) assay. Characteristics of the case and control cohorts are given in Table 5.1 (page 126). The female participants represented 68.78% of the cases, and 71.58% of the controls. A higher percentage of females than males in the case cohort is consistent with the reported higher prevalence of FECD in females [177]. Importantly, there was no statistically significant difference in gender distribution ($p=0.555$) between cases and controls. This is critical for reliability of data because a significant difference in gender between cases and controls would introduce a bias. The controls were significantly older than cases ($p=0.023$) by design to reduce the likelihood of yet to manifest disease in controls.

Table 5.1 Characteristics of the case and control cohorts. Student's t-test and a chi-square test were used for comparing age and gender in cases and controls, respectively. SD = standard deviation.

Variables	Cases	Controls	P-value
Total number of participants	189	183	NA
Total number of female participants (%)	130 (68.78%)	131 (71.58%)	0.555
Average age in years ± SD	69.86 ± 11.20	76.56 ± 8.69	0.023

The results of genotyping of the *TCF4* TGC trinucleotide repeat in individual cases and controls are summarised in Table 5.2. The expanded TGC repeat alleles were dichotomised such that $TGC_{\geq 40}$ was considered an expanded allele, and $TGC_{<40}$ was considered normal allele. Out of the cases screened for the TGC repeat alleles, 57 had two normal alleles of same lengths, and 35 with normal alelles of different lengths (range: 11 – 38 repeat units). The remaining cases harboured the expanded allele, of whom 5 had two expanded alelles of the same length, and 92 with the expanded alleles of different lengths (range: 11 – 115 repeat units).

Table 5.2 Dichotomisation of the TGC repeat alleles in the *TCF4* gene between FECD cases and normal controls.

Cohort	Number of TGC repeat alleles with repeat length of	
	<40	**≥40**
Cases	184	194
Controls	348	18

The number of controls with the normal allele was 174, representing ~ 95% of the screened controls. Of these, 20 had both alleles of the same length, and 154 with normal alleles of different repeat lengths, ranging from 11 to 39. Only 9 of the controls were identified to have the expanded allele, and all were of different lengths (range: 11 – 83 repeat units).

Table 5.3 Distribution of the TGC repeat length in the *TCF4* gene in FECD cases and normal controls.

Cohort	**Number of in dividuals with same or different TGC repeat length**		
	Same (<40)	**Same (≥40)**	**Different (<40/≥40)**
Cases	57	5	35/92
Controls	20	0	154/9

Plotting of the distribution of the expanded TGC repeat risk allele between cases and controls showed that the majority of cases have TGC repeat length of ≥40 (Figure 5.1, page 128). It also revealed that about 50% of the controls have TGC repeat length of ~20 units. Consistently, Mann-Whitney U test revealed that the TGC repeat length is significantly longer in cases than controls ($p = 0.0005$).

DNA samples from four cases and four controls with the shortest normal alleles of the same length (12 repeats per individual), and three cases and controls with the longest two expanded alleles of same lengths (cases: 76, 83 and 84 repeats; controls: 18 repeats each individual) were sequenced by the Sanger method for confirmation of repeat length.

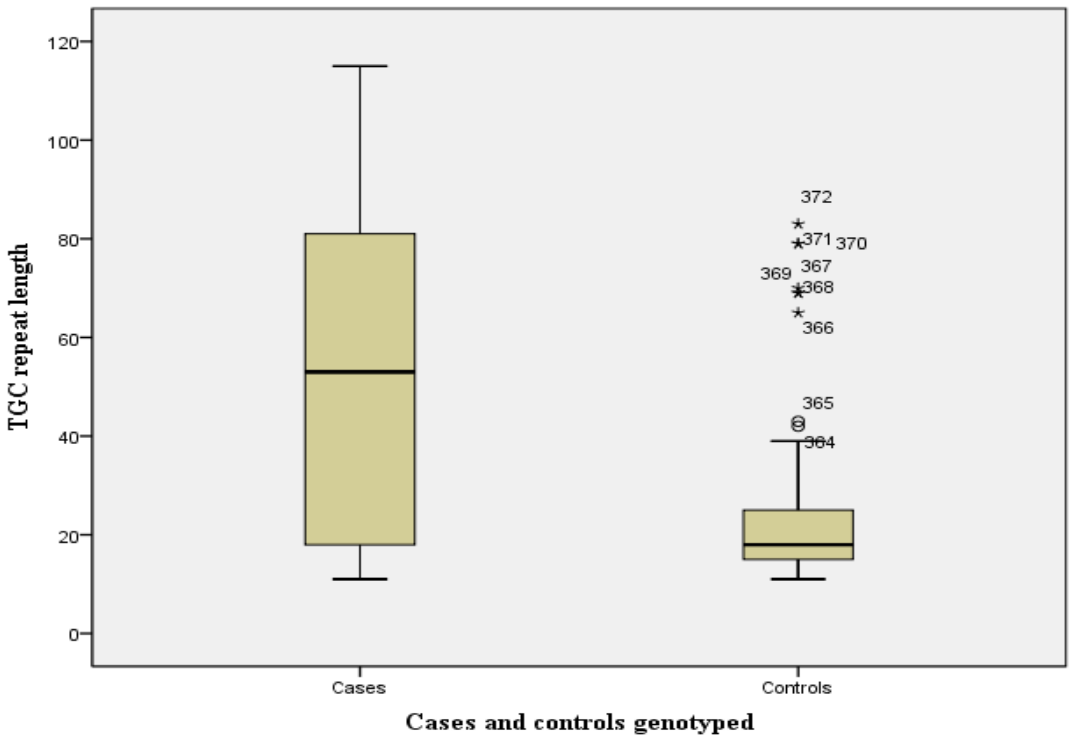


Figure 5.1 Distribution of TGC repeat length between FECD cases and controls. The boxes or rectangles represent second and third quartiles. The lower and upper quartiles are shown as horizontal bends of the rectangles. The dark line in the middle of a box indicates the mean TGC repeat length. The bottom and top wiskers respectively represent 10^{th} and 90^{th} percentiles, respectively. Plotted numbers are individual controls with TGC repeat length ≥40 repeat unit and are outliers.

Analysis of the PCR products from the selected cases and controls by agarose gel electrophoresis (Figure 5.2, page 129) showed PCR products of the expected sizes for the TGC repeat lengths. Sequenced DNA products confirmed the earlier detected TGC repeat length in each of the cases and controls, thus supporting data from the STR assay. Chromatograms from sequencing and STR electropherogram tracings of representative cases and controls are shown in Figures 5.3 (page 130) and 5.4 (page 131), respectively.

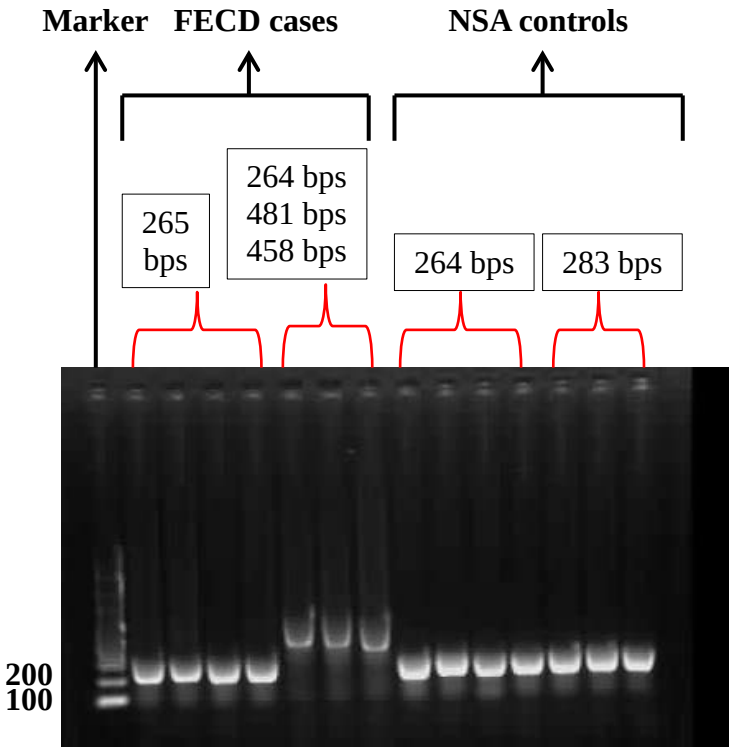


Figure 5.2 Agarose gel electrophoresis of PCR products of the *TCF4* TGC repeat in FECD cases and controls carrying the shortest or the longest repeat alleles of same length. Representative data from 7 FECD cases and 7 normal controls are shown.

Chi-square test indicated a significant association of the expanded allele (>40 TGC repeats) in the *TCF4* gene with FECD under the allelic model ($P = 2.58 \times 10^{-22}$; OR (95% CI) = 15.66 (7.79 – 31.49)). This result replicates the findings of previous studies in Caucasians [102, 105]. Logistic regression analysis revealed that the expanded TGC repeat is associated with FECD independent of age and gender ($P = 2.09 \times 10^{-14}$; OR (95% CI) = 18.26 (8.67 – 38.46)). In addition, conditional analysis on the previously associated SNP rs613872 showed the association of the TGC repeat as partially independent of the SNP ($P = 9.74 \times 10^{-10}$; OR (95% CI) = 10.76 (5.02 – 23.05)); evident from the reduced p-value of association of the TGC repeat allele with FECD.

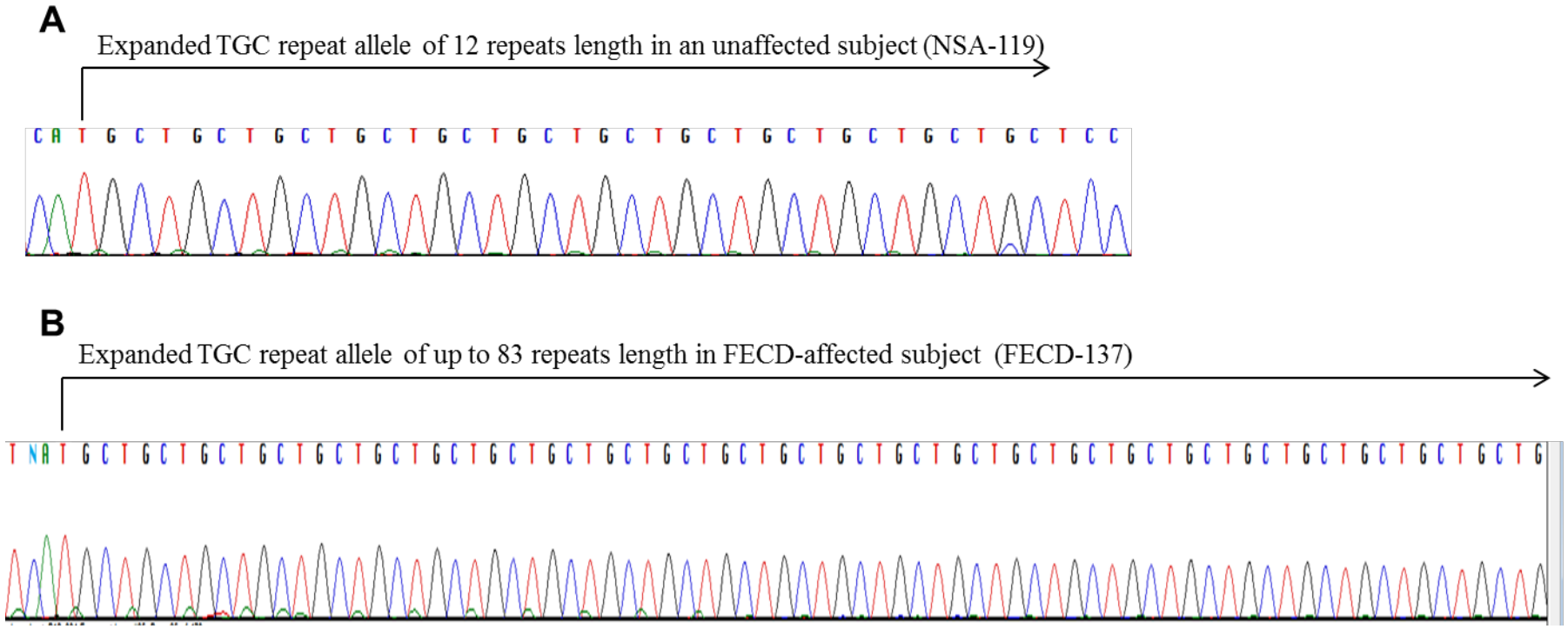


Figure 5.3 Representative chromatograms from Sanger sequencing of an unaffected individual (**A**) and a FECD-affected individual (**B**) homozygous for a short and long allele, respectively.

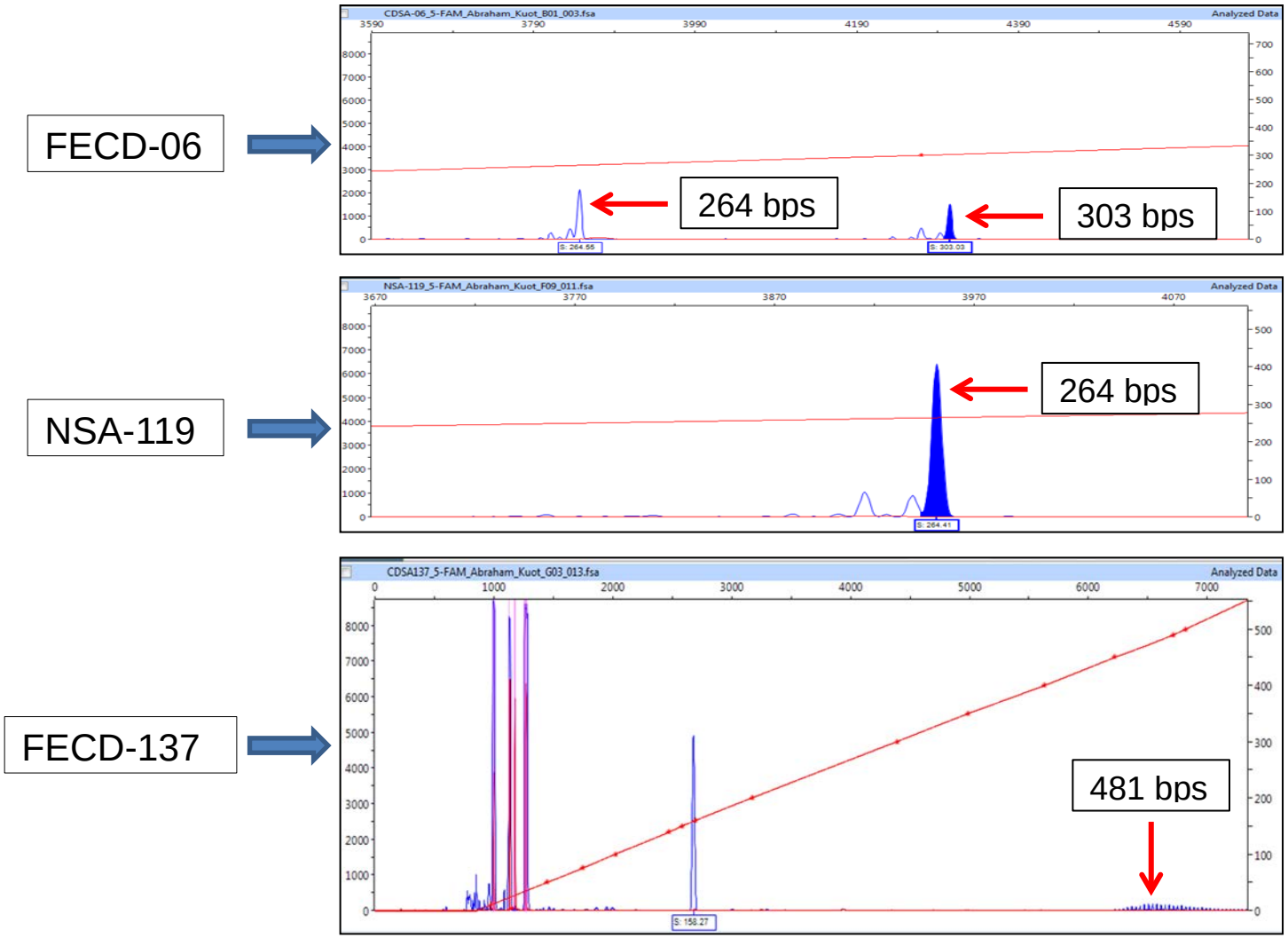


Figure 5.4 Representative electropherograms of TGC repeat allele sizes detected by STR assay in two FECD patients (FECD-06 and FECD-137, top and bottom panels respectively), and in a control individual (NSA, middle panel).

5.4 DISCUSSION

In this study, I aimed to determine association between FECD and the expanded TGC repeat polymorphism in *TCF4* in Australian cases. The study revealed that the expanded TGC repeat is significantly associated with the disease in the Australian cases. This association is more significant than the rs613872 SNP with the disease, reported in Chapter 4. Conditional analysis of the rs613872 SNP with the expanded TGC repeat allele suggests a partially independent association of both polymorphisms with FECD, implying that each could independently contribute to the pathogenesis of the disease. Collectively, these findings replicate previous findings by other groups in American and Chinese populations [102, 105].

FECD is the first ocular disease to be associated with trinucleotide repeats. Pathogenic expansion of trinucleotide repeat sequences occurs in several neurodegenerative and neuromuscular diseases such as FRDA and myotonic dystrophy type 1 [104, 233, 234]. Additionally, pathogenesis of Parkinson's and Alzheimer's diseases have been genetically associated with common SNPs in the *TCF4* gene [235]. However, there are no studies reported to have investigated any association of Parkinson's and Alzheimer's diseases with pathogenic expansion of the trinucleotide repeat in *TCF4*.

Pathogenic expansion of both GAA repeat alleles in the first intron of the *FXN* gene in the majority of FRDA-affected cases causes the disease by interfering with transcriptional elongation, resulting in significant decrease in FXN protein levels [221]. The disease can also occur in a minority of patients as a result of expansion of one GAA repeat allele and presence of a point mutation in the second GAA allele

[221]. Either way, the transcription inhibition causes loss of function of *FXN*, and consequently pathogenesis of FRDA [236].

Although both FECD-associated expanded TGC repeat allele in *TCF4* and FRDA-pathogenic GAA repeat allele in *FXN* are intronic [104], there are distinct differences in pathogenesis of the two diseases and in functions of the genes encoding the expanded repeat alleles. Therefore, transcription inhibition may be involved in the pathogenesis of FECD.

RNA-mediated gain-of-function mechanism has been shown to underlie myotonic dystrophy type 1 [104]. This disease is caused by an expanded CTG repeat allele of ~60 to 100 repeat units in the 3'UTR of the *DMPK* gene [104]. Independent studies have demonstrated that the expanded CTG repeats affect alternative splicing of the *DMPK* gene [104, 237]. The mutant mRNA transcript, DMPK $(CUG)_{n,}$ is normally transcribed but remain untranslated [238, 239]. They are thus retained in the nucleus, where they form hairpin structures [239, 240]. These structures exhibit a toxic dominant gain-of-function abnormality by sequestering and accumulating RNA-binding regulatory proteins, such as the alternative splicing regulator muscleblind-like 1 (MBNL1) and CUG triplet repeat RNA-binding protein 1 (CUG-BP1) [241]. This upsets the fine balance of cellular processes (6 – 8) and leads to retention of the RNA and the regulatory proteins in the nucleus, where they form nuclear foci [242]. The nuclear foci cause cellular toxicity, which leads to the pathogenesis of myotonic dystrophy type 1 [237].

Significant common findings have been reported between myotonic dystrophy type 1 and FECD that suggest the possibility of a common mechanism in the pathogenesis of the two diseases. Firstly, FECD was recently identified in American patients with

myotonic dystrophy type 1 disease [232]. Secondly, both the expanded TGC repeat and the pathogenic CTG repeat are trinucleotide repeat alleles, and located in non-coding regions in the *TCF4* and *DMPK* genes, respectively. Thus it is possible that a RNA-mediated mechanism may underlie the pathogenesis of FECD due to the expanded TGC repeat length in the *TCF4* gene.

In conclusion, this study has independently replicated genetic association of the expanded TGC repeat polymorphism in the *TCF4* gene with advanced FECD. It is the first study to observe this association in Caucasian Australian cases. The expanded TGC repeat allele and the rs613872 polymorphisms are partially, independently associated with FECD in Australian cases. The expanded TGC repeat allele may be contributing to the pathogenesis of FECD via a RNA-mediated mechanism in the corneal endothelium. FECD has been found in patients with myotonic dystrophy type 1 [232], which also supports an RNA-mediated mechanism.

CHAPTER 6
Identification of differentially abundant proteins between FECD-affected and unaffected Descemet's membrane

6.1 INTRODUCTION

Abnormally thickened DM is one of the three main histopathological features of FECD [35, 206]. This occurs as a result of protein accumulation in the DM [73]. Accumulated proteins are suggested to originate from excessive expression and secretion by the corneal endothelial cells [46, 121, 243]. Previous studies investigated differential expression of proteins between diseased and normal DM [73, 114]. Their aim was to uncover pathways/mechanisms involving differentially expressed proteins, and their potential implication in FECD pathogenesis. Through those studies, CLU and TGFBI proteins were identified to be overexpressed in the CE+DM complex from FECD patients compared to normal controls [114]. We reported differential distribution of the CLU and TGFBI proteins in diseased corneas [93], supporting a role of these proteins in the disease process.

CLU is a molecular chaperone and is implicated in protecting cells from effects of physiological stress caused by aging, oxidative stress and apoptosis, which are potential mechanisms underlying the pathophysiology of FECD [69, 133, 134]. TGFBI is a secreted extracellular matrix protein that mediates cell adhesion by interacting with major protein components of DM, including collagens [114]. In FECD, corneal endothelial cells lose cell-cell and cell-DM contacts, and proteins in the DM are shown to be dysregulated [73, 114, 139, 244].

Most recent molecular studies have shown additional differentially expressed proteins in the DM, as well as TGFBI and CLU, between the affected and normal

individuals [131, 132, 204]. These proteins include several types of collagen, agrin, biglycan, and laminins. Identification of differentially expressed proteins in the DM between FECD-affected and unaffected tissue suggest abnormal synthesis and/or deposition of these proteins in the disease. These findings also support the involvement of extracellular matrix alterations in the pathophysiology of FECD [114, 130].

As knowledge of dysregulated proteins can shed light on the disease mechanism [73, 245], and there were not many reports of such studies in FECD at the commencement of the present study, I aimed to identify additional differentially abundant proteins between affected and unaffected DM. Label-free quantitative mass spectrometry, nUPLC-MSE (nanoscale ultra-performance liquid chromatography-mass spectrometry), approach was employed to investigate differentially abundant proteins between FECD-affected and normal DM. In addition, immunohistochemistry was used to profile expression patterns of identified differentially abundant protein.

6.2 MATERIALS AND METHODS

6.2.1 Comparative mass spectrometry analysis

Proteins from FECD-affected and unaffected Descemet's membrane were extracted, and buffer exchanged for a nUPLC-MSE compatible buffer as described in Chapter 2, section 2.3.3.2. Label-free mass spectrometry of the extracted proteins was performed by our collaborators Dr. Maurizio Ronci and Dr. Peter Urbani from the University of Rome, Italy. To perform nUPLC-MS^E, proteins were trypsin digested as follows. 20µL of solution corresponding to a protein amount of 15.26µg were subjected to tryptic digestion as previously described [246]. Briefly, the extract was

reduced and alkylated by adding sequentially 2 µL of 100 mM DTT (60 min at 37°C) and 2.5 µL of 200 mM iodoacetamide (IAA; 60 min at RT). A volume of 0.5µL of DTT (10 min at 37°C) was added before the trypsin to avoid protease alkylation. Finally, 1µL of sequencing grade trypsin (Promega, Madison, WI, USA) at 0.5 mg/mL was added and let react for 16 hours at 37°C. The reaction was quenched by adding 1µL of 10% formic acid (30 min at 37°C). The samples were diluted with aqueous formic acid 0.1% to a final peptide concentration of 0.5 mg/mL and MassPrep yeast enolase digestion standard (SwissProt P00924; Waters, Milford, MA, USA) was added as internal standard to a final concentration of 100 fmol/uL.

Chromatographic separation was achieved on a nanoACQUITY UPLC System (Waters) by injecting 2 µL of sample per run. The samples were loaded on a 5 µm Symmetry C_{18} trapping column 180 µm × 20 mm (Waters) and separated on a 1.7 µm BEH 130 C_{18} Nano Ease 75 µm × 250 mm LC column (Waters) at a flow rate of 250 nL/min using a gradient from 3 to 40% CH_3CN in 145 min. The lock mass ([Glu1]-Fibrinopeptide B, Sigma, 500 fmol/mL) was delivered from the auxiliary pump of the UPLC at a constant flow rate of 600 nL/min.

Separated peptides were introduced into the hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-Tof Premier, Waters) through the nano ESI interface. The instrument was programmed to step between low (4 eV) and high (15–40 eV) collision energies in the collision cell, using a scan time of 1.5 s per function over a mass range of 50–1990 m/z.

Data were acquired using the Waters proprietary data-independent parallel parent and fragment ion acquisition mode (Expression - MS^E). This method does not use a fixed ion transmission window of the first mass analyser before collision induced dissociation [247]. Continuum LC-MS data from three technical replicates for each sample were processed for qualitative and quantitative analysis using the software ProteinLynx Global Server v. 2.4 (PLGS, Waters Corp.).

Qualitative identification of proteins was obtained using the embedded ion accounting algorithm of PLGS 2.4 (Waters), searching the UniProt KB/Swiss-Prot Protein Knowledgebase release 2013_08, 24-July-13, consisting of 540732 entries and 192091492 residues, abstracted from 221115 references and using the Human taxonomical restriction (20266 sequences), to which the sequence of *S. cerevisiae* Enolase was appended (UniProtKB/Swiss-Prot AC: P00924) [246, 248]. The search parameters included: automatic tolerance for precursor ions and for product ions, at least 3 fragment ions matched per peptide, and 7 fragment ions matched per protein, at least 2 peptides matched per protein, 1 missed cleavage allowed, carbamidomethylation of cysteine as fixed modification and oxidation of methionine as variable modification. The false positive rate (FPR) was fixed below 4% for protein identification and the concentration of the calibration protein (internal standard Enolase from Yeast) was set to 200 fmol [247].

The label-free quantitative analysis was performed on three technical replicates available for each experimental condition. Within the differential analysis, the EMRT clusters tables (list of peptide Exact Masses paired to their Retention Times) and the Protein tables were generated upon normalization to the most reproducible

peptides of yeast Enolase (Swiss-Prot AC: P00924) for retention time and intensity (m/z 975.56, m/z 1038.59, m/z 1088.65, m/z 1435.75 and m/z 1988.035). Quantitative analysis was performed based on 164086 molecular spectral features using the EMRT cluster annotation. The differentially expressed proteins dataset was filtered by considering only those identifications from the alternate scanning LC- MS^E data exhibiting a good replication rate (at least 6 out of 9 injections per condition, 66.7%) and with $p < 0.05$ for the relative protein fold change (two-tailed Student's t-test). The significance of regulation level specified at 30%, which is typically 2–3 times higher than the estimated error on the intensity measurement, was used as a threshold to identify significant up- (≥1.3-fold) or down-regulation (≤0.7-fold) [247].

6.3 RESULTS

6.3.1 Identification of differentially abundant proteins between FECD-affected and unaffected Descemet's membrane

In this study, nUPLC-MS^E isotope free shotgun profiling was used to identify differentially abundant proteins between FECD-affected and normal corneal Descemet's membrane. Three pairs of diseased and normal specimens from Caucasian Australians were used. The disease specimens were from patients aged 64-78 and control specimens were from donors aged 72-96 years. The controls were older than cases by design to reduce the likelihood of yet to manifest disease in controls. The analysis of quantitative differences in protein abundance between FECD and normal samples was performed by the Expression Analysis Software PLGS (Waters Corp.), using peptide ion peak intensities observed in the low collision energy mode in triplicates of each sample. The method has been

extensively described [29, 32-34] and depends on the relationship between MS signal and the corresponding protein concentration in the peptide analyte signal from each EMRT (Exact Mass Retention Time) cluster. The EMRT cluster component directly reflects the corresponding protein extract concentration. Results of proteins identified in the three sample pairs are shown in Table 6.1 (page 141 – 142). A total of 55 proteins were identified. Of these, 18 proteins were identified both in the disease and control samples, 15 in only disease samples, and 22 in only control samples.

In order to quantify identified proteins for differential abundance, the protein dataset was filtered by considering only those identifications from the alternative scanning LC-MSE data exhibiting a good replication rate (i.e., at least 6 out of 9 injections per condition, 66.7%). Additionally, a p-value of ≤ 0.05 (two-tailed Student's t-test) was set as a significant threshold for relative fold-change of each protein to be considered differentially abundant. A total of 8 proteins met these criteria (Table 6.2, page 143).

TGFBI, CLU, IGHG1, and APOE were quantified in both disease and control samples. However quantification of POTEI, HIST1H4A, ACTBL2 and ALB proteins were only in the control samples, suggesting that they are unique to normal Descemet's membrane. Ratios for the TGFBI, CLU, IGHG1, and APOE proteins were calculated using the relative protein fold-changes of individual proteins in FECD Descemet's membrane versus control Descemet's membrane (Table 6.2).

Table 6.1: Proteins identifid in FECD-affected and/or control Descemet's membrane by nUPLC-MS^{E}**.** The name of each protein, its corresponing symbol and accession number are given. The score indicates measurement of the degree of match between identified peptides of a protein and their experimental MS/MS spectra. The peptide that displayed the best score was given the highest rank and assumed as the identification result. Ratio represents signal intensities of the identified protein peptides in FECD-affected versus unaffected Descemet's membrane.

Protein name	Protein Symbol	Accession #	Score	Ratio (FECD versus normal)
Collagen alpha 3 IV chain	COL4A3	Q01955	99.62	1.38
Keratocan	KERA	O60938	109.67	0.97
C-Type Lectin Domain Family 11, Member A	CLEC11A	Q9Y240	359.89	0.91
Transforming growth factor beta induced protein	TGFBI	Q15582	1135.53	0.84
Serine protease HTRA1	HTRA1	Q92743	227.26	0.83
Clusterin	CLU	P10909	125.29	0.81
Ig gamma 2 chain C region	IGHG2	P01859	438.26	0.79
Actin cytoplasmic 1	ACTB	P60709	468.76	0.76
Complement component C9	C9	P02748	176.98	0.7
Fibulin 5	FBLN5	Q9UBX5	272.85	0.7
Ig alpha 1 chain C region	IGHA1	P01876	185.2	0.62
Ig kappa chain C region	IGKC	P01834	547.01	0.47
Collagen alpha 1 VIII chain	COL8A1	P27658	92.44	0.46
Ig gamma 3 chain C region	IGHG3	P01860	325.79	0.41
Ig gamma 1 chain C region	IGHG1	P01857	413.86	0.4
Apolipoprotein E	APOE	P02649	536.23	0.3
Actin gamma enteric smooth muscle	ACTG2	P63267	674.26	0.27
POTE ankyrin domain family member F	POTEF	A5A3E0	377.87	0.21
Transmembrane protein 179	TMEM179	Q6ZVK1	193.43	Controls only
Actin related protein 3C	ACTR3C	Q9C0K3	173.01	FECD only
Homeobox protein Hox B2	HOXB2	P14652	73.88	Controls only
Eukaryotic translation initiation factor 5A	EIF5A	P63241	112.33	Controls only
Extracellular superoxide dismutase	SOD3	P08294	108.53	Controls only
Apolipoprotein D	APOD	P05090	112.68	Controls only

Ig alpha 2 chain C region	IGHA2	P01877	93.1	Controls only
HAUS augmin like complex subunit 3	HAUS3	Q68CZ6	126.16	FECD ony
Keratin type I cytoskeletal 20	KRT20	P35900	97	Controls only
Lipid phosphate phosphohydrolase 2	PPAP2C	O43688	94.56	FECD only
Fibulin 1	FBLN1	P23142	144.71	Control only
Tachykinin 3	TAC3	Q9UHF0	147.51	FECD only
Vimentin	VIM	P08670	138	FECD only
Alpha 1 antitrypsin	SERPINA1	P01009	170.57	Controls only
Transcription factor 25	TCF25	Q9BQ70	73.16	FECD only
Collagen alpha 1 IV chain	COL4A1	P02462	79.83	FECD only
Actin cytoplasmic 2	ACTG1	P63261	468.76	Controls only
Prostaglandin H2 D isomerase	PTGDS	P41222	193.47	Controls only
Thrombospondin 4	THBS4	P35443	111.56	Controls only
GTP binding protein SAR1a	SARIA	Q9NR31	87.48	FECD only
Protein FAM90A1	FAM90A1	Q86YD7	91.71	Controls only
mRNA decapping enzyme 1A	DCP1A	Q9NPI6	69.44	Controls only
Leucine rich repeat and fibronectin type III N3	LRFN3	Q9BTN0	83.57	FECD only
E3 ubiquitin protein ligase MARCH11	MARCH11	A6NNE9	70.44	Controls only
TPT1 like protein	TPT1-Like protein	Q56UQ5	142.72	FECD only
Actin alpha skeletal muscle	ACTA1	P68133	668.77	FECD only
Cyclin D1 binding protein 1	CCNDBP1	O95273	79.51	Controls only
Putative beta actin like protein 3	POTEKP	Q9BYX7	290.45	Controls only
Serum amyloid P component	APCS	P02743	441.77	Controls only
Collagen alpha 5 IV chain	COL4A5	P29400	122.17	FECD only
Cadherin 12	CDH12	P55289	64.08	FECD only
Reticulocalbin 3	RCN3	Q96D15	143.26	FECD only
Putative uncharacterized protein	LINC00588	Q9Y4M8	161.14	FECD only
POTE ankyrin domain family member I	POTEI	P0CG38	186.8	Controls only
Histone H4	HIST1H4A	P62805	529.96	Controls only
Beta actin like protein 2	ACTBL2	Q562R1	200.06	Controls only
Serum albumin	ALB	P02768	176.2	Controls only

Table 6.2: Differentially abundant proteins identified between FECD-affected and/or control Descemet's membrane by nUPLC-MS^{E}.

