Is there still a role for genome-wide association studies? Investigating the genetic architecture of primary open-angle glaucoma and associated phenotypes

Cristina Venturini

Institute of Ophthalmology
Faculty of Brain Science
University College of London

A thesis submitted in partial fulfilment of the requirements for the degree of *Doctor of Philosophy* in Life Sciences at the University College London.

September 2015

Declaration

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

Cristina Venturini *Date*

Abstract

Despite the success of GWASs in the recent years, the genetic architecture of complex diseases, such as primary open-angle glaucoma is yet to be understood. The purpose of this thesis is to gain a better understanding of primary open-angle glaucoma using genetic association studies on several quantitative traits, such as intraocular pressure and optic nerve measurements. Collaborative efforts through extensive meta-analysis will help to increase the power of genetic association studies to identify novel variants for intraocular pressure and optic nerve parameters. In addition, gene-set enrichment analysis on the results of these meta-analyses may also help to shed a light on the mechanisms underlying glaucoma pathogenesis. Intraocular pressure and optic nerve cupping are not independent and they act together to determine the disease. Genetic association studies that combine information from different quantitative traits may help to identify novel genes affecting glaucoma susceptibility. In addition, epistatic interactions may also play a role in determining glaucoma endophenotypes variation and may help to identify novel loci affecting primary open-angle glaucoma susceptibility. Finally, primary open-angle glaucoma has been also associated with Alzheimer's disease, particularly the decrease of retinal ganglion cells. Elucidating the genetic underlying retinal nerve fibre layer and its relationship with Alzheimer's disease may help to discover and understand neurodegenerative mechanisms affecting primary open-angle glaucoma.

Acknowledgements

Firstly, I would like to thank my primary supervisor Mr. Viswanathan for the opportunity to do this project and for his insightful comments and encouragement through these years. I would like to express my sincere gratitude to my supervisor, Dr. Hysi Pirro for the continuous support of my PhD study, for his patience and motivation. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better mentor for my study.

My sincere thanks also goes to Prof. Hammond, who provided me the opportunity to join his team, for his time and support. I have really enjoyed to be part of the ”Eye Team” where enthusiasm for research has always been contagious and motivational for me.

I would also like to thank Prof. Rahi for her contributions of time and ideas and she has provided an excellent example as a successful woman and professor.

I gratefully acknowledge the funding sources (Allergan) and all the people working for the Grand Challenge UCL who made my Ph.D. work possible.

The members of the Department of Twins Research have contributed immensely to my studies. I thank Prof. Spector for the opportunity of joining the department and all the amazing people working there. I thank my fellow labmates, Katie Williams, for the support, lunches, chats and all the fun time we have had together, Abhishek Nag, for all the nice meals and for encouraging my sport side and Kate Yonova for the stimulating discussions and support. I am grateful to all my colleagues/friends in the Department of Twins Research. In particular, I would like to thank Leonie Roos, for the friendship and the support especially through the thesis writing. My thanks to all my friends including Pei-Chien Tsai and Idil Yet for all the nice time spent together and precious friendship,

and Craig Glastonbury for all the stimulating discussions over coffee breaks.

I would like to thank Pete Jones from the Institute of Ophthalmology (UCL) for making available his templates and his help in Latex.

A very special thanks goes to all my good friends, in primis Rosilari for her friendship, Skype chats, scientific discussions and amazing time together. Her support has been endless and I cannot ask for a better friend. A big thanks goes to Karin, for her friendship and all the fun we have had, as well as the support through these years. Thanks also to Sara, because even if we are far away, we keep supporting and encouraging each other. I would like to thanks all my wonderful friends in London for the amazing time: Mika, Rachele, Manuela, Andrea, Dagmara, Jon and Ben.

A very special thanks goes to my boyfriend and best friend, Sam, without whose love, encouragement and spiritual support, I would not have finished this thesis. I would like to thank him for always being there for me and providing, with Candy, a loving home to come back every day. He has helped me immensely listening to my science ideas, reading my drafts (sorry must have been so boring!) and trying to keep my stress level under control! Thank you!

Vorrei ringraziare la mia famiglia per tutto il sostegno in questi anni. Un grande grazie va ai miei genitori che mi hanno cresciuta con amore e fatto nascere la passione per la scienza, li ringrazio per avermi supportato e sopportato tutti questi anni quando i miei studi sembravano interminabili e poco comprensibili! Ringrazio le mie sorelle, Cecilia e Michela, per esserci sempre state e per aver creduto in me. Grazie a miei nipoti, Elisa, Marco e Roderik per avermi sempre fatto sorridere anche nei momenti peggiori!

Contents

List of Figures

List of Tables

Key of abbreviations

AD	Alzheimer's disease
BMES	Blue Mountains Eye Study
CCT	Central corneal thickness
CDR	Cup/disc ratio
GWAS	Genome-wide association study
GO	Gene Ontology
IGGC	International Glaucoma Genetics Consortium
IOP	Intraocular pressure
KEGG	Kyoto Encyclopedia of Genes and Genomes
LD	Linkage disequilibrium
NTG	Normal tension glaucoma
POAG	Primary open-angle glaucoma
RNFL	Retinal nerve fibre layer thickness
SE	Spherical equivalent
SNP	Single nucleotide polymorphism
VCD	Vertical cup diameter
VCDR	Vertical cup/disc ratio
VDD	Vertical disc diameter

CHAPTER 1

Introduction

Here I review the previous work in the genetic epidemiology of glaucoma and present the aim and objectives of my thesis.

1.1 Introduction to primary open-angle glaucoma (POAG)

Glaucoma is one of the leading causes of irreversible blindness worldwide (Quigley 1996). It is a heterogeneous group of degenerative disorders that have in common a degeneration of retinal ganglion cells, a characteristic excavation of the optic disc and irreversible visual field loss (Foster *et al.* 2002). There are many types of glaucoma, but the most common include primary open-angle glaucoma (POAG), chronic angle-closure (ACG), and exfoliation glaucoma (XFG). Primary open-angle glaucoma (POAG) is the predominant form of glaucoma (Weinreb & Khaw 2004) and is characterised by an open angle without a secondary explanatory cause (Leske 2007).

1.1.1 Anatomy and physiology

Intraocular pressure is regulated by a balance between the secretion and drainage of aqueous humour, which provides nutrients to the iris, lens and cornea. The fluid is secreted posterior to the iris by the ciliary body and then flows anteriorly to the anterior chamber.

Axons of retinal ganglion cells comprise the retinal nerve fibre layer, the innermost layer of the retina. These axons converge on the optic disc and form the optic nerve. The optic disc is about 1-5 mm in diameter and vertically oval. The convergence of the axons forms a central depression in the disc, known as the cup.

Glaucoma is characterised by the slow, progressive degeneration of retinal ganglion cells leading to an enlargement of the optic disc cup. The pathophysiology is not fully understood. The level of intraocular pressure may be related to the death of retinal ganglion cells and optic nerve fibres in some patients. Although no obstruction can be seen with clinical examination, resistance to aequous outflow through the trabecular meshwork is increased in patients with this form of glaucoma. When pressure increases the pressure gradient across the lamina cibrosa also increases. As a result, the lamina cibrosa and the retinal ganglion cell axons undergo deformation and mechanical stress (Weinreb & Khaw 2004).

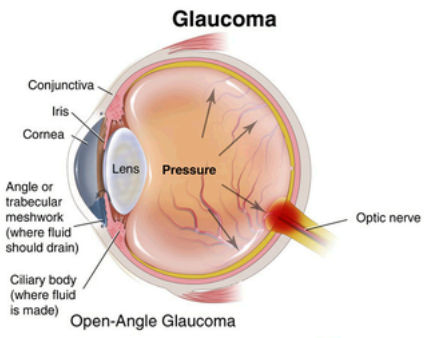


Figure 1.1: The figure shows a summary of the pathophysiology of POAG

1.1.2 Epidemiology of glaucoma

Studies report prevalence estimates of around 1-3% in Europe, 1-4% in Asia and 2-3% in Australia (Rudnicka *et al.* 2006; Quigley & Broman 2006). Prevalence is much higher in Africa, with prevalence estimates varying from 1% in Nigeria (Murdoch *et al.* 2001) to over 8% in Ghana (Ntim-Amponsah *et al.* 2004). High prevalence of 7-9% were found in African-Caribbean populations, originating from West Africa where prevalence is also very high (Mason *et al.* 1989) (see Figure 1.2). As a consequence of this variability, the public health impact of glaucoma blindness also differs across ancestral groups. Glaucoma is the second most common cause of blindness in the world (responsible for 12% of the global blindness (Resnikoff *et al.* 2004)), however it is the commonest cause in African-descent populations, with one-third of blindness due to POAG (Hyman *et al.* 2001; Congdon *et al.* 2004), probably also due to health care deficiencies. The incidence of POAG is about 0.1 to 0.2% per year in populations of European origin (Leske 2007). A higher incidence of 0.5% per year was found in black participants of the Barbados Eye Studies (Leske *et al.* 2007).

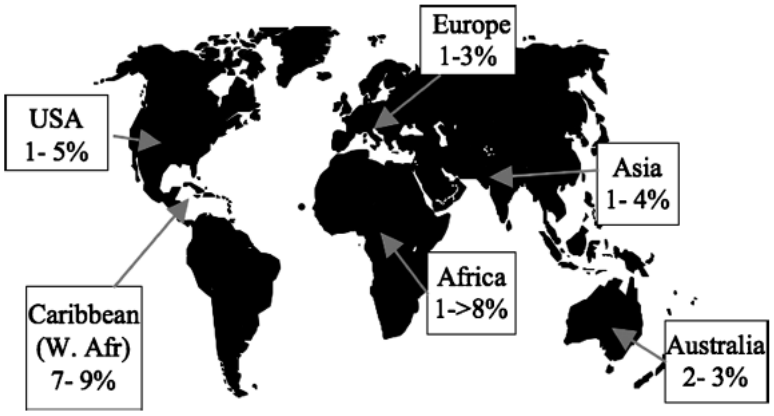


Figure 1.2: Primary open-angle glaucoma global prevalence estimates: this figure shows the POAG global prevalence estimates for adult older than 40 years old. This figure has been published elsewhere (Leske 2007).

Risk factors for POAG include demographic, ocular and systemic factors. Age is a well known risk factor for POAG. The disease rarely occurs before 40 years old and the risk doubles after age 60 as compared with the 40 to 49 years old age group (Leske *et al.* 2008). In a meta-analysis, prevalence rose steadily with age, especially in European-derived populations: for each decade increase in age, the odds ratio was 2.05 (CI: 1.91-2.18, 95%) in European population, 1.61 (1.53 to 1.70) in African and 1.57 (1.46 to 1.68) in Asian populations (Rudnicka *et al.* 2006).

An elevated intraocular pressure (IOP) is a strong and established risk factor for POAG. In the Barbados nine-year study, results showed that elevated IOP is a major contributor to POAG development, with IOP levels $\geq$ 21 mmHg increasing the 9-year relative risk to at least 5 as shown in Figure 1.3 (Nemesure *et al.* 2007). IOP level remains the only major POAG risk factor that is known to be modifiable. For that reason, there has been long-standing interest in IOP-lowering treatment to decrease risk and progression of POAG (Czudowska *et al.* 2010; Kass *et al.* 2002). Several clinical trials have demonstrated the effectiveness of lowering IOP to reduce visual field loss and a meta-analysis treatment is associated with approximately a 40% decreased risk of glaucoma progression (Maier *et al.* 2005). However, IOP has limitations as predictor of POAG. An estimated 20-50% of all patients with all the symptoms of the disease have IOP within the normal range (normal tension glaucoma), whereas many individuals with elevated IOP do not have any sign of POAG (this condition is referred to as ocular hypertension) (Leske 2007).

Another ocular risk factor has been identified in thinner central corneal thickness (CCT) (Leske *et al.* 2008). The mechanisms for an association

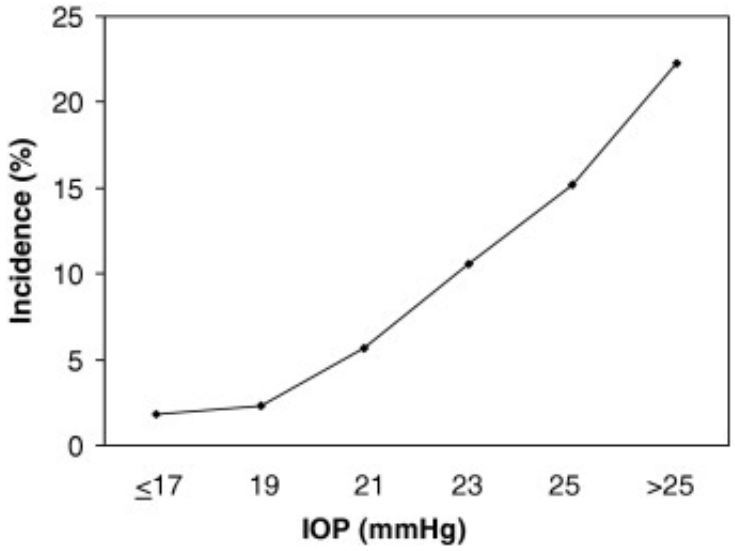


Figure 1.3: Incidence of POAG and IOP: this figure graphically represents the relationship between incidence of POAG (y axis) and baseline IOP (x axys) in the Barbados Eye Studies. Overall, when IOP was evaluated as a continuous variable, the risk of OAG increased by 12% with each 1-mmHg increase in IOP. This figure has been published elsewhere (Nemesure *et al.* 2007).

between CCT and POAG have yet to be determined. The association was initially reported in the Ocular Hypertension Treatment Study, where thin CCT was a strong factor predicting the development of POAG in individuals with ocular hypertension (Gordon *et al.* 2002). While CCT has a well-known effect on IOP measurement (a thin central cornea may lead to an underestimation of IOP), the relationship between CCT and POAG remains unclear.

Associations with spherical error (SE) have been reported in several studies that report myopia is likely to increase POAG risk (Leske 2007). For example, myopic individuals in the Blue Mountains Eye study had a twofold to threefold increased risk of glaucoma compared with those non myopic (Mitchell *et al.* 1999). In particular, high myopia has been correlated with glaucomatous optic nerve damage, for instance in the Beijing Study high myopic eyes, defined as -6D had a higher correlation

with glaucomatous damage (Xu *et al.* 2007).

Large optic discs are mentioned as another risk factor, however larger discs are associated with larger optic cups, which could be identified as an early sign of disease(Leske 2007).

Several systemic and metabolic conditions are known to have profound effects on the structure of the eye. Vascular dysregulation has been suspected as a risk factor for glaucoma (Flammer *et al.* 2002). A vascular role is suggested by the reported associations with vasospasm, migraine headaches, and with abnormalities in ocular blood flow and circulatory auto-regulation (Grieshaber & Flammer 2005). Blood pressure has been associated with IOP, however neither systolic nor diastolic blood pressure has been related to POAG (Tielsch *et al.* 1995a; Bonomi *et al.* 2000). Diabetes has also a well known association with higher IOP (Tielsch *et al.* 1995b), but the relationship with POAG remains unclear (Leske 2007).

Neurological conditions have been associated with POAG, for instance higher prevalence of glaucoma in Alzheimer's patients in comparison to the general population has been reported (Bayer *et al.* 2002). It has been hypothesised that POAG and Alzheimer's disease may share common pathophysiological mechanisms or risk factors i.e. neuro-inflammation (Tsilis *et al.* 2014). In addition, the brain has the same embryonic origin than the retina and interestingly, a thicker retinal nerve fibre layer was associated with better cognitive performance in healthy young individuals (Koolwijk *et al.* 2009).

1.2 Review of previous work on genetics of POAG

Genetics may also represent a risk factor for developing and progressing POAG. Research in the last decade has shown that genes have a key role in the pathogenesis of POAG. The identification of disease causing genes for POAG is important for at least two main reasons. First, the identification of genes and their biological pathways may elucidate the pathophysiological mechanisms that are presently little understood. This may then provide new directions for the development of POAG therapy. Second, knowing the genes altering risk of onset or progression of POAG may enable to create diagnostic and prognostic DNA based tests. This may be especially valuable for POAG, as many patients are diagnosed only after significant and irreversible visual field damage has occurred. Early diagnosis and treatment may prevent or delay this damage. In addition, many suspects who repeatedly attend the glaucoma clinic unnecessarily might be spared the burden of regular surveillance based on their DNA test. In the last two decades important progress has been made in unravelling the genetics of POAG through advanced in genetic epidemiology. Genetic epidemiology investigates how genes influence disease risk in human populations. It differs from classical epidemiology by its explicit consideration of genetic factors and it differs from medical genetics by its emphasis on population-based studies.

1.2.1 Genetic susceptibility of POAG

In broad terms, common eye diseases, including POAG, are caused by environmental exposures in genetically predisposed individuals. Complex diseases are caused by a combination of genetic, environmental

and lifestyle factors, most of which have not yet been identified. Epidemiological studies have shown a generally high prevalence in African individuals, together with an earlier age of presentation (Rudnicka *et al.* 2006; Mitchell *et al.* 1996; Leske *et al.* 1994; Foster *et al.* 2000). Ancestral differences, even though they might reflect different environmental exposure (i.e. lack of health system in African countries), suggest that POAG onset and progression may be affected by genetic factors.

In addition, family history is another well-known risk factor for POAG, with family members clearly having increased risk across populations (Wolfs *et al.* 1998). Up to 50% glaucoma patients have a positive family history (Tielsch JM *et al.* 1994). More over, first degree relatives of an affected individual have a 22% risk of developing POAG compared with 2.3% risk in family members of controls (Wolfs *et al.* 1998). The overall risk of developing POAG amongst first degree relatives of an affected individual is increased 3-9 fold (Wolfs *et al.* 1998; Tielsch JM *et al.* 1994).

Despite the evidence that family history plays a role in POAG onset and progression, only one study (in 1987) determined the heritability for POAG (Teikari 1987). Heritability is defined as the proportion of the overall variation in a population that is ascribable to genetic variation amongst individuals (Sanfilippo *et al.* 2010). Generally, when estimating heritability, genetic effects are assumed to be additive and this is called ”narrow sense” heritability. Additive genetic effects occur when alleles at a locus on a chromosome add up to constitute a phenotype (Sanfilippo *et al.* 2010). Heritability can be calculated in different ways. The classical twin model is a well-known way to calculate heritability. In this approach, we determine whether identical twins (monozygotic, MZ) raised together are

more similar to each other for a given trait than non identical (dizygotic twins, DZ) raised together. The rationale is that if a genetic basis exists we would expect MZ twins to be more similar as they share 100% of their DNA, compared with DZ twins who share only 50% on average. Family studies could also include analysis of phenotypic similarity to familial relationship of greater complexity.

Other methods to calculate heritability are based on regression and correlation of intrapair trait values. For instance, with parent-child data we could regress the trait value of the child on the average value of the parents and the regression coefficient (or the slope) provides an estimate of the narrow-sense heritability (Sanfilippo *et al.* 2010). Alternatively, in sibling and twin study designs it is possible to calculate the intrapair correlations and twice of the residual correlation provides an estimate of the narrow-sense heritability (Sanfilippo *et al.* 2010).
POAG heritability, calculated using twin concordance ratio (correlation method), was estimated to be 13% (Teikari 1987), however, this is likely to be an underestimation, as a large proportion of glaucoma goes undiagnosed and the concordances were calculated from national registries rather than individual screening and ascertainment.

Studying familial aggregation and heritability in POAG has some limitations. Ascertainment of age-related condition, such as POAG, is complicate in families as parents are usually not alive when an individual is diagnosed. Younger siblings may not have the disease, however cannot be considered unaffected as they may develop POAG later in life. Recruiting a large number of cases is expensive and it can be used only for that

particular disease. Moreover categorising an individual as affected or non-affected is arbitrary and therefore susceptible to misclassification. An important strategy for gene discovery in POAG is to dissect the disease into its constituent parts or ”endophenotypes”. An endophenotype has been described as a quantitative biological trait that is associated with the disease (it should be associated with the causes rather than effects of disorders), it is heritable and it should vary continuously in the general population (Cannon & Keller 2006). Within families, endophenotype and disease co-segregate (Gottesman & Gould 2003). Rather than studying POAG, a better approach is to study POAG endophenotypes that are cup-to-disc ratio (CDR) or vertical cup-to-disc ratio (VCDR), rim area, retinal nerve fibre layer (RNFL) thickness, intraocular pressure (IOP) and central corneal thickness (CCT). These intermediate traits are thought to have a simpler genetic architecture, can be measured on a continuous or quantitative scale, and may be more sensitive measures of important aspects of the disease process (Sanfilippo *et al.* 2010). Moreover, quantitative traits can also be studied in individuals without POAG, which means that this approach greatly increases the number of individuals available for genetic studies. Quantitative traits studies are also cost-efficient since you can have many phenotypes for each individual.

Much effort has gone into investigating IOP variation as it is considered the most important risk factor for POAG. Heritability estimates range from 0.29 to 0.67 for IOP (see Figure 1.4) with a pooled estimate of 0.55 (Sanfilippo *et al.* 2010). Thinner central corneal measurements may increase the risk for POAG, so much attention has been dedicated to understand CCT variation. CCT showed a heritability estimate of 0.95, which suggests that 95% of its variance is explained by the effects of genes

(Hougaard *et al.* 2003; Toh *et al.* 2005). Heritability estimates for CCT are comparable to other highly heritable human traits such as fingerprint (Sanfilippo *et al.* 2010). However, CCT may not represent a true POAG endophenotype as poor genetic correlation has been found between CCT and POAG, compared with IOP (Charlesworth *et al.* 2010). Evaluation of structural integrity of the optic nerve is important for the assessment of POAG. Heritability estimates for CDR range from 0.48 to 0.79 and from 0.48 to 0.82 for RNFL thickness (Hougaard *et al.* 2003; Koolwijk *et al.* 2007; Chang *et al.* 2005; Haines *et al.* 2005; Levene *et al.* 1970)(see Figure 1.5).

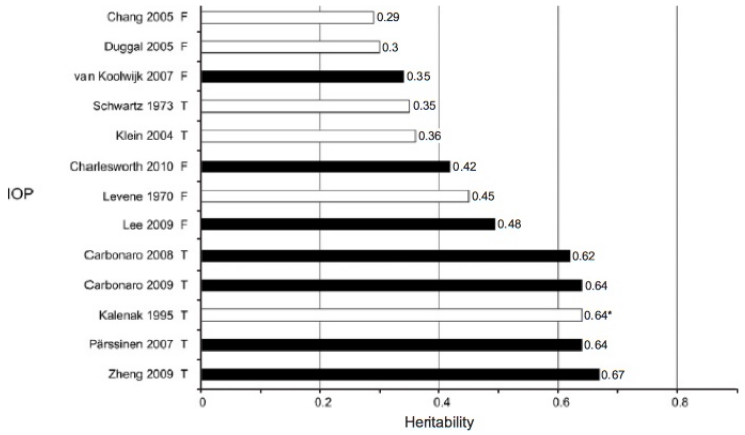


Figure 1.4: Heritability estimates for IOP: this graph shows the heritability estimates for IOP in different studies. Colours in the bars represent different methods: black bars represent heritability determined by structural equation modelling (i.e.classical twins studies) and white bars by regression/correlation methods. This graph has been published elsewhere (Sanfilippo *et al.* 2010).

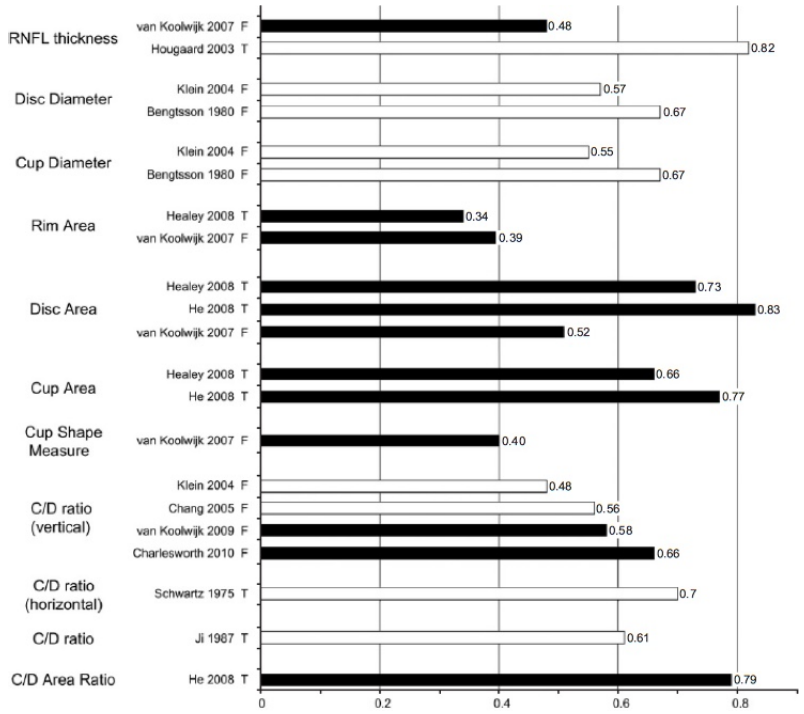


Figure 1.5: Heritability estimated for optic nerve head parameters: this graph shows heritability estimates for optic disc parameters in different studies. Colours in the bars represent different methods: black bars represent heritability determined by structural equation modelling (i.e. classical twins studies) and white bars by regression/correlation methods. This graph has been published elsewhere (Sanfilippo *et al.* 2010).

1.2.2 Linkage analysis

Identifying genes that affect common disorders, such as POAG, is challenging. The most commonly used method to identify genes that are transmitted in a Mendelian pattern is genetic linkage analysis. This method uses families where multiple members are affected. The power to detect linkage is proportional to the number of individuals affected within families and the total number of families. Once linkage to a specific chromosomal locus has been identified, analysis of the genes within that region is conducted. The rationale is that two loci lying close together have a higher probability of being inherited together. The further apart two loci are, the higher the chance of a recombination event occurring between them during meiosis. Therefore, the probability of two loci segregating

together is a measure of the genetic distance between them. Thus, the probability of a genetic marker and the disease segregating together is a measure of the distance between the marker and the disease gene.

Before the advent of current large-scale genotyping platforms, family-based linkage studies have traditionally played an important role to locate genes altering disease of phenotypic trait susceptibility. Linkage analysis provide a tool to identify the chromosomal location (locus) of a disease without prior knowledge of the biological mechanisms.

Traditionally linkage analysis has been performed to study monogenic diseases in large families with multiple affected members. This approach is called parametric or model-based linkage analysis. The aim is to discover how often two loci are separated by meiotic recombination. The rationale is that two loci lying close together have a higher probability of being inherited together. The further apart two loci are, the higher the chance of a recombination event occurring between them during meiosis. Therefore, the probability of two loci segregating together is a measure of the genetic distance between them. Thus, the probability of a genetic marker and the disease segregating together is a measure of the distance between the marker and the disease gene.

This approach requires an assumption of the genetic model, in which the mode of inheritance, the disease and marker allele numbers and frequencies, and the penetrance of the disease genotype need to be specified. As long as an adequate model can be assumed, parametric linkage provides a powerful method to locate a disease gene. However, it has failed to find the more common genes underlying complex diseases, for which a valid gene model cannot be specified. The shift towards the genetics of complex diseases has therefore led to the development of new

methods of linkage analysis that are nonparametric or model-free. Both parametric and nonparametric linkage approaches have been used to identify the chromosomal locations of POAG susceptibility genes.

Fourteen chromosomal loci for POAG (GLC1A-N) are listed in HUGO Genome Nomenclature Committee. Three genes associated with glaucoma have been identified within these loci, including myocilin on chromosome 1 (Sheffield *et al.* 1993), optineurin (Rezaie *et al.* 2002) on chromosome 10 and *WDR36* (Monemi *et al.* 2005) on chromosomes 5 (see Figure 1.6).

Myocilin (*MYOC*), on chromosome 1, was the first gene identified for POAG in the GLC1A locus (Sheffield *et al.* 1993). It is transmitted as an autosomal dominant Mendelian trait and mutations in the gene have been found in patients with juvenile or early-adult forms of POAG. These patients have a severe clinical phenotype including elevated IOP (Fingert *et al.* 1999). How mutations in *MYOC* gene lead to POAG is still unclear. It has been proposed that these mutations may alter protein function and promote an increase of IOP and loss of retinal ganglion cells. However, studies using mouse model with the *Tyr437His* mutation have produced conflicting results: one study reported a normal phenotype (Gould *et al.* 2006), whereas other studies reported an increase of the IOP and loss of retinal ganglion cells (Senatorov *et al.* 2006; Zhou *et al.* 2008). MYOC protein is normally released into the aqueous humour from inside of the cells and release of MYOC protein into the extracellular space in the presence of glaucoma mutations is significantly reduced (Jacobson *et al.* 2001).

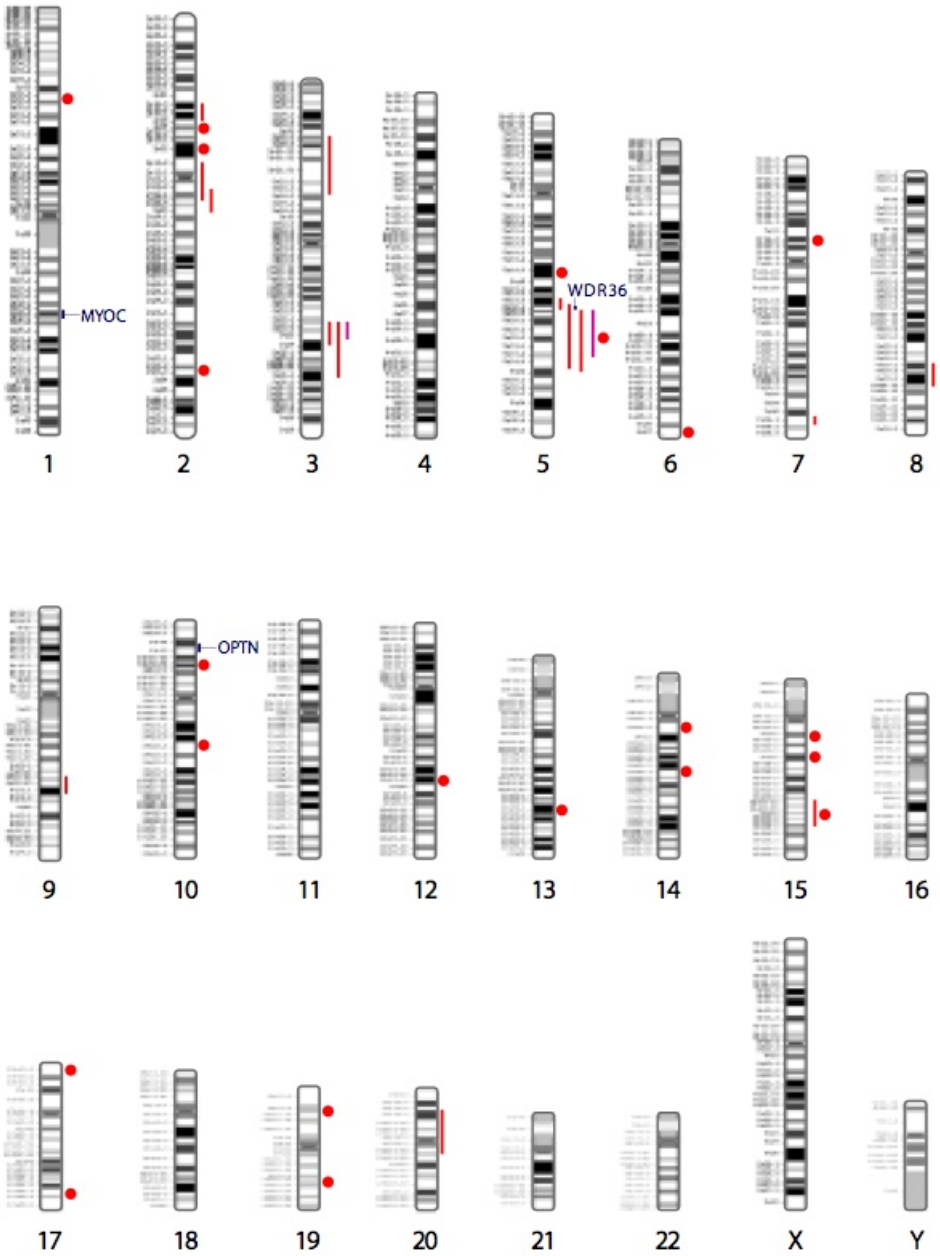


Figure 1.6: Chromosomal regions linked to POAG: this figure has been published elsewhere (Venturini *et al.* 2014a). The red lines represent linkage regions that have been identified by analyses in single, large pedigrees. Line boundaries have been determined by haplotype analyses. Red dots represent maximum LOD scores from population based (family) studies.

OPTN, on chromosome 10, was the second gene identified for POAG through linkage analysis (Rezaie *et al.* 2002). *OPTN* associated POAG is characterised by normal IOP as seen in patients with normal tension glaucoma (Rezaie *et al.* 2002). The main mutation in *OPTN* gene linked with POAG is E50 K (Allingham *et al.* 2009). Patients carrying this particular mutation have more severe POAG phenotype, such as earlier onset and more advanced optic nerve cupping. The molecular mechanisms related to *OPTN* mutations are still unclear, however it has been hypothesised that the OPTN protein may exert an effect on retinal ganglion cells by increasing their susceptibility to premature cell death (apoptosis) (Allingham *et al.* 2009).

WDR36 is another gene identified for POAG though linkage analysis (Monemi *et al.* 2005). Hauser *et al.* have suggested that *WDR36* mutations may influence susceptibility to POAG rather than cause the disease (Hauser *et al.* 2006). *WDR36* mutations may directly affect axon growth of the retinal ganglion cells and lead to progressive retinal degeneration with normal IOP in transgenic mice (Chi *et al.* 2010).

Mutations in *NTF4* gene were reported in 1.7% of POAG patients in the European population through linkage analysis (Pasutto *et al.* 2009). It was hypothesised that these mutations may affect either NTF4 dimer stability or the interaction between NTF4 dimer and its receptor TrkB (Pasutto *et al.* 2009). However, the role of *NTF4* gene in POAG remains unclear as other studies have failed to replicate this result (Liu & Allingham 2011).

1.2.3 Association studies

In the last decade, population-based association studies have become increasingly popular in different medical fields, including eye research. In principle linkage and association are two different phenomena. As previously described, linkage is a specific genetic relationship between loci, on the other hand association is simply a statistical statement about the co-occurrence of alleles or phenotypes. Allele A is associated with the disease D if the individuals with D also have A more often (or less often) than would be expected from the individual frequencies of D and A in the population. Linkage creates association within families, but it does not itself produce any association in the general population where individuals are unrelated. For example, if a locus X is linked to a disease locus, within a family where the disease mutations is segregating, we would expect affected people to have the same allele A for the X locus, but over the whole population the distribution of X alleles is just the same in people with the disease or not. However, if two ”unrelated” individuals with the disease D have actually inherited it from a distant common ancestor, they may well also tend to share particular ancestral alleles at loci closely linked to D.

Association studies in comparison with linkage design are generally more powerful for complex disease influenced by multiple genes and they allow a more precise localisation of the region of interest. Originally, association studies were conceived around the ”common disease-common variants” (CDCV) hypothesis stating that if a heritable disease/trait is common in the population (greater than 1-5%), then the causative genetic factors are also likely to be common (Lander 1996). As a ramification, the total genetic risk must be spread across multiple common genetic factors

with small effect.

Initially, association studies were performed for candidate regions that were selected for their biological interest and understanding of the disease. Here I review the literature for candidate and in particular genome-wide association analysis performed for POAG and endophenotypes.

1.2.3.1 Candidate gene analyses

Many candidate genetic association studies for POAG have been performed. Candidate genes have been selected through different strategies such as genes involved in mechanisms that are assumed to be involved in POAG pathogenesis, genes identified with linkage studies and genes involved in the pathogenesis of related diseases. Many of these studies have had inconsistent results and the role of these genes in the aetiology of POAG is still controversial.

In the first category, association studies of genes regulating ocular blood flow (nitric oxide synthase and endothelin-1 related genes) (Ishikawa *et al.* 2005; Logan *et al.* 2005; Motallebipour *et al.* 2005), aqueous humour outflow (renin-angiotensin system genes) (Bunce *et al.* 2005; Hashizume *et al.* 2005), apoptosis (tumour-protein P53 gene) (Dimasi *et al.* 2005; Lin *et al.* 2002), immune system (interleukin 1β and tumour necrosis factor α genes) (Lin *et al.* 2003, 2003), and neurodegeneration (apolipoprotein E gene) (Copin *et al.* 2002; Lam *et al.* 2006; Mabuchi *et al.* 2005; Vickers *et al.* 2002) were performed. Given the identification of *MYOC* as a disease-causing gene, Mukhopadhyay *et al.* used a bio-informatics approach to search for myocilin-related proteins that had a conserved olfactomedin domain and were expressed in the eye (Mukhopadhyay

et al. 2004). They thus identified the Noelin 1 and 2 genes as potential candidates for POAG. One association study has since been performed for the Noelin 2 gene (*OLFM2*) in Japanese subjects (Funayama *et al.* 2006). A possible disease causing mutation was identified and common genetic variants were suggested to contribute to the glaucoma phenotype by interacting with the optineurin gene.

