University of Alberta

Study of the molecular cause of anophthalmia in a consanguineous pedigree

by

Azam Khorshidi

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Medical Sciences- Medical Genetics

Fall 2012

Edmonton, Alberta

Abstract

Anophthalmia is a genetically heterogeneous congenital disorder. By using homozygosity mapping in six individuals with anophthalmia from a consanguineous family, five homozygous regions more than one megabase (Mb) in size were identified, which together encompassed 18 Mb. Sequencing of high-priority candidate genes failed to identify the causative mutation. Alternatively, whole exome sequencing of one affected individual revealed a homozygous missense mutation (c.39T>A [p.Ala13Val]) in *TNIP3*, located on homozygous interval on chromosome 4q26-28.1. This mutation was not present in 140 control individuals, single-nucleotide polymorphism databases, or the 1000 Genomes database. There were also several other potential variants elsewhere in the genome which their pathogenesity could not be ruled out, indicating that in heterogeneous diseases a single exome sequencing data is not enough to isolate the pathogenic variant with high confidence. Exome sequencing of more individuals in this family hold the promise to identify mutant gene responsible for the anophthalmia phenotype.

ACKNOWLEDGEMENTS

First I would like to thank my supervisor, Dr. Ordan Lehmann for his enthusiasm, encouragement and support. I am thankful to Dr. Michael A. Walter and Dr. Andrew Wasckiewicz for their help and accepting to sit on my graduate committee.

I would like to thank my colleagues in the Department of Medical Genetics, friends and family for their ongoing support and encouragement, without which this work would not have been accomplished.

Table of Contents

LIST OF TABLES

Chapter 1

Chapter 2

Supplementary Tables of Chapter 2

Chapter 3

Appendices

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendices

Abbreviations

AD: Autosomal Dominant

AM: Anophthalmia Microphthalmia

AR: Autosomal Recessive

BAF: B-Allele Frequency

BMP: Bone Morphogenic Protein

Bp: Base Pair

C: Cornea

CB: Ciliary Body

CNV: Copy Number Variation

DNA: Deoxyribonucleic Acid

dNTP: Deoxynucleotide Triphosphate

dOS: Dorsal Optic Stalk

ECL: Efficient Chemiluminescence

ePCR : Emulsion PCR

GCL: Ganglion Cell Layer

GFP: Green Florescent Protein

HLA: Human Leukocyte Antigen

HTG: High Tension Glaucoma

I: Iris

IBD: Identical By Descent

IBS: Identical By State

InDels: Insertions/Deletions

INL: Inner Nuclear Layer

IOP: Intra Ocular Pressure

kDa: kilo Dalton

LE: Lens Epithelium

LF: Lens Fibre Cells

LOD: Logarithm Of The Odds

LOH: Loss Of Homozygosity

LV: Lens Vesicle

MAC: Microphthalmia Anophthalmia Coloboma

Mb: Mega Base

MIM: Mendelian Inheritance In Man

MO: Morpholino

NGS: Next Generation Sequencing

NR: Neural Retina

NTG: Normal Tension Glaucoma

ON: Optic Nerve

ONL: Outer Nuclear Layer

OS: Optic Stalk

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

pM: Pico Molar

PMSF: Phenylmethanesulfonyl Fluoride

POAG: Primary Open Angle Glaucoma

RA: Retinoic Acid

RFLPs: Restriction Fragment Length Polymorphisms

RNA: Ribonucleic Acid

ROH: Runs Of Homozygosity

RP: Retinitis Pigmentosa

RPCs: Retinal Progenitor Cells

RPE: Retinal Pigmented Epithelium

SHH: Sonic Hedgehog

SNP: Single Nucleotide Polymorphisms

SOLiD: Sequencing By Oligonucleotide Ligation And Detection

STRs: Short Tandem Repeats

TM: Trabecular Meshwork

TNF-α: Tumor Necrosis Factor-α

UTR: Untranslated Region

VNTRs: Variable Number Of Tandem Repeats

vOS: Ventral Optic Stalk

WES: Whole-Exome Sequencing

Chapter 1:

General introduction

Introduction

The visual system is undoubtedly the most sensitive and delicate organ system we possess. It is a principle conduit for acquiring external sensory information and is responsible for four-fifths of all the information our brain receives–which is probably why we rely on our eyesight more than any other sense. However, the number of individuals with major eye diseases is increasing and vision loss is becoming a major public health problem. As blindness leads to loss of independence and reduced quality of life, the importance of studying eye disease is undeniable. Congenital blindness, which is present in childhood, occurs in the developed world at a frequency of 2 to 3 per 1000 live births [1]. A survey of McKusick Catalogs in 1985 showed that the eye was involved in 27% of the 2811 phenotype recorded by Mendelian Inheritance in Man (MIM) and ranked eye as the fourth most common organ in man affected by genetic diseases [2].

This project is an attempt to provide a comprehensive genetic study of a rare and somewhat extreme congenital ocular anomalies (anophthalmia or absent eyes) in a consanguineous family. This disorder is caused by perturbation of key steps in early eye morphogenesis and provides a readily recognizable marker of aberrant ocular development. We hypothesize that milder variants in such genes also result in less extreme phenotypes, and thus that characterizing such genes will provide insight into a diverse spectrum of human disease.

Overview of the eye morphogenesis

The major sensory organ of the head, the eye, is essentially a highly specialized extension of the brain and develops in a series of distinct morphological stages. The human eye develops from distinct tissues: the lens and the cornea are derived from the surface ectoderm, but the retina and the epithelial layers of the iris and ciliary body are from the anterior neural plate [3]. Although documenting the stage of human eye development is difficult due to difficulty of obtaining human embryonic and fetal material for analysis, information has been gathered from a series of embryos in the Carnegie institute collection [4]. This information was then supplemented by the study of eye development in animal models in murine, chick, amphibian and fish [5-6].

Formation of the optic vesicle and lens induction

In vertebrates, the first morphological sign of eye formation is the bilateral evagination of the diencephalon which leads to the formation of the optic vesicles (up to E11 in mouse, 33 days of gestation in human). (Figure 1.1A) As development continues, the budding optic vesicles extend towards the overlying surface ectoderm. The two tissues, then, come into close physical contact and exchange inductive signals, which will, in turn, result in the thickening of the surface ectoderm and formation of the lens placode (Figure 1.1B). Concurrent invagination of lens placode and the distal portion of the optic vesicle lead to the formation of a double-layered optic cup. This process also contributes to the formation of lens vesicle through the process of invagination of surface ectoderm, which then develop into lens proper (Figure 1.1C). [7-8].

Optic fissure closure

The optic stalk, which maintains the connection between the optic cup and the forebrain, undergo invagination along its ventral side to form a choroid fissure. Between the fifth and the seventh week of gestation, the laterally growing edges of this fissure begin to fuse in the middle of the long axis, and continue proximally (toward optic stalk and distally (towards the future iris). The closure of this fissure provides a route for the entry of hyaloid artery into the eye and an exit channel for projecting axons. Incomplete closure results in coloboma. Finally, eye development is completed with the later formation of the cornea and sclera.

Organization of the retina

After the completion of first major phase of morphogenesis, the individual layers of the eye start to adopt their characteristics of their mature structure. The thinner, outer layer of the optic cup–the prospective pigmented retinal epithelium–begins to express pigment. The inner layer—future neural retina—thickens and undergoes extensive cell proliferation to generate retinal precursor cells. Later in the development, progenitor cells undergo differentiation and ultimately give rise to a variety of retinal neurons (Figure 1.2). The ganglion cells are first born neurons, followed by cone photoreceptors, horizontal and amacrine cells. Rod photoreceptor and bipolar cells are generated later [9].

Figure 1.1 Schematic overview of the eye morphogenesis, derived from Chow RL *et al.*, Annu Rev Cell Dev Biol **17**:255-96 [10]. In panels *A–D*, presumptive or differentiated eye tissues are color-coded in the following manner: *blue*, lens/cornea;*green*, neural retina; *yellow*, retinal pigmented epithelium (RPE); *purple*, optic stalk; *red*, ventral forebrain/prechordal mesenchyme; *grey*, mesenchyme.

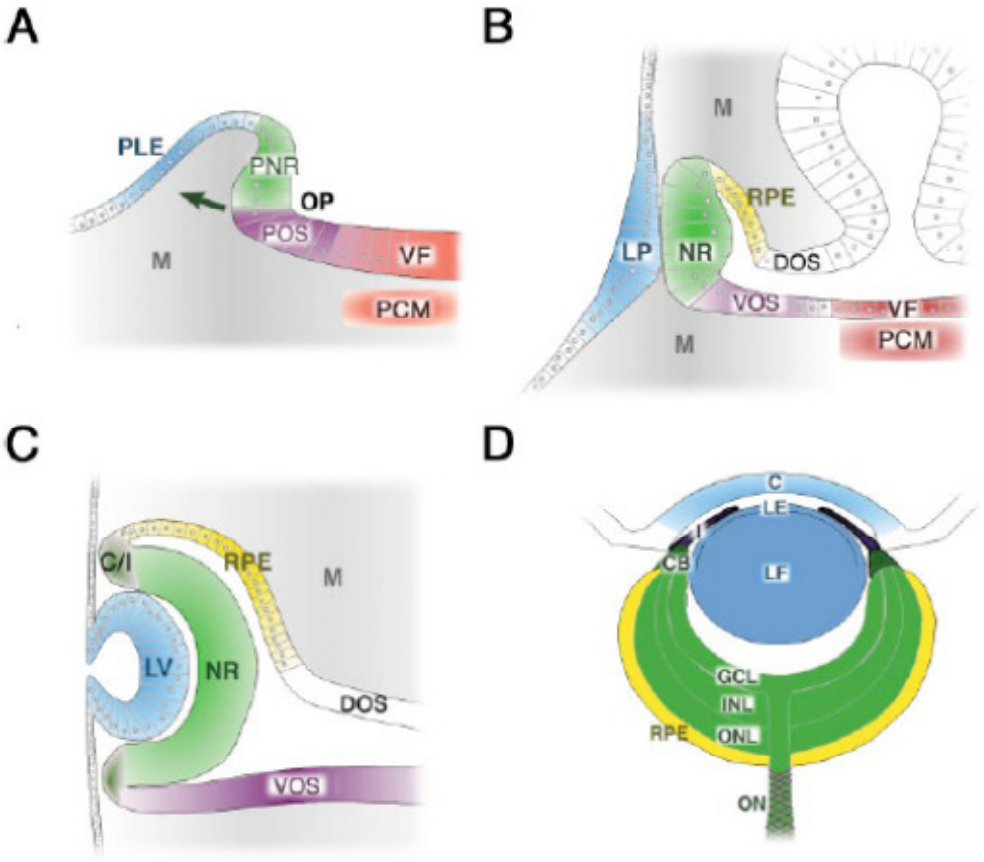
A. Evagination of the forebrain and formation of the optic pit.

B. Contact between the neural ectoderm of the optic vesicle and the surface ectoderm results in induction of the lens placode.

C. The invaginating lens placode forms the lens vesicle (LV) that pinches off the surface ectoderm.

D. Mature eye

Abbreviations: C: cornea, **LE**: lens epithelium, **LF**, lens fibre cells, **I**: iris, **CB**: ciliary body, **GCL**: ganglion cell layer; **INL**: inner nuclear layer; **ONL**: outer nuclear layer; **ON**: optic nerve



Molecular events underlying the eye morphogenesis

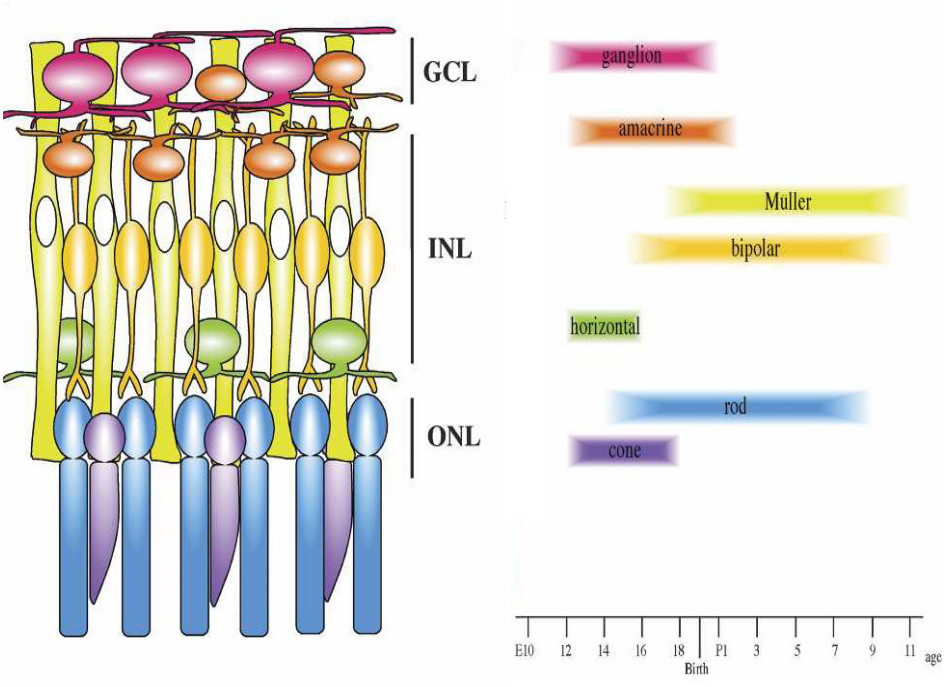
Although distinct morphological stages in eye development have long been noted and characterized extensively, it remained a challenge to categorize the molecular mechanisms responsible for them. Several studies have demonstrated that all of these developmental processes in the eye are controlled by a complex network of regulatory genes, many of which have been highly conserved throughout evolution [11]. These genes are mainly involved in transcription and signaling events.

Molecular events underlying optic vesicle evagination

Before any morphological sign of an optic vesicle formation, several transcription factors begin to be expressed in the eye field, which is a patch of neuroepithelium destined to give rise to the presumptive optic vesicle and its derivatives. *Pax6* and *Rx* are the first two genes expressed in the neuroepithelial sheet and play crucial roles in the specification of the ocular fate [12]. *Six3*, *Optx* and *Otx2* are expressed later in development and have thus been defined as a second wave of gene expression necessary for retinal development (Figure 1.3A &B).

Evagination of the optic vesicles entails dramatic changes in cell behaviour and movement of the retinal progenitor cells. In fish, retinal progenitor cells (RPCs) first show a directed movement toward the midline followed by a subsequent active migration outward, which results in formation of an evaginating optic vesicles [13].

Figure 1.2 Retinal cell fates, derived from Harada T *et al.*, Genes Dev **21**(4):367-78. Three major divisions—GCL, INL, and ONL—give rise to seven retinal cell types that arise from common multipotent progenitors in a fixed order. The RGC is the first neuronal cell type and Müller glia appear last [14].



Rx is a homeodomain transcription factor that is responsible for some of the segregative behaviours of the RPCs. It has been shown that Rx inhibits the canonical Wnt signaling pathway and results in inhibition of the posterior fate of the anterior neural plate and activates non canonical Wnt, which is required for the control of RPCs migratory behaviour [15]. In addition, *Rx* is necessary for the expression of *Optx2*, a gene that controls the proliferation of the RPCs in the eye field [16].

Rx is first expressed in optic vesicle, but its expression will be later restricted to proliferative neuronal retina [12]. Misexpression of *Rx* gene in zebrafish induces ectopic retinal tissue formation in the forebrain, indicating its role in controlling the proliferation of retinal cells [17]. Null mice for *Rx*, on the other hand, fail to develop optic vesicle and, hence, do not develop eyes [18].

Pax6, a transcription factor containing a paired box and a paired-like homeobox, was the first protein identified to be required for the maintenance and normal development of all regions of both developing and adult eye [19]. This gene is considered as "master control gene" as its expression from various animal phyla are capable of inducing ectopic eye development [20].

Pax6 is required for the formation of lens placode, which is necessary for optic cup formation. Also the high level of *pax6* expression in the margins of the optic vesicles neural epithelium indicates its requirement for the iris formation. *Pax6* is also expressed in all cells of the developing retina and in amacrine and ganglion

cells of the adult retina indicating its requirement for normal functioning of these cells.

The encoded protein of this gene is highly conserved in evolution and contains two DNA-binding motifs, including a homeodomain and a paired domain, which are required for directing the transcription factor to its target sites in the genome [21]. In humans, those rare individuals homozygous for mutations in Pax6 are born without eyes and die shortly afterward, [22], while heterozygotes have an array of human diseases including congenital absence of iris (aniridia), Peter's anomaly, and cataracts in humans. Null mice for the murine orthologue of this gene are anophthalmic and show severe brain malformation [23].

Six3 is expressed in the eye field and later throughout the optic vesicle. Disturbances of *Six3* expression result in inhibition of lens placode invagination at early stages, which indicates that this regulatory factor is necessary for normal eye morphogenesis [24]. The ectopic expression of the mouse gene Six3 in the embryos of Medaka fish induces the formation of additional lens and retina [25].

Otx2 is a member of the *orthodenticle*-related family of transcription factors. This gene is expressed in the entire ectoderm before gastrulation. However, its downregulation in the eye field showed that it does not affect early eye specification. Later in development, *Otx2* shows an initial broad expression in optic vesicle and then preferential expression in the retinal pigmented epithelium. In the *Otx2*-/- mice, the expression of Mitf and tyrosinase, which are RPE-specific

genes, is lost and the outer layer of the optic cup will proceed to differentiate as an 'ectopic' neural retina [26].

Molecular evens underlying compartment formation in the developing optic cup

During optic vesicle evagination, the neuroepithelium is bipotential meaning that the presumptive RPE can differentiate into NR, and vice-versa. The homeobox gene, *Chx10* and the bHLH transcription factor *Mitf*, are the two key players in the specification of optic vesicle into RPE and NR (Figure 1.3C&D).

The expression of *Chx10* can be detected in the presumptive neuroretina of the optic vesicle and continues to be expressed in all proliferating retinal progenitor cells where it regulates their proliferation [27-28]. In *Chx10* null mice, all differentiated retinal cell type are present, but the number of proliferating progenitor cells is reduced resulting in the reduction of whole eye growth and microphthalmia. This finding suggests that *Chx10* is not important for specifying cell fates within the neural retina, but is important for the proliferation of the retinal progenitor cells [28-29].

In mice, extraocular mesenchyme produces an activin-like factor, which induces the expression of *Mitf* in the entire optic vesicle. Upon lens induction, *Mitf* is extinguished in the presumptive neural retina, and concurrently, *Chx10* expression is activated in this region [30]. The expression of *Mitf* will be restricted to RPE region where, this gene, in cooperation with *Otx2*, transactivates the expression of pigment genes that are required for the determination of the RPE cell fate [26].

Mitf is essential for the normal proliferation and differentiation of the RPE layer and its ablation in mice results in small eye phenotype due to abnormalities and partial loss of the RPE progenitor cells [31].

FGF signaling is required for RPE maintenance and neural retina formation. In the presence of FGF in the cultured chick optic vesicle, presumptive RPE can transdifferentiate into neural retina. On the other hand, the addition of FGF2 neutralizing antibodies in the explant culture can inhibit neural retinal development. Later in the development, FGF secreted from the lens placode, activating ERK, which in turn induces the expression of *Vsx2* and *Sox2*. Expression of the latter two genes results in NR specification.

Sonic hedgehog signaling (*Shh*) is also required for RPE maintenance in the optic cup and its reduction, as shown in chick and mice, although not affecting RPE specification, results in loss of RPE marker expression, increased proliferation of the RPE and transdifferentiation into retina.

Molecular events underlying optic cup and lens morphogenesis

The expression of *Six* 3 in the surface ectoderm activates *Pax6* and *Sox2*. The expression of these three genes is required for the thickening and invagination of the lens placode (Figure 1.3E & F). Several lines of evidence indicate that *BMP* signaling is important at this stage of eye morphogenesis. In *Bmp7* null mice, *Sox2* and *Pax6* expression is abnormal and, as a result, eye development is arrested at the optic vesicle stage. In addition, Bmp signaling is disrupted in *Lhx2* mutant mice which exhibit eye-less phenotype. At least some, but not all aspects

of the eye phenotype in *Lhx2* mice can be rescued by treatment of the *Lhx2* mutant explants with exogenous *Bmps*.

Retinoic acid (RA) is another important signaling molecule that has a role in optic cup morphogenesis. Two enzymes that synthesize retinoic acid, Retinaldehyde dehydrogenase 2 and 3, are expressed in mesenchyme and RPE, respectively [32]. The synthesized RA is required for morphogenic movement that leads to ventral invagination of the optic vesicle to generate an optic cup [33]. Immediately after optic cup formation RA acts to stimulate apoptosis in the perioptic mesenchyme needed to correctly generate the cornea and eyelids [34]. After initial establishment of RPE in the optic vesicle, cells of this layer cease to proliferate and begin to differentiate. Maintenance of the RPE is required for proper eye development and lamination of the retina. Therefore, controlling of the RPE fate is regulated by several important signaling pathways, including *Shh, Bmp* and *Wnt* (Figure 1.3F).

Developmental abnormalities of the eye

Ocular abnormalities are frequent in human population. As in other organs of the body, developmental abnormalities of the eye arise from disruption of critical processes during eye morphogenesis [35]. In the previous section, I described the major stages during eye morphogenesis and requirement of an organized expression of developmental genes at the right time and the right place for normal eye development. Alterations in such genes can lead to severe disorders that become apparent at birth or shortly afterwards. Depending on the stage of

development affected, these abnormalities can cause significant visual impairment, or they may have only cosmetic significance [36]. The most common causes of paediatric blindness have genetic aetiology though certain environmental toxins and teratogens (drugs or viruses) may also be included [37-38]. Ocular abnormalities can be isolated or part of complex multisystem syndromes which in the latter case they can be harbingers of serious "hidden" abnormalities in other organs such as the heart, brain, or vascular system [39].

Due to the complexity of eye development with numerous opportunities for disruption, the range of structural eye malformation is diverse and encompasses anophthalmia (absent eye) to irregularly shaped pupils. Anophthalmia together with microphthalmia and coloboma forms an interrelated spectrum of congenital eye abnormalities which are collectively referred to as MAC (Figure 1.4) [40]. These structural abnormalities are classified as major congenital anomalies and are responsible for approximately 15% to 20% of paediatric blindness and severe visual impairment worldwide [40-42].

Anophthalmia

Anophthalmia refers to absence of the ocular structures in an orbit as the result of a deficiency in the development of the primary optic vesicle, although adnexal elements are normal and rudiments of optic vesicle-derived structures can be found by serial histopathological sectioning [43]. The presence of ocular tissue remnant may suggest regression of a partially developed optic vesicle [8, 44]. Anophthalmia is rare, with an estimated incidence of 1 case per 100,000 live births [45-46] . This disorder is a serious problem not only due to the absent of an

effective eye, but also because it causes cosmetic disfigurement and psychosocial challenge in affected children.

Different modes of inheritance are described for anophthalmia, including recessive, dominant, and X-linked, though many of these cases appear to occur more or less sporadically in the population [47]. Heritable causes of anophthalmia include chromosome abnormalities and single gene disorders. The non-Mendelian inheritance pattern of anophthalmia indicates that complex mechanisms are involved in the pathogenesis of the disease, including multigenic inheritance or environmental causes [48]. Environmental factors play a contributory role; according to a recent study, the frequency of this disorder appears to be dependent on geographic and ethnic variations. For instance the frequency of anophthalmia were lower in the United Kingdom than that quoted for other developed countries [49].

In 75% of cases, anophthalmia affects both eyes. In the remaining cases, where pathology is unilateral, patients with anophthalmia have a poor visual capacity of the fellow eye, perhaps due to the high incidence of associated developmental anomalies of the fellow eye [50]. In terms of association with systemic diseases, patients with unilateral anophthalmia tend to display ipsilateral facial anomalies, whereas patients with bilateral anophthalmia are characterised mainly by intracranial anomalies [50]. As mentioned before, anophthalmia may be isolated or associated with a broader syndrome (Table 1.1) [51-57]. Isolated anophthalmia is generally an autosomal recessive condition [58-59].

Figure 1.3 Summary of transcription factors in early eye development, derived from Fuhrmann S, Curr Top Dev Biol **93**:61-84 [60].

A. Factors from surrounding tissues regulate patterning of the neural retina and RPE in the vertebrate optic vesicle, which then expresses the specific transcription factors *Vsx2* and *Mitf*, respectively.

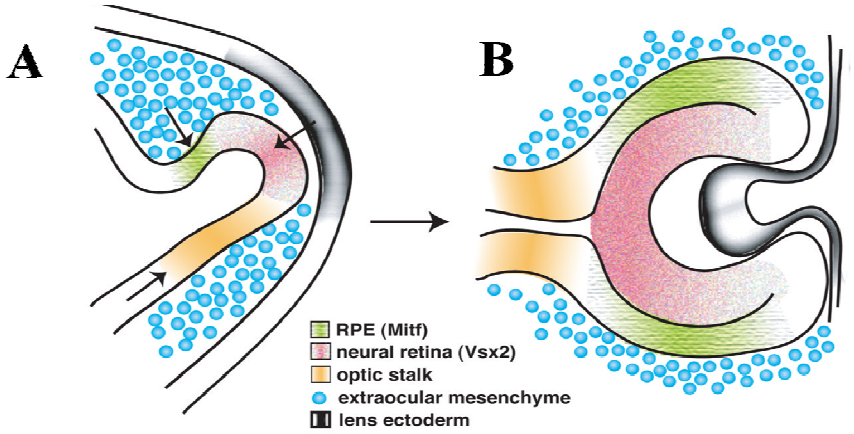
B. Invagination of the distal optic vesicle (presumptive retina) and the overlying lens placode results in formation of the optic cup and lens vesicle.

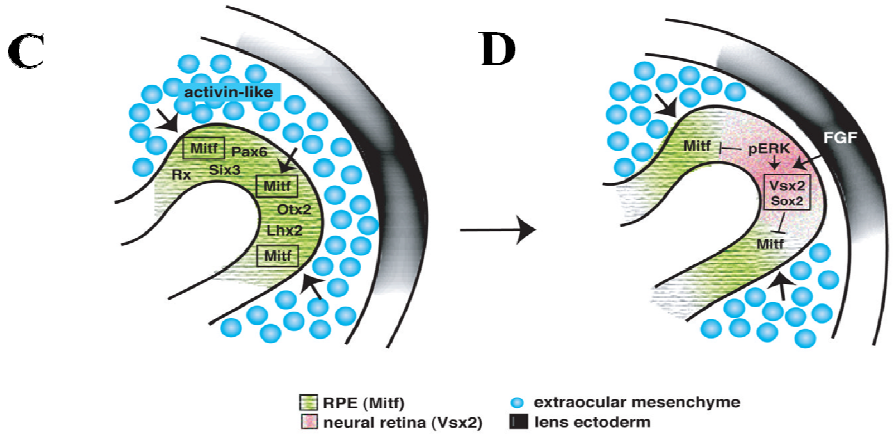
C. RPE specification in mouse (early optic vesicle): the extraocular mesenchyme, induces *Mitf* expression in the entire optic vesicle.

D. Retina specification in mouse and chick (late optic vesicle): subsequentially, activation of ERK, potentially through FGF secreted from the lens ectoderm, induces/maintains *Vsx2* and *Sox2* expression in the distal optic vesicle to promote retina development, which requires *Vsx2* mediated suppression of *Mitf*.

E. Invagination of the lens placode requires correct specification of the lens ectoderm that is dependent on *Six3*-mediated maintenance and activation of *Pax6* and *Sox2*, respectively. FGF and BMP signaling may be also required for lens induction.

F. Maintenance of the RPE in the optic cup by several signaling pathways.


A
B
RPE (Mitf)
neural retina (Vsx2)
optic stalk
extraocular mesenchyme
lens ectoderm


C
activin-like
Mitf
Pax6
Rx
Six3
Mitf
Otx2
Lhx2
Mitf
D
FGF
Mitf
pERK
Vsx2
Sox2
Mitf
RPE (Mitf)
neural retina (Vsx2)
extraocular mesenchyme
lens ectoderm

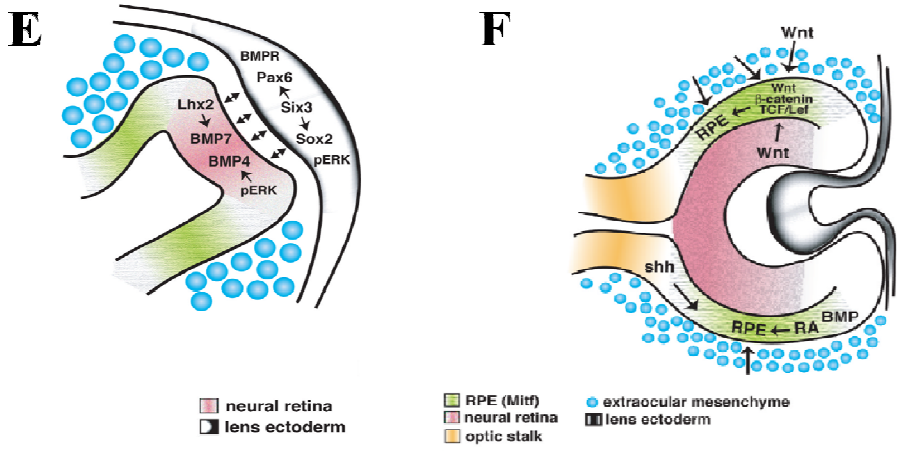

E
BMPR
Pax6
Six3
Lhx2
Sox2
BMP7
pERK
BMP4
pERK
neural retina
lens ectoderm
F
Wnt
Wnt
β-catenin
TCF/Lef
RPE
Wnt
shh
RPE ← RA
BMP
RPE (Mitf)
neural retina
optic stalk
extraocular mesenchyme
lens ectoderm

Table 1.1 Clinical Syndromes Associated with Anophthalmia [61].

Syndromes	**# MIM**	**Inheritance**	**Characteristics in Addition to Anophthalmia**
Fryns	60076	AR	Orofacial clefting, uterine abnormalities, ear abnormalities, neural tube defects
Oculocerbral-cutaneous	164180	AD	Orbital cysts, focal dermal hypoplasia, cerebral malformations, cleft lip/palate
Lenz microphthalmia	309800	X-linked	Microphthalmos, mental retardation, distal limb abnormalities, microcephaly, orofacial clefting, tooth & skeletal anomalies, hearing loss, GU malformations, imperforate anus
Matthew-Wood	601186	Sporadic	Pulmonary hypoplasia
Waardenburg	206920	AR	Syndactyly/other distal limb abnormalities, mental retardation, skeletal anomalies
14q22–23 del	607932	Sporadic	Polydactyly, pituitary hypoplasia, ?SIX6 hemizygosity
	605856		Growth & mental retardation, callosal agenesis, heminasal hypoplasia, atypical clefting, external ear abnormalities,
	60092	Sporadic	Esophageal atresia

Microphthalmia

Microphthalmia refers to a small size of the ocular globe. This condition is more common than anophthalmia, having a birth prevalence of 3 per 100,000 live births [45-46]. The severity of the disorder is highly variable ranging from extreme microphthalmia, which is essentially very similar to anophthalmia, to those with almost normal eye [44, 62]. Extreme microphthalmia may result from failure in very early stages of embryonic life, at the time of optic cup formation. On the other hand, less severe microphthalmia may result from defects occurring at the time of optic fissure closure, as many cases of microphthalmia have been associated with coloboma.

Microphthalmia is bilateral in most cases with the exception of isolated microphthalmia, which is usually unilateral. Asymmetric reduction of the volume of the eyes is common in bilateral cases [63]. Microphthalmia may represent a primary ocular developmental abnormality, or occur in association with syndromes that include non-ocular abnormalities such as cardiac defects, facial clefts, microcephaly, and hydrocephaly [36, 62]. Microphthalmia can be isolated or familial, or can occur in a number of single genes, chromosomal and multisystem malformation syndromes. For example about 25% of patients with microphthalmia/coloboma cases are associated with CHARGE syndrome (coloboma, heart defects, choanal atresia, retarded growth and development, genital malformations and ear anomalies) [43]. The empiric risk of recurrence in a sibling is 2% if both parents are unaffected and increases to 14% if one parent is affected.

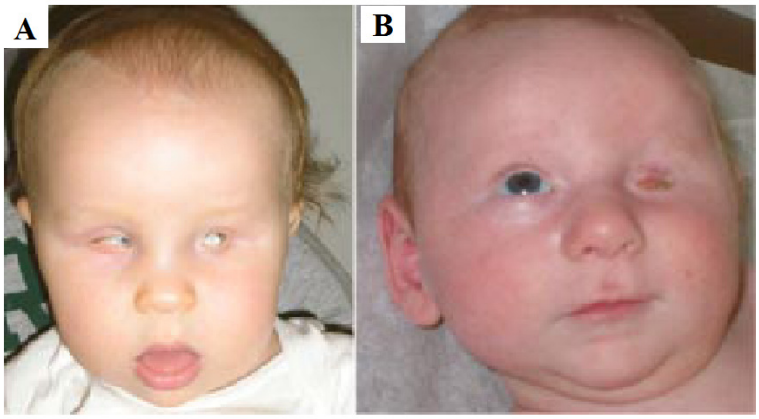
Figure 1.4 Clinical representation of severe malformation of the eye including microphthalmia, anophthalmia and coloboma.

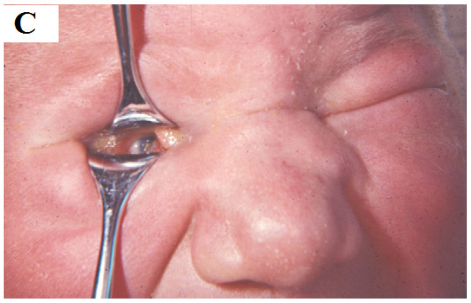
A. Patient with left clinical anophthalmia and right microphthalmia, derived from Schneider A *et al.*, Am J Med Genet A **149A**(12):2706-15 [64].

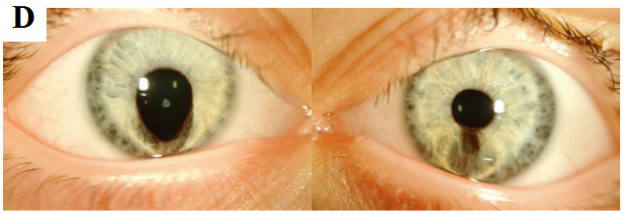
B. Patient with left clinical anophthalmia and normal right eye, derived from derived from Schneider A *et al.*, Am J Med Genet A **149A**(12):2706-15 [64].

C. Microphthalmia and short palpebral fissure in infant with trisomy 13, derived from Guercio JR *et al.*, Otolaryngol Clin North Am **40**(1):113-40, vii [39].

D. Left and right eyes of a patient with typical uveal colobomas. The right iris shows a full thickness defect inferonasally and a keyhole pupil while the left eye has a round pupil and only thinning of the inferior iris, derived from Traboulsi *et al.*, Pediatric Ophthalmology 2009 [63].