Protein name	**Accession #**	**Score**	**Ratio**
Transforming growth factor beta induced protein	Q15582	1135.53	0.84
Clusterin	P10909	125.29	0.81
Ig gamma 1 chain C region	P01857	413.86	0.4*
Apolipoprotein E	P02649	536.23	0.3*
POTE ankyrin domain family member 1	P0CG38	186.8	N/A
Histone H4	P62805	529.96	N/A
Beta actin like protein 2	Q562R1	200.06	N/A
Serum albumin	P02768	176.2	N/A

*$p \leq 0.05$. N/A, not applicable because protein was unique to control samples.

Only APOE and IGHG1 showed differential abundance in the affected compared to unaffected DM (Table 6.2). Both the proteins were found to have lower abundance in the affected Descemet's membrane than in the unaffected. TGFBI and CLU, respectively, showed 0.84 and 0.81 fold change, suggesting no variation in relative abundances of these proteins in the Desecemt's membrane between FECD-affected and unaffected individuals.

A previous study by Poulsen et al [132] also identified APOE, TGFBI, and CLU by label-free MS as some of the most abundant proteins in both the normal and FECD DM. However, they were not found to be differentially regulated [132]. Quantification of TGFBI and CLU by Poulsen and colleagues [132] using iTRAQ did however reveal both proteins to have higher relative abundance in FECD-affected DM when compared to the normal DM.

Since differential abundance of APOE and IGHG1 in FECD is a novel finding, further investigation of distribution of the proteins in between FECD-affected and unaffected corneas was of interest. The data obtained from the gene expression study reported in Chapter 7 in this thesis showed that *APOE* mRNA level in FECD-affected corneal endothelium is lower than in normal corneal endothelium. The gene expression microarray did not include probe/s for the *IGHG1* gene. Thus, only APOE was prioritised for investigation. Initially, Western blotting was attempted to validate lower relative abundance of APOE in FECD-affected Descemet's membrane lysate. However, it was not found to be sensitive enough to detect the protein (data not shown). Thus, as an alternative, immunohistochemistry was performed on sections of corneas from individuals with advanced FECD and normal individuals to determine any differences in distribution of the protein. This approach was based on our previous study that revealed differential distribution of TGFBI and CLU proteins between the affected and normal corneas [93].

6.3.2 Distribution of APOE in FECD-affected and normal corneas

Immunohistochemical analysis of APOE was performed on sections of three pairs of FECD-affected and normal corneas. APOE-positive immunolabelling was observed in the corneal epithelium, stroma, DM and CE in FECD and normal corneas (Figure 6.1, page 145). Distribution of the labelling seemed more in the basal cells than in the apical cells in the epithelium of diseased cornea. In contrast, the labelling was consistently distributed throughout the epithelium in normal cornea. In Bowman's layer, no APOE labelling was observed in either FECD or normal corneas. APOE labelling was observed in the stroma in both normal and FECD corneas, but labelling in the stroma in FECD cornea was less intense. Both disease and normal corneas revealed strong but differentially distributed labelling for APOE

in the DM. The protein distributed solely in the anterior face of the DM in affected corneas. Comparatively, wider distribution of APOE labelling was observed in the DM in normal corneas.

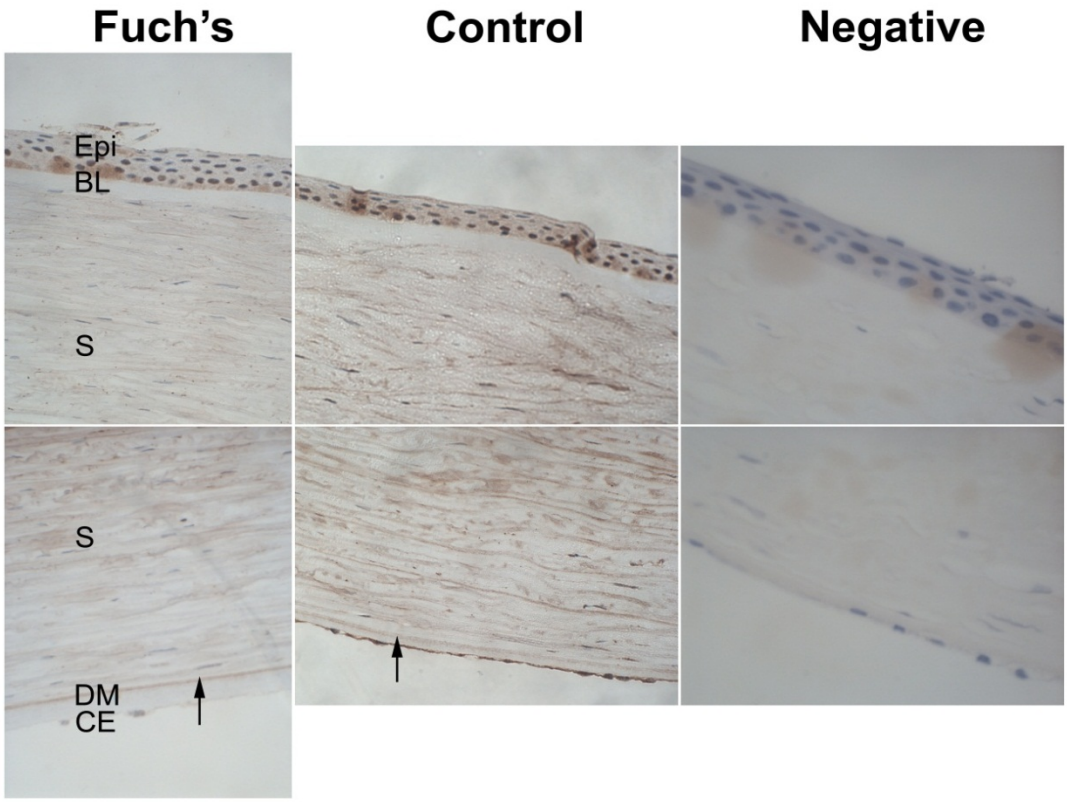


Figure 6.1 Immunohistochemical labelling for Apolipoprotein E (APOE) in sections of corneas from FECD-affected (left panels) and normal (middle panels) individuals. A negative control (no primary antibody) is shown in the right panels. APOE was detected using mouse monoclonal anti-human APOE antibody. Similar distribution of APOE-positive labelling (brown) can be seen throughout the epithelium (Epi) in both FECD and normal cornea. APOE labelling is absent in the Bowman's layer (BL) in both normal and FECD cornea. Distribution of APOE can be seen in the corneal stroma (S) in both FECD and normal cornea, but labelling in the stroma in normal cornea is more intense. The protein is distributed solely in the anterior face of the Descemet's membrane (DM) in FECD cornea. In contrast, wider distribution of APOE labelling can be seen in the DM in normal cornea. In the normal cornea, strong positive labelling of APOE is observed throughout the corneal endothelium (CE). No labelling in the negative sections indicates the specificity of the mouse anti-

human APOE primary antibody. Representative images from experiments in three independent pairs of corneas from FECD patients and controls are presented. Images are at 400× magnification. Epi, Epithelium; BL, Bowman layer; S, Stroma; DM, Descemet's membrane; CE, corneal endothelium; kDa, kilodalton

The DM appeared approximately twice as thick in the FECD cornea than the in normal cornea. This observation is consistent with abnormal thickening of the DM in the disease. Strong APOE-positive labelling was consistently observed throughout the corneal endothelium in normal corneas, but loss of endothelial cells made it difficult to ascertain similar labelling in FECD affected corneas. In conclusion, differential distribution of APOE between FECD and control DM supports down-regulation of the protein in the disease. In addition, positive immunolabelling for the protein in the endothelium, stroma and epithelium suggests its expression by the cells in these layers, and its role in the cornea.

6.3.3 Relative expression of *APOE* transcript between FECD-affected and normal corneal endothelium

Descemet's membrane is post-mitotically laid by the corneal endothelium. Therefore, a relative abundance of APOE in the Descemet's membrane can correlate with its expression and secretion by the corneal endothelium. As a result, we investigated expression of the gene in corneal endothelium from affected and unaffected individuals by quantitative RT-PCR. This analysis revealed down-regulation of *APOE* in FECD corneal endothelium compared to normal corneal endothelium (Figure 6.2, page 147; $p = 0.05$). This finding is in agreement with the result of the microarray gene expression study reported in Chapter 7 in this thesis, and, more importantly, correlates with the reduced relative abundance of APOE

protein in FECD-affected Descemet's membrane. Overall, we found altered regulation of the *APOE* gene in the corneal endothelium in FECD.

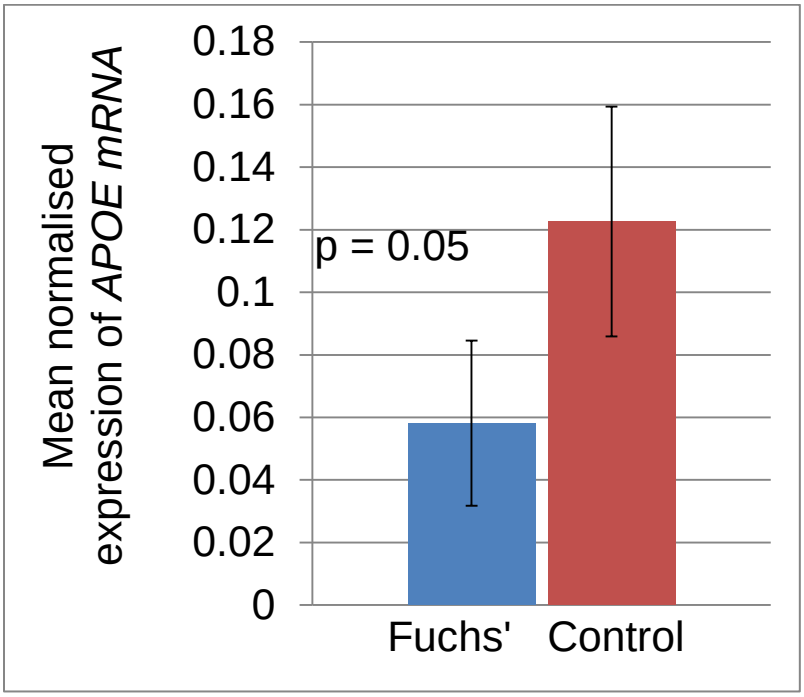


Figure 6.2 Relative expression of *APOE* mRNA levels in FECD-affected and normal corneal endothelium. Relative mRNA expression levels were determined by quantitative RT-PCR. Relative expression was normalised against expression of *ACTB* reference gene. Data are presented as mean normalised expression ; error bars indicate standard deviation. Statistical analysis was performed using Student's *t*-test.The p-value for significant differential expression between affected and unaffected corneal endothelium was at $p \leq 0.05$.

6.4 DISCUSSION

The aim of this study was to identify proteins with differential relative abundance between FECD-affected and unaffected DM. Label-free nUPLC-MS^E was used to achieve this aim. Label-free nUPLC-MS^E is the latest high throughput comparative proteomic technique for identification and quantification of proteins in a specific tissue of interest under two or more biological conditions.

In the present study, nUPLC-MSE identified 55 proteins in FECD and control DM. However, only 8 of these proteins met the set criteria for quantification analysis for differential relative abundance between the affected and unaffected DM. Out of the 8 proteins, only APOE and IGHG1 showed significant differential abundance in FECD DM. Both proteins were down-regulated in the disease. To the best of my knowledge, this study is the first to report lower relative abundance of APOE and IGHG1 in FECD-affected DM.

While this study was in progress, Poulsen et al [132] published findings from their study on comparative proteomic analysis of components of DM between FECD-affected and unaffected individuals. The label-free quantitative and isobaric tags for relative and absolute quantification (iTRAQ) techniques were employed. Both techniques independently identified and quantified dozens of proteins in both FECD-affected and unaffected DM, using the same tissue as used in this study [132].

Whereas Poulsen and colleagues identified 48 proteins from the diseased and normal DM by label-free method, only 10 were significantly differentially abundant [132]. Of these differentially abundant proteins, AGRN, APOD and KERA showed higher

abundance, whereas COL2A1, COL3A1, COL4A4, FBN1, KRT1, KRT9 and KRT10 showed lower abundance in the disease [132].

In contrast to the present study, Poulsen and colleagues identified APOE as most abundant protein in both diseased and normal DM, but it was not found to be significantly differentially abundant [132]. Furthermore, IGHG1 was not identified in either FECD-affected or unaffected DM in the study by Poulsen et al [132]. Although APOD, KERA and COL4A3 were identified in both diseased and normal DM in the present study, these proteins were not significantly differentially regulated. Observed differences in the results between this study and that of Poulsen and colleagues [132] could be accounted for by variation in the number and characteristics of samples used, protocols followed for protein extraction and digestion, and techniques and methods used for peptide analyses, identification and quantification.

In my study, three pairs of DM specimens from three FECD patients and three healthy deceased donor individuals were used. In contrast, Poulsen et al [132] analysed ten FECD-affected DM and four age-matched controls with pseudophakic bullous keratoplasty. It is likely that analyses of a larger number of specimens in the study by Poulsen et al [132] may have increased the number of proteins identified to be differentially abundant. To further understand the importance of my findings, immunohistochemistry was performed for APOE on sections of corneas from individuals with advanced FECD and normal individuals to investigate differences in distribution of the protein.

Positive immunolabelling for APOE indicates its presence in all the corneal layers, except Bowman's membrane, in both the diseased and normal corneas. The

Bowman's membrane is composed of only collagenous extracellular matrix proteins secreted by the anterior stromal keratocytes [3, 11]. Lack of positive APOE immunolabelling in this layer might be due to undetectable levels of the protein. APOE labelling in the corneal epithelium, stroma, Descemet's membrane, and corneal endothelium indicates for the first time that it plays an important functional role in these corneal layers.

The corneal epithelium and stroma are not primarily affected during the onset of FECD. However, they are affected in the advanced stage of the disease [177]. The putative lack of observed differential distribution of APOE in the epithelium and stroma might be related to the fact that they are not the primary disease-affected corneal layers.

Corneal endothelium and its basal membrane, Descemet's membrane, are the main layers affected during the onset of the disease [35]. Normal DM is made up of two distinct layers, the anterior and the posterior [14]. The anterior DM, or anterior banded layer, consists of banded collagens and is laid down *in utero* [14]. The posterior DM, normally referred to as a posterior non-banded collagen layer , is progressively secreted by the corneal endothelium throughout life [14].

In FECD, the PNBL is significantly attenuated, or completely lost, and is replaced by an additional abnormal banded collagen layer called posterior collagenous layer (PCL), in the extra posterior DM [249]. The PCL is hypothesised to be secreted when CECs are stressed by damage or disease [15], thus leading to abnormally thickened DM in FECD. Interestingly, my work revealed that APOE distributed solely in the ABL in FECD-affected corneas, suggesting that the loss of the protein in the Descemet's membrane may be related to the disease. This proposal correlates

with lower abundance of the protein in diseased DM, as revealed by Label-free nUPLC-MS^{E}.

Since protein components of the DM are produced by the CE [15], I investigated the latter for expression of *APOE* mRNA between individuals with advanced FECD and normal individuals. Consistent with our proteomic and immunohistochemical findings, relative gene expression analysis showed down-regulation of *APOE* in FECD CE, and suggests that altered regulation of *APOE* in CE in FECD may underlie observed lower abundance of the protein in the disease-affected DM. Similarly, previous studies revealed a correlation between over-expression of the *TGFBI* and *CLU* genes in the CE with the TGFBI and CLU proteins in the DM in FECD [114].

APOE has been reported to indirectly regulate expression of *CLU* and *TGFBI* genes through *FOS*. Over-expression of *APOE* decreases *FOS* expression [250]. In rat 208F cells, *FOS* increases expression of *CLU* and *TGFBI* [251-253]. This suggests that lower expression of *APOE* can lead to over-expression of *FOS*, and subsequent, abnormal over-expression of *CLU* and *TGFBI* in FECD. In line with these findings, lower expression of *APOE* in CE in the disease might underlie overexpression of *FOS*, *CLU* and *TGFBI* in CE, and correlating higher abundance of TGFBI and CLU in DM in the disease [114, 131, 132].

APOE is a multifunctional, low-density lipoprotein receptor (LDL-R) ligand that primarily serves as a lipid transporter. It also serves as an antioxidant [254], and a regulator in the innate immune system. It is a 299-amino acid long protein that consists of two independently folded domains that are linked by a protease-sensitive loop [255]. The first is the amino-terminal domain (residues 1 – 191) that has an

ordered structure consisting of four amphipathic α helix bundles (i.e. helix 1 – 4) [256]. The segment containing helix 4 has a high affinity binding site for heparan-sulfate proteoglycans (HSPGs) with lysine- and arginine-rich residues interacting with the negatively charged carboxylate and sulphate groups of proteoglycans [257, 258]. Furthermore, residues 243 – 272 of the APOE are thought to bear heparin-binding site [257].

HSPGs are basement membrane proteins, and major components of the DM [3]. In addition, they are secreted by the CE throughout life and are constituents of the extracellular matrix (ECM) in the PNBL [35]. My study has revealed the presence of APOE in the ABL and PNBL in normal corneas, supporting uterine and postnatal expression and secretion of the protein by CE, and functional interactions with HSPGs in the DM. I hypothesise that APOE plays a role in maintenance of functional integrity of the normal DM by providing high affinity binding site for HSPGs. Down-regulation of the *APOE* gene, and the APOE protein in FECD-affected CE and DM respectively, could result in dysregulated binding of HSPGs, and ultimately alterations in ECM of the DM in late-onset disease; as previously observed [131, 132].

APOE has been implicated in many diseases including age-related macular degeneration (ARMD), Alzheimer's disease (AD), and Atherosclerosis [254, 259]. Oxidative stress significantly contributes to the pathogenesis of these diseases [260, 261], as well as being the leading pathway in the pathogenesis of FECD. An oxidant-antioxidant imbalance, which is proposed to be a major contributor in generation of the oxidative stress has been also reported in the corneal endothelium in FECD [69]. In addition, under-expression of

peroxiredoxins and thioredoxin-dependent antioxidants in corneal endothelium of patients with FECD compared to age-matched controls has been reported [130, 245]. Thus, I hypothesise that dysregulation of *APOE* found in this study is likely to lead to an oxidant-antioxidant imbalance in the corneal endothelium in the disease, resulting in oxidative stress, which would contribute to pathogenesis of FECD. Further investigations are required to investigate this hypothesis.

The present study has some limitations. Due to scarcity of the CE+DM specimens, only three pairs of disease-affected and unaffected DM samples were used for comparative quantitative analysis. Analyses of a larger number of specimens may have increased the number of proteins identified and quantified. Another limitation was that a quantitative validation technique could not be used because of feasibility issues with Western blot.

In conclusion, the present study has revealed for the first time lower abundance of APOE and IGHG1 proteins in FECD-affected DM. Lower abundance of APOE in the disease-affected DM correlates with differential distribution of the protein in the DM, and down-regulation of the gene transcript in corneal endothelium in the affected individuals compared with unaffected individuals. Collectively, these findings suggest altered regulation of the APOE in FECD, which may contribute to disease pathogenesis.

CHAPTER 7
Identification of differentially expressed genes between FECD-affected and unaffected corneal endothelium

7.1 INTRODUCTION

Corneal endothelium secretes protein components of the Descemet's membrane [243]. Reported findings from previous molecular studies [73, 114, 154] have consistently shown correlation in expression of differentially regulated genes and proteins in corneal endothelium and Descemet's membrane, respectively, of FECD patients compared with unaffected individuals. We therefore hypothesised in this study that investigation of comparative transcriptomes between disease-affected and unaffected corneal endothelium can reveal novel differentially expressed genes, which can shed light on the molecular mechanisms underlying the pathogenesis of FECD.

To date, only one study has been performed to investigate genome-wide expression of genes between affected and unaffected corneal endothelium using serial analysis of gene expression (SAGE) [127]. Therefore, there was a need for an independent gene expression study to identify additional differentially expressed genes using a different approach. In the present study, I aimed to identify differentially expressed genes between FECD-affected and unaffected corneal endothelium using microarray analysis. Unlike a candidate gene approach that is normally limited to a small number of pre-determined genes, microarray analysis provides a snapshot of global gene expression changes without making assumptions about the relevance of each gene to the condition being investigated [262, 263]. Microarray results were

validated by quantitative RT-PCR. Some novel pathways were identified that shed light on the mechanism of FECD.

7.2 MATERIALS AND METHODS

7.2.1 Microarray analysis

RNA extraction was performed from pooled FECD-affected CE+DM specimens (n= 3 per pool) and three individual normal specimens, as described in Chapter 2 (section 2.3.2). A garose gel electrophoresis, performed as detailed in section 2.4.1.3 in Chapter 2, was used to assess RNA integrity and any residual DNA contamination. Microarray analysis was performed at the Australian Genome Research Facility (AGRF Ltd; The Walter and Eliza Hall Institute of Medical Research, VIC, Australia).

The quality of total RNA samples was ascertained using the Agilent 2100 Bioanalyser 2100 (Agilent Technologies, Santa Clara, CA, USA), following the NanoChip protocol, at the AGRF. A total of 500ng RNA was labelled with biotin using the Ambion® Total Prep RNA amplification kit (Catalogue number: IL1791; Life Technologies). The quantity of labelled product was ascertained using the Agilent 2100 Bioanalyser (Agilent Technologies) according to the NanoChip protocol.

A total of 1.5µg of biotin-labelled cRNA was used for hybridisation to the *Sentrix* Human HT-12 v3.0 Gene Expression BeadChip (Illumina Inc., San Diego, California, USA), after preparation of a probe cocktail (cRNA at 0.05µg/µl) that included GEX-HYB hybridisation buffer (Illumina Inc.). A hybridisation volume of

30 µl was prepared for each sample, and loaded into a single array. Six different labelled samples were loaded in duplicate into 12 individual arrays on the BeadChip.

Hybridisation was performed for 16 hours at 58°C in an oven with a rocking platform. After hybridisation, the chip was washed following the protocol outlined in the manual (Illumina Inc.), to remove unhybridised transcripts. Upon completion of the washing, the chip was coupled with streptavidin-Cy3 and scanned in the iScan scanner (Illumina Inc.) to obtain measures of gene expression. The scanner operating software, GenomeStudio (Illumina Inc.), converted the signal on the array into a TXT file for analysis.

Raw probes signal intensity values were background-adjusted using the Illumina® BeadStudio GX Module software (Illumina Inc.). Expression level was adjusted to 50% of the background level (estimated from negative control genes) to avoid negative expression values. Background-corrected expression values were then normalised with $\log_2$ transformation and variance stabilisation using the lumiR package of R Bioconductor (www.Bioconductor.org) to remove systematic/technological noise present in the data, to improve the detection of differentially expressed genes. Probe sets where all values for all samples fell below expression detection threshold were removed.

Normalised data was subsequently analysed in Partek Genomics Suite 6.6 (www.Partek.com) using one-way ANOVA (analysis of variance) to calculate significance of gene expression variation between tested sample groups. Sample replication was taken into account during the analysis. This ANOVA was performed with p-value adjustment for all genes identified. A fold change was calculated as

mean ratio. Probes with an unadjusted p-value of 0.05 or less and an absolute fold change of 1.5 or more were defined as differentially expressed.

A two-dimensional PCA (principal component analysis) was performed on genes with adjusted p-values that met the cut-off values indicated by ANOVA to identify factors that explain the most important variations between the datasets. Genes identified were then filtered with CV (covariance) >20%, and a three-dimensional PCA was conducted on the filtered genes. Three lists of differentially expressed genes were generated with up- or down-regulated log fold change of 1.2 and p-values filters of <0.001, 0.01 and 0.05.

Hierarchical cluster analysis of the three genes lists, Squared Euclidean and Average Distance parameters were performed to discover groups of genes with similar expression patterns in the data. In addition, volcano plots for the three gene lists were generated by plotting the negative log of the p-values (on the y-axis) against the log of the fold-changes (on the x-axis) between the cases and normal controls datasets.

In the microarray data analysis performed at the AGRF, the total number of probes available on the chip were considered for analysis. The unexpressed probes were not removed and thus could increase the likelihood of false positive results. To increase reliability of the results, the data was re-analysed using the *Limma* R-Statistics program [264], by considering only the probes expressed in corneal endothelium. The analysis was conducted through collaboration with Dr. Mark Corbett, Neurogenetics Research Laboratory, SA Pathology, Women's and Children's Hospital, Adelaide, Australia.

In brief, background subtracted expression values produced using Illumina BeadStudio software (Illumina Inc.) were quantile normalised using the *lumi* package of R Bioconductor [265]. Probe sets where all values fell below expression detection threshold were removed, and subsequent values analysed for significant differential expression taking into account sample replication [266]. Boxplots were used to visualise the distribution of the un-normalised and normalised data. Normalised gene expression levels were analysed for significant differential expression using the empirical Bayes moderated t-statistics [264]. The resulting p-values were adjusted by a multiple testing procedure to account for false discovery rate.

7.2.2 Quantitative RT-PCR analysis

Quantitative RT-PCR was performed to validate the prioritised differentially expressed genes in independent samples of CE+DM tissue from affected and unaffected individuals as described in Chapter 2 (See Section 2.4.3) with the following exceptions. Three pairs of pooled CE+DM specimens from FECD patients (n = 5 per pool) and normal controls (n = 2 per pool) were used for RNA extraction. For cDNA synthesis 0.471 µg of total RNA was reverse-transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen) and random hexamer primers, as detailed in Chapter 2 (Section 2.4.3.2). Gene-specific pre-designed quantitative RT-PCR primers listed in Table 2.2, Chapter 2, were used to perform quantitative RT-PCR as described in Chapter 2 Section 2.4.3.3. Data were analysed using the StepOne Plus software, as described in Chapter 2 Section 2.4.3.4.

7.2.3 Bioinformatic analysis

In order to prioritise differentially expressed genes for validation, Ingenuity Pathway Analysis (IPA) was performed using the Ingenuity® Knowledge Base. The analysis

was performed using the core analysis option, and by considering only direct relationships among the genes. Furthermore, pie charts were drawn to show distribution of the genes by their functions or localisation in the corneal endothelium. IPA was also performed to visualise functional relationships among the differentially expressed genes.

7.3 RESULTS

7.3.1 Characteristics of participants for expression analysis

For the microarray analysis CE+DM samples from 9 FECD patients (mean age 65.1 ± 5.5 years; male-to-female ratio 1:2) and 3 normal deceased donors (mean age 79.3 ± 6.7 years; male-to-female ratio 2:1) were used. Student's t-test respectively showed p values of 0.310 and 0.066 for differences in mean age and gender between cases and controls. For validation of differentially expressed genes CE+DM samples from 15 FECD patients (mean age 69.5 ± 12.6 years; male-to-female ratio 1.14:1) and 6 normal deceased donors (mean age 87.2 ± 5.9 years; male-to-female ratio 1:2) were used. According to the Student t-test, p values of differences in mean age and gender between cases and controls were 0.407 and 0.082, respectively. Thus there was no difference in mean age and gender between cases and controls from whom specimens were used for microarray analysis and validation of findings by quantitative RT-PCR.

7.3.2 Quality assessment for RNA samples

Initial analysis of total RNA samples by agarose gel electrophoresis to assess their quality for use in microarray analysis showed good RNA integrity with minimum residual DNA contamination (Figure 7.1, page 160). Before microarray analysis,

RNA samples were quality assessed using the Agilent Bioanalyser 2100 at the *AGRF* and passed QC.

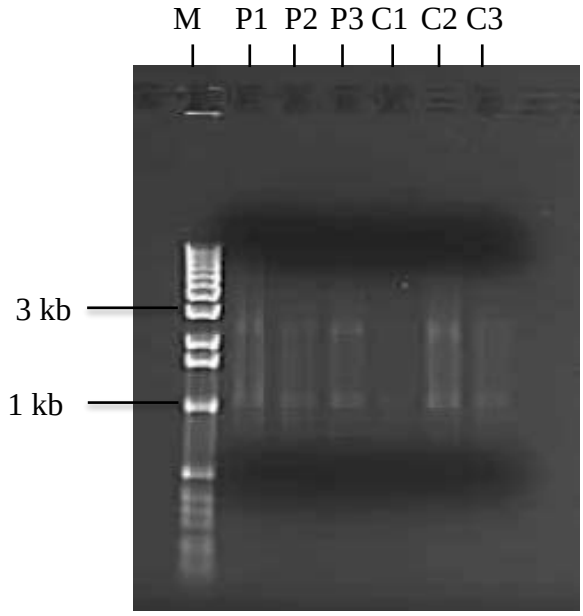


Figure 7.1 Agarose gel electrophoresis of FECD case and control RNA samples used for microarray analysis. P1 to P3 represent total RNA extracted from three pooled FECD cases (n = 3 per pool). C1 to C3 represent RNA samples extracted from three independent normal individuals. M = marker DNA molecular size standard. The 18S and 28S ribosomal RNAs are marked with the arrows.

7.3.3 Microarray analysis

Microarray analysis was performed using the Illumina HumanHT-12 v3 Gene Expression array. The array contained 48,804 probes, representing 29,265 genes. Each sample was analysed in duplicate. Principal component analysis showed clear distinction between FECD samples and control samples, and a close similarity in technical duplicates within each biological group (Figure 7.2, page 161). One biological replicate from the control group was different from the other two

replicates, but was still considered as acceptable because it was clearly distinct from the cases.

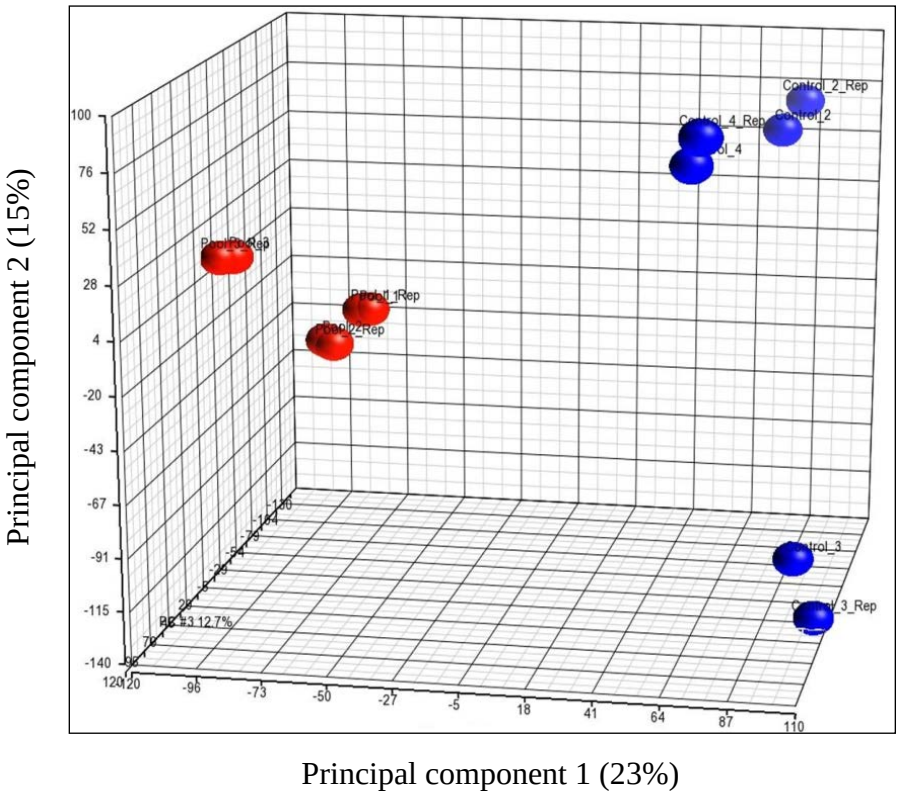


Figure 7.2 Principal component analysis of FECD-affected (red dots) and control (blue dots) samples. Each sample was analysed in duplicate. FECD samples are clearly distinguishable from control samples. There is close similarity in technical duplicates within each biological group. One of the biological replicates in the control group varied from the other two replicates.

