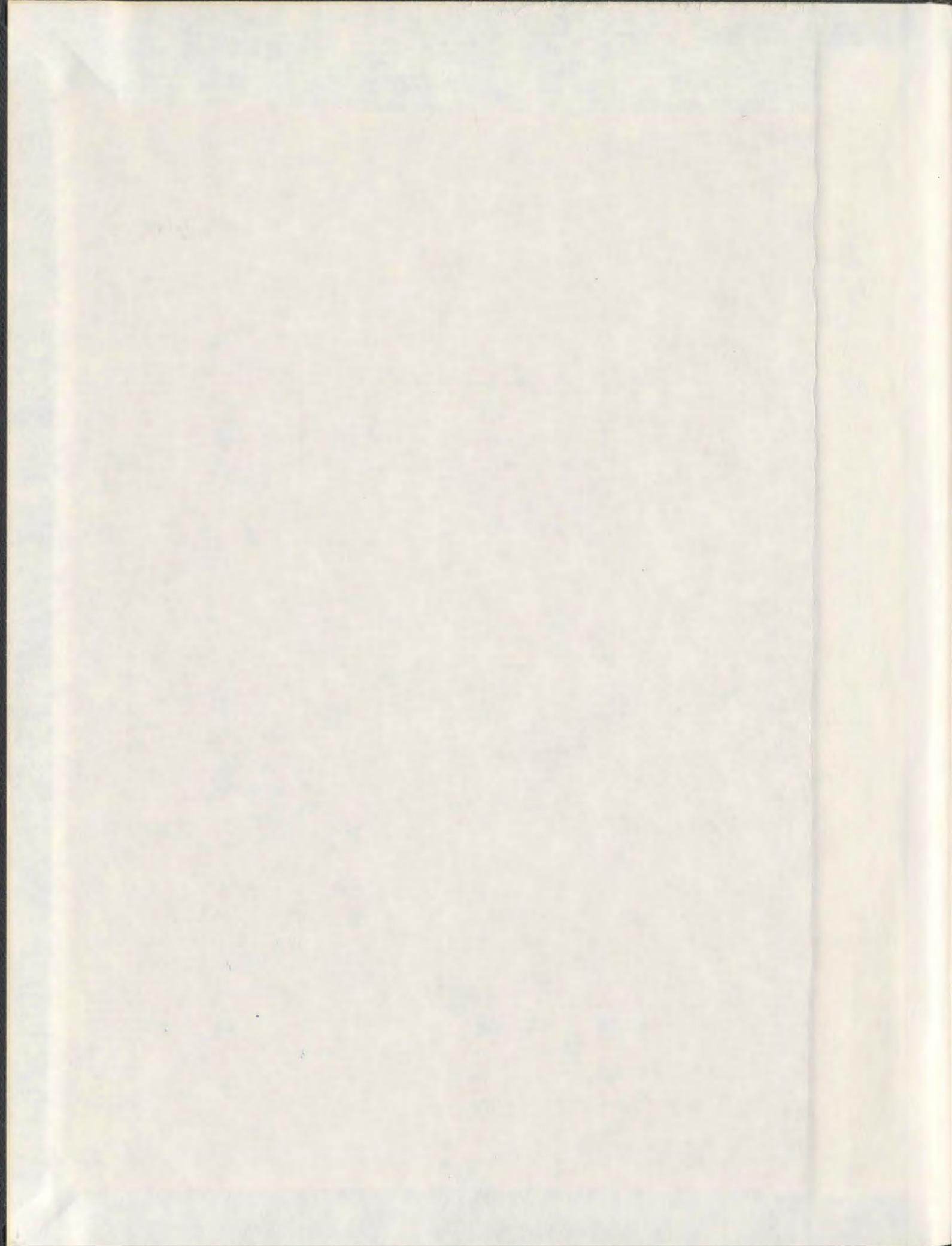


THE GENETIC CHARACTERIZATION OF
MENDELIAN OCULAR DISORDERS IN THE
POPULATION OF NEWFOUNDLAND AND LABRADOR

LANCE P. DOUCETTE



**The Genetic Characterization of Mendelian Ocular Disorders in the Population of
Newfoundland and Labrador**

By

© Lance P. Doucette

A thesis submitted to the School of Graduate Studies (SGS) in partial fulfillment of the requirements for the degree of Doctorate of Philosophy in the Faculty of Medicine, Discipline of Genetics, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada.

2012

Abstract

Background

Recently, ocular genetics have shown the first successes in genetic therapies, and treatment of genetic diseases making identification of disease genes of great importance. Gene discovery is most successful through the study of genetic founder populations, such as that of Newfoundland and Labrador.

Objective

The objective of this thesis was to identify disease genes in three forms of Mendelian ocular disease: anterior segment dysgenesis (ASD), achromatopsia (ACHM), and microphthalmia-dwarfism (MDW). This was undertaken to find novel genes and mutations to further our understanding genetic pathways involved in each condition.

Results

Of the 11 families recruited for this study, 8 were solved through identification of pathogenic variants. The ASD phenotype was discovered to be caused by a novel mutation in *FOXE3*, seven ACHM families through mutations in *CNGA3* and *CNGB3*, one ACHM family was found to actually have a rare disease called Jalili Syndrome through a novel mutation in *CNNM4*, and two MDW families helped determine a putative disease locus on 16q21.

Conclusions

The identification of seven mutations (two novel, five previously described) have solved the genetic etiology in eight of eleven families providing insight into the disease pathways for these families. This allows for genetic counseling and the possibility for genetically based therapies in the future.

Acknowledgements

There are so many people I'd like to thank for assistance, support, friendship, and love during my PhD program, arguably one of the most exciting and grueling stages of my life. Mostly, I would like to thank:

My Mom for being my support system through one of the most difficult and trying periods of my life. You raised me to be a kind, hardworking individual, and you've been an absolute inspiration to me with your strength, your advice, and your wisdom. I love you.

The memory of my late Father. You were so strong during the hardest of times, and I like to think that I took some of that strength. Your memory keeps me going even when I think things couldn't get any worse. Thank you, Dad.

The Young Lab Staff Dante, Jim, and Tammy. Thanks so much for all your support, your help, your jokes, and your fun. You guys make the Young lab what it is, and you help all the students grow into better people.

The Young Lab Students Nancy, Jessica, David, Amy, and Nelly. There were lots of students that came through the lab during my time there, but you five stick out the most as the biggest influences, the greatest of labmates, and the best of friends. Nancy, my predecessor, thank you for teaching me most of what I know, encouraging me, challenging me, and showing me that it's possible to get REALLY excited about future research. You're going to make a great PI, and I miss you dearly. Jessica, you've been with me in the Young lab for years, we've had a lot of fun together, and you made grad school a little more bearable just by being around, smiling, and laughing. I wish you the best of luck. David, we've been best friends for years. It was a pleasure to work with my good friend from childhood, I love you buddy, you'll make a fantastic Doctor. Amy, you weren't around long, but we bonded quickly, and had a tonne of laughs at conferences and in the lab. Another great doctor in the making. And Nelly, you've been in the Young lab about as long as I have, which is saying something! You've been a fantastic colleague, and co-worker. I'm also very proud to call you my friend. I hope we can keep in touch and do some collaborating in the future ☺

My grad school friends, Chris Butt and Jen Shea. You two were the best support system I could have ever asked for. It's so nice to know that on those days that experiments weren't working, or your supervisor was driving you crazy that someone was right there with you in the trenches. It has been an absolute pleasure knowing you both, and I really hope we stay in touch.

Terry-Lynn Young, You've been the best supervisor I could ever have asked for. Thank you for your knowledge, your wisdom, and your help. You had faith in me when even I didn't. You picked me up when I was down, and made me feel like a scientist when I felt like an idiot. We've had some great times, and some hard times, but you've always been on my side. I'll always treasure the relationship we have. Thank you.

Jane Green, my adopted supervisor. You have been so good to me over the past few years. Your knowledge of...well...everything! It has been such a pleasure to be around, and it's been so much fun working with you on the eye projects. Thank you, like Terry, for being on my side, and getting excited about the things I was doing. I couldn't have done this without you.

My roommates and friends Brian and Juanita for the fun, the relaxation, the jokes, and the fat cat (I love you, fuzzy buddy). Brian and Juanita, I love you guys, and I'll miss you so much when I leave. I wish you both the best of luck in life.

All my other friends of which there are too many to name, but here goes: Paul Ryan, David Richardson, Steve Winter, David Ross, David Neil, Mike Reid, Dane Sheppard...Apologies if I've missed anybody. You've all been awesome. I wouldn't be the person I am without you folks. I'm sorry to have to group you all here, but I love each and every one of you.

The European Journal of Human Genetics for publishing "A novel, non-stop mutation in *FOXE3* causes an autosomal dominant form of variable anterior segment dysgenesis including Peters anomaly." 2011 Mar;19(3):293-9

Ophthalmic Genetics for publishing "Molecular Genetics of Achromatopsia in Newfoundland Reveal Genetic Heterogeneity, Founder Effects and the First Cases of Jalili Syndrome in North America." 2013, Jan ePub ahead of print.

My thesis examiners Drs Darren O'Reilly (Memorial University), Sevtap Savas (Memorial University), and Andrew Orr (University of Dalhousie) for all the helpful comments and criticisms of this document. Your revisions have greatly improved this thesis, and I appreciate the time and effort put into it.

Table of Contents

Abstract.....	2
Background.....	2
Objective.....	2
Results.....	2
Conclusions.....	2
Acknowledgements.....	3
List of Tables.....	10
List of Figures.....	11
List of Abbreviations and Symbols.....	13
Chapter 1: General Introduction.....	16
1.1 Aims of this study.....	16
1.2 Anatomy and Physiology.....	16
1.2.1 Anterior Chamber.....	18
1.2.2 Retinal Tissues.....	20
1.2.3 Retinal Structure.....	26
1.3 Formation of the Eye: Involved genes and pathways.....	28
1.4 Vision Loss and Ocular Disease.....	32
1.5 Why Study The Genetics Of Ocular Disease?.....	33
1.5.1 Historical Contributions of Ocular Genetics.....	33
1.5.2 Recent Contributions of Ocular Genetics Studies.....	34
1.6 How to Study Ocular Disease.....	37
1.6.1 Founder Effect and Founder Populations.....	37
1.6.2 The Newfoundland Population.....	39
1.6.3 Gene discovery using a family-based approach.....	43
1.6.4 Genetic Analysis of Families.....	45
1.7 Summary and Study goals.....	53
Chapter 2: A Novel non-stop mutation in FOXE3 causes an autosomal dominant form of variable anterior segment dysgenesis including Peters Anomaly.....	54
2.1 Summary.....	55
2.2 Introduction.....	56

2.2.1 What is anterior segment dysgenesis?	56
2.2.2 The Genetics of ASD	57
2.2.3 Current Study	62
2.3 Materials and Methods.....	63
2.3.1 Clinical Investigation of Proband and Relatives.....	63
2.3.2 Selection and Screening of Functional Candidate Genes	66
2.3.3 DNA Isolation and Primer design.....	67
2.3.4 PCR Setup.....	67
2.3.5 Bidirectional Sequencing.....	68
2.3.6 RNA Isolation/cDNA synthesis.....	69
2.4 Results.....	71
2.4.1 Long-term Clinical Follow-up	71
2.4.2 Candidate Gene Screen Reveals Causative Variant in FOXE3	76
2.5 Discussion	90
2.6 Acknowledgements.....	95
2.7 Conflict of Interest	95
Chapter 3: A population-based study of achromatopsia in a Canadian founder population reveals genetic heterogeneity, a novel mutation in <i>CNNM4</i> , and the first family with Jalili Syndrome in North America.....	96
3.1 Summary	97
3.1 Introduction.....	98
3.1.1 What is Achromatopsia?.....	98
3.1.2 Clinical Presentation and diagnosis of Achromatopsia	99
3.1.3 Molecular Genetics of Achromatopsia	102
3.1.4 Aims of this work.....	105
3.3 Materials and Methods.....	106
3.3.1 Patients and Ophthalmological Examination.....	106
3.3.2 Functional Candidate Gene Selection and Screening	106
3.3.3 Variant and Haplotype Analysis.....	107
3.3.4 Cloning.....	107
3.3.5 Mini-Prep.....	108

3.3.6 Exome Sequencing.....	109
3.4 Results.....	110
3.4.1 Patients and Clinical Examination.....	110
3.4.2 Candidate Gene Sequencing.....	110
3.4.3 CNGB3 mutations revealed in families 1442, 1492, 1713, 1734.....	110
3.4.4 CNGA3 mutations revealed in families 1491, 1723, and 1726.....	111
3.4.5 Haplotype Construction and Analysis.....	117
3.4.6 Unsolved Families (Families 1734 & 0094).....	121
3.5 Discussion.....	136
3.6 Author Contributions.....	143
3.7 Acknowledgements.....	143
3.8 Conflict of Interest Statement.....	143
Chapter 4: A Rare Form of Microphthalmia Associated with Proportionate Dwarfism (MDW) is Linked to a Large Region on Chromosome 16.....	144
Credit and Significant Contributions.....	145
4.1 Summary.....	146
4.2 Introduction.....	147
4.2.1 What is Microphthalmia, and how is it caused?.....	147
4.2.2 The Genetics of Microphthalmia.....	148
4.2.3 Hallermann-Streiff Syndrome.....	154
4.2.4 Current Study.....	155
4.3 Materials and Methods.....	156
4.3.1 Patients and Clinical Observations.....	156
4.3.2 SNP Genotyping.....	156
4.3.3 Homozygosity Haplotyping.....	156
4.3.4 Two-point and Multipoint Linkage Analysis.....	157
4.3.5 Sanger Sequencing of Positional Candidate Genes.....	158
4.3.6 Variant Analysis.....	159
4.4 Results.....	160
4.4.1 Patients and Clinical Observation.....	160
4.4.2 Homozygosity Haplotyping (HH) Analysis.....	169

4.4.3 Linkage Analyses.....	171
4.4.4 Sanger Sequencing of Positional Candidate Genes	174
4.4.5 Haplotype Construction and Critical Region Reduction	181
4.4.6 Exome Sequencing.....	189
4.4.7 CNV Analysis	190
4.5 Discussion	191
4.5.1 MDW and HSS	191
4.5.2 Linkage Analyses and HH mapping	192
4.5.3 Positional Candidate Gene Selection and Sequencing.....	193
4.5.4 Exome Sequencing and Analysis.....	196
4.5.5 Other Analyses and Future Directions	197
Chapter 5: General Discussion and Conclusions	202
References.....	209
Appendix A: Polymerase Chain Reaction (PCR) setup protocols.....	223
Appendix B: TouchDown (TD) Thermocycling Program.....	224
Appendix C: 1x ABI BigDye Terminator (BDT) 3.1 Kit Cocktail for Sanger Sequencing	225
Appendix D: Thermocycling Protocols for Cycle Sequencing (ABI.SEQ)	226
Appendix E: Cycle Sequencing DNA Precipitation Protocol.....	227
Appendix F: Sequences and Tms for all Primers used for PCR and bidirectional sequencing in the Anterior Segment Dysgenesis project.....	229
Appendix G: cDNA Synthesis Protocol.	235
Appendix H: Sequences and Tms for all Primers used for PCR and bidirectional sequencing in the achromatopsia project.	236
Appendix I: Setup Protocol for the TOPO TA Cloning kit (Invitrogen) and schematic of the PCR 4-TOPO plasmid used for cloning experiments in the achromatopsia project.	240
Appendix J: Sequencing results from cloning experiments using 1734.1 surrounding exons 17-18 of PDE6C.	242
Appendix K: All cutoff values for the Homozygosity Haplotyping study carried out in collaboration with the Genome Center of McGill University.	245

Appendix L: List of primer sequences and melting temperatures (T_m) for genes sequenced in the critical region on chromosome 16 in the MDW project, namely *ARL2BP*, *BBS2*, *C16ORF57*, *MMP15*, and *HERPUDI*. 253

Appendix M: Primer sequences and T_ms for all non-coding genes (miR, SNORA, and SNORD) genes..... 257

List of Tables

Table 2.1 – Clinical description of Family 0023 members	66
Table 2.2 – Variants discovered through sequencing of functional candidate genes .	79-80
Table 3.1 – Clinical description of affected individuals from seven achromatopsia families.	115
Table 3.2 – Clinical description of affected individuals from family including ophthalmological and dental descriptions	116
Table 3.3 – List of discovered variants through sequencing of <i>CNGA3</i> , <i>CNGB3</i> , <i>GNAT2</i> , and <i>PDE6C</i>	117
Table 3.4 – Pathogenic variants discovered in all seven solved families with achromatopsia and Jalili syndrome	118
Table 3.5 – Number of variants discovered through exome sequencing of samples from family 0094 and applied filters	137
Table 4.1 – Clinical description of all affected family members from families 0066 and 1499.	167
Table 4.2 – Variants discovered by sequencing of 23 positional candidate genes .	177-179
Table 4.3 – Haplotype of the 16q21 region and critical region determination in affected members of families 0066 and 1499	184-186
Table 4.4 – Haplotype of the 16q21 region and critical region determination in unaffected members of families 0066 and 1499	187-189

List of Figures

Figure 1.1 – Schematic of the human eye.....	17
Figure 1.2 – Schematic of a Rod photoreceptor.....	22
Figure 1.3 – Vitamin A cycle.....	24
Figure 1.4 – Schematic of the human Retina.....	27
Figure 1.5 – Illustration of the embryological formation of the eye.....	29
Figure 2.1 – Pedigree of Family 0023.....	65
Figure 2.2 – Ocular phenotype of the proband from family 0023.....	73
Figure 2.3 – Haplotype analysis of <i>B3GALTL</i> in family 0023.....	81
Figure 2.4 – Haplotype analysis of <i>CYP11B1</i> in family 0023.....	82
Figure 2.5 – Haplotype analysis of <i>PAX6</i> in family 0023.....	83
Figure 2.6 – Segregation analysis of a <i>GJA8</i> variant in family 0023.....	84
Figure 2.7 – Haplotype analysis of <i>PITX2</i> in family 0023.....	85
Figure 2.8 – Segregation analysis of a <i>PITX3</i> variant in family 0023.....	86
Figure 2.9 – Haplotype analysis of <i>FOXC1</i> in family 0023.....	87
Figure 2.10 – Segregation analysis of a <i>CRYAA</i> variant in family 0023.....	88
Figure 2.11 – Haplotype analysis of <i>FOXE3</i> in family 0023.....	89
Figure 2.12 – cDNA Sequencing Results.....	90
Figure 3.1 – Pedigrees of the eight families recruited for the ACHM project.....	114
Figure 3.2 – Haplotype analysis of <i>CNGB3</i> mutations in four families.....	120
Figure 3.3 – Haplotype analysis of <i>CNGA3</i> mutations in three families.....	122
Figure 3.4 – Sequence surrounding exon 17 of <i>PDE6C</i>	124
Figure 3.5 –Agarose gel showing amplification of two samples using three primers (Primer 1, Primer 2, and Primer 3).....	126
Figure 3.6 –Agarose gel stained with SYBR safe showed amplification of 1734.1 and a control sample 0094.4 with Primer 1 and <i>PDE6C17+18-R</i>	128
Figure 3.7 – Agarose gel showing amplification of various primer combinations.....	130
Figure 3.8 – Haplotype Analysis of <i>CNGB3</i> in Family 0094.....	133
Figure 3.9 – Haplotype Analysis of <i>CNGA3</i> in Family 0094.....	134
Figure 3.10 – Haplotype Analysis of <i>PDE6C</i> in Family 0094.....	135
Figure 3.11 – Haplotype Analysis of <i>GNAT2</i> in Family 0094.....	136
Figure 3.12 – Protein Schematics of <i>CNGA3</i> , <i>CNGB3</i> , and <i>CNNM4</i>	140

Figure 4.1 – Circularized Pedigree of Family 0066	163
Figure 4.2 – Condensed Pedigree of Family 0066.....	164
Figure 4.3 – Circularized Pedigree of Family 1499.....	165
Figure 4.4 – Condensed Pedigree of Family 1499.....	166
Figure 4.5 – Results from Homozygosity Haplotyping Analysis.....	172
Figure 4.6 – Results from Multipoint Linkage Analysis	175
Figure 4.7 – Results of cDNA experiment with <i>NDRG4</i> variant.....	181
Figure 4.8 – Results of cDNA experiment with <i>SETD6</i> variant.....	182
Figure 4.9 – Picture of the 16q21 critical region from the UCSC Genome browser.....	190

List of Abbreviations and Symbols

AAPC	Attenuated Adenomatous Polyposis Coli
AAV	Adeno-Associated Viral Vector
ABCA4	<i>ATP-Binding Cassette A4</i>
AD	Autosomal Dominant
AR	Autosomal Recessive
ARVC/D	Arrythmogenic Right Ventricular Cardiomyopathy/Dysplasia
APC	<i>Adenomatous Polyposis Coli</i>
ASD	Anterior Segment Dysgenesis
B3GALTL	<i>Beta-1,3-Glucosyltransferase Like</i>
BBS	Bardet-Biedl Syndrome
BCOR	<i>BCL6 Corepressor</i>
bHLH	Basic Helix-Loop-Helix
BMP4	<i>Bone Morphogenic Protein 4</i>
cDNA	Complementary DNA
CDH23	<i>Cadherin 23</i>
CHD7	<i>Chromodomain Helicase DNA-binding protein 7</i>
CNGA3	<i>Cyclic Nucleotide Gated Channel A3</i>
CNGB3	<i>Cyclic Nucleotide Gated Channel B3</i>
CNNM2	<i>Cyclin M2</i>
CNNM4	<i>Cyclin M4</i>
CRYAA	<i>Crystallin Alpha-a</i>
cM	<i>CentiMorgan</i>
CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1
DKFZp762I194	<i>cDNA DKFZp762I194</i>
DNA	Deoxyribonucleic Acid
dbSNP	SNP Database
DMD	<i>Duchenne Muscular Dystrophy</i>
ECM	Extracellular Matrix
ERG	Electroretinogram
F8	<i>Factor VIII</i>
FOXC1	<i>Forkhead Box C1</i>
FOXE3	<i>Forkhead Box E3</i>
GJA1	<i>Gap-Junction A1</i>
GJA8	<i>Gap-Junction A8</i>
GNAT2	Guanine Nucleotide Alpha Transducin
HH	Homozygosity Haplotyping

IL1RN	Interleukin-1 Receptor Antagonist
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HSN2	Hereditary Sensory Neuropathy type II
HSS	Hallermann-Streiff syndrome
IBD	Identical (Identity) by Descent
IOP	Intraocular Pressure
L	Long
LCA	Lebers Congenital Amaurosis
LHON	Lebers Hereditary Optic Neuropathy
LOD	Logarithm of the Odds
M	Medium
MEN1	<i>Multiple Endocrine Neoplasia type 1</i>
MITF	<i>Microphthalmia Associated Transcription Factor</i>
MKKS	<i>McKusick-Kaufman/Bardet-Biedl syndromes putative chaperonin</i>
MSH2	<i>mutS Homolog 2</i>
mtDNA	Mitochondrial DNA
mRNA	Messenger RNA
NDRG4	<i>N-myc downregulator gene</i>
NGS	Next Generation Sequencing
OCT	Ocular Coherence Tomography
ODDD	Oculodentodigital dysplasia
OMIM	Online Mendelian Inheritance in Man
OPN1LW	Opsin 1 Long Wave
OPN1MW	Opsin 1 Medium Wave
OTC	<i>Ornithine Transcarbamylase</i>
OTX2	<i>Orthodenticle homeobox 2</i>
PAX6	<i>Paired box homeodomain</i>
PCDH15	<i>Protocadherin-15</i>
PCR	Polymerase Chain Reaction
PID	Pedigree Identification
PITX2	Pituitary Homeobox 2
PITX3	Pituitary Homeobox 3
RAX	<i>Retina and Anterior Neural Fold Homeobox</i>
RBP1	<i>Retinol Binding Protein 1</i>
RCHH	Region of Common Homozygous Haplotype
RNA	Ribonucleic Acid

<i>ROM1</i>	<i>Rod Outer Segment 1</i>
RP	Retinitis Pigmentosa
RPE	Retinal Pigment Epithelium
<i>RPE65</i>	<i>Retinal Pigment Epithelium 65</i>
<i>SETD6</i>	<i>SET-Domain 6</i>
SIFT	Sorting Intolerant from Tolerant
<i>SIX6</i>	<i>Sine Oculis Homeobox 6</i>
SNP	Single Nucleotide Polymorphism
<i>SOX2</i>	<i>SRY-BOX2</i>
STGD	Stargardt Macular Dystrophy
<i>STRA6</i>	<i>Stimulated by Retinoic Acid 6</i>
<i>SMOC1</i>	<i>Sparc-related Modular Calcium Binding Protein 1</i>
<i>TCAG</i>	Toronto Center for Applied Genomics
<i>TMEM43</i>	Transmembrane 43
UTR	Untranslated Region
VA	Visual Acuity
<i>VSX2</i>	<i>Visual System Homeobox Gene 2</i>
<i>WFS1</i>	Wolframin-Syndrome 1
<i>WHO</i>	World Health Organization

Chapter 1: General Introduction

1.1 Aims of this study

This thesis is an extension of research started by Dr. Jane Green over 30 years ago. Dr. Green reviewed the registration records maintained by the Newfoundland and Labrador division of the Canadian National Institute for the Blind (CNIB) and through the Ocular Genetics Clinic obtained a plethora of family-based information and pedigrees with the goal of pursuing genetic studies to determine the molecular etiology of these visual conditions. This thesis examines the genetics of three ocular diseases in the population of Newfoundland, from the earlier work of Dr. Green, to identify genes and mutations involved in the pathogenic mechanisms of these ocular conditions. These conditions included (1) Anterior Segment Dysgenesis (ASD), a malformation of the anterior chamber of the eye (iris, cornea, and lens), (2) Achromatopsia, a cone photoreceptor defect of the retina, and (3) Microphthalmia-Dwarfism (MDW), a syndromic form of microphthalmia (small eyes causing blindness). This introduction discusses the anatomy and structure of the eye, important pathways involved in its formation and function, and the importance of identifying disease genes in these conditions.

1.2 Anatomy and Physiology

The eye is a complex organ comprised of many structures which function to translate light energy into sight. These structures exist in two important major sections of the eye including the anterior segment (cornea, uvea, lens), and the posterior chamber, consisting of the sclera and the light sensitive retinal tissue and the retinal pigment epithelium (RPE) (Figure 1.1).

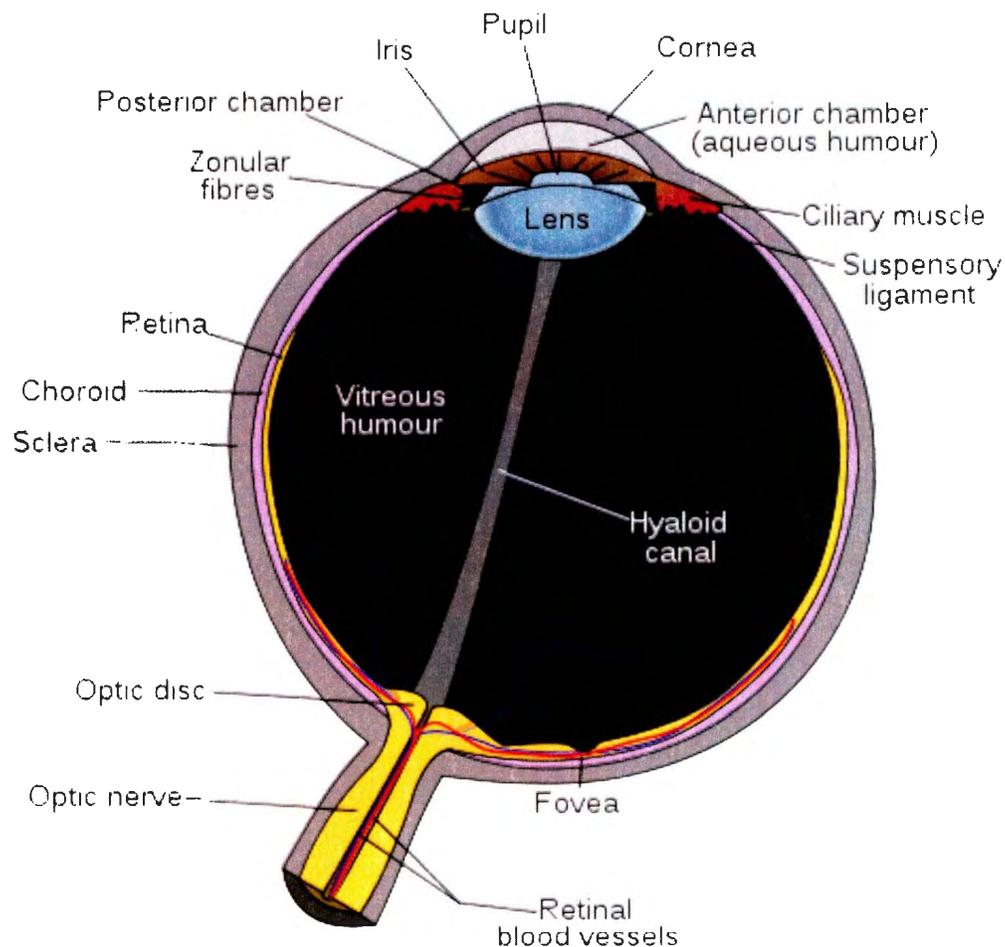


Figure 1.1: Cross-section of the human eye illustrating the location of various ocular structures. The anterior chamber of the eye can be seen at the top of the diagram and consists of the lens, iris, cornea and various ligaments for eye and lens movement. The chamber posterior to the anterior segment is filled with a gelatinous substance known as vitreous humour which nourishes the various structures of the eye. The retina, which translates light energy to sight, is the inner layer of the posterior segment of the eye. The choroid (middle) is a vascular layer which nourishes the retina, and the outer layer is the tough fibrous sheath known as the sclera which provides the structure of the eye. The Hyaloid canal is a small transparent canal running through the vitreous humour, from the lens to the optic nerve, and is used to help focus sight (accommodation). This image was obtained from <http://en.wikipedia.org/wiki/Eye> and has been released into the public domain for free use. (http://en.wikipedia.org/wiki/File:Schematic_diagram_of_the_human_eye_en.svg)

1.2.1 Anterior Chamber

The second chapter and first research project of this thesis is concerned with the genetics of a malformation of the anterior segment of the eye, ASD. The anterior segment (anterior chamber) refers to the three layers at the front of the eye which consist of the cornea, the uvea, and the lens. These structures are important in protection and nourishment of the eye, as well as focusing light onto the retina, where it is translated into sight through a biochemical pathway known as the phototransduction cascade.

The cornea, a thin transparent, avascular layer, is the outermost structure of the eye and is continuous with the sclera, the white fibrous outer sheath of the eye that provides its globular structure (Figure 1.1). The cornea is the most important structure in refraction of light onto the retina, and is the first line of defense against infectious agents, foreign objects, and physical injury.^{1,2} The structure of the cornea is organized into three cellular layers (Epithelium, Stroma, and Endothelium) which maintain clarity of the cornea, necessary for the passage of light,³⁻⁵ and two membranous layers, Bowman's Membrane and Descemet's membrane, which maintain the cornea's three-dimensional structure.⁶

The second component of the anterior chamber, the uvea, is a pigmented, vascular layer consisting of the iris, and the ciliary body, and is continuous with the choroid (which exists in the posterior chamber, an area directly behind the anterior chamber). The iris is a pigmented ring of intrinsically controlled (i.e. not voluntary) smooth muscle which regulates the amount of light entering the eye.⁷ During times of high luminescence, the iris constricts using a set of muscles called the circular (or constrictor) muscles to

form a smaller ring, making the opening of the pupil smaller, allowing minimal amounts of light to enter.⁸ During times of low light, the iris relaxes using the radial (or dilator) muscles, increasing the opening of the pupil, allowing more light into the eye.⁸ An extension of the iris, the ciliary body, is responsible for the production of aqueous humour, a clear substance comprised of mucopolysacchrides.⁷ This fluid is responsible for the nourishment of the cornea and lens.² The flow of this fluid through the anterior chamber also maintains the pressure within the eye, called the intraocular pressure (IOP). The aqueous humour flows around the anterior chamber, and eventually out through the trabecular meshwork, a tissue continuous with the ciliary body which regulates the flow of aqueous humour to maintain IOP. Malformations of these structures can cause elevated IOP and can lead to damage or death of the optic nerve, a condition known as glaucoma.⁹ The third component of the uvea, the choroid, located between the sclera and retinal tissues, contains blood vessels and capillaries to provide oxygen and nourishment to the outer layers of the retina. The boundaries of this layer begin at the uvea, but extend to the retinal pigment epithelium (RPE) at the posterior section of the eye (Figure 1.1).

Lastly, the lens is important in light focusing and is a biconvex structure located at the back of the anterior chamber (Figure 1.1). The lens is comprised of a layer of cuboidal epithelial cells, and a population of elongated fiber cells.¹⁰ The majority of the dry weight of the lens structure is comprised of a number of heat-shock like proteins known as crystallins. These proteins create a stable, gelatinous like material which is translucent due to the lack of cellular organelles (including nuclei) which allows the passage of light. This lack of nuclei in lens fiber cells means that no cell division occurs

after formation, making some fiber cells as old as the organism itself. The exception to this are epithelial cells of the lens, which retain nuclei and continue to slowly form lens fiber cells in the outer most regions of the lens, throughout the lifetime of the organism.

Mutations in genes involved in the proper formation and function of the anterior chamber and its structures can cause various malformations. One such genetic condition, ASD, is further discussed in Chapter 2.

1.2.2 Retinal Tissues

The third chapter of this thesis examines a condition (achromatopsia) affecting the light sensitive tissue of the eye, referred to as the retina. The retina, the inner layer of the posterior segment of the eye (Figure 1.1), is responsible for translating the energy of incoming photons into vision. This is accomplished through a biochemical process known as ‘the phototransduction cascade’ which involves a number of complex enzymatic processes within the retinal photoreceptors (rods and cones), and the RPE. The electric signal produced by these processes is then transmitted via the neuronal cells (bipolar, horizontal, amacrine, ganglion, and Muller glial cells).

(i) Photoreceptors (Rods and Cones)

The photoreceptors of the retina are the cells responsible for the biochemistry involved in phototransduction and are located in the outer nuclear layer of the retina, adjacent to the retinal pigment epithelium (RPE). These highly specialized cells come in

two subtypes: rods and cones. Both cells are similar in structure, containing outer and inner segments, though they carry out different responses termed 'scotopic' (rod-mediated) and 'photopic' (cone-mediated) responses.¹¹ The rod cells of the retina are by and large the most prominent of the photoreceptors (approximately 120 million), by about a margin of 20:1 in comparison to their counterpart, the cones.¹² The rod cells are extremely sensitive to small amounts of light and are thus specialized for low light vision but have low spatial resolution and because of their increased numbers in the periphery of the retina provide side (peripheral) vision.¹³ The cone cells, however, have high spatial resolution, but are comparatively insensitive to incoming light. Thus, the cones are responsible for high resolution and colour vision, and sight in well-lit conditions where photons are plentiful. The threshold for activation of the photopic responses in these cells is much higher than that of rods, so cones are only engaged in times of sufficient light abundance.¹²

Both rods and cones are organized into an inner and outer segment containing photopigments which absorb photons (Figure 1.2).¹⁴ The inner segment of the cell contains the components necessary for normal metabolic function (endoplasmic reticulum, Golgi apparatus, ribosomes, and many mitochondria required for powering the cell).¹⁵ The outer segment, connected to the inner segment by a modified cilium,¹⁶ is organized into a number of flat membranous structures called the photoreceptor discs which are produced and secreted by the connecting cilium.¹⁶ These discs are the main functional unit of the human photoreceptor and contain the photopigment, rhodopsin. Rhodopsin is a member of the G-protein coupled receptor (GPCR) family of proteins, and

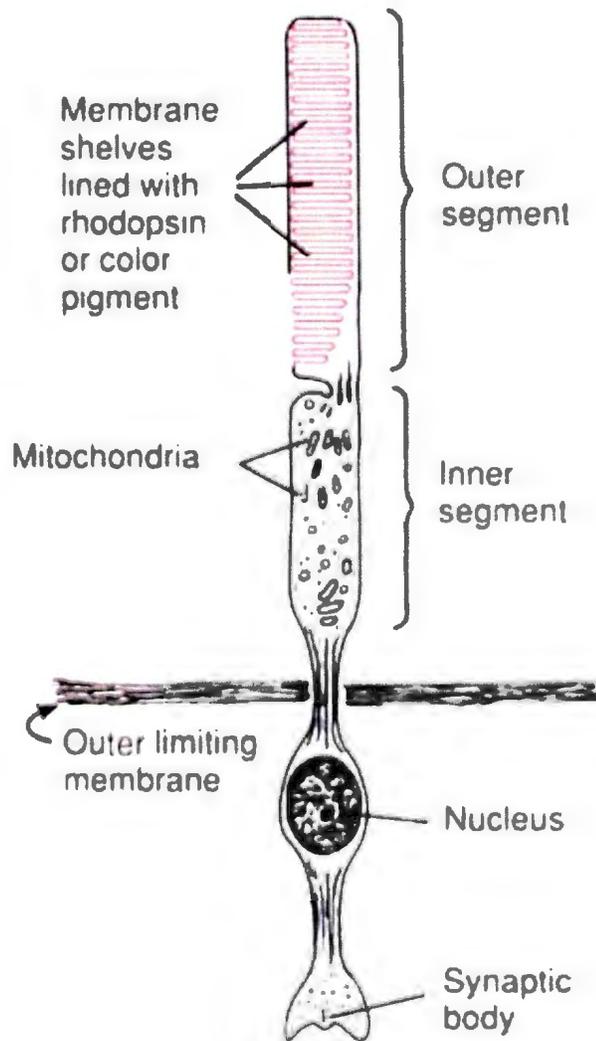


Figure 1.2: Image of a human rod photoreceptor cell illustrating the presence of the outer segment discs (top), and in descending order: the inner segment, nucleus, and synaptic body which terminates on neuronal cells. The outer segment discs contain the components necessary for the biochemistry of phototransduction. This image was obtained from en.wikipedia.org/wiki/Photoreceptor_cell, and has been released into the public domain for free use. (<http://en.wikipedia.org/wiki/File:Rod%26Cone.jpg>)

is activated by exposure to light. As photons are absorbed by Rhodopsin, cis-retinal, a vitamin-A derivative covalently bonded to the rhodopsin molecule, is converted to all-trans-retinal which is then hydrolyzed from the Rhodopsin molecule. The all-trans retinal then enters a cycle to replenish the cis-retinal molecule, resetting the photoreceptor cell to allow for another response (Figure 1.3). The isomerization of cis-retinal to all-trans retinal creates a conformational change in the GPCR, and initiates the phototransduction cascade.¹⁷ The cascade begins with activation of guanine nucleotide transducin molecules (GNAT) which in turn activate phosphodiesterases (PDE). These PDE enzymes reduce the amount of cyclic guanine monophosphate (cGMP) within the cell, causing closure of a channel called the cyclic-nucleotide gated channel (CNG). When this channel closes, the cell hyperpolarizes, and transmits a signal to the neuronal cells through the bipolar and horizontal neurons. Malformation of the CNG channel, or mutation of the GNAT or PDE class of enzymes can cause a number of conditions involving improper function of the photoreceptors. One such condition, achromatopsia, will be discussed in detail in Chapter 3 of this thesis.

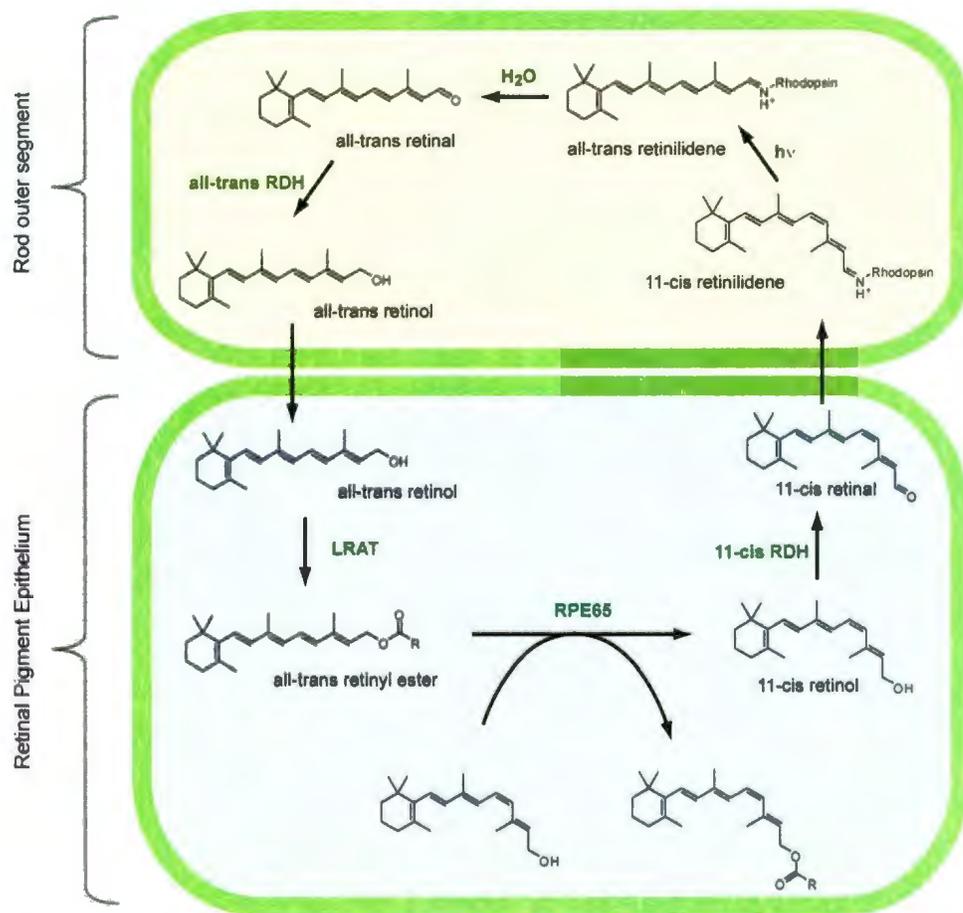


Figure 1.3: Schematic of the biochemical recycling of cis-retinal in the photoreceptor cells. 11-cis retinal is covalently bonded to an opsin molecule forming 11-cis retinilidene (Rhodopsin) and when stimulated by a photon isomerizes to all-trans retinilidene. A hydrolysis step then removes the molecule from rhodopsin, forming all-trans retinal. This is then transported to the Retinal Pigment Epithelium (RPE) where lecithin-retinol acyltransferase (LRAT) esterifies the molecule to all-trans retinyl ester. RPE65, an enzyme of the RPE then isomerizes the retinal molecule to 11-cis retinol. 11-cis retinol is then oxidized via 11-cis Retinol dehydrogenase (11-cis RDH) to regenerate the 11-cis retinal molecule. The newly recycled 11-cis retinal is then transported back to the outer segment of the photoreceptor to be covalently bonded to the photopigment opsin (forming rhodopsin) and resetting the cycle. This image was obtained from http://en.wikipedia.org/wiki/Visual_phototransduction and is published under the GNU Free Documentation License.

(ii) Neuronal Cells

Besides the rods and cones, there are four types of neurons (horizontal, bipolar, amacrine, and ganglion) and one glial cell (Muller Glia) which exist in the mature retina. These cell types (including the photoreceptors) originate from a single progenitor cell population, and are organized into precise layers at the back of the eye.¹⁸

The photoreceptors terminate on neurons in the inner nuclear layer. When the photoreceptors generate a neuronal signal, the bipolar cells, which are postsynaptic (occur after) to the photoreceptors,¹³ transmit the signal to the Ganglion cells, the 'output cells' of the retina.^{18: 19} The bipolar cells also receive signals from the horizontal cells and amacrine cells, which are interneurons (connecting neurons) between the photoreceptor and bipolar cells.¹⁸ These interneurons function to transmit the neural signals from the photoreceptors to the bipolar cells, and ultimately to the ganglion cells. The seventh cell type within the retina is the Muller glial cells which exist throughout the entire thickness of the retina. These cells, unlike their counterparts, are not neuronal, but structural in nature and are responsible for maintaining retinal cell homeostasis.²⁰ It has also been suggested that, due to their shape, these cells funnel light to the photoreceptors,²¹ and are also involved in retinal repair upon injury by dedifferentiation into multipotent progenitor stem cells.²²

1.2.3 Retinal Structure

All the cells described in the previous section make up the light sensitive tissue known as the retina. These cells are organized into a variety of structures within the retina important in the transduction of light energy into vision. These structures include the optic disc and cup which form the end of the optic nerve, and the macula/fovea (Figure 1.3).

The fovea and macula are important avascular areas located in the center of the retina predominated by photoreceptor cells. The macula is a large darkened spot with a small invagination about 1.0mm long in the center referred to as the fovea (*fovea centralis*) (Figure 1.4). These two structures contain the largest proportion of cone photoreceptors, and are therefore responsible for high visual acuity, central and distance vision, and colour discrimination. Towards the periphery of the retina, the number of cone photoreceptors decrease and number of rod photoreceptors increase, thus allowing low resolving peripheral vision. Loss or damage to the photoreceptors, and/or damage to the macula and fovea results in a loss of colour vision, high visual acuity, and central vision such as is seen in a group of progressive diseases called cone-dystrophies, or stationary cone disorders such as achromatopsia. Diseases which cause loss or damage of the rod photoreceptors or damage to the peripheral regions of the retina results in tunnel vision, and night blindness such as is seen with Retinitis Pigmentosa, a progressive condition which affects the rod photoreceptors. The retinal malfunction that this thesis examines is a form of cone malformation, called achromatopsia.

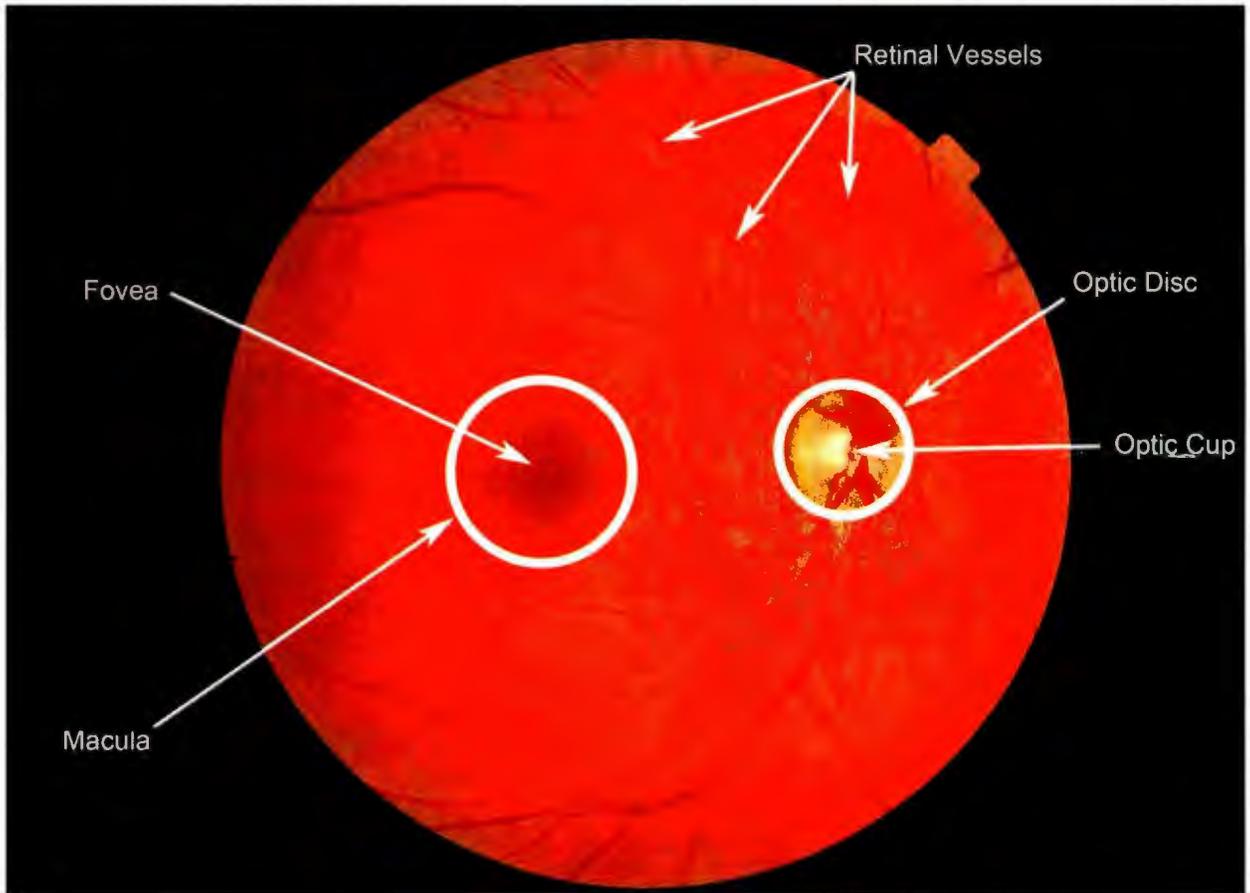


Figure 1.4: View of the retina of a right eye (oculus dexter; OD) upon funduscopy (view of the back of the eye using an ophthalmoscope). Important structures are indicated by white arrows and circles. The large dark area in the middle of the retina is called the macula, while the invagination (indicated by the darkest area) in the center of the macula is called the fovea. The fovea and macula contain only the cone photoreceptors responsible for high visual acuity. The large yellow spot to the right of the retina is called the optic disc, while the center is the optic cup. The optic disc is the area where all the retinal neurons become the optic nerve. The retinal vessels which nourish the retina with constant blood and oxygen supply enter the retina through the optic disc.

1.3 Formation of the Eye: Involved genes and pathways

The formation of the eye is a complex process mediated by a number of interactions between tissues that eventually become the retina, and those that become the anterior chamber. During embryogenesis, the optic vesicle is formed from the ventral forebrain neuroepithelium (Figure 1.5 shown in purple)²³ which then makes contact with the surface head ectoderm and begins to thicken, creating the lens placode (Figure 1.5 shown in green).^{23:24} During this process, the lens placode and distal optic vesicle invaginate to create the lens pit and optic cup (Figure 1.5), which gives rise to the neural retina and retinal pigment epithelium (RPE). The lens placode further invaginates to become the lens pit,²⁵ and later forms the lens vesicle, which dissociates from the surface ectoderm to become the mature lens.²³ These tissue-tissue interactions are mediated by transcription factors which determine the fate of each developing cell, many of which were discovered through studies of *Xenopus laevis* (African clawed Frog) which has informed a large portion of vertebrate retina development.

During development of the *Xenopus* nervous system, an area of the neural plate (i.e. the precursor of the neural tube which forms the spinal cord and brain) becomes designated as the 'eye field' and consists of a population of multipotent stem cells. The formation of the eye field is mediated by the inhibition of the bone morphogenic family of proteins (BMP) which are responsible for orchestration of tissue formation, and the canonical Wnt signals (a signaling pathway involved in embryogenesis).^{26:27} This inhibition is carried out by *Noggin*, which inhibits BMP in high doses. Once the eye field is established, a number of other transcription factors such as Pax6, Six3, and Rx1,

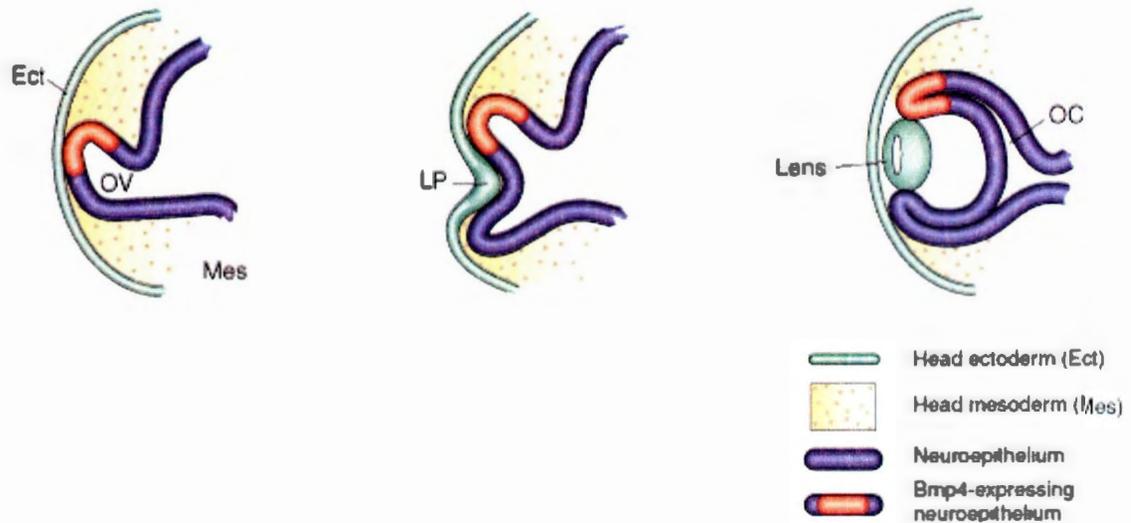


Figure 1.5 – Image of the developing eye. The purple area shows the neuroepithelium forming the optic vesicle (OV), which makes contact with the head ectoderm to form the Lens Placode (LP). The LP and OV invaginate to form the mature lens and optic cup (OC) respectively. (From Molly Weaver & Brigid Hogan. “*Powerful ideas driven by simple tools: lessons from experimental embryology*” *Nature Cell Biology* 3, E165 - E167 (2001). Reprinted by permission of the Nature Publishing Group, License #: 2953131304891).

termed 'eye field transcription factors' (EFTF) are activated. These EFTFs assist in proliferation of multipotent progenitor cells while surrounding areas are undergoing cellular differentiation. This proliferation maintains a large pool of cells necessary to generate all the retinal cell types at different stages of development.²⁸ As the eye field becomes specified, Six3 activates the Sonic hedgehog (Shh) pathway, which then defines the ventral diencephalon of the brain. Following this, the eye field designates two areas which will become the optic vesicles, which then invaginate to become the optic cup (Figure 1.4). At this stage, the retinal progenitor cells continue to proliferate and will generate the six types of retinal neurons (rods, cones, bipolar, amacrine, ganglion and horizontal). The retinal neurons develop in a time-dependent fashion with ganglion cells developing first, followed by cones, horizontal cells, amacrine cells, rods, bipolar cells, and Muller glial cells. The ability of a cell to differentiate into a particular neuronal cell is dependent on the time at which a progenitor cell exits the cell cycle. Early progenitors are competent to give rise to early-born cells, while late progenitors generate late-born cells. Typically, Ganglion cells are the first to develop followed by cones, horizontal cells, amacrine cells, rods, bipolar cells, and finally Muller glia. Whether or not a cell will exit the cell cycle is regulated by transcription factors or signaling molecules such as *Rx1*, *Vsx2* or *Notch*, which maintain cells in a proliferative, undifferentiated state by down regulating genes which promote cell-cycle inhibition, and neurogenesis.²⁸ Down-regulation of *Rx1*, *Vsx2*, or *Notch*, and activation of genes which promote cellular differentiation, such as *Hairy2*, *Zic2* and *p27Xic*, allow cells to exit the cell cycle and differentiate.

Mutations in the genes involved in these various processes create malformations in the eye's structure and function, causing visual defects. A widely studied example is *PAX6*, a gene originally cloned in the small eye mouse (*sey*), and is a homologue of the eyeless gene (*ey*) seen in *Drosophila*. This evolutionary conservation indicates that this gene has a highly important role in formation of the eye, even in lower organisms. *PAX6* is a homeobox transcription factor which has been called a 'master control' gene in eye morphogenesis, due to its ability to control a large number of genes during eye formation. *PAX6* was also the first gene discovered to play a role in human anterior segment disorders. During embryogenesis, *PAX6* is expressed in the tissue from the surface ectoderm and neuroectoderms including the corneal epithelium, lens, and optic cup. Heterozygous mutations in this gene have been associated with a number of conditions, especially those of the anterior segment such as aniridia (complete or partial absence of the iris), congenital cataracts, and/or Peters anomaly (corneal opacities, and keratolenticular adhesions) which will be discussed in Chapter 2.²⁹⁻³¹ *PAX6* has also been shown to play a role in the formation of retinal neurons, as *Pax6* knockout mice only develop amacrine cells, suggesting that Pax6 is necessary for production of all retinal neurons except amacrine cells. This is accomplished by down-regulating various basic helix-loop-helix (bHLH) transcription factors, such as Hes bHLH, involved in the production of the other six neuron types.²⁸

1.4 Vision Loss and Ocular Disease

According to the World Health Organization (WHO) ocular disease causing vision loss affects 285 million people worldwide of which 39 million are blind. This number is predicted to increase as the world population grows and ages.^{32: 33} This large prevalence of visual loss can have serious social and economic implications on a global scale. A study by the WHO in 2003 showed that the worldwide economic burden (loss of productivity) for those affected with blindness was \$19 billion in the year 2000 and was predicted to grow to \$50 billion by 2020 if no medical intervention (research/treatments) was carried out. A similar figure was proposed for those with residual vision and this was predicted to grow from \$42 billion to \$110 billion from 2000-2020.³³ These large economic losses are coupled with the social stigma that goes with visual loss such as difficulty finding meaningful employment and lowered self-esteem/self-worth.

A large number of diseases related to vision loss have been shown to have a genetic component. These conditions can be both complex (caused by gene-environment interactions), or Mendelian (caused by a single gene). Complex conditions pose particular challenges when studying ocular conditions as the environmental factors seem to play a large role and can be difficult to elucidate. Mendelian diseases (sometimes called monogenic disorders) are conditions caused by mutation of a single gene, and can be passed on from generation to generation. As only one gene is suspected to be causing the disease being studied, Mendelian conditions are much easier to study than complex conditions, especially through the study of families where inheritance of a single gene mutation through generations can be observed. An example of Mendelian eye disease

which has been studied through a family-based approach is retinitis pigmentosa (RP). This condition affects 1 in 4000 individuals in the U.S. and is caused by mutations in over 50 genes and loci. Mendelian conditions of the eye can be nonsyndromic, such as RP which only affects the retina, or can be syndromic, which affect the eye and other systems. An example of syndromic eye disease is Usher Syndrome, a condition which has RP as well as hereditary hearing loss. These types of conditions are also caused by mutations of single genes, though the genes can be involved in different processes, creating multiple system malformations. The focus of this thesis is to examine nonsyndromic and syndromic Mendelian eye disorders, with the purpose of discovering genes involved the pathogenic mechanisms of the ocular condition.

1.5 Why Study The Genetics Of Ocular Disease?

1.5.1 Historical Contributions of Ocular Genetics

The study of ocular genetics has always been at the forefront of medical and experimental genetics. Conditions of the eye can present early in life, or can develop over time, and due to the eye's external location, can be more easily observed in conditions with a structural change (i.e. cataracts). Conditions with a functional/internal change, such as those which affect the retina, are more difficult to diagnose than those with external changes due to ease of observation, though still remain easier in comparison to other conditions such as various cancers in terms of detection. The relative ease of phenotyping has made the genetic study of ocular diseases very attractive to researchers,

especially when seen in families. As a result, the field of ocular genetics has contributed a great deal to the scientific community through discovery of various genetic phenomena. These phenomena include X-linked inheritance, discovered through the study of colour blindness,³⁴ mitochondrial inheritance through study of Lebers hereditary optic neuropathy (OMIM 535000), Knudson's two hit hypothesis in cancer genetics via Retinoblastoma,³⁵ and digenic inheritance through study of RP.³⁶

1.5.2 Recent Contributions of Ocular Genetics Studies

As shown in the previous section, the study of ocular disease has always been important in the advancement of genetics and these studies continue to advance the science to this day. The hope and goal of any genetically-based research endeavor is to ultimately find a way to treat, or at the very least, attenuate symptoms of a particular genetic condition. Recently, the study of ocular conditions has taken this goal and made it a reality through the first successful attempts at genetic therapies.

In 1990, the first FDA approved tests were carried out to cure a genetic disease called adenosine deaminase severe combined immunodeficiency (ADA-SCID) in two young patients. The results from these experiments were promising but only temporary. In 1999, tests were carried out to cure a genetic disease called ornithine transcarbamylase (OTC) deficiency (OMIM 311250). OTC deficiency is a defect in nitrogen metabolism which is usually fatal at birth, but less severe cases can be managed with proper diet. The cure for this condition came via an adenoviral vector carrying a normal copy of the *OTC* gene, a process termed 'gene therapy'. The theory behind this work is that the adenoviral

vector would use the patient's normal translational machinery to produce a functional OTC enzyme. Unfortunately, following injection of the adenoviral vector the patient, Jesse Gelsinger, suffered anaphylactic shock and passed away.³⁷ This tragedy was a major setback for many scientists in the field of genetic therapies, which then came under intense scrutiny from the public and government.³⁸

Recently, the field of gene therapy was revisited in human subjects for the first time since the death of Jesse Gelsinger through a study of a retinal condition called Leber Congenital Amaurosis (LCA). LCA is an autosomal recessive condition which presents at birth or early in life with severe vision loss or total blindness and absent photoreceptor response on electroretinogram (ERG) testing.³⁹ This condition affects approximately 1/80000 individuals worldwide⁴⁰ and is genetically heterogeneous with 18 associated genes.^{41,42} One of these genes, *RPE65*, has been extensively studied in mouse, dog, and humans. *RPE65* produces an enzyme of the RPE which produces 11-cis retinol in the vitamin A cycle (Figure 1.3), a molecule essential in normal retinal function.³⁹ In 2008, two trials, one in Philadelphia, USA⁴³ and one in London, England,⁴⁴ were published back-to-back in the New England Journal of Medicine. These studies also used an adenovirus-associated vector (AAV) with a wildtype copy of *RPE65*, in hopes of producing a normally functioning protein. After administration of an AAV to six patients with LCA and mutations in *RPE65* (one null mutation, five missense mutations) improved the vision in all tested individuals.

These trials have been successful in the eye as it is an easily accessible, compartmentalized, and immune privileged organ making it a very unique organ for gene-delivery.⁴⁵ These characteristics have made it a great first step towards generalized genetic therapies. The LCA trials have ushered in a new era of retinal therapies and have led to a number of new exciting animal studies involving gene therapy for various ocular disorders, in hopes of developing other human trials. These conditions include a study of *CNGA3* and *CNGB3* gene therapy in achromatopsia in dogs,⁴⁶ and mice,^{47: 48} which have had promising results. All these exciting experiments have created a flurry of other clinical trials for various conditions. A quick search of 'gene therapy' on the clinicaltrials.gov website reveals hundreds of planned, ongoing, and completed trials using gene therapies, and it will not be unsurprising to see many more added in the coming years.

The advances made through gene therapies have been brought about by the long standing research of genetic ocular conditions. Without knowing the particular genes, mutations, and pathways involved with each individual disease, development of a treatment would be impossible. Treatment of patients with ocular disease will depend wholly on knowing which condition and which gene mutation the patient has, which will lead to a personalized form of medicine. This makes the discovery of gene mutations involved in ocular disorders all the more important for clinical practice, and for treatment development. Discovery of novel genes and mutations, such as is done in this thesis, will allow us to gain more knowledge of the pathology of the condition, and the various pathways involved, paving the way for treatments like those listed above.

1.6 How to Study Ocular Disease

As discussed in the previous section, to eventually develop a treatment for a particular ocular disease, one must first understand the pathological mechanisms involved. An effective method of discovering disease genes is through studies of monogenic ocular conditions. Due to the large number of monogenic (single gene) ocular disorders that have been identified, a very effective method for the study of these conditions is via family-based or population-based approaches. Founder populations such as that of the Mennonite, Hutterite, Icelandic, or as is used in this study, the Newfoundland population, offer unique resources for the study of monogenic disorders through availability of large pedigrees, and potential founder effects.