Coloboma

Coloboma denotes 'defect' and is a term applied to a variety of eye disorders wherein a portion of the eye is absent due to failure of closure of the embryonic choroidal fissure [65]. Depending on the extent of involvement, iris, ciliary body, inferior choroid, and/or optic nerve head may be affected [66-69]. Eyes with coloboma are usually microphthalmic although they may be of normal size. In the area of defective closure, a cyst may form, producing microphthalmia with cyst [63, 67]. The relative sizes of the ocular globe and the cyst are variable and, on occasion, the cyst may be much larger than the ocular globe [62].

The incidence of coloboma depends upon the population studied ranging from 0.75 in China [70], 2.6 in the USA [71], 0.7 in France [72], 0.5 in Spain [73], and 0.41 in Hungary [74] per 10,000 births. Although in the overall this condition is rare, it comprises a significant proportion of blindness in children (up to 10% of blind children in Europe) [75].

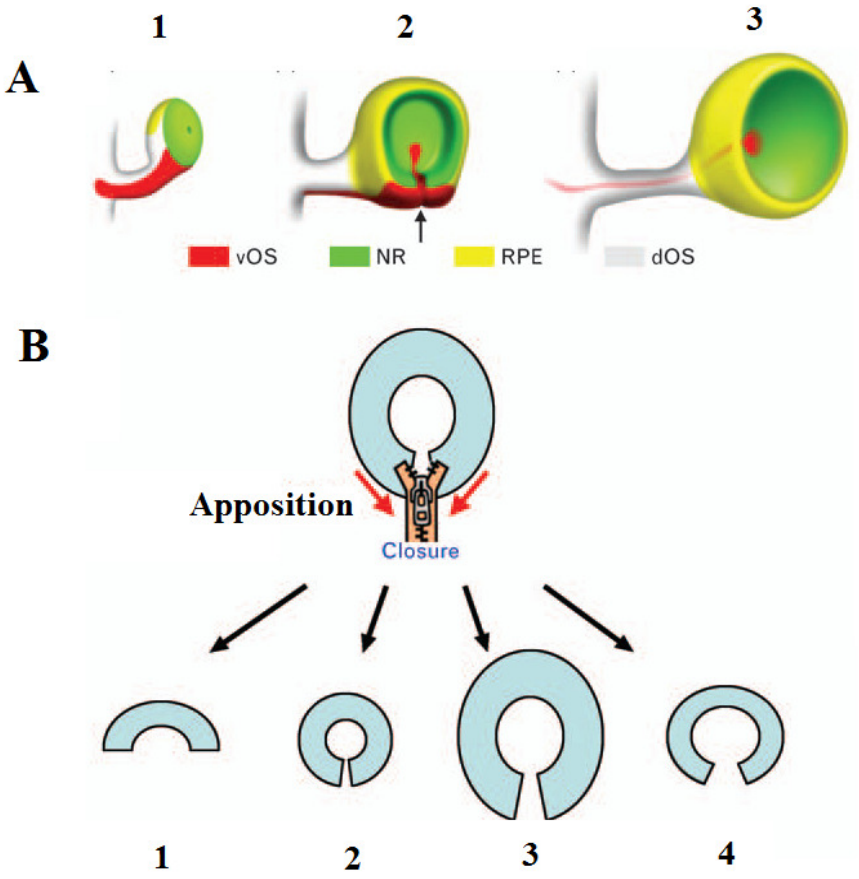
As to pathogenesis, Chang *et al.* have proposed four mechanisms for coloboma (Figure 1.5). They explained that apposition (approximation of the two lips of the optic fissure) and fissure closure are two essential events that must occur normally during development. Thus, in a mutated developing optic cup with slower growth, the edges of optic fissure might not come close enough to fuse. Instead, a different mutation may affect cell-to-cell adhesion, preventing proper fusion even if the edges of optic cup were close enough [76].

Figure 1.5 Embryonic eye development and the two processes required to avoid a coloboma: apposition and closure [76-77].

A. Embryonic eye development, derived from Mui SH *et al.*, Genes Dev **19**(10):1249-59 [77]. **1** Optic vesicle evagination. **2**. Optic cup formation. **3.** Optic fissure closure.

B. Apposition and closure, derived from Chang L *et al.*, Curr Opin Ophthalmol **17**(5):447-70 [76]. **1.** A poor apposition due to small size of the optic cup; a large coloboma in a small eye results. **2.** Closure failure in a small optic cup; a small coloboma in a small eye results. **3.** When optic cup has normal size but optic fissure fails to close; a small coloboma in a normal-sized eye results. **4.** When there is variable degree of failure in optic fissure in a normal sized optic cup; variable-sized coloboma with variable eye size results.

Abbreviation: vOS: ventral optic stalk , **NR**: neural retina, **RPE**: retinal pigmented epithelium, **dOS**: dorsal optic stalk.



Etiology of the MAC

MAC is a very heterogeneous condition, exhibiting diverse patterns of inheritance and variable severity, with both genetic and environmental factors assumed to be the potential causes of the disease. Genetic causes may be due to major chromosomal abnormalities or mutations in a number of developmental genes. However, almost 60-70% of patients do not have an identified genetic etiology for their birth defect, which may suggest that new genes or pathways remain to be identified.

Environmental causes of MAC

Only 20% of congenital eye malformations are the result of Mendelian inheritance and the rest sporadic forms of MAC cases are of unknown aetiology [78-81]. A number of potential environmental causes have been predicted in the pathogenesis of the disease though there is only limited evidence in their support. Descriptive epidemiological studies have shown a marked geographical variation in the prevalence of congenital eye anomalies. The general pattern for these anomalies suggests that they are more common in Asia than Africa or Latin America. Although this might be related to genetic factors (e.g. consanguinity), the role of environmental factors or interaction between the two is also probable [81].

The strongest evidence in support for the role of environmental causes is maternal infections with, rubella [82], varicella [83], cytomegaloviruses [84] and toxoplasmosis [85]. There are also reasonable evidence to support a correlation between using thalidomide and alcohol during pregnancy on the one hand, and

manifestation of ocular abnormalities in infants, on the other hand [86]. For example up to 90% of children whose mothers have misused alcohol during pregnancy, show ocular manifestations [87]. Another important environmental cause of ocular abnormalities is maternal vitamin A deficiency. It has been shown that in South India, where 50% of women are suffering from vitamin A deficiency, 16% of infants are born with coloboma [88].

Heritable causes of MAC

Although a significant genetic contribution to the aetiology of MAC is suspected, the identification of the molecular basis of these conditions has been shown in only a few cases where a detectable chromosomal abnormalities or a familial clustering of these conditions is observed [89]. Cytogenetic analysis and linkage mapping have provided some clue to where to look for the candidate genes. This has led to the identification of several loci and genes in association with MAC. The large number of potential candidate genes and loci indicates the genetic and locus heterogeneity of the diseases and the need for several intact genes and networks to achieve a normal eye development [89]. In addition to chromosomal abnormalities and single gene mutations, abnormal copy number variation can also contribute to the pathogenesis of the MAC [90-92].

Chromosomal abnormalities in A/M

Causative chromosome aberrations, including aneuploidy, deletion, duplication and balanced *de novo* translocation, are found in an estimated 25-30% of MAC cases. The significance of these chromosomal anomalies is their help in identifying regions in which candidate genes for MAC can be found [40].

Trisomy 13 (Patau syndrome), trisomy 18 (Edwards syndrome) and trisomy 9 mosaic syndrome are the most common chromosome aneuploidies, which have been associated with anophthalmia/microphthalmia (A/M) and other eye anomalies [93-95]. Using conventional cytogenetic techniques various chromosomal rearrangements have been also reported in the literature, including del7p15.1–p21.1 [96] and del14q22.1q23.2 [97] - where *SIX6* has been mapped to this region and identified as a possible candidate gene for anophthalmia [98]. In addition, several cases of microdeletion and microduplication have been reported by using microarray techniques, including deletion of 16p11.2 [99]and 16 q11.2q12.2 [100] and duplication of 10q24.31 and 15q11.2q13.1 [90].

Monogenic causes of MAC

A subset of heritable forms of MAC originates from mutations in particular genes [43]. These mutations display a high risk of familial transmission either in autosomal recessive or autosomal dominant patterns [101]. Thus far, several genetic mutations have been linked to MAC, amongst which the *PAX6* gene was the first to be identified [22], though *SOX2* is a major causative gene [102]. Other reported mutations associated with MAC include *RAX* gene leading to anophthalmia in humans [103], and loss of function mutation in the *OTX2* and *CHX10* genes related to microphthalmia [104]. Mutation in the above mentioned genes show a considerable phenotypic heterogeneity, either inter- or intrafamilial, indicating that additional genetic or environmental factors play a part in modifying the final ocular phenotype [105-106].

SOX2

Mutation in the gene *SOX2* accounts for 4-20% of bilateral MAC cases [64, 102, 107-109]. Although most individuals with mutation in *SOX2* suffer from severe bilateral MAC, a wide range of ocular phenotypes has been reported, including unilateral A/M to cataracts, coloboma, papillary defects with hypermetropia and retinal dystrophy [64, 108-112]. These variable phenotypes of *SOX2* mutation in human match different ocular phenotypes observed in the mouse with gene dosage allelic series of *Sox2* mutations [110, 113]. Nevertheless, some individuals with *SOX2* mutation but with no ocular phenotype have been also reported [114]. *SOX2* mutations typically follow autosomal dominant inheritance with most cases result from sporadic new mutation [115]. In some cases, familial recurrence has been observed -- even with clinically unaffected parents -- indicating the possibility of germ line mosaicism [116].

PAX6

PAX6 was the first gene implicated in anophthalmia and perhaps the most extensively studied gene in the eye development. In human, heterozygous mutations are associated with aniridia, and rare cases with homozygous mutation have been observed in association with severe craniofacial abnormalities, anophthalmia and malformations of the nervous system [22]. Although *PAX6* mutations in anophthalmia cases are rare, a plausible collaborative role between *PAX6* and *SOX2* has been demonstrated [117]. *PAX6* and *SOX2* interaction drives formation of the lens placode through γ-crystallin, which is required for lens development [44, 117-118]. Therefore, it seems that failure in lens induction is the

primary cause of microphthalmia in patients with mutation in *PAX6* and *SOX2* genes.

OTX2

OTX2 is another important transcription factor important for eye and brain development in vertebrates [104, 119-120]. Mutations in *OTX2* contribute to 3% of MAC cases [91, 121] and can be associated with other systemic anomalies including growth retardation [122] and pituitary abnormalities [123-124]. In mice, homozygous deletion of *Otx2* causes prenatal death and heterozygous mutation results in a range of malformation including anophthalmia, microphthalmia, holoprosencephaly, anencephaly and short nose [122, 125]. Similar to *SOX2*, affected individuals can inherit the disease from phenotypically normal parents due to mosaicism or incomplete penetrance [121].

RAX

RAX, a conserved homeobox gene in vertebrate, has been shown to be important for eye development by regulation of retinal progenitor cell proliferation. Molecular analysis of 75 A/M cases led to the identification of a nonsense mutation in *RAX* in an individual with A/M and sclerocornea [103]. Since then, mutations in the *RAX* gene have been detected in other A/M cases as well as in patients with coloboma [126-127]. Systemic malformation is usually absent with *RAX* mutation except for two cases who were exhibited with septum pellucidum agenesis, cortical atrophy, optic nerve hypoplasia, hydrocephalus and congenital hip dislocation in addition to anophthalmia [101].

CHX10

CHX10 is required for the regulation of retinal progenitor cell proliferation, which is known to be important for eye development based on its pattern of expression and strong sequence conservation in vertebrates [28, 128-129]. Although *CHX10* was first described and identified in two families with non-syndromic microphthalmia, other ocular malformations have also been reported, including cataract, iris abnormalities, retinal dystrophy and coloboma [130-132]. To date, no extraocular finding is reported in *CHX10* mutations [133].

TGF-β superfamily

Recently mutation in members of TGF-β superfamily has been identified in patients with eye anomalies. *GDF3* [134], *GDF6* [135], and *BMP4* [136] are examples of such genes, which regulate cell proliferation and death. In addition they can be involved in dorso-ventral patterning, bone formation and organogenesis [137-139].

The role of *BMP4* in ocular disease was first described in two anophthalmic patients with *de novo* deletion of chromosome 14 (del14q22.3-q23.2 and del14q22.2q-q23.1) [48, 140] encompassing *BMP4* and *OTX2* genes. The following report of a patient with sclerocornea and congenital glaucoma who had deletion in *BMP4* but not *OTX2,* suggested the involvement of *BMP4* in the pathogenesis of MAC [141]. In mice, *Bmp4* is important for lens induction and cooperate with *Pax6* and *Bmp7* during eye development. Heterozygote mice show a variety of ocular abnormalities including A/M due to failure in lens induction, anterior segment dysgenesis and optic nerve hypoplasia [48, 142].

GDF6 is a member of growth and differentiation factor superfamily that acts through SMAD signaling [135, 143]. Implication of *GDF6* in ocular disease was described in a patient with bilateral chorioretinal coloboma, who had a deletion in chromosome 8 encompassing *GDF6* and 30 other genes [135]. Subsequent screening of 489 patients with MAC revealed four missense mutations [144]. Moreover, a recent study showed that *GDF6* is mutated in 8% of MAC cases (4/50) [101]. In *Gdf6* knockout mice, variable and asymmetric ocular defects have been observed, including optic cup excavation and microphthalmia [144].

GDF3, another member of *BMP* family, is also a key developmental regulator of eye formation. Recently, it has been shown that mutations in *GDF3* contribute to 1.5% of MAC cases, and knockdown of dvr1, an ortholog of *GDF1*/*GDF3*, in zebrafish cause ocular abnormalities [134].

Other significant genes

There are also several other genes which have been associated with MAC (Table 1.1). *SIX6* was described in an anophthalmic patient with 14q22.3-q23 deletion as the candidate gene [98]. Interstitial deletion in this region was also reported in three other A/M families. Nevertheless, subsequent screening of 173 A/M cases did not identify any significant sequence alteration for this gene suggesting that mutation in *SIX6* is not a frequent cause of A/M [145-146]. *LHX2* is a transcription factor necessary for eye development [147]. In mice, it regulates the expression of other eye developmental genes such as *Rax*, *Sox2*, *Chx10* and *Otx2* [148-149]. However, it seems that mutation in *LHX2* is uncommon in human(s) as a recent genetic screening of 70 A/M patients did not detect any significant

variants [99]. *SMOC1,* which has been identified as a causative gene for Waardenburg anophthalmia syndrome, is another significant gene [150-151]. In addition, mutations in *TMX3* [152], *GLI2* [153], *STRA6* [154-155], *FOXE3* [156], *HCCS* [157], *SHH* [158-160] and *BCOR* [161] have been linked to the MAC phenotype. There are also some genes specifically linked to coloboma, such as *VAX2* [162], *PAX2* [163-165], *SIX3* [166], *PTCH* [167], *ZFHX1B* [168] and *MAF1* [169].

Copy number variation (CNV)

The eye is a sensitive organ for gene dosage of several developmental genes, including *BMP4*, *PAX6* and *FOXC*; increase or decrease in the level of expression in such genes causes ocular malformations [142, 170-171]. There is increasing evidence that gene copy-number variations are associated with several human disease -- including ocular abnormalities -- as well as disease susceptibility, which make these structural variants as significant as single nucleotide mutations in the study of human diseases [172].

An autosomal aneuploidy, named trisomy 21 or Down syndrome was one of the earliest structural copy number variations to be linked to a human phenotype [173]. Since then, several duplications or deletions that affect sizable chunks of chromosome were identified in association with a number of inherited disorders as well as in normal individuals [174]. The first genome-wide mapping of structural variation in the human genome, primarily, were carried out by Steve Scherer's group and another by Michael Wigler's group in 2006 using array-based comparative genomic hybridization approaches to measure the occurrence of copy

variants across the genome [175-176]. Both teams observed numerous submicroscopic chromosomal alterations in hundreds of places in the human genome, including areas coding for disease-related genes [175-176].

These quantitative genomic variants were eventually named CNVs and were defined as a segment of DNA larger than 1 kilobase presenting at a variable copy number as compared to a reference genome [175]. Altogether, CNVs cover approximately 12% of the human genome [175].

Importance of CNVs in relation to human diseases was proven by showing the fact that almost 300 known disease causing genes overlapped with CNVs [175]. Upon these findings, scientists immediately began to speculate that CNVs might have phenotypic consequences and be associated to certain diseases. In some cases, this association was discovered by searching for CNVs in known disease causing genes, such as α-synuclein in Parkinson disease or APP in familial Alzheimer's disease [177-179]. There are also some cases of CNVs that was discovered by association studies including CNV in chemokine ligand 3-like 1 (*CCL3L1*) and increased risk of HIV-1 infection, rheumatoid arthritis and diabetes [180-181]. Similarly, copy number variations were identified with spinal muscle atrophy and DiGeorge syndrome [182].

Table 1.2 Proposed list of candidate genes for anophthalmia, microphthalmia and coloboma**,** derived from Raca G *et al.* Mol Genet Metab **100**(2):184-92 [183].

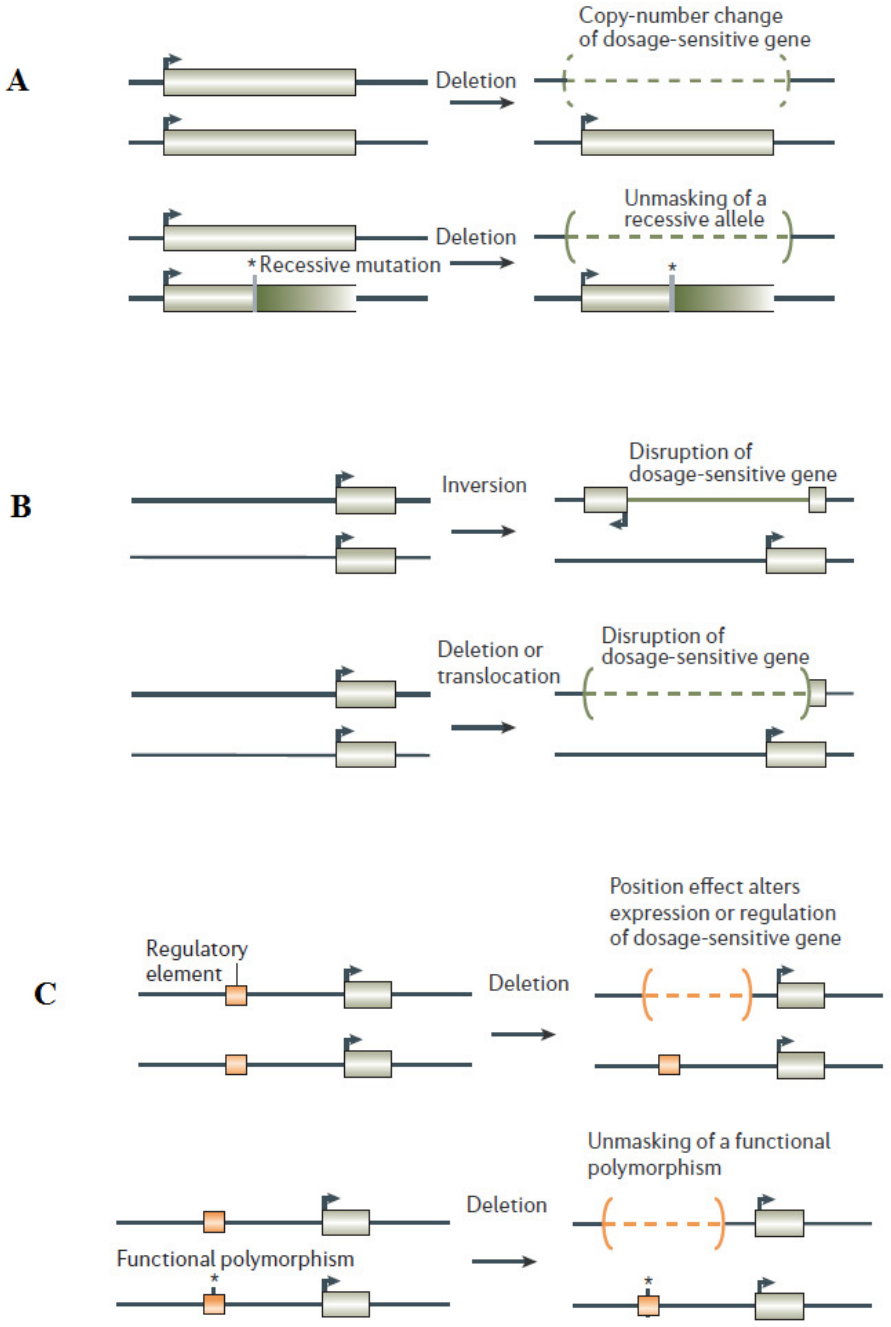
Implicated in syndromic anophthalmia, microphthalmia and coloboma	Implicated in non-syndromic anophthalmia, microphthalmia and coloboma	Code for proteins on the SHH signaling pathway	Code for proteins on the WNT signaling pathway	Other
DPYD, SIX3, RAB3GAP, ZFHX1B, ALG3, PITX2, *POMT1, KIAA1279, RBP4, PAX2, PAX6,* *PTPN11, CREBBP, SALL1 NDP, MKS1, SALL4,* *PQBP1, BCOR, PORCN, IGBP1, FLNA, CHD7,* *CC2D2A, HMX1, CLDN19, LRP2, TFAP2A, IGBP1,* *IKBKG, B3GALTL, GDF6, SOX2,OTX2, JAG1,* *BMP4, MITF, HESX1*	*CHX10, MAF,* *SIX6, RAX, GDF3*	*GLI2, GLI3,* *SMO,VAX1,* *VAX2, SUFU,* *DRM, SHH, PTCH*	*DVL1, WNT1, WNT16, WNT5A, WNT5B, WNT7B,* *WNT7A, WNT2, WNT8A, WNT10A, WNT6, WNT8B,* *WNT3, WNT3A, WNT9A, WNT9B, WNT11, WNT4,* *WNT2B, LRP5, LRP6, FZD6, FZD9, FZD2, FZD1, FZD7,* *FZD5, FZD10, AXIN2, GSK3A, APC, CTNNB1, CER1*	*CRYAA, CRYAB, CRYBA1,* *CRYBA2, CRYBA4, CRYBB1,* *CRYBB2, CRYBB3, CRYGA,* *CRYGB, CRYGC, CRYGD, CRYGS,* *CRYZ* *SOX10, ZNF703, ZNF503* *CRX, FOXG1,BMP7, CHD2,* *DLX1, DLX2, TBX2,TBX5, FGF8,* *HES1, LHX1*
40 loci	4 loci	8 loci	32 loci	28 loci

CNVs are speculated to cause diseases through several mechanisms, as shown in Figure 1.6. First, CNVs can result in insertion or deletion of a dosage sensitive gene, which in turn causes alteration in gene expression. A recessive allele, which is not normally expressed, may also become unmasked due to deletion of the normal allele (Figure 1.6.A). Structural variants can also disrupt the expression of a gene through inversions, deletions, or translocations (Figure 1.6.B). In addition, structural variants can affect a gene's expression indirectly by altering the potential regulatory elements of a gene which is called “position effect” (Figure 1.6.C) [184-185].

In keeping with the association of CNVs and human diseases, sensitivity to dosage changes of key regulatory genes, have been reported in some cases of severe abnormalities of the eye [142, 170-171, 186]. Dosage sensitive genes might be signaling molecules whose function depends on partial or variable occupancy of a receptor, or transcriptional regulators which are needed in quantity in rate limiting steps in developmental events [187].

Eye development has shown to be extremely sensitive to change in dosage of *PAX6* gene; either a too low or too high level of this gene results in abnormalities of the eye [170, 188]. Heterozygous mutations in *Pax6* cause Small eye (*Sey*) phenotype in the mouse [23, 171, 189], and Peters’ Anomaly and aniridia in humans which is characterized by a varying degree of iris hypoplasia, corneal opacification, cataract and glaucoma [190-191].

Figure 1.6 Influence of structural variants on phenotype, derived from Feuk L *et al.*, Nat Rev Genet **7**(2):85-97. A. Duplication or deletion event encompassing dosage-sensitive genes can cause disease (upper panel; a deletion is shown here). Alternatively, deletion of a gene can unmask a recessive mutation on the homologous chromosome (lower panel). B. Inversion (upper panel), translocation or deletion (lower panel) can disrupt expression of a dosage-sensitive gene. C. Structural variants that are located at a distance from dosage-sensitive genes can affect expression through position effects [184].



Homozygous *PAX6* knockout mice exhibit anophthalmia and severe brain abnormalities; they die shortly after birth. Homozygous *PAX6* mutations in human are rare but exhibit a phenotype very similar to one in Pax6 null mice [22, 192]. In addition, mice that carry multiple copies of *PAX6* on a wild-type background were found to have eye abnormalities which highlight the need for a correct gene dosage for normal eye development [188]. Similarly, microscopic or submicroscopic deletions of both *PITX2* and *FOXC1* as well as duplication of 6p25 encompassing *FOXC1* are known to cause Axenfeld-Rieger malformations (ARM) affecting structures in the anterior eye segment [186, 193-195].

Copy number variation studies have also been used to identify loci for congenital eye malformations, including deletion of 8q12 in CHARGE syndrome and identification of *CHD7* [196], 8q21.2-q22.1 deletion in a patient with chorioretinal coloboma and the discovery of *GDF6* [135]. Also a 2.7 Mb deletion at chromosome 18q22.1in a microphthalmic patient, and 6p24.3 deletion in branchio-oculo facial syndrome led to the identification of *TMX3* and *TFARP2A*, respectively [152, 197].

A recent study reported the use of comparative genomic hybridization to detect copy number abnormalities in genes responsible for MAC. In a cohort of 32 patients, they detected 7 nonpolymorphic copy number changes, amongst which the duplication in *PAX2* gene in a patient with renal coloboma seemed to be associated with the patient's phenotype while the rest of the changes appeared to be clinically insignificant [90].

Genetic approaches to identify disease genes in MAC

As described in the previous section, the molecular basis of MAC is still poorly understood, although a large genetic component is suspected in the aetiology of MAC. In order to understand the mechanisms underlying the molecular basis of these disorders, an important step is to identify the entire set of disease-associated genes. A variety of approaches have been used in the attempt to map and identify genes causing MAC, including linkage analysis, candidate gene screening, and homozygosity mapping. Each of these methods has its own strengths and weaknesses.

Linkage mapping

Despite genetic and phenotypic heterogeneity, linkage analysis has been extensively used in the identification of several potential DNA regions for MAC cases where these disorders segregate in a Mendelian fashion. Examples are mapping syndromic clinical anophthalmia to 15q23-q25.1 locus and discovery of *STRA6* [198], mapping bilateral microphthalmia, cataract, and coloboma to 14q24.3 locus and discovery of *CHX10* [199], and linkage to Xp21.2-p11.4 loci in a family with Lenz microphthalmia syndrome and identification of *BCOR* [200-201].

This technique basically looks for the cosegregation of alleles for genetic markers with a disease phenotype in a family (Figure 1.7). If two markers were close together, a recombination event would be unlikely to separate them through meiosis; thus, these markers tend to be transmitted as a block through a pedigree.

In order to yield positive results in linkage scan, at least 10 informative meioses is required. A meiosis is informative if it can be identified whether or not the gamete is recombinant. The proportion of gametes that are recombinant is called recombination fraction (θ) which can vary between 0 and 0.5 in case of complete linkage and no linkage, respectively.

$$\text{Recombination fraction } (\theta) = \frac{\text{Recombinant meioses}}{\text{Recombinant meioses + Nonrecombinant Meioses}}$$

Evidence for or against linkage between a disease locus and a marker can be statistically estimated by the logarithm of the odds (LOD score) which is calculated as the ratio of obtaining the test data if the two loci are indeed linked, to the likelihood of observing the same data purely by chance. Positive LOD scores (>3) favor the presence of linkage, whereas negative LOD scores (<-2) indicate that linkage is less likely.

$$\text{LOD score} = \frac{\text{likelihood if recombination fraction } \theta = \theta_1}{\text{likelihood if recombination fraction } \theta = 0.5}$$

Performing linkage analysis requires genetic markers, which are polymorphic DNA sequences with known location on a chromosome, to enable us to follow a chromosomal segment through a pedigree. Different types of genetic marker are available (Table1.1), but microsatellite markers and single nucleotide

polymorphisms (SNP) are commonly used because they can be scored easily and cheaply using readily available material (e.g. blood cells).

Linkage studies being performed on two markers at the same time, refers to two point linkage analysis which is a first step in the detection of potential disease loci. Since the chromosomal segments identified by two point analysis are fairly large, multipoint linkage analysis can be used to evaluate the linkage of multiple markers to a disease in a previously identified locus [202]. If these calculations are carried out under the assumption of a specific mode of inheritance for the trait locus, linkage analysis is referred to as parametric. In contrast, non-parametric linkage analysis does not require a genetic model and is based on shared chromosomal segments in affected individuals.

Pitfalls of linkage analysis

Linkage analysis is standard method for identifying disease gene; however, it may run into problem. The major pitfalls are:

- Prone to error: errors can happen during sampling, genotyping, evaluating the phenotype, assumption of the mode of inheritance and the presence of genetic and locus heterogeneity.
- Pedigree size limitation: The resolution of linkage mapping is limited by the number of available meioses. For example in a simple autosomal recessive disease each affected individual contributes ~ 0.3 to the LOD score, so requiring at least 10 informative meioses to obtain LOD score of 3 (Table 1.2) [203].

Figure 1.7 Schematic presentation of the concept of linkage mapping, derived from Lidral AC *et al.*, Birth Defects Res A Clin Mol Teratol **70**(12):893-901. A hypothetical marker co-segregate with the disease status in this family, as shown by the occurrence of the 1 allele in all affected individuals and never in any of the unaffected individuals [204]. Black circles or squares represent individuals with the disease phenotype. White circles or squares represent unaffected individuals.

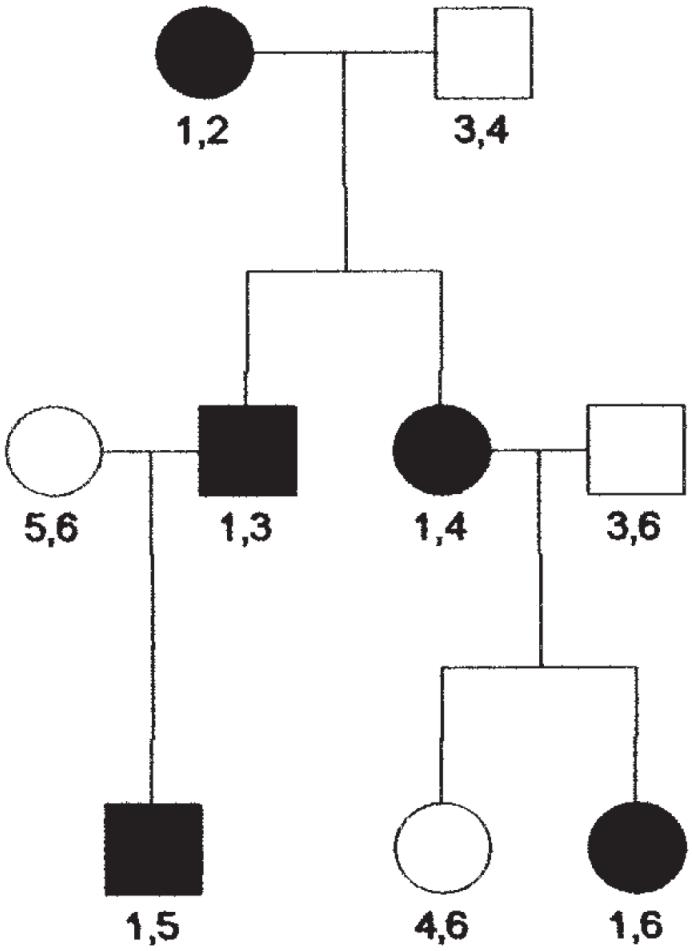


Table1.3 Development of human genetic markers. (Strachan and Read 2004).

Type of marker	No. Of Loci	Features
Blood groups (1910-1960)	~ 20	May need fresh blood, rare antisera Genotype not always inferred from phenotype because of dominance
Electrophoretic mobility variants of serum proteins (1960-75)	~30	May need fresh serum, specialized assays, limited polymorphism
HLA tissue types (1970-)	1 (Haplotype)	One linked set, highly informative, can only be tested for linkage to 6p21.3
DNA RFLPs (1975-)	$> 10^5$ (potentially)	Two allele markers, maximum heterozygosity 0.5, initially required Southern blotting but no PCR, Easy physical localization
Minisatellites (VNTRs) (1985-)	$> 10^4$ (potentially)	Many alleles, highly informative, type by Southern blotting, easy physical localization, tend to cluster near end of chromosomes
Microsatellites (STRs) (1989-)	$> 10^5$ (potentially)	Many alleles, highly informative, type by automated multiplex PCR, easy physical localization, Distributed throughout genome
Single Nucleotide Polymorphisms (SNPs)	$> 4 \times 10^6$	Less informative than Microsatellites, Can be typed on a very large scale by automated equipments without gel electrophoresis

Table 1.4 Number of phase-known, fully informative meioses needed to detect linkage at various value of recombination fraction θ [203].

θ	**Number of informative meiosis needed for LOD = 3**	**Expected LOD per meiosis**
0.00	10	0.30
0.01	11	0.28
0.02	12	0.26
0.05	14	0.21
0.10	19	0.16
0.20	36	0.08

Candidate gene approach

As opposed to linkage mapping, which is an unbiased search of the entire genome without any knowledge of the underlying mechanisms of the disease, candidate gene approach look for certain genes based on their expression pattern and/or their known biochemical functions in association with the biology of the disease being investigated [205]. Examples of candidate genes that are likely to be implicated in human eye disease are *SIX6* [146], *HES1*[206], *RX* [126], and the *EYA* [207] which primarily identified through their function in mammalian eye development. Nevertheless, there are some exceptions in that some genes may have different function in different species. For example, mouse mutants for Mitf are microphthalmic, but mutation in orthologous gene in human causes Waardenburg syndrome [208-209]. Likewise, there are too many genes with plausible function in eye development while there are no valid selection criteria. Therefore, this approach would not be the best one in identification of novel genes in MAC [45].