Consistently, boxplots of microarray intensity from non-normalised data showed an observable difference in dye intensity between the case and control groups, and a reasonable similarity among replicates within each group (Appendix, Figure 1, page 212). The raw data was successfully quantile normalised as evident from boxplots of normalised data (Appendix, Figure 2, page 212). One-way ANOVA revealed several differentially expressed genes (logFC = 1.2, $p \leq 0.05$) in corneal endothelium between

cases and controls. The volcano plot shows the distribution of the full gene set in Figure 7.3. Even under very strict differential expression cut-off of logFC of 10 and associated p-value of 0.001, 156 genes remained differentially expressed (Figure 7.4, page 163).

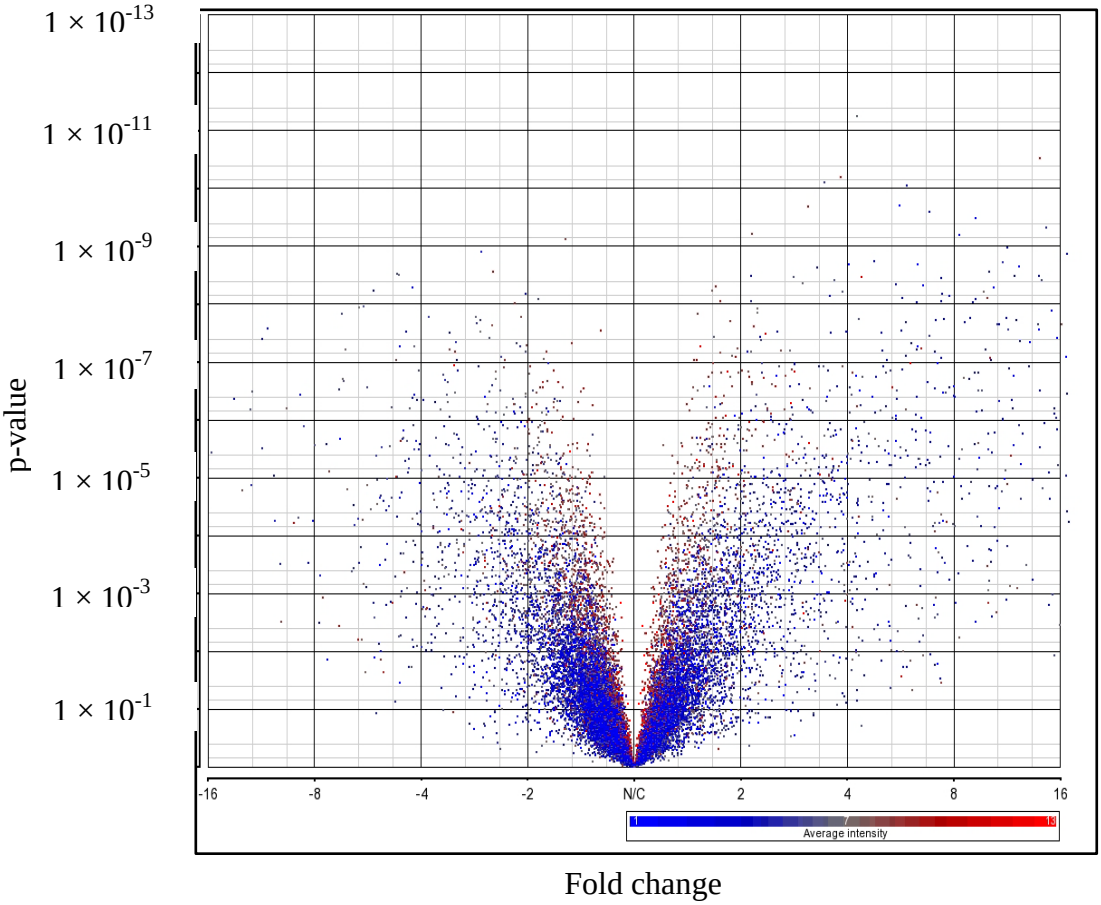


Figure 7.3 Volcano plot of 29,265 genes showing distribution of their expression in corneal endothelium between FECD cases and controls. Each point represents a gene plotted as a function of fold change (Log2 (fold change), x-axis) and statistical significance (-Log 10(p-value), y-axis). "-" indicates underexpression and "+" indicates overexpression. Red dots depict genes with higher average signal intensity in the array. Blue dots depict genes with lesser average signal intensity in the array. N/C represents neutral/constant expression.

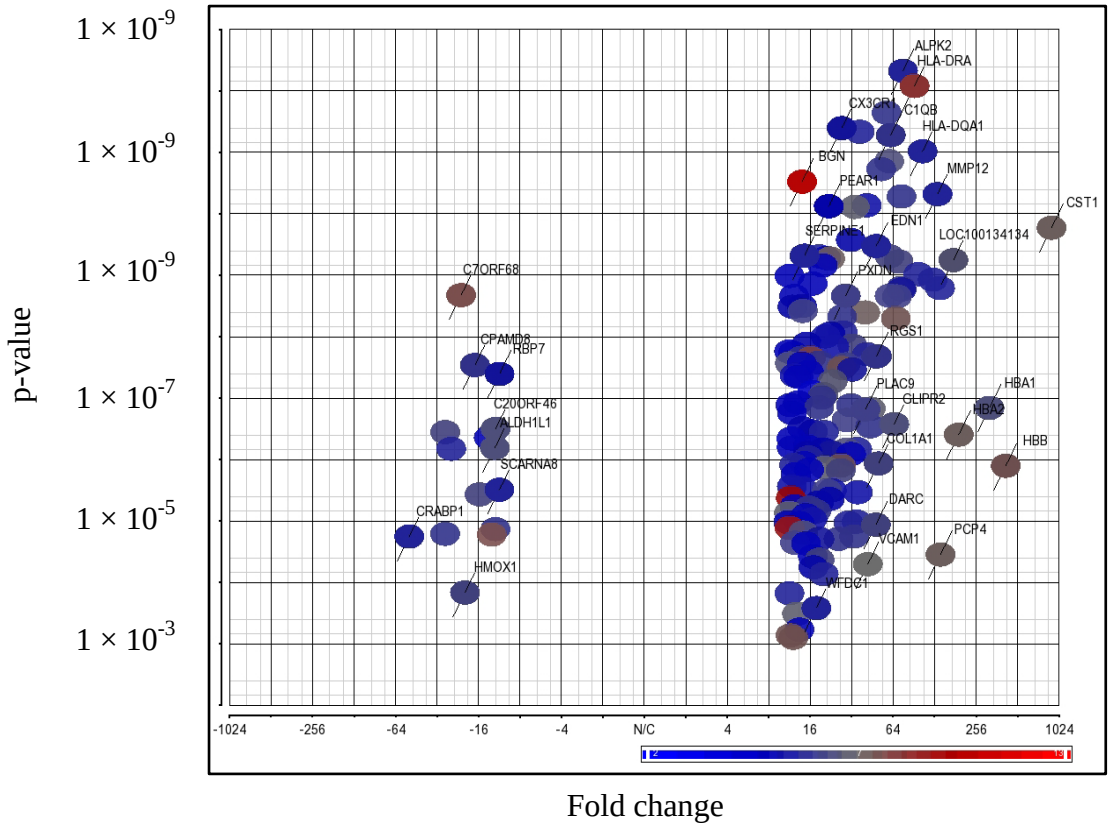


Figure 7.4 Volcano plot of 156 differentially expressed genes (logFC = 10, p = 0.001) between FECD-affected and unaffected corneal endothelium. Each point represents a gene plotted as a function of fold change (Log2 (fold change), x-axis) and statistical significance (-Log 10(p-value), y-axis). "-" and "+" respectively refer to under and over expression. Red and blue dots respectively depict genes with higher and lower average signal intensity in the array. N/C represents neutral/constant expression.

Since the microarray data analysis conducted at the AGRF involved all probes on the array, there was a likelihood for false positive differential expression of the analysed probes. To avoid this, the data was re-analysed considering only the probes representing genes expressed in corneal endothelium. The 22450 expressed probes

were thus re-analysed. The volcano plot of expression distribution of these probes is shown in Figure 7.5.

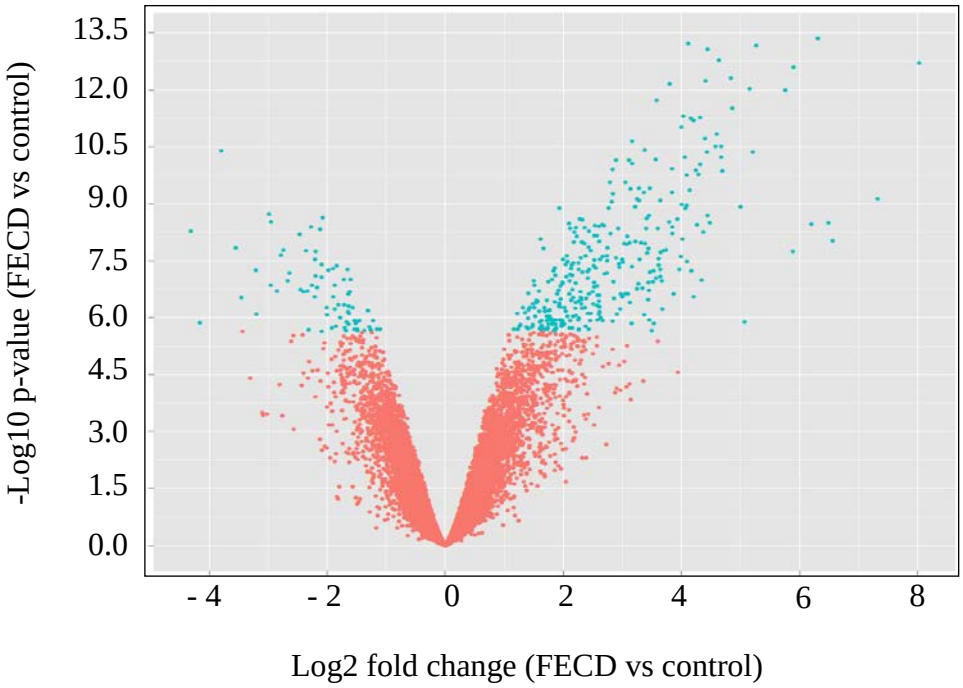


Figure 7.5 Volcano plot of microarray data from the second analysis showing distribution of expression of 22450 probes representing expressed genes in corneal endothelium between FECD-affected and unaffected samples. Each dot represents a probe plotted as a function of fold change (Log2 (fold change), x-axis) and statistical significance (-Log 10 (p-value), y-axis). Cyan dots represent significantly differentially expressed probes between FECD corneal endothelium and control corneal endothelium, at Bonferroni-corrected p-value of <0.05. Orange dots represent all the probes with no variation in expression between affected and unaffected corneal endothelium.

Following Bonferroni multiple testing correction, 396 genes showed significant differential expression with fold change of ≥1.2 in up- or down-regulation and adjusted p≤0.05. The heat map (Figure 7.6, page 165) illustrates the spread of gene

expression in CE between diseased and normal tissue. This is clearly evident from the heat map of these 396 genes (Figure 7.6). This number was too large for further analysis, therefore a fold-change cut-off of >1.5 up- or down-regulation was applied to generate a new list of differentially expressed genes. A total of 135 genes (represented by 161 probes) were thus found to be differentially expressed (Table 7.1 and Appendix Table 4, on page 166 and 213 – 216, respectively). Majority of these genes (n = 123) were over-expressed in FECD compared to control samples.

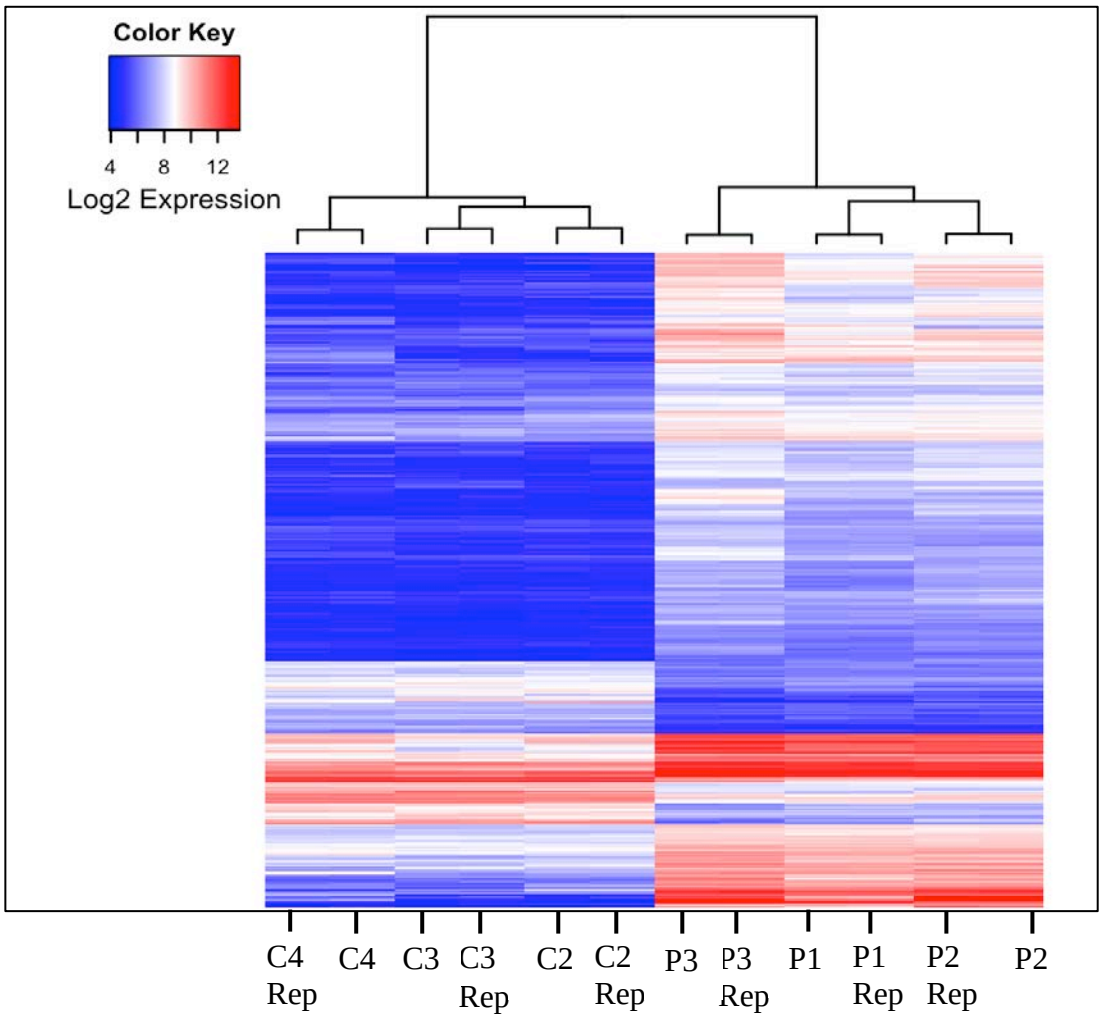


Figure 7.6 Heat map of hierarchical clustering of the 396 differentially expressed genes between FECD-affected and normal corneal endothelium, with >1.5 fold-change in up- or down-regulation and adjusted p≤0.05, based on Holm-Bonferroni

multiple testing correction. The hierarchical clustering of the differentially expressed genes confirmed the relationship observed in the principal component analysis, Figure 7.2, between FECD samples and control samples; as it demonstrates a clear distinction between the two biological groups. Red and blue colours respectively indicate overexpressed and underexpressed genes.

Table 7.1 The top 20 differentially expressed genes in FECD-affected corneal endothelium compared to unaffected corneal endothelium. Genes with log fold-change of ≥3.16 in up- or down-regulation, and Bonferroni-adjusted p-value $\leq 10^{-5}$ are listed. Symbol of each gene and its corresponding accession number in the NCBI database are given. The log fold-change and adjusted p-values are also given.

Gene symbol	**Accession #**	**Log Fold Change**	**Bonferroni adjusted *p* value**
HLA-DRA	NM_019111	6.31	4.28×10^{-8}
CX3CR1	NM_001337	4.11	5.74×10^{-8}
ALPK2	NM_052947	4.44	8.38×10^{-8}
CSF1R	NM_005211	4.63	1.56×10^{-7}
CST1	NM_001898	8.02	1.89×10^{-7}
HLA-DQA1	XM_936128	4.83	4.73×10^{-7}
C1QB	NM_000491	4.41	5.75×10^{-7}
BGN	NM_001711	3.80	6.61×10^{-7}
HLA-DPA1	NM_033554	5.15	9.32×10^{-7}
MMP12	NM_002426	5.76	9.92×10^{-7}
TSHR	NM_000369	3.58	1.86×10^{-6}
SPP1	NM_001040058	4.03	4.75×10^{-6}
C1QC	NM_172369	4.31	5.30×10^{-6}
CLIC6	NM_053277	4.15	5.59×10^{-6}
ITGB2	NM_000211	4.20	6.37×10^{-6}
EDN1	NM_001955	3.99	9.19×10^{-6}
TYROBP	NM_003332	4.59	1.41×10^{-5}
NOX4	NM_016931	4.40	1.84×10^{-5}
CD74	NM_001025159	3.16	2.22×10^{-5}
ALOX5AP	NM_001629	4.57	3.04×10^{-5}

All the 135 DE genes were categorised into subgroups based on their biological functions and/or localisation in cells, as derived from Kyoto Encyclopaedia of Genes

and Genomes (KEGG). The subgroups are presented as pie charts (Figures 7.7 and 7.8 (page 168)) showing their predominant functions and location in the endothelium.

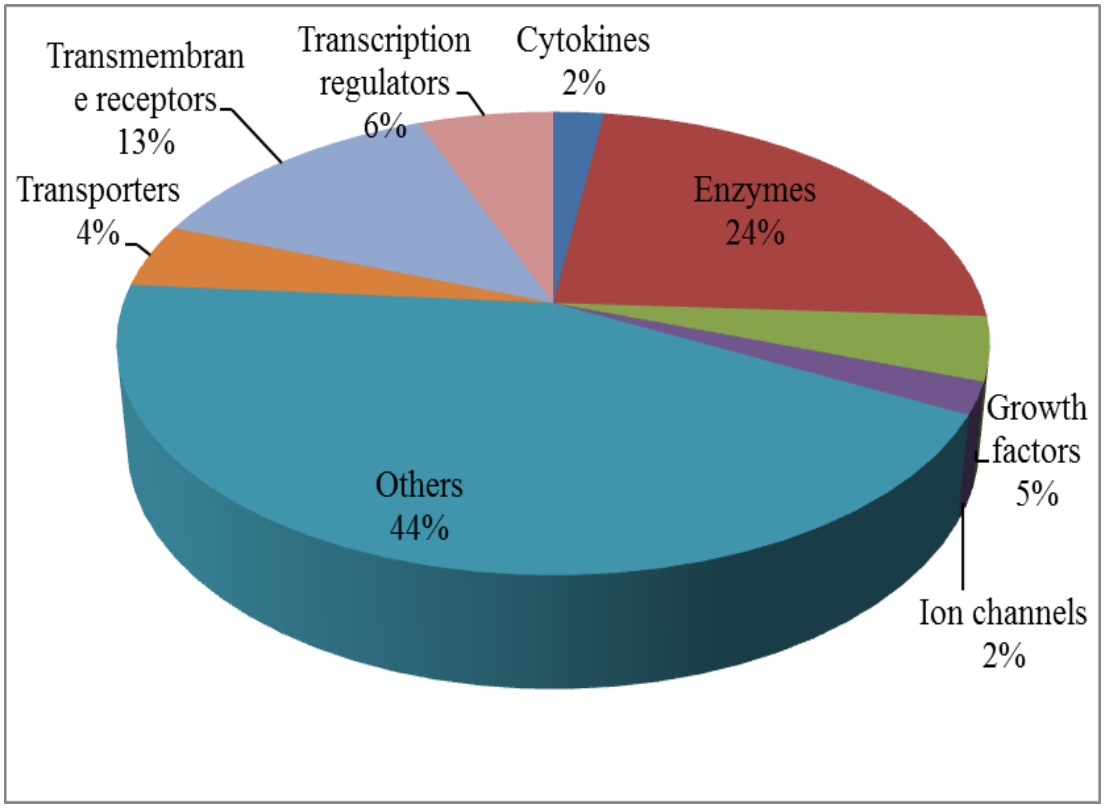


Figure 7.7 Distribution of 135 differentially expressed genes based on their biological functions, as derived from the KEGG knowledge Base. Genes responsible for general cellular functions (collectively referred to as others) are overrepresented among the differentially expressed genes between FECD-affected and unaffected corneal endothelium. The second most over represented genes encode for enzymes.

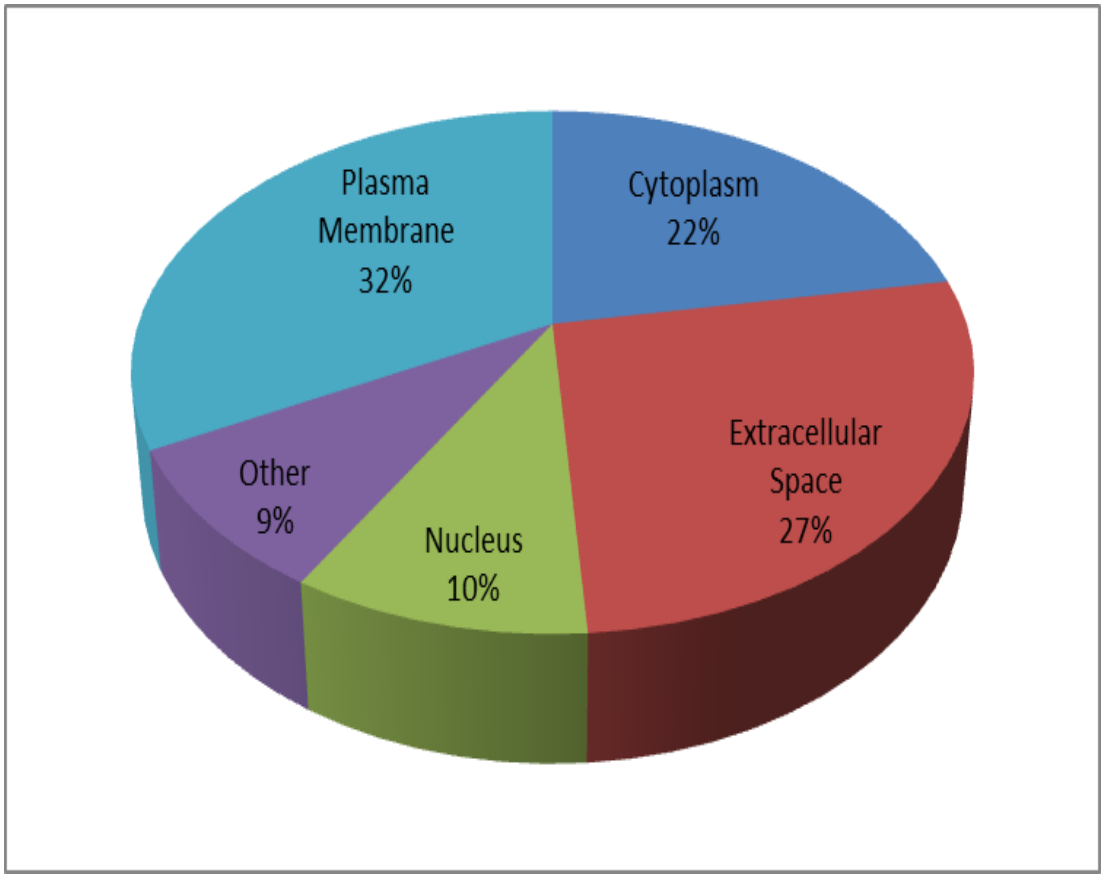


Figure 7.8 Distribution of 135 differentially expressed genes based on their sub-cellular location in cells as per KEGG knowledge Base. Genes responsible for plasma membrane and extracellular proteins are overrepresented in the list of the differentially expressed genes between FECD-affected and unaffected corneal endothelium.

7.3.4 Prioritisation of differentially expressed genes for validation

As many differentially expressed genes were identified, they were prioritised for validation. For this, seven previously reported FECD-related genes (*COL8A2*, *CLU*, *LOXHD1*, *SLC4A11*, *TCF4*, *TGFBI*, and *ZEB1*) were added to the 135 genes and all the 142 genes were analysed using IPA. Core analysis was performed considering only direct relationships among the genes being analysed. The top 20 genes from the 135 differentially expressed genes were then ranked based on the hierarchy of the

networks in which they were included, and functional networks they shared with the 7 FECD-related genes included in the analysis. Thus, 10 top-ranked genes were prioritised for validation. In the order in which they were prioritised, these genes included *EDN1*, *SPP1*, *HLA-DRA*, *CSF1R*, *BGN*, *CX3CR*1, *ALPK2*, *CLIC6*, *NOX4* and *CST1*. The genes were all over-expressed in FECD-affected corneal endothelium compared to normal corneal endothelium. In order to account for any bias that may arise from validation of unidirectional, differentially expressed genes, two top under-expressed genes (*CPAMD8* and *PPP1R1B*), which respectively ranked 38 and 79 in the list of the 135 DE genes, were included for validation. Therefore, a total of 12 differentially expressed genes were prioritised for validation.

7.3.5 Demonstration of specificity of primers used for quantitative RT-PCR

For each analysed gene in this study, cDNA was amplified on a StepOne Plus real-time PCR system (Applied Biosystems) using RT^2SYBR^2Green master mix (SABiosciences) and gene-specific forward and reverse primer mix (0.5μM final concentration each). DNA polymerase was activated at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds and annealing and extension at specific temperature (°C) for 1 minute. Analysis of data using the StepOne Plus software verified melt curves for the amplified genes to be specific. Additionally, all the primers used for the 12 DE prioritised genes were confirmed to be specific and amplified a single PCR product. For examples, representative agarose gel electrophoreses for *ACTB* and *GAPDH* reference genes, and a differentially expressed *HLA-DRA* showed specific, amplified single bands of the PCR products (Figure 7.9, page 170).

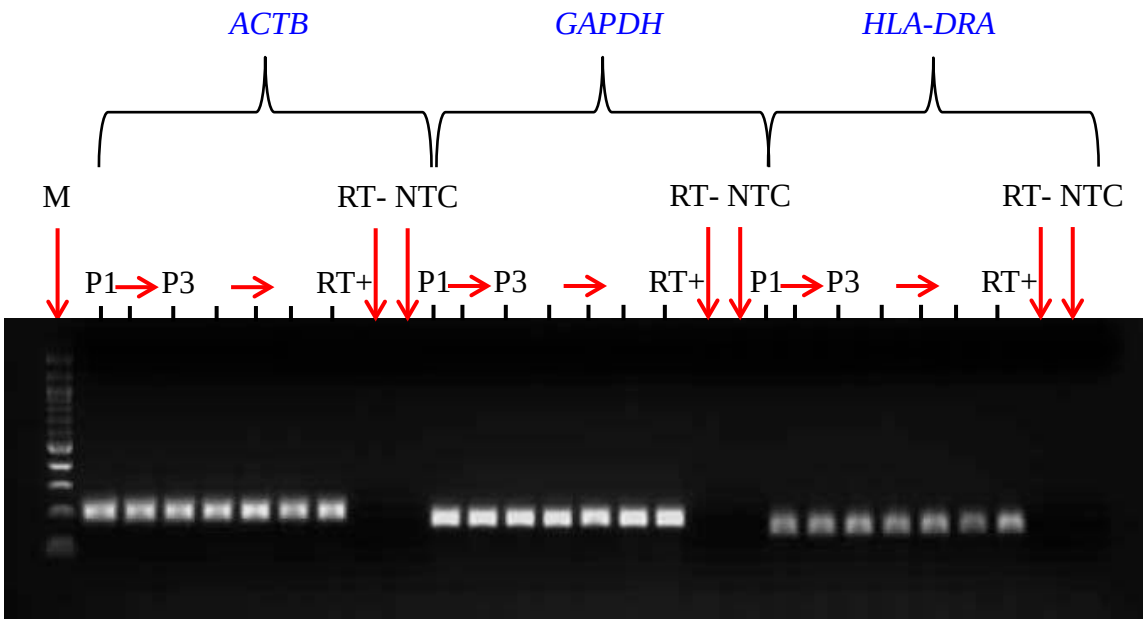


Figure 7.9 Agarose gel electrophoresis of products of *ACTB*, *GAPDH* and *HLA-DRA* genes from quantitative RT-PCR analysis. PCR products are shown for each gene of interest and found to be of the expected sizes. P1 to P3 are PCR products from cDNA from FECD RNA pool 1 to 3, respectively. C1 to C3 are PCR products from cDNA from control RNA control 1 to 3, respectively. RT+ refers to a standard cDNA sample. RT- refers to a standard cDNA sample without reverse transcriptase. NTC = no template control. M = DNA ladder (100bp) ladder.

7.3.6 Determination of amplification efficiencies of primers used for quantitative RT-PCR

For each gene of interest, a standard curve of its quantitative RT-PCR assay was generated as described in Chapter 2 section 2.4.4.3, and used to deterimne the *E* as shown for the representative gene *CST1* (Figure 7.10, page 171). The *E* values for the 12 DE genes prioritised for validation ranged from ~ 88 – 110% (Table 7.2, page 172). The R^2 values ranged from 0.9712 to 0.9994 (Table 7.2). The *E* values for all the genes were within the acceptable range of 85 to 105%, according to the MIQE

(minimum information for publication of quantitative RT-PCR experiments), with the exception of the *E* value for *ALPK2*, which was slighly higher [267, 268]. All assays were considered suitable for calculating relative expression levels of the genes in FECD-affected and unaffected CE.

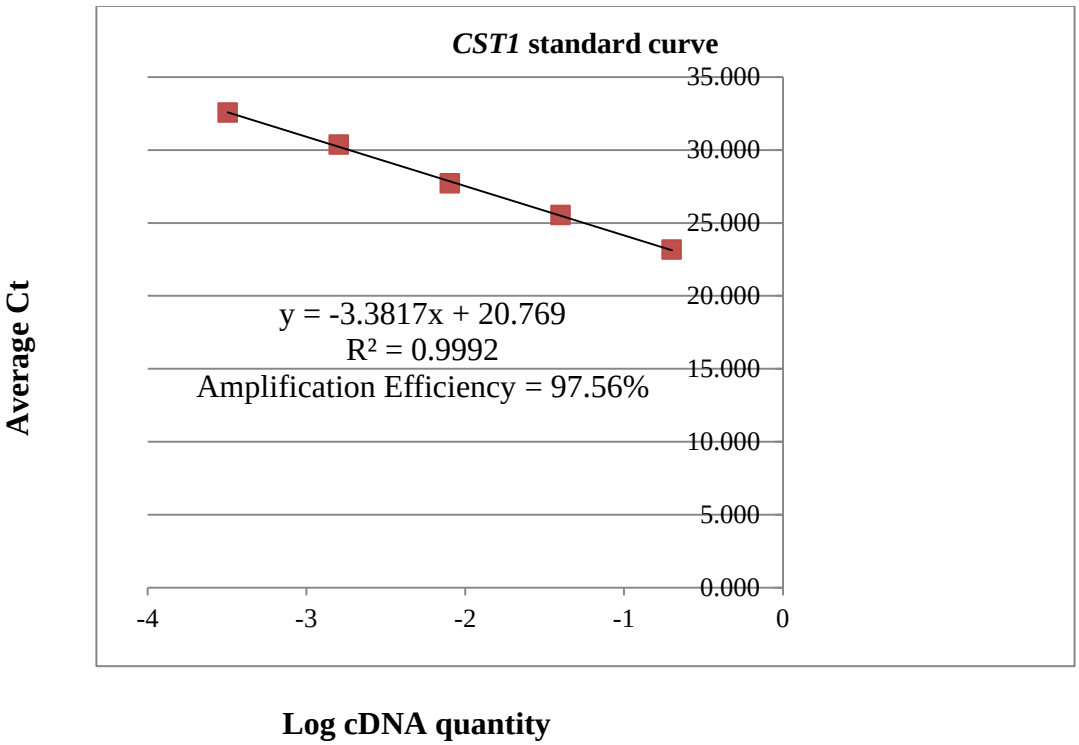


Figure 7.10 The standard curve for *CST1* quantitative RT-PCR assay. Quantitative RT-PCR was performed using five volume/volume five-fold serial dilutions of the standard cDNA, starting with 1:20 dilution (See Chapter 2, section 2.4.4.2). The average Ct value of duplicate amplification reactions from each dilution was plotted against the arbitrary log cDNA concentration. The gradient of the regression line joining the data points was used to calculate amplification efficiency of the assay.