For the last category, genetic association studies were conducted for *OPA1*, the gene responsible for autosomal dominant optic atrophy (Delettre *et al.* 2000; Alexander *et al.* 2000). Like glaucoma, optic atrophy is a progressive optic neuropathy caused by degeneration of retinal ganglion cells. Due to the absence of raised IOP, *OPA1* was hypothesised to be most likely associated with normal tension glaucoma. Genetic variants in the *OPA1* gene have indeed been associated with normal tension glaucoma, but not with high tension glaucoma in Caucasian patients (Aung *et al.* 2002; Powell *et al.* 2003). A study in Japanese subjects confirmed this, but showed that within the group of high tension glaucoma patients an *OPA1* variant was significantly related to the age at the time of diagnosis (Mabuchi *et al.* 2005). The association between *OPA1* variants and normal tension glaucoma could not be replicated in Korean and African American subjects (Woo *et al.* 2004; Yao *et al.* 2006).

1.2.3.2 Genome-wide association analyses

In a genome-wide association study (GWAS), a dense set of markers along the whole genome is assessed for association with a disease or quantitative trait.

In 2007, the first GWAS of glaucoma yielded two common exonic variants in *LOXL1* that explained a significant proportion of the cases of exfoliation glaucoma in a European population (Thorleifsson *et al.* 2007). Exfoliation glaucoma is a form of secondary glaucoma in which aqueous humour outflow is obstructed by fibrillar extracellular (exfoliative) material. The data suggested that *LOXL1* was particularly associated with the accumulation of microfibrillar deposits in the anterior segment of the eye (exfoliation syndrome) rather than with the consequent onset of secondary glaucoma. Subsequent studies in populations of Caucasian, African and Asian ancestry confirmed the association of *LOXL1* with exfoliation glaucoma and did not support any association with other subtypes of glaucoma, such as primary open-angle glaucoma (Williams *et al.* 2010; Chen *et al.* 2009; Challa *et al.* 2008).

A study in 1,263 affected cases and 34,877 controls identified a common variant near *CAV1* and *CAV2* genes (Thorleifsson *et al.* 2010) associated with POAG. Both genes are expressed in the trabecular meshwork as well as in retinal ganglion cells. A recent study in a US Caucasian population confirmed the association of the *CAV1-CAV2* gene region with POAG and further suggested that this association predominated in women (Wiggs *et al.* 2011).

Burdon *et al.* conducted a GWAS in a cohort of patients with severe glaucomatous visual field loss and identified susceptibility loci for POAG at *TMCO1* and *CDKN2B-AS1* genes (Burdon *et al.* 2011). They replicated their findings in an independent cohort of patients with advanced POAG, and also in two cohorts of patients with less severe disease, suggesting that the approach of selecting extreme phenotypes may be effective in identifying

susceptibility genes for complex diseases.

In addition to these case-control studies, population-based GWASs have identified chromosomal regions for POAG-related quantitative traits (Table 1.1). Two GWASs in Caucasian populations independently identified a locus for optic disc size near the *ATOH7* gene at chromosome 10 (Macgregor *et al.* 2010; Ramdas *et al.* 2010). Although *ATOH7* had already been implicated in retinal ganglion cell formation in animal studies, the study by Macgregor *et al.* was the first to associate this gene with optic nerve pathology in humans (Macgregor *et al.* 2010). A subsequent GWAS in 4,445 Singaporean individuals of Indian and Malay ancestry confirmed the association of *ATOH7* with optic disc area, suggesting that there are shared genetic pathways which determine optic disc size across different ethnicities (Khor *et al.* 2011).

Another chromosomal region of interest is the *CDKN2Bas* locus which was associated with VCDR by Ramdas *et al.* (Ramdas *et al.* 2010). Both genes are involved in the cell cycle regulation. This locus was replicated by a GWAS in patients with severe glaucomatous visual field loss, indicating that the same gene may influence both the normal variance in VCDR and the risk of developing severe glaucoma (Burdon *et al.* 2011). These findings support the hypothesis that studies of quantitative traits, such as VCDR, in healthy individuals may assist in identifying susceptibility genes for glaucoma. To assess whether the genetic variants identified for optic disc parameters were associated with glaucoma, a meta-analysis of 6 case-control studies (total N = 3,161 glaucoma cases and 42,837 controls) was performed (Ramdas *et al.* 2011). Of the 8 variants evaluated, statistically significant associations with glaucoma were found for

CDKN2Bas (identified for VCDR), *SIX1/SIX6* (identified for VCDR), and *ATOH7* (identified for both optic disc area and VCDR). A subsequent study in a US Caucasian sample of 539 glaucoma cases and 336 controls confirmed these associations for *CDKN2Bas* and *SIX1/6* (Fan *et al.* 2011).

For IOP, two loci were initially identified by a GWAS (van Koolwijk *et al.* 2012). The first locus overlapped with the *TMCO1* locus, also identified in cases with severe glaucomatous visual field loss. This finding again indicated that the same genetic factors may be responsible for both IOP variation and POAG, supporting the use of quantitative traits to identify genes for POAG. The second locus was *GAS7*, a gene previously implicated in cell remodelling. *TMCO1* and *GAS7* genes were associated with glaucoma in a meta-analysis of 4 case-control studies in populations of European ancestry. Both genes were found to be highly expressed in the ocular tissues involved in glaucoma and to functionally interact with known glaucoma disease genes.

A recent meta-analysis of GWAS of CCT involving more than 20,000 individuals of European and Asian ancestry, identified 16 new loci (Cirulli & Goldstein 2010). The genes involved are *COL5A1, AVGR8, FOX01, AKAP13, ZNF469, FNDC3B* and *LRRK1-CHSY1* for the European ancestry populations and *COL5A1, AKAP13* and *ZNF469* for the Asian ancestry groups. The loci associated with CCT in European and Asian populations together explained 8.3% of the variance in Europeans and 7% in Asians. The relevance of CCT-associated loci for glaucoma was evaluated, using independent case-control sets, showing that *FNDC3B* locus was associated

also with POAG.

Gene region	Location	Associated trait	Reference
CDC7 / TGFBR3	1p22.1	Optic disc size	Ramdas WD, PLoS Genet 2010;6:e1000978
		Optic disc size	Khor CC, Hum Mol Genet 2011;20:1864-1872
TMCO1	1q24.1	Advanced POAG	Burdon KP, Nat Genet 2011;43:574-578
		IOP	Van Koolwijk LME, PLoS Genet 2012
ZP4	1q43	POAG	Nakano M, Proc Natl Acad Sci USA 2009;106:12838-12842
SRBD1	2p21	NTG	Meguro A, Ophthalmology 2010;117:1331-1338
FNDC3B	3q26.31	central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
		POAG	Lu Y, Nat Genet 2013; 45:155-163
ELOVL5	6p21.1-p12.1	NTG	Meguro A, Ophthalmology 2010;117:1331-1338
CAV1 / CAV2	7q31	POAG	Thorleifsson G, Nat Genet 2010;42:906-909
CDKN2A / CDKN2B / CDKN2B-AS1	9p21.3	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
		Advanced POAG	Burdon KP, Nat Genet 2011;43:574-578
COL5A1	9q34.2	central corneal thickness	Vitart V, Hum Mol Genet 2010;19:4304-4311
		central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
PLXDC2	10p12.31	POAG	Nakano M, Proc Natl Acad Sci USA 2009;106:12838-12842
ATOH7	10q21.3-q22.1	Optic disc size and VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
		Optic disc size	Macgregor S, Hum Mol Genet 2010;19:2716-2724
		Optic disc size	Khor CC, Hum Mol Genet 2011;20:1864-1872
SCYL1 / LTBP3	11q13.1	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
DKFZp762A217	12q21.31	POAG	Nakano M, Proc Natl Acad Sci USA 2009;106:12838-12842
AVGR8	13q12.11	central corneal thickness	Vitart V, Hum Mol Genet 2010;19:4304-4311
		central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
DCLK1	13q13.3	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
FOXO1	13q14.1	central corneal thickness	Lu Y, PLoS Genet 2010;6:e1000947
		central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
SIX1 / SIX6	14q23.1	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
LOXL1	15q24.1	Exfoliation glaucoma	Thorleifsson G, Science 2007;317:1397-1400
AKAP13	15q25.3	central corneal thickness	Vitart V, Hum Mol Genet 2010;19:4304-4311
		central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
LRRK1-CHSY1	15q26.3	central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
SALL1	16q12.1	Optic disc size	Ramdas WD, PLoS Genet 2010;6:e1000978
ZNF469	16q24.2	central corneal thickness	Lu Y, PLoS Genet 2010;6:e1000947
		central corneal thickness	Vitart V, Hum Mol Genet 2010;19:4304-4311
		central corneal thickness	Lu Y, Nat Genet 2013; 45:155-163
GAS7	17p13.1	IOP	Van Koolwijk LME, PLoS Genet 2012
BCAS3	17q23.2	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
CHEK2	22q12.1	VCDR	Ramdas WD, PLoS Genet 2010;6:e1000978
CARD10	22q13.1	Optic disc size	Khor CC, Hum Mol Genet 2011;20:1864-1872

Table 1.1: Genome-wide association studies of POAG and related quantitative traits: the tables shows a summary of the genes identified for POAG and POAG-related traits

1.3 Limits in existing knowledge and aims of this thesis

GWAS represent a powerful tool and have identified hundreds of genetic variants associated with several complex diseases, including POAG, providing valuable insights into the complexities of their genetic architecture (Hindorff *et al.* 2009). However, most variants identified so far confer a small increment of in risk and explain only a small proportion of heritability. Many explanations for this missing heritability have been suggested, including limitations in the design of early GWAS (i.e. imprecise phenotypes), much larger number of variants of smaller effect yet to be found, rarer and structural variants with larger effect sizes that are poorly detectable in genotyping arrays, low power to detect gene-gene interactions and inadequate accounting for shared environment among relatives (Manolio *et al.* 2009; Maher 2008). Consensus is however lacking on approaches to research to examine the ”dark matter” of GWAS (dark matter because we know it exists, we can detect its influence, but we cannot ”see” it).

GWAS have shown that complex diseases cannot be explained by a limited number of common variants of moderate effects. However, GWAS will probably remain an efficient way of investigating missing heritability, because their signals may define the genomic regions where rare and structural variants are likely to cluster. The value of future studies could be enhanced in various ways, including the increase of sample sizes through meta-analyses of comparable data, expanding studies to non-European ancestry populations, improving phenotyping and including more quantitative and precise phenotypes. Other approaches may also include investigating gene-gene and gene-environment interactions and

enhancing the investigation of the X chromosome (Manolio *et al.* 2009).

Following the previous considerations, the main purpose of my PhD project is to gain insight into the genetic aetiology of POAG and related endophenotypes, using GWAS techniques. The rationale is to make the most out of existing GWAS data and samples using different approaches. The specific questions are:

- Which are the genetic variants underlying POAG endophenotypes in the general population and the biological mechanisms underlying them??

- Are POAG endophenotypes correlated with each other and if so is it possible to combine POAG endophenotypes in order to unravel novel shared genetic loci?

- Are genetic variants for POAG endophenotypes interacting with other loci or between each other to determine IOP and VCDR?

- Is RNFL thickness correlated with other POAG endophenotypes? Are there genetic variants affecting RNFL thickness? Do RNFL thickness and Alzheimer's disease have a shared genetics?

CHAPTER 2

Materials and methods

Here I describe materials and methods used in my PhD project.

2.1 Populations

To achieve the objectives of this thesis two cohorts have been used: the Blue Mountains Eye Study (BMES) and the TwinsUK.

2.1.1 The Blue Mountains Eye Study

The BMES is a population-based survey of vision and common eye diseases in the Blue Mountains region, Australia. This area has a stable and homogeneous population, representative of Australians for income and other socioeconomic status measurements. This study was approved by the Western Sydney Area Health Service Human Ethics Committee. Written, informed consent was obtain from all participants. A door-to-door census of the study region was conducted using maps developed by the Australian

Bureau of Statistics. All residents with birth-dates before January 1, 1943, were invited to attend a detailed eye examination at the local clinic. Of the 4433 eligible people, 3654 participated in the study between January 1992 and January 1994, of which 57% were male and 43% females, all aged 49-97 years old (BMES-1). The second BMES study followed up participants from the original cohort, five years later, from 1997-1999. From the original 3,654 people who participated at baseline, 2,334 attended eye examinations for the second data collection period. The third BMES study followed up participants from the baseline cohort, ten years after the original data collection, in 2002-2004. From the 3,654 people who participated at baseline, 1,952 attended interviews for the third data collection period.

2.1.1.1 Phenotyping in the BMES cohort

At the clinic visit, a detailed questionnaire was administered including demographic data, medications, family history, and medical history of systemic disorders. Problems with vision, past eye disease or eye treatment, and ocular symptoms were included. Applanation tonometry using a Goldmann tonometer (Haag-Streit, Bern, Switzerland) was performed using a drop of Fluress (Barnes-Hind). A single measurement was taken, but repeated if judged unreliable. After pupil dilation with tropicamide 1.0% and phenylephrine 10%, stereoscopic retinal and optic disc photographs were taken using a Zeiss fundus camera (FF3 - Carl Zeiss, Oberkochen, Germany) and Kodachrome 25 film processed by Kodak (Eastman Kodak, Rochester, NY). Each participant had stereotopic 30° colour retinal photographs taken centred on the optic disc, macula, and nonstereo photographs of modified lateral macula, upper temporal arcade

and lower temporal arcade. Thirty-five millimetre slide transparencies were mounted in clear plastic sheets, allowing close apposition of the stereo pairs. Stereoscopic assessment of the optic disc in all participants was then conducted by an ophthalmologist, who assessed the disc and field results. Optic disc parameters subsequently were measured from the stereo optic disc photographs using a Donaldson stereoviewer with a plastic template placed under one the stereo pair. Vertical cup-disc ratios were calculated from disc measurements, excluding the peripapillary halo. In measuring vertical diameters, the longest diameter in a range between clock hours of 11 and 1 to 5 and 7 was taken and then used to measure both disc and optic cup. The optic cup was determined by its contour. The path of vessels helped to define the contour of the neuroretinal rim. The presence of rim thinning and extension of the cup to the superior or inferior rim of the disc was noted. Other signs recorded included presence and extent of peripapillary atrophy, hemorrhage crossing the disc margin, or a notch in the cup. All photographs were graded by one or two graders. A summary of the characteristics of the BMES population is shown in Table 2.1

For the BMES cohort, retinal nerve fibre layer thickness (RNFL) measurements for the right eye was available. Several technologies allow the measurements of the retinal ganglion layers and in particular of RNFL layer. For instance, HRT (Heildenberg Retina Tomograph) is a confocal scanning laser that takes images of serial sections of the optic nerve to create a map of topographical sections of the optic nerve head and the surrounding retina. HRT can measure the changes and loss of RNFL (Belyea *et al.* 2014). Subjective refraction results were used to set initial scan focus. The HRT cylindrical lenses were adapted for subjects with astigmatism > -1.0 diopter (D). In brief, a 3-dimensional topographic

image was constructed from multiple focal planes arranged axially along the ONH. An average of 3 consecutive scans were obtained and aligned to compose a single mean topography for analysis. The disc margin was defined as the inner edge of the Elschnig ring and an experienced examiner outlined the optic disc contour on the topographic and reflectance images. ONH parameters were analysed by HRT 3 software automatically. Eyes with a topography standard deviation (TSD) higher than 50 mm were excluded from analysis because of the unreliability of measurement. Only subjects with eligible HRT scans for right eyes were included in the study.

2.1.1.2 Genetic data in the BMES cohort

During the first follow-up, BMES, DNA was collected for 2,983 individuals and sent to the Wellcome Trust Sanger Institute to be genotyped on the Illumina Human 660-W Quad array platform as part of the Wellcome Trust Case Control Consortium 2 (WTCC2). This is a custom chip designed by WTCCC2 comprising Human 550. 2765 samples passed quality control tests for DNA concentration (Invitrogen PicoGreen) and quality (X-Y intensity analysis).

2.1.2 The TwinsUK registry cohort

The TwinsUK adult twin registry based at St.Thomas' Hospital in London is a volunteer cohort of over 12,000 twins from the general population (Spector & Williams 2006; Moayyeri *et al.* 2013). Twins largely volunteered unaware of the eye studies, gave fully informed consent under a protocol reviewed by the St.Thomas' Hospital Local Research Ethics Committee. About 83% of the registry is female (mean age of 55 years). It contains

now 51% MZ and 49% DZ twins. The majority of twins have been seen in a clinical research unit where clinical tests are performed and blood, urine, DNA and adipose tissue (in a subset) are collected and stored at -80°C.

2.1.2.1 Phenotyping in the TwinsUK cohort

Intraocular pressure (IOP) was measured with a non-contact air-puff tonometer. The Ocular Response Analyser (ORA, Reichert, Buffalo, NY) ejects an air impulse in order to flatten the cornea, which is detected by an electro-optical collimation system. The mean IOP was calculated from 4 readings (2 from each eye) for each participant. VCDR in the subjects was measured from stereo disc photographs using the Nidek-3DX stereo camera, with digitised images scanned from Polaroid images and StereoDx stereoscopic planimetric software (StereoDx) using a Z-screen (StereoGraphics Corp) and software obtained from James Morgan from Cardiff University software, Wales.

2.1.2.2 Genetic data in the TwinsUK cohort

Subjects were genotyped in two different batches of approximately the same size, using two genotyping platforms from Illumina: 300K Duo for and HumanHap610-Quad arrays. Whole genome imputation of the genotypes was performed using HapMap2 (www.hapmap.org) haplotypes. Stringent quality control (QC) measures were implemented, including minimum genotyping success rate (>95%), Hardy-Weinberg equilibrium (P>E-06), minimum MAF (>1%) and imputation quality score (>0.7). Subjects of non Caucasian ancestry were excluded from the analysis.

Pheno	BMES	TwinsUK
Age	66.19 (± 9.77),49-97	56.5(± 11.8), 16-83
Disc size	150.3 mm (± 16.35), 66.5-219mm (diameter)	2.58 mm^2 (± 0.6 mm^2),0.58-5.3 mm^2 (area)
CDR	0.42 (± 0.13), 0.09-0.94	0.33(± 0.1), 0.035-0.7
IOP	18.04 (± 2.63), 8.00-34.5 mmHg	15.55 (± 3.18), 6.5-37 mmHg
Gender (% men)	57%	2%

Table 2.1: Characteristics of the BMES and TwinsUK cohorts: the table shows the characteristics of the BMES and TwinsUK cohort

2.1.3 The International Glaucoma Genetics Consortium

Both BMES and TwinsUK cohorts are part of the International Glaucoma Genetics Consortium (IGGC) which puts together the efforts of several studies in order to unravel the genetic of POAG and associated quantitative traits. For these cohorts, the summary statistics for several analysis were available, however all raw phenotypic and genotypic data were analysed separately by each group. I will describe briefly the main characteristics of each cohort:

ALIENOR

The Alienor study is a population-based study in residents of Bordeaux, France. The 963 participants, aged 73 years or more, were recruited from an ongoing population-based study (3C Study). They underwent an ophthalmological examination, including a recording of ophthalmological history, measures of visual acuity, refraction, two 45 non mydriatic colour retinal photographs (one centred on the macula, the other centred on the optic disc), measures of intraocular pressure and central corneal thickness and break-up time test. This research followed the tenets of the Declaration

of Helsinki. Participants gave written consent for the participation in the study. Individuals were genotyped at the French national centre for genotyping (CNG) using Illumina Human 610-Quad BeadChip and then imputed up to Hapmap2 CEU release 22 as reference panel.

Australian Twin Studies (BATS and TEST)

The Australian Twin Eye Study comprises participants examined as part of the Twins Eye Study in Tasmania (TEST) or the Brisbane Adolescent Twins Study (BATS) (Mackey *et al.* 2009). The Australian cohorts were genotyped on the Illumina Human Hap610W Quad array, with part of the sample typed alongside the TwinsUK cohort and the remainder typed as a separate contract with DeCODE genetics. Imputation was done with reference to HapMap release 22 CEU using MACH (http://www.sph.umich.edu/csg/abecasis/MACH/). In BATS data from 1,152 people, from 517 families, were included in the analyses. For TEST, 686 individuals from 350 families were included. The mean IOP of both eyes was used as outcome variable.

Beijing Eye Study

The Beijing Eye Study (BES) is a population-based cohort of Han Chinese in the rural region and in the urban region of Beijing in North China (Xu *et al.* 2005, 2008).The Medical Ethics Committee of the Beijing Tongren Hospital approved the study protocol and all participants gave informed consent, according to the Declaration of Helsinki. At baseline in 2001, 4,439 individuals out of 5,324 eligible individuals aged 40

years or older participated (response rate: 83.4%). In the year 2006, the study was repeated by re-inviting all participants from the survey from 2001 to be re-examined. Out of the 4,439 subjects examined in 2001, 3,251 (73.2%) subjects returned for the follow-up examination in 2006. All study participants underwent an ophthalmic examination including refractometry, pneumotonometry, slit-lamp biomicroscopy and photography of the cornea, lens, optic disk and macula. Genotyping was performed in 2,929 individuals using Illumina Human610-Quad BeadChip in 988 subjects. Imputation was performed using IMPUTE v2.2.2 7 on post-QC SNPs. The HapMap Phase II panel (build36, release 22 db126 JPT+CHB HapMap panel) was used for the imputation.

Erasmus Rucphen Family Study (ERF)

The Erasmus Rucphen Family (ERF) Study is a family-based cohort in a genetically isolated population in the southwest of the Netherlands with over 3,000 participants aged between 18 and 86 years (Aulchenko *et al.* 2004; Pardo *et al.* 2005). Cross-sectional examination took place between 2002 and 2005. The IOP was measured with Goldmann applanation tonometry (Haag-Streit, Bern, Switzerland). All measurements in these studies were conducted after the Medical Ethics Committee of the Erasmus University had approved the study protocols and all participants had given a written informed consent in accordance with the Declaration of Helsinki. DNA was genotyped on one of four different platforms (Illumina 610k, Illumina 300K, Illumina 370K and Affymetrix 250K) and then imputed up to HapMap 2 build 36.

Framingham Eye Study

The Framingham Eye Study (FES) was nested within the Framingham Heart Study (FHS, http://www.framinghamheartstudy.org), which began its first round of extensive physical examinations in 1948 by recruiting 5,209 men and women from the town of Framingham, MA, USA (Hm *et al.* 1979). Surviving participants from the original cohort returned for biennial exams, which continue to the present. A total of 2,675 FHS participants were also examined as part of the FES between 1,973 and 1,975. The FES was designed to evaluate ocular characteristics of examinees such as: senile cataract, age-related macular disease, glaucoma, and retinopathy. Between 1989 and 1991, 1,603 offspring of original cohort participants also received ocular examinations. Genotyping was conducted as part of the NHLBI Framingham SNP Health Association Resource (SHARe). This sub-study contains genotype data for approximately 550000 SNPs (Affymetrix 500K mapping arrays [Mapping250kNsp and Mapping250KSty] plus Affymetrix 50K supplemental human gene-focused array) in over 9200 FHS participants. Genotype imputation to the HapMap-II reference panel (CEU population release 22, NCBI build 36) was carried out in a two-step process using the Markov Chain Haplotyping (MACH version 1.0.16.a) software.

Gutenberg Health Studies (GHS I, GHS II)

The GHS is a population-based, prospective, observational cohort study in the Rhine-Main Region in midwestern Germany with a total of 15,010

participants and follow-up after five years. The study sample is recruited from subjects aged between 35 and 74 years at the time of the exam. The sample was drawn randomly from local governmental registry offices and stratified by gender, residence (urban and rural) and decade of age. Within GHS, DNA was extracted from buffy-coats from EDTA blood samples. Genetic analysis was conducted in the first 5,000 study participants. For these, 3,463 individuals were genotyped in 2008 (GHS I) and further 1,439 individuals in 2009 (GHS II). Genotyping was performed for GHS I and GHS II using the Affymetrix Genome-Wide Human SNP Array 6.0. Genotypes were called using the Affymetrix Birdseed-V2 calling algorithm.Imputation of missing genotypes was performed using IMPUTE software v2.1.012 and HapMap release 24, NCBI Build 36.

Orkney Complex Disease Study (ORCADES)

The Orkney Complex Disease Study (ORCADES) is a population-based, cross-sectional study in the Scottish archipelago of Orkney, including 1,285 individuals with eye measurements. The study received approval from relevant ethics committees in Scotland and followed the tenets of the Declaration of Helsinki. Genotypes were generated using HumanHap 300v2 or 370CNV-Quad Illumina SNP arrays. Genotypes were determined using the Illumina BeadStudio software following the manufacturerâs standard recommendations. Imputation of allele dosage for over 2 million SNPs on the 22 autosomal chromosomes with reference to HapMap CEU build 36 release 22 was performed using the software MACH v1.0.15.

RAINE (Western Australian Pregnancy Cohort 20 year follow up Eye Study)

The Western Australian Pregnancy Cohort (Raine) Study is an ongoing prospective cohort study of pregnancy, childhood, adolescence and young adulthood in Perth, Western Australia. At the initiation of the study, 2,900 pregnant women were recruited at 16-18 weeks' gestation from the state's largest public women's hospital and surrounding private practises for a randomised clinical trial investigating effects of intensive ultrasound and Doppler studies in pregnancy outcomes. Following this study, the offspring of the recruited individuals have been evaluated in detail during childhood and adolescence. At the 20-year review of the cohort, Raine participants underwent a comprehensive ocular examination for the first time. DNA samples and consents for GWAS studies were available from the previous assessments. Genotype data was generated using the genome-wide Illumina 660 Quad Array at the Centre for Applied Genomics (Toronto, Ontario, Canada). Imputation was performed with MACH v1.0.16 (http://www.sph.umich.edu/csg/yli/mach/index.html) software using the CEU samples from HapMap phase2 build 36 release 22 (http://hapmap.ncbi.nlm.nih.gov/index.html.en).

Rotterdam Study (RS-I, RS-II, RS-III)

The Rotterdam Study is a prospective population-based cohort study in the elderly living in Ommoord, a suburb of Rotterdam, the Netherlands (Hofman *et al.* 2011). In brief, the Rotterdam Study consists of 3 independent cohorts: RSI, RS-II, and RS-III. Participants underwent multiple physical examinations with regular intervals from 1991 to present. All measurements in RS-I, RS-II and RS-III were conducted after the

Medical Ethics Committee of the Erasmus University had approved the study protocols and all participants had given a written informed consent in accordance with the Declaration of Helsinki. DNA was extracted from blood leucocytes according to standard procedures. Genotyping of SNPs was performed using the Illumina Infinium II HumanHap550 chip v3.0 array (RS-I); the HumanHap550 Duo Arrays and the Illumina Human610-Quad Arrays (RS-II), and the Human 610 Quad Arrays Illumina (RS-III). Data were imputed up to NCBI build 36, HapMap release.

Singapore cohorts (SIMES, SINDI, SCES)

Singapore Malay Eye Study (SIMES)

SIMES is a population-based prevalence survey of Malay adults aged 40 to 79 years living in Singapore that was conducted between August of 2004 and June of 2006 (Foong *et al.* 2007). From a Ministry of Home Affairs random sample of 16,069 Malay adults in the Southwestern area, an age-stratified random sampling strategy was used in selecting 1400 from each decade from age 40 years onward (40-49, 50-59, 60-69, and 70-79 years). The 4,168 eligible participants from the sampling frame, while 3,280 (78.7%) participated. Genome-wide genotyping was performed in 3,072 individuals. Total of 3,072 DNA samples were genotyped using the Illumina Human 610 Quad Beadchips.

Singapore Indian Eye Study (SINDI)

SINDI is a population-based survey of major eye diseases in ethnic Indians aged 40 to 80 years living in the South-Western part of Singapore and was conducted from August 2007 to December 2009. In brief, 4,497 Indian adults were eligible and 3,400 participated. Genome-wide genotyping was performed in 2,953 individuals. The Illumina Human610 Quad Beadchips was used for genotyping all DNA samples from SINDI (n=2,593).

Singapore Chinese Eye Study (SCES)

Similar to SINDI, the Singapore Chinese Eye Study (SCES) is a population-based cross-sectional study of eye diseases in Chinese adults 40 years of age or older residing in the southwestern part of Singapore. The methodology of the SCES study has been described in details previously. Between 2009 and 2011, 3,353 (72.8%) of 4,605 eligible individuals underwent a comprehensive ophthalmologic examination, using the same protocol as SINDI. Genome-wide genotyping using was done in a subset of 1,952 SCES participants using Illumina Human610-Quad BeadChip. After phenotype and genotype QC, 1,723 individuals were left for the analysis.

POAG case-control cohorts

ANZRAG

The Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG) recruits cases of advanced glaucoma Australia-wide through ophthalmologist referral. The cohort also included participants enrolled in

the Glaucoma Inheritance Study in Tasmania (GIST) that met the criteria for ANZRAG. Advanced Primary Open Angle Glaucoma was defined as best-corrected visual acuity worse than 6/60 due to primary open angle glaucoma, or a reliable 24-2 Visual Field with a mean deviation of worse than -22db or at least 2 out of 4 central fixation squares affected with a Pattern Standard Deviation of $< 0.5\%$. The less severely affected eye was also required to have signs of glaucomatous disc damage. Worst ever recorded intraocular pressure was obtained from the patientâs medical record. Due to the advanced nature of the glaucoma, all participants were on medical treatment to lower intraocular pressure and/or had undergone surgery for the same. DNA extracted from peripheral whole blood was genotyped on Illumina Human1M-Omni or OmniExpress arrays. SNPs with a mean BeadStudio GenCall score < 0.7 were excluded. Controls used for this analysis were historic, primarily unexamined controls, from several sources. The Blue Mountains Eye Study (described above) were included. Known cases of glaucoma were excluded from this dataset. Additional controls from the Wellcome Trust Case Control Consortium 1958 British Birth Cohort genotyped on Illumina arrays and Caucasian samples from the Illumina iControl datasets were used. Imputation was calculated with reference to HapMap release 22 CEU using MACH.

deCode

The Icelandic primary open angle glaucoma (POAG) cases were identified from a list of participants in the Reykjavik Eye Study and from a list compiled by Icelandic ophthalmologists in 2008 that included patients 55-86 years old at the time of diagnosis, all meeting either structural (glaucomatous optic neuropathy) or functional (glaucomatous visual field

defects) criteria of glaucoma or both. For visual fields measurements, Octopus 123 perimeter (Haag-Streit AG, Koniz Switzerland) was used. Intra ocular pressure (IOP) was not a part of the definition. On gonioscopy the angles were found to be open and normal in appearance. The combined list includes 598 individuals, 290 men and 308 women with POAG. The diagnosis of exfoliation syndrome (XFS) was specifically evaluated and if detected the participant was excluded from the study. The control group included 98,670 Icelandic individuals without know history of glaucoma. 533 of the POAG samples and 85,689 of the controls samples were assayed with the Illumina HumanHap300, HumanCNV370, HumanHap610, HumanHap1M, HumanHap660, Omni-1, Omni 2.5 or Omni Express bead chips at deCODE genetics and genotypes for about 34 million sequence variants were imputed into the chip typed individuals based on a training set of 2,230 whole genome sequenced Icelanders.

MEEI

491 cases and 351 controls collected from the Massachusetts Eye and Ear Infirmary glaucoma clinic and comprehensive ophthalmology clinics were recruited for this study. All cases and controls were residents of the continental United States and were of mainly European ancestry, which was confirmed by both self-identification and genetic markers. Primary open angle glaucoma (POAG) cases were defined as individuals for whom reliable visual field (VF) tests show characteristic VF defects consistent with glaucomatous optic neuropathy. Individuals were classified as affected if the VF defects were reproduced on a subsequent test or if a single qualifying VF was accompanied by a cup-disc ratio (CDR) of 0.7 or more in at least one eye. The examination of the ocular anterior

segment did not show signs of secondary causes for elevated IOP such as exfoliation syndrome or pigment dispersion syndrome and the filtration structures were deemed to be open based on clinical measures. Elevation of IOP was not a criterion for inclusion; however, 67% of cases did have a history of elevated IOP ($\geq$ 22 mm Hg) measured in a clinical setting (typically between the hours of 8AM and 5PM) and were classified as high-pressure glaucoma (HPG). Genotyping was performed using the Illumina Human660WQuadv1 array and 495,132 SNPs passed quality control filters. Illumina's BeadStudio and GenomeStudio and Autocall software along with genotype cluster definitions based on study samples were used to generate genotyping calls. Imputation was performed with IMPUTE2 using the March 2012 1000 genomes as a reference panel.

NEIGHBOR

2,170 cases and 2,347 controls collected from 12 sites throughout the United States were genotyped for the NEIGHBOR study. Primary open angle glaucoma (POAG) cases were defined as individuals for whom reliable visual field (VF) tests show characteristic VF defects consistent with glaucomatous optic neuropathy. Individuals were classified as affected if the VF defects were reproduced on a subsequent test or if a single qualifying VF was accompanied by a cup-disc ratio (CDR) of 0.7 or more in at least one eye. The examination of the ocular anterior segment did not show signs of secondary causes for elevated IOP such as exfoliation syndrome or pigment dispersion syndrome and the filtration structures were deemed to be open based on clinical measures. Elevation of IOP

was not a criterion for inclusion; however, 67% of cases did have a history of elevated IOP ($\geq$ 22 mm Hg) measured in a clinical setting (typically between the hours of 8AM and 5PM) and were classified as high-pressure glaucoma (HPG). Genotyping was performed using the Illumina Human660WQuadv1 array and 523,528 SNPs passed quality control filters. Allele cluster definitions for each SNP were determined using Illumina GenomeStudio Genotyping Module version 1.7.4, GenTrain version 1.0 and the combined intensity data from 99.9% of the samples. The resulting cluster definitions were used on all samples. Imputation was performed with IMPUTE2 using the March 2012 1000 genomes as a reference panel.

Study	n	Mean age(sd)	Age Range	% men	Mean VCDR(sd)	VCDR range
BATS	966	20.2(3.8)	13-34	46.1	0.46(0.13)	0.09-0.75
BMES	1656	66.2(9.8)	49-97	57	0.43(0.13)	0.09-0.94
ERF	2131	47.3(14)	18-85	44.4	0.31(0.2)	0.00-0.87
GHS-I	783	55.9(10.9)	35-74	52.7	0.42(0.09)	0.22-0.86
GHS-II	485	55.1(10.9)	35-74	49.1	0.42(0.11)	0.17-0.90
Glaugen	341	64.4(11.1)	40-89	46.3	0.3(0.1)	0.10-0.70
Neighbor	2050	69.3(11.3)	35-97	43.9	0.32(0.13)	0.00-0.90
RAINE	1046	20(0.4)	18-22	51.3	0.28(0.2)	0.00-0.78
RS-I	5322	68(8.4)	55-99	41.6	0.5(0.13)	0.05-0.87
RS-II	2054	64.9(7.8)	55-98	45.8	0.5(0.13)	0.10-0.86
RS-III	1962	56.1(5.5)	45-89	43.9	0.29(0.21)	0.00-1.00
TEST	376	20.9(17.4)	5-79	41.5	0.44(0.12)	0.09-0.88
TwinsUK	1922	56.9(11.6)	16-83	2.1	0.34(0.11)	0.04-0.70
BES	727	58.6(9.3)	45-86	35	0.43(0.13)	0.00-0.77
SCES	1726	58.8(9.6)	44-86	51.2	0.34(0.17)	0.00-0.95
SINDI	2026	58(10)	43-84	51.2	0.46(0.23)	0.00-1.00
SIMES	2305	59.1(11)	40-80	49.4	0.39(0.23)	0.00-1.00

Table 2.2: Summary of demographic and phenotypes characteristics for each cohort for VCDR

Abbreviation used in article	Cohort name	N	Mean age (SD)	% males	Mean mmHg IOP (SD)	IOP measurement method	Genotyping Platform(s)	Imputation Method
ALIENOR	Antioxydants, Lipides Essentiels, Nutrition et maladies OculaiRes	939	80.20 (4.44)	38.02	14.21 (2.72)	Non contact tonometer (Kowa KT800)	Illumina Human 610-Quad BeadChip	
BATS	Brisbane Adolescent Twins Study	1152	20.10 (0.08)	46	15.78(2.88)	TONO-PEN XL (Reichert, Inc. New York, USA)	Illumina 610/660 Quad Array	Mach
BES	Beijing Eye Study	805	60.42(10.04)	40.4	15.73 (3.03)	Non-contact pneumotonometer	Illumina Human Quad 610 chip	IMPUTE
BMES	Blue Mountains Eye Study	1667	66.19(9.77)	57	16.04(2.63)		Illumina Human 660W Quad	IMPUTE
ERF	Erasmus Rucphen Family Study	2591	49.05(14.31)	45	15.09(3.00)	Goldmann applaration tonometry (Haag-Streit, Bern, Switzerland)	Illumina 6K, Illumina 318K, Illumina 370K, Affymetrix 250K	Mach
Framingham	Framingham Eye Study	2455	57.38 (12.39)	44.8	13.6 (3.32)	Goldmann applaration tonometry	Affymetrix 250k_Nsp 250k_Sty HuGeneFocused50K	Mach
GHS I	Gutenberg Health Study I	2727	55.9(10.9)	52	14.19(2.81)	Non-contact tonometer	Affymetrix Genome-Wide Human SNP 6.0 Array	IMPUTE
GHS II	Gutenberg Health Study II	1130	55.10(10.90)	49.6	13.90(2.7)	Non-contact tonometer	Affymetrix Genome-Wide Human SNP 6.0 Array	IMPUTE
ORCADES	Orkney Complex Disease Study	474	56.5(14.5)	38	15.38(2.80)			
RAINE	Western Australian Pregnancy (Raine) Cohort 20 year follow up Eye Study	1025	20.0 (0.43)	51	15.42(3.21)	Icare TAO1i Tonometer, Icare Finland Oy, Helsinki, Finland	Illumina 660 Quad Array	Mach
RS-I	Rotterdam Study-I	5782	68.76(8.92)	41	14.72(3.34)	Goldmann applaration tonometry (Haag-Streit, Bern, Switzerland)	Illumina Infinium II HumanHap500chip v3.0 array	Mach
RS-II	Rotterdam Study-II	2116	64.81(7.91)	46	14.2(3.1)	Goldmann applaration tonometry (Haag-Streit, Bern, Switzerland)	HumanHap550 Duo Arrays + Human610-Quad Arrays Illumina	Mach
RS-III	Rotterdam Study-III	2038	56.15(5.81)	44	13.61(2.96)	Goldmann applanation tonometry (Haag-Streit, Bern, Switzerland)	Human 610 Quad Arrays Illumina	Mach
SCES	Singapore Chinese Eye Study	1884	58.51 (9.53)	51	14.60 (3.21)	Goldmann Tonometry	Illumina Human Quad 610 chip	IMPUTE
SIMES	Singapore Malay Eye Study	2537	59.09 (11.04)	49	15.42 (3.76)	Goldmann Tonometry	Illumina Human Quad 610 chip	IMPUTE
SINDI	Singapore Indian Eye Study	2535	58.04 (10.01)	51	15.84 (2.87)	Goldmann Tonometry	Illumina Human Quad 610 chip	IMPUTE
TEST	Twins Eye Study in Tasmania	663	25.64(18.80)	40	15.84(3.12)	TONO-PEN XL (Reichert, Inc. New York, USA)	Illumina 610/660 Quad Array	Mach
TwinsUK	Twins UK Cohort Study	2776	56.46 (11.83)	2	15.55 (3.18)	Ocular Response Analyser	300K Duo, HumanHap610-Quad array	IMPUTE

Table 2.3: This table has been published elsewhere (Hysi *et al.* 2014). Summary of the demographic and phenotypes characteristics for each cohort for intraocular pressure

2.2 Statistical analyses

To date several methods have been used to investigate the genetic of complex disease such as POAG. I present here the statistical analyses used and developed for the purpose of this thesis.