Homozygosity mapping

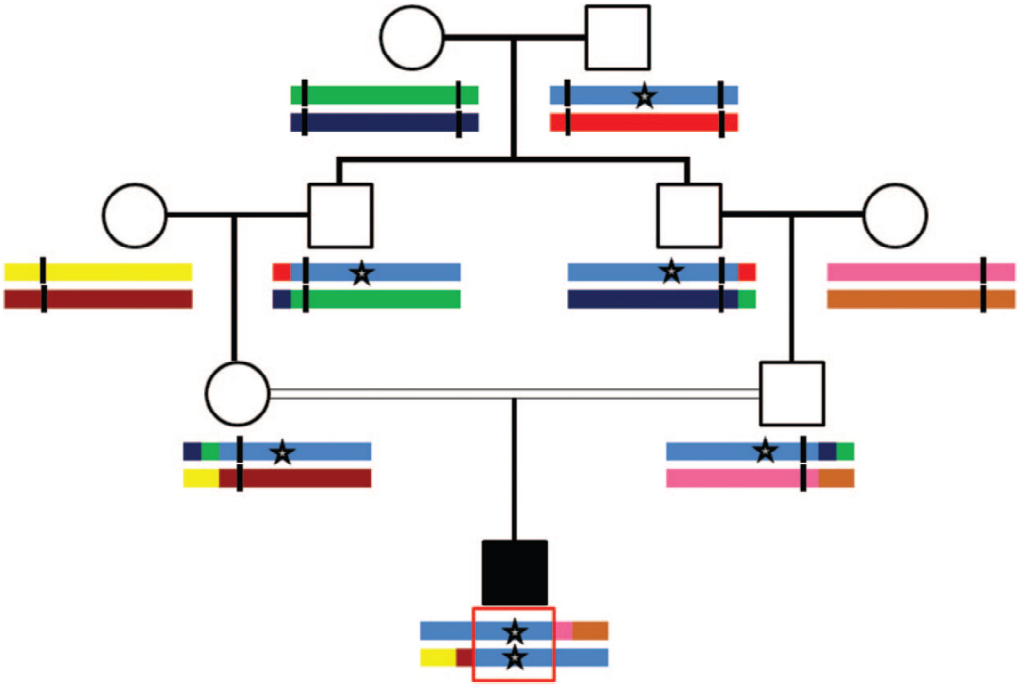
Homozygosity mapping in consanguineous families was proposed in 1987 and is a strong method for locating the genes associated with rare recessive diseases especially when families available for study are limited in number and size [210]. The idea behind the homozygosity mapping is that the offspring of consanguineous unions inherit both copies of pathogenic allele from a single common ancestor (Figure 1.8). In addition to the disease allele, autosomal DNA regions in the vicinity of the mutation are likely to be homozygous by descent too.

This makes it possible to map the disease locus by genome-wide searching for the regions of the genome that are homozygous in all affected individuals or identical by descent (IBD) [211].

The use of this theory in practice awaited the construction of the first whole genome map of microsatellite which led to the successful mapping of the disease loci for Friedreich ataxia and alkaptonuria to 8q and 3q2, respectively. Recently, the development of high density SNP platforms has provided an extremely simple and quick way to map autozygous segments, consequently, paved the way for identification of several disease-related variants [212].

To consider any homozygous segment as IBD, the size of the segment and the density of the marker within that region do matter. Evidently, any two copies of the DNA segment might be homozygous without being reflective of a common ancestor. These segments are called identical by state (IBS) and might be independently introduced to the genetic pool of the population. In order to distinguish between IBD and IBS, most homozygosity mapping programs use the threshold of 1 megabase (Mb) for the size [213]. This is because the probability for a segment to be homozygous only by chance is adversely correlated with the size of that segment. Likewise, the number and the level of heterozygosity of the markers are important. For example, a segment with only one marker with the heterozygosity of 0.5 has a 50% chance of being homozygous by chance in any two individuals in the population. On the other hand, 3 markers with heterozygosity of 0.25 would be expected to be homozygous by chance with the

Figure 1.8 Schematic presentation of the concept of identical by descent derived from Alkuraya FS *et al.*, Genet Med **12**(12):765-71. The pedigree showing a first-cousin marriage producing a child affected with a recessive disease, due to homozygosity by descent for a disease-causing allele [213].



probability of (1- 0.25) (1- 0.25) (1-0.25) = 42%. Therefore, markers with low levels of heterozygosity can be informative at high density. The fact that SNPs are highly frequent in the genome, with a range of 10^4-10^6, as compared to ~ 400 microsatellite panel, has made them a preferred choice for markers despite their low level of heterozygosity [211, 213].

How much coverage is required? The answer depends on the degree of consanguinity of the family being studied. If the mutation is introduced in a pedigree very early, recombination events might lead to breakage of the ancestral haplotype, therefore, higher density of markers is required to capture such a segment [214-216]. For example if we take 5% as cut-off for type I error, one would need 41 consecutive homozygous SNPs (with accepted heterozygosity of 0.35) on a 250,000 platform (which is equal to 0.5 Mb of DNA) to call a homozygous segment reliable $(1\text{-}0.35)^{x}$ 250,000= 0.05, X= 41. Nevertheless, this equation is correct if SNPs segregate independently but it is not due to linkage disequilibrium indicating that SNPs that are close together are likely to segregate together [217]. Therefore, 1-2 Mb cut-off is usually considered as a realistic sensitivity in homozygosity mapping [215, 218].

Limitations of homozygosity mapping

While homozygosity mapping is an efficient approach for mapping recessive genes in inbred families, the presence of both genetic and allelic heterogeneity might complicate the analysis and lead to false positive results [219-226]. As a case in point, Benayoun *et al*. (2009) reported the failure of homozygosity mapping in detection of causative loci in two extended families with retinitis

pigmentosa (RP). In one family, they failed to find any shared homozygous genomic interval among three affected siblings. However, analysis of all known loci for RP revealed that all three individuals were compound heterozygote for *CRB1* gene, indicating the presence of allelic heterogeneity. On the other hand, in the second family haplotype analysis around the known loci for RP revealed homozygous mutation in *RDH12* but only in 14 of 17 affected individuals, which indicates the presence of genetic heterogeneity [220]. For some diseases, such heterogeneity may occur even within the same consanguineous family [226].

Moreover, it is important to know the degree of relatedness of the family under study because it has been shown that the presence of hidden consanguinity in a pedigree would result in a higher degree of homozygosity than expected based on the degree of relatedness of the parents [227]. In some cases, background inbreeding information is not reflected in the abbreviated pedigrees which can potentially lead to the identification of false positive loci and reduces the significance of the LOD score calculations [228-230].

Similar to linkage studies in small families, homozygosity mapping can detect several large intervals some of which may be unrelated to the disease [213]. These IBSs intervals can be problematic especially when dealing with smaller segments, because as mentioned earlier, the length of homozygous regions are inversely correlated with the probability of sharing two haplotypes by chance [215]. Regardless of the IBD or IBS status of the identified putative intervals, these blocks may contain hundreds of genes, amongst which the identification of

causative variants requires extensive in silico analysis of candidate genes in the region. Nevertheless, the usefulness of these bioinformatics tools in real time is limited due to our incomplete understanding of the human biology [213, 231-233].

Aim of the study

We recently ascertained a consanguineous family in which several individuals are affected by anophthalmia. We hypothesized that the phenotype of affected siblings is suggestive of a rare autosomal recessive disorder and used Single Nucleotide Polymorphism (SNP) arrays to genotype family members. The aims of the research presented in this thesis were:

1. To identify disease-loci for anophthalmia in this pedigree by performing large-scale homozygosity mapping and to find candidate gene(s) implicated in this disorder by mutational analysis. This work is described in chapter 2.
2. To investigate copy number changes in the genome of affected individuals and to determine the significance of genes contained within the affected genomic regions. This work is described in chapter 2.
3. To investigate the presence of all possible exonic variants in the genome of one affected individual by exome sequencing. This work is described in chapter 3.
4. An overview of the findings, some limitations of this study and future directions are discussed in chapter 4.

References:

1. Robinson GC, Jan JE, Kinnis C. Congenital ocular blindness in children, 1945 to 1984. Am J Dis Child 1987; 141(12):1321-4.

2. Costa T, Scriver CR, Childs B. The effect of Mendelian disease on human health: a measurement. Am J Med Genet 1985; 21(2):231-42.

3. Graw J. Eye development. Curr Top Dev Biol 2010; 90:343-86.

4. O'Rahily R MF. Developmental stages in human embryos. Washington, DC, Carnegie Institution 1987.

5. Pei YF, Rhodin JA. The prenatal development of the mouse eye. Anat Rec 1970; 168(1):105-25.

6. Zygar CA, Cook TL, Grainger RM, Jr. Gene activation during early stages of lens induction in Xenopus. Development 1998; 125(17):3509-19.

7. Edward DP, Kaufman LM. Anatomy, development, and physiology of the visual system. Pediatr Clin North Am 2003; 50(1):1-23.

8. van Heyningen V. Developmental eye disease--a genome era paradigm. Clin Genet 1998; 54(4):272-82.

9. Cepko CL, Austin CP, Yang X, Alexiades M, Ezzeddine D. Cell fate determination in the vertebrate retina. Proc Natl Acad Sci U S A 1996; 93(2):589-95.

10. Chow RL, Lang RA. Early eye development in vertebrates. Annu Rev Cell Dev Biol 2001; 17:255-96.

11. Gehring WJ, Ikeo K. Pax 6: mastering eye morphogenesis and eye evolution. Trends Genet 1999; 15(9):371-7.

12. Mathers PH, Jamrich M. Regulation of eye formation by the Rx and pax6 homeobox genes. Cell Mol Life Sci 2000; 57(2):186-94.

13. Rembold M, Loosli F, Adams RJ, Wittbrodt J. Individual cell migration serves as the driving force for optic vesicle evagination. Science 2006; 313(5790):1130-4.

14. Harada T, Harada C, Parada LF. Molecular regulation of visual system development: more than meets the eye. Genes Dev 2007; 21(4):367-78.

15. Martinez-Morales JR, Wittbrodt J. Shaping the vertebrate eye. Curr Opin Genet Dev 2009; 19(5):511-7.

16. Zuber ME, Perron M, Philpott A, Bang A, Harris WA. Giant eyes in Xenopus laevis by overexpression of XOptx2. Cell 1999; 98(3):341-52.

17. Zhang L, Mathers PH, Jamrich M. Function of Rx, but not Pax6, is essential for the formation of retinal progenitor cells in mice. Genesis 2000; 28(3-4):135-42.

18. Furukawa T, Kozak CA, Cepko CL. rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina. Proc Natl Acad Sci U S A 1997; 94(7):3088-93.

19. Koroma BM, Yang JM, Sundin OH. The Pax-6 homeobox gene is expressed throughout the corneal and conjunctival epithelia. Invest Ophthalmol Vis Sci 1997; 38(1):108-20.

20. Gehring WJ. The master control gene for morphogenesis and evolution of the eye. Genes Cells 1996; 1(1):11-5.

21. Treisman J, Harris E, Desplan C. The paired box encodes a second DNA-binding domain in the paired homeo domain protein. Genes Dev 1991; 5(4):594-604.

22. Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. Nat Genet 1994; 7(4):463-71.

23. Hill RE, Favor J, Hogan BL, Ton CC, Saunders GF, Hanson IM, Prosser J, Jordan T, Hastie ND, van Heyningen V. Mouse small eye results from mutations in a paired-like homeobox-containing gene. Nature 1991; 354(6354):522-5.

24. Zhu CC, Dyer MA, Uchikawa M, Kondoh H, Lagutin OV, Oliver G. Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. Development 2002; 129(12):2835-49.

25. Oliver G, Loosli F, Koster R, Wittbrodt J, Gruss P. Ectopic lens induction in fish in response to the murine homeobox gene Six3. Mech Dev 1996; 60(2):233-9.

26. Martinez-Morales JR, Signore M, Acampora D, Simeone A, Bovolenta P. Otx genes are required for tissue specification in the developing eye. Development 2001; 128(11):2019-30.

27. Liu IS, Chen JD, Ploder L, Vidgen D, van der Kooy D, Kalnins VI, McInnes RR. Developmental expression of a novel murine homeobox gene (Chx10): evidence for roles in determination of the neuroretina and inner nuclear layer. Neuron 1994; 13(2):377-93.

28. Horsford DJ, Nguyen MT, Sellar GC, Kothary R, Arnheiter H, McInnes RR. Chx10 repression of Mitf is required for the maintenance of mammalian neuroretinal identity. Development 2005; 132(1):177-87.

29. Burmeister M, Novak J, Liang MY, Basu S, Ploder L, Hawes NL, Vidgen D, Hoover F, Goldman D, Kalnins VI, Roderick TH, Taylor BA, Hankin MH, McInnes RR. Ocular retardation mouse caused by Chx10 homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation. Nat Genet 1996; 12(4):376-84.

30. Bora N, Defoe D, Smith SB. Evidence of decreased adhesion between the neural retina and retinal pigmented epithelium of the Mitfvit (vitiligo) mutant mouse. Cell Tissue Res 1999; 295(1):65-75.

31. Yasumoto K, Amae S, Udono T, Fuse N, Takeda K, Shibahara S. A big gene linked to small eyes encodes multiple Mitf isoforms: many promoters make light work. Pigment Cell Res 1998; 11(6):329-36.

32. Molotkov A, Molotkova N, Duester G. Retinoic acid guides eye morphogenetic movements via paracrine signaling but is unnecessary for retinal dorsoventral patterning. Development 2006; 133(10):1901-10.

33. Mic FA, Molotkov A, Molotkova N, Duester G. Raldh2 expression in optic vesicle generates a retinoic acid signal needed for invagination of retina during optic cup formation. Dev Dyn 2004; 231(2):270-7.

34. Duester G. Keeping an eye on retinoic acid signaling during eye development. Chem Biol Interact 2009; 178(1-3):178-81.

35. Graw J. The genetic and molecular basis of congenital eye defects. Nat Rev Genet 2003; 4(11):876-88.

36. Levin AV. Congenital eye anomalies. Pediatr Clin North Am 2003; 50(1):55-76.

37. Dollfus H, Verloes A. Dysmorphology and the orbital region: a practical clinical approach. Surv Ophthalmol 2004; 49(6):547-61.

38. Epstein CJ. The new dysmorphology: application of insights from basic developmental biology to the understanding of human birth defects. Proc Natl Acad Sci U S A 1995; 92(19):8566-73.

39. Guercio JR, Martyn LJ. Congenital malformations of the eye and orbit. Otolaryngol Clin North Am 2007; 40(1):113-40, vii.

40. Verma AS, Fitzpatrick DR. Anophthalmia and microphthalmia. Orphanet J Rare Dis 2007; 2:47.

41. Gilbert C, Foster A. Childhood blindness in the context of VISION 2020--the right to sight. Bull World Health Organ 2001; 79(3):227-32.

42. Hornby SJ, Gilbert CE, Rahi JK, Sil AK, Xiao Y, Dandona L, Foster A. Regional variation in blindness in children due to microphthalmos, anophthalmos and coloboma. Ophthalmic Epidemiol 2000; 7(2):127-38.

43. Ragge NK, Subak-Sharpe ID, Collin JR. A practical guide to the management of anophthalmia and microphthalmia. Eye (Lond) 2007; 21(10):1290-300.

44. Fitzpatrick DR, van Heyningen V. Developmental eye disorders. Curr Opin Genet Dev 2005; 15(3):348-53.

45. Morrison D, FitzPatrick D, Hanson I, Williamson K, van Heyningen V, Fleck B, Jones I, Chalmers J, Campbell H. National study of microphthalmia, anophthalmia, and coloboma (MAC) in Scotland: investigation of genetic aetiology. J Med Genet 2002; 39(1):16-22.

46. Forrester MB, Merz RD. Descriptive epidemiology of anophthalmia and microphthalmia, Hawaii, 1986-2001. Birth Defects Res A Clin Mol Teratol 2006; 76(3):187-92.

47. Quaranta-Leoni FM. Congenital anophthalmia: current concepts in management. Curr Opin Ophthalmol 2011; 22(5):380-4.

48. Bakrania P, Efthymiou M, Klein JC, Salt A, Bunyan DJ, Wyatt A, Ponting CP, Martin A, Williams S, Lindley V, Gilmore J, Restori M, Robson AG, Neveu MM, Holder GE, Collin JR, Robinson DO, Farndon P, Johansen-Berg H, Gerrelli D, Ragge NK. Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between the BMP4 and hedgehog signaling pathways. Am J Hum Genet 2008; 82(2):304-19.

49. Shah SP, Taylor AE, Sowden JC, Ragge NK, Russell-Eggitt I, Rahi JS, Gilbert CE. Anophthalmos, microphthalmos, and typical coloboma in the United Kingdom: a prospective study of incidence and risk. Invest Ophthalmol Vis Sci 2011; 52(1):558-64.

50. Schittkowski MP, Guthoff RF. Systemic and ophthalmological anomalies in congenital anophthalmic or microphthalmic patients. Br J Ophthalmol 2010; 94(4):487-93.

51. Fryns JP. Fryns syndrome: a variable MCA syndrome with diaphragmatic defects, coarse face, and distal limb hypoplasia. J Med Genet 1987; 24(5):271-4.

52. Seller MJ, Davis TB, Fear CN, Flinter FA, Ellis I, Gibson AG. Two sibs with anophthalmia and pulmonary hypoplasia (the Matthew-Wood syndrome). Am J Med Genet 1996; 62(3):227-29.

53. Sandler D, Mancuso A, Becker T, Zori R, Hellrung J, Silverstein J, Burton V, Hamosh A, Williams C. Association of anophthalmia and esophageal atresia. Am J Med Genet 1995; 59(4):484-91.

54. Traboulsi EI, Nasr AM, Fahd SD, Jabbour NM, Der Kaloustian VM. Waardenburg's recessive anophthalmia syndrome. Ophthalmic Paediatr Genet 1984; 4(1):13-8.

55. Ahmad ME, Dada R, Dada T, Kucheria K. 14q(22) deletion in a familial case of anophthalmia with polydactyly. Am J Med Genet A 2003; 120A(1):117-22.

56. Guion-Almeida ML, Richieri-Costa A. New syndrome of growth and mental retardation, structural anomalies of the central nervous system, and first branchial arch, anophthalmia, heminasal a/hypoplasia, and atypical clefting: report on four Brazilian patients. Am J Med Genet 1999; 87(3):237-44.

57. Moog U, Kruger G, Stengel B, De Die-Smulders C, Dykstra S, Bleeker-Wagemakers E. Oculocerebrocutaneous syndrome: a case report, a follow-up, and differential diagnostic considerations. Genet Couns 1996; 7(4):257-65.

58. Hesselberg C. Congenital bilateral anophthalmia. Acta Ophthalmol (Copenh) 1951; 29(2):183-9.

59. Rosenbaum J. Congenital bilateral anophthalmos. Can Med Assoc J 1919; 9(3):254-6.

60. Fuhrmann S. Eye morphogenesis and patterning of the optic vesicle. Curr Top Dev Biol 2010; 93:61-84.

61. Brian P. Brooks EIT. Duane's Ophthalmology: Lippincott Williams & Wilkins; 2006.

62. Gujar SK, Gandhi D. Congenital malformations of the orbit. Neuroimaging Clin N Am 2011; 21(3):585-602, viii.

63. Traboulsi AVRaEI. Congenital Ocular Malformations. Pediatric Ophthalmology 2009.

64. Schneider A, Bardakjian T, Reis LM, Tyler RC, Semina EV. Novel SOX2 mutations and genotype-phenotype correlation in anophthalmia and microphthalmia. Am J Med Genet A 2009; 149A(12):2706-15.

65. Onwochei BC, Simon JW, Bateman JB, Couture KC, Mir E. Ocular colobomata. Surv Ophthalmol 2000; 45(3):175-94.

66. Gregory-Evans CY, Williams MJ, Halford S, Gregory-Evans K. Ocular coloboma: a reassessment in the age of molecular neuroscience. J Med Genet 2004; 41(12):881-91.

67. Warburg M. Classification of microphthalmos and coloboma. J Med Genet 1993; 30(8):664-9.

68. I M. Typical colobomata (retina, chroid and iris). Developmental abnormalities of the eye. Philadelphia: Lippincott; 1957. p. 74- 93.

69. Pagon RA. Ocular coloboma. Surv Ophthalmol 1981; 25(4):223-36.

70. Lu BX. [An analysis of 193 cases of congenital intraocular colobomas]. Zhonghua Yan Ke Za Zhi 1989; 25(6):357-9.

71. Porges Y, Gershoni-Baruch R, Leibu R, Goldscher D, Zonis S, Shapira I, Miller B. Hereditary microphthalmia with colobomatous cyst. Am J Ophthalmol 1992; 114(1):30-4.

72. Stoll C, Alembik Y, Dott B, Roth MP. Epidemiology of congenital eye malformations in 131,760 consecutive births. Ophthalmic Paediatr Genet 1992; 13(3):179-86.

73. Bermejo E, Martinez-Frias ML. Congenital eye malformations: clinical-epidemiological analysis of 1,124,654 consecutive births in Spain. Am J Med Genet 1998; 75(5):497-504.

74. Vogt G, Puho E, Czeizel AE. A population-based case-control study of isolated ocular coloboma. Ophthalmic Epidemiol 2005; 12(3):191-7.

75. EI. T. Colobomatous microphthalmia, anophthalmia, and associated malformation syndromes. New York: Oxford University Press; 1999.

76. Chang L, Blain D, Bertuzzi S, Brooks BP. Uveal coloboma: clinical and basic science update. Curr Opin Ophthalmol 2006; 17(5):447-70.

77. Mui SH, Kim JW, Lemke G, Bertuzzi S. Vax genes ventralize the embryonic eye. Genes Dev 2005; 19(10):1249-59.

78. Fraser FC, Skelton J. Possible teratogenicity of maternal fever. Lancet 1978; 2(8090):634.

79. Spagnolo A, Bianchi F, Calabro A, Calzolari E, Clementi M, Mastroiacovo P, Meli P, Petrelli G, Tenconi R. Anophthalmia and benomyl in Italy: a multicenter study based on 940,615 newborns. Reprod Toxicol 1994; 8(5):397-403.

80. Hornby SJ, Ward SJ, Gilbert CE. Eye birth defects in humans may be caused by a recessively-inherited genetic predisposition to the effects of maternal vitamin A deficiency during pregnancy. Med Sci Monit 2003; 9(11):HY23-6.

81. Hornby SJ, Ward SJ, Gilbert CE, Dandona L, Foster A, Jones RB. Environmental risk factors in congenital malformations of the eye. Ann Trop Paediatr 2002; 22(1):67-77.

82. Khandekar R, Al Awaidy S, Ganesh A, Bawikar S. An epidemiological and clinical study of ocular manifestations of congenital rubella syndrome in Omani children. Arch Ophthalmol 2004; 122(4):541-5.

83. Andreou A, Basiakos H, Hatzikoumi I, Lazarides A. Fetal varicella syndrome with manifestations limited to the eye. Am J Perinatol 1995; 12(5):347-8.

84. Hittner HM, Desmond MM, Montgomery JR. Optic nerve manifestations of human congenital cytomegalovirus infection. Am J Ophthalmol 1976; 81(5):661-5.

85. Suhardjo, Utomo PT, Agni AN. Clinical manifestations of ocular toxoplasmosis in Yogyakarta, Indonesia: a clinical review of 173 cases. Southeast Asian J Trop Med Public Health 2003; 34(2):291-7.

86. Miller MT, Stromland K. Teratogen update: thalidomide: a review, with a focus on ocular findings and new potential uses. Teratology 1999; 60(5):306-21.

87. Stromland K, Pinazo-Duran MD. Ophthalmic involvement in the fetal alcohol syndrome: clinical and animal model studies. Alcohol Alcohol 2002; 37(1):2-8.

88. Rahi JS, Sripathi S, Gilbert CE, Foster A. Childhood blindness due to vitamin A deficiency in India: regional variations. Arch Dis Child 1995; 72(4):330-3.

89. Bardakjian TM, Schneider A. The genetics of anophthalmia and microphthalmia. Curr Opin Ophthalmol 2011; 22(5):309-13.

90. Raca G, Jackson CA, Kucinskas L, Warman B, Shieh JT, Schneider A, Bardakjian TM, Schimmenti LA. Array comparative genomic hybridization analysis in patients with anophthalmia, microphthalmia, and coloboma. Genet Med 2011; 13(5):437-42.

91. Wyatt A, Bakrania P, Bunyan DJ, Osborne RJ, Crolla JA, Salt A, Ayuso C, Newbury-Ecob R, Abou-Rayyah Y, Collin JR, Robinson D, Ragge N. Novel heterozygous OTX2 mutations and whole gene deletions in anophthalmia, microphthalmia and coloboma. Hum Mutat 2008; 29(11):E278-83.

92. Bardakjian TM, Kwok S, Slavotinek AM, Schneider AS. Clinical report of microphthalmia and optic nerve coloboma associated with a de novo

microdeletion of chromosome 16p11.2. Am J Med Genet A 2010; 152A(12):3120-3.

93. Kouassi FX, Koffi KV, Safede K, Cochard C, Cochener B. [Congenital anophthalmias: a case of trisomy 13]. J Fr Ophtalmol 2006; 29(4):e10.

94. Kallen B, Robert E, Harris J. The descriptive epidemiology of anophthalmia and microphthalmia. Int J Epidemiol 1996; 25(5):1009-16.

95. Ginsberg J, Soukup S, Bendon RW. Further observations of ocular pathology in trisomy 9. J Pediatr Ophthalmol Strabismus 1989; 26(3):146-9.

96. Hoover-Fong JE, Cai J, Cargile CB, Thomas GH, Patel A, Griffin CA, Jabs EW, Hamosh A. Facial dysgenesis: a novel facial syndrome with chromosome 7 deletion p15.1-21.1. Am J Med Genet A 2003; 117A(1):47-56.

97. Bessant DA, Khaliq S, Hameed A, Anwar K, Mehdi SQ, Payne AM, Bhattacharya SS. A locus for autosomal recessive congenital microphthalmia maps to chromosome 14q32. Am J Hum Genet 1998; 62(5):1113-6.

98. Gallardo ME, Lopez-Rios J, Fernaud-Espinosa I, Granadino B, Sanz R, Ramos C, Ayuso C, Seller MJ, Brunner HG, Bovolenta P, Rodriguez de Cordoba S. Genomic cloning and characterization of the human homeobox gene SIX6 reveals a cluster of SIX genes in chromosome 14 and associates SIX6 hemizygosity with bilateral anophthalmia and pituitary anomalies. Genomics 1999; 61(1):82-91.

99. Desmaison A, Vigouroux A, Rieubland C, Peres C, Calvas P, Chassaing N. Mutations in the LHX2 gene are not a frequent cause of micro/anophthalmia. Mol Vis 2010; 16:2847-9.

100. Lemyre E, Lemieux N, Decarie JC, Lambert M. Del(14)(q22.1q23.2) in a patient with anophthalmia and pituitary hypoplasia. Am J Med Genet 1998; 77(2):162-5.

101. Gonzalez-Rodriguez J, Pelcastre EL, Tovilla-Canales JL, Garcia-Ortiz JE, Amato-Almanza M, Villanueva-Mendoza C, Espinosa-Mattar Z, Zenteno JC. Mutational screening of CHX10, GDF6, OTX2, RAX and SOX2 genes in 50 unrelated microphthalmia-anophthalmia-coloboma (MAC) spectrum cases. Br J Ophthalmol 2010; 94(8):1100-4.

102. Fantes J, Ragge NK, Lynch SA, McGill NI, Collin JR, Howard-Peebles PN, Hayward C, Vivian AJ, Williamson K, van Heyningen V, FitzPatrick DR. Mutations in SOX2 cause anophthalmia. Nat Genet 2003; 33(4):461-3.

103. Voronina VA, Kozhemyakina EA, O'Kernick CM, Kahn ND, Wenger SL, Linberg JV, Schneider AS, Mathers PH. Mutations in the human RAX homeobox gene in a patient with anophthalmia and sclerocornea. Hum Mol Genet 2004; 13(3):315-22.

104. Ragge NK, Brown AG, Poloschek CM, Lorenz B, Henderson RA, Clarke MP, Russell-Eggitt I, Fielder A, Gerrelli D, Martinez-Barbera JP, Ruddle P, Hurst J, Collin JR, Salt A, Cooper ST, Thompson PJ, Sisodiya SM, Williamson KA, Fitzpatrick DR, van Heyningen V, Hanson IM. Heterozygous mutations of OTX2 cause severe ocular malformations. Am J Hum Genet 2005; 76(6):1008-22.

105. Schimmenti LA, Cunliffe HE, McNoe LA, Ward TA, French MC, Shim HH, Zhang YH, Proesmans W, Leys A, Byerly KA, Braddock SR, Masuno M, Imaizumi K, Devriendt K, Eccles MR. Further delineation of renal-coloboma syndrome in patients with extreme variability of phenotype and identical PAX2 mutations. Am J Hum Genet 1997; 60(4):869-78.

106. Sanyanusin P, Schimmenti LA, McNoe TA, Ward TA, Pierpont ME, Sullivan MJ, Dobyns WB, Eccles MR. Mutation of the gene in a family with optic nerve colobomas, renal anomolies and vesicoureteral reflux. Nat Genet 1996; 13(1):129.

107. Ragge NK, Lorenz B, Schneider A, Bushby K, de Sanctis L, de Sanctis U, Salt A, Collin JR, Vivian AJ, Free SL, Thompson P, Williamson KA, Sisodiya SM, van Heyningen V, Fitzpatrick DR. SOX2 anophthalmia syndrome. Am J Med Genet A 2005; 135(1):1-7; discussion 8.

108. Bakrania P, Robinson DO, Bunyan DJ, Salt A, Martin A, Crolla JA, Wyatt A, Fielder A, Ainsworth J, Moore A, Read S, Uddin J, Laws D, Pascuel-Salcedo D, Ayuso C, Allen L, Collin JR, Ragge NK. SOX2 anophthalmia syndrome: 12 new cases demonstrating broader phenotype and high frequency of large gene deletions. Br J Ophthalmol 2007; 91(11):1471-6.

109. Reis LM, Tyler RC, Schneider A, Bardakjian T, Semina EV. Examination of SOX2 in variable ocular conditions identifies a recurrent deletion in microphthalmia and lack of mutations in other phenotypes. Mol Vis 2010; 16:768-73.

110. Kelberman D, Rizzoti K, Avilion A, Bitner-Glindzicz M, Cianfarani S, Collins J, Chong WK, Kirk JM, Achermann JC, Ross R, Carmignac D, Lovell-

Badge R, Robinson IC, Dattani MT. Mutations within Sox2/SOX2 are associated with abnormalities in the hypothalamo-pituitary-gonadal axis in mice and humans. J Clin Invest 2006; 116(9):2442-55.

111. Wang P, Liang X, Yi J, Zhang Q. Novel SOX2 mutation associated with ocular coloboma in a Chinese family. Arch Ophthalmol 2008; 126(5):709-13.

112. Mihelec M, Abraham P, Gibson K, Krowka R, Susman R, Storen R, Chen Y, Donald J, Tam PP, Grigg JR, Flaherty M, Gole GA, Jamieson RV. Novel SOX2 partner-factor domain mutation in a four-generation family. Eur J Hum Genet 2009; 17(11):1417-22.

113. Taranova OV, Magness ST, Fagan BM, Wu Y, Surzenko N, Hutton SR, Pevny LH. SOX2 is a dose-dependent regulator of retinal neural progenitor competence. Genes Dev 2006; 20(9):1187-202.

114. Chassaing N, Gilbert-Dussardier B, Nicot F, Fermeaux V, Encha-Razavi F, Fiorenza M, Toutain A, Calvas P. Germinal mosaicism and familial recurrence of a SOX2 mutation with highly variable phenotypic expression extending from AEG syndrome to absence of ocular involvement. Am J Med Genet A 2007; 143(3):289-91.

115. Stark Z, Storen R, Bennetts B, Savarirayan R, Jamieson RV. Isolated hypogonadotropic hypogonadism with SOX2 mutation and anophthalmia/microphthalmia in offspring. Eur J Hum Genet 2011; 19(7):753-6.

116. Faivre L, Williamson KA, Faber V, Laurent N, Grimaldi M, Thauvin-Robinet C, Durand C, Mugneret F, Gouyon JB, Bron A, Huet F, Hayward C, Heyningen V, Fitzpatrick DR. Recurrence of SOX2 anophthalmia syndrome with gonosomal mosaicism in a phenotypically normal mother. Am J Med Genet A 2006; 140(6):636-9.

117. Kondoh H, Uchikawa M, Kamachi Y. Interplay of Pax6 and SOX2 in lens development as a paradigm of genetic switch mechanisms for cell differentiation. Int J Dev Biol 2004; 48(8-9):819-27.

118. Kamachi Y, Uchikawa M, Tanouchi A, Sekido R, Kondoh H. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. Genes Dev 2001; 15(10):1272-86.

119. Simeone A, Acampora D, Mallamaci A, Stornaiuolo A, D'Apice MR, Nigro V, Boncinelli E. A vertebrate gene related to orthodenticle contains a homeodomain of the bicoid class and demarcates anterior neuroectoderm in the gastrulating mouse embryo. EMBO J 1993; 12(7):2735-47.

120. Larsen KB, Lutterodt MC, Mollgard K, Moller M. Expression of the homeobox genes OTX2 and OTX1 in the early developing human brain. J Histochem Cytochem 2010; 58(7):669-78.

121. Schilter KF, Schneider A, Bardakjian T, Soucy JF, Tyler RC, Reis LM, Semina EV. OTX2 microphthalmia syndrome: four novel mutations and delineation of a phenotype. Clin Genet 2011; 79(2):158-68.

122. Ashkenazi-Hoffnung L, Lebenthal Y, Wyatt AW, Ragge NK, Dateki S, Fukami M, Ogata T, Phillip M, Gat-Yablonski G. A novel loss-of-function mutation in OTX2 in a patient with anophthalmia and isolated growth hormone deficiency. Hum Genet 2010; 127(6):721-9.

123. Tajima T, Ohtake A, Hoshino M, Amemiya S, Sasaki N, Ishizu K, Fujieda K. OTX2 loss of function mutation causes anophthalmia and combined pituitary hormone deficiency with a small anterior and ectopic posterior pituitary. J Clin Endocrinol Metab 2009; 94(1):314-9.

124. Dateki S, Kosaka K, Hasegawa K, Tanaka H, Azuma N, Yokoya S, Muroya K, Adachi M, Tajima T, Motomura K, Kinoshita E, Moriuchi H, Sato N, Fukami M, Ogata T. Heterozygous orthodenticle homeobox 2 mutations are associated with variable pituitary phenotype. J Clin Endocrinol Metab 2010; 95(2):756-64.

125. Ang SL, Jin O, Rhinn M, Daigle N, Stevenson L, Rossant J. A targeted mouse Otx2 mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. Development 1996; 122(1):243-52.

126. Lequeux L, Rio M, Vigouroux A, Titeux M, Etchevers H, Malecaze F, Chassaing N, Calvas P. Confirmation of RAX gene involvement in human anophthalmia. Clin Genet 2008; 74(4):392-5.

127. London NJ, Kessler P, Williams B, Pauer GJ, Hagstrom SA, Traboulsi EI. Sequence alterations in RX in patients with microphthalmia, anophthalmia, and coloboma. Mol Vis 2009; 15:162-7.

128. Dhomen NS, Balaggan KS, Pearson RA, Bainbridge JW, Levine EM, Ali RR, Sowden JC. Absence of chx10 causes neural progenitors to persist in the adult retina. Invest Ophthalmol Vis Sci 2006; 47(1):386-96.

129. Dorval KM, Bobechko BP, Ahmad KF, Bremner R. Transcriptional activity of the paired-like homeodomain proteins CHX10 and VSX1. J Biol Chem 2005; 280(11):10100-8.