1.6.1 Founder Effect and Founder Populations

The founder effect is a phenomenon seen when a number of separately identified individuals share a single common ancestor often found through genealogical or archival studies. This effect can be caused by genetic bottlenecks created through natural disasters (killing off a large portion of the population), or migration of a small group of individuals from a larger population to create a founder population. If a greater proportion of the new smaller population contains a rare pathogenic variant than the larger source population, it will become more frequent as the population expands. This can result in a higher than normal prevalence of a disease phenotype. This situation can be exacerbated by limitation of intermarriage with other groups due to differences in culture, religion, language, or geography.⁴⁹ If a rare mutation is present in one of the founders, it may become relatively

common as the population expands. As a result of the founder effect, individuals carrying mutations causing rare diseases may share a common ancestor⁴⁹ and the mutation will be part of a disease-associated haplotype (a collection of alleles that are located close together on the same chromosome). With successive generations, meiotic recombinations will slowly reduce the size of the disease associated haplotype and only markers closest to the disease allele (i.e. are linked) will be shared amongst affected individuals of a founder population. This phenomenon of alleles that are physically close to one another on the chromosome and are inherited together is termed linkage disequilibrium (LD). An example of a population used in the context of ocular disease is that of the Pingelap Islands. This small island in the Pacific Ocean was home to a small population of about 200 people, until a massive storm known as Typhoon Lengkieki decimated the island in approximately 1775, killing all but about 20 individuals (including 9 males).⁵⁰ It is believed that this event created a genetic bottleneck in the Pingelap population, increasing the proportion of carriers for an autosomal recessive condition called achromatopsia (discussed in Chapter 3) creating an approximate disease prevalence of about 10% amongst its now 250 individuals (as opposed to 0.003% in the USA). A study of this island population had found that the increased frequency of achromatopsia could be traced back to a single individual, named Nanmwarki Mwahuele. As founder populations, such as the Pingelapese, are of great use in gene discovery studies, the subjects and genetic conditions outlined in this thesis will concentrate on, and use individuals from the genetic isolates of Newfoundland and Labrador, Canada.

1.6.2 The Newfoundland Population

The province of Newfoundland and Labrador on the East Coast of Canada includes the island of Newfoundland and the adjacent mainland, Labrador. The island of Newfoundland was originally discovered by an Italian navigator, Giovanni Caboto (John Cabot), in 1497. The island of Newfoundland is a collection of genetic isolates, separated by geographical and religious isolation, with a population of 509,348 (Newfoundland and Labrador Statistics Agency: <http://www.stats.gov.nl.ca/> accessed June 29th, 2012). The original approximately 20,000 immigrants from south-west England (Protestant), and southern Ireland in 1760 (Roman Catholic) settled in individual protected bays and coasts, along the coastal regions of the island to have access to the large fish stock. The original population had expanded to approximately 200,000 by 1890.⁵¹ Due to the differing religious beliefs there was also segregation between the Protestant and Irish Catholic settlers.⁵² This has resulted in multiple separate founder clusters in the various bays and inlets surrounding the island. The existence of multiple founder effects and the large family size in the Newfoundland population has created a unique resource for studying genetic conditions and has led to the discovery of a number of founder mutations and novel genes.

(i) Founder Mutation Discoveries in Newfoundland

A number of studies in the Newfoundland population have identified identical mutations shared by multiple separate families, assumed (or proven) to be founder mutations. In the 1980s-1990s, a study of Multiple Endocrine Neoplasia type I (MEN1; OMIM 131100), was carried out on the Burin Peninsula/Fortune Bay region of

Newfoundland. *MEN1* is an autosomal dominant cancer syndrome primarily affecting organs involved in the endocrine system, namely the pituitary, pancreas, and parathyroid glands. This condition is highly penetrant with over 94% penetrance by age 50, and is caused by mutations in the *MEN1* gene. The Newfoundland study identified a “founder” nonsense mutation in the *MEN1* gene in four large multiplex families.⁵³ Archival studies have so far demonstrated a common ancestor to three of the four families. Also in the 1980s-1990s, Attenuated Adenomatous Polyposis Coli (AAPC; OMIM 611731), an autosomal dominant cancer predisposition which can lead to a variable number of adenomatous polyps in the colon and rectum, was studied in five families from the North-East coast of Newfoundland. This study led to the discovery of a splice mutation in exon four of the *APC* gene across all five families.⁵⁴ A common ancestor has subsequently been identified. Hereditary Nonpolyposis Colorectal Cancer (HNPCC; Lynch Syndrome OMIM 120345) has been extensively studied in the Newfoundland population. A linkage study of two families from Newfoundland identified the first HNPCC gene, *MSH2*, a mismatch DNA repair gene.⁵⁵ Currently, 17 independently identified families from the same part of the province have this founder mutation. Four of these families have a demonstrated common ancestor.

(ii) Gene Discoveries from the Newfoundland population

As well as multiple founder mutation discoveries, the Newfoundland population has proven to be a resource to find novel genes, or novel associations of already known genes. Arguably the greatest accomplishment from a study of the Newfoundland

population has been that of a syndromic form of hereditary vision loss called Bardet-Biedl Syndrome (BBS; OMIM 209900). BBS is an autosomal recessive condition characterized by retinal dystrophy, renal abnormalities, polydactyly, mental retardation, and mild obesity with a prevalence in Newfoundland of 1 in 17500.⁵⁶ BBS is a genetically heterogeneous condition with 17 known loci or genes currently identified.⁵⁷ The first locus, *BBS1*, was originally described in 1994,⁵⁸ but was refined to a 1cM region in 1999 through a study of *BBS1*-linked Newfoundland families.⁵⁹ This refinement eventually led to the discovery of the *BBS1* gene in 2002.⁶⁰ Secondly, the *BBS3* locus, described in 1994,⁶¹ was refined to 3cM using a Newfoundland pedigree, leading to the discovery of the gene in 2004.⁶² The *BBS5* locus was identified through homozygosity mapping of a large Newfoundland family in 1999.⁶³ The gene at this locus (*DKFZp762I194*) was then discovered in 2004.⁶² Finally, the *BBS6* locus and gene (*MKKS*) were discovered through a genome-wide scan of a large Newfoundland pedigree in 2000 and was the first gene to be described for BBS.⁶⁴ The discoveries made in BBS through studying the Newfoundland population were very interesting, as one would think that a disease with such multi-system malformations as BBS, at an increased frequency as was seen in Newfoundland, would have a single genetic cause in a founder population. These studies revealed that the Newfoundland population is actually made up of multiple distinct genetics isolates rather than a single isolate.

More recently, a discovery was made in the Newfoundland population which has had massive implications for clinical diagnoses and treatment in patients with Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia (ARVC/D). ARVC

(OMIM 107970) is a very severe autosomal dominant condition of the heart where the myocardium (heart muscle) becomes replaced with fibrofatty tissue, creating an aberration in the electrical activity of the heart. The first presentation of this disease is often sudden cardiac death, and affects males more frequently and at an earlier age than females. The age at onset in men is between 40-50 years of age, and in women between 50-60 years of age. ARVC is genetically heterogeneous with 11 genetic loci and eight genes currently known.⁶⁵⁻⁶⁹ The eighth gene was elucidated through a study of Newfoundland families with a more severe form of ARVC than in other populations. The locus for the Newfoundland form of ARVC/D was originally mapped to 3p23⁷⁰ in an eight generation family from Newfoundland first identified in the 1980s and was called *ARVD5*.⁷¹ For many years after the mapping of the *ARVD5* locus, a disease associated haplotype was used to assist in diagnosis of patients with ARVC, as the gene was not known. However, in 2008, a study of 15 families who were known to have ARVC (14 families plus the original one from the 1980s) were used to refine the *ARVD5* locus to 2.36Mb on 3p25.⁷² Cascade sequencing of the positional candidate genes in this region identified a missense mutation (p.S358L) in *TMEM43* which segregated with the disease phenotype in all 15 families.⁷²

The discovery of genes causing BBS and ARVC are arguably the two largest contributions from studies of the Newfoundland population, though other novel genes and associations have been found as well. These include a study of non-syndromic hearing loss which identified mutations in the *WFS1* gene as an autosomal recessive cause of hearing loss, which was previously thought to only cause Wolfram-Syndrome

(OMIM 222300).⁷³ Other examples of gene discoveries from using the Newfoundland population include *MSH2* in Lynch Syndrome (the first identified gene for this condition),⁷⁴ *HSN2* in hereditary sensory and autonomic neuropathy type 2,⁷⁵ and *IL1RN* in interleukin 1 receptor antagonist deficiency.⁷⁶

1.6.3 Gene discovery using a family-based approach

An effective method of studying monogenic conditions, especially in a founder population, is via a family based method. This approach has yielded great success in the past 20 years, and allows for multiple comparisons of variants within and between families, as well as the advantage of segregation analysis. Segregation analysis simply allows a researcher to track the inheritance pattern of a particular variant or marker genotype through a family. Thus, it is important to carefully select the disease of interest, and to thoroughly define the minimal diagnostic criteria for the disease, so that phenotyping can be done easily and accurately (though this is not always an easy task). The pedigree structure can then give insight into the mode of inheritance (autosomal dominant, recessive, X-linked, mitochondrial, or complex). An inaccurate diagnosis of the disease, either through phenocopies (individuals with the disease, but from another cause), variable age of onset, variable expression, and incomplete or reduced penetrance, or ill-defined diagnostic criteria can severely skew the results of segregation analysis, and either overlook or miss a very important result.

Phenocopies, in the case of family studies, are members of a family that show the disease phenotype, but because of some extraneous factor (i.e. environment), rather than

genetics. A prime example of phenocopies is illustrated by the formation of cataracts in the anterior segment of the eye. If a researcher is studying a family with early onset/congenital cataracts, it is possible that there may be some family members with cataracts not caused by a genetic predisposition. For example, the use of some corticosteroids, and exposure to UV radiation has also been associated with cataract formation. Thus, it becomes possible that some members of the studied family have cataracts from environmental exposure. It becomes important to carefully review medical records and exclude individuals who may possibly have the phenotype in question for a different reason.

Variable expression is when members of the same family or people who share the same disease have varying severities of the condition or different tissues or organs involved. This can be problematic when trying to distinguish affected family members, from unaffected family members, as it may be possible that some individuals are affected, though they have a subclinical phenotype (very mild affection which may go unnoticed during examinations). An example of variable expression will be discussed in Chapter 2 in a study of anterior segment dysgenesis.

Penetrance is defined as the probability that a person with a particular deleterious mutation will express the associated disease or phenotype. For example, achromatopsia, which is discussed in Chapter 3 of this thesis, is a fully penetrant autosomal recessive condition.⁷⁷ This means that if a patient has a mutation on both copies of a gene known to cause achromatopsia, they will express the phenotype 100% of the time. An example of a

low penetrance ocular condition is retinoblastoma, an aggressive eye cancer occurring in childhood.⁷⁸ Studies have found that truncating (null) mutations (nonsense mutations, splice mutations) of *RBI* cause a high penetrance of retinoblastoma, whereas non-truncating mutations (missense) confer a lower chance of developing retinal tumours. It has been shown that an inherited null mutation, accompanied by subsequent loss of heterozygosity through somatic mutation (Knudson's Two Hit Hypothesis), causes high penetrance retinoblastoma.⁷⁹ The presence of a missense mutation leaves residual function of the *RBI* gene, lowering the chance for tumour formation, creating a low penetrance condition.⁷⁹ Thus, it is necessary to have a sense of the degree of penetrance in the disease being studied, as otherwise interesting mutations can easily be overlooked if the expected phenotype is not present in all affected individuals. It may be possible that a mutation is indeed causative, but is a mutation of reduced penetrance.

1.6.4 Genetic Analysis of Families

(i) Functional Candidate Gene Screening

Once a disease, minimal diagnostic criteria, pedigree, and mode of inheritance are carefully laid out, one can begin searching for the genetic mutation causing the disease of interest in the family (or families). Typically, the first step in these cases is to begin searching for the putative mutation in genes already known to cause the condition (referred to as a 'candidate gene approach'). Chapter 2 of this thesis describes such an approach by sequencing genes known to cause an anterior segment dysgenesis phenotype. This approach is also highlighted in the study of achromatopsia in Chapter 3,

where affected individuals from multiple families were screened for mutations in the four known achromatopsia genes. This initial approach can save time and money in the long term as some cases may be able to be solved through the screening of previously associated genes.

(ii) Genome Wide/Chromosomal Scans

If a pathogenic variant is not found in previously described genes, or the disease in question is incredibly rare and has no previously associated genes or regions, one can begin a genome-wide search for the mutation in question. A genome-wide analysis is typically conducted by using a number of markers which span each of the 22 autosomes (or the X and/or Y chromosome in the case of sex-linked inheritance). These markers can either be microsatellites (repetitive elements of the genome), or Single Nucleotide Polymorphisms (SNPs). The genotypes at each of these markers are then determined. Microsatellite markers are areas of the genome that contain repetitive sequences and often vary in length/size between individuals. As a result, these markers are considered poly-allelic, and have many different genotypes, whereas SNPs are biallelic, containing one of two alleles (e.g. either a C nucleotide base, or a G nucleotide base). Microsatellites can provide more information than SNPs due to being poly-allelic, but are less evenly distributed across the genome (~16,000 microsatellites) than SNPs. Although SNPs do not provide as much information as microsatellites due to their biallelic nature, there are currently microarray chips which allow the genotyping of millions of SNPs across the genome in a cheap and timely manner.

(iii) Linkage and Haplotype Analysis

Once the genotypes of a set of markers have been determined, an analysis particularly effective in family studies can be carried out, called 'Linkage Analysis'. The purpose of this analysis is to link the individuals with a disease phenotype with a genotype at a particular marker (or markers), giving an idea as to where in the genome the disease gene is located. The position of the various SNPs and microsatellites are then determined (i.e. what chromosomal region) by comparison to the Human Genetic Map via various databases (UCSC Genome Browser,⁸⁰ Ensembl,⁸¹ etc) and can then be placed in the order that they exist on the corresponding chromosome. The ordering of alleles in this manner is referred to as a haplotype and represents the physical position of these markers on a chromosome. During meiosis, chromosomes have the opportunity to recombine at matching regions on homologous chromosomes. This recombination rearranges the alleles of various genes on either chromosome, and is responsible for a great deal of the variation seen in humans. Despite this shuffling of genetic information, areas of chromosomes that are located physically close to one another tend not to recombine, and are said to be in linkage disequilibrium. The distance between two markers is determined by the likelihood of recombination between the two regions and is measured in centimorgans (cM). If two loci are 1 cM apart, then the recombination rate between them is 1% and is generally considered to be approximately one million basepairs of the human genome. This is called a genetic map, or genetic distance, as opposed to a physical map or distance which is measured in basepairs. Recombination fraction (θ) is another expression of the frequency of recombination between two

markers, and is directly related to genetic distance. A recombination fraction of 0.0 means that the two markers are tightly linked, and will remain together after the recombination process whereas a recombination fraction of 0.5 means that the two markers are not linked, and will be subject to frequent recombination (50%). Thus, the goal of linkage analysis is to determine whether a genotype is tightly linked ($\theta = 0$) to a disease locus.

Linkage analysis can be divided into two types: parametric and nonparametric studies. Parametric analysis requires knowledge of the phenotype in question, specifically how rare the disease allele is (common or rare disease), if phenocopies are commonly seen, the correct mode of inheritance, or penetrance. Non-parametric analysis does not require any of these inputs and tests for polymorphic marker genotypes that are identical between all of the affected members of the family, called Identity By Descent (IBD). Both these analyses use a statistical measure of linkage, referred to as the Logarithm of the Odds ratio (LOD). The LOD score is a computation of two probabilities during the analysis: 1) The probability that the observed phenotypes in a family are due to genetic linkage between loci. 2) The probability of observing the phenotypes in a family assuming there is no linkage between loci ($\theta = 0.5$). A LOD score of 3 ($\theta = 0$) at a marker indicates a 1:1000 chance of this marker not being linked to the disease gene as $\log(1000) = 3$, and is considered the threshold for linkage. The minimum threshold for linkage exclusion is -2, meaning that any LOD score below -2 can be rejected from further experiments.

These analyses allow a researcher to determine the position on a specific chromosome that a disease gene may reside. This region is called the 'critical region' and is then studied in more detail. The use of both parametric and non-parametric analysis will be explained and used in Chapter 4 of this thesis.

(iv) Critical Region Refinement

Once the critical region of the genome has been determined, this region can be 'refined' or made smaller. Generally, a smaller number of widely spaced markers or SNPs are used for the initial linkage analysis, to maximize speed and efficiency of computations. Thus, the critical region may be very large in the first steps of analysis. To refine this region, a number of markers spaced closer together than those in the initial scan are genotyped. This step is made easier by analyzing as many individuals in a single (or multiple) family(ies) as possible. This allows a higher resolution haplotype to be constructed. This haplotype can then be visualized on a pedigree to search for any recombinations within the critical region that may reduce its size.

(v) Positional Candidate Gene Sequencing

Genes within the now refined critical region are termed 'positional candidates'. Once the critical region has been defined, the positional candidate genes are prioritized for screening. For example, if one is studying a hearing loss family, any genes involved in the functioning or formation of the inner ear would be prioritized, whereas genes involved in retinal function, cancer, or immune system function may be put lower on the

priority list. One method for screening genes is the Sanger Sequencing method, which involves the incorporation of dideoxy nucleotide chain terminators. This method was first introduced by Frederick Sanger in 1977,⁸² and has become the gold standard for sequencing. Once the list of positional candidates has been made, Sanger sequencing is performed using primers surrounding the gene's coding regions (i.e. areas of the gene that code for the corresponding protein). A smaller number of affected and unaffected family members can be used for this initial sequencing screen, to maximize efficiency. This will, ideally, identify a number of potentially pathogenic variants.

(vi) Variant analysis and Validation

After the screening of positional candidate genes, a number of variants will likely be identified. These variants undergo a two-fold analysis: 1) the variants will undergo segregation analysis, to look for correlation between the variant and the phenotype (and can be added to the critical region haplotype to potentially reduce the size of the critical region), and 2) The variants will be analyzed for any potential effect on the translated protein. These effects include nonsense mutations (premature stop codon), missense mutations (mutated amino acid), or splice variations which may alter the mature RNA transcript. Any mutations which are of particular interest will be screened in the remainder of family members. This step will exclude a number of variants depending on the number of available family members, leaving a smaller number of variants to analyze. Next, variants of interest can be screened in ethnically-matched population controls. Variants which are found to occur frequently in the study population can be excluded

from further analysis, as they are likely polymorphic variants (SNPs) rather than disease causing mutations. *In silico* analyses are also available to further characterize some types of mutations (i.e. SIFT, PolyPhen-2 for missense mutations, or BDGP splice prediction tool for splice site mutations). In order to further characterize splice site variations, RNA can be extracted from immortalized cell lines, or whole blood and then used to determine any altered splicing patterns. Ideally, if a variant which appears to be rare (based on database information and population controls), segregates with the disease, and has predictive deleterious effects upon *in silico* analysis, functional studies can then be carried out to determine the putative effect on the protein.

(vii) Next Generation and Exome Sequencing

Another, more recent, approach to gene discovery projects has been developed in the last five years. This new technology, called Next Generation Sequencing (NGS) or Massively Parallel Sequencing, allows for whole genome sequencing with a very rapid throughput. The major application of NGS technologies in the past two to three years has been that of exome sequencing. This technology is a revolutionary method to sequence all the coding regions of the genome (~1% of the genome) relatively quickly, while excluding the intronic and intergenic regions.⁸³ Exome sequencing has had many successes over the past 3 years, and has proven its utility in gene discovery for a variety of diseases including infantile hepatopathy, non-small cell lung carcinoma, inflammatory bowel disease,⁸⁴ and retinitis pigmentosa to name but a few. It has also shown great utility as a diagnostic tool, and has great advantages over linkage studies as no previous

information (though information on penetrance and mode of inheritance can ease analysis) is required for analysis. Despite these numerous advantages, analysis of exome sequencing data is not without its difficulties. Once exome sequencing has been carried out on a sample, a list of hundreds of thousands of variants is generated which can be difficult to deal with. To ease further analysis, this large list is filtered for common variants by comparing the exome dataset to previously established databases such as the 1000 Genomes project, dbSNP, the HapMap project, or the Exome Variant Server. These filters will remove many common, or already discovered SNPs and non-pathogenic variants, leaving a list of rare variations. This smaller list can then be filtered for variants with potentially pathogenic effects (nonsense, missense, or potential splice variations). This process is made easier by sequencing the exomes of multiple patient samples with the same disease, particularly from a single family. If multiple family members are available for analysis, a segregation analysis of potentially pathogenic variants discovered through exome sequencing can be performed. This is accomplished by confirming the variant via Sanger sequencing, then tracing the putative mutation through family members.

As previously stated, exome sequencing can also be used as a diagnostic tool if the disease phenotype is not well known, or if a particular feature of the phenotype is overlooked upon examination. An example of exome sequencing as a diagnostic tool is outlined in Chapter 3 of this thesis through a study of a Newfoundland family with achromatopsia.

1.7 Summary and Study goals

This thesis is an extension of research started by Dr. Jane Green over 30 years ago. Dr. Green reviewed the blind registration records maintained by the Newfoundland and Labrador division of the Canadian National Institute for the Blind (CNIB) and through ocular genetics clinics, obtained a plethora of family based information and pedigrees. This research project is a culmination of these three decades of work to finally solve the genetic basis for eleven of these families with three conditions. These conditions include Anterior Segment Dysgenesis (ASD; 1 family), Achromatopsia (ACHM; 8 families), and Microphthalmia-Dwarfism (MDW; 2 families). The discovery of novel genes and variants through this study can, in the future, provide patient care for these families as genetic and molecular therapies are developed and tested. This thesis approaches all three conditions from different standpoints, using all the techniques described in this introduction such as functional candidate screening, linkage analysis, and exome sequencing.

Chapter 2: A Novel non-stop mutation in FOXE3 causes an autosomal dominant form of variable anterior segment dysgenesis including Peters Anomaly

Lance Doucette¹, Jane S. Green¹, Bridget Fernandez¹, Gordon J Johnson¹, Patrick Parfrey¹, and Terry-Lynn Young¹

¹ Faculty of Medicine, Memorial University of Newfoundland

A shorter version of this chapter is published in the European Journal of Human Genetics 19: 293-299 (2011)

2.1 Summary

Anterior segment dysgenesis (ASD) is a spectrum of disorders which affect the anterior ocular chamber consisting of the cornea, iris, and lens. Clinical studies on a Newfoundland family over the past 30 years show that 11 relatives have a variable ocular phenotype ranging from microcornea to Peters anomaly, a severe form of ASD, segregating as an autosomal dominant trait. To determine the molecular etiology of the variable ASD in this family, we sequenced nine functional candidate genes associated with ASD (*PAX6*, *PITX2*, *PITX3*, *FOXC1*, *CYP11B1*, *CRYAA*, *GJA8*, *B3GALTL* and *FOXE3*) and identified 44 variants. Segregation analysis revealed a point mutation (c.959 G>T) in *FOXE3*, which codes for a transcription factor involved in the formation of the lens and surrounding structures, co-segregating with the variable dominant ocular phenotype. This novel mutation substitutes the stop codon for a leucine residue, predicting the addition of 72 amino acids to the C-terminus of FOXE3 (p.X359L). Two recent reports have also identified non-stop mutations in *FOXE3* in patients with variable ocular phenotypes and have predicted an extended protein. Although *FOXE3* is a lens-specific gene, we successfully isolated cDNA from lymphoblasts of an affected family member and our sequencing results show that the mutant c.959T allele is absent, suggesting that it may be degraded at the RNA level. Though preliminary, our results challenge the notion that an extended *FOXE3* protein causes ASD and instead suggests a mechanism of haploinsufficiency in the case of non-stop mutations. This study adds to several reports that suggest that autosomal dominant mutations within *FOXE3* cause ASD and has important clinical utility especially for the diagnosis of mildly affected patients.

2.2 Introduction

2.2.1 *What is anterior segment dysgenesis?*

Anterior segment dysgenesis (ASD) is a clinically heterogeneous phenotype and is a spectrum of various malformations of the anterior chamber generally affecting the cornea, iris, and lens. Examples of these disorders include microcornea, congenital cataracts, scleralization of the cornea, and a severe condition known as Peters anomaly.⁸⁵

Microcornea is a mild form of ASD and patients are considered to be affected if their cornea measurement is <11mm. This condition is rarely seen in isolation and is usually associated with other ASD phenotypes such as congenital cataracts (cataract-microcornea syndrome OMIM 116150), or Peters anomaly (OMIM 604229). Peters anomaly (PA) is a congenital condition, originally described in 1906,⁸⁶ and is characterized by an opaque cornea (called corneal leukoma), cornea-lens, and cornea iris adhesions and absent Descemet's membrane. PA was originally thought to be sporadic, though it has been described in both autosomal recessive and autosomal dominant inheritance patterns. This condition can be seen either in isolation, or as part of a syndromic condition such as Peters Plus Syndrome (Krause-Kivlin Syndrome OMIM 261540), a condition characterized by the ocular phenotype of Peters anomaly, but also extraocular features such as short stature, short fingers, and mental retardation.

2.2.2 The Genetics of ASD

ASD is a complex trait and can be caused by mutations in many different genes.²⁹

In addition, it seems as though the same mutation can cause many different phenotypes, suggesting other modifier or stochastic events are important for disease progression. The majority of the genes involved in the pathogenesis of ASD are important transcription factors involved in neural crest cell migration. These genes include *PAX6*,³¹ *PITX2*,^{87:88} *FOXC1*,⁸⁹ *FOXE3*^{90:91} and *PITX3*,^{92:93} which have been associated with disorders such as aniridia, Axenfeld-Rieger syndrome, Peters anomaly, and congenital cataracts. Genes that encode structural components of the eye, such as the crystallins (i.e. *CRYAA*)⁹⁴ or the gap junction gene *GJA8*,⁹⁵ have also been associated with various forms of ASD such as congenital cataracts and microcornea-cataract syndrome respectively. *CYP11B1*, a member of the cytochrome P450 family has been implicated in the pathogenesis of primary congenital glaucoma,^{96:97} as well as Peters anomaly.⁹⁸ *B3GALTL* has been shown to be the only gene involved in Peters Plus syndrome (Krause-Kivlin syndrome; OMIM #261540), a syndromic form of Peters anomaly.⁹⁹

(i) *PAX6* (OMIM 607108)

PAX6 is considered a 'master control gene' involved in the development of the brain and ocular structures.¹⁰⁰ *PAX6* homologues have also been described in Small Eye (*Sey*) mice, and in eyeless *Drosophila* (*ey*), indicating the conservation of this important gene across various species. This gene is also essential in forming the lens placode in mammalian development. Heterozygous mutations of *PAX6* have been implicated in many ASD conditions such as aniridia, cataracts, coloboma, and Peters anomaly. In very

rare cases, homozygous *PAX6* mutations have been shown to be involved in anophthalmia (lack of eyes), as well as lethal neonatal brain malformations, indicating the importance of gene dosage in this gene. This data also suggested that *PAX6* plays a role in the migration and differentiation of neural progenitor cells.³⁰

(ii) *PITX2* (OMIM 601542)

The *Pituitary Homeobox 2* gene (*PITX2*), located on 4q25, was isolated in 1996 through studies of a rare autosomal dominant condition called Axenfeld-Rieger Syndrome (ARS; OMIM #180500).¹⁰¹ This condition is characterized by malformations of the anterior chamber, which consists of microcornea, corneal opacities, iris hypoplasia, and posterior embryotoxon, in which 50% of patients develop glaucoma. Systemic anomalies are also described, namely dental hypoplasia, excess periumbilical skin and maxillary hypoplasia.¹⁰² *PITX2* is known to be expressed in the areas affected in ARS (umbilicus, mandibular epithelium, and the periocular mesenchyme) and is only expressed in the pituitary gland, though few cases of ARS show any signs of pituitary malfunction.

(iii) *PITX3* (OMIM 602669)

The *Pituitary Homeobox 3* (*PITX3*) gene, located on 10q25, is a member of the bicoid class homeobox transcription factors, like its relative, *PITX2*. When *PITX2* was isolated in 1996, this prompted a search for other genes of this nature and human *PITX3* was subsequently discovered in 1997.¹⁰³ *PITX3* was then sequenced in 80 patients with

ASD phenotypes and a 17-bp duplication was identified in family with an ASD phenotype (corneal opacities, iris-lens adhesions) and a missense mutation in one family with autosomal dominant cataracts.⁹³ More recently a 1-bp deletion in the coding region of *PITX3* was found in a family of English origin with autosomal dominant posterior polar cataract.¹⁰⁴

(iv) *FOXC1* (OMIM 601090)

The *Forkhead Box Domain C1* (*FOXC1*), a single exon gene located on 6q25.3, is a member of a large group of transcription factors with a common DNA binding domain termed the 'forkhead domain'. This gene was isolated in 1994 as part of a search for forkhead domain genes in humans, as the forkhead domain was already known to *Drosophila*, rats, metazoans, and yeast.¹⁰⁵ Since its discovery, mutations in *FOXC1* have been implicated in various ocular disorders such as primary congenital glaucoma,^{106: 107} ARS, Peters anomaly^{108: 109} and most recently, in the progression/invasion of certain types of breast cancer through regulation of genes like *NF-κB* (an important transcription factor in immune function/responses),¹¹⁰ or *MMP17* (responsible for extracellular matrix remodeling).¹¹¹

(v) *FOXE3* (OMIM 601094)

The *Forkhead Box Domain E3* (*FOXE3*), a single exon gene located on 1p33 was discovered in 1995¹¹² and was later found to be highly expressed in the lens vesicle and lens placode during development, but only in the lens epithelium post-natally.¹¹³ This

transcription factor is primarily responsible for the differentiation and proliferation of cells within the lens, and mutations can affect the formation of this important structure. Mutations within human *FOXE3* have been shown to cause aphakia (lack of lens development) when in the homozygous state^{114, 115} as well as other conditions such as sclerocornea, microphthalmia, and optic disc coloboma.¹¹⁶ It has also been shown to cause various forms of ASD when inherited in an autosomal dominant manner (i.e. heterozygous transmission) such as congenital cataracts and posterior embryotoxon.⁹¹

(vi) *CRYAA* (OMIM 123580)

The *Crystallin alpha-A* (*CRYAA*) gene, located on 2q22.3, is one of a large group of crystallin genes, responsible for formation and maintenance of transparency of the mammalian lens. These proteins share high sequence homology with the heat shock protein (Hsp) family of proteins, which form aggregates to sequester harmful proteins upon injury to protect the cell under stressful conditions.¹¹⁷ Mutations of *CRYAA* have been shown to be the cause of autosomal dominant congenital cataract in a number of cases,¹¹⁶⁻¹¹⁸ and one report showed that mutation of this gene can affect the ability of *CRYAA* to form aggregates, causing microcornea-cataract syndrome (OMIM 116150).¹¹⁹ Interestingly, it has been suggested that due to its position on chromosome 21, *CRYAA* may play a role in the cataracts seen in Down Syndrome patients (trisomy 21).

(vii) *GJA8* (OMIM 600897)

The *Gap Junction Alpha 8* (*GJA8*) gene, located on 1q21.2, produces a gap junction protein, connexin 50 (Cx50) which is present in all cells of the lens, first being synthesized in the lens epithelium.¹²⁰ When lens epithelial cells differentiate into secondary fiber cells they lose their intracellular organelles, which otherwise might cause scattering of incoming light, in order to maintain transparency in the lens.¹²¹ Due to lack of organelles, the inner lens fiber cells, have developed a complex network of connexin proteins to facilitate proper metabolism, and to transport small metabolites or secondary messengers between cells.^{122: 123} Mutations of *GJA8* have been associated with zonular pulverulent cataracts,¹²⁴ nuclear progressive cataracts,¹²⁵ as well as cataract-microcornea syndrome.¹²⁶

(viii) *CYP1B1* (OMIM 601771)

CYP1B1, isolated in 1994 and located on 2p22.2, is a member of the Cytochrome P450 superfamily of genes which are responsible for xenobiotic and drug metabolism.¹²⁷ *CYP1B1* has been primarily associated with cases of primary congenital glaucoma,^{128: 129} and is suspected to be involved in the metabolism of a yet-unknown molecule involved in proper eye development.⁹⁷ Other studies have also implicated mutations in *CYP1B1* in patients with Peters anomaly.^{98: 130} This gene is also involved in the hydroxylation of 17-beta-estradiol, a compound postulated to be involved in breast cancers.

(ix) *B3GALTL* (OMIM *610308)

Beta-1,3-Galactosyltransferase-Like (B3GALTL) was cloned in 2003, and codes for a 498 amino acid protein¹³¹ responsible for the synthesis of a disaccharide found on repetitive elements (thrombospondin type-1 repeats) of many biologically important proteins.¹³² This gene has been associated with a rare autosomal recessive condition termed Peters Plus Syndrome (AKA Krause-Kivlin Syndrome; OMIM 261540). This condition is characterized by the anterior segment dysgenesis phenotype of Peters anomaly but with other extraocular features such as disproportionate short stature, mental retardation, craniofacial malformations, and cleft lip/palate. The association of *B3GALTL* with Peters plus syndrome was first shown in a study using array-based comparative genomic hybridization (aCGH) which identified an interstitial deletion surrounding *B3GALTL*.⁹⁹ Since then, a number of other mutations have been described in *B3GALTL* as causing Peters plus syndrome.¹³³⁻¹³⁵

2.2.3 Current Study

We previously reported a proband diagnosed with congenital Peters anomaly and a family history of a variable form of autosomal dominant ASD, including microcornea, congenital cataracts, and scleralization of the iris.¹³⁶ This study was undertaken to identify the genetic etiology of ASD in this Newfoundland family. Throughout the course of this study, I not only identify the causative gene but also present the progression of the ocular phenotype observed over a 30-year period and describe the functional candidate approach used to determine the causative gene.

2.3 Materials and Methods

2.3.1 *Clinical Investigation of Proband and Relatives*

We studied a 4-generation family from the island of Newfoundland with mild to severe forms of ASD segregating as an autosomal dominant trait (Figure 2.1). The proband was diagnosed at birth in 1979 with Peters anomaly.¹³⁶ Long-term clinical and ophthalmic examination of family members has continued since that time (Table 2.1) and has included visual acuity measurements, refraction, direct and indirect ophthalmoscopy, slit lamp examination, tonometry, gonioscopy, corneal diameter measurements, axial length measurements (ultrasound) and external eye photographs as previously described.¹³⁶ Thirty-one of 66 known family members participated in this molecular genetics study (though clinical examinations were performed on a total of 52 family members), and of these, we have identified 11 cases of ASD across six sibships (Figure 2.1). Due to the variability of ocular phenotypes seen in this family, we (JG, GJJ) conducted repeated, complete slit lamp examination of the cornea, iris, angle and lens, and measurement of the corneal diameter on a number of individuals. These examinations provided evidence for minimal diagnostic criteria which included either corneal scleralization, lens opacities, strands in the angle or small corneal diameter (<11mm). This study was approved by the Human Investigations Committee, Memorial University, St. John's, Newfoundland, Canada (HIC #02.116).

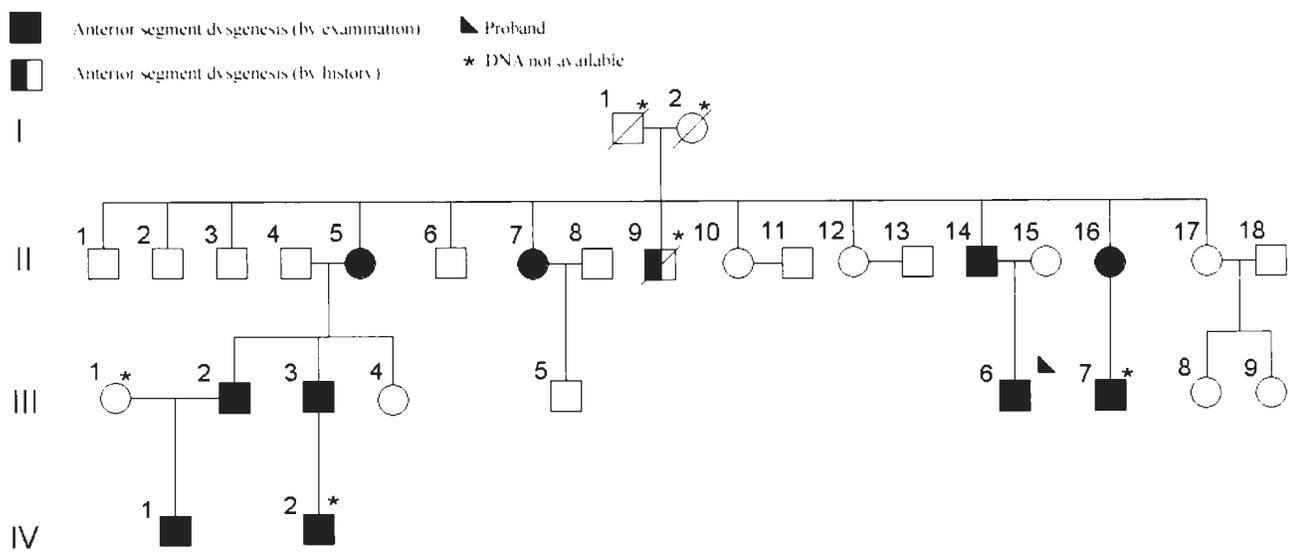


Figure 2.1: A four generation Newfoundland family segregating an autosomal dominant form of anterior segment dysgenesis (ASD). The proband (III-6; arrow) has congenital Peters anomaly. The phenotype in this family ranges from a mild anterior segment phenotype such as microcornea, to a very severe form called Peters anomaly.

Table 2.1 – Clinical examination of affected members from Family 0023

PID	Age at last examination	Cornea	Corneal Diameter (mm)	Lens	Axial Length (mm)	Visual Acuity
III-6	30	Right - Peters Anomaly Left - Peters Anomaly	R - 9.25 L - 9.5	Cataract and adhesion from lens to cornea	R - 20.1 L - phthisical	R - 6/30 L - NLP
II-14	57	Moderate scleralization	R - 10 L - 10	PSC and nuclear cataracts. Extractions at 39 and 40	R - 21.2 L - 20.9	R - 6/7.5 ³ L - 6/9
II-5	68	Moderate scleralization	R - 10 L - 10	Cataract extractions at 36 and 38	R - 21.5 L - 22.2	NA
II-7	65	Moderate scleralization	R - 8.5 L - 8.5	Cataract extractions at 32 and 36	R - 21.3 L - Enucleated	HM NLP
II-16	59	NA	R - MC L - MC	Cataract extractions at 27 and 29	NA	NA
III-7	20	Mild scleralization	R - NA L - NA	AC, PSC, diffuse SC cataracts. Extractions at 16	NA	NA
III-2	41	Mild scleralization	R - 10.5 L - 10.5	AC, PSC Cataract extractions at 40 years	R - 21.8 L - 22.3	R - 6/9 L - 6/6
III-3	39	Moderate scleralization	R - 9.5 L - 9.75	AP, Nuclear, PSC cataracts. Cataract Extractions at 21	R - 21.8 L - 22.8	R - 6/9 L - 6/9
IV-2	10	Left corneal opacity	NA	AP cataracts	NA	R - 6/6 L - 6/9
IV-1	6	Central Corneal opacities	R - NA L - 9	AP cataract extractions at 8 months	NA	R - 6/18 L - 6/18

PID = Pedigree Identifier **AC** = Anterior cortical; **AP** = Anterior polar; **PSC** = Posterior subcapsular; **SC** = Subcapsular; **MC** = Microcornea; **HM** = Hand movements **NLP** = No light perception; **NA** = Not available

2.3.2 Selection and Screening of Functional Candidate Genes

Genes involved in the structural development of the eye or with previous association with ASD, such as Peters anomaly and related disorders (i.e. Axenfeld-Rieger syndrome, Krause-Kivlin syndrome) were sequentially screened for mutations, namely *PAX6*, *PITX2*, *FOXC1*, *PITX3*, *CYP1B1*, *B3GALTL*, *CRYAA*, *GJA8*, and *FOXE3*. *PAX6* encodes a transcription factor which has been implicated in aniridia and Peters anomaly.^{31: 137} *PITX2* (*RIEG1*) is primarily associated with Axenfeld-Rieger syndrome, a syndromic form of ASD with similar ocular phenotypes,⁸⁸ though some mutations have been identified as causing non-syndromic Peters anomaly.⁸⁷ *PITX2* interacts with the transcription factor *FOXC1*^{89: 138} and mutations in *FOXC1* produce a similar phenotype as seen in *PITX2* mutants. *PITX3* has been associated with autosomal dominant congenital cataracts and Peters anomaly.^{92: 93: 139} Mutations in *B3GALTL*, a galactosyl transferase gene, cause Peters plus syndrome, which consists of Peters anomaly, short stature, cleft lip, cleft palate and mental retardation.¹³³ Though there have been no reports of incomplete or atypical forms of Krause-Kivlin syndrome showing only the ocular phenotype, *B3GALTL* was screened in this family to check for possible hypomorphic alleles which may cause the non-syndromic phenotype seen in this family. *CRYAA*, a member of the crystallin family which constitutes the major structural protein in lens fiber cells, has been associated with autosomal dominant cataracts, microcornea, and corneal opacities.^{94: 140} *GJA8* encodes a gap junction protein which regulates ion and crystallin concentrations within the lens fiber cells.⁹⁵ This gene has been associated with cataract-microcornea syndrome, and various other forms of cataracts. *CYP1B1* is a member of the cytochrome P450 superfamily, and mutations in this gene have been

attributed to primary congenital glaucoma, and Peters anomaly.^{29: 98: 130} Finally, *FOXE3* is a lens-specific transcription factor which has been implicated in cases of ASD including autosomal dominant Peters anomaly,⁹⁰ congenital cataracts,⁹¹ and autosomal recessive primary congenital primary aphakia.¹¹⁴

2.3.3 DNA Isolation and Primer design

Genomic DNA was isolated from peripheral leukocytes by DG¹⁴¹ and screened for mutations in the proband and five relatives with ocular phenotypes representing the full range of phenotypes found in this family (Figure 2.1: (Pedigree Identifier (PID) II-5; II-14; III-2; III-6; IV-1), and two unaffected individuals (PIDs: II-10; II-15). Oligonucleotide primers were designed using the Primer 3 software (<http://frodo.wi.mit.edu/>) to amplify all coding regions, and UTRs of the following genes: *PAX6* (NM_001604), *PITX2* (NM_153426), *PITX3* (NM_005029), *CYP11B1* (NM_000104), *FOXC1* (NM_001453), *B3GALTL* (NM_194318), *CRYAA* (NM_000394), *GJA8* (NM_005267), and *FOXE3* (NM_012186).

2.3.4 PCR Setup

Polymerase chain reaction (PCR) was set up to amplify targets from genomic DNA (gDNA) using the KapaTaq PCR system. DNA samples were diluted to 10 ng/uL and 1 uL was added to a 19 uL of KapaTaq master mix (Appendix A). Betaine (1.75mM), a PCR enhancer which acts by disrupting hydrogen bonds to reduce secondary structure in DNA, and assist in denaturation was added to each master mix at 25% per

volume. Dimethylsulfoxide (DMSO) was used in instances where a sequence was particularly difficult to amplify due to CG rich composition at 5% per volume (i.e. *FOXC1* and *FOXE3* are CG rich genes which can be difficult to amplify). DMSO acts to disrupt the three hydrogen bonds between C-G bases, like that of Betaine and was used in combination with Betaine if the initial PCR reaction did not work. All plates were run on a TouchDown 54 (TD) program which begins at 64 degrees Celsius, to provide high specificity of primer annealing, and over five cycles drops to 54 degrees Celsius (Appendix B). Amplicons were run on a 1% agarose gel stained with SYBR Safe, which binds to DNA and fluoresces under UV light, (Invitrogen by Life Technologies) for size verification. Primer sequences for each gene can be found in Appendix F. Amplicons were purified using 50% S300HR sephacryl (Amersham Biosciences) and Multiscreen HTS filter plates (Millipore Corporation). Filter plates were prepared by addition of 300 uL of Sephacryl, suspended in Tris-EDTA (TE) buffer, to the plates, these filter plates were then placed on a 'catch plate' and were spun at 3000 rotations per minute (rpm) for 5 minutes. The contents of the catch plate were discarded, and the filter plate was placed on a 96 well PCR plate. DNA samples were added to each well of the filter plate, and spun for 5 minutes at 3000rpm.

2.3.5 Bidirectional Sequencing

Purified DNA (1 uL) was added to 19 uL of BigDye Terminator Kit v 3.1 (Applied Biosystems) master mix (setup can be found in Appendix C) in a 96 well LightCycler sequencing plate. The plates were placed on a GeneAmp thermocycler and

run on the program ABI.SEQ (cycling conditions can be found in Appendix D). Once cycling was complete, precipitation was carried out as an extra purification step to remove excess salts and proteins. Sixty-five uL of 95% ethanol, and 5 uL of 125 mM EDTA was added to each well of the LightCycler sequencing plates, which were then placed at -20 degrees Celsius for at least 30 minutes. The plate was then spun at 3000rpm for 30 mins. The ethanol/EDTA solution was then decanted from the plate, and was spun upside down on a paper towel up to 200rpm to remove excess solution. A wash step was then performed using 150uL of 70% ethanol. The plate was spun at 3000rpm for 15 minutes and the ethanol decanted and removed from the plate. The plate was allowed to dry, then 15 uL of HiDi Formamide was added to each well, and then placed on a GeneAmp thermocycler at 95 degrees Celsius for 2 minutes, then placed on ice. The plates were run on an ABI 3130xl DNA Analyzer. Electropherograms were inspected manually for quality and imported into Mutation Surveyor software v 3.2 (Transition Technologies) to detect sequence variants. All sequencing variants were checked for co-segregation with the ASD phenotype in the family. Only one allele (*FOXE3* c. 959G>T) co-segregated with the disease phenotype, so the population frequency of this variant was determined using population controls from Newfoundland.¹⁴²

2.3.6 RNA Isolation/cDNA synthesis

One mL of Trizol reagent was added to 1.0×10^6 Epstein Barr virus (EBV)-transformed B lymphocytes from one affected individual (PID IV-1), in a 2.0 mL Eppendorf tube. The solution was pipetted up and down to mix the solution, and was

allowed to incubate at room temperature for 5 minutes. Two-hundred uL of chloroform was added to the tube and the tube was shaken vigorously for 15 seconds then incubated at room temperature for three minutes. The tubes were centrifuged at 12,000x g for 10 minutes at 4 degrees Celsius, and when removed were separated into an aqueous (top), interphase (middle), and organic layers (bottom). The aqueous phase was transferred to a 1.5 mL Eppendorf tube and the other phases discarded. Five-hundred uLs of isopropanol were added to the the aqueous phase to precipitate the RNA from solution. The tube was incubated at room temperature for 10 minutes and was then centrifuged at 12,000 x g at four degrees Celsius for 10 minutes. A waxy pellet was observed on the bottom of the tube post-centrifugation, and the isopropanol solution was decanted from the tube and the pellet allowed to dry at room temperature. The pellet was resuspended in 25 uL of deionized water and was followed by addition of 2 uL of DNaseI (Ambion) followed by incubation at 37 degrees Celsius for 30 minutes to remove any gDNA contamination. Complementary DNA (cDNA) synthesis was performed using the Superscript III cDNA synthesis kit (Invitrogen by Life Technologies). Up to 5 ug of RNA (depending on the yield of the RNA extraction) was added to 8 uL of a mastermix cocktail prepared as per Appendix G using oligo dT primers. PCR was carried out using primers surrounding the stop codon of *FOXE3* (FOXE3-cDNA-F and -R) and was analyzed by size fractionation using agarose gel electrophoresis and direct sequencing. A set of primers designed to amplify the intron between exons 5 and 6 of *GAPDH* (a widely expressed gene responsible for function of energy production via glycolysis) were used as an internal control to check for the presence of genomic DNA contamination. These primers amplify

a 170 bp fragment with a cDNA template, and a 278 bp fragment using a genomic template. Primers and experimental conditions are listed in Appendix F.

2.4 Results

2.4.1 Long-term Clinical Follow-up

The proband (PID III-6) is a 33-year old male born at 35 weeks gestation (Figure 2.1). Bilateral dense corneal opacities were noted at birth and were initially attributed to maternal prenatal infection. Examination under anesthetic showed corneal opacification of the proband's right eye from the temporal side past the midline and hazy cornea nasally (Table 2.1; Figure 2.2A). Fine adhesions from the collarette of the iris to the cornea obscured the angle. The proband's left cornea was more densely opaque with a large central adhesion from the cornea to the lens and peripheral adhesions from the iris to the cornea (Figure 2.2B). Both globes were of normal size but corneal measurements were 10mm horizontally and 9mm vertically ($N \geq 11$ mm). At six months of age the proband had a corneal transplant of the more severely affected eye (left) and optical iridectomy of the right eye. Pathology review of the left corneal button indicated absent Descemet's membrane, partial absence of Bowman's membrane and thinning of the central cornea consistent with Peters anomaly. The corneal graft was rejected 6 weeks post-operatively and the eye became phthisical following subsequent attempts at re-grafting.

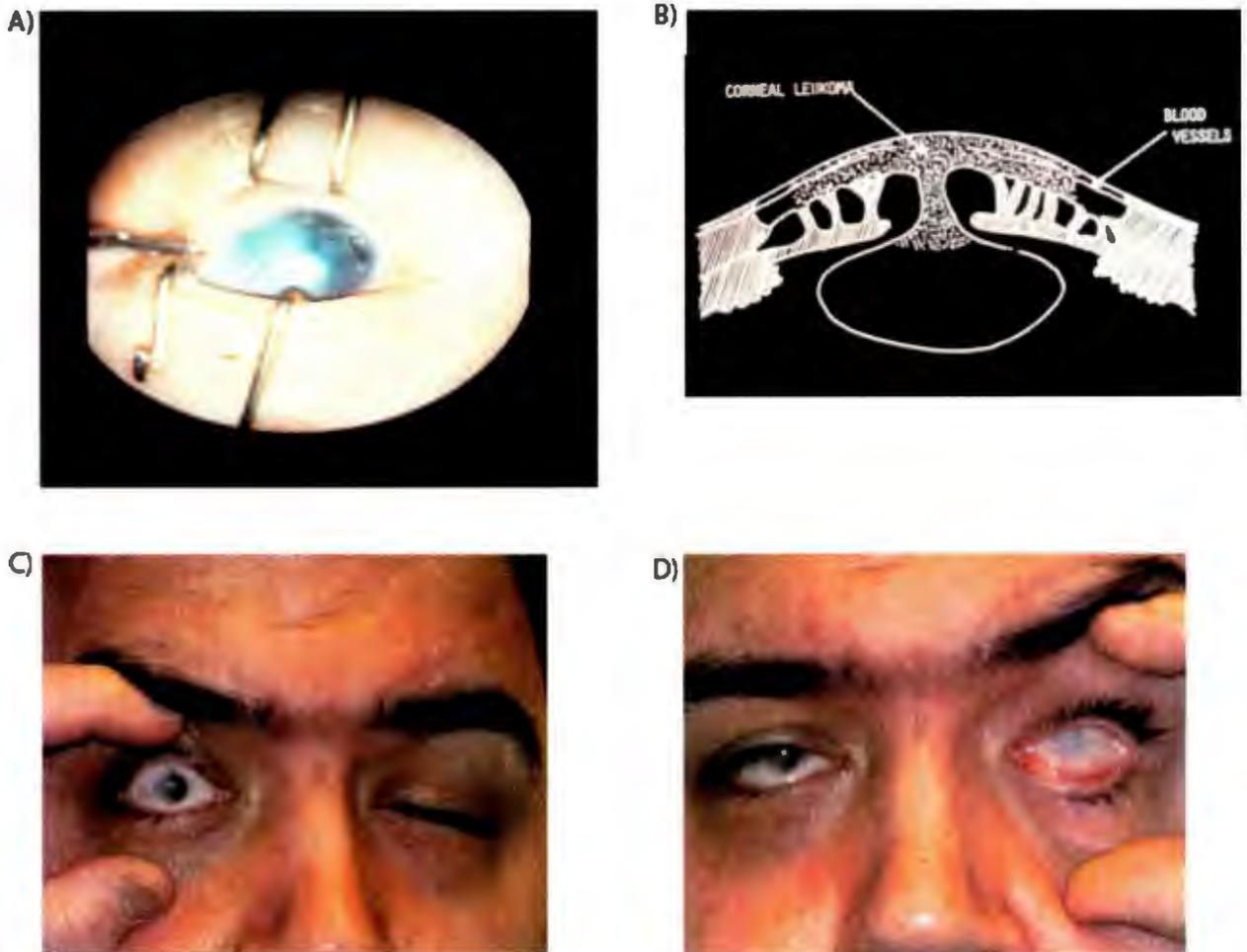


Figure 2.2: Eye phenotype of the proband (PID III-6) over time. **A)** Picture of the proband's right eye at 7 weeks showing corneal opacities consistent with Peters anomaly. **B)** Surgical drawing of the proband's eye showing keratolenticular adhesion, and corneal leukoma (Pictures A and B used with permission). **C)** Proband's right eye at 30 years of age. **D)** Proband's phthisical left eye at 30 years of age.

The proband was re-examined at 30 years of age by a clinical geneticist (BF; Figures 2.2C & 2.2D) and found to be of normal intelligence. His height was 172 cm (25%ile); ratio of upper to lower segment 0.98; arm span 170 cm; weight 107kg (>95%ile); and head circumference was 62 cm (this information was not available for other relatives). There was no facial dysmorphism apart from features related to a phthisical left eye due to post-operative complications: the left eyebrow was lower and the left palpebral fissure was shorter than that of the right eye. His palate was intact with a single uvula, and the philtrum was well-developed (philtrum length 2 cm, 50-75%ile). Ears were large (length of 7.3 cm), but were normal in position and contour. The only other minor physical anomaly was clinodactyly of the right 2nd and 3rd fingers and of the left 2nd finger. The lack of extra-ocular phenotypes seen in this patient allowed us to rule out Krause-Kivlin syndrome or other related syndromes as a possible diagnosis for this individual.

The proband's father (PID II-14) was examined at 27 years of age (the year the proband was born) and shown to have small posterior subcapsular and central nuclear cataracts, first noted at age seven (Table 2.1). His visual acuity was 6/7.5⁻³ in the right eye and 6/9 in the left eye. He had bilateral microcornea (corneal diameter of 10 mm) and scleralization and vascularization of the cornea, particularly superiorly. Upon gonioscopy, fine iris processes were noted extending over the trabecular meshwork. Subsequently, he had cataract extractions at ages 39 and 40 because of decreasing visual acuity.

Three paternal aunts (PIDs II-5; II-7; II-16) who had lens opacities documented in childhood had cataract extractions prior to 1979 (in their twenties or thirties) when mature cataracts developed. All three individuals presented with microcornea (corneal diameters of 8.5 mm – 10.5 mm), mild to moderate scleralization of the cornea with varying degrees of vascularization, particularly superiorly and inferiorly (Table 1). Three paternal cousins (PIDs III-2; III-3; III-7) had cataract extractions at ages 16-40 years. The cataracts were originally described as anterior polar, anterior cortical, and nuclear and posterior subcapsular cataracts. Six of these seven affected family members had favorable results post-operatively although one had vitreous hemorrhage, choroidal detachment and temporary hypotony of one eye. Individual PID II-7 had serious post-operative complications including retinal detachments with failed repair, failed corneal grafting and enucleation of a painful blind eye with hand movement vision only in the remaining eye (Table 2.1).

The youngest affected family members are PID IV-1 (6 years old) and PID IV-2 (10 years old) (Figure 2.1). Individual PID IV-2 had anterior polar cataracts detected at birth but has not yet required cataract extraction. Individual PID IV-1 had a complicated pre and postnatal course with extreme premature birth at 24 weeks gestation. He was hospitalized for 4 ½ months during which time he had necrotizing enterocolitis with perforation of the bowel, and successful treatment of retinopathy of prematurity. Because of the size and central location of his anterior polar cataracts, he had cataract extractions at 8 months of age (Table 2.1). He also had bilateral iridectomies due to central corneal opacities. His central acuity is recorded as 6/18 but his course continues to be complicated with a recent diagnosis of autism.

A paternal uncle who died of pneumonia at 6 months of age in the 1940s was likely affected, as he had 'white eyes' and was registered blind with the Canadian National Institute for the Blind (PID II-9). Eight other paternal aunts and uncles and 15 paternal cousins had normal corneal diameters, and clear corneas and lenses upon examination (although some at older ages had a prominent arcus suggesting elevated cholesterol levels).

Interestingly, there is no evidence that either of the grandparents had any form of ASD (DNA not available). For instance, the proband's paternal grandfather (PID I-1) was examined and had corneal diameters of 11 mm, mild corneal scleralization, inferior arcus, and faint lens opacities including anterior polar specks, and anterior and posterior cortical spokes but retained 6/7.5 and 6/6 visual acuity at 77 years of age. Two of his brothers had clear corneas and lenses in their 60s or 70s. The proband's paternal grandmother (PID I-2) had corneal diameter of 11.5 mm, mild central endothelial corneal changes (Fuch's dystrophy), arcus senilis, and faint dot and spoke lens opacities when examined at 71 years. She subsequently had mature cataracts extracted at ages 78 and 81. Upon examination, her brother, sister, niece, and nephew all had

clear corneas and lenses, and were considered unaffected. The phenotypes seen in the grandparents can all be attributed to the aging process. It is possible that one had a subclinical phenotype or that gonadal mosaicism was present.

2.4.2 Candidate Gene Screen Reveals Causative Variant in *FOXE3*

Screening of the nine functional candidate genes in seven individuals (five affected; two unaffected PIDS: III-6, II-14; II-5; III-2; IV-1; II-10; II-15) revealed 44 sequence variants (Table 2.2). When possible, variants were arranged into manually constructed haplotypes to check for segregation of a disease haplotype (Figures 2.3-2.10). Single variants were also checked for segregation. This analysis showed co-segregation (Figure 2.10; Yellow Haplotype) between the ASD trait and a non-stop mutation within the transcription factor gene *FOXE3* (c.959 G>T; p.X320L; Figure 3B) which caused the elimination of the functional opal stop codon (UGA) at the 3' end of *FOXE3*. The next available stop codon is 213 bps downstream (in the 3' UTR) predicting the addition of 72 amino acid residues to the C-terminus of *FOXE3*.

We then genotyped all available members of the extended family and found that the variant co-segregated with the disease phenotype (Table 2.2). This variant was not detected in 141 ethnically-matched controls. To test whether this mutation would be present in the mRNA of *FOXE3*, RNA was isolated from a lymphoblast cell line of individual PID IV-1 and converted to cDNA. *FOXE3* is a lens-specific gene, found only in the lens epithelium,⁹⁰ as a result this approach was not guaranteed to succeed. However, we were able to amplify a 458 bp cDNA product from *FOXE3* surrounding the c.959 G> T mutation (Figure 2.12A). Direct sequencing revealed that the c.959 G>T mutation was absent in the cDNA (Figure 2.12B), suggesting that

the mRNA transcribed from the 'non-stop' allele may be degraded before being translated, or it is possible that the RNA is not transcribed at all.

Table 2.2: Summary of sequencing variants identified in nine functional candidate genes sequenced in five affected members with a range of ASD from mild phenotype to Peters anomaly, and 2 unaffected members. The novel pathogenic variant discovered in *FOXE3* (c.959 G>T: p.X320L) is shown in bold.

Gene	Sequence Variants	Location of Variant	Protein Effect	dbSNP ID	Segregation
PAX6 (11p13)	c.-4393 G<A	5' UTR	None	Not listed	No
	c.1074 + 107 C<T	Intronic	None	rs3026384	No
	c.1075 - 174 G<A	Intronic	None	rs2071754	No
	c.1225 + 44 T<G	Intronic	None	rs3026393	No
	c.1746 + 21 delA	Intronic	None	Not listed	No
PITX2 (4q25)	c.-485 C<G	5' UTR	None	rs2739200	No
	c.629 - 105 C>A	Intronic	None	Not listed	No
	c.767 - 81 A>C	Intronic	None	Not listed	No
PITX3 (10q24.32)	c.285 C<T	Coding	I95I	rs2281983	No
FOXC1 (6p25.3)	c.1123_1124 ins CGG	Coding	ins 375 G	rs71807729	No
	c.1361 ins CGG	Coding	ins 456 G	Not listed	No
CYP11B1 (2p22.2)	c.-12 C<T	5' UTR	None	rs2617266	No
	c.142 C<G	Coding	R48G	rs10012	No
	c.355 G<T	Coding	A119S	rs1056827	No
	c.1294 G<C	Coding	V432L	rs1056836	No
	c.1347 T<C	Coding	D449D	rs1056837	No
	c.1358 A<G	Coding	N453S	rs1800440	No
	c.1633 + 210ins T	3' UTR	None	Not listed	No
	c.1633 + 350 C>A	3' UTR	None	Not listed	No
	c.1633 + 693 G>T	3' UTR	None	Not listed	No
	c.1633 + 809 del A	3' UTR	None	Not listed	No
	c.1633 + 875 A>G	3' UTR	None	Not listed	No
	c.1633 + 2408 A<T	3' UTR	None	Not listed	No
B3GALTL (13q12.3)	c.71 -5 del T	Intronic	None	Not listed	No
	c.271 - 58 delT	Intronic	None	Not listed	No
	c.271 - 67 T>C	Intronic	None	Not listed	No
	c.347 + 20 C<G	Intronic	None	Not listed	No
	c.348 T<C	Coding	H116H	Not listed	No
	c.596 + 20 delC	Intronic	None	Not listed	No
	c.597 - 168 delA	Intronic	None	Not listed	No
	c.780 + 58 A<G	Intronic	None	Not listed	No
	c.781 -124 G<C	Intronic	None	Not listed	No
	c.781 -31 ins CATA	Intronic	None	Not listed	No
	c.1065 - 142 T<C	Intronic	None	rs1041073	No
	c.1108 G<A	Coding	E370K	Not listed	No
	c.1497 + 31 G>T	3' UTR	None	Not listed	No
	c.1497 + 37 A>G	3' UTR	None	Not listed	No
	c.1497 + 127 G>A	3' UTR	None	Not listed	No
	c.1497 + 938 delT	3' UTR	None	Not listed	No
	c.1497 + 1548 C>T	3' UTR	None	Not listed	No

CRYAA (21q22.3)	c.6 C<T	Coding	D2D	Not listed	No
GJA8 (1q21.1)	c.1365 + 26 delA	Intronic	None	Not listed	No
FOXE3 (1p33)	c.959 G>T	Coding	X320L	Not listed	Yes
	c.960 + 72 T>C	3'UTR	None	rs2820969	No
	c.960 + 77 A>G	3'UTR	None	rs6666370	No

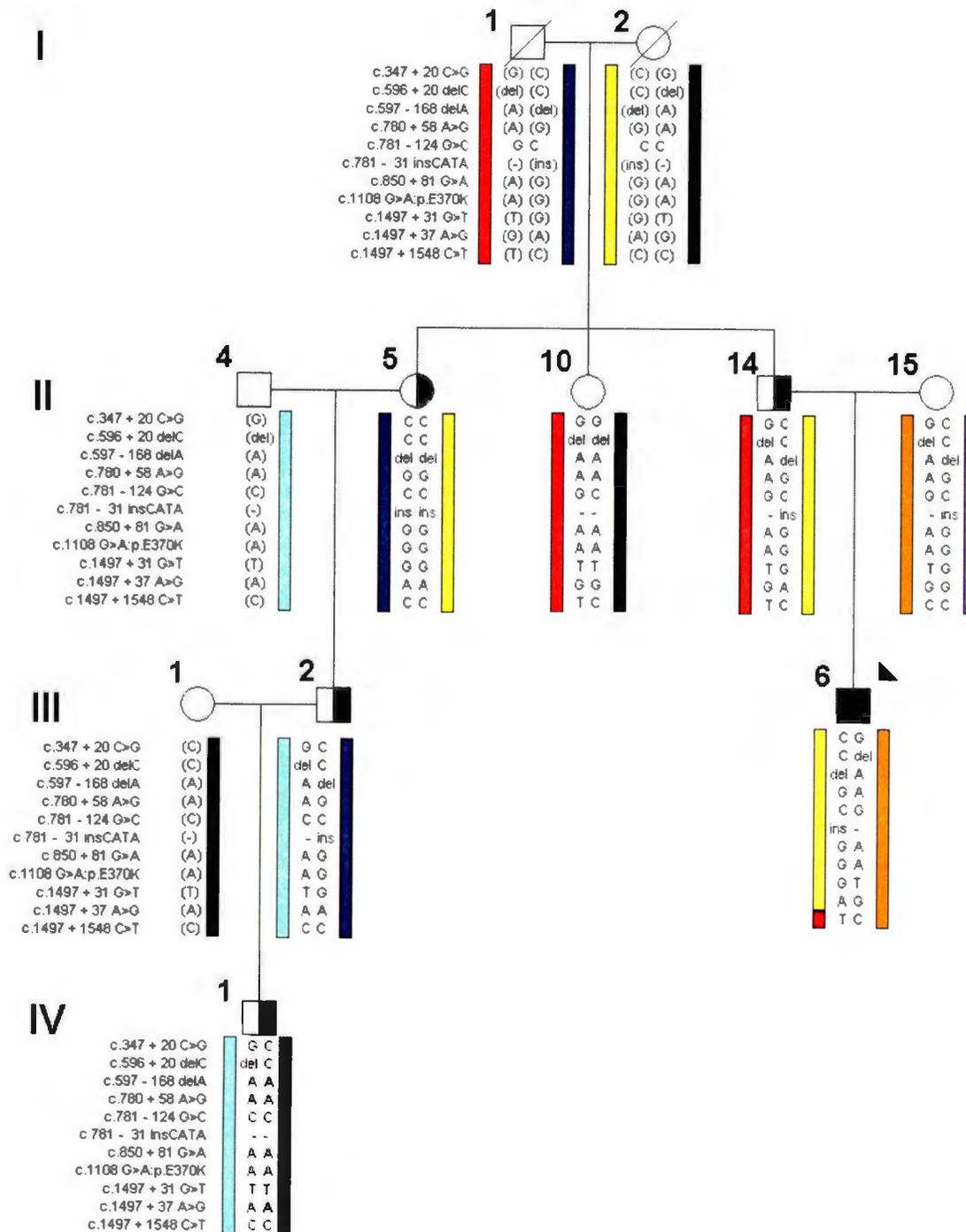


Figure 2.3 – Haplotype analysis of *B3GALTL* variants in family 0023. The proband is indicated by an arrow to the top left. Full shaded individuals have Peters anomaly whereas half-shaded individuals have other ASD phenotypes. Haplotypes were manually constructed using variants discovered through direct sequencing. Exclusion of *B3GALTL* as the causative gene in family 0023 can be observed through the lack of a common haplotype between all five affected individuals. The yellow haplotype is shared by three affected individuals (II-5, II-14, and III-6) but is not passed onto III-2 or IV-2, thus excluding *B3GALTL* as the causative gene in family 0023.