2.2.1 Imputation and quality control for the BMES cohort

Genotyping data from the BMES cohort were imputed up to 1000Genomes Pilot data (2010) (http://www.1000genomes.org/) using IMPUTE2 (Howie *et al.* 2009). Imputation methods estimate unobserved genotypes

by combining a reference panel of individuals genotyped at a dense set of polymorphic sites (i.e. SNPs) with a study sample collected from a genetically similar population and genotyped at a subset of these sites. A population genetic model is used to extrapolate allelic correlations measured in the reference panel that can predict unobserved genotypes in the study sample. The aim is to expand the set of SNPs that can be tested for association and can enhance true association signals and facilitate meta-analysis.
All other cohorts included in this thesis were imputed to HapMap2 when I have started my Phd course, for this reason only the SNPs also available in the HapMap 2 data-set were selected from the 1000G imputation. This step lead to a total of 2,617,903 SNPs.

Quality control was then performed at SNP and individual level using Plink v1.07 (http://pngu.mgh.harvard.edu/ purcell/plink/). At the SNP level quality control included excessive missing genotype, identification of SNPs showing a significant deviation from Hardy-Weinberg equilibrium (Hardy Weinberg p-value less than 0.001) and removal of markers with very low allele frequency (less than 4%). Removing these low quality SNPs was a key process in further analyses in order to reduce false positives and identify the true association. Quality control was also performed in the samples. Classically it involves identification of individuals with discordant sex information, outlying missing genotype and heterozygosity rate and duplicated and related individuals. In addition, homogeneity of the population was checked in order to avoid issue due to ancestry structure using principal components. The principal components are calculated so that the first principal component accounts for as much variation as possible in the data. In the PCA model of ancestry detection,

the observations are the individuals and the potentially correlated variables are the markers. I calculated the principal component in the BMES data using pruned genome-wide genotype, that is when only independent (non in LD) markers are included. This is was done in Plink v1.07. Figure 2.1 shows the plot of the first two principal components against each other. The data were then merged with four reference panels from HapMap Phase 2 data from Europe (CEU), Asia (CHB+ JPT) and Africa (YRI). Figure 2.2 shows that the BMES cohort clusters together with the CEU population from HapMap2.

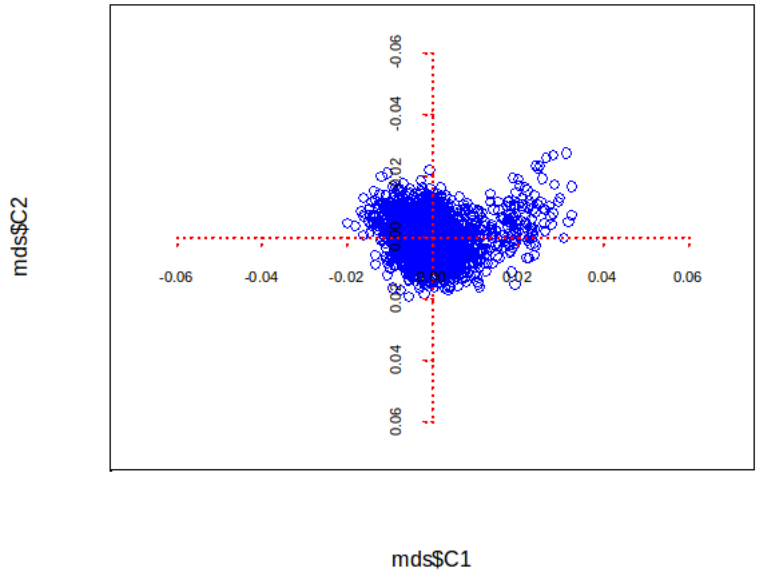


Figure 2.1: First and second principal components for the BMES cohort

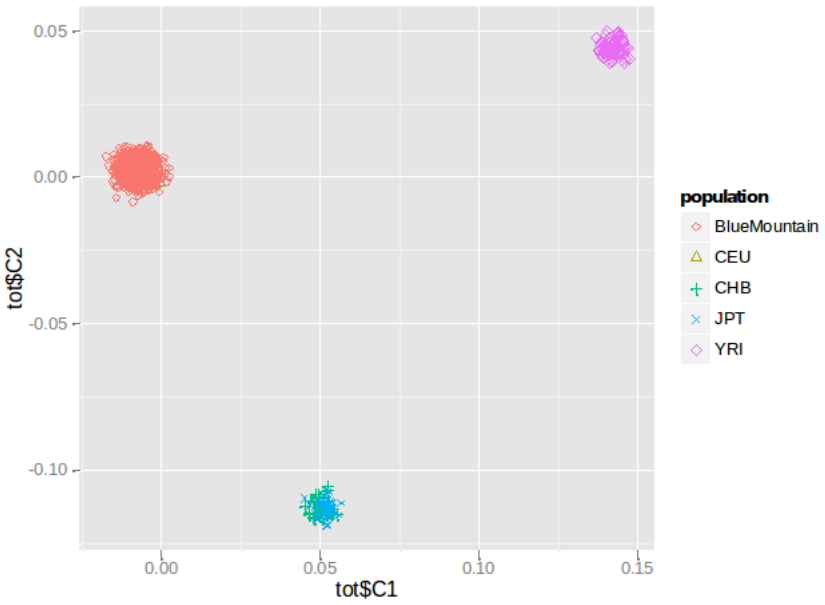


Figure 2.2: First and second principal components for the BMES cohort plotted with HapMap2 reference panels

2.2.2 Genome-wide association studies

Following quality control, genome-wide association analyses were performed in the BMES and TwinsUK cohorts.

The analysis of genome-wide association data is a series of single-locus tests, examining each SNP independently for association to the phenotype. Quantitative traits are generally analysed using linear regression model approaches where the phenotype of interest is the outcome (y) and the genotype for a locus is the predictor(x) as in:

$$y = X\beta + \epsilon \tag{2.1}$$

The measure of association between x and y is given by $\beta 1$, defined as the amount of change that occurs with one unit change in x. If x is an indicator for the presence of a variant allele at a given SNP locus and y is the IOP level, then $\beta 1$ is the difference in the mean IOP between individuals

with or without this variant allele. Importantly, $\beta 1$ is a measure of linear association, capturing information on the extent to which the relationship between x and y is a straight line. A linear regression can also include multiple independent variables becoming a multivariable linear regression model. For this additive model, the null hypothesis is the no association between the genotype and the trait, given by $h0 : \beta 1 = 0$. The calculation of the parameter can be done with application of the Wald test or likelihood ratio test.

All GWAS for the BMES in this thesis are performed using Plink v1.07 (http://pngu.mgh.harvard.edu/ purcell/plink/). An example command is:

```
for i in {1..22}
do
plink --bfile mydata_chromosome$i --linear --ci 0.95
--pheno mydata_pheno --pheno-name IOP --covar mydata_cov
--covar-name AGE,SEX --out results
done
```

The outcome of the model were POAG endophenotypes (IOP, optic disc measurements and RNFL thickness) and the predictor was the genotype. All analysis were adjusted for age and sex. GWAS for the TwinsUK were conducted in a similar way, however the analyses needed to take into account the family structure. This is was done using MERLIN (http://csg.sph.umich.edu/abecasis/Merlin/).

Manhattan and quantile-quantile plots were produced with R

(https://www.r-project.org/). The QQ plot shows the expected distribution of association test statistics (X-axis) across the million SNPs compared to the observed values (Y-axis).

Genomic control was calculated for each GWAS using a R package called GenABEL (http://www.genabel.org/). Genomic control is a standard quality-control measure for GWAS and meta-analysis and the concept behind is that apart from a small number of SNPs that show true association with the trait, the test statistics for other SNPs should follow the distribution under the null hypothesis of no association between a SNP and a trait. Test statistics will be inflated across the whole genome in case of an artificial difference in allele frequencies for example due to population stratification, cryptic relatedness and errors.

Regional association plots of GWAS results were produced using a web programme called LocusZoom (http://locuszoom.sph.umich.edu/locuszoom/)

2.2.3 Meta-analysis of genome-wide association studies

Inverse variance weighted fixed-effect meta-analyses were performed using either METAL software (http://www.sph.umich.edu/csg/abecasis/metal/index.html) or GWAMA (http://www.well.ox.ac.uk/gwama/). For IGGC meta-analyses, summary association statistics for each variant and phenotype of interest (IOP and VCDR) were generated from individual level information locally at each participating team. All study effect estimates were corrected using genomic control and were oriented to the positive strand of the NCBI Build 36 reference sequence of the human genome. After meta-analysis SNPs that

were available in less than 16 populations and/or were showing large heterogeneity, defined as I2>75% were removed. Heterogeneity is an important concept in meta-analysis and it exists when the true effects differ between studies and may be detectable if the variation between the results of the studies is greater than expected by chance. It is therefore important to quantify the extent of heterogeneity among a collection of studies. I2 describes the percentage of variability in point estimates that is due to heterogeneity rather than sampling error (Higgins & Thompson 2002).

2.2.3.1 Gene expression

As the majority of the common genetic variants that are associated with complex traits map to non-coding regions and may thus alter gene regulation (Pennisi 2011), gene expression data were used to investigate whether regions identified through GWAS and meta-analyses were related to cis-regulation (in the vicinity of the gene) of mRNA. This analysis was performed following IGGC meta-analysis on IOP. mRNA expression of selected variants was investigated across multiple tissues: adipose, lymphoblastoid cell lines (LCL) and skin. The analysis was performed in 856 mono and dyzigotic twins from the TwinsUK cohort from the MuTHER (Multiple Tissue Human Expression Resource) project (Grundberg *et al.* 2012).

2.2.4 Gene annotation and gene-set enrichment analysis

Results for each GWAS obtained were annotated using the ENSEMBL database (www.ensembl.org). Each gene was initially assigned the significance of the best associated Single Nucleotide Polymorphism (SNP) within its transcription-coding protein and the $\pm$100 kb flanking its

transcription coding site. The choice of the extent of the flanking region is often subjective. This choice was made on the basis of evidence that most cis- regulation occurs within this distance for most genes (Dixon *et al.* 2007). In case of conflict (a SNP eligible for more than one gene) the gene closest to the SNP was used for annotation purposes and in the rare instances where a SNP was inside two genes' transcript start sequences, both genes were assigned the significance of the SNP in question as long as this was the strongest associated SNP within their sequences. Only genes whose best-associated SNP passed a pre-set significance threshold of $p \leq 0.001$ were included in further enrichment analysis.

Gene set enrichment analyses (GSEA) were carried out using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) Bioinformatics tool (Huang *et al.* 2008) to determine if genes selected on the basis of their GWAS significance participate in specific annotation categories. The rationale of GSEA is that if a biological process is abnormal in a given study, the co-functioning genes should have a higher potential (enriched) to be selected as a relevant group by GWAS. I focused on functional categories for the enrichment of the three classes of the Gene Ontology (GO) database (biological processes, cellular components and molecular functions) (http://www.geneontology.org/) and also the Kyoto Encyclopedia of Genes and Genomes (KEGG) categories. I then investigated the evidence that any of these categories were over-represented in the GWAS results. Fisher's exact tests were conducted to assess the actual representation of these annotation categories compared to expectations under the null hypothesis of non enrichment and randomness of each gene set in the GWAS results.

2.2.4.1 Meta-analysis of gene-set enrichment results

Fisher's exact p-values were meta-analysed using a method to combine p-value developed by Fisher. I included in the meta-analyses only those results for which fold-change enrichment's values were in the same direction for all cohorts, all greater than 1 or all smaller than 1, and removed from the analyses functional categories that showed trends with less clear directionality.

2.2.5 Multivariate analysis

Several methodologies have been used to combine POAG quantitative traits in a GWAS setting including epidemiological and genetic correlation, dimension reduction methods and regression models.

2.2.5.1 Correlation

A total of 2,302 from the BMES was included in the analysis. A list of several factors putatively linked with POAG was compiled. This list included: IOP, vertical disc diameter (VDD), vertical cup diameter (VCD), VCDR, CCT, refractive error and age. All variables were standardised in order to have all mean equal to zero and standard deviation equal to 1. A correlation analysis was carried out in the BMES sample using R (http://www.r-project.org/) and Pearson's coefficient was obtained.

2.2.5.2 Genetic correlation

I used a linear mixed model to obtain estimates of the genetic covariance and hence genetic correlation between POAG traits as implemented in GCTA software (Yang *et al.* 2011) (Genome Complex trait analysis). This technique allows to estimate the genetic variance of each trait and that genetic covariance between two traits that can be captured by all SNPs (Lee *et al.* 2012). A set of 2,301 individuals with information on a panel of 2,617,903 SNPs from the BMES was analysed.

2.2.5.3 Principal factor analysis

A total of 5,796 individuals of Caucasian origin were taken from two different cohorts: 2,302 from the BMES and 3,493 from the TwinsUK. An iterated factor analysis was carried out aiming to capture the main variation from POAG phenotypes. IOP, VDD, VCD, CDR, CCT measurements and age were included. The analysis was run including all the cohorts in order to have a phenotype that I could compare across populations. The analysis was run in STATA version 12 (Stata Corp. College Station, TX) using ”ipf” command. The most relevant factors were identified, using a threshold of an eigenvalue of 1 (Kaiser’s criterion) as the criterion to include principal factors in subsequent stages of the analysis.

2.2.5.4 Regression models

There are several approaches to perform a multivariate test of association with multiple phenotypes. For instance, one could simply test the global null hypothesis of no association with either phenotype using MANOVA treating

the two phenotypes as a bivariate normal response and the single genetic variant as an explanatory variable (Stephens 2013).

A set of 2,302 of individuals from the BMES and a set of 1,508 from the TwinsUK were analysed. In the TwinsUK one twin of each pair needed to be excluded to avoid issue created by family structure. A bivariate regression model was implemented including IOP and VCDR. The multivariate linear model is:

$$Y_{(nxm)} = X_{(nxk+1)}B_{(k+1xm)} + E_{(nxm)}$$

where Y is a matrix of *n* observation on *m* response variable; X is a model matrix with columns for *k+1* regressors, B is a matrix of regression coefficients, one column for each response variable; and E is a matrix of errors. There are several commonly employed multivariate test statistics that can be used: Pillai-Bartlett Trace, Hotelling-Lawley Trace, Wilk's Lambda and Roy's Maximum Root. This multivariate linear model was fit in R with the **lm** function, where the left-hand side of the model is a matrix of responses with a column for IOP and a column for CDR and each row and observation; the right-hand side of the model are the same as for an univariate linear model. The response matrix is composed from individual response variables via the **cbind** function. First I fitted a multivariate linear model to the data, treating IOP and CDR as responses:

```
1 mod <- lm(cbind(IOP, CDR) ~ SNP + age + gender
```

We then use the **manova** function in R to test the null hypothesis of no association with either IOP and CDR:

```
test <- manova(mod)
```

I then use the function **summary** to extract the Pillai-Bartlett Trace p-values. Results from the BMES and TwinsUK were then combined using Fisher's combine probability method as described elsewhere (Fisher 1992).

Another multivariate regression model known as generalised estimating equation (GEE) was implemented. This approach includes generalised linear models that take into account a possible unknown correlation between outcomes. A set of 2,302 individuals from the BMES and 1,508 from the TwinsUK cohort were selected. The GEE implemented in SMAT R package was applied. Specifically in Schifano *et al.* paper (Schifano *et al.* 2013) when multiple phenotypes are positively correlated and measure the same underlying trait in the same direction (after transformation), the authors have implemented an inverse probability weighted (IPW) regression in order to test and estimate the shared common effect of SNPs on multiple continuous secondary phenotypes. As an estimating-equation-based approach, it accounts for arbitrary correlation among multiple phenotypes and is robust to departure from normality.

2.2.6 Snp by snp interaction model

An interaction model was implemented to take into account of snp by snp interaction in a regression framework.

Two lists of 17 SNPs associated with VDCR and 10 SNPs associated with IOP in the IGGC meta-analysis described in Chapter 3 were compiled.

The model was:

$$VCDR = \beta 1SNP1 + \beta 2SNP2 + \beta 3SNP1xSNP2 \quad (2.2)$$

and:

$$IOP = \beta 1SNP1 + \beta 2SNP2 + \beta 3SNP1xSNP2 \quad (2.3)$$

where SNP1 is one of the variant included in the lists and SNP2 in a SNP included in all genome (imputed to HapMap2 for a total of $>$ 2,500,000 variants). This produced a total of $>17*$2,500,000 interaction models for VCDR and $>12*$2,500,000 for IOP. All analyses were adjusted for age and sex. The analysis was performed in Plink v1.07 (command –epistasis). Family structure was not taken into account, thus only unrelated individuals could be used. For example, for the TwinsUK, I selected one individual from each pair at random to be included in further analyses.

2.2.7 Investigating the shared genetics between two diseases

Several methods can be used to investigate the genetic link between two diseases or quantitative traits. A possible genetic overlap between the

RNFL and Alzheimer's disease (AD) was investigated. AD information was not available either in the BMES or TwinsUK cohorts, thus a list of candidate genetic variants for AD was create through literature search. In particular, results from the Alzheimer's Disease Genetics Consortium (ADGC) on late-onset Alzheimer's disease (Naj *et al.* 2011) were used. The meta-analysis included a total of 18,832 cases and 35,597 controls and genotypes were imputed up to HapMap 2. Summary statistics for all SNPs with a p-value$<$0.0001 were publicly available. 435 genetic variants were included in the list. A more recent meta-analysis for AD was conducted in 2013, however results from the meta-analysis in 2011 were preferred because all results with a p-value $<$0.0001 were included in the paper and data were imputed in HapMap 2 rather than 1000G. This AD-associated SNPs list was used to check whether there was an overlap with RNFL meta-analysis and if the effect sizes from the two meta-analyses were correlated. All these analyses and related graphs were performed in R (http://www.r-project.org/).

2.2.7.1 Polygenic risk score

To evaluate whether AD-associated SNPs have an aggregated role on RNFL thickness, polygenic risk score analysis was performed. Variation across associated loci for AD was summarised into quantitative scores and then these scores were related to RNFL thickness variation in the BMES population. Two sets of SNPs were used to generate aggregate risk scores. The first risk score included 37 SNPs, selecting from the 435 AD-associated SNPs according to linkage disequilibrium (LD). The second score was created including 16 AD-associated SNPs that are also significant in the RNFL meta-analysis conducted and are independent between each other.

For each individual in the BMES cohort I calculated the number of score alleles they possessed, each weighted by the log of the odds ratio from the AD meta-analysis. The score is simply a sum across SNPs of the number of reference alleles (0,1,2) and that SNPs multiplied by the score for that SNP. The process of selecting SNPs according to linkage disequilibrium and the calculation of the risk scores was done using PRSice software (Euesden *et al.* 2015). PRSice is a software package written in R, including wrappers for PLINK2 (Chang *et al.* 2015) to minimise computational time.

2.2.7.2 Multi-SNP risk score

Large sample sizes are required to identify genetic association and likely the BMES cohort alone does not have enough power to pinpoint genetic overlap between RNFL thickness variation and AD through the polygenic risk score. Given the small sample size in the BMES cohort (960 individuals), a multi-SNP score analysis, which allows summary results from both RNFL thickness meta-analysis and AD meta-analysis, is appealing. The method has been proposed to investigate the link between adiponectin levels and Type 2 Diabetes (Dastani *et al.* 2012). Using this methodology, the average effect of AD risk alleles on RNFL thickness could be approximated, allowing the use of meta-analysis summary statistics rather than individual level data as used normally in polygenic risk score calculation. The weighted sum of the individual SNP coefficients leads also to an estimate of the explained variance. The genetic risk score was calculated using results from AD consortium and the RNFL thickness meta-analysis described in this chapter. The risk score was weighted using the AD effects for the 16 independent SNPs, previously

selected to be significant in both AD and RNFL thickness meta-analysis and not in LD. This analysis was performed with the R package ”gtx” (https://cran.r-project.org/web/packages/gtx/index.html)

Plots were produced using basic graph, ggplot2 and routines in R language (http://www.r-project.org/). Ggplot2 is a plotting system for R, which provides a powerful model of graphics (http://ggplot2.org/).

CHAPTER 3

Genome-wide association studies and glaucoma endophenotypes

In this chapter I will present my work aimed to unravel the genetic architecture of POAG endophenotypes using GWAS, meta-analysis of GWAS and gene-set enrichment analysis.

STATEMENT OF CONTRIBUTION TO THIS RESEARCH:

I performed quality control and GWAS for the BMES study for POAG endophenotypes. I was involved in the conceptualisation and experimental design of the project on *ATOH7* gene and I performed all the analysis for the BMES. I am first author of the manuscript that has been published in the British Journal of Ophthalmology (Venturini *et al.* 2014b).

For the IGGC meta-analysis, I partecipated in the meta-analysis and performed all analyses required for the BMES. I took part in revising the manuscript and producing some of the graphs of the original publications in Nature Genetics (see reference (Hysi *et al.* 2014) and (Springelkamp *et al.* 2014)).

I was involved in the experimental design and performed gene-set enrichment analysis for each cohort and meta-analysed the results. Results were previously presented at ARVO meeting 2014.

3.1 Background

Studying an age-related complex disease like POAG presents several challenges. An important strategy for gene discovery in POAG is to dissect the disease into its associated endophenotypes, that are quantitative traits associated with the disease (Cannon & Keller 2006). Optic disc parameters and IOP are considered endophenotypes for POAG, they can be measured on a quantitative scale in large data set of healthy individuals and may be more sensitive measures of salient aspects of POAG pathogenesis.

Optic nerve parameters are important aspects of POAG assessment and treatment. Evaluation of the structural integrity of the optic nerve head is part of the assessment of POAG. Specifically, the optic disc is the location where the axons of the retinal ganglion cells leave the eye and form the optic nerve (O'Neill *et al.* 2009). Glaucomatous optic neuropathy can be recognised by changes in the morphology of the optic disc, caused by loss of retinal ganglion cells. Clinically, the optic nerve can be quantitatively assessed measuring different parameters such as the optic disc size and the cup/disc ratio (CDR). The CDR is a parameter commonly used in clinical glaucoma management (Kwon *et al.* 2009). An unusually large CDR is a significant determinant of glaucoma (Keltner *et al.* 2006) and an increasing CDR may be a sign of glaucoma progression (Hsu *et al.* 2010). On the other hand, the disc size, measured as the area or the diameter of the disc, has a controversial role in POAG assessment. It has been suggested that a larger optic disc may suffer more from IOP-related stress (Bellezza *et al.* 2000). However, it has been argued that a larger optic disc may have more anatomical reserve due to a higher number of nerve fibres (Hoffmann *et al.* 2007), and thus larger optic disc may be protective against POAG.

Elevated IOP is the main risk factor for POAG and the only one that can be modifiable by treatments (Czudowska *et al.* 2010; Kass *et al.* 2002). Currently, IOP-lowering is the only treatment for POAG and has been shown to reduce visual field loss and thus decrease POAG progression (Maier *et al.* 2005).

Several genes have been associated with optic disc measurements and IOP and some of them were also implicated in POAG (as described in Chapter 1). For example, the *ATOH7* gene has been associated with both VCDR and POAG (Burdon *et al.* 2011; Macgregor *et al.* 2010; Ramdas *et al.* 2010, 2011). This gene is expressed in the retina where it controls photoreceptor development and, in addition, it has a key role in ocular embryogenesis (Khor *et al.* 2011). Another example is represented by the *TMCO1* gene which has been associated with both IOP and POAG susceptibility (Burdon *et al.* 2011; van Koolwijk *et al.* 2012).

Despite the encouraging successes with GWAS, we still lack a clear understanding of the biological mechanisms and molecular events controlling POAG endophenotypes and consequently the risk of POAG. The genetic associations reported to date explain only a small percentage of the trait heritability, as GWAS need even larger samples to achieve sufficient power to identify further variants. Using current sample sizes, true positive association signals may be hidden among the random noise of false positive results (Yang *et al.* 2010).

In this chapter, I will present my work, in which I aimed to investigate the genetic architecture of POAG endophenotypes using genome-wide

association studies. GWAS for optic disc measurements and IOP have already led to discoveries about genes and pathways about POAG providing new biological insight. The aim of this chapter was to confirm and identify novel genetic factors for POAG endophenotypes through extensive meta-analyses and interpretation of GWAS results. Specifically the objectives were:

- to identify novel genetic variants affecting POAG endophenotypes through GWAS in the BMES and also to clarify the role of *ATOH7* gene in optic nerve measurements

- to meta-analyse the results from the BMES with several other European and Asian cohorts from IGGC

- to identify functional categories of genes associated with IOP using gene-set enrichment analysis from GWAS results from IGGC

3.2 Methods

Quality control and genome-wide association studies were performed on IOP and VDCR in the BMES cohort as described in Chapter 2.2.2. In brief, average between the two eyes was used as phenotype for both VCDR and IOP and a linear regression was performed adjusted for age and sex.
I performed further analysis in a candidate region including *ATOH7* gene in the BMES and TwinsUK cohorts. *ATOH7* was defined as the open reading frame $\pm$100 kbp. This candidate region covers from 560k bp to 761k bp on chromosome 10 based on NCBI36/hg18 assembly. I performed association

analysis for the *ATOH7* gene region and three different linear regressions were performed:

- the first linear regression had disc size (diameter for the BMES and area for TwinsUK) as output and genotypes as predictor adjusted on age, sex and IOP
- the second linear regression had CDR as output and genotypes as predictor adjusted on age, sex and IOP
- finally the third model had CDR as output, genotypes as predictor and adjustments were done for age, sex, IOP and disc size.

Meta-analysis of GWAS results with European and Asian cohorts were performed as part of the IGGC and more details are described in Chapter 2.2.3. Analysis for VCDR were also adjusted for disc size and those for IOP were adjusted for CCT.
Following IOP meta-analysis, a gene-set enrichment analysis for GWAS for each cohort was carried out. Results were then meta-analysed using Fisher's combined probability method. More details are described in the Chapter 2.2.4).

3.3 Results

3.3.1 Genome-wide association analyses in the BMES for POAG endophenotypes

The first aim of this chapter was to investigate common genetic variants underlying POAG endophenotypes in the BMES cohort.

2,617,903 SNPs were analysed for a total of 1,677 individuals from the BMES data set. The average age in the BMES cohort was 66.2 (SD=9.8, range=49-97) and 57% of them were males.

3.3.1.1 Association analyses for VCDR

The mean VCDR was 0.43 (SD=0.13) and the phenotype was normally distributed (Figure 3.1).

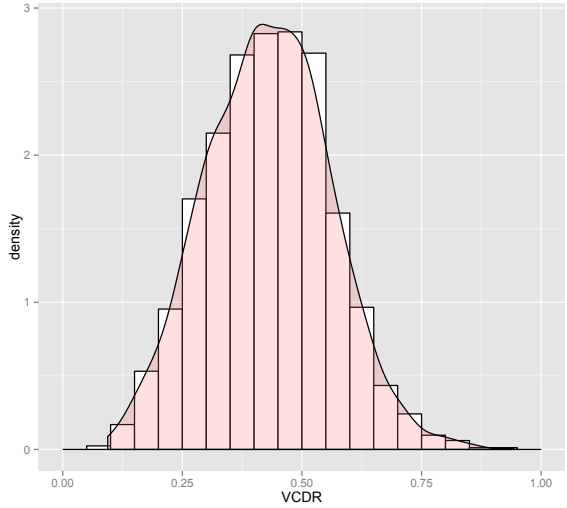


Figure 3.1: Distribution of VCDR in the BMES cohort: the figure shows the histogram and the density curve of the VCDR phenotype.

A GWAS on VCDR in the BMES sample was carried out and the inflation factor α was 1.006, showing the absence of population substructure. Results were plotted in a quantile-quantile plot (Q-Q plot) as shown in Figure 3.2.

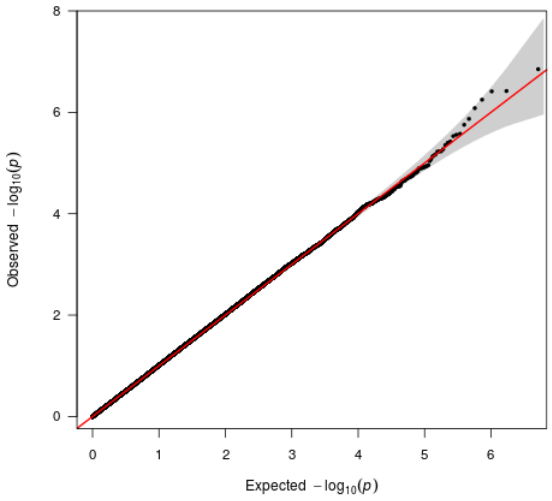


Figure 3.2: Q-Q plot for the GWAS for VCDR in the BMES cohort

No loci passed Bonferroni correction as shown in Figure 3.3. However, a number of loci of suggestive significance were identified in the VCDR GWAS for the BMES data set (Figure 3.3). Several genetic signals were pinpointed on chromosome 6, for example at rs287832 and rs1757824 (respectively β=0.026, SE=0.005, P=3.786E-07 and β=0.026, SE=0.005, P=3.856E-07) which are close to the *TPBG* transcript. This gene encodes a leucine-rich transmembrane glycoprotein that may be involved in cell adhesion. The effect size for these loci was 0.026, which means that each copy of the risk alleles (in this case the minor alleles: A for rs287832 and G for rs1757824) increased VCDR of 0.026 mm. The minor allele was 0.27 for both variants.

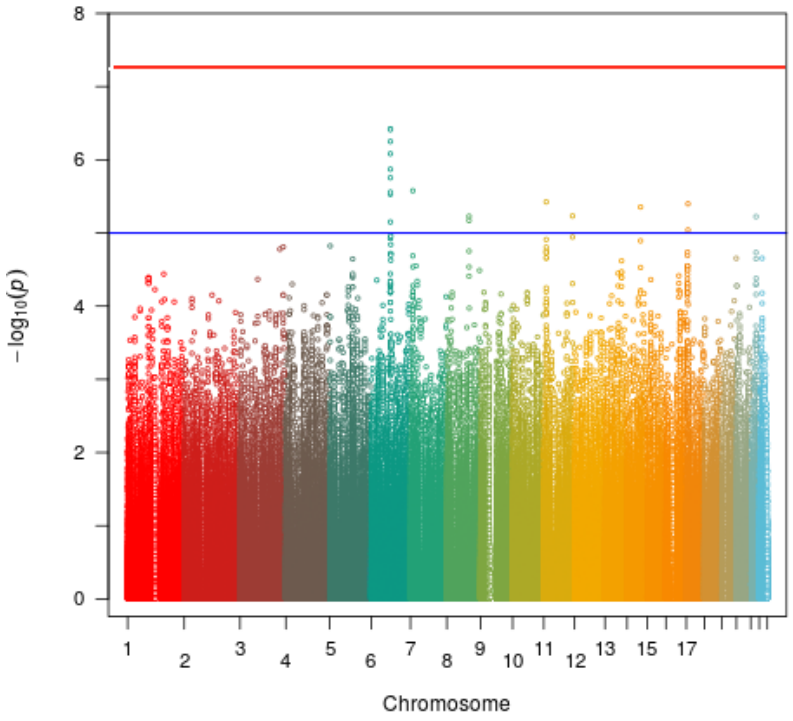


Figure 3.3: Manhattan plot of the results from the GWAS for VCDR in the BMES cohort: the 22 autosomes are plotted along the x-axis, whereas the values in the y-axis denote the $\log 10$ transformed p-values from the GWAS in the BMES cohort with VCDR observed for any of the SNPs

3.3.1.2 Clarifying the role of ATOH7 gene in optic disc measurements

Following the GWAS on VCDR in the BMES cohort, I focused on a candidate region, including *ATOH7* gene (Venturini *et al.* 2014b). A set of 1,677 individuals from the BMES data set was analysed. After adjustments for confounds, a strong signal was found at rs7916697 (Allele=G, β=3.78, SE=0.66, P=1.68E-08) for VDD. In addition, rs7916697 was significant to a lesser degree for VCDR (Allele=G, β=0.02, SE=0.005, P=2.43E-04) adjusted on age, sex and IOP. However, no association signal at this SNP was found when the analysis for VCDR was adjusted on age, sex, IOP and VDD (Allele=G, β=0.007, SE=0.04, P=0.15) as shown in Figure 3.4. Results are shown in Table 3.1.

	A1	Effect BMES	SE BMES	P BMES	Effect TwinsUK	SE TwinsUK	P TwinsUK
Disc size	G	3.779	0.66	1.676E-08	0.15	0.02	1.84E-08
CDR	G	0.019	0.005	2.43E-04	0.01	0.004	0.01
CDR adjusted on disc size	G	0.007	0.004	0.148	0.005	0.004	0.24

Table 3.1: The results of the three association analyses for rs7916697 for the BMES and TwinsUK cohorts: the table shows the results for the association analysis for rs7916697 lying on *ATOH7* gene. Effect, effect size; SE, standard error; P, p-value for the association. This table has been published elsewhere (see reference (Venturini *et al.* 2014b))

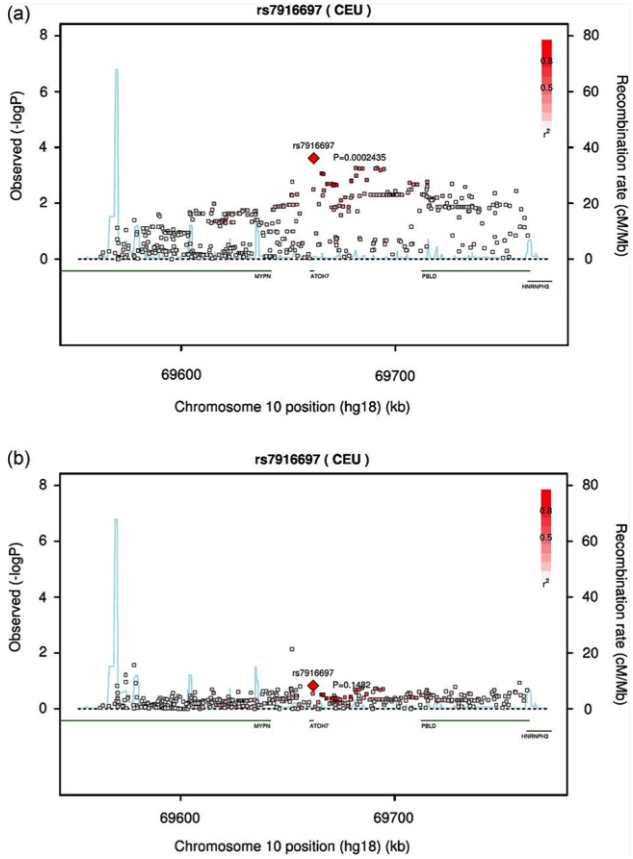


Figure 3.4: Regional association plots for VCDR (A) and VCDR adjusted for disc diameter (B) for the results in the BMES cohort: this figure shows the regional association plots for analysis on VCDR and VCDR adjusted for disc diameter in the BMES cohort for *ATOH7* gene region (Venturini *et al.* 2014b).

The results showed that adjusting for disc size decreased the magnitude of the effect. In the BMES cohort, each additional copy of the susceptibility allele for rs7916697 increased the VCDR by an average of 0.019 units when the analysis was not adjusted on disc size. When VDD was added

as covariate, the increase was only of 0.007 units. Replication analyses were carried out in 1,922 individuals from the TwinsUK cohort. The signal at rs7916697 for disc size was replicated (Allele=G, β=0.15, SE=0.02, P=1.84E-08). The association for this SNP was significant to a lesser degree for CDR (Allele=G, β=0.01, SE=0.004, P=0.01), but was not significant for CDR additionally adjusted on disc size (Allele=G, β=0.005, SE=0.004, P=0.24).

3.3.1.3 Association analyses on IOP

Similar to the VCDR analysis, another GWAS was performed for IOP level. 1,667 individuals from the BMES cohort were analysed with 16.04 mmHg (SD=2.63) average IOP and 66.19 (SD=9.77) mean age (Figure 3.5).