130. Human microphthalmia associated with mutations in the retinal homeobox gene CHX10. Percin EF, Ploder LA, Yu JJ, Arici K, Horsford DJ, Rutherford A, Bapat B, Cox DW, Duncan AMV, Kalnins VI, Kocak-Altintas A, Sowden JC, Trabousli E, Sarfarazi M, McInnes RR.*(1) Nat Gen 2000;25:397-401. Am J Ophthalmol 2001; 131(1):156-7.

131. Bar-Yosef U, Abuelaish I, Harel T, Hendler N, Ofir R, Birk OS. CHX10 mutations cause non-syndromic microphthalmia/ anophthalmia in Arab and Jewish kindreds. Hum Genet 2004; 115(4):302-9.

132. Iseri SU, Wyatt AW, Nurnberg G, Kluck C, Nurnberg P, Holder GE, Blair E, Salt A, Ragge NK. Use of genome-wide SNP homozygosity mapping in small pedigrees to identify new mutations in VSX2 causing recessive microphthalmia and a semidominant inner retinal dystrophy. Hum Genet 2010; 128(1):51-60.

133. Bardakjian T, Weiss A, Schneider AS. Anophthalmia / Microphthalmia Overview. 1993.

134. Ye M, Berry-Wynne KM, Asai-Coakwell M, Sundaresan P, Footz T, French CR, Abitbol M, Fleisch VC, Corbett N, Allison WT, Drummond G, Walter MA, Underhill TM, Waskiewicz AJ, Lehmann OJ. Mutation of the bone morphogenetic protein GDF3 causes ocular and skeletal anomalies. Hum Mol Genet 2010; 19(2):287-98.

135. Asai-Coakwell M, French CR, Berry KM, Ye M, Koss R, Somerville M, Mueller R, van Heyningen V, Waskiewicz AJ, Lehmann OJ. GDF6, a novel locus for a spectrum of ocular developmental anomalies. Am J Hum Genet 2007; 80(2):306-15.

136. Reis LM, Tyler RC, Schilter KF, Abdul-Rahman O, Innis JW, Kozel BA, Schneider AS, Bardakjian TM, Lose EJ, Martin DM, Broeckel U, Semina EV. BMP4 loss-of-function mutations in developmental eye disorders including SHORT syndrome. Hum Genet 2011; 130(4):495-504.

137. Belecky-Adams T, Adler R. Developmental expression patterns of bone morphogenetic proteins, receptors, and binding proteins in the chick retina. J Comp Neurol 2001; 430(4):562-72.

138. Nistala H, Lee-Arteaga S, Siciliano G, Smaldone S, Ramirez F. Extracellular regulation of transforming growth factor beta and bone morphogenetic protein signaling in bone. Ann N Y Acad Sci 2010; 1192:253-6.

139. Huang XF, Chai Y. TGF-ss signalling and tooth development. Chin J Dent Res 2010; 13(1):7-15.

140. Nolen LD, Amor D, Haywood A, St Heaps L, Willcock C, Mihelec M, Tam P, Billson F, Grigg J, Peters G, Jamieson RV. Deletion at 14q22-23 indicates a contiguous gene syndrome comprising anophthalmia, pituitary hypoplasia, and ear anomalies. Am J Med Genet A 2006; 140(16):1711-8.

141. Hayashi S, Okamoto N, Makita Y, Hata A, Imoto I, Inazawa J. Heterozygous deletion at 14q22.1-q22.3 including the BMP4 gene in a patient with psychomotor retardation, congenital corneal opacity and feet polysyndactyly. Am J Med Genet A 2008; 146A(22):2905-10.

142. Chang B, Smith RS, Peters M, Savinova OV, Hawes NL, Zabaleta A, Nusinowitz S, Martin JE, Davisson ML, Cepko CL, Hogan BL, John SW. Haploinsufficient Bmp4 ocular phenotypes include anterior segment dysgenesis with elevated intraocular pressure. BMC Genet 2001; 2:18.

143. Hanel ML, Hensey C. Eye and neural defects associated with loss of GDF6. BMC Dev Biol 2006; 6:43.

144. Asai-Coakwell M, French CR, Ye M, Garcha K, Bigot K, Perera AG, Staehling-Hampton K, Mema SC, Chanda B, Mushegian A, Bamforth S, Doschak MR, Li G, Dobbs MB, Giampietro PF, Brooks BP, Vijayalakshmi P, Sauve Y, Abitbol M, Sundaresan P, van Heyningen V, Pourquie O, Underhill TM, Waskiewicz AJ, Lehmann OJ. Incomplete penetrance and phenotypic variability characterize Gdf6-attributable oculo-skeletal phenotypes. Hum Mol Genet 2009; 18(6):1110-21.

145. Aijaz S, Clark BJ, Williamson K, van Heyningen V, Morrison D, Fitzpatrick D, Collin R, Ragge N, Christoforou A, Brown A, Hanson I. Absence of SIX6 mutations in microphthalmia, anophthalmia, and coloboma. Invest Ophthalmol Vis Sci 2004; 45(11):3871-6.

146. Gallardo ME, Rodriguez De Cordoba S, Schneider AS, Dwyer MA, Ayuso C, Bovolenta P. Analysis of the developmental SIX6 homeobox gene in patients with anophthalmia/microphthalmia. Am J Med Genet A 2004; 129A(1):92-4.

147. Porter FD, Drago J, Xu Y, Cheema SS, Wassif C, Huang SP, Lee E, Grinberg A, Massalas JS, Bodine D, Alt F, Westphal H. Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. Development 1997; 124(15):2935-44.

148. Tetreault N, Champagne MP, Bernier G. The LIM homeobox transcription factor Lhx2 is required to specify the retina field and synergistically cooperates with Pax6 for Six6 trans-activation. Dev Biol 2009; 327(2):541-50.

149. Yun S, Saijoh Y, Hirokawa KE, Kopinke D, Murtaugh LC, Monuki ES, Levine EM. Lhx2 links the intrinsic and extrinsic factors that control optic cup formation. Development 2009; 136(23):3895-906.

150. Okada I, Hamanoue H, Terada K, Tohma T, Megarbane A, Chouery E, Abou-Ghoch J, Jalkh N, Cogulu O, Ozkinay F, Horie K, Takeda J, Furuichi T, Ikegawa S, Nishiyama K, Miyatake S, Nishimura A, Mizuguchi T, Niikawa N, Hirahara F, Kaname T, Yoshiura K, Tsurusaki Y, Doi H, Miyake N, Furukawa T, Matsumoto N, Saitsu H. SMOC1 is essential for ocular and limb development in humans and mice. Am J Hum Genet 2011; 88(1):30-41.

151. Abouzeid H, Boisset G, Favez T, Youssef M, Marzouk I, Shakankiry N, Bayoumi N, Descombes P, Agosti C, Munier FL, Schorderet DF. Mutations in the SPARC-related modular calcium-binding protein 1 gene, SMOC1, cause waardenburg anophthalmia syndrome. Am J Hum Genet 2011; 88(1):92-8.

152. Chao R, Nevin L, Agarwal P, Riemer J, Bai X, Delaney A, Akana M, JimenezLopez N, Bardakjian T, Schneider A, Chassaing N, Schorderet DF, FitzPatrick D, Kwok PY, Ellgaard L, Gould DB, Zhang Y, Malicki J, Baier H, Slavotinek A. A male with unilateral microphthalmia reveals a role for TMX3 in eye development. PLoS One 2010; 5(5):e10565.

153. Bertolacini C, Ribeiro-Bicudo L, Petrin A, Richieri-Costa A, Murray J. Clinical findings in patients with GLI2 mutations - phenotypic variability. Clin Genet 2012; 81(1):70-5.

154. West B, Bove KE, Slavotinek AM. Two novel STRA6 mutations in a patient with anophthalmia and diaphragmatic eventration. Am J Med Genet A 2009; 149A(3):539-42.

155. White T, Lu T, Metlapally R, Katowitz J, Kherani F, Wang TY, Tran-Viet KN, Young TL. Identification of STRA6 and SKI sequence variants in patients with anophthalmia/microphthalmia. Mol Vis 2008; 14:2458-65.

156. Reis LM, Tyler RC, Schneider A, Bardakjian T, Stoler JM, Melancon SB, Semina EV. FOXE3 plays a significant role in autosomal recessive microphthalmia. Am J Med Genet A 2010; 152A(3):582-90.

157. Wimplinger I, Shaw GM, Kutsche K. HCCS loss-of-function missense mutation in a female with bilateral microphthalmia and sclerocornea: a novel gene for severe ocular malformations? Mol Vis 2007; 13:1475-82.

158. Bakrania P, Ugur Iseri SA, Wyatt AW, Bunyan DJ, Lam WW, Salt A, Ramsay J, Robinson DO, Ragge NK. Sonic hedgehog mutations are an

uncommon cause of developmental eye anomalies. Am J Med Genet A 2010; 152A(5):1310-3.

159. Belloni E, Muenke M, Roessler E, Traverso G, Siegel-Bartelt J, Frumkin A, Mitchell HF, Donis-Keller H, Helms C, Hing AV, Heng HH, Koop B, Martindale D, Rommens JM, Tsui LC, Scherer SW. Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. Nat Genet 1996; 14(3):353-6.

160. Schimmenti LA, de la Cruz J, Lewis RA, Karkera JD, Manligas GS, Roessler E, Muenke M. Novel mutation in sonic hedgehog in non-syndromic colobomatous microphthalmia. Am J Med Genet A 2003; 116A(3):215-21.

161. Hilton E, Johnston J, Whalen S, Okamoto N, Hatsukawa Y, Nishio J, Kohara H, Hirano Y, Mizuno S, Torii C, Kosaki K, Manouvrier S, Boute O, Perveen R, Law C, Moore A, Fitzpatrick D, Lemke J, Fellmann F, Debray FG, Dastot-Le-Moal F, Gerard M, Martin J, Bitoun P, Goossens M, Verloes A, Schinzel A, Bartholdi D, Bardakjian T, Hay B, Jenny K, Johnston K, Lyons M, Belmont JW, Biesecker LG, Giurgea I, Black G. BCOR analysis in patients with OFCD and Lenz microphthalmia syndromes, mental retardation with ocular anomalies, and cardiac laterality defects. Eur J Hum Genet 2009; 17(10):1325-35.

162. Barbieri AM, Broccoli V, Bovolenta P, Alfano G, Marchitiello A, Mocchetti C, Crippa L, Bulfone A, Marigo V, Ballabio A, Banfi S. Vax2 inactivation in mouse determines alteration of the eye dorsal-ventral axis, misrouting of the optic fibres and eye coloboma. Development 2002; 129(3):805-13.

163. Favor J, Sandulache R, Neuhauser-Klaus A, Pretsch W, Chatterjee B, Senft E, Wurst W, Blanquet V, Grimes P, Sporle R, Schughart K. The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney. Proc Natl Acad Sci U S A 1996; 93(24):13870-5.

164. Weston CR, Wong A, Hall JP, Goad ME, Flavell RA, Davis RJ. JNK initiates a cytokine cascade that causes Pax2 expression and closure of the optic fissure. Genes Dev 2003; 17(10):1271-80.

165. Cunliffe HE, McNoe LA, Ward TA, Devriendt K, Brunner HG, Eccles MR. The prevalence of PAX2 mutations in patients with isolated colobomas or colobomas associated with urogenital anomalies. J Med Genet 1998; 35(10):806-12.

166. Wallis DE, Roessler E, Hehr U, Nanni L, Wiltshire T, Richieri-Costa A, Gillessen-Kaesbach G, Zackai EH, Rommens J, Muenke M. Mutations in the homeodomain of the human SIX3 gene cause holoprosencephaly. Nat Genet 1999; 22(2):196-8.

167. Hahn H, Wicking C, Zaphiropoulous PG, Gailani MR, Shanley S, Chidambaram A, Vorechovsky I, Holmberg E, Unden AB, Gillies S, Negus K, Smyth I, Pressman C, Leffell DJ, Gerrard B, Goldstein AM, Dean M, Toftgard R, Chenevix-Trench G, Wainwright B, Bale AE. Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome. Cell 1996; 85(6):841-51.

168. Gregory-Evans CY, Vieira H, Dalton R, Adams GG, Salt A, Gregory-Evans K. Ocular coloboma and high myopia with Hirschsprung disease associated with a novel ZFHX1B missense mutation and trisomy 21. Am J Med Genet A 2004; 131(1):86-90.

169. Jamieson RV, Perveen R, Kerr B, Carette M, Yardley J, Heon E, Wirth MG, van Heyningen V, Donnai D, Munier F, Black GC. Domain disruption and mutation of the bZIP transcription factor, MAF, associated with cataract, ocular anterior segment dysgenesis and coloboma. Hum Mol Genet 2002; 11(1):33-42.

170. Davis-Silberman N, Kalich T, Oron-Karni V, Marquardt T, Kroeber M, Tamm ER, Ashery-Padan R. Genetic dissection of Pax6 dosage requirements in the developing mouse eye. Hum Mol Genet 2005; 14(15):2265-76.

171. Hanson IM, Fletcher JM, Jordan T, Brown A, Taylor D, Adams RJ, Punnett HH, van Heyningen V. Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. Nat Genet 1994; 6(2):168-73.

172. Dear PH. Copy-number variation: the end of the human genome? Trends Biotechnol 2009; 27(8):448-54.

173. Lejeune J, Turpin R, Gautier M. [Mongolism; a chromosomal disease (trisomy)]. Bull Acad Natl Med 1959; 143(11-12):256-65.

174. Buckland PR. Polymorphically duplicated genes: their relevance to phenotypic variation in humans. Ann Med 2003; 35(5):308-15.

175. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. Large-scale copy number polymorphism in the human genome. Science 2004; 305(5683):525-8.

176. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. Nat Genet 2004; 36(9):949-51.

177. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, Lincoln S, Crawley A, Hanson M, Maraganore D, Adler C, Cookson MR, Muenter M, Baptista M, Miller D, Blancato J, Hardy J, Gwinn-Hardy K. alpha-Synuclein locus triplication causes Parkinson's disease. Science 2003; 302(5646):841.

178. Sleegers K, Brouwers N, Gijselinck I, Theuns J, Goossens D, Wauters J, Del-Favero J, Cruts M, van Duijn CM, Van Broeckhoven C. APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. Brain 2006; 129(Pt 11):2977-83.

179. Cabrejo L, Guyant-Marechal L, Laquerriere A, Vercelletto M, De la Fourniere F, Thomas-Anterion C, Verny C, Letournel F, Pasquier F, Vital A, Checler F, Frebourg T, Campion D, Hannequin D. Phenotype associated with APP duplication in five families. Brain 2006; 129(Pt 11):2966-76.

180. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'Connell R J, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science 2005; 307(5714):1434-40.

181. McKinney C, Merriman ME, Chapman PT, Gow PJ, Harrison AA, Highton J, Jones PB, McLean L, O'Donnell JL, Pokorny V, Spellerberg M, Stamp LK, Willis J, Steer S, Merriman TR. Evidence for an influence of chemokine ligand 3-like 1 (CCL3L1) gene copy number on susceptibility to rheumatoid arthritis. Ann Rheum Dis 2008; 67(3):409-13.

182. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. Nature 2006; 444(7118):444-54.

183. Raca G, Jackson C, Warman B, Bair T, Schimmenti LA. Next generation sequencing in research and diagnostics of ocular birth defects. Mol Genet Metab 2010; 100(2):184-92.

184. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet 2006; 7(2):85-97.

185. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. Copy number variation: new insights in genome diversity. Genome Res 2006; 16(8):949-61.

186. Lehmann OJ, Ebenezer ND, Jordan T, Fox M, Ocaka L, Payne A, Leroy BP, Clark BJ, Hitchings RA, Povey S, Khaw PT, Bhattacharya SS. Chromosomal duplication involving the forkhead transcription factor gene FOXC1 causes iris hypoplasia and glaucoma. Am J Hum Genet 2000; 67(5):1129-35.

187. Fisher E, Scambler P. Human haploinsufficiency--one for sorrow, two for joy. Nat Genet 1994; 7(1):5-7.

188. Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, van Heyningen V, Hastie ND. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. Cell 1996; 86(1):71-82.

189. Ton CC, Miwa H, Saunders GF. Small eye (Sey): cloning and characterization of the murine homolog of the human aniridia gene. Genomics 1992; 13(2):251-6.

190. Glaser T, Walton DS, Maas RL. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. Nat Genet 1992; 2(3):232-9.

191. Jordan T, Hanson I, Zaletayev D, Hodgson S, Prosser J, Seawright A, Hastie N, van Heyningen V. The human PAX6 gene is mutated in two patients with aniridia. Nat Genet 1992; 1(5):328-32.

192. Hodgson SV, Saunders KE. A probable case of the homozygous condition of the aniridia gene. J Med Genet 1980; 17(6):478-80.

193. D'Haene B, Meire F, Claerhout I, Kroes HY, Plomp A, Arens YH, de Ravel T, Casteels I, De Jaegere S, Hooghe S, Wuyts W, van den Ende J, Roulez F, Veenstra-Knol HE, Oldenburg RA, Giltay J, Verheij JB, de Faber JT, Menten B, De Paepe A, Kestelyn P, Leroy BP, De Baere E. Expanding the spectrum of

FOXC1 and PITX2 mutations and copy number changes in patients with anterior segment malformations. Invest Ophthalmol Vis Sci 2011; 52(1):324-33.

194. Nishimura DY, Searby CC, Alward WL, Walton D, Craig JE, Mackey DA, Kawase K, Kanis AB, Patil SR, Stone EM, Sheffield VC. 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 2001; 68(2):364-72.

195. Zhang HZ, Li P, Wang D, Huff S, Nimmakayalu M, Qumsiyeh M, Pober BR. FOXC1 gene deletion is associated with eye anomalies in ring chromosome 6. Am J Med Genet A 2004; 124A(3):280-7.

196. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 2004; 36(9):955-7.

197. Milunsky JM, Maher TA, Zhao G, Roberts AE, Stalker HJ, Zori RT, Burch MN, Clemens M, Mulliken JB, Smith R, Lin AE. TFAP2A mutations result in branchio-oculo-facial syndrome. Am J Hum Genet 2008; 82(5):1171-7.

198. Pasutto F, Sticht H, Hammersen G, Gillessen-Kaesbach G, Fitzpatrick DR, Nurnberg G, Brasch F, Schirmer-Zimmermann H, Tolmie JL, Chitayat D, Houge G, Fernandez-Martinez L, Keating S, Mortier G, Hennekam RC, von der Wense A, Slavotinek A, Meinecke P, Bitoun P, Becker C, Nurnberg P, Reis A, Rauch A. Mutations in STRA6 cause a broad spectrum of malformations including anophthalmia, congenital heart defects, diaphragmatic hernia, alveolar capillary dysplasia, lung hypoplasia, and mental retardation. Am J Hum Genet 2007; 80(3):550-60.

199. Ferda Percin E, Ploder LA, Yu JJ, Arici K, Horsford DJ, Rutherford A, Bapat B, Cox DW, Duncan AM, Kalnins VI, Kocak-Altintas A, Sowden JC, Traboulsi E, Sarfarazi M, McInnes RR. Human microphthalmia associated with mutations in the retinal homeobox gene CHX10. Nat Genet 2000; 25(4):397-401.

200. Ng D, Hadley DW, Tifft CJ, Biesecker LG. Genetic heterogeneity of syndromic X-linked recessive microphthalmia-anophthalmia: is Lenz microphthalmia a single disorder? Am J Med Genet 2002; 110(4):308-14.

201. Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, Hadley DW, Tifft C, Zhang L, Wilkie AO, van der Smagt JJ, Gorlin RJ, Burgess SM, Bardwell VJ, Black GC, Biesecker LG. Oculofaciocardiodental and Lenz

microphthalmia syndromes result from distinct classes of mutations in BCOR. Nat Genet 2004; 36(4):411-6.

202. Halpern J, Whittemore AS. Multipoint linkage analysis. A cautionary note. Hum Hered 1999; 49(4):194-6.

203. Jonathan L. Haines MAP-V. Genetic analysis of complex diseases Second ed. Hoboken, New Jersey: John Wiley & Sons. Inc. ; 2006.

204. Lidral AC, Murray JC. Genetic approaches to identify disease genes for birth defects with cleft lip/palate as a model. Birth Defects Res A Clin Mol Teratol 2004; 70(12):893-901.

205. Kwon JM, Goate AM. The candidate gene approach. Alcohol Res Health 2000; 24(3):164-8.

206. Votruba M, Payne A, Moore AT, Bhattacharya SS. Dominant optic atrophy: exclusion and fine genetic mapping of the candidate gene, HRY. Mamm Genome 1998; 9(10):784-7.

207. Azuma N, Hirakiyama A, Inoue T, Asaka A, Yamada M. Mutations of a human homologue of the Drosophila eyes absent gene (EYA1) detected in patients with congenital cataracts and ocular anterior segment anomalies. Hum Mol Genet 2000; 9(3):363-6.

208. Takeda K, Takemoto C, Kobayashi I, Watanabe A, Nobukuni Y, Fisher DE, Tachibana M. Ser298 of MITF, a mutation site in Waardenburg syndrome type 2, is a phosphorylation site with functional significance. Hum Mol Genet 2000; 9(1):125-32.

209. Tachibana M, Perez-Jurado LA, Nakayama A, Hodgkinson CA, Li X, Schneider M, Miki T, Fex J, Francke U, Arnheiter H. Cloning of MITF, the human homolog of the mouse microphthalmia gene and assignment to chromosome 3p14.1-p12.3. Hum Mol Genet 1994; 3(4):553-7.

210. Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 1987; 236(4808):1567-70.

211. Alkuraya FS. Homozygosity mapping: one more tool in the clinical geneticist's toolbox. Genet Med 2010; 12(4):236-9.

212. Gibbs JR, Singleton A. Application of genome-wide single nucleotide polymorphism typing: simple association and beyond. PLoS Genet 2006; 2(10):e150.

213. Alkuraya FS. Autozygome decoded. Genet Med 2010; 12(12):765-71.

214. Hosking FJ, Papaemmanuil E, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA, Allan JM, Taylor M, Tomlinson IP, Greaves M, Houlston RS. Genome-wide homozygosity signatures and childhood acute lymphoblastic leukemia risk. Blood 2010; 115(22):4472-7.

215. Genin E, Todorov AA, Clerget-Darpoux F. Optimization of genome search strategies for homozygosity mapping: influence of marker spacing on power and threshold criteria for identification of candidate regions. Ann Hum Genet 1998; 62(Pt 5):419-29.

216. Enciso-Mora V, Hosking FJ, Houlston RS. Risk of breast and prostate cancer is not associated with increased homozygosity in outbred populations. Eur J Hum Genet 2010; 18(8):909-14.

217. The International HapMap Project. Nature 2003; 426(6968):789-96.

218. Hildebrandt F, Heeringa SF, Ruschendorf F, Attanasio M, Nurnberg G, Becker C, Seelow D, Huebner N, Chernin G, Vlangos CN, Zhou W, O'Toole JF, Hoskins BE, Wolf MT, Hinkes BG, Chaib H, Ashraf S, Schoeb DS, Ovunc B, Allen SJ, Vega-Warner V, Wise E, Harville HM, Lyons RH, Washburn J, Macdonald J, Nurnberg P, Otto EA. A systematic approach to mapping recessive disease genes in individuals from outbred populations. PLoS Genet 2009; 5(1):e1000353.

219. Frishberg Y, Ben-Neriah Z, Suvanto M, Rinat C, Mannikko M, Feinstein S, Becker-Cohen R, Jalanko H, Zlotogora J, Kestila M. Misleading findings of homozygosity mapping resulting from three novel mutations in NPHS1 encoding nephrin in a highly inbred community. Genet Med 2007; 9(3):180-4.

220. Benayoun L, Spiegel R, Auslender N, Abbasi AH, Rizel L, Hujeirat Y, Salama I, Garzozi HJ, Allon-Shalev S, Ben-Yosef T. Genetic heterogeneity in two consanguineous families segregating early onset retinal degeneration: the pitfalls of homozygosity mapping. Am J Med Genet A 2009; 149A(4):650-6.

221. Fendri K, Kefi M, Hentati F, Amouri R. Genetic heterogeneity within a consanguineous family involving the LGMD 2D and the LGMD 2C genes. Neuromuscul Disord 2006; 16(5):316-20.

222. Lezirovitz K, Pardono E, de Mello Auricchio MT, de Carvalho ESFL, Lopes JJ, Abreu-Silva RS, Romanos J, Batissoco AC, Mingroni-Netto RC. Unexpected genetic heterogeneity in a large consanguineous Brazilian pedigree presenting deafness. Eur J Hum Genet 2008; 16(1):89-96.

223. Ducroq D, Shalev S, Habib A, Munnich A, Kaplan J, Rozet JM. Three different ABCA4 mutations in the same large family with several consanguineous loops affected with autosomal recessive cone-rod dystrophy. Eur J Hum Genet 2006; 14(12):1269-73.

224. Laurier V, Stoetzel C, Muller J, Thibault C, Corbani S, Jalkh N, Salem N, Chouery E, Poch O, Licaire S, Danse JM, Amati-Bonneau P, Bonneau D, Megarbane A, Mandel JL, Dollfus H. Pitfalls of homozygosity mapping: an extended consanguineous Bardet-Biedl syndrome family with two mutant genes (BBS2, BBS10), three mutations, but no triallelism. Eur J Hum Genet 2006; 14(11):1195-203.

225. Pannain S, Weiss RE, Jackson CE, Dian D, Beck JC, Sheffield VC, Cox N, Refetoff S. Two different mutations in the thyroid peroxidase gene of a large inbred Amish kindred: power and limits of homozygosity mapping. J Clin Endocrinol Metab 1999; 84(3):1061-71.

226. Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, Cideciyan AV, Haider N, Stone EM, Sheffield VC, Wright AF. Pitfalls in homozygosity mapping. Am J Hum Genet 2000; 67(5):1348-51.

227. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet 1999; 65(6):1493-500.

228. Leutenegger AL, Labalme A, Genin E, Toutain A, Steichen E, Clerget-Darpoux F, Edery P. Using genomic inbreeding coefficient estimates for homozygosity mapping of rare recessive traits: application to Taybi-Linder syndrome. Am J Hum Genet 2006; 79(1):62-6.

229. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Malik Sharif S, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF. Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet 2006; 78(5):889-96.

230. Carr IM, Flintoff KJ, Taylor GR, Markham AF, Bonthron DT. Interactive visual analysis of SNP data for rapid autozygosity mapping in consanguineous families. Hum Mutat 2006; 27(10):1041-6.

231. Tiffin N, Kelso JF, Powell AR, Pan H, Bajic VB, Hide WA. Integration of text- and data-mining using ontologies successfully selects disease gene candidates. Nucleic Acids Res 2005; 33(5):1544-52.

232. Kohler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. Am J Hum Genet 2008; 82(4):949-58.

233. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res 2009; 37(Web Server issue):W305-11.

Chapter 2:

Homozygosity mapping in an anophthalmic pedigree provides evidence of additional genetic heterogeneity

A version of this chapter has been published

Azam Khorshidi, Laurie Russell, Steven Bamforth, Garry Drummond, Royce Johnson and Ordan J Lehmann. 2012 *Ophthalmic Genetics*. 1-9.

Summary

Purpose: Anophthalmia is a heterogeneous developmental disorder characterized by absent eyes whose diverse etiology encompasses chromosomal and monogenic aberrations, as well as environmental causes. Since the molecular basis has been defined in only a small proportion of cases and extending this offers potential to enhance understanding of key steps in ocular development, a consanguineous anophthalmic pedigree was investigated using homozygosity mapping.

Methods: DNA samples from six individuals, two anophthalmic, were genotyped with an array featuring approximately 620,000 single nucleotide polymorphisms (SNPs) in order to identify homozygous or copy number variant (CNV) regions. Candidate genes located in regions of identity by descent (IBD) defined by homozygosity mapping were subsequently screened by direct sequencing.

Results: Genotyping identified five homozygous intervals (4q26-28.1, 13q12.11, 14q22.1-22.2, 15q26.2-26.3 and 19q13.12) larger than 1 Mb that do not correspond with the known loci and which contain a total of 205 annotated genes. No CNVs were identified that segregated with the disease phenotype, and sequencing of 5 candidate genes (*PRDM5, FGF2, SOS2, POU2F2* and *CIC*) did not identify any mutations.

Conclusions: Although constrained by the pedigree's size, the homozygosity mapping approach employed in this study extends the locus heterogeneity of anophthalmia. The results indicate that a novel molecular cause remains to be determined in this pedigree with the causative gene likely located within one of the five IBD regions.

Introduction

The human eye develops through a coordinated series of embryonic events beginning with formation of the optic vesicle prior to closure of the neural tube. The proximity of the optic vesicle to the surface ectoderm first induces formation of the lens placode with subsequent interaction between the optic vesicle and lens placode resulting in invagination and formation of the double-layered optic cup. The lens vesicle is formed by detachment of the lens placode from the surface ectoderm, and closure of the fissure in the optic cup's ventral axis permits creation of the future spherical globe [1-2]. Perturbation of any of these steps results in a spectrum of disorders encompassing microphthalmia (small eyes), anophthalmia (absent eyes), and coloboma, hereafter referred to as MAC [3-4]. Although the prevalence of MAC is low (30 per 100,000 births), they represent important causes of pediatric blindness and are present in up to 13% of such cases [5-10].

Colobomata arise from defects in optic fissure closure and their variable severity is reflected in the range of different tissues involved, including iris, lens, retina, and optic nerve [11-12]. Similar to colobomata, the precise mechanisms underlying microphthalmia and anophthalmia are incompletely understood. Defects in steps prior to optic cup formation as well as degeneration of the optic vesicle have been proposed to account for anophthalmia [3], and in keeping with the profound phenotype, it is presumed that perturbation of early and critical

ocular developmental steps may result in completely absent eyes or rudimentary cysts [3, 8].

The etiology of MAC is complex with both environmental and genetic factors involved [9, 13]. Environmental causes of altered intrauterine eye development include viral infection (rubella [14], varicella [15] and cytomegalovirus [16]), vitamin A deficiency [17], and exposure to alcohol, warfarin or thalidomide [18], during early pregnancy [13]. However, heritable factors are thought to underlie the majority of MAC cases with a 10% sibling recurrence risk identified in a large population based study [5]. The spectrum of genetic anomalies includes: karyotypically-visible chromosomal abnormalities (e.g. trisomy 13 [19] and 18 [20] being the most frequently observed), smaller copy number variations, as well as gene mutation [8], and at the present time the genetic basis of ~ 15-20% of MAC cases has been identified.

Some two thirds of MAC cases arise sporadically, while the remainder show evidence of familial recurrence. The phenotype may be restricted to ocular mal-development, or alternatively be syndromic with systemic malformations that can include: cardiac defects, facial defects, microcephaly and hydrocephaly [5-6, 21-22]. The marked phenotypic heterogeneity is illustrated by unilateral or bilateral cases, as well as frequent incomplete penetrance [23]. A broad range of inheritance patterns have been documented including autosomal recessive [24],

dominant [21], and X-linked inheritance [25], although in the majority of cases the inheritance is undetermined [5, 8].

Several loci associated with MAC have been identified including six for isolated microphthalmia (MCOP1-6), ten for syndromic microphthalmia (MCOPS1-10) and twenty seven for coloboma [26]. So far, genetic studies have identified mutations in 20 genes in rare syndromic or isolated forms of MAC (Figure 2.2), including *SOX2* [27], *OTX2* [28], *PAX6* [29], *CHX10* [30], *BMP4* [31], *CHD7* [32], *RAX* [33] *GDF6* [34], *GDF3* [35], *FOXE3* [36], and *SMOC1* [37] . However, since mutations in the known genes underlie <20% of cases, the majority of patients have an unidentified genetic etiology [27].

Characterization of mutations in developmental disorders such as MAC is important for the insight it provides into key steps in eye development and for revealing the genetic basis of congenital ocular malformation. In addition, identification of causative genes for rare disorders can serve as a model for understanding systemic diseases [38], as illustrated by *IFR6* (interferon regulatory factor 6) in which variants both cause van der Woude syndrome, a rare disorder characterized by lower lip pits, orofacial defects and hypodontia (prevalence 1/100 000), as well as cleft lip (prevalence 1/700) [39]. Such examples highlight the possibility that milder variants in MAC-causing genes may result in less extreme phenotypes, and thus identifying such genes may provide insight into a range of ocular diseases.

Human disease genes have been identified with a variety of methods including linkage analysis [40], genome-wide association studies [41], homozygosity mapping [42], and candidate gene screening [43]. Each technique has relative advantages and disadvantages as illustrated by linkage analysis, which relies on co-segregation of genetic markers with a disease phenotype, and in turn requires large pedigrees. In a dominantly inherited disorder, each informative meiosis would contribute a maximum LOD score of 0.3, necessitating at least eleven informative meiosis events to produce significant linkage [44]. In contrast, homozygosity mapping takes advantage of consanguineous relationships [42], utilizing the principle that the chromosomal segment surrounding the disease-causing variant is present in both the maternal and paternal alleles and thus manifests as homozygous interval. These intervals, which are termed identical by descent (IBD) [42], are readily detectable by genome-wide analysis with either microsatellites markers or increasingly high-density SNP microarrays [45]. Homozygosity mapping has been widely used to map a variety of ocular [46-50] and non-ocular disorders [51-56], and due to its requirement for much smaller pedigrees, provides an effective means of identifying recessive loci.

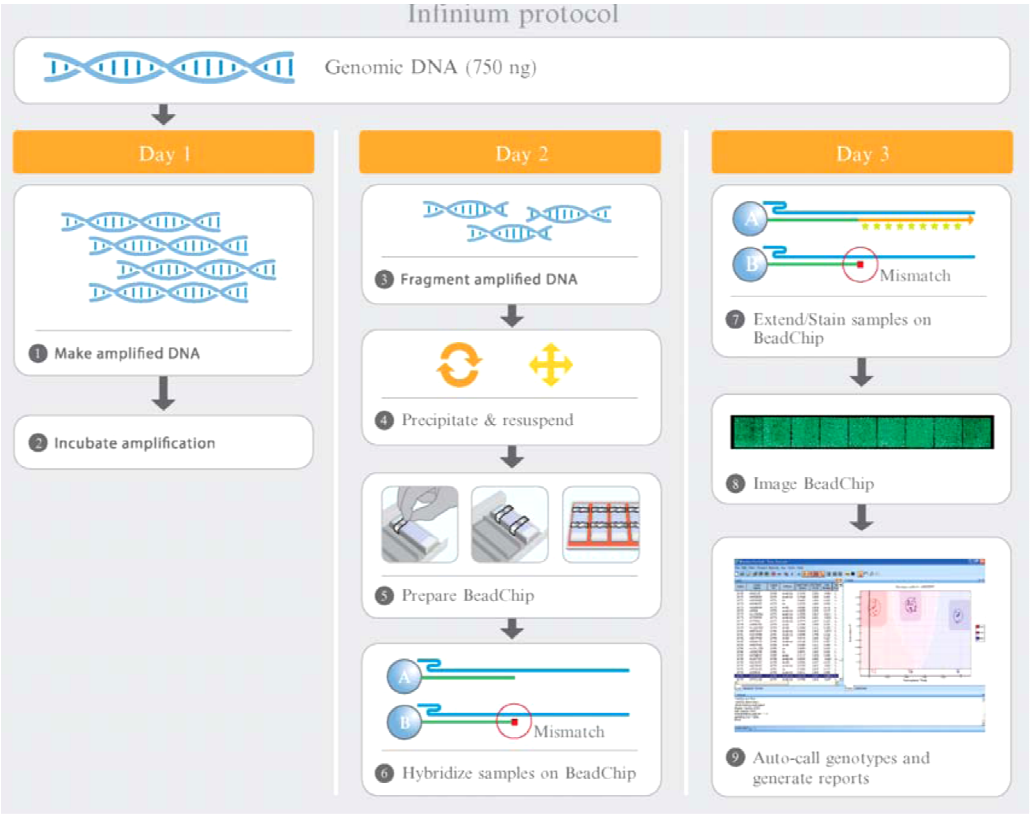
The goal of this study was to map homozygous regions in a consanguineous anophthalmia pedigree and subsequently find the causative mutation. With analyses performed to identify the causative locus, several IBD regions were defined and promising candidate genes within them were screened for mutation.