Table 7.2 Amplification efficiencies and R^2 values derived from the standard curves of assays for the candidate genes for validation using gene-specific quantitative RT-PCR assays.

Primer pair	Amplification efficiency (%)	R^2 value
ALPK2	110.81	0.9992
BGN	95.47	0.9997
CLIC6	88.37	0.9955
CPAMD8	91.56	0.9964
CSF1R	105.81	0.9829
CST1	97.56	0.9992
CX3CR1	98.80	0.9939
EDN1	102.63	0.9718
HLA-DRA	84.83	0.9787
NOX4	98.85	0.9907
PPP1R1B	88.12	0.9941
SPP1	98.70	0.9994

7.3.7 Validation of prioritised differentially expressed genes

Comparison of mRNA expression levels of each prioritised gene between FECD-affected and unaffected corneal endothelium was performed in three pairs of RNA samples extracted from pools of DSAEK specimens from diseased (n = 3 pools with 5 specimens per pool) and normal controls (n = 3 pools with 2 specimens per pool). Each sample was analysed in duplicate. A total of 5 reference genes (*ACTB*, *GAPDH*, *HMBS*, *HPRT1*, and *UBC*) were found to be the most stably expressed genes in the human corneal endothelium according to the microarray data. Following evaluation of expression of the 5 genes in cDNA samples derived from FECD-affected and unaffected corneal endothelium by quantitative RT-PCR, the *ACTB* gene was observed to have the least variation in its expression between diseased and normal control corneal endothelim. It was therefore used as the reference gene in this study. Quantitative RT-PCR analysis showed significant overexpression of *ALPK2*, *BGN*, *CLIC6*, *CST1*, *CX3CR1*, *EDN1*, *HLA-DRA*, *NOX4*, and under-expression of

CPAMD8 and *PPP1R1B* between FECD-affected and unaffected corneal endothelium (Figure 7.11 and 7.12, on pages 174-175 and 176-177, respectively). In addition, the analysis showed a trend for overexpression of the *CSF1R* and *SPP1* genes in diseased CE compared to normal corneal endothelium, but the results were not statistically significant (Figure 7.12).

In keeping with the microarray results, quantitative RT-PCR analysis confirmed significant differential expression for *ALPK2*, *BGN*, *CLIC6*, *CPAMD8*, *CST1*, *CX3CR1*, *EDN1*, *HLA-DRA*, *NOX4*, and *PPP1R1B* genes between FECD-affected and unaffected corneal endothelium. Although *CPAMD8* and *PPP1R1B* respectively ranked 38th and 79th in list of the 135 differentially expressed genes identified by microarray analysis, they were still confirmed as significantly under-expressed (fold change = 0.7, $p \leq 0.05$) in the affected corneal endothelium compared to unaffected corneal endothelium. These findings indicate the robustness of the microarray data, and strongly suggest that most of the top 79 genes in the list of the 135 differentially expressed genes would be differentially regulated in corneal endothelium in the disease.

Based on their biological functions, the identified dysregulated genes fall into 7 functional groups. These functions are DNA repair (*ALPK2*), extracellular matrix assembly (*BGN*), protein degradation (*CLIC6* and *CST1*), cellular senescence (*CST1*), production of reactive oxygen species (*EDN1* and *NOX4*), glycogen catabolism (*PPP1R1B*), and immune response (*BGN*, *CPAMD8*, *CX3CR1*, *EDN1*, and *HLA-DRA*).

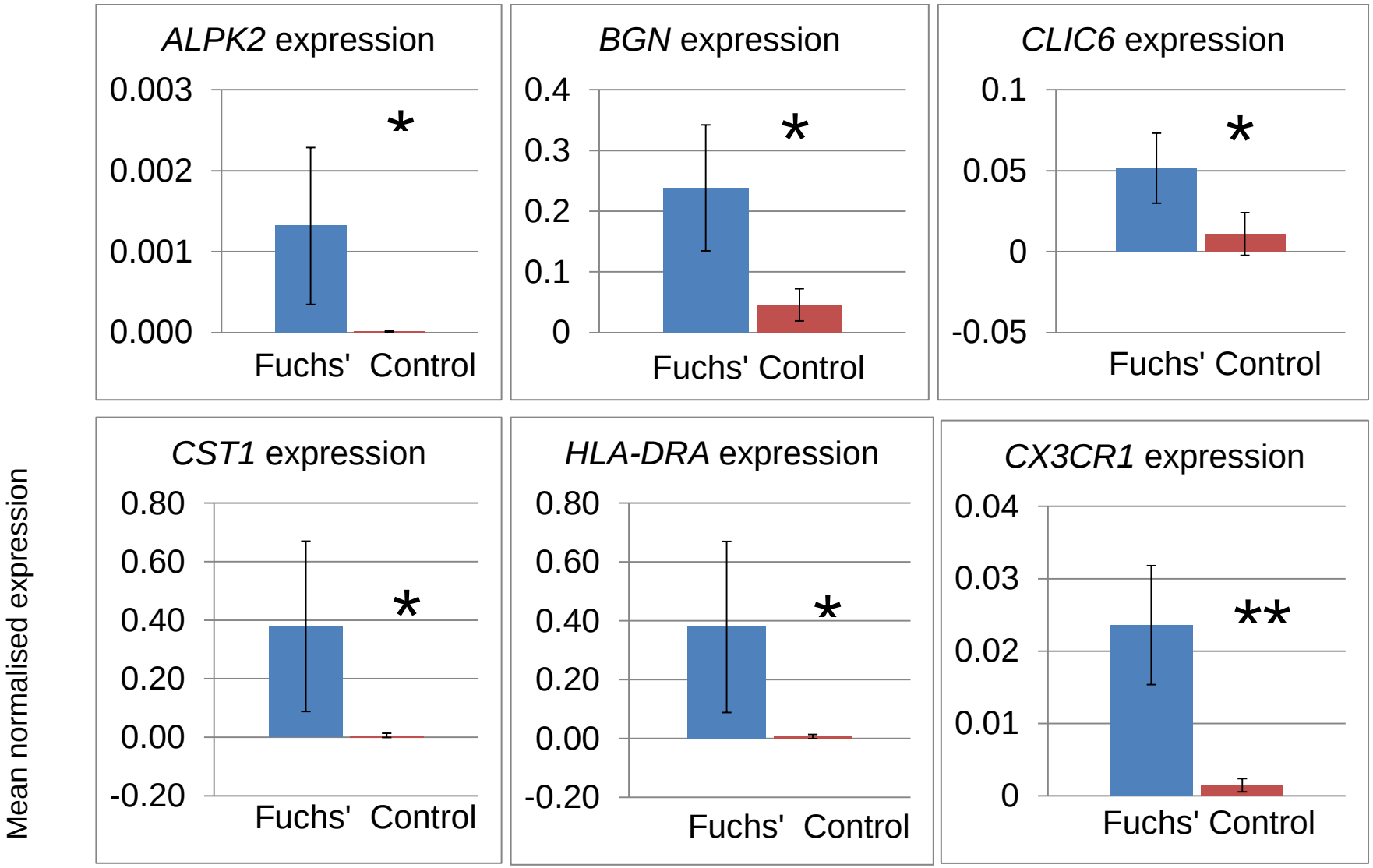

ALPK2 expression
0.003
0.002
0.001
0.000
*
Fuchs'
Control
BGN expression
0.4
0.3
0.2
0.1
0
*
Fuchs'
Control
CLIC6 expression
0.1
0.05
0
-0.05
*
Fuchs'
Control
CST1 expression
0.80
0.60
0.40
0.20
0.00
-0.20
*
Fuchs'
Control
HLA-DRA expression
0.80
0.60
0.40
0.20
0.00
-0.20
*
Fuchs'
Control
CX3CR1 expression
0.04
0.03
0.02
0.01
0
**
Fuchs'
Control
Mean normalised expression

Figure 7.11 Results of validation of differentially expressed genes in human corneal endothelium between FECD-affected and unaffected individuals. Quantification was performed in pooled RNA from cases (n = 3 pools of 5 DSAEK specimens per pool) and controls (n = 3 pools of 2 DSAEK specimens per pool) by quantitative RT-PCR. Data was analysed by Q-gene method. Relative expression levels were normalised to the reference gene *ACTB*, and presented as mean normalised expression (y-axis). Error bars indicate standard deviation of the mean. Student's *t*-test was used to analyse statistical significance;*$P<0.05$, **$P<0.01$, ***$P<0.001$. DSAEK, Descemet's membrane stripping automated endothelial keratoplasty.

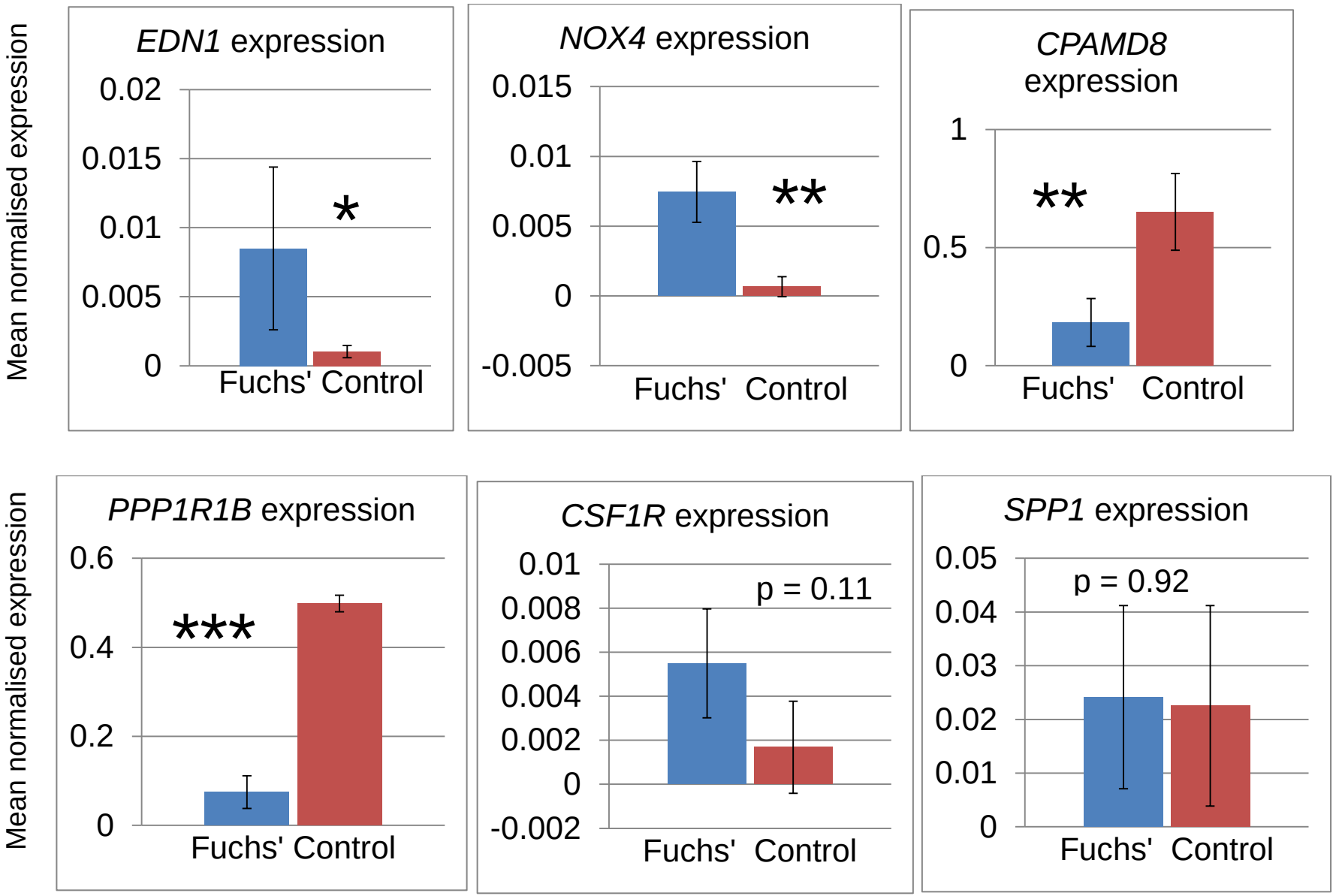

Mean normalised expression
EDN1 expression
0.02
0.015
0.01
0.005
0
*
Fuchs'
Control
NOX4 expression
0.015
0.01
0.005
0
-0.005
**
Fuchs'
Control
CPAMD8 expression
1
0.5
0
**
Fuchs'
Control
Mean normalised expression
PPP1R1B expression
0.6
0.4
0.2
0
***
Fuchs'
Control
CSF1R expression
0.01
0.008
0.006
0.004
0.002
0
-0.002
p = 0.11
Fuchs'
Control
SPP1 expression
0.05
0.04
0.03
0.02
0.01
0
p = 0.92
Fuchs'
Control

Figure 7.12 Results of validation of differentially expressed genes in human corneal endothelium between FECD-affected and unaffected individuals. Quantification was performed in pooled RNA from cases (n = 3 pools of 5 DSAEK specimens per pool) and controls (n = 3 pools of 2 DSAEK specimens per pool) by quantitative RT-PCR. Data was analysed by Q-gene method. Relative expression levels were normalised to the reference gene *ACTB*, and presented as mean normalised expression (y-axis). Error bars indicate standard deviation of the mean. Student's *t*-test was used to analyse statistical significance; **P*<0.05, ***P*<0.01, ****P*<0.001. DSAEK, Descemet's membrane stripping automated endothelial keratoplasty.

7.3.9 Summary of results

Through microarray analysis, 135 corneal endothelial genes were identified as differentially expressed between FECD-affected and unaffected individuals. Quantitative RT-PCR analysis confirmed differential expression of 10 genes in corneal endothelium in the disease.

7.4 DISCUSSION

The aim of this study was to identify differentially expressed genes in corneal endothelium between FECD-affected and unaffected individuals. The microarray analysis results showed a large number of genes (123) to be up-regulated in corneal endothelium in affected individuals when compared with normal individuals. The majority of the identified differentially expressed genes were observed to code for plasma membrane (32%) and extracellular proteins (27%) (Figure 7.7 and 7.8). The other genes encode cytoplasmic (22%), nuclear (10%) and other (9%) proteins (Figure 7.7 and 7.8). These results suggest predominant involvement of genes responsible for transcribing membranous and extracellular matrix proteins in FECD mechanism.

Validation of 12 prioritised genes confirmed 10 of the genes to be significantly differentially expressed in FECD. Once again, seven of the validated genes encode proteins that localise to the plasma membrane (*CLIC6, CX3CR1* and *HLA-DRA*) and extracellular space (*BGN*, *CPAMD8*, *CST1* and *EDN1*). Proteins encoded by the remaining genes are located to the cytoplasm (*NOX4* and *PPP1R1B*) and nucleus (*ALPK2*).

ALPK2 (*Alpha-kinase 2*) is a protein-coding gene with little known about its biological function. The encoded protein belongs to the serine/threonine kinases subgroup of the protein kinase superfamily. ALPK2 contains one alpha-type protein kinase and two Ig-like domains. It uses the former to recognise phosphorylation sites that are surrounded by peptides with an alpha-helical conformation. The Ig-like domains allow the protein to interact with other proteins or ligands [269]. Since proteins containing the Ig-like domain are involved in diverse functions [269], changes in the encoding genes can cause a variety of disorders that affect different body systems.

ALPK2 has been implicated for the first time in FECD. It was shown to promote expression of DNA repair-related genes [269]. DNA damage reportedly occurs in FECD-affected corneal endothelium [140, 141]. Observed overexpression of *ALPK2* in this study may be an indicator of increased DNA-repair activity in FECD.

ALPK2 is located at chromosomal locus 18q21.31 within the FECD-linked locus, *FCD2* (18q21.2 – q21.32) [80]. The *TCF4* gene is also located within this locus, but does not account for the linkage signal [94], thus suggesting involvement of as yet unknown gene/s in the disease. It will be of interest to know whether *ALPK2* is the disease causing gene at this locus.

DNA damage in FECD is partly caused by increased ROS in the affected endothelium [140, 141]. In this study, *EDN1* (*Endothelin 1*) and *NOX4* (*NADPH oxidase 4*), which generate ROS, were found to be overexpressed in FECD-affected corneal endothelium compared to the unaffected corneal endothelium. It is likely that up-regulation of both these genes in FECD contribute to increased production of

ROS that damages the corneal endothelial cells/DNA, leading to the pathogenesis of the disease.

The protein encoded by the *EDN1* gene belongs to a family of endothelins, which are cytokines that affect physiology and pathophysiology of immune system [270]. In response to pro-inflammatory cytokines, bacterial toxins, hypoxia, and oxidised low-density lipoprotein (LDL), vascular endothelial cells were shown to up-regulate the expression of *EDN1* [271]. Furthermore, the gene has been shown to contribute to chronic diseases with an inflammatory component such as atherosclerosis, hypertension end-organ injury, and renal failure [271], as well as diabetes/insulin resistance, and allograft rejection [272].

Atherosclerosis, hypertension and diabetes are associated with increased risk for FECD development. Additionally, oxidised LDL was shown to contribute to corneal endothelium damage in FECD, especially in the absence or insufficient presence of antioxidants. In this project, APOE (Apolipoprotein E) antioxidant was identified to have lower relative abundance in FECD-affected DM, which can result in compromised DM integrity. The gene was also found to be under-expressed in the corneal endothelium in FECD compared to normal. It is possible that observed over-expression of *EDN1* and under-expression of the *APOE* in the disease leads to antioxidant-oxidant imbalance in the affected endothelium, resulting in an increased oxidative stress. This hypothesis is consistent with the generally proposed mechanism of oxidative stress in FECD [69].

NOX4 encodes the NOX4 protein, which is a member of the NOX family of enzymes that functions as the catalytic subunit of the NADPH oxidase complex. The protein is located in non-phagocytic cells where it acts as an oxygen sensor and catalyses the

reduction of molecular oxygen to various ROS. NOX4 has been implicated in numerous biological functions including signal transduction. It regulates signalling cascades probably through inhibition of phosphatases. In addition, it has been suggested to play a role in apoptosis and lipopolysaccharide-mediated activation of NFkB. Both the apoptosis and NFkB pathways are implicated in the pathogenesis of FECD [134]. In light of these, over-expression of *NOX4* may contribute to FECD by increased generation of ROS, leading to endothelial damage over time, and subsequent endothelial cell apoptosis.

PPP1R1B (*protein phosphatase 1, regulatory (inhibitor) subunit 1B*) is an inhibitor of protein phosphatase 1. Its main function is to prevent glycogen synthesis through reduction of protein phosphatase 1 activity. Observed under-expression of *PPP1R1B* in my study could result in increased activation of protein phosphatase 1, which subsequently favours an increased glycogen synthesis. This may result in lower levels of glucose in the corneal endothelial cells, leading to glucose deprivation to the cells. In support of this proposal, a previous study of serial analysis of gene expression between FECD-affected and unaffected corneal endothelium revealed under-expression of genes involved in energy production in the affected endothelium, thus suggesting lower energy levels in the disease. Prolonged energy deprivation could lead to progressive death of corneal endothelial cells in FECD.

Similar to the *PPP1R1B* gene, *CPAMD8* (*C3 and PZP-like, alpha-2-macroglobulin domain containing 8*) is under-expressed in the disease. *CPAMD8* belongs to the complement component-3/alpha-2-macroglobulin (A2M) family of proteins involved in innate immunity and damage control [273]. The expression of *CPAMD8* was shown to be markedly up-regulated in U251 glioma and RD (rhabdomyosarcoma)

cells after treatment with interleukin-1β and interleukin-6, respectively [273]. It was then concluded that expression of *CPAMD8* may be enhanced in certain situations of immune challenge, but in a localised manner.

CPAMD8 expression has been observed in multiple normal human tissues, including the kidney, brain, heart, liver and small intestine [273]. However, the present study is the first to report expression of the gene in human corneal endothelium. The expression of *CPAMD8* in several tissues, including those that are outside the area of adaptive immunity and its high conservation in several species [273] suggests that it has very important biological function in the body. Therefore, its dysregulation in a specific tissue of interest can cause a disease related to that tissue.

CPAMD8 has been identified as a susceptibility locus for Crohn's disease (CD) in the Ashkenazi Jewish [274]. CD is a complex, inflammatory bowel disorder resulting from an inappropriate inflammatory response to intestinal microbiota in genetically susceptible individuals [275, 276]. In FECD, increased complement activation has been reported [149], suggesting an involvement of immune response in its pathogenesis. Thus, it could be hypothesised that observed dysregulation of *CPAMD8* in the present study may indicate the role of an immune response in FECD. It is likely that down-regulation of the gene in corneal endothelium of FECD patients predispose them to weaker immune response towards damaging immune stimulants, thus exposing corneal endothelial cells to injury and subsequently apoptosis. Consistent with this hypothesis, immune-regulatory genes *CX3CR1*, *EDN1* and *HLA-DRA* are overexpressed in FECD.

The *CX3CR1* (*Chemokine (C-X3-C motif) receptor 1*) gene transcribes CX3CR1 receptor for the CX3CL1 (C-X3-C motif ligand 1) chemokine. This receptor

mediates both adhesive and migratory functions of the CX3CL1. It is also involved in activation of various signalling pathways, such as mitogen-activated protein kinase (MAPK), phospholipase C (PLC), and phosphatidyl inositol-3 kinase (PI3K), leading to varied functional outcomes including polarization, adhesion and chemotaxis.

MAPK, PLC, and PI3K pathways have not been implicated in FECD pathogenesis. However, cell-cell adhesion, and polarization of CECs are compromised in the disease. Similarly, there is disruption in the adhesion between corneal endothelial cells and extracellular matrix proteins of the Descemet's membrane in FECD [132], resulting in disorganisation of structural integrity of the membrane. It is more likely that observed dysregulation of the *CX3CR1* gene in the disease-affected corneal endothelium contributes to abnormal adhesion and polarisation of corneal endothelial cells in FECD.

HLA-DRA (*Major histocompatibility complex, class II, DR alpha*) encodes a single-pass type I membrane protein, which is expressed in antigen presenting cells. HLA-DRA plays a central role in the immune system by presenting extracellular or membranous peptides through exogenous antigen presentation pathway for degradation in the lysosomes. The exogenous antigens must compete with those derived from endogenous components for degradation. Autophagy, which is a source of endogenous peptides, and autophagosomes constitutively fuse with MHC class II loading compartments to facilitate protein degradation.

Altered autophagy pathway has been observed in a knock-in mouse model ($Col^{Q455K/Q455K}$) of early-onset disease [153], and suggested to be a mechanism underlying FECD. In addition, activated unfolded protein response has been demonstrated to occur in the disease [126]. Consistently, observed up-regulation of

HLA-DRA in corneal endothelium in FECD in the present study implies an increased activity for protein degradation [153], thus supporting the role of activated unfolded protein response in the pathogenesis of the disease.

BGN (*biglycan*) encodes BGN, a member of a small leucine-rich repeat proteoglycan (SLRP) family found in a variety of extracellular matrix tissues including bone, and endothelium [277, 278]. The functions of biglycan appear to be dependent on a particular microenvironment and the organ in question [277, 279]. When the protein is proteolytically released from the ECM, it acts as a danger signal signifying tissue stress or injury [131, 279]. Additionally, BGN plays a role in immune response by stimulating multifunctional pro-inflammatory signalling pathways linking the innate to the adaptive immune response [280-283]. These biological functions, and observed over-expression *BGN* in FECD suggest that stress and immune response pathways may be involved in the disease.

BGN expression has been shown to be altered by growth factors and certain pathologic conditions [284-286]. Increased levels of *BGN* were previously reported in skeletal muscle of Duchenne muscular dystrophy (DMD) patients [283]. DMD is a lethal X-linked recessive disorder caused by mutations in the dystrophin gene, resulting in lack of assembly of the dystrophin-glycoprotein complex [287]. It is characterised by repeated cycles of muscle damage and regeneration, which eventuate into depletion of regenerative myogenic cells and loss of regenerative ability [287]. FECD shares similar disease processes as DMD, as in dysregulation of extracellular matrix proteins in the endothelial-secreted Descemet's membrane, and gradual degeneration of the endothelial layer [35, 131].

BGN was reported to be over-expressed in Descemet's membrane in FECD [131], consistent with the elevation of *BGN* mRNA observed in this study. Oxidative stress is a well-established pathway involved in the pathogenesis of FECD[69]. Therefore, correlation in the up-regulation of *BGN* in the diseased endothelium and DM is consistent with the presence of oxidative stress or tissue injury in FECD patients.

CST1 (*Cystatin 1*) transcribes a cystatin SN secretory peptide, which belongs to the type 2 cystatin superfamily; that include *CST2*, *CST3*, *CST4*, *CSTP1*, and *CSTP2* [288-291]. The cystatin proteins inhibit the proteolytic activity of cysteine proteases [292, 293], which are widely expressed in tissues, and have numerous functions; including inflammatory tissue destruction and remodelling, modulation of immune response, and induction of migration of monocytes and cancer cells [292-295].

Little is known about the specific biological role of *CST1*. Its expression is reported in submandibular gland, gall bladder, and uterus [289]. In the present study, I have shown for the first time expression of *CST1* in the corneal endothelium. In vitro studies suggest that the gene might have potent inhibitory activity towards lysosomal cysteine proteases [296-300]. Keppler *et al* [288] observed accumulation of CST1 protein within the endosomal-lysosomal comparts of cells from *ras*-mediated senescent human fibroblasts. Its expression was also shown to be highly associated with cellular senescence [288]. These findings have led to the suggestion that induction and accumulation of CST1 in senescent cells might impact on lysosomal proteolysis [301-303].

Impaired lysosomal protein degradation is considered as a key feature of the senescent phenotype [304], and may contribute to various age-related diseases like FECD [303]. Consistently, altered autophagy, activated unfolded protein response,

and premature senescence processes have been demonstrated in corneal endothelium in FECD [70]. Therefore, it can be hypothesised that observed upregulation of *CST1* in the disease in this study may indicate a dysregulation in the lysosomal protein degradation pathway in the corneal endothelium of FECD patients.

CLIC6 (chloride intracellular channel 6) encodes the CLIC6 protein. CLIC6 is a member of the chloride intracellular channel (CLIC) family of proteins [305], and is primarily expressed in intracellular membranes in late endosomes, lysosomes, and osteoclast ruffled membrane [306]. Disruption of *CLIC6* expression results in neurodegeneration, characterised by pathological features of lysosomal storage disease as shown in knock-out mouse model [306]. In this model, neurons displayed intracellular, electron-dense deposits that stained for lysosomal marker proteins and the subunit c of ATP-synthase, a protein that is typically accumulated in a subset of human lysosomal storage disease called neuronal ceroid lipofuscinosis [306].

Accumulation of abnormal protein aggregates is reported in rough endoplasmic reticulum in FECD-affected endothelium [107]. Whether this accumulation is related to dysfunctional endosomal-lysosomal pathway, due to dysregulation of *CLIC6*, remains to be determined. The fact that *CLIC6, CST1* and *HLA-DRA* are involved in endosomal-lysosomal pathway and over-expressed in this study implicates the role of this pathway in FECD mechanism.

In summary, this study has identified many processes involving dysregulated genes in corneal endothelium in FECD. These processes may collectively be contributing to the pathogenesis of the disease. Five out of the ten DE genes are involved in immune response, making it the leading dysregulated pathway in FECD in this study. My study is the first to clearly implicate immune response in FECD. It is likely that

dysregulation of the immune system and inflammatory processes in FECD contribute to corneal endothelium injury, and ultimately apoptosis [134, 307]. Endosomal-lysosomal pathway is another major process highlighted in the present study. This pathway is important for protein degradation and turnover in living cells. Its impairment can lead to accumulation of cytotoxic protein aggregates that are harmful to the corneal endothelium. This hypothesis is consistent with previous findings of activated unfolded protein response, altered autophagy, and abnormal protein build-up in corneal endothelium in FECD [126]. Premature senescence is another key process that has been implicated in FECD in the present study, consistent with a previous independent study [153]. Involvement of oxidant-antioxidant imbalance, cellular energy deprivation, DNA damage and repair, and abnormal cellular adhesion, polarisation and growth in corneal endothelium in FECD have been reported in past studies [35, 69, 70, 140]. My study has identified novel dysregulated genes with functional roles in these disease-implicated pathways.

One of the main strengths of this study was the use of Illumina microarrays for identifying differentially expressed genes. This is because they provide a snapshot of global gene expression changes without making assumptions about the relevance of each gene to the condition being investigated [262, 263]. The main limitation of gene expression microarrays is that analytical artefacts in datasets may overshadow true biological variations involving small fold changes in gene expression between conditions being tested [308]. This limitation was likely minimised by normalisation of the dataset. Additional limitation is the use of surgical tissues, which come from end-stage disease. As a consequence, identified variations in some of the dysregulated gene may be related to the pathophysiological effects of FECD than causes. Furthermore, while some of the gene expression changes may be specific to

FECD, it is possible that some are due to a general cellular response to a pathological insult. Comparison of dysregulated genes in other diseases involving the corneal endothelium with those in FECD in the future can indicate the disease specific dysregulated genes. Nevertheless, the identified differentially expressed genes have still contributed in revealing the processes involved in the pathophysiology of FECD.

In conclusion, this study has identified several differentially expressed genes between FECD-affected and unaffected corneal endothelium. Of the 12 DE genes that underwent validation, 10 were confirmed to be significantly dysregulated. The dysregulated genes are involved in antioxidant-oxidant imbalance, DNA damage/repair, dysregulated cell cycle, and abnormal cellular adhesion, polarisation and growth pathways known to be involved in FECD, and in endosomal-lysosomal, and immune response novel pathways.

CHAPTER 8
FINAL DISCUSSION

FECD is a disease of major clinical importance across the world, as it causes progressive vision loss, and ultimately blindness, if left untreated [10]. Corneal grafting is the effective treatment for FECD. Currently, the disease is one of the leading indications for corneal transplantation [51]. In Australia, ~6% of all corneal grafts are performed in FECD patients annually [47]. Although many corneal grafts perform well for many years, there is a high failure rate as well as significant risk of surgical complications [47]. In addition, corneal grafting depends upon the availability of cadaveric donor tissue.

The mechanism of FECD is poorly understood. Previous studies have shown it as a complex disease with strong genetic heterogeneity, much of which remains undetermined [39, 76, 78, 79, 83]. Additionally, molecular studies suggest involvement of oxidative stress, apoptosis, protein misfolding, activated unfolded protein response, altered autophagy, and accelerated senescence in the disease pathophysiology [69, 79, 126, 153, 307, 309]. However, the reported pathways do not provide a complete understanding of the disease mechanism.

This project aimed to advance the understanding of molecular basis of FECD. The specific aims were to identify genetic causes of FECD in Caucasian Australian cases, determine relative abundance of proteins between disease-affected and unaffected Descemet's membrane, and identify dysregulated genes between disease-affected and unaffected corneal endothelium.

Mutation screening of the *LOXHD1*, *SLC4A11* and *ZEB1* genes in 128 cases by Ion Torrent Next-Generation Sequencing identified three novel mutations in *LOXHD1* and *SLC4A11* as the possible causes of FECD in three cases. This finding implies that the screened genes contribute to the disease in a very small proportion of Caucasian Australian cases. My finding is consistent with previous findings in other non-Australian populations [78, 83]. Thus, I concluded that much of the genetic determinants for FECD, particularly in the Australian cases, remain to be identified.

To identify additional novel genes associated with FECD, GWAS was employed. As it is more cost effective, DNA pooling technique was used for the discovery of novel loci associated with the disease. The successful utility of this method was demonstrated in previous studies [172, 188]. Overall, the pooling strategy, combined with follow-up individual genotyping, was succesful in identifying *TCF4* as the major susceptibility gene for development of FECD in Caucasian Australians.

I observed that individuals with risk alleles at the associated SNPs in *TCF4* are at least three times more likely to develop FECD than the non-carriers. These observations are consistent with previous findings about association of *TCF4* with FECD [92, 94]. However, my study is the first to report *TCF4* as the major genetic factor associated with the risk of developing the disease in Australia.

TCF4 maps inside the *FCD2* locus that has been linked with late-onset FECD in multiple families [80]. However, variations in the *TCF4* gene do not appear to account for the linkage signal [94]. In addition, genetic association studies have recently revealed significant association between FECD and an intronic TGC repeat polymorphism in *TCF4* in Caucasian Americans [102, 105], as well as in Chinese population [106, 310]. The TGC repeat is located in the same intron, 3, as the

rs613872 SNP, but with stronger association than the latter [102]. Haplotype association tests for linkage disequilibrium between the rs613872 and TGC repeat polymorphisms have however shown that the two are independently associated with the disease [105]. These findings suggest that each of the polymorphisms may separately contribute to the pathogenesis of FECD. In fact, the most recent study has shown for the first time that the expanded TGC repeat polymorphism contributes to FECD through the mechanism of RNA toxicity and mis-splicing [108].