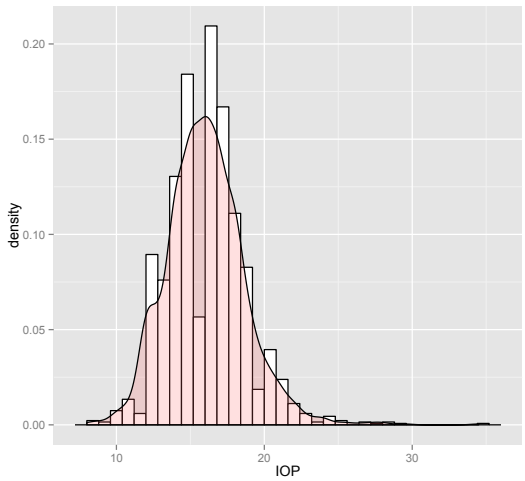


Figure 3.5: Distribution of IOP in the BMES cohort: the figure shows the histogram and the density curve of the VCDR phenotype in mmHg scale.

The genomic inflation factor (α) was calculated and was estimated to be 0.99 for the BMES cohort, suggesting within-study control (Figure 3.6).

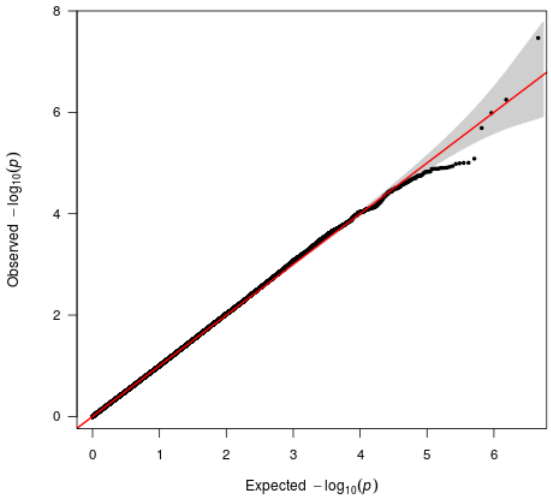


Figure 3.6: Quantile-quantile plot for the GWAS for IOP in the BMES cohort

GWAS results on IOP suggested that several loci were associated with IOP in the BMES cohort (Figure 3.7). One signal crossed the genome-wide significance at rs1571173 on chromosome 20 (β=1.135, SE=0.2, P=3.415E-08). For each copy of the risk allele (A) IOP increased of 1.135 mmHg. This SNP was imputed in the BMES with an high imputation score (0.914) and minor allele frequency of 0.06. The variant lies on an intergenic region and the closest gene is *DOK5*. The protein encoded by this gene is a member of the DOK family of membrane proteins, which are adapter proteins involved in signal transduction. It interacts with phosphorylated receptor tyrosine kinases to mediate neurite outgrowth and activation of the MAP kinase pathway. Near the same gene, another variant was identified with suggestive significance (rs1571172: β=1.052, SE=0.2, P=5.626E-07). Another suggestive genetic locus was found at rs4792947 on chromosome 17 (β=0.571, SE=0.116, P=1.031E-06). This is a genotyped variant (minor allele frequency = 0.25) which lies on an intergenic region. However, this region shows high regulatory activity and it is close to several genes such as *SLC25A39,RNDC3A, SCC4A1, GPATCH8,*

HDAC5, G6PC3.

A further GWAS on IOP level adjusted on CCT measurements was conducted. The results for the BMES cohort did not change after adjusting for CCT. The two loci near to *DOK5* gene did not lose in strength and significance (rs1571173: β=1.127, SE=0.204, P=4.312E-08 and rs1571172: β=1.044, SE=0.209, P=6.856E-07).

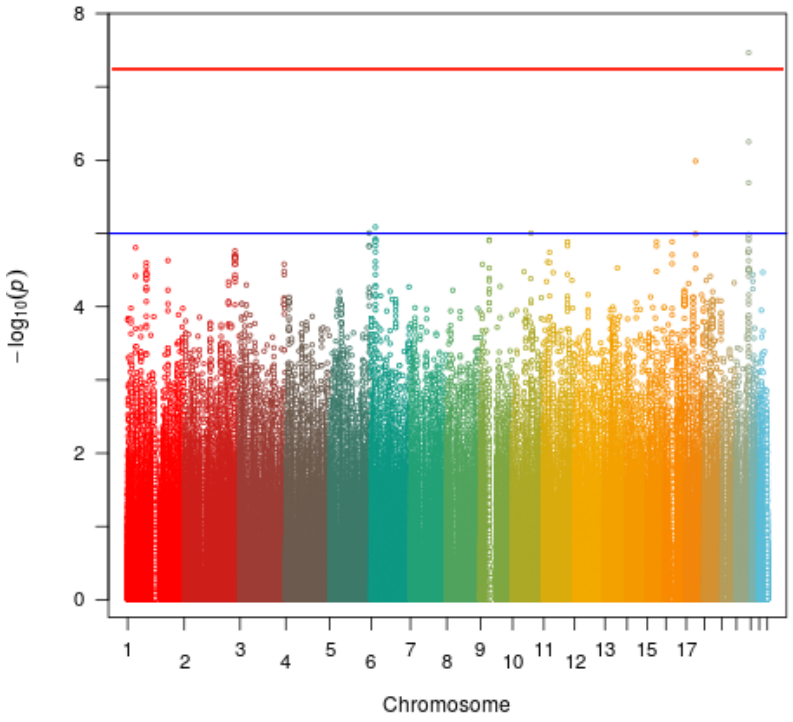


Figure 3.7: Manhattan plot of the results from the GWAS for IOP in the BMES cohort: the 22 autosomes are plotted along the x-axis, whereas the values in the y-axis denote the $\log 10$ transformed p-values from the GWAS in the BMES cohort with IOP observed for any of the SNPs. The blue line represents a suggestive significance (P=1E-05).

3.3.2 Meta-analysis of genome-wide association studies for POAG endophenotypes

A large meta-analysis for POAG endophenotypes with data from several studies from Europe, USA, Australia and Asia, as part of the IGGC (more details on the IGGC, see Chapter 2) was performed to investigate genetic factors underlying POAG.

For VCDR meta-analysis, 21,094 individuals from 10 European-derived populations were used in the meta-analysis. Replication was performed in 6,784 individuals from four Asian cohorts. 2.5 million SNPs imputed to HapMap2 were analysed. More details were described in Chapter 2.2.3. For VCDR meta-analysis, the overall inflation factors (λ) was calculated and was 1.05.

In the meta-analysis of European cohorts, 440 genome-wide significant SNPs across 15 different regions were identified (Table 3.2). Genome-wide significance was assessed using Bonferroni correction.

SNP	CHR	Gene	A1	MAF Euro	Effect Euro	SE Euro	P Euro	MAF Asians	Effect Asians	SE Asians	P Asians	Effect Com	SE Com	P Com	P Het
rs4658101	1	CDC7/TGFBR3	a	0.18	0.015	0.002	8.80E-14	0.14	0.016	0.005	3.13E-03	0.015	0.002	1.06E-15	0.17
rs2623325	3	COL8A1	a	0.09	0.018	0.003	7.05E-09	0.17	0.011	0.005	1.46E-02	0.016	0.003	6.61E-10	0.70
rs17658229	5	DUSP1	c	0.05	-0.02	0.004	8.06E-09	0	-0.086	0.133	5.17E-01	-0.02	0.004	8.06E-09	0.59
rs17756712	6	EXOC2	g	0.17	0.01	0.002	1.98E-08	0.12	0.011	0.005	1.76E-02	0.01	0.002	1.13E-09	0.72
rs7865618	9	CDKN2BAS	g	0.43	-0.013	0.001	2.80E-20	0.15	-0.021	0.005	8.11E-06	-0.013	0.001	4.97E-24	0.70
rs1900005	10	ATOH7	a	0.23	-0.019	0.002	7.21E-31	0.26	-0.01	0.004	2.08E-02	-0.018	0.002	5.51E-31	8.5E-05
rs7072574	10	PLCE1	a	0.33	0.009	0.002	6.17E-09	0.37	0.007	0.003	4.80E-02	0.009	0.001	1.02E-09	0.26
rs1346	11	SSSCA1	t	0.19	-0.014	0.002	2.54E-15	0.16	0.003	0.005	5.23E-01	-0.012	0.002	4.89E-13	0.15
rs4936099	11	ADAMTS8	c	0.42	-0.009	0.002	6.38E-09	0.06	-0.007	0.009	4.15E-01	-0.009	0.002	4.61E-09	0.68
rs11168187	12	RPAP3	g	0.16	-0.009	0.002	2.96E-08	0.18	-0.005	0.004	2.80E-01	-0.009	0.002	2.96E-08	0.99
rs10862688	12	TMTC2	g	0.45	0.008	0.001	1.24E-11	0.55	0.004	0.003	2.48E-01	0.008	0.001	1.49E-11	0.02
rs4901977	14	SIX1/6	t	0.31	0.01	0.002	1.98E-11	0.51	0.017	0.003	2.64E-07	0.011	0.001	2.13E-16	0.20
rs1345467	16	SALL1	g	0.27	0.01	0.002	2.70E-12	0.11	0.011	0.006	5.53E-02	0.01	0.001	4.19E-13	0.25
rs6054374	20	BMP2	t	0.42	-0.009	0.002	1.79E-08	0.73	0.001	0.004	8.66E-01	-0.007	0.001	1.69E-07	0.08
rs1547014	22	CHEK2	t	0.31	-0.013	0.001	2.98E-18	0.15	-0.013	0.004	4.26E-03	-0.013	0.001	4.77E-20	0.39
rs301801	1	RERE	c	0.33	0.008	0.001	1.61E-07	0.13	0.012	0.005	2.59E-02	0.008	0.001	1.66E-08	0.52
rs868153	6	HSF2	g	0.36	-0.007	0.001	5.08E-06	0.38	-0.013	0.003	1.44E-04	-0.007	0.001	1.39E-08	0.79
rs5756813	22	CARD10	g	0.39	0.006	0.001	1.60E-05	0.33	0.017	0.004	1.71E-06	0.008	0.001	7.73E-09	0.20

Table 3.2: Summary of the results of the meta-analysis of genome-wide association studies for VCDR: this table shows the summary of the results of the meta-analysis of genome-wide association studies for VCDR that showed genome-wide significance in 21,094 European and 6,784 Asian individuals. Gene is based on reference NCBI build 37; A1, reference allele; MAF, minor allele frequency; effect, effect size on VCDR based on A1; SE, standard error of the effect size; P het, test for heterogeneous effects between European and Asian samples. This table has been published elsewhere (Springelkamp *et al.* 2014).

Genetic factors identified in the Caucasian meta-analysis were checked for replication in the Asian cohorts. Eight of these genetic loci were nominally significant at $p < 0.05$ and the direction and magnitude of the effects was consistent with the Caucasian sample (Table 3.2). In addition, five of the seven loci that did not reach nominal significance in those of Asian descent had a similar effect in the same direction.

A combined meta-analysis of the European and Asian populations was also performed and resulted in three additional genome-wide-significant loci on chromosomes 1, 6 and 22 (Table 3.2, Figure 3.8). The level of heterogeneity across the samples is shown in Table 3.2, suggesting that for most variants the difference between studies was not significant.

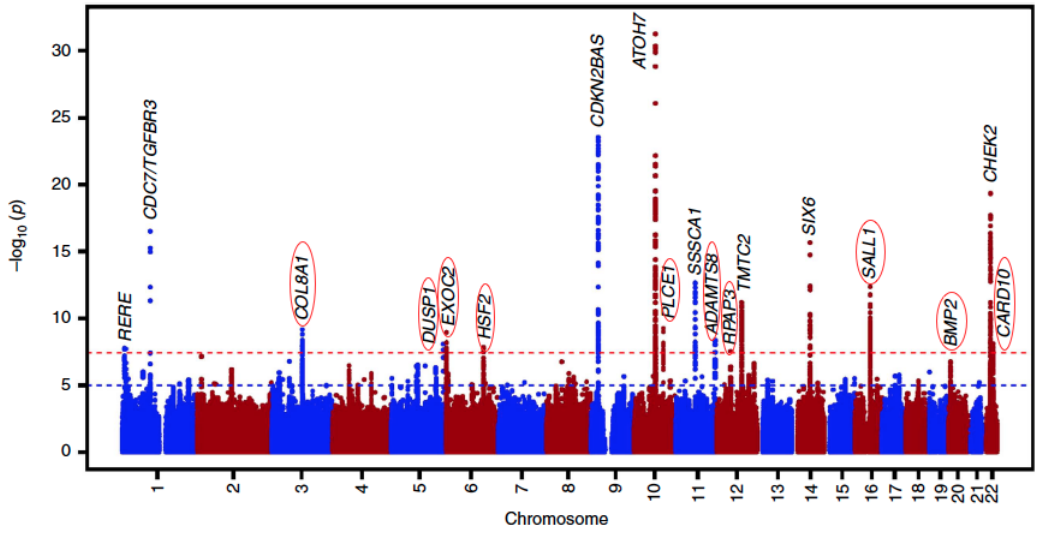


Figure 3.8: This figure has been published elsewhere (see reference (Springelkamp *et al.* 2014)). Manhattan plot of the GWAS meta-analysis for vertical cup-disc ratio in the combined analysis (n=27,878). The plot shows the $-\log 10$ transformed P values for all SNPs. The red-dotted horizontal line represents the genome-wide significance threshold of P<5E-08; the blue-dotted line indicates P value of 1E-05. Red circles indicate novel loci identified for VCDR.

Ten novel loci were identified for VCDR and lies next to/within *COL8A1, DUSP1, EXOC2, PLCE1, ADAMTS8, RPAP3,SALL1, BMP2, HSF2 CARD10* genes. The allele frequency was not different across different cohorts. A number of loci identified for VCDR in the combined meta-analysis have been previously associated with optic disc area: *CDC7/TGFBR3, ATOH7, SALL1 and CARD10* (Ramdas *et al.* 2010; Khor *et al.* 2011). In particular, in the combined meta-analysis rs1900005 was the most significant signal (β=-0.018, SE=0.002, P=5.51E-31) that lies within *ATOH7* gene. The effect size (β) was -0.018, meaning that each copy of A allele decreased VCDR of 0.018 mm in average. The second signal was identified at rs7865618 within *CDKN2BAS* gene. The effect size was -0.013 (SE=0.001) and the p-value was 4.97E-24.

Further analyses included GWAS results adjusted for optic nerve disc

size (Table 3.3). For *CDC7-TGFBR3* locus (rs4658101), adjusting for disc area reduced the significance in the IGGC meta-analysis (p=3.48E-11 to p=9.00E-03). As observed for BMES and TwinsUK cohorts, effect sizes reduced also for the variants lying in *ATOH7* gene. However, in the meta-analysis, there was still some significant disc-area independent effect (P=1.60E-06). For the ten novel loci, there was no change in association significance suggesting that these genetic factors affect primarily CDR and not the disc area.

SNP	CHR	Gene	A1/A2	Effect Unadj	SE Unadj	P Unadj	Effect adj disc size	SE adj disc size	P adj disc size
rs4658101	1	CDC7/TGFBR3	a/g	0.014	0.002	4.26E-13	0.005	0.002	9.04E-03
rs2623325	3	COL8A1	a/c	0.018	0.003	1.50E-10	0.015	0.003	2.70E-08
rs17658229	5	DUSP1	c/t	-0.020	0.004	3.56E-08	-0.017	0.004	8.35E-07
rs17756712	6	EXOC2	g/a	0.010	0.002	6.08E-09	0.008	0.002	3.19E-06
rs7865618	9	CDKN2BAS	g/a	-0.014	0.002	7.04E-21	-0.014	0.001	9.56E-27
rs1900005	10	ATOH7	a/c	-0.018	0.002	4.44E-29	-0.007	0.002	1.60E-06
rs7072574	10	PLCE1	a/g	0.009	0.002	2.98E-10	0.007	0.001	1.38E-06
rs1346	11	SSSCA1	t/a	-0.012	0.002	7.21E-12	-0.010	0.002	2.25E-08
rs4936099	11	ADAMTS8	c/a	-0.009	0.002	4.66E-08	-0.008	0.001	5.93E-09
rs11168187	12	RPAP3	g/a	-0.009	0.002	6.17E-07	-0.008	0.002	2.04E-06
rs10862688	12	TMTC2	g/a	0.007	0.001	4.32E-09	0.005	0.001	6.66E-06
rs4901977	14	SIX1/6	t/c	0.012	0.002	7.56E-17	0.014	0.001	8.89E-23
rs1345467	16	SALL1	g/a	0.010	0.002	1.44E-10	0.007	0.002	2.35E-05
rs6054374	20	BMP2	t/c	-0.007	0.002	2.20E-06	-0.007	0.001	5.65E-09
rs1547014	22	CHEK2	t/c	-0.013	0.002	1.15E-17	-0.011	0.002	3.23E-13
rs301801	1	RERE	c/t	0.008	0.002	6.42E-07	0.008	0.002	1.49E-07
rs868153	6	HSF2	g/t	-0.008	0.002	2.24E-07	-0.007	0.001	2.76E-06
rs5756813	22	CARD10	g/t	0.009	0.002	3.32E-09	0.007	0.001	2.62E-09

Table 3.3: SNPs that showed genome-wide significant (P<5E-08) association with VCDR in subjects of European and Asian ancestry combined, adjusted for disc size: the table shows results for the meta-analysis on VCDR asjusted for disc size. Only individuals with both information on VCDR and disc size were included in this comparison. A1 = reference allele; A2 = other allele; Effect = effect size on VCDR based on allele A1; Chr= chromosome; SE = standard error

Likewise, a meta-analysis for IOP as part of the IGGC was conducted. The IOP meta-analysis included 35,296 individuals whose 27,558 were of European ancestry and 7,738 of Asian ancestry. 2.5 million SNPs imputed

to HapMap2 were analysed.

The results for each cohort showed that the inflation factors α ranged between 0.992 and 1.043. There were 145 SNPs that crossed conventional genome-wide significance threshold (p<5E-08) (Dudbridge & Gusnanto 2008)(Table 3.4). All of these SNPs clustered around seven separate regions of the genome (Figure 3.9). The most significant signal was found at rs9913911 (β=-0.179, SE=0.026, P=1.03E-11), lying within *GAS7* gene. G allele was protective, leading to a decrease of 0.179 mmHg of IOP level. *GAS7* gene has been previously identified for IOP (van Koolwijk *et al.* 2012). Two variants were identified at rs4656461 and rs7555523 (respectively β=0.228, SE=0.039, P=6.51E-09 and β=0.235, SE=0.039, P=2.19E-09), lying within *TMCO1* gene. This gene has also been previously associated with IOP (Burdon *et al.* 2011; van Koolwijk *et al.* 2012). There was also an association for IOP level with variants within *CAV1* and *CAV2* genes (P=1.87E-11 for rs10258482) which had previously been associated with POAG (Thorleifsson *et al.* 2010). The region on chromosome 11 was novel for IOP and included several genes such as *AGBL2, SPI1* and *PTPRJ* (best p=1.04E-11 for rs747782). Other novel loci on chromosome 9 included *ABCA1* gene (P=2.80E-11 for rs2472493) and *ABO* (P=3.08E-11 for rs8176743). Another novel signal was found at rs6445055 on chromosome 3 within *FNDC3B* gene (P=4.19E-08).

CHR	BP	SNP	A1/A2	Gene	Effect	SE	P	P Heterogeneity	I2
1	165687205	rs4656461	G/A	TMCO1	0.228	0.039	6.51E-09	0.46	0
1	165718979	rs7555523	C/A	TMCO1	0.235	0.039	2.19E-09	0.55	0
3	171992387	rs6445055	A/G	FNDC3B	-0.177	0.03	4.19E-08	0.17	0.24
7	116150095	rs10258482	A/C	CAV1	0.196	0.029	1.87E-11	0.81	0
7	116150952	rs10262524	A/C	CAV1	0.186	0.029	9.69E-11	0.67	0
9	107695848	rs2472493	G/A	ABCA1	0.159	0.024	2.80E-11	4E-05	0.66
9	136131415	rs8176743	T/C	ABO	0.261	0.039	3.08E-11	0.53	0
11	47468545	rs12419342	C/T	RAPSN	0.153	0.026	4.77E-09	0.75	0
11	47940925	rs747782	C/T	NUP160	PTPRJ	0.203	0.03	1.04E-11	0.95
11	47969152	rs1681630	T/C	PTPRJ	0.144	0.026	1.69E-08	0.6	0
11	48004369	rs7946766	T/C	PTPRJ	0.23	0.035	2.71E-11	0.35	0.09
17	10031183	rs9913911	G/A	GAS7	-0.179	0.026	1.03E-11	4E-04	0.61

Table 3.4: Results for the IGGC meta-analysis for IOP from the general population cohorts: this table shows the results for the association with IOP from the general population cohorts for SNPs significant at a multiple testing correction level ($p<5E-08$). This table has been published elsewhere (Hysi *et al.* 2014).

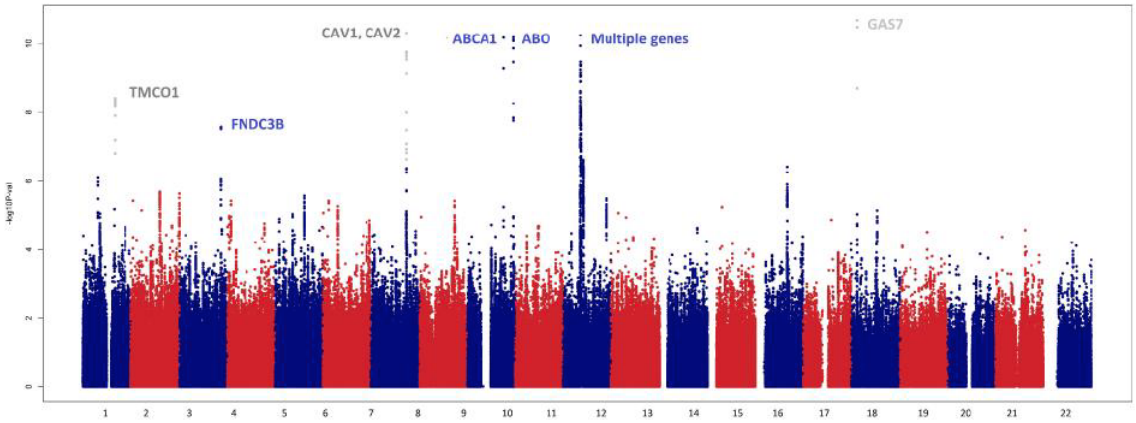


Figure 3.9: This figure has been published elsewhere (see reference (Hysi *et al.* 2014)). Manhattan plot of the results from the meta-analyses of results from 18 multi-ethnic cohorts from the IGGC. The 22 autosomes are plotted along the x-axis, whereas the values in the y-axis denote the $\log 10$ transformed p-values from the meta-analysis of association with IOP observed for any of the SNPs. Loci previously associated with IOP are shown in grey.

I also investigated whether the genetic signal near to *DOK5* gene identified in IOP GWAS in the BMES cohort was confirmed in the IGGC meta-analysis.

Even though rs1571173 did not reach genome-wide significance in this larger sample, it showed nominal significance (β=0.09, SE=0.04, P=0.042).

Following the meta-analysis, the effect on the clinical outcome of the SNPs identified of POAG was investigated in four independent cohorts. These included 4,284 POAG cases and 95,560 controls. Novel associations to IOP were also confirmed in POAG for variants lying on *ABCA1* gene (P=4.15E-09 for rs2472493), *FNDC3B* gene (P=0.03 for rs6445055) and on the chromosome 11 cluster (p=0.008 for rs12419342). Association for known loci for POAG was also confirmed: *TMCO1* (P=1.34E-16 for rs7555523), *CAV1/CAV2* (P=6.27E-09 for rs10258482) and *GAS7* (P=5.22E-13 for rs12150284). All alleles associated with an increase of IOP were also associated with an increased risk of POAG (see Table 3.5). The only gene that was not associated with POAG susceptibility was *ABO*.

CHR	BP	SNP	A1/A2	Gene	OR	95% CIs	95% CIs	P
1	165687205	rs4656461	G/A	TMCO1	1.38	1.28	1.5	2.55E-15
1	165718979	rs7555523	C/A	TMCO1	1.4	1.3	1.52	1.34E-16
3	171992387	rs6445055	A/G	FNDC3B	0.92	0.85	0.99	0.03
7	116150095	rs10258482	A/C	CAV1	1.2	1.13	1.28	6.27E-09
7	116150952	rs10262524	A/C	CAV1	1.2	1.13	1.28	1.39E-08
9	107695848	rs2472493	G/A	ABCA1	1.24	1.16	1.34	4.15E-09
9	136131415	rs8176743	T/C	ABO	1.07	0.96	1.19	0.2
11	47468545	rs12419342	C/T	RAPSN	1.09	1.02	1.16	0.008
11	47940925	rs747782	C/T	NUP160/PTPRJ	1.03	0.96	1.11	0.36
11	47969152	rs1681630	T/C	PTPRJ	1.06	0.99	1.12	0.08
11	48004369	rs7946766	T/C	PTPRJ	1.03	0.95	1.12	0.43
17	10031183	rs9913911	G/A	GAS7	0.8	0.75	0.85	2.98E-13

Table 3.5: Results for the association of IOP identified SNPs with POAG in a case-control validation meta-analyses: this table shows the results for POAG case-control meta-analysis for the SNPs identified in the IOP meta-analysis. CHR: chromosome; BP: base-pair position; A1: effect allele; A2: other allele; Gene: closest gene/transcript; OR: odds ratio; 95% CIs: 95% confidence intervals; P= p-value of the meta-analysis. This table has been published somwehere else (Hysi *et al.* 2014).

Following this analysis, a linear regression was performed to investigate whether effect sizes for IOP meta-analysis were linearly related to effect sizes on POAG case-control analysis (Do *et al.* 2013). The findings suggested that the strength on SNP effect on IOP levels was related with their effects on the risk of POAG (β=0.788, SE=0.264, P=0.031).

Gene expression for IOP identified loci was investigated (as described in Chapter 2.2.3.1) in 856 mono and dyzigotic twins from the TwinsUK cohort (Table 3.6).

CHR	BP	SNP	A1/A2	Gene	Adipose	LCL	Skin	Probe	Gene
1	165687205	rs4656461	G/A	TMCO1	0.004	0.12	0.003	1793829	TMCO1
1	165718979	rs7555523	C/A	TMCO1	0.39	0.0001	0.05	1761804	ALDH9A1
3	171992387	rs6445055	A/G	FNDC3B	-	-	-	-	
7	116150095	rs10258482	A/C	CAV1	-	-	-	-	-
7	116150952	rs10262524	A/C	CAV1	5.79E-16	8.54E-05	3.91E-13	1687583	CAV1
9	107695848	rs2472493	G/A	ABCA1	0.19	3.67E-05	0.36	1766054	ABCA1
9	136131415	rs8176743	T/C	ABO	-	-	-	-	-
11	47468545	rs12419342	C/T	RAPSN	0.002	4.32E-08	0.0003	1696463	SPI1
11	47940925	rs747782	C/T	NUP160	0.36	-	-	-	-
11	47969152	rs1681630	T/C	PTPRJ	0.006	2.72E-10	0.002	1696463	SPI1
11	48004369	rs7946766	T/C	PTPRJ	0.66	2.02E-05	0.0066	1688627	AGBL2
17	10031183	rs9913911	G/A	GAS7	-	-	-	-	-

Table 3.6: eQTL effects observed for the IOP associated SNPs in three different tissues: eQTL effects observed for the IOP associated SNPs in three different tissues (adipose, LCL and skin) from 849 individuals from the TwinsUK cohort. This table has been published elsewhere (Hysi *et al.* 2014).

Significant eQTL association was observed for rs4656461 and rs7555523 (P=0.003 and P=0.0001 with *TMCO1* and *ALDH9A1* transcript expressions in skin and LCLs respectively), rs2024211 (P=5.43E-16 and 3.84E-13 with *CAV1* in adipose and skin tissues respectively), rs2472493 with *ABCA1* (P=3.67E-05 in LCLs) and rs1681630 with the *SPI1* expression on chromosome 11 (P=2.72E-10 in LCLs) (Table 3.6).

CCT is known to be related with IOP, therefore, additional analyses were performed for IOP adjusted on age, sex and also CCT measurements (Tonnu *et al.* 2005). The results were similar in the meta-analysis with 19,563 individuals with CCT measurements. Rs6445055 within *FNDC3B* gene was previously associated with CCT (Lu *et al.* 2013), however it maintained nominal significance for IOP after adjustment (P= 9.87E-04, β= -0.121

compared to -0.177 prior to adjustment for CCT), suggesting that this locus has some independent effect on CCT.

3.3.3 Common mechanisms underlying IOP identified in functional analysis of gene lists from GWAS results

Following IOP GWAS and meta-analysis, the results for each GWAS from the IGGC were annotated using Ensembl as described in Chapter 2.2.4. After annotation the number of genes from each annotation did not exceed 2,000 unique genes or transcripts. Gene set enrichment analysis was carried out for each cohort and meta-analysed the results. I focused on Gene Ontology (GO) database, which describes gene products in three structured ontologies in terms of their associated biological processes, cellular components and molecular functions. For this analysis, a total of 14,721 entries from the GO database were investigated: 9,818 biological processes, 1,208 cellular component entries, and 3,695 molecular functions. In the discovery cohorts of European descent, several GO entries that were highly significantly enriched across all cohorts were identified. There was a remarkable overlap of results between all data sets, both in term of strong probabilities for significant enrichment and highly correlated values of fold enrichment for the respective terms. Meta-analysed results for each GO category are shown in Table 3.7, Table 3.8 and Table 3.9.

GO Term	GO ID	European meta-analysis p-value	Asian meta-analysis p-value
biological adhesion	GO:0022610	1.91E-20	1.39E-04
cell adhesion	GO:0007155	1.43E-19	1.47E-04
cell-cell adhesion	GO:0016337	1.03E-11	N.S
neuron development	GO:0048666	4.37E-11	0.01
neuron differentiation	GO:0030182	7.09E-10	0.001
cell projection organization	GO:0030030	3.42E-09	N.S.
cell motion	GO:0006928	9.34E-09	3.26E-04
homophilic cell adhesion	GO:0007156	1.5x10-08	1.47x10-04

Table 3.7: Gene-set enrichment meta-analysis results: Gene ontology (GO) biological process: this table shows the results of the gene-set enrichment meta-analysis for biological processes for both European and Asian cohorts. Go Term: gene ontology term; GO ID: gene ontology identification number

GO Term	GO ID	European meta-analysis p-value	Asian meta-analysis p-value
plasma membrane part	GO:0044459	9.76E-18	2.33E-08
plasma membrane	GO:0005886	2.32E-16	7.65E-05
cell junction	GO:0030054	2.68E-15	1.9E-05
synapse	GO:0045202	9.14E-10	3.34E-05
cell projection	GO:0042995	1.42E-08	0.02

Table 3.8: Gene-set enrichment meta-analysis results: Gene ontology (GO) cellular component: this table shows the results of the gene-set enrichment meta-analysis for cellular components for both European and Asian cohorts. Go Term: gene ontology term; GO ID: gene ontology identification number

Term	GO ID	European meta-analysis p-value	Asian meta-analysis p-value
calcium ion binding	GO:0005509	6.56x10-20	1.23x10-04

Table 3.9: Gene-set enrichment meta-analysis results: Gene ontology (GO) molecular function: this table shows the results of the gene-set enrichment meta-analysis for molecular functions for both European and Asian cohorts. Go Term: gene ontology term; GO ID: gene ontology identification number

Across the different GO categories, strong suggestions for association between adhesion pathways and IOP were found in the discovery cohorts of European descent. Among the most associated GO entries were biological

adhesion (P= 1.91E-20), cell adhesion (P= 1.44E-19) and plasma membrane (P= 9.76E-18), cell junction (P= 2.68E-15), and calcium ion binding (P= 6.57E-20). Most of these results suggest the importance of cellular adhesion and calcium channels in IOP regulation. GO entries are not independent between each other and a simplified scheme of the associated GO entries taking into account their dependencies is shown in Figure 3.10

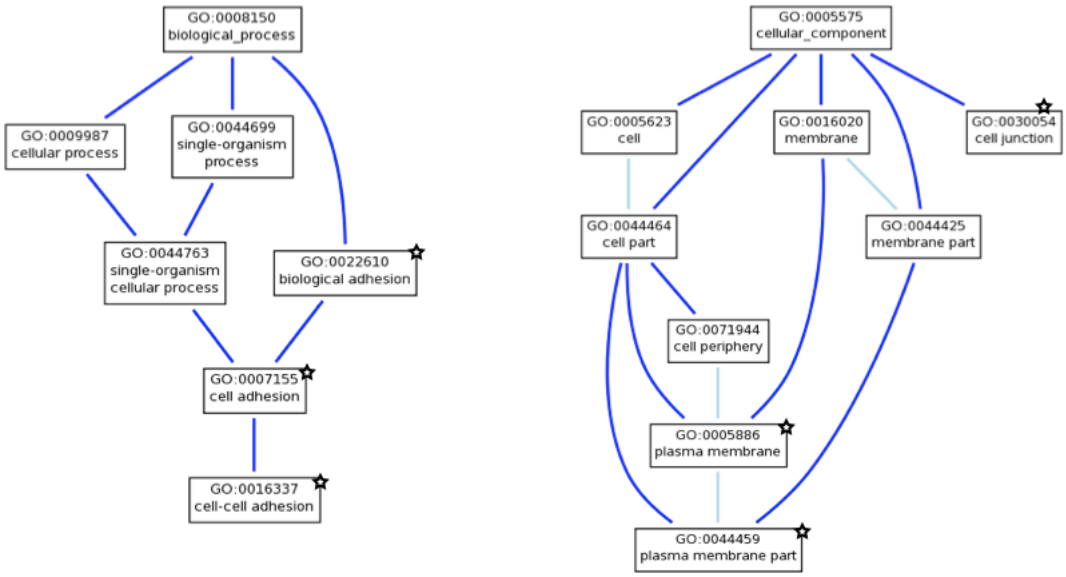


Figure 3.10: Scheme of the main Gene Ontology entries and dependencies significantly enriched in the meta-analysis: the figure shows the GO entries identified in the meta-analysis (marked with a star) and how they are related between each other.

The importance of cellular adhesion was confirmed in the Asian replication populations. Biological and cell adhesion showed significant enrichment in the four Asians cohorts (P= 1.39E-04 and P= 1.47E-04 respectively). In addition, plasma membrane (P= 2.34E-08), cell junction (P= 2.68E-15) and calcium ion binding (P=1.23E-04) were also replicated in the Asian cohorts, suggesting a role of cellular adhesion in plasma membrane and calcium channels in IOP regulation.

3.4 Discussion

In this chapter I investigated the genetic architecture of POAG endophenotypes, in particular VCDR and IOP, in order to better understand the mechanisms underlying POAG.

Firstly, I investigated common genetic variants underlying POAG endophenotypes in the BMES cohort. For VCDR, no loci crossed genome-wide significance, however a few suggestive regions were identified. A region on chromosome 6 near to *TPBG* gene was identified, encoding a leucine-rich transmembrane glycoprotein that may be involved in cell adhesion (Murakami *et al.* 2010). Another signal significant for VCDR was a locus within *ATOH7*, a known gene for disc size (Ramdas *et al.* 2010; Macgregor *et al.* 2010; Khor *et al.* 2011). The role *ATOH7* in optic disc phenotypes variation was investigated further. The association for a locus within *ATOH7* gene was strong for disc size, however it was totally lost for VDCR adjusted on disc size. Therefore, these results suggest that *ATOH7* gene has a key role in disc size variation, however its association with CDR depends on the correlation between the two measurements. *ATOH7* gene encodes a member of the basic helix-loop-helix family of the transcription factors with similarity to Drosophila atonal gene that controls photoreceptor development. Studies in mice suggest that this gene plays a central role in retinal ganglion cell and optic nerve formation (Mao *et al.* 2011). CDR is a well known POAG endophenotype, whereas the relationship between the disease and disc size is weak and not well understood. Larger optic discs may suffer more from IOP-related stress, however the size of the disc affects the likelihood of clinical diagnosis of glaucoma, thus providing a source of bias (Crowston *et al.* 2004).

For IOP in the BMES cohort, I identified a significant region nearby *DOK5* gene. The protein encoded by this gene is a member of the DOK family of membrane proteins, which are adapter proteins involved in signal transduction. The encoded protein interacts with phosphorylated receptor tyrosine kinases to mediate neurite outgrowth and activation of the MAP kinase pathway. In addition, it may have a role in apoptosis (Pan *et al.* 2013) and cardiomyocyte differentiation (Wen *et al.* 2009). The gene has been also associated with obesity and type-2 diabetes (Tabassum *et al.* 2010).

Larger sample sizes are needed to increase the power to detect genetic variants underlying POAG endophenotypes. For this purpose, the BMES GWAS results were meta-analysed with several European and Asian populations from the IGGC. For VCDR meta-analysis, 18 loci were associated with VCDR and ten of them were novel. Most of them were associated independently of the disc size. The results suggest a role of cellular stress response. For instance, *DUSP1* expression is induced by oxidative/heat stress and growth factor and it is involved in oxidative stress and environmental stress response (Keyse & Emslie 1992). It may also play a role in negative regulation of cellular proliferation and in the immune response (Cáceres *et al.* 2013). In addition, also *HSF2* is part of the cellular stress response pathway. This gene is part of the HSF family of transcription factors that bind specifically to the heat-shock promoter element and activate transcription. Heat shock transcription factors activate heat-shock response genes under conditions of heat or other stresses and its role is crucial in the maintenance of neural survival, also in retinal ganglion cells (Kwong *et al.* 2006). It is also involved in

neuronal migration (El Fatimy *et al.* 2014). Another process involved in CDR variation is exocytosis. *EXOC2* gene encodes a protein part of the exocyst complex, which is involved in directing exocytic vesicles to the plasma membrane (Lipschutz & Mostov 2002). This mechanism is also important for neuronal degeneration (Coleman & Yao 2003). *PLCE1* is an interesting gene for VCDR as it encodes for a phospholipase enzyme that may affect cell growth and differentiation and cancer (Wang *et al.* 2010; Wang *et al.* 2014; Abnet *et al.* 2010).