Material and methods

Genotyping: Genomic DNA from the six family members was extracted from venous blood samples using standard methodology (QIAGEN GmbH, Hilden, Germany). Genotyping was subsequently performed using the Illumina Infinium HD Human610-Quad BeadChip comprising 592,000 SNPs and 28,000 non-polymorphic copy number variation probes (intensity only), covering most of the common SNPs and CNVs with median marker spacing of 2.7 kb (Illumina, Inc. San Diego, CA, USA) (Figure 2.1) . All the samples were typed using bead array technology (deCODE genetics Inc., Reykjavík, Iceland). Briefly, 1,125 ng genomic DNA is incubated for 20 to 24 hours at 37 ^{0}C for whole genome amplification. Amplified DNA is fragmented, precipitated and resuspended before hybridization to Human610-Quad BeadChip containing 50 - mer capture probes to SNP loci for 16–24 hr at 48°C. After hybridization of the target to the bead, single base extension followed by signal amplification utilizing Tecan instruments is used for scoring SNPs [57]. Intensity data for all samples is analyzed with GenomeStudio genotyping module (Illumina Inc., San Diego, USA). This study was approved by the University of Alberta Hospital Health Research Ethics Board, and informed consent was obtained from all participants.

MRI imaging: cerebral MRI images, including orbital sections, were obtained from one affected individual on a 3T Trio scanner (Siemens) with image processing performed with FMRIB software (FLS).

Figure 2.1 SNP genotyping on DNA arrays workflow (taken from illumina website).



Ocular histology: Tissue removed intra-operatively during the implantation of an orbital prosthesis was fixed in 10% neutral buffered formalin and embedded in paraffin for histological analysis. Sections were stained with haematoxylin and eosin and examined using light microscopy.

Homozygosity mapping: SNP data were analyzed using GenomeStudio (Illumina Inc., San Diego, USA) with identified non-Mendelian errors, excluded from analysis [58]. Homozygosity mapping analysis was next performed with PLINKv1.02 and analysis restricted to regions of homozygosity > 1 Mb in size and containing > 100 consecutive homozygous SNPs. To allow centromeric and SNP-poor regions to be algorithmically excluded from analysis, at least 1 SNP / 50 kb was required for defining a homozygous run. The autosomes were scanned for homozygosity using these criteria, with at most five missing genotypes and one heterozygote call permitted per run of homozygosity. The total percentage of the genome included in runs of homozygosity (ROHs) for each individual was calculated by dividing the sum of homozygous length by a factor of 2,788, which represents the extent of the autosomal SNP coverage on the Illumina 610-Quad SNP array.

CNV analysis: CNV detection was performed with software that analyzed the 'B-allele frequency' (allelic intensity ratios) and 'Log R ratio' (normalized total intensity) data for each individual simultaneously (Figure 2.2) (GenomeStudio, Illumina Inc., San Diego, USA). The B allele frequency represents the ratio of

values for the two alleles at each SNP as determined by the hybridization intensity of the fluorescent dyes used (Cy5 (green) A allele; Cy3 (red) B allele). The total fluorescent intensity signal provided by the copy number and SNP data is used to calculate the log of the observed to the expected intensities (log R ratio). Any deviation from zero ($\log_2 1$) provides evidence of copy number changes, with for example, a log R ratio of approximately -1 ($\log_2$ of 50% signal decrease = -1) expected with a hemizygyous deletion.

Linkage analysis: Two-point LOD scores were calculated using a PERL script for two-/multi-point linkage analyses (easyLINKAGE), assuming an autosomal recessive inheritance, full penetrance, and a disease allele frequency of 0.001 as the default option of the program [59].

In silico analysis: To determine which of the several regions identified by homozygosity mapping contained the causative gene, selected candidates were evaluated based upon their expression patterns [60], known ortholog function [61], as well as homology from syntenic regions in multiple model organisms (murine [62], zebrafish [17] and Drosophila [63]).

DNA sequencing: After prioritization of candidate genes, primers were designed to amplify exons and splice sites (http://primer3.sourceforge.net), with sequences and PCR conditions listed (Supplementary Table 2.1). Amplicons were sequenced on an ABI Prism 3100 capillary sequencer (Applied Biosystems, Foster City, CA, USA) and analyzed using Sequencher 4.5 (GeneCodes, Madison, WI, USA)

Figure 2.2 Schematic representation of known MAC loci (red) with the causative genes shown together with IBD intervals identified by homozygosity mapping (blue) – note *PITX2* and *OTX2* lie 3.12 Mb and 4.5 Mb respectively from the start of their respective homozygous intervals. Inset below, the critical intervals on chromosomes 4, 14 and 19 where candidate genes selected for analysis are depicted (bold) [64] (chromosome ideogram is taken from web: http://www.biologia.uniba.it/rmc/0-internal-images/z-ideograms/ideograms.html).

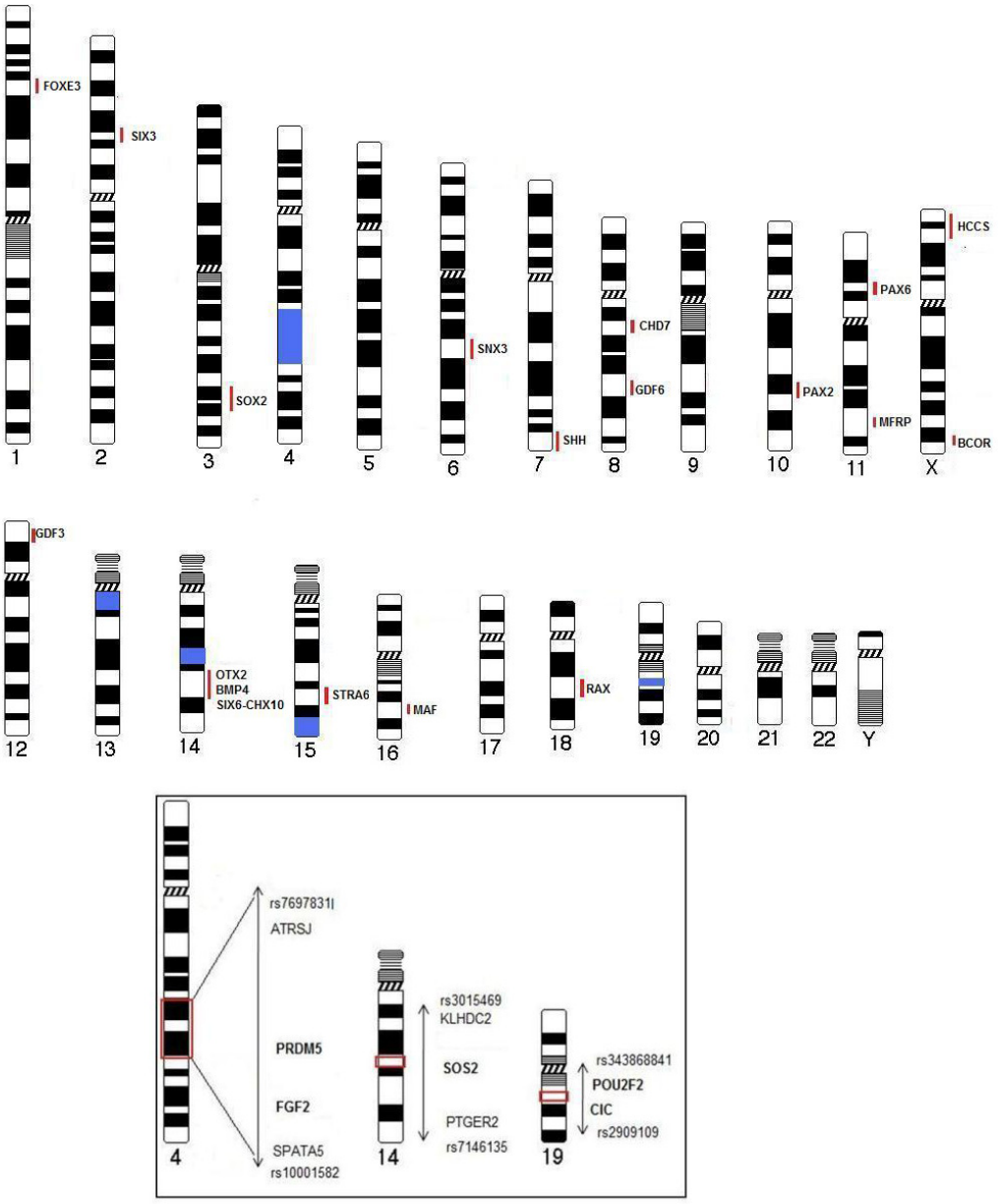
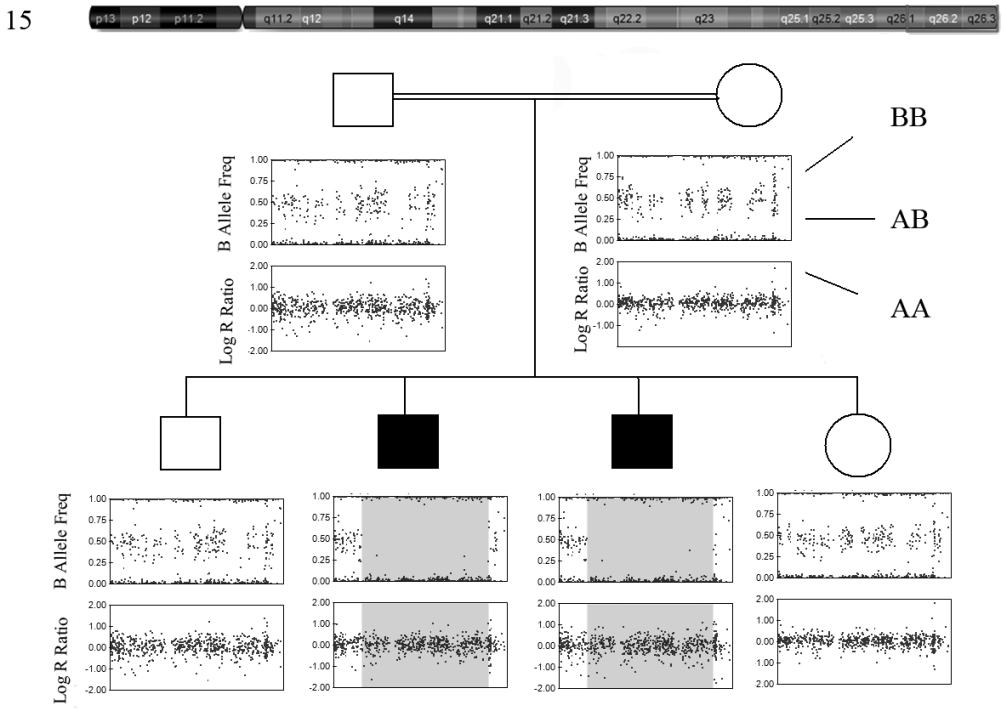


Figure 2.3 SNP data illustrating extended region of homozygosity on chromosome 15q26.2-15q26.3 with homozygous regions indicated by gray shading, taken from [64]. Each panel comprises two plots: upper (B allele frequency) of individual genotypes (AA, BB or AB); lower (log R ratio) of copy number data for each SNP. The absence of AB genotypes across a large region (shaded), in the absence of altered copy number, indicates a region of homozygosity.



Results

The parents are healthy first cousins of Kurdish extraction, of normal intelligence and with 20/20 vision. Their ocular examination was unremarkable, as was that of the two unaffected children (also 20/20 acuity with normal corneal diameters). In contrast, the proband who was born in Iraq following an uncomplicated 40 weeks gestation, was noted to have anophthalmia. An MRI examination, at the age of 9 years, demonstrated that in addition to empty orbits, part of the intra-orbital optic nerves were absent and the chiasm was hypoplastic (Figure 2.4). A second affected male born at 38 weeks (in Canada) was noted, during surgery to implant a prosthesis, to have a small residual orbital cystic structure that was removed and subsequently analyzed histologically. Although no recognizable structures were evident macroscopically, histological sections revealed a dense collagenous wall, lined by a thick disorganized layer of retinoglial tissue (Figure 2.5). Some sections exhibited associations of epithelium and pigment containing cells with loose fibrovascular tissue suggestive of uveal tissue; however, there was no evidence of lenticular or corneal development (Figure 2.5).

Figure 2.4 Axial cerebral MRI images illustrating bilateral absent globes together with absence of part of the intraorbital optic nerve (arrow) and a hypoplastic optic chiasm in one affected individual (A) compared to normal (B). Taken from [64].

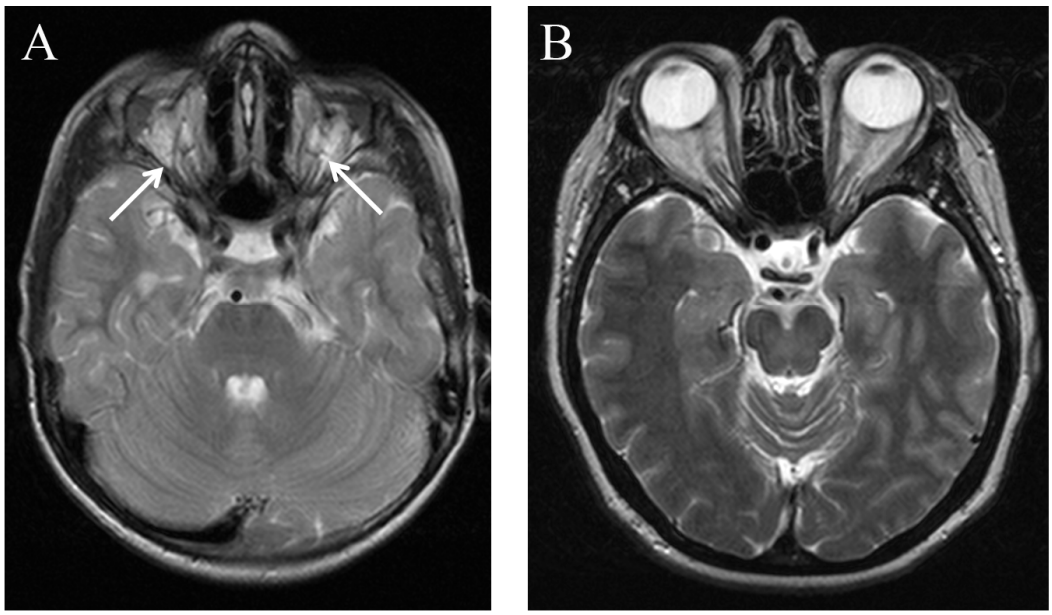
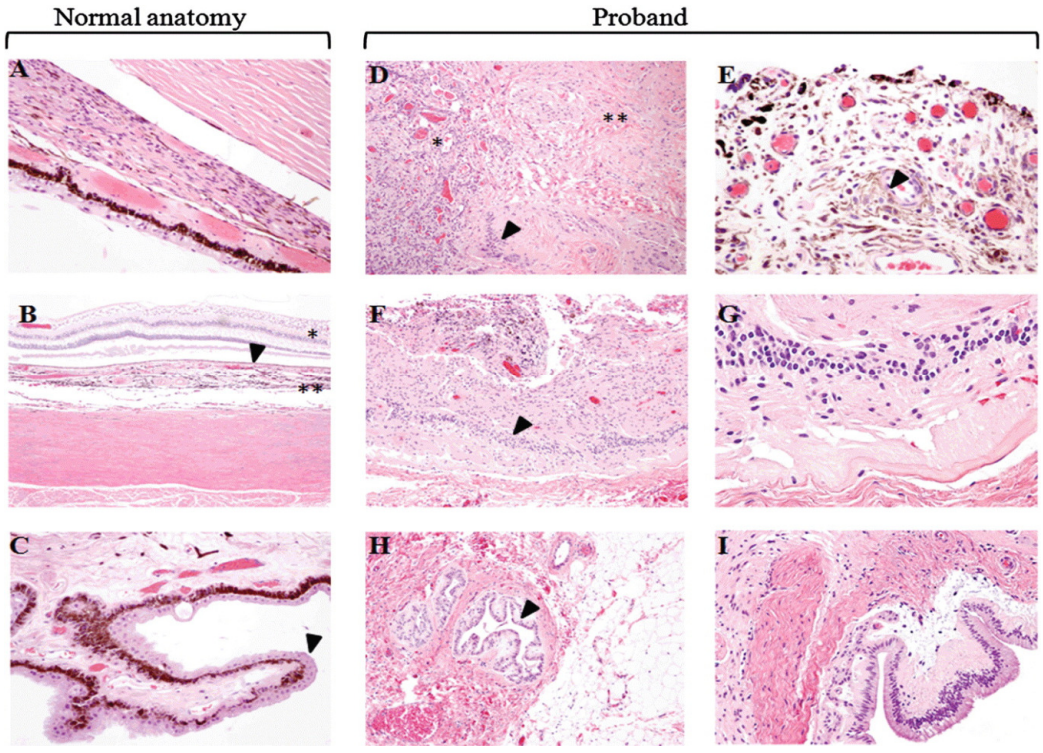


Figure 2.5 Histological features of patient's excised congenital orbital cyst demonstrating rudimentary ocular tissues, taken from [64]. In contrast to the normal anatomy of the ciliary body (A), retina (asterisk), RPE (arrowhead), and choroid (double asterisk) (B), and ciliary process (arrowhead) (C), the patient's orbital cyst exhibits a cyst wall with vascularized connective tissue (asterisk) in dense fibrous tissue (double asterisk) and neuroglial structures (arrowhead) (D) with a layer of pigmented epithelial cells resembling choroid and RPE (E). In addition, neuroglial tissue is present with one distinct retinal layer (arrowhead) but without any recognizable ganglion cells or photoreceptors (F, G). An additional feature is the presence of neuroepithelial processes (arrowhead) in connective tissue with adjacent adipose tissue (H, I). [All sections stained with hematoxylin-eosin and imaged at the following magnifications: x40 (B); x100 (D, F, H); x200 (A, C, I); x400 (E, G)



SNP analysis identified 169 homozygous tracts > 1 Mb in length, comprising 11 to 37 segments per individual. Notably the parents had lower levels of homozygosity than their children (Table 2.1A), with parental homozygosity (≤1%) being substantially less than that observed in the unaffected or affected offspring (3.4-4.4%). Regions that were exclusively homozygous in the two affected siblings were selected by pairwise estimate leading to the identification of 5 IBD regions that segregated with disease status (4q26-28.1, 13q12.11, 14q22.1-22.2, 15q26.2-26.3 and 19q13.12) (Table 2.1B). These 5 regions comprise a total of 18.3 Mb and contain 205 known genes, open reading frames, pseudogenes and gene coding hypothetical proteins (Supplementary Table 2.1). Repetition of the analysis to identify smaller IBD regions (>300kb) again did not identify any that overlapped known MAC loci. CNV analysis revealed several regions of altered copy number in affected individuals, of which two, on 6p and 8p, were homozygously deleted (Table 2.2). However since these two regions do not segregate with disease (being also present in the parents), and did not contain any genes, they were interpreted as being not causative of the anophthalmia phenotype.

Table 2.1 **A**. Summary of the number and extent of homozygous regions in each individual, **B**. Summary of the homozygous regions common to the two affected individuals that segregate with the disease phenotype. Taken from [64].

A)

Individual	Number of segments	Mb	Percentage of the homozygosity in autosomes
Father	18	27.5	1.0
Mother	11	14.1	0.5
Affected 1	37	122.4	4.4
Affected 2	36	94.8	3.4
Unaffected 1	30	90.8	3.6
Unaffected 2	37	107.7	3.9

B)

Chromosomal location	Start position (Mb)	End position (Mb)	Size (Mb)	Number of genes and predicted transcripts	Candidate genes screened
4q26-4q28.1	114.65	124.58	9.93	82	*PRDM5* *FGF2*
13q12.11	18.53	19.55	1.02	10	-
14q22.1-14q22.2	50.23	52.80	2.57	37	*SOS2*
15q26.2-15q26.3	95.94	99.36	3.42	23	-
19q13.12	41.54	42.86	1.32	53	*POU2F2* *CIC*

Table 2.2 CNVs identified that were common to both affected individuals. Copy number value (CN value) for the regions with no copy number change is 2. Values of 0 and 1 show loss of one or both copy (deletion) and values of 3 and 4 represent gain of one or two copy (duplication). Taken from [64].

Chromosome	Start position (Mb)	End position (Mb)	Size (Mb)	CN type	CN value	Present in parents	Present in unaffected
6	32.56	32.59	0.28	Deletion	0	Yes	Yes
8	39.35	39.49	0.14	Deletion	0	Yes	Yes
15	18.28	20.09	1.80	Deletion	1	Yes	No
XY	88.49	92.17	3.82	Duplication	3	Yes	No

Linkage analysis was next performed to assist with determining which of the five IBD regions was associated with disease with two-point linkage analysis. Although constrained by the small pedigree size, the maximum LOD score (0.84 at recombination frequency θ=0), was observed for two of the five IBD regions (chromosomes 4 and 14) with lower LOD scores (0.25) for the other three (chromosomes 13, 15 and 19).

Bioinformatic analysis of the 205 genes in the IBD intervals (Supplementary Table 2.2) did not reveal any that were preferentially expressed in the eye (UniGene) nor that were associated with ocular defects in Drosophila, zebrafish or murine models. Accordingly, five growth or transcription factors, known to be involved in development, were selected for sequencing: *PRDM5, FGF2, CIC, POU2F2,* and *SOS2*. Two SNPs in exons 6 and 11 of *PRDM5* were identified (dbSNP #343192 and #12499000), which occur at high prevalence in normal individuals, however no sequence variations that could be considered to be disease causing were identified in these five genes.

Discussion

The key finding of this research was identification of five shared autozygous blocks that do not correspond with known MAC loci (Figure 2.2). Although based on their size (1-10 Mb) it cannot be predicted which interval contains the causative gene, the novelty of all five regions indicates the existence of additional genetic heterogeneity for MAC. Other important aspects of this study include utilizing both genotyping and copy variation approaches for comprehensive analysis, plus integration of clinical studies to provide magnetic resonance imaging and histological data on a specific anophthalmia subtype. The MRI findings illustrate extensive structural changes to the visual pathway in anophthalmia, extending beyond absent globes to encompass the intra-orbital optic nerves as well as hypoplasia of the optic chiasm (Figure 2.4). Equally, histological examination which identified epithelial cells resembling RPE and immature neuroglial tissue in the proband's intra-orbital cyst (Figure 2.5), accords with arrested development of a primitive optic vesicle.

This study is predicated on the assumption of autosomal recessive inheritance. Although a reasonable supposition in view of the consanguinity and confirmed unaffected status of both parents, alternatives such as autosomal dominant with incomplete penetrance or X-linked inheritance, cannot be excluded. Theoretically, progeny of a first cousin marriage share 1/16th (0.0625) of their genomes [65], a proportion that varies between siblings due to random recombination events [66]. Identification of a larger number of homozygous segments, that are of greater

extent, in affected progeny compared to their parents (Table 2.1A) plus the absence of any unexpected increases in the percentage of homozygosity in children, excludes any hidden consanguinity and lends support for autosomal recessive inheritance.

Genomic structural variation has a major role in the pathogenesis of human disease [67], and in view of this and the eye's known sensitivity to alterations in gene dosage, the possibility of a causative CNV was actively explored. As evident from Figure 2.3, the SNP platform utilized provides a sensitive means of identifying regions of homozygosity, which in this case was unassociated with copy number variation. Since no CNV was identified that segregated with the disease phenotype (Table 2.2), this indicates that any potential CNV is either below the array's resolution (~ 27 kb), or more likely, the phenotype is attributable to by variants at the base pair level. Accordingly, extensive bioinformatic analysis was undertaken to identify potential candidates genes within the five IBD regions. Of the 205 encompassed transcripts, the five strongest candidates (comprising 76 amplicons) were sequentially sequenced, starting with *PRDM5* due to its involvement in *Wnt* signalling. *PRDM5* is a member of a family of transcriptional regulators that regulate cell growth and differentiation, and which modulate aspects of both canonical and non-canonical *Wnt* signalling [68-69]. *FGF2* is expressed in surface ectoderm overlying the optic vesicle, controls division of the neuro-epithelium into neuronal retina and RPE [70], and was selected due to these factors as well as its involvement in lens

morphogenesis [71], retinal cellular proliferation [72], and retinal ganglion cell axonal guidance [73]. However, no disease-causing variants were identified in *PRDM5* and *FGF2*, or three other candidate genes (*CIC, POU2F2,* and *SOS2*).

Since this study has not identified the causative variants, it is worthwhile reviewing factors that may have limited research progress. In addition to the previously discussed possibility of an unanticipated mode of inheritance, either locus or allelic heterogeneity would compromise the utility of homozygosity mapping [74]. In addition, the small pedigree size severely limits the usefulness of linkage approaches in either identifying the causative locus or prioritizing the number of IBD intervals identified by homozygosity mapping [75]. A final potential concern is the existence of variants in either intronic or distal regulatory sequences of the candidate genes screened. In this context, two main approaches remain to define the molecular basis of this family's anophthalmia. The first involves SNP genotyping of a cohort of MAC probands, especially from consanguineous communities, to identify regions of homozygosity that could refine the disease-causing interval. A second more direct approach is either targeted [76] or exomic next generation sequencing [77]. By focusing analysis on the regions of interest, the former minimizes the need for extensive bioinformatic filtering of the large number of identified variants [78], although it necessitates a bespoke and consequently expensive targeted capture array [79]. The rapidly declining cost of whole exome sequencing makes the latter more attractive,

especially due to the ability to bioinformatically extract variants corresponding to the five IBD regions [77, 79], followed if need be by subsequent studies of variants elsewhere in the genome. We anticipate that such future next generation sequencing of this pedigree will permit substantive progress defining the causative variant and lead, as we have predicted, to extending the genetic heterogeneity of anophthalmia.

References

1. O'Rahilly R. The prenatal development of the human eye. Exp Eye Res 1975; 21(2):93-112.

2. Graw J. Eye development. Curr Top Dev Biol 2010; 90:343-86.

3. Guercio JR, Martyn LJ. Congenital malformations of the eye and orbit. Otolaryngol Clin North Am 2007; 40(1):113-40, vii.

4. Fitzpatrick DR, van Heyningen V. Developmental eye disorders. Curr Opin Genet Dev 2005; 15(3):348-53.

5. Morrison D, FitzPatrick D, Hanson I, Williamson K, van Heyningen V, Fleck B, Jones I, Chalmers J, Campbell H. National study of microphthalmia, anophthalmia, and coloboma (MAC) in Scotland: investigation of genetic aetiology. J Med Genet 2002; 39(1):16-22.

6. Kallen B, Robert E, Harris J. The descriptive epidemiology of anophthalmia and microphthalmia. Int J Epidemiol 1996; 25(5):1009-16.

7. Graw J. The genetic and molecular basis of congenital eye defects. Nat Rev Genet 2003; 4(11):876-88.

8. Verma AS, Fitzpatrick DR. Anophthalmia and microphthalmia. Orphanet J Rare Dis 2007; 2:47.

9. Hornby SJ, Gilbert CE, Rahi JK, Sil AK, Xiao Y, Dandona L, Foster A. Regional variation in blindness in children due to microphthalmos, anophthalmos and coloboma. Ophthalmic Epidemiol 2000; 7(2):127-38.

10. Kallen B, Tornqvist K. The epidemiology of anophthalmia and microphthalmia in Sweden. Eur J Epidemiol 2005; 20(4):345-50.

11. Chang L, Blain D, Bertuzzi S, Brooks BP. Uveal coloboma: clinical and basic science update. Curr Opin Ophthalmol 2006; 17(5):447-70.

12. Gregory-Evans CY, Williams MJ, Halford S, Gregory-Evans K. Ocular coloboma: a reassessment in the age of molecular neuroscience. J Med Genet 2004; 41(12):881-91.

13. Miller MT. Ocular teratology. Observations, speculations, questions, principles reaffirmed. Eye (Lond) 1992; 6 (Pt 2):177-80.

14. Givens KT, Lee DA, Jones T, Ilstrup DM. Congenital rubella syndrome: ophthalmic manifestations and associated systemic disorders. Br J Ophthalmol 1993; 77(6):358-63.

15. Prober CG, Gershon AA, Grose C, McCracken GH, Jr., Nelson JD. Consensus: varicella-zoster infections in pregnancy and the perinatal period. Pediatr Infect Dis J 1990; 9(12):865-9.

16. Frenkel LD, Keys MP, Hefferen SJ, Rola-Pleszczynski M, Bellanti JA. Unusual eye abnormalities associated with congenital cytomegalovirus infection. Pediatrics 1980; 66(5):763-6.

17. Hornby SJ, Ward SJ, Gilbert CE. Eye birth defects in humans may be caused by a recessively-inherited genetic predisposition to the effects of maternal vitamin A deficiency during pregnancy. Med Sci Monit 2003; 9(11):HY23-6.

18. Papst W. [Thalidomide and Congenital Abnormalities of the Eye]. Ber Zusammenkunft Dtsch Ophthalmol Ges 1964; 65:209-15.

19. Kouassi FX, Koffi KV, Safede K, Cochard C, Cochener B. [Congenital anophthalmias: a case of trisomy 13]. J Fr Ophtalmol 2006; 29(4):e10.

20. Parker SE, Mai CT, Canfield MA, Rickard R, Wang Y, Meyer RE, Anderson P, Mason CA, Collins JS, Kirby RS, Correa A. Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004-2006. Birth Defects Res A Clin Mol Teratol 2010; 88(12):1008-16.

21. Morle L, Bozon M, Zech JC, Alloisio N, Raas-Rothschild A, Philippe C, Lambert JC, Godet J, Plauchu H, Edery P. A locus for autosomal dominant colobomatous microphthalmia maps to chromosome 15q12-q15. Am J Hum Genet 2000; 67(6):1592-7.

22. Bessant DA, Anwar K, Khaliq S, Hameed A, Ismail M, Payne AM, Mehdi SQ, Bhattacharya SS. Phenotype of autosomal recessive congenital microphthalmia mapping to chromosome 14q32. Br J Ophthalmol 1999; 83(8):919-22.

23. Fryns JP. Autosomal dominant simple microphthalmos: incomplete penetrance and variable expression in a large family. J Med Genet 1995; 32(4):326.

24. Bessant DA, Khaliq S, Hameed A, Anwar K, Mehdi SQ, Payne AM, Bhattacharya SS. A locus for autosomal recessive congenital microphthalmia maps to chromosome 14q32. Am J Hum Genet 1998; 62(5):1113-6.

25. Graham CA, Redmond RM, Nevin NC. X-linked clinical anophthalmos. Localization of the gene to Xq27-Xq28. Ophthalmic Paediatr Genet 1991; 12(1):43-8.

26. Porges Y, Gershoni-Baruch R, Leibu R, Goldscher D, Zonis S, Shapira I, Miller B. Hereditary microphthalmia with colobomatous cyst. Am J Ophthalmol 1992; 114(1):30-4.

27. Fantes J, Ragge NK, Lynch SA, McGill NI, Collin JR, Howard-Peebles PN, Hayward C, Vivian AJ, Williamson K, van Heyningen V, FitzPatrick DR. Mutations in SOX2 cause anophthalmia. Nat Genet 2003; 33(4):461-3.

28. Ragge NK, Brown AG, Poloschek CM, Lorenz B, Henderson RA, Clarke MP, Russell-Eggitt I, Fielder A, Gerrelli D, Martinez-Barbera JP, Ruddle P, Hurst J, Collin JR, Salt A, Cooper ST, Thompson PJ, Sisodiya SM, Williamson KA, Fitzpatrick DR, van Heyningen V, Hanson IM. Heterozygous mutations of OTX2 cause severe ocular malformations. Am J Hum Genet 2005; 76(6):1008-22.

29. Azuma N, Yamaguchi Y, Handa H, Hayakawa M, Kanai A, Yamada M. Missense mutation in the alternative splice region of the PAX6 gene in eye anomalies. Am J Hum Genet 1999; 65(3):656-63.

30. Ferda Percin E, Ploder LA, Yu JJ, Arici K, Horsford DJ, Rutherford A, Bapat B, Cox DW, Duncan AM, Kalnins VI, Kocak-Altintas A, Sowden JC, Traboulsi E, Sarfarazi M, McInnes RR. Human microphthalmia associated with mutations in the retinal homeobox gene CHX10. Nat Genet 2000; 25(4):397-401.

31. Bakrania P, Efthymiou M, Klein JC, Salt A, Bunyan DJ, Wyatt A, Ponting CP, Martin A, Williams S, Lindley V, Gilmore J, Restori M, Robson AG, Neveu MM, Holder GE, Collin JR, Robinson DO, Farndon P, Johansen-Berg H, Gerrelli D, Ragge NK. Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between the BMP4 and hedgehog signaling pathways. Am J Hum Genet 2008; 82(2):304-19.

32. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 2004; 36(9):955-7.

33. Voronina VA, Kozhemyakina EA, O'Kernick CM, Kahn ND, Wenger SL, Linberg JV, Schneider AS, Mathers PH. Mutations in the human RAX homeobox gene in a patient with anophthalmia and sclerocornea. Hum Mol Genet 2004; 13(3):315-22.

34. Asai-Coakwell M, French CR, Berry KM, Ye M, Koss R, Somerville M, Mueller R, van Heyningen V, Waskiewicz AJ, Lehmann OJ. GDF6, a novel locus for a spectrum of ocular developmental anomalies. Am J Hum Genet 2007; 80(2):306-15.

35. Ye M, Berry-Wynne KM, Asai-Coakwell M, Sundaresan P, Footz T, French CR, Abitbol M, Fleisch VC, Corbett N, Allison WT, Drummond G, Walter MA, Underhill TM, Waskiewicz AJ, Lehmann OJ. Mutation of the bone morphogenetic protein GDF3 causes ocular and skeletal anomalies. Hum Mol Genet 2010; 19(2):287-98.