In this project, I identified significant association between TGC repeat expansion in *TCF4* and FECD in Caucasian Australian cases. Association of the expanded TGC repeat allele is stronger than that of the rs613872 SNP, and independent of age and gender, the two main risk factors for the disease. The two polymorphisms showed partial independent association with FECD, suggesting some kind of interaction. These findings are the first to be reported in Australian cases of FECD, and replicate previous findings from other groups [102, 105].

Following the FECD-associated SNPs identified through GWAS in a Caucasian American case-control cohort, I have identified *LAMC1*, *ATP1B1* and *KANK4* as novel genetic factors involved in the pathogenesis of the disease in Caucasian Australians. Variations in *LAMC1* were identified to be associated with FECD in the GWAS using pooled DNA, but the association was below the threshold for genome-wide significance. Both *ATP1B1* and *KANK4* were not revealed in the GWAS using pooled DNA.

The main difference between the GWAS performed in the present study and by our collaborators is the number of FECD cases used in each study. In the Australian study, 94 cases were studied compared to 2500 cases studied in the American study.

As a result, my study was under-powered to identify association of lower effect sizes with the disease.

The effects of the risk alleles at the FECD-associated SNPs rs1200114 and rs1200115 near *ATP1B1* increase the chance of developing the disease by 1.4 times in carriers compared with non-carriers. Similarly, carriers of the risk alleles at the rs2296212 or rs20560, respectively, in the *LAMC1* gene are 1.4 times more likley to develop FECD than non-carriers. In contrast, disease-associated SNPs rs6424883, rs1413386 or rs3768617 in *LAMC1* harbour protective alleles against susceptibility to the pathogenesis of FECD and confer 40% protection in carriers than non-carriers. The *KANK4* gene SNPs rs79742895 and rs12082238 associated with FECD, with the risk of or protection against susceptibility to the disease for carriers compared with non-carriers. Whereas the risk allele at the SNP rs79742895 increases the likelihood of developing FECD by 2.4 times in carriers compared with non-carriers, the allele at the rs12082238 confers protection against the disease development by about 27% in carriers compared to non-carriers. Therefore, the effect of the *LAMC1*, *ATP1B1*, and *KANK4* in contributing to the risk of developing FECD is less than that of the *TCF4* gene. Larger population based or case-control studies would be required to accurately calculate the combined contribution of these SNPs to FECD risk in the general population.

To further analyse the relationship between FECD and the *ATP1B1* and *LAMC1* genes, I investigated expression of both genes at the mRNA and protein levels in FECD-affected corneal endothelium and cornea, respectively, compared to those unaffected. Expression analysis of *LAMC1* showed no difference in gene expression between diseased and normal corneal endothelium, or in the distribution of protein

expression in the corneal endothelium, stroma and epithelium between diseased and normal corneas. Expression analysis of *ATP1B1* revealed down-regulation of the gene in diseased corneal endothelium compared to normal endothelium. In addition, ATP1B1 was present and evenly distributed throughout the corneal endothelium between diseased and normal corneas.

Down-regulation of *ATP1B1* and no variation in the regulation of *LAMC1* in FECD are consistent with the most recent findings from other researchers, who also reported downregulation of *ATP1B1* and no difference in expression of *LAMC1* in FECD-affected corneal endothelium compared to the unaffected [204, 311]. However, my study is the first to simultaneously reveal both genetic association and dysregulation of the *ATP1B1* gene in the disease. Furthermore, it is the first to describe distribution of ATP1B1 protein in the affected and unaffected corneas; and of LAMC1 in corneal stroma, Bowman's layer and epithelium of diseased cornea compared to normal cornea.

Observed down-regulation of the *ATP1B1* transcript may lead to dysregulation of electrochemical gradients homeostasis, as observed in the disease [35, 138], and subsequent reduction in sodium-potassium pump function. The reduction in the pump activity could contribute to accumulation of fluid in the corneal stroma in the disease [311].

Comparison of the relative abundance of proteins between FECD-affected and unaffected Descemet's membrane by label-free mass spectrometry revealed reduced relative abundance of APOE in the disease. To the best of our knowledge, this study is the first to report down-regulation of the protein in disease-affected Descemet's membrane. Furthermore, the protein had not been implicated in the disease before.

To confirm these results, expression of APOE was further investigated in corneal sections from FECD-affected and unaffected corneas by immunohistochemistry.

Immunohistochemistry validated APOE as a component of the epithelium, stroma, Descemet's membrane and endothelium in FECD and normal corneas. It revealed strong but differentially distributed labelling for the protein in Descemet's membrane between diseased and normal corneas. My work revealed that APOE distributed solely in the anterior banded layer of the Descemet's membrane in the affected cornea compared to wider distribution of the protein throughout the Descemet's membrane in unaffected cornea. This finding correlates with the finding of down-regulation of the protein in the disease.

Since protein components of the Descemet's membrane are produced by the corneal endothelium [15, 312, 313], I investigated the latter for expression of *APOE* mRNA between diseased and normal corneal endothelium. Consistent with the proteomic and immunohistochemical findings, relative gene expression analysis showed a correlating down-regulation of *APOE* in FECD corneal endothelium. This finding suggests that down-regulation of *APOE* transcript in corneal endothelium in the disease may underlie observed lower abundance of the protein in FECD-affected Descemet's membrane. Previous studies revealed a similar correlation between over-expression of the *TGFBI* and *CLU* genes in the corneal endothelium with the TGFBI and CLU proteins in the Descemet's membrane in FECD [114], suggesting involvement of these proteins in the pathophysiology of the disease.

APOE is a multifunctional, low-density lipoprotein receptor ligand that primarily serves as a lipid transporter [69]. It also serves as an antioxidant, inflammation regulator in the innate immune system, and a ligand for extracellular matrix

protein binding in the Descemet's membrane [254, 255, 257, 258]. These functions implicate APOE in homeostatic control and maintenance of Descemet's membrane proteins, and oxidant-antioxidant pathway known to be involved in the pathogenesis of FECD [69], as well as in the novel immune response pathway, implicated in the disease.

Investigation of dysregulated genes between FECD-affected and unaffected corneal endothelium by microarray analysis revealed 135 differentially expressed genes; the majority of which were upregulated in the disease. Validation of 12 prioritised genes by quantitative RT-PCR confirmed significant up-regulation of *ALPK2*, *BGN*, *CLIC6*, *CST1*, *CX3CR1*, *EDN1*, *HLA-DRA*, and *NOX4*, and significant down-regulation of *CPAMD8* and *PPP1R1B* in FECD. Biologically, the validated genes perform functions involving DNA repair (*ALPK2*), extracellular matrix assembly (*BGN*), protein degradation (*CLIC6* and *CST1*), cellular senescence (*CST1*), production of reactive oxygen species (*EDN1* and *NOX4*), glycogen catabolism (*PPP1R1B*), and immune response (*BGN*, *CX3CR1*, *EDN1*, *HLA-DRA* and *CPAMD8*). Whilst corneal endothelial damage and DNA repair, dysregulated electrochemical gradient homeostasis, premature senescence, and oxidant-antioxidant imbalance have been reported in FECD [69, 140, 141, 153, 311], involvement of an immune response in the disease is a novel finding.

The studies performed in this project were designed with a view that their findings could collectively predict common mechanism/s or novel pathways, which underlie the pathogenesis of FECD. Using IPA, I therefore explored molecular relationships and interactions among the 17 genes/proteins found to be involved in the disease in this project, and 25 genes/proteins (Appendix, Table 5, page 217) reported in

previous independent studies [93, 114, 130-132, 245, 311]. The analysis was performed using the core analysis option, and by considering only direct relationships among the genes.

IPA revealed involvement of the analysed genes in two functional networks (Table 8.1). The first network contained 25 of the analysed genes, and involved functions related to development and cellular morphology. The second network had 13 genes, and was associated with functions related to cancer and gastrointestinal or inflammatory disease.

Table 8.1 Ingenuity pathway analysis-derived functional networks involving 42 novel and reported genes/proteins involve in FECD. The novel genes/proteins were identified in the present study.

Network ranking	Associated top functions	Genes/proteins involved	Score
1	Development, Tissue morphology	*AGRN, APOD, APOE, BGN, CLU, FN1, COL16A1, COL1A1, COL3A1, COL4A1, CSF1R, DICER1, EDN1, FBN1, ITGA4, JUN, KERA, LAMC1, MT-CO2, PRDX2, PRDX3, PRDX5, TCF4, TGFBI, ZEB1*	67
2	Cancer Gastrointestinal disease Inflammatory disease	*ALPK2, ATP1B1, CLIC6, CPAMD8, CST1, CX3CR1, DRAM1, HLA-DRA, KANK4, LOXHD1, NOX4, PPP1R1B, SLC4A11*	29

The likelihood that the assembly of the focused genes in networks 1 and 2 could be due to random chance alone would approximately be 10^{-67} (score of 67) and 10^{-29} (score of 29), respectively (Table 8.1). Therefore, both networks have highly significant scores, thus greater confidence in the associated functions.

Involvement of the analysed genes in the functions associated with the two networks was consistent with their involvement in the top molecular and cellular functions derived from this IPA (Table 8.2, page 198). These functions related to free radical scavenging, molecular transport/biochemistry, and cellular signalling or movement. Additionally, these functions were supported by the top canonical pathways represented by some of the analysed genes (Table 8.3, page 199).

Table 8.2 Ingenuity pathway analysis-derived top molecular and cellular functions involving 42 genes/proteins implicated in FECD.

Function ranking	Associated top functions	p-value range	Genes involved
1	Free Radical Scavenging	1.83×10^{-3} - 2.74×10^{-9}	*APOE, CSF1R, EDN1, FN1, ITGA4, JUN, NOX4, PRDX2, PRDX3, PRDX5*
2	Molecular Transport	1.92×10^{-3} - 2.74×10^{-9}	*AGRN, APOD, APOE,CLU, CSF1R, CX3CR1, DICER1, EDN1, FN1, ITGA4, JUN, NOX4, PPP1R1B, PRDX2, PRDX3, PRDX5*
3	Cell Signalling	3.64×10^{-4} - 8.08×10^{-9}	*AGRN, APOE, ATP1B1, BGN, CLU, COL1A1, COL3A1, COL4A1, CSF1R, CX3CR1, DICER1 EDN1, FBN1, FN1, HLA-DRA, ITGA4, JUN, LAMC1, NOX4, PPP1R1B, PRDX2,TGFBI,ZEB1*
4	Small Molecule Biochemistry	1.92×10^{-3} - 8.08×10^{-9}	*AGRN, APOD, APOE, CLU, DICER1, EDN1, FN1, ITGA4, JUN, NOX4, PPP1R1B, PRDX2, PRDX3, PRDX5*
5	Cellular Movement	1.79×10^{-3} - 2.77×10^{-8}	*APOE, BGN, CLU, COL1A1, COL3A1, COL4A1, CSF1R, CX3CR1, ,DICER1, DRAM1, EDN1, FBN1, FN1, ITGA4, JUN, LAMC1, NOX4, PRDX2,TGFB1, ZEB1*

Table 8.3 Top five canonical pathways represented by some of the 42 genes/proteins involved in FECD. The novel genes/proteins were identified in the present study. The overlap column indicates the percentage of number of analysed genes (numerator in the bracket) in the given canonical pathway divided by the total genes (denominator in the bracket) found in the Ingenuity Knowledge base to represent that pathway.

Name	P-value	Overlap	Involved genes
Atherosclerosis Signalling	1.26×10^{-7}	4.8% (6/124)	*APOD, APOE, CLU, COL1A1, COL3A1,ITGA4*
Hepatic Fibrosis/Hepatic Stellate Cell Activation	1.98×10^{-6}	3.0% (6/198)	*COL16A1, COL1A1, COL3A1, COL4A1, EDN1, FN1*
Agrin Interactions at Neuromuscular Junction	9.15×10^{-6}	5.8% (4/69)	*AGRN, ITGA4, JUN, LAMC1*
Phagosome Maturation	8.11×10^{-5}	3.3% (4/120)	*HLA-DRA, NOX4, PRDX2, PRDX5*
IL-12 Signalling and Production in Macrophages	1.28×10^{-4}	3.0% (4/135)	*APOD, APOE, CLU, JUN*

Specifically, the revealed canonical pathways included processes related to inflammatory and immune responses [314-316], aberrant accumulation of extracellular matrix proteins [317, 318], localisation and maintenance of cell-cell or cell-extracellular matrix adhesion, cellular morphology, assembly and organisation [319-322], and removal of apoptotic cells. The last process features signalling pathways involving Rho-GTPases such as Ras, and their effector molecules – the c-JUN family members [320, 321]. With the exception of inflammation/immune response, the processes associated with the canonical pathways correlate with involvement of altered expression, excessive secretion and accumulation of extracellular matrix proteins in Descemet's membrane [35, 131], abnormal morphology and apoptosis of corneal endothelial cells [13, 15], increased oxidative stress [69, 127, 309, 323], and dysregulation of *JUN* and functional activities related to Rho-GTPases in FECD pathophysiology [324-326].

To further investigate the inter-relationships of the analysed genes in the networks, known interactions among the genes were retrieved from the Ingenuity Knowledge Base and illustrated in Figures 8.1 (page 201) and 8.2 (page 202), respectively. The genes studied in the present project are presented as coloured, with up- and down-regulated genes in FECD respectively indicated in red/deep orange and green. The genes depicted in light orange/green were not observed to be differentially regulated in the corneal endothelium in the disease in this project.

As can be seen in Figure 8.1, *TCF4* is a direct or indirect upstream regulator of the majority of genes associated with or altered in FECD. It reportedly promotes expression of its downstream effectors (Figure 8.1).

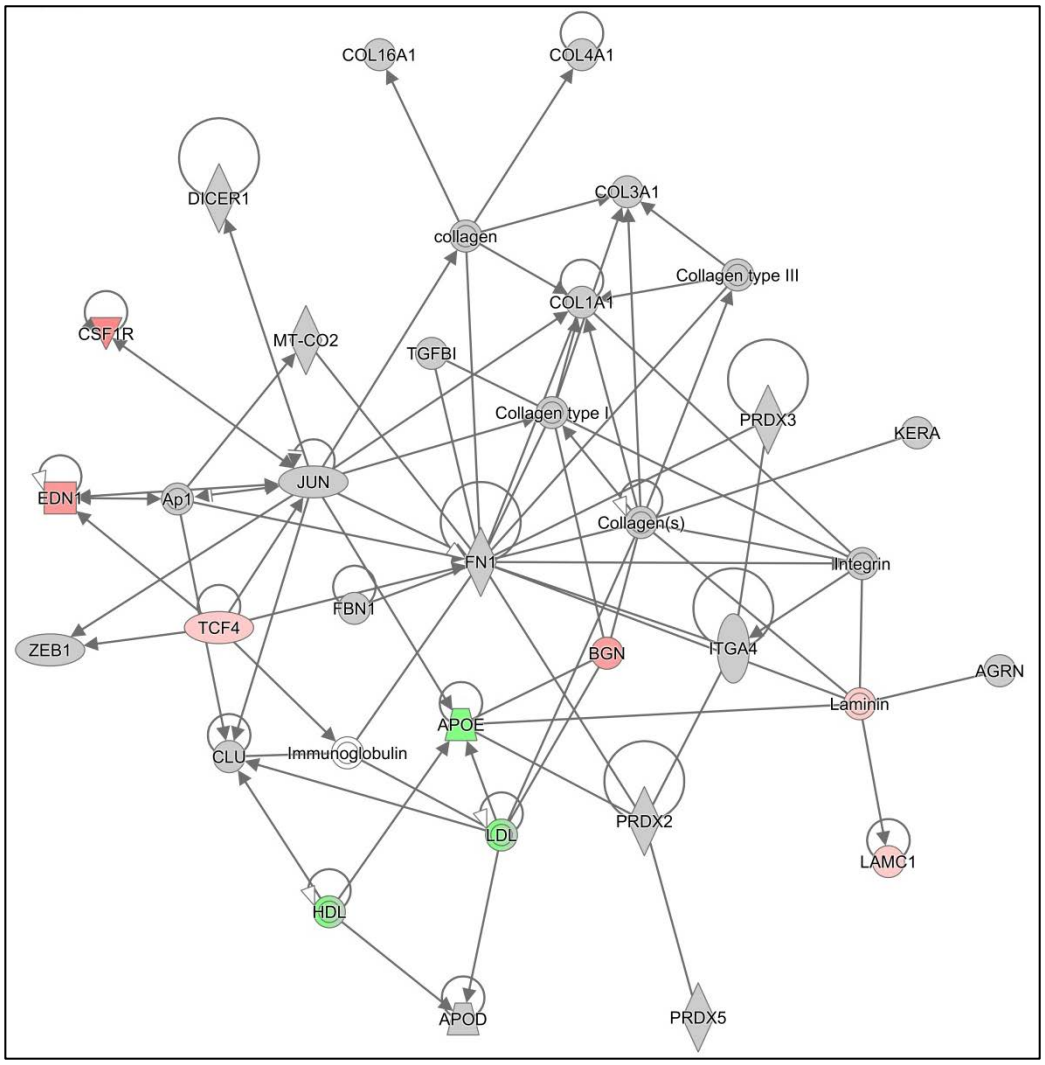


Figure 8.1. Known interactions among genes and/or proteins (n = 25) identified to be involved in FECD in this project and previous independent studies in the disease in Network 1. The analysis was conducted using Ingenuity Pathway Analysis. The genes in colour are from this project. The up- and down-regulated genes in corneal endothelium in FECD are indicated in red and green, respectively. Genes coloured light orange are not differentially regulated. Genes from previous independent studies on FECD are indicated in grey. An arrow indicates activation or regulation of expression of a gene, the direction of arrow indicates the direction of relationship. A line represents protein-to-protein interaction between molecules.

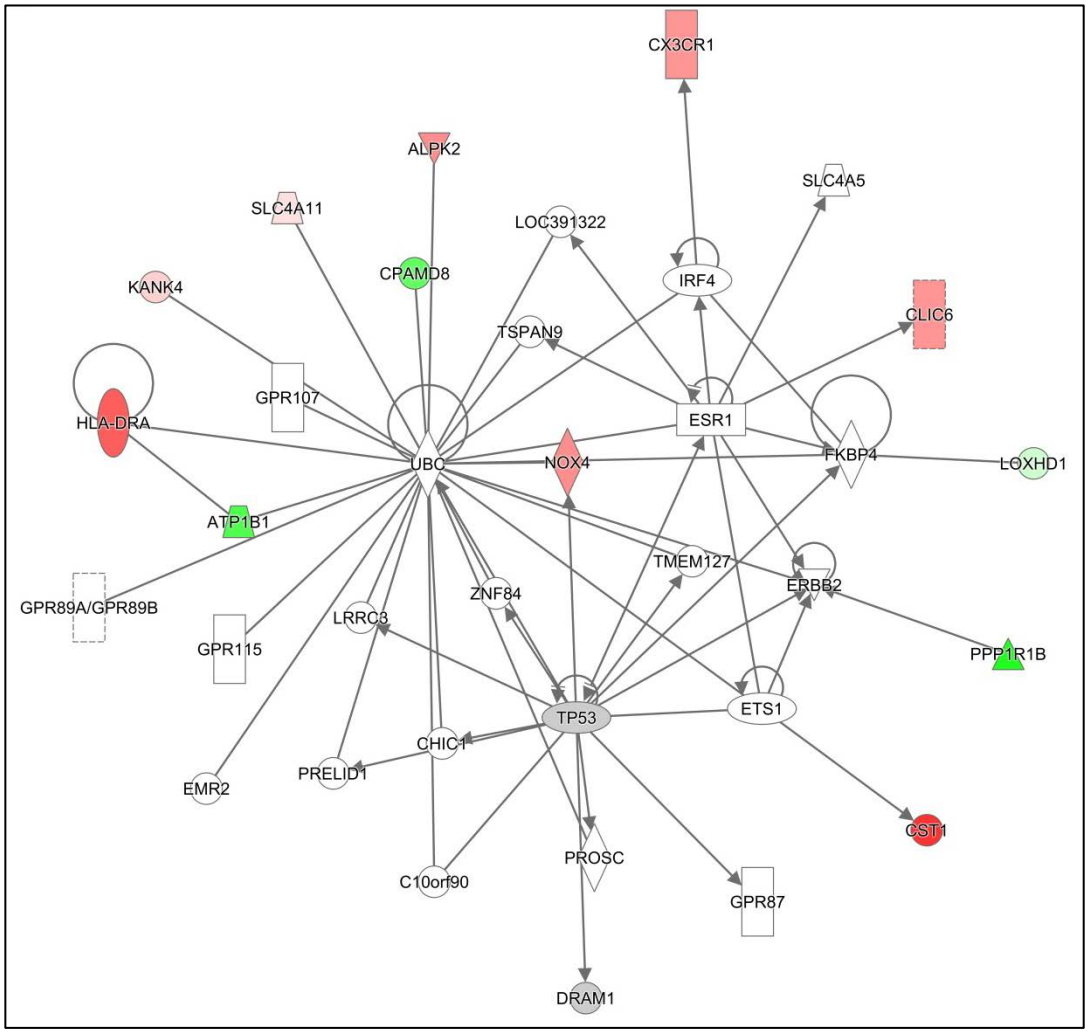


Figure 8.2 Known interactions among genes and/proteins (n = 16) identified to be involved in FECD in this project and previous independent studies on the disease in Network 2. The analysis was conducted using Ingemuity pathway analysis. The genes in colour are from this project. The up- and down-regulated genes in corneal endothelium in FECD are indicated in red and green, respectively. The genes highlighted in grey have been implicated in FECD in previous independent studies. The unhighlighted genes/molecules in the network share similar associated functions related to the network with FECD-implicated genes. An arrow indicates activation or regulation of expression of a gene, the direction of arrow indicates the direction of relationship. A line represents protein-to-protein interaction between molecules.

Through reported *in vitro* assays, the *TCF4* gene has been shown to activate the expression of the downstream genes, indicated by the Network 1, through complexing with Beta-catenin [327-331]. Beta-catenin regulates genes involved in embryogenesis and tissue homeostasis [332]. It can also contribute to initiation and progression of colon cancer [333, 334]. Genes encoding members of the basic helix-loop-helix family of proteins, including TCF4 and ZEB1, represent one of the largest groups of Beta-catenin target genes [332, 333]. These findings are consistent with observed functions involving Network 1.

Further evidence for involvement of the *TCF4* gene in FECD could be inferred from reported up-regulation of *JUN* in the disease, both at the transcript and protein levels, in a previous independent study [154]. JUN reportedly increases expression of its downstream effectors shown in Figure 8.1, except *APOE*, which it has been been demonstrated to repress. Interestingly, down-regulation of *APOE* transcript and protein levels has been observed in FECD in the present project. These observations suggest that an increased activation of *JUN* by *TCF4* might be contributing to down-regulation of the *APOE* gene and up-regulation of *DICER1*, *EDN1*, *ZEB1*, and genes encoding extracellular matrix proteins in FECD.

Figure 8.2 shows the inter-relationships of the analysed genes in Network 2. Results of known interactions among the genes revealed TP53 transcription factor as the common upstream regulator of majority of the genes in the network. It promotes expression of the *NOX4* [335] and *DRAM1* genes [336-339]. In addition, previous independent studies have shown TP53 as an activator of *DICER1* expression [340, 341]. Importantly, overexpression of *TP53* and its downstream effectors, *DRAM1* and *DICER1*, have been reported in FECD-affected corneal endothelium at the

mRNA and protein levels [153, 154, 204], suggesting central role for dysregulated TP53 and *DRAM1*-driven autophagy pathways in the pathophysiology of the disease. Consistently, my project has revealed upregulation of novel genes *NOX4, CLIC6* and *CST1,* which have shown inter-relationship with these pathways in Network 2. It is likely that observed up-regulation of the *NOX4*, *CLIC6*, and *CST1* genes in corneal endothelium of FECD patients in this project is through dysregulated TP53 and/or autophagy pathways.

Additionally, activation of the TP53 reportedly stimulates expression of the FK506 Binding Protein 4 (*FKBP4*) [342, 343], Estrogen receptor 1 (*ESR1*) [344-350], V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog (*ETS1*) [351, 352], V-Erb-B2 Avian Erythroblastic Leukemia Viral Oncogene Homolog (*ERBB2*) [353, 354], and V-Myc Avian Myelocytomatosis Viral Oncogene Homolog (*MYC*) [355-359]. In contrast, TP53 suppresses expression of the *Ubiquitin C* (*UBC)* gene in a murine model [342, 360, 361]. The FKBP4, ESR1, ETS1, ERBB2 and MYC transcriptional regulators play roles in cellular immunoregulation, protein folding and trafficking, regulation of gene expression, including *MT-CO2* or *COX2*, *CDKN2A*, cytokines and chemokines, transformation, and apoptosis. These functions are consistent with observed dysregulation of the *LOXHD1*, *CLIC6*, *CST1*, *PPP1R1B*, and *FBN1* in previous independent studies [78, 131, 132], or in the present project. The UBC protein regulates ubiquitination, which has been associated with protein degradation via endoplasmic reticulum-associated degradation (ERAD) or proteasome pathway, DNA repair, activation of the NF-KB factor, regulation of cell cycle or cellular signalling pathways, and kinase modification. These functions are relevant to the reported roles of *ALPK2, ATP1B1, CPAMD8, HLA-DRA, KANK4, SLC4A11*, *DICER1* and *PRDX5* in normal phsyiological system or in the pathophysiology of

FECD [362-368]. Based on the observed findings in FECD by previous studies and the present project, the pathogenesis of the disease possibly arises through multiple mechanisms triggered largely in part by pathogenic genetic variants in the carriers. These possibly contributing mechanisms are summarised in the schematic shown in Figure 8.3.

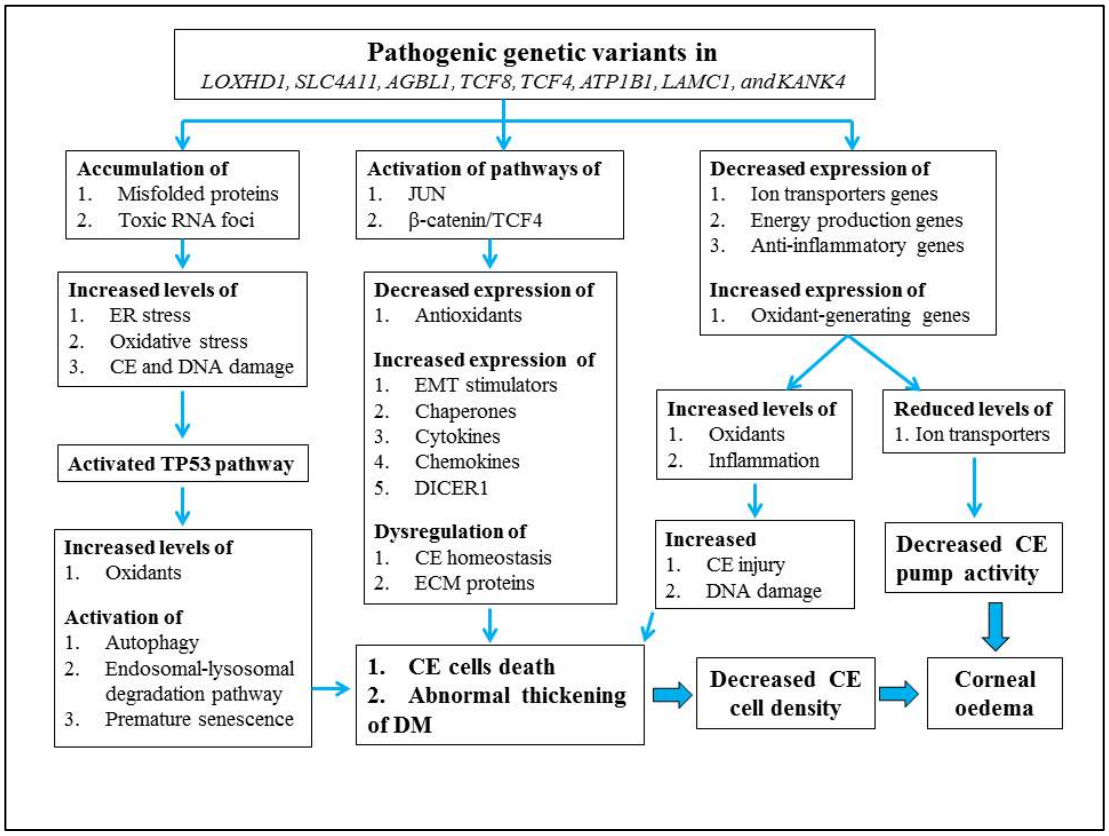


Figure 8.3 Summary of the mechanisms potentially contributing to the pathophysiology of Fuchs' endothelial corneal dystrophy. Arrows point to the direction of the effects of the preceeding events. Thick arrows at the bottom right cornea signify the converging effects of the processes leading to the clinical symptoms of Fuchs' endothelial corneal dystrophy. **Abbreviations:** *AGBL1, ATP/GTP binding protein-like 1*; *ATP1B1, ATPase, Na+/K+ transporting , beta 1 polypeptide*; CE, corneal endothelium; DICER1, Dicer 1, ribonuclease type III; DM, Desecemet's membrane; ER, endoplasmic reticulum; ECM, extracellular matrix; EMT, Endothelial/Epithelial-to-mesenchymal transition; *JUN, Jun proto-oncogene*; *KANK4, KN motif and ankyrin repeat domain 4*; *LAMC1, Laminin, gamma 1*; *LOXHD1, Lipoxygenase homology domains 1*; *SLC4A11, Solute carrier family 4, sodium borate transporter, member 11*; *TCF4, Trancription factor 4*; *TCF8, Transcription factor 8*

Briefly, pathogenic mutations in known FECD-causing or associated genes, and possibly in yet to be identifed genes, would lead to (1) accumulation of toxic mutant proteins and RNA foci, (2) activation of JUN and β-catenin/TCF4 complex pathways, and (3) concurrent increase and decrease in expression of oxidant-generating genes, and of ion transporters, energy production and immune response control genes, respectively, in corneal endothelium of FECD patients.

Increased levels of mutant protein aggregates and toxic RNA foci results in damage of corneal endothelial cells or DNA, and endoplasmic reticulum and oxidative stress. These events trigger activation of the TP53 pathway, leading to further increased levels of oxidants, and activation of autophagy, endosomal-lysosomal degradation, and premature senesence pathways. These pathways facilitate removal of toxic proteins and RNA foci in FECD-affected corneal endothelium, and apoptosis of the injured cells.

Activation of JUN and β-catenin/TCF4 complex pathways likely leads to decreased expression of antioxidants, simultaneous increase in expression of the endothelial-to-mesenchymal transition activators, chaperones, cytokines, chemokines, and DICER1 (a mediator of post-transcriptional gene silencing). Furthermore, stimulation of JUN and β-catenin/TCF4 complex would result in dysregulation of the corneal endothelium homeostasis and expression of extracellular matrix proteins. These processes could contribute to corneal endothelial cell injury and subsequent death, and possibly excessive secretion and accumulation of extracellular matrix proteins in Descemet's membrane in FECD patients.

Similarly, overexpression of oxidant-generating genes, and concurrent underexpression of the genes coding for energy production, and control of immune

response and inflammation would lead to increased levels of oxidants and inflammation, causing progressive corneal endothelial cells injury and DNA damage, and ultimately apoptosis of the harmed cells. Alternatively, decreased expression of genes encoding ion transporters, such as the ATP1B1 protein could result in reduced levels of the ion transporters in corneal endothelium of patients with FECD, leading to reduced sodium pump activity.

Collectively, the reduction in sodium-potassium pump activity and corneal endothelial cell density could in synergy contribute to the clinical symptom of oedema in the corneal stroma in FECD patients. Future studies to better understand pathways underlying the reduction in sodium-potassium activity and/or corneal endothelial cell density in the disease may assist in devising strategies to prevent, slow, or minimise the clinical symptoms associated with FECD.

Based on the findings from this project, future studies may include the following. Functional studies are needed to better understand the molecular mechanism of RNA toxicity and mis-splicing due to the expanded *TCF4* TGC repeat in corneal endothelium in FECD. In light of the robust replication of genetic association of *TCF4* with FECD in multiple ethnicities, large effect sizes of its risk alleles, and the role of the expanded TGC in the gene, the associated variants can be used to predict the risk of developing the disease. This would help in earlier detection of FECD, which might lead to early intervention. Associations of the novel *ATP1B1*, *KANK4* and *LAMC1* genes with the risk of developing the disease in Caucasian Australians and Americans require further replication in other Caucasian and non-Caucasian populations. Given consistent findings of the involvement of *ATP1B1*, at the genetic and molecular levels, in FECD, future studies are needed to detect pathogenic

variants by sequencing the whole *ATP1B1* gene. Three novel mutations identified in the *LOXHD1* and *SLC4A11* genes in Caucasian Australian cases of FECD warrant functional analysis to ascertain their impact in the pathogenesis of the disease in the future. Identification of these mutations also suggests further mutation screening in the two genes in new Australian cases of FECD. This might be less useful since few Australian cases of the disease are observed to carry FECD-causing mutations in *LOXHD1* and *SLC4A11*. In this regard, further functional studies are required for differentially expressed genes identified in this project. This is to further understand the roles of autophagy, endosomal-lysosomal degradation, cellular senescence, oxidant-antioxidant imbalance, endothelial or epithelial-to-mesenchymal transition, inflammation or immune response and their relationships with the TP53, JUN and β-catenin/TCF4 complex pathways, or abnormal secretion of extracellular matrix proteins. Some of the differentially expressed genes identified by microarray analysis in this project have not been validated by an independent method; these need validation by quantitative RT-PCR. This will further reveal novel differentially expressed genes involved in FECD.