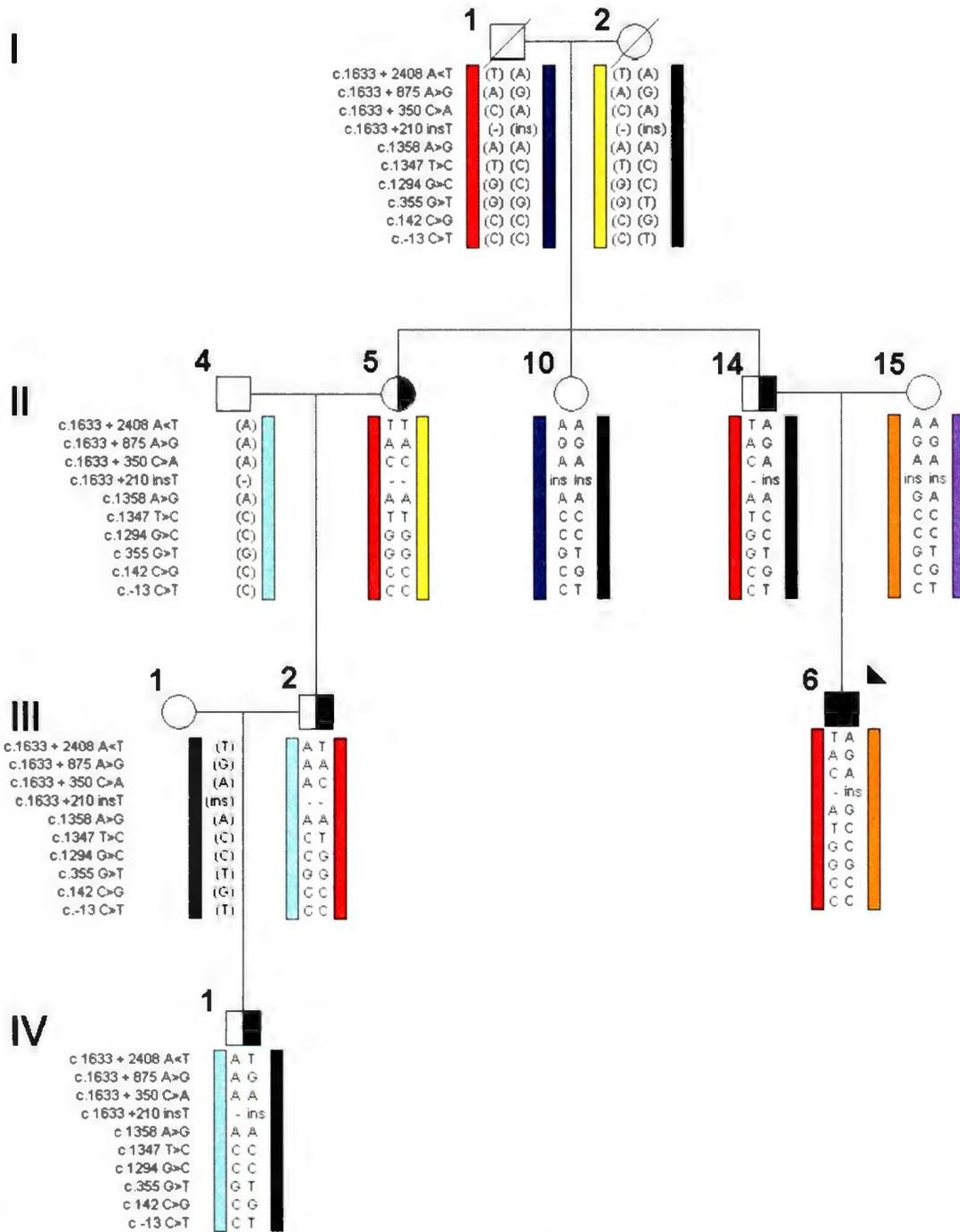


Figure 2.4 – Haplotype analysis of *CYP1B1* in family 0023. *CYP1B1* can be excluded as the causative gene in family 0023 as the red haplotype which is shared amongst four of the affected individuals is not passed onto IV-1.

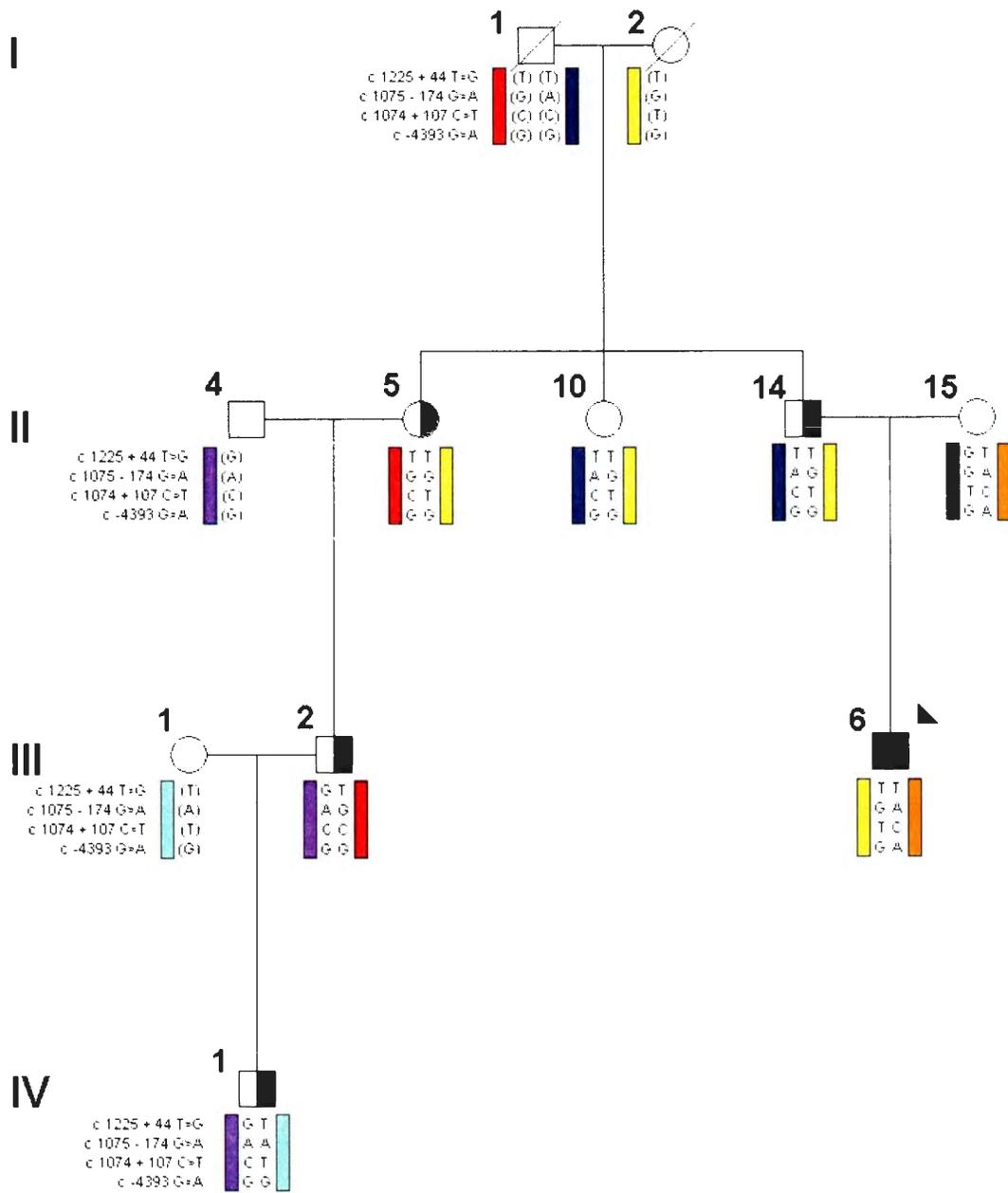


Figure 2.5 – Haplotype analysis of *PAX6* in family 0023. *PAX6* can be excluded in this family as no affected individuals share common haplotypes. The yellow haplotype exists in three affected individuals, but is also present in unaffected individual II-10. The red haplotype only exists in two affected individuals (III-2 and II-5) but in no others.

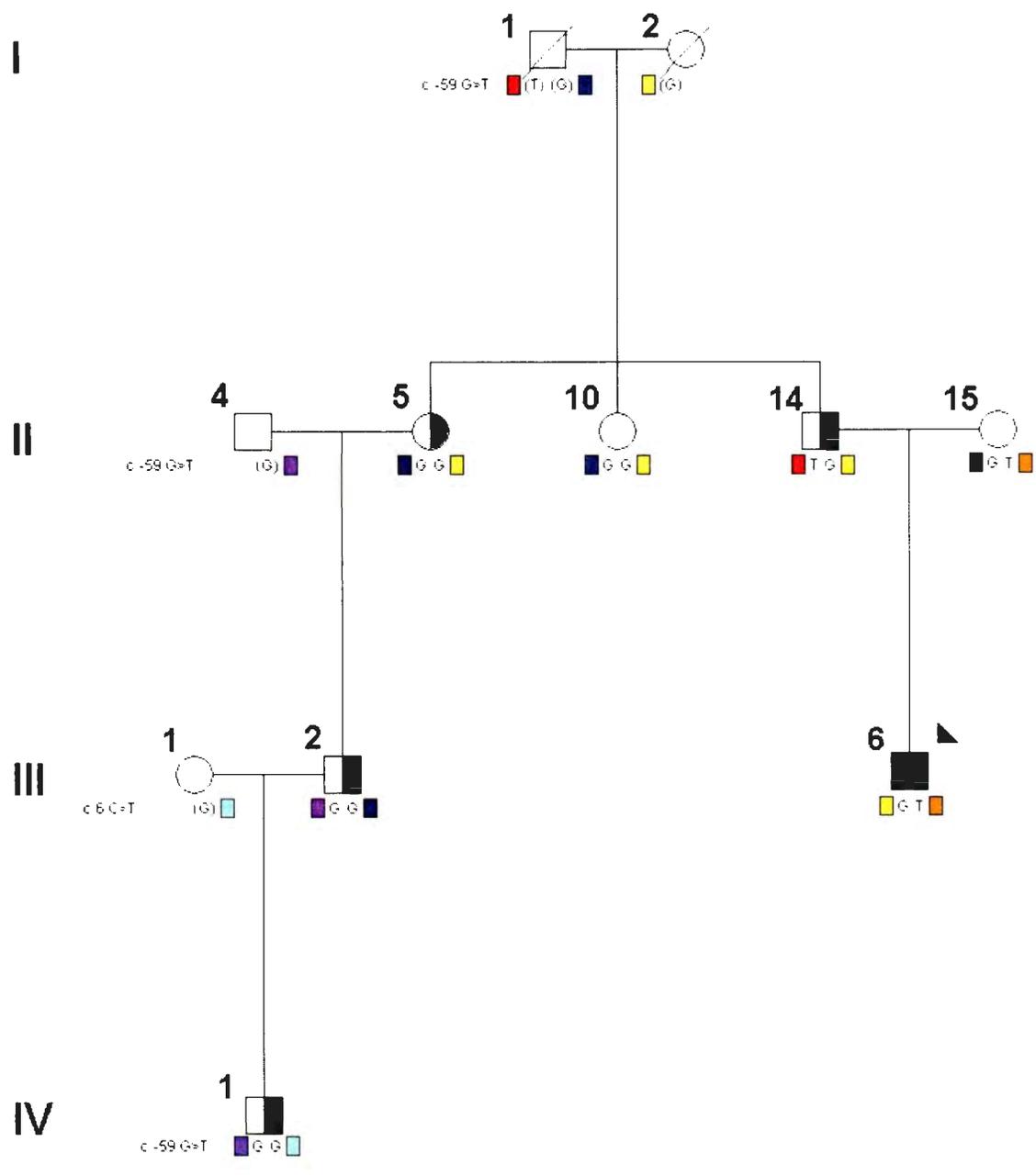


Figure 2.6 – Segregation analysis of the c.-59 G>T found in *GJA8* within family 0023. This variant is excluded as the mutant T allele only appears in individuals II-14 and III-6, but is also carried by the probands unaffected mother II-15.

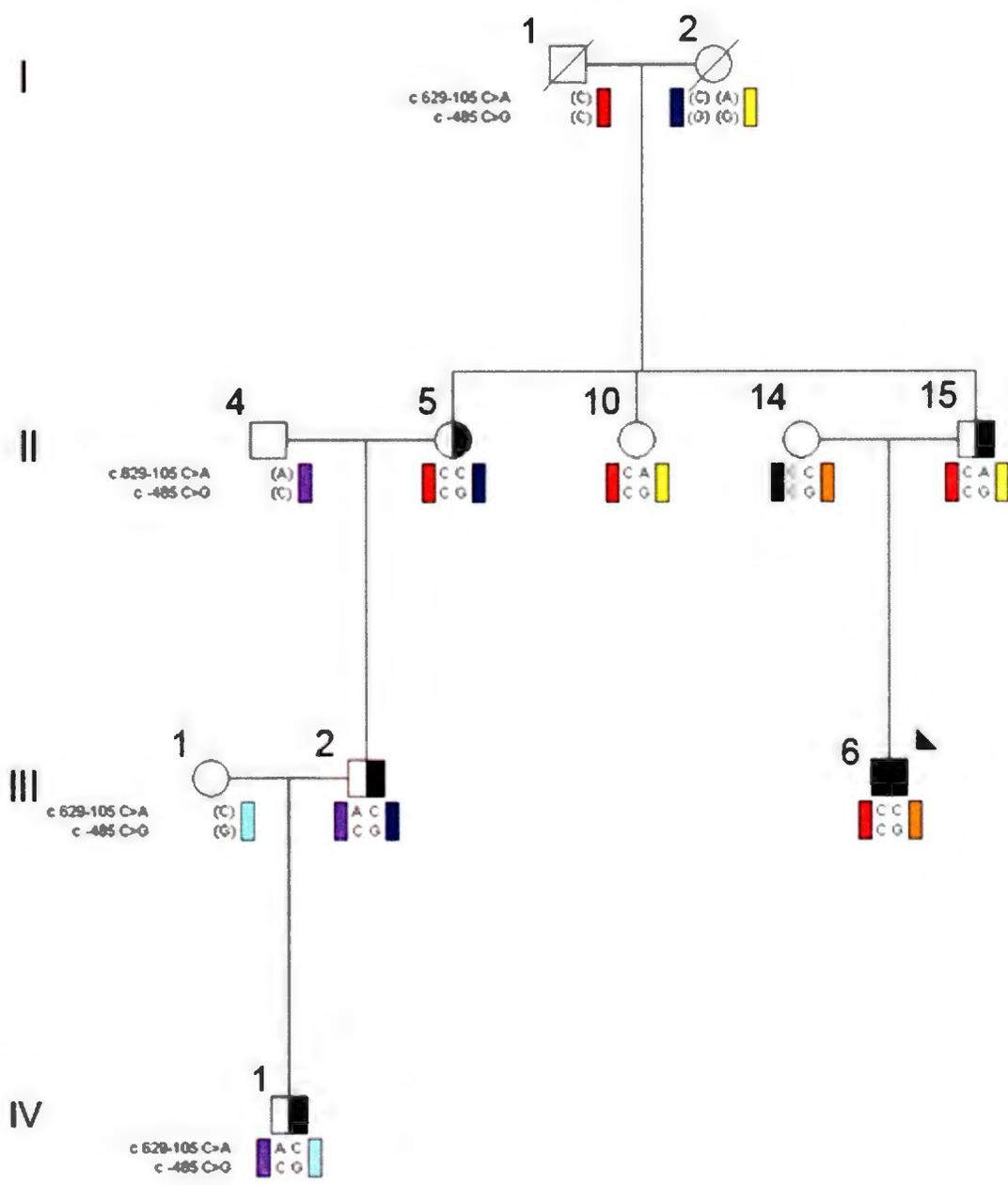


Figure 2.7 – Haplotype analysis of *PITX2* variants in family 0023. Exclusion of *PITX2* can be seen as none of the affected members appear to share a common coloured haplotype.

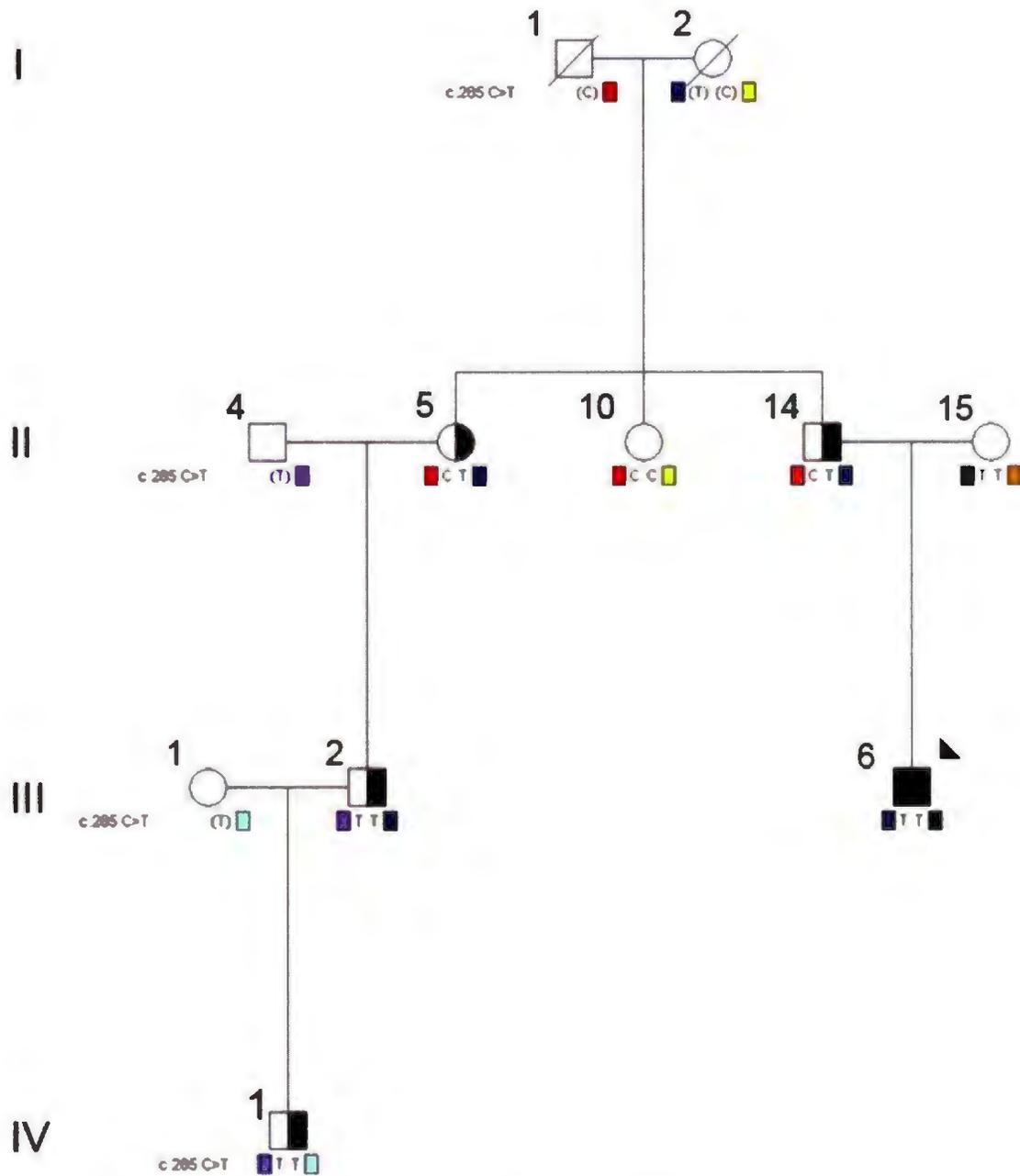


Figure 2.8 – Segregation analysis of c.285 C>T in *PITX3* found in family 0023. This variant shows no segregation with the disease phenotype and was thus excluded *PITX3* as the causative gene in this family.

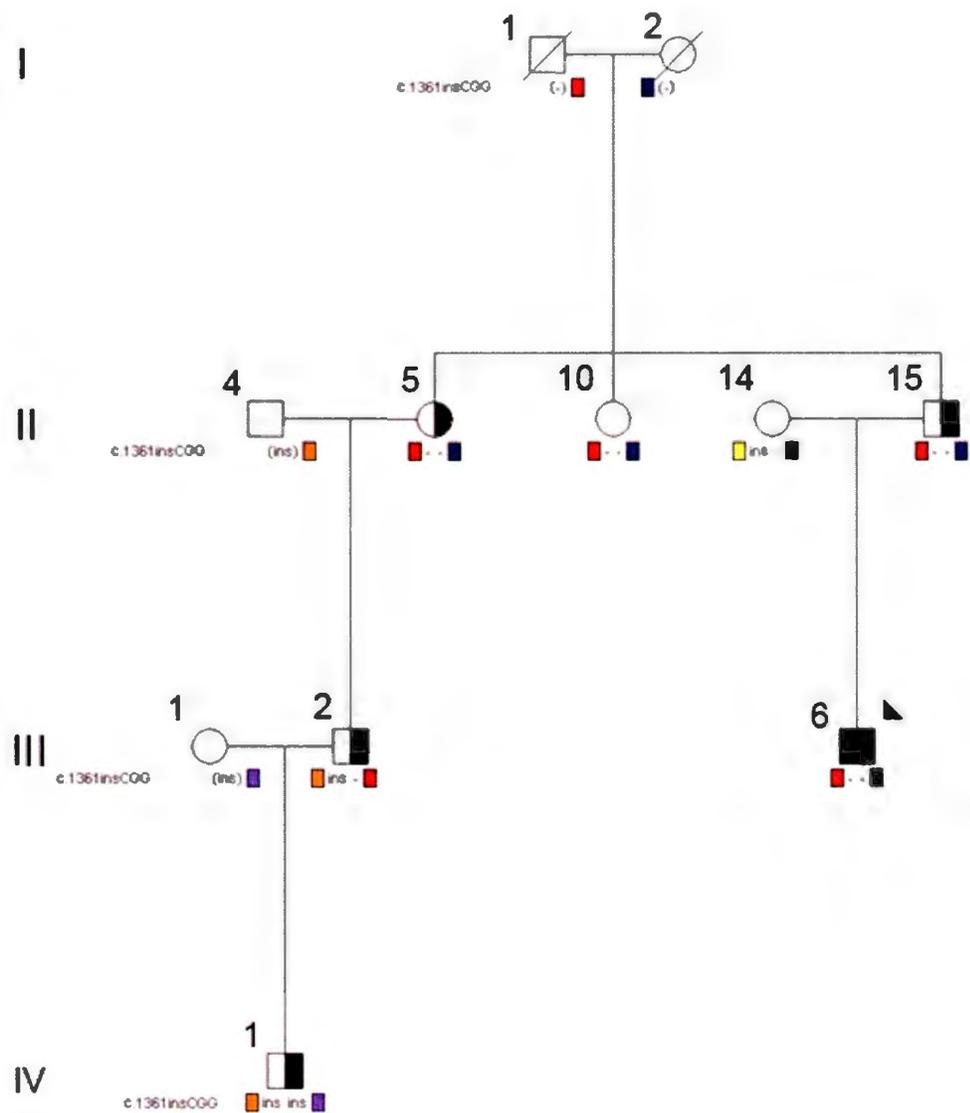


Figure 2.9 – Segregation analysis of an insertion in *FOXC1*, c.1361insCGG. This analysis excludes *FOXC1* as the causative gene as three affected individuals do not have this mutation, while the unaffected mother of the proband does.

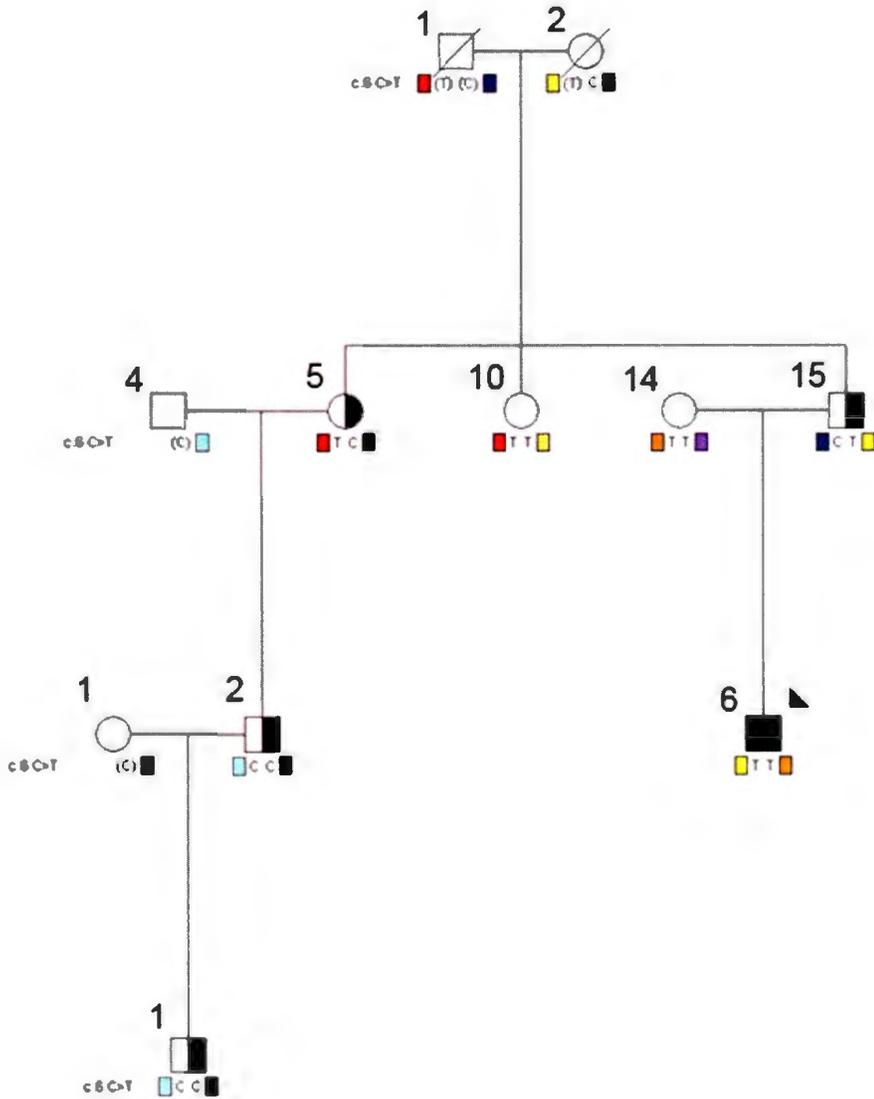


Figure 2.10 – Segregation analysis of the c.6C>T variant found in *CRYAA* within members of family 0023. As this variant only exists in three of the five affected individuals, and is a homozygote in the proband’s unaffected mother and the proband, *CRYAA* can be excluded as the causative gene in this family.

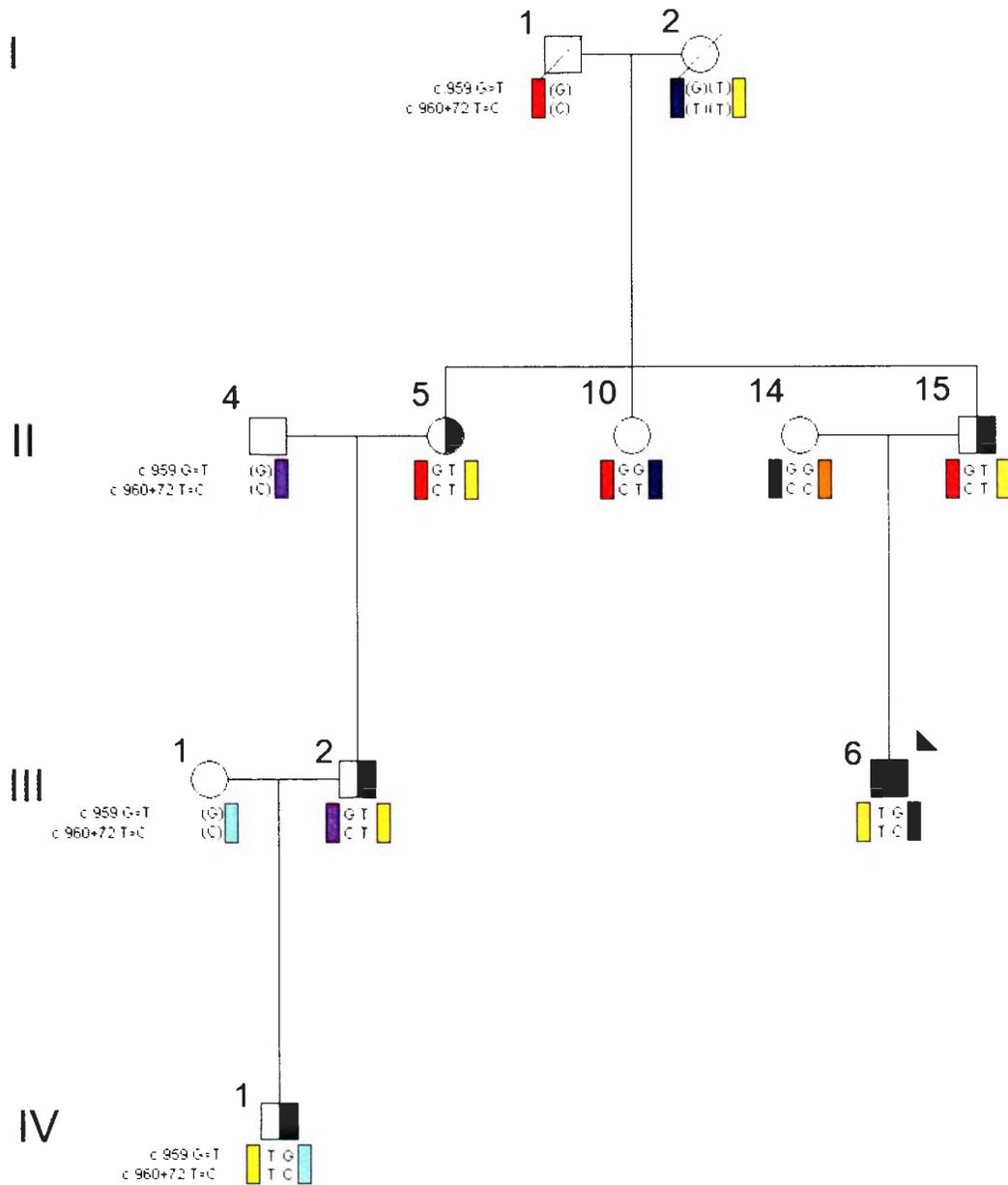


Figure 2.11 – Haplotype analysis of *FOXE3* variants in family 0023. A disease associated haplotype (yellow) can be seen segregating in an autosomal dominant manner with the affected individuals of Family 0023. This data provides evidence that *FOXE3* is causing the variable ASD phenotype seen in this family.

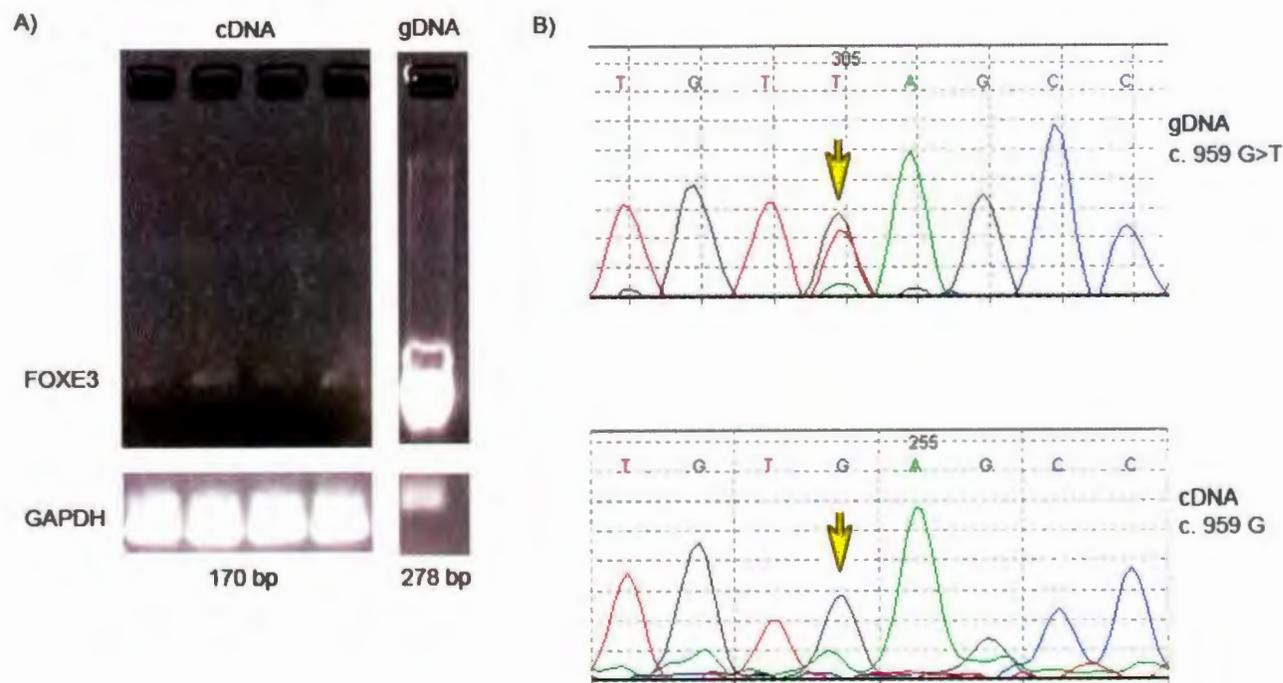


Figure 2.12: **A)** Agarose gel (1%) stained with SYBR Safe showing PCR amplification of *FOXE3* from cDNA from a lymphoblastoid cell line of PID IV-1 (in quadruplicate). The internal control of *GAPDH* amplifying a 170 bp fragment to check for presence of cDNA (and absence of gDNA) is shown below. A genomic control is shown to the right, and the 278bp fragment amplified from *GAPDH* exclusively in gDNA. **B) (Top)** Electropherogram of the novel, nonstop mutation in *FOXE3* (c.959 G>T: p.X320L) identified in all affected family members. **(Bottom)** Electropherogram of the region surrounding the stop codon of *FOXE3* from the cDNA of an affected patient (PID IV-1) isolated from a lymphoblast cell line.

2.5 Discussion

Peters anomaly is a rare congenital disorder often regarded as a sporadic condition with a low risk of recurrence and attributed to aneuploidy or fetal alcohol syndrome. Inherited cases associated with malformations of the anterior chamber of the eye have also been described. In 1986, we first reported a family with an autosomal dominant form of anterior segment dysgenesis with the proband exhibiting Peters anomaly.¹³⁶ This four generation family from Newfoundland has had extensive clinical assessment over the course of 30 years, revealing a progression of the ASD phenotype in affected family members. For example, the father of the proband (PID II-14) had mild cataracts at the age of 7 and required cataract extractions at the ages of 39 and 40 due to decreasing visual acuity. Other affected family members (PIDs III-3; III-7; II-16) have also shown a non-static phenotype with mild cataracts noted at young ages progressing to mature cataracts causing decreased visual acuity in their 20s-40s, making extractions necessary.

The anterior chamber is formed by three successive waves of embryonic cells derived from the neural crest. Aberration in the migration of these cells can cause anterior chamber defects such as microcornea, corneal opacities, or congenital cataracts.¹⁴³ Using a functional candidate gene approach, we looked for mutations in genes previously associated with ASD or Peters anomaly in this family and identified a novel non-stop mutation in *FOXE3* (c.959 G>T: p.X320L; Figure 2.12B) which segregated with the variable ocular phenotype, and was absent from 282 ethnically matched chromosomes. Though rare, non-stop mutations are reported in a number of cases involving various diseases such as non-classic 3-beta-HSD congenital adrenal hyperplasia,¹⁴⁴ and excessive hemorrhaging due to mutation of coagulation Factor X.¹⁴⁵

The normal interaction of mouse *Foxe3* with genes responsible for cellular differentiation and proliferation provides an explanation for the presence of small lenses and other eye malformations seen in humans. In mice, *Foxe3* encodes a DNA-binding transcription factor expressed during the formation of the lens placode and assists in the formation of the lens itself.¹¹³ Mutations in *Foxe3* reduce the ability of the transcription factor to bind DNA, causing formation of a small lens.⁹⁰ Sometimes the anterior lens epithelium does not separate from the cornea, resulting in keratolenticular adhesions,⁹⁰ which we observed in the proband with Peters anomaly. One explanation is that this adhesion is caused by a dysregulation of genes responsible for apoptosis within the corneal stalk which is thought to degrade during lens morphogenesis as a result of apoptosis.¹⁴⁶ An alternative explanation is that mutation of *Foxe3* may cause a dysregulation of cadherin proteins such as E-cadherin, which are present in the corneal stalk.¹⁴⁶ *Foxe3* also controls the expression of *Cryaa* within the developing lens, and dysfunction of *Cryaa* expression reduces the solubility of the crystallin protein complex causing crystallization and potential cataract formation. *Foxe3* is responsible for the down-regulation of *Prox1* which controls a gene responsible for blocking cell-cycle progression (*Cyclin dependant kinase inhibitor; Cdkn1c*). As *Foxe3* expression becomes dysregulated, *Prox1* expression increases, subsequently increasing *Cdkn1c* expression causing a reduction of cellular proliferation in the anterior lens.^{113; 147} *Foxe3* also controls expression of *Platelet derived growth factor receptor alpha (Pdgfra)* which is responsible for lens fiber differentiation within the anterior lens epithelium.¹¹³ Post-natally, *FOXE3* is expressed exclusively in the anterior lens epithelium which is the only site of cell-proliferation within the lens. This restricted cellular expression is consistent with the eye only phenotype observed in *FOXE3* mutation carriers.^{113; 146}

Mutations in *FOXE3* and its involvement in human ASD and/or Peters anomaly were first identified by Semina *et al.* who described two patients with apparent autosomal dominant posterior embryotoxon and congenital cataracts caused by a heterozygous single nucleotide insertion which altered the five terminal amino acids and added an additional 111 amino acids to the *FOXE3* protein.⁹¹ A second report of *FOXE3* and its involvement in ASD and/or Peters anomaly was published by Ormestad *et al.*, who identified a heterozygous missense mutation (p.R90L) coding for the Forkhead DNA-binding domain of the *FOXE3* protein of a single affected individual with familial Peters anomaly, though no other DNA samples from the family were available to confirm segregation.⁹⁰ Recently, Iseri *et al.* described four other families containing *FOXE3* mutations, c.21_24del: p.M71IfsX216, c.146G>C: p.G49A, c.244A>G: p.M82V, c.958T>C: p.X320ArgextX72, all of which are hypothesized to be pathogenic.¹⁴⁸ Homozygous mutations within *Foxe3/FOXE3* can also cause autosomal recessive congenital aphakia within mice (*dyl* mice)¹¹³ and humans respectively.¹¹⁴ Also, mice containing heterozygous mutations within *Foxe3* showed histological ocular malformations upon investigation,⁹⁰ further showing the ability of *FOXE3* to cause dominant and recessive ocular conditions. It is interesting to note that mutations of *FOXE3* seem to cause a spectrum of malformations, all falling under the umbrella of ASD, which range from mild (microcornea, and/or congenital cataracts) to very severe (Peters anomaly). That a single mutation can give rise to a phenotype of variable severity is unsurprising and has been documented in a number of conditions, especially those of autosomal dominant inheritance. It is suspected that mutations in *FOXE3* cause this variable phenotype either due to stochastic effects during cell migration in development, or potentially via modifier effects. It remains possible that different individuals carry protective or damaging variants in genes which interact with *FOXE3* which may

exacerbate or ameliorate the expression of the disease, giving rise to the variable phenotype observed in family 0023.

Two recent reports (Iseri *et al.* (2009)¹⁴⁸ and Brémond-Gignac *et al.* (2010)¹⁴⁹) identify similar non-stop mutations within *FOXE3* (c.958T>C: p.X320ArgextX72 and c.959G>C: p.X320SerextX72 respectively) in patients with ocular diseases. Iseri *et al.* reported a family with autosomal dominant ASD¹⁴⁸ where the proband was diagnosed with unilateral Peters anomaly and in four family members segregating congenital cataracts, microphthalmia, and iris coloboma. Brémond-Gignac *et al.* reported a single patient with congenital cataract and no family history of ocular disease.¹⁴⁹ Both of these studies predict a similar 72 amino acid extension of the *FOXE3* protein, though no functional analyses were performed to confirm this. It is possible that an extended C-terminus of the *FOXE3* protein may interfere with post-translational modifications, as the C-terminus is a common site for these types of alterations. This has been illustrated in another gene of the FOX family, called *FOXA2*, which has been shown to have a transcriptional activation domain in its C-terminal.¹⁵⁰ In order to completely elucidate the disease mechanism underlying ASD, functional studies must be done to confirm that an extended protein is produced, as our results suggest that these non-stop RNA species may be degraded. The mutation identified in the Newfoundland family (c.959 G>T) occurs in the stop codon of *FOXE3* and according to our preliminary data from cDNA (lymphoblastoid cell line) of individual PID IV-1, the mRNA transcript containing the non-stop allele is degraded (or not transcribed) suggesting haploinsufficiency of *FOXE3* (Figure 2.12), though the possibility of an extended protein cannot be ruled out without protein expression data. An alternative interpretation of the data is that the RNA containing the c.959 T allele may be more unstable than that of the wildtype, making it under-represented in total RNA. This could also explain why

the mutation was not detected by direct sequencing of cDNA, thus quantitative and translation based studies are required to confirm the absence of both the mutant RNA and the extended protein. It would be ideal to undertake quantitative and protein expression studies in a lens epithelial cell line, or a cell line more closely related to ocular structures, as the expression pattern of *FOXE3* in these cells may differ highly from that within a lymphoblastoid cell line.

Semina *et al.* (2001) also described an insertion which caused the addition of 111 amino acids to the *FOXE3* protein.⁹¹ It is possible that this mutation described by Semina *et al.*, as well as the non-stop mutations described by Iseri *et al.*, and Brémond-Gignac *et al.* would also undergo the same type of mRNA degradation that we observe. Quantitative studies of the *FOXE3* c.959 G>T mutation within human cells would be necessary to validate our hypothesis of RNA degradation.

In summary, our results implicate *FOXE3* in the pathology of ASD and corroborate other recent reports of ASD cases with *FOXE3* mutations. Therefore, mutations within *FOXE3* may explain currently unsolved ASD cases so *FOXE3* should be screened. Cataract extraction in individuals with *FOXE3* mutations is more likely to result in post-operative complications so identifying mutations carriers can affect clinical management. Discovery of genes and mutations causing ASD in families is of particular clinical relevance. A molecular diagnosis of ASD cases through mutation screening can provide accurate risks of recurrence, especially in light of mild or subclinical phenotypes which may show progression over time such as those seen in the family in this study.

2.6 Acknowledgements

We thank all family members for their longstanding participation in this study. This work was financially supported by: The Canadian National Institute for the Blind (CNIB), The E.A Baker Foundation, The Canadian Foundation for Innovation (New Opportunities Fund #9384; Leaders Opportunity Fund #13120); the Janeway Children's Hospital Foundation; Memorial University Opportunities Fund; Genome Canada Competition III Award (Atlantic Medical Genetics and Genomics Initiative: AMGGI).

2.7 Conflict of Interest

The authors declare no conflict of interest.

Chapter 3: A population-based study of achromatopsia in a Canadian founder population reveals genetic heterogeneity, a novel mutation in *CNNM4*, and the first family with Jalili Syndrome in North America.

Lance Doucette¹, Jane Green¹, Coleman Black¹, Gordon J Johnson¹, Dante Galutira¹, Jeremy Schwartzentruber² and Terry-Lynn Young¹

¹ Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada A1B 3V6. ² McGill University and Génome Québec Innovation Centre, Montréal, Canada H3A 1A4

A shorter version of this chapter has been published in Ophthalmic Genetics on January 30th, 2013.

3.1 Summary

Achromatopsia is a rare (1/30,000 live births) retinal condition affecting the cone photoreceptors and presents with a loss of visual acuity, lack of colour vision, nystagmus, and photophobia. We obtained DNA from eight families (0094, 1442, 1491, 1492, 1713, 1723, 1726, and 1734) from Newfoundland, Canada and screened available members for mutations within the four achromatopsia genes *CNGA3*, *CNGB3*, *GNAT2*, and *PDE6C*. This approach solved the genetic etiology of six of the eight families and yielded a total of 27 variants, five of which were pathogenic: two in *CNGB3* (c.1148delC and c.886_896del11insT) and three within *CNGA3* (c.848 G>A: p.R283Q, c.1279 T>C: p.R427C, and c.1580 T>G: p.L527R). Two mutations, *CNGB3* c.1148delC, and *CNGA3* c.1580 T>G occurred in more than one family. Haplotype analysis of these eight families showed a common haplotype surrounding *CNGB3* c.1148delC and *CNGA3* c.1580 T>G across multiple families, suggesting that these mutations reside on founder haplotypes. One patient from family 1734 had only one mutation in *CNGB3* (c.1148delC) identified. The unsolved family (0094) was subjected to exome sequencing revealing seven homozygous variants exclusive to the affected individuals. Interestingly, amongst them was a novel mutation (c.1555 C>T: p.R519X) in *CNNM4*, the only gene known to cause Jalili Syndrome, a rare form of cone-rod dystrophy with a dental malformation, amelogenesis imperfecta. Review of archived records revealed tooth abnormalities in family members with achromatopsia consistent with that of amelogenesis imperfecta; and represents the first report of Jalili syndrome in North America. This study highlights the genetic/allelic heterogeneity of achromatopsia, even within a founder population and highlights the importance of exome sequencing coupled with accurately maintained records as a diagnostic tool.

3.1 Introduction

3.1.1 *What is Achromatopsia?*

Achromatopsia, also referred to as rod monochromacy (OMIM #262300; OMIM #8216900; OMIM +139340; OMIM*600827), is a rare autosomal recessive (AR), fully penetrant, congenital vision disorder affecting approximately 1 in every 30,000 individuals.¹⁵¹ This condition affects the cone photoreceptors, resulting in a loss of high visual acuity, and colour vision from birth. In the past, this condition was thought to be stationary (the severity of the condition remaining the same over time), though recent evidence suggests that achromatopsia may be a progressive condition through physical changes in the retina/fovea, and the cone photoreceptors, as well as eventual rod involvement. Achromatopsia can also be categorized into complete (typical) and incomplete (atypical). The incomplete form of this condition, also called dyschromatopsia, has symptoms similar to complete achromatopsia, though with less severe visual dysfunction.¹⁵² Although being known for a number of years, achromatopsia was particularly well-defined in a founder of the Pingelap Island, off the coast of Papua, New Guinea in the Micronesian Atolls.

The Pingelapese

The Pingelap island of the Micronesian Atolls small island in the Pacific Ocean was home to a small population of about 200 people, until a massive storm known as Typhoon Lengkieki decimated the island in 1775, killing all but about 20 individuals (including 9 males).⁵⁰ It is believed that this event created a genetic bottleneck in the Pingelap population with an increase in prevalence of carriers of achromatopsia in the remaining population creating an approximate disease prevalence of about 10% amongst its now 250 individuals (as opposed to

the 0.003% in the U.S.A). It is believed that an individual, Nanmwarki Mwahuele, is the ancestor through three wives, of all achromatopsia carriers of the Pingelap.¹⁵³ This population was used to map the *ACHM3* locus and to eventually identify *CNGB3*.¹⁵⁴⁻¹⁵⁶

3.1.2 Clinical Presentation and diagnosis of Achromatopsia

Patients with achromatopsia have reduced visual acuity, loss of colour vision, congenital pendular nystagmus, central scotoma (loss of central vision), and a severe sensitivity to light (termed photophobia).¹⁵⁷ This condition can exist as either complete or incomplete achromatopsia. Complete achromatopsia is considered to be the more severe form of the disease and is associated with total loss of colour vision and severely reduced visual acuity. This loss varies from case to case but tends to be greater than 20/200 (considered legally blind). Incomplete achromatopsia presents similarly to complete achromatopsia but is less severe, with partial remaining colour vision, and better overall visual acuity (not usually greater than 20/80). Both types of achromatopsia are caused by malfunction of the three types of cone photoreceptors: short wave (red), medium wave (green), and long Wave (blue).

Achromatopsia can be diagnosed through a number of ophthalmic observations and tests. These tests include, but are not limited to, electroretinography (ERG), ocular coherence tomography (OCT), colour vision testing (Farnsworth D15 colour arrangement, Ishihara colour test, Hardy-Rand-Ritler colour test), and in some instances physical changes can be seen upon funduscopy.

(i) Electroretinography (ERG)

The gold standard for diagnosis of achromatopsia is an ERG which measures the response of photoreceptors to various stimuli. Full field ERG measures the entire retinal response to a light stimulus but is broken down into two types of responses, scotopic (rod response) where dim blue stimulus is used and photopic (cone response) when bright white light or a flicker stimulus is used.¹⁵⁸ Typically, during an ERG test, the pupil is dilated and the eyes are dark-adapted for up to 30 minutes. These conditions are designed to be able to distinguish between the rod and cone responses. Rods respond more actively at low light conditions, and also to blue light stimuli, while cones respond to bright light stimuli and rapidly flashing stimuli but do not respond in the dark. To measure scotopic responses, a dim blue or white light flashes during dark adaptation. Photopic responses are measured under light adapted conditions (used to desensitize or 'bleach' the rods) using bright white flash stimuli, or a flicker test at 30 Hz (regular repeating pattern of flashing light).

In complete achromatopsia the dark adapted flash stimuli show normal amplitudes upon testing, which indicate a lack of rod involvement. However, the flicker responses to stimuli and light adapted flash stimuli show flat amplitudes, indicating sole involvement of the cones. Incomplete achromatopsia shows a normal dark-adapted response, but also shows residual photopic responses, which suggests a partial functioning of the cone photoreceptors. These patients also present with a higher visual acuity (typically better than 20/80), and less severe photophobia/nystagmus than patients presenting with complete achromatopsia.¹⁵⁸

(ii) Ocular Coherence Tomography (OCT)

A second test used to identify achromatopsia in a patient with reduced vision is Optical Coherence Tomography (OCT). This test determines the functioning of the outer retina, and the retinal pigment epithelium. This is done by measuring the resting potential between the front and back of the eye (termed the corneo-fundal potential) which changes depending on the illumination of the retina.¹⁵⁹ This technology allows medical professionals to get an idea about the 3D structure and thickness of the retinal layers.

Testing of achromatopsia patients using OCT has shown that the cone photoreceptors lose both the inner and outer segments of the cell. This loss of cone cells apparently increases with age, suggesting that achromatopsia is not as stationary a disease as was once thought.¹⁶⁰

(iii) Colour Vision Testing

A hallmark feature of achromatopsia is the lack or loss of colour vision in affected patients. There are three types of colour vision loss, all involving the different cone photoreceptors (red, blue, green). Protan, is a defect of the long wave (red) photoreceptors, and patients are described as Red-blind, or Red-weak depending on the severity of their loss. Deutan loss is the malfunction of the medium wave (green) photoreceptors, and patients are called green-blind or green-weak. Tritan, is the malfunction of the short wave (blue) photoreceptors, and this type of loss causes individuals to be blue-blind, or blue-weak.

Colour loss in achromatopsia affects all three types of photoreceptor and is thus known as "total colour blindness". This lack of colour discrimination can be determined by a number of visual tests such as the Farnsworth color arrangement test, or the Ishihara colour test. The Farnsworth colour test requires patients to arrange a set of coloured discs based on the shade of

the discs. The order that the patient arranges the colours in is then charted, and assessed. Those without the ability to see colours often arrange the discs according to darkness instead of shade.

Physical changes in the retina and fovea are seen in some cases of achromatopsia. Most patients have a normal fundus;¹⁶¹ though some show some subtle changes in the appearance of the macula. These changes may present as a negative foveal reflex (no response from the fovea when illuminated), macular pigment changes and atrophy,¹⁶² or narrowing of the retinal vessels.¹⁶³ Individuals diagnosed with complete achromatopsia show no recordable cone response upon electroretinogram (ERG), absent colour vision and poor visual acuity whereas individuals with incomplete achromatopsia show residual cone function, better colour vision and higher visual acuity.¹⁶²

3.1.3 Molecular Genetics of Achromatopsia

Achromatopsia has previously been shown to be caused by mutations in one of four genes (*CNGA3*,¹⁶⁴ *CNGB3*,¹⁵⁷ *GNAT2*,¹⁶⁵ *PDE6C*¹⁶⁶) involved in the functioning of cone photoreceptors which are responsible for photopic responses, trichromatic color vision, and high-resolution visual acuity.¹⁵⁵ More recently, a study of *PDE6H* has revealed a nonsense mutation in three individuals of two families from the Netherlands, and Belgium and identified *PDE6H* as the fifth achromatopsia locus.¹⁶⁷

(i) *CNGA3* (*ACHM2*; OMIM #216900)

CNGA3 (OMIM *600053) creates the alpha subunit of the cyclic nucleotide cGMP-gated cation channel (CNG) found in the photoreceptor cells (the function of the CNG channel is described on page 22 of this thesis). The *ACHM2* locus was originally mapped in 1997 using a

Jewish Iranian kindred to Chr2p11.1-q12.¹⁶⁸ Just a year later, in 1998, the gene for *ACHM2* was identified to be *CNGA3*¹⁶⁹ which was previously found to be involved in retinitis pigmentosa.¹⁷⁰ *CNGA3* pairs with *CNGB3* in a two by two fashion (two A3 subunits, two B3 subunits) to form a heterotetramer,¹⁷¹ in which *CNGA3* is the major functional subunit. Mutations of *CNGA3* account for approximately 25% of achromatopsia cases worldwide,¹⁶¹ and over 50 pathogenic missense variants have been discovered in this gene. This gene has also been associated with cone dystrophy,¹⁶¹ autosomal recessive retinitis pigmentosa,¹⁷⁰ and Lebers congenital amaurosis.¹⁷²

(ii) *CNGB3* (*ACHM3*; OMIM #262300)

The *ACHM3* locus was first mapped in 1999 to a 6.5 cM interval located on Chr8:q21-q22 by using an Irish population¹⁵⁶ and was confirmed in the previously mentioned Pingelap population.¹⁵⁵ This locus was then refined in 2000 to a 1.4 cM region and the causative gene, *CNGB3*, was determined by using the Pingelapese population.¹⁵⁷ *CNGB3* (OMIM *605080) was also found to be causative in a case of achromatopsia previously thought to be due to a Robertsonian Translocation on chromosome 14 and uniparental disomy,¹⁷³ a locus previously referred to as *ACHM1*.¹⁷⁴ Currently, mutations of *CNGB3* are known to cause approximately 45-50% of achromatopsia cases,¹⁷⁵ the majority of which are null alleles such as nonsense or frameshift mutations. The most common of these mutations is c.1148delC: p.T383IfsX13, which is considered to be a founder mutation accounting for up to 80% of *CNGB3* mutant alleles.¹⁷³ The translated protein is the beta subunit of the CNG channel, and pairs with its counterpart *CNGA3* in a heterotetrameric fashion.¹⁷¹

(iii) *GNAT2* (*ACHM4*; OMIM #613856)

ACHM4 was mapped to chromosome 1p13 in 2002, and the causative gene was found to be *GNAT2* (OMIM 139340).¹⁶⁵ Mutations in this gene are thought to cause a small proportion of achromatopsia cases (~2%).¹⁷³ The protein coded by *GNAT2* is Guanine Nucleotide-binding protein Alpha-Transducin 2, a protein involved in the phototransduction cascade by activating phosphodiesterases (PDE) to regulate the amount of Cyclic Guanosine Monophosphate (cGMP) in the photoreceptor cell. Upon light stimulation an exchange of Guanosine Diphosphate (GDP) for Guanosine Triphosphate (GTP) on the transducin molecule occurs.¹⁷⁶ This exchange activates GNAT, and allows removal of inhibitory gamma domains from the PDE proteins, activating the PDE and lowers the amount of available cGMP in the cell. As cGMP concentrations lower in the cell, the CNG closes, causing hyperpolarization of the cell which then creates a neural response.

(iv) *PDE6C* (*ACHM5* OMIM *600827)

PDE6C encodes a member of the phosphodiesterase family of enzymes present in the photoreceptors cells which is responsible for regulation of cGMP levels. This gene was mapped in 1995 to 10q24 through a study using fluorescence in situ hybridization (FISH).¹⁷⁷ Recently, *PDE6C* was shown to cause achromatopsia through a study in the *cpfl1* mouse strain which exhibited a phenotype much like that of achromatopsia patients with lack of cone function, and rapid cone degeneration, though in humans the lack of cone function is present at birth.¹⁶⁶ The human homologue, *PDE6C*, is a cone-specific phosphodiesterase which works in conjunction with the transducin *GNAT2* to cause hydrolysis and subsequent reduction of cGMP after light stimulation.¹⁶⁶ Chang *et al.* tested achromatopsia patients who were negative for mutations in the three previously described genes and was able to identify mutations in *PDE6C* that caused achromatopsia in four families.

(v) *PDE6H* (ACHM6 OMIM *601190)

PDE6H encodes a second protein which belongs to the phosphodiesterase family of enzymes, the inhibitory gamma subunit of the PDE complex. The PDE complex is a homodimer comprised of two alpha-prime subunits (coded by *PDE6C*) and its activity is inhibited by the binding of the gamma subunit. As GDP is exchanged for GTP on the GNAT protein, the gamma subunit is sequestered allowing for activation of the PDE complex. The cDNA for *PDE6H* was originally isolated in 1996 from retinal tissue, and mutations were discovered to cause retinal cone dystrophy in 2005. More recently, Kohl *et al.* (2012) discovered a nonsense mutation in two siblings from Belgium, and in one individual from the Netherlands with incomplete achromatopsia.¹⁶⁷

3.1.4 Aims of this work

This chapter describes eight families from the Canadian province of Newfoundland and Labrador with AR achromatopsia. The aim of this study was to identify the genetic defects in these eight families through bidirectional Sanger sequencing of functional candidate genes previously associated with achromatopsia, namely *CNGA3*, *CNGB3*, *GNAT2*, and *PDE6C*. This was undertaken in hopes of identifying novel mutations in these four genes, or ideally to identify a novel achromatopsia locus through further studies involving individuals not solved by sequencing these four genes. Due to its relatively recent discovery, *PDE6H* was not examined during the course of this thesis.

3.3 Materials and Methods

3.3.1 Patients and Ophthalmological Examination

Families diagnosed with achromatopsia were identified through review of CNIB records or through referral to Ocular Genetics Clinics in St. John's or rural hospitals. Diagnosis of achromatopsia was confirmed through ophthalmological examination, visual field testing, colour vision testing, and ERG when possible. DNA was extracted from whole blood taken from participating patients. This study was approved by the Ethics Committee, Faculty of Medicine, Memorial University, St. John's Newfoundland, Canada. (HIC#06.157).

3.3.2 Functional Candidate Gene Selection and Screening

Genomic DNA from all individuals was screened for mutations in the previously described achromatopsia genes *CNGA3*, *CNGB3* and *GNAT2*. These three genes were selected for priority screening as they are known to cause the majority of achromatopsia cases worldwide. If no pathogenic variants were identified, the probands were then sequenced for *PDE6C*. DNA was isolated as previously described.¹⁴¹ Oligonucleotide primers were designed using Primer 3 software (<http://frodo.wi.mit.edu/primer3/>) and can be found in Appendix H. PCR, purification, and bidirectional sequencing for each gene was set up as per section 2.3.3 of this thesis. Due to its relatively recent association with achromatopsia, *PDE6H* was not studied in this thesis.

3.3.3 Variant and Haplotype Analysis

Intragenic haplotypes were manually constructed, where SNPs were available, to check for segregation of a disease-associated haplotype. Bioinformatic analysis to assess pathogenicity of missense variants was carried out using SIFT,¹⁷⁸ PolyPhen,¹⁷⁹ and PANTHER.¹⁸⁰

3.3.4 Cloning

Primers PDE6C17+18-F and -R were used to amplify DNA fragments and these were purified as per section 2.3.3. The amplicons were quantified using a NanoDrop spectrophotometer and 4 uL of the amplicons was added to the TOPO TA Cloning Kit mastermix (mastermix protocol can be found in Appendix I) to insert the purified fragment into a pCRTM4-TOPO® vector. The mixture was incubated at room temperature for 30 minutes and the entire plasmid solution was transferred to a 2 mL Eppendorf tube containing 50 uL of TOP10 Chemically Competent E.Coli cells and incubated on ice for 30 minutes. The bacteria were heat shocked at 42 degrees Celsius for 30 seconds to stimulate plasmid uptake and immediately transferred to ice where 250 ul sterile Super Optimal broth with Catabolite repression (S.O.C) medium. One-hundred and 150 uL of the bacterial solution were spread on LB/Agar plates with 100 ug/mL Ampicillin using a sterile steel spreader, and the plates grown overnight at 37 degrees Celsius.

3.3.5 Mini-Prep

After 24 hours of growth, the LB/Agar plates were examined for *E. Coli* colony growth. Colonies were picked from the plate using a pipette tip and placed in a 15 mL centrifuge tube with 4 mL of LB broth+Ampicillin. These tubes were placed on a shaker overnight at 37 degrees Celsius. After 24 hours of growth, the tubes were examined for a cloudy appearance indicating sufficient bacterial growth. Two mLs of the bacterial culture was used with the Qiagen Miniprep kit to isolate plasmids from the bacterial cells for subsequent sequencing. The bacteria were pelleted in a 2 mL Eppendorf tube at 13,000 rpm for five minutes, and the pellet was resuspended in 250 uL of Buffer P1 (a resuspension buffer). Two-hundred fifty uLs of Buffer P2 (to bind DNA to the membrane) was added to the tube and inverted to mix the solution followed by addition of 350 uL Buffer N3. A white precipitate was observed and the tubes centrifuged at 13,000 rpm. The supernatant was transferred to a QIAprep spin column and centrifuged at 13,000 rpm for 1 minute, the flow-through discarded and 750 uL of Buffer PE (a wash buffer) was added to the filter membrane of the QIAprep column. The column was spun at 13,000 rpm and the flow through discarded; this step was repeated once again to dry the membrane. To elute the now purified plasmid from the filter membrane, 20 uL of water was added to the membrane and was allowed to incubate at room temperature for 1 minute. The QIAprep column was then transferred to a 1.5 mL Eppendorf tube, and was spun at 13,000 rpm for 1 minute to isolate the plasmid. The solution was then quantified using a Nanodrop Spectrophotometer. The plasmid was then sequenced as per section 2.3.3 using a forward and reverse primer called M13-Forward, and M13-Reverse which were specific to the bacterial plasmid from the TOPO TA Cloning Kit.