36. Reis LM, Tyler RC, Schneider A, Bardakjian T, Stoler JM, Melancon SB, Semina EV. FOXE3 plays a significant role in autosomal recessive microphthalmia. Am J Med Genet A 2010; 152A(3):582-90.

37. Okada I, Hamanoue H, Terada K, Tohma T, Megarbane A, Chouery E, Abou-Ghoch J, Jalkh N, Cogulu O, Ozkinay F, Horie K, Takeda J, Furuichi T, Ikegawa S, Nishiyama K, Miyatake S, Nishimura A, Mizuguchi T, Niikawa N, Hirahara F, Kaname T, Yoshiura K, Tsurusaki Y, Doi H, Miyake N, Furukawa T, Matsumoto N, Saitsu H. SMOC1 is essential for ocular and limb development in humans and mice. Am J Hum Genet 2011; 88(1):30-41.

38. Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. Nat Rev Genet 2006; 7(4):277-82.

39. Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med 2004; 351(8):769-80.

40. Weeks DE, Lange K. The affected-pedigree-member method of linkage analysis. Am J Hum Genet 1988; 42(2):315-26.

41. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008; 9(5):356-69.

42. Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 1987; 236(4808):1567-70.

43. Collins A. Approaches to the identification of susceptibility genes. Parasite Immunol 2009; 31(5):225-33.

44. Ott J. Analysis of Human Genetic Linkage Johns Hopkins University Press, Baltimore; 1991.

45. Morrow EM, Yoo SY, Flavell SW, Kim TK, Lin Y, Hill RS, Mukaddes NM, Balkhy S, Gascon G, Hashmi A, Al-Saad S, Ware J, Joseph RM, Greenblatt R, Gleason D, Ertelt JA, Apse KA, Bodell A, Partlow JN, Barry B, Yao H, Markianos K, Ferland RJ, Greenberg ME, Walsh CA. Identifying autism loci and genes by tracing recent shared ancestry. Science 2008; 321(5886):218-23.

46. den Hollander AI, Lopez I, Yzer S, Zonneveld MN, Janssen IM, Strom TM, Hehir-Kwa JY, Veltman JA, Arends ML, Meitinger T, Musarella MA, van den Born LI, Fishman GA, Maumenee IH, Rohrschneider K, Cremers FP, Koenekoop RK. Identification of novel mutations in patients with Leber congenital amaurosis and juvenile RP by genome-wide homozygosity mapping with SNP microarrays. Invest Ophthalmol Vis Sci 2007; 48(12):5690-8.

47. Littink KW, Koenekoop RK, van den Born LI, Collin RW, Moruz L, Veltman JA, Roosing S, Zonneveld MN, Omar A, Darvish M, Lopez I, Kroes HY, van Genderen MM, Hoyng CB, Rohrschneider K, van Schooneveld MJ, Cremers FP, den Hollander AI. Homozygosity mapping in patients with cone-rod dystrophy: novel mutations and clinical characterizations. Invest Ophthalmol Vis Sci 2010.

48. Iseri SU, Wyatt AW, Nurnberg G, Kluck C, Nurnberg P, Holder GE, Blair E, Salt A, Ragge NK. Use of genome-wide SNP homozygosity mapping in small pedigrees to identify new mutations in VSX2 causing recessive microphthalmia and a semidominant inner retinal dystrophy. Hum Genet 2010; 128(1):51-60.

49. Pereiro I, Valverde D, Pineiro-Gallego T, Baiget M, Borrego S, Ayuso C, Searby C, Nishimura D. New mutations in BBS genes in small consanguineous families with Bardet-Biedl syndrome: detection of candidate regions by homozygosity mapping. Mol Vis 2010; 16:137-43.

50. Thiadens AA, den Hollander AI, Roosing S, Nabuurs SB, Zekveld-Vroon RC, Collin RW, De Baere E, Koenekoop RK, van Schooneveld MJ, Strom TM, van Lith-Verhoeven JJ, Lotery AJ, van Moll-Ramirez N, Leroy BP, van den Born LI, Hoyng CB, Cremers FP, Klaver CC. Homozygosity mapping reveals PDE6C

mutations in patients with early-onset cone photoreceptor disorders. Am J Hum Genet 2009; 85(2):240-7.

51. Paisan-Ruiz C, Scopes G, Lee P, Houlden H. Homozygosity mapping through whole genome analysis identifies a COL18A1 mutation in an Indian family presenting with an autosomal recessive neurological disorder. Am J Med Genet B Neuropsychiatr Genet 2009; 150B(7):993-7.

52. Knight HM, Maclean A, Irfan M, Naeem F, Cass S, Pickard BS, Muir WJ, Blackwood DH, Ayub M. Homozygosity mapping in a family presenting with schizophrenia, epilepsy and hearing impairment. Eur J Hum Genet 2008; 16(6):750-8.

53. Garshasbi M, Motazacker MM, Kahrizi K, Behjati F, Abedini SS, Nieh SE, Firouzabadi SG, Becker C, Ruschendorf F, Nurnberg P, Tzschach A, Vazifehmand R, Erdogan F, Ullmann R, Lenzner S, Kuss AW, Ropers HH, Najmabadi H. SNP array-based homozygosity mapping reveals MCPH1 deletion in family with autosomal recessive mental retardation and mild microcephaly. Hum Genet 2006; 118(6):708-15.

54. Fiskerstrand T, Houge G, Sund S, Scheie D, Leh S, Boman H, Knappskog PM. Identification of a gene for renal-hepatic-pancreatic dysplasia by microarray-based homozygosity mapping. J Mol Diagn 2010; 12(1):125-31.

55. Schraders M, Lee K, Oostrik J, Huygen PL, Ali G, Hoefsloot LH, Veltman JA, Cremers FP, Basit S, Ansar M, Cremers CW, Kunst HP, Ahmad W, Admiraal RJ, Leal SM, Kremer H. Homozygosity mapping reveals mutations of GRXCR1 as a cause of autosomal-recessive nonsyndromic hearing impairment. Am J Hum Genet 2010; 86(2):138-47.

56. Walsh T, Shahin H, Elkan-Miller T, Lee MK, Thornton AM, Roeb W, Abu Rayyan A, Loulus S, Avraham KB, King MC, Kanaan M. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. Am J Hum Genet 2010; 87(1):90-4.

57. Gunderson KL, Steemers FJ, Ren H, Ng P, Zhou L, Tsan C, Chang W, Bullis D, Musmacker J, King C, Lebruska LL, Barker D, Oliphant A, Kuhn KM, Shen R. Whole-genome genotyping. Methods Enzymol 2006; 410:359-76.

58. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81(3):559-75.

59. Lindner TH, Hoffmann K. easyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses. Bioinformatics 2005; 21(3):405-7.

60. Pagon RA. Ocular coloboma. Surv Ophthalmol 1981; 25(4):223-36.

61. Fraser FC, Skelton J. Possible teratogenicity of maternal fever. Lancet 1978; 2(8090):634.

62. Spagnolo A, Bianchi F, Calabro A, Calzolari E, Clementi M, Mastroiacovo P, Meli P, Petrelli G, Tenconi R. Anophthalmia and benomyl in Italy: a multicenter study based on 940,615 newborns. Reprod Toxicol 1994; 8(5):397-403.

63. Miller MT, Stromland K. Teratogen update: thalidomide: a review, with a focus on ocular findings and new potential uses. Teratology 1999; 60(5):306-21.

64. Khorshidi A, Russell L, Bamforth S, Drummond G, Johnson R, Lehmann OJ. Homozygosity mapping in an anophthalmic pedigree provides evidence of additional genetic heterogeneity. Ophthalmic Genet 2012.

65. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet 1999; 65(6):1493-500.

66. Carothers AD, Rudan I, Kolcic I, Polasek O, Hayward C, Wright AF, Campbell H, Teague P, Hastie ND, Weber JL. Estimating human inbreeding coefficients: comparison of genealogical and marker heterozygosity approaches. Ann Hum Genet 2006; 70(Pt 5):666-76.

67. Lupski JR. Genomic rearrangements and sporadic disease. Nat Genet 2007; 39(7 Suppl):S43-7.

68. Deng Q, Huang S. PRDM5 is silenced in human cancers and has growth suppressive activities. Oncogene 2004; 23(28):4903-10.

69. Schneider R, Bannister AJ, Kouzarides T. Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci 2002; 27(8):396-402.

70. Nguyen M, Arnheiter H. Signaling and transcriptional regulation in early mammalian eye development: a link between FGF and MITF. Development 2000; 127(16):3581-91.

71. McAvoy JW, Chamberlain CG, de Iongh RU, Richardson NA, Lovicu FJ. The role of fibroblast growth factor in eye lens development. Ann N Y Acad Sci 1991; 638:256-74.

72. Guilarducci-Ferraz CV, da Silva GM, Torres PM, Dos Santos AA, de Araujo EG. The increase in retinal cells proliferation induced by FGF2 is mediated by tyrosine and PI3 kinases. Neurochem Res 2008; 33(5):754-64.

73. Tanaka T, Saika S, Ohnishi Y, Ooshima A, McAvoy JW, Liu CY, Azhar M, Doetschman T, Kao WW. Fibroblast growth factor 2: roles of regulation of lens cell proliferation and epithelial-mesenchymal transition in response to injury. Mol Vis 2004; 10:462-7.

74. Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, Cideciyan AV, Haider N, Stone EM, Sheffield VC, Wright AF. Pitfalls in homozygosity mapping. Am J Hum Genet 2000; 67(5):1348-51.

75. Wong FL, Cantor RM, Rotter JI. Sample-size considerations and strategies for linkage analysis in autosomal recessive disorders. Am J Hum Genet 1986; 39(1):25-37.

76. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. Nat Methods 2010; 7(2):111-8.

77. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR. Genome-wide in situ exon capture for selective resequencing. Nat Genet 2007; 39(12):1522-7.

78. Rehman AU, Morell RJ, Belyantseva IA, Khan SY, Boger ET, Shahzad M, Ahmed ZM, Riazuddin S, Khan SN, Friedman TB. Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. Am J Hum Genet 2010; 86(3):378-88.

79. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 2008; 26(10):1135-45.

Supplementary Table 2.1. Primers used to amplify coding region of five candidate genes**.** 1) PRDM5, 2) FGF2, 3) POU2F2, 4) SOS2, 5) CIC.

1)

***PRDM5* (Exons)**	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	**Annealing Temp (^{0}C)**
1	CAACTTCGGGCCAATCAG	GTGGCGCAGCGAGTAAAG	60
2	TTGATTTATTGTACTAATCCTTGTTCA	CATTAAAATAAGTCATATCCTCACCTC	57
3	CCATTGTAAACTCAGGAGCCTTT	CCCCCTCACAGTGCATAGTA	58
4	GAAACAAATGAATGGTTCAGCA	TCTCCCAGATCACCAATTAAAGA	59
5	TGAGCATTTCTTGGCTCTGA	GCTGAATGGAATAAACTGTGTTAGC	59
6	TCCAGTTTTTGTTTTCTTTTCTTTC	GCCATTTCATAACTCAGACACAA	59
7	TTTCTTCCTGTTCCGCTTTG	AACGACCCTCCAACGACTC	60
8	TTGTTTTTCCTCTTAATAGTCTCATTT	TCTTTTCACATCAAAATAACTTGG	57
9	CACTTTTGCTTGAGAGCTGTG	GCCCCTGATTACCACTCTTAG	57
10	TGAGATTGTCTCCACTCATTTTT	AAATCACAGCAGCAAATTCAT	57
11	TCCTCCTGACTCAGTGTTTTTC	TCTCAACTGCCTGAATCGTG	58
12	GGACCCTGCTATTGCTTCTT	TGTTCAAACTAACAGAAGACTTCCA	58
13	GCCTTTTTATCCCTTTCCTTTT	CTTGGAAGCATGTGATTTCTCT	58
14	GCACAGGACTGGTTCTGCTA	TCATGATCAATATTAATGAAACAGAAT	57
15	TCATTCACTGGCAATTTTGG	GGAAGACACTATGGGGGAAA	58
16	TGGTTGGTCTTTGGGTCTTT	GGATTCATATTAGGAGCCCTTC	58

2)

***FGF2* (Exons)**	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	**Annealing Temp (^{0}C)**
1	CAGAAAACCCGAGCGAGTAG	GAGAACCCACGAAATGGAAA	60
2	CATCTGCTCCCTGTCACTCA	AACGTGTGAAAACAGAAAGTATGC	59
3	GCTGGTTGAGCAGAATAGGC	TGACCAATTATCCAAACTGAGC	59

3)

***POU2F2* (Exons)**	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	**Annealing Temp (^{0}C)**
1	TGGTTTCCTCCCTGAGTGAC	AAATGTGCTGCCTGTCCTCT	60
2-3	CTCCTGGATTCCCCTCCTT	CTGCCCTCCTACCATAGGC	60
4	CTCTGTGCCGCCTATGGTAG	GCTCTCTCAGCCCTTGGAC	60
5	CACAGGTGGGCATTCTCTCT	ATGAGGGGTGGCTGGTTT	60
6-7	CAGCCAGGACTCAGGCTATC	TGGGTCAAATGGAAAGGAGA	59
8	CTGAGCCCCATCCTGGTC	CTCGCGGCATCTATCAACTG	61
9	AAGGCCCCACCCTGACTT	GGGAGACGTGAGCATGAGA	59

10	TCCTGGAGAAGTGGCTCAAC	CAGGCAGAGGGCTCGTTAG	60
11	GAAGAGTTTTCTAGCGGTGAGG	GTAGGGTGGGCTTCACACAG	59
12	GGTCCTACAGGGAGCATCAC	TCCCACTGTCTCTCCCATTC	59
13	AGTAGGAGCTGACCCCAGGT	GGGACCTGCCAACATAACTG	60
14	CCCAGTTATGTTGGCAGGTC	GGGTATGAAGAGGCAAGCAA	60
15	GCTGGAAGGCAAAGGTCTC	CTCCCTTGTCACTCCTGCTC	59

4)

SOS2	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	**Annealing Temp (^{0}C)**
1	ACCCAGACAAAGGAGGAGGA	GACCTGCTCCCCGTTAGAAG	61
2	AAGATTCTCTTTGACTTTTGTTCAAT	ACAGTGGTTGAGTGACTTTTTGA	58
3	TCTGCTAGCATTAAGAATAGAGATCAA	CACACTAACACAGACCATGCTACA	59
4	TGACAGTGCTACATTTTGCAATTA	CACAAAAAGATTGAGGCAGAAA	59
5	CTCCCAAAGTGCTGGGATTA	CTGGGCAACAGAGCAAGACT	60
6	GAGCAGGGCTATTCCATAGG	CAGCATTTAGGGCTTTAATGAGA	58
7	CCTGGGATCAAGTGATAATAATCTG	ACGAGGATTCCAGAATTACAAA	57
8	TGCTGTTGCATATCCTTGATG	GACTGGTACTGTGTGTCTCCAAA	59
9	TTGGGGGAAGAATCAAACAA	TGACAAGCACAACTTTCAATAATGT	60
10-1	CCATTTGGGTTGCTAGCTGT	GCTGCCATCCAGTTGTTTTT	60
10-2	GACTCGGCTTCCAGGTTACA	TTTTCCCCAAGCCTCAAATA	60
11	GGAGGAAGGCCTCCAAATA	ACTGCCAATTTAATGCCAAA	58
12	CCACTTTTCGTGGAAGTGGT	CACATAAAATGTGTGTGTGTGTG	57
13	TCGTGGCATTAGAGAAGGAAG	TCACACAAATTTGAACATAAACTCA	58
14	GGCTGGGTGAAAGTTAAGGA	GCACTCATGAATAATGTGCAAAA	58
15	TTCATCTGTCTTTTGGGTTAATTG	GCTTAGAAAAGTTTCTTCACCTCA	58
16	TTTAAGCTAAAATTCTGGCTACTCA	TGAAGCAAGGATGAATTTACAAAG	58
17	AAAAATTGTTTTCCAGATAATTTGC	TGCCTCTGAAGACTGCTCTC	58
18	ACTAGATCAGTTGTTTTCCCAAA	CCTGTTGTGCGATCCAATAG	57
19	AGCATTAAATAAGTAATTTTCAGCAA	CACCTAGGAACAGCCATTCAA	57
20	GCACGCTTAAGAACTGTATGCT	TGATAAGAATTATCTGAGACACAGGA	58
21	CATGGCAAAAGATGATGATAAAA	TTGCCAAAACTCAAAATACTCAAA	58
22	CAAATGCCATGGTAGAAGTATCG	TGGTAGCCAAAGTATAATTCACAGA	60
23-1	CCAGATTAGTTTATTTTGAGCCTCT	GGTACCCTTGGAGAGGGTGT	58
23-2	TCGGCCTCCAGAACACTTTA	GATCAGTAGCATTTTTGTAAGAGCA	60

5)

CIC	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	**Annealing Temp (^{0}C)**
1	CGAGGGAGGAGGGAGAGG	ATACCCCCACCACCGTTACT	60
2	GTGGGGAAGAGCTTGAGTTG	TGGGGACACAGAGGTAGATCA	59
3	GGTCCAGGTGGCCTAGGAG	GCAGCAGGGACAGTGAGG	60
4-5	GTGGGATGGCCAGAGACAT	ACCCCACCAGAAGAAGCTG	60
6	TCCTCCCAGCTTCTTCTGGT	AAGCAAGGGGGTAGCTGGT	61
7	TGCTGGTGACAGGCTTACTG	CAGGATAGGTGGCAGAGAGG	59
8	ACCTGTCCAGGGCAGCAG	AAAAGAAAACAAAGGGGCCTA	59
9	CTGAGATCCAGGCTCCAGAC	CCCGGGAGAAACAGACATT	59
10-1	GCTATCGAGGTGTCGGAGTG	TCTTGCTCCTCCTTGTTTGG	60
10-2	CAAGAGACCCGAAAGTGTGG	CTGGGCAATGAACTGGACA	60
10-3	GCCTGCCACTGTCACTAACC	GGCAAGAAAGGAAAAGGTTACA	60
11	TGTTTGGCTCCCTTGTAACC	GCCAACATCCAGCAGGTAGA	59
12-13	TCCCACTTGAGGTCTTGGTC	AGTTAGGGCCCGGACTGC	60
14	GGCCCTAACTTGGTCTCCTG	CCTGACTGGAGCCCATCC	61
15	CGCCACAGGTAGGTGTCAG	CAGAGGGGACCCTTTAACCT	59
16	ACAGCCTTCTCAAGGGGTCT	GAAACTCAGGCAACTCAGCA	59
17-18	GTGTTAGGGTGGCGGGAGT	GAGACCCACTCCTCTTGCAG	60
19	TAGGTTGCCCTGTGACTGTG	GGGAGCACAGGATGAGATGT	59
20	TCACTCGGGTGGGACTTATC	TTACCCGGGAGGAGATAACC	59

Supplementary Table 2.2 list of all annotated genes in the five IBD regions.

Symbol	Position	Description
CAMK2D	4	calcium/calmodulin-dependent protein kinase II delta
ARSJ	4	arylsulfatase family, member J
UGT8	4q26	UDP glycosyltransferase 8
MIR577	4	microRNA 577
LOC100131828	4	similar to CBF1 interacting corepressor
NDST4	4q25-q26	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 4
MRPS33P3	4q26	mitochondrial ribosomal protein S33 pseudogene 3
LOC100128462	4	similar to brix domain containing 1
LOC100131611	4	similar to phosphoglycerate mutase 1 (brain)
KRT18P21	4	keratin 18 pseudogene 21
LOC645368	4	similar to eukaryotic translation initiation factor 3, subunit 12
RLFP	4q26	rearranged L-myc fusion pseudogene
LOC100288825	4	hypothetical LOC100288825
MIR1973	4q26	microRNA 1973
LOC100288861	4	hypothetical LOC100288861
LOC100129008	4	similar to cullin 4A
LOC344978	4	similar to actinin, alpha 4
TRAM1L1	4	translocation associated membrane protein 1-like 1
RPSAP35	4	ribosomal protein SA pseudogene 35
NT5C3P1	4	5'-nucleotidase, cytosolic III pseudogene 1
LOC100288955	4	hypothetical protein LOC100288955
NDST3	4	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 3
LOC100132656	4	FK506 binding protein 4, 59kDa pseudogene
SNHG8	4q26	small nucleolar RNA host gene 8 (non-protein coding)
SNORA24	4q26	small nucleolar RNA, H/ACA box 24
PRSS12	4q28.1	protease, serine, 12 (neurotrypsin, motopsin)
LOC100128177	4	similar to NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)
LOC100288991	4	hypothetical LOC100288991
CEP170L	4	centrosomal protein 170kDa-like
LOC729218	4	hypothetical LOC729218

LOC100289025	4	similar to hCG2014315
LOC100132769	4	hypothetical LOC100132769
LOC729227	4	similar to hCG1984118
METTL14	4	methyltransferase like 14
SEC24D	4	SEC24 family, member D (S. cerevisiae)
SYNPO2	4	synaptopodin 2
MYOZ2	4q26-q27	myozenin 2
LOC100289084	4	similar to mitochondrial ribosomal protein L42
MRPL42P1	4	mitochondrial ribosomal protein L42 pseudogene 1
USP53	4	ubiquitin specific peptidase 53
C4orf3	4	chromosome 4 open reading frame 3
FABP2	4q28-q31	fatty acid binding protein 2, intestinal
GK7P	1q41	glycerol kinase 7 pseudogene
GK6P	4	glycerol kinase 6 pseudogene
LOC100128874	4	similar to hCG1739111
LOC100131884	4	hypothetical LOC100131884
LOC100128460	4	similar to hCG1793472
FLJ14186	4	hypothetical LOC401149
LOC645513	4	hypothetical LOC645513
PDE5A	4q25-q27	phosphodiesterase 5A, cGMP-specific
LOC730456	4	hypothetical LOC730456
MAD2L1	4q27	MAD2 mitotic arrest deficient-like 1 (yeast)
LOC100289220	4	hypothetical LOC100289220
SAR1P3	4	SAR1 homolog (S. cerevisiae) pseudogene 3
PRDM5	4q25-q26	PR domain containing 5
LOC100129988	4	hypothetical LOC100129988
C4orf31	4	chromosome 4 open reading frame 31
TNIP3	4	TNFAIP3 interacting protein 3
QRFPR	4q27	pyroglutamylated RFamide peptide receptor
ANXA5	4q28-q32	annexin A5
TMEM155	4	transmembrane protein 155
LOC100192379	4	hypothetical LOC100192379
EXOSC9	4	exosome component 9
CCNA2	4q25-q31	cyclin A2
BBS7	4	Bardet-Biedl syndrome 7
TRPC3	4	transient receptor potential cation channel, subfamily C, member 3
KIAA1109	4	KIAA1109
ADAD1	4	adenosine deaminase domain containing 1 (testis-specific)
IL2	4q26-q27	interleukin 2
IL21	4q26-q27	interleukin 21

LOC729338	4	CETN4 pseudogene
BBS12	4	Bardet-Biedl syndrome 12
LOC727709	4	DRAM pseudogene
RPL34P12	4q26	ribosomal protein L34 pseudogene 12
FGF2	4q26-q27	fibroblast growth factor 2 (basic)
RPS26P23	4q26	ribosomal protein S26 pseudogene 23
NUDT6	4q26	nudix (nucleoside diphosphate linked moiety X)-type motif 6
SPATA5	4	spermatogenesis associated 5
COILP2	4q27-q28	coilin pseudogene 2
SPRY1	4	sprouty homolog 1, antagonist of FGF signaling (Drosophila)
TRNAC-GCA	4	transfer RNA cysteine (anticodon GCA)
LOC285419	4	hypothetical LOC285419
LOC729501	13	zinc finger protein ENSP00000344568-like
LOC729524	13	fem-1 homolog a pseudogene
LOC283523	13	telomeric repeat binding factor (NIMA-interacting) 1 pseudogene
LOC729535	13	hypothetical LOC729535
LOC654337	13q11	brain cytoplasmic RNA 1, pseudogene
LOC100101111	13q11	glycosyltransferase 8 domain containing 3 pseudogene
LOC100131550	13	similar to calponin 2
LOC100132598	13	similar to zinc finger protein 165
LOC645626	13	coiled-coil domain containing 29-like
LOC400094	13	sorting nexin 19 pseudogene
KLHDC2	14	kelch domain containing 2
SDCCAG1	14q22	serologically defined colon cancer antigen 1
RN7SL3	14	RNA, 7SL, cytoplasmic 3
RN7SL2	14	RNA, 7SL, cytoplasmic 2
ARF6	14q21.3	ADP-ribosylation factor 6
C14orf182	14	chromosome 14 open reading frame 182
PDLIM1P	14	PDZ and LIM domain 1 pseudogene
RPS15AP2	14	ribosomal protein S15a pseudogene 2
LOC196913	14	hypothetical protein LOC196913
C14orf138	14	chromosome 14 open reading frame 138
SOS2	14q21	son of sevenless homolog 2 (Drosophila)
L2HGDH	14	L-2-hydroxyglutarate dehydrogenase
ATP5S	14	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit s
CDKL1	14	cyclin-dependent kinase-like 1 (CDC2-related kinase)
MAP4K5	14q11.2-q21	mitogen-activated protein kinase kinase kinase kinase 5

ATL1	14	atlastin GTPase 1
SNRPGP	14	small nuclear ribonucleoprotein polypeptide G pseudogene
SAV1	14q13-q23	salvador homolog 1 (Drosophila)
ZNF405P	14	zinc finger protein 405 pseudogene
NIN	14	ninein (GSK3B interacting protein)
ABHD12B	14	abhydrolase domain containing 12B
PYGL	14q21-q22	phosphorylase, glycogen, liver
MRP63P9	14q22.1	mitochondrial ribosomal protein 63 pseudogene 9
TRIM9	14	tripartite motif-containing 9
TMX1	14	thioredoxin-related transmembrane protein 1
SETP2	14	SET pseudogene 2
FRMD6	14	FERM domain containing 6
OR7E105P	14q21	olfactory receptor, family 7, subfamily E, member 105 pseudogene
OR7E106P	14q21	olfactory receptor, family 7, subfamily E, member 106 pseudogene
OR7E159P	14	olfactory receptor, family 7, subfamily E, member 159 pseudogene
GNG2	14q21	guanine nucleotide binding protein (G protein), gamma 2
C14orf166	14	chromosome 14 open reading frame 166
NID2	14q21-q22	nidogen 2 (osteonidogen)
COX5AP2	14	cytochrome c oxidase subunit Va pseudogene 2
PTGDR	14	prostaglandin D2 receptor (DP)
LOC100131689	14	similar to KIAA0020
PTGER2	14q22	prostaglandin E receptor 2 (subtype EP2), 53kDa
LOC145820	15	hypothetical protein LOC145820
RNU2P1	15q26.2	RNA, U2 small nuclear pseudogene 1
NR2F2	15q26	nuclear receptor subfamily 2, group F, member 2
MIR1285-1	15q26.2	microRNA 1285-1
LOC100288183	15	hypothetical protein LOC100288183
LOC728800	15	similar to FLJ00402 protein
LOC100289303	15	hypothetical LOC100289303
LOC390643	15	similar to phosphoglycerate mutase 1 (brain)
RPL31P55	15	ribosomal protein L31 pseudogene 55
LOC388181	15	hypothetical LOC388181
SPATA8	15	spermatogenesis associated 8
LOC91948	15	hypothetical LOC91948
ARRDC4	15	arrestin domain containing 4
FAM169B	15	family with sequence similarity 169, member B

IGF1R	15q26.3	insulin-like growth factor 1 receptor
LOC145814	15	pyroglutamyl-peptidase 1-like
SYNM	15	synemin, intermediate filament protein
TTC23	15	tetratricopeptide repeat domain 23
LRRC28	15	leucine rich repeat containing 28
HSP90B2P	15q26.3	heat shock protein 90kDa beta (Grp94), member 2 (pseudogene)
LOC100289465	15	similar to serologically defined breast cancer antigen NY-BR-40
RPL7P5	15q26.3	ribosomal protein L7 pseudogene 5
MEF2A	15q26	myocyte enhancer factor 2A
CYP2A7P1	19q13.2	cytochrome P450, family 2, subfamily A, polypeptide 7 pseudogene 1
CYP2G2P	19q13.2	cytochrome P450, family 2, subfamily G, polypeptide 2 pseudogene
CYP2A13	19q13.2	cytochrome P450, family 2, subfamily A, polypeptide 13
CYP2F1	19q13.2	cytochrome P450, family 2, subfamily F, polypeptide 1
CYP2T3P	19	cytochrome P450, family 2, subfamily T, polypeptide 3 pseudogene
RPL36_7_1657	19	ribosomal protein L36 pseudogene
CYP2S1	19q13.1	cytochrome P450, family 2, subfamily S, polypeptide 1
AXL	19q13.1	AXL receptor tyrosine kinase
TRNAK38P	19	transfer RNA lysine 38 (anticodon UUU) pseudogene
HNRNPUL1	19	heterogeneous nuclear ribonucleoprotein U-like 1
CCDC97	19	coiled-coil domain containing 97
TGFB1	19q13.1	transforming growth factor, beta 1
B9D2	19	B9 protein domain 2
TMEM91	19q13.2	transmembrane protein 91
EXOSC5	19q13.1	exosome component 5
BCKDHA	19q13.1-q13.2	branched chain keto acid dehydrogenase E1, alpha polypeptide
B3GNT8	19	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 8
ATP5SL	19	ATP5S-like
C19orf69	19	chromosome 19 open reading frame 69
LOC644330	19	tropomyosin 3 pseudogene
PLEKHA3P1	19	pleckstrin homology domain containing, family A member 3 pseudogene 1
CEACAM21	19	carcinoembryonic antigen-related cell adhesion molecule 21
CEACAMP3	19q13.2	carcinoembryonic antigen-related cell adhesion molecule pseudogene 3

LOC100286885	19	hypothetical LOC100286885
CEACAM4	19q13.2	carcinoembryonic antigen-related cell adhesion molecule 4
LOC100286912	19	hypothetical LOC100286912
CEACAM7	19q13.2	carcinoembryonic antigen-related cell adhesion molecule 7
CEACAM5	19q13.1-q13.2	carcinoembryonic antigen-related cell adhesion molecule 5
CEACAM6	19q13.2	carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen)
CEACAM3	19q13.2	carcinoembryonic antigen-related cell adhesion molecule 3
LOC645001	19	similar to heterogeneous nuclear ribonucleoprotein A1
LYPD4	19	LY6/PLAUR domain containing 4
DMRTC2	19	DMRT-like family C2
RPS19	19q13.2	ribosomal protein S19
CD79A	19q13.2	CD79a molecule, immunoglobulin-associated alpha
ARHGEF1	19q13.13	Rho guanine nucleotide exchange factor (GEF) 1
LOC390937	19	similar to hCG2040171
RABAC1	19	Rab acceptor 1 (prenylated)
ATP1A3	19q13.31	ATPase, Na+/K+ transporting, alpha 3 polypeptide
GRIK5	19q13.2	glutamate receptor, ionotropic, kainate 5
ZNF574	19	zinc finger protein 574
POU2F2	19	POU class 2 homeobox 2
DEDD2	19	death effector domain containing 2
ZNF526	19	zinc finger protein 526
GSK3A	19	glycogen synthase kinase 3 alpha
LOC100132272	19	hypothetical LOC100132272
ERF	19q13	Ets2 repressor factor
CIC	19q13.2	capicua homolog (Drosophila)
PAFAH1B3	19q13.1	platelet-activating factor acetylhydrolase, isoform Ib, subunit 3 (29kDa)
PRR19	19	proline rich 19
TMEM145	19	transmembrane protein 145
MEGF8	19q12	multiple EGF-like-domains 8

Chapter 3:

Exome analysis using next-generation sequencing data

Introduction

In the previous chapter, I described the use of homozygosity mapping in consanguineous pedigree with anophthalmia which led to the identification of five novel IBD regions. Since genomic regions linked to anophthalmia by homozygosity mapping housed hundreds of genes, the identification of the pathogenic allele from such large number of candidates with conventional Sanger sequencing was expensive and time-consuming. With the advent of new technologies in massively parallel sequencing, analyzing the whole genome or the protein-coding portion of the genome is now feasible and it has led to the detection of several causative genetic mutations in diseases [1-2]. Contrary to traditional gene mapping approaches, which are unable to detect all forms of genomic variation, next generation sequencing (NGS) has the capability to identify all types of genetic variation at base-pair resolution throughout the human genome [3].

Accordingly, a method for whole-exome sequencing (WES) with the SOLiD4 DNA-sequencing platform was used to explore the possibility of the presence of both known and novel variants in the genome of one affected individual in this pedigree. The method is more cost effective than whole genome sequencing, as it allows a large number of samples to be processed simultaneously, reduces the amount of data to be analyzed per sample and hence the computational power and time required for analysis [4-5].

Material & methods

The whole procedure of exome sequencing and data analysis was performed at The Center for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario.

Library preparation

Exome libraries of genomic DNAs were generated using the Agilent SureSelect Target Enrichment System protocol (adapted from SOLiD Fragment Library construction kit (Life Technologies) and compatible with SOLiD multiplexed paired end sequencing (Agilent Technologies; Version 1.3, January 2011). Three micrograms of genomic DNA was sheared using a Covaris (Woburn,MA) S2 to a target peak size of 250 bp. Fragmented DNA was end repaired using end polishing reagents, followed by ligation of SOLiD P1 and IA adapters. The ligated template was loaded into a 2% agarose size-select Invitrogen E-gel and selected at the 200 base pair band from the TrackIt 50-bp DNA ladder (Life Technologies; P/N 10488043). The size-selected libraries underwent nick translation and library amplification. The Agilent 2100 Bioanalyzer DNA 1000 chip was used to confirm the libraries' fragment length. A successful pre-capture library shows a peak size at around 250-270 bp.

Exome capture

Biotynilated RNA oligonucleotide baits were hybridized with sheared DNA at 65 ^{0}C for 24 hours in a GeneAmp PCR system 9700 (Applied Biosystems). Captured

fragments were removed from solution via Dynal MyOne streptavidin T1 beads (Life Technologies; P/N 65602) and subsequently eluted and purified using AMPure beads (Agencourt, P/N A63881). The captured library was then subjected to one more PCR amplification using primers targeting the SOLiD anchors in a GenAmp PCR System 9700 (Applied Biosystem). Resulting libraries were quantified via Agilent 2100 DNA Bioanalyzer before proceeding to SOLiD platform library preparation (described below).

High throughput sequencing on a SOLiD 4 system

Prior to emulsion PCR (ePCR) an equimolar pool of six samples is prepared and diluted to a final concentration of 0.5 pM to be used in the ePCR. Emulsification, amplification and bead enrichment are performed on the EZ Bead system following the manufacture's instruction (Life technologies; P/N 4448417). Following emulsion breaking and subsequent washing, enrichment for template beads was conducted using reagents provided in the EZ Bead Enricher E80 Reagent and Accessories kits (P/N 4452725 and 4453073) . The enriched beads were modified with a 3′ amino group for surface attachment, and prepared for deposit on glass surface of the sequencing slide. Approximately 778 million beads are loaded onto the slide, and paired-end sequencing is performed with the SOLiD ToP Paired End Sequencing kit-BC fragment Library MM50/35/5 (P/N 4459181) following Applied Biosystem SOLiD 4 System Instrument Operation Guide (Applied Biosystems; P/N 4448379).