Seven regions were identified through IGGC meta-analysis for IOP. A signal within *ABCA1* gene were identified. This gene encodes for a membrane-associated protein, member of the super family of ATP-binding cassette (ABC) trasnporters and it functions as a cholesteral efflux pump in the cellular lipid removal pathway. The gene has been previously associated with Alzheimer's disease (Furney *et al.* 2011). In addition, *ABCA1* gene is expressed in many tissues (Denis *et al.* 2003) and its expression in leukocytes is significantly up-regulated in glaucoma patients (Yeghiazaryan *et al.* 2005).The meta-analysis pinpointed a number of SNPs within *ABO* blood group gene and IOP. However, this association was not confirmed in the POAG case-control meta-analysis and this might be due to type I error in the first meta-analysis or insufficient power to detect IOP SNPs in the case-control meta-analysis which also includes NTG patients. The findings showed an association between IOP and variants within a large region on chromosome 11 and eQTL analysis underlined *SPI1* and *AGBL2* as possible interesting candidates. *SPI1* gene encodes an ETS-domain transcription factor that activates gene expression during myeloid and B-lymphoid cell development. The protein encoded by *SPI1* gene binds to a purine-rich sequence known as the PU-box found near the promoters

of target genes, and regulates their expression in coordination with other transcription factors and co-factors. The regulation of the transcription factor is disrupted in leukemia (Prange *et al.* 2014) and it may also be involved in viability and function of human brain microglia (Smith *et al.* 2013). *AGBL2* gene is a latexin-interacting protein that regulates tubulin tyrosination cycle and it may be involved in cell proliferation in gastric (Zhu *et al.* 2014) and breast cancer (Zhang *et al.* 2014).

Interestingly, the signal for *DOK5* identified in the BMES cohort for IOP was replicated, albeit at nominal significance, in the IGGC meta-analysis, showing that apoptosis might be involved in IOP regulation.

eQTL analyses pinpointed *ALDH9A1* gene as a possible candidate for IOP, given its location downstream of the *TMCO1*-associated variant. *ALDH9A1* encodes a protein that belongs to the aldehyde dehydrogenase family of proteins is expressed in the ciliary body (Janssen *et al.* 2012).

After the meta-analysis on IOP level, I investigated potential biological categories affecting the IOP control among general populations through gene-set enrichment analysis. Aqueous humour outflow pathway is a key determinant of IOP. Fluid flow in the anterior chamber maintains IOP and supplies nutrients to the non-vascularised cornea, lens, and trabecular meshwork (TM) (Tian *et al.* 2000). I identified calcium ion binding as a significantly enriched GO term in the molecular process category in both European and Asian samples. The effect of calcium on IOP has been studied previously and, in particular, several investigators have focused on how calcium channel blockers (CCBs) affect IOP in humans and animals. CCBs have a wide spectrum of action and are primarily used for the treatment

of hypertension and angina. High levels of systolic blood pressure are known to be associated with higher IOP in several studies (Klein & Klein 1981). One of the first studies investigating the association of CCBs with IOP conducted by Abelson and colleagues (Abelson & Smith 1988) found that local application of verapamil, an L-type calcium blocker, significantly decreased the IOP in both eyes in human participants. In addition, many studies have reported the effect of CCBs on IOP after topical administration, but not with systemic administration (Mayama 2014). However, the extent of the IOP drop after topical CCBs seems to be very small, and has been reported to be approximately 2 mmHg after topical verapamil in humans (Abreu *et al.* 1998). CCBs may affect IOP through several mechanisms. For instance, they may alter the aqueous humour inflow by interfering with the gap junctions between the pigmented and pigmented ciliary epithelium (Abelson & Smith 1988) and cation transport in the non pigmented ciliary epithelial cells (Mito *et al.* 1993). In addition, it has been shown that human and bovine trabecular meshwork cells express voltage activated L-type calcium channels (Steinhausen *et al.* 2000). Finally, CCBs may increase the conventional outflow facility via trabecular meshwork (Erickson *et al.* 1995).

I found several GO terms related to biological and cellular adhesion, gap junction and plasma membrane enriched for IOP. Adhesion molecules, such as integrins, selectins and the immunoglobin super family, play a key role in blood pressure and in particular in the vascular tone, i.e. the contractile activity of vascular smooth muscle cells (SMCs) in the walls of small arteries and arterioles. For their role in the vascular tone, adhesion molecules are relevant not only for hypertension, but also for a number of artery disease processes (i.e. atherosclerosis) (Jang *et al.*

1994). Cell adhesion molecules exert their activity mostly in cell-cell interaction and cell-extracellular matrix interaction. High BP increases stress on the endothelium, and may result in endothelial cell activation or even damage. Several studies have reported increased level of adhesion molecules, including sVCAM-1, sICAM-1 and sE-selectin in the presence of hypertension. Increased expression of endothelial cell adhesion molecules may contribute to the development of hypertension due to failure of the endothelium in regulating vascular tone (Schram & Stehouwer 2005). The physiology of the correlation between BP and IOP is still speculative and several mechanisms have been proposed (Klein & Klein 1981). Bill *et al.* demonstrated that variation in systolic BP resulted in small changes in aqueous humour flow formation, possibly related to increased capillary pressure in the ciliary body (Bill 1973). Blood pressure may also affect the episcleral venous pressure, which is important in regulating the aqueous humour outflow across the TM and the Schlemm's canal (Bill 1975; Klein *et al.* 2005). Given this correlation, the processes that affect vascular tone, such as cellular adhesion, are likely to play a role in aqueous humour outflow facility and thus in regulating IOP. In fact, a number of factors such as cell shape, volume, contractility and adhesion to neighbouring cells and to the extracellular matrix (ECM) could affect resistance by altering the dimensions or direction of flow pathways (Tian *et al.* 2000). Alteration of the actin cytoskeleton, which is responsible for cell adhesion, might play a role in modifying these factors and influence outflow facility within the TM.

GO terms are not independent and they can overlap and show different aspect of the same process. It is well known that Ca2+ is involved nearly in every aspect of cellular life, including muscle contraction, cell shape, and adhesion (Clapham 2007). In particular, Ca2+ affects cadherins and

integrins, families of molecules critical for cell adhesion (Bamji 2005).

Genetic studies provide evidence of the involvement of cellular adhesion and calcium channel activity in IOP regulation. Genes encoding proteins involved in adhesion are of key importance for the maintenance of the TM architecture and regulation of outflow. RNA expression studies in TM showed that proteins involved in adherents and gap junctions, which permit the mechanical attachments between adjacent cells and are important in biological adhesion, are present in TM tissue (Borrás 2003). In addition, genes previously associated with IOP seem to be involved in adhesion processes and calcium channel activity. *GAS7*, for example, is involved in cell remodelling and in the reorganization of microfilaments, which are a part of the cytoskeleton (Wentz-Hunter *et al.* 2004). *MYOC*, encoding for myocilin, a protein found in the TM extracellular matrix tissue, has been previously associated with POAG. Calcium ions are important biological factors that have several roles in extracellular matrix proteins and recently a novel, high affinity Ca2+ binding site within the OLF domain of myocilin has been identified (Donegan *et al.* 2012). *TMC01*, another gene associated with IOP and POAG, is highly expressed in the TM (Liton *et al.* 2006). *CAV1/CAV2* have been previously associated with POAG and also with IOP in the IGGC meta-analysis. Caveolae are highly abundant in mammalian cells and they form stable membrane domains in the plasma membrane and respond to plasma membrane changes. Caveolae can be found in smooth-muscle cells, fibroblasts, endothelial cells and adipocytes and are implicated in a number of signal transduction events, including calcium signalling. Caveolae also seem to have a role in the regulation of cell adhesion, focal contacts and extracellular matrix interactions (Parton & Simons 2007). Caveolins from genetically modified

mice also suggest that they may have a role in cardiovascular functions, including endothelial barrier function, regulation of nitric oxide synthesis, cholesterol metabolism, and cardiac function (Gratton *et al.* 2004).

3.5 Conclusion

In summary, I investigated the genetic architecture of VCDR and IOP through association studies. I clarified the role of *ATOH7* gene in VCDR, showing that its main effect is on disc size. Through IGGC meta-analysis, ten novel loci were identified for VCDR that underlie the role of neuronal cell growth and cellular response to stress as key mechanisms in determining VCDR variation and possibly POAG. IGGC meta-analysis also identified a number of novel loci involved in IOP regulation and POAG susceptibility. These genes are involved in biological and cellular adhesion in the plasma membrane and cell junctions and also to calcium ion binding. Identifying genes and mechanisms underlying IOP and VCDR variation will have an impact on the future investigation of potential therapy for POAG.

CHAPTER 4

Quantitative traits underlying POAG and combination of phenotypic information in genetic analysis

In this chapter I will describe and discusses the relationship amongst POAG quantitative traits and whether and how we could combine them to better understand the genetic architecture of POAG.

STATEMENT OF CONTRIBUTION TO THIS RESEARCH:

I designed this study and I performed all the analysis described in this chapter.

4.1 Background

Genome-wide association studies are a common tool in searching for genetic factors underlying human traits and conditions (Hindorff *et al.* 2009). Often a complex disease is studied through analyses of its underlying quantitative traits, especially in cases where the disease represents only the extremes of a smoothly distributed trait, qualitatively present in the general population, or when studying quantitative traits correlated with

the presence of the disease. In addition, quantitative traits studies are useful where recruitment of a large number of affected cases is impractical or extremely difficult (such as late onset diseases).

To date, most genetic studies have focused on single traits. Single-phenotype association studies had an enormous success and they identified thousands of genetic variants associated with a large number of traits and conditions. However, large-scale clinical and epidemiological studies often include data collection for multiple traits. For example studies on POAG include measurements on the same individuals for IOP, optic nerve head measurements and other ocular and non ocular traits. POAG-related traits taken alone are only imperfect predictors of the disease status. For instance, IOP is considered the main risk factor and often chosen as proxy in genetic analysis, however IOP is not even part of the criteria used for clinical diagnosis. Moreover, the signs of glaucomatous optic nerve damage may occur at any pressure level and only 5-10% of individuals with high pressure have such damage (Leske 1983). In addition, from a statistical point of view, POAG traits are correlated between each other and this correlation is not taken into consideration in single traits analysis and this might lead to loss of information and thus power in GWAS.

Many methods have been proposed to test the association between a genetic variant and multiple phenotypes. Firstly, it is possible to implement a polygenic score approach to assess whether pleiotropy exists between two phenotypes without pointing to any particular variants. A polygenic approaches uses all or a large proportion of SNPs genome wide to establish genetic overlap between two phenotypes. Secondly, multivariate analyses

jointly analyse more than one trait in a unified framework and test for the association of multiple phenotypes with a genetic variant. The aim of these analyses is to identify loci associate with multiple traits and these associations are termed cross-phenotype (CP) associations (Solovieff *et al.* 2013). These analyses are well suited for studies in which subjects are phenotyped across various traits since multivariate methods require that all phenotypes are available on the same individual. This also allows the investigation of the correlations between the traits themselves. Numerous approaches have been proposed for testing the association between a genetic variant and multiple phenotypes, especially for correlated traits.Some approaches include a dimension reduction techniques on the phenotypes before testing the association with the genetic variant. For example principal component analysis and principal factor analysis extracts linear combinations of the phenotypes (i.e. principal components or factors) that explain the most variance in the data that can be used as new phenotypes in a GWAS. Other approaches, for continuous phenotypes, are based on a multivariate regression framework, for example a multivariate analysis of variance (MANOVA). This model is used when there are two or more dependent variables and it helps to understand whether a change in the independent variable (i.e. genotypes) have a significant effect on the dependent variables (Solovieff *et al.* 2013). Other methods extend the regression framework, using variations of the generalised estimating equations (GEE), which are generalised linear models that account for a possible unknown correlation between outcomes.

Advantages of performing combined analysis of correlated traits include increased power to detect loci and increased precision of parameter

estimation (Jiang & Zeng 1995; Zhu & Zhang 2009). Moreover, combined analyses will allow to explore pleiotropy (one locus influencing multiple correlated traits) as well as to investigate endophenotypes intermediate between a gene and a trait. Furthermore, multivariate analysis avoids multiple testing penalty incurred in analysing each phenotype separately (Yang & Wang 2012).

So far there has been little discussion about whether and how these methods could be implemented in glaucoma's genetics research. Following the above considerations, the aim of this project was to explore whether pleiotropy exists between POAG endophenotypes and to identify genetic cross-phenotypes loci within POAG. To achieve this, the specific objectives were:

- to explore epidemiological correlation amongst POAG endophenotypes
- to investigate genetic correlation amongst POAG endophenotypes
- to use dimension reduction approaches in a GWAS setting
- to explore multivariate regression models to accommodate multiple phenotypes in GWAS settings.

4.2 Methods

More details are described in Chapter 2.2.5. To establish the amount of correlation between POAG endophenotypes in the BMES cohort, a list of several factors linked to POAG was compiled including IOP, vertical disc diameter (VDD), vertical cup diameter(VCD), VCDR, CCT and spherical

equivalent (SE). All variables were standardised and a correlation analysis was carried out to obtain Pearson's coefficients as described in Chapter 2.2.5.1.

A linear mixed model was used to calculate the genetic correlation for POAG traits in the BMES cohort as implemented in GCTA (more details Chapter 2.2.5.2).

An iterated factor analysis was carried out including IOP, VDD, VCD, CDR, CCT and age in the BMES and TwinsUK cohorts (Chapter 2.2.5.3). The most relevant factors were selected based on the eigenvalue's value. The selected factors were then used as outcome in GWAS analyses in the BMES and TwinsUK populations. A meta-analysis was then run with the results from the two data sets.

A multivariate linear model (MANOVA) was implemented including IOP and VCDR as outcomes as described in Chapter 2.2.5.4. Results from the BMES and TwinsUK cohorts were then meta-analysed with Fisher's combining probability method.

A multivariate model based on generalised estimating equation (GEE), which takes into account of correlation between outcomes, was implemented in the BMES and TwinsUK cohorts and results were then meta-analysed (Chapter 2.2.5.4).

4.3 Results

4.3.1 Correlation between POAG quantitative traits

2,302 Caucasian individuals from the BMES cohort with measurements for IOP, VDD, VCD, VCDR, SE, CCT and age were used in the correlation analysis. Measurements of the optic disc nerve head are correlated between each other: the VCDR is correlated with the VCD (Pearson's correlation coefficient = 0.64) and with the CDR (Pearson's correlation coefficient = 0.42). The VCD and the VCDR were highly correlated (Pearson's correlation coefficient = 0.96). In addition, the VDD is also correlated with SE (Pearson's correlation coefficient = 0.21). IOP was correlated with CCT (Pearson's correlation coefficient = 0.09) and with CDR (Pearson's correlation coefficient = 0.10) (Table 4.1).

	AGE	IOP	SE	VDD	VCD	VCDR	CCT
AGE	1.00	0.07	0.32	0.03	0.04	0.05	-0.11
IOP	0.07	1.00	-0.04	-0.07	0.12	0.17	0.09
SE	0.32	-0.04	1.00	0.21	0.04	-0.02	-0.01
VDD	0.03	-0.07	0.21	1.00	0.64	0.42	-0.02
VCD	0.04	0.12	0.04	0.64	1.00	0.96	-0.03
VCDR	0.05	0.17	-0.02	0.42	0.96	1.00	-0.03
CCT	-0.11	0.09	-0.01	-0.02	-0.03	-0.03	1.00

Table 4.1: Correlation for POAG endophenotypes. The table shows Pearson's correlation coefficients between several POAG phenotypes: IOP, VDD, VCD, VCDR and CCT. Other variables, usually included in POAG analysis, were incorporated in the analysis.

4.3.2 Genetic correlation between POAG quantitative traits

Given the epidemiological relationship between POAG traits, genetic correlation was investigated. 2,302 individuals with information on a panel of 2,617,903 SNPs from the BMES were analysed. Firstly, I calculated a

SNP based heritability for each trait as shown in Table 4.2. All SNPs used explained a total of 19% of the variance of IOP, 16% of the variance of VCDR and 11% for VCD. The SNPs seem to account for a very low proportion of VDD's variance. 88% of the variance of CCT might be explained by all SNPs. Figure 4.1 shows a summary of the SNP-based heritability for POAG traits.

Phenotype	Proportion of variance explained by SNPs	SE
IOP	0.19	0.17
VCDR	0.16	0.17
VCD	0.11	0.17
VDD	0.0005	0.17
CCT	0.88	0.32

Table 4.2: SNP-based heritability: the table shows the heritability calculated with several SNPs for all POAG traits (proportion of variance explained by SNPs) and stardard errors (SE).

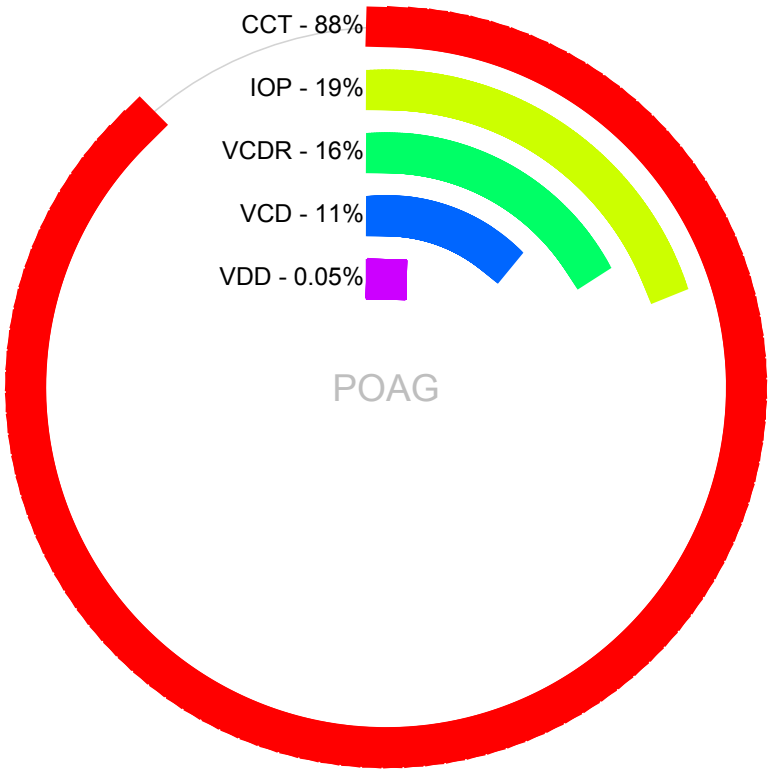


Figure 4.1: SNP based heritability for POAG traits. The figure shows the heritability calculated with the SNPs for POAG quantitative traits

The mixed model analysis yielded a SNP-based correlation of -0.64 (SE=0.97) for IOP and VCDR (Table 4.3 and Figure 4.2). This means that 61% of the epidemiological correlation between the two traits could be explained by the SNPs. The negative sign of the correlation is more difficult to explain. It might indicate that genetic factors for IOP and CDR are negatively correlated with some genes that increase one trait and decrease the another one. In addition, IOP and CCT are negatively correlated with an estimate of -0.56 (SE=0.57). As expected, the genetic factors for VCDR

and VCD are highly correlated (estimate close to 1, with a SE of 0.11).

Phenotypes	Genetic correlation	SE
IOP-VCDR	-0.64	0.97
IOP-CCT	-0.37	0.38

Table 4.3: Genetic correlation between POAG phenotypes: the table shows the genetic correlation between IOP and VCDR and IOP and CCT with standard errors (SE).

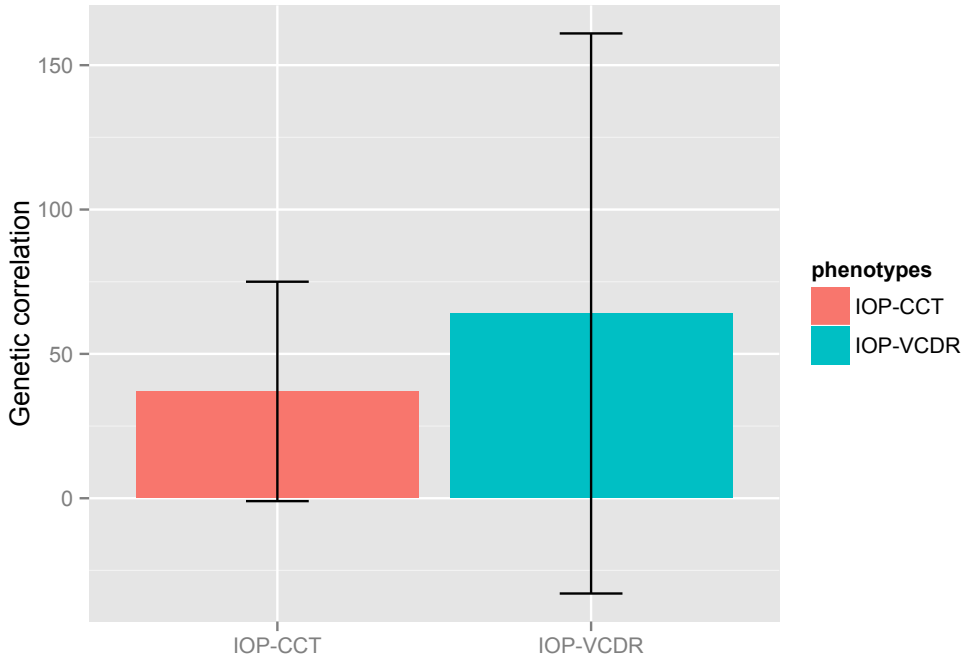


Figure 4.2: Barplot for the genetic correlation between POAG traits: the figure shows the barplot of the genetic correlation and standard errors for POAG quantitative traits

4.3.3 Principal factor analysis in POAG quantitative traits

4.3.3.1 Calculating principal factors

A set of 2,302 individuals from the BMES cohort and 3,493 individuals from the TwinsUK cohort were analysed. The principal factors analysis

produced five different factors. Only the first principal factor had an eigenvalue greater than 1 (2.17) and alone explained 68% of the variance. The second factor was also kept for further analyses. The eigenvalue was 0.48 and the variance explained was of 0.15. Factor loadings represent how much a factor explains a variable. In Figure 4.3 the factor loadings for each variable for the first factor are plotted against those for the second factor, showing that the first factor summarises well the data for the optic disc variables, whereas the second factor describes IOP and CCT and VCDR.

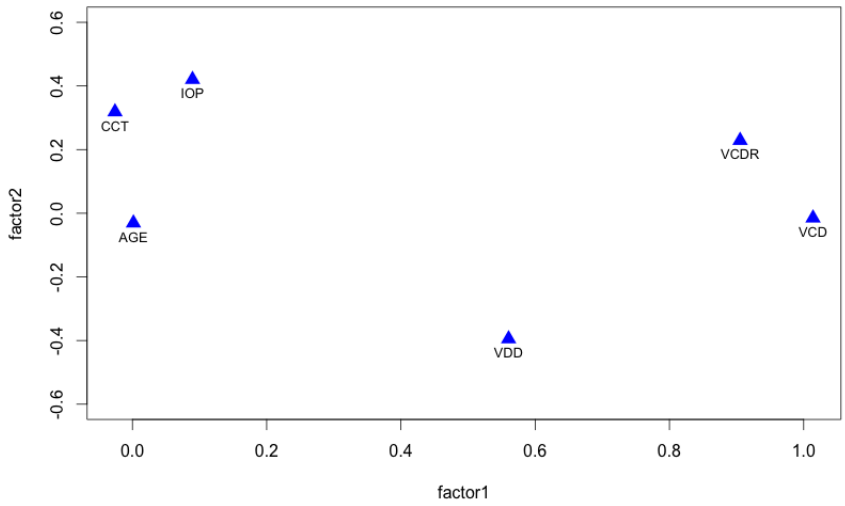


Figure 4.3: Loading factors of the principal factor analysis: the figure shows the loading factors for each variable for the first factor (factor1) plotted against those for the second factor (factor2). Factor1 has a loading factor of $>$0.8 for VCD and VCDR, meaning that it explains a large amount of these variables. On the other hand, factor2 has a high loading factor for IOP and CCT.

In addition, correlation analysis between POAG traits and the first principal factor showed that the principal factor correlated well with VCD (0.9999) and VCDR (0.86). The second factor was correlated with IOP (0.53) and CCT (0.43) as shown in Figure 4.4.

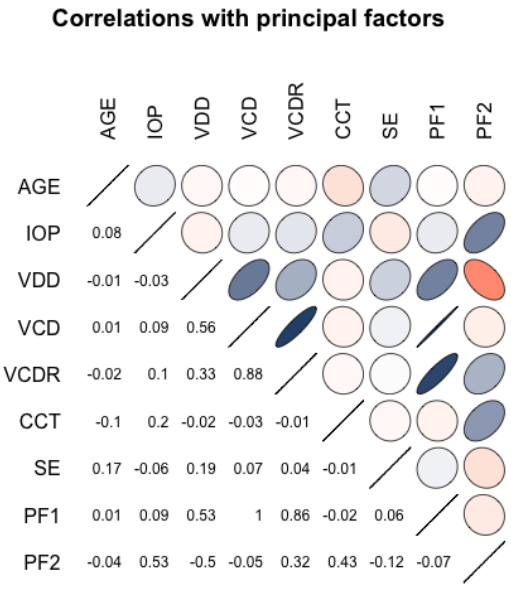


Figure 4.4: Correlation between POAG endophenotypes and principal factors: the figure is a graphical display of the correlation matrix, using blue for positive and red for negative correlations. The ellipses represent the level of correlation.

4.3.3.2 Genome wide association studies with the principal factors

GWAS analyses were performed for the two first principal factors in the BMES (n=2,303) and in the TwinsUK (n=3,494) data sets on a panel of more than 2 million SNPs imputed to HapMap2. An inverse-based meta-analysis was then performed with GWAS results from the two cohorts.

The top results for both principal factors are shown in Table 4.4. Figure 4.5 displays -log10 p values for the first principal factor across all SNPs traditional passing quality control. Quantile-quantile plot for the first principal factor p-values is shown in Figure 4.6.

Phenotype	SNP	A1	MAF	Meta P	Meta Effect	Meta SE	BMES P	BMES Effect	BMES SE	TwinsUK P	TwinsUK Effect	TwinsUK SE	Gene
Pf1	Rs11017216	A	0.08	1.079E-07	0.34	0.06	0.010	0.25	0.10	1.468E-06	0.40	0.08	GLRX3
Pf1	Rs11017215	A	0.08	1.364E-07	0.33	0.06	0.012	0.24	0.10	1.460E-06	0.40	0.08	GLRX3
Pf1	Rs12097296	T	0.27	1.346E-06	0.17	0.03	6.839E-04	0.18	0.05	5.093E-04	0.15	0.04	RGS7
Pf1	Rs12593257	A	0.42	1.520E-06	0.15	0.03	0.001	0.15	0.04	3.245E-04	0.15	0.04	LOC102724973
Pf1	Rs10762198	A	0.38	1.652E-06	-0.17	0.04	4.923E-05	-0.21	0.05	0.005	-0.13	0.05	PBLD
Pf1	Rs7916697	A	0.42	1.937E-06	-0.17	0.04	1.516E-04	-0.20	0.05	0.003	-0.14	0.05	ATOH7
Pf1	Rs7972528	T	0.37	2.529E-06	-0.15	0.03	0.004	-0.13	0.05	1.390E-04	-0.16	0.04	TMTC2
Pf1	Rs729679	T	0.44	2.669E-06	-0.15	0.03	0.002	-0.14	0.04	3.066E-04	-0.15	0.04	LOC102724973
Pf1	Rs1115988	A	0.15	2.727E-06	-0.15	0.03	0.002	-0.14	0.05	4.204E-04	-0.15	0.04	LINC00290
Pf1	Rs11777845	T	0.08	2.884E-06	-0.24	0.05	2.182E-04	-0.30	0.08	0.002	-0.20	0.07	ASPH
Pf2	Rs12031354	A	0.40	4.915E-07	0.11	0.02	5.642E-06	0.14	0.03	0.011	0.08	0.03	UHMK1
Pf2	Rs920581	A	0.11	2.099E-06	-0.16	0.03	0.002	-0.15	0.05	3.447E-04	-0.17	0.05	SCAMP1
Pf2	Rs7555190	A	0.42	2.672E-06	0.10	0.02	1.005E-04	0.12	0.03	0.006	0.09	0.03	C1orf111
Pf2	Rs12912995	T	0.19	2.854E-06	-0.14	0.03	0.003	-0.11	0.04	1.242E-04	-0.17	0.04	IQGAP1
Pf2	Rs4910524	A	0.13	2.926E-06	-0.16	0.03	0.009	-0.12	0.05	4.937E-05	-0.19	0.05	OR51E1
Pf2	Rs292654	T	0.47	3.353E-06	-0.10	0.02	0.001	-0.10	0.03	7.978E-04	-0.11	0.03	STRA8
Pf2	Rs2057671	T	0.11	4.751E-06	0.16	0.04	2.553E-04	0.17	0.05	0.005	-0.15	0.06	JAZF1
Pf2	Rs17063793	T	0.43	4.980E-06	0.10	0.02	3.765E-04	-0.11	0.03	0.004	-0.10	0.04	LINC00347
Pf2	Rs4722750	T	0.11	5.185E-06	0.16	0.04	2.939E-04	0.17	0.05	0.005	0.15	0.06	JAZF1
Pf2	Rs292641	C	0.49	5.219E-06	0.10	0.02	0.001	0.10	0.03	8.004E-04	0.11	0.03	STRA8

Table 4.4: Results GWAS for the first two principal factors: the table shows the top hits for the meta-analysis of GWAS performed on the first two principal factors in the BMES and TwinsUK cohorts. Pf1: first principal factor, pf2: second principal factor, A1: effect allele, MAF= minor allele frequency, Meta P = p-value from the meta-analysis, Meta Effect= β from the meta-analysis, Meta SE= standard error from the meta-analysis, BMES P= BMES GWAS p-value, BMES Effect = BMES GWAS effect size, BMES SE= BMES GWAS standard error, TwinsUK P= TwinsUK GWAS p-value, TwinsUK Effect = TwinsUK GWAS effect size, TwinsUK SE = TwinsUK GWAS standard error, Gene: closest transcript/gene

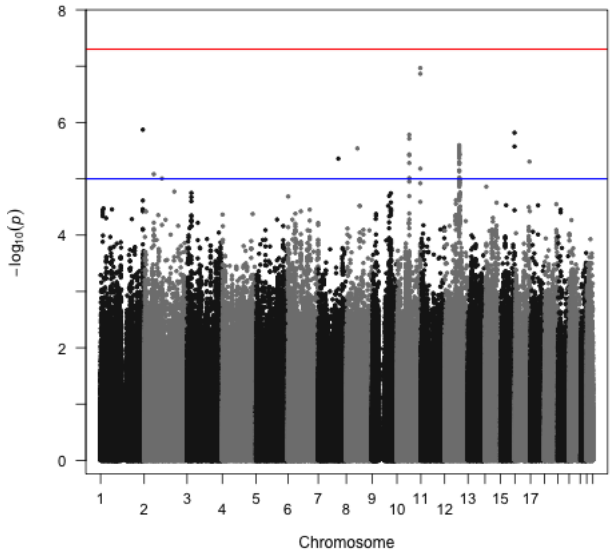


Figure 4.5: Manhattan plot for the meta-analysis of GWAS for the first principal factor: the figure shows the -log10 of the p-values for the meta-analysis of GWAS results for the first principal factor

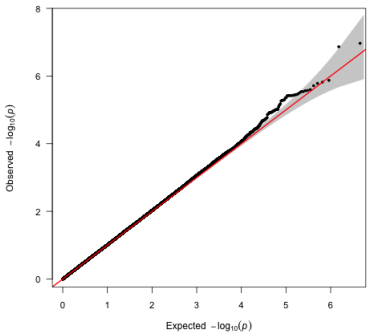


Figure 4.6: Q-Q plot for the meta-analysis of GWAS for the first principal factor: the figure shows the quantile-quantile plot of the p-values from the first principal factor results

For the first principal factor, the strongest signal was found on chromosome 10 at rs11017216 (P= 1.079E-07, effect=0.335, SE= 0.063). Figure 4.7 shows the forest plot for this SNP where the effects in both cohorts follow the same direction and the confidence intervals largely overlapped. This SNP lies on an intergenic region (chr10:132,004,985) and there are no genes within 100k bp. However, the closest gene is *GLRX3* (chr10: 131,824,652-131,872,774).

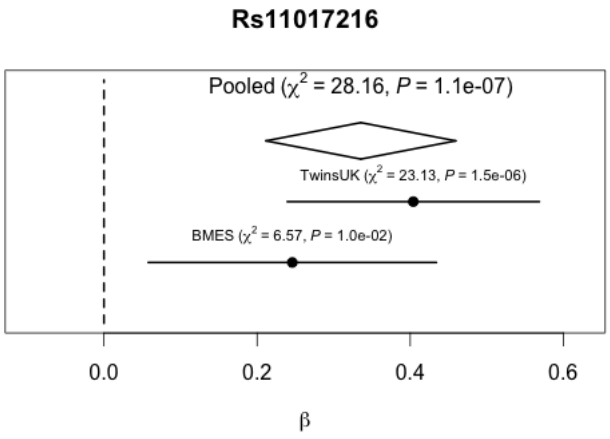


Figure 4.7: Forest plot for rs11017216 in the meta-analysis for the first principal component: the figure shows a graphical representation (forest plot) of the top hit rs11017216 meta-analysis results. The vertical line represent no effect. Effect sizes and confidence intervals for each cohort are plotted.

Another signal was found at rs12097296 (P=1.346E-06, effect=0.165, SE=0.034). The effects for both cohorts are positive and the confidence intervals overlapped. This SNP lies within the sequence region of *RGS7* (regulator of G-protein signaling 7), which is a protein-coding gene. Another genetic signal was identified at rs12593257 on chromosome 15 which lies in an intergenic region (chr15: 98,320,782-chr15:98,321,282) with an effect of 0.15 and standard error of 0.031 (P=1.520E-06). This region shows a regulatory activity and the closest gene are *LOC102724973* and *ADAMTS17*. Another signal for the first principal component has been found at rs7916697 on chromosome 10 with an effect size of -0.17. This SNP lies within the region of *ATOH7* gene.

Figure 4.8 shows the -log 10 p-values for the second principal factor across all genome. Figure 4.9 shows the quantile-quantile plot for the results. The strongest signal was found at rs12031354 on chromosome 1 with an effect size of 0.111 (SE=0.022 and P=4.915E-06). This SNP lies on an intergenic region (chr1:161,128,721-161,129,221) and the closest gene is C1orf110. In addition, we identified another signal at rs920581 on chromosome 5 with an effect of -0.158 (SE=0.033, P=2.099E-06). This SNPs lies on an intergenic region (chr5:78,009,982-78,060,081) which is highly conserved across species and has a high regulatory activity. The closest gene is *LHFPL2*. Other signals for the second principal factor were identified within *IQGAP1, OR51E1* and *JAZF1* genes.

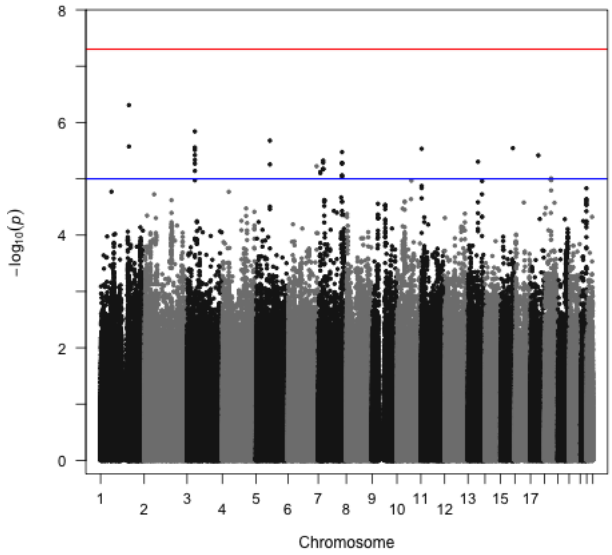


Figure 4.8: Manhattan plot for the meta-analysis of GWAS for the second principal factor: the figure shows the -log10 of the p-values for the meta-analysis of GWAS results for the second principal factor.

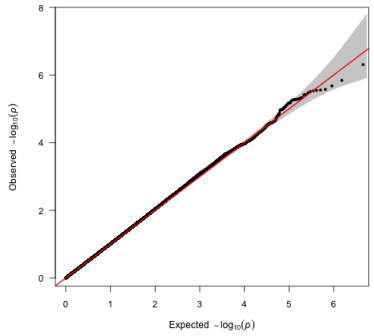


Figure 4.9: Q-Q plot for the meta-analysis of GWAS for the second principal factor: the figure shows the quantile-quantile plot of the p-values from the second principal factor meta-analysis results

4.3.4 Multivariate regression models in POAG quantitative traits

Rather than reduce the dimensions of the outcome, multivariate analyses jointly analyse more than one phenotype in an unified framework and test for association of multiple phenotypes with a genetic variant. Two

methods were selected: multivariate analysis of variance (MANOVA) and a generalised estimating equations (GEE) method.