3.3.6 Exome Sequencing

DNA samples from the unsolved family (0094) were subjected to whole exome sequencing using three DNA samples from Family 0094 (two affected, one unaffected). Standard manufacturer protocols were followed to perform target capture with the Illumina TruSeq exome enrichment kit and sequencing of 100 bp paired end reads on Illumina HiSeq. Approximately 10 Gb of sequence was generated for each subject such that >90% of the coding bases of the exome defined by the consensus coding sequence (CCDS) project were covered by at least 20 reads. Adaptor sequences were removed and reads were quality trimmed using the Fastx toolkit and then a custom script to ensure that only read pairs with both mates present were subsequently used. Reads were aligned to hg19 with BWA,¹⁸¹ and duplicate reads were marked using Picard and excluded from downstream analyses. Single nucleotide variants (SNVs) and short insertions and deletions (indels) were called using samtools pileup and varFilter¹⁸² with the base alignment quality (BAQ) adjustment disabled, and were then quality filtered to require at least 20% of reads supporting the variant call for SNVs (15% for indels). Variants were annotated using both Annovar¹⁸³ and custom scripts to identify whether they affected protein coding sequence, and whether they had previously been seen in dbSNP131, the 1000 genomes pilot release (Nov. 2010), or in approximately 240 exomes previously sequenced at the McGill Genome Innovation Center.

3.4 Results

3.4.1 Patients and Clinical Examination

We obtained 21 blood samples from consenting family members for DNA extraction (13 from affected; eight from unaffected individuals) across eight families identified with autosomal recessive achromatopsia. Diagnosis of achromatopsia was confirmed through ophthalmological examination, visual field testing, colour vision testing, and ERG when possible in a total of 13 individuals across the eight families. Clinical descriptions of affected family members can be found in Table 3.1, and Table 3.2.

3.4.2 Candidate Gene Sequencing

Screening of the three known achromatopsia genes, *CNGA3*, *CNGB3*, and *GNAT2* yielded 19 variants across all seven families (Table 3.3). Homozygous or compound heterozygous pathogenic variants were discovered in either *CNGA3* or *CNGB3* in six of the eight families (Table 3.4).

3.4.3 *CNGB3* mutations revealed in families 1442, 1492, 1713, 1734

Affected individuals in two multiplex families, 1442 and 1492 (Figure 3.1), were homozygous for the most common mutation in *CNGB3*, c.1148delC: p.T383IfsX13. This mutation was also discovered in the proband of Family 1713, though it was compound heterozygous with a previously described 10 basepair deletion in exon 7 of *CNGB3*, c.886_896del11insT: p.T296YfsX9.¹⁸⁴ There was no parental DNA from this family to check for segregation of both alleles, but no mutant alleles were found in either *CNGA3* or *GNAT2*. The

individual from Family 1734 was noticed to have one copy of the c.1148delC: p.T383IfsX13 mutation, though no other mutations were discovered in *CNGB3*.

3.4.4 CNGA3 mutations revealed in families 1491, 1723, and 1726

Affected members of Family 1491 (Figure 3.1) were homozygous for the recently described c.1580 T>G: p.L527R mutation¹⁸⁵ which resides in the cGMP binding site of the CNGA3 protein. The p.L527R mutation was also observed in Family 1723 as a compound heterozygote with a previously described mutation c.1279 C>T: p.R427C (Table 3.4). We also observed c.1279 C>T: p.R427C as a compound heterozygote with c.848 G>A: p.R283Q in the proband of Family 1726 (Table 3.4). Parental DNA was unavailable for both Families 1723 and 1726.

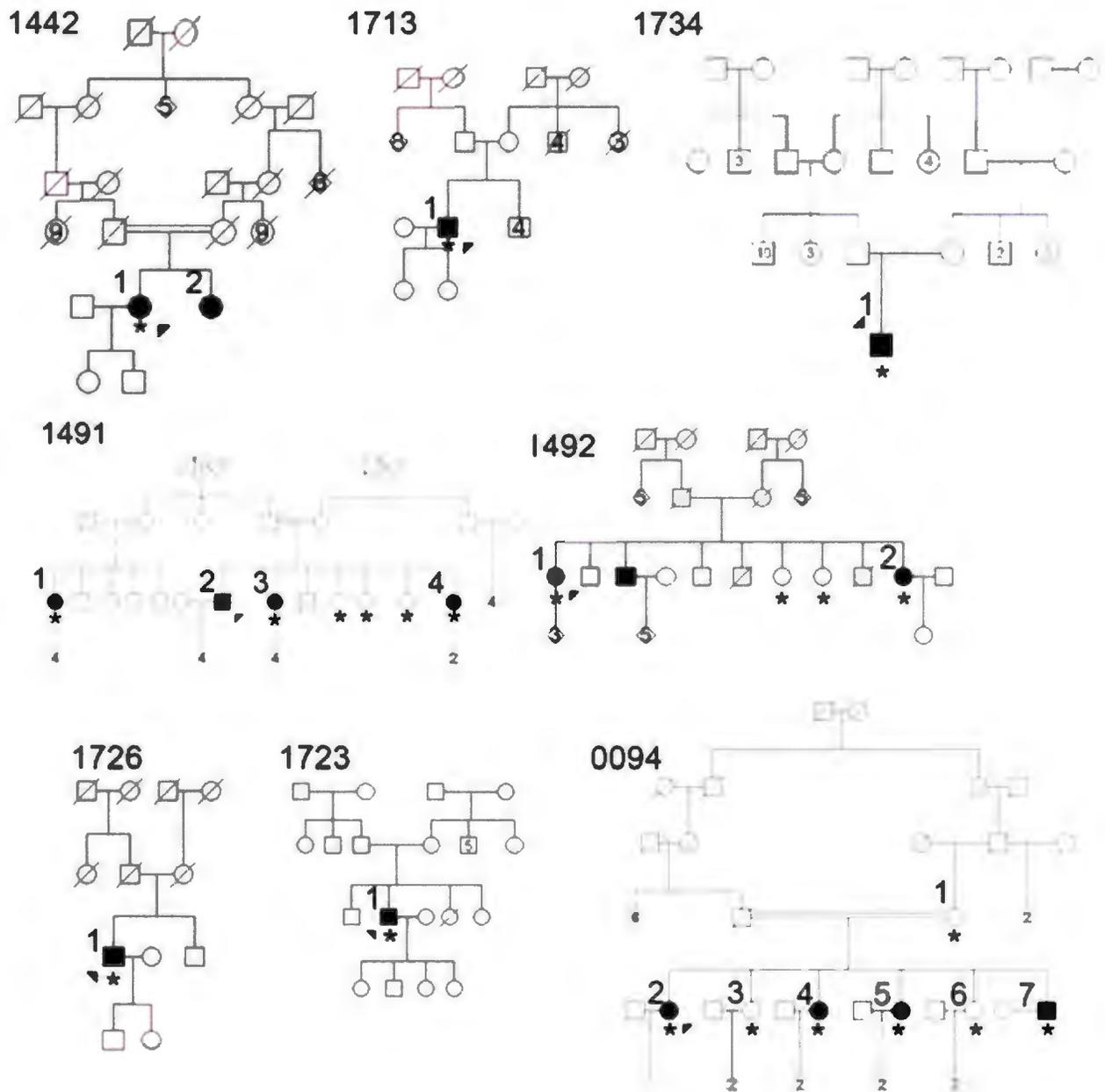


Figure 3.1 Pedigrees of the eight families (1442, 1491, 1492, 1713, 1726, 1723, 1734 and 0094) diagnosed with achromatopsia. Numbers to the top right of the symbols correspond to the clinical information in Tables 3.1 and 3.2. An asterisk at the bottom of a symbol indicates that that DNA was contributed by this individual.

	YOB	Visual Acuity (VA)		Nystagmus	Colour Vision	Visual Fields	ERG Responses
		First VA	Recent VA				
R1442.1	1944	1982 20/200 OD 20/200 OS	2006 20/200 OD 20/400 OS	Yes	1985 Farns D15 - Abn Ishihara - Nil HRR - Nil	ND	ND
R1442.2	1948	1985 20/200 OD 20/200 OS	ND	Yes	1985 Farns D15 - Abn Ishihara - Nil HRR - Nil	ND	ND
R1491.1	1938	1996 20/200 OD 20/200 OS	2007 20/400 OD 20/400 OS	No	ND	ND	ND
R1491.2	1937	1984 20/200 OD 20/200 OS	ND	No	1984 Farns D15 - Nil Ishihara - Nil HRR - Nil	1984 Central Scotoma OD/OS	ND
R1491.3	1938	1974 20/300 OD 20/400 OS	2006 20/400 OD 20/400 OS	No	ND	ND	ND
R1491.4	1952	1982 20/200 OD 20/200 OS	2005 20/400 OD 20/400 OS	Yes	1982 Farns D15 - Nil Ishihara - Nil HRR - Nil	1982 Peripheral Constriction OD/OS	2005 Rod - Reduced and Delayed Cone - Absent
R1492.1	1943	1994 20/200 OD 20/200 OS	1995 20/200 OD 20/200 OS	Yes	1994 Farns D15 - Abn Ishihara - Nil HRR - Nil	1994 Peripheral Constriction OD/OS	1994 Rod - Reduced Cone - Absent
R1492.2	1958	1988 20/200 OD 20/200 OS	2006 20/200 OD 20/200 OS	Yes	2002 Farns D15 - Nil Ishihara - Nil HRR - Nil	ND	ND
R1713.1	1953	2009 20/200 OD 20/200 OS	ND	No	2009 Farns D15 - Abn	2009 Normal	2009 Rod - Reduced Cone - Absent
R1723.1	1966	1983 20/200 OD 20/200 OS	2010 20/200 OD 20/200 OS	Yes	1983 Farns D15 - Abn Ishihara - Nil HRR - Abn	2010 R - Relative peripheral loss L - Central scotoma; relative peripheral loss	2010 Rod - Normal Cone - Low amplitude Irregular wave pattern
R1726.1	1936	2008 20/200 OD 20/200 OS	2010 20/200 OD 20/200 OS	Yes	ND	ND	2010 Rod - Reduced and Delayed Cone - Absent
R1734.1	1975	1981 20/400 OD 20/300 OS	1990 CF OD CF OS	Yes	1987 Ish - Abn HRR - Abn Farns D15 - Abn	1982 Central Scotoma OD/OS	1982 Rod - Normal Cone - Absent

Table 3.1: Clinical description of affected individuals from seven of the eight families diagnosed with achromatopsia. **YOB:** Year Of Birth, **OD:** Oculus Dexter (Right eye), **OS:** Oculus Sinister (Left eye) **CF:** Counting Fingers, **HM:** Hand Movements, **ND:** Not Done, **HRR:** Hardy-Rand-Ritler Colour Test, **Farns D15:** Farnsworth D15 Colour Arrangement.

	YOB	Status	CNIB Reg.	Visual Acuity (VA)		Nystagmus	Colour Vision	Visual Fields	ERG Responses	Dental Abnormalities
				First VA	Recent VA					
0094.2	1959	Affected	1973	1987 20/300 OD 20/300 OS Myopia	2007 CF OD CF OS	1973 Yes 1987 No	1987 Ishihara – No Response HRR – No Response Farns D-15 - Abnormal	1987 Relative Peripheral Loss	1987 Rod-Borderline Cone-Absent	1973 – 22 teeth extracted, 'severe enamel dysplasia'. 1980 – 11 teeth extracted
0094.3	1960	Unaffected	NR	1982 20/20 ⁻¹ OD 20/25 ⁺¹ OS	2011 20/20 ⁻³ OD 20/25 ⁺¹ OS	No	Farns D-15 – normal	ND	ND	1987 - Upper teeth extracted. 2010 - lower teeth extracted
0094.4	1961	Affected	1973	1995 Required escort b/c of VA	2011 HM 8 ft OD/OS Bone spicules	1973 Yes	ND	ND	ND	1973 – 32 teeth extracted, 'Severe enamel dysplasia'. 1980 – 10 secondary teeth extracted
0094.5	1963	Affected	1973	1982 20/200 OD 20/200 OS	2011 HM 8 ft OD/OS Maculopathy	1973 Yes 1982 Yes	1982 Ishihara – No Response HRR – No Response Farns D-15 - Abnormal	1987 Relative Peripheral Loss	1987 Rod-Normal Cone-Abnormal	1973 - 'Poor dentition and defective enamel' noted. 1984 - 28 teeth extracted
0094.6	1964	Unaffected	NR	2007 20/15 OD 20/15 OS	2010 20/25 OD 20/20 ⁻² OS	No	ND	ND	ND	2011 - 4 front teeth on plate, 2 other teeth decayed
0094.7	1966	Affected	1973	1982 20/200 ⁺¹ OU Myopia	ND	1973 Yes 1982 Yes	1982 Ishihara – No Response HRR – No Response Farns D-15 - Abnormal	ND	ND	1973 – 'severe enamel dysplasia' noted 1977 - irregular teeth with brown discolouration, marked wearing, jumbled eruption pattern 1979 – 17 teeth extracted

Table 3.2: Clinical description of family 0094 from archived medical records, showing ophthalmological and dental examinations. **YOB:** Year Of Birth **NR:** Not Registered, **OD:** Oculus Dexter (Right eye), **OS:** Oculus Sinister (Left eye), **CF:** Counting Fingers, **HM:** Hand Movement, **ND:** Not Done, **HRR:** Hardy-Rand-Ritler Colour Test, **Farns D15:** Farnsworth D15 Colour Arrangement Test.

Gene	Accession #	Variant	dbSNP ID	Minor Allele Frequency (%)	Disease Segregation	SIFT Score	Polyphen-2
GNAT2	NM_005272	c.-32 A>G	rs2304355	11.2	No	NA	NA
		c.119-69 C>T	rs6658641	11.3	No	NA	NA
		c.546 G>A: p.T182T	rs1799875	43.5	No	NA	NA
CNGA3	NM_001298	c.848 G>A: p.R283Q	rs104893614	N/A	Yes	Deleterious	Probably Damaging
		c.1279 T>C: p.R427C	NA	N/A	Yes	Deleterious	Probably Damaging
		c.1580 T>G: p.L527R	NA	N/A	Yes	Deleterious	Probably Damaging
CNGB3	NM_019098	c.702 T>G: p.C234W	rs6471482	6.8	Yes	Tolerated	Benign
		c.892 A>C: p.T298P	rs4961206	34.1	Yes	Deleterious	Benign
		c.886_896del11insT: p.T296YfsX9		N/A	Yes	NA	NA
		c.919 A>G: p.I307V	rs13265557	5.7	No	Tolerated	Benign
		c.1148delC: p.T383fsX13		N/A	Yes	NA	NA
		c.1179-78 A>T	rs34307542	6.8	No	NA	NA
		c.1179-38 T>C	rs3735969	4.8	No	NA	NA
		c.1320+56 G>A	rs67463707	7	No	NA	NA
		c.1320+239 G>A	rs67858232	6.8	No	NA	NA
		c.1481-145 A>C	rs13258590	33.6	No	NA	NA
		c.2430+1303 G>A	rs17683284	4.3	No	NA	NA
		c.2430+1639 C>A	rs990192	43.7	No	NA	NA
c.2430+1881 T>C	rs990193	14.8	No	NA	NA		
PDE6C	NM_006204	c.252 G>A: p.L84L	rs1131978	16.7	No	NA	NA
		c.723+53 G>A	rs1223306	48.1	No	NA	NA
		c.808 T>A: p.S270T	rs701865	42.8	No	Tolerated	Benign
		c.1120-60 T>C	rs1856563	3.4	No	NA	NA
		c.1270-7 A>G	rs616522	29.5	No	NA	NA
		c.1380 C>G: p.T460T	rs3737228	27.3	No	NA	NA
		c.1629+81 C>G	rs2275227	15.2	No	NA	NA
		c.2284-72 A>G	rs10882298	38.7	No	NA	NA

Table 3.3: All variants discovered through the screen of *CNGA3*, *CNGB3*, *GNAT2*, and *PDE6C*. Variants considered pathogenic are shown in bold print.

Family	Gene	Allele 1	Allele 2
0094	CNNM4	c.1555 C>T: p.R519X	c.1555 C>T: p.R519X
1442	CNGB3	c.1148delC: p.T383IfsX13	c.1148delC: p.T383IfsX13
1492	CNGB3	c.1148delC: p.T383IfsX13	c.1148delC: p.T383IfsX13
1713	CNGB3	c.1148delC: p.T383IfsX13	c.886_896del11insT: p.T296YfsX9
1491	CNGA3	c.1580 T>G: p.L527R	c.1580 T>G: p.L527R
1723	CNGA3	c.1580 T>G: p.L527R	c.1279 C>T: p.R427C
1726	CNGA3	c.848 G>A: p.R283Q	c.1279 C>T: p.R427C

Table 3.4: Pathogenic variants discovered in all seven families.

3.4.5 Haplotype Construction and Analysis

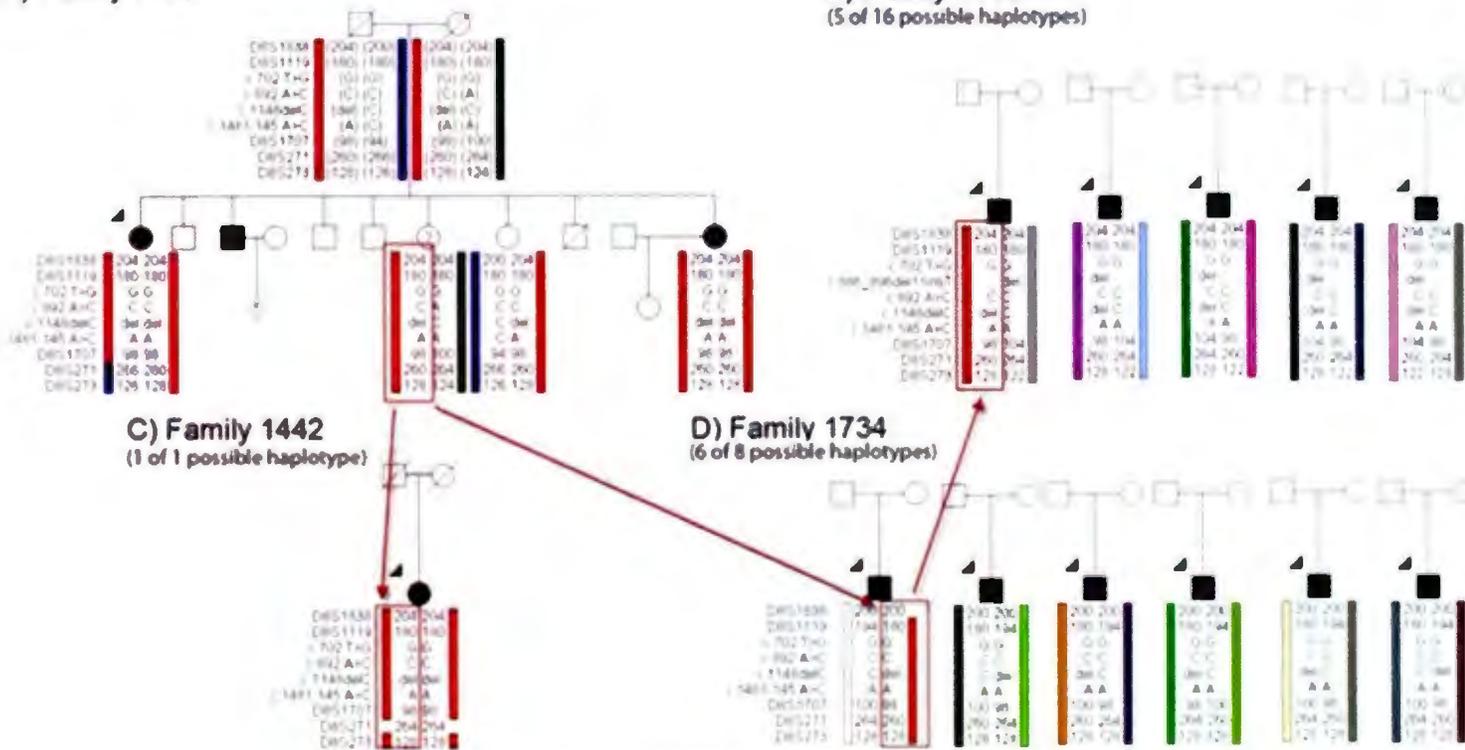
All variants discovered through Sanger sequencing as well as genotypes from flanking microsatellite markers were used to create haplotypes in each family with mutations in *CNGA3* and *CNGB3*. Examination of these haplotypes allowed for interfamilial comparisons, and determination of founder haplotypes/mutations.

(i) *CNGB3* Haplotypes

The haplotype surrounding the most common *CNGB3* mutation, c.1148delC, was constructed using family 1492, which has the most affected individuals with this mutation (Figure 3.2). This haplotype (shown in red) was then used to construct the haplotypes in families 1442, 1713 and 1734. It can be noted that small differences in genotypes exist, denoted by the light green portions of the haplotypes as compared to that of 1492. The individual from Family 1734 was heterozygous for the c.1148delC mutation, though a second mutation was not found. The c.1148delC founder haplotype can be seen in red, though it should be noted that though he appears to have the same haplotype as Family 1492 the genotype at *D8S271* (260/264) can be switched to have the same haplotype as 1442. To confirm the proper orientation of the haplotype, parental samples would need to be obtained to perform segregation analysis of this marker.

A) Family 1492

B) Family 1713
(5 of 16 possible haplotypes)



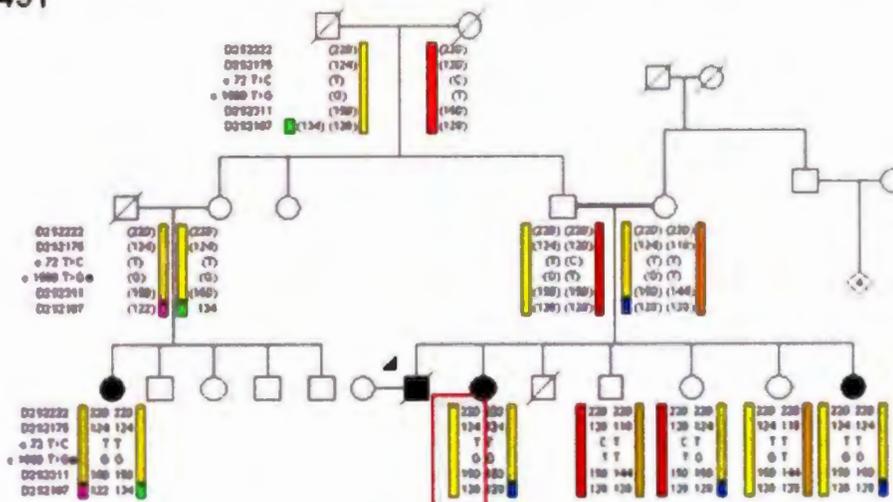
Marker	Genomic Coordinates (GRCh37/hg19)
D8S1838	84624454
D8S1119	27171877
D8S1707	87888313
D8S271	88518245
D8S273	88911428

Figure 3.2: Haplotypes of family members carrying *CNGB3* mutations. **A)** The haplotype surrounding the c.1148delC mutation was first constructed using Family 1492 and is denoted as the red haplotype which was homozygous between the two affected individuals. A recombination was observed between markers *D8S1707* and *D8S273* in the proband (indicated by the arrow). **B)** Possible haplotypes from Family 1713. We were limited to information on one individual from Family 1713, so all possible haplotypes are shown. The red box indicates that this red haplotype shows most similarity to the red haplotype constructed in Family 1492. **C)** Like Family 1713, only one individual was available from this family. Only one haplotype was possible based on the multiple homozygous genotypes. The red box and arrow show the similarity to the red haplotype from family 1492. **D)** Possible genotype combinations from the individual of Family 1734. The red box and arrow show the similarities to the red haplotype seen in the three other families. This data suggests that the c.1148delC mutation may exist on a common founder haplotype.

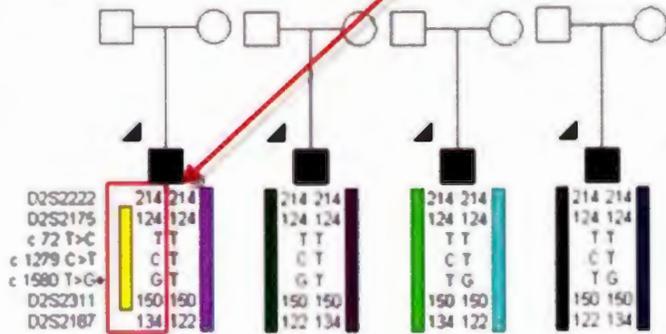
(ii) CNGA3 Haplotypes

Four microsatellite markers were genotyped to construct haplotypes surrounding each mutation. Affected members of Family 1491 shared a homozygous haplotype surrounding the c.1580 T>G mutation (Figure 3.3) which showed some intra-familial variation at the most 3' marker *D2S2187*. Assuming that family 1723 and 1491 are sharing the same haplotype, as they have the same mutation, the same yellow haplotype was also seen in Family 1723, in which the affected individual was a compound heterozygote for c.1580 T>G and c.1279 C>T. As the yellow haplotype could already be inferred from Family 1491, this made the construction of the haplotype surrounding c.1279 C>T possible (shown in light blue). The affected family member from Family 1726 did not show a common haplotype with Family 1723, despite harbouring the c.1279 C>T mutation as one allele.

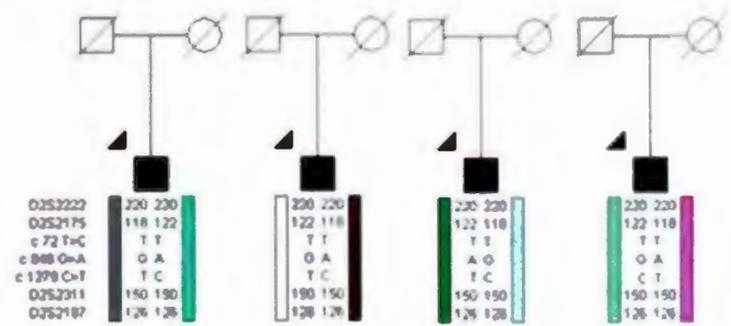
A) Family 1491
c.1580 T>G



B) Family 1723
(4 of 4 possible haplotypes)
c.1279 C>T / c.1580 T>G



C) Family 1726
(4 of 4 possible haplotypes)
c.848 G>A / c.1279 C>T



Marker	Genomic Coordinates (GRCh37/hg19)
D2S2222	98299348
D2S2175	98784684
D2S2311	99016618
D2S2187	99105207

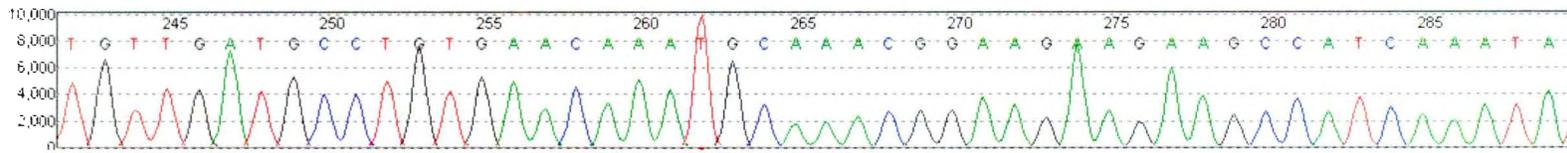
Figure 3.3: *CNGA3* haplotypes constructed using four microsatellite markers, one SNP and the pathogenic variants discovered through initial screening. **A)** The haplotype surrounding the c.1580 T>G variant was done first, as it is homozygous in Family 1491, which has the most individuals available for haplotype construction. The disease haplotype surrounding c.1580 T>G is denoted in yellow, though there are slight differences at particular markers (i.e. *D2S2187* has genotypes of 122/126/128/134) denoted by differing colours. **B)** Possible haplotypes from family 1723. One sample was obtained from this family, and four haplotypes were possible. A similar haplotype surrounding c.1580 T>G is seen in family 1723 despite differences in the two most terminal markers (*D2S2187* and *D2S2222*) and is denoted with a red box and arrow from Family 1492. The c.1279 C>T mutation was also seen in family 1723. **C)** Four possible haplotypes based on information from the individual of family 1726. None of the haplotypes from this family match the possible ones from 1723 who both carry the c.1279 C>T mutation.

3.4.6 Unsolved Families (Families 1734 & 0094)

(i) Family 1734

Screening of the three previously associated achromatopsia genes *CNGA3*, *CNGB3*, and *GNAT2* revealed one heterozygous pathogenic mutation (c.1148delC) in *CNGB3* though a second pathogenic variant was not identified in this gene. This sample was then sequenced for *PDE6C*, the fourth known achromatopsia gene. This screen identified four confirmed homozygous SNPs (rs1223306, rs1856563, rs616322, and rs10882298) all of which were excluded from further analysis due to their high frequency in the general population (Table 3.3), and one sequence anomaly. The sequence surrounding exons 17 and 18 of *PDE6C* showed an odd sequence which in the forward strand, started with high background but at nucleotide c.2259T became of better quality (Figure 3.4). This pattern was consistent across multiple attempts. It was hypothesized that this anomaly was being caused by an insertion/deletion event to cause the sequence to appear shifted, followed by a second insertion/deletion event to bring the sequence back into frame. To examine the nature of this sequence anomaly, this amplicon was inserted into a pCRTM4-TOPO® vector after amplification using primers PDE6C17+18-F and PDE6C17+18-R and cloned into TOP10 Chemically Competent E.Coli cells as per manufacturers protocol and grown overnight on LB/Agar plates. Six colonies were chosen from these plates, and the plasmid isolated and purified. Sequencing of the plasmids and inserts revealed one clone with a sequence matching that of exons 17 and 18 of *PDE6C* (Appendix I pink). Two other clones were noted to have sequence which belonged to neither *PDE6C*, nor the plasmid, except for a 20 bp sequence consistent with the primer PDE6C17+18-R (Appendix I). The remaining three clones showed small fragments of this aberrant sequence from partial ligations. The entire length of this unknown sequence was run through UCSC's Blast-Like-Alignment-

PDE6C (NM_006204) Reference Sequence



Family 1734.1 Sequence

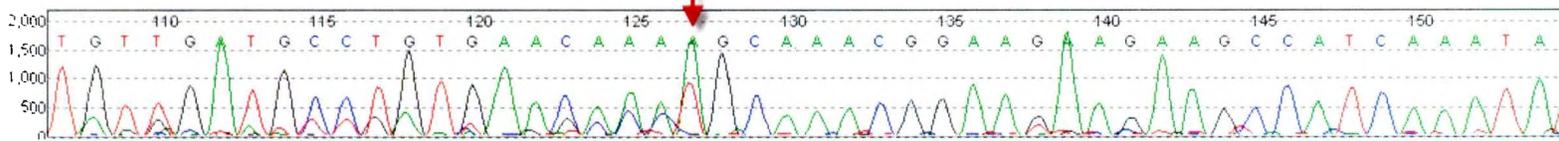


Figure 3.4 – Sequence surrounding exon 17 of *PDE6C* illustrating the presence of a sequence anomaly. (Top) Reference sequence of *PDE6C* (NM_006204.3) obtained from the NCBI Nucleotide database. (Bottom) Experimental sequence obtained from sequencing the proband of family 1734. The sequence to the left of the red arrow shows high background, whereas the sequence to the right of the red arrow shows much cleaner sequence.

Tool (BLAT) to determine the origin of this sequence. This analysis revealed a 99.2% sequence similarity to an intergenic region on chromosome 8q24.13, making it unlikely to be caused by two insertion/deletion events, as described above. To explore if this sequence anomaly was being caused by a potential translocation event from Chr8, three primers were designed:

Primer 1) 1734.1-Chr8-F (GTGGAGGTTCCCAAACCTTGA) which according to BLAT was specific to the intergenic region on Chr8q24.13

Primer 2) 1734.1-PDE6C-R (ATTCCCAATTCATGGGAG) which according to BLAT was specific to intron 18 of *PDE6C*

Primer 3) 1734.1-Chr8-843bp-R (CTGGAACTGGGTAAACAAGCAG) which was specific to Chr8q24.13.

It was hypothesized that if a true translocation event had occurred, that Primers 1+2 would generate a band upon PCR amplification, indicating that the Chr8 region was in close proximity to the Chr10 region. Primer 3 was used as a control for the Chr 8 region and was hypothesized to amplify a band 843 bp in size when combined with Primer 1. PCR was carried out using a sample from the proband of 1734, and a control DNA sample from family 0094 (0094.4) who did not show this sequence anomaly upon sequencing of *PDE6C*. Results from the experiment showed no amplification in either the experimental sample or the control using Primers 1+2, but Primers 2+3 showed the expected band in both the sample and control (Figure 3.5). This suggests that the sequence anomaly in Figure 3.4 is not due to a translocation event.

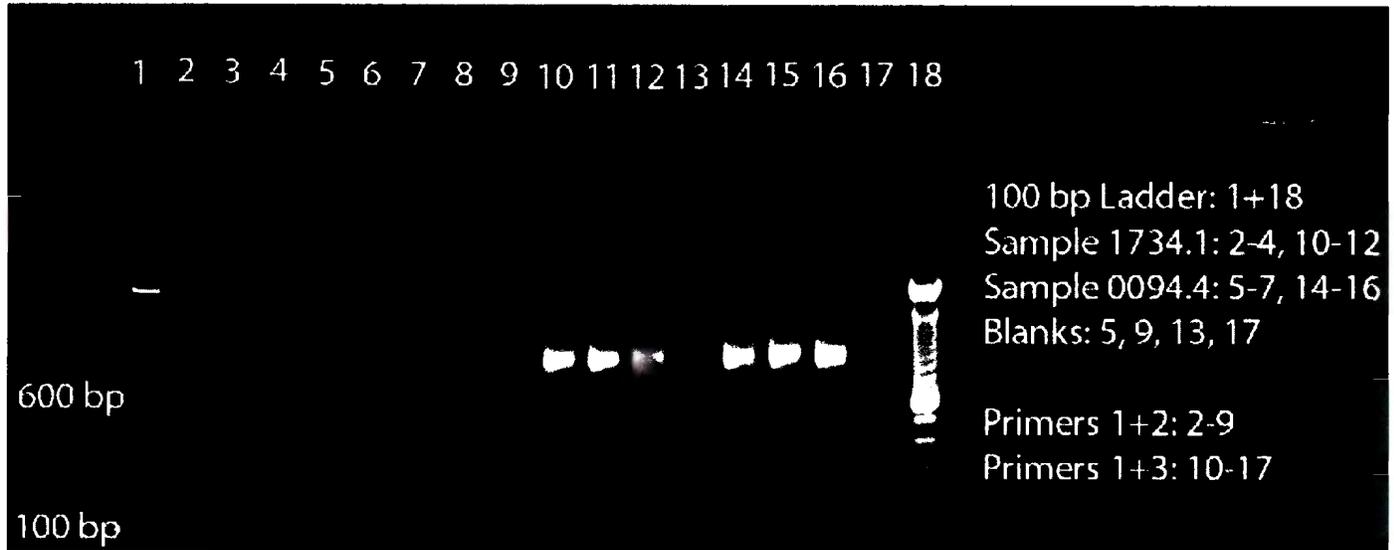


Figure 3.5 – 1% Agarose gel stained with SYBR safe showing PCR amplification of two samples (1734.1 and 0094.4) using three primers. Primer 1 = 1734.1-Chr8-F, Primer 2 = 1734.1-PDE6C-R, Primer 3 = 1734.1-Chr8-843bp-R. No amplification was seen using Primers 1+2 in either sample, suggesting that the sequence anomaly isn't from a translocation event. The expected band for Primer 1+3 was obtained in both samples.

Upon receiving the primers, a note was attached from the manufacturer that Primer 2 (1734.1-PDE6C-R) contained significant secondary structure which may interfere with some assays. It is possible that no amplification would be seen due to this fact, and that the sequence anomaly may still be due to a translocation event. To once again test this hypothesis, the chromosome 8 specific primer, Primer 1 (1734.1-Chr8-F) was paired with the reverse primer PDE6C-17+18-R from the initial sequencing of *PDE6C*. It was hypothesized that if amplification occurred in sample 1734.1 but not in sample 0094.4 that it would be possible that there was indeed a translocation event present. This experiment showed amplification in both samples (Figure 3.6), suggesting that perhaps the reverse primer (Primer 2) was binding to chromosome 8, even though BLAT analysis suggested it was specific to *PDE6C*. Next, two things were examined:

- 1) If the secondary structure of Primer 2 is sufficient to stop amplification
- 2) If the primer PDE6C-17+18-F had the potential to bind to other regions non-specifically, like its complementary primer PDE6C-17+18-R.

To examine the first question, PDE6C17+18-F was paired with Primer 2. It was hypothesized that if amplification occurred, then the secondary structure was insufficient to stop PCR amplification, and that the result from Figure 3.6 suggested there was no translocation event in 1734.1. To answer the second question, PDE6C-17+18-F was paired with Primer 3 (1734.1-Chr8-843bp-R) which was shown to bind to multiple sites upon BLAT analysis. It was hypothesized that if amplification occurred, that PDE6C-17+18-F was also a nonspecific primer but would amplify a specific band when paired with a more specific primer (as seen in Figure 3.6). The results from this experiment showed that a single band amplified in Lane 7 of the gel

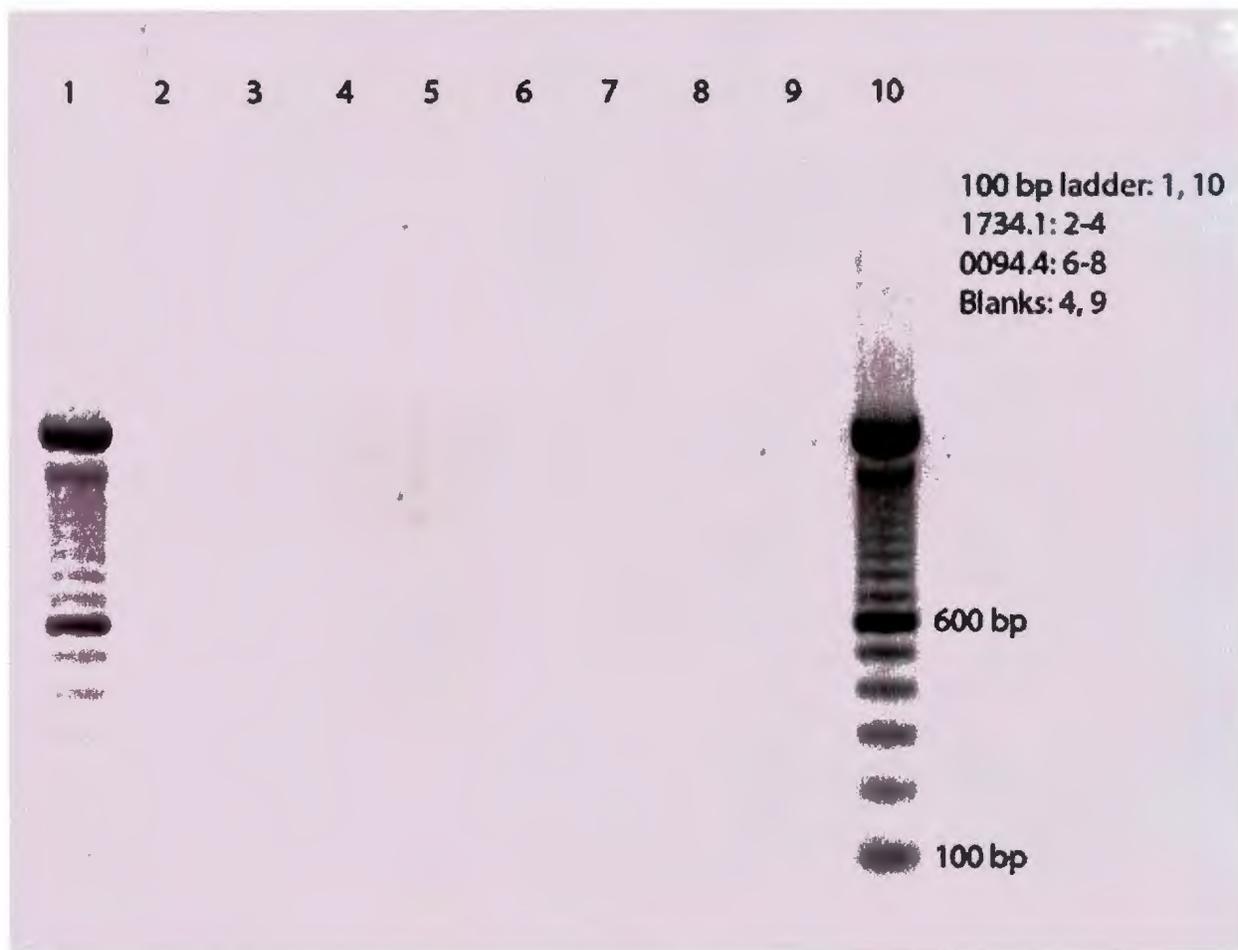


Figure 3.6 – 1% Agarose gel stained with SYBR safe showed amplification of 1734.1 and a control sample 0094.4 with Primer 1 and PDE6C17+18-R. Amplification was very weak, thus to better view the bands, the colour was inverted in this figure.

(Figure 3.7) suggesting that Primer 3 can in fact produce amplification, though not efficiently (as there was no amplification in the other lanes. The experiment addressing the second question showed three bands in total (Figure 3.7), suggesting that PDE6C-17+18-F has binding sites other than *PDE6C*. It was also noteworthy that PDE6C-17+18-F and -R were difficult to amplify and sequence in all tested samples. Therefore, a second primer set surrounding *PDE6C* exons 17 and 18 was designed and used to sequence this region in Family 0094 and Family 1734, which showed no pathogenic variants in any tested sample.

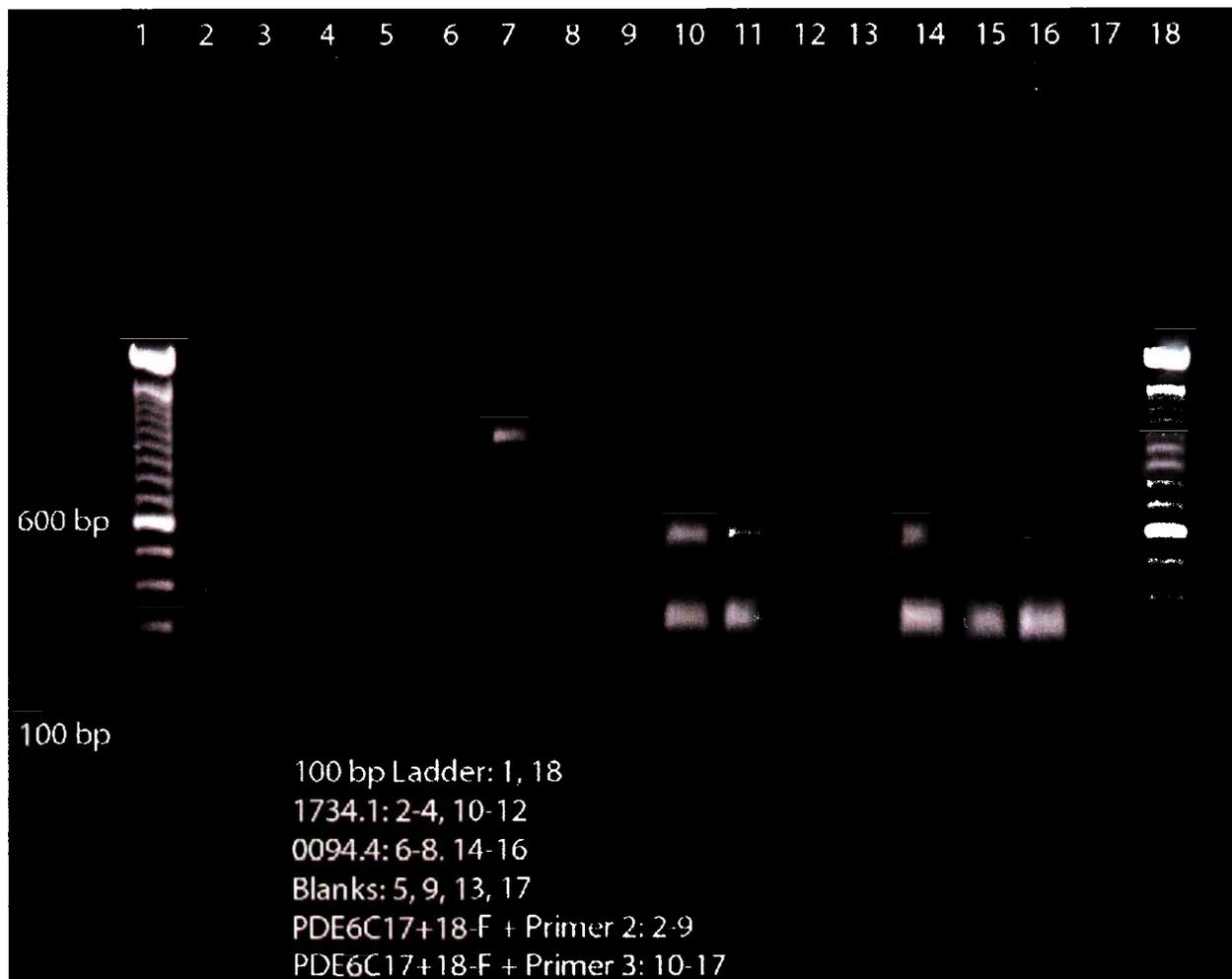


Figure 3.7 – 1% agarose gel stained with SYBR safe. Lanes 2-9 illustrate the combination of Primers 17+18-F and Primer 2, which was known to have significant secondary structure. The amplified band in lane 7 but in no other suggests that the reverse can amplify, but not efficiently. Lanes 10-17 illustrate the combination of PDE6C-17+18-F and Primer 3, which is known to bind to multiple regions of the genome (though produces a single band when paired with a specific primer, see Figure 3.6). The amplification of three bands in both the experimental sample (1734.1) and the control sample (0094.4) suggests that the PDE6C-17+18-F primer has binding sites other than PDE6C.

(ii) *Family 0094*

No pathogenic mutations were found in the affected members of Family 0094 in *CNGA3*, *CNGB3*, *GNAT2* or *PDE6C*. To check for segregation of a disease associated haplotype in *CNGA3*, *CNGB3*, *GNAT2* or *PDE6C*, which might suggest a regulatory mutation missed on initial sequencing, we constructed intragenic haplotypes with discovered SNPs as well as microsatellite markers, however no haplotypes co-segregated with the phenotype in this way (Figures 3.8-3.11)

Exome sequencing was then carried out at the McGill University Genome Centre on pooled genomic DNA from two affected siblings and on genomic DNA from the unaffected parent. Pooling of DNA was done to both reduce costs, and to enrich for regions of the genome shared amongst affected individuals to ease identification of a pathogenic variant. Target capture with Illumina TruSeq exome enrichment kit and sequencing of 100 bp paired-end reads on Illumina HiSeq yielded approximately 10 Gb of sequence for each DNA sample such that >90% of the coding bases of the exome defined by the consensus coding sequence (CCDS) project were covered by at least 20 reads. Sequence reads were aligned and variants were called and filtered as previously described.¹⁸⁶ Exome sequencing yielded 11,292 non-synonymous coding variants in the affected pool versus 10,399 in the unaffected parental exome at ~80x coverage. We filtered the variants in the affected pool to remove those seen in more than one of ~200 unrelated exomes sequenced at Genome Quebec, which left 255 variants. Of these 255 variants, only seven were homozygous in the affected pool (one nonsense), which would be expected with consanguinity and an autosomal recessive mode of inheritance (Table 3.5). These seven variants included a novel nonsense mutation (c.1555 C>T: p.R519X) in *CNNM4* (*Ancient Conserved Protein Domain 4*; *ACPD4*; NM_020184), the only gene known to cause Jalili Syndrome. We verified co-segregation of c.1555 C>T: p.R519X with achromatopsia in Family

0094 by Sanger Sequencing, and this variant was not present in 236 ethnically-matched control chromosomes.¹⁸⁷

A search of archival medical records from the 1970s for documentation of dental abnormalities revealed that each of the four siblings diagnosed with achromatopsia had markedly abnormal dentition as young children with enamel dysplasia, brown discolouration of teeth and irregular tooth eruption. These four siblings had undergone surgical extraction of all primary and secondary teeth (Table 3.2). At the time of the ophthalmological examination in 1982, all siblings had dentures which were not uncommon in rural NL as dental care in many outport communities was unavailable and extraction was a common treatment for common dental caries. Ocular re-examination of two of the siblings with achromatopsia in 2011 showed a further reduction of visual acuity to hand movements (HM) at 8 feet (previously 20/200 in IV-4) with additional macular and peripheral retinal changes noted (Table 3.3), revealing a progressive retinal phenotype, consistent with cone-rod dystrophy associated with Jalili syndrome, rather than achromatopsia which tends to be a stationary condition.

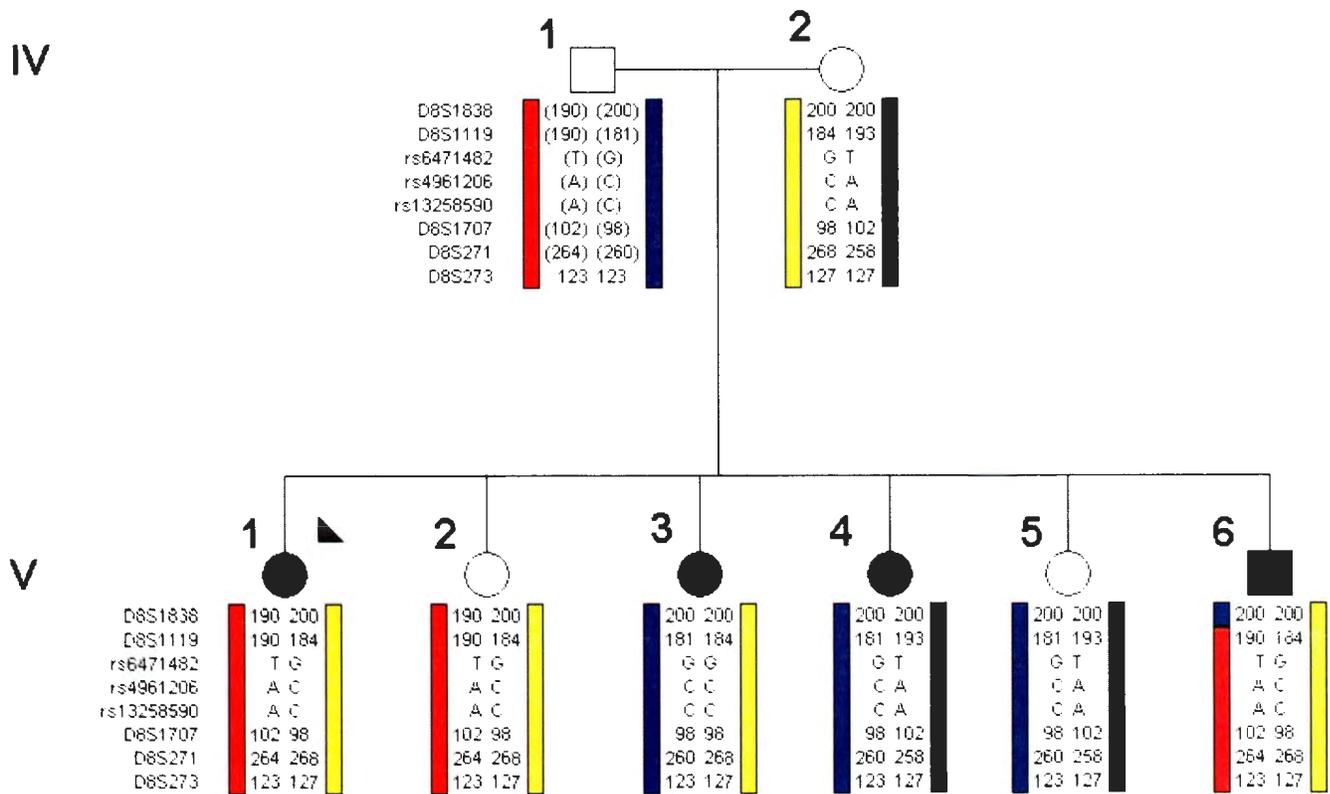


Figure 3.8: Partial pedigree of family 0094 with haplotypes of *CNGB3* variants and marker genotypes. There are no common haplotype combinations within all four affected individuals. For example, PID IV-1 shares common haplotypes with her unaffected sister PID IV-2, which is not shared by any other sibling. PID IV-3 and IV-4 both share a blue haplotype though they have yellow or green haplotypes as their second haplotypes. This data, along with the fact that no novel mutations have been found in family 0094 within *CNGB3*, excludes *CNGB3* as the causative gene in this family.

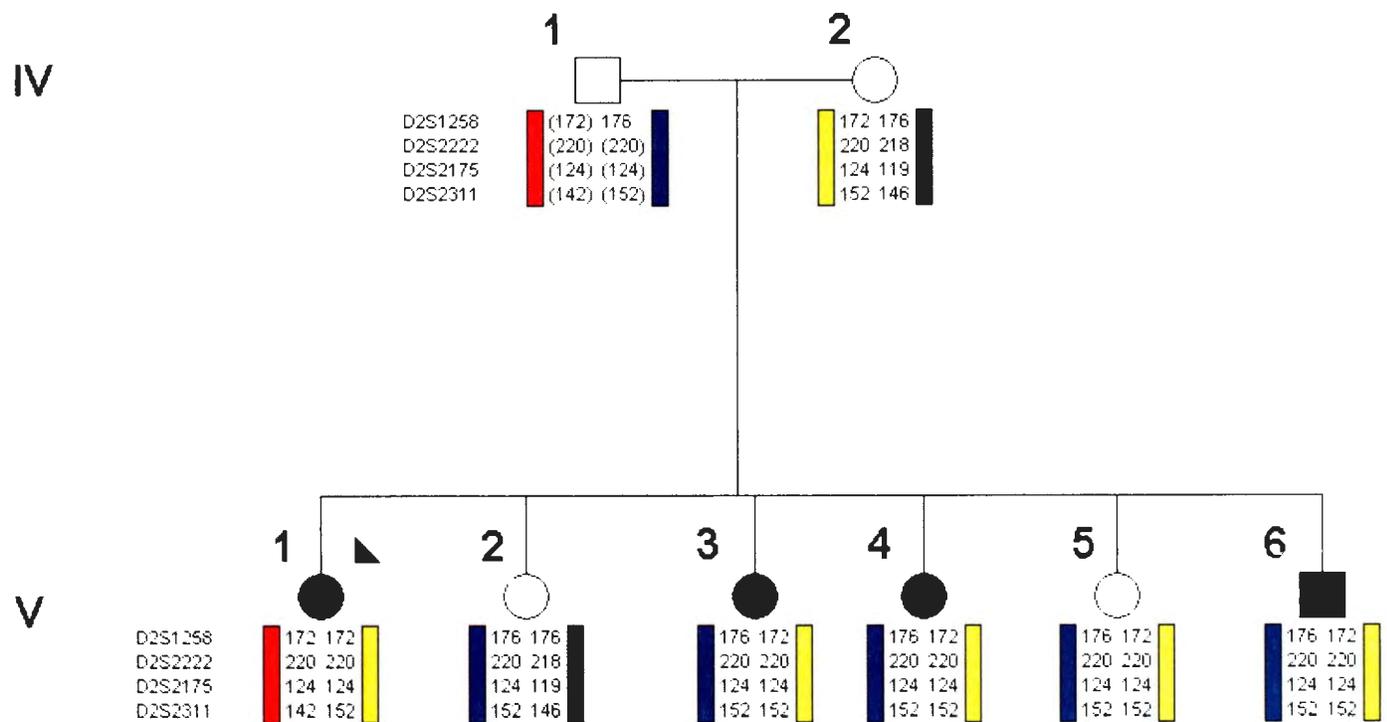


Figure 3.9: Partial pedigree of family 0094 with haplotypes of *CNGA3* marker genotypes. No SNPs or novel variants were discovered in *CNGA3* within Family 0094 members, suggesting a highly conserved structure and sequence within the gene and protein. Three affected individuals (PID IV-3, IV-4, and IV-6) both share both the blue and yellow haplotypes, but this is excluded as one unaffected individual (PID IV-5) also shares these haplotypes. The fourth affected individual (PID IV-1) has a red and a yellow haplotype, providing further evidence for the exclusion of *CNGA3* in this family.

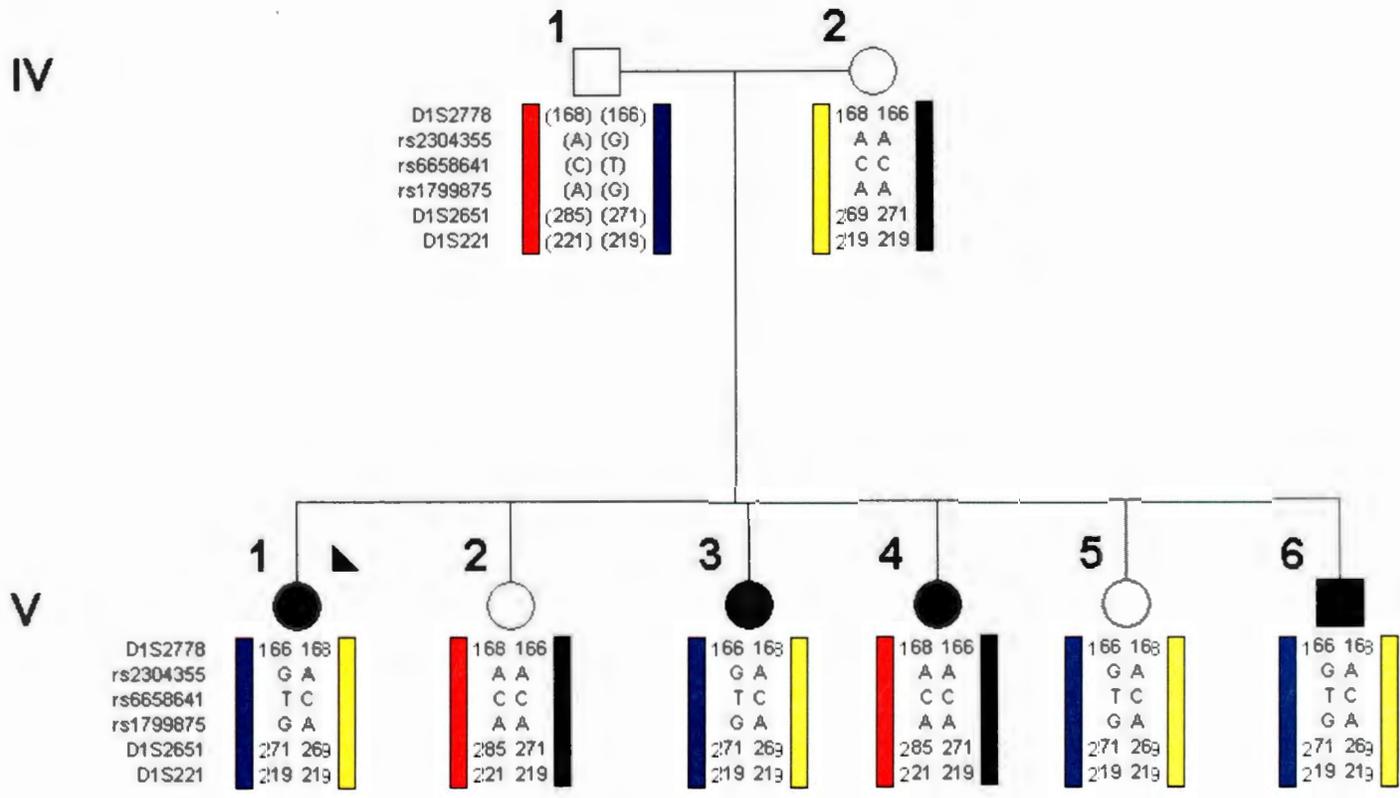


Figure 3.11: Haplotypes of variants and markers surrounding and within the *GNAT2* gene. Three of the affected individuals (PID IV-1, IV-3, and IV-6) share blue and yellow haplotypes though the unaffected sister (IV-5) also shares these same haplotypes. As well, individual IV-4 have two different (red and green) haplotypes providing further evidence for exclusion of *GNAT2* as the causative gene in this family.

Coding variants in Affected sample pool	Coding variants in Unaffected Sample	Variants seen in ≤ 2 previous exomes	Homozygous Variants seen in ≤ 2 previous exomes	Homozygous variants seen in ≤ 2 previous exomes, and absent or heterozygous in Unaffected sample	Number of Genes with previous association to retinal disease
11,292	10,399	255	7	7	1

Table 3.5: Table of variant filtering from the exome dataset. As family 0094 segregated an autosomal recessive form of disease, the exome dataset was initially filtered for homozygous variations. One nonsense mutation was revealed to be previously associated with a retinal phenotype, c.1555 C>T: p.R519X in *CNNM4*.

3.5 Discussion

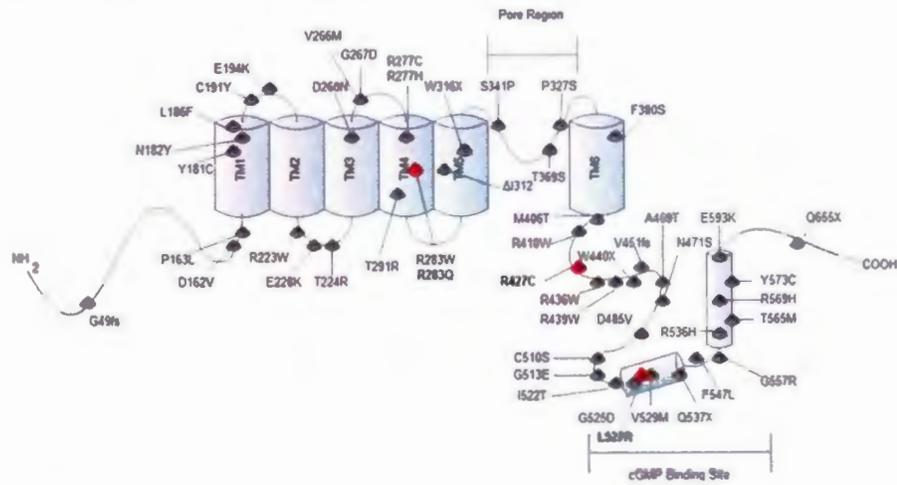
Mutations in *CNGA3* and *CNGB3* have been shown to account for a majority of achromatopsia cases worldwide. In this study of the Newfoundland population, mutations of these two genes solved six out of the seven recruited families. Two families (1442 and 1492) were homozygous for the most common mutation in *CNGB3* c.1148delC: p.T383IfsX13. A third family (1713) also had the c.1148delC: p.T383IfsX13 mutation but it was a compound heterozygote with a 10 bp indel, c.886_896del11insT: p.T296YfsX9 (Table 3.4). It is likely that both these mutations undergo nonsense mediated mRNA decay as they introduce a premature stop codon early in the open reading frame, and the mutation exists >50bp from the terminal exon.¹⁸⁸ The c.1148delC mutation is the most frequent mutation causing achromatopsia worldwide, and this was also seen in the achromatopsia patients of the Newfoundland population. Of the 20 alleles that were studied in patients with achromatopsia, mutations in *CNGB3* mutations accounted for nine of them (45%), while of these nine mutations, c.1148delC accounted for eight of them (89%).

The other three families (1491, 1723, and 1726) were solved by mutations in *CNGA3*. Affected members of Family 1491 were homozygous for the previously reported c.1580 T>G: p.L527R¹⁸⁵ mutation which exists in the cGMP binding site of the *CNGA3* protein (Figure 3.12). Conservation analysis by Lam *et al.* show that the L527 residue is highly conserved from *Homo sapiens* to *Drosophila melanogaster* thus suggesting that this residue is located within a functionally important domain of the *CNGA3* subunit. Furthermore, *in silico* analysis using SIFT, Polyphen-2, and PANTHER (Table 3.3) suggest that the L527R variant is highly damaging. As the mutation exists in the cGMP binding site, it is likely that this mutation impedes the ability of CNG to bind cGMP, disrupting the closure of the cyclic nucleotide gated channel (CNG) of the cone photoreceptor, though no functional analysis has been done on this

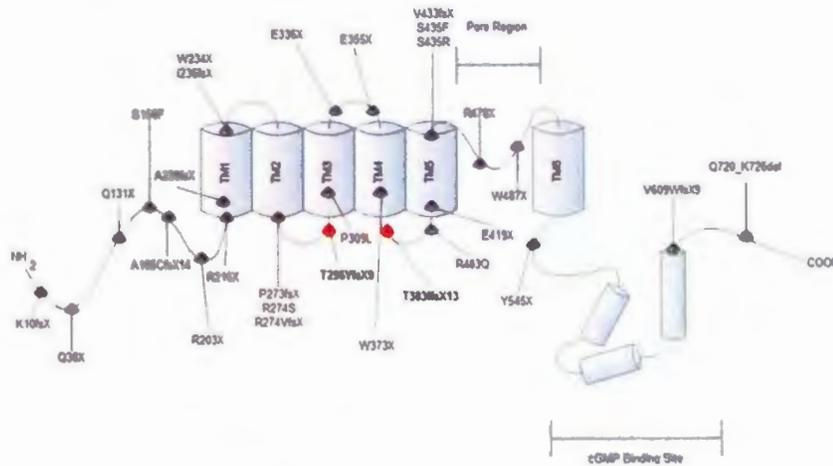
mutation to date. The p.L527R mutation was also seen in the single affected member of Family 1723, as a compound heterozygote with c.1279 C>T: p.R427C. Previous studies of p.R427C show that this mutation impedes the trafficking of CNGA3 to the cellular membrane.¹⁸⁹ It is possible that due to the relatively close proximity of p.R427C to p.L527R, that p.L527R may also act in this manner. The sixth family, 1726, also had the c.1279 C>T: p.R427C mutation as a compound heterozygote with c.848 G>A: p.R283Q. Of the 20 mutant alleles analyzed in this study, 10 of them were in *CNGA3* (50%), and the most common mutation seen was the c.1580 T>G: p.L527R mutation with seven out of these 10 (70%). This is of course biased as we have more information and samples on family 1491, than on the other two families 1723 and 1726.