Primary data analysis and read map

The image data collected by SOLiD 4 is analyzed using Applied Biosystems corona pipline to produce sequence data in colour space, in which each colour represents two consecutive DNA bases on the DNA sequence). Next, paired end reads were pooled and then mapped as single reads to the reference human genome (UCSC's hg18), excluding unordered sequence and alternate haplotypes, using BFAST and BFAST implementation of BWA version 0.6.5a. To remove reference bias introduced by the aligners due to coulrspace conversions, GATK version 1.05506 base quality score recalibration is used with the default parameters optimized for SOLiD dataset.

Filtering SNPs and InDels

The SNP and indel reports were filtered against the Single Nucleotide Polymorphism Database and the 1000-Genome Project to identify novel variants in the sequenced sample. Variants were further filtered as being non-coding (intronic or intergenic) or coding (within an exonic region) and then based upon their conservation or putative effect on the encoded protein: synonymous, non-synonymous, nonsense, splice-site, 5′ or 3′ UTR.

Validation of potential variants

All homozygous variants that were identified by exome sequencing and included in the pedigree regions of homozygosity were validated by traditional Sanger sequencing. Additionally, family inheritance patterns of the identified variants

were examined. The frequencies of the potential variant were further determined by direct sequencing of controls.

Results

In total, 28,943 SNPs were identified, with 2,138 (7.3%) of the SNPs being novel (Figure 3.1). The majority of the identified SNPs (17,540) were within the coding regions of the genome (11,488 heterozygous and 6049 homozygous variants). These variants included synonymous SNPs and SNPs with potential functional impact on the protein function (Table 3.1). After rejecting the involvement of any known MAC genes in this data set, I applied a series of exclusion steps to uncover the causative mutation.

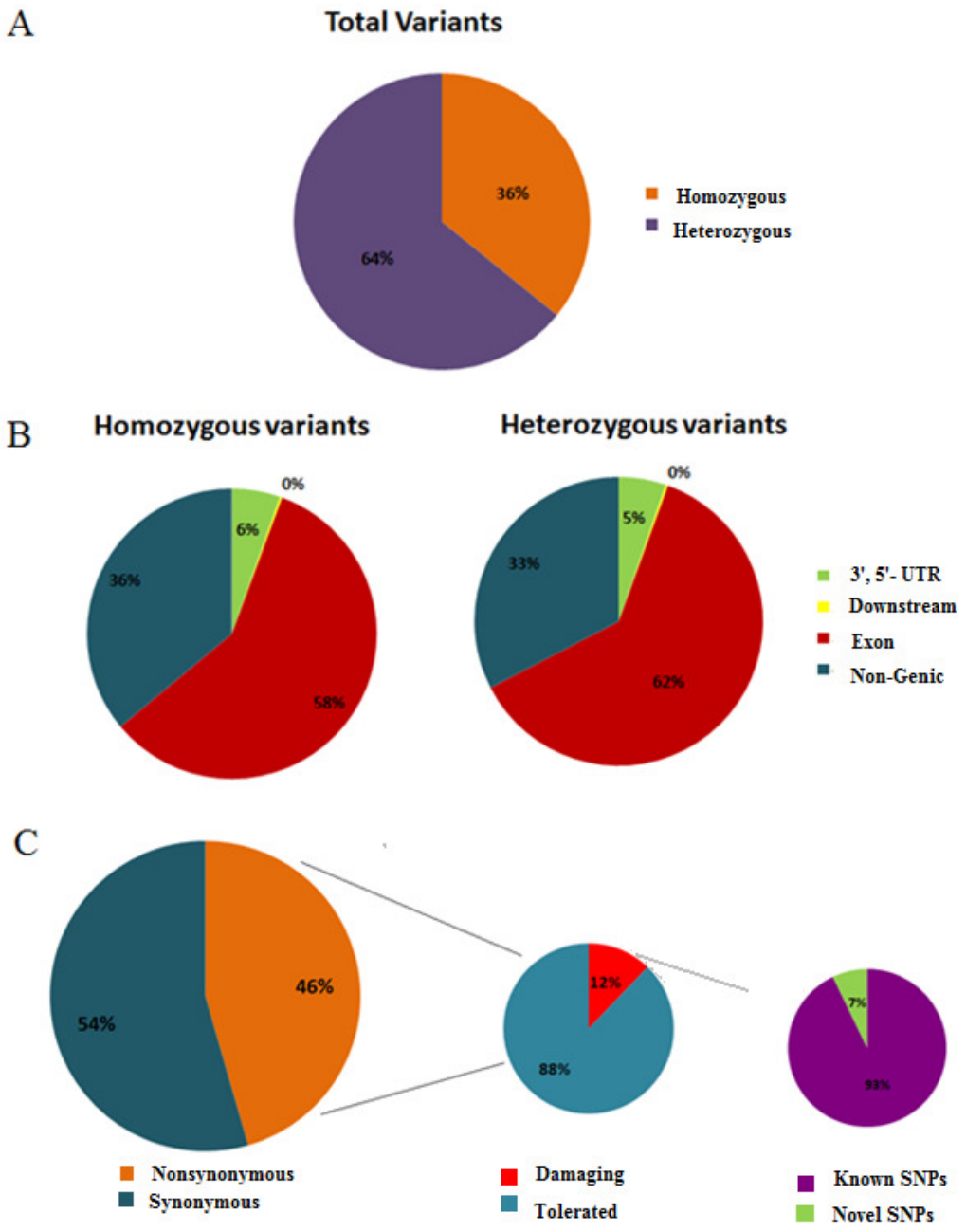
Filtering variants

The primary filter, which was used to exclude bystander variations, was based on variant function—namely, if the variant affects coding regions or other non-coding intronic regions. So, I rejected all non-coding variants (11,403) as well as synonymous variants (8,766). Moreover, considering the autosomal recessive mode of inheritance, I rejected variants that were heterozygote and synonymous, reducing the candidate list to 2,769 positions, out of which 336 positions were variants predicted to be damaging to the protein function. For the final step, variations that were documented in dbSNP or in the 1000 Genomes project were eliminated.

Novel variants

After filtering out the dbSNP and 1000 genome common variants, 24 variants remained, amongst which 13 were predicted to be damaging and 11 were tolerated (Table 3.1). As a complementary approach, MutationTaster [6] and SIFT [7] were

Figure 3.1 distribution of variants in next generation sequencing SNP report. A) total number of variants and the distribution of homozygous and heterozygous variants. B) Distribution of coding and non-coding variants in homozygous and heterozygous category. C) proportion of the genome of the proband that contained novel damaging homozygous variants.



used to evaluate disease-causing potential of variants. However the results of predictions were not consistent between these two tools (Table 3.1), therefore, I trusted on SIFT. In addition, in silico analysis were performed to prioritize variants based upon the expression patterns of candidate genes [6], their known ortholog function [7], as well as homology from syntenic regions in multiple model organisms (murine [8], zebrafish [9] and Drosophila [10]).

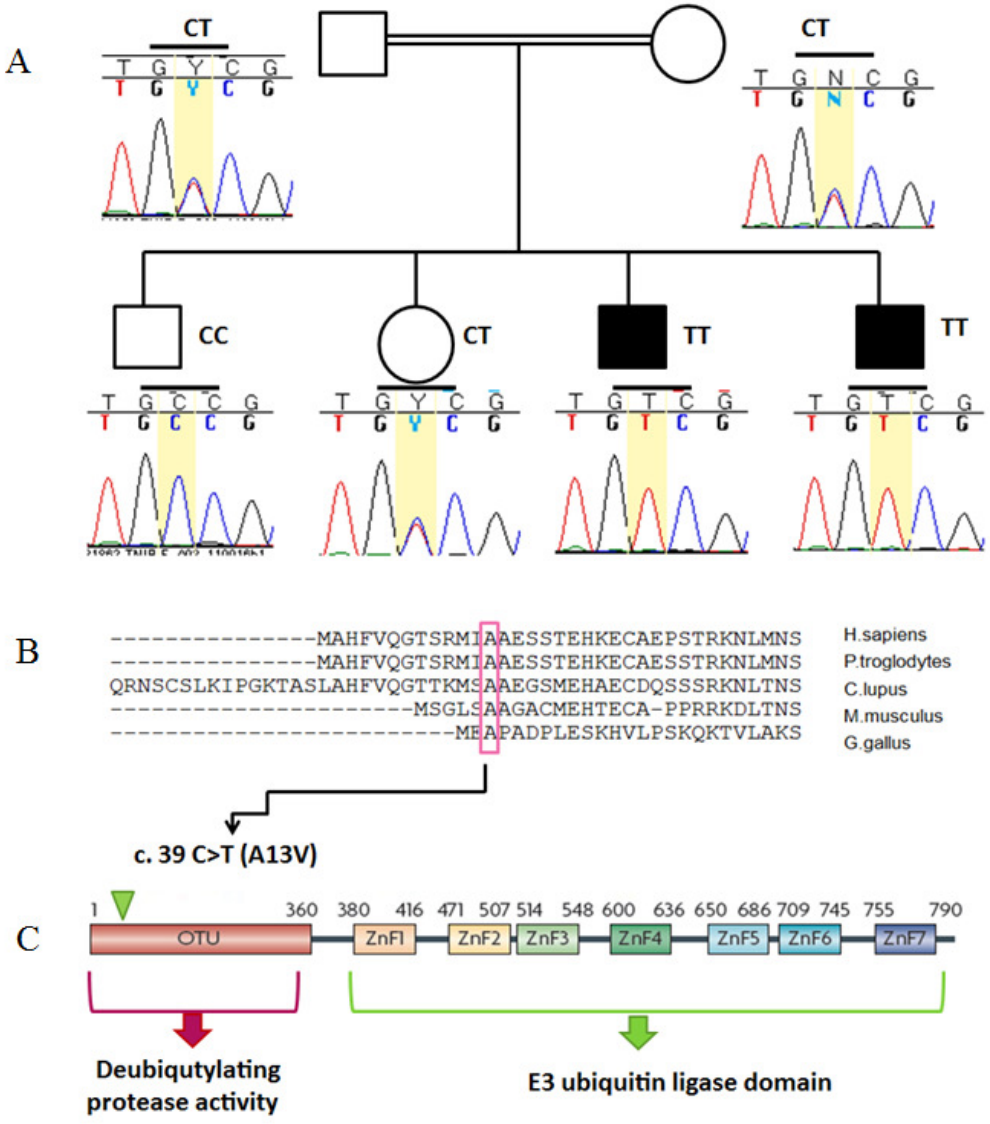
Two genes, namely *TNIP3* and *LLRC16B* (on the IBD regions on chromosome 4 and 14, respectively), contained novel homozygous variants. Sanger sequencing confirmed the homozygous presence of the former variant in both affected individual and its presence as a heterozygous change in both parents (Figure 3.2). The latter variant, however, was a sequencing error. In addition to these two variants, a few other SNPs were also checked with Sanger sequencing (*FARSA*, *MFSD10* and *USHBP1*); however, they were false positive excluding their potential involvement in this disease.

The variant in *TNIP3* is an A13V (c.39 C>T) change, which is located in the OTU domain of the protein. The domain, in turn, has a deubiquitylating activity. Screening of the 144 normal controls revealed the absence of this variant in healthy people.

Table 3.1 Complete list of homozygous nonsynonymous variants

Gene	Description	Substitution	Prediction (SIFT)
MFSD10	Major facilitator superfamily domain-containing protein 10	R116G	Damaging
SLC4A2	Solute carrier family 4 member 2	P1081L	Damaging
FARSA	Phenylalanine--tRNA ligase alpha chain	P247L	Damaging
PXDNL	Peroxidasin-like protein Precursor	P740S	Damaging
NAGK	N-acetyl-D-glucosamine kinase	V194A	Damaging
RBBP7	Retinoblastoma-binding protein 7	R37H	Damaging
KCNH3	Potassium voltage-gated channel subfamily H member 3	P856R	Damaging
LRRC16B	Leucine-rich repeat-containing protein 16B	R158W	Damaging
TNIP3	TNFAIP3-interacting protein 3	A13V	Damaging
USHBP1	USH1C-binding protein 1	P146L	Damaging
CTBP2	C-terminal-binding protein 2	N567K	Damaging
CTBP2	C-terminal-binding protein 2	R697W	Damaging
CTBP2	C-terminal-binding protein 2	G568V	Damaging
CER1	Cerberus Precursor	N184K	Tolerated
L3MBTL	Lethal(3)malignant brain tumor-like protein	S190R	Tolerated
VTN	Vitronectin Precursor	R20W	Tolerated
HLA-DQA1	HLA class II histocompatibility antigen	G78R	Tolerated
AC069029.1	Pyroglutamyl-peptidase 1-like protein	E185Q	Tolerated
AC022098.3	Paralemmin-3	E206K	Tolerated
FBXW8	F-box/WD repeat-containing protein 8	G514R	Tolerated
PJA1	E3 ubiquitin-protein ligase Praja1	A429T	Tolerated
IFT140	Intraflagellar transport protein 140 homolog	A437V	Tolerated
CTBP2	C-terminal-binding protein 2	A715V	Tolerated
PRAM1	PML-RARA-regulated adapter molecule 1	V100F	Tolerated

Figure 3.2 TNIP3 mutational analysis in anophthalmia family. A) Sequencing of *TNIP3* shows the carrier status of both parents and the homozygous mutation c.39-C>T of both affected siblings. B) Sequence comparison of *TNIP3* in different species. The amino acid that is changed is in the pink box. C) Domain structure of *TNIP3* protein and location of the variant (green arrow).



Insertions/Deletions (InDels)

In addition to SNPs, the genome of the affected child had several insertion deletion and frameshift mutations. To consider InDels as causative sequence variation, the same criteria, as described for SNPs, were used for excluding the spurious variants. For frameshift mutations, the locations of the sequence change were also important because those changes in the beginning of the sequence are probably more damaging than those variations near the end of the sequence.

Of a total of 5,822 indels identified, only 1,764 were in coding regions. Approximately, 93% of theses coding InDels were heterozygous. Of the remaining 61 homozygous substitutions, 11 were in-frame deletions amongst which, 5 variants were known SNPs and one was located on 3'-UTR (Table 3.2A). 15 insertions, ranging from 3-12 bp, were also identified (Table 3.2B). There were also 33frameshift mutations, in transcripts with no protein product, 3′and 5′ UTR, as well as exonic region of protein coding transcripts (Table 3.2C).

Table 3.2 Complete list of InDel homozygous substitutions**.** A) In-frame amino acid deletion. B) In-frame amino acid insertion. C) frameshift mutations

A)

Ref gene	**Location**	**Number of exons**	**InDel**	**Note**
OR14A16	Exon 1	1	-AGG	Known SNP
CELA1	5-UTR	8	-AAGGAC	Known SNP
LRRC49	Exon 11	16	-CAA	-
USP36	Exon 17	20	-TTTTTC	-
ZNF2	Exon 4	4	-GCG	-
NEFH	Exon 4	4		Known SNP
VEGFC	3-UTR	7	-TCA	-
BAT2				-
HLA-DQA1				-
C6orf164	Exon 1	2	-GAT	Known SNP
GPR112	Exon 9	17	-GAT	Known SNP

B)

Ref gene	**Location**	**Number of exons**	**InDel**
SERINC2	Exon 9	10	+CAG
DNHD1	Exon 21	43	+TGCCCTACTGCA
REC8	Exon 9	20	+GAA
EME1	Exon 2	9	+AGC
C18orf25	Exon 4	5	+CTG
RTTN	Exon 7	49	+CTC
OR7C2	Exon 1	1	+ATC
ZNF714	Exon 5	5	+ATA
TMEM37	Exon 2	2	+GTGTGC
LTF	Exon 2	17	+CTT
MUC13	Exon 2	12	+AAG
HAVCR1	Exon 3	8	+GTT
AKAP12	Exon 3	4	+GAG
PCLO	Exon 5	25	+TCA
C8orf59	3-UTR	4	+AACATT

C)

Ref gene	Location	Number of exons	InDel	Note
HRNR	3-UTR			
ZNF518A				No protein product
MTCH2	5-UTR			No protein product
MMP12				No protein product
VPS11				No protein product
EMG1				No protein product
NR2E3	Last nt in exon 8	11	-C	
CNTNAP4	Last nt in exon 1	25	+T	
KRT24	31st nt in exon 2 (83)	8	-A	
HSH2D	Last nt in exon 8	9	-A	
LIG1	104th nt in exon 22 (145)	28	-CA	
VSIG10L	3-UTR			
SIGLEC12	230th nt in exon 1 (483)	8	+C	
ZNF761				No protein product
ZNF274	Last nt in exon 4	8	+G	
TNFAIP6	Last nt in exon 6	6	-A	
PLK1S1				No protein product
C22orf46	3-UTR			
SCAMP1				No protein product
HAVCR1	160th nt in exon 3 (333)	9	-T	Membrane receptor for hepatitis A virus
CYFIP2	72nd nt in exon 4 (77)	33	+C	
SLC22A1	213th nt in exon7 (215)	10	-TGGTAAGT	
AOAH	199th nt in exon 22 (274)	22	+T	
ATG9B				No protein product
UBXN8				No protein product
ADAM5P				No protein product
PRKDC	First nt in exon 31(131)	87	+G	

Discussion

Whole exome sequencing followed by a systematic analysis of the SNPs and InDels data with stringent filtering criteria, such as mode of inheritance, conservation, and loss of function prediction, were used to isolate the disease causing mutations in consanguineous anophthalmia pedigree. The only variation that passed the multiple exclusion criteria was Ala13Val in *TNIP3* gene. Nevertheless, there were multiple nonsynonymous homozygous variants and insertion/deletion/frameshift sequence alterations which any of them could be a hypothetical suspect and the proof of their innocence need further biochemical/biological characterization (Table 3.1).

For the first step of our analysis, the possibility that the disorder is caused by known MAC genes was evaluated. Looking for known mutations in previously identified disease associated genes is the most straightforward method to identify causal mutations in an individual [11]. However, I could not find any homozygous variation or compound heterozygous variations in known MAC genes indicating that the disorder is caused by a new autosomal gene and further emphasize on the genetic heterogeneity of anophthalmia.

In view of the fact that IBD intervals have an overwhelming probability to hold the disease mutations in inbred families [12], I analyze the variants in the IBD

intervals in the next step. Theoretically the longest segment is likely to carry the disease mutation [13] because the length of homozygous runs is conversely correlated with the number of recombination events from the common ancestor [14]. The longest homozygous region in our analysis was on chromosome 4q26-28.1, which bears 11variants. A few heterozygote variants were also identified in this region (3 out of 11 variants). Change in the ratio of variant reads to the reference reads due to the presence of pseudogenes is one possibility for observing a heterozygote variant in a homozygous block. [15]. Alternatively, these variants might be just the sequencing artefact and their existence needs to be validated in the first place.

All the variants were found in the human variation depositories except *TNIP3*. This variant is our best candidate so far. In addition to the novelty of this variant and its absence in normal control, its segregation perfectly matches with our hypothesis of the autosomal recessive mode of inheritance in this pedigree (Figure 3.2). Moreover, the variant is in a conserved domain of the protein (Figure 3), and, predicted to be damaging to the protein function.

TNIP3 or *A20* is a cytoplasmic zinc finger protein which was first discovered in 1990 as a main response-gene of Tumor Necrosis factor-α (TNF-α) [16]. As an important regulator of inflammation, *TNIP3* downregulates the *NF-kappa B* signaling by its ubiquitin-editing activity [17]. This gene contains 8 transcripts

and 5 protein isoforms. The isoform 1 contains 11 exons and encodes a protein of approximately 39 kDa (325 amino acids). *TNIP3* has anti-apoptotic activity and the null mice for this gene exhibits severe inflammation in multiple organs as a result of failure in downregulating the *NF-kappa B* signaling activity and increased apoptosis [18]. Given that *A20* attenuates *NF-kappa B* signaling and Tumor *TNF-α* mediated programmed cell death, it would logically follow that reduced *TNIP3* activity due to mutation might promote cell death, therefore, it may propose *TNIP3* as a plausible candidate gene in this pedigree.

Although *TNIP3* seems as a potential candidate, all other nonsynonymous variants elsewhere in the genome of the proband should be evaluated for their pathogenesity. The reason is that homozygosity mapping may have not picked the true disease-causing intervals due to genetic heterogeneity of anophthalmia and because the mode of inheritance might have been selected wrongly [19-20].

In evaluating these variants, priority was given to those which were nonsynonymous as they tend to be disruptive to proteins and, hence, of large effect. Of the 24 nonsynonymous homozygous variants that passed the threshold, and specifically the predicted damaging variants, none of them turned out to be potential casual genes based on their function or expression pattern in the eye. Five of them were sequencing artifact, however, to exclude the rest of them as being pathological changes further validations are required.

The majority of homozygous frameshift mutations seem not to be pathogenic as they were on regions without any genes, transcripts with no protein product, and 3`-UTR. In addition, 8 of frameshift substitutions contained nucleotide insertions or deletions in the beginning or at the end of exons. These are perhaps technical artifacts that might have arisen during sequencing or alignment of sequence reads. In the remaining 5 frameshift mutations on *KRT24* (type 1 keratin family), *SIGLEC12* (immunoglobin superfamily), *HAVCR1* (receptor for hepatitis A virus), *LIG1* (DNA ligase) and *AOH*, expression and functional data provided few clues to prioritize the list.

In conclusion, our study showed the difficulties in distinguishing the true pathogenic mutations from background polymorphism when the sequencing data for only one individual is available. Further sequencing data from the second affected sibling as well as unaffected individual is likely to lead to the isolation of the causative gene in this pedigree

Reference:

1. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010; 42(1):30-5.

2. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009; 106(45):19096-101.

3. Mardis ER. A decade's perspective on DNA sequencing technology. Nature 2011; 470(7333):198-203.

4. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet 2011; 48(9):580-9.

5. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking Mendelian disease using exome sequencing. Genome Biol 2011; 12(9):228.

6. Pagon RA. Ocular coloboma. Surv Ophthalmol 1981; 25(4):223-36.

7. Fraser FC, Skelton J. Possible teratogenicity of maternal fever. Lancet 1978; 2(8090):634.

8. Spagnolo A, Bianchi F, Calabro A, Calzolari E, Clementi M, Mastroiacovo P, Meli P, Petrelli G, Tenconi R. Anophthalmia and benomyl in Italy: a multicenter study based on 940,615 newborns. Reprod Toxicol 1994; 8(5):397-403.

9. Hornby SJ, Ward SJ, Gilbert CE. Eye birth defects in humans may be caused by a recessively-inherited genetic predisposition to the effects of maternal vitamin A deficiency during pregnancy. Med Sci Monit 2003; 9(11):HY23-6.

10. Miller MT, Stromland K. Teratogen update: thalidomide: a review, with a focus on ocular findings and new potential uses. Teratology 1999; 60(5):306-21.

11. Raca G, Jackson C, Warman B, Bair T, Schimmenti LA. Next generation sequencing in research and diagnostics of ocular birth defects. Mol Genet Metab 2010; 100(2):184-92.

12. Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 1987; 236(4808):1567-70.

13. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Malik Sharif S, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF. Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet 2006; 78(5):889-96.

14. Clark AG. The size distribution of homozygous segments in the human genome. Am J Hum Genet 1999; 65(6):1489-92.

15. Walsh T, Lee MK, Casadei S, Thornton AM, Stray SM, Pennil C, Nord AS, Mandell JB, Swisher EM, King MC. Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. Proc Natl Acad Sci U S A 2010; 107(28):12629-33.

16. Dixit VM, Green S, Sarma V, Holzman LB, Wolf FW, O'Rourke K, Ward PA, Prochownik EV, Marks RM. Tumor necrosis factor-alpha induction of novel gene products in human endothelial cells including a macrophage-specific chemotaxin. J Biol Chem 1990; 265(5):2973-8.

17. Hymowitz SG, Wertz IE. A20: from ubiquitin editing to tumour suppression. Nat Rev Cancer 2010; 10(5):332-41.

18. Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. Science 2000; 289(5488):2350-4.

19. Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, Cideciyan AV, Haider N, Stone EM, Sheffield VC, Wright AF. Pitfalls in homozygosity mapping. Am J Hum Genet 2000; 67(5):1348-51.

20. Laurier V, Stoetzel C, Muller J, Thibault C, Corbani S, Jalkh N, Salem N, Chouery E, Poch O, Licaire S, Danse JM, Amati-Bonneau P, Bonneau D, Megarbane A, Mandel JL, Dollfus H. Pitfalls of homozygosity mapping: an extended consanguineous Bardet-Biedl syndrome family with two mutant genes (BBS2, BBS10), three mutations, but no triallelism. Eur J Hum Genet 2006; 14(11):1195-203.

Chapter 4:

General discussion

In the preceding chapters, I have presented and discussed the combined use of homozygosity mapping, CNV analysis and exome sequencing in the deep study of a rare congenital eye disease- anophthalmia- in a consanguineous pedigree. The key finding of this study was the identification of five novel loci in this pedigree with high density SNP microarray, which indicates that additional loci/gene is involved in the pathogenesis of this condition in this family. Moreover, this technique helped us narrow areas of interest for further directed sequencing. In keeping with the homozygosity mapping data, our systematic analysis of whole exome sequencing (WES) data indicates that the disorder is caused by a new autosomal gene as I exclude the presence of any variants in known genes associated with MAC. However, isolating the pathogenic mutation from the large number of background polymorphisms was challenging and even after applying computational filters on the identified variants, multiple candidate genes remained indicating that WES data for a single subject is not enough to specifically identify causal variants. In this chapter, I will discuss some concern in using WES in a genetically heterogeneous condition, like our case, and will provide some suggestions for future directions.

The power of combining homozygosity mapping with WES as a strategy for unravelling the molecular cause of the recessive disease is in its ability to help exclude irrelevant parts of the exome or genome prior to the application of other computational filters [1-3]. In this setting, each proband was required to have two mutated allele in the same gene and identification of such variants in IBD segments are highly efficient at even low level of per-base coverage (e.g., 5X

coverage) because there should not be too many rare homozygous protein-altering variants in these segments [4]. Therefore, theoretically, data from only one affected individual will suffice to find the pathogenic mutation. This approach has been used successfully in a number of WES projects and related studies to identify or confirm novel disease genes [5-12].

However, there are two concerns regarding the use of homozygosity mapping data for filtering WES variants that should be considered to avoid misinterpretation of the data: First, the frequency of alleles in variation depositories, like dbSNP or 1000 genome, for non-Western societies- where consanguineous marriages are common- are less represented, which reduce the power of eliminating common polymorphisms [13-15]. Therefore, the phenotypically normal people who are carrier for a recessive disease gene will be genotyped and the pathogenic allele will be deposited in the database [16]. Second, focusing on variants in runs of homozygosity in genetically heterogeneous diseases, as in anophthalmia case, could be problematic as well [17], due to the possibility of the occurrence of separate *de novo* mutation(s) in different genes or the same gene but in a different location (compound heterozygote). This means that the causative mutation(s) could be found anywhere in the genome. The reason for assuming the possibility of the presence of *de novo* mutation is that it has recently been shown that the rate of *de novo* mutations in humans is higher than previously thought. The per generation mutation rate in humans has been estimated at between 7.6×10^{9} and 2.2×10^{8}, or roughly 50 to 100 new mutations per new born, which corresponds to 0.86 *de novo* amino acid altering mutations [18-19]. Hence, it seems logical to

hypothesize that *de novo* mutations in this genetically heterogeneous condition might be common. However, allowing for this model will inflate candidate lists, and more sophisticated approaches will need to conduct candidate prioritization in a single WES data set.

In addition to probable flaws in selecting the mode of inheritance and, consequently, the wrong filtering criteria, WES itself has several inherent limitations, which should be considered in any exome sequencing studies. The first concern relates to sequencing coverage, which is not the same across the genome and this may result in missing many regions of interest [20]. In addition using the current platforms, some regions of the genome are difficult to align to the reference genome, including repetitive or GC rich regions which can potentially lead to missed variants or an excess of variant calls [21]. On the other hand, a subset of genes appeared enriched for novel variants, e.g. CDC27, perhaps due to misalignment of the sequence reads or the presence of highly similar sequences in the genome (pseudogenes) which may mistakenly result in an excess of variant calls for a gene. For those genes that are known to be duplicated or have paralogues in the genome, these variants can be removed from consideration but for duplicated sequences that are not identified in human reference genome, these spurious variants will not be removed and can lead to false positive variants unrelated to disease. Finally, the length of the gene can affect the number of variant calls. For example, in a study on Kabuki syndrome, MUC16 was the only gene that was common in 10 affected individuals. However, it turned to be a

wrong candidate because the long length of this gene was the reason for a higher number of variant calls for this gene across all individuals [22] .

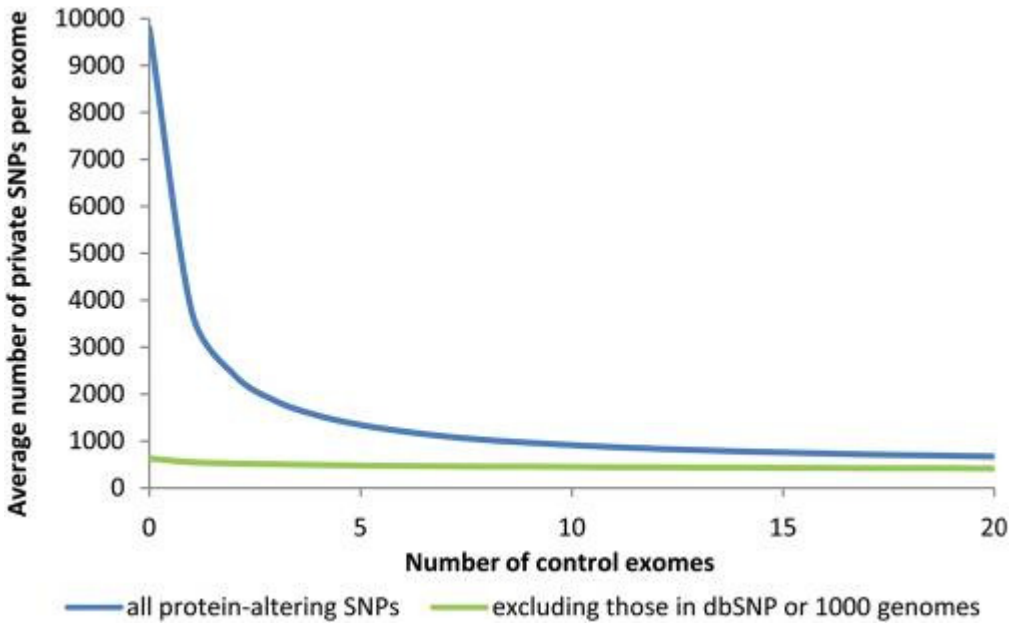
To deal with the above mentioned issues, one option is to sequence more affected individuals or perform case-parent trios analysis strategy to identify potentially pathogenic [23-24]. The reason is that under the assumption of random distribution of variants in the population, only the true pathogenic allele will be present in all affecteds (Figure 4.1). For example if 5% of 20,000 target genes show rare novel variants, in one individual, 1000 genes will contain sequence variants. However, if we consider two affecteds at the same time, 5% of 1000 genes or 50 candidates will be predicted to have the same variants in both individuals. The number of candidate genes will be reduced to less than one, if we sequence a third individual [2]. Therefore, this approach will enable us to identify rare polymorphism versus true disease causing mutation.

For future work, DNA samples of the second affected and one unaffected sibling have been sent for exome sequencing. Analyzing three sets of exome data in this pedigree is beneficial as it allows for population-matched controls, especially considering that this family comes from the Middle East and the allele frequency for this ethnic group may be under-represented in the current variation databases [21-22, 25]. In addition, it would be easier to control for technical artifacts that may arise during the sequencing or alignment of sequence reads, so it is predicted that the three exome data may provide near-exact identification of false-positive candidate genes and, therefore, reduce the number of candidate genes for further analysis.

Another option for future work is to perform WGS. In this study, the analysis was restricted to coding variants. However, it is possible that mutation in regulatory elements and distant enhancers be present and affect the phenotype [26]; even though the majority of mutations associated with congenital diseases have been found in or near exonic regions. There are some reports that have shown the association of intronic mutations with hereditary diseases [27]. Therefore, there will be a need to perform whole genome sequencing to gain a broader view of possible genetic alterations that might be underlying the anophthalmia phenotype in this pedigree.

Overall, the discovery of the genetic basis of anophthalmia will substantially expand our understanding of biology of this rare disease, in general, and lessons learned from this rare disease can be also relevant to common diseases.

Figure 4.1 Effect of increasing number of control exomes on private variants observed in a single exome, with and without the use of dbSNP and 1000 genomes data, derived from Ng SB *et al.*, Hum Mol Genet **19**(R2):R119-24 [16]. The number of private mutations observed in an individual from sequential addition of control exomes was averaged from 10 000 permutations of 21 published exomes of non-African ancestry [21-22, 25, 28].



References:

1. Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 1987; 236(4808):1567-70.

2. Robinson PN, Krawitz P, Mundlos S. Strategies for exome and genome sequence data analysis in disease-gene discovery projects. Clin Genet 2011; 80(2):127-32.

3. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008; 322(5903):881-8.

4. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009; 106(45):19096-101.

5. Rehman AU, Morell RJ, Belyantseva IA, Khan SY, Boger ET, Shahzad M, Ahmed ZM, Riazuddin S, Khan SN, Friedman TB. Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. Am J Hum Genet 2010; 86(3):378-88.

6. Nikopoulos K, Gilissen C, Hoischen A, van Nouhuys CE, Boonstra FN, Blokland EA, Arts P, Wieskamp N, Strom TM, Ayuso C, Tilanus MA, Bouwhuis S, Mukhopadhyay A, Scheffer H, Hoefsloot LH, Veltman JA, Cremers FP, Collin RW. Next-generation sequencing of a 40 Mb linkage interval reveals TSPAN12 mutations in patients with familial exudative vitreoretinopathy. Am J Hum Genet 2010; 86(2):240-7.

7. Volpi L, Roversi G, Colombo EA, Leijsten N, Concolino D, Calabria A, Mencarelli MA, Fimiani M, Macciardi F, Pfundt R, Schoenmakers EF, Larizza L. Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. Am J Hum Genet 2010; 86(1):72-6.

8. Lalonde E, Albrecht S, Ha KC, Jacob K, Bolduc N, Polychronakos C, Dechelotte P, Majewski J, Jabado N. Unexpected allelic heterogeneity and spectrum of mutations in Fowler syndrome revealed by next-generation exome sequencing. Hum Mutat 2010; 31(8):918-23.

9. Johnston JJ, Teer JK, Cherukuri PF, Hansen NF, Loftus SK, Chong K, Mullikin JC, Biesecker LG. Massively parallel sequencing of exons on the X

chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. Am J Hum Genet 2010; 86(5):743-8.