4.3.4.1 Multivariate analysis of variance (MANOVA)

A set of 2,302 of individuals from the BMES cohort and a set of 1,508 from the TwinsUK cohort were analysed. In the TwinsUK dataset only one twin of each pair was included to avoid family structure. A bivariate regression model, also called multivariate analysis of variance (MANOVA), was implemented including IOP and VCDR. The results from the two cohorts were combined using Fisher's combining probability method.

No SNPs crossed the line of genome-wide significance (Figure 4.10). However, there are some suggestive signals (Table 4.5). The strongest signal was identified at rs2902210 on chromosome 1 (P=1.366E-06). This SNP lies within *SEC16B/RGPR/KIAA1926* genes. Another signal was found at rs7792197 (P=3.001E-06) on chromosome 7 and it lies in an intergenic region. The closest genes are *LOC100128464, LOC100419447, EIF4A1P13*. Other signals were found near other genes such as *DOCK11P1, RRAGAP1, LOC101927351, LOC101927395, LOC101927418, C14orf28, LOC401770 and KLHL28*. The -log10 of the Fisher's combined p-values across all genome are shown in Figure 4.10 and Figure 4.11 shows the quantile-quantile plot for the Fisher's pvalues.

Marker Name	Chr	Position	P-value	Gene
rs2902210	1	177936039	1.366E-06	SEC16B
rs7792197	7	88432198	3.001E-06	LOC100128464,LOC100419447,EIF4A1P13
rs17353068	7	88433028	4.081E-06	LOC100128464,LOC100419447,EIF4A1P13
rs1593278	14	44841227	4.467E-06	DOCK11P1,RRAGAP1,LOC101927351,LOC101927395,LOC101927418,C14orf28,LOC401770,KLHL28
rs17092468	14	44842318	4.467E-06	DOCK11P1,RRAGAP1,LOC101927351,LOC101927395,LOC101927418,C14orf28,LOC401770,KLHL28
rs17115642	14	44844397	4.467E-06	DOCK11P1,RRAGAP1,LOC101927351,LOC101927395,LOC101927418,C14orf28,LOC401770,KLHL28
rs11965318	6	82778474	4.485E-06	TPBG
rs2494174	1	210197611	5.366E-06	ST13P19,SYT14,SERTAD4,SERTAD4-AS1
rs2485901	1	210197416	6.075E-06	ST13P19,SYT14,SERTAD4,SERTAD4-AS1
rs1865273	18	40427300	6.236E-06	LINC01477

Table 4.5: Fisher's combined results for the bivariate analysis for IOP and CDR: the table shows the results from the meta-analysis for the bivariate analysis using IOP and VCDR in the BMES and TwinsUK cohorts. Chr: chromosome, Position: base-pair position of the SNP, P-value: p-value of the bivariate test, Gene: closest transcript/gene

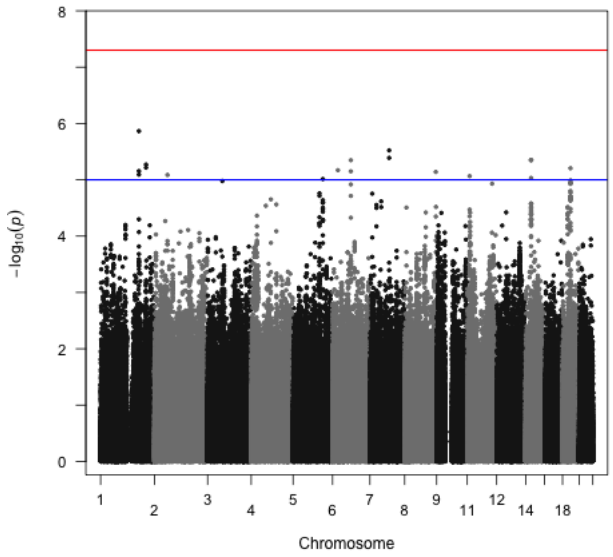


Figure 4.10: Manhattan plot for Fisher's combined results for the bivariate model

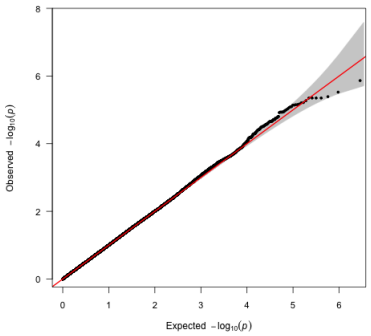


Figure 4.11: Q-Q plot Fisher's combined p-values for the bivariate model

4.3.5 GEE regression model

Another multivariate regression model was applied on the same 2,302 of individuals from the BMES and a set of 1,508 from TwinsUK cohort. This method is an extension of GEE models as described by Schifano *et al.* (Schifano *et al.* 2013). The main difference between this model and the previous MANOVA analysis is that GEE model can take into account correlation between traits. After running the regression model, results for BMES and TwinsUK cohorts were meta-analysed using Fisher's combined probability method.

Figure 4.12 displays -log10 Fisher's combined p values for GEE approach across all SNPs passing quality control. Quantile-quantile plot for the SMAT combined p-values is shown in Figure 4.13.

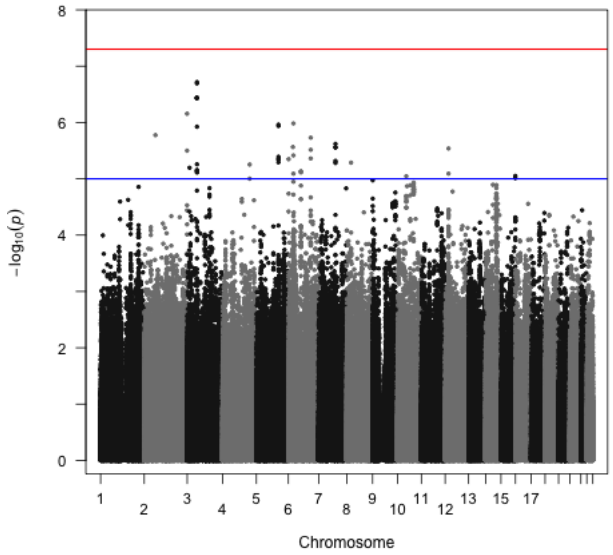


Figure 4.12: Manhattan plot of the meta-analysed results for GEE regression model: the figure shows the -log10 of the p-values obtained from Fisher's meta-analysis for GEE model across all genome.

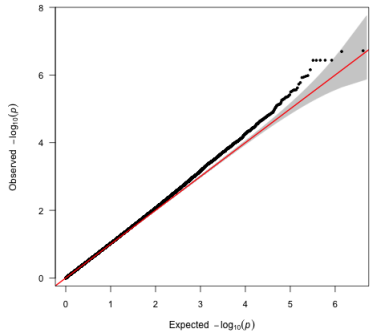


Figure 4.13: Q-Q plot of the meta-analysed results for GEE regression model: the figure shows the Q-Q plot of the p-values from Fisher's meta-analysis for SMAT model.

No SNPs passed genome-wide significance, however there were 8 SNPs that reached p<E-07 (see Table 4.6). Most of these SNPs lie within *CACNA2D3* gene. In addition, another signal reached p<E-07 which lies within *SNED1* gene.

Marker Name	Chr	Position	Gene	Chisq	P-value	TwinsUK pvalue	BMES pvalue
rs11130451	3	55078388	CACNA2D3	36.866	1.919E-06	4.102E-08	0.240
rs17257847	3	55085117	CACNA2D3	36.756	2.022E-06	4.683E-08	0.222
rs17257660	3	55081044	CACNA2D3	35.506	3.657E-06	8.132E-08	0.239
rs4955928	3	55081482	CACNA2D3	35.506	3.657E-06	8.132E-08	0.239
rs4955929	3	55081645	CACNA2D3	35.506	3.657E-06	8.132E-08	0.239
rs4955925	3	55079930	CACNA2D3	35.492	3.681E-06	8.132E-08	0.241
rs4955926	3	55079974	CACNA2D3	35.492	3.681E-06	8.132E-08	0.241
rs7571117	2	241040133	SNED1	34.136	6.987E-06	4.135E-06	0.093
rs2859365	6	28423688	ZSCAN23	33.295	1.039E-06	0.020	2.876E-06
rs11241800	5	125312879	LOC101927421	33.185	1.094E-06	0.341	1.820E-06

Table 4.6: Results of the meta-analysed results for GEE regression model: the table shows the meta-analysed results for the GEE regression model: chr:chromosome, Position: the base-pair position, Gene: the closest gene, Chisq: the chi-square of the Fisher's combined method, P-value: the p-value of the Fisher's meta-analysis, TwinsUK pvalue: the p-value from the regression model for the TwinsUK cohort, BMES pvalue: the p-value form the regression model for the BMES cohort

4.4 Discussion

Although association analyses have been successful in identifying thousands of variants associated with many traits, including POAG, most GWAS have been conducted using the simple statistical models focused on single phenotype, leaving out interactions and inter-relationships with putative biologic significance. Methods for analysing multivariate phenotypes can accommodate different degrees of dependence between several POAG phenotypes in integrated models. The aim of this project was to explore whether pleiotropy exists between POAG endophenotypes and thus to identify genetic cross-phenotypes loci within POAG. To achieve this, firstly I established and confirmed the correlation amongst POAG traits and I then explored several ways to combine these phenotypes in a multivariate framework.

Firstly, epidemiological correlation between POAG traits was analysed. Results were in line with previous studies. The relationship between IOP and CCT (Doughty & Zaman 2000) and between optic disc measurements, including disc size and VCDR (Quigley HA *et al.* 1990) was confirmed. More interesting is the link between IOP and CDR. The relationship between IOP and the glaucomatous damage is not clear, however critical to understand and treat the disease.

Mere phenotypic correlation between quantitative traits does not guarantee the presence of pleiotropy and shared genes for several POAG traits as the correlation may due to environmental factors. Interestingly, a genetic correlation of 64% was identified for IOP and VCDR, showing a shared genetic determination between the two traits and thus suggesting the existence of pleitropic factors underlying both phenotypes.

Principal factor analysis was used to reduce the dimensions of the traits and calculate factors that can be used in a GWAS setting. Meta-analysis of GWAS in two Caucasian Populations using principal factors identified suggestive signals lying close or within *GLRX3, RGS7, LOC102724973, ADAMTS17* and *ATOH7* genes. *GLRX3* gene encodes a member of the glutaredoxin family. The encoded protein binds to and modulates the function of protein kinase C theta and may also inhibit apoptosis, playing a role in cellular growth (Cheng *et al.* 2011). Regulator of G protein signalling (RGS) protein are appreciated modulators of neurotransmitter actions. Many neurotransmitters and hormones rely on G proteins to activate intracellular signalling pathways that regulate nearly all aspects of cell and organ physiology (Rose *et al.* 2000). *RGS7* gene is expressed predominantly in the brain and it attenuates signal transduction in

mammalian cells. In addition, it may play a role in synaptic vesicle exocytosis (Hunt *et al.* 2003) and may play an important role in the rapid regulation of neuronal excitability (Saitoh *et al.* 1999). *RGS7* gene may also has a role in the eye and especially in the retina (Cabrera *et al.* 1998). In addition, a genetic signal (rs12593257) for the first principal factor was identified on chromosome 15. This SNP lies in an intergenic region (chr15: 98,320,782-chr15:98,321,282) that shows a regulatory activity and the closest genes are *LOC102724973* and *ADAMTS17*. The latter encodes a member of the ADAMT protein family. The function of this protein has not been determined, however truncating mutations within this gene family produce typical eye anomalies described in Weill-Marchesani syndrome, which is characterised by high myopia and secondary glaucoma due to shallowing of anterior chamber angles (Morales *et al.* 2009). Another signal for the first principal factor has been found at rs7916697 on chromosome 10. This SNP lies within the region of *ATOH7* gene which is a well known gene for optic disc parameters and glaucoma (Ramdas *et al.* 2011; Springelkamp *et al.* 2014).

For the second principal factor, the strongest signal has been found at rs12031354 on chromosome 1. This SNP lies on an intergenic region (chr1:161,128,721-161,129,221). The closest gene is *C1orf110* which is a protein binding gene. It is worth noticing that another gene in the proximity is *RGS4*, which is a member of the RGS (regulator G protein signaling) and it may inhibit signal transduction and includes pathways similar to RGS7 previously identified for the first principal factor. *RGSA4* has been previously associated with schizophrenia (Morris *et al.* 2004) and seems to be specific to the neural system, modulation the function of

dopamine and glutamate receptors (De Vries *et al.* 2000).

Multivariate regression models can directly take into account two or more outcomes without the need of calculate factors first. Results for bivariate regression using MANOVA with IOP and CDR showed that no SNPs crossed the line of genome-wide significance. However, there are some suggestive interesting signals which lie within/close to *SEC16B* and *SYT14* genes. *SEC16B* is involved in the organisation of transitional endoplasmatic reticulum (ER) sites and protein export (Watson *et al.* 2006) and also it has been associated with obesity (Hotta *et al.* 2009). *SYT14* gene is a member of the synaptotagmin gene family and encodes a protein similar to other family members that mediate membrane trafficking in synaptic transmission. The encoded protein is a calcium-independent synaptotagmin. Mutations in this gene are a cause of autosomal recessive spinocerebellar ataxia-11 (SCAR11), and a t(1;3) translocation of this gene has been associated with neurodevelopmental abnormalities (Sailer & Houlden 2012). Alternatively spliced transcript variants encoding multiple isoforms have been observed for this gene, and a pseudogene of this gene is located on the long arm of chromosome 4.

Since POAG traits are correlated between each other, it may be more appropriate to use a regression model that can take into account the correlation between the traits such as GEE.Using a GEE regression model, other suggestive signals were identified in the meta-analysis between BMES and TwinsUK cohorts. *CACNAD3* is a member of the alpha-2/delta subunit family that modulates the expression and the function of the voltage-gated calcium channel complex. The protein encoded is expressed in spiral ganglion neurons and auditory brain stem nuclei and it might be

required for normal acoustic response (Pirone *et al.* 2014). In addition, this gene seems to be involved in pain sensitivity. Mice mutant exhibit impaired behavioural heat pain sensitivity and SNPs variants in humans were associate with reduced sensitivity to acute noxious heat and chronic back pain (Neely *et al.* 2010). *SNED1* is a protein-coding gene and GO annotations include calcium ion binding. An important paralog of this gene is *DLK2*, which is an epidermal growth factors-like protein and has been identified as a modulator of adipogenesis in vitro (Rivero *et al.* 2012).

Despite identifying a few suggestive signals, the results of applying multivariate models on glaucoma endophenotypes are far from being definitive. Mechanisms underlying POAG, in particular the role of the high IOP and the changes in the optic disc nerve head in the pathophysiology, have not been determined yet. The cupping of the optic disc arises in response to IOP. It has been hypothesised that an increase of IOP raises the pressure gradient across the lamina cibrosa and leads to a deformation of retinal ganglion cell axons due to the mechanical stress (Bellezza *et al.* 2003; Pena *et al.* 2001). Animal models provide evidence for this hypothesis. For example, monkeys have been shown to develop glaucomatous changes in the optic nerve when their IOP is raised experimentally (Crawford *et al.* 2001). However, there is a certain grade of in-dependency amongst mechanisms that regulate IOP and the cupping in the optic nerve head. The retina is dependent on its blood supply for metabolic needs, and local ischaemia-hypoxia, probably due to dysfunction of blood flow regulation, it has been implicated. In particular vascular factors has been proven important for normal tension glaucoma, where vascular dysregulation increases the impact of ischemia affecting also the optic nerve without high IOP (Mi *et al.* 2014). Other factors

that contribute independently to IOP to death of retinal ganglion cells and optic nerve fibres may be excessive stimulation of the glutamatergic system (i.e. N-methyl-D-aspartate subtypes) (Dreyer *et al.* 1996), poorly functioning cellular pumps and glutamate transporters, oxidative stress and formation of free radicals, inflammatory cytokines in and aberrant immunity (Schwartz 2003).

Of course there are a number of limitations that might have diminished the chance of identifying shared signals between POAG endophenotypes. First of all, it is often very difficult to give a meaningful interpretation of multivariate analyses. For instance, effect sizes are difficult to interpret since they were represented as a matrix (for regression models) or not informative such as in principal factor analysis. Moreover, multivariate approaches need complete observations for each individuals leading to to the exclusions of subjects that miss even only one phenotype and thus decreasing the total sample size. GWAS need thousands of individuals to identify significant loci, considering genetic variants affecting multiple phenotypes are likely to have a small effect. In addition, the use of HapMap2 data might make difficult to pick up rare genetic signals because focused on common variants which are known to have very small effect on phenotypes.

4.5 Conclusion

This study has investigated the correlation between POAG endophenotypes and the existence of genetic loci affecting both IOP and VCDR. The link between POAG endophenotypes have been confirmed through

epidemiological and genetic correlation. However, this study has identified a very small number of suggestive signals shared between IOP and VCDR in a GWAS setting and the role of these variants is not clear. Therefore, these findings suggest that IOP and VCDR might underlying different processes of the disease. The contribution of this study has to be confirmed with future association studies, however it has gone some way towards enhancing our understanding of POAG mechanisms.

CHAPTER 5

Detecting gene-gene interactions in POAG endophenotypes in a meta-analysis of multi-ancestry cohorts

In this chapter I will describe and discuss my research on epistatic interaction on POAG, in particular I investigated whether CDR and IOP known loci interacted with other loci in the genome to affect CDR and IOP.

STATEMENT OF CONTRIBUTION TO THIS RESEARCH:

I helped to design the study and I performed all the analyses described in this chapter for the BMES and TwinsUK cohort. I participated in leading the meta-analysis in the IGGC (i.e. writing the protocol for the analysis) and in performing the meta-analysis.

5.1 Background

GWAS have been a successful tool to identify common genetic factors associated with POAG phenotypes, though the variants identified only explain a small portion of the heritability (Manolio *et al.* 2009). Usually GWAS use an additive framework where each allele that affects a trait acts in an independent, non linear, cumulative manner. However, a line of thought in the literature supports the idea that the mechanisms of gene

action, thus genetic architecture of complex traits, are much more complex than the additive model (Carlborg & Haley 2004; Phillips 2008; Moore *et al.* 2006; Wright 1931). Epistasis has been suggested to answer part of the missing heritability problem (Manolio *et al.* 2009; Frazer *et al.* 2009; Eichler *et al.* 2010; Zuk *et al.* 2012). As heritability in a narrow sense (h2) accounts for additive variance, epistasis should be part of the unknown (non additive) heritability (Visscher *et al.* 2008). However, it has been shown that epistasis might also contribute to h2, not only by statistical illusion of additive variance (where confounding non additive and common environmental affects twin study calculation of h2, but also creating real additive variation as marginal effects from higher order genetic interaction (Hemani *et al.* 2013).

Epistasis is very common in nature and is a key in characterising the genetic basis of complex disease. For example it has been described in model organisms such as budding yeast (Costanzo *et al.* 2010) and in Drosophila (Huang *et al.* 2012). Epistasis seems to have an evolutionary role for example maintaining deleterious variants under selection (Hemani *et al.* 2013). At the macroevolutionary scale, it has been shown that advantageous substitution in one specie is often found deleterious in others, supporting the idea that effect on fitness depends upon not only another gene but all the genetic background (Breen *et al.* 2012).

Gene-gene interactions are also described in human studies, yet there is little empirical exploration of the role of epistasis in complex traits in humans. Interaction between HLA-C and ERAP1 loci has been reported for psoriasis (Strange *et al.* 2010) and polymorphisms of ERAP1 only affect ankylosing spondylitis risk in HLA-B27-positive individuals (Evans *et al.*

2011). Hemani *et al.* (Hemani *et al.* 2014) identified 501 interactions between common SNPs and expression of 238 genes in a cohort of 846 individuals with 7,339 gene expression levels measured in peripheral blood. They also demonstrated physical organisation of interacting loci within a cell, suggesting a mechanism by which biological function can lead to epistatic genetic variance, since different chromosomal regions spatially co-localise in the cell through chromatin interaction (Hemani *et al.* 2014). Brown *et al.* reported 26 replicated variance eQTL and 57 replicated cis-epistatic interactions, which explain 16% of the variance of BMI using RNA-sequence data from lymphoblastoid cell lines (LCLs) from the TwinsUK cohort (Brown *et al.* 2014). At the time of writing, only one study investigated epistasis interaction in glaucoma (Hu *et al.* 2015). The authors build a human phenotype network (HPN) and used GLAUGEN glaucoma GWAS dataset and apply HPN as a biological knowledge-based filter to prioritise genetic variants. After that, they used a statistical epistasis network (SEN) to identify significant connected network of pairwise epistatic interactions among the prioritised SNPs. After functional annotation, Hu *at al.* were able to link glaucoma to several complex diseases (such as T2D, obesity, Alzheimer's disease, cataracts and myopia).

Epistasis is defined as interaction between different genes (Cordell 2002). Two different definitions of epistasis exist: biological and statistical interaction. In 1909, Bateson (Bateson 2007) described the "masking effect" where a variant or allele prevents another one from manifesting its effect in a similar manner to a dominance effect. This concept is similar to biological protein-protein interaction and it has been described as "biological interaction". Studying the biological interactions between genes might exploit the qualitative nature of the mechanism of action of a factor

affected by the presence/absence of another. Another definition of epistasis was proposed by Fisher (Cordell 2002) and it has been called ”statistical interaction”. Fisher describes epistasis as deviation from additivity in the effect of alleles at different loci with respect to their contribution to a quantitative trait. The statistical interaction is basically a deviation from a linear model. It is worthy to underlie that statistical interaction does not necessarily imply interaction at a biological level (Greenland 2009; Gibson 1996; VanderWeele 2009; Cordell 2009).

Several approaches have been used to study epistatic interactions, including regression framework, information theory and entropy, two locus versus high order interactions approaches, machine learning, data-mining, recursive partioning, multifactorial dimension reduction and Bayesian model of selection (Cordell 2009).

Following the previous point, the aim of this chapter was to investigate epistatic interactions in POAG phenotypes. The objectives of this chapter were:

- perform snp by snp interaction analyses to identify novel variants that interact with known ones to alter VCDR in the general population and to elucidate whether known VCDR-associated SNPs interact together to affect the phenotype

- perform snp by snp interaction analyses to identify novel variants that interact with known ones to alter IOP in the general population and to elucidate whether known IOP-associated SNPs interact together to affect the phenotype

5.2 Methods

An interaction model was implemented as described in Chapter 2.2.6.
A list of 17 SNPs associated with VDCR in the IGGC meta-analysis described in Chapter 3 was compiled. The interaction model was performed for each pair of SNPs where the first SNP was on of the 17 SNPs in the list and the second SNP was one included in HapMap2. This produced a total of $>17*2{,}500{,}000$ interaction models. Likewise, for IOP, 12 SNPs were selected for their association in the IGGC meta-analysis (Chapter 3) for a total of $>12*2{,}500{,}000$ interactions. All analyses were adjusted for age and sex. Fifteen cohorts from IGGC participated in the analysis and the snp-snp genome-wide interaction analysis was performed for each cohort. Quality control and meta-analysis was then carried out (Chapter 2.2.6 and Chapter 2.2.3).

5.3 Results

The aim is to investigate epistatic interactions amongst the known glaucoma quantitative traits hits (in association with IOP and CDR) and all other HapMap2 SNPs.

5.3.1 Cup to disc ratio

Fourteen cohorts from the IGGC for a total of 24,629 individuals were included in the analysis.Quality control and analysis for each cohort were performed independently by each group and I was responsible for the BMES and the TwinsUK cohorts. However, before the meta-analysis,

another quality control was necessary to assess the quality of the results received.

Results were grouped in 17 different GWAS that included interactions from one of the 17 SNPs included in the VCDR list against all the genome in HapMap2. Genomic inflation factors were then checked for each of the 17 meta-analysis (Figure 5.1), showing a high variability of λ that varies from 0.01 to 7.43. To understand whether this variability was due to specific cohorts, genomic inflation was double-checked for each cohort for the 17 analysis. As Figure 5.2 shows, one particular cohort, Gutenberg Health Study 1 (GHS1) had a wide range of λ values including the extreme values of 0.01 and 7.43 and therefore was excluded for further analysis.

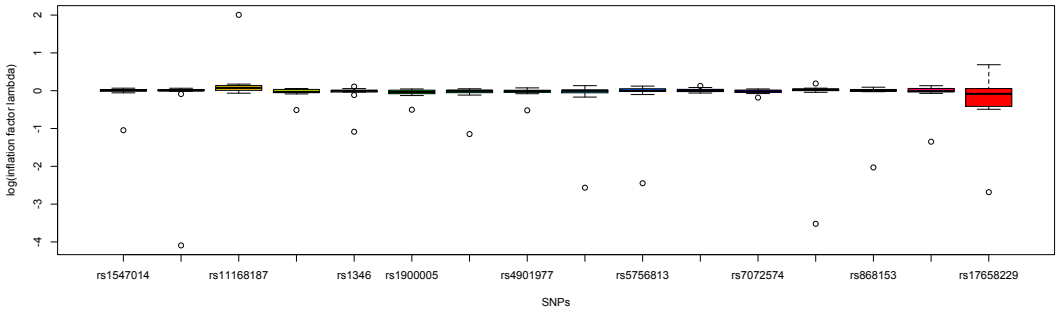


Figure 5.1: VCDR genomic inflation factor λ for each SNPs for all cohorts: this figure shows the 17 VCDR-related SNPs included in the analysis (x-axis) the log of the inflation factors λ in all cohorts (y-axis).

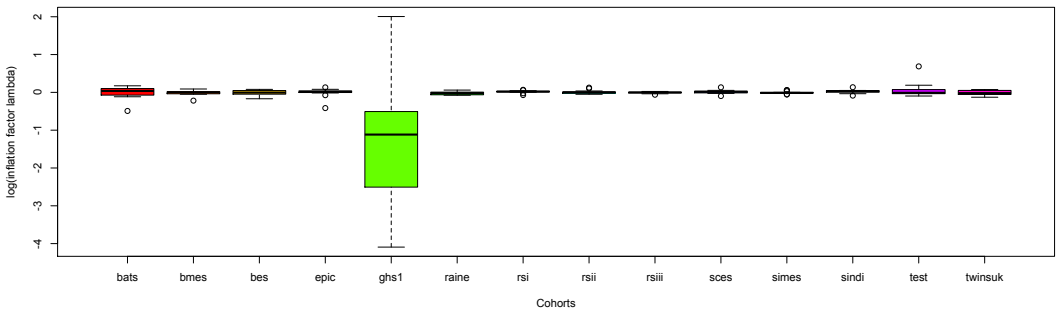


Figure 5.2: CDR genomic inflation factor λ for each cohort for all SNPs: the figure shows the 14 cohorts included in the analysis (x-axis) and the log of the inflation factors λ for all SNPs (y-axis).

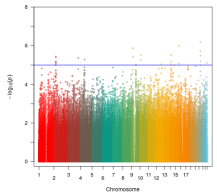
Multiple testing correction was necessary due to the multiple tests performed. Conventionally for GWAS with imputed HapMap 2 data, Bonferroni genome-wide significance threshold is represented by 5E-08. However, in gene-gene interaction analysis the number of tests increase and in this case Bonferroni threshold is 1.17E-09 (considering an average of 2,500,000 SNPs times 17 SNPs). No interaction passed Bonferroni correction (Figure 5.3. However, several loci reached suggestive significance (Table 5.1). I identified rs13244205, rs3801350, rs7793287 on chromosome 7 that interacted with rs7072574 on chromosome 10 (respectively β=0.0155, SE=0.002, P=4.19E-08, β=0.0127, SE=0.002, P=1.36E-07, β=0.0127, SE=0.002, P=1.40E-07). The three SNPs lies within/next to *PDE1C* transcript and rs7072574 on *PLCE1* gene. Another interesting suggestive interaction was between rs579343 (located in *PITX2* gene) on chromosome 4 and rs17658229 within *DUSP1* gene on chromosome 5 (β=-0.040, SE=0.007, P=1.55E-07). Interactions were also found for rs13071457 and rs4658101 (*EFHB* and *CDC7/TGFBR3* genes), rs2029203 and rs1345467 (*PCDH7* and *SALL1* genes), rs11610668 and rs4901977 (*C2CD5* and *SIX1/6* genes).

Chr1	SNP1	Gene1	Chr2	SNP2	Gene2	β	SE	P	I2
10	rs7072574	PLCE1	7	rs13244205	PDE1C	0.015	0.002	4.19E-08	0.00
10	rs7072574	PLCE1	7	rs3801350	PDE1C	0.012	0.002	1.36E-07	0.00
10	rs7072574	PLCE1	7	rs7793287	PDE1C	0.012	0.002	1.40E-07	0.00
5	rs17658229	DUSP1	4	rs579343	PITX2	-0.040	0.007	1.55E-07	0.00
1	rs4658101	CDC7/TGFBR3	3	rs13071457	EFHB	0.022	0.004	1.78E-07	0.00
16	rs1345467	SALL1	4	rs2029203	PCDH7	0.018	0.003	2.77E-07	0.35
14	rs4901977	SIX1/6	12	rs11610668	C2CD5	0.028	0.005	3.43E-07	0.00
1	rs7072574	PLCE1	2	rs13432997	OSR1/WDR35	0.017	0.003	4.15E-07	0.00
6	rs868153	HSF2	12	rs17801352	PPFIBP1	-0.017	0.003	4.66E-07	0.12
12	rs11168187	RPAP3	4	rs9992273	PCDH7	0.018	0.003	4.77E-07	0.36

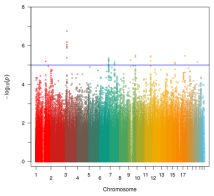
Table 5.1: Most significant meta-analysis results for VCDR snp-snp interaction GWAS: the table shows the most significant results for all the 17 meta-analysis for VCDR. CHR1: chromosome for the SNP in the list, SNP1: SNP in the VCDR-associated variants' list, Gene1: closest gene in the list, CHR2: chromosome for the SNP in the genome that showed interaction, SNP2: SNP in the genome that showed interaction, Gene2: closest gene for the SNP in the genome that showed interaction

The interactions between the 17 SNPs was also investigated to check whether the most significant variants for VCDR worked independently or had a joint effect on the phenotype. Table 5.2 showed that rs301801 and rs17658229 are interacting with several other VCDR SNPs.

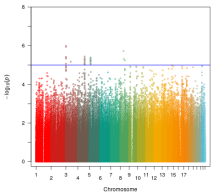
In particular, rs17658229 located in *DUSP1* gene interacted with the variants in *SSCA1*, *RPAP3*, *PLCE1* and *TMTC2* genes. *TMTC2* gene also interacted with *CHEK2* gene. In addition, the genetic variants within *RERE* gene had significant interactions with SNPs located in *SALL1*, *CDKN2BAS* and *CDC7/TGFBR3* genes (Figure 5.4). Also the variant in *SIX1/6* gene seems to interact with the one in *CARD10*. Figure 5.4 and Figure 5.4 are graphical representations of the interactions.



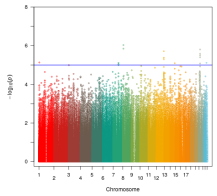
(a) rs10862688



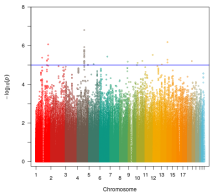
(b) rs4658101



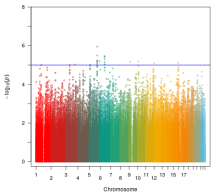
(c) rs1900005



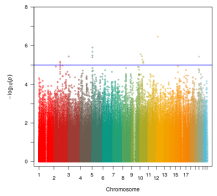
(d) rs1345467



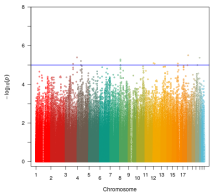
(e) rs17658229



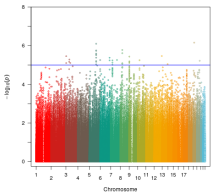
(f) rs2623325



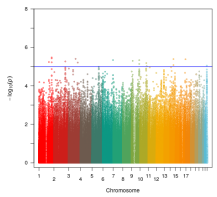
(g) rs4901977



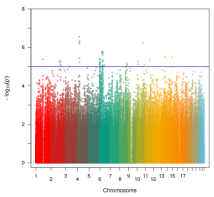
(h) rs5756813



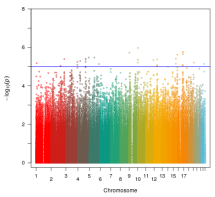
(i) rs301801



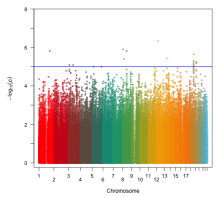
(j) rs1346



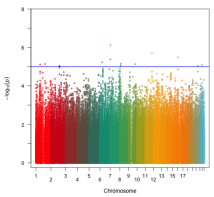
(k) rs11168187



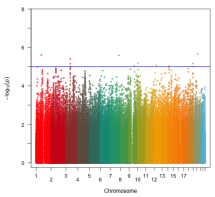
(l) rs1547014



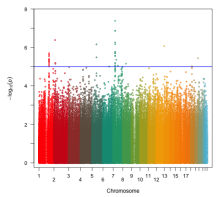
(m) rs868153



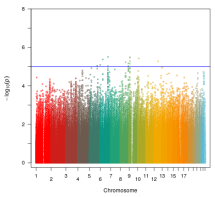
(n) rs7865618



(o) rs6054374



(p) rs7072574



(q) rs4936099

Figure 5.3: Manhattan plots for epistatic meta-analyses with VCDR SNPs: the figure shows the 17 different meta-analysis performed for each of the 17 VCDR-associated SNP interacting with all the genome. The rs number for each plot represents the different 17 VDCR-related SNPs.

Snp1	Snp2	Effect	SE	P
rs301801	rs4658101	0.01	0.00	0.04
rs301801	rs1345467	0.01	0.00	0.01
rs301801	rs7865618	0.01	0.01	0.00
rs2623325	NA	NA	NA	
rs17658229	rs10862688	-0.01	0.01	0.01
rs17658229	rs1346	0.02	0.01	0.02
rs17658229	rs7072574	-0.01	0.01	0.01
rs17658229	rs11168187	-0.01	0.01	0.05
rs868153	NA	NA	NA	
rs4936099	NA	NA	NA	
rs10862688	rs1547014	0.01	0.00	0.02
rs4901977	rs5756813	-0.00	0.00	0.05
rs6054374	NA	NA	NA	

Table 5.2: SNP-SNP interactions between VCDR-related SNPs: this table shows the interactions between VCDR-associated variants. Effect, SE and P refer to effect sizes, standard errors and p-values of the interaction model.

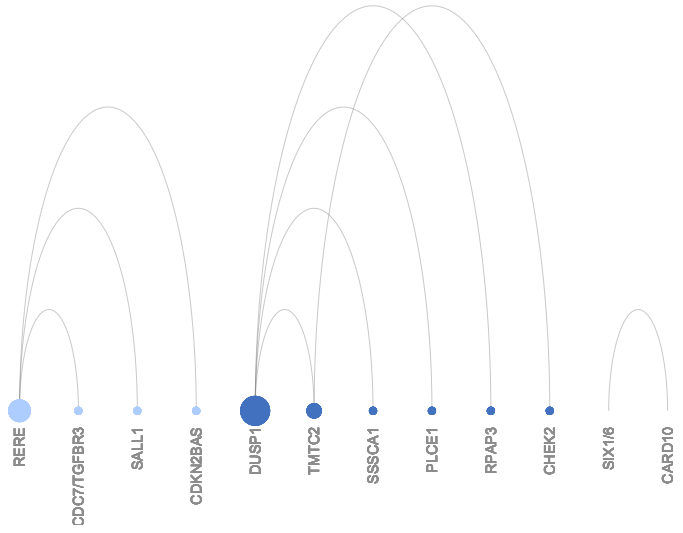


Figure 5.4: SNP-SNP interactions between VCDR genes: this figure shows three separate clusters of interaction between the 17 variants associated with VCDR. The figure only shows the name of the gene in which the variant is located.

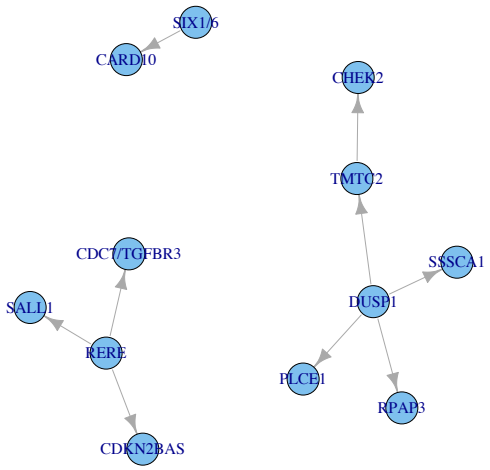


Figure 5.5: SNP-SNP interactions between VCDR genes network: this figure shows genes' clusters for VCDR as networks. Variants are represented with the names of the closest genes.

5.3.2 Intraocular pressure

Similar analyses were performed for IOP. Fifteen cohorts from the IGGC for a total of 26,816 individuals were included in the analysis. Each group was responsible for quality control and analysis for their own cohort, I performed the analyses for the BMES and TwinsUK cohorts.

As for VCDR, another quality control was necessary to assess the quality of the results received. 12 separate analyses were performed for each of the 12 IOP-associated SNPs. Genomic inflation factors were checked for each of the 12 SNPs interaction analysis (Figure 5.6), showing a high variability of λ that varies from 0.2 to 6.645. High variability was due mainly to the Asian-ancestry populations 5.7. For this reason, Beijing Eye Study (BES), Singapore Chinese Eye Study (SCES) and Singapore Malay Eye Study (SIMES) were excluded for further analysis.