Functional analysis of the p.R283Q mutation has shown that this mutation affects the ability of the CNG channel within HEK293 cells to conduct calcium flow.¹⁸⁹ The combination of p.R427C and p.R283Q in this patient potentially causes achromatopsia through reduction of functional CNGA3 subunits. This is most likely accomplished through the combination of aberrant trafficking and reduction of calcium influx. It is interesting to note that there are a multitude of pathogenic missense variants in *CNGA3* which are known to cause achromatopsia, whereas the mutations in *CNGB3* tend to be null alleles such as nonsense or frameshift mutations (Figure 3.12A and 3.12B), suggesting that the CNGA3 protein is more readily affected by mutation. This lends to the notion that the CNGA3 subunit is the major functional subunit of the CNG channel, and that mild mutations of *CNGB3* can be compensated for by a functional CNGA3 subunit.

A) CNGA3



B) CNGB3



C) CNNM4

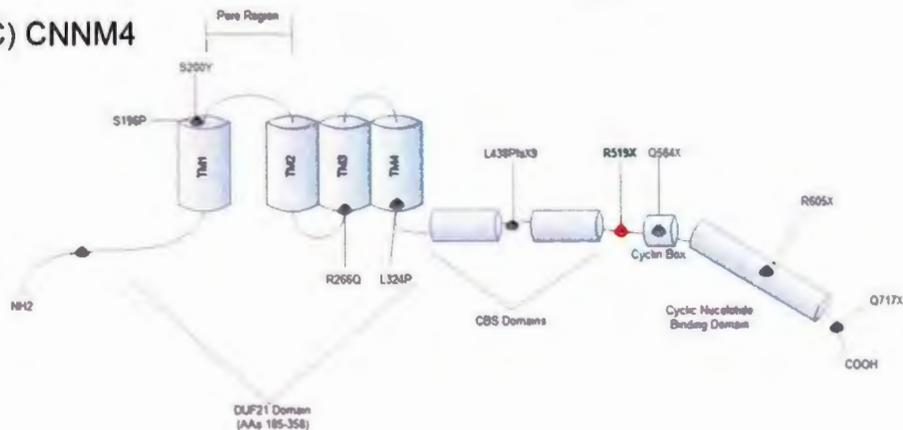


Figure 3.12: Protein schematic of **A) CNGA3** and **B) CNGB3** showing all functional domains and previously identified mutations and their approximate position within the protein. **C)** Hypothetical protein structure of CNNM4 based on the structure of its homologue CNNM2. All the previously described mutations and their approximate position in the various domains of the protein are shown. The novel nonsense mutation p.R519X is shown in bold print and its position between the two CBS domains of CNNM4, is indicated with a red marker.

To examine the potential existence of a founder effect within *CNGA3* and *CNGB3* in the Newfoundland population, haplotypes surrounding each mutation were constructed. This revealed a common haplotype (Figure 3.2) surrounding the c.1148delC mutation in *CNGB3* across all four families who have this mutation (1442, 1492, 1713, and 1734). Small differences can also be seen in the haplotype which suggests that either there are many haplotypes with this mutation in existence (suggesting a recurrent mutation), or that the haplotype has mutated over time. The latter notion seems to be the most likely, as it has been suggested that c.1148delC:p.T383IfsX is a founder mutation in populations other than Newfoundland.¹⁷³ It is likely this mutation is an ancient mutation, and that the differences seen in the haplotypes in this study are most likely due to haplotype mutating over the course of time. As we were only able to obtain DNA from one individual of Family 1442 and 1734, all possible genotype combinations are shown in Figure 3.2. In order to confirm that the haplotype surrounding c.1148delC is indeed similar to that seen in Family 1492; parental DNA would need to be obtained, though this conclusion seems likely.

The haplotypes surrounding the three mutations of *CNGA3* were also examined. First, the haplotype surrounding c.1580 T>G in Family 1492 was examined, as we had the most samples available for genotyping, and we knew that family members were homozygous for this mutation. This haplotype (Figure 3.3) segregates with the disease phenotype except for the most terminal marker *D2S2187*, which varies significantly between affected individuals. Like the situation for the *CNGB3* haplotypes, we were only able to obtain one sample from families 1723, and 1726, and all possible genotype combinations are shown in Figure 3.3. The disease haplotype is also observed in Family 1723, who was compound heterozygous for c.1580 T>G and c.1279 C>T. The common haplotype seen in families 1491 and 1723 surrounding the c.1580 T>G mutation

suggests that these two families may share a common ancestor. The only other report of this mutation was in a population of European origin and was identified through exome sequencing. As the Newfoundland population is of Northern European origin, it would be interesting to examine if the c.1580 T>G mutation was on a common haplotype, suggesting a founder mutation of European origin. Interestingly, the c.1279 C>T mutation, which was shared between families 1723 and 1726 appeared to be on two separate haplotypes, suggesting that this mutation has occurred on separate genetics backgrounds and could potentially be a recurrent mutation. This would need to be confirmed on samples from various populations.

Two unsolved families, 1734 and 0094 remained after the initial screens of *CNGA3*, *CNGB3*, *GNAT2*, and *PDE6C*. The sample from Family 1734 showed only one mutation in *CNGB3*, c.1148delC and a second mutation was unable to be found. Screening of *PDE6C* showed an interesting sequence anomaly (Figure 3.4), which was at first hypothesized to be due to an insertion/deletion causing a frameshift, followed by a second insertion/deletion renewing the reading frame. Cloning experiments revealed that this was not the case in this sample, and showed that one strand in fact showed 99.2% sequence similarity to a region on chromosome 8. It was then hypothesized that this abnormal looking sequence was caused by a potential translocation event from chromosome 8 to chromosome 10, and that we had sequenced the breakpoint of insertion. A number of primer combinations revealed that the primers PDE6C-17+18-F and -R were in fact binding to other regions of the genome, and that this result was most likely due to the primers binding to this region in Chr8. It was also noteworthy that this region was particularly difficult to amplify, and sequence using these primers and that sequences from family 0094 did not show this particular variation, but sequences surrounding this region were of poor quality. A new primer set was designed and sequenced in all samples which

showed much more efficient amplification and sequencing results, revealing no pathogenic variations. Unfortunately, this case remains unsolved though it remains likely that the second mutation is to be found in a controlling region surrounding *CNGB3* (enhancer, repressor, splice variation) as the subject contains one mutation in this region already (c.1148delC).

The most interesting result of this study comes from the remaining unsolved family 0094 who showed no pathogenic variants, or common haplotypes amongst affected individuals in *CNGA3*, *CNGB3*, *GNAT2*, or *PDE6C* (Figure 3.8-3.11). Exome sequencing of individuals from this family revealed a novel nonsense mutation (c.1555 C>T; p.R519X) in *CNNM4*, the only gene known to cause Jalili syndrome, an extremely rare syndromic form of cone-rod dystrophy. The presence of a homozygous nonsense mutation in *CNNM4*, documentation of abnormal dentition in the archived medical records and the progressive nature of the retinal phenotype support the revised diagnosis of Jalili syndrome in Family 0094. The dental phenotype was missed in the initial assessment as the siblings were young adults when seen, and had dentures. Dental extraction and replacement with dentures was a common treatment for common dental caries in rural Newfoundland due to lack of routine dental services, thus the dental phenotype was missed at first examination.

CNNM4 shows high sequence similarity to the cyclin family of genes, and although it has no cyclin activity *in vivo*, may play a role in metal ion homeostasis, as it interacts with the metal ion chaperone COX11¹⁹⁰ and shows sequence similarity to a known magnesium transporter CNNM2.¹⁹¹ Others have shown that *CNNM4* is expressed in both the neural retina and ameloblasts of developing teeth, supporting a role for *CNNM4* in both the phototransduction cascade and in the biomineralization of teeth.¹⁹¹ Based on the amino acid sequence similarity to CNNM2, we used both the CNNM2 crystal structure¹⁹² and the known domains of CNNM4

(Transmembrane, CBS, DUF21, and Cyclic nucleotide binding domains)¹⁹¹ to produce a hypothetical structure of the 775 amino acid CNNM4 protein (Figure 3.12C), on which we have indicated the nine previously seen mutations.^{187; 191; 193-196} The novel transition (c.1555 C>T) identified in Family 0094 is located within exon 3 of *CNNM4*, and the resulting nonsense mutation (p.R519X) resides between the cyclin and CBS domains of the CNNM4 protein (Figure 3.12C). It is predicted, that because the c.1555 C>T mutation occurs >50 bases from the terminal exon, the *CNNM4* mRNA transcript would undergo non-sense mediated mRNA decay.¹⁸⁸

Whole exome/whole genome approaches hold promise for clinical diagnostics, particularly when targeted gene screening has failed to yield the underlying etiology. In this report, the molecular diagnosis was critical to making the correct clinical diagnosis. There will be times when grouping clinical features in individuals assumed to have only one disorder may not be appropriate as it is also possible to have a combined phenotype caused by mutations in more than one gene. This was recently illustrated in a study which describes a patient with Jalili syndrome and Neurofibromatosis type 1, due to mutations in both *CNNM4* and *NF1*.¹⁹⁶

This study used a combination of functional candidate gene screening for cases of achromatopsia, whole exome sequencing, and a review of archived medical records to identify a novel disease-causing mutation in *CNNM4* and a heretofore-unrecognized diagnosis of Jalili syndrome. We describe a number of disease causing mutations in *CNGA3* and *CNGB3* in achromatopsia as well as the first case of Jalili syndrome in North America and show the value of next generation sequencing as a diagnostic tool. In combination with accurate medical records this molecular data can be used not only to explain unsolved disease cases, but also to confirm or revise diagnoses.

3.6 Author Contributions

Study design was conceived by: T.L.Y., J.G., L.D., C.B. Ophthalmological investigations carried out by: J.G. and G.J.J. Data was generated and reviewed by: T.L.Y., J.G., L.D, C.B. and D.G. Exome sequencing was carried out by J.S. Manuscript was prepared by: T.L.Y., L.D., and J.G.. All parts of this study were supervised by T.L.Y. and J.G.

3.7 Acknowledgements

We gratefully thank family members for participating in this study. This work was supported by grants from the Canadian Foundation for Innovation (New Investigator Award no. 9384; Leaders Opportunity Fund no.13120) and Genome Canada (Atlantic Medical Genetics and Genomics Initiative) to TLY. The authors also gratefully acknowledge financial support from the CNIB E.A Baker Foundation for the Prevention of Blindness, the Janeway Children's Hospital Foundation, Memorial University IRIF and the Government of Newfoundland and Labrador. Exome sequencing was carried out by McGill University and Génome Québec Innovation Centre, Montréal, Canada.

3.8 Conflict of Interest Statement

The authors declare no conflict of interest.

Chapter 4: A Rare Form of Microphthalmia Associated with Proportionate Dwarfism (MDW) is Linked to a Large Region on Chromosome 16.

Credit and Significant Contributions

Study design was conceived by Terry-Lynn Young, Jane Green, and Lance Doucette. Ophthalmological and clinical investigations were carried out by Jane Green, Bridget Fernandez and Gordon Johnson. Sanger sequencing data was generated and reviewed by McGill Genome Center, Lance Doucette and Dante Galutira. Microsatellite genotype data generation and haplotype construction was carried out by Lance Doucette. SNP genotyping and homozygosity haplotyping analyses were performed by McGill Genome Center. Linkage analysis was performed by Nicole Roslin of the Toronto Center for Applied Genomics. Copy Number Variation (CNV) analysis was carried out by Alexandre Monpetit of McGill University. Exome sequencing was carried out by McGill Genome Center, and variant filtering analysis was performed by Jeremy Schwartzentruber. RNA and cDNA experiments were designed and carried out by Lance Doucette. Karyotype analysis of patient IV-2 of family 0066 was carried out through Eastern Health. All parts of this study were supervised by Terry-Lynn Young, and Jane Green.

4.1 Summary

This thesis chapter describes two families (0066 and 1499) from the South coast of Newfoundland with a Hallerman-Streiff like syndrome, which we refer to as Microphthalmia-Dwarfism (MDW). The patients presented at birth with microphthalmia, a reduced axial length of the ocular globe and developed cataracts and corneal opacities at an early age, causing a loss of visual acuity. These patients were also noted, over time, to have short stature with proportionate limbs, as well as a progressive loss of hair (hypotrichosis), frontal bossing, and pointed nose due to loss of facial fat deposits. The phenotype has been described to be similar to that of the very rare Hallermann-Streiff Syndrome (Francois Dyscephalic Syndrome), a condition characterized by dyscephaly (malformation of the cranium and bones of the face), hypotrichosis, microphthalmia, cataracts, beaked nose, micrognathia, and proportionate short stature (dwarfism). It is possible that the MDW phenotype described here may be part of a spectrum including Hallermann-Streiff syndrome and the causative genes in these two conditions could play roles in similar or related pathways. We obtained a total of 31 DNA samples (five from affected individuals, 26 from unaffected individuals) across these two families and performed Homozygosity Haplotyping (HH) and Linkage analyses using SNP genotypes from a 610K Illumina SNP array. These analyses suggested two overlapping regions on Chr16q21 (Linkage: 30cM, HH: 8cM). Twenty-three genes were chosen in the overlapping 8cM region of 16q21 as determined by the HH analysis and bidirectionally sequenced. Ninety-eight variants were discovered in this screen, though none were determined to be pathogenic, but in combination with microsatellite markers were used to reduce the critical region to ~3cM. Exome sequencing was performed, and filtering for variants with emphasis on the 3cM critical region revealed no pathogenic variants shared between these two families.

4.2 Introduction

4.2.1 *What is Microphthalmia, and how is it caused?*

Microphthalmia is a condition which exists as part of a spectrum of disorders involving the formation of the globe of the eye (a detailed description of embryological globe formation can be found in Chapter 1 of this thesis). This spectrum includes anophthalmia (clinically defined as the lack of eyes) and coloboma, caused by improper closure of the choroidal fissure.¹⁹⁷ Combined, these malformations have a global incidence of 2/10,000 live births, can affect one (unilateral) or both eyes (bilateral), and 1/3 of individuals with these malformations have extraocular features.¹⁹⁸ During embryogenesis, the formation of the eye is mediated by five broad embryological steps: 1) Optic vesicle outgrowth, 2) Optic lens/cup formation 3) Optic fissure closure 4) anterior and posterior chamber formation, and 5) and functional maturation.¹⁹⁹ Interruption of any of these processes can lead to the many malformations involving the anterior chamber or to the globe as a whole. These processes are regulated by a number of important gene-gene and tissue-tissue interactions, controlled by important transcription factors expressed during embryogenesis.

4.2.2 The Genetics of Microphthalmia

The spectrum of anophthalmia and microphthalmia is caused by a wide range of transcription factors involved in the development of many components of the eye, nervous system, and other embryological structures. These genes include, but are not limited to *OTX2*, *PAX6*, *SOX2*, *RAX*, *VSX2*, *BCOR*, *CHD7*, *BMP4*, *FOXE3*, *STRA6*, *SMOC1*, *MITF* and *SIX6*.

(i) *OTX2* (OMIM 600037)

The *Orthodenticle homeobox 2* gene (*OTX2*) is a five exon gene located on 14q21-22 coding for a 297 amino acid protein. This transcription factor is a member of the bicoid sub-family of homeodomain containing proteins, and is thought to play a role in the formation of various sensory organs, and brain structures. Mutation of *OTX2* has been associated with both anophthalmia and microphthalmia in an autosomal dominant form, particularly in syndromic microphthalmia 5 (MCOPS5; OMIM #610125) which consists of various ocular phenotypes such as microphthalmia, coloboma, microcornea, or cataracts, as well as extraocular features such as developmental delay, joint laxity, and seizures. This condition has shown examples of both variable expression and incomplete penetrance in the literature.¹⁹⁷

(ii) *SOX2* (OMIM 184429)

The *SRY-BOX2* (*SOX2*) gene is quite possibly the most clinically relevant gene to studies of microphthalmia, being mutated in 10-20% of patients with bilateral defects.¹⁹⁹ This single exon gene located on 3q26.3–q27 encodes a 317 amino acid gene that works with a large number of other transcription factors such as *PAX6* and *OTX2* due to its low binding affinity to DNA via a C-terminal transactivation domain.¹⁹⁷ This very important gene is expressed in embryonic stem cells, and is responsible for early organogenesis in the central nervous system,¹⁹⁷ and in ocular tissues. In *Xenopus* (frog) studies, it has been shown to be involved in cellular proliferation as well as the differentiation of cells into non-neuronal cell types in the eye.²⁰⁰ Mutation of this gene is associated with syndromic Microphthalmia 3 (MCOPS3; OMIM #206900), which includes many extraocular features (micropenis, cryptorchidism, motor disabilities, neurocognitive delays, febrile seizures etc.) as well as microphthalmia/anophthalmia or coloboma.

(iii) *PAX6* (OMIM 607108)

For information pertaining to *PAX6*, please refer to section 2.2.2 of this thesis.

(iv) *RAX* (OMIM 601881)

The *Retina and Anterior Neural Fold Homeobox (RAX)* is a three exon gene located on 18q21.32 which encodes a 346 amino acid transcription factor. This transcription factor was originally isolated from a *Xenopus* cDNA library and is thought to be responsible for the establishment and proliferation of retinal progenitor stem cells.²⁰¹ *RAX* contains a conserved sequence upstream of its promotor region which has been shown to be a binding site for OTX2, and SOX2 which act synergistically to upregulate *RAX* expression.²⁰² Through sequencing of patients with sclerocornea and anophthalmia in 2004, this gene was also shown to be involved in the pathogenesis of the A/M spectrum. Two missense mutations in *RAX* occurred in the DNA binding domains.²⁰³

(v) *VSX2* (AKA *CHX10* OMIM 142993)

The *Visual System Homeobox Gene 2 VSX2* (AKA *CHX10*) was originally mapped to chromosome 14 in 1989. This gene produces a transcription factor of the homeobox family of factors, which is expressed in the developing retina²⁰⁴ and is generally thought to be a transcriptional repressor.²⁰⁵ Mutations of this gene have been associated with autosomal recessive anophthalmia/microphthalmia with and without ocular coloboma in humans,²⁰⁵ mice,²⁰⁶ and zebrafish.²⁰⁷

(vi) *BCOR* (OMIM 300485)

BCL6 Corepressor (BCOR), located on Xp11.4, was cloned in 1990²⁰⁸ is known to repress transcription of *BCL6* (OMIM 109565),²⁰⁹ a transcriptional repressor involved in formation of the germinal center in lymph nodes. Mutations of this gene have been associated with oculofaciocardiodental (OFCD) syndrome, a condition which includes microphthalmia/anophthalmia, anterior segment dysgenesis, with malformations of the facial structures (long philtrum, broad nasal tip, short nose etc), heart conditions (atrial and ventricular septal defect), and dental malformations.²¹⁰

(vii) *CHD7* (OMIM 608892)

Chromodomain Helicase DNA-binding protein 7 (CHD7), located on 8q12.1 - 8q12.2, is a transcriptional regulator which binds elements in the nucleoplasm and is involved in the formation of ribosomal RNA.²¹¹ Most interestingly, *CHD7* has been found to be expressed in undifferentiated neuroepithelium, which is instrumental in the formation of the eye.²¹² Mutations of *CHD7* have been associated with a condition known as CHARGE syndrome (Coloboma, Heart anomaly, choanal Atresia, Retardation, Genital hypoplasia, and Ear anomalies; OMIM #214800) which can include optic disc coloboma, and/or microphthalmia.^{213; 214}

(vii) *BMP4* (OMIM *112262)

Bone Morphogenic Protein 4 (BMP4), located on chromosome 14q22.2, is an important molecule in embryological signaling and is involved in the formation and development of the mesoderm, teeth, limbs, and also in bone induction/fracture repair.²¹⁵ Consequently, mutations in *BMP4* have been associated with a large number of disorders, including fibrodysplasia ossificans progressive,²¹⁶ SHORT (Stature, Hyperextensibility of joints, Ocular depression, Rieger anomaly, Teething anomaly),²¹⁷ and Orofacial cleft.^{215; 217}

(ix) *FOXE3* (OMIM *601094)

For information pertaining to *FOXE3*, please refer to section 2.2.2 of this thesis.

(x) *STRA6* (OMIM *610745)

Stimulated by Retinoic Acid 6 is a 20 exon gene located on 15q24.1 which produces a membrane receptor for *Retinol Binding Protein 1 (RBP1)* and mediates cellular uptake of vitamin A.²¹⁸ This gene is thought to act in a time-dependent manner, but is also upregulated by increasing concentrations of retinoic acid in the cell.²¹⁹ Mutations in this gene have been associated with a syndromic form of microphthalmia, syndromic microphthalmia-9, which is often fatal at birth (or soon after) and includes respiratory malformations (unilobar lungs, pulmonary hypoplasia), cardiac malformations (atrial septal defect), and other organ system malformations (malrotation of the kidney).

(xi) *SMOC1* (OMIM *606488)

Sparc-related Modular Calcium Binding Protein 1 was identified in 2002 through screening of a human fetal brain cDNA library.²²⁰ Mutations of *SMOC1* have been associated with Waardenburg Anophthalmia syndrome, a condition characterized by ocular malformations, as well as limb anomalies.²²¹ These findings suggest that *SMOC1* has a role in both eye and limb growth and formation, though the exact role that it plays is currently unknown.

(xii) *MITF* (OMIM *156845)

Microphthalmia Associated Transcription Factor belongs to a class of transcription factors known as basic helix-loop-helix (bHLH) leucine zipper proteins. This class of transcription factors is involved in proper cellular formation, such as in the neural crest, which are important in ocular formation.²²² *MITF* has been an extensively studied gene, and is known to be involved in the formation of melanocytes which are important in the pigmentation of hair, skin, and eyes, and hearing function in the cochlea.²²³ The mouse *MITF* homologue, *mi*, has been shown to cause reduced eye size, and early onset deafness through changes in pigmentation in both ocular and ear structures.²²² Mutations in *MITF* have been shown to cause Waardenburg Syndrome type II (OMIM 193510) which, much like the mouse model, shows early onset hearing loss, with microphthalmia.

(xiii) *SIX6* (OMIM 606326)

The *Sine Oculis Homeobox 6* (*SIX6*) belongs to the SIX family of genes which are homologous to the 'sine oculis' (so) genes found in *Drosophila*. These genes produce a protein located in the nucleus which is required for development of the eye. Deletions of areas surrounding *SIX6*, on chromosome 14q23 have been observed in patients with bilateral clinical anophthalmia, and pituitary anomalies,²²⁴ though very few mutations in the *SIX6* gene itself have been found (one mutation thought to be pathogenic is p.T165A).²²⁵

4.2.3 Hallermann-Streiff Syndrome

Hallermann-Streiff syndrome (HSS; Francois Dyscephalic Syndrome) was a condition originally described in 1893 by Aubry, though was redescribed by Hallerman in 1948 and Streiff in 1950, which expanded the phenotype and separated it from other similar disorders (progeria, and mandibulofacial dysostosis).²²⁶ This condition is a very rare syndrome characterized by dyscephaly (craniofacial malformations), hypotrichosis (thinning of the hair on the head and pubic regions), microphthalmia, cataracts, beaked nose, micrognathia (small jaw), and proportionate short stature (dwarfism).²²⁶ Other manifestations do exist for this condition, though are not necessarily seen in all HSS patients, such as cardiovascular abnormalities,²²⁷ progeria,²²⁸ dental abnormalities,²²⁹ and respiratory problems.²³⁰ HSS is typically thought to be a sporadic condition,²²⁶ though cases of HSS-related syndromes, or atypical HSS inherited in families have been described, though the mode of inheritance of HSS is still not agreed upon. Recently, a condition with some overlapping features of HSS, called Oculodentodigital dysplasia (ODDD OMIM: 164200) was shown to be caused by mutations in the gap junction protein, CX43.²³¹ The gene for this connexin protein, *GJA1*, was analyzed in a patient with an overlapping phenotype of ODDD and HSS, and mutations were found at a highly conserved Arginine residue. A second

patient was analyzed for mutations in *GJAI*, who had a full blown HSS phenotype, though no mutations were found.²³²

4.2.4 Current Study

This thesis chapter will examine a never before reported form of syndromic Microphthalmia which is herein referred to as Microphthalmia-Dwarfism (MDW). MDW was noted in two families from the South coast of Newfoundland in an apparent autosomal recessive fashion. Patients from these two families had microphthalmia, proportionate short stature, hypotrichosis, frontal bossing of the forehead, as well as a pointed nose due to the lack of fat deposits. This condition interestingly shows overlap with a similar condition called Hallermann-Streiff syndrome (HSS; OMIM 234100), a condition characterized by all the aforementioned features but can also have micrognathia (small jaw), dental abnormalities, heart defects, and premature aging (progeria). HSS has previously been associated with *GJAI* in one patient.²³¹

4.3 Materials and Methods

4.3.1 Patients and Clinical Observations

We obtained DNA from 29 individuals as per section 2.3.3 on page 66 (five affected, 24 unaffected) from two families who reside on the southern coast of the island of Newfoundland, Canada. We were also able to obtain detailed medical records on the five affected individuals who contributed DNA, and two additional affected members from family 0066 who were deceased at time of sample collection.

4.3.2 SNP Genotyping

Twenty-nine DNA samples (including DNA from all five affected individuals) were outsourced to the Genome Center at McGill University, Montreal, Quebec to undergo SNP genotyping on a 610K Illumina SNP array.

4.3.3 Homozygosity Haplotyping

The SNP data was then analyzed using the homozygosity haplotyping (HH) method as described by Miyazawa *et al.* (2007).²³³ This is a non-parametric method where all SNPs with a heterozygous allele are discarded from the dataset and only those SNPs with a homozygous genotype are kept. This allows for a 'Region of Common Homozygous Haplotype' (RCHH) to be constructed which, due to the lack of heterozygosity, is already in phase. This method is most useful for distantly related affected relatives, as siblings would share a large amount of homozygous haplotypes (HHs) through inheritance of the same parental material. Distantly related affected individuals are more likely to share common HHs without interference from sharing the same parental DNA.

4.3.4 Two-point and Multipoint Linkage Analysis

The SNP genotype data from the Illumina chip was given to biostatisticians at The Centre for Applied Genomics (TCAG) in Toronto, Canada. A number of techniques were used to determine LOD scores for these families. A model of autosomal recessive inheritance and a disease allele frequency of 0.001 were used for these analyses.

(i) Theoretical Maximum LOD Scores

The statisticians at TCAG generated a theoretical maximum LOD score for both families. This was performed using a simulated marker which segregates with the autosomal recessive disease phenotype, and a LOD score is calculated from this simulation. This score indicates the maximum possible score given perfect conditions and can be used as a method to evaluate the power of a pedigree before undertaking linkage analyses.

(ii) Multipoint Linkage Analyses

Multipoint linkage analysis was carried out on both families using a program called Merlin. This analysis required pruning of the pedigrees, which removed extraneous members from the pedigrees in order to maximize effectiveness and reduce time for analysis. The pedigree for Family 1499 was reduced to the 10 individuals from which we obtained DNA, across four generations. The pedigree for Family 0066 was reduced to those with samples, and necessary individuals to connect them. Relationships between pedigrees were also analyzed.

(iii) Full pedigree analyses

In order to perform multi-point linkage analyses using the entire pedigrees, a program called SimWalk2 was used. SimWalk2 does not evaluate the LOD score exactly at the SNP positions, but instead provides scores at positions between the markers. Results from this analysis were summarized by taking the average of the LOD scores on the left and right of each SNP. The set of SNPs used was pruned for quality, minor allele frequency, and linkage disequilibrium, and then every third marker was selected (for speed).

The statistician used an approximate parametric analysis using Markov chain Monte Carlo as implemented in SimWalk2. Parameters included an autosomal recessive model, fully penetrant, with a disease allele frequency of 0.0001. In order to reduce the amount of time for analysis using large pedigrees, a total of 6617 SNPs were analyzed on the autosomes (the X chromosome was not analyzed).

4.3.5 Sanger Sequencing of Positional Candidate Genes

Seven samples (five affected, and two unaffected) were selected for screening of positional candidate genes (Fam 0066: IV-4, IV-13, and V-2 Fam 1499: III-2, IV-1, and IV-2). We used the smaller 8cM region as determined by the HH analysis to select 23 positional and functional candidate genes for initial Sanger sequencing (Table 4.2). Primers were custom designed for all exons of *ARL2BP*, *BBS2*, *C16ORF57*, *MMP15*, and *HERPUD1* (Appendix L) using Primer3 Software (<http://frodo.wi.mit.edu/>). All primers were checked for specificity using the UCSC *in silico* PCR tool from the UCSC Genome Bioinformatics Database (<http://genome.ucsc.edu/>). PCR was set up as per Appendix A1 using a Touchdown 54 program (Appendix B) and size fragmented on a 1% agarose gel stained with SYBR safe for confirmation of amplification. Bidirectional Sequencing of all amplicons was carried out using the protocol

outlined in Section 2.3.3 of this thesis and was run on an ABI 3130xl or ABI 3730. The other 18 genes were sequenced at the Genome Center in McGill University on an ABI 3730.

4.3.6 Variant Analysis

All sequences were manually inspected using Mutation Surveyor 4.0.4 (Transition Technologies) and were double checked by two individuals (LD and DG). All identified variants were checked in dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>) and minor allele frequencies (MAF) were recorded from dbSNP and the 1000 Genome Project (www.1000genomes.org). Variants with no rs numbers from dbSNP, or with low MAFs, and which could have a potential effect on protein function (i.e. nonsense mutation, missense mutation, splice variation) were checked for segregation on families 0066 and 1499. *In silico* analyses of each missense mutation and splice variation were carried out using SIFT (Sorting Intolerant From Tolerant),²³⁴ PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>), and PANTHER (Protein ANalysis THrough Evolutionary Relationships).¹⁸⁰

4.4 Results

4.4.1 Patients and Clinical Observation

Seven affected individuals from both families (five from 0066 and two from 1499) (Figures 4.1-4.4) underwent clinical and ophthalmological assessment (Table 4.1). All the affected individuals were consistent in the diagnoses of microphthalmia, short stature, cataracts and other forms of anterior segment dysgenesis (glaucoma, corneal opacities/clouding), and the older affected individuals showed hypotrichosis.

(i) 0066 V-4

This female member of family 0066 was born in 1945 and is currently still alive. In 1971, she was described as having NLP in both eyes, microcornea, microphthalmia, severe corneal opacities, bilateral glaucoma, and later had a bilateral enucleation and ocular implants placed. In 2005, she was noted to be 119cm tall (<<3rd %ile, 50th % for a 6.5 year old), and weighed 36.1 kg (<3rd %ile; 50th %ile for 6.5 year old). Pictures from when V-4 was younger showed that she had a full head of hair, but now at 60 years of age shows marked hypotrichosis, suggesting that this portion of the MDW phenotype may be later onset. Currently, she requires 2 liters of oxygen per day due to severe Chronic Obstructive Pulmonary Disease (COPD), and has generalized osteopenia with scoliosis. She is of normal intelligence, and communicates well though she has mild to moderate hearing loss related to frequent infections. A cytogenetic analysis was carried out on a sample from V-4, showing no major chromosomal abnormalities.

Symbol definitions	
□	○ Unaffected
■	● Microphthalmia-dwarfism
◻	◉ Carrier
◻?	◉? Possibly affected
◻	◉ other eye disease

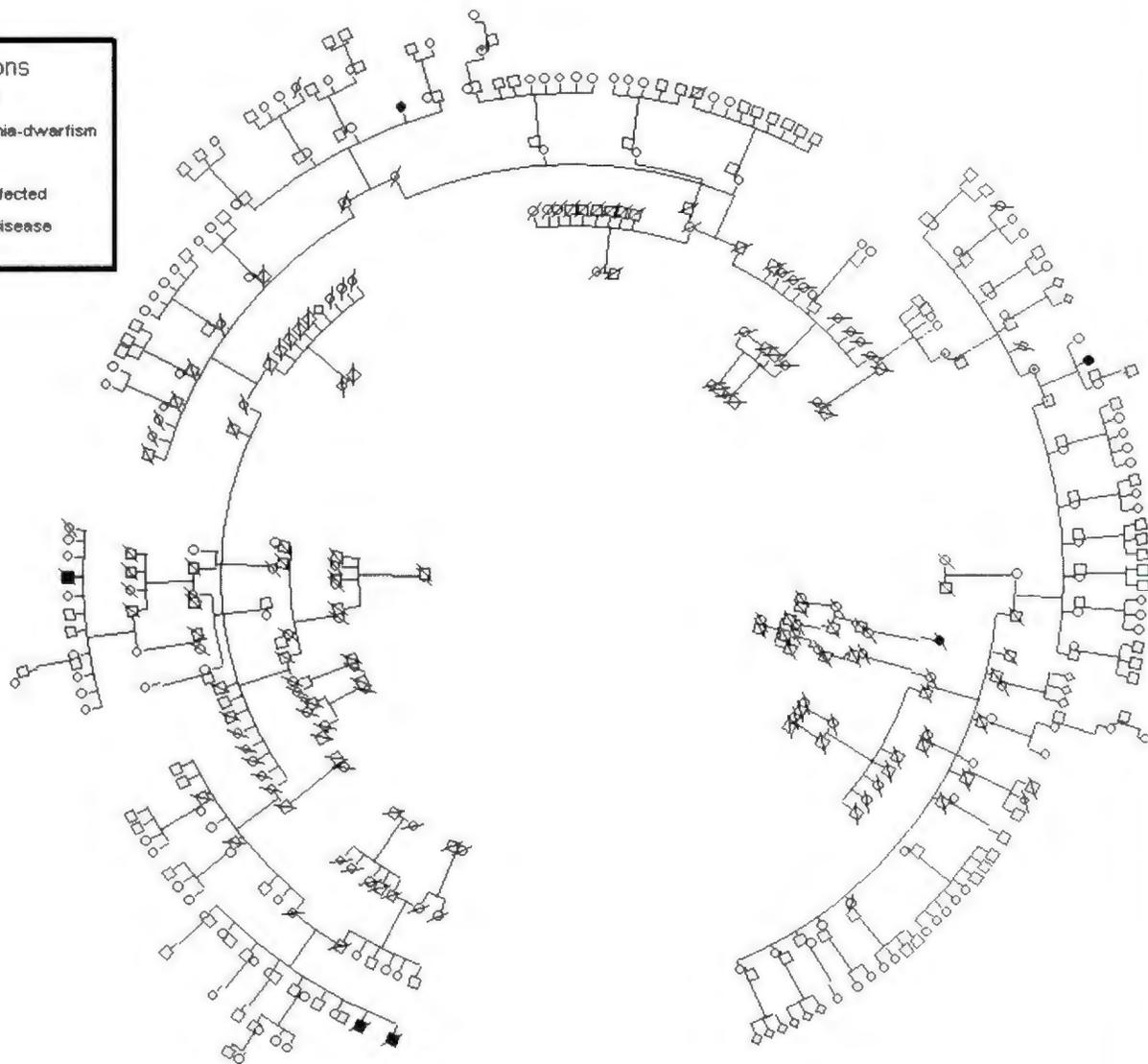


Figure 4.1 – Circularized pedigree of Family 0066 with microphthalmia-dwarfism from Newfoundland. All the patients have unaffected parents, consistent with an autosomal recessive mode of inheritance.

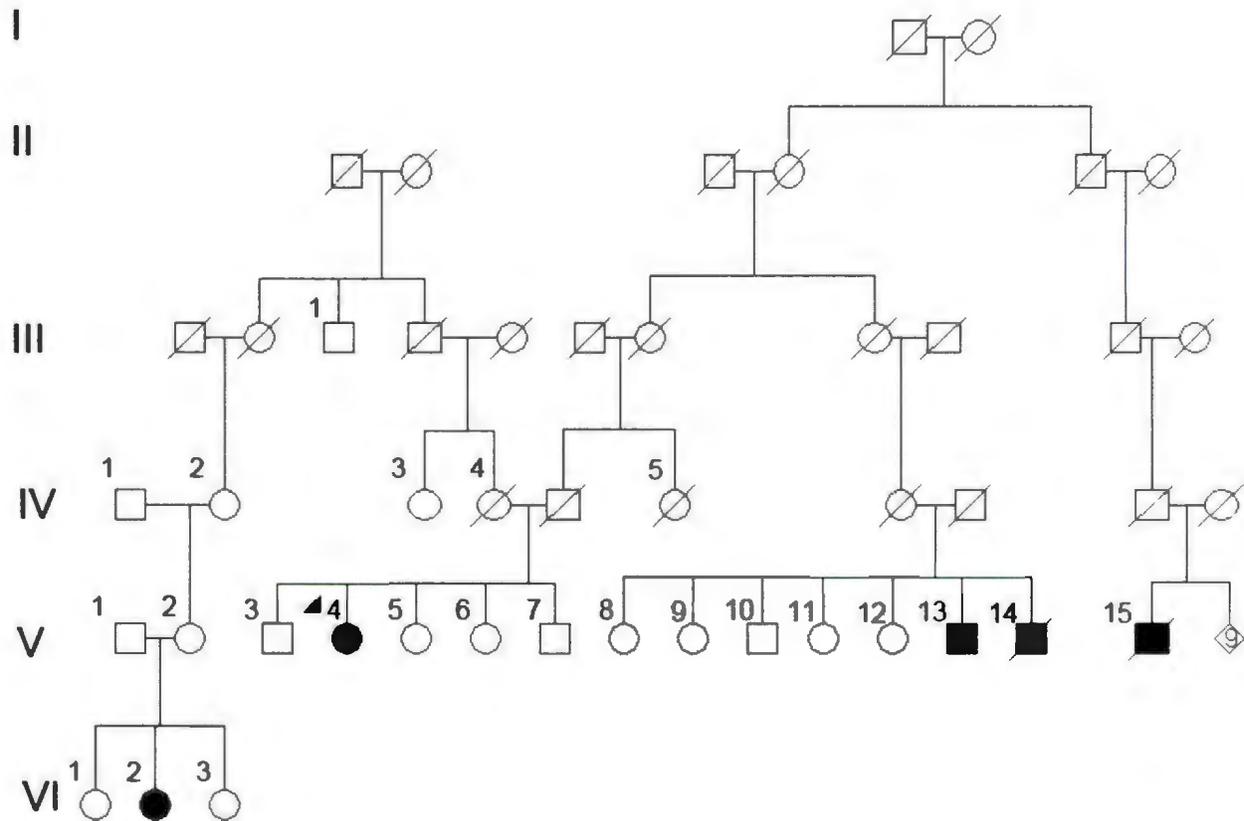


Figure 4.2 – Condensed pedigree of Family 0066 showing the four affected individuals (shaded) across three sibships. Individuals with numbers to the upper left of their symbol contributed DNA to this research project. We were able to obtain DNA and information on family members over four generations. Individuals who contributed DNA to this project are denoted by a number to the upper left of their symbol (except for V-15 and V-14, for whom medical records were available, but no DNA).

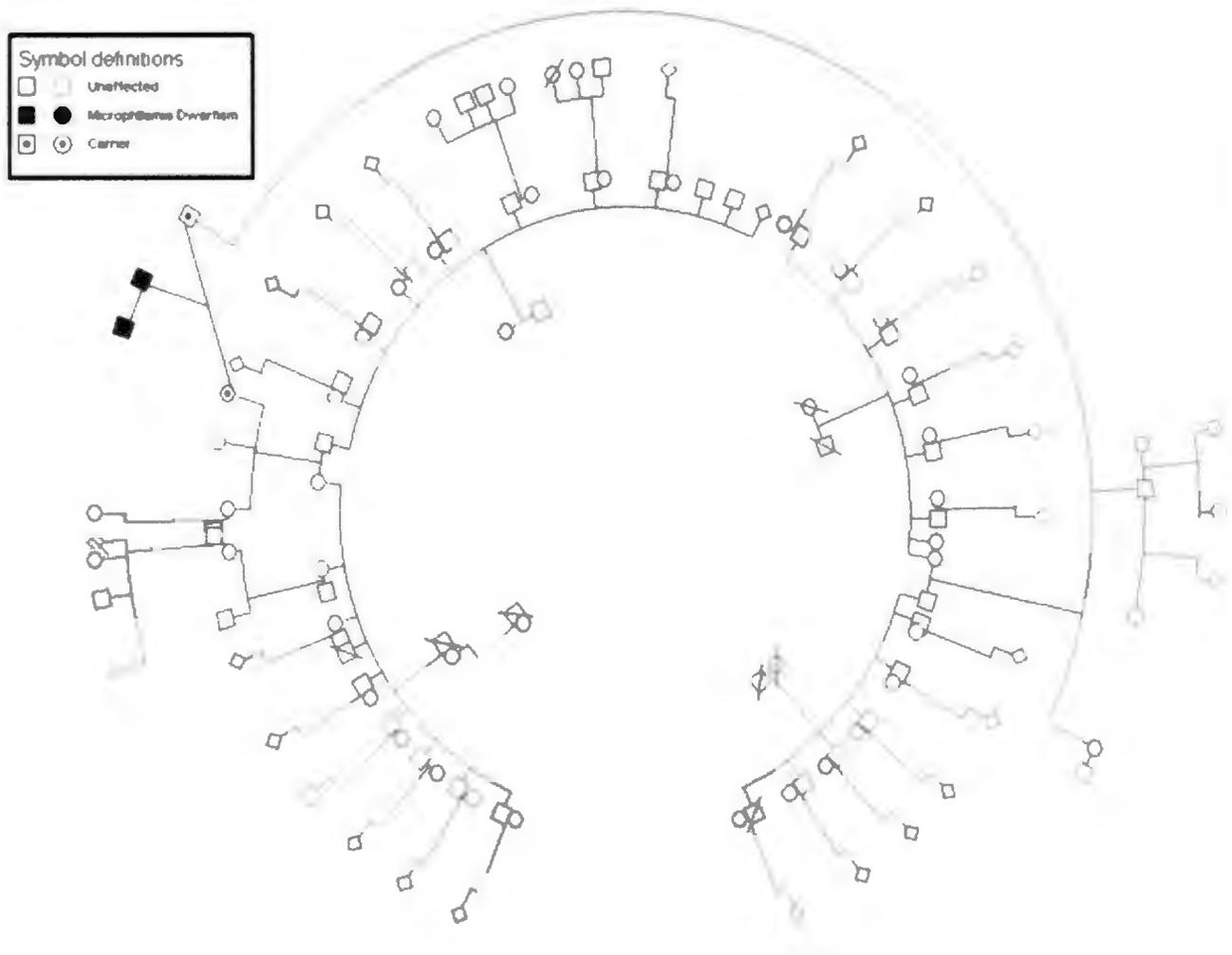


Figure 4.3 – Circularized pedigree of Family 1499. This pedigree illustrates a large number of individuals with only two affected brothers.

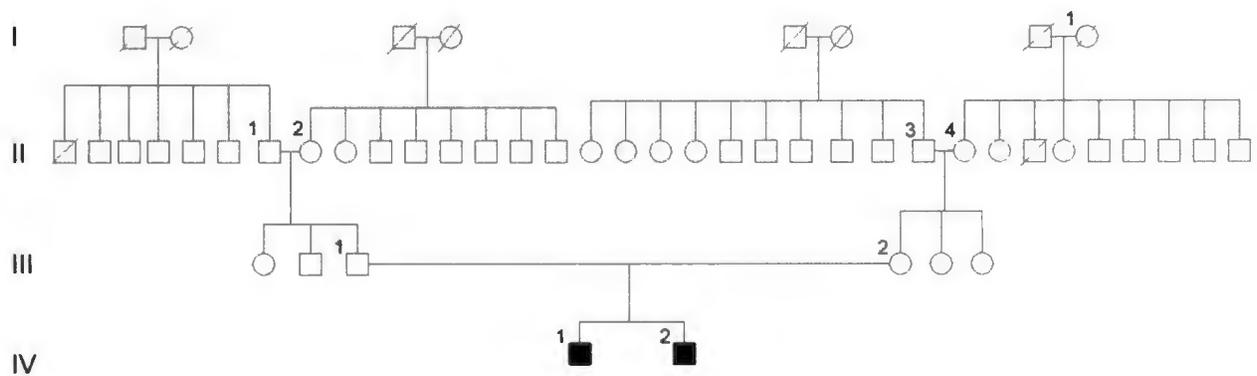


Figure 4.4 – Condensed pedigree of Family 1499 highlighting the one affected generation from the circular pedigree. Individuals with numbers to the top left of their symbol have contributed DNA to this research project, thus giving potential information from four generations of individuals.

	0066 V-4	0066 V-13	0066 V-14	0066 V-15	0066 VI-2	1499 V-1	1499 V-2
YOB	1945	1940	1931	1949	1978	1998	1993
Gender	Female	Male	Male	Male	Female	Male	Male
Microphthalmia	+	+	+	+	+	+	+
Cataracts	+	+	+	+	+	+	+
Anterior Segment Dysgenesis (Corneal Opacities/Clouding, Glaucoma)	+	+	+	+	+	+	+
Short Palpebral Fissures	+	+	+	+	+	+	+
Nystagmus	NA	+	+	+	+	+	+
Proportionate Short Stature	+	+	+	+	+	+	+
Brachycephaly	+	+	+	+	+	+	+
Beaked/Pointed Nose	+	+	+	+	+	+	+
Micrognathia	-	-	-	-	-	-	-
Frontal Bossing	+	+	+	+	+	+	+
Skeletal Malformations (Kyphosis, Scoliosis, etc)	+	+	+	-	-	-	-
Hypogonadism/Irregular Periods	+	+	+	+	+	+	+
Hypotrichosis	+	+	+	+	-	-	-
Skin Atrophy	-	+	+	+	-	-	-
Respiratory Issues (COPD, Bronchitis)	+	-	+	+	-	-	-
Dental Abnormalities*	NA	NA	NA	NA	-	+	-
Hearing Loss	+	+	+	-	-	+	+

NA: Not available. *: Most patients had dentures at time of observation.

Table 4.1 – Clinical information of the affected members from Families 0066 and 1499. Conditions noted with a + means the individual showed this phenotype, whereas a – symbol means that they did not. Listed conditions include the clinical features associated with Hallermann-Streif Syndrome including dyscephaly, hypotrichosis, microphthalmia, cataracts, beaked nose, micrognathia, and proportionate short stature. All affected individuals have microphthalmia, short stature, beaked nose, frontal bossing, short palpebral fissures, and an anterior segment phenotype (including cataracts, nystagmus, and/or corneal opacities). Only the older individuals exhibit Hypotrichosis, respiratory issues, and frequent infections.

(ii) 0066 V-13

This male member of family 0066 was born in 1940 and passed away in 1994. Clinical examinations showed frontal bossing of the forehead, hypotrichosis of both the head and pubic region, and microphthalmia. Ophthalmological examination in the late 70s showed no light perception (NLP) in the left eye, indicating total blindness, and light perception (LP) in the right eye. He was noted to have microphthalmia, severe optic nerve damage due to glaucoma, cataracts in the right eye, and horizontal/vertical nystagmus bilaterally. He was measured to be 127cm tall with a head circumference less than the second percentile at age 49. Further examination of medical records revealed chronic bronchitis and COPD along with skeletal dysplasia, hypogonadism, atrophic skin, and 50% hearing loss due to frequent ear infections. At the time of death, he was described 71kg in weight, though limbs were noted to be of proportionate size, consistent with proportionate short stature. He was also noted to suffer from Diabetes Mellitus, and Diabetes Insipidus indicating further endocrine system problems.

(iii) 0066 V-14

This male member of family 0066 was an older brother to 0066 IV-13, being born in 1931 and passing away in 1981. Unfortunately, we were unable to collect DNA from this individual as he had died before we were able to obtain a sample, though we were able to obtain detailed medical records. In 1989, it was noted that he had only light perception in his right eye, and no light perception (NLP) in his left eye. He had been diagnosed with glaucoma and microcornea at age 14 V-14, like his younger brother, was described as having COPD and chronic bronchitis, which contributed to his death. He was also noted to have hypotrichosis of the head and pubic region, dry atrophic skin, hypogonadism, severe osteoporosis, kyphosis, scoliosis, and hearing loss due to frequent infection. A small mandible was also observed, but no

mention of micrognathia was seen in the reports. At the time of death, he was 150 cm tall and 75 kg in weight.

(iv) 0066 V-15

This male member of family 0066 was born in 1949 and died in 1988. Ophthalmological examination showed that this individual had bilateral cataracts, nystagmus, microphthalmia, and NLP in both eyes. An autopsy report described V-15 as being 150 cm in height and 60-70 kg in weight. DNA was unavailable from this individual.

(v) 0066 VI-2

This female member of family 0066 was born in 1978, and is currently living. Ophthalmological investigations showed bilateral microphthalmia, short palpebral fissures, and bilateral cataracts. She had no vision in her right eye, with LP and colour discrimination in her left eye. At the age of 27, she was noted to be 138.5 cm tall (<3rd %ile), and weighed 39.5 kg (<5th percentile) with a head circumference of 57 cm (>97th %ile). She had no hypotrichosis, no skin atrophy, and no hearing loss or dental/jaw abnormalities. She is of normal intelligence and currently holds a Bachelor degree in Psychology. A karyotype was carried out on a sample from this individual which showed a normal chromosome count.

(vi) 1499 IV-1

One of the two affected brothers of Family 1499, IV-1, was born in 1993 and is currently still alive. At birth, he was noticed to have horizontal pendular nystagmus, shallow anterior angles, bilateral corneal opacities, macular hypoplasia, and correctopia (displacement of the pupil). His birth weight (7lbs, 6oz) and length were normal, but at five years of age he was 96.52 cm tall (3rd %ile), weighed 34.5 kg (3rd %ile), and had a head circumference of 52.5 cm (50th %ile). He was of normal intelligence, and is currently attending school using assistive technologies because of his poor vision. He has had normal pubescent growth with facial and pubic hair growth (though has hypogonadism) and has not shown any hypotrichosis. Some dental abnormalities were noticed as he had thin enamel and soft teeth. Human growth hormone tests were carried out to determine if this was the cause of his short stature, though these tests came back normal. At 16 years, he was 151.5 cm tall (<3rd %ile), and weighed 87.9 kg (90-97th %ile). He still did not exhibit any hypotrichosis at this age.

(vii) 1499 IV-2

The younger brother of 1499 IV-1, 1499 IV-2, was born in 1998 with normal birth weight (7lbs, 4oz). In 2010, he was 134 cm tall (3rd %ile), weighed 33.2 kg (25-50%ile) and had a head circumference of 55 cm (+1 SD). He was severely photophobic, and had microphthalmia, short palpebral fissures, and had multiple surgeries to remove cataracts and to alleviate symptoms of glaucoma. His visual acuity at 9 years of age was 20/800 in his right eye, and NLP in his left eye. His nose was described as thin, delicate, and pointed with a short philtrum which was thin but well developed. There were no observed dental malformations, and he had normal dentition. He was also noted to have hypogonadism. Like his brother, he was not noted to have hypotrichosis, or any respiratory issues, and was of normal intelligence.

4.4.2 Homozygosity Haplotyping (HH) Analysis

The genotype data from the Illumina 610K SNP array underwent HH analysis by bioinformaticians at the McGill Genome Center in Montreal, Quebec, Canada. By using a theoretical cutoff of >3 centimorgans we identified a homozygous haplotype shared across all five affected individuals, on chromosome 16q12.2-21 (Figure 4.5). This region ranges from Chr16:55661080-62017481 and contains 136 genes according to the BIOMART tool from the Ensembl Genome Browser (<http://useast.ensembl.org/index.html>).



Figure 4.5 - HH analysis from the SNP genotyping of Families 0066 and 1499. A cutoff of >3cM reveals two regions of Region of Common Homozygous Haplotype (RCHH). The region on chromosome 16 spans 8.5cM and is bordered by rs289710 and rs219568. A second region can also be noted on chromosome 3 bordered by rs11922267 and rs4602330 which spans 2.3 cM.

4.4.3 Linkage Analyses

(i) Theoretical Maximum LOD score

Theoretical LOD scores were calculated at the Toronto Center for Applied Genetics (Nicole Roslin) for both family 1499 and 0066 to gauge the power of both pedigrees for linkage analyses. Family 1499 was a straight forward analysis containing four generations of people all linked through 10 DNA samples. The theoretical maximum LOD score for this family was 0.602. The second pedigree, 0066, was more complicated due to the size of the pedigree, and the number of individuals between family members which had no genotypes. There are three individuals in family 0066 with DNA who are known to be affected. Analysis also suggested a very distant relationship between 1499 II-1 and 0066 IV-13, though the exact relationship was unclear. There were no inbreeding loops known, though it is likely that several of the founders are distantly related due to the geographic location on the sparsely populated South coast of Newfoundland.

Under a fully penetrant autosomal recessive model, a disease causing allele must be entering Family 0066 at least four times, complicating this analysis. To alleviate the conditions here FastLINK 3.00 was used to simulate a completely informative multi-allelic marker which was linked to the disease locus. Each replicate was analyzed using Superlink 1.7 under the same model as the generating model. The disease causing allele was assumed to have a frequency of 0.0001. To avoid the ambiguity of having un-genotyped individuals in the pedigree, everyone was assumed to have genotypes. This method provided a maximum LOD score for Family 0066 of 3.957 after 1000 simulations. Since LOD scores can be additive when a disease is believed to be caused by the same locus in two or more families, the total LOD score for families 0066 and 1499 was calculated to be 4.559

(ii) Multipoint Linkage Analyses

Multipoint parametric analysis was carried out using Merlin. Unfortunately, this places computational limits on the analysis due to the size of the pedigrees. Both pedigrees were pruned and analysis was run using an autosomal recessive model with full penetrance and a disease allele frequency of 0.0001. This analysis revealed a LOD score of 4.539 ($\theta = 0$) at SNP rs1875233 on chromosome 16q12.2, only 0.020 away from the theoretical maximum LOD. The next highest LOD score (0.403) was seen on chromosome 18 at rs4594329 (Figure 4.6)

SimWalk2 does not evaluate the LOD score exactly at the SNP positions, but instead provides scores at positions between the markers. These results were summarized by taking the average of the LOD scores on the left and right of each SNP. The set of SNPs used here was after pruning for minor allele frequency and linkage disequilibrium, and then every third marker was selected (for speed).

Since SimWalk2 does not provide LOD scores at each SNP, obligate recombinants cannot be observed. Therefore, LOD scores were calculated at each SNP using a singlepoint analysis. This used an exact procedure, and the complete pedigrees were used. However, since it was a singlepoint analysis, the LOD scores were quite a bit smaller. The maximum LOD score was ~2.2. However, the obligate recombinants are noted with a LOD score of -1000, and so the linked region is bounded by rs2024477 and rs4887996. This is, however, a large region of about 33 cM. The region bounded by rs4131945 to rs4387626 (about 19 cM), linkage is never excluded, meaning that the LOD score never drops below -2. The set of SNPs used here was after pruning for minor allele frequency and linkage disequilibrium.

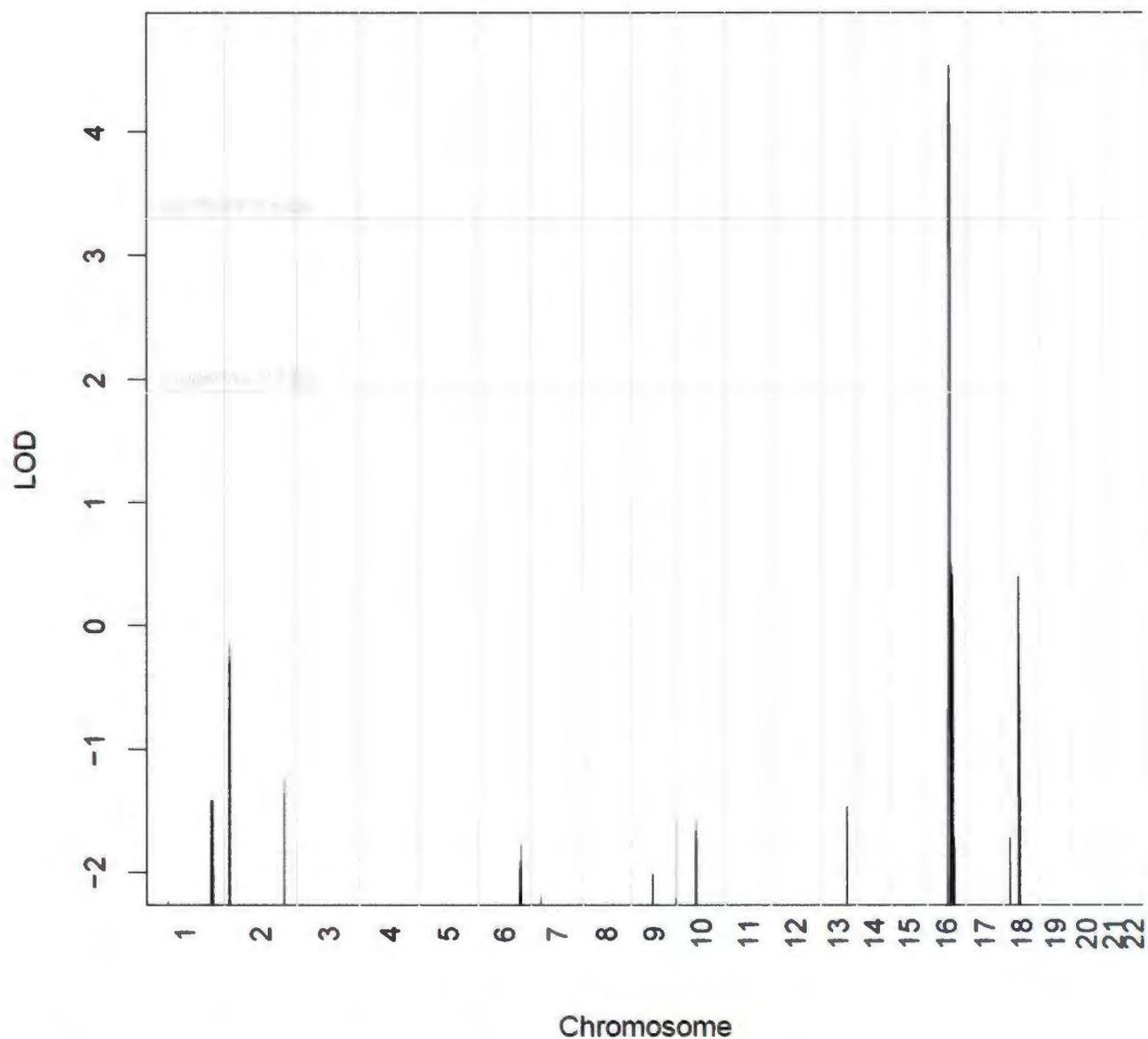


Figure 4.6 - Multipoint linkage analysis performed using Merlin and trimmed pedigrees. A total of 6627 SNPs were used to perform the analysis. A model of autosomal recessive inheritance and a disease allele frequency of 0.0001 was used. One region exhibited a maxLOD score of 4.539 ($\theta = 0$) at SNP rs1875233 on chromosome 16q12.2. The second highest LOD score from this analysis was 0.403 at rs4594329 on chromosome 18.

4.4.4 Sanger Sequencing of Positional Candidate Genes

Initially, we used the smaller region on chromosome 16q21 bound by rs289710 and rs219568 as determined by the HH study which spanned 8.5 cM and contained 136 genes. We selected 23 genes based on function, previous association with disease, and potential involvement in the MDW phenotype (Table 4.2). In addition to the coding genes in this region, primers were designed to amplify *SNORA46* and *SNORA50* along with 10 other non-coding genes including micro RNAs (miRNAs) and other snoRNAs (Appendix M) which existed in the much larger linkage region. Sequencing of these 12 additional non-coding genes revealed no variants in any of the DNA from affected individuals or from two unaffected family members (0066: III-4 1499: III-2). These genes were sequenced either at Memorial University or the Genome Center at McGill University and 98 variants were discovered upon analysis (Table 4.2). Three genes (*NDRG4*, *SETD6*, and *GOT2*) contained mutations which were of particular interest.

(i) *NDRG4* c.873+3 G>A

This mutation was of interest due to its proximity to the intron exon boundary of exon 14 within *NDRG4*. Mutations close to this boundary have the potential to disrupt proper splicing patterns by either exon skipping, or intron inclusion. *NDRG4* 873+3 G>A is seen in the 1000 Genome Project's database, but shows a Minor Allele Frequency (MAF) of 0.02% in a British and Puerto Rican population (Accessed June, 2012), and has never been seen as a homozygote according to the individual genotypes within the database. This mutation was found in one of 142 (0.7%) screened control chromosomes from the island of Newfoundland. The potential effect of this variant was tested by designing three primers, which when multiplexed, would

Table 4.2 – Table of all sequencing variants found in the 23 selected genes from the HH region. All variants are named according to HGVS guidelines. Columns include position of the mutation within the gene (intron, exon, or UTR), any effect on the protein, dbSNP ID (if available), Minor Allele Frequency (MAF) from the 1000 Genome Project, whether the variant segregates in an AR fashion with the MDW phenotype, and SIFT/PolyPhen-2 Scores for missense variants.