In conclusion, I have identified through genetic studies *TCF4* as the major genetic factor for development of FECD in Caucasian Australian. Additionally, *ATP1B1*, *KANK4* and *LAMC1* were identified as novel genes associated with the risk of developing the disease in Caucasian Australian population. Molecular studies on the *ATP1B1* and *LAMC1* have shown lower expression of the former in FECD, and the presence of the latter in the corneal endothelium, stroma and epithelium. Through proteomic approach, APOE was identified for the first time to have reduced relative abundance in the Desecemet's membrane in FECD. Molecular studies revealed a correlation in differential expression of the protein and gene in the corneal

endothelium in the disease. These findings have revealed APOE as an important novel gene in the pathophysiology of FECD. Gene expression studies have identified several novel genes with differential expression between FECD-affected and unaffected corneal endothelium. Biological functions of some of these genes suggest their involvement in protein synthesis, electrochemical gradient homeostasis, premature senescence and oxidant-antioxidant imbalance pathways known to be involved in FECD. Other genes have shown for the first time an involvement of inflammation or immune response in corneal endothelium in FECD patients. Ingenuity Pathway Analysis of all the genes/proteins identified to be involved in FECD in this project and in past studies further revealed regulation of FECD-implicated genes by the *TCF4* gene. Additionally, the new genes are seen to participate in the known TP53 and JUN pathways, and inflammatory/ immune response. Thus, the aims of this project were successfully achieved and have led to novel findings, and advanced understanding of the molecular basis of FECD.

APPENDIX

Table 1. Summary of the defined corneal dystrophies: modes of inheritance, genetic loci and the identified gene are indicated. The table has been derived from Klintworth 2009 [10].

CD phenotype	Mode of inheritance	Genetic locus	Causative gene
Superficial CD			
Meesman	AD	12q13	*KRT3*
		17q12	*KRT12*
Stocker-Holt	AD	17q12	*KRT12*
Reis-Bucklers	AD	5q31	*TGFBI*
Thiel-Behnke	AD	5q31	*TGFBI*
		10q23-q24	Unknown
Gelatinous droplike	AR	1p32	*TACSTD2*
Subepithelial mucinous	AD	Unknown	Unknown
Lisch epithelial dystrophy	XR	Xp22.3	Unknown
Epithelial recurrence erosion	AD	Unknown	Unknown
Stromal CD			
Macular	AR	16q22	*CHST6*
Granular type I and type II	AD	5q31	*TGFBI*
Lattice type I and variants	AD	5q31	*TGFBI*
Lattice type II	AD	9q34	*GSN*
Fleck dystrophy	AD	2q35	*PIP5K3*
Schnyder	AD	1p34.1-p36	*UBIAD1*
Posterior amorphous	AD	Unknown	Unknown
Congenital stromal	AD	12q13.2	*DCN*
Posterior CD			
FECD (early-onset)	AD	1p34.3	*COL8A2*
FECD (late-onset)	AD	13pTel-13q12.13	Unknown
		18q21.2-q21.23	Unknown
		20p13-p12	*SLC4A11*
		10p11.2	*ZEB1*
		18q21.1	*LOXHD1*
		15q25.3	*AGBL1*
PPCD type 1	AD	20P11.2	Unknown
PPCD type 2	AD	1p34.3-p32.3	*COL8A2*
PPCD type 3	AD	10p11.2	*ZEB1*
CHED type 1	AD	20P11.2-q11.2	Unknown
CHED type 2	AR	20p13-p12	*SLC4A11*
XECD	XR	Unknown	Unknown

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; XR, X-linked recessive; CD, corneal dystrophy; CHED, congenital hereditary endothelial dystrophy; XECD, X-linked endothelial corneal dystrophy; FECD, Fuchs' endothelial corneal dystrophy; PPCD, Posterior polymorphous corneal dystrophy

Table 2. Set-up of a standard PCR mix.

Reagent	**10µl volume**	**20µl volume**	**Final Concentration**
MQH_2O	5.7	13.4	-
10X PCR buffer	1.0	2	1X
dNTP (4mM)	0.25	0.5	400 µM
Forward primer (10µM)	0.5	1.0	0.05 µM
Reverse primer (10µM)	0.5	1.0	0.05 µM
Hotstar Taq Polymerase	0.05	0.1	0.5U
Template DNA (20ng/µL)	2	2	40 ng/µL

Table 3. Results of quantification of amplified, barcoded FECD amplicon libraries using Qubit® 2.0 Flurometer. The table indicates the number of each FECD DNA pool, and the corresponding barcode and an estimated DNA concentration in ng/mL.

FECD DNA Pool #	Barcode used	Qubit® Concentration (ng/mL)	
		Primer pool 1	Primer pool 2
3	7	476	308
4	8	729	492
5	9	379	256
6	10	496	405
7	11	427	547
8	12	379	615
9	13	634	894
10	14	587	572
11	15	1260	1080
12	2	1180	936
13	3	974	598
14	4	1280	520
15	5	980	738
16	6	1140	800
17	7	946	785
18	8	911	1000

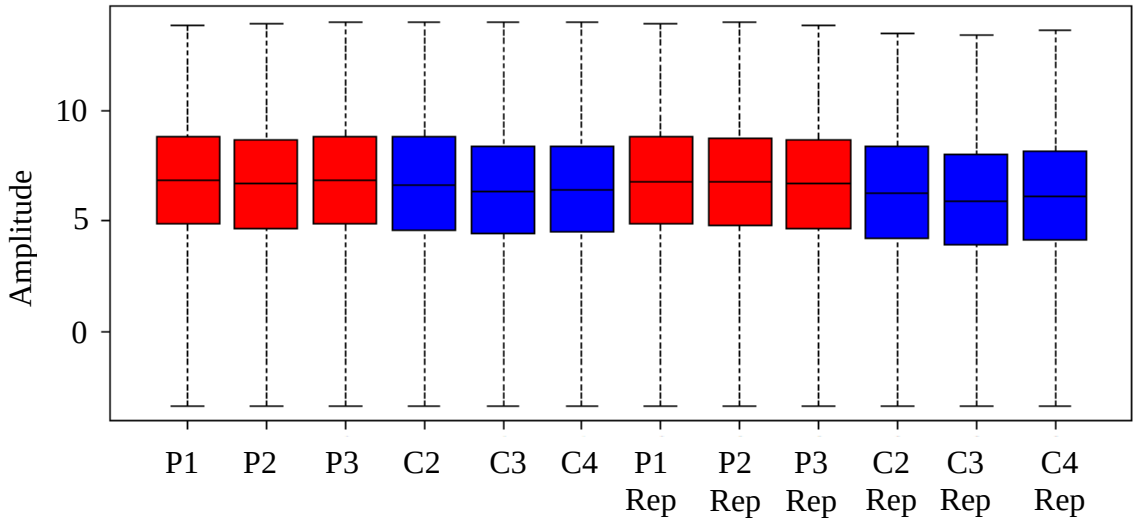


Figure 1 Boxplots of non-normalised microarray data intensity for FECD (P1 – 3) and control (C2 – 4) samples. Rep = replication. P1, pool1; P2, pool2, P3, pool3; C1, control1; C2, control2; C3, control3.

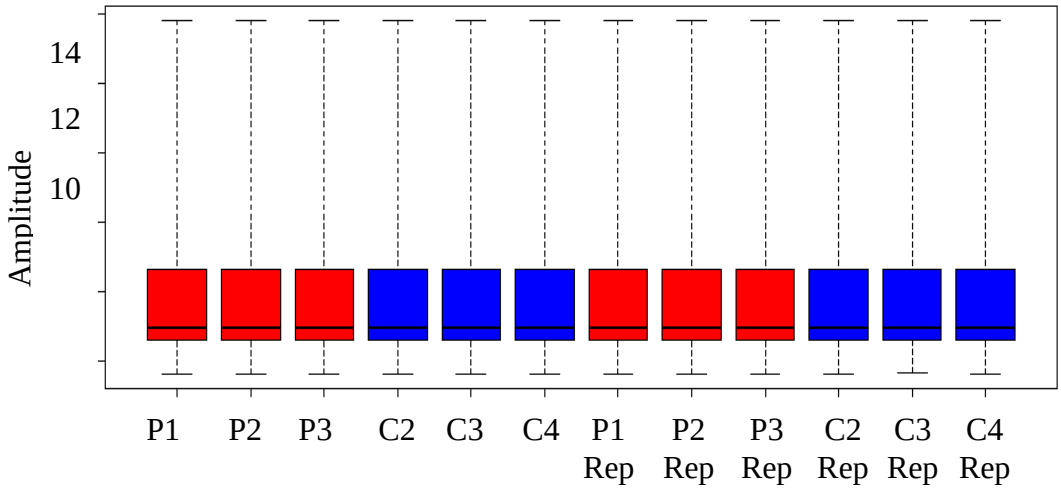


Figure 2 Boxplots of normalised microarray data intensity for FECD (P1 – 3) and control (C2 – 4) samples. Rep means replication. P1, pool1; P2, pool2, P3, pool3; C1, control1; C2, control2; C3, control3.

Table 4. The list of 135 differentially expressed genes in FECD-affected corneal endothelium compared to unaffected corneal endothelium with up- or down-regulated log fold-change threshold of of ≥1.5, and Bonferroni-adjusted p-value ≤0.05. For each gene, the symbol, accession number, log fold change and adjusted p-value have been given.

Gene symbol	Accession #	Log Fold Change	Bonferroni adjusted *p* value
HLA-DRA	NM_019111	6.31	4.28×10^{-8}
CX3CR1	NM_001337	4.11	5.74×10^{-8}
ALPK2	NM_052947	4.44	8.38×10^{-8}
CSF1R	NM_005211	4.63	1.56×10^{-7}
CST1	NM_001898	8.02	1.89×10^{-7}
HLA-DQA1	XM_936128	4.83	4.73×10^{-7}
C1QB	NM_000491	4.41	5.75×10^{-7}
BGN	NM_001711	3.80	6.61×10^{-7}
HLA-DPA1	NM_033554	5.15	9.32×10^{-7}
MMP12	NM_002426	5.76	9.92×10^{-7}
TSHR	NM_000369	3.58	1.86×10^{-6}
SPP1	NM_001040058	4.03	4.75×10^{-6}
C1QC	NM_172369	4.31	5.30×10^{-6}
CLIC6	NM_053277	4.15	5.59×10^{-6}
ITGB2	NM_000211	4.20	6.37×10^{-6}
EDN1	NM_001955	3.99	9.19×10^{-6}
TYROBP	NM_003332	4.59	1.41×10^{-5}
NOX4	NM_016931	4.40	1.84×10^{-5}
CD74	NM_001025159	3.16	2.22×10^{-5}
ALOX5AP	NM_001629	4.57	3.04×10^{-5}
LYZ	NM_000239	3.38	3.72×10^{-5}
HLA-DMB	NM_002118	4.43	4.28×10^{-5}
KRT81	NM_002281	5.21	4.31×10^{-5}
EGR1	NM_001964	4.05	5.67×10^{-5}
TNFRSF11B	NM_002546	4.67	5.82×10^{-5}
FOS	NM_005252	3.56	6.50×10^{-5}
SLCO2B1	NM_007256	3.11	6.86×10^{-5}
PEAR1	NM_001080471	2.89	6.96×10^{-5}
LRRC32	NM_005512	3.84	1.18×10^{-4}
SLC40A1	NM_014585	2.83	1.24×10^{-4}
GLIPR2	NM_022343	4.69	1.34×10^{-4}
GPC3	NM_004484	4.29	1.63×10^{-4}
HLA-DRB6	NR_001298	4.09	1.74×10^{-4}
HSPA6	NM_002155	3.05	2.65×10^{-4}

CD93	NM_012072	2.79	2.71×10^{-4}
ST6GAL1	NM_003032	3.28	3.73×10^{-4}
MS4A6A	NM_152851	3.46	3.74×10^{-4}
AIF1	NM_032955	3.13	3.89×10^{-4}
RGS1	NM_002922	4.14	4.23×10^{-4}
MARCKS	NM_002356	3.84	5.00×10^{-4}
CYBB	NM_000397	2.83	5.30×10^{-4}
NEB	NM_004543	3.25	7.28×10^{-4}
HBB	NM_000518	7.32	7.30×10^{-4}
CTHRC1	NM_138455	3.64	7.90×10^{-4}
CEBPA	NM_004364	3.28	7.98×10^{-4}
SLC2A5	NM_003039	2.82	8.59×10^{-4}
BMP4	NM_130851	4.00	1.03×10^{-3}
COL5A1	NM_000093	4.09	1.09×10^{-3}
CYP24A1	NM_000782	5.00	1.17×10^{-3}
ABCC3	NM_003786	3.21	1.19×10^{-3}
TUBB2B	NM_178012	4.06	1.26×10^{-3}
CD68	NM_001251	2.76	1.29×10^{-3}
SERPINE2	NM_006216	1.93	1.30×10^{-3}
CPAMD8	NM_015692	-2.99	1.84×10^{-3}
PLAU	NM_002658	3.45	1.99×10^{-3}
PAPSS2	NM_004670	4.44	2.03×10^{-3}
VEGFA	NM_003376	-2.08	2.27×10^{-3}
FCER1G	NM_004106	3.41	2.36×10^{-3}
VENTX	NM_014468	2.27	2.49×10^{-3}
HCLS1	NM_005335	2.32	2.58×10^{-3}
DIO2	NM_001007023	-2.95	3.00×10^{-3}
FXYD5	NM_144779	3.79	3.00×10^{-3}
KRT7	NM_005556	4.48	3.02×10^{-3}
HBA1	NM_000558	6.49	3.05×10^{-3}
BAMBI	NM_012342	2.09	3.26×10^{-3}
IGF1	NM_000618	4.26	3.54×10^{-3}
HBEGF	NM_001945	2.66	3.56×10^{-3}
TUBB6	XM_940079	2.68	3.67×10^{-3}
NT5E	NM_002526	2.53	3.74×10^{-3}
CD14	NM_001040021	2.60	3.81×10^{-3}
PDK4	NM_002612	-2.27	3.95×10^{-3}
VCAM1	NM_001078	3.84	4.00×10^{-3}
CD163	NM_203416	2.21	4.13×10^{-3}
ENC1	NM_003633	2.36	4.18×10^{-3}
TRIB1	NM_025195	2.93	4.27×10^{-3}
GPR65	NM_003608	2.11	4.36×10^{-3}

MCOLN2	NM_153259	3.60	4.39×10^{-3}
PPP1R1B	NM_181505	-4.32	5.23×10^{-3}
BMP6	NM_001718	2.43	5.23×10^{-3}
MEGF6	NM_001409	2.19	5.37×10^{-3}
COL4A1	NM_001845	4.37	5.44×10^{-3}
PRSS23	NM_007173	2.80	5.93×10^{-3}
CDKN2B	NM_078487	3.84	5.96×10^{-3}
SHOX	NM_006883	-2.47	6.07×10^{-3}
CHN2	NM_004067	2.52	6.61×10^{-3}
ANXA1	NM_000700	2.56	6.63×10^{-3}
KRT80	NM_182507	3.08	7.06×10^{-3}
EPB41L3	NM_012307	2.23	7.08×10^{-3}
MATN3	NM_002381	4.02	8.50×10^{-3}
HCST	NM_001007469	1.61	8.50×10^{-3}
HBA2	NM_000517	6.56	9.09×10^{-3}
SERPINE1	NM_000602	2.61	9.63×10^{-3}
COL18A1	NM_130445	3.22	9.67×10^{-3}
RAC2	NM_002872	2.31	1.00×10^{-2}
KAL1	NM_000216	3.31	1.05×10^{-2}
ESM1	NM_007036	3.12	1.07×10^{-2}
S100A11	NM_005620	3.03	1.07×10^{-2}
S100A10	NM_002966	3.38	1.10×10^{-2}
CLEC5A	NM_013252	2.35	1.11×10^{-2}
HAMP	NM_021175	3.07	1.12×10^{-2}
IGFBP7	NM_001553	2.43	1.34×10^{-2}
NAPSB	NR_002798	2.75	1.39×10^{-2}
CLEC4GP1	NR_002931	-3.55	1.42×10^{-2}
COL6A2	NM_001849	3.44	1.50×10^{-2}
PAMR1	NM_015430	1.66	1.51×10^{-2}
MME	NM_000902	3.74	1.59×10^{-2}
ALDH1L1	NM_012190	-2.74	1.64×10^{-2}
FCGR2A	NM_021642	2.46	1.68×10^{-2}
RBP7	NM_052960	-2.36	1.68×10^{-2}
PCP4	NM_006198	5.88	1.74×10^{-2}
HAVCR2	NM_032782	2.73	2.26×10^{-2}
SYK	NM_003177	2.05	2.28×10^{-2}
PLAC9	NM_001012973	3.99	2.38×10^{-2}
LPAR5	NM_020400	2.29	2.56×10^{-2}
FGD2	NM_173558	2.89	2.62×10^{-2}
PXDN	NM_012293	3.61	2.77×10^{-2}
SPARCL1	NM_004684	-2.21	2.80×10^{-2}
CCL2	NM_002982	1.99	2.82×10^{-2}

ALOX5	XM_001127464	2.62	2.83×10^{-2}
TMEM158	NM_015444	2.52	3.04×10^{-2}
HLA-DRB4	NM_021983	4.09	3.25×10^{-2}
LRRC17	NM_005824	2.16	3.56×10^{-2}
FHL2	NM_201557	2.25	3.59×10^{-2}
LAPTM5	NM_006762	2.94	3.71×10^{-2}
SCARNA11	NR_003012	-2.10	3.79×10^{-2}
FILIP1L	NM_014890	2.04	3.88×10^{-2}
IGFBP4	NM_001552	3.61	3.97×10^{-2}
FEZ1	NM_005103	2.14	4.05×10^{-2}
PLAT	NM_000930	2.51	4.26×10^{-2}
TSC22D3	NM_004089	-1.84	4.26×10^{-2}
RNASE1	NM_198235	3.16	4.43×10^{-2}
SLC7A5	NM_003486	2.95	4.62×10^{-2}
BCAT1	NM_005504	1.85	4.80×10^{-2}
PLCB2	NM_004573	2.55	4.83×10^{-2}
DHDH	NM_014475	-1.91	4.95×10^{-2}

Table 5. The list of the genes/proteins involved in FECD reported in the literature.

Gene symbol	References
AGBL1	[79]
AGRN	[131, 132]
APOD	[127, 132]
CDKN1A	[70]
CDKN2A	[70]
CLU	[93, 114, 131, 132]
COL1A1	[132, 204]
COL3A1	[131]
COL4A1	[204]
COL16A1	[131]
DICER1	[204]
DRAM1	[153]
FBN1	[131]
FN1	[131]
ITGA4	[131]
JUN	[154]
KERA	[132]
COX2	[127, 154]
PRDX2	[69, 130]
PRDX3	[130]
PRDX5	[69, 130]
TGFBI	[93, 114, 131, 132]
TP53	[369]
ZEB1	[83]

REFERENCES

1. Forrester, JB, Dick, AD, McMenamin, PG, et al., *Anatomy of the eye and orbit*, in *The Eye: Basic Sciences In Practice*. 2002, Elsevier Limited: London. p. 15-58
2. Reinhard, T and Larkin, F, eds. *Cornea and External Eye Disease*. Essentials in Ophthalmology, ed. Krieglstein, GW, RN 2008, Springer: London, Freiburg.
3. Asbell, P and Brocks, D, *Cornea Overview*, in *Encyclopedia of the Eye*, Darlene, AD, Editor. 2010, Academic Press: Oxford. p. 522-531.
4. Hassell, JR and Birk, DE, *The molecular basis of corneal transparency.* Experimental Eye Research, 2010. **91**(3): p. 326-335.
5. Barishak, YR, *Embryology of the eye and its adnexae.* Developments in Ophthalmology, 1992. **24**: p. 1-142.
6. Graw, J, *Eye development.* Current Topics in Developmental Biology, 2010. **90**: p. 343-386.
7. Stepp, MA, *Corneal Epithelium: Cell Biology and Basic Science*, in *Encyclopedia of the Eye*, Darlene, AD, Editor. 2010, Academic Press: Oxford. p. 435-441.
8. Davanger, M and Evensen, A, *Role of the pericorneal papillary structure in renewal of corneal epithelium.* Nature, 1971. **229**(5286): p. 560-561.
9. Huang, AJ and Tseng, SC, *Corneal epithelial wound healing in the absence of limbal epithelium.* Investigative Ophthalmology and Visual Science, 1991. **32**(1): p. 96-105.

10. Klintworth, G, *Corneal dystrophies.* Orphanet Journal of Rare Diseases, 2009. **4**(1): p. 7.

11. Hecker, LA, McLaren, JW, Bachman, LA, et al., *Anterior Keratocyte Depletion in Fuchs Endothelial Dystrophy.* Archives of Ophthalmology, 2011. **129**(5): p. 555-561.

12. Szentmary, N, Szende, B, and Suveges, I, *Epithelial cell, keratocyte, and endothelial cell apoptosis in Fuchs' dystrophy and in pseudophakic bullous keratopathy.* European Journal of Ophthalmology, 2005. **15**(1): p. 17-22.

13. Bourne, WM, *Biology of the corneal endothelium in health and disease.* Eye, 2000. **17**(8): p. 912-918.

14. Johnson, DH, Bourne, WM, and Campbell, RJ, *The Ultrastructure of Descemet's Membrane: I. Changes With Age in Normal Corneas.* Archives of Ophthalmology, 1982. **100**(12): p. 1942-1947.

15. Bourne, WM, *Biology of the corneal endothelium in health and disease.* Eye, 2003. **17**(8): p. 912-918.

16. Mergler, S and Pleyer, U, *The human corneal endothelium: new insights into electrophysiology and ion channels.* Progress in Retinal and Eye Research, 2007. **26**(4): p. 359-378.

17. Bonanno, JA, *Identity and regulation of ion transport mechanisms in the corneal endothelium.* Progress in Retinal and Eye Research, 2003. **22**(1): p. 69-94.

18. Bonanno, JA and Srinivas, SP, *Regulation of Corneal Endothelial Function*, in *Encyclopedia of the Eye*, Darlene, AD, Editor. 2010, Academic Press: Oxford. p. 21-27.

19. Whikehart, DR, *Corneal Endothelium: Overview*, in *Encyclopedia of the Eye*, Darlene, AD, Editor. 2010, Academic Press: Oxford. p. 424-434.

20. Hartsock, A and Nelson, WJ, *Adherens and Tight Junctions: Structure, Function and Connections to the Actin Cytoskeleton.* Biochimica et Biophysica Acta, 2008. **1778**(3): p. 660-669.

21. Edelhauser, HF, *The balance between corneal transparency and edema: the Proctor Lecture.* Investigative Ophthalmology and Visual Science, 2006. **47**(5): p. 1754-1767.

22. Maurice, DM, *Passive ion fluxes across the corneal endothelium.* Current Eye Research, 1985. **4**(4): p. 339-349.

23. Joyce, NC, *Proliferative capacity of the corneal endothelium.* Progress in Retinal and Eye Research, 2003. **22**(3): p. 359-389.

24. Murphy, C, Alvarado, J, and Juster, R, *Prenatal and postnatal growth of the human Descemet's membrane.* Investigative Ophthalmology and Visual Science, 1984. **25**(12): p. 1402-1415.

25. Bourne, WM, Nelson, LR, and Hodge, DO, *Central corneal endothelial cell changes over a ten-year period.* Investigative Ophthalmology and Visual Science, 1997. **38**(3): p. 779-782.

26. Hollingsworth, J, Perez-Gomez, I, Mutalib, HA, et al., *A population study of the normal cornea using an in vivo, slit-scanning confocal microscope.* Optometry and Vision Science, 2001. **78**(10): p. 706-711.

27. Edelhauser, HF, *The resiliency of the corneal endothelium to refractive and intraocular surgery.* Cornea, 2000. **19**(3): p. 263-273.

28. O'Neal, MR and Polse, KA, *Decreased endothelial pump function with aging.* Investigative Ophthalmology and Visual Science, 1986. **27**(4): p. 457-463.

29. Bourne, WM, Nelson, LR, and Hodge, DO, *Continued endothelial cell loss ten years after lens implantation.* Ophthalmology, 1994. **101**(6): p. 1014-1022; discussion 1022-1013.

30. Bourne, WM, *Cellular changes in transplanted human corneas.* Cornea, 2001. **20**(6): p. 560-569.

31. Schultz, RO, Matsuda, M, Yee, RW, et al., *Corneal endothelial changes in type I and type II diabetes mellitus.* American Journal of Ophthalmology, 1984. **98**(4): p. 401-410.

32. Wu, SC, Jeng, S, Huang, SC, et al., *Corneal endothelial damage after neodymium:YAG laser iridotomy.* Ophthalmic Surgery and Lasers, 2000. **31**(5): p. 411-416.

33. Konowal, A, Morrison, JC, Brown, SV, et al., *Irreversible corneal decompensation in patients treated with topical dorzolamide.* American Journal of Ophthalmology, 1999. **127**(4): p. 403-406.

34. Stainer, GA, Akers, PH, Binder, PS, et al., *Correlative microscopy and tissue culture of congenital hereditary endothelial dystrophy.* American Journal of Ophthalmology, 1982. **93**(4): p. 456-465.

35. Adamis, AP, Filatov, V, Tripathi, BJ, et al., *Fuchs' endothelial dystrophy of the cornea.* Survey of Ophthalmology, 1993. **38**(2): p. 149-168.

36. Mandell, RB, Polse, KA, Brand, RJ, et al., *Corneal hydration control in Fuchs' dystrophy.* Investigative Ophthalmology and Visual Science, 1989. **30**(5): p. 845-852.

37. Weiss, JS, *Corneal dystrophies: molecular genetics to therapeutic intervention--Fifth ARVO/Pfizer Ophthalmics Research Institute Conference.*

Investigative Ophthalmology and Visual Science, 2010. **51**(11): p. 5391-5402.

38. Feldman, BH and Afshari, NA, *Corneal Dystrophies,* in *Encyclopedia of the Eye,* Darlene, AD, Editor. 2010, Academic Press: Oxford. p. 416-423.

39. Riazuddin, SA, Zaghloul, NA, Al-Saif, A, et al., *Missense mutations in TCF8 cause late-onset Fuchs corneal dystrophy and interact with FCD4 on chromosome 9p.* The American Journal of Human Genetics, 2010. **86**(1): p. 45-53.

40. Stone, EM, Mathers, WD, Rosenwasser, GOD, et al., *Three autosomal dominant corneal dystrophies map to chromosome 5q.* Nature Genetics, 1994. **6**(1): p. 47-51.

41. Munier, FL, Korvatska, E, Djemai, A, et al., *Kerato-epithelin mutations in four 5q31-linked corneal dystrophies.* Nature Genetics, 1997. **15**(3): p. 247-251.

42. Hirano, K, Klintworth, GK, Zhan, Q, et al., *Beta ig-h3 is synthesized by corneal epithelium and perhaps endotheliumin Fuchs' dystrophic corneas.* Current Eye Research, 1996. **15**(9): p. 965-972.

43. Clout, NJ and Hohenester, E, *A model of FAS1 domain 4 of the corneal protein beta(ig)-h3 gives a clearer view on corneal dystrophies.* Molecular Vision, 2003. **9**: p. 440-448.

44. Biswas, S, Munier, FL, Yardley, J, et al., *Missense mutations in COL8A2, the gene encoding the alpha2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy.* Human Molecular Genetics, 2001. **10**(21): p. 2415-2423.

45. Eghrari, AO and Gottsch, JD, *Fuchs' corneal dystrophy.* Expert review of ophthalmology, 2010. **5**(2): p. 147-159.

46. Elhalis, H, Azizi, B, and Jurkunas, UV, *Fuchs endothelial corneal dystrophy.* The ocular surface, 2010. **8**(4): p. 173-184.

47. Williams, KA, Lowe, MT, Keane, MC, et al., *The Australian Corneal Graft Registry 2012 Report*. 2012, Flinders University: Adelaide.

48. Raecker, ME, Erie, JC, Patel, SV, et al., *Long-term keratometric changes after penetrating keratoplasty for keratoconus and Fuchs Endothelial dystrophy.* American Journal of Ophthalmology, 2009. **147**(2): p. 227-233.

49. Williams, KA, Lowe, MT, Bartlett, CM, et al., *The Australian Corneal Graft Registry: 2007 Report*. 2007, Flinders University: Adelaide.

50. Price, MO, Fairchild, KM, Price, DA, et al., *Descemet's Stripping Endothelial Keratoplasty: Five-Year Graft Survival and Endothelial Cell Loss.* Ophthalmology, 2011. **118**(4): p. 725-729.

51. Price, MO, Gorovoy, M, Benetz, BA, et al., *Descemet's Stripping Automated Endothelial Keratoplasty Outcomes Compared with Penetrating Keratoplasty from the Cornea Donor Study.* Ophthalmology, 2010. **117**(3): p. 438-444.

52. Godeiro, KD, Coutinho, AB, Pereira, PR, et al., *Histopathological diagnosis of corneal button specimens: an epidemiological study.* Ophthalmic Epidemiology, 2007. **14**(2): p. 70-75.

53. Li, X, Wang, L, Dustin, L, et al., *Age distribution of various corneal diseases in China by histopathological examination of 3112 surgical specimens.* Investigative Ophthalmology and Visual Science, 2014. **55**(5): p. 3022-3028.

54. Xie, L, Song, Z, Zhao, J, et al., *Indications for penetrating keratoplasty in north China.* Cornea, 2007. **26**(9): p. 1070-1073.

55. Sheldon, CA, McCarthy, JM, and White, VA, *Correlation of clinical and pathologic diagnoses of corneal disease in penetrating keratoplasties in Vancouver: a 10-year review.* Canadian Journal of Ophthalmology, 2012. **47**(1): p. 5-10.

56. Galvis, V, Tello, A, Gomez, AJ, et al., *Corneal Transplantation at an Ophthalmological Referral Center in Colombia: Indications and Techniques (2004-2011).* The open ophthalmology journal, 2013. **7**: p. 30-33.

57. Zare, M, Javadi, MA, Einollahi, B, et al., *Changing indications and surgical techniques for corneal transplantation between 2004 and 2009 at a tertiary referral center*. Middle East African journal of ophthalmology, 2012. **19**(3): p. 323-329.

58. Omar, N, Bou Chacra, CT, and Tabbara, KF, *Outcome of corneal transplantation in a private institution in Saudi Arabia.* Clinical Ophthalmology (Auckland, N.Z.), 2013. **7**: p. 1311-1318.

59. Ghosheh, FR, Cremona, FA, Rapuano, CJ, et al., *Trends in penetrating keratoplasty in the United States 1980-2005.* International Ophthalmology, 2008. **28**(3): p. 147-153.

60. Santo, RM, Yamaguchi, T, Kanai, A, et al., *Clinical and histopathologic features of corneal dystrophies in Japan.* Ophthalmology, 1995. **102**(4): p. 557-567.

61. Afshari, NA, Pittard, AB, Siddiqui, A, et al., *Clinical study of Fuchs corneal endothelial dystrophy leading to penetrating keratoplasty: a 30-year experience.* Archives of Ophthalmology, 2006. **124**(6): p. 777-780.

62. Krachmer, JH, Purcell, JJ, Jr., Young, CW, et al., *Corneal endothelial dystrophy. A study of 64 families.* Archives of Ophthalmology, 1978. **96**(11): p. 2036-2039.

63. Rosenblum, P, Stark, WJ, Maumenee, IH, et al., *Hereditary Fuchs' Dystrophy.* American Journal of Ophthalmology, 1980. **90**(4): p. 455-462.

64. Louttit, MD, Kopplin, LJ, Igo, RP, Jr., et al., *A multicenter study to map genes for Fuchs endothelial corneal dystrophy: baseline characteristics and heritability.* Cornea, 2012. **31**(1): p. 26-35.

65. Williams, K, Lowe, M.T., Keane, M.C., Jones, V.J., Loh, R.S., Coster, D.J. , *The Australian Corneal Graft Registry 2012 Report*. 2012, Flinders University of South Austalia: Adelaide.