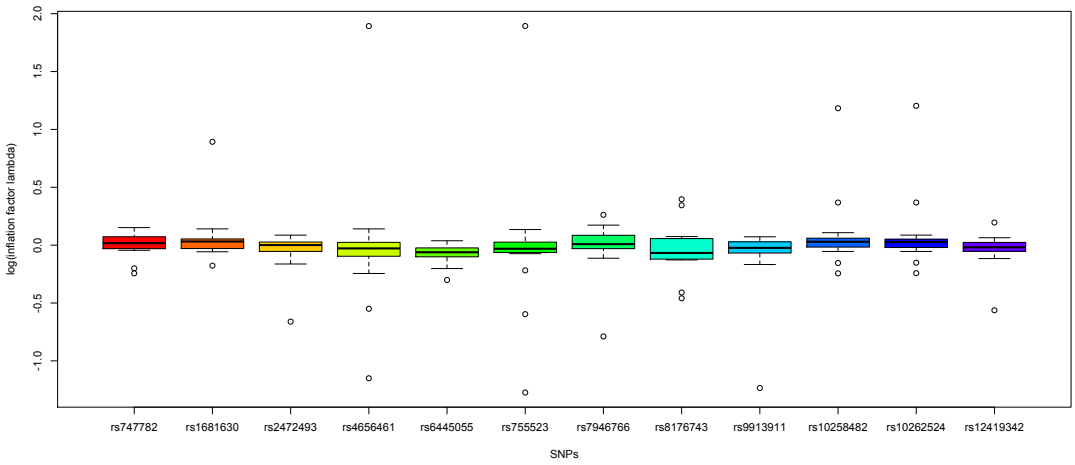


Figure 5.6: IOP genomic inflation factor λ for each SNPs for all cohorts: x-axis shows the 12 SNPs included in the analysis and y-axis the log of the inflation factors λ in all cohorts.

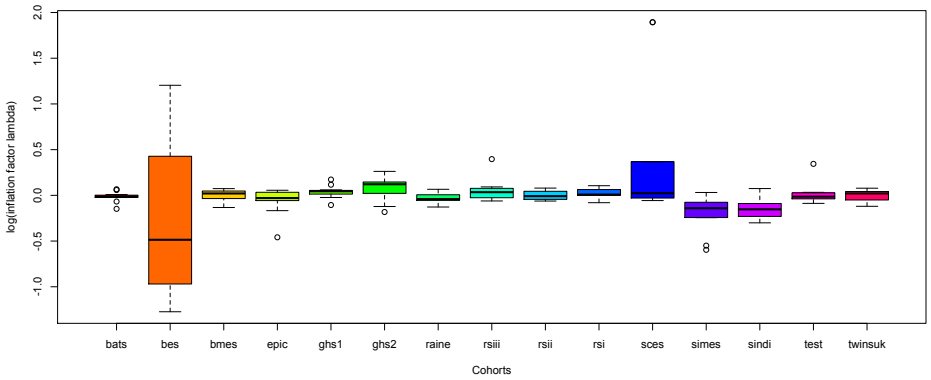
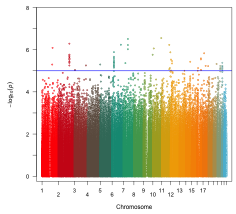


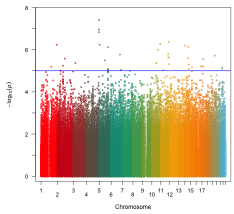
Figure 5.7: IOP genomic inflation factor λ for each cohort for all SNPs: x-axis shows the 15 cohorts included in the analysis and y-axis the log of the inflation factors λ for all SNPs.

Multiple testing correction was necessary due to the high number of test performed and Bonferroni threshold was equal to 1.17E-09. No interaction passed Bonferroni correction (Figure 5.8). However, several loci reached suggestive significance (Table 5.3). The most significant interaction was between rs2211675 (located within *SCAF4/SOD1* gene) and rs7946766 (within *PTPRJ* gene) (β=-0.335, SE=0.058, P=1.27E-08). Another interesting interaction was between rs17607561 (*SLC3A1* gene) and

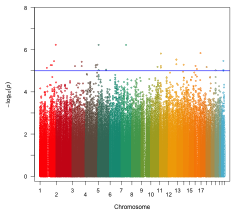
rs8176743 (*ABO* gene) (β=0.626, SE=0.111, P=1.75E-08). Interestingly, the variant located in *CAV1* gene (rs10258482) had significant interactions with varianys lying within *TBC1D1* gene on chromosome 4 and *TOX2* gene on chromosome 20 (respectively β= -0.330, SE=0.059, P=2.63E-08 and β=-0.303, SE=0.054, P=2.77E-08). The variant within *PTPRJ* gene also interacted with the one within *SORBS2* gene (rs1401571, β=0.251, SE= 0.045, P=3.96E-08). Rs4656461 and rs7555523 both near *TMCO1* gene had a significant interaction with rs17124680 (*GNG2* gene) (respectively β=0.455, SE=0.083, P=4.71E-08 and β=0.454, SE=0.083, P=5.26E-08). Finally, the variant in *ABO* gene (rs8176743) had a suggestive interaction with the variant within *DOCK7* gene (β=-0.432, SE=0.080, P=8.09E-08).



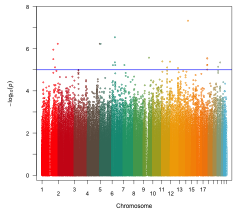
(a) rs747782



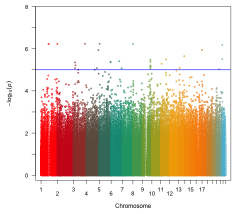
(b) rs1681630



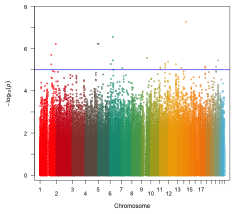
(c) rs2472493



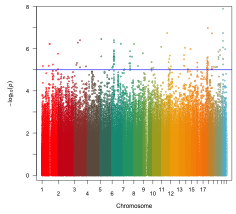
(d) rs4656461



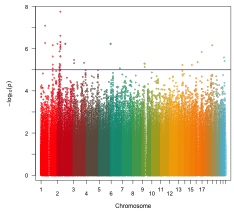
(e) rs6445055



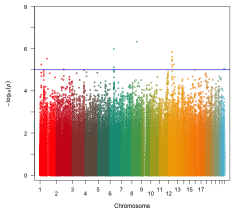
(f) rs7555523



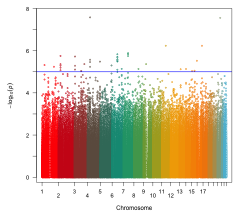
(g) rs7946766



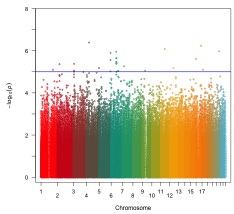
(h) rs8176743



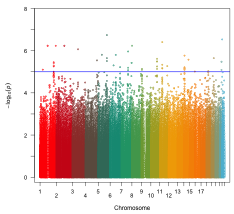
(i) rs9913911



(j) rs10258482



(k) rs10262524



(l) rs12419342

Figure 5.8: Manhattan plots for epistatic meta-analyses with IOP SNPs: the figure shows the 12 different meta-analysis performed for each of the 12 IOP-associated SNP interacting with all the genome. The rs number for each plot represents the different 12 IOP-related SNPs.

Chr1	Gene1	SNP1	Chr2	SNP2	Gene2	Effect	SE	P	I2
11	PTPRJ	rs7946766	21	rs2211675	SCAF4/SOD1	-0.335	0.059	1.27E-08	0.00
9	ABO	rs8176743	2	rs17607561	SLC3A1	0.627	0.111	1.75E-08	0.00
7	CAV1	rs10258482	4	rs1364953	TBC1D1	-0.330	0.059	2.63E-08	0.00
7	CAV1	rs10258482	20	rs17753711	TOX2	-0.304	0.055	2.77E-08	0.31
11	PTPRJ	rs1681630	4	rs1401571	SORBS2	0.252	0.046	3.96E-08	0.18
1	TMCO1	rs4656461	14	rs17124680	GNG2	0.456	0.083	4.71E-08	0.35
1	TMCO1	rs7555523	14	rs17124680	GNG2	0.454	0.083	5.26E-08	0.33
9	ABO	rs8176743	1	rs10493333	DOCK7	-0.433	0.081	8.09E-08	0.03
11	PTPRJ	rs7946766	17	rs8075325	SDK2	0.398	0.075	1.04E-07	0.38
11	PTPRJ	rs1681630	4	rs4862626	SORBS2	0.253	0.048	1.11E-07	0.14
11	PTPRJ	rs1681630	4	rs6814242	SORBS2	0.239	0.045	1.48E-07	0.20
11	PTPRJ	rs7946766	21	rs2833737	SCAF4/SOD1	-0.308	0.059	1.81E-07	0.27
11	PTPRJ	rs7946766	11	rs10897770	LOC101928944	-0.394	0.075	1.82E-07	0.00
11	RAPSN	rs12419342	5	rs283763	P4HA2	-0.300	0.057	1.82E-07	0.00
11	PTPRJ	rs7946766	18	rs2006776	RAB27B	-0.389	0.075	1.87E-07	0.39

Table 5.3: Top results for IOP interaction meta-analysis: CHR1, SNP1 and Gene1 refers to the chromosomes, snnps and genes for the 12 IOP-related SNPs whereas CHR2, SNP2 and Gene2 to the SNPs in all genome that interacted with known gene for IOP

As for VCDR, I checked whether IOP SNPs interacted amongst each other. Variants located within *PTPRJ* gene (rs747782 and rs1681630) significantly interacted with rs9913911 within *GAS7* gene and rs6445055 within *FNDC3B* gene (respectively β=0.115, SE=0.056, P=0.042,β= -0.062, SE=0.028, P=0.030). Variants located in *ABCA1* and *TMCO1* genes also had a significant interaction (rs2472493-rs4656461: β=-0.138, SE=0.066, P=0.037). Graphical representations are shown in Figure 5.9 and Figure 5.10.

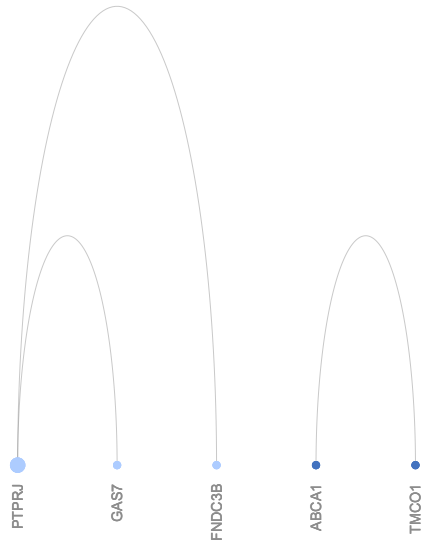


Figure 5.9: SNP-SNP interactions between IOP genes: this figure shows two separate clusters of interaction between the 12 variants associated with IOP. The figure only shows the name of the gene in which the variant is located.

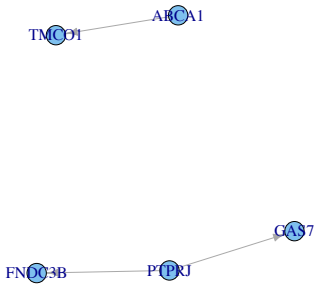


Figure 5.10: SNP-SNP interactions between IOP genes network: this figure shows genes' clusters for IOP as networks. Variants are represented with the names of the closest genes.

5.4 Discussion

The aim of this chapter was to investigate the epistatic interactions between genetic variation associated with POAG endophenotypes in order to identify novel variants that interact with known ones to alter VCDR and IOP.

Several suggestive genetic variants were identified for VCDR. *PLCE1* is a known gene for VCDR and it encodes for a phospholipase enzyme that might affects cell growth and cell differentiation, oesophageal (Wang *et al.* 2010; Wang *et al.* 2014) and gastric cancer (Abnet *et al.* 2010). I found a suggestive significant interaction between a variant within this gene and a variant within *PDE1C* trascript on chromosome 7. This gene is involved in calcium-calmodulin complex (Repaske *et al.* 1992), it regulates soluble adenylyl cyclase/ cAMP signalling and lysosome-mediated collagen 1 protein degradation, in particular regulating collagen homeostasis during pathological vascular remodelling (Cai *et al.* 2011). Interestingly, pde1c has male-specific expression pattern in the central nervous system and affect fertility in *Drosophila melanogaster* (Morton *et al.* 2010). Another interesting interaction affecting VCDR variation was between variants located in *DUSP1* and *PITX2* genes. DUSP1 gene is known to affect VCDR and it is a MAPK negative regulator influencing the regulation of the host immune response (i.e. for poxvirus infection) (Cáceres *et al.* 2013). The protein of this gene is inducible by oxidative stress and heat shock and may play a role in environmental stress response and participates in the negative regulation of cell proliferation (Keyse & Emslie 1992). *PITX2* gene is a member of the homeobox protein family, whose transcription factor plays a critical role in early development, especially of the anterior segment of the eye (Gage *et al.* 2005). Mutations in *PITX2* gene are involved in Axenfeld Rieger syndrome which affects mostly the eye but also

has systemic manifestations. Knock-out mice die from cardiac pathology and have several eye defects, including loss of extraocular muscles (Gage *et al.* 2005). In zebrafish models, alteration of pitx2 produces disruption of neural crest migration into the craniofacial region, malformation of the jaw and corneal endothelium iris stroma. It also acts with *CYP1B1* gene and retinoic acid to specify a population of neural crest-derived cells for migration and differentiation (Bohnsack *et al.* 2012). Interestingly it also seems to be involved in atrial fibrillation with *CAV1*, a POAG gene (Martin *et al.* 2015). Another interesting interaction was found between variants located in *CDC7/TGFBR3*, which are involved in cellular growth and cancer, and *EFHB* genes. The latter is involved in calcium-ion binding and it is associated with type 1 diabetes (Howson *et al.* 2009). Another gene for VCDR *SALL1* had a significant interaction with *PCDH7* gene. *SALL1* encodes for a zinc finger transcriptional repressor involved in ocular development, neural tube closure in mice (Böhm *et al.* 2008) and a congenital anomaly-mental retardation syndrome that causes central nervous system defects and cortical blindness (Vodopiutz *et al.* 2013). *PCDH7* encodes an integral membrane protein, involved in cell-cell recognition and adhesion, calcium-ion binding and axon guidance (Leung *et al.* 2013). This gene has also been identified through GWASs associated with sleep (Ollila *et al.* 2014), epilepsies (on Complex Epilepsies 2014) and musical aptitude and ear development (Oikkonen *et al.* 2015). It has been also involved in cell division in post-natal and adult mice' amygdala (Hertel *et al.* 2012). Finally, I identified another novel gene, *C2CD5* whose variant had a significant interaction with a variant within *SIX1/6* gene. *C2CD5* encodes an insulin-stimulated glucose transporter.

To investigate whether known VCDR genes interact with each other, I focused on interactions between the 17 SNPs associated with VCDR variation. The interaction is not affected by the strength or significance of these variants with VCDR, but it is only based on the statistical interaction between the two variants. The SNP in *DUSP1* gene interacted with variants located in *PLCE1*, *SSSCA1*, *RPAP3* and *TMTC2* genes, which also is connected with *CHEK2* gene. These interactions underline the role of cell growth and cell differentiation in VCDR variation, supporting the findings in our previous GWAS meta-analysis paper (Springelkamp *et al.* 2014). The SNP within *RERE* gene had significant interactions with variants located in *SALL1*, *CDKN2BAS* and *CDC7/TGFBR3* genes, suggesting an influence during embrionic development on VCDR variation. Indeed, *RERE* gene is involved in eye and cerebellum development (Kim *et al.* 2013; Kim & Scott 2014), *SALL1* gene in neural tube development (Böhm *et al.* 2008) and *CDC7/TGFBR3* gene in cell growth.

Suggestive interactions between IOP genetic variations and others in all genome were identified, allowing the discovery of new potential genes that may affect IOP variation in the general population. The SNP in *PTPRJ* gene had suggestive interactions with variants located in *SCAF4/SOD1* and *SORBS2* genes. *PTPRJ* gene is involved in a variety of cell processes including cell growth, differentiation, mitotic cycle and oncogenic transformations. In particular, it might play a role in cancer cell proliferation (Ortuso *et al.* 2013) and, interestingly, in endothelial cell permeability induced by vascular endothelial growth factor (Spring *et al.* 2014). The role of *SCAF* gene is mostly unknown and its possible role in IOP variation not clear. It has been associated with diabetic nephropathy (Alwohhaib *et al.* 2014). *SOD1* is a known gene for ocular colobomas ,

brain malformation and Alzheimer's disease (Spisak *et al.* 2014) and it is well studied for amyotrophic lateral sclerosis (Rotunno *et al.* 2014). This gene, which is ubiquitously expressed and which genomic sequence is highly conserved across species, binds copper and zinc ions and forms a homodimer. Its role is of defence of oxygen toxicity (Saccon *et al.* 2013). Lack of sod1 increased the vulnerability to stress of motor neurons leading to cell death and axon injury (Saccon *et al.* 2013). Interestingly, knock-out mice exhibit a contractile vascular phenotype with impairment of endothelial function during ageing (Didion *et al.* 2006). *SORBS2* gene encodes an adapter protein that assemble signalling complexes in stress fibres and it is localised in epithelial and cardiac muscle cells: it is released from damaged cardiac tissue into the bloodstream upon lethal acute myocardial infarction (Kakimoto *et al.* 2013). It has been also associated with diabetic nephropathy (Nakatani *et al.* 2011). A variant within *ABO*, blood group gene, had showed suggestive interactions with variants within SLC3A1 gene, which encodes a renal amino acid transporter and it is involved in cystinuria (Miyamoto *et al.* 1996), and *DOCK7* gene. The latter is crucial for axon formation (Pinheiro & Gertler 2006), neural polarisation (Watabe-Uchida *et al.* 2006) and encepholopathy and cortical blindness (Perrault *et al.* 2014). Another suggestive interaction was between variants located in *TMCO1* and *GNG2* genes. The latter is important for signalling mechanisms across membranes. It is involved in proliferation of malignant melanoma cells in humans (Yajima *et al.* 2012). The knockdown of the G protein gng2 blocks the normal angiogenic process in developing zebrafish embryos. In particular, loss of function of this protein inhibits the ability of VEGF (vascular endothelium growth factor) to promote the angiogenic sprouting of blood vessels (Leung *et al.* 2006). Finally, suggestive interactions between variants located in *CAV1* and *TOX2* and

TBC1D1 genes were identified. *TOX2* gene regulates human natural killer cell development by controlling T-BET expression (Vong *et al.* 2014) and it has been associated with dyastolic blood pressure in a gene-environment association study (Basson *et al.* 2014). *TBC1D1* gene is crucial for cell growth regulation and differentiation and it is important in human skeletal muscle (Cartee 2010).

I also checked whether IOP associated genes interacted with each other. A variant within *PTPRJ* gene had nominal significant interaction with variants *GAS7* and *FNDC3B* genes, suggesting a role of cell growth and neuronal differentiation in IOP.

There are potential limitations of the study. Firstly, statistical interaction does not necessarily imply interaction at the biological level. Functional epistasis (i.e. protein-protein interaction) is a ubiquitous component of the pathogenesis of any disease, however this does not mean this epistasis will be detected as statistical interaction, especially if genetic variations examined are surrogates for true causal variants because of linkage disequilibrium (Cordell 2009). Despite this limitation, allowing for interaction can lead to improved power to detect genetic factors underlying the trait or disease (Cordell 2002). Another limitation is that gene-gene interaction does not strictly help to explain the remain missing heritability we cannot explain with GWAS. GWAS try to describe narrow sense heritability, which is made up of additive variance. By definition, since gene-gene interaction is a deviation of the additive model, it would not help to explain narrow sense heritability but should be part of the unknown heritability (Visscher *et al.* 2008). However, Herani *et al.* has

suggested the epistasis could also contribute to additive variance by statistical illusion (confounding non additive and environmental effects in twin study calculation) and creating real additive variation as marginal effects (Hemani *et al.* 2013). In this study, Bonferroni correction was used to filter significant results, even though tests are not independent of each other and thus it may be too conservative. Other corrections methods, such as permutation, are computationally prohibitive and the choice of specificity over sensitivity was made. This study also focused on two-locus interaction rather than high-order interactions. However an exhaustive search of all three-way or four-way or any higher level interactions seems impractical in a genome-wide and meta-analysis settings.

5.5 Conclusion

To summarise, findings suggest that mechanisms such as cell growth, neuronal differentiation, calcium channels and axon guidance might alter VCDR, thus playing a role in the pathogenesis of POAG. The role of neural development was confirmed also in IOP variation, providing evidence for a link between increased IOP and optic nerve alterations in POAG. IOP is also affected by vascular and systemic processes, including vascular endothelium mechanism, blood pressure and diabetes. To my knowledge this is the first study investigating epistatic interactions in POAG endophenotypes at the meta-analysis level. The present findings extends our knowledge of mechanisms underlying VCDR and IOP, thus POAG, and will serve for future studies to better understand POAG mechanisms.

CHAPTER 6

Novel genetic loci and Alzheimer's disease genes influence retinal nerve fibre layer thinning: an European ancestry meta-analysis of 10,502 individuals

In this chapter I will describe and discuss my research on RNFL thickness genetics in the BMES cohort and other European-ancestry populations. I will also investigate the overlap between RNFL thickness and Alzheimer's disease.

STATEMENT OF CONTRIBUTION TO THIS RESEARCH:

I designed the study and performed all the analysis described in this chapter. This work will be presented at ASHG 2015 as a poster.

6.1 Background

The retina is the neurosensory component of the eye and it is constituted by different cell elements and layers to meet all functional needs. The ganglion cell axons travel towards the optic nerve head within the retinal nerve fibre layer (RNFL). Ganglion cell axons, as all the other axons in the brain, are supported by glia, in particular astrocytes and Muller cells (Hildebrand & Fielder 2011).

Several technologies allow the measurements of the retinal ganglion layers and in particular of RNFL layer. For instance, HRT (Heildenberg Retina Tomograph) is a confocal scanning laser that takes images of serial sections of the optic nerve to create a map of topographical sections of the optic nerve head and the surrounding retina. HRT can measure the changes and loss of RNFL (Belyea *et al.* 2014).

Optical coherence tomography (OCT) has revolutionised ophthalmology and it has become a very important tool in clinical practice. Time-domain OCT has been recently replaced by the newer Spectral-Domain OCT technology that has faster acquisition time and higher resolution. OCT provides high-resolution images of the peripapillary RNFL, macular volume, macular ganglion cell layer and optic nerve (Rebolleda *et al.* 2015). Both technologies are quick and efficient, allowing a sensitive and non-invasive measurement of the axonal integrity in the retina.

RNFL evaluation is a useful tool in early glaucoma diagnosis (Jonas & Dichtl 1996). Glaucoma involves the loss of axons of the ganglion cells and RNFL loss can be detected even before functional loss (Kass *et al.* 2002). Airaksinen *et al.* studied 51 individuals with glaucoma, 52 with ocular hypertension and 29 controls. Reduction of nerve fibres was seen in 48 of 51 individuals with glaucoma, 27 of the 52 with high intraocular pressure, but only 7 in the controls (Airaksinen *et al.* 1984). Similarly, Sammer *et al.* checked 1,400 eyes divided into those with glaucomatous with visual field loss and controls and showed that assessing RNFL thickness has a 80% sensitivity (proportion of glaucomatous eyes with nerve fibre loss) and 94% specificity (proportion of normal control eyes without defects) (Sommer *et al.* 1984).

Family history of glaucoma is a risk factor for thinner RNFL, together

with age, male gender and axial length (Rougier *et al.* 2015). Genetics is likely to play a role in RNFL thickness variation. Heritability ranges between 40% and 49% in different sectors of RNFL in 1552 Dutch individuals (Koolwijk *et al.* 2007). However, only a very few genetic studies have been published to this date. Axenovich *et al.* conducted a linkage and association studies on a family-based Dutch population (ERF) with a total of 1,434 (for linkage study) and 1,488 (for association study) individuals (Axenovich *et al.* 2011). Linkage study on RNFL thickness identified a peak region on chromosome 14 which include *SIX1*, a gene associated with cup/disc ratio (Ramdas *et al.* 2010). *DCLK3* gene on chromosome 3 was identified by the same authors through GWAS. This gene is a member of the doublecortin family, which is involved in neuronal migration, neurogenesis and eye-receptor development (Axenovich *et al.* 2011).Further genetic studies have focused on the role of *SIX1/SIX6* in RNFL thickness. A recent quantitative trait locus (QTL) analysis in 231 participants from the Diagnostic Innovations in Glaucoma Study and African Descent and Glaucoma Evaluation Study investigated rs10483727 (*SIX1/SIX6*) and found that each copy of the T allele decreased the thickness of RNFL by -0.16 μm. Likewise Cheng *et al.* have investigated the variant rs33912345 (*SIX6*) in 2,129 eyes from 1,243 individuals of the SCES study and showed that the C allele of the *SIX6* His141 decreases the RNFL thickness by 1.44 μm (Cheng *et al.* 2015). *SIX6* gene is expressed in neural retina and optic nerve during embryonic development (Conte *et al.* 2010). Missense variants in *SIX6* gene have been shown to reduce eye size and optic nerve volume in a zebrafish model (Iglesias *et al.* 2014).

Besides the application in ophthalmology, RNFL thickness has also been investigated as a possible bio-marker in various neurological conditions.

Alzheimer's disease (AD) is the most common degenerative dementia and causes a progressive decline in several cognitive functions (Blennow *et al.* 2006). AD affects 13% of people over 65 years old and 30%-50% over 80 years old (Hebert *et al.* 2003). The disease manifests with progressive decline in memory and other cognitive functions, leading to loss of autonomy.

Interestingly, visual impairment is one of the earliest complaints in AD patients (Valenti 2010). Visual functions such as contrast sensitivity in lower spatial frequency (Cronin-Golomb *et al.* 1991), motion perception (Thiyagesh *et al.* 2009; Mapstone *et al.* 2008; Gilmore *et al.* 1994), visual field (Armstrong 1996; Trick *et al.* 1995) and colour discrimination (Cronin-Golomb *et al.* 1993) decreased in AD, showing a similarity with the visual deficits in glaucoma (Quigley *et al.* 1982; Shabana *et al.* 2003). Patients with AD and glaucoma show a more rapid and aggressive glaucomatous visual field loss (Bayer & Ferrari 2002). Both AD and glaucoma affect the visual pathway, but start in different region along the neural pathway. AD might start in the visual association area (McKee *et al.* 2006) whereas glaucoma has its initial damage in the optic nerve. In addition, AD treatments have been shown to have an impact on glaucoma. ChEI (cholinesterase inhibitors) treatment, commonly prescribed for AD, lowers IOP and seems to be protective for retinal ganglion cells (Estermann *et al.* 2006; Miki *et al.* 2006). Memantine is a neuroprotective drug for AD and also have implications for the visual system. Memantine has been investigated also for the treatment of glaucoma and in monkeys with glaucoma and it slows down the progression of cell loss in lateral geniculate nucleus (LGN) compared with those animals not treated (Yücel *et al.* 2006).

Epidemiological studies have shown the link between the two diseases. Two groups whose 112 with AD and 774 without have been compared and 25.9% of the people with AD developed also glaucoma, but only 5.2% of the group without AD. The authors suggested that the optic nerve might be less resistant to IOP levels in AD patients (Bayer *et al.* 2002). Likewise, Tamura *et al.* investigated 172 AD patients in Japan compared with 156 controls and found that those with AD has a greater rate of glaucoma (23.8%) compared with the controls (9.9%) (Tamura *et al.* 2006).Even though AD and glaucoma might be two different degenerative processes affecting the visual pathway, surely some structures and functions are similar, for example there is evidence that AD affects the RNFL layer. Parisi *et al.* compared RNFL thickness in 17 AD and 14 control individuals, showing that RNFL is thinner in participants with AD (Parisi *et al.* 2001). Another study showed a likewise result with decrease of RNFL thickness in AD and also correlated the thickness of RNFL with severity of cognitive function loss (Iseri *et al.* 2006). Interestingly, also mild-cognitive impairment (MCI), a condition characterised by slight, but noticeable decline in cognitive skills and associated with AD risk, showed a decrease of RNFL thickness. Paquet *et al.* analysed 14 mild AD, 23 MCI and 15 controls and found a thinner RNFL in AD and MCI in comparison with controls (Paquet *et al.* 2007).

Loss of nerve fibre layer tissue in the retina and optic nerve may be an early bio marker for AD and may appear even before any hippocampus damage, which is the brain structure that impacts memory (Valenti 2011). Given that AD diagnosis is complicated and can be only confirmed in post-mortem examination, RNFL assessment may become an useful and precise tool (and also non invasive) and it is been showed to be at least as sensitive and specific as current methods for AD diagnosis (Larrosa *et al.*

2014).

Considering the previous points, the objectives of this chapter are:

- to clarify the relationship of RNFL thickness with POAG endophenotypes in order to understand whether RNFL thickness can be used as a POAG quantitative trait itself
- to investigate common genetic factors and related biological pathways underlying RNFL thickness variation in the general population given the lack of genetic studies on RNFL thickness so far
- to explore the genetic link between RNFL thickness and AD in order to identify possible shared biological mechanisms that can firstly elucidate the relationship with POAG and also can help to clarify the role of RNFL thickness as a marker of AD

6.2 Methods

To assess the link between RNFL thickness and other POAG traits correlation and linear regression models were used in the BMES cohort. RNFL thickness was available in 960 individuals for the right eye and details about the measurement are described in Chapter 2.1.1.

A GWAS using RNFL thickness as outcome was then performed in the BMES cohort adjusting for age and sex as described in Chapter 2.2.2. Inverse-based fixed effect meta-analysis was performed in 10,502 individuals from seven European-derived cohorts: BMES cohort, Gutenberg Health Study 1 and 2, Erasmus Rucphen Family Study and Rotterdam 1,2,3.

To investigate a possible genetic link between RNFL thickness and AD several analyses were performed. A list of 435 AD-associated variants was compiled from the Alzheimer Disease Genetics Consortium as described in Chapter 2.2.7. A second list of 50 SNPs that were associated both with AD and RNFL thickness was compiled. Effect sizes from AD and RNFL meta-analyses were compared (Chapter 2.2.7). To evaluate whether AD-associated variants SNPs have an aggregate role on RNFL thickness in the BMES cohort, two polygenic risk score analyses were performed. The first risk score included 37 SNPs, selected from the list of 435 variants associated with AD , whereas the second one included 16 independent SNPs selected from the 50 SNPs associated with both AD and RNFL thickness (Chapter 2.2.7.1). Regression analysis were performed with the scores as predictors in separated models and RNFL thickness as outcome, adjusting for age and sex. These models were compared using r-squares and ANOVA with the null model of a regression including only age and sex as predictors.

Given the small sample size in the BMES cohort (960 individuals), a multi-SNP score analysis, which allows summary results from both RNFL thickness meta-analysis and AD meta-analysis, was performed. The risk score was weighted using the AD effects for the 16 independent SNPs, previously selected to be significant in both AD and RNFL thickness meta-analysis and not in LD (Chapter 2.2.7.2).

6.3 Results

6.3.1 RNFL thickness and POAG endophenotypes in the BMES cohort

A total of 960 individuals from the BMES cohort had genotypic and phenotypic information and were used for further analysis. RNFL thickness measurements were available in mm only in the right eye. The distribution was approximately normally distributed (Figure 6.1) with a mean of 0.23 mm (SD=0.07) and median 0.22 mm and the values range from 0.02 mm to 0.83 mm. Given this, raw data were used for further analyses and no transformation was applied.

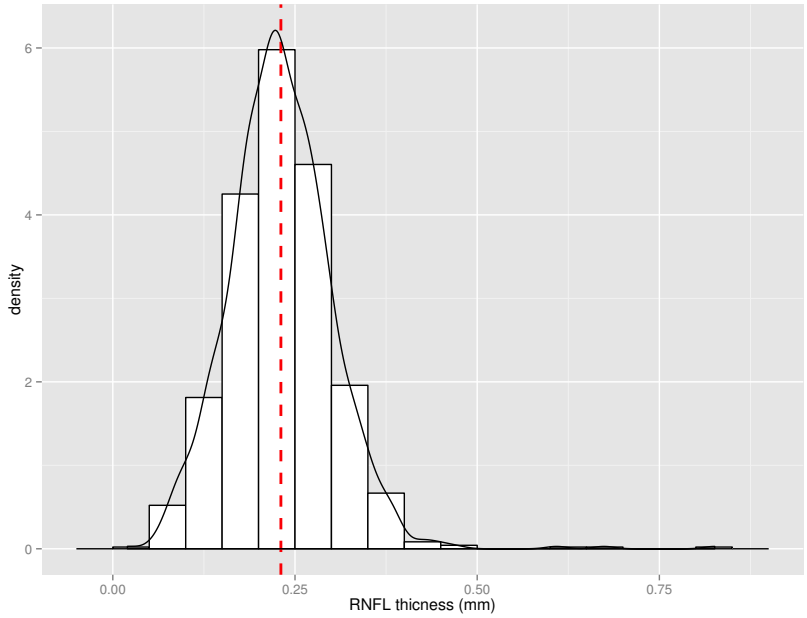


Figure 6.1: RNFL thickness distribution in the BMES cohort: the figure displays the histrograms and the density curve for the RNFL thickness phenotypes measured in mm. The vertical red line represent the mean of the distribution.

To better understand the role of RNFL in POAG pathogenesis, the epidemiological link between RNFL thickness and POAG quantitative traits was explored. Thus, correlation between RNFL thickness and the other

POAG phenotypes was investigated in the BMES cohort. The highest correlation was found with optic disc measurements: VCDR (-0.144) and VCD (-0.134) (Table 6.1), showing that an increase of the diameter of the cup disc is associated with thinner RNFL. The correlation between RNFL and IOP was -0.075 and 0.035 with refractive error.

Phenotypes	Pearson coefficient
IOP	-0.075
VCDR	-0.144
SE	0.035
VDD	-0.020
VCD	-0.134
CCT	0.001

Table 6.1: Correlation between RNFL thickness and POAG endophenotypes in the BMES cohort: the table shows the Pearson's coefficients for the correlation analysis in 960 individuals analysed.

To further investigate the link between RNFL and POAG endophenotypes, a univariate linear model was fitted with RNFL thickness as outcome and POAG quantitative traits as predictors. All analyses were adjusted for age and sex. VCDR and VCD were associated (respectively P= 0.00002 and P=0.00007) even though with a small effect on RNFL thickness (β=-0.07 for VCDR and β=-0.0004 for VCD). An increase of IOP was also significantly associated with thinner RNFL (P=0.02) with a small effect size (β=-0.002). In addition, refraction was shown to be associated with RNFL thickness (β=0.002 and P=0.03).

Phenotype	Estimate	SE	P
IOP	-0.002	0.0008	0.020
VCDR	-0.076	0.0180	2.19E-05
VDD	-8.77E-05	1.42E-04	0.53
SE	0.002	0.0012	0.03
VCD	-3.94E-04	9.88E-05	7.10E-05
CCT	-4.23E-06	8.21E-05	0.95

Table 6.2: Results for linear models fitted with RNFL thickness as outcome and POAG endophenotypes as predictors in the BMES cohort: this tables shows the results for a linear regression on RNFL thickness. Estimate: effect size, SE: standard error, P: p-value

6.3.2 The genetics of RNFL thickness

In order to explore potential genetic influences of RNFL thickness in the general population, a genetic association study was conducted. Firstly, a GWAS for RNFL thickness in the BMES was performed, then a meta-analysis with other European ancestry populations was carried out.

6.3.2.1 GWAS for RNFL thickness in the BMES population

A set of 960 individuals with information for a panel of 2,297,195 SNPs was analysed. All of these SNPs passed stringent quality control criteria as described in the Chapter 2.

No SNPs reached genome-wide significance for RNFL thickness (Figure 6.2). However, suggestive association signals have been found in several new loci. The SNP that reached the highest significance is rs1154119 (P=2.41E-07). This SNP lies within the genomic sequence of *SLC25A21* gene (chromosome 14, Figure 6.3). The second SNP that reached the highest significance is rs2493383 (P=9.82E-06), located in an intergenic

region on chromosome 6 (Figure 6.4). Another suggestive locus is on chromosome 14 (rs6571668, P=1.67E-06, MAF= 0.44) close to *SNX6* gene (Figure 6.5). This gene is involved in several stage of intracellular trafficking and it promotes lysosomal degradation of *CDKN1B*.