Gene	Variant	Gene Position	Protein Effect	dbSNP ID	MAF dbSNP	Disease Segregation	SIFT Score	PolyPhen Score
HERPUD1	c.1-146 G>A	1 (5' UTR)	NA	rs37025	26.20%	Yes	NA	NA
	c.1-86 T>G	1 (5' UTR)	NA	rs37024	41.80%	No	NA	NA
	c.554+23 G>C	Intron 5	NA	NA	NA	No	NA	NA
	c.554+64 A>G	Intron 5	NA	rs2133783	38.30%	Yes	NA	NA
CPNE2	c.180+86 A>T	Intron 2	NA	rs8053682	41.30%	No	NA	NA
	c.180+87 G>C	Intron 2	NA	rs8047463	41.13%	Yes	NA	NA
	c.1061+319 G>T	Intron 11	NA	NA	NA	No	NA	NA
	c.1302+39 C>G	Intron 14	NA	rs2305693	7.86%	Yes	NA	NA
ARL2BP	No Variants		NA	NA	NA	NA	NA	NA
CIAPIN1	c.-131 G>A	5' UTR	NA	NA	NA	NA	NA	NA
	c.154 C>G	Exon 2	Q52E	rs11557674	3.29%	No	Tolerated	Benign (0.069)
	c.157+16 C>T	Intron 2	NA	rs72780616	3.29%	No	NA	NA
POLR2C	c.137-24 A>G	Intron 2	NA	rs223839	48.95%	Yes	NA	NA
	c.174 T>C	Exon 3	V150V	rs4937	27.47%	Yes	NA	NA
	c.387+8 C>T	Intron 5	NA	rs601194	46.90%	Yes	NA	NA
	c.439+100 G>A	Intron 6	NA	rs223874	26.51%	Yes	NA	NA
CCDC102A	c.286 C>T	Exon 2	R96W	rs12935069	4.20%	Yes	Tolerated	Benign (0.00)
	c.585+82 A>G	Intron 2	NA	rs8045906	48.40%	Yes	NA	NA
	c.825 G>A	Exon 4	A275A	rs28446687	18.37%	Yes	NA	NA
GPR114	c.-7 G>A	5' UTR	NA	rs1004363	43.05%	Yes	NA	NA
	c.140+181 C>T	Intron 3	NA	rs1859259	40.22%	Yes	NA	NA
	c.140+269 G>A	Intron 3	NA	rs1859260	42.96%	Yes	NA	NA
	c.846 C>T	Exon 9	R282R	rs9937918	46.44%	Yes	NA	NA
CCDC135	c.151 G>A	Exon 2	D51N	rs55645458	1.55%	No	Tolerated	Probably Damaging (0.967)
	c.557 G>C	Exon 5	C186S	rs7196016	28.15%	Yes	Tolerated	Possibly Damaging (0.585)
	c.859-26 C>T	Intron 6	NA	rs12933056	17.78%	Yes	NA	NA
	c.1077+22 G >C	Intron 7	NA	rs2122486	17.69%	Yes	NA	NA
	c.1298 C>T	Exon 10	P433L	rs3809611	19.70%	Yes	Deleterious	Probably Damaging (0.826)
	c.2296 T>C	Exon 16	C766R	rs2923147	38.44%	Yes	Tolerated	Benign (0.00)
KIFC3	c.381+105 T>C	Intron 4	NA	rs2965789	32.31%	Yes	NA	NA
	c.699 T>C	Exon 6	L233L	rs2967165	33.82%	Yes	NA	NA
	c.765+32 G>A	Intron 6	NA	rs2965790	47.35%	Yes	NA	NA

	c.765+62 T>C	Intron 6	NA	rs2967166	36.01%	Yes	NA	NA
	c.939+72 T>A	Intron 7	NA	NA	NA	Yes	NA	NA
	c.940-44 A>G	Intron 7	NA	rs3803590	33.82%	Yes	NA	NA
	c.1803 C>T	Exon 14	D601D	rs2967172	47.30%	Yes	NA	NA
	c.2003-21 T>A	Intron 15	NA	rs2965799	44.42%	Yes	NA	NA
<i>C16ORF57</i>	c.99-36 G>A	Intron 2	NA	rs3743560	43.60%	Yes	NA	NA
	c.265+43 C>G		NA	NA	NA	No	NA	NA
	c.504-60 C>T	Intron 5	NA	rs4784022	49.68%	Yes	NA	NA
<i>CCL22</i>	c.5 A>C	Exon 1	D2A	rs4359426	7.90%	Yes	Tolerated	Benign (0.00)
<i>KATNB1</i>	c.289+65 A>G	Intron 4	NA	rs2967152	28.56%	Yes	NA	NA
	c.704+4 C>T	Intron 9	NA	rs2965797	43.28%	Yes	NA	NA
	c.726C>T	Exon 10	D242D	rs2965798	47.07%	Yes	NA	NA
<i>ZNF319</i>	c.1536 G>A	Exon 2	K512K	rs3743556	35.33%	Yes	NA	NA
	c.1749+20 A>G	Exon 2		rs3743555	35.33%	Yes	NA	NA
<i>MMP15</i>	c.213-17 C>T	Exon 3	NA	rs2241774	22.12%	No	NA	NA
<i>C16ORF80</i>	No variants	NA	NA	NA	NA	NA	NA	NA
<i>CCDC113</i>	c.628-249 G>C	Intron 5	NA	rs11076235	49.36%	Yes	NA	NA
	c.628-242 C>G	Intron 5	NA	rs11076236	49.36%	Yes	NA	NA
	c.628-117 C>T	Intron 5	NA	rs11076237	49.31%	Yes	NA	NA
<i>PRSS54</i>	c.-6-16 G>A	Intron 2	NA	rs2288013	9.37%	No	NA	NA
	c.75 T>C	Exon 3	Y25Y	rs2288012	25.96%	Yes	NA	NA
	c.263+56 A>G	Intron 4	NA	rs3803582	9.73%	No	NA	NA
	c.263+60 T>C	Intron 4	NA	rs3803583	8.55%	No	NA	NA
	c.263+61G>A	Intron 4	NA	NA	NA	No	NA	NA
	c.264-112 A>G	Intron 4	NA	rs4784916	31.17%	Yes	NA	NA
	c.544 A>G	Exon 6	S182G	rs3815803	54.6%	Yes	Tolerated	Probably Damaging (76.7)
	c.718 G>A	Exon 7	V240I	rs1052276	9.69%	No	Tolerated	Benign (0.122)
	c.883 A>G	Exon 7	T295A	rs2241414	9.69%	No	Tolerated	Benign (0.000)
<i>NDRG4</i>	c.223+90 G>A	Intron 4	NA	rs2427787	30.62%	Yes	NA	NA
	c.612+13 A>G	Intron 9	NA	rs2271948	41%	Yes	NA	NA
	c.612+77 A>G	Intron 9	NA	rs40185	100%*	Yes	NA	NA
	c.672 G>A	Exon 10	V224V	rs17821543	4.07%	Yes	NA	NA
	c.873+3 G>A	Intron 14	NA	rs142344402	0.002%	Yes	NA	NA
	c.1062 A>G	Exon 16	S354S	rs42945	21.34%	Yes	NA	NA
<i>SETD6</i>	c.193 T>C	Exon 2	L65L	rs4784046	26.14%	Yes	NA	NA
	c.263-64_263-50dup	Intron 3	NA	NA		Yes	NA	NA
	c.902-77 G>A	Intron 7	NA	rs2280397	45.98%	Yes	NA	NA
	c.1338 G>A	Exon 9	E446E	rs3607	25.50%	Yes	NA	NA
	c.*26 G>A	3' UTR	NA	rs3169293	4.16%	Yes	NA	NA
<i>CNOT1</i>	c.103-32 C>G	Intron 2	NA	rs34014750	4.16%	Yes	NA	NA
	c.1216-69 G>A	Intron 11	NA	rs72790275	25.50%	No	NA	NA
	c.1584+22 dup T	Intron 11	NA	NA	NA	Yes	NA	NA

	c.1704+110 G>A	Intron 14	NA	rs9936144	22.53%	No	NA	NA
	c.2332+99 G>A	Intron 18	NA	rs41445346	20.48%	No	NA	NA
	c.2891+43 C>T	Intron 21	NA	rs1568399	46.12%	Yes	NA	NA
	c.2919 G>A	Exon 22	Q973Q	rs11866002	46.12%	Yes	NA	NA
	c.3201+124 G>T	Intron 23	NA	rs7184114	35.29%	No	NA	NA
	c.4006+12 A>G	Intron 29	NA	rs2241570	46.16%	Yes	NA	NA
	c.4434+90 G>A	Intron 31	NA	rs41260	21.70%	Yes	NA	NA
	c.4716 C>T	Exon 34	Y1572Y	rs11540994	46.16%	Yes	NA	NA
	c.4801-83 G>A	Intron 34	NA	rs246195	22.35%	Yes	NA	NA
	c.5415-61 G>A	Intron 38	NA	rs2244453	20.70%	Yes	NA	NA
	c.5646+88 T>G	Intron 39	NA	NA	NA	No	NA	NA
	c.*82 G>A	3'UTR	NA	rs17821549	4.16%	Yes	NA	NA
<i>GIN53</i>	c.198 T>G	Exon 3	L66L	NA		Yes	NA	NA
<i>GOT2</i>	c.213 T>C	Exon 2	N71N	rs257636	32.08%	Yes	NA	NA
	c.228 T>G	Exon 2	V76V	rs14221	29.71%	Yes	NA	NA
	c.246+252 C>T	Intron 2	NA	rs257634	41.91%	Yes	NA	NA
	c.562 G>A	Exon 5	G188S	rs11076256	3.98%	No	Tolerated	NA Possibly Damaging (0.727)
	c.816 C>T	Exon 7	C272C	rs1058192	28.47%	Yes	NA	NA
	c.853+100_853+103del	Intron 7	NA	NA		Yes	NA	NA
	c.1037 T>G	Exon 9	NA	rs30842	26.19%	Yes	NA	NA
<i>SLC38A7</i>	c.883+19 delC	Intron 8	NA	NA		Yes	NA	NA
	c.1286+51 T>C	Intron 11	NA	rs9923128	34.69%	Yes	NA	NA

amplify a 232 bp fragment in cDNA, and a 675 bp fragment in gDNA (Figure 4.7). If the c.873+3 mutation caused an intron inclusion, a fragment of 350 bp would be amplified, whereas if the mutation caused exon skipping we would see no amplification, as the reverse primer exists in exon 15, which would be skipped. PCR using a touchdown thermocycler program revealed a 232 bp fragment in two affected family members, and in one unaffected control that did not have the *NDRG4* mutation, and a 675 bp in gDNA controls (Figure 4.7).

(ii) *SETD6 c.263-64_263-50dup15*

This mutation was of interest as it appeared to be a microsatellite expansion of a CCT repeat towards the middle of intron 2 and is not reported in dbSNP. It was hypothesized that this could potentially alter splicing patterns by affecting the position of the branch site towards the center of the intron. A set of primers was designed which would amplify a 1000 bp fragment in gDNA, and a 635 bp fragment in cDNA. PCR at TD54 using 5% DMSO showed a 635 bp fragment in the two affected control cDNAs and one unaffected cDNA (Figure 4.8). The cDNA samples from affected family members shown in lanes 2-4 (1499 IV-1) and 6-8 (1499 IV-2) show the same size fragment as that of the unaffected control sample who lacks this mutation in lanes 10-12 (Figure 4.8). No cell lines were available for testing from Family 0066. Interestingly, there was a second band weakly amplified in the cDNA samples between the 400-500 bp ladder standard, that did not amplify in the genomic DNA. This band corresponds to a second isoform of *SETD6* (NM_024860) where 218 bps are spliced out, leaving a 457 bp fragment. *SETD6* is not known to be involved in eye development.

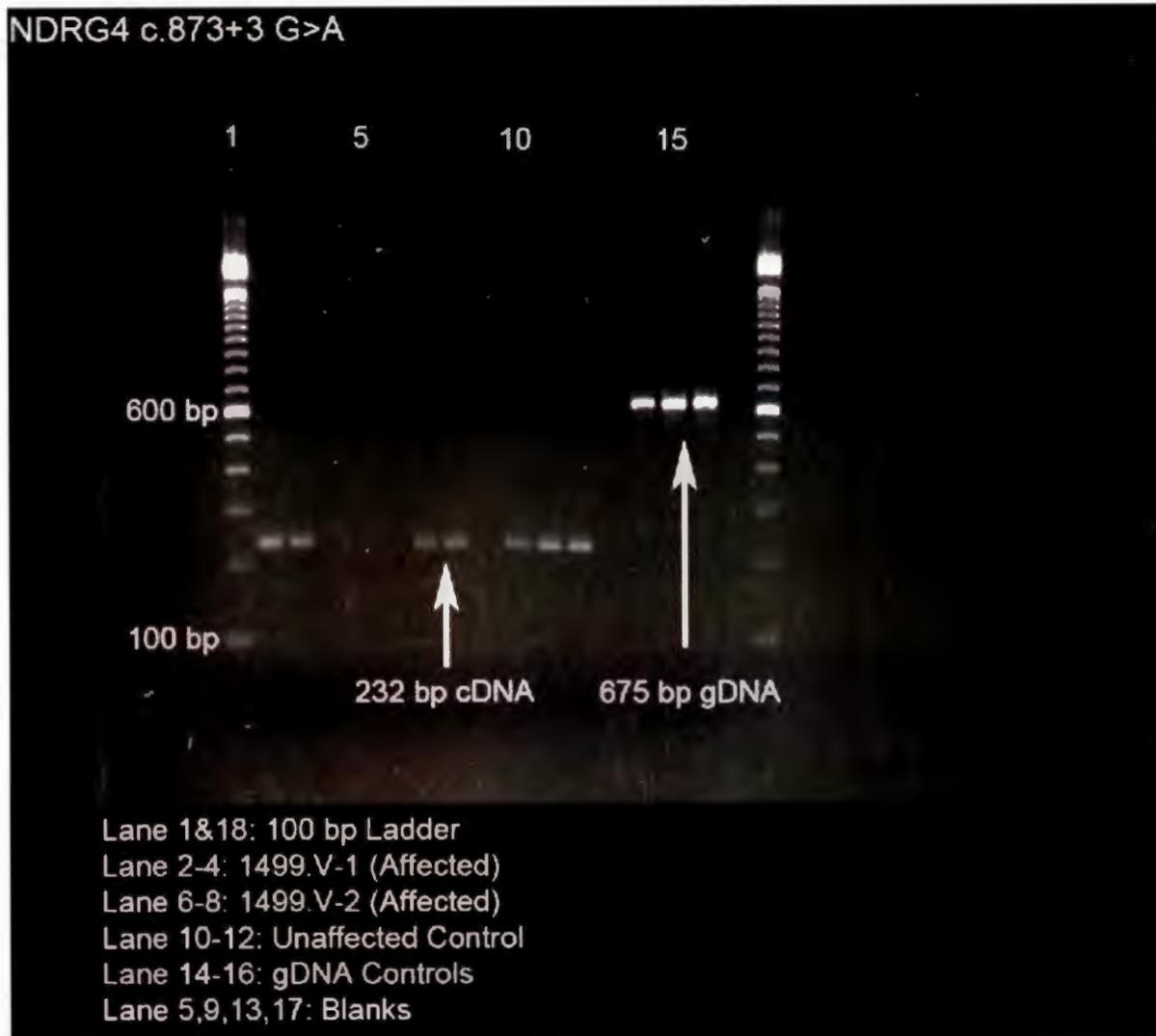


Figure 4.7 - 2% Agarose gel run at 100V for two hours. Lanes 1 and 18 contain a 100 bp ladder used to determine fragment size. Lanes 2-3 contain cDNA from a Lymphoblastoid cell line from individual IV-1 of family 1499 in triplicate. Lanes 6-8 contain cDNA from a Lymphoblastoid cell line of affected individual IV-2 from family 1499, and lanes 10-12 contain cDNA from an unaffected individual from another family with hereditary deafness (and does not have the NDRG4 mutation). All cDNA samples show a band at the expected 232bp position, but do not show the 675 bp fragment seen in gDNA (lanes 14-16), which would suggest gDNA contamination. If the c.873+3 mutation caused splice alterations the bands from the affected samples would be 350bp (intron inclusion) or would not have amplified (exon skipping) due to the position of the primers. The bands in the affected individuals coincide with that of the unaffected control sample, suggesting that the c.873+3 mutation in NDRG4 has no discernible effect on splice variation.

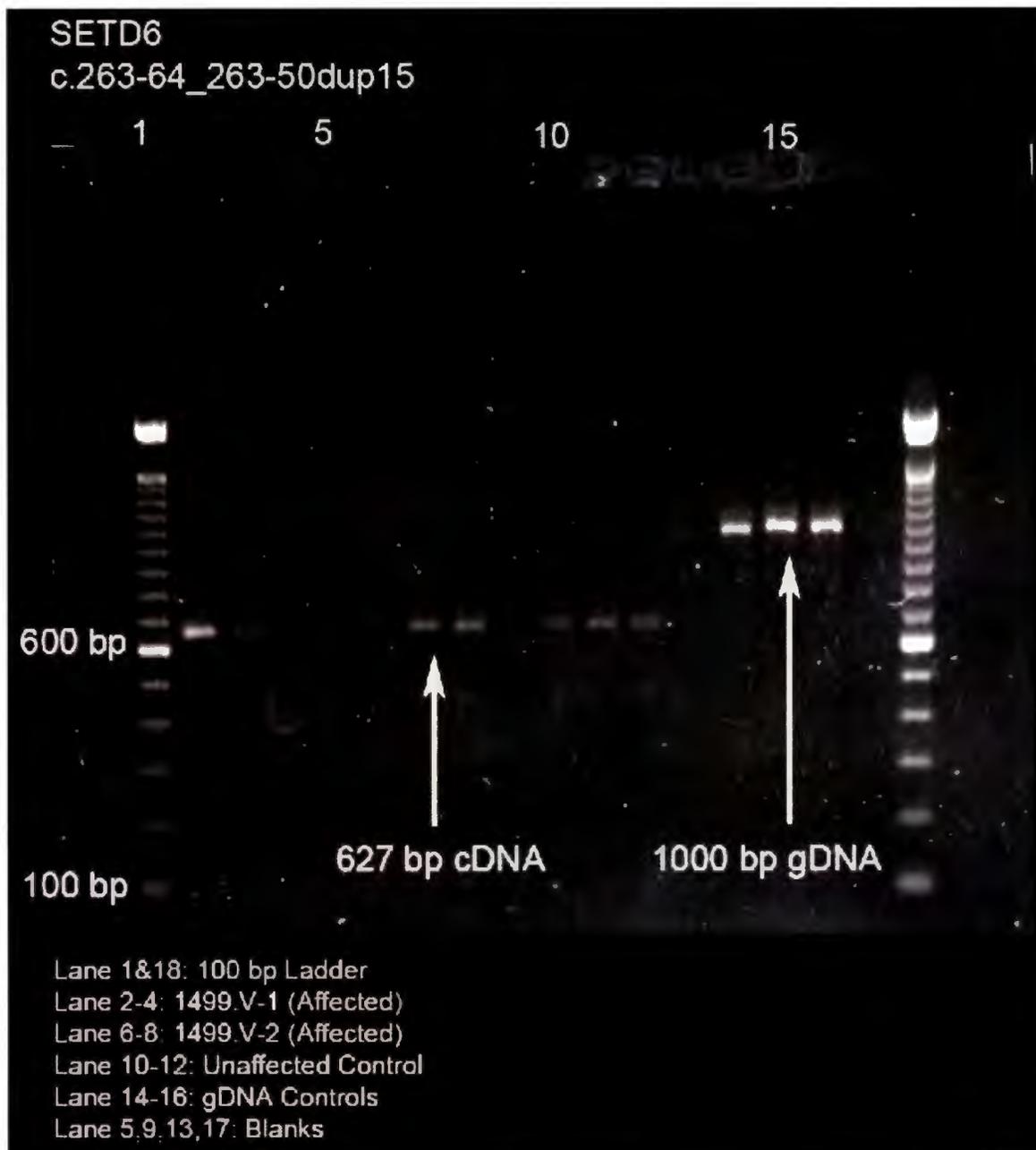


Figure 4.8 - 1% agarose gel ran at 100 volts for 1 hour. Primers were designed within exon 12 and exon 15 of *SETD6*, which would amplify a 1000 bp fragment in gDNA, and a 675 bp fragment in cDNA. The cDNA samples (run in triplicate) seen on the gel all show a fragment between the 600 and 700 bp fragments on the ladder which corresponds to the expected 675 bp fragment but did not show the 1000 bp fragment existent in gDNA (seen in lanes 14-16) which eliminates the possibility of gDNA contamination. The bands from the sample lanes (2-4 and 6-8) show a size consistent with that of the cDNA control (lanes 6-8), suggesting that the *SETD6* duplication has no discernible effect on splicing in this assay. An unexpected band between 400-500 bp can be faintly seen under each cDNA sample.

(iii) *GOT2* c.853+100_853+103del4

This four basepair deletion was not found in dbSNP and occurred towards the center of an intron. It was suspected that this could affect the consensus sequence surrounding the branch point required for proper intron splicing. One-hundred forty two control chromosomes were screened from the Newfoundland population and found this mutation in 115 of them (~80%). Due to the high frequency of this variant in the Newfoundland population, no further studies involving this variant were carried out.

4.4.5 Haplotype Construction and Critical Region Reduction

Analysis of positional candidate genes within the HH region revealed 98 variants in total. These genotypes were used along with genotypes from microsatellite markers surrounding the HH region (16q12-21) to construct haplotypes amongst the five affected individuals, and two parents. Analysis of these haplotypes reduced the critical region to a region between a variant in *HERPUDI* c.554+23 G>C and *D16S3132* which spans approximately 3cM, decreased from the original 8cM as denoted from the HH analysis (Tables 4.3 and 4.4).

Table 4.3 - Table showing all variants and microsatellite marker genotypes in the five affected individuals across both families, used to define the critical region on chromosome 16q21. The critical region is highlighted in yellow and can be defined between markers D16S673-D16S3132 where the genotypes differ between all the affected individuals. This haplotype spans a region of ~3 cM and is shared in a homozygous fashion between all five affected individuals and is heterozygous in two unaffected parents seen in Table 4.4.

Gene	Variant	SNP#	Family 0066						Family 1499			
			IV-13		IV-4		V-2		IV-1		IV-2	
			Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
<i>D16S673</i>			274	278	278	200	200	278	200	294	200	294
<i>HERPUD1</i>	c.1-146 G>A	rs37025	A	A	A	A	A	A	A	A	A	A
	c.1-86 T>G	rs37024	T	T	T	G	T	G	T	G	T	G
	c.554+23 G>C		G	G	G	C	G	C	G	G	G	G
	c.554+64 A>G	rs2133783	A	A	A	A	A	A	A	A	A	A
<i>CPNE2</i>	c.180+86 A>T	rs8053682	A	A	A	A	A	A	A	A	A	A
	c.180+87 G>C	rs8047463	C	C	C	C	C	C	C	C	C	C
	c.1302+39 C>G	rs2305693	G	G	G	G	G	G	G	G	G	G
<i>CCL22</i>	c.5 A>C	rs4359426	C	C	C	C	C	C	C	C	C	C
<i>CIAPIN1</i>	c.-131 G>A		A	A	A	A	A	A	A	A	A	A
	c.154 C>G	rs11557674	C	C	C	C	C	C	C	C	C	C
	c.157+16 C>T	rs72780616	C	C	C	C	C	C	C	C	C	C
<i>POLR2C</i>	c.137-24 A>G	rs223839	G	G	G	G	G	G	G	G	G	G
	c.174 T>C	rs4937	C	C	C	C	C	C	C	C	C	C
	c.387+8 C>T	rs601194	T	T	T	T	T	T	T	T	T	T
	c.439+100 G>A	rs223874	A	A	A	A	A	A	A	A	A	A
<i>D16S3057</i>			195	195	195	195	195	195	195	195	195	195
<i>CCDC102A</i>	c.286 C>T	rs12935069	T	T	T	T	T	T	T	T	T	T
	c.585+82 A>G	rs8045906	G	G	G	G	G	G	G	G	G	G
	c.825 G>A	rs28446687	A	A	A	A	A	A	A	A	A	A
	c.1140 A>G	rs28756858	G	G	G	G	G	G	G	G	G	G
	c.1258+14 T>A	rs28649020	A	A	A	A	A	A	A	A	A	A
<i>GPR114</i>	c.-7 G>A	rs1004363	A	A	A	A	A	A	A	A	A	A
	c.140+181 C>T	rs1859259	T	T	T	T	T	T	T	T	T	T
	c.140+269 G>A	rs1859260	A	A	A	A	A	A	A	A	A	A
	c.846 C>T	rs9937918	T	T	T	T	T	T	T	T	T	T
<i>CCDC135</i>	c.151 G>A	rs5564545	G	G	G	G	G	G	G	G	G	G
	c.557 G>C	rs7196016	C	C	C	C	C	C	C	C	C	C
	c.859-26 C>T	rs12933056	T	T	T	T	T	T	T	T	T	T
	c.1077+22 G >C	rs2122486	C	C	C	C	C	C	C	C	C	C
	c.1298 C>T	rs3809611	T	T	T	T	T	T	T	T	T	T
	c.2296 T>C	rs2923147	C	C	C	C	C	C	C	C	C	C
<i>KATNB1</i>	c.289+65 A>G	rs2967152	G	G	G	G	G	G	G	G	G	G
	c.704+4 C>T	rs2965797	T	T	T	T	T	T	T	T	T	T
	c.726C>T	rs2965798	T	T	T	T	T	T	T	T	T	T
<i>KIFC3</i>	c.381+105 T>C	rs2965789	C	C	C	C	C	C	C	C	C	C
	c.699 T>C	rs2967165	C	C	C	C	C	C	C	C	C	C
	c.765+32 G>A	rs2965790	A	A	A	A	A	A	A	A	A	A
	c.765+62 T>C	rs2967166	C	C	C	C	C	C	C	C	C	C

	c.940-44 A>G	rs3803590	G	G	G	G	G	G	G	G	G
	c.1803 C>T	rs2967172	T	T	T	T	T	T	T	T	T
	c.2003-21 T>A	rs2965799	A	A	A	A	A	A	A	A	A
ZNF319	c.1536 G>A	rs3743556	A	A	A	A	A	A	A	A	A
	c.1749+20 A>G	rs3743555	G	G	G	G	G	G	G	G	G
C16ORF57	c.99-36 G>A	rs3743560	A	A	A	A	A	A	A	A	A
	c.265+43 C>G		C	C	C	C	C	C	C	C	C
	c.504-60 C>T	rs4784022	T	T	T	T	T	T	T	T	T
MMP15	c.213-17 C>T	rs2241774	C	C	C	C	C	C	C	C	C
CCDC113	c.628-249 G>C	rs11076235	C	C	C	C	C	C	C	C	C
	c.628-242 C>G	rs11076236	G	G	G	G	G	G	G	G	G
	c.628-117 C>T	rs11076237	T	T	T	T	T	T	T	T	T
PRSS54	c.-6-16 G>A	rs2288013	G	G	G	G	G	G	G	G	G
	c.75 T>C	rs2288012	C	C	C	C	C	C	C	C	C
	c.263+56 A>G	rs3803582	A	A	A	A	A	A	A	A	A
	c.263+60 T>C	rs3803583	T	T	T	T	T	T	T	T	T
	c.263+61G>A		G	G	G	G	G	G	G	G	G
	c.264-112 A>G	rs4784916	G	G	G	G	G	G	G	G	G
	c.544 A>G	rs3815803	G	G	G	G	G	G	G	G	G
	c.718 G>A	rs1052276	G	G	G	G	G	G	G	G	G
	c.883 A>G	rs2241414	A	A	A	A	A	A	A	A	A
D16S3038			215	215	215	215	215	215	215	215	215
GINS3	c.198 T>G		G	G	G	G	G	G	G	G	G
NDRG4	c.223+90 G>A	rs2427787	A	A	A	A	A	A	A	A	A
	c.612+13 A>G	rs2271948	G	G	G	G	G	G	G	G	G
	c.612+77 A>G	rs40185	G	G	G	G	G	G	G	G	G
	c.672 G>A	rs17821543	A	A	A	A	A	A	A	A	A
	c.873+3 G>A		A	A	A	A	A	A	A	A	A
	c.1062 A>G	rs42945	G	G	G	G	G	G	G	G	G
SETD6	c.193 T>C	rs4784046	C	C	C	C	C	C	C	C	C
	c.263-64_263-50dup		dup								
	c.902-77 G>A	rs2280397	A	A	A	A	A	A	A	A	A
	c.1338 G>A	rs3607	A	A	A	A	A	A	A	A	A
	c.*26 A>G	rs3169293	G	G	G	G	G	G	G	G	G
CNOT1	c.103-32 C>G	rs34014750	G	G	G	G	G	G	G	G	G
	c.976 A>G		A	A	A	A	A	A	A	A	A
	c.1216-69 G>A	rs72790275	G	G	G	G	G	G	G	G	G
	c.1584+22 dup T		dup								
	c.1704+110 G>A	rs9936144	G	G	G	G	G	G	G	G	G
	c.2332+99 G>A	rs41445346	G	G	G	G	G	G	G	G	G
	c.2891+43 C>T	rs1568399	T	T	T	T	T	T	T	T	T
	c.2919 G>A	rs11866002	A	A	A	A	A	A	A	A	A
	c.3201+124 G>T	rs7184114	G	G	G	G	G	G	G	G	G
	c.4006+12 A>G	rs2241570	G	G	G	G	G	G	G	G	G
	c.4434+90 G>A	rs41260	A	A	A	A	A	A	A	A	A
	c.4716 C>T	rs11540994	T	T	T	T	T	T	T	T	T
	c.4801-83 G>A	rs246195	A	A	A	A	A	A	A	A	A
	c.5415-61 G>A	rs2244453	A	A	A	A	A	A	A	A	A
	c.5646+41 A>G		A	A	A	A	A	A	A	A	A
	c.5646+88 T>G		T	T	T	T	T	T	T	T	T

	c.*82 G>A	rs17821549	A	A	A	A	A	A	A	A	A	A
SLC38A7	c.883+19 delC		del									
	c.1286+51 T>C	rs9923128	C	C	C	C	C	C	C	C	C	C
GOT2	c.213 T>C	rs257636	C	C	C	C	C	C	C	C	C	C
	c.228 T>G	rs14221	G	G	G	G	G	G	G	G	G	G
	c.246+252 C>T	rs257634	T	T	T	T	T	T	T	T	T	T
	c.562 G>A	rs11076256	G	G	G	G	G	G	G	G	G	G
	c.816 C>T	rs1058192	T	T	T	T	T	T	T	T	T	T
	c.853+100_853+103del		del									
	c.1037 T>G	rs30842	G	G	G	G	G	G	G	G	G	G
D16S494			188	188	188	188	188	188	188	188	188	188
D16S3094			169	169	169	169	169	169	169	169	169	169
D16S3132			264	272	272	282	272	282	264	264	264	264
D16S3089			193	193	193	193	193	193	193	183	193	183

Table 4.4 - Table illustrating the haplotype presented in Table 4.3 within unaffected parents. The portion shared by the affected individuals is highlighted in yellow, and can be seen in the heterozygous state in these two parental samples.

Gene	Variant	SNP#	Family 0066		Family 1499	
			V-2		III-1	
			Allele 1	Allele 2	Allele 1	Allele 2
<i>D16S673</i>			280	200	278	294
<i>HERPUD1</i>	c.1-146 G>A	rs37025	G	A	G	A
	c.1-86 T>G	rs37024	G	T	T	G
	c.554+23 G>C		C	G	G	G
	c.554+64 A>G	rs2133783	A	A	G	A
<i>CPNE2</i>	c.180+86 A>T	rs8053682	A	A	T	A
	c.180+87 G>C	rs8047463	C	C	G	C
	c.1302+39 C>G	rs2305693	G	G	G	G
<i>CCL22</i>	c.5 A>C	rs4359426	C	C	C	C
<i>CIAPIN1</i>	c.-131 G>A		A	A	G	A
	c.154 C>G	rs11557674	C	C	G	C
	c.157+16 C>T	rs72780616	C	C	T	C
<i>POLR2C</i>	c.137-24 A>G	rs223839	G	G	A	G
	c.174 T>C	rs4937	C	C	C	C
	c.387+8 C>T	rs601194	T	T	C	T
	c.439+100 G>A	rs223874	A	A	A	A
<i>D16S3057</i>			195	195	191	195
<i>CCDC102A</i>	c.286 C>T	rs12935069	T	T	T	T
	c.585+82 A>G	rs8045906	G	G	G	G
	c.825 G>A	rs28446687	A	A	A	A
	c.1140 A>G	rs28756858	G	G	G	G
	c.1258+14 T>A	rs28649020	A	A	A	A
<i>GPR114</i>	c.-7 G>A	rs1004363	A	A	A	A
	c.140+181 C>T	rs1859259	T	T	T	T
	c.140+269 G>A	rs1859260	A	A	A	A
	c.846 C>T	rs9937918	C	T	C	T
<i>CCDC135</i>	c.151 G>A	rs5564545	G	G	A	G
	c.575 G>C	rs7196016	G	C	G	C
	c.859-26 C>T	rs12933056	C	T	C	T
	c.1077+22 G >C	rs2122486	G	C	G	C
	c.1298 C>T	rs3809611	C	T	C	T
	c.2296 T>C	rs2923147	T	C	T	C
<i>KATNB1</i>	c.289+65 A>G	rs2967152	A	G	A	G
	c.704+4 C>T	rs2965797	C	T	C	T
	c.726C>T	rs2965798	C	T	C	T
<i>KIFC3</i>	c.381+105 T>C	rs2965789	T	C	T	C
	c.699 T>C	rs2967165	T	C	T	C
	c.765+32 G>A	rs2965790	G	A	G	A
	c.765+62 T>C	rs2967166	T	C	T	C
	c.940-44 A>G	rs3803590	A	G	A	G
	c.1803 C>T	rs2967172	C	T	C	T
	c.2003-21 T>A	rs2965799	T	A	T	A

ZNF319	c.1536 G>A	rs3743556	G	A	G	A
	c.1749+20 A>G	rs3743555	A	G	A	G
C16ORF57	c.99-36 G>A	rs3743560	G	A	G	A
	c.265+43 C>G		C	C	G	C
	c.504-60 C>T	rs4784022	C	T	C	T
MMP15	c.213-17 C>T	rs2241774	C	C	T	C
CCDC113	c.628-249 G>C	rs11076235	C	C	G	C
	c.628-242 C>G	rs11076236	G	G	C	G
	c.628-117 C>T	rs11076237	T	T	C	T
PRSS54	c.-6-16 G>A	rs2288013	G	G	A	G
	c.75 T>C	rs2288012	C	C	C	C
	c.263+56 A>G	rs3803582	A	A	G	A
	c.263+60 T>C	rs3803583	T	T	C	T
	c.263+61G>A		G	G	A	G
	c.264-112 A>G	rs4784916	G	G	G	G
	c.544 A>G	rs3815803	G	G	A	G
	c.718 G>A	rs1052276	G	G	A	G
	c.883 A>G	rs2241414	A	A	G	A
D16S3038			215	215	215	215
GINS3	c.198 T>G		T	G	T	G
NDRG4	c.223+90 G>A	rs2427787	A	A	A	A
	c.612+13 A>G	rs2271948	A	G	A	G
	c.612+77 A>G	rs40185	G	G	G	G
	c.672 G>A	rs17821543	G	A	G	A
	c.873+3 G>A		G	A	G	A
	c.1062 A>G	rs42945	A	G	A	G
SETD6	c.193 T>C	rs4784046	T	C	T	C
	c.263-64_263-50dup		wt	dup	wt	dup
	c.902-77 G>A	rs2280397	G	A	G	A
	c.1338 G>A	rs3607	G	A	G	A
	c.*26 A>G	rs3169293	A	G	A	G
CNOT1	c.103-32 C>G	rs34014750	C	G	C	G
	c.976 A>G		G	A	G	A
	c.1216-69 G>A	rs72790275	A	G	A	G
	c.1584+22 dup T		wt	dup	wt	dup
	c.1704+110 G>A	rs9936144	A	G	A	G
	c.2332+99 G>A	rs41445346	A	G	A	G
	c.2891+43 C>T	rs1568399	C	T	C	T
	c.2919 G>A	rs11866002	G	A	G	A
	c.3201+124 G>T	rs7184114	T	G	T	G
	c.4006+12 A>G	rs2241570	A	G	A	G
	c.4434+90 G>A	rs41260	G	A	G	A
	c.4716 C>T	rs11540994	C	T	C	T
	c.4801-83 G>A	rs246195	G	A	G	A
	c.5415-61 G>A	rs2244453	G	A	G	A
	c.5646+41 A>G		G	A	G	A
	c.5646+88 T>G		G	T	G	T
	c.*82 G>A	rs17821549	G	A	G	A
SLC38A7	c.883+19 delC		del	del	wt	del
	c.1286+51 T>C	rs9923128	C	C	C	C

<i>GOT2</i>	c.213 T>C	rs257636	C	C	C	C
	c.228 T>G	rs14221	G	G	G	G
	c.246+252 C>T	rs257634	T	T	T	T
	c.562 G>A	rs11076256	G	G	A	G
	c.816 C>T	rs1058192	C	T	T	T
	c.853+100_853+103del		wt	del	wt	del
	c.1037 T>G	rs30842	T	G	G	G
<i>D16S494</i>			178	188	178	188
<i>D16S3094</i>			169	169	169	169
<i>D16S3132</i>			278	282	268	264
<i>D16S3089</i>			181	193	187	183

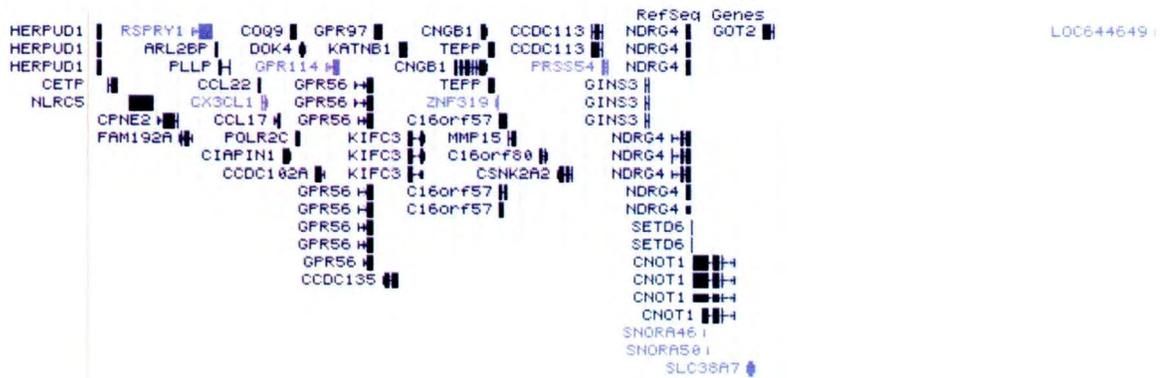


Figure 4.9: Screen capture from the UCSC Genome Bioinformatics database of the critical region between markers *HERPUD1* c.554+23 G>C - *D16S3132* (Human Feb 2009 GRCh37/hg19 Assembly). This region spans approximately 3 cM and contains a small cluster of 71 genes followed by a rather large intergenic region with a sole pseudogene *LOC644649*.

4.4.6 Exome Sequencing

(i) Sample Pooling Strategy

Following the Sanger sequencing and haplotype analysis, we sent away 8 genomic DNA (gDNA) samples (five affected, three unaffected) for whole exome sequencing to the Genome Center at McGill University. Genomic DNA samples from three affected individuals from Family 0066 were sent away and pooled, with the assumption that these three samples would contain the same autosomal recessive mutation and similarly the two gDNA samples from Family 1499 were pooled.

(ii) Variant Filtering

Exome sequencing was carried out as per Section 3.3.4 of this thesis. The exome dataset was filtered for rare variants, which would be consistent with an autosomal recessive mode of inheritance (homozygous or compound heterozygous), shared between both Families 0066 and 1499, and within the chromosome 16 critical region determined through linkage, HH analysis, and critical region reduction.

As a double check, the exome dataset was checked for mutations in genes related to microphthalmia and its associated syndromes (*OTX2*, *PAX6*, *SOX2*, *RAX*, *VSX2*, *BCOR*, *CHD7*, *BMP4*, *FOXE3*, *STRA6*, *SMOC1*, *MITF*, and *SIX6*), as well as *GJA1*, which is associated with ODDD and HSS-like phenotypes. No pathogenic variants were found during this analysis.

0066 Filtering

Filtering for variants in the chromosome 16 critical region revealed 314 variants, 271 of which were common SNPs. Of the remaining 43 variants, there were three missense mutations, five synonymous mutations and no truncating mutations (frameshifting deletions, or nonsense mutations). The three missense mutations were in *CNGBI* (p.A961V), *ZNF319* (p.K512K), and *PRSS54* (p.S182G). These variants had a minor allele frequency (MAF) of 2.2%, 66.8%, and 54.6% respectively from the 1000 Genomes Project.

1499 Filtering

Filtering for variants in the chromosome 16 critical region revealed 294 variants, of which 261 were common SNPs. Of the remaining 33 variants, there were two missense variants, five synonymous variants, and no truncating mutations. The two missense mutations were in *CNGBI* (p.A961V) and *PRSS54* (p.S182G) having a 2.2% and 54.6% MAF respectively from the 1000 Genomes Project.

4.4.7 CNV Analysis

To examine the hypothesis of CNVs in this region, we asked a bioinformatician at Genome Quebec (Alexandre Monpetit) to analyze the SNP data from the Illumina 610K Quad Chip for CNVs in this region. This quick analysis revealed that the individuals were sharing a large homozygous region, as was previously suspected, but the copy numbers of each SNP were normal, suggesting that this area does not contain any CNVs of interest.

4.5 Discussion

4.5.1 MDW and HSS

Microphthalmia is a condition caused by the malformation of the optic vesicle during embryological development, and is part of a broader spectrum of disorders including clinical anophthalmia and coloboma. This study presents a syndromic form of this spectrum, which we refer to as microphthalmia-dwarfism (MDW) and which has not previously been described. It is not surprising to find rare conditions associated with microphthalmia, as up to 1/3 of microphthalmia cases are associated with extraocular features.¹⁹⁸

Interestingly, the individuals with MDW bear great similarity to individuals with Hallermann-Streiff syndrome (HSS; OMIM 234100). HSS is an autosomal recessive condition consisting of brachycephaly with microphthalmia, proportionate short stature, frontal bossing, hypotrichosis, cataracts, and a pointed/beaked nose, all of which are seen in the two families with MDW. Despite the similarities of these conditions, the individuals with MDW do not show a progeroid (premature aging) component like those of some HSS patients,²²⁸ and do not exhibit micrognathia (small jaw). It is quite possible that both HSS and MDW are caused by mutations of the same (or functionally similar) gene(s), which would account for the similarities between these conditions. Interestingly, assessment of patient 0066 IV-4 provides detailed clinical data over the course of many years. In her earlier years, pictures showed that she had a full head of hair, but at the age of 60 she had marked hypotrichosis, though always blind, having microphthalmia and cataracts. Her facial structure and appearance was certainly not that of classic HSS in the pictures from her youth, though more recent pictures in her 50s and 60s show a much more similar looking phenotype to HSS as described by Cohen in 1991. This very interesting observation suggests that perhaps the MDW phenotype observed in Families 0066 and 1499 is indeed part of the HSS spectrum, but a milder, progressive form of the disease.

Other, younger, patients (1499 IV-1, IV-2 and 0066 V-2) from these two families also do not show the same features as their older relatives, such as hypotrichosis, respiratory issues (COPD, chronic bronchitis), hearing loss, and frequent infection. However, they may develop these phenotypes later in life, again suggesting that this condition may be a more mild progressive form of HSS.

4.5.2 Linkage Analyses and HH mapping

Elucidating the genetics of MDW posed to be a potentially interesting and challenging endeavor as this condition seems to be related to HSS, a condition with little to no knowledge of monogenetic involvement. We began these analyses through genotyping all available samples from families 0066 and 1499 on a 610K Illumina SNP array in hopes of identifying a putative locus through linkage analyses. This particular array was chosen over a more densely populated chip, due to the lower cost and the lower number of markers required for linkage analysis. The reduced number of SNPs eases the speed of analysis, but still provides sufficient coverage due to the large number of SNPs available in comparison to microsatellites. This strategy proved quite useful as we were able to determine a multi-point linkage on chromosome 16 of 4.535 on band q12-21 which spanned approximately 30 cM. This large region contained over 300 genes, which is not an ideal situation for Sanger sequencing of positional candidate genes (as Sanger sequencing can only sequence a maximum of ~700-1000 bp at a time).

While the linkage analyses were being performed, a HH analyses was also carried out using samples from affected individuals of both families. This analysis revealed the same, but smaller region on chromosome 16q12-21 which spanned ~8.5 cM between SNPs rs289710 and rs219568. This, combined with the linkage data, suggested that the affected members from

families 0066 and 1499 were either sharing a common homozygous region on chromosome 16, or were potentially carrying a heterozygous deletion of this entire region (this is discussed further in section 4.5.5). This region was also significantly smaller than the 30 cM region as indicated by the linkage analysis, and contained only 71 genes.

4.5.3 Positional Candidate Gene Selection and Sequencing

For ease, positional candidate gene selection was based on the smaller region as determined by the HH analysis presented in section 4.5.2. Of the 71 genes in this region (Human Feb 2009 GRCh37/hg19 Assembly), we chose 23 genes for preliminary direct sequencing based on function and previous disease association. Genes associated with i) cellular proliferation/growth or ii) the cell cycle (*NDRG4*, *ARL2BP*), iii) were transcription factors or proteins related to DNA replication/transcription (*CCDC102A*, *POLR2C*, *CCDC135*, *ZNF319*, *CCDC113*, *CNOT1*, *GINS3*), iv) a cellular function involving proper function/growth (*HERPUD1*, *CIAPIN1*, *GPR114*, *KIFC3*, *KATNBI*, *MMP15*, *CPNE1*, *SLC38A7*, *PRSS54*), v) or mitochondrial function/energy production (*GOT2*) were chosen. Other genes were chosen based on previous disease associations such as *C16ORF57* in poikiloderma/neutropenia, or if they had no previous association but could be plausibly involved such as *SETD6* (histone methylation) or *C16ORF80* (no known function or association).

Sequencing of these 23 genes revealed 94 variants in total (Table 4.2), 82 with a previous dbSNP number. Eleven of these variants were missense mutations and all underwent *in silico* analysis using SIFT (Sorting Intolerant From Tolerant), PolyPhen, and PANTHER. The scores from these analyses (when a missense mutation), along with minor allele frequencies (MAF) from the 1000 genome project were used to determine putative pathogenicity of each variant.

Three variants, c.878+3 G>A in *NDRG4*, c.263-64_263-50dup in *SETD6*, and c.853+100_853+103del in *GOT2* were of interest due to low MAF and potential splicing effects.

The *GOT2* variant, c.853+100_853+103del, was a four bp deletion of GTGT in intron 7. This variant was chosen for further analysis as it had no previous rs number from dbSNP, and also showed segregation with the MDW phenotype. One hundred and forty two chromosomes from the NL population were tested for this mutation, and 115 carried this variant. (~80%). The high frequency at which c.853+100_853+103del occurred in the Newfoundland population excluded this variant from any further analysis.

The c.873+3 G>A transition in *NDRG4* occurs close to exon 14, making it a potential splice variant and is a previously described SNP (rs142344402). This SNP only shows a 0.2% frequency in all genomes (0.006 AMR and 0.004 EUR) according to the 1000 Genomes project (accessed June 17th, 2012), and has never been described in the homozygous state. Segregation analysis of this variant was carried out amongst all remaining family members, and appeared to segregate in an autosomal recessive fashion with the MDW phenotype. The rare occurrence of this SNP, the segregation of the allele, its potential for splice variation and the fact that it has never been seen as a homozygote made this variant interesting to study. This variant was first checked in normal population controls from an ongoing colorectal study, and was found in 1 of 142 control chromosomes (0.7%) indicating that this variant is also rare in the Newfoundland population. To search for a potential effect of this variant in the MDW families, three primers were designed to amplify a 325 bp fragment in cDNA and a 675 bp fragment in gDNA through multiplex PCR. Two lymphoblastoid cell lines from family 1499 of the proband (IV-1) and his brother (IV-2) were available for testing. The results from this experiment showed a fragment size consistent with the expected 325 bp which was confirmed in an unaffected control from a

separate project. This suggests that the c.873+3 G>A variant is likely benign, having no negative effect on splice variation in *NDRG4*, and is simply riding on a rare haplotype in the Newfoundland population. This was further confirmed through *in silico* analyses of the potential splice variation using BDGP (Donor scores G:0.85 A:0.97) and NetGene2 (Donor scores: G 0.95, A: 1.00), both of which predict a slightly stronger splice donor due to this mutation, but no negative effects.

The mutation in *SETD6* (c.263-64_263-50dup) was also tested for potential splice variation as it was a duplication towards the center of intron two. This experiment utilized cDNA from individuals IV-1 and IV-2 from family 1499, and used two primers which would amplify a 675 bp fragment in cDNA, and a 1000 bp fragment in gDNA. This experiment was analyzed using gel electrophoresis and showed a band at the expected 675 bp fragment in cDNA, which also showed up in an unaffected control sample. Interestingly, there appeared to be second fragment in between the 400-500 bp ladder controls. This band was weak, but appeared in multiple iterations of this experiment, and also showed up in the control sample. According to the UCSC Genome Browser, there are currently two known isoforms of *SETD6* (NM_001160305 and NM_02486). Upon inspection, it appeared that isoform NM_02486 is missing a total of 218 bp from exon 2 that the longer isoform, NM_001160305, includes in its open reading frame (ORF). This isoform would thus produce a band of 457 bp, and accounts for the second band seen on the gel in Figure 4.8. These results suggest that both isoforms of *SETD6* are expressed in lymphoblastoid cells lines, though based on the intensity of the observed bands in Figure 4.8 the longer isoform (NM_001160305) is expressed more readily. The presence of the 675 bp fragment in all three samples (two affecteds, one control) suggests that the 15 bp

duplication (c. c.263-64_263-50dup) is unlikely to cause any splice variation in the MDW patients.

4.5.4 Exome Sequencing and Analysis

Once variant analysis had been carried out from the Sanger sequencing data, we outsourced DNA from families 0066 and 1499 to the Genome Centre at McGill University for exome sequencing. We employed a pooling strategy in which we created an affected and unaffected pool for each of the two families. This was done as it was expected that all the affected individuals would share the same homozygous, or compound heterozygous mutation. We also filtered the variant list for those genes in the critical region on chromosome 16q21. Filtering for the critical region in both families 0066 and 1499 revealed two shared missense mutations, *CNGBI* (p.A961V), *PRSS54* (p.S182G). The *PRSS54* variant can be easily excluded as this mutation has been seen in 54.6% of samples from the 1000 Genomes project, making it likely that this is a common polymorphism. The *CNGBI* variant, while rare can, also be excluded as this gene functions solely in the rod photoreceptors, and mutations are known to cause autosomal recessive retinitis pigmentosa.

Unfortunately, no rare variants of potential causative effects were shared between the affected members of both families in this critical region on chromosome 16. This is an interesting result, as we initially expected the causative mutation to easily fall out during filtering, using these criteria. This data led us to believe that perhaps the mutation was not in a coding gene, but perhaps a non-coding one. This region contained two genes which fit these criteria, *SNORA46* and *SNORA50*. *SNORA 46* and *SNORA50* are H/ACA small nucleolar RNA (snoRNA) genes which are transcribed from their host gene (the gene in which they reside),

CNOT1. The H/ACA class of snoRNA molecules is responsible for a process known as pseudouridylation, which creates a post-transcriptionally modified nucleoside, pseudouridine.²³⁵ Pseudouridine is important in the structures of various RNA molecules such as small nuclear RNAs (snRNAs), or transfer RNAs (tRNAs), though the exact mechanism in which they work is unknown.

4.5.5 Other Analyses and Future Directions

The lack of any pathogenic variants found in the critical chromosome 16 region is a very interesting result. We had expected exome sequencing to reveal the mutation with relative ease, especially with the convincing locus on 16q21 revealed through linkage, HH, and critical region analyses. Certainly, with exome sequencing being a rather new and unperfected technology, the causative gene may have been overlooked in analysis, or have been missed in the sequencing itself. The second possible conclusion from this data falls in line with the hypothesis of the causative mutation potentially being in a non-coding gene, or an intergenic region.

Recently, a microdeletion on 16q11.2-q12.1 has been shown to cause a case of Lenz microphthalmia,²³⁶ thus we examined the possibility of copy number variation (CNV) as the cause of MDW. The haplotypes shown in Table 4.3 and 4.4 show that the yellow haplotype segregates with the disease but is heterozygous in two unaffected parents. This makes the presence of a large heterozygous deletion (which would give the appearance of a homozygous haplotype) in the affected family members unlikely. To further examine the hypothesis of CNVs in this region, we asked a bioinformatician at Genome Quebec (Alexandre Monpetit) to analyze the SNP data from the Illumina 610K Quad Chip for CNVs in this region. This quick analysis revealed that the individuals were sharing a large homozygous region, as was previously

suspected, but the copy numbers of each SNP were normal, suggesting that this area does not contain any CNVs of interest. It is certainly possible that a deleterious CNV may exist between probes in the 610K SNP Illumina Quad-Chip that we used for this study. This possibility could be explored through further studies using a higher density SNP array, or array comparative genomic hybridization (aCGH) which are sensitive arrays designed to pick up variations in copy numbers.

Figure 4.9 depicts the critical region from the MDW study according to the UCSC Genome Browser, and contains a small gene cluster (71 genes) followed by a particularly large intergenic region of about 2 cM. Another possibility for this study is that the mutation lies in the large 2 cM intergenic region, and may be a point mutation creating or disrupting an enhancer or repressor site. This phenomenon has been examined in multiple association studies of prostate cancer which have identified associations with multiple SNPs in a gene desert on 8q24. This region has been shown to be an enhancer region for a tumour suppressor gene called *MYC*. Functional studies of this region have shown that this region interacts with the *MYC* promoter in both prostate and colorectal cancer cell lines through a conformational change of the DNA.²³⁷ To fully examine this hypothesis, this large intergenic region would need to be sequenced through next generation sequencing technologies to identify any potentially pathogenic variants.

It is also very interesting to note that the critical region on chromosome 16 appears to contain a number of genes which explain (or could potentially explain) components of the MDW phenotype. For example, *C16ORF57* causes both poikiloderma (a condition of the skin which can cause thinning of the skin leading to prominent veins, and discolouration of the skin), and neutropenia, a lack of neutrophils in the blood stream leading to frequent infections.²³⁸ This made this gene an exciting candidate as the individuals from the MDW family have both

atrophic/dry skin with prominent veins, as well as frequent infections. Other genes in this region are involved in other aspects of immune function (*NLR5*), cholesterol and fat metabolism (CETP), which is interesting as individuals with MDW lose facial fat deposits over time. It is possible that this cluster of 71 genes is being controlled by a common promoter/regulatory element, and that disruption of this region would cause dysregulation of gene expression, resulting in the MDW phenotype. Though, again, further analysis would be required to confirm this through Next Generation Sequencing (NGS) or aCGH in the case of CNVs to identify the putative mutation. Quantitative expression studies would then need to be conducted upon discovery of a variant to help confirm pathogenicity, and regulation.

This project was very interesting, as it illustrated the strength of so called 'old' techniques such as linkage analysis, and also highlighted the limitations of newer technologies like exome sequencing. Though powerful, NGS is currently in the infancy stages of becoming mainstay tools for genetic analyses as the sheer amount of data is rather difficult to deal with. It also creates a problem when coverage of particular areas is low either from random error, or is particularly hard to sequence due to the nature of the region (i.e. CG rich areas, repetitive regions, or areas near the centromere). This project shows that again, while powerful, exome sequencing is not the 'magic bullet' solution to all conditions thought to be Mendelian or monogenetic in nature. This is not entirely due to limitations of NGS, but that in some cases, diseases thought to be Mendelian or monogenic in nature, may in fact be more complex than originally thought. It is certainly likely that as researchers are better able to handle the sheer amount of raw sequence data produced by the NGS platforms, that we'll begin to see more and more conditions solved through mutations in non-coding areas.

It's interesting to note that these regions which were once thought to be 'junk DNA' or 'evolutionary leftovers' can in fact be playing a much larger role than originally thought. A consortium called the Encyclopedia of DNA Elements (ENCODE), started in 2003, aims to identify the function of these so called 'junk DNA' regions. This group has published numerous papers on the location and functions of these DNA elements, and suggests that up to 80% of the human genome is involved in biological function, and not solely the 1-2% of coding DNA originally thought to control the majority of biological processes.²³⁹ Recently, much excitement has been generated about the function of non-coding transcripts (RNA species) in organization of chromatin in the nucleus. It has been shown that 274 tRNA genes show localization to the nucleolus of Yeast. It is suspected that this clustering has a major impact on the spatial organization of the genome. Secondly, repetitive elements of the genome such as Short Interspersed Nuclear Elements (SINE) have been shown to contain RNA Polymerase III (Pol III) binding sites, and it has been hypothesized that these species participate in chromatin remodeling within the cell.²⁴⁰

Other non-coding elements code for multiple types of regulatory components such as binding sites for transcription factors, regulatory RNA species such as microRNA molecules (miRNA), or long non-coding RNA (lncRNA), or trans and cis acting enhancers. The two RNA molecules (miRNA) and (lncRNA) have been implicated in various conditions such as cardiovascular disease, various cancers, and neurodegenerative conditions (such as Alzheimer's, Parkinson's, and Huntington's).²⁴¹ Interestingly, and perhaps most applicable to the hypothesis that a regulatory mutation in the non-coding region of chromosome 16 for MDW, is the discovery of a regulatory element on 8q24 for prostate cancer. It was found that a SNP within a 'gene desert' of chromosome 8 (rs6983267) which had been previously associated with prostate

cancer in a number of association studies was within a cis-acting enhancer of the gene *MYC*, a known proto-oncogene (a gene that when mutated can promote cancer growth).²³⁷ While prostate cancer is a complex trait, and MDW seems to be a recessive condition, the possibility remains that perhaps a mutation in this intergenic region of chromosome 16 could be having a negative impact on the expression of nearby genes, and perhaps giving rise to the MDW phenotype.

Chapter 5: General Discussion and Conclusions

The goal of this thesis was to use the genetic isolates of Newfoundland and Labrador to identify disease genes in various ocular disorders. These conditions included ASD, a variable autosomal dominant defect of the anterior chamber of the eye including Peters anomaly; achromatopsia, an autosomal recessive malfunction of the cone photoreceptors; and MDW an autosomal recessive syndromic form of microphthalmia. The use of genetic isolates and founder populations for gene discovery projects has been very successful in the past,⁴⁹ and the population of Newfoundland and Labrador is no exception in that regard. Researchers in NL have successfully identified a number of founder mutations, and have had past successes in gene discovery projects.^{53-55; 59; 72; 142} This thesis highlights the advantages of using the genetic isolate of Newfoundland for gene discovery projects, and also examines some of the limitations involved with the current perception of monogenic diseases.

ASD is an autosomal dominant condition of the anterior ocular chamber, consisting of various malformations of the iris, the cornea, and the lens. This thesis examined a single multiplex family from Newfoundland which segregated a variable form of this condition. Affected individuals carried a spectrum of disorders ranging from mild phenotypes such as microcornea, to more severe forms such as Peters Anomaly. ASD is also a very genetically heterogeneous condition with over a dozen known genes and loci. The variable expression and the genetic heterogeneity of this disorder made this a potentially challenging study. As highlighted in the general introduction, this project began with the sequencing of functional candidate genes which had been previously associated with ASD. Nine genes were chosen and sequenced in this fashion, and we identified a novel mutation in *FOXE3* (c.959 G>T). Up to this point, only a handful of publications had described *FOXE3* mutations in ASD. The gene product,

FOXE3, is known to be a lens specific transcription factor responsible for proper lens formation during embryogenesis. The mutation discovered through this method (c.959 G>T) was of great interest as it was a rare type of mutation, referred to as a non-stop mutation (a mutation of the stop codon in a gene). This type of mutation is rarely seen, and the exact effects of its action in *FOXE3* have simply been hypothesized to cause an extended protein though no functional studies have been carried out to confirm this. To further examine this hypothesis, we used cDNA from an affected individual to identify that the c.959 G>T mutation was in fact absent in cDNA according to Sanger sequencing. This preliminary data suggests that the mutation discovered in family 0023 is degraded at the mRNA level. Of course this result is strictly based on the limitations surrounding Sanger sequencing, which is unable to determine quantitative levels of cDNA, and thus mRNA. To further examine this hypothesis, quantitative studies using PCR based methods could be carried out to quantify the amount of the mutant transcript (if any).

The second project described in this thesis concerns a rare autosomal recessive condition known as achromatopsia. Achromatopsia is a retinal condition affecting the cone photoreceptors, and is caused by mutations in one of four genes, *CNGA3*, *CNGB3*, *GNAT2*, and *PDE6C*. Much like the situation with ASD, all four previously associated genes were sequenced in individuals from eight families from Newfoundland. Six of these families were solved through this method and affected individuals had mutations in *CNGA3* and *CNGB3* (three with mutations in *CNGA3*, and three with mutations in *CNGB3*) while an individual of a seventh; unsolved family had one mutation in *CNGB3* identified. The possibility of founder mutations was also evaluated, and it was found that the c.1148delC mutation in *CNGB3* appears to be on the same disease associated haplotype in all four families (1442, 1492, 1713, and 1734). The limitation of this conclusion lay in the fact that only one affected sample was available from families 1442, 1713 and 1734.

Although, only one sample from an affected individual was available from Family 1442, all marker/SNP genotypes were homozygous and thus a haplotype was able to be established for this individual. A second founder mutation was also found in *CNGA3* within this population, c.1580T>G: p.L527R. We compared haplotypes across the two families who harboured this mutation (1491, and 1723), though only one sample was available from family 1723. This showed a common partial haplotype surrounding this mutation suggesting that this may be a recurrent founder population in the population of Newfoundland. Interestingly, this mutation has also been described by Lam et al in 2011 through exome sequencing of a single patient of 'European Descent'. The Newfoundland population is known to have been peopled by a group of individuals of Northern European descent, thus it would be interesting to obtain a sample from the patient described by Lam et al, and compare haplotypes. This could provide further evidence for p.L527R as a founder mutation, and could provide evidence that it originated in Europe and was carried into the Newfoundland population during migration.

All the haplotype data discussed above is based around a rather large assumption: that all the mutations being examined were on a common haplotype. Most of the samples obtained were from a single individual in a family and without at least three generations of data, or a large number of siblings and parents across at least two generations, constructing a haplotype is problematic. The assumption made here was that the mutation is shared on a common haplotype, and when markers surrounding *CNGB3/CNGA3* were genotyped, they could be arranged to produce a similar haplotype. Of course, this is a huge assumption, and parental DNA would be required in order to confirm the segregation of these marker genotypes.

Interestingly, one family (1734) showed one mutation in *CNGB3* (c.1148delC), but a second mutation was unable to be found in any of the four genes. A sequence anomaly was

observed surrounding exons 17 and 18 in *PDE6C* (Figure 3.4) though it was determined through a series of cloning experiments and primer tests, that the primers used to amplify these exons were not as specific as *in silico* analysis had suggested, and had amplified a region on Chr8. It is also possible that the conditions for amplification were not stringent enough, allowing the primers to bind in other locations. A second set of primers was designed to amplify this region, and no pathogenic variants were found in the patient sample from Family 1734. It remains possible that the second mutation for this individual is in *CNGB3*, but is unable to be picked up by sequencing the coding regions of this gene. The second mutation may be in a regulatory domain of *CNGB3*, such as the promotor, or perhaps in one of the larger intronic regions creating a splice variation in a branch site. To evaluate the possibility of a splice mutation somewhere in *CNGB3*, a cell line could be obtained, and the entire *CNGB3* cDNA transcript amplified and compared to a normal control. Any variation in the size of the cDNA transcript could provide hints towards the mechanism of disease in this individual. A second approach to solve this individual would be to use whole genome or targeted next generation sequencing to identify the second mutation, or potentially another gene entirely. The limitation of this method lies in the massive amounts of data that is produced. If mutations are found in a regulatory region, it becomes difficult to distinguish which variant is having an effect without functional studies on each individual variant found. My suspicion is that the former situation is going to be the most likely, where a splice variation of *CNGB3* is the second mutation which was missed through the initial screening of coding regions in functional candidate genes.

Even more interesting was family 0094, a large family with achromatopsia and no pathogenic mutations in any of the four known achromatopsia genes. Rather than using linkage analyses for this project, we decided to use NGS technology and used exome sequencing to solve

the etiology of achromatopsia in this family. The results from the exome sequencing were intriguing, as after filtering for common variations, a transition, c.1555 C>T, was seen in *CNNM4*, a gene associated with Jalili Syndrome, a disorder with achromatopsia or cone-rod dystrophy and defective tooth enamel (amelogenesis imperfecta). The transition mutation caused a premature stop at codon 519 (p.R519X) and prompted a review of archived medical records. These records revealed that in fact, these patients had teeth like that of Jalili Syndrome patients (brown discolouration, soft and weak enamel) which were extracted at a young age. This was not recognized at the time of ophthalmological examination in the 1980s as all the family members had dentures, a common ‘treatment’ for dental caries in areas of Newfoundland where access to dental services was scarce. Their ocular phenotype was also progressive, resembling that of cone-rod dystrophy, rather than achromatopsia. This portion of the study provided an interesting story, as well as highlighted the importance of exome sequencing technologies as a diagnostic tool. More interestingly still, if no archived records had been available, we could not have confirmed the dental phenotype, and would potentially have made a claim that mutation of *CNNM4* causes achromatopsia. This highlights the importance of combining these new technologies with old, accurate records, which can refine diagnoses and has allowed us to present the first cases of Jalili Syndrome in North America.

Chapter 4 examined, in my opinion, the most interesting of the three diseases, MDW. This apparently autosomal recessive condition showed great similarity to that of HSS, with microphthalmia, short stature, frontal bossing and hypotrichosis. The interesting thing about the presence of the phenotype in these two families is that the younger individuals show very little to no resemblance to the facial appearance of HSS described by Cohen in 1991,²²⁶ though one of the affected females from family 0066 who is currently age 60 shows very marked similarities to

HSS. This, along with the clinical description of the younger individuals from both families suggests that this is a form of HSS which progresses over time. Some cases of HSS are known to be caused by mutations in *GJAI*, and it would not be surprising if the mutation in this family is in a gene in the same pathway or process that *GJAI* is involved in. The studies of HH, and linkage showed great promise by using the 610K SNP Illumina SNP array by revealing a putative locus on Chromosome 16q12-21, which was then refined to 16q21 through addition of microsatellite markers, and sequencing of positional candidate genes. This now 3 cM region contains 71 genes, and a large gene desert of about 1-2 cM in size.