66. Eghrari, AO, McGlumphy, EJ, Iliff, BW, et al., *Prevalence and severity of fuchs corneal dystrophy in Tangier Island.* American Journal of Ophthalmology, 2012. **153**(6): p. 1067-1072.

67. Cross, HE, Maumenee, AE, and Cantolino, SJ, *Inheritance of Fuchs' Endothelial Dystrophy.* Archives of Ophthalmology, 1971. **85**(3): p. 268-272.

68. Matsuda, M, Yee, RW, and Edelhauser, HF, *Comparison of the corneal endothelium in an American and a Japanese population.* Archives of Ophthalmology, 1985. **103**(1): p. 68-70.

69. Jurkunas, UV, Bitar, MS, Funaki, T, et al., *Evidence of oxidative stress in the pathogenesis of fuchs endothelial corneal dystrophy.* American Journal of Pathology, 2010. **177**(5): p. 2278-2289.

70. Matthaei, M, Meng, H, Meeker, AK, et al., *Endothelial Cdkn1a (p21) overexpression and accelerated senescence in a mouse model of Fuchs*

endothelial corneal dystrophy. Investigative Ophthalmology and Visual Science, 2012. **53**(10): p. 6718-6727.

71. Mok, JW, Kim, HS, and Joo, CK, *Q455V mutation in COL8A2 is associated with Fuchs' corneal dystrophy in Korean patients.* Eye, 2008. **23**(4): p. 895-903.

72. Gottsch, JD, Sundin, OH, Liu, SH, et al., *Inheritance of a novel COL8A2 mutation defines a distinct early-onset subtype of fuchs corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2005. **46**(6): p. 1934-1939.

73. Gottsch, JD, Zhang, C, Sundin, OH, et al., *Fuchs corneal dystrophy: aberrant collagen distribution in an L450W mutant of the COL8A2 gene.* Investigative Ophthalmology and Visual Science, 2005. **46**(12): p. 4504-4511.

74. Kabosova, A, Azar, DT, Bannikov, GA, et al., *Compositional differences between infant and adult human corneal basement membranes.* Investigative Ophthalmology and Visual Science, 2007. **48**(11): p. 4989-4999.

75. Aldave, AJ, Rayner, SA, Salem, AK, et al., *No pathogenic mutations identified in the COL8A1 and COL8A2 genes in familial Fuchs corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2006. **47**(9): p. 3787-3790.

76. Vithana, EN, Morgan, PE, Ramprasad, V, et al., *SLC4A11 mutations in Fuchs endothelial corneal dystrophy.* Human Molecular Genetics, 2008. **17**(5): p. 656-666.

77. Riazuddin, SA, Iqbal, M, Wang, Y, et al., *A splice-site mutation in a retina-specific exon of BBS8 causes nonsyndromic retinitis pigmentosa.* The American Journal of Human Genetics, 2010. **86**(5): p. 805-812.

78. Riazuddin, SA, Parker, DS, McGlumphy, EJ, et al., *Mutations in LOXHD1, a recessive-deafness locus, cause dominant late-onset Fuchs corneal dystrophy.* The American Journal of Human Genetics, 2012. **90**(3): p. 533-539.

79. Riazuddin, SA, Vasanth, S, Katsanis, N, et al., *Mutations in AGBL1 cause dominant late-onset Fuchs corneal dystrophy and alter protein-protein interaction with TCF4.* The American Journal of Human Genetics, 2013. **93**(4): p. 758-764.

80. Sundin, OH, Broman, KW, Chang, HH, et al., *A common locus for late-onset Fuchs corneal dystrophy maps to 18q21.2-q21.32.* Investigative Ophthalmology and Visual Science, 2006. **47**(9): p. 3919-3926.

81. Riazuddin, SA, Eghrari, AO, Al-Saif, A, et al., *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, 2009. **50**(12): p. 5667-5671.

82. Meadows, DN, Eghrari, AO, Riazuddin, SA, et al., *Progression of Fuchs corneal dystrophy in a family linked to the FCD1 locus.* Investigative Ophthalmology and Visual Science, 2009. **50**(12): p. 5662-5666.

83. Riazuddin, SA, Vithana, EN, Seet, LF, et al., *Missense mutations in the sodium borate cotransporter SLC4A11 cause late-onset Fuchs corneal dystrophy.* Human Mutation, 2010. **31**(11): p. 1261-1268.

84. Liu, Y, El-Naggar, S, Darling, DS, et al., *Zeb1 links epithelial-mesenchymal transition and cellular senescence.* Development, 2008. **135**(3): p. 579-588.

85. Vandewalle, C, Van Roy, F, and Berx, G, *The role of the ZEB family of transcription factors in development and disease.* Cellular and Molecular Life Sciences, 2009. **66**(5): p. 773-787.

86. Hurt, EM, Saykally, JN, Anose, BM, et al., *Expression of the ZEB1 (deltaEF1) transcription factor in human: additional insights.* Molecular and Cellular Biochemistry, 2008. **318**(1-2): p. 89-99.

87. Soumittra, N, Loganathan, SK, Madhavan, D, et al., *Biosynthetic and functional defects in newly identified SLC4A11 mutants and absence of COL8A2 mutations in Fuchs endothelial corneal dystrophy.* Journal of Human Genetics, 2014. **59**(8): p. 444-453.

88. Risch, N and Merikangas, K, *The future of genetic studies of complex human diseases.* Science, 1996. **273**(5281): p. 1516-1517.

89. Dean, M, *Approaches to identify genes for complex human diseases: lessons from Mendelian disorders.* Human Mutation, 2003. **22**(4): p. 261-274.

90. Wang, WY, Barratt, BJ, Clayton, DG, et al., *Genome-wide association studies: theoretical and practical concerns.* Nature reviews. Genetics, 2005. **6**(2): p. 109-118.

91. Burdon, KP, Macgregor, S, Hewitt, AW, et al., *Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1*. Nature Genetics, 2011. **43**(6): p. 574-578.

92. Baratz, KH, Tosakulwong, N, Ryu, E, et al., *E2-2 protein and Fuchs's corneal dystrophy.* New England Journal of Medicine, 2010. **363**(11): p. 1016-1024.

93. Kuot, A, Hewitt, AW, Griggs, K, et al., *Association of TCF4 and CLU polymorphisms with Fuchs' endothelial dystrophy and implication of CLU and TGFBI proteins in the disease process.* European Journal of Human Genetics, 2012. **20**(6): p. 632-638.

94. Riazuddin, SA, McGlumphy, EJ, Yeo, WS, et al., *Replication of the TCF4 intronic variant in late-onset Fuchs corneal dystrophy and evidence of independence from the FCD2 locus.* Investigative Ophthalmology and Visual Science, 2011. **52**(5): p. 2825-2829.

95. Li, Y-J, Minear, MA, Rimmler, J, et al., *Replication of TCF4 through Association and Linkage Studies in Late-Onset Fuchs Endothelial Corneal Dystrophy.* PloS One, 2011. **6**(4): p. e18044.

96. Stamler, JF, Roos, BR, Wagoner, MD, et al., *Confirmation of the association between the TCF4 risk allele and Fuchs endothelial corneal dystrophy in patients from the Midwestern United States.* Ophthalmic Genetics, 2013. **34**(1-2): p. 32-34.

97. Thalamuthu, A, Khor, CC, Venkataraman, D, et al., *Association of TCF4 gene polymorphisms with Fuchs' corneal dystrophy in the Chinese.* Investigative Ophthalmology and Visual Science, 2011. **52**(8): p. 5573-5578.

98. Wang, KJ, Jhanji, V, Chen, J, et al., *Association of transcription factor 4 (TCF4) and protein tyrosine phosphatase, receptor type G (PTPRG) with corneal dystrophies in southern Chinese.* Ophthalmic Genetics, 2014. **35**(3): p. 138-141.

99. Nanda, GG, Padhy, B, Samal, S, et al., *Genetic Association of TCF4 Intronic Polymorphisms, CTG18.1 and rs17089887, With Fuchs' Endothelial Corneal Dystrophy in an Indian Population.* Investigative Ophthalmology and Visual Science, 2014. **55**(11): p. 7674-7680.

100. Lau, LC, Ma, L, Young, AL, et al., *Association of Common Variants in TCF4 and PTPRG with Fuchs' Corneal Dystrophy: A Systematic Review and Meta-Analysis.* PloS One, 2014. **9**(10): p. e109142.

101. Breschel, TS, McInnis, MG, Margolis, RL, et al., *A novel, heritable, expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1.* Human Molecular Genetics, 1997. **6**(11): p. 1855-1863.

102. Wieben, ED, Aleff, RA, Tosakulwong, N, et al., *A Common Trinucleotide Repeat Expansion within the Transcription Factor 4 (TCF4, E2-2) Gene Predicts Fuchs Corneal Dystrophy.* PloS One, 2012. **7**(11): p. e49083.

103. Kiliszek, A and Rypniewski, W, *Structural studies of CNG repeats.* Nucleic Acids Research, 2014. **42**(13): p. 8189-8199.

104. Orr, HT and Zoghbi, HY, *Trinucleotide repeat disorders.* Annual Review of Neuroscience, 2007. **30**: p. 575-621.

105. Mootha, VV, Gong, X, Ku, HC, et al., *Association and familial segregation of CTG18.1 trinucleotide repeat expansion of TCF4 gene in Fuchs' endothelial corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2014. **55**(1): p. 33-42.

106. Xing, C, Gong, X, Hussain, I, et al., *Trans-ethnic Replication of Association of CTG18.1 Repeat Expansion of TCF4 Gene with Fuchs Corneal Dystrophy in Chinese Implies Common Causal Variant.* Investigative Ophthalmology and Visual Science, 2014.

107. Zhu, AY, Eberhart, CG, and Jun, AS, *Fuchs endothelial corneal dystrophy: a neurodegenerative disorder?* JAMA ophthalmology, 2014. **132**(4): p. 377-378.

108. Du, J, Aleff, RA, Soragni, E, et al., *RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy.* Journal of Biological Chemistry, 2015. **290**(10): p. 5979-5990.

109. Wright, AF and Dhillon, B, *Major progress in Fuchs's corneal dystrophy.* New England Journal of Medicine, 2010. **363**(11): p. 1072-1075.

110. Murre, C, Bain, G, van Dijk, MA, et al., *Structure and function of helix-loop-helix proteins.* Biochimica et Biophysica Acta, 1994. **1218**(2): p. 129-135.

111. Sobrado, VR, Moreno-Bueno, G, Cubillo, E, et al., *The class I bHLH factors E2-2A and E2-2B regulate EMT.* Journal of Cell Science, 2009. **122**(Pt 7): p. 1014-1024.

112. Eger, A, Aigner, K, Sonderegger, S, et al., *DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells.* Oncogene, 2005. **24**(14): p. 2375-2385.

113. Jurkunas, UV, Bitar, MS, Rawe, I, et al., *Increased clusterin expression in Fuchs' endothelial dystrophy.* Investigative Ophthalmology and Visual Science, 2008. **49**(7): p. 2946-2955.

114. Jurkunas, UV, Bitar, M, and Rawe, I, *Colocalization of increased transforming growth factor-beta-induced protein (TGFBIp) and Clusterin in Fuchs endothelial corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2009. **50**(3): p. 1129-1136.

115. Li, YJ, Minear, MA, Qin, X, et al., *Mitochondrial polymorphism A10398G and Haplogroup I are associated with Fuchs' endothelial corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2014. **55**(7): p. 4577-4584.

116. Jackson, AJ, Robinson, FO, Frazer, DG, et al., *Corneal guttata: a comparative clinical and specular micrographic study.* Eye (Lond), 1999. **13 (Pt 6)**: p. 737-743.

117. Chiou, AG, Kaufman, SC, Beuerman, RW, et al., *Confocal microscopy in cornea guttata and Fuchs' endothelial dystrophy.* British Journal of Ophthalmology, 1999. **83**(2): p. 185-189.

118. Fuchs, E, *Dystrophia epithelialis corneae.* Albrecht von Graefes Archiv für Ophthalmologie, 1910. **76**(3): p. 478-508.

119. Muragaki, Y, Jacenko, O, Apte, S, et al., *The alpha 2(VIII) collagen gene. A novel member of the short chain collagen family located on the human chromosome 1.* Journal of Biological Chemistry, 1991. **266**(12): p. 7721-7727.

120. Sage, H and Iruela-Arispe, ML, *Type VIII collagen in murine development. Association with capillary formation in vitro.* Annals of the New York Academy of Sciences, 1990. **580**: p. 17-31.

121. Levy, SG, Moss, J, Sawada, H, et al., *The composition of wide-spaced collagen in normal and diseased Descemet's membrane.* Current Eye Research, 1996. **15**(1): p. 45-52.

122. Illidge, C, Kielty, C, and Shuttleworth, A, *Type VIII collagen: heterotrimeric chain association.* International Journal of Biochemistry and Cell Biology, 2001. **33**(5): p. 521-529.

123. Stephan, S, Sherratt, MJ, Hodson, N, et al., *Expression and supramolecular assembly of recombinant alpha1(viii) and alpha2(viii) collagen homotrimers.* Journal of Biological Chemistry, 2004. **279**(20): p. 21469-21477.

124. Wilson, SE and Bourne, WM, *Fuchs' Dystrophy.* Cornea, 1988. **7**(1): p. 2-18.

125. Bergmanson, JP, Sheldon, TM, and Goosey, JD, *Fuchs' endothelial dystrophy: a fresh look at an aging disease.* Ophthalmic and Physiological Optics, 1999. **19**(3): p. 210-222.

126. Engler, C, Kelliher, C, Spitze, AR, et al., *Unfolded protein response in fuchs endothelial corneal dystrophy: a unifying pathogenic pathway?* American Journal of Ophthalmology, 2010. **149**(2): p. 194-202.e192.

127. Gottsch, JD, Bowers, AL, Margulies, EH, et al., *Serial analysis of gene expression in the corneal endothelium of Fuchs' dystrophy.* Investigative Ophthalmology and Visual Science, 2003. **44**(2): p. 594-599.

128. Chung, DW, Frausto, RF, Ann, LB, et al., *Functional impact of ZEB1 mutations associated with posterior polymorphous and Fuchs' endothelial corneal dystrophies.* Investigative Ophthalmology and Visual Science, 2014. **55**(10): p. 6159-6166.

129. Grillet, N, Schwander, M, Hildebrand, MS, et al., *Mutations in LOXHD1, an evolutionarily conserved stereociliary protein, disrupt hair cell function in mice and cause progressive hearing loss in humans.* The American Journal of Human Genetics, 2009. **85**(3): p. 328-337.

130. Jurkunas, UV, Rawe, I, Bitar, MS, et al., *Decreased expression of peroxiredoxins in Fuchs' endothelial dystrophy.* Investigative Ophthalmology and Visual Science, 2008. **49**(7): p. 2956-2963.

131. Weller, JM, Zenkel, M, Schlotzer-Schrehardt, U, et al., *Extracellular matrix alterations in late-onset Fuchs' corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2014. **55**(6): p. 3700-3708.

132. Poulsen, ET, Dyrlund, TF, Runager, K, et al., *Proteomics of Fuchs' endothelial corneal dystrophy support that the extracellular matrix of Descemet's membrane is disordered.* Journal of Proteome Research, 2014. **13**(11): p. 4659-4667.

133. Kim, JH, Kim, JH, Jun, HO, et al., *Protective effect of clusterin from oxidative stress-induced apoptosis in human retinal pigment epithelial cells.* Investigative Ophthalmology and Visual Science, 2010. **51**(1): p. 561-566.

134. Borderie, VM, Baudrimont, M, Vallee, A, et al., *Corneal endothelial cell apoptosis in patients with Fuchs' dystrophy.* Investigative Ophthalmology and Visual Science, 2000. **41**(9): p. 2501-2505.

135. Fuxe, J, Vincent, T, and Garcia de Herreros, A, *Transcriptional crosstalk between TGF-beta and stem cell pathways in tumor cell invasion: role of EMT promoting Smad complexes.* Cell Cycle, 2010. **9**(12): p. 2363-2374.

136. Zaniolo, K, Bostan, C, Rochette Drouin, O, et al., *Culture of human corneal endothelial cells isolated from corneas with Fuchs endothelial corneal dystrophy.* Experimental Eye Research, 2012. **94**(1): p. 22-31.

137. Buddi, R, Lin, B, Atilano, SR, et al., *Evidence of oxidative stress in human corneal diseases.* Journal of Histochemistry and Cytochemistry, 2002. **50**(3): p. 341-351.

138. Borboli, S and Colby, K, *Mechanisms of disease: Fuchs' endothelial dystrophy.* Ophthalmology clinics of North America, 2002. **15**(1): p. 17-25.

139. Zhang, C, Bell, WR, Sundin, OH, et al., *Immunohistochemistry and electron microscopy of early-onset fuchs corneal dystrophy in three cases with the same L450W COL8A2 mutation.* Transactions of the American Ophthalmological Society, 2006. **104**: p. 85-97.

140. Czarny, P, Kasprzak, E, Wielgorski, M, et al., *DNA damage and repair in Fuchs endothelial corneal dystrophy.* Molecular Biology Reports, 2013. **40**(4): p. 2977-2983.

141. Czarny, P, Seda, A, Wielgorski, M, et al., *Mutagenesis of mitochondrial DNA in Fuchs endothelial corneal dystrophy.* Mutation Research, 2014. **760**: p. 42-47.

142. Klenkler, B and Sheardown, H, *Growth factors in the anterior segment: role in tissue maintenance, wound healing and ocular pathology.* Experimental Eye Research, 2004. **79**(5): p. 677-688.

143. Maatta, M, Tervahartiala, T, Harju, M, et al., *Matrix metalloproteinases and their tissue inhibitors in aqueous humor of patients with primary open-angle glaucoma, exfoliation syndrome, and exfoliation glaucoma.* Journal of Glaucoma, 2005. **14**(1): p. 64-69.

144. Bramsen, T and Stenbjerg, S, *Fibrinolytic factors in aqueous humour and serum from patients with Fuchs' dystrophy and patients with cataract.* Acta Ophthalmologica, 1979. **57**(3): p. 470-476.

145. Wilson, SE, Bourne, WM, Maguire, LJ, et al., *Aqueous humor composition in Fuchs' dystrophy.* Investigative Ophthalmology and Visual Science, 1989. **30**(3): p. 449-453.

146. Richardson, MR, Segu, ZM, Price, MO, et al., *Alterations in the aqueous humor proteome in patients with Fuchs endothelial corneal dystrophy.* Molecular Vision, 2010. **16**: p. 2376-2383.

147. Bramsen, T and Ehlers, N, *Bullous keratopathy (Fuchs' endothelial dystrophy) treated systemically with 4-trans-amino-cyclohexano-carboxylic acid.* Acta Ophthalmologica, 1977. **55**(4): p. 665-673.

148. Tuberville, AW and Wood, TO, *Aqueous humor protein and complement in pseudophakic eyes.* Cornea, 1990. **9**(3): p. 249-253.

149. Fust, A, Csuka, D, Imre, L, et al., *The role of complement activation in the pathogenesis of Fuchs' dystrophy.* Molecular Immunology, 2014. **58**(2): p. 177-181.

150. Olsen, T, *Is there an association between Fuchs' endothelial dystrophy and cardiovascular disease?* Graefes Archive for Clinical and Experimental Ophthalmology, 1984. **221**(5): p. 239-240.

151. Stehouwer, M, Bijlsma, WR, and Van der Lelij, A, *Hearing disability in patients with Fuchs' endothelial corneal dystrophy: unrecognized co-pathology?* Clinical Ophthalmology (Auckland, N.Z.), 2011. **5**: p. 1297-1301.

152. Jun, AS, Meng, H, Ramanan, N, et al., *An alpha 2 collagen VIII transgenic knock-in mouse model of Fuchs endothelial corneal dystrophy shows early endothelial cell unfolded protein response and apoptosis.* Human Molecular Genetics, 2012. **21**(2): p. 384-393.

153. Meng, H, Matthaei, M, Ramanan, N, et al., *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, 2013. **54**(3): p. 1887-1897.

154. Matthaei, M, Hu, J, Meng, H, et al., *Endothelial cell whole genome expression analysis in a mouse model of early-onset Fuchs' endothelial corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2013. **54**(3): p. 1931-1940.

155. Jumblatt, MM and Willer, SS, *Corneal endothelial repair. Regulation of prostaglandin E2 synthesis.* Investigative Ophthalmology and Visual Science, 1996. **37**(7): p. 1294-1301.

156. Jumblatt, MM, *Autocrine regulation of corneal endothelium by prostaglandin E2.* Investigative Ophthalmology and Visual Science, 1994. **35**(6): p. 2783-2790.

157. Jumblatt, MM and Paterson, CA, *Prostaglandin E2 effects on corneal endothelial cyclic adenosine monophosphate synthesis and cell shape are mediated by a receptor of the EP2 subtype.* Investigative Ophthalmology and Visual Science, 1991. **32**(2): p. 360-365.

158. Shaulian, E and Karin, M, *AP-1 as a regulator of cell life and death.* Nature Cell Biology, 2002. **4**(5): p. E131-136.

159. Meng, Q and Xia, Y, *c-Jun, at the crossroad of the signaling network.* Protein Cell, 2011. **2**(11): p. 889-898.

160. Hopfer, U, Fukai, N, Hopfer, H, et al., *Targeted disruption of Col8a1 and Col8a2 genes in mice leads to anterior segment abnormalities in the eye.* FASEB Journal, 2005. **19**(10): p. 1232-1244.

161. Liu, Y, Peng, X, Tan, J, et al., *Zeb1 mutant mice as a model of posterior corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2008. **49**(5): p. 1843-1849.

162. Liskova, P, Tuft, SJ, Gwilliam, R, et al., *Novel mutations in the ZEB1 gene identified in Czech and British patients with posterior polymorphous corneal dystrophy.* Human Mutation, 2007. **28**(6): p. 638.

163. Mehta, JS, Vithana, EN, Tan, DT, et al., *Analysis of the posterior polymorphous corneal dystrophy 3 gene, TCF8, in late-onset Fuchs endothelial corneal dystrophy.* Investigative Ophthalmology and Visual Science, 2008. **49**(1): p. 184-188.

164. Lopez, IA, Rosenblatt, MI, Kim, C, et al., *Slc4a11 Gene Disruption in Mice: cellular targets of sensorineuronal abnormalities.* Journal of Biological Chemistry, 2009. **284**(39): p. 26882-26896.

165. Tuft, SJ, Williams, KA, and Coster, DJ, *Endothelial repair in the rat cornea.* Investigative Ophthalmology and Visual Science, 1986. **27**(8): p. 1199-1204.

166. Dimasi, DP, Burdon, KP, Hewitt, AW, et al., *Genetic investigation into the endophenotypic status of central corneal thickness and optic disc parameters in relation to open-angle glaucoma.* American Journal of Ophthalmology, 2012. **154**(5): p. 833-842.e832.

167. Awadalla, MS, Burdon, KP, Souzeau, E, et al., *Mutation in TMEM98 in a large white kindred with autosomal dominant nanophthalmos linked to 17p12-q12.* JAMA ophthalmology, 2014. **132**(8): p. 970-977.

168. Burdon, KP, Macgregor, S, Bykhovskaya, Y, et al., *Association of polymorphisms in the hepatocyte growth factor gene promoter with keratoconus.* Investigative Ophthalmology and Visual Science, 2011. **52**(11): p. 8514-8519.

169. Mitchell, P, Smith, W, Attebo, K, et al., *Prevalence of age-related maculopathy in Australia. The Blue Mountains Eye Study.* Ophthalmology, 1995. **102**(10): p. 1450-1460.

170. Mitchell, P, Smith, W, Attebo, K, et al., *Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study.* Ophthalmology, 1996. **103**(10): p. 1661-1669.

171. Barratt, BJ, Payne, F, Rance, HE, et al., *Identification of the sources of error in allele frequency estimations from pooled DNA indicates an optimal*

experimental design. Annals of Human Genetics, 2002. **66**(Pt 5-6): p. 393-405.

172. Sham, P, Bader, JS, Craig, I, et al., *DNA Pooling: a tool for large-scale association studies.* Nature reviews. Genetics, 2002. **3**(11): p. 862-871.

173. Purcell, S, Neale, B, Todd-Brown, K, et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* The American Journal of Human Genetics, 2007. **81**(3): p. 559-575.

174. Simon, P, *Q-Gene: processing quantitative real-time RT–PCR data.* Bioinformatics, 2003. **19**(11): p. 1439-1440.

175. Muller, PY, Janovjak, H, Miserez, AR, et al., *Processing of gene expression data generated by quantitative real-time RT-PCR.* Biotechniques, 2002. **32**(6): p. 1372-1374, 1376, 1378-1379.

176. Laemmli, UK, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-685.

177. Hamill, CE, Schmedt, T, and Jurkunas, U, *Fuchs endothelial cornea dystrophy: a review of the genetics behind disease development.* Seminars in Ophthalmology, 2013. **28**(5-6): p. 281-286.

178. Vithana, EN, Morgan, P, Sundaresan, P, et al., *Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2).* Nature Genetics, 2006. **38**(7): p. 755-757.

179. Lunter, G, *Probabilistic whole-genome alignments reveal high indel rates in the human and mouse genomes.* Bioinformatics, 2007. **23**(13): p. i289-296.

180. Wu, S, Romfo, CM, Nilsen, TW, et al., *Functional recognition of the 3[prime] splice site AG by the splicing factor U2AF35.* Nature, 1999. **402**(6763): p. 832-835.

181. Lamond, AI, Konarska, MM, and Sharp, PA, *A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation.* Genes and Development, 1987. **1**(6): p. 532-543.

182. Reed, R, *The organization of 3' splice-site sequences in mammalian introns.* Genes and Development, 1989. **3**(12b): p. 2113-2123.

183. Reed, R, *Mechanisms of fidelity in pre-mRNA splicing.* Current Opinion in Cell Biology, 2000. **12**(3): p. 340-345.

184. Li, Q, Wojciechowski, R, Simpson, CL, et al., *Genome-wide association study for refractive astigmatism reveals genetic co-determination with spherical equivalent refractive error: the CREAM consortium.* Human Genetics, 2015. **134**(2): p. 131-146.

185. Osman, W, Low, SK, Takahashi, A, et al., *A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma.* Human Molecular Genetics, 2012. **21**(12): p. 2836-2842.

186. Rahmioglu, N, Nyholt, DR, Morris, AP, et al., *Genetic variants underlying risk of endometriosis: insights from meta-analysis of eight genome-wide association and replication datasets.* Human Reproduction Update, 2014. **20**(5): p. 702-716.

187. Spencer, CCA, Su, Z, Donnelly, P, et al., *Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip.* PLoS Genetics, 2009. **5**(5): p. e1000477.

188. Lu, Y, Dimasi, DP, Hysi, PG, et al., *Common genetic variants near the Brittle Cornea Syndrome locus ZNF469 influence the blinding disease risk factor central corneal thickness.* PLoS Genetics, 2010. **6**(5): p. e1000947.

189. Brown, KM, Macgregor, S, Montgomery, GW, et al., *Common sequence variants on 20q11.22 confer melanoma susceptibility.* Nature Genetics, 2008. **40**(7): p. 838-840.

190. Macgregor, S, Zhao, ZZ, Henders, A, et al., *Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays.* Nucleic Acids Research, 2008. **36**(6): p. e35.

191. Macgregor, S, Visscher, PM, and Montgomery, G, *Analysis of pooled DNA samples on high density arrays without prior knowledge of differential hybridization rates.* Nucleic Acids Research, 2006. **34**(7): p. e55.

192. Kuot, A, *Fuchs' Endothelial Dystrophy: Molecular Genetics and Proteomics Analyses*, in *Department of Ophthalmology*. 2010, Flinders University of South Australia. p. 62.

193. Rose, AM and Valdes, R, Jr., *Understanding the sodium pump and its relevance to disease.* Clinical Chemistry, 1994. **40**(9): p. 1674-1685.

194. Suhail, M, *Na, K-ATPase: Ubiquitous Multifunctional Transmembrane Protein and its Relevance to Various Pathophysiological Conditions.* Journal of Clinical Medicine Research, 2010. **2**(1): p. 1-17.

195. Ljubimov, AV, Bartek, J, Couchman, JR, et al., *Distribution of individual components of basement membrane in human colon polyps and adenocarcinomas as revealed by monoclonal antibodies.* International Journal of Cancer, 1992. **50**(4): p. 562-566.

196. Ljubimov, AV, Burgeson, RE, Butkowski, RJ, et al., *Human corneal basement membrane heterogeneity: topographical differences in the expression of type IV collagen and laminin isoforms.* Laboratory Investigation, 1995. **72**(4): p. 461-473.

197. Champliaud, MF, Virtanen, I, Tiger, CF, et al., *Posttranslational modifications and beta/gamma chain associations of human laminin alpha1 and laminin alpha5 chains: purification of laminin-3 from placenta.* Experimental Cell Research, 2000. **259**(2): p. 326-335.

198. Sanchez-Tillo, E, Lazaro, A, Torrent, R, et al., *ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1*. Oncogene, 2010. **29**(24): p. 3490-3500.

199. Terraz, C, Toman, D, Delauche, M, et al., *delta Ef1 binds to a far upstream sequence of the mouse pro-alpha 1(I) collagen gene and represses its expression in osteoblasts.* Journal of Biological Chemistry, 2001. **276**(40): p. 37011-37019.

200. Frisch, SM, *E1a induces the expression of epithelial characteristics.* Journal of Cell Biology, 1994. **127**(4): p. 1085-1096.

201. Rosenbloom, KR, Dreszer, TR, Pheasant, M, et al., *ENCODE whole-genome data in the UCSC Genome Browser*. Nucleic Acids Research, 2010. **38**(Database issue): p. D620-625.

202. Phelan, ML, Sif, S, Narlikar, GJ, et al., *Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits.* Molecular Cell, 1999. **3**(2): p. 247-253.

203. Kwon, CS and Wagner, D, *Unwinding chromatin for development and growth: a few genes at a time.* Trends in Genetics, 2007. **23**(8): p. 403-412.

204. Matthaei, M, Hu, J, Kallay, L, et al., *Endothelial cell microRNA expression in human late-onset Fuchs' dystrophy.* Investigative Ophthalmology and Visual Science, 2014. **55**(1): p. 216-225.

205. Grant, DS and Leblond, CP, *Immunogold quantitation of laminin, type IV collagen, and heparan sulfate proteoglycan in a variety of basement membranes.* Journal of Histochemistry and Cytochemistry, 1988. **36**(3): p. 271-283.

206. Kenney, MC, Labermeier, U, Hinds, D, et al., *Characterization of the Descemet's membrane/posterior collagenous layer isolated from Fuchs' endothelial dystrophy corneas.* Experimental Eye Research, 1984. **39**(3): p. 267-277.

207. Barros, CS, Franco, SJ, and Muller, U, *Extracellular matrix: functions in the nervous system.* Cold Spring Harbor Perspectives in Biology, 2011. **3**(1): p. a005108.

208. Yamaguchi, M, Ebihara, N, Shima, N, et al., *Adhesion, migration, and proliferation of cultured human corneal endothelial cells by laminin-5.* Investigative Ophthalmology and Visual Science, 2011. **52**(2): p. 679-684.

209. Murray, P and Edgar, D, *Regulation of programmed cell death by basement membranes in embryonic development.* Journal of Cell Biology, 2000. **150**(5): p. 1215-1221.

210. Smyth, N, Vatansever, HS, Murray, P, et al., *Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation.* Journal of Cell Biology, 1999. **144**(1): p. 151-160.

211. Nde, PN, Simmons, KJ, Kleshchenko, YY, et al., *Silencing of the laminin gamma-1 gene blocks Trypanosoma cruzi infection.* Infection and Immunity, 2006. **74**(3): p. 1643-1648.

212. Barwe, SP, Kim, S, Rajasekaran, SA, et al., *Janus model of the Na,K-ATPase beta-subunit transmembrane domain: distinct faces mediate alpha/beta assembly and beta-beta homo-oligomerization.* Journal of Molecular Biology, 2007. **365**(3): p. 706-714.

213. Shoshani, L, Contreras, RG, Roldan, ML, et al., *The polarized expression of Na+,K+-ATPase in epithelia depends on the association between beta-subunits located in neighboring cells.* Molecular Biology of the Cell, 2005. **16**(3): p. 1071-1081.

214. Vagin, O, Tokhtaeva, E, and Sachs, G, *The role of the beta1 subunit of the Na,K-ATPase and its glycosylation in cell-cell adhesion.* Journal of Biological Chemistry, 2006. **281**(51): p. 39573-39587.

215. Barwe, SP, Skay, A, McSpadden, R, et al., *Na,K-ATPase beta-subunit cis homo-oligomerization is necessary for epithelial lumen formation in mammalian cells.* Journal of Cell Science, 2012. **125**(Pt 23): p. 5711-5720.

216. Schmedt, T, Silva, MM, Ziaei, A, et al., *Molecular bases of corneal endothelial dystrophies.* Experimental Eye Research, 2012. **95**(1): p. 24-34.