10. Wang JL, Yang X, Xia K, Hu ZM, Weng L, Jin X, Jiang H, Zhang P, Shen L, Guo JF, Li N, Li YR, Lei LF, Zhou J, Du J, Zhou YF, Pan Q, Wang J, Li RQ, Tang BS. TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing. Brain 2010; 133(Pt 12):3510-8.

11. Estrada-Cuzcano A, Neveling K, Kohl S, Banin E, Rotenstreich Y, Sharon D, Falik-Zaccai TC, Hipp S, Roepman R, Wissinger B, Letteboer SJ, Mans DA, Blokland EA, Kwint MP, Gijsen SJ, van Huet RA, Collin RW, Scheffer H, Veltman JA, Zrenner E, den Hollander AI, Klevering BJ, Cremers FP. Mutations in C8orf37, Encoding a Ciliary Protein, are Associated with Autosomal-Recessive Retinal Dystrophies with Early Macular Involvement. Am J Hum Genet 2011.

12. Huang L, Szymanska K, Jensen VL, Janecke AR, Innes AM, Davis EE, Frosk P, Li C, Willer JR, Chodirker BN, Greenberg CR, McLeod DR, Bernier FP, Chudley AE, Muller T, Shboul M, Logan CV, Loucks CM, Beaulieu CL, Bowie RV, Bell SM, Adkins J, Zuniga FI, Ross KD, Wang J, Ban MR, Becker C, Nurnberg P, Douglas S, Craft CM, Akimenko MA, Hegele RA, Ober C, Utermann G, Bolz HJ, Bulman DE, Katsanis N, Blacque OE, Doherty D, Parboosingh JS, Leroux MR, Johnson CA, Boycott KM. TMEM237 Is Mutated in Individuals with a Joubert Syndrome Related Disorder and Expands the Role of the TMEM Family at the Ciliary Transition Zone. Am J Hum Genet 2011; 89(6):713-30.

13. Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, Nickerson DA. Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. Nat Genet 2003; 33(4):518-21.

14. Via M, Gignoux C, Burchard EG. The 1000 Genomes Project: new opportunities for research and social challenges. Genome Med 2010; 2(1):3.

15. Bittles A. Consanguinity and its relevance to clinical genetics. Clin Genet 2001; 60(2):89-98.

16. Parker SE, Mai CT, Canfield MA, Rickard R, Wang Y, Meyer RE, Anderson P, Mason CA, Collins JS, Kirby RS, Correa A. Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004-2006. Birth Defects Res A Clin Mol Teratol 2010; 88(12):1008-16.

17. Koziell A, Grech V, Hussain S, Lee G, Lenkkeri U, Tryggvason K, Scambler P. Genotype/phenotype correlations of NPHS1 and NPHS2 mutations in

nephrotic syndrome advocate a functional inter-relationship in glomerular filtration. Hum Mol Genet 2002; 11(4):379-88.

18. Lynch M. Rate, molecular spectrum, and consequences of human mutation. Proc Natl Acad Sci U S A 2010; 107(3):961-8.

19. Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science 2010; 328(5978):636-9.

20. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet 2011; 48(9):580-9.

21. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010; 42(1):30-5.

22. Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 2010; 42(9):790-3.

23. Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. A de novo paradigm for mental retardation. Nat Genet 2010; 42(12):1109-12.

24. Robinson PN. Whole-exome sequencing for finding de novo mutations in sporadic mental retardation. Genome Biol 2010; 11(12):144.

25. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009; 461(7261):272-6.

26. Lettice LA, Hill AE, Devenney PS, Hill RE. Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly. Hum Mol Genet 2008; 17(7):978-85.

27. Cooper DN, Chen JM, Ball EV, Howells K, Mort M, Phillips AD, Chuzhanova N, Krawczak M, Kehrer-Sawatzki H, Stenson PD. Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. Hum Mutat 2010; 31(6):631-55.

28. Ng SB, Nickerson DA, Bamshad MJ, Shendure J. Massively parallel sequencing and rare disease. Hum Mol Genet 2010; 19(R2):R119-24.

Thesis Appendices:

Studying the Role of *GDF3* in Pathogenesis of Glaucoma

Introduction

Glaucoma, is the second leading cause of permanent blindness worldwide [1]. Clinical features include abnormal optic nerve head, decreased sensitivity of visual function in the mid- peripheral field, with progressive loss that eventually involves central and peripheral fields. This condition is frequently accompanied by elevated intra ocular pressure (IOP); however, high IOP is neither necessary nor sufficient for the onset of the disease [2]. Vision loss results from the degeneration of the axons in the optic nerve head that causes irreversible blindness if untreated.

In Primary Open Angle Glaucoma (POAG), as the most frequently occurring form of the disease, elevated IOP results from an increase in aqueous humor outflow resistance in the trabecular meshwork (TM). So far, 25 loci have been linked to POAG while only three underlying genes, *Myocilin* [3], *Optineurin* [4] and *WDR36* [5] are identified. Even though there is strong evidence for genetic influence in glaucoma pathogenesis, only a small proportion of the disease has a characterized genetic abnormality and the pathogenic mechanism remains largely unknown.

In the last couple of years, members of the TGF-β superfamily came into the focus of glaucoma research. TGF-β ligands which are known to regulate a plethora of important developmental processes, contain more than 30 members and are classified into two branches: TGF-β/activin/Nodal that signals through Smads 2/3 and BMP/GDF that signals through Smad 1/5/8.[6]. In general, BMP signals are mediated by binding to and bringing together type I and type II serine-

threonine kinase receptor on the cell surface. This will allow receptor II to phosphorylate the receptor I, which in turn, recruits the Smad proteins to convey the signal from the cell surface to nucleus. Smad 1, 2, 3, 5 and 8 (receptor regulated Smad) are directly activated by receptor I and undergo homotrimerization and formation of heteromeric complex with Smad 4 (Co-mediator Smad). This complex is, then, subsequently translocated into the nucleous and regulates the transcription of target genes [6]. In contrast, activation of TGF-β receptors by their ligands causes the activation and phosphorylation of Smad2 and 3, followed again by heterodimerization with Smad4 and translocation into the nucleus.

Several studies have examined the importance of BMP signals in pathogenesis of the glaucoma. Wordinger *et al.* described the expression of BMP-2, BMP-4, BMP-5, and BMP-7 and the BMP receptors in two tissues involved in glaucoma pathogenesis; optic nerve head and trabecular meshwork [7]. Amongst them BMP4 is of special interest since the deficiency of BMP-4 leads to abnormalities of the optic nerve and an elevated IOP [8]. These reports testify the importance of BMP signals in maintaining normal homeostasis in the ocular tissues associated with glaucoma, and also suggest that alterations in growth factor/growth factor receptors may play a role in glaucoma pathogenesis.

In this project, I selected GDF3 for study in the pathogenesis of glaucoma. GDF3, also called Vgr-2, was first isolated from mouse embryonic day (E) 6.5 cDNA embryonic library by homology to *Xenopus* VG1 (57% amino acid identity, 79% similarity) [9]. It was placed in the BMP/GDF subclass of the TGF-β family for

its sequence homology. GDF3 protein has six of the seven cysteines that are present in other TGF-β ligands. This missing cysteine is involved in intermolecular interaction between TGF-β ligands.

Regarding the function of GDF3, there are two contradictory reports in the literature. Brivanlou (2006) examined the function of GDF3 in the context of stem cell differentiation and reported that this gene is one of the several key factors responsible for maintaining the pluripotency state in stem cells. They found that the transition from pluripotency to differentiated state is accompanied by reduction in GDF3 while there is an increase in BMP signals and concluded that GDF3 is an inhibitor of its own subfamily, BMP signaling [10]. On the other hand, Chen showed that GDF3 is poorly processed in cell culture and that it needs an EGF-CFC co-receptor called Cripto, which is the putative receptor for Nodal ligands, to signal through a Nodal-like pathway. These combined findings raise the intriguing possibility that GDF3 acts as a bi-functional protein, to regulate the balance between the two modes of TGF-β signaling.

Very little is known about GDF3 and its role in disease, however, the association of large deletion and duplication of chromosome 12p, where GDF3 is located, with ocular disease suggests that this gene might be involved in ocular diseases. Recently, Ye (2010) showed the importance of this gene in the development of vertebrates' eyes by demonstrating the contribution of GDF3 to 1.5% of microphthalmia, anophthalmia and coloboma (MAC), a spectrum of congenital ocular disorders that cause paediatric vision loss. Since the mutation of GDF6, as another BMP, results in MAC and we have an increasing body of evidence that

indicate it also causes a milder late onset disease (glaucoma), I hypothesized that other members of the TGFβ family, including GDF3, could play contributing roles in glaucoma. Thus, probably mutations in the GDF-3 gene can cause cells in retina to become vulnerable to death and, subsequently, cause glaucoma either directly by inducing neuronal loss or indirectly by causing death of the other cells in the eye. In light of this finding, I screened a panel of glaucoma patients and found two mutations. Interestingly, these two mutations have been also found in the MAC panel. Western blot showed reduced protein production for both mutant and luciferase assay showed decreased protein activity for one of them [11]. For further study the association of these mutations and glaucoma disease, morpholino inhibition of dvr1, an ortholog of GDF1/GDF3, was pursued to investigate knockdown phenotypes and see the effect on the density of retinal ganglion cells (RGCs).

Material and methods

PCR

Primers were designed for 2 coding exons of GDF3 for screening of patients with Normal Tension Glaucoma (NTG) and High Tension Glaucoma (HTG) (Table 1). The primers were designed using Primer 3 (V.04.0).

Cycling temperature: 1) 95 ^{0}C- 5 min, 2) 95 ^{0}C- 30 sec, 3) 57 ^{0}C- 45 sec, 4) 72 ^{0}C- 1 min, 5) repeat (2-4) 35 times, 6) 72 ^{0}C- 5 min, 7) Final hold at 10 ^{0}C.

Chemical reagents: each reaction, carried out in a total of 25 µl, had a final PCR buffer concentration of 1X (10X is 200 mM Tris, pH 8.4, 500 mM KCL; Invitrogen), dNTP concentrations of 0.2 mM (Invitroge), glycerol concentration of 10% v/v (Anachemia), formamide concentration of 5% (Sigma) and 0.5 U of Taq polymerase (New England BioLabs).

Amplified PCR products were validated by gel electrophoresis-ethidium bromide stained 1% agarose gel. After validation, the PCR products were purified (Sigma Spin Post-Reaction Clean-Up Columns) prior to sequencing.

Sequencing

The sequencing reaction was done with the Applied Biosystems BigDye Terminator v3.1 Cycle sequencing kit and unicorporated dyes were removed with ethanol. Sequencing was run on the AApplied bisystems (ABI) 3130/XL (The Applied Genomic Centre, Alberta, Canada) or 3737/XL (CHUL Research Center,

Quebec, Canada). Genetic Analyzer with the same primer pairs used to amplify GDF3 amplicons.

Restriction enzyme digestion

Variants that were identified in GDF3 were validated with restriction enzyme (R.E.) digestion. Two restriction enzymes, *HpaII* (New England BioLab; recognition sequence-CCGG) and *TsprI* (New England BioLab; recognition sequence-), were used for validating variants R195Q and R266C, respectively. 10 µl of PCR products (at concentration of ~ 100ng/µl) was incubated with 1µl of restriction enzymes (20,000U/ml) at 37^0C for 16 hour. The digested products were analyzed with ethidium bromide stained 1% agarose gel electrophoresis.

Western blot

1. **Cell transfection**

Monkey kidney cells COS-7 were cultured to 80% confluency before transfection in DMEM plus 10% FBS at 378. For each V5-tagged-GDF3 plasmid, 4 mg of DNA was transfected and transiently expressed in COS-7 cells by FuGENE (Roche Diagnosis, IN, USA) in a ratio of 2:1 fugene to µg DNA. Briefly, 4µg of each vector plus 1µg of LacZ added to an eppendorf tube containing 400 µl of DMEM. 12 µl of FuGENE then added to each tube followed by another 400 µl of DMEM. The mixture kept in hood for 40 min. Cell were transfected by dropping vector mix in spiral around plates and incubated at 37^0C for 48 h.

2. Protein extraction

Both cell lysate and culturing medium was collected at 48 h post-transfection. Briefly, media (5ml) was pipette into nalgene tube. 5 ml PBS 1X were added to each plate. 8 ml aceton were added into nalgene tubes and vortexed to mix and incubated at -80^{0}C for 1h. To extract protein from cells, previous PBS was aspirated and 1 ml of a mixture of PBS plus 0.1 PMSF were added to each plate and scraped into 1.5 ml chilled eppendorf. Tubes were centrifuged at 3000 rpm at 4^{0}C for 5 min to collect cells. Supernatant was aspirated and cell pallet was resuspended in 200 µl cold lysis buffer +pic (2µl) + 0.1M PMSF (1µl) per tube. Incubate at 4^{0}C for 1 h on shaker. Cells were sonicated for 10 sec and centrifuged at 13000xg at 4^{0}C for 5 min.

3. Western assay

Proteins from cell lysate and medium were separated by 15%SDS-PAGE, transferred to 0.45 mm nitrocellulose membranes (BioRad, CA, USA), and blotted by mouse anti-V5 antibody (1:10 000, Invitrogen, ON, Canada) and subsequently by HRP conjugated anti-mouse IgG (1:5000, Jackson Labs). They were also blotted by anti phospho SMAD 1/5/8. Western blot analyses were resolved by ECL reactions (efficient chemiluminescence, ThermoScientific, IL, USA). Alpha tubulin and secreted alkaline phosphatase were used as loading controls for cytosol GDF3 and secreted GDF3, respectively.

Morpholino Inhibition

Three types of morpholino were designed to zebrafish dvr1; a translation blocking, a splice blocking and a mismatch morpholino. Morpholino injections often increase p53 activation due to non-specific cell death. This is considered non-specific so in order to correct this, 6ng of each dvr1 morpholino was co-injected with 1ng p53 morpholino. Controls were fish injected with 1ng p53 MO. Fish were phenotyped at 5 days for ocular size and any other morphological abnormalities.

Table A.1 List of primers

Amplicons	**Forward primer (5'-3')**	**Forward primer (5'-3')**	**Annealing Temp (^{0}C)**
1	GCAATTCACACTTGATTATCTTACATC	TCAGTAATTGTCATTTCCTATTGTCC	57
2-1	GTATATGTGGGTGTCTGGCATTTC	CAGGTCCCGGAAGTTAATGAATAG	57
2-2	AGACACCTGTGCCAGACTAAGATG	TATGATTATTAGGGCTCCAGGATG	57

Table A.2 List of morpholino

Morpholino	Sequence
Translation Blocking	AGGCTCTGAGGAGGACTAAGAACAT
Splice Blocking	GCTCTGAGGAGGACCAAGAACATTA
Mismatch	UCUGGCTUAUTGAGUAUAGGAUAUTAAGUTUACAT

Results

Human data

Screening of 156 NTG and 96 HTG patients yielded two variants of GDF3. The first is a heterozygous change at G585A resulting in substitution of Glutamine (Q) to Arginine (R) at position 195 (Table 1). This change is present in the pre-prodomain and is excised from the mature protein. The second variant is a heterozygous change (C796T) that substitutes Arginine with cysteine (C) at amino acid position 266. This variant is in mature domain. These two mutations were partially conserved between the species (Figure 1).

480 normal controls have been screened by a former lab member for these variants. I also screened 130 more normal controls. Screening was done with *Hpa*II and *Tsp*RI restriction enzyme digestion for R195 and R266C mutation, respectively. These two mutations were absent from 610 normals and t-test showed that distribution of these mutations in POAG patients and their absence from normal controls was statistically significant (t-test, p= 0.02).

Phenotype of the morphants

Morphants showed several developmental abnormalities that were distinguishable after 3 days of post-fertilization, including reduced eye size, curled or kinked tails and pericardial edema (Figure 2). They also had problem swimming. However, not all of the embryos showed these phenotypes and a fraction of them looked normal. Furthermore, interesting to note was abnormal heart morphology. The heart of the control embryos underwent looping and the ventricle and

Table 1 Sequenced GDF3 Heterozygous variants found in NTG and HTG disease panel, frequency in patients and controls and their P value.

	Mutation	Prevalence in POAG	Prevalence in controls	P value
	R266C	1/252	0/610	
	R195Q	1/252	0/610	
Total	2	2/252	0/610	0.02

Figure 1. GDF3 variants in HTG/NTG panel. Upper panel: Chromatograms for the two identified variants in the HTG/NTG patients. Red box shows the location of nucleotide change. Lower panel: Location of mutations in GDF3's protein and their conservation between species (Taken from [11]).

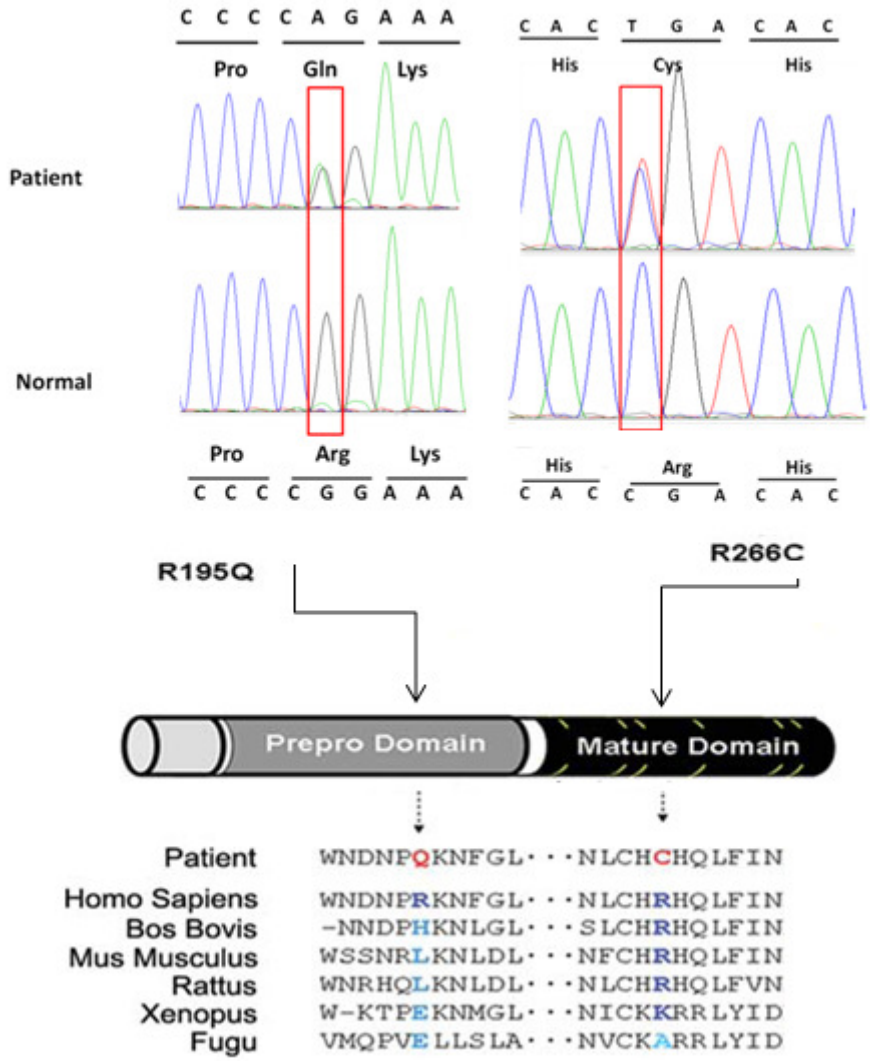
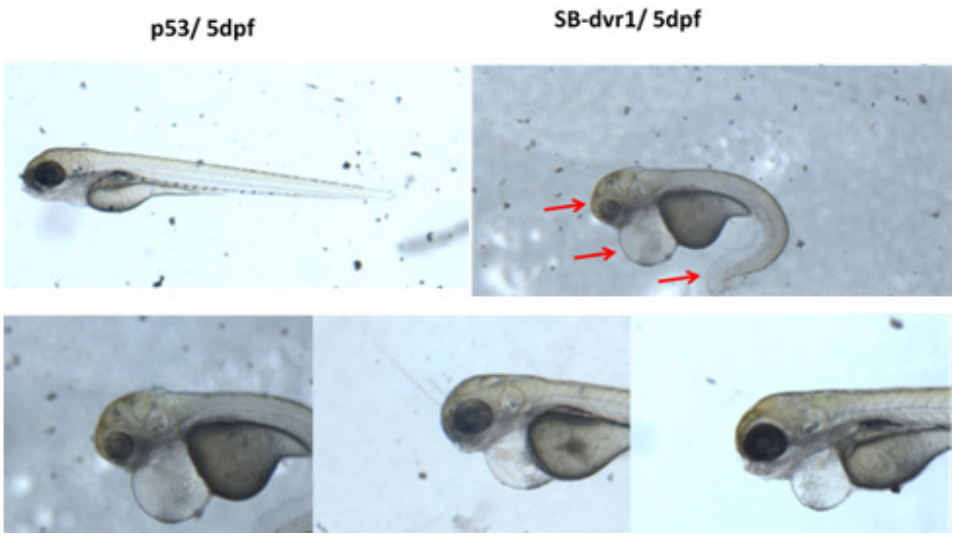


Figure 2. Reduced ocular size, skeletal defects and pericardiac edema are seen in morphants vs. controls. B. Morphants show a range of reduction in eye size vs. control



atrium were overlapped. However, the morphant showed looping failure, resulting in the linear heart. The tubular hearts in the morphants were able to contract but not as effectively as the control heart (Figure 2).

RGCs quantification

A transgenic fish in which the regulatory regions of the zebrafish isl2b gene (formerly isl3; Zebrafish Information Network) were used to drive expression of GFP in RGCs (*isl-2b: GFP).* In addition to RGCs it also labels small clusters of neurons in the forebrain and dorsal diencephalon. To be able to count the cells in the RGC layer, nuclei were stained with Hoechst. Pictures were taken from the whole eye using confocal microscope and GFP intensity was quantified by ImageJ analysis (Figure 3). For each eye zstack were taken and the middle section was analyzed. Next, the surface area of RGCs was selected and the intensity was measured using image J. To make sure that the effect of reduction in eye size is not counted in total intensity, the mean intensity values were calculated by dividing the intensity by surface area. Those MO-treated embryos that were severely affected showing no GFP signal or no retinal lamination as well as those who looked normal, were excluded from our analysis (Figure 4). Eventually, statistical analysis was conducted using student t-test on a sub group of test and control samples (n=8).

Western assay

Protein expression of Smad2/3, or Smad1/5/8 was analysed by Western blotting of cell extracts after 48 h. Since GDF3 is inhibitor of BMP signaling, a decrease in the level of Smad 1/5/8 were expected. However, no change can be detected in the level of either Smads between different GDF3 mutants expressing cells or the cells that expressed wild type GDF3.

Figure 3. Confocal images of whole flat mounted eyes and the effects of dvr1 morpholino inhibition on GFP signal. Eyes were fixed and processed at 5dpf. GFP Intensity was quantified using Image J around RGCs boundaries. Averages are indicated below each picture.

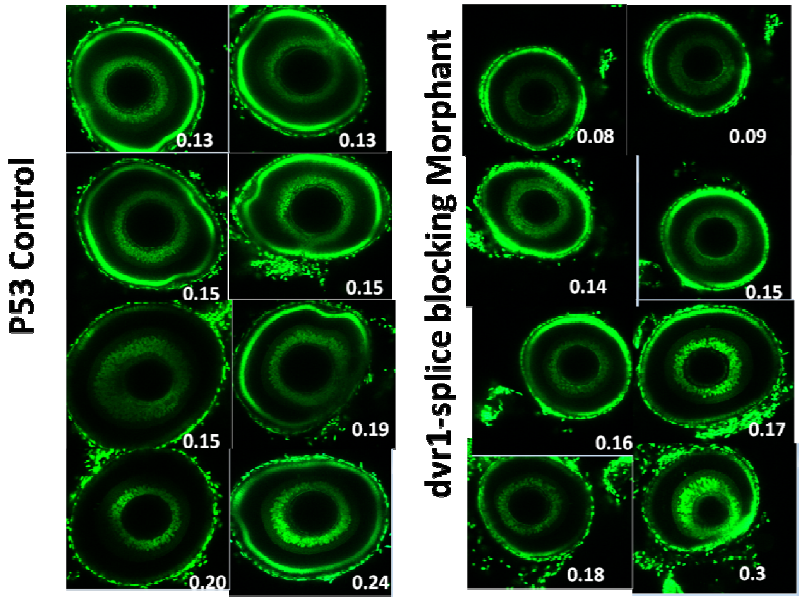


Figure 4. GFP intensity measurement in morphants and controls. Average intensity was calculated by dividing intensity value by the area. The graph shows the mean intensity for each individual eye.

P53 Control (1ng)

Area	Intensity	Intensity/Area
8741.87	1114.10	0.13
8369.80	1125.80	0.13
8610.12	1300.38	0.15
10234.79	1557.73	0.15
8774.71	1314.63	0.15
8402.64	1637.55	0.19
6684.1	1358.20	0.20
8630.98	2098.19	0.24

Splice blocking dvr1 (6ng) + 1ng p53

Area	Intensity	Intensity/Area
7879.51	629.61	0.08
7144.64	671.28	0.09
10113.85	1441.68	0.14
6151.69	903.59	0.15
5743.68	945.68	0.16
6251.37	1069.76	0.17
7493.53	1354.41	0.18
8342.37	2503.96	0.30

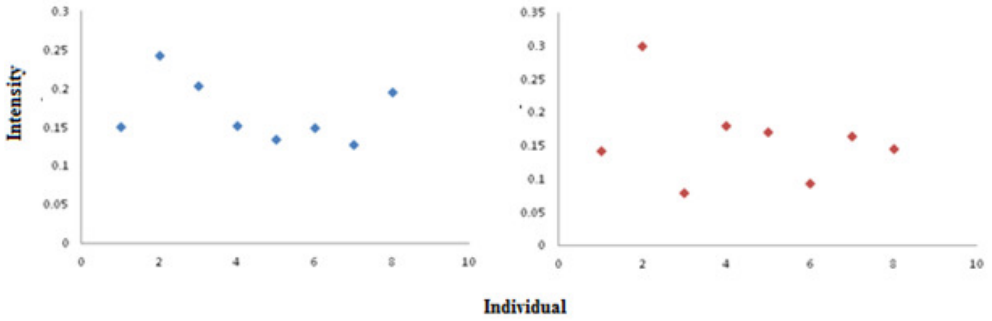
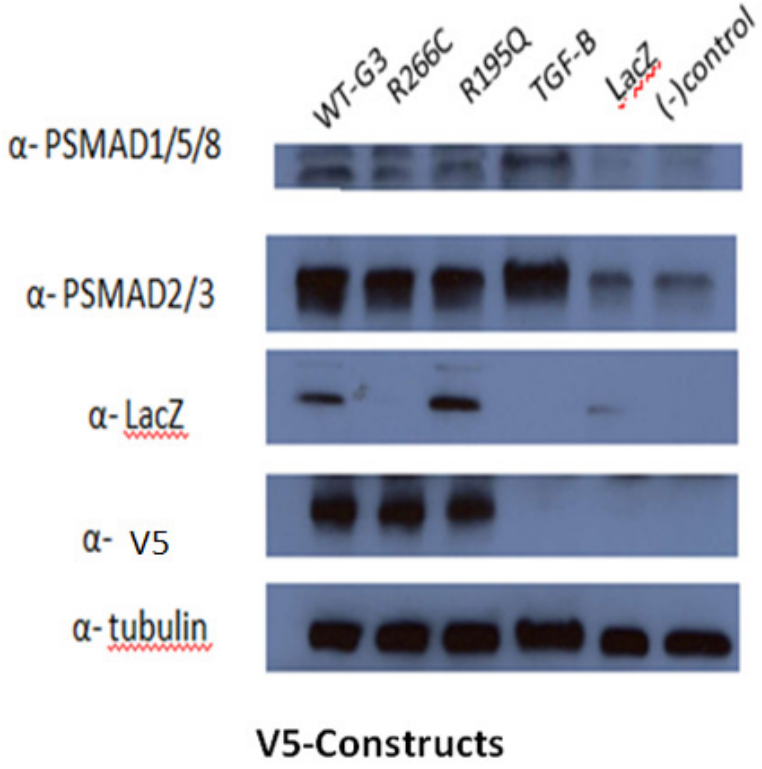


Figure 5. Immunoblot analysis of pSMAD 158 and pSMAD 2/3 in whole cell lysate extract from COS-7 cells transfected with different GDF3 constructs at 48 h post transfection. TGF-beta was used as positive control.



Discussion

Previous data and screening results showed that GDF3 is not only an important determinant in MAC, but also it may play a role in other retinal diseases (Figure 1). The fact that developmental genes have multiple functions and are used for multiple purpose, makes them fascinating candidates for study in human disease since they can be involved in early as well as late onset disease. However, the data from this chapter showed that it is difficult to make a genotype- phenotype correlation in the case of complex diseases, like glaucoma.

Different degrees of eye size reduction were observed in morpholino group (Figure2) which can be explained to some extent by variability in MO's injection volume. So, the dose of MO in some embryos might not be at the appropriate level to give the phenotype. In addition, morpholino injection into the yolk has the disadvantage that the efficiency of knockdown of the target gene is not 100% and to obtain a reliable result examining a larger numbers of embryos and careful interpretation of the data is necessary. Also, another possibility for incomplete penetrance of phenotypes might be the presence of other TGF-β ligands with redundant function to dvr1 that can compensate for the absence of this gene in knockdown fish. Regarding the heart defect, this observation is in line with previous in situ data in zebrafish that showed the highest expression of dvr1 in the tail and heart of the embryo 18 hpf [11]. So, it is not surprising that knocking down this gene produces abnormalities in the heart. Understanding which of the

two orthologs of dvr1—whether GDF3 or GDF1—caused this phenotype, merit further investigation.

Quantification of the GFP signal revealed no significant difference in splice blocking morphants versus p53 injected embryos at 5 dpf (Figure 3, 4). Although the size of the eye in morphant is considerably smaller than the controls' eye, it seems that this difference is not resulted by specific loss of cells in the RGC layer and it is a delay in the growth or maturation of the eye that affect all individual cell type in the retina. Here I present some explanation for the observed data. First, these data are representative of the knockdown experiment at 5 days post fertilization and the presence of change at earlier or later time points during development cannot be excluded. Moreover, there might be changes in either density or differentiation of other cell types of the retina even though the size and the shape of the retinal cells seemed unaffected. In addition, glaucoma is a late onset disease and using MO, which will dilute out after a while, would not be a perfect method to see the long term effect of this gene. And last but not least is the possibility of the presence of other TGF-β ligand with redundant function to dvr1 that compensate its lack.

For western blot experiment, again no conclusion could be made as there were no difference between mutant and wild type GDF3 in inducing the expression of Smads proteins. This can be due to endogenous GDF3 activity which confounded

the results or due to antibody nonspecific binding to both phosphorylated and nonphosphorylated Smads. In addition, the expected decrease in the level of GDF mutants were not observed which cast shadow on the validity of the test.

This study had several drawbacks. First of all, the animal model was not chosen correctly because there is no exact homologue of GDF3 in Zebrafish and the gene dvr1 corresponds to GDF1 and GDF3 in human with only 42% identity [12]. Hence, if there was a difference in retina after MO knockdown experiment, it would still be difficult to discern how much of the phenotype results from the knockdown of GDF3 and how much from knockdown of GDF1. In this context, using mammalian models including rats or mice that have GDF3 gene would result in more promising data. Secondly, it is known that Zebrafish have the ability to regenerate its RGCs after injury [13-14]. Therefore, it is possible that subtle change in RGCs would be healed at the time of our investigation. Thirdly, the transient knockdown experiment was conducted in this study [15]. However, to be able to see the long term effect of mutation it would be worthwhile to undertake a different approach.

Reference

1. Quigley HA. Number of people with glaucoma worldwide. Br J Ophthalmol 1996; 80(5):389-93.

2. Ray K, Mukhopadhyay A, Acharya M. Recent advances in molecular genetics of glaucoma. Mol Cell Biochem 2003; 253(1-2):223-31.

3. Stone EM, Fingert JH, Alward WL, Nguyen TD, Polansky JR, Sunden SL, Nishimura D, Clark AF, Nystuen A, Nichols BE, Mackey DA, Ritch R, Kalenak JW, Craven ER, Sheffield VC. Identification of a gene that causes primary open angle glaucoma. Science 1997; 275(5300):668-70.

4. Rezaie T, Child A, Hitchings R, Brice G, Miller L, Coca-Prados M, Heon E, Krupin T, Ritch R, Kreutzer D, Crick RP, Sarfarazi M. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science 2002; 295(5557):1077-9.

5. Monemi S, Spaeth G, DaSilva A, Popinchalk S, Ilitchev E, Liebmann J, Ritch R, Heon E, Crick RP, Child A, Sarfarazi M. Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. Hum Mol Genet 2005; 14(6):725-33.

6. Ducy P, Karsenty G. The family of bone morphogenetic proteins. Kidney Int 2000; 57(6):2207-14.

7. Wordinger RJ, Agarwal R, Talati M, Fuller J, Lambert W, Clark AF. Expression of bone morphogenetic proteins (BMP), BMP receptors, and BMP associated proteins in human trabecular meshwork and optic nerve head cells and tissues. Mol Vis 2002; 8:241-50.

8. Zode GS, Clark AF, Wordinger RJ. Bone morphogenetic protein 4 inhibits TGF-beta2 stimulation of extracellular matrix proteins in optic nerve head cells: role of gremlin in ECM modulation. Glia 2009; 57(7):755-66.

9. Jones CM, Simon-Chazottes D, Guenet JL, Hogan BL. Isolation of Vgr-2, a novel member of the transforming growth factor-beta-related gene family. Mol Endocrinol 1992; 6(11):1961-8.

10. Levine AJ, Brivanlou AH. GDF3, a BMP inhibitor, regulates cell fate in stem cells and early embryos. Development 2006; 133(2):209-16.

11. Ye M, Berry-Wynne KM, Asai-Coakwell M, Sundaresan P, Footz T, French CR, Abitbol M, Fleisch VC, Corbett N, Allison WT, Drummond G, Walter MA, Underhill TM, Waskiewicz AJ, Lehmann OJ. Mutation of the bone morphogenetic protein GDF3 causes ocular and skeletal anomalies. Hum Mol Genet 2010; 19(2):287-98.

12. Andersson O, Bertolino P, Ibanez CF. Distinct and cooperative roles of mammalian Vg1 homologs GDF1 and GDF3 during early embryonic development. Dev Biol 2007; 311(2):500-11.

13. Bernhardt RR, Tongiorgi E, Anzini P, Schachner M. Increased expression of specific recognition molecules by retinal ganglion cells and by optic pathway glia accompanies the successful regeneration of retinal axons in adult zebrafish. J Comp Neurol 1996; 376(2):253-64.

14. Sperry RW. Patterning of central synapses in regeneration of the optic nerve in teleosts. Physiol Zool 1948; 21(4):351-61.

15. Bedell VM, Westcot SE, Ekker SC. Lessons from morpholino-based screening in zebrafish. Brief Funct Genomics 2011; 10(4):181-8.