Surprisingly, exome sequencing of affected patients from both families revealed no pathogenic variants after filtering for rare variations within the chromosome 16 critical region. This result prompted sequencing of all *miRNAs*, *SNORA*, and *SNORD* genes within the region, which would not be covered by exome sequencing. Only two *SNORA* genes resided in the critical region (though a total of 10 from the larger linkage region were sequenced as well) and no mutations were found. Some of the genes in the critical region have functions in the immune system, skin conditions, retinal and eye conditions etc, which when taken together make sense as the causative genes for this condition. However, since no mutations were found in these genes, this leads me to believe that the mutation of interest is in fact in this large gene desert telomeric to the cluster of 71 genes (shown in Figure 4.9). A mutation in the large intergenic region could be that of an enhancer or repressor, which is skewing the expression and function of these genes. To fully evaluate this hypothesis, however, whole genome technologies would need to be employed to gather variant information from this region. Alternatively, a sequence capture of the 2 cM gene desert could be performed. The issue with this is again, the large amount of data, but it is likely that a large number of variants will be discovered in this region, and determining

which variant is having an effect is not an easy task. Another interesting assay would be to pick some of the genes in this region and to quantitatively check their expression against a normal unaffected individual using qPCR based approaches. This may provide some clue as to the mechanism of disease in these families.

Lastly, these projects highlight the usefulness of many types of analysis, and the use of a genetic isolate with large family sizes such as Newfoundland to identify many rare conditions. I used functional candidate gene, linkage, HH, and NGS approaches to solve a total of eight families (one ASD, seven achromatopsia) and determined one putative locus for a rare progressive form of HSS, MDW. These studies contain great importance to the medical management of these families, as gene therapy trials are showing great promise for diseases such as achromatopsia. By knowing the gene and mutation present in each family, these treatments can eventually be implemented, increasing the quality of life of these patients, and leading to a new era of medicine.

References

1. DelMonte, D.W., and Kim, T. (2011). Anatomy and physiology of the cornea. *J Cataract Refract Surg* 37, 588-598.
2. Levin, L.A., and Kaufman, P.L. (2011). *Adler's physiology of the eye : clinical application.*(Edinburgh ; New York: Saunders/Elsevier).
3. Jester, J.V., Moller-Pedersen, T., Huang, J., Sax, C.M., Kays, W.T., Cavangh, H.D., Petroll, W.M., and Piatigorsky, J. (1999). The cellular basis of corneal transparency: evidence for 'corneal crystallins'. *J Cell Sci* 112 (Pt 5), 613-622.
4. Srinivas, S.P. (2012). Cell signaling in regulation of the barrier integrity of the corneal endothelium. *Exp Eye Res* 95, 8-15.
5. Beebe, D.C., and Coats, J.M. (2000). The lens organizes the anterior segment: specification of neural crest cell differentiation in the avian eye. *Dev Biol* 220, 424-431.
6. Wolff, E., and Warwick, R. (1976). *Eugene Wolff's Anatomy of the eye and orbit : including the central connexions, development, and comparative anatomy of the visual apparatus.*(Philadelphia: Saunders).
7. Yanoff, M., Duker, J.S., and Augsburger, J.J. (2004). *Ophthalmology.*(St. Louis, MO: Mosby).
8. Sherwood, L. (2004). *Human physiology : from cells to systems.* In. (Australia ; Belmont, CA, Thomson/Brooks/Cole.
9. Yanoff, M., and Fine, B.S. (2002). *Ocular pathology.*(Philadelphia: Mosby).
10. Rong, P., Wang, X., Niesman, I., Wu, Y., Benedetti, L.E., Dunia, I., Levy, E., and Gong, X. (2002). Disruption of Gja8 (alpha8 connexin) in mice leads to microphthalmia associated with retardation of lens growth and lens fiber maturation. *Development* 129, 167-174.
11. Vaughan, D., Asbury, T., and Lange Medical Books/McGraw-Hill. (2011). *Vaughan & Asbury's general ophthalmology.* In. (New York, Lange Medical Books/McGraw-Hill), p v.190.
12. Mustafi, D., Engel, A.H., and Palczewski, K. (2009). Structure of cone photoreceptors. *Prog Retin Eye Res* 28, 289-302.
13. Bloomfield, S.A., and Dacheux, R.F. (2001). Rod vision: pathways and processing in the mammalian retina. *Prog Retin Eye Res* 20, 351-384.
14. Fatt, I., and Weissman, B.A. (1992). *Physiology of the eye : an introduction to the vegetative functions.*(Boston: Butterworth-Heinemann).
15. Besharse, J.C., and Pfenninger, K.H. (1980). Membrane assembly in retinal photoreceptors I. Freeze-fracture analysis of cytoplasmic vesicles in relationship to disc assembly. *J Cell Biol* 87, 451-463.
16. Rohlich, P. (1975). The sensory cilium of retinal rods is analogous to the transitional zone of motile cilia. *Cell Tissue Res* 161, 421-430.
17. Jastrzebska, B., Tsybovsky, Y., and Palczewski, K. (2010). Complexes between photoactivated rhodopsin and transducin: progress and questions. *Biochem J* 428, 1-10.
18. Poche, R.A., and Reese, B.E. (2009). Retinal horizontal cells: challenging paradigms of neural development and cancer biology. *Development* 136, 2141-2151.
19. Okawa, H., and Sampath, A.P. (2007). Optimization of single-photon response transmission at the rod-to-rod bipolar synapse. *Physiology (Bethesda)* 22, 279-286.
20. Newman, E.A. (1996). Acid efflux from retinal glial cells generated by sodium bicarbonate cotransport. *J Neurosci* 16, 159-168.
21. Franze, K., Grosche, J., Skatchkov, S.N., Schinkinger, S., Foja, C., Schild, D., Uckermann, O., Travis, K., Reichenbach, A., and Guck, J. (2007). Muller cells are living optical fibers in the vertebrate retina. *Proc Natl Acad Sci U S A* 104, 8287-8292.

22. Bernardos, R.L., Barthel, L.K., Meyers, J.R., and Raymond, P.A. (2007). Late-stage neuronal progenitors in the retina are radial Muller glia that function as retinal stem cells. *J Neurosci* 27, 7028-7040.
23. Fuhrmann, S. (2010). Eye morphogenesis and patterning of the optic vesicle. *Curr Top Dev Biol* 93, 61-84.
24. Ogino, H., and Yasuda, K. (2000). Sequential activation of transcription factors in lens induction. *Dev Growth Differ* 42, 437-448.
25. Gunhaga, L. (2011). The lens: a classical model of embryonic induction providing new insights into cell determination in early development. *Philos Trans R Soc Lond B Biol Sci* 366, 1193-1203.
26. Bassett, E.A., and Wallace, V.A. (2012). Cell fate determination in the vertebrate retina. *Trends Neurosci* 35, 565-573.
27. Cepko, C.L., Austin, C.P., Yang, X., Alexiades, M., and Ezzeddine, D. (1996). Cell fate determination in the vertebrate retina. *Proc Natl Acad Sci U S A* 93, 589-595.
28. Andreazzoli, M. (2009). Molecular regulation of vertebrate retina cell fate. *Birth Defects Res C Embryo Today* 87, 284-295.
29. Chavarria-Soley, G., Michels-Rautenstrauss, K., Caliebe, A., Kautza, M., Mardin, C., and Rautenstrauss, B. (2006). Novel CYP1B1 and known PAX6 mutations in anterior segment dysgenesis (ASD). *J Glaucoma* 15, 499-504.
30. Glaser, T., Jepeal, L., Edwards, J.G., Young, S.R., Favor, J., and Maas, R.L. (1994). PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet* 7, 463-471.
31. Hanson, I.M., Fletcher, J.M., Jordan, T., Brown, A., Taylor, D., Adams, R.J., Punnett, H.H., and van Heyningen, V. (1994). Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. *Nat Genet* 6, 168-173.
32. WHO. (2012). Visual impairment and blindness Fact Sheet. In. (
33. Frick, K.D., and Foster, A. (2003). The magnitude and cost of global blindness: an increasing problem that can be alleviated. *Am J Ophthalmol* 135, 471-476.
34. Jaeger, W. (1992). Horner's law. The first step in the history of the understanding of X-linked disorders. *Ophthalmic Paediatr Genet* 13, 49-56.
35. Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68, 820-823.
36. Kajiwara, K., Berson, E.L., and Dryja, T.P. (1994). Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. *Science* 264, 1604-1608.
37. Teichler Zallen, D. (2000). US gene therapy in crisis. *Trends Genet* 16, 272-275.
38. (2000). Controversy of the year. Biomedical ethics on the front burner. *Science* 290, 2225.
39. Koenekoop, R.K. (2008). Successful RPE65 gene replacement and improved visual function in humans. *Ophthalmic Genet* 29, 89-91.
40. Stone, E.M. (2007). Leber congenital amaurosis - a model for efficient genetic testing of heterogeneous disorders: LXIV Edward Jackson Memorial Lecture. *Am J Ophthalmol* 144, 791-811.
41. Chiang, P.W., Wang, J., Chen, Y., Fu, Q., Zhong, J., Yi, X., Wu, R., Gan, H., Shi, Y., Barnett, C., et al. (2012). Exome sequencing identifies NMNAT1 mutations as a cause of Leber congenital amaurosis. *Nat Genet* 44, 972-974.
42. Koenekoop, R.K. (2004). An overview of Leber congenital amaurosis: a model to understand human retinal development. *Surv Ophthalmol* 49, 379-398.
43. Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh, E.N., Jr., Mingozzi, F., Bennicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., et al. (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* 358, 2240-2248.

44. Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., et al. (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* 358, 2231-2239.
45. Liu, M.M., Tuo, J., and Chan, C.C. (2011). Republished review: Gene therapy for ocular diseases. *Postgrad Med J* 87, 487-495.
46. Komaromy, A.M., Alexander, J.J., Rowlan, J.S., Garcia, M.M., Chiodo, V.A., Kaya, A., Tanaka, J.C., Acland, G.M., Hauswirth, W.W., and Aguirre, G.D. (2010). Gene therapy rescues cone function in congenital achromatopsia. *Hum Mol Genet* 19, 2581-2593.
47. Pang, J.J., Deng, W.T., Dai, X., Lei, B., Everhart, D., Umino, Y., Li, J., Zhang, K., Mao, S., Boye, S.L., et al. (2012). AAV-mediated cone rescue in a naturally occurring mouse model of CNGA3-achromatopsia. *PLoS One* 7, e35250.
48. Carvalho, L.S., Xu, J., Pearson, R.A., Smith, A.J., Bainbridge, J.W., Morris, L.M., Fliesler, S.J., Ding, X.Q., and Ali, R.R. (2011). Long-term and age-dependent restoration of visual function in a mouse model of CNGB3-associated achromatopsia following gene therapy. *Hum Mol Genet* 20, 3161-3175.
49. Neuhausen, S.L. (2000). Founder populations and their uses for breast cancer genetics. *Breast Cancer Res* 2, 77-81.
50. Brody, J.A., Hussels, I., Brink, E., and Torres, J. (1970). Hereditary blindness among Pingelapese people of Eastern Caroline Islands. *Lancet* 1, 1253-1257.
51. Rahman, P., Jones, A., Curtis, J., Bartlett, S., Peddle, L., Fernandez, B.A., and Freimer, N.B. (2003). The Newfoundland population: a unique resource for genetic investigation of complex diseases. *Hum Mol Genet* 12 Spec No 2, R167-172.
52. Mannion, J.J., and Memorial University of Newfoundland. Institute of Social and Economic Research. (1977). *The Peopling of Newfoundland : essays in historical geography.*(St. John's, Nfld.: Institute of Social and Economic Research, Memorial University of Newfoundland).
53. Olufemi, S.E., Green, J.S., Manickam, P., Guru, S.C., Agarwal, S.K., Kester, M.B., Dong, Q., Burns, A.L., Spiegel, A.M., Marx, S.J., et al. (1998). Common ancestral mutation in the MEN1 gene is likely responsible for the prolactinoma variant of MEN1 (MEN1Burin) in four kindreds from Newfoundland. *Hum Mutat* 11, 264-269.
54. Spirio, L., Green, J., Robertson, J., Robertson, M., Otterud, B., Sheldon, J., Howse, E., Green, R., Groden, J., White, R., et al. (1999). The identical 5' splice-site acceptor mutation in five attenuated APC families from Newfoundland demonstrates a founder effect. *Hum Genet* 105, 388-398.
55. Kopciuk, K.A., Choi, Y.H., Parkhomenko, E., Parfrey, P., McLaughlin, J., Green, J., and Briollais, L. (2009). Penetrance of HNPCC-related cancers in a retrolective cohort of 12 large Newfoundland families carrying a MSH2 founder mutation: an evaluation using modified segregation models. *Hered Cancer Clin Pract* 7, 16.
56. Green, J.S., Parfrey, P.S., Harnett, J.D., Farid, N.R., Cramer, B.C., Johnson, G., Heath, O., McManamon, P.J., O'Leary, E., and Pryse-Phillips, W. (1989). The cardinal manifestations of Bardet-Biedl syndrome, a form of Laurence-Moon-Biedl syndrome. *N Engl J Med* 321, 1002-1009.
57. Cox, K.F., Kerr, N.C., Kedrov, M., Nishimura, D., Jennings, B.J., Stone, E.M., Sheffield, V.C., and Iannaccone, A. (2012). Phenotypic expression of Bardet-Biedl syndrome in patients homozygous for the common M390R mutation in the BBS1 gene. *Vision Res*.
58. Leppert, M., Baird, L., Anderson, K.L., Otterud, B., Lupski, J.R., and Lewis, R.A. (1994). Bardet-Biedl syndrome is linked to DNA markers on chromosome 11q and is genetically heterogeneous. *Nat Genet* 7, 108-112.

59. Young, T.L., Woods, M.O., Parfrey, P.S., Green, J.S., Hefferton, D., and Davidson, W.S. (1999). A founder effect in the newfoundland population reduces the Bardet-Biedl syndrome I (BBS1) interval to 1 cM. *Am J Hum Genet* 65, 1680-1687.
60. Mykytyn, K., Nishimura, D.Y., Searby, C.C., Shastri, M., Yen, H.J., Beck, J.S., Braun, T., Streb, L.M., Cornier, A.S., Cox, G.F., et al. (2002). Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. *Nat Genet* 31, 435-438.
61. Sheffield, V.C., Carmi, R., Kwitek-Black, A., Rokhlina, T., Nishimura, D., Duyk, G.M., Elbedour, K., Sunden, S.L., and Stone, E.M. (1994). Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. *Hum Mol Genet* 3, 1331-1335.
62. Fan, Y., Esmail, M.A., Ansley, S.J., Blacque, O.E., Boroevich, K., Ross, A.J., Moore, S.J., Badano, J.L., May-Simera, H., Compton, D.S., et al. (2004). Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome. *Nat Genet* 36, 989-993.
63. Young, T.L., Penney, L., Woods, M.O., Parfrey, P.S., Green, J.S., Hefferton, D., and Davidson, W.S. (1999). A fifth locus for Bardet-Biedl syndrome maps to chromosome 2q31. *Am J Hum Genet* 64, 900-904.
64. Katsanis, N., Beales, P.L., Woods, M.O., Lewis, R.A., Green, J.S., Parfrey, P.S., Ansley, S.J., Davidson, W.S., and Lupski, J.R. (2000). Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet-Biedl syndrome. *Nat Genet* 26, 67-70.
65. Heuser, A., Plovie, E.R., Ellinor, P.T., Grossmann, K.S., Shin, J.T., Wichter, T., Basson, C.T., Lerman, B.B., Sasse-Klaassen, S., Thierfelder, L., et al. (2006). Mutant desmocollin-2 causes arrhythmogenic right ventricular cardiomyopathy. *Am J Hum Genet* 79, 1081-1088.
66. Pilichou, K., Nava, A., Basso, C., Boffagna, G., Bauce, B., Lorenzon, A., Frigo, G., Vettori, A., Valente, M., Towbin, J., et al. (2006). Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy. *Circulation* 113, 1171-1179.
67. Sen-Chowdhry, S., Syrris, P., and McKenna, W.J. (2005). Genetics of right ventricular cardiomyopathy. *J Cardiovasc Electrophysiol* 16, 927-935.
68. Dokuparti, M.V., Pamuru, P.R., Thakkar, B., Tanjore, R.R., and Nallari, P. (2005). Etiopathogenesis of arrhythmogenic right ventricular cardiomyopathy. *J Hum Genet* 50, 375-381.
69. MacRae, C.A., Birchmeier, W., and Thierfelder, L. (2006). Arrhythmogenic right ventricular cardiomyopathy: moving toward mechanism. *J Clin Invest* 116, 1825-1828.
70. Ahmad, F., Li, D., Karibe, A., Gonzalez, O., Tapscott, T., Hill, R., Weilbaecher, D., Blackie, P., Furey, M., Gardner, M., et al. (1998). Localization of a gene responsible for arrhythmogenic right ventricular dysplasia to chromosome 3p23. *Circulation* 98, 2791-2795.
71. Thiene, G., Nava, A., Corrado, D., Rossi, L., and Pennelli, N. (1988). Right ventricular cardiomyopathy and sudden death in young people. *N Engl J Med* 318, 129-133.
72. Merner, N.D., Hodgkinson, K.A., Haywood, A.F., Connors, S., French, V.M., Drenckhahn, J.D., Kupprion, C., Ramadanova, K., Thierfelder, L., McKenna, W., et al. (2008). Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene. *Am J Hum Genet* 82, 809-821.
73. Young, T.L., Ives, E., Lynch, E., Person, R., Snook, S., MacLaren, L., Cater, T., Griffin, A., Fernandez, B., Lee, M.K., et al. (2001). Non-syndromic progressive hearing loss DFNA38 is caused by heterozygous missense mutation in the Wolfram syndrome gene WFS1. *Hum Mol Genet* 10, 2509-2514.
74. Leach, F.S., Nicolaidis, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215-1225.

75. Lafreniere, R.G., MacDonald, M.L., Dube, M.P., MacFarlane, J., O'Driscoll, M., Brais, B., Meilleur, S., Brinkman, R.R., Dadvivas, O., Pape, T., et al. (2004). Identification of a novel gene (HSN2) causing hereditary sensory and autonomic neuropathy type II through the Study of Canadian Genetic Isolates. *Am J Hum Genet* 74, 1064-1073.
76. Aksentijevich, I., Masters, S.L., Ferguson, P.J., Dancey, P., Frenkel, J., van Royen-Kerkhoff, A., Laxer, R., Tedgard, U., Cowen, E.W., Pham, T.H., et al. (2009). An autoinflammatory disease with deficiency of the interleukin-1-receptor antagonist. *N Engl J Med* 360, 2426-2437.
77. Kohl, S., and Hamel, C.P. (2011). Clinical utility gene card for: achromatopsia. *Eur J Hum Genet* 19.
78. Dimaras, H., Kimani, K., Dimba, E.A., Gronsdahl, P., White, A., Chan, H.S., and Gallie, B.L. (2012). Retinoblastoma. *Lancet* 379, 1436-1446.
79. Abouzeid, H., Schorderet, D.F., Balmer, A., and Munier, F.L. (2009). Germline mutations in retinoma patients: relevance to low-penetrance and low-expressivity molecular basis. *Mol Vis* 15, 771-777.
80. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res* 12, 996-1006.
81. Flicek, P., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., et al. (2012). Ensembl 2012. *Nucleic Acids Res* 40, D84-90.
82. Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-5467.
83. Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461, 272-276.
84. Christodoulou, K., Wiskin, A.E., Gibson, J., Tapper, W., Willis, C., Afzal, N.A., Upstill-Goddard, R., Holloway, J.W., Simpson, M.A., Beattie, R.M., et al. (2012). Next generation exome sequencing of paediatric inflammatory bowel disease patients identifies rare and novel variants in candidate genes. *Gut*.
85. Reis, L.M., and Semina, E.V. (2011). Genetics of anterior segment dysgenesis disorders. *Curr Opin Ophthalmol* 22, 314-324.
86. Peters, A. (1906). Ueber angeborene Defektbildung der Descemetischen Membran. *Klin Monatsbl Augenheilkd* 44, 27-40, 105-119.
87. Doward, W., Perveen, R., Lloyd, I.C., Ridgway, A.E., Wilson, L., and Black, G.C. (1999). A mutation in the RIEG1 gene associated with Peters' anomaly. *J Med Genet* 36, 152-155.
88. Hjalt, T.A., Amendt, B.A., and Murray, J.C. (2001). PITX2 regulates procollagen lysyl hydroxylase (PLOD) gene expression: implications for the pathology of Rieger syndrome. *J Cell Biol* 152, 545-552.
89. Nishimura, D.Y., Searby, C.C., Alward, W.L., Walton, D., Craig, J.E., Mackey, D.A., Kawase, K., Kanis, A.B., Patil, S.R., Stone, E.M., et al. (2001). A spectrum of FOXC1 mutations suggests gene dosage as a mechanism for developmental defects of the anterior chamber of the eye. *Am J Hum Genet* 68, 364-372.
90. Ormestad, M., Blixt, A., Churchill, A., Martinsson, T., Enerback, S., and Carlsson, P. (2002). Foxe3 haploinsufficiency in mice: a model for Peters' anomaly. *Invest Ophthalmol Vis Sci* 43, 1350-1357.
91. Semina, E.V., Brownell, I., Mintz-Hittner, H.A., Murray, J.C., and Jamrich, M. (2001). Mutations in the human forkhead transcription factor FOXE3 associated with anterior segment ocular dysgenesis and cataracts. *Hum Mol Genet* 10, 231-236.
92. Burdon, K.P., McKay, J.D., Wirth, M.G., Russell-Eggit, I.M., Bhatti, S., Ruddle, J.B., Dimasi, D., Mackey, D.A., and Craig, J.E. (2006). The PITX3 gene in posterior polar congenital cataract in Australia. *Mol Vis* 12, 367-371.

93. Semina, E.V., Ferrell, R.E., Mintz-Hittner, H.A., Bitoun, P., Alward, W.L., Reiter, R.S., Funkhauser, C., Daack-Hirsch, S., and Murray, J.C. (1998). A novel homeobox gene PITX3 is mutated in families with autosomal-dominant cataracts and ASMD. *Nat Genet* 19, 167-170.
94. Gu, F., Luo, W., Li, X., Wang, Z., Lu, S., Zhang, M., Zhao, B., Zhu, S., Feng, S., Yan, Y.B., et al. (2008). A novel mutation in AlphaA-crystallin (CRYAA) caused autosomal dominant congenital cataract in a large Chinese family. *Hum Mutat* 29, 769.
95. Devi, R.R., and Vijayalakshmi, P. (2006). Novel mutations in GJA8 associated with autosomal dominant congenital cataract and microcornea. *Mol Vis* 12, 190-195.
96. Sarfarazi, M., Akarsu, A.N., Hossain, A., Turacli, M.E., Aktan, S.G., Barsoum-Homsy, M., Chevrette, L., and Sayli, B.S. (1995). Assignment of a locus (GLC3A) for primary congenital glaucoma (Buphthalmos) to 2p21 and evidence for genetic heterogeneity. *Genomics* 30, 171-177.
97. Stoilov, I., Akarsu, A.N., and Sarfarazi, M. (1997). Identification of three different truncating mutations in cytochrome P4501B1 (CYP1B1) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the GLC3A locus on chromosome 2p21. *Hum Mol Genet* 6, 641-647.
98. Vincent, A., Billingsley, G., Priston, M., Williams-Lyn, D., Sutherland, J., Glaser, T., Oliver, E., Walter, M.A., Heathcote, G., Levin, A., et al. (2001). Phenotypic heterogeneity of CYP1B1: mutations in a patient with Peters' anomaly. *J Med Genet* 38, 324-326.
99. Lesnik Oberstein, S.A., Kriek, M., White, S.J., Kalf, M.E., Szuhai, K., den Dunnen, J.T., Breuning, M.H., and Hennekam, R.C. (2006). Peters Plus syndrome is caused by mutations in B3GALTL, a putative glycosyltransferase. *Am J Hum Genet* 79, 562-566.
100. Gehring, W.J. (1996). The master control gene for morphogenesis and evolution of the eye. *Genes Cells* 1, 11-15.
101. Semina, E.V., Reiter, R., Leysens, N.J., Alward, W.L., Small, K.W., Datson, N.A., Siegel-Bartelt, J., Bierke-Nelson, D., Bitoun, P., Zabel, B.U., et al. (1996). Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome. *Nat Genet* 14, 392-399.
102. Fitch, N., and Kaback, M. (1978). The Axenfeld syndrome and the Rieger syndrome. *J Med Genet* 15, 30-34.
103. Semina, E.V., Reiter, R.S., and Murray, J.C. (1997). Isolation of a new homeobox gene belonging to the Pitx/Rieg family: expression during lens development and mapping to the aphakia region on mouse chromosome 19. *Hum Mol Genet* 6, 2109-2116.
104. Berry, V., Francis, P.J., Prescott, Q., Waseem, N.H., Moore, A.T., and Bhattacharya, S.S. (2011). A novel 1-bp deletion in PITX3 causing congenital posterior polar cataract. *Mol Vis* 17, 1249-1253.
105. Pierrou, S., Hellqvist, M., Samuelsson, L., Enerback, S., and Carlsson, P. (1994). Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J* 13, 5002-5012.
106. Nishimura, D.Y., Swiderski, R.E., Alward, W.L., Searby, C.C., Patil, S.R., Bennet, S.R., Kanis, A.B., Gastier, J.M., Stone, E.M., and Sheffield, V.C. (1998). The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25. *Nat Genet* 19, 140-147.
107. Chakrabarti, S., Kaur, K., Rao, K.N., Mandal, A.K., Kaur, I., Parikh, R.S., and Thomas, R. (2009). The transcription factor gene FOXC1 exhibits a limited role in primary congenital glaucoma. *Invest Ophthalmol Vis Sci* 50, 75-83.
108. Weisschuh, N., Wolf, C., Wissinger, B., and Gramer, E. (2008). A novel mutation in the FOXC1 gene in a family with Axenfeld-Rieger syndrome and Peters' anomaly. *Clin Genet* 74, 476-480.
109. Honkanen, R.A., Nishimura, D.Y., Swiderski, R.E., Bennett, S.R., Hong, S., Kwon, Y.H., Stone, E.M., Sheffield, V.C., and Alward, W.L. (2003). A family with Axenfeld-Rieger syndrome and Peters

- Anomaly caused by a point mutation (Phe112Ser) in the FOXC1 gene. *Am J Ophthalmol* 135, 368-375.
110. Wang, J., Ray, P.S., Sim, M.S., Zhou, X.Z., Lu, K.P., Lee, A.V., Lin, X., Bagaria, S.P., Giuliano, A.E., and Cui, X. (2012). FOXC1 regulates the functions of human basal-like breast cancer cells by activating NF-kappaB signaling. *Oncogene*.
 111. Sizemore, S.T., and Keri, R.A. (2012). The Forkhead Box Transcription Factor FOXC1 Promotes Breast Cancer Invasion by Inducing Matrix Metalloprotease 7 (MMP7) Expression. *J Biol Chem*.
 112. Larsson, C., Hellqvist, M., Pierrou, S., White, I., Enerback, S., and Carlsson, P. (1995). Chromosomal localization of six human forkhead genes, freac-1 (FKHL5), -3 (FKHL7), -4 (FKHL8), -5 (FKHL9), -6 (FKHL10), and -8 (FKHL12). *Genomics* 30, 464-469.
 113. Blixt, A., Mahlapuu, M., Aitola, M., Pelto-Huikko, M., Enerback, S., and Carlsson, P. (2000). A forkhead gene, FoxE3, is essential for lens epithelial proliferation and closure of the lens vesicle. *Genes Dev* 14, 245-254.
 114. Valleix, S., Niel, F., Nedelec, B., Algros, M.P., Schwartz, C., Delbosc, B., Delpech, M., and Kantelip, B. (2006). Homozygous nonsense mutation in the FOXE3 gene as a cause of congenital primary aphakia in humans. *Am J Hum Genet* 79, 358-364.
 115. Anjum, I., Eiberg, H., Baig, S.M., Tommerup, N., and Hansen, L. (2010). A mutation in the FOXE3 gene causes congenital primary aphakia in an autosomal recessive consanguineous Pakistani family. *Mol Vis* 16, 549-555.
 116. Ali, M., Buentello-Volante, B., McKibbin, M., Rocha-Medina, J.A., Fernandez-Fuentes, N., Koga-Nakamura, W., Ashiq, A., Khan, K., Booth, A.P., Williams, G., et al. (2010). Homozygous FOXE3 mutations cause non-syndromic, bilateral, total sclerocornea, aphakia, microphthalmia and optic disc coloboma. *Mol Vis* 16, 1162-1168.
 117. Zhang, L., Zhang, Y., Liu, P., Cao, W., Tang, X., and Su, S. (2011). Congenital anterior polar cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. *Mol Vis* 17, 2693-2697.
 118. Sun, W., Xiao, X., Li, S., Guo, X., and Zhang, Q. (2011). Mutation analysis of 12 genes in Chinese families with congenital cataracts. *Mol Vis* 17, 2197-2206.
 119. Zhang, L.Y., Yam, G.H., Tam, P.O., Lai, R.Y., Lam, D.S., Pang, C.P., and Fan, D.S. (2009). An alphaA-crystallin gene mutation, Arg12Cys, causing inherited cataract-microcornea exhibits an altered heat-shock response. *Mol Vis* 15, 1127-1138.
 120. Dahm, R., van Marle, J., Prescott, A.R., and Quinlan, R.A. (1999). Gap junctions containing alpha8-connexin (MP70) in the adult mammalian lens epithelium suggests a re-evaluation of its role in the lens. *Exp Eye Res* 69, 45-56.
 121. Bassnett, S., and Beebe, D.C. (1992). Coincident loss of mitochondria and nuclei during lens fiber cell differentiation. *Dev Dyn* 194, 85-93.
 122. Goodenough, D.A. (1992). The crystalline lens. A system networked by gap junctional intercellular communication. *Semin Cell Biol* 3, 49-58.
 123. Mathias, R.T., Rae, J.L., and Baldo, G.J. (1997). Physiological properties of the normal lens. *Physiol Rev* 77, 21-50.
 124. Polyakov, A.V., Shagina, I.A., Khlebnikova, O.V., and Evgrafov, O.V. (2001). Mutation in the connexin 50 gene (GJA8) in a Russian family with zonular pulverulent cataract. *Clin Genet* 60, 476-478.
 125. Willoughby, C.E., Arab, S., Gandhi, R., Zeinali, S., Luk, D., Billingsley, G., Munier, F.L., and Heon, E. (2003). A novel GJA8 mutation in an Iranian family with progressive autosomal dominant congenital nuclear cataract. *J Med Genet* 40, e124.
 126. Hu, S., Wang, B., Zhou, Z., Zhou, G., Wang, J., Ma, X., and Qi, Y. (2010). A novel mutation in GJA8 causing congenital cataract-microcornea syndrome in a Chinese pedigree. *Mol Vis* 16, 1585-1592.

127. Sutter, T.R., Tang, Y.M., Hayes, C.L., Wo, Y.Y., Jabs, E.W., Li, X., Yin, H., Cody, C.W., and Greenlee, W.F. (1994). Complete cDNA sequence of a human dioxin-inducible mRNA identifies a new gene subfamily of cytochrome P450 that maps to chromosome 2. *J Biol Chem* 269, 13092-13099.
128. Bejjani, B.A., Lewis, R.A., Tomey, K.F., Anderson, K.L., Dueker, D.K., Jabak, M., Astle, W.F., Otterud, B., Leppert, M., and Lupski, J.R. (1998). Mutations in CYP1B1, the gene for cytochrome P4501B1, are the predominant cause of primary congenital glaucoma in Saudi Arabia. *Am J Hum Genet* 62, 325-333.
129. Belmouden, A., Melki, R., Hamdani, M., Zaghoul, K., Amraoui, A., Nadifi, S., Akhayat, O., and Garchon, H.J. (2002). A novel frameshift founder mutation in the cytochrome P450 1B1 (CYP1B1) gene is associated with primary congenital glaucoma in Morocco. *Clin Genet* 62, 334-339.
130. Vincent, A., Billingsley, G., Priston, M., Glaser, T., Oliver, E., Walter, M., Ritch, R., Levin, A., and Heon, E. (2006). Further support of the role of CYP1B1 in patients with Peters anomaly. *Mol Vis* 12, 506-510.
131. Heinonen, T.Y., Pasternack, L., Lindfors, K., Breton, C., Gastinel, L.N., Maki, M., and Kainulainen, H. (2003). A novel human glycosyltransferase: primary structure and characterization of the gene and transcripts. *Biochem Biophys Res Commun* 309, 166-174.
132. Hess, D., Keusch, J.J., Oberstein, S.A., Hennekam, R.C., and Hofsteenge, J. (2008). Peters Plus syndrome is a new congenital disorder of glycosylation and involves defective Omicron-glycosylation of thrombospondin type 1 repeats. *J Biol Chem* 283, 7354-7360.
133. Reis, L.M., Tyler, R.C., Abdul-Rahman, O., Trapane, P., Wallerstein, R., Broome, D., Hoffman, J., Khan, A., Paradiso, C., Ron, N., et al. (2008). Mutation analysis of B3GALTL in Peters Plus syndrome. *Am J Med Genet A* 146A, 2603-2610.
134. Dassie-Ajdid, J., Causse, A., Poidvin, A., Granier, M., Kaplan, J., Burglen, L., Doummar, D., Teisseire, P., Vigouroux, A., Malecaze, F., et al. (2009). Novel B3GALTL mutation in Peters-plus Syndrome. *Clin Genet* 76, 490-492.
135. Aliferis, K., Marsal, C., Pelletier, V., Doray, B., Weiss, M.M., Tops, C.M., Speeg-Schatz, C., Lesnik, S.A., and Dollfus, H. (2010). A novel nonsense B3GALTL mutation confirms Peters plus syndrome in a patient with multiple malformations and Peters anomaly. *Ophthalmic Genet* 31, 205-208.
136. Green, J.S., and Johnson, G.J. (1986). Congenital cataract with microcornea and Peters' anomaly as expressions of one autosomal dominant gene. *Ophthalmic Paediatr Genet* 7, 187-194.
137. Churchill, A.J., and Yeung, A. (2005). A compound heterozygous change found in Peters' anomaly. *Mol Vis* 11, 66-70.
138. Sowden, J.C. (2007). Molecular and developmental mechanisms of anterior segment dysgenesis. *Eye* 21, 1310-1318.
139. Berry, V., Yang, Z., Addison, P.K., Francis, P.J., Ionides, A., Karan, G., Jiang, L., Lin, W., Hu, J., Yang, R., et al. (2004). Recurrent 17 bp duplication in PITX3 is primarily associated with posterior polar cataract (CPP4). *J Med Genet* 41, e109.
140. Hansen, L., Yao, W., Eiberg, H., Kjaer, K.W., Baggesen, K., Hejtmanck, J.F., and Rosenberg, T. (2007). Genetic heterogeneity in microcornea-cataract: five novel mutations in CRYAA, CRYGD, and GJA8. *Invest Ophthalmol Vis Sci* 48, 3937-3944.
141. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16, 1215.
142. Woods, M.O., Hyde, A.J., Curtis, F.K., Stuckless, S., Green, J.S., Pollett, A.F., Robb, J.D., Green, R.C., Croitoru, M.E., Careen, A., et al. (2005). High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. *Clin Cancer Res* 11, 6853-6861.
143. Harissi-Dagher, M., and Colby, K. (2008). Anterior segment dysgenesis: Peters anomaly and sclerocornea. *Int Ophthalmol Clin* 48, 35-42.

144. Pang, S., Wang, W., Rich, B., David, R., Chang, Y.T., Carbutaru, G., Myers, S.E., Howie, A.F., Smillie, K.J., and Mason, J.I. (2002). A novel nonstop mutation in the stop codon and a novel missense mutation in the type II 3beta-hydroxysteroid dehydrogenase (3beta-HSD) gene causing, respectively, nonclassic and classic 3beta-HSD deficiency congenital adrenal hyperplasia. *J Clin Endocrinol Metab* 87, 2556-2563.
145. Ameri, A., Machiah, D.K., Tran, T.T., Channell, C., Crenshaw, V., Fernstrom, K., Khachidze, M., Duncan, A., Fuchs, S., and Howard, T.E. (2007). A nonstop mutation in the factor (F)X gene of a severely haemorrhagic patient with complete absence of coagulation FX. *Thromb Haemost* 98, 1165-1169.
146. Medina-Martinez, O., and Jamrich, M. (2007). Foxe view of lens development and disease. *Development* 134, 1455-1463.
147. Wigle, J.T., Chowdhury, K., Gruss, P., and Oliver, G. (1999). Prox1 function is crucial for mouse lens-fibre elongation. *Nat Genet* 21, 318-322.
148. Iseri, S.U., Osborne, R.J., Farrall, M., Wyatt, A.W., Mirza, G., Nurnberg, G., Kluck, C., Herbert, H., Martin, A., Hussain, M.S., et al. (2009). Seeing clearly: the dominant and recessive nature of FOXE3 in eye developmental anomalies. *Hum Mutat* 30, 1378-1386.
149. Bremond-Gignac, D., Bitoun, P., Reis, L.M., Copin, H., Murray, J.C., and Semina, E.V. (2010). Identification of dominant FOXE3 and PAX6 mutations in patients with congenital cataract and aniridia. *Mol Vis* 16, 1705-1711.
150. Pani, L., Overdier, D.G., Porcella, A., Qian, X., Lai, E., and Costa, R.H. (1992). Hepatocyte nuclear factor 3 beta contains two transcriptional activation domains, one of which is novel and conserved with the Drosophila fork head protein. *Mol Cell Biol* 12, 3723-3732.
151. Rojas, C.V., Maria, L.S., Santos, J.L., Cortes, F., and Allende, M.A. (2002). A frameshift insertion in the cone cyclic nucleotide gated cation channel causes complete achromatopsia in a consanguineous family from a rural isolate. *Eur J Hum Genet* 10, 638-642.
152. Genead, M.A., Fishman, G.A., Rha, J., Dubis, A.M., Bonci, D.M., Dubra, A., Stone, E.M., Neitz, M., and Carroll, J. (2011). Photoreceptor structure and function in patients with congenital achromatopsia. *Invest Ophthalmol Vis Sci* 52, 7298-7308.
153. Hussels, I.E., and Morton, N.E. (1972). Pingelap and Mokil Atolls: achromatopsia. *Am J Hum Genet* 24, 304-309.
154. Kohl, S., Baumann, B., Broghammer, M., Jagle, H., Sieving, P., Kellner, U., Spegal, R., Anastasi, M., Zrenner, E., Sharpe, L.T., et al. (2000). Mutations in the CNGB3 gene encoding the beta-subunit of the cone photoreceptor cGMP-gated channel are responsible for achromatopsia (ACHM3) linked to chromosome 8q21. *Hum Mol Genet* 9, 2107-2116.
155. Winick, J.D., Blundell, M.L., Galke, B.L., Salam, A.A., Leal, S.M., and Karayiorgou, M. (1999). Homozygosity mapping of the Achromatopsia locus in the Pingelapese. *Am J Hum Genet* 64, 1679-1685.
156. Milunsky, A., Huang, X.L., Milunsky, J., DeStefano, A., and Baldwin, C.T. (1999). A locus for autosomal recessive achromatopsia on human chromosome 8q. *Clin Genet* 56, 82-85.
157. Sundin, O.H., Yang, J.M., Li, Y., Zhu, D., Hurd, J.N., Mitchell, T.N., Silva, E.D., and Maumenee, I.H. (2000). Genetic basis of total colourblindness among the Pingelapese islanders. *Nat Genet* 25, 289-293.
158. Fishman, G.A., and Sokol, S. (1990). *Electrophysiologic testing in disorders of the retina, optic nerve, and visual pathway.* (San Francisco, CA: American Academy of Ophthalmology).
159. Brown, M., Marmor, M., Vaegan, Zrenner, E., Brigell, M., and Bach, M. (2006). *ISCEV Standard for Clinical Electro-oculography (EOG) 2006.* *Doc Ophthalmol* 113, 205-212.
160. Thiadens, A.A., Somervuo, V., van den Born, L.I., Roosing, S., van Schooneveld, M.J., Kuijpers, R.W., van Moll-Ramirez, N., Cremers, F.P., Hoyng, C.B., and Klaver, C.C. (2010). Progressive loss of

- cones in achromatopsia: an imaging study using spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci* 51, 5952-5957.
161. Wissinger, B., Gamer, D., Jagle, H., Giorda, R., Marx, T., Mayer, S., Tippmann, S., Broghammer, M., Jurklies, B., Rosenberg, T., et al. (2001). CNGA3 mutations in hereditary cone photoreceptor disorders. *Am J Hum Genet* 69, 722-737.
 162. Thiadens, A.A., Slingerland, N.W., Roosing, S., van Schooneveld, M.J., van Lith-Verhoeven, J.J., van Moll-Ramirez, N., van den Born, L.I., Hoyng, C.B., Cremers, F.P., and Klaver, C.C. (2009). Genetic etiology and clinical consequences of complete and incomplete achromatopsia. *Ophthalmology* 116, 1984-1989 e1981.
 163. Traboulsi, E.I. (2012). *Genetic diseases of the eye.*(New York: Oxford University Press).
 164. Kohl, S., Marx, T., Giddings, I., Jagle, H., Jacobson, S.G., Apfelstedt-Sylla, E., Zrenner, E., Sharpe, L.T., and Wissinger, B. (1998). Total colourblindness is caused by mutations in the gene encoding the alpha-subunit of the cone photoreceptor cGMP-gated cation channel. *Nat Genet* 19, 257-259.
 165. Aligianis, I.A., Forsheew, T., Johnson, S., Michaelides, M., Johnson, C.A., Trembath, R.C., Hunt, D.M., Moore, A.T., and Maher, E.R. (2002). Mapping of a novel locus for achromatopsia (ACHM4) to 1p and identification of a germline mutation in the alpha subunit of cone transducin (GNAT2). *J Med Genet* 39, 656-660.
 166. Chang, B., Grau, T., Dangel, S., Hurd, R., Jurklies, B., Sener, E.C., Andreasson, S., Dollfus, H., Baumann, B., Bolz, S., et al. (2009). A homologous genetic basis of the murine cpfl1 mutant and human achromatopsia linked to mutations in the PDE6C gene. *Proc Natl Acad Sci U S A* 106, 19581-19586.
 167. Kohl, S., Coppieters, F., Meire, F., Schaich, S., Roosing, S., Brennenstuhl, C., Bolz, S., van Genderen, M.M., Riemsdag, F.C., Lukowski, R., et al. (2012). A Nonsense Mutation in PDE6H Causes Autosomal-Recessive Incomplete Achromatopsia. *Am J Hum Genet* 91, 527-532.
 168. Arbour, N.C., Zlotogora, J., Knowlton, R.G., Merin, S., Rosenmann, A., Kanis, A.B., Rokhlina, T., Stone, E.M., and Sheffield, V.C. (1997). Homozygosity mapping of achromatopsia to chromosome 2 using DNA pooling. *Hum Mol Genet* 6, 689-694.
 169. Wissinger, B., Jagle, H., Kohl, S., Broghammer, M., Baumann, B., Hanna, D.B., Hedels, C., Apfelstedt-Sylla, E., Randazzo, G., Jacobson, S.G., et al. (1998). Human rod monochromacy: linkage analysis and mapping of a cone photoreceptor expressed candidate gene on chromosome 2q11. *Genomics* 51, 325-331.
 170. Dryja, T.P., Finn, J.T., Peng, Y.W., McGee, T.L., Berson, E.L., and Yau, K.W. (1995). Mutations in the gene encoding the alpha subunit of the rod cGMP-gated channel in autosomal recessive retinitis pigmentosa. *Proc Natl Acad Sci U S A* 92, 10177-10181.
 171. Peng, C., Rich, E.D., and Varnum, M.D. (2004). Subunit configuration of heteromeric cone cyclic nucleotide-gated channels. *Neuron* 42, 401-410.
 172. Wang, X., Wang, H., Cao, M., Li, Z., Chen, X., Patenia, C., Gore, A., Abboud, E.B., Al-Rajhi, A.A., Lewis, R.A., et al. (2011). Whole-exome sequencing identifies ALMS1, IQCB1, CNGA3, and MYO7A mutations in patients with Leber congenital amaurosis. *Hum Mutat* 32, 1450-1459.
 173. Wiszniewski, W., Lewis, R.A., and Lupski, J.R. (2007). Achromatopsia: the CNGB3 p.T383fsX mutation results from a founder effect and is responsible for the visual phenotype in the original report of uniparental disomy 14. *Hum Genet* 121, 433-439.
 174. Pentao, L., Lewis, R.A., Ledbetter, D.H., Patel, P.I., and Lupski, J.R. (1992). Maternal uniparental isodisomy of chromosome 14: association with autosomal recessive rod monochromacy. *Am J Hum Genet* 50, 690-699.
 175. Kohl, S., Varsanyi, B., Antunes, G.A., Baumann, B., Hoyng, C.B., Jagle, H., Rosenberg, T., Kellner, U., Lorenz, B., Salati, R., et al. (2005). CNGB3 mutations account for 50% of all cases with autosomal recessive achromatopsia. *Eur J Hum Genet* 13, 302-308.

176. Morris, T.A., and Fong, S.L. (1993). Characterization of the gene encoding human cone transducin alpha-subunit (GNAT2). *Genomics* 17, 442-448.
177. Piriev, N.I., Viczian, A.S., Ye, J., Kerner, B., Korenberg, J.R., and Farber, D.B. (1995). Gene structure and amino acid sequence of the human cone photoreceptor cGMP-phosphodiesterase alpha' subunit (PDEA2) and its chromosomal localization to 10q24. *Genomics* 28, 429-435.
178. Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31, 3812-3814.
179. Ramensky, V., Bork, P., and Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 30, 3894-3900.
180. Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S., et al. (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 31, 334-341.
181. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760.
182. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
183. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38, e164.
184. Thiadens, A.A., Roosing, S., Collin, R.W., van Moll-Ramirez, N., van Lith-Verhoeven, J.J., van Schooneveld, M.J., den Hollander, A.I., van den Born, L.I., Hoyng, C.B., Cremers, F.P., et al. (2010). Comprehensive analysis of the achromatopsia genes CNGA3 and CNGB3 in progressive cone dystrophy. *Ophthalmology* 117, 825-830 e821.
185. Lam, K., Guo, H., Wilson, G.A., Kohl, S., and Wong, F. (2011). Identification of variants in CNGA3 as cause for achromatopsia by exome sequencing of a single patient. *Arch Ophthalmol* 129, 1212-1217.
186. Hood, R.L., Lines, M.A., Nikkel, S.M., Schwartzentruber, J., Beaulieu, C., Nowaczyk, M.J., Allanson, J., Kim, C.A., Wieczorek, D., Moilanen, J.S., et al. (2012). Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. *Am J Hum Genet* 90, 308-313.
187. Polok, B., Escher, P., Ambresin, A., Chouery, E., Bolay, S., Meunier, I., Nan, F., Hamel, C., Munier, F.L., Thilo, B., et al. (2009). Mutations in CNNM4 cause recessive cone-rod dystrophy with amelogenesis imperfecta. *Am J Hum Genet* 84, 259-265.
188. Graham, S.V. (2003). Nonsense-mediated decay breaks the circle? *Biochem J* 373, e5-6.
189. Koeppen, K., Reuter, P., Kohl, S., Baumann, B., Ladewig, T., and Wissinger, B. (2008). Functional analysis of human CNGA3 mutations associated with colour blindness suggests impaired surface expression of channel mutants A3(R427C) and A3(R563C). *Eur J Neurosci* 27, 2391-2401.
190. Guo, D., Ling, J., Wang, M.H., She, J.X., Gu, J., and Wang, C.Y. (2005). Physical interaction and functional coupling between ACDP4 and the intracellular ion chaperone COX11, an implication of the role of ACDP4 in essential metal ion transport and homeostasis. *Mol Pain* 1, 15.
191. Parry, D.A., Mighell, A.J., El-Sayed, W., Shore, R.C., Jalili, I.K., Dollfus, H., Bloch-Zupan, A., Carlos, R., Carr, I.M., Downey, L.M., et al. (2009). Mutations in CNNM4 cause Jalili syndrome, consisting of autosomal-recessive cone-rod dystrophy and amelogenesis imperfecta. *Am J Hum Genet* 84, 266-273.
192. de Baaij, J.H., Stuver, M., Meij, I.C., Lainez, S., Kopplin, K., Venselaar, H., Mueller, D., Bindels, R.J., and Hoenderop, J.G. (2012). Membrane topology and intracellular processing of Cyclin M2 (CNNM2). *J Biol Chem*.

193. Downey, L.M., Keen, T.J., Jalili, I.K., McHale, J., Aldred, M.J., Robertson, S.P., Mighell, A., Fayle, S., Wissinger, B., and Inglehearn, C.F. (2002). Identification of a locus on chromosome 2q11 at which recessive amelogenesis imperfecta and cone-rod dystrophy cosegregate. *Eur J Hum Genet* 10, 865-869.
194. Jalili, I.K. (2010). Cone-rod dystrophy and amelogenesis imperfecta (Jalili syndrome): phenotypes and environs. *Eye (Lond)* 24, 1659-1668.
195. Michaelides, M., Bloch-Zupan, A., Holder, G.E., Hunt, D.M., and Moore, A.T. (2004). An autosomal recessive cone-rod dystrophy associated with amelogenesis imperfecta. *J Med Genet* 41, 468-473.
196. Zobor, D., Kaufmann, D.H., Weckerle, P., Sauer, A., Wissinger, B., Wilhelm, H., and Kohl, S. (2011). Cone-rod dystrophy associated with amelogenesis imperfecta in a child with neurofibromatosis type 1. *Ophthalmic Genet*.
197. Schilter, K.F., Schneider, A., Bardakjian, T., Soucy, J.F., Tyler, R.C., Reis, L.M., and Semina, E.V. (2011). OTX2 microphthalmia syndrome: four novel mutations and delineation of a phenotype. *Clin Genet* 79, 158-168.
198. Morrison, D., FitzPatrick, D., Hanson, I., Williamson, K., van Heyningen, V., Fleck, B., Jones, I., Chalmers, J., and Campbell, H. (2002). National study of microphthalmia, anophthalmia, and coloboma (MAC) in Scotland: investigation of genetic aetiology. *J Med Genet* 39, 16-22.
199. Ragge, N.K., Lorenz, B., Schneider, A., Bushby, K., de Sanctis, L., de Sanctis, U., Salt, A., Collin, J.R., Vivian, A.J., Free, S.L., et al. (2005). SOX2 anophthalmia syndrome. *Am J Med Genet A* 135, 1-7; discussion 8.
200. Van Raay, T.J., Moore, K.B., Iordanova, I., Steele, M., Jamrich, M., Harris, W.A., and Vetter, M.L. (2005). Frizzled 5 signaling governs the neural potential of progenitors in the developing *Xenopus* retina. *Neuron* 46, 23-36.
201. Mathers, P.H., Grinberg, A., Mahon, K.A., and Jamrich, M. (1997). The *Rx* homeobox gene is essential for vertebrate eye development. *Nature* 387, 603-607.
202. Danno, H., Michiue, T., Hitachi, K., Yukita, A., Ishiura, S., and Asashima, M. (2008). Molecular links among the causative genes for ocular malformation: *Otx2* and *Sox2* coregulate *Rax* expression. *Proc Natl Acad Sci U S A* 105, 5408-5413.
203. Voronina, V.A., Kozhemyakina, E.A., O'Kernick, C.M., Kahn, N.D., Wenger, S.L., Linberg, J.V., Schneider, A.S., and Mathers, P.H. (2004). Mutations in the human *RAX* homeobox gene in a patient with anophthalmia and sclerocornea. *Hum Mol Genet* 13, 315-322.
204. Ferda Percin, E., Ploder, L.A., Yu, J.J., Arici, K., Horsford, D.J., Rutherford, A., Bapat, B., Cox, D.W., Duncan, A.M., Kalnins, V.I., et al. (2000). Human microphthalmia associated with mutations in the retinal homeobox gene *CHX10*. *Nat Genet* 25, 397-401.
205. Dorval, K.M., Bobechko, B.P., Ahmad, K.F., and Bremner, R. (2005). Transcriptional activity of the paired-like homeodomain proteins *CHX10* and *VSX1*. *J Biol Chem* 280, 10100-10108.
206. Burmeister, M., Novak, J., Liang, M.Y., Basu, S., Ploder, L., Hawes, N.L., Vidgen, D., Hoover, F., Goldman, D., Kalnins, V.I., et al. (1996). Ocular retardation mouse caused by *Chx10* homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation. *Nat Genet* 12, 376-384.
207. Barabino, S.M., Spada, F., Cotelli, F., and Boncinelli, E. (1997). Inactivation of the zebrafish homologue of *Chx10* by antisense oligonucleotides causes eye malformations similar to the ocular retardation phenotype. *Mech Dev* 63, 133-143.
208. Nagase, T., Kikuno, R., Nakayama, M., Hirose, M., and Ohara, O. (2000). Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 7, 273-281.

209. Huynh, K.D., Fischle, W., Verdin, E., and Bardwell, V.J. (2000). BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev* 14, 1810-1823.
210. Ng, D., Thakker, N., Corcoran, C.M., Donnai, D., Perveen, R., Schneider, A., Hadley, D.W., Tiffet, C., Zhang, L., Wilkie, A.O., et al. (2004). Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. *Nat Genet* 36, 411-416.
211. Zentner, G.E., Hurd, E.A., Schnetz, M.P., Handoko, L., Wang, C., Wang, Z., Wei, C., Tesar, P.J., Hatzoglou, M., Martin, D.M., et al. (2010). CHD7 functions in the nucleolus as a positive regulator of ribosomal RNA biogenesis. *Hum Mol Genet* 19, 3491-3501.
212. Sanlaville, D., Etchevers, H.C., Gonzales, M., Martinovic, J., Clement-Ziza, M., Delezoide, A.L., Aubry, M.C., Pelet, A., Chemouny, S., Cruaud, C., et al. (2006). Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J Med Genet* 43, 211-217.
213. Chousou, P.A., Jennings, A., and Balmain, S. (2012). CHARGE syndrome - A rare combination of cardiac and endocrine disease. *Int J Cardiol*.
214. Hittner, H.M., Hirsch, N.J., Kreh, G.M., and Rudolph, A.J. (1979). Colobomatous microphthalmia, heart disease, hearing loss, and mental retardation--a syndrome. *J Pediatr Ophthalmol Strabismus* 16, 122-128.
215. Bakrania, P., Efthymiou, M., Klein, J.C., Salt, A., Bunyan, D.J., Wyatt, A., Ponting, C.P., Martin, A., Williams, S., Lindley, V., et al. (2008). Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between the BMP4 and hedgehog signaling pathways. *Am J Hum Genet* 82, 304-319.
216. Connor, J.M. (1996). Fibrodysplasia ossificans progressiva -- lessons from rare maladies. *N Engl J Med* 335, 591-593.
217. Reis, L.M., Tyler, R.C., Schilter, K.F., Abdul-Rahman, O., Innis, J.W., Kozel, B.A., Schneider, A.S., Bardakjian, T.M., Lose, E.J., Martin, D.M., et al. (2011). BMP4 loss-of-function mutations in developmental eye disorders including SHORT syndrome. *Hum Genet* 130, 495-504.
218. Kawaguchi, R., Yu, J., Honda, J., Hu, J., Whitelegge, J., Ping, P., Wiita, P., Bok, D., and Sun, H. (2007). A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A. *Science* 315, 820-825.
219. Bouillet, P., Sapin, V., Chazaud, C., Messaddeq, N., Decimo, D., Dolle, P., and Chambon, P. (1997). Developmental expression pattern of Stra6, a retinoic acid-responsive gene encoding a new type of membrane protein. *Mech Dev* 63, 173-186.
220. Vannahme, C., Smyth, N., Miosge, N., Gosling, S., Frie, C., Paulsson, M., Maurer, P., and Hartmann, U. (2002). Characterization of SMOC-1, a novel modular calcium-binding protein in basement membranes. *J Biol Chem* 277, 37977-37986.
221. Abouzeid, H., Boisset, G., Favez, T., Youssef, M., Marzouk, I., Shakankiry, N., Bayoumi, N., Descombes, P., Agosti, C., Munier, F.L., et al. (2011). Mutations in the SPARC-related modular calcium-binding protein 1 gene, SMOC1, cause waardenburg anophthalmia syndrome. *Am J Hum Genet* 88, 92-98.
222. Fuse, N., Yasumoto, K., Takeda, K., Amae, S., Yoshizawa, M., Uono, T., Takahashi, K., Tamai, M., Tomita, Y., Tachibana, M., et al. (1999). Molecular cloning of cDNA encoding a novel microphthalmia-associated transcription factor isoform with a distinct amino-terminus. *J Biochem* 126, 1043-1051.
223. Tachibana, M., Takeda, K., Nobukuni, Y., Urabe, K., Long, J.E., Meyers, K.A., Aaronson, S.A., and Miki, T. (1996). Ectopic expression of MITF, a gene for Waardenburg syndrome type 2, converts fibroblasts to cells with melanocyte characteristics. *Nat Genet* 14, 50-54.
224. Bennett, C.P., Betts, D.R., and Seller, M.J. (1991). Deletion 14q (q22q23) associated with anophthalmia, absent pituitary, and other abnormalities. *J Med Genet* 28, 280-281.

225. Gallardo, M.E., Rodriguez De Cordoba, S., Schneider, A.S., Dwyer, M.A., Ayuso, C., and Bovolenta, P. (2004). Analysis of the developmental SIX6 homeobox gene in patients with anophthalmia/microphthalmia. *Am J Med Genet A* 129A, 92-94.
226. Cohen, M.M., Jr. (1991). Hallermann-Streiff syndrome: a review. *Am J Med Genet* 41, 488-499.
227. Friede, H., Lopata, M., Fisher, E., and Rosenthal, I.M. (1985). Cardiorespiratory disease associated with Hallermann-Streiff syndrome: analysis of craniofacial morphology by cephalometric roentgenograms. *J Craniofac Genet Dev Biol Suppl* 1, 189-198.
228. Stone, W.C. (1975). Hallermann-Streiff syndrome and progeria. *South Med J* 68, 1139-1141.
229. Tuna, E.B., Sulun, T., Rosti, O., El Abdallah, F., Kayserili, H., and Aktoren, O. (2009). Craniofacial manifestations in Hallermann-Streiff syndrome. *Cranio* 27, 33-38.
230. Robinow, M. (1991). Respiratory obstruction and cor pulmonale in the Hallermann-Streiff syndrome. *Am J Med Genet* 41, 515-516.
231. Paznekas, W.A., Karczeski, B., Vermeer, S., Lowry, R.B., Delatycki, M., Laurence, F., Koivisto, P.A., Van Maldergem, L., Boyadjiev, S.A., Bodurtha, J.N., et al. (2009). GJA1 mutations, variants, and connexin 43 dysfunction as it relates to the oculodentodigital dysplasia phenotype. *Hum Mutat* 30, 724-733.
232. Pizzuti, A., Flex, E., Mingarelli, R., Salpietro, C., Zelante, L., and Dallapiccola, B. (2004). A homozygous GJA1 gene mutation causes a Hallermann-Streiff/ODDD spectrum phenotype. *Hum Mutat* 23, 286.
233. Miyazawa, H., Kato, M., Awata, T., Kohda, M., Iwasa, H., Koyama, N., Tanaka, T., Huqun, Kyo, S., Okazaki, Y., et al. (2007). Homozygosity haplotype allows a genomewide search for the autosomal segments shared among patients. *Am J Hum Genet* 80, 1090-1102.
234. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4, 1073-1081.
235. Kiss, A.M., Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol* 24, 5797-5807.
236. Bardakjian, T.M., Schneider, A.S., Ng, D., Johnston, J.J., and Biesecker, L.G. (2009). Association of a de novo 16q copy number variant with a phenotype that overlaps with Lenz microphthalmia and Townes-Brocks syndromes. *BMC Med Genet* 10, 137.
237. Wasserman, N.F., Aneas, I., and Nobrega, M.A. (2010). An 8q24 gene desert variant associated with prostate cancer risk confers differential in vivo activity to a MYC enhancer. *Genome Res* 20, 1191-1197.
238. Tanaka, A., Morice-Picard, F., Lacombe, D., Nagy, N., Hide, M., Taieb, A., and McGrath, J. (2010). Identification of a homozygous deletion mutation in C16orf57 in a family with Clericuzio-type poikiloderma with neutropenia. *Am J Med Genet A* 152A, 1347-1348.
239. Kuehn, B.M. (2012). Massive study compilation illuminates regulatory role of non-gene-encoding DNA. *JAMA* 308, 1419-1420.
240. Lunyak, V.V., and Atallah, M. (2011). Genomic relationship between SINE retrotransposons, Pol III-Pol II transcription, and chromatin organization: the journey from junk to jewel. *Biochem Cell Biol* 89, 495-504.
241. Saito, Y., and Saito, H. (2012). MicroRNAs in cancers and neurodegenerative disorders. *Front Genet* 3, 194.

Appendix A: Polymerase Chain Reaction (PCR) setup protocols
1x Polymerase Chain Reaction (PCR) Protocol w/ 25% Betaine

<u>Reagent</u>	<u>Volume (uL)</u>
dH ₂ O	9.92
KapaTaq Buffer	2
25mM Betaine	5
KapaTaq dNTPs	0.4
Forward Primer (10uM)	0.8
Reverse Primer	0.8
KapaTaq Taq Pol.	0.08
<u>10 ng/uL DNA Sample</u>	<u>1</u>
Total Volume	20

1x Polymerase Chain Reaction (PCR) Protocol w/ 25% Betaine and 5% Dimethylsulfoxide (DMSO)

<u>Reagent</u>	<u>Volume (uL)</u>
dH ₂ O	8.92
DMSO	1
KapaTaq Buffer	2
25mM Betaine	5
KapaTaq dNTPs	0.4
Forward Primer	0.8
Reverse Primer	0.8
KapaTaq Taq Pol.	0.08
<u>10 ng/uL DNA Sample</u>	<u>1</u>
Total Volume	20

Appendix B: TouchDown (TD) Thermocycling Program

94°C for 5 mins

5 cycles of:

94°C Denaturation for 30s

64°C Annealing minus 2°C per cycle for 30s

72°C Extension for 30s

30 cycles of:

94°C Denaturation for 30s

54°C Annealing for 30s

72°C Extension for 30s

72°C Final Extension for 7 mins

Hold at 4°

Appendix C: 1x ABI BigDye Terminator (BDT) 3.1 Kit Cocktail for Sanger Sequencing

<u>Reagent</u>	<u>Volume (uL)</u>
dH ₂ O	15.68
BDT 5x Buffer	2
BDT Sequencing Mix	0.5
Primer (10uM)	0.32
<u>Purified PCR Sample</u>	<u>1</u>
Total Volume	20

Appendix D: Thermocycling Protocols for Cycle Sequencing (ABI.SEQ)

94°C for 1 min

25 cycles of:

96°C Denaturation for 10s

50°C Annealing for 5s

60°C Extension for 4 mins

Hold at 4°

Appendix E: Cycle Sequencing DNA Precipitation Protocol

This step was performed after Cycle sequencing was completed, as a second purification step before being placed on the ABI 3130xl or ABI 3730.

Step 1) DNA precipitation

- Add 65uL of 95% Ethanol (EtOH) to each well
- Add 5uL of 125mM Ethylenediaminetetraacetic acid to each well
- Let precipitate for 15 mins to overnight in dark
 - Can place at -20°C if preferred or if not using plate for a few days.

Step 2) Ethanol mixture Removal

- Place plate in centrifuge, spin at 3000 RPM for 30 mins
- Remove plate from centrifuge and decant ethanol mixture onto a dry paper towel by inverting the plate.
- Leave plate inverted on paper towel and place back in centrifuge. Spin up to 200 RPM, and then immediately stop the spinning.
- Remove and discard paper towel.

Step 3) Rinse step

- Add 150uL of 70% EtOH to each well, and place in centrifuge.
- Spin plate at 3000 RPM for 5 mins.
- Remove plate from centrifuge and decant EtOH mixture onto a paper towel
- Leave plate inverted on paper towel and place back in centrifuge. Spin up to 200 RPM, and then immediately stop the spinning.

- Remove and discard paper towel. Let plate dry in dark and uncovered for 20 mins

Step 4) Sample Resuspension

- Add 15 uL of Hi-Dye Formamide (HDF) to each well
- Place plate in thermocycler on 'denat' program
 - 95°C for 2 mins
 - Hold at 4°C

Appendix F: Sequences and Tms for all Primers used for PCR and bidirectional sequencing in the Anterior Segment Dysgenesis project.