217. Zhu, Y, Kakinuma, N, Wang, Y, et al., *Kank proteins: a new family of ankyrin-repeat domain-containing proteins.* Biochimica et Biophysica Acta, 2008. **1780**(2): p. 128-133.

218. Kakinuma, N, Zhu, Y, Wang, Y, et al., *Kank proteins: structure, functions and diseases.* Cellular and Molecular Life Sciences, 2009. **66**(16): p. 2651-2659.

219. Kakinuma, N, Roy, BC, Zhu, Y, et al., *Kank regulates RhoA-dependent formation of actin stress fibers and cell migration via 14-3-3 in PI3K-Akt signaling.* Journal of Cell Biology, 2008. **181**(3): p. 537-549.

220. Tojkander, S, Gateva, G, and Lappalainen, P, *Actin stress fibers--assembly, dynamics and biological roles.* Journal of Cell Science, 2012. **125**(Pt 8): p. 1855-1864.

221. Kozlowski, P, de Mezer, M, and Krzyzosiak, WJ, *Trinucleotide repeats in human genome and exome.* Nucleic Acids Research, 2010. **38**(12): p. 4027-4039.

222. Cagnoli, C, Michielotto, C, Matsuura, T, et al., *Detection of large pathogenic expansions in FRDA1, SCA10, and SCA12 genes using a simple fluorescent repeat-primed PCR assay.* The Journal of molecular diagnostics : JMD, 2004. **6**(2): p. 96-100.

223. Pumpernik, D, Oblak, B, and Borstnik, B, *Replication slippage versus point mutation rates in short tandem repeats of the human genome.* Molecular genetics and genomics : MGG, 2008. **279**(1): p. 53-61.

224. Toth, G, Gaspari, Z, and Jurka, J, *Microsatellites in different eukaryotic genomes: survey and analysis.* Genome Research, 2000. **10**(7): p. 967-981.

225. Kelkar, YD, Tyekucheva, S, Chiaromonte, F, et al., *The genome-wide determinants of human and chimpanzee microsatellite evolution.* Genome Research, 2008. **18**(1): p. 30-38.

226. Madsen, BE, Villesen, P, and Wiuf, C, *Short tandem repeats in human exons: a target for disease mutations.* BMC Genomics, 2008. **9**: p. 410.

227. Borstnik, B and Pumpernik, D, *Tandem repeats in protein coding regions of primate genes.* Genome Research, 2002. **12**(6): p. 909-915.

228. Weissenbach, J, Gyapay, G, Dib, C, et al., *A second-generation linkage map of the human genome.* Nature, 1992. **359**(6398): p. 794-801.

229. Ellegren, H, *Microsatellites: simple sequences with complex evolution.* Nature reviews. Genetics, 2004. **5**(6): p. 435-445.

230. Mirkin, SM, *Expandable DNA repeats and human disease.* Nature, 2007. **447**(7147): p. 932-940.

231. Pearson, CE, Nichol Edamura, K, and Cleary, JD, *Repeat instability: mechanisms of dynamic mutations.* Nature reviews. Genetics, 2005. **6**(10): p. 729-742.

232. Gattey, D, Zhu, AY, Stagner, A, et al., *Fuchs endothelial corneal dystrophy in patients with myotonic dystrophy: a case series.* Cornea, 2014. **33**(1): p. 96-98.

233. Brusco, A, Gellera, C, Cagnoli, C, et al., *Molecular genetics of hereditary spinocerebellar ataxia: mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families.* Archives of Neurology, 2004. **61**(5): p. 727-733.

234. Nelson, DL, Orr, HT, and Warren, ST, *The Unstable Repeats - Three Evolving Faces of Neurological Disease.* Neuron, 2013. **77**(5): p. 825-843.

235. Forrest, MP, Hill, MJ, Quantock, AJ, et al., *The emerging roles of TCF4 in disease and development.* Trends in Molecular Medicine, 2014. **20**(6): p. 322-331.

236. Santos, R, Lefevre, S, Sliwa, D, et al., *Friedreich ataxia: molecular mechanisms, redox considerations, and therapeutic opportunities.* Antioxidants and redox signaling 2010. **13**(5): p. 651-690.

237. Echeverria, GV and Cooper, TA, *RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity.* Brain Research, 2012. **1462**: p. 100-111.

238. Miller, JW, Urbinati, CR, Teng-Umnuay, P, et al., *Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy.* EMBO Journal, 2000. **19**(17): p. 4439-4448.

239. Fardaei, M, Rogers, MT, Thorpe, HM, et al., *Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells.* Human Molecular Genetics, 2002. **11**(7): p. 805-814.

240. Mankodi, A, Urbinati, CR, Yuan, QP, et al., *Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2.* Human Molecular Genetics, 2001. **10**(19): p. 2165-2170.

241. Konieczny, P, Stepniak-Konieczna, E, and Sobczak, K, *MBNL proteins and their target RNAs, interaction and splicing regulation.* Nucleic Acids Research, 2014. **42**(17): p. 10873-10887.

242. O'Rourke, JR and Swanson, MS, *Mechanisms of RNA-mediated disease.* Journal of Biological Chemistry, 2009. **284**(12): p. 7419-7423.

243. Bourne, WM, Johnson, DH, and Campbell, RJ, *The Ultrastructure of Descemet's Membrane: III. Fuchs' Dystrophy.* Archives of Ophthalmology, 1982. **100**(12): p. 1952-1955.

244. Iwamoto, T and DeVoe, AG, *Electron microscopic studies on Fuchs'combined dystrophy. I. Posterior portion of the cornea.* Investigative Ophthalmology, 1971. **10**(1): p. 9-28.

245. Gottsch, JD, Seitzman, GD, Margulies, EH, et al., *Gene expression in donor corneal endothelium.* Archives of Ophthalmology, 2003. **121**(2): p. 252-258.

246. Pieroni, L, Finamore, F, Ronci, M, et al., *Proteomics investigation of human platelets in healthy donors and cystic fibrosis patients by shotgun nUPLC-*

MSE and 2DE: a comparative study. Molecular bioSystems, 2011. **7**(3): p. 630-639.

247. Vissers, JP, Langridge, JI, and Aerts, JM, *Analysis and quantification of diagnostic serum markers and protein signatures for Gaucher disease.* Molecular & Cellular Proteomics, 2007. **6**(5): p. 755-766.

248. Silva, JC, Denny, R, Dorschel, CA, et al., *Quantitative proteomic analysis by accurate mass retention time pairs.* Analytical Chemistry, 2005. **77**(7): p. 2187-2200.

249. Waring, GO, 3rd, *Posterior collagenous layer of the cornea. Ultrastructural classification of abnormal collagenous tissue posterior to Descemet's membrane in 30 cases.* Archives of Ophthalmology, 1982. **100**(1): p. 122-134.

250. Kim, WS, Kim, HJ, Lee, ZH, et al., *Apolipoprotein E inhibits osteoclast differentiation via regulation of c-Fos, NFATc1 and NF-kappaB.* Experimental Cell Research, 2013. **319**(4): p. 436-446.

251. Miyamoto-Sato, E, Fujimori, S, Ishizaka, M, et al., *A comprehensive resource of interacting protein regions for refining human transcription factor networks.* PloS One, 2010. **5**(2): p. e9289.

252. Jin, G and Howe, PH, *Transforming growth factor beta regulates clusterin gene expression via modulation of transcription factor c-Fos.* European Journal of Biochemistry, 1999. **263**(2): p. 534-542.

253. Johnston, IM, Spence, HJ, Winnie, JN, et al., *Regulation of a multigenic invasion programme by the transcription factor, AP-1: re-expression of a down-regulated gene, TSC-36, inhibits invasion.* Oncogene, 2000. **19**(47): p. 5348-5358.

254. Greenow, K, Pearce, NJ, and Ramji, DP, *The key role of apolipoprotein E in atherosclerosis.* Journal of Molecular Medicine (Berlin, Germany), 2005. **83**(5): p. 329-342.

255. Weisgraber, KH, *Apolipoprotein E: structure-function relationships.* Advances in Protein Chemistry, 1994. **45**: p. 249-302.

256. Wilson, C, Wardell, MR, Weisgraber, KH, et al., *Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E.* Science, 1991. **252**(5014): p. 1817-1822.

257. Weisgraber, KH, Rall, SC, Jr., Mahley, RW, et al., *Human apolipoprotein E. Determination of the heparin binding sites of apolipoprotein E3.* Journal of Biological Chemistry, 1986. **261**(5): p. 2068-2076.

258. Libeu, CP, Lund-Katz, S, Phillips, MC, et al., *New insights into the heparan sulfate proteoglycan-binding activity of apolipoprotein E.* Journal of Biological Chemistry, 2001. **276**(42): p. 39138-39144.

259. Sharma, S, Chataway, T, Burdon, KP, et al., *Identification of LOXL1 protein and Apolipoprotein E as components of surgically isolated pseudoexfoliation material by direct mass spectrometry.* Experimental Eye Research, 2009. **89**(4): p. 479-485.

260. Matus, S, Glimcher, LH, and Hetz, C, *Protein folding stress in neurodegenerative diseases: a glimpse into the ER.* Current Opinion in Cell Biology, 2011. **23**(2): p. 239-252.

261. Federico, A, Cardaioli, E, Da Pozzo, P, et al., *Mitochondria, oxidative stress and neurodegeneration.* Journal of the Neurological Sciences, 2012. **322**(1-2): p. 254-262.

262. Russo, G, Zegar, C, and Giordano, A, *Advantages and limitations of microarray technology in human cancer.* Oncogene, 2003. **22**(42): p. 6497-6507.

263. Wildsmith, SE and Elcock, FJ, *Microarrays under the microscope.* Molecular Pathology, 2001. **54**(1): p. 8-16.

264. Smyth, GK, *Linear models and empirical bayes methods for assessing differential expression in microarray experiments.* Statistical Applications in Genetics and Molecular Biology, 2004. **3**: p. Article3.

265. Du, P, Kibbe, WA, and Lin, SM, *lumi: a pipeline for processing Illumina microarray.* Bioinformatics, 2008. **24**(13): p. 1547-1548.

266. Smyth, GK, Michaud, J, and Scott, HS, *Use of within-array replicate spots for assessing differential expression in microarray experiments.* Bioinformatics, 2005. **21**(9): p. 2067-2075.

267. Bustin, SA, Benes, V, Garson, JA, et al., *The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.* Clinical Chemistry, 2009. **55**(4): p. 611-622.

268. Bustin, SA, *Why the need for qPCR publication guidelines?--The case for MIQE.* Methods, 2010. **50**(4): p. 217-226.

269. Yoshida, Y, Tsunoda, T, Doi, K, et al., *ALPK2 is crucial for luminal apoptosis and DNA repair-related gene expression in a three-dimensional colonic-crypt model.* Anticancer Research, 2012. **32**(6): p. 2301-2308.

270. Smith, TP, Haymond, T, Smith, SN, et al., *Evidence for the endothelin system as an emerging therapeutic target for the treatment of chronic pain.* Journal of Pain Research, 2014. **7**: p. 531-545.

271. Molet, S, Furukawa, K, Maghazechi, A, et al., *Chemokine- and cytokine-induced expression of endothelin 1 and endothelin-converting enzyme 1 in endothelial cells.* Journal of Allergy and Clinical Immunology, 2000. **105**(2 Pt 1): p. 333-338.

272. Cyr, CR, Rudy, B, and Kris, RM, *Prolonged desensitization of the human endothelin A receptor in Xenopus oocytes. Comparative studies with the human neurokinin A receptor.* Journal of Biological Chemistry, 1993. **268**(35): p. 26071-26074.

273. Li, ZF, Wu, XH, and Engvall, E, *Identification and characterization of CPAMD8, a novel member of the complement 3/alpha2-macroglobulin family with a C-terminal Kazal domain.* Genomics, 2004. **83**(6): p. 1083-1093.

274. Kenny, EE, Pe'er, I, Karban, A, et al., *A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci.* PLoS Genetics, 2012. **8**(3): p. e1002559.

275. Eckburg, PB and Relman, DA, *The role of microbes in Crohn's disease.* Clinical Infectious Diseases, 2007. **44**(2): p. 256-262.

276. Lapaquette, P, Brest, P, Hofman, P, et al., *Etiology of Crohn's disease: many roads lead to autophagy.* Journal of Molecular Medicine (Berlin, Germany), 2012. **90**(9): p. 987-996.

277. Chen, S and Birk, DE, *The regulatory roles of small leucine-rich proteoglycans in extracellular matrix assembly.* The FEBS journal, 2013. **280**(10): p. 2120-2137.

278. Couchman, JR and Pataki, CA, *An introduction to proteoglycans and their localization.* Journal of Histochemistry and Cytochemistry, 2012. **60**(12): p. 885-897.

279. Brandan, E, Cabello-Verrugio, C, and Vial, C, *Novel regulatory mechanisms for the proteoglycans decorin and biglycan during muscle formation and muscular dystrophy.* Matrix Biology, 2008. **27**(8): p. 700-708.

280. Schaefer, L, Babelova, A, Kiss, E, et al., *The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages.* Journal of Clinical Investigation, 2005. **115**(8): p. 2223-2233.

281. Babelova, A, Moreth, K, Tsalastra-Greul, W, et al., *Biglycan, a danger signal that activates the NLRP3 inflammasome via toll-like and P2X receptors.* Journal of Biological Chemistry, 2009. **284**(36): p. 24035-24048.

282. Frey, H, Schroeder, N, Manon-Jensen, T, et al., *Biological interplay between proteoglycans and their innate immune receptors in inflammation.* The FEBS journal, 2013. **280**(10): p. 2165-2179.

283. Nastase, MV, Young, MF, and Schaefer, L, *Biglycan: a multivalent proteoglycan providing structure and signals.* Journal of Histochemistry and Cytochemistry, 2012. **60**(12): p. 963-975.

284. Burch, ML, Yang, SN, Ballinger, ML, et al., *TGF-beta stimulates biglycan synthesis via p38 and ERK phosphorylation of the linker region of Smad2.* Cellular and Molecular Life Sciences, 2010. **67**(12): p. 2077-2090.

285. Ungefroren, H, Lenschow, W, Chen, WB, et al., *Regulation of biglycan gene expression by transforming growth factor-beta requires MKK6-p38 mitogen-activated protein Kinase signaling downstream of Smad signaling.* Journal of Biological Chemistry, 2003. **278**(13): p. 11041-11049.

286. Genovese, F, Barascuk, N, Larsen, L, et al., *Biglycan fragmentation in pathologies associated with extracellular matrix remodeling by matrix metalloproteinases.* Fibrogenesis Tissue Repair, 2013. **6**(1): p. 9.

287. De Palma, C, Perrotta, C, Pellegrino, P, et al., *Skeletal muscle homeostasis in duchenne muscular dystrophy: modulating autophagy as a promising therapeutic strategy*. Frontiers in Aging Neuroscience, 2014. **6**: p. 188.

288. Keppler, D, Zhang, J, Bihani, T, et al., *Novel expression of CST1 as candidate senescence marker*. Journals of Gerontology. Series A: Biological Sciences and Medical Sciences, 2011. **66**(7): p. 723-731.

289. Dickinson, DP, Zhao, Y, Thiesse, M, et al., *Direct mapping of seven genes encoding human type 2 cystatins to a single site located at 20p11.2.* Genomics, 1994. **24**(1): p. 172-175.

290. Henskens, YM, Veerman, EC, and Nieuw Amerongen, AV, *Cystatins in health and disease.* Biological Chemistry Hoppe-Seyler, 1996. **377**(2): p. 71-86.

291. Dickinson, DP, Thiesse, M, and Hicks, MJ, *Expression of type 2 cystatin genes CST1-CST5 in adult human tissues and the developing submandibular gland.* DNA and Cell Biology, 2002. **21**(1): p. 47-65.

292. Barrett, AJ, *The cystatins: a diverse superfamily of cysteine peptidase inhibitors.* Biomedica Biochimica Acta, 1986. **45**(11-12): p. 1363-1374.

293. Turk, V and Bode, W, *The cystatins: protein inhibitors of cysteine proteinases.* FEBS Letters, 1991. **285**(2): p. 213-219.

294. Koblinski, JE, Ahram, M, and Sloane, BF, *Unraveling the role of proteases in cancer.* Clinica Chimica Acta, 2000. **291**(2): p. 113-135.

295. Lah, TT, Babnik, J, Schiffmann, E, et al., *Cysteine proteinases and inhibitors in inflammation: their role in periodontal disease.* Journal of Periodontology, 1993. **64**(5 Suppl): p. 485-491.

296. Al-Hashimi, I, Dickinson, DP, and Levine, MJ, *Purification, molecular cloning, and sequencing of salivary cystatin SA-1.* Journal of Biological Chemistry, 1988. **263**(19): p. 9381-9387.

297. Baron, A, Barrett-Vespone, N, and Featherstone, J, *Purification of large quantities of human salivary cystatins S, SA and SN: their interactions with the model cysteine protease papain in a non-inhibitory mode.* Oral Diseases, 1999. **5**(4): p. 344-353.

298. Baron, A, DeCarlo, A, and Featherstone, J, *Functional aspects of the human salivary cystatins in the oral environment.* Oral Diseases, 1999. **5**(3): p. 234-240.

299. Hiltke, TR, Lee, TC, and Bobek, LA, *Structure/function analysis of human cystatin SN and comparison of the cysteine proteinase inhibitory profiles of human cystatins C and SN.* Journal of Dental Research, 1999. **78**(8): p. 1401-1409.

300. Tseng, CC, Tseng, CP, Levine, MJ, et al., *Differential effect toward inhibition of papain and cathepsin C by recombinant human salivary cystatin SN and its variants produced by a baculovirus system.* Archives of Biochemistry and Biophysics, 2000. **380**(1): p. 133-140.

301. Ivy, GO, Roopsingh, R, Kanai, S, et al., *Leupeptin causes an accumulation of lipofuscin-like substances and other signs of aging in kidneys of young rats: further evidence for the protease inhibitor model of aging.* Annals of the New York Academy of Sciences, 1996. **786**: p. 12-23.

302. Ivy, GO, Schottler, F, Wenzel, J, et al., *Inhibitors of lysosomal enzymes: accumulation of lipofuscin-like dense bodies in the brain.* Science, 1984. **226**(4677): p. 985-987.

303. Terman, A and Brunk, UT, *Aging as a catabolic malfunction.* International Journal of Biochemistry and Cell Biology, 2004. **36**(12): p. 2365-2375.

304. Dodson, M, Darley-Usmar, V, and Zhang, J, *Cellular metabolic and autophagic pathways: traffic control by redox signaling.* Free Radical Biology and Medicine, 2013. **63**: p. 207-221.

305. Griffon, N, Jeanneteau, F, Prieur, F, et al., *CLIC6, a member of the intracellular chloride channel family, interacts with dopamine D(2)-like receptors.* Brain Research: Molecular Brain Research, 2003. **117**(1): p. 47-57.

306. Jentsch, TJ, *Chloride and the endosomal-lysosomal pathway: emerging roles of CLC chloride transporters.* Journal of Physiology, 2007. **578**(Pt 3): p. 633-640.

307. Li, QJ, Ashraf, MF, Shen, DF, et al., *The role of apoptosis in the pathogenesis of Fuchs endothelial dystrophy of the cornea.* Archives of Ophthalmology, 2001. **119**(11): p. 1597-1604.

308. Dunning, MJ, Barbosa-Morais, NL, Lynch, AG, et al., *Statistical issues in the analysis of Illumina data.* BMC Bioinformatics, 2008. **9**: p. 85.

309. Matthaei, M, Zhu, AY, Kallay, L, et al., *Transcript profile of cellular senescence-related genes in Fuchs endothelial corneal dystrophy.* Experimental Eye Research, 2014. **129**: p. 13-17.

310. Wieben, ED, Aleff, RA, Eckloff, BW, et al., *Comprehensive Assessment of Genetic Variants Within TCF4 in Fuchs' Endothelial Corneal Dystrophy.* Investigative Ophthalmology and Visual Science, 2014. **55**(9): p. 6101-6107.

311. Jalimarada, SS, Ogando, DG, and Bonanno, JA, *Loss of ion transporters and increased unfolded protein response in Fuchs' dystrophy.* Molecular Vision, 2014. **20**: p. 1668-1679.

312. Wulle, KG, *Electron microscopy of the fetal development of the corneal endothelium and Descemet's membrane of the human eye.* Investigative Ophthalmology, 1972. **11**(11): p. 897-904.

313. Murphy, C, Alvarado, J, Juster, R, et al., *Prenatal and postnatal cellularity of the human corneal endothelium. A quantitative histologic study.* Investigative Ophthalmology and Visual Science, 1984. **25**(3): p. 312-322.

314. Mankad, R, *Atherosclerotic vascular disease in the autoimmune rheumatologic patient.* Current atherosclerosis reports, 2015. **17**(4): p. 497.

315. Hartman, J and Frishman, WH, *Inflammation and atherosclerosis: a review of the role of interleukin-6 in the development of atherosclerosis and the potential for targeted drug therapy.* Cardiology in Review, 2014. **22**(3): p. 147-151.

316. Bekkering, S, Joosten, LA, van der Meer, JW, et al., *The epigenetic memory of monocytes and macrophages as a novel drug target in atherosclerosis.* Clinical Therapeutics, 2015. **37**(4): p. 914-923.

317. Duval, F, Moreno-Cuevas, JE, Gonzalez-Garza, MT, et al., *Protective mechanisms of medicinal plants targeting hepatic stellate cell activation and extracellular matrix deposition in liver fibrosis.* Chinese Medicine, 2014. **9**(1): p. 27.

318. Duval, F, Moreno-Cuevas, JE, Gonzalez-Garza, MT, et al., *Liver fibrosis and protection mechanisms action of medicinal plants targeting apoptosis of hepatocytes and hepatic stellate cells.* Advances in Pharmacological Sciences, 2014. **2014**: p. 373295.

319. Darabid, H, Perez-Gonzalez, AP, and Robitaille, R, *Neuromuscular synaptogenesis: coordinating partners with multiple functions.* Nature Reviews: Neuroscience, 2014. **15**(11): p. 703-718.

320. Ferraro, E, Molinari, F, and Berghella, L, *Molecular control of neuromuscular junction development.* Journal of cachexia, sarcopenia and muscle, 2012. **3**(1): p. 13-23.

321. Tezuka, T, Inoue, A, Hoshi, T, et al., *The MuSK activator agrin has a separate role essential for postnatal maintenance of neuromuscular synapses.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(46): p. 16556-16561.

322. Witzemann, V, *Development of the neuromuscular junction.* Cell and Tissue Research, 2006. **326**(2): p. 263-271.

323. Sagoo, P, Chan, G, Larkin, DF, et al., *Inflammatory cytokines induce apoptosis of corneal endothelium through nitric oxide.* Investigative Ophthalmology and Visual Science, 2004. **45**(11): p. 3964-3973.

324. Koizumi, N, Okumura, N, and Kinoshita, S, *Development of new therapeutic modalities for corneal endothelial disease focused on the proliferation of corneal endothelial cells using animal models.* Experimental Eye Research, 2012. **95**(1): p. 60-67.

325. Koizumi, N, Okumura, N, Ueno, M, et al., *Rho-associated kinase inhibitor eye drop treatment as a possible medical treatment for Fuchs corneal dystrophy.* Cornea, 2013. **32**(8): p. 1167-1170.

326. Koizumi, N, Okumura, N, Ueno, M, et al., *New therapeutic modality for corneal endothelial disease using Rho-associated kinase inhibitor eye drops.* Cornea, 2014. **33 Suppl 11**: p. S25-31.

327. Sanchez-Tillo, E, de Barrios, O, Siles, L, et al., *beta-catenin/TCF4 complex induces the epithelial-to-mesenchymal transition (EMT)-activator ZEB1 to regulate tumor invasiveness.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(48): p. 19204-19209.

328. Kim, TH, Xiong, H, Zhang, Z, et al., *[beta]-Catenin activates the growth factor endothelin-1 in colon cancer cells.* Oncogene, 2004. **24**(4): p. 597-604.

329. Sen, M, *Wnt signalling in rheumatoid arthritis.* Rheumatology, 2005. **44**(6): p. 708-713.

330. Carter, CJ, *eIF2B and oligodendrocyte survival: where nature and nurture meet in bipolar disorder and schizophrenia?* Schizophrenia Bulletin, 2007. **33**(6): p. 1343-1353.

331. Verzi, MP, Hatzis, P, Sulahian, R, et al., *TCF4 and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(34): p. 15157-15162.

332. Herbst, A, Jurinovic, V, Krebs, S, et al., *Comprehensive analysis of beta-catenin target genes in colorectal carcinoma cell lines with deregulated Wnt/beta-catenin signaling.* BMC Genomics, 2014. **15**: p. 74.

333. Kolligs, FT, Nieman, MT, Winer, I, et al., *ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with beta-catenin defects and promotes neoplastic transformation.* Cancer Cell, 2002. **1**(2): p. 145-155.

334. Kolligs, FT, Bommer, G, and Goke, B, *Wnt/beta-catenin/tcf signaling: a critical pathway in gastrointestinal tumorigenesis.* Digestion, 2002. **66**(3): p. 131-144.

335. Chen, J and Sadowski, I, *Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(13): p. 4813-4818.

336. Crighton, D, Wilkinson, S, and Ryan, KM, *DRAM links autophagy to p53 and programmed cell death.* Autophagy, 2007. **3**(1): p. 72-74.

337. Balaburski, GM, Hontz, RD, and Murphy, ME, *p53 and ARF: unexpected players in autophagy.* Trends in Cell Biology, 2010. **20**(6): p. 363-369.

338. Kerley-Hamilton, JS, Pike, AM, Li, N, et al., *A p53-dominant transcriptional response to cisplatin in testicular germ cell tumor-derived human embyronal carcinoma.* Oncogene, 2005. **24**(40): p. 6090-6100.

339. Laforge, M, Limou, S, Harper, F, et al., *DRAM triggers lysosomal membrane permeabilization and cell death in CD4(+) T cells infected with HIV.* PLoS Pathogens, 2013. **9**(5): p. e1003328.

340. Burrows, AE, Smogorzewska, A, and Elledge, SJ, *Polybromo-associated BRG1-associated factor components BRD7 and BAF180 are critical regulators of p53 required for induction of replicative senescence.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(32): p. 14280-14285.

341. Muller, PA, Trinidad, AG, Caswell, PT, et al., *Mutant p53 regulates Dicer through p63-dependent and -independent mechanisms to promote an invasive phenotype.* Journal of Biological Chemistry, 2014. **289**(1): p. 122-132.

342. Burnum, KE, Hirota, Y, Baker, ES, et al., *Uterine deletion of Trp53 compromises antioxidant responses in the mouse decidua.* Endocrinology, 2012. **153**(9): p. 4568-4579.

343. Galigniana, MD, Harrell, JM, O'Hagen, HM, et al., *Hsp90-binding immunophilins link p53 to dynein during p53 transport to the nucleus.* Journal of Biological Chemistry, 2004. **279**(21): p. 22483-22489.

344. Sayeed, A, Konduri, SD, Liu, W, et al., *Estrogen receptor alpha inhibits p53-mediated transcriptional repression: implications for the regulation of apoptosis.* Cancer Research, 2007. **67**(16): p. 7746-7755.

345. Qin, C, Nguyen, T, Stewart, J, et al., *Estrogen up-regulation of p53 gene expression in MCF-7 breast cancer cells is mediated by calmodulin kinase IV-dependent activation of a nuclear factor kappaB/CCAAT-binding transcription factor-1 complex.* Molecular Endocrinology, 2002. **16**(8): p. 1793-1809.

346. Berger, CE, Qian, Y, Liu, G, et al., *p53, a target of estrogen receptor (ER) alpha, modulates DNA damage-induced growth suppression in ER-positive breast cancer cells.* Journal of Biological Chemistry, 2012. **287**(36): p. 30117-30127.

347. Elangovan, S, Ramachandran, S, Venkatesan, N, et al., *SIRT1 is essential for oncogenic signaling by estrogen/estrogen receptor alpha in breast cancer.* Cancer Research, 2011. **71**(21): p. 6654-6664.

348. Konduri, SD, Medisetty, R, Liu, W, et al., *Mechanisms of estrogen receptor antagonism toward p53 and its implications in breast cancer therapeutic response and stem cell regulation.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(34): p. 15081-15086.

349. Shirley, SH, Rundhaug, JE, Tian, J, et al., *Transcriptional regulation of estrogen receptor-alpha by p53 in human breast cancer cells.* Cancer Research, 2009. **69**(8): p. 3405-3414.

350. Liu, W, Konduri, SD, Bansal, S, et al., *Estrogen receptor-alpha binds p53 tumor suppressor protein directly and represses its function.* Journal of Biological Chemistry, 2006. **281**(15): p. 9837-9840.

351. Kim, E, Gunther, W, Yoshizato, K, et al., *Tumor suppressor p53 inhibits transcriptional activation of invasion gene thromboxane synthase mediated by the proto-oncogenic factor ets-1*. Oncogene, 2003. **22**(49): p. 7716-7727.

352. Do, PM, Varanasi, L, Fan, S, et al., *Mutant p53 cooperates with ETS2 to promote etoposide resistance.* Genes and Development, 2012. **26**(8): p. 830-845.

353. Scian, MJ, Stagliano, KE, Deb, D, et al., *Tumor-derived p53 mutants induce oncogenesis by transactivating growth-promoting genes.* Oncogene, 2004. **23**(25): p. 4430-4443.

354. Lin, SC, Lee, KF, Nikitin, AY, et al., *Somatic mutation of p53 leads to estrogen receptor alpha-positive and -negative mouse mammary tumors with high frequency of metastasis.* Cancer Research, 2004. **64**(10): p. 3525-3532.

355. Grandori, C, Cowley, SM, James, LP, et al., *The Myc/Max/Mad network and the transcriptional control of cell behavior*. Annual Review of Cell and Developmental Biology, 2000. **16**: p. 653-699.

356. Chen, D, Kon, N, Zhong, J, et al., *Differential effects on ARF stability by normal versus oncogenic levels of c-Myc expression.* Molecular Cell, 2013. **51**(1): p. 46-56.

357. Ceballos, E, Delgado, MD, Gutierrez, P, et al., *c-Myc antagonizes the effect of p53 on apoptosis and p21WAF1 transactivation in K562 leukemia cells.* Oncogene, 2000. **19**(18): p. 2194-2204.

358. Tanikawa, J, Nomura, T, Macmillan, EM, et al., *p53 suppresses c-Myb-induced trans-activation and transformation by recruiting the corepressor mSin3A.* Journal of Biological Chemistry, 2004. **279**(53): p. 55393-55400.

359. Suh, SS, Yoo, JY, Nuovo, GJ, et al., *MicroRNAs/TP53 feedback circuitry in glioblastoma multiforme.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(14): p. 5316-5321.

360. Cui, D, Li, L, Lou, H, et al., *The ribosomal protein S26 regulates p53 activity in response to DNA damage.* Oncogene, 2014. **33**(17): p. 2225-2235.

361. Fu, X, Yucer, N, Liu, S, et al., *RFWD3-Mdm2 ubiquitin ligase complex positively regulates p53 stability in response to DNA damage.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(10): p. 4579-4584.

362. Emanuele, MJ, Elia, AE, Xu, Q, et al., *Global identification of modular cullin-RING ligase substrates.* Cell, 2011. **147**(2): p. 459-474.

363. Walseng, E, Furuta, K, Bosch, B, et al., *Ubiquitination regulates MHC class II-peptide complex retention and degradation in dendritic cells.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(47): p. 20465-20470.

364. Kim, W, Bennett, EJ, Huttlin, EL, et al., *Systematic and quantitative assessment of the ubiquitin-modified proteome.* Molecular Cell, 2011. **44**(2): p. 325-340.

365. Shi, Y, Chan, DW, Jung, SY, et al., *A data set of human endogenous protein ubiquitination sites.* Molecular & Cellular Proteomics, 2011. **10**(5): p. M110.002089.

366. Yoshimura, SH, Iwasaka, S, Schwarz, W, et al., *Fast degradation of the auxiliary subunit of Na+/K+-ATPase in the plasma membrane of HeLa cells.* Journal of Cell Science, 2008. **121**(Pt 13): p. 2159-2168.

367. Danielsen, JM, Sylvestersen, KB, Bekker-Jensen, S, et al., *Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level.* Molecular & Cellular Proteomics, 2011. **10**(3): p. M110.003590.

368. Povlsen, LK, Beli, P, Wagner, SA, et al., *Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass.* Nature Cell Biology, 2012. **14**(10): p. 1089-1098.

369. Azizi, B, Ziaei, A, Fuchsluger, T, et al., *p53-regulated increase in oxidative-stress--induced apoptosis in Fuchs endothelial corneal dystrophy: a native tissue model.* Investigative Ophthalmology and Visual Science, 2011. **52**(13): p. 9291-9297.