PAX6

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	ggcgtgctttgcataaa	61	249	No
	Reverse	tggtgatggctcaagtgtgt	60.2		
2	Forward	ttcctcaccacactcttt	59.5	418	No
	Reverse	ctcctcgtggaaacttctc	60		
3	Forward	agtgggatccgaacttgcta	59.7	576	No
	Reverse	tgagggtcactcagctctgc	61.8		
4	Forward	gaggttgagtggatcaattcct	59.4	388	No
	Reverse	cagtatcgagaagagccaagc	59.2		
5	Forward	ttgtcctttatgtatcgatagca	60	466	No
	Reverse	gggtccataattagcatcgtttac	60		
6	Forward	cactttaagcaaggtcagcaca	60.5	468	No
	Reverse	tcgctactctcggtttactacca	60.3		
7	Forward	aaagtccaagtctggacaatc	60.5	399	No
	Reverse	aggtaaagaggagagagcattgg	60.3		
8	Forward	gagatgggtgactgtgtcttca	60.2	310	No
	Reverse	agaggaaatggttgggagagtag	60		
9	Forward	aagaaggctgacagttaccttg	61	392	No
	Reverse	caaagggcctggctaaat	61.3		
10	Forward	gtgggaaagtcttccaagtacag	60.4	379	No
	Reverse	cagagcatttagcagactgaacc	60.4		
11	Forward	tttcttagagacagaggtgcttg	60.1	447	No
	Reverse	cagatgtgaaggaggaaactgag	60.3		
12	Forward	cagtgctaccaaccaattccac	60.7	405	No
	Reverse	gattgactgtctccgactgact	59.8		
13	Forward	cataggcagctttcttagctg	59.8	442	No
	Reverse	cccataagaccaggagattctgt	60.7		
14	Forward	ttccatgtctgtttctcaaagg	59.2	310	No
	Reverse	aagtccattccttccccagt	59.8		

PITX2

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	gttggcgcagattgcttt	60	734	No
	Reverse	aagccacaaatcacctacgg	60		
2	Forward	ccgcttcttacagccttct	60.9	247	No
	Reverse	ctggcgatttggttctgatt	60.1		

3	Forward	cagctcttcacggcttct	60.7	387	No
	Reverse	gctgccttcacattctctc	60		
4-1	Forward	gaggccagggtgtgtgag	60.3	670	No
	Reverse	ctgcatactggcaagcactc	59.6		
4-2	Forward	cagcctgagactgaaagcaa	59.3	572	No
	Reverse	ctcgcaagcgaaaaatccta	60.5		
4-3	Forward	aacaaatgtttgactggatatgaca	59.7	489	No
	Reverse	aggaggggagaaagaatcca	60		

PITX3

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	tactccctccgagagtcc	60.3	236	No
	Reverse	tttagggattccaaggtcca	59.34		
2	Forward	gggaatttacgaggaaacgc	60.8	327	No
	Reverse	gacctctaagccactcgc	60		
3	Forward	agcgagtggcttaggaggtc	60.9	466	No
	Reverse	agtgggagcagaggctgg	61.6		
4-1	Forward	tctcgtttattgaccgagg	61.2	612	No
	Reverse	cgtgctgtttggctttgag	60.6		
4-2	Forward	cgtgtcctgccttatgc	60.2	601	No
	Reverse	cgctgtagacagtgggg	60.3		

CYP1B1

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	ccctccttctaccagtcct	59.6	569	No
	Reverse	cgtaacggttctgcaatct	60.1		
2-1	Forward	tctccagagagtcagctccg	60.8	786	No
	Reverse	gggtcgtcgtggctgtag	60.3		
2-2	Forward	atggctttcgccactac	58.7	359	No
	Reverse	tatggagcacacctcacctg	59.7		
3-1	Forward	catgcaaggcctattacagga	60.1	588	No
	Reverse	gcctttgcccactgaaaa	61.1		
3-2	Forward	gaccagtgaaagtggcctaa	60.1	493	No
	Reverse	cacacctcacctgatggaca	60.6		

3-3	Forward	atgagcctgcgaaaatgaat	52.8	426	No
	Reverse	cacacctcacctgatggaca	60.6		
3-4	Forward	cagttcctcttttgctgct	59.7	318	No
	Reverse	tcttgattcccacaaaaa	60.3		
3-5	Forward	caccaaacttacacaaa	58.4	333	No
	Reverse	tgagcaactgactttatgattatcc	59.9		
3-6	Forward	tgattgaaggatgataaggaaa	58.6	362	No
	Reverse	tgaaatcagtacattatggtcaca	57.1		
3-7	Forward	ttgcttgccaaagtacagaa	58.6	321	No
	Reverse	gaaaccctgcttcattcca	60.1		
3-8	Forward	ccagccaagctttaattatgtg	60	512	No
	Reverse	tctgccttccacaggagaat	59.8		
3-9	Forward	tggaaatgaagcagggttc	60.1	469	No
	Reverse	ttccaacagctccaaga	58.5		
3-10	Forward	gggacagaactcccattacaa	58.9	345	No
	Reverse	cccacacacacatacacacaa	58.7		
3-11	Forward	ttggtagctgggaaagcat	59.7	552	No
	Reverse	ggcaagcctgctttgtgtag	61		
3-12	Forward	tcaaagcatggacatttagaaga	58.9	371	No
	Reverse	aggcatccagattggttcat	59.4		
3-13	Forward	caggcttcccagtacattt	60.1	372	No
	Reverse	tccagcctccaaattcagtt	59.7		
3-14	Forward	atgtaggcccagtcgtcatc	60	420	No
	Reverse	ggctaagttctgggacatgaa	59.2		
3-15	Forward	ccggagagagaatgtatttgc	58.8	441	No
	Reverse	tgactgtagctattatgcacaca	60.7		
3-16	Forward	ctgtgacacaactgtgtattaa	57.8	254	No
	Reverse	tttgtaatggtgtcccagt	57.8		

FOXC1

Exon	Direction	Sequence	Tm	Size	DMSO
1-1	Forward	cccgactcggactcggc	68.8	428	Yes
	Reverse	aagcgtccatgatgaactgg	64.5		
1-2	Forward	ccaaggacatggtgaagc	61.5	710	Yes
	Reverse	ctgaagccctggctatggt	60.2		
1-3	Forward	atcaagaccgagaacggtacg	61.4	635	Yes

	Reverse	gtgaccggaggcagagagta	60.4		
1-4	Forward	taccactgcaacctgcaagc	62.4	517	Yes
	Reverse	gggttcgatttagttggct	60.5		

CRYAA

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	ccttaatgcctccattctgc	59.7	395	No
	Reverse	cagagtccatcgctctccac	61		
2	Forward	acatggcacgtttgatttc	60.8	176	No
	Reverse	ccctctcccacctctcagt	59.2		
3_1	Forward	gcagcttctctggcatgg	60.7	437	No
	Reverse	TAAGCTCTCCTGGCTGCTCT	59.5		
3_2	Forward	GGCGGGGTGTCTGTCTTC	61.7	486	No
	Reverse	agacagcaccagcagacg	58.6		

B3GALTL

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	gtcagacgcctcgaagga	60.1	621	No
	Reverse	ggaccaagaccgaaagg	60.4		
2	Forward	cagcaatgcagctgcttaaa	60.3	406	No
	Reverse	aaccaagtggatcagccttaaa	60		
3	Forward	aatcttctcccatgtgctg	60.1	289	No
	Reverse	gctgggccaagacagaata	60.2		
4	Forward	ttgtcttagaattacatacgaattga	57.7	383	No
	Reverse	tgagtcagaggagtattctcgt	57.2		
5	Forward	ggtgcagagtgtagaggttc	59.9	538	No
	Reverse	agagcgagactctgtctccaa	59.3		
6	Forward	cctgaactagtgttattatggaga	57.5	537	No
	Reverse	tggctcattataagctctgtcc	58.5		
7	Forward	ccaccttctattggtcttacatga	59.4	332	No
	Reverse	ttaccacacactgcctctgg	59.7		
8	Forward	tgtgttagtcttgctgacacttc	58.6	281	No
	Reverse	aagcacctaattcgctggag	59.5		
9	Forward	tggttcagcctactctctatgg	58.5	357	No
	Reverse	accaactgcacatgaatgc	59.6		
10	Forward	agtgtagcaggcaggtccag	60.5	405	No
	Reverse	tcacttcaacatccgaattgtc	60		

11	Forward	ggaagcaagcaactcttggga	60.5	404	No
	Reverse	gaaccaatcagtatctgagagaagg	59.7		
12	Forward	ttggcatgtaagctcttagaaca	59.1	349	No
	Reverse	tggtccttaagaggattgggtca	60		
13	Forward	ctgccactctcaccacacaa	60.9	537	No
	Reverse	ccattcacagagtactcacgtacag	60.2		
14	Forward	gctactattgcctctcagtctgc	59.7	385	No
	Reverse	tggagcctgtcaacacagtt	59.3		
15	Forward	tcctatagccaatgttaggtagtgaa	59.5	563	No
	Reverse	ctgctgtatcacagcttgag	58.7		
15.1	Forward	tgctgtgctcacaacacttg	59.6	388	No
	Reverse	ccattcacaatatccatgtcaa	58.2		
15.2	Forward	ttcctgttctgtctcttcattg	57.1	353	No
	Reverse	gcaatacgaagctctagcagaa	58.9		
15.3	Forward	tgaagtctgtaatcatggtggtt	58.5	379	No
	Reverse	tgatcaccatgctgtccttc	59.6		
15.4	Forward	gaggcttctgctagagcttcg	59.9	720	No
	Reverse	actgcaacacgtgaggctcg	59.9		
15.5	Forward	ggacagcatggtgatcagg	60.1	427	No
	Reverse	actgcaacacgtgaggctcg	59.9		
15.6	Forward	cagacctcacgtgttgagct	59.9	347	No
	Reverse	gccagcttcaatcatgtcaa	59.8		
15.7	Forward	gctgcctacaccagtgga	59.8	563	No
	Reverse	aagacaatatctggtcctggaa	57.7		
15.8	Forward	gctggcatccgtatatgaaga	60.1	469	No
	Reverse	aagacaatatctggtcctggaa	57.7		
15.9	Forward	tgggtgaagcctaaccacaag	57.8	686	No
	Reverse	gaggatatatacaacaagaggtctca	58.3		
15.1	Forward	tgtgtggaataactggaatgtca	60.3	326	No
	Reverse	catctctgattatgaatggaagaca	59.5		
15.11	Forward	tctagaatgcctaggagcagaa	58.4	307	No
	Reverse	tggatgcctggagaatatca	59		

CRYAA

Exon	Direction	Sequence	Tm	Size	DMSO
1.1	Forward	gttctggcaacttggaagg	59.7	547	No
	Reverse	cagccggaacttcttagtgc	60		
1.2	Forward	atggaggagaagcgcaaa	59.5	563	No

	Reverse	gtctccaccatcccaacct	59.8		
1_3	Forward	agaaagccaagggtatcag	58.6	599	No
	Reverse	cagatcatgttgccacctt	58.6		

FOXE3

Exon	Direction	Sequence	Tm	Size	DMSO
1-1	Forward	ttgggaatgatccaaaggag	59.86	584	Yes
	Reverse	tgagcgcatgtacgagtag	60.18		
1-2	Forward	gggctgggagaggaaattag	60.03	819	Yes
	Reverse	aggctgtcgacgctgaac	60.15		
1-3	Forward	ccatctaccgcttcatcacc	60.48	682	Yes
	Reverse	gcgaggctcacaggtag	60.73		
1-4	Forward	cgacagcctggtgaacct	59.82	559	Yes
	Reverse	gctcacaagtggaggagaa	60.39		
1-5	Forward	cctggactcaaactcctgct	59.45	559	Yes
	Reverse	acctaaaggcccaccaactc	60.36		
1-6	Forward	ctgccctcagcttcttct	59.69	300	Yes
	Reverse	aataaagctcaccgctctcc	58.27		

Appendix G: cDNA Synthesis Protocol.

1. Mix and briefly centrifuge each component before use.

2. Combine the following in a 0.2- or 0.5-ml tube:

Component	Amount
RNA (up to 5ug)	n uL
50uM Oligo dT primers	1uL
10 mM dNTP mix	1uL
DEPC-Treated Water	up to 10uL

3. Incubate at 65°C for 5 min, then place on ice for at least 1 min.

4. Prepare the following cDNA Synthesis Mix, adding each component in the indicated order:

Component	1 Rxn
10X RT buffer	2 µl
25 mM MgCl ₂	4 µl
0.1 M DTT	2 µl
RNaseOUT (40 U/µl)	1 µl
SuperScript III RT (200 U/µl)	1 µl

5. Add 10 µl of cDNA Synthesis Mix to each RNA/primer mixture, mix gently, and collect by brief centrifugation. Incubate for 50 min at 50°C

6. Terminate the reactions at 85°C for 5 min. Chill on ice.

7. Collect the reactions by brief centrifugation. Add 1 µl of RNase H to each tube and incubate for 20 min at 37°C.

8. cDNA synthesis reaction can be stored at -20°C or used for PCR immediately.

Appendix H: Sequences and Tms for all Primers used for PCR and bidirectional sequencing in the achromatopsia project.

CNGA3

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	cttcctaagatggcaaac	54.9	576	No
	Reverse	ctgcgtgtggttctctaact	55.6		
2	Forward	agagacagcagagggtgtgc	60.6	385	No
	Reverse	ctgccctctgacctgtgg	60.4		
3	Forward	ggtggctttccctgctaag	59.8	300	No
	Reverse	ttttcctgagaggtcactacctt	59.7		
4	Forward	cagacagagagggaggaga	59.5	374	No
	Reverse	gtgctgcatccaacaggat	61.1		
5	Forward	gaggacctgttggacag	60.6	384	No
	Reverse	ggagaaaccgaggacaata	60.1		
6	Forward	agggccattccatattaca	59.1	300	No
	Reverse	atccaccatgctgggtctc	60.9		
7	Forward	tcagagtgcatttctgtag	53.9	577	No
	Reverse	caaggtggaccagtagagac	54.7		
8a	Forward	cttgaccgcacagagac	55.1	500	No
	Reverse	tctcatccaccgtcttctt	55.6		
8b	Forward	ggagtatctctttgtggtcgtaga	58.9	600	No
	Reverse	aggctgtagtagccaatgctg	59.5		
8c	Forward	ctggggattatatctgcaag	54.9	558	No
	Reverse	cctctgtttttagcatcc	55		
8d	Forward	aaggacctgaggagaaagt	54.6	552	No
	Reverse	gcctaaaaacacctgcag	56.1		
8e	Forward	gtgtgactgctgagagaac	55.8	590	No
	Reverse	aggggatgtctcttctccta	55.4		
8f	Forward	ggtatcacctcgtgtgttct	55	512	No
	Reverse	ctgcaatctgtgagctttc	54.5		
8g	Forward	gagtcaggcccttaggtga	60.2	500	No
	Reverse	ttttgaggaggaaaagaggaga	59.5		
8h	Forward	agaggaagcctcaagatcc	55.9	500	No
	Reverse	cctatcctttctaggaagc	54.8		

CNGB3

Exon Direction Sequence Tm Size DMSO

1	Forward Reverse	tgctgtaggaattaaccag tgggccaggtaaaactact	55.4 54.9	382	No
2	Forward Reverse	gtggaaaatggtgttatgttt tttcatcagacagcacattt	54.1 54.2	245	No
3	Forward Reverse	aagcagcatcttttagctc aaaggggagagtggatattt	55 54.3	300	No
4	Forward Reverse	aacctaatgttgaattgtgc tttgggagatccaaactaaa	54.4 54.9	389	No
5	Forward Reverse	ggtgtttggttaagaattca tgatcctttgcaggttattt	54.4 57.9	387	No
6	Forward Reverse	cctctatggcaaattactc tgctgttactttttagctc	54 53.4	399	No
7	Forward Reverse	gaaccaaccaagagaaacag gcagaaacttcaggcttacc	54.8 54.3	280	No
8	Forward Reverse	ttggtcaacttttacagacatc ttgggaaaaattaagaatttga	54.5 54.9	300	No
9	Forward Reverse	aaaagacatatgaatgaaatta ccaaagctgaaattatctct	50.6 53	300	No
10	Forward Reverse	acagttcacaaatccaaagc atagcatttaccagccattg	55.3 55.4	365	No
11	Forward Reverse	tgtttgccagtgtgtct aaagaagaaccagacagatttt	56.2 54	390	No
12	Forward Reverse	tgaataaaatcaagtaatcagg aggatcatttagttgttttcaa	51.6 54.9	369	No
13	Forward Reverse	ggtatggaggtccaatagaa ttgttaaacataaggcaaaaa	54.1 54.6	374	No
14	Forward Reverse	catagccattggcagttaat tgaactctgagagcacgtta	55.4 54.6	385	No
15	Forward Reverse	tgttctttgcttcttctcatt ctttggtgaagggatgaata	55.1 55.1	400	No
16	Forward Reverse	cctcacctctagggatgat agttgttcattaagcatatctcac	55.7 54.5	397	No
17	Forward Reverse	aatgtaggtagccattgaaact aaatcatcccagtgctgaa	54.4 55.5	392	No
18a	Forward Reverse	gtctgtcttgggtgatct ctgcatgaaatcacactctc	54.9 54.1	596	No
18b	Forward Reverse	gtcaaagaaaaggctaagca gaaagcttcatgatctctgc	55 55.1	593	No
18c	Forward Reverse	cttctttggcactctaaagc ttgtgggatatgtcaatct	54.6 55.3	560	No
18d	Forward	aaccaacttaaagacaccaa	54.1	559	No

	Reverse	ttgattgtacatggtttgc	54.5		
18e	Forward	gctggtgtcggaaagtaataa	55.4	564	No
	Reverse	aagcataaattgcaatggtc	55.3		
18f	Forward	caggatttaaggaatatgg	53.8	586	No
	Reverse	gggaagtaacatggcaaaa	55.1		

GNAT2

Exon	Direction	Sequence	Tm	Size	DMSO
1	Forward	gtctttacccttgccata	55.2	542	No
	Reverse	agaggagaggaaaaatgacc	54.9		
2	Forward	cagtcacccctatgtattctcc	55.3	299	No
	Reverse	gcattggaatcagatcctt	55		
3	Forward	ttgagaaagcagtagcaaca	54.8	389	No
	Reverse	agcagggtgggattttagtta	54.1		
4	Forward	cttcactggatactgcttcc	54.9	482	No
	Reverse	cagatctcccatagcatta	55.3		
5	Forward	aggtttgaaggaaagaatcc	54.9	400	No
	Reverse	gggagaagcagaacagtaat	53.6		
6	Forward	cctttggtgaagcctaactt	56.1	400	No
	Reverse	gagacattgattggtctgct	55.2		
7	Forward	gctgttctctgtcttgcttt	54.9	394	No
	Reverse	gtaactgtgcccaaggttc	55.6		

PDE6C

Exon	Direction	Sequence	Tm	Size	DMSO
1.1	Forward	gccttctgtcacatcccaag	60.66	363	No
	Reverse	ACCTGGGTCAGCTCAGAGAA	59.99		
1.2	Forward	TACCTGGAGGAGAACCCTCA	59.65	547	No
	Reverse	ttgcaaactaatgccaaa	60.11		
2	Forward	tgtcatcatccctagatctacca	58.54	267	No
	Reverse	aaatttcctgagcctcaac	58.24		
3	Forward	TTGCTGTGATCATGGCAGTT	60.27	400	No
	Reverse	ggacacgtgggactgaagtt	60.01		
4	Forward	tttatttgatgcacctcatgttt	58.56	278	No
	Reverse	ggcgatgtgagatgagacaa	59.79		
5	Forward	ggtaggtcatggaacagtgtgt	58.85	180	No
	Reverse	tttagcacctgacacatgc	59.72		
6+7	Forward	caagttcccaatgaaaaatg	59.3	397	No
	Reverse	ggatgcattcttgccttcac	58.3		
8	Forward	ctcactcccaatgcgttctt	60.25	214	No
	Reverse	gaggagaggcgagagcatc	60.2		

9	Forward	ctgggtgaagggtacctgaa	59.96	447	No
	Reverse	tggtgcagtttatagaagtcagga	58.15		
10	Forward	tcatggtgcatggtacacttt	58.91	486	No
	Reverse	ccctccttccctgtctctc	58.75		
11	Forward	atcttcccgcctcttctc	60.69	391	No
	Reverse	accagaggcgacaaagatga	60.8		
12	Forward	ctgtgcaaactgcatcatcc	60.27	496	No
	Reverse	ggtagaaggctgtggcaatc	60.00		
13	Forward	tggctctgcttactgatgt	59.58	462	No
	Reverse	accggtgctcaaggatactg	60.13		
14	Forward	aaggttcagtcttaccacaca	59.61	425	No
	Reverse	ggaagcttgggagggaaatag	60.03		
15	Forward	tgaggaggagagaagagagaca	59.11	371	No
	Reverse	gctacttccacttggtcttcc	60.30		
16	Forward	agaggggttgcttagtcca	58.79	490	No
	Reverse	ggcaggcatctgtcttgaa	59.93		
17+18	Forward	tctggcctgtagactttgc	60.40	547	No
	Reverse	gggcctaatacaattgcctac	59.82		
17+18 New	Forward	aaacccttctgagccaaagt	59.3	400	No
	Reverse	catcatcattgccttaggaa	59.9		
19	Reverse	ggaagaaaaagctccccttc	60.53	411	No
	Reverse	ttggtcaggatggtctcgat	60.47		
20	Forward	cagtgagccgagattgtgg	60.41	354	No
	Reverse	gtgataccccaatggctttc	59.25		
21	Forward	acttcataggcccgtggtta	59.45	409	No
	Reverse	tctgcaccagtaaagcaggt	58.52		
22	Forward	ccactgcacctaaccagat	59.99	635	No
	Reverse	ttggttaagcaccagctctgt	59.65		

Appendix I: Setup Protocol for the TOPO TA Cloning kit (Invitrogen) and schematic of the PCR 4-TOPO plasmid used for cloning experiments in the achromatopsia project.

Reagent*	Volume
Fresh PCR product	0.5–4 μL
Salt Solution	1 μL
Water	add to a total volume of 5 μL
TOPO® vector	1 μL
Final Volume	6 μL

*Store all reagents at -20°C when finished. Salt solutions and water can be stored at room temperature or 4°C .

1. Mix the reaction gently and incubate for **5 minutes** at room temperature ($22\text{--}23^{\circ}\text{C}$).

Note: For most applications, 5 minutes will yield sufficient colonies for analysis. Depending on your needs, the length of the TOPO®-cloning reaction can be varied from 30 seconds to 30 minutes. For routine subcloning of PCR products, 30 seconds may be sufficient. For large PCR products (greater than 1 kb) or if you are TOPO®-cloning a pool of PCR products, increasing the reaction time will yield more colonies.

2. Place the reaction on ice

LacZα initiation codon

M13 Reverse priming site | T3 priming site

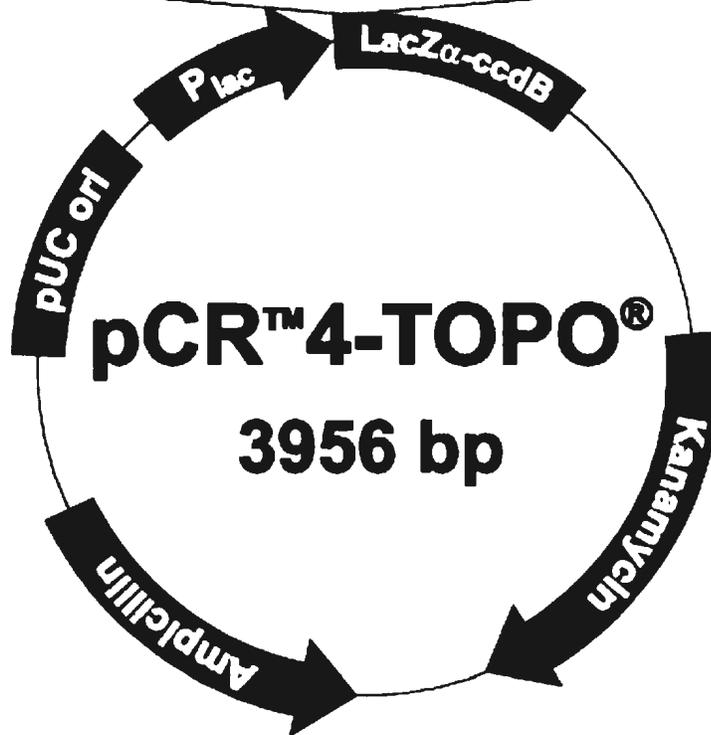
201 CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTCA GAATTAACCC TCACTAAAGG
 GTGTGTCCTT TGTCGATACT GGTACTAATG CCGTTCGAGT CTTAATTGGG AGTGATTTC

SpeI | PstI | PmeI | EcoRI | EcoRI | NotI

261 GACTAGTCCT GCAGGTTTAA ACGAATTTCG CCTT PCR Product AAGGGC GAATTCGCGG
 CTGATCAGGA CGTCCAAATT TGCTTAAGCG GGA TTCCCG CTTAAGCGCC

T7 priming site | M13 Forward (-20) priming site

311 CCGCTAAATT CAATTCGCCC TATAGTGAGT CGTATTACAA TTCACTGGCC GTCGTTTTAC
 GGCGATTTAA GTTAAGCGGG ATATCACTCA GCATAATGTT AAGTGACCGG CAGCAAATG



Appendix J: Sequencing results from cloning experiments using 1734.1 surrounding exons 17-18 of PDE6C.

 = pCR4.0 EcoR1 Site

 = pCR4.0 Pme I site

 = pCR4.0 Plasmid Sequence

 = Matches PDE6C Sequence (Appears that this is the reverse primer in all cases)

 = Neither Plasmid nor PDE6C Sequence (Seems to be related to Chr8)

 = Poor quality sequence

1F Sequence:

TTAGCGGCCCGGAATTCGCCCTTGGCCCTGTAGACTTTGCAGGGTACAGCCCCACTCCAAGCTTCCTTCATAGGAT
 GGCATTGAGTGTCTGTGGCTTTTCCAGGTGCATGGTGCAAGCTGTCAGTGGATCTACCATTCTGGGGTCTGGAGG
 ACAATAAGGCACTCTCAGGCAGACCTGTGTTTGTAGGTCTGGAGGACAGGTCTACTAGGGAATGCCCCAGTGGGG
 ACTCTGTGTGGGCTCTCCAAACCCACATTTCCCTTCTGCACTGGCCTAGCCAAGGTTCTCCATGAGGGCCCTGCCCC
 TGCAGCAAACCTTCTGCCTGGACATCCAGGCATTTCCATACATCCTCTGAAATCTAGGTGGAGGTTCCCAAACCTGA
 ATTCTTGACTTCTGTGCACCCACAGGCTCAACACAATGTGCAAGCTGTCAGGGCTTGGGACTTGCACCCTCTGAAG
 CCATGGCCTGAGCTGTACCTTGGCCTCTTTAGCCACAGCTAAAAGACACAGGGCACCAAGTCTGAGACTGCCCA
 AAGCCACAAGGCCCTGGGCCCCACCATGAAACATTTTTCTCCTAGGTCTCTGGGCCTGTGATGGGAGAAGCT
 GCCATGAAGACCTCTGACATGCCTTGGAAACATTTCCCATTTGTCAGGCAATTTGATTAGGCCCAAGGGCGAAT
 TCGTTTAAACCTGCAGACTAGTCCCTTAGTGAGGGTTAATCTGAGCTTGGCGTAATCATGGTCATAGCTGTTTCC
 CTGTGTGAAATTGTTATCCGCTCACAAATCCACACCACATACGAGCCCGGAAGCATAAAGTGTAAAGCCTGGGGTG
 CCTAATGAGTGAGCTAACTCACATTAATTGCCGTGCGCTCACTGCCCGCTT

TATGAGATGTTGATTAAGCCCA

1R Sequence:

TAGTCCCTGCAGGTTTAAAGGAATTCGCCCTTGGGCTAATCAAATTGCCTAGACAATGGGGAAAATGTTTCCAAGG
 CATGTCAGAGGTCTTCATGGCAGCTTCTCCATCACAGGCCAGAGACCTAGGAGGAAAAAATGGTTTTCATGGGT
 GGGGCCAGGGCCTTGTGGCTTTGGGCAGTCTCAGGACTTGGTGCCCTGTGTCTTTAGCTGTGGCTAAAAGAGG
 CCAAGGTACAGCTCAGGCCATGGCTTCAAGGGTCAAGTCCAAGCCCTGACAGCTTGCACATTGTGTTGAGCC
 TGTGGGTGCACAGAAGTCAAGAATTCAAGTTTGGGAACCTCCACCTAGATTTAGAGGATGTATGGAAATGCCTG
 GATGTCCAGGCAGAAGTTTGTGTCAGGGGCAGGGCCCTCATGGAGAACCTTGGCTAGGCCAGTGCAGAAGGGAA
 ATGTGGGTTTGGAGAGCCACACAGAGTCCCCACTGGGGCATTCCCTAGTAGACCTGTCCTCCAGACCTCAAACAC
 AGGTCTGCCTGAGAGTGCCTTATTGTCTCCAGACCCAGAATGGTAGATCCACTGACAGCTTGCACCATGCACCT
 GGAAAAGCCACAGACTCAATGCCATCCTATGAAGGAAGCTTGGAGTGGGGCTGTACCCTGCAAAGTCTACAGG
 GCCAAGGGCGAATTCGCGGCCGCTAAATTCATTCGCTTATAGTGAGTCTGATTACAAATTCATGGCCGCTGTTT
 TACAACGTGTGACTGGGAAAACCTTGGCGTTACCCAACTTAATCGG

2F Sequence:

GTAGGGGGTGGTGGCCCAAATATAAGGGCAATTGATTTAGCGGCCGC GAATTCGCCCTTGGCCCTGTAGACTTTG
CAGGGTAAGAAGGTAAGAGCTTCCTTCATAGGATGGCATTGAGTGTCTGAGGAAATTACCAGGTGCATGGTGCAA
GCTGTCAGTGGATCTACCATTCTGGGGTCTGGAGGACAATAAGGCACTCTCAGGCAGACCTGTGTTGAGGTCTG
GAGGACAGGTCTACTAGGGAATGCCCCAGTGGGGACTCTGTGTGGGCTCTCAAACCCACATTTCCCTTCTGCACT
GGCCTAGCCAAGGTTCTCCATGAGGGCCCTGCCCTGCAGCAAATTCTGCCTGGACATCCAGGCATTTCCATACA
TCCTCTGAAATCTAGGTGGAGGTTCCCAAATTGAATTCTTGACTTCTGTGCACCCACAGGCTCAACACAATGTGCA
AGCTGTCAGGGCTTGGGACTTGACCCTCTGAAGCCATGGCCTGAGCTGTACCTTGGCCTCTTTAGCCACAGCTA
AAAGACACAGGGCACCAAGTCTGAGACTGCCAAAGCCACAAGGCCCTGGGCCCCACCCATGAAACCATTTTTTC
CTCCTAGGTCTCTGGGCCTGTGATGGGAGAAGCTGCCATGAAGACCTCTGACATGCCTTGAAACATTTCCCAT
TGCTTAGGCAATTTGATTAGGCCCAAGGGCGAATTCGTTTAAACCTTGCAGGACTAGTCCCTTATGAGGGTTAA
TTCTGAGCTTGGCGTAAATCATGGTCATAGCTGTTTCTGTGTGAATTGTTATCCGCTCACATCCACACAACATAC
GAGCCCGAAGCATAAGTGAAAAGCCTGGGGGTGCCTAATGAAGTGAGCCTACCTCACCTAATGGCGTTGCCGC
TCAATTGCCCCGCTTCCAATTCCGAAAACCCTGTCTGCCAGCTGCATTTATGGAATCCG

2R Sequence:

AGTCTGCAGGTTTAAACGAATTCGCCCTTGGGCGTAATCCCTTGCCTAGACAATGGGGAAAATGTTTCCAAGGC
ATGTCAGAGGTCTTCATGGCAGCTTCTCCATCACAGGCCAGAGACCTAGGAGGAAAAAATGTTTCTTGGGTG
GGGCCAGGGCCTTGTGGCTTGGGCAGTCTCAGGACTTGGGGCCTGTGTTTTTAACTGTGGTAAAAAAAAGGC
AAGGGTACCCTAAGGCCGGGTTTTCAAAGGGGGAAGTCCCAACCC

3F Sequence:

TAGCGGCCGC GAATTCGCCCTTGGGCTAATCAAATTGCCTAGGCCAAAGTCAACAGGGCCAGAAAGGGCGAATTC
GTTTAAACCTTGCAGGACTAGTCCCTTATGAGGGTTAATTCTGAGCTTGGCGTAATCATGGTCATAGCTGTTTCT
GTGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCT
AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCT
GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCATTTGGGCGCTCTCCGCTTCTCGCTCACTGA
CTCGCTGCGCTCGGTGTTGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCCGTAATACGGTTATCCACAGA
ATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCG
TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAA
ACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGGAAGCGTGGGCGCTTTCATAGCTCACGCTGTAGTATCTC
AGTTCGGTGTAGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCCGACCGCTGCGCCTTATC
CGGTAACATCGTCTGAGTCCA

3R Sequence:

TAGTCCCTGCAGGTTTAAACGAATTCGCCCTTCTGGCCCTGCACACTTGCCTAGGCAATTTGATTAGGCCCAAGGG
CGAATTCGGGCGCTAAATTCATTCGCCCTATAGTGAGTCGTATTACAATCACTGGCCGTCGTTTTACAACGTC
GTGALTGGGAAMACCTGGGGTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAG
CGAAGAGGCCCGCACCGATCGCCCTCCCAACAGTTGCGCAGCCTATACGTACGGCAGTTAAGGTTTACACCTAT
AAAAGAGAGAGCGTTATCGTCTGTTGTGGATGTACAGAGTGATATTATTGACACGCCGGGGCGACGGATGGTG
ATCCCCCTGGCCAGTGCACGTCTGCTGTACAGATAAAGTCTCCCGTGAACCTTACCCGGTGGTGATATCGGGGATG
AAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCA
GCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGGAATATAAATG

5F Sequence:

CCCCAAAAAAGGGGGCGAAATGAATTTAGCGGCCGCGAATTCGCCCTTGGCCCTAATCAAAATTCCTAGTGTG
TCACATAATATTGTAATTTTTAGAAATCTGAGGTTGCCTGAGATCTTTAGGTGGACATTTACATG

5R Sequence:

TAGTCCCTGCAGGTTTAAACGAATTCGCCCTTCTGGCCCTGTACACAAAAGAAATGTGGACATCAGGATGCAAAGT
GGGAACGTGACATCTGAAAGGCATGTATTTAAGATATTTTTCTCTCTTTTCATAAGCTAAAATGTTGCTCACAGC
TGTATCTTTCTAGGAAGAGGACCATGTTCCAAAAAATTGTTGATGCCTGTGAACAAATGCAAACGGAAGAAGAA
GCCATCAAATATGTAAGTGTGATCCAACCAAGAAAGAGATTATCATGTAGGTAGTTGAAATTGATTTCTCTCTTG
TTTTTTGTAATAAATAACTTATTTTGTACTTATTTCTATTTTAAAATTTTTGTTTTCTTCTAAGGGCAATGATGATGA
CGGCATGTGACTTGTCTGCTATTACCAAGCCCTGGGAGGTGCAAAAGTCAGGTGAGTCCAGTTAAGTTTTCTTTCT
GTCACACTGTCAGTACTGCCTAGGTCCCATGTAATGTCCACTAAAGATCTCAGGCAACCTCAGAATTCGAAA
AATTACAAATATTAGTGAGACTAGGCAATTTGATTAGGCCCAAGGGCGAATTCGGGGCGGCTAAATTCATTCC
CCCCATAGTGAGTCGTATTACAATCACTGGCCCGTCGTTTTTACAACGTCGTGALTGGGAAAAACCCCTGGGGT
ACCCAAACTTAAATCGCCTTGACGACATTCACCTTCCCAAGCTGGCGTAATACCGAAAAAGGCCCGCACCGATC
GCCCTTCCCAACAGTTGCCAGCCAATCCTTACGGCATTAGGGTTACCCCTAATAAAAAAAAAAACCCTTTTCC
TCTGTTTTGGGATGGACAGAATGAAATTATTGACCCCGGGGCGGACCGGAGGGGGTTCCCTCTGGCCGAGG
GCACATCTGGCCGGCCAATCAAATCCCCCGGGGACACTTCCCCGGGGGGGGTGCAATATCCGGGGGGATAA
AAGCCGGGGCGCATATGATGGACCCCCCATACTGGCCCCGCGTGTGCCCGGCTCTCCCTCTTATACCCGGGA
AAAAATTGGGGTCTGACACTCACCTCCCC

Appendix K: All cutoff values for the Homozygosity Haplotyping study carried out in collaboration with the Genome Center of McGill University.

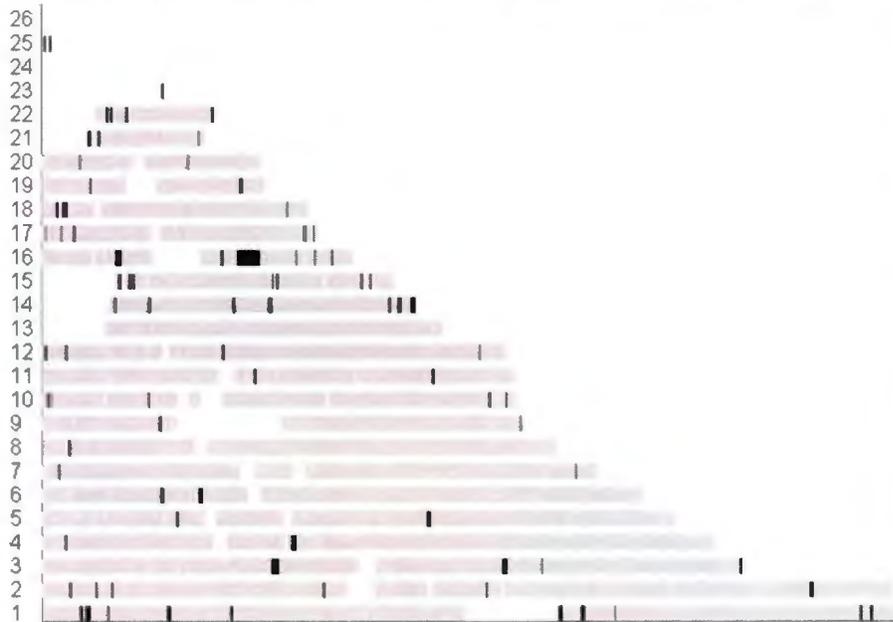


Figure 2. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a cutoff of 1 cM for 5 affected individual

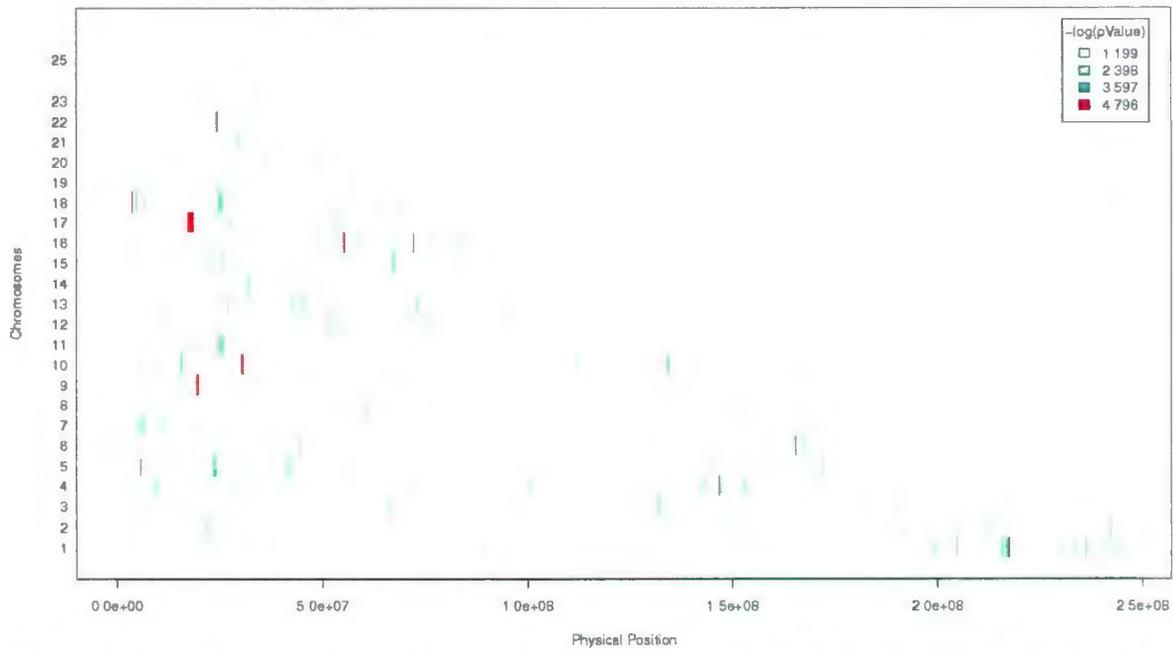


Figure 3. Identification of candidate regions for 5 affected individuals and 24 controls for two Microphthalmia dwarfism pedigrees using HH approach for a cutoff of 1 cM.

Cutoff of 1.5 cM



Figure 4. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a **cutoff of 1.5 cM** for **5 affected individuals**

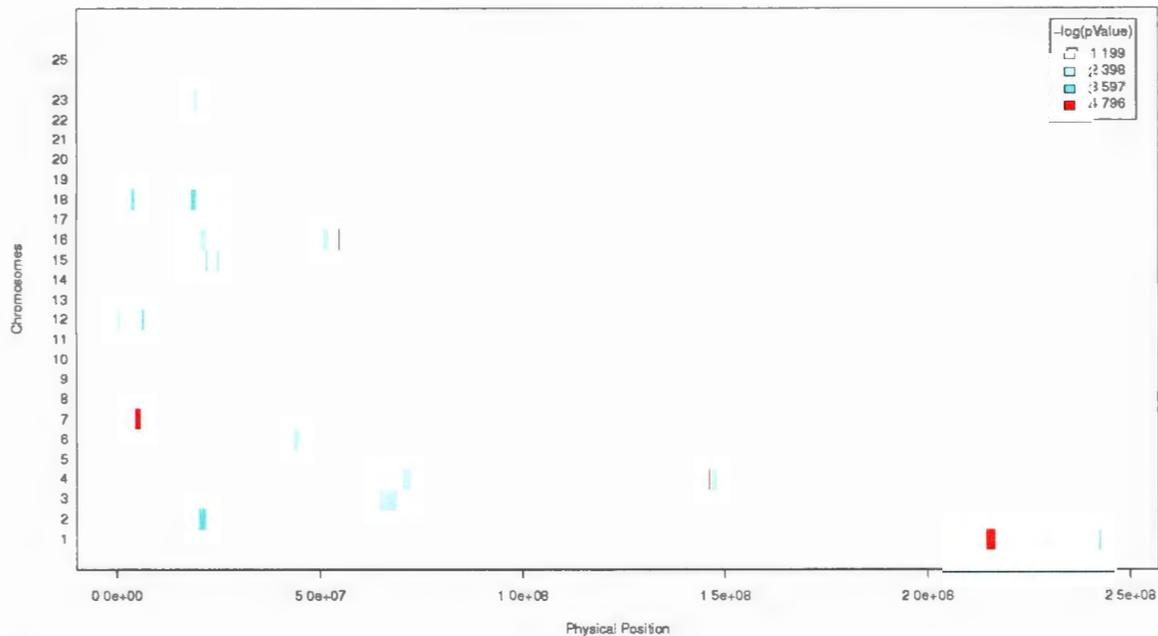


Figure 5. Identification of candidate regions for **5 affected individuals and 24 controls** for two Microphthalmia dwarfism pedigrees using HH approach for a **cutoff of 1.5 cM**.

Cutoff of 2 cM



Figure 6. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a **cutoff of 2 cM** for **5 affected individuals**

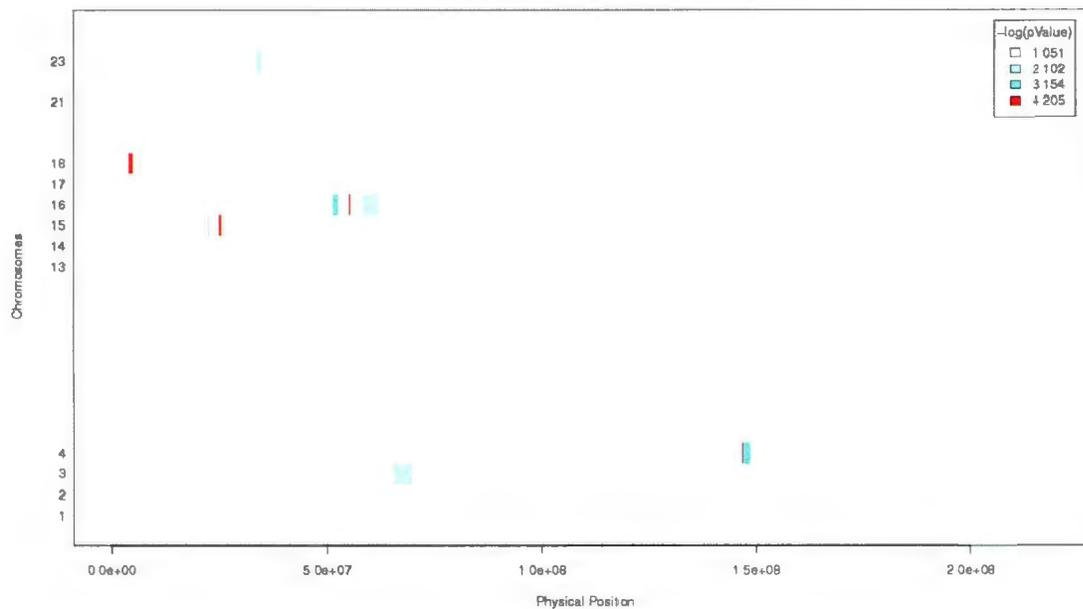


Figure 7. Identification of candidate regions for **5 affected individuals and 24 controls** for two Microphthalmia dwarfism pedigrees using HH approach for a **cutoff of 2 cM**.

Cutoff of 3 cM



Figure 8. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a **cutoff of 3 cM** for **5 affected individuals**

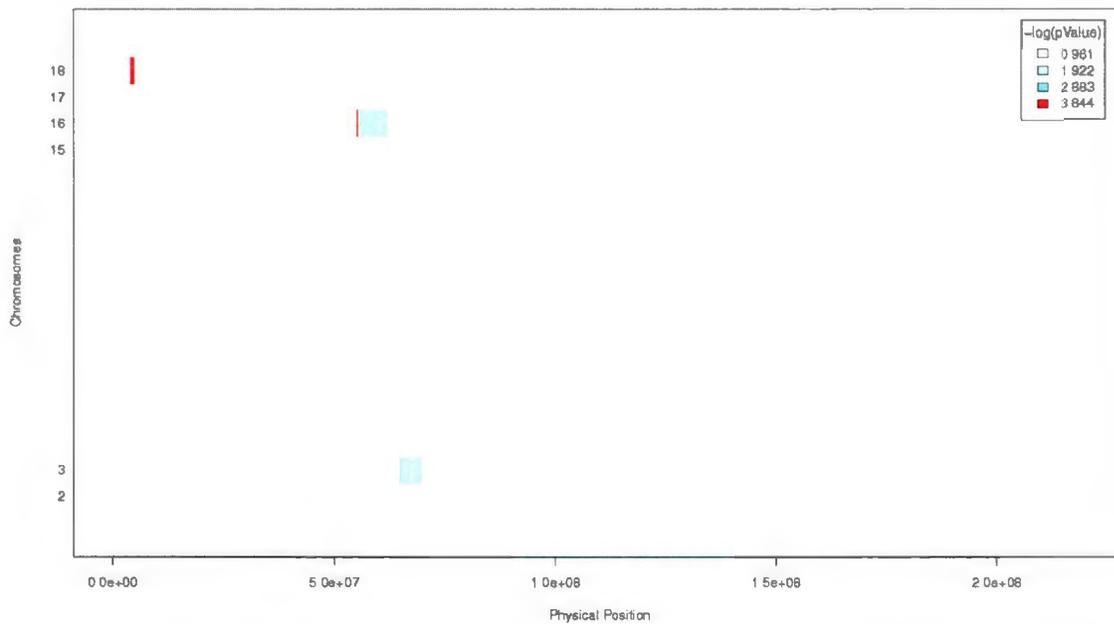


Figure 9. Identification of candidate regions for **5 affected individuals and 24 controls** for two Microphthalmia dwarfism pedigrees using HH approach for a **cutoff of 3 cM**.

Cutoff of 4 cM



Figure 10. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a **cutoff of 4 cM** for **5 affected individuals**

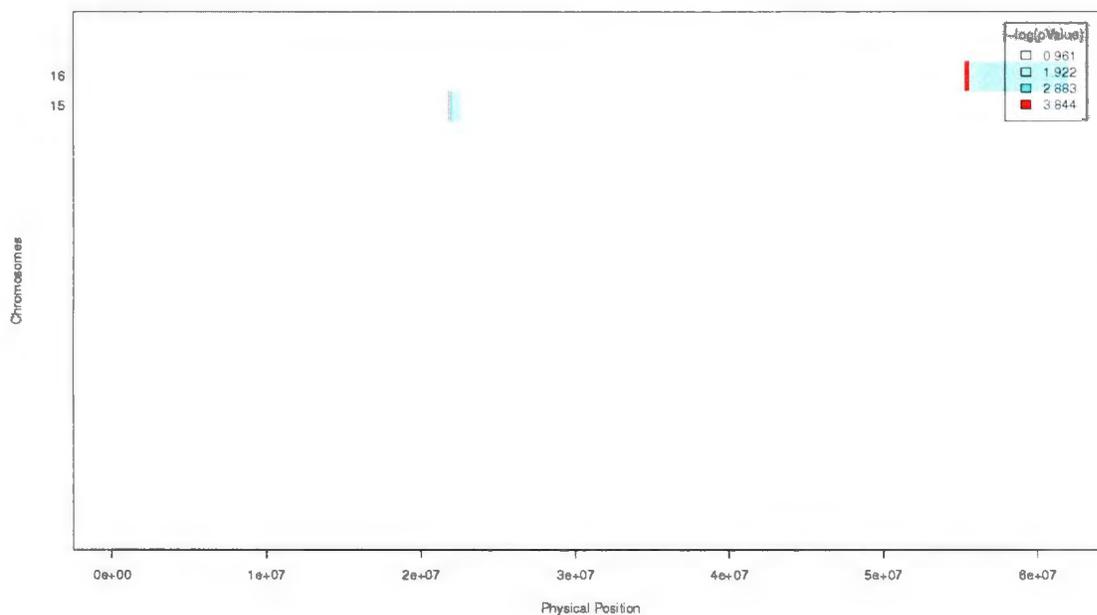


Figure 11. Identification of candidate regions for **5 affected individuals and 24 controls** for two Microphthalmia dwarfism pedigrees using HH approach for a **cutoff of 4 cM**.

Cutoff of 5 cM



Figure 12. Identification of candidate regions for two Microphthalmia dwarfism pedigrees using the HH approach for a **cutoff of 5 cM** for **5 affected individuals**

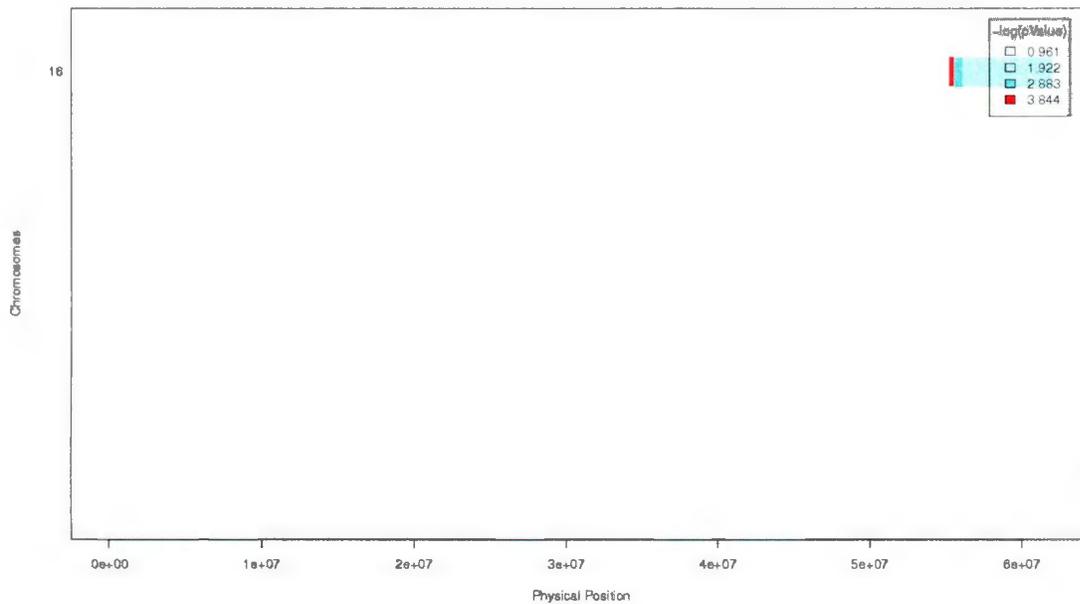


Figure 13. Identification of candidate regions for **5 affected individuals and 24 controls** for two Microphthalmia dwarfism pedigrees using HH approach for a **cutoff of 5 cM**.

Cutoff of 1 cM

There are 96 RCHH, found on all chromosomes except chromosomes 13, 24, and 26, which may include false positives RCHH (see table 1). 60 of them, on all chromosomes except 13, 23, 24, and 26, are longer than or equal to 1 cM (see table 2). When we included controls, there are 59 RCHH found on chromosomes 1, 4, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, and 22 with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 3). Only 7 of them are longer than or equal to 1 cM and we find them on chromosomes 9, 10, 11, 17, and 18 (see table 4).

Cutoff of 1.5 cM

There are 12 RCHH, found on chromosomes 3, 12, 14, 15, 16, 17, 18, and 21 which may include false positives RCHH (see table 5). 10 of them, on chromosomes 3, 12, 14, 15, 16, 17, 18, and 21, are longer than or equal to 1.5 cM (see table 6). When we included controls, there are 19 RCHH found on chromosomes 1, 4, 7, 12, 15, 16, and 18 with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 7). Only 2 of them are longer than or equal to 1.5 cM and we find them on chromosomes 1, and 7 (see table 8).

Cutoff of 2 cM

There are 6 RCHH, found on chromosomes 3, 14, 15, 16, 17, and 21 (see table 9). 5 of them, on chromosomes 3, 14, 16, 17, and 21, are longer than or equal to 2 cM (see table 10). When we included controls, there are 10 RCHH found on chromosomes 4, 15, 16, and 18 with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 11). Only 2 of them are longer than or equal to 2 cM and we find them on chromosomes 15, and 18 (see table 12).

Cutoff of 3 cM

There are only 2 RCHH, found on chromosomes 3 and 16 (see table 13). 1 of them, on chromosome 16, is longer than or equal to 3 cM (see table 14). When we included controls, there are 5 RCHH found on chromosomes 16, and 18 with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 15). Only 1 of them is longer than or equal to 3 cM and we find it on chromosome 18 (see table 16).

Cutoff of 4 cM

There is only 1 RCHH, found on chromosome 16 (see table 17) and it is longer than 4 cM (see table 18). When we included controls, there are 3 RCHH found on chromosomes 15, and 16 with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 19), but they are smaller than 1 cM.

Cutoff of 5 cM

There is only 1 RCHH, found on chromosome 16 (see table 21) and it is longer than 4 cM (see table 22). When we included controls, there are 2 RCHH found on chromosome 16, with a $-\log_{10}(p_value)$ higher than or equal to 3 (see table 23), but they are smaller than 1 cM.

Appendix L: List of primer sequences and melting temperatures (T_m) for genes sequenced in the critical region on chromosome 16 in the MDW project, namely *ARL2BP*, *BBS2*, *C16ORF57*, *MMP15*, and *HERPUD1*.

ARL2BP

Exon 1	Sequence	T _m	Size
Forward	gtggccatggtgacaggag	62.62	599
Reverse	tgctgacagctgagggttg	61.2	
Exon 2	Sequence	T _m	Size
Forward	tgaagatgggtgatgggaag	60.86	300
Reverse	aattgcaggcagcttgatgt	60.81	
Exon 3	Sequence	T _m	Size
Forward	cctccagagactgatcattgag	60.25	392
Reverse	caaaggctgtgcagaggaa	60.12	
Exon 4	Sequence	T _m	Size
Forward	ttggtatccttgccactcag	58.72	400
Reverse	gagccaatgctgatcttt	57.82	
Exon 5	Sequence	T _m	Size
Forward	tgacagagcaagaccagtc	58.97	397
Reverse	ccagccacaacagaatagca	59.86	
Exon 6	Sequence	T _m	Size
Forward	tagggcaggccagtattgag	60.23	440
Reverse	AGGTCCAGGAGGGTTAGGTC	59.42	

BBS2

Exon 1	Sequence	T _m	Size
Forward	GTTCAACCCGCAGGAGTAAA	60.1	674
Reverse	TGCCTGATGACAAGAAAACG	59.8	
Exon 2	Sequence	T _m	Size
Forward	ccactggattccgtgattt	60.86	566
Reverse	gcaaaatctccatgtgtttacc	60.81	
Exon 3	Sequence	T _m	Size
Forward	gggccagttaatgtgtgagg	60.4	485
Reverse	atgcatcgtttctccattt	59.4	
Exon 4	Sequence	T _m	Size
Forward	tcttttggtagccttaagcattt	59.5	500

Reverse	caccatggccttgatctttg	59.1	
Exon 5	Sequence	Tm	Size
Forward	agagggagcatttctgagca	60.1	480
Reverse	tgctttatgtattcaactatatgcaa	60.2	
Exon 6+7	Sequence	Tm	Size
Forward	ttccccactaatggattgga	60.1	599
Reverse	tgtgcagctcttgacaatgat	59.5	
Exon 8	Sequence	Tm	Size
Forward	ttgaaaactgctagactaaatgacct	59	374
Reverse	tccccaactttggtgaattt	59.3	
Exon 9	Sequence	Tm	Size
Forward	tcaccaaagtggggagaac	59.9	368
Reverse	ttctgacaatggcaaaaagg	58.7	
Exon 10	Sequence	Tm	Size
Forward	gctggaggctctgtctttg	60.1	387
Reverse	acaggagaatggcatgaacc	59.9	
Exon 11	Sequence	Tm	Size
Forward	accacgcccagctaattttt	60.8	494
Reverse	ttttacaggtttcaaactgagga	59.2	
Exon 12	Sequence	Tm	Size
Forward	gatgaccagaagtggcttt	59.1	369
Reverse	aagccaccccccaagtagag	60.5	
Exon 13	Sequence	Tm	Size
Forward	ccttcatatctagaggcacattaa	59.2	300
Reverse	aatctatgccctctcattcc	59.4	
Exon 14	Sequence	Tm	Size
Forward	agctgctggaataatgcag	58.1	432
Reverse	tgcaaaaattctgaaaacatca	59.6	
Exon 15	Sequence	Tm	Size
Forward	taagcgaacaggggaaagaa	59.8	396
Reverse	ttgctgggtcatagaggaca	59.2	
Exon 16	Sequence	Tm	Size
Forward	caaagcaaactagcagttactcg	58.5	400
Reverse	gctactttcagccccaatatg	58.7	
Exon 17	Sequence	Tm	Size
Forward	ttttcaataattccctcaccaa	59.7	740
Reverse	ccagcccacattttgtttt	59.8	

C16ORF57

Exon 1	Sequence	Tm	Size
---------------	-----------------	-----------	-------------

Forward	gggctgaatcttcaccactg	60.66	695
Reverse	gccactctggagctttgtct	59.6	
Exon 2	Sequence	Tm	Size
Forward	cccaatgagacaataactggaga	59.05	590
Reverse	tcctcgtgaacagggagaat	59.65	
Exon 3	Sequence	Tm	Size
Forward	agacataggctgctgtccaag	59.51	484
Reverse	aaggtcacggaggatgagaa	59.65	
Exon 4	Sequence	Tm	Size
Forward	gttactcgggctgggtgat	59.94	384
Reverse	gctgcctttatagtcatagagttgg	59.75	
Exon 5	Sequence	Tm	Size
Forward	agagttggctcatggagcag	60.55	400
Reverse	cagaaggctggcattcagag	61.08	
Exon 6	Sequence	Tm	Size
Forward	cagtcataagcaaggaccag	58.31	500
Reverse	gggcaaagatctgtgtt	57.64	

HERPUD1

Exon 1	Sequence	Tm	Size
Forward	ctgctaggataacccactcc	58.62	695
Reverse	gaggaggccacaggacag	59.31	
Exon 2+3	Sequence	Tm	Size
Forward	aaagtgtttcccgtgatgc	59.98	572
Reverse	agcactaacggaacaatctgc	59.4	
Exon 4	Sequence	Tm	Size
Forward	tttgtcacgttcagaggtg	57.28	451
Reverse	tttgtgtactaaaaggctgtg	57.16	
Exon 5	Sequence	Tm	Size
Forward	catgggtgcgtcatacagtc	59.99	376
Reverse	cactcatccctccattctga	58.59	
Exon 6	Sequence	Tm	Size
Forward	tcaaaggcatactactactgaaa	59.63	376
Reverse	cagagcgagactccatctca	59.23	
Exon 7	Sequence	Tm	Size
Forward	cagggtgctgctgtgattag	59.47	300
Reverse	aaatggtaaaccgcaagctg	60.13	
Exon 8	Sequence	Tm	Size
Forward	ctatccagcctagggcatca	60.19	500
Reverse	TGTATCACGGCTTCACGTTT	59.2	

MMP15

Exon 1	Sequence	Tm	Size
Forward	CTCGGGCTTGGGAATTTG	61.52	472
Reverse	ccaagtccgcttgcttcc	61.89	
Exon 2	Sequence	Tm	Size
Forward	gaggatgatgacggctcagat	60.23	397
Reverse	ccaaaccatttggcatctct	59.93	
Exon 3	Sequence	Tm	Size
Forward	ctctgggcactcagtcttc	59.99	457
Reverse	gtgtgtggctggctcatca	60.76	
Exon 4	Sequence	Tm	Size
Forward	gataatgacgggcttgttg	60.33	578
Reverse	tggtagggacaacagactgaa	60.13	
Exon 5	Sequence	Tm	Size
Forward	agcaagggcaggctctct	60.25	399
Reverse	tgtgggaagtcactaatcct	59.61	
Exon 6	Sequence	Tm	Size
Forward	tctgctgtagatggacaggatt	60.14	486
Reverse	tctagagctgagggcttaatgg	60	
Exon 7	Sequence	Tm	Size
Forward	ccactgcctgcctcagtatc	60.82	379
Reverse	gaccagacatgcaatgcaac	60.13	
Exon 8+9	Sequence	Tm	Size
Forward	gggccccagaagagcattc	60.71	596
Reverse	atcacagtgccgggtgac	60.56	
Exon 10	Sequence	Tm	Size
Forward	cagtctggctcctgcaagtt	60.59	658
Reverse	AGCCATCTCAGAACCACAGG	60.26	

Appendix M: Primer sequences and Tms for all non-coding genes (miR, SNORA, and SNORD) genes.

Gene	Direction	Sequence	Tm	GC%	Size
miR3935	Forward	gccctgacgtccattg	60.05	55.56	461
	Reverse	gggcatgtttgacagaga	61.05	50	
miR138-2	Forward	gggacccagattccaccata	61.88	55	390
	Reverse	ttgaaagctgccagggtgatt	60.78	45	
miR328	Forward	accacctgtcgaagtct	60.94	55	394
	Reverse	gactcctgccccattctct	60.6	55	
SNORD111	Forward	ggtgccttttaggaatcagg	58.65	50	232
	Reverse	aacccacagctcaaaattc	59.03	45	
SNORD111B	Forward	tggaaatgggtgttttcaa	58.84	35	226
	Reverse	tcattagccttattctctgaaatagat	57.47	29.63	
miR1972	Forward	ttgagatggggtcttgctatg	60.08	47.62	224
	Reverse	ggctcgaactcctgacctc	58.82	60	
miR140	Forward	ctcctctccaggctctgct	59.83	63.16	239
	Reverse	taaaccagcaaggggatgct	59.93	50	
miR1538	Forward	tccgttctgtttctctc	60.23	50	395
	Reverse	gcttctctcggctctgtctc	59.43	60	
SNORA70D	Forward	aaaagctgtggagataacattca	59.23	33.33	246
	Reverse	cagagacatgaacatccatctga	60.13	43.48	
SNORD71	Forward	ggcacagataagctaggcaaa	59.51	47.62	224
	Reverse	tgctacctgctgctaacag	59.24	55	