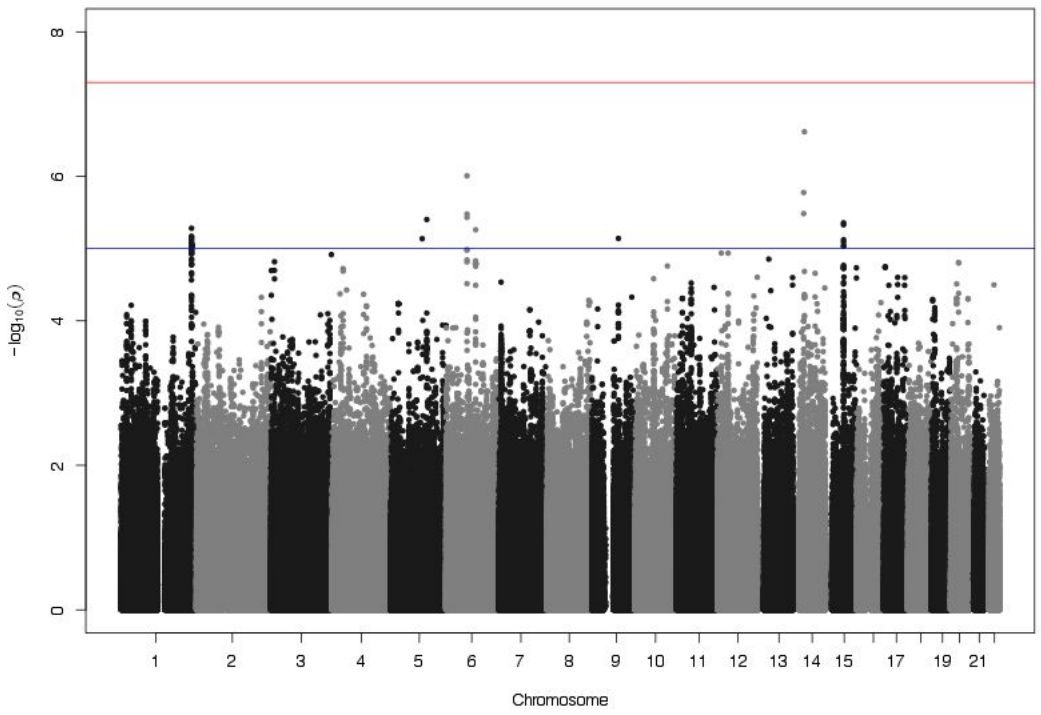


Figure 6.2: Manhattan plot for RNFL thickness in the BMES population: the 22 autosomes are plotted along the x-axis, whereas the values in the y-axis denote the -log 10 transformed p-values from the GWAS in the BMES cohort on RNFL thickness

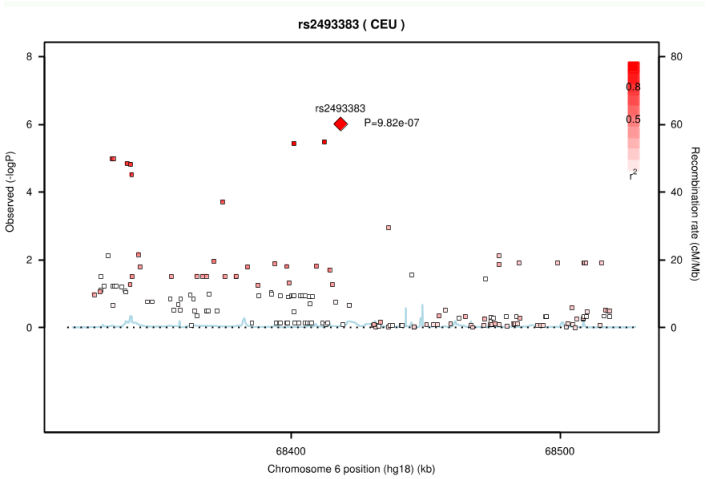


Figure 6.3: Regional association plot for rs2493383: the figure shows the regional information of rs2493383 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

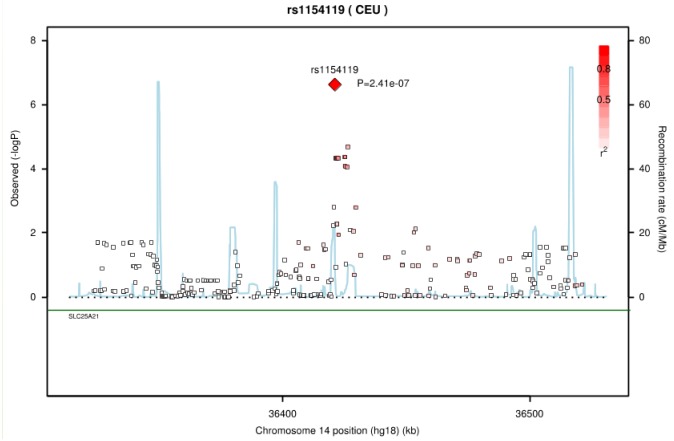


Figure 6.4: Regional association plot for rs1154119: the figure shows the regional information of rs1154119 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

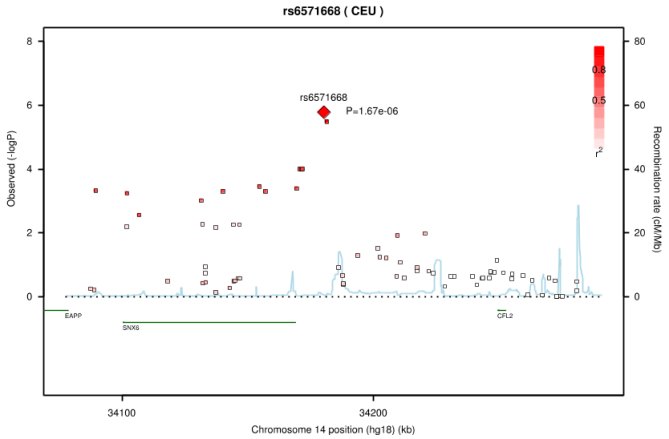


Figure 6.5: Regional association plot for rs6571668: the figure shows the regional information of rs6571668 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

6.3.2.2 Meta-analysis for RNFL thickness

As larger sample sizes are needed to unravel common variants underlying complex diseases, GWAS results on RNFL thickness for 10,502 individuals from 7 populations of European ancestry were meta-analysed. The meta-analysis included BMES, Gutenberg Health Study 1 and 2, and four Dutch studies, Rotterdam 1,2 and 3 and Erasmus Rucphen Family study.

The results showed that the genomic inflation factors λ in individuals studies ranged between 0.9-1.01, indicating within-study control of population structure (Figure 6.6).

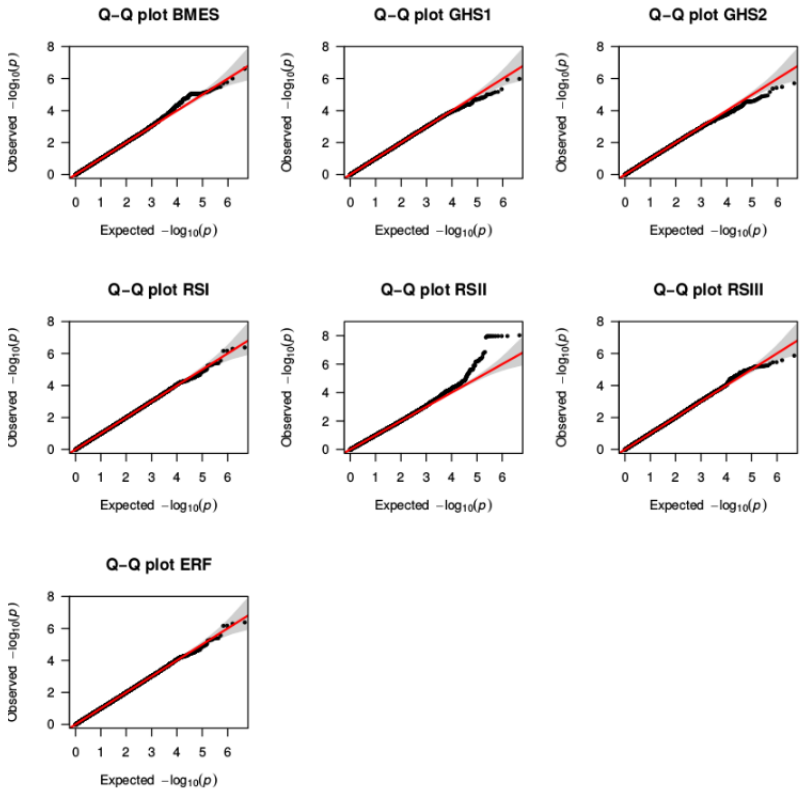


Figure 6.6: Q-Q plot for all the 7 cohorts included in the meta-analysis: the figure shows the -log10 of the p-values for each study (in black) plotted with the expected value of no association (red line). In grey are the confidence intervals.

Four SNPs crossed the conventional genome-wide significance threshold of P=5E-08, as shown in the Manhattan plot Figure 6.7.

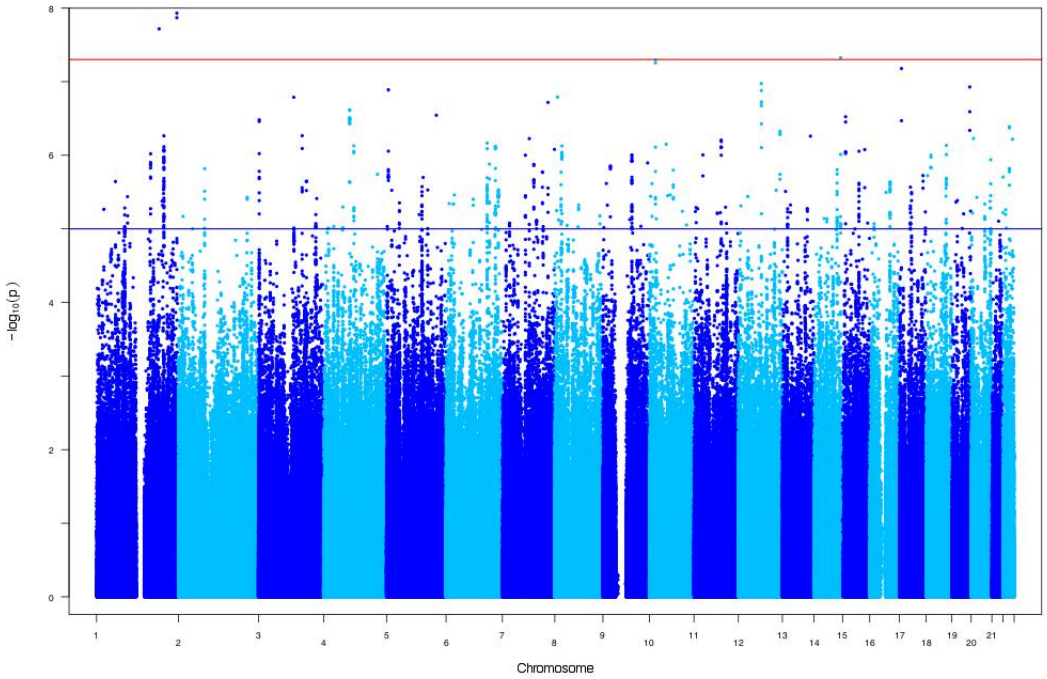


Figure 6.7: Manhattan plot for RNFL thickness meta-analysis results: the 22 autosomes are plotted along the x-axis, whereas the values in the y-axis denote the $\log 10$ transformed p-values from the meta-analysis on RNFL thickness. The red line represents the genome-wide significance line (P=5E-08), whereas the blue line shows the suggestive significance line (1E-05).

Rs	Chr	Position	A1	Eaf	Effect	SE	P	Dir	Gene
rs7554059	1	244460825	G	0.890	0.014	0.002	1.17E-08	++++-+?	C1orf100
rs6429459	1	244460570	C	0.944	0.015	0.003	1.35E-08	++++-+?	C1orf100
rs12087686	1	190481726	C	0.054	0.013	0.002	1.92E-08	++++-+?	BRINP3
rs11625568	14	101230885	A	0.959	-0.014	0.003	4.72E-08	- - - - - +?	DLK1
rs1277751	10	18522260	T	0.832	0.007	0.001	5.04E-08	+++++++	CACNB2
rs1277752	10	18522474	A	0.833	0.007	0.001	5.54E-08	+++++++	CACNB2
rs1966778	17	6434680	C	0.943	0.012	0.002	6.63E-08	+++++++	PITPNM3
rs12322558	12	70080152	G	0.108	-0.008	0.002	1.06E-07	- - - - - + -	BEST3
rs11879136	19	54760690	T	0.250	0.008	0.001	1.18E-07	++++++?	LILRB5
rs2913316	5	5597977	G	0.727	-0.006	0.001	1.29E-07	- - - - - + -	NA

Table 6.3: Most significant variants for RNFL thickness meta-analysis: the table shows the results of the RNFL thickness meta-analysis. Rs: marker name, Chr: chromosome, A1: effect allele, Eaf: effect allele frequency, Effect: effect size, SE: standard error, P= p-value, Dir: direction of the effect for all cohorts, Gene: closest gene

As shown in Table 6.3, the strongest signal was found at rs7554059 and rs6429459 (respectively P=1.17E-08 and P=1.35E-08) on chromosome 1. The minor allele frequency of these variants is respectively of 0.43 and 0.33 in the European population. The variants lie on an intergenic region, however they show regulatory activity and the closest genes are *C1orf100* (chromosome 1 open reading frame 100) and *ZBTB18/ZNF238* (zinc finger and BTB domain containing 18). On the same chromosome (1), another signal was identified at rs12087686 (minor allele frequency= 0.14, P=1.92E-08) in high regulatory intergenic region close to *BRINP3* (bone morphogenetic protein/retinoic acid inducible neural-specific 3) transcript. *BRINP3* is protein coding gene involved in neuronal cell proliferation by negative regulation of the cell cycle transition. On chromosome 14, I pinpointed a significant genetic signal at rs11625568 lying nearby *DLK1* transcript. This gene encodes a transmembrane protein containing six epidermal growth factor repeats, which is involved in the differentiation of several cell types, including adipocytes. It is also thought to be a tumor suppressor. Other interesting signals were found on chromosome 10: rs1277751 and rs1277752 (respectively P= 5.04E-08 and P=5.54E-08) with a very consistent effects' direction amongst all cohorts. These SNPs are both lying within *CACNB2* transcript (calcium channel, voltage-dependent, beta 2). This gene encodes a subunit of a voltage-dependent calcium channel protein that is a member of the voltage-gated calcium channel super-family. Additionally,rs1966778 on chromosome 17 was identified (P=6.63E-08) with again a high consistency of direction of effects for all cohorts. This variant lies within *PITPNM3* gene (Phosphatidylinositol Transfer Protein, Membrane-Associated 3). This gene encodes a member of a family of membrane-associated phosphatidylinositol transfer domain-containing proteins. Other suggestive signal was found

at rs12322558, located within *BEST3* gene (vitelliform macular dystrophy 2-like 3 isoform), which belongs to the bestrophin family of anion channels. Finally, rs11879136 was identified on chromosome 19 within *LILRB5* gene (leukocyte immunoglobulin-like receptor). This gene is a member of the leukocyte immunoglobulin-like receptor (LIR) family, which is found in a gene cluster at chromosomal region 19q13.4.

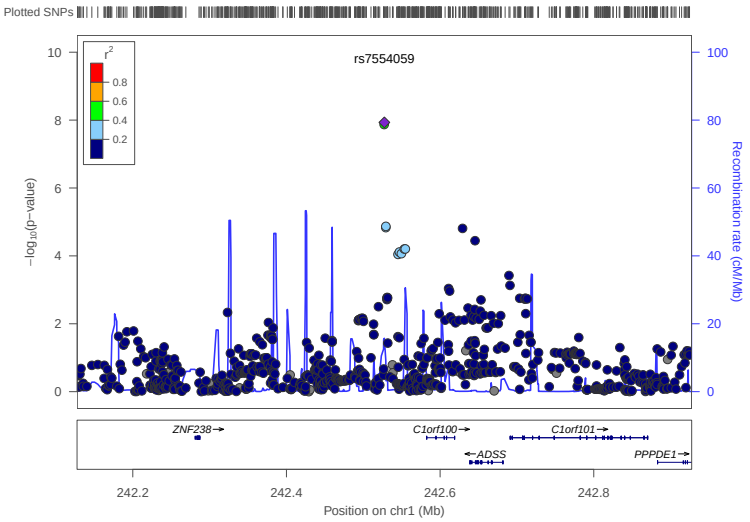


Figure 6.8: Regional plot for rs7554059: the figure shows the regional information of rs7554059 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

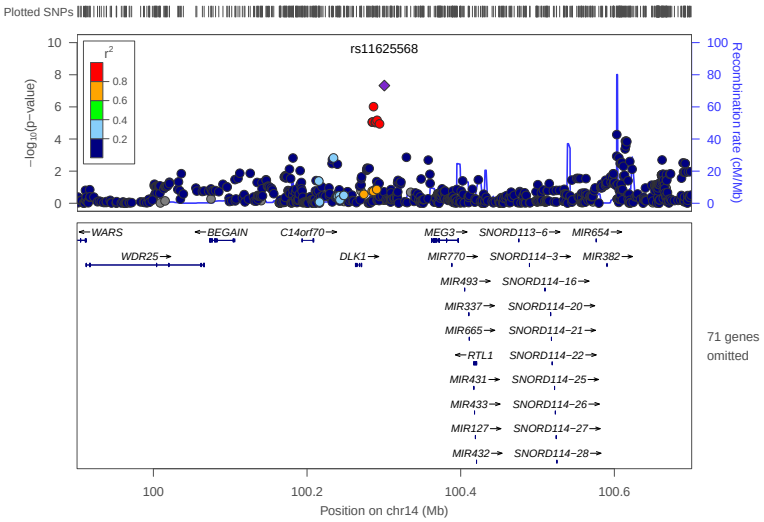


Figure 6.9: Regional plot for rs11625568:the figure shows the regional information of rs11625568 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

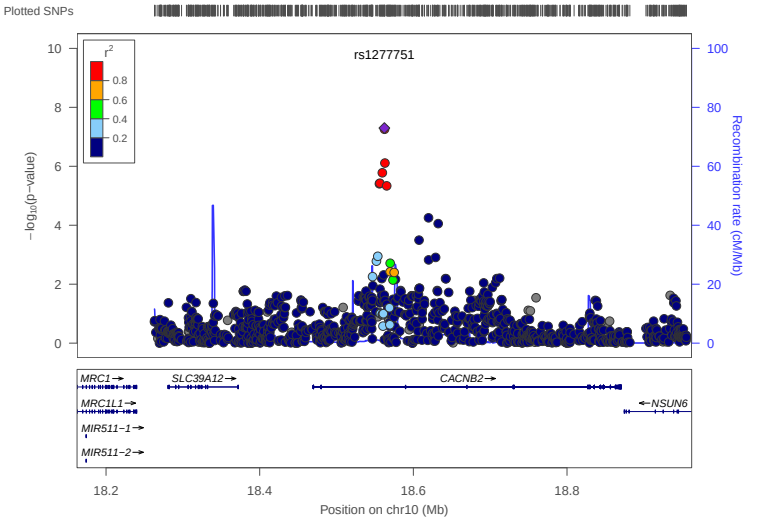


Figure 6.10: Regional plot for rs1277751: the figure shows the regional information of rs1277751 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

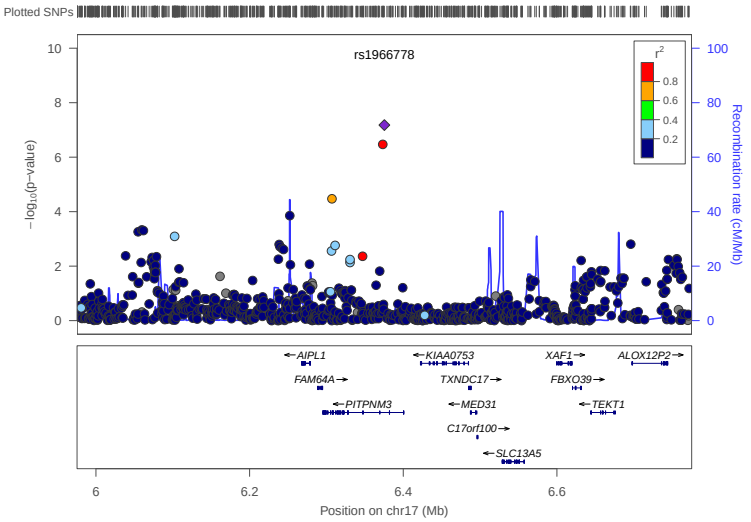


Figure 6.11: Regional plot for rs1966778: the figure shows the regional information of rs1966778 such as the strength of the association (y-axis) relative to genomic position, local linkage disequilibrium, represented by the red colour and recombination patterns (blue lines).

Given the relationship between RNFL thickness and POAG, the ten most significant SNPs from RNFL thickness meta-analysis were checked for magnitude of effects and significance in IOP and VCDR meta-analysis (Chapter 3). For VCDR only rs12322558 (*BEST3*) reached significance in the IGGC results (see Table 6.4). However, for IOP none of signal identified in RNFL meta-analysis reached nominal significance in IGGC results (Table 6.5).

Rs	A1	Eaf	Effect	SE	P
rs12087686	G	0.932	0.003	0.003	0.332603
rs6429459	C	0.944	0.003	0.004	0.424393
rs7554059	G	0.903	0.003	0.003	0.346288
rs11625568	A	0.957	-0.000	0.004	0.930275
rs1277752	A	0.829	-0.002	0.002	0.379816
rs1277751	C	0.219	0.000	0.002	0.937081
rs1966778	C	0.945	-0.002	0.003	0.582203
rs12322558	A	0.884	-0.007	0.002	0.000211
rs11879136	T	0.248	-0.000	0.002	0.873969
rs2913316	G	0.740	0.002	0.002	0.121232

Table 6.4: The most significant variants identified in the RNFL thickness meta-analysis in the VCDR meta-analysis: Rs: marker name, A1: effect allele, SE: standard error, P: p-value

Rs	A1	Eaf	Effect	SE	P
rs12087686	G	0.917	0.002	0.047	0.962
rs6429459	C	0.853	0.017	0.057	0.771
rs7554059	G	0.812	-0.075	0.051	0.135
rs11625568	A	0.951	-0.050	0.058	0.384
rs1277751	C	0.214	-0.002	0.032	0.957
rs1277752	A	0.777	0.007	0.031	0.827
rs1966778	C	0.940	-0.053	0.066	0.424
rs12322558	A	0.871	-0.040	0.037	0.278
rs11879136	T	0.244	-0.008	0.036	0.815
rs2913316	G	0.733	0.027	0.029	0.337

Table 6.5: The most significant variants identified in the RNFL thickness meta-analysis in the IOP meta-analysis: Rs: marker name, A1: effect allele, SE: standard error, P: p-value

To further elucidate potential mechanisms that cause RNFL thickness variation in the general population, I carried out a gene-set enrichment analysis of the results obtained after the GWAS meta-analysis. A total of 2,424 unique genes or transcripts were identified. With the exception of the BioCarta pathways, in which no significant enrichment was observed,

high level of entry enrichment were found in all other databases.

Amongst the most associated GO entries were cell and biological adhesion (P=9.99E-11 and P=1.09E-10), plasma membrane (P=3.23E-06) and voltage ion channel activity (p=2.87E-06) (Tables 6.6, 6.7, 6.8). In addition, several KEGG pathways were identified, for instance ECM-receptor interaction (P=1.50E04) and focal adhesion (P=5.65E-04) (Table 6.9). These findings suggest that adhesion in the plasma membrane and ion channels (i.e. calcium channels) may play a role in RNFL thickness variation.

Category	Term	PValue	Fold Enrichment	Bonferroni
GOTERM_BP_FAT	GO:0007155~cell adhesion	9.99E-11	1.85	3.52E-07
GOTERM_BP_FAT	GO:0022610~biological adhesion	1.09E-10	1.84	3.85E-07
GOTERM_BP_FAT	GO:0048568~embryonic organ development	8.37E-08	2.57	2.95E-04
GOTERM_BP_FAT	GO:0048562~embryonic organ morphogenesis	1.19E-06	2.64	4.19E-03
GOTERM_BP_FAT	GO:0030030~cell projection organization	2.05E-06	1.88	7.20E-03
GOTERM_BP_FAT	GO:0048598~embryonic morphogenesis	6.86E-06	1.92	2.39E-02
GOTERM_BP_FAT	GO:0000904~cell morphogenesis involved in differentiation	8.03E-06	2.04	2.79E-02

Table 6.6: GO biological processes identified with gene-set enrichment analysis: the table shows the GO biological processes significantly enriched for RNFL thickness meta-analysis.

Category	Term	PValue	Fold Enrichment	Bonferroni
GOTERM_CC_FAT	GO:0044459~plasma membrane part	3.23E-06	1.30	1.74-03
GOTERM_CC_FAT	GO:0005886~plasma membrane	6.31E-06	1.20	3.39E-03
GOTERM_CC_FAT	GO:0031224~intrinsic to membrane	8.10E-06	1.14	4.35E-03
GOTERM_CC_FAT	GO:0031012~extracellular matrix	4.88E-05	1.77	2.59E-02
GOTERM_CC_FAT	GO:0042995~cell projection	5.43E-05	1.51	2.88E-02
GOTERM_CC_FAT	GO:0016021~integral to membrane	7.32E-05	1.13	3.86E-02

Table 6.7: GO cellular components identified with gene-set enrichment analysis: the table shows the GO cellular component significantly enriched for RNFL thickness meta-analysis.

Category	Term	PValue	Fold Enrichment	Bonferroni
GOTERM_MF_FAT	GO:0005244~voltage-gated ion channel activity	2.87E-06	2.25	3.34E-03
GOTERM_MF_FAT	GO:0022832~voltage-gated channel activity	2.87E-06	2.25	3.34E-03
GOTERM_MF_FAT	GO:0005509~calcium ion binding	2.97E-06	1.50	3.46E-03
GOTERM_MF_FAT	GO:0019198~transmembrane receptor protein phosphatase activity	7.79E-06	6.24	9.05E-03
GOTERM_MF_FAT	GO:0022843~voltage-gated cation channel activity	3.30E-05	2.29	3.77E-02

Table 6.8: GO molecular functions identified with gene-set enrichment analysis:the table shows the GO molecular functions significantly enriched for RNFL thickness meta-analysis.

Category	Term	PValue	Fold Enrichment	Bonferroni
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	1.50E-04	2.60	2.65E-02
KEGG_PATHWAY	hsa04510:Focal adhesion	5.65E-04	1.85	9.62E-02
KEGG_PATHWAY	hsa05412:Arrhythmogenic right ventricular cardiomyopathy (ARVC)	1.14E-03	2.44	1.84E-01
KEGG_PATHWAY	hsa05414:Dilated cardiomyopathy	1.41E-03	2.25	2.23E-01
KEGG_PATHWAY	hsa05410:Hypertrophic cardiomyopathy (HCM)	1.47E-03	2.31	2.32E-01
KEGG_PATHWAY	hsa04010:MAPK signaling pathway	1.04E-02	1.51	8.46E-01
KEGG_PATHWAY	hsa00534:Heparan sulfate biosynthesis	2.68E-02	2.94	9.92E-01
KEGG_PATHWAY	hsa04340:Hedgehog signaling pathway	2.85E-02	2.14	9.94E-01
KEGG_PATHWAY	hsa03410:Base excision repair	3.61E-02	2.49	9.99E-01

Table 6.9: KEGG pathways identified with gene-set enrichment analysis: the table shows the KEGG pathways significantly enriched for RNFL thickness meta-analysis.

6.3.3 Investigating the genetic link between RNFL thickness and Alzheimer's disease

Given the epidemiological relationship between RNFL thickness and AD, my interest was to investigate a possible genetic overlap between the RNFL and the disease. A list of 435 genetic variants associated with AD was compiled. Of these SNPs, 50 reached significance in the RNFL meta-analysis (Table 6.10). Interestingly, some of the most important genes associated with AD such as *APOE, PVRL2, PICALM* and *CLU* were also associated with RNFL thickness.

SNP	Gene	chr	A1_rnfl	eaf	beta	se	p_rnfl	A1_ad	OR	p_ad
rs2814022	C10orf107	10	A	0.371362	3.83E-03	1.01E-03	1.55E-04	A	0.96	3.11E-02
rs2787723	C10orf107	10	T	0.372088	3.83E-03	1.01E-03	1.56E-04	T	0.96	3.55E-02
rs2814020	C10orf107	10	T	0.371629	3.79E-03	1.01E-03	1.78E-04	T	0.96	3.55E-02
rs12357683	C10orf107	10	A	0.371204	3.71E-03	1.01E-03	2.38E-04	A	0.96	4.57E-02
rs12356284	C10orf107	10	T	0.371328	3.69E-03	1.01E-03	2.58E-04	T	0.96	4.34E-02
rs7250924	GEMIN7	19	C	0.594076	-3.59E-03	1.12E-03	1.31E-03	T	0.9	2.08E-06
rs7093099	LOC254312	10	C	0.807219	-4.17E-03	1.32E-03	1.60E-03	T	0.91	5.96E-04
rs17444979	LOC254312	10	C	0.81772	-4.11E-03	1.31E-03	1.71E-03	T	0.91	4.17E-04
rs11767557	EPHA1	7	T	0.796575	-3.62E-03	1.23E-03	3.23E-03	C	0.87	2.43E-07
rs3813498	FOXO3	6	T	0.817217	3.56E-03	1.27E-03	5.01E-03	C	0.92	8.87E-04
rs1469704	GEMIN7	19	T	0.534567	-2.67E-03	1.00E-03	7.73E-03	C	0.91	3.72E-06
rs2060250	GEMIN7	19	A	0.53726	-2.68E-03	1.01E-03	7.76E-03	G	0.91	2.11E-06
rs1871047	PVRL2	19	A	0.611251	-2.65E-03	1.02E-03	9.61E-03	G	0.85	1.08E-14
rs12416248	PRKCQ	10	C	0.347019	-2.69E-03	1.05E-03	1.01E-02	C	0.95	2.16E-02
rs519825	PVRL2	19	C	0.38486	-2.60E-03	1.01E-03	1.02E-02	C	1.17	6.49E-14
rs1898895	PICALM	11	C	0.67258	2.66E-03	1.06E-03	1.18E-02	T	0.87	5.68E-11
rs567075	PICALM	11	C	0.674696	2.67E-03	1.06E-03	1.19E-02	T	0.87	3.14E-11
rs377702	PVRL2	19	A	0.384104	-2.53E-03	1.01E-03	1.20E-02	A	1.17	3.36E-14
rs416041	PVRL2	19	G	0.384699	-2.54E-03	1.01E-03	1.20E-02	G	1.17	5.52E-14
rs659023	PICALM	11	G	0.62312	2.50E-03	1.02E-03	1.46E-02	A	0.88	2.78E-10
rs4803763	PVRL2	19	G	0.701436	2.89E-03	1.19E-03	1.48E-02	C	1.51	2.90E-67
rs1871046	PVRL2	19	T	0.615781	-2.50E-03	1.03E-03	1.52E-02	C	0.85	8.45E-15
rs542126	PICALM	11	G	0.610389	2.47E-03	1.02E-03	1.57E-02	A	0.89	5.02E-09
rs561655	PICALM	11	A	0.645149	2.47E-03	1.04E-03	1.75E-02	G	0.87	7.00E-11
rs10402271	BCAM	19	G	0.338378	-2.49E-03	1.05E-03	1.80E-02	G	1.33	8.72E-44
rs12703526	EPHA1	7	T	0.51799	-2.44E-03	1.04E-03	1.86E-02	G	0.91	1.02E-05
rs2927480	PVRL2	19	C	0.336457	-2.55E-03	1.09E-03	1.88E-02	C	1.36	2.69E-47
rs1237999	PICALM	11	A	0.640026	2.40E-03	1.03E-03	1.98E-02	G	0.88	3.29E-10
rs526904	PICALM	11	C	0.6397	2.39E-03	1.03E-03	2.00E-02	T	0.88	2.31E-10
rs497816	PICALM	11	T	0.639748	2.39E-03	1.03E-03	2.01E-02	C	0.88	2.41E-10
rs543293	PICALM	11	G	0.661299	2.39E-03	1.04E-03	2.23E-02	A	0.87	7.91E-11
rs7359852	BCAM	19	C	0.337846	-2.45E-03	1.08E-03	2.34E-02	C	1.36	3.86E-47
rs586274	PICALM	11	G	0.657823	2.33E-03	1.04E-03	2.53E-02	T	0.87	6.14E-11
rs9480861	LACE1	6	T	0.394723	-2.22E-03	1.01E-03	2.83E-02	T	0.92	1.22E-04
rs725206	PRKCQ	10	T	0.347032	-2.21E-03	1.03E-03	3.21E-02	T	0.95	1.72E-02
rs472486	PICALM	11	G	0.637036	2.20E-03	1.03E-03	3.22E-02	A	0.88	6.34E-10
rs573167	PICALM	11	A	0.655714	2.21E-03	1.04E-03	3.32E-02	G	0.87	1.31E-10
rs17125924	FERMT2	14	A	0.900731	3.46E-03	1.63E-03	3.36E-02	G	1.12	8.22E-04
rs1532278	CLU	8	C	0.620158	2.22E-03	1.05E-03	3.46E-02	T	0.89	8.34E-08
rs439401	APOE	19	C	0.649678	-2.39E-03	1.13E-03	3.47E-02	T	0.65	1.65E-73
rs17125944	FERMT2	14	T	0.901269	3.44E-03	1.63E-03	3.49E-02	C	1.12	7.61E-04
rs519961	PICALM	11	C	0.656701	2.15E-03	1.04E-03	3.92E-02	T	0.87	1.20E-10
rs1332027	LACE1	6	A	0.931492	4.07E-03	1.99E-03	4.08E-02	G	0.89	1.33E-03
rs846848	CEACAM22P	19	T	0.857897	3.23E-03	1.59E-03	4.17E-02	G	0.88	1.13E-05
rs7110631	PICALM	11	G	0.68077	2.14E-03	1.06E-03	4.33E-02	C	0.87	3.53E-10
rs7941541	PICALM	11	A	0.677676	2.12E-03	1.06E-03	4.57E-02	G	0.87	5.57E-10
rs12152750	LOC285627	5	A	0.751868	-2.28E-03	1.14E-03	4.66E-02	G	1.08	5.41E-04
rs2984	RPS6KA2	6	T	0.635023	-2.13E-03	1.08E-03	4.80E-02	C	1.06	1.07E-02
rs541458	PICALM	11	T	0.678247	2.08E-03	1.06E-03	4.95E-02	C	0.88	1.43E-09
rs17056877	LOC285627	5	G	0.754413	-2.24E-03	1.14E-03	4.98E-02	A	1.09	3.28E-04

Table 6.10: List of AD-associated SNPs that are significant in RNFL meta-analysis: A1_rnfl: effect allele in the RNFL analysis, eaf: effect allele frequency, beta= effect in RNFL analysis, SE: standard error, p_rnfl: p-value of RNFL analysis, A1_ad: effect allele in AD analysis, OR: odds ratio, p_ad: p-value in AD analysis.

Amongst these 50 SNPs, 16 SNPs were not in linkage disequilibrium. The correlation between their effects size on AD and RNFL thickness was examined (Figure 6.12). The correlation between AD and RNFL thickness effect sizes was -0.25 (P=0.33). The negative sign suggests that variants that increase RNFL thickness also diminished the risk of AD. For example, the allele of the variant within *APOE* gene is protective for RNFL thickness (effect size calculated on T allele is 0.002) and reduce the risk of AD (the beta is -1.81, which is the -log10 of an odd ratio of 0.65).

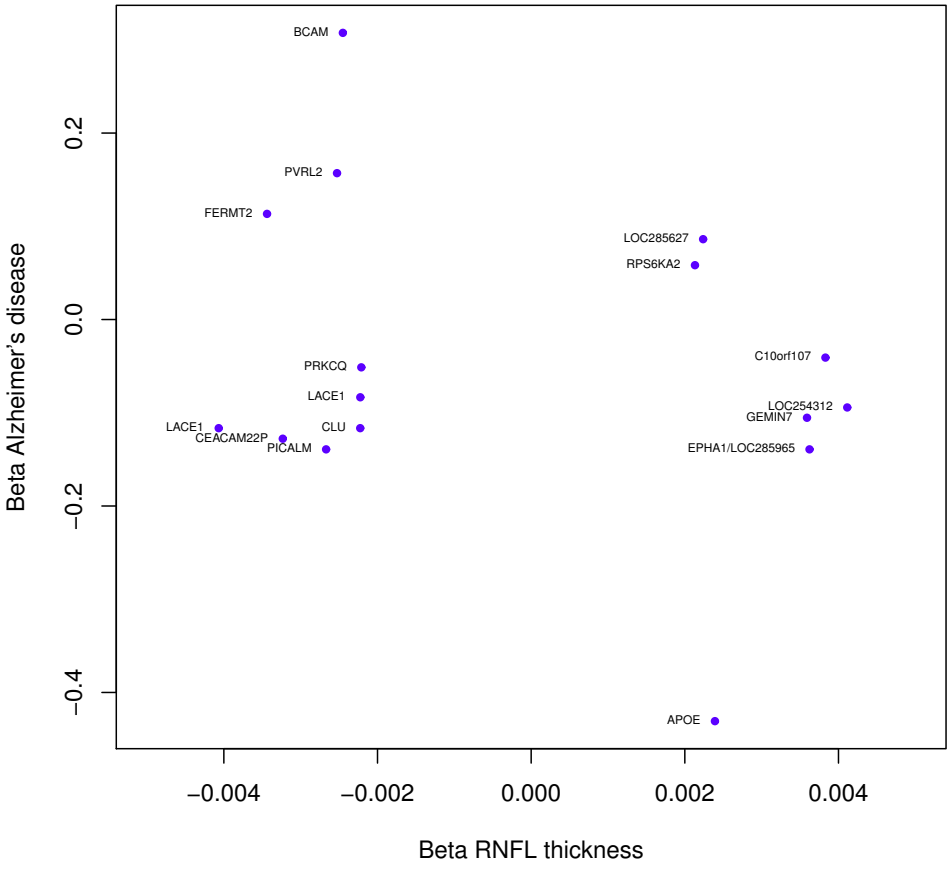


Figure 6.12: Plot of the β for RNFL thickness against AD: this figure shows the beta for RNFL thickness (x-axis) plotted against those for AD (y-axis). Beta for AD were calculated using the -log10 of the odds ratio.

Polygenic score analysis was performed in the BMES cohort. From 435 SNPs, 37 independent SNPs (not in LD) were included with association p<0.0001 in the AD joint meta-analysis and used to calculate the polygenic score. The weighted score was not significantly associated with RNFL thickness (β=-0.1955, SE=0.18, P=0.29) in the BMES cohort. The change of RNFL thickness for the quantiles of the polygenic score are shown in Figure 6.13. The overall r-square of the model was 0.014 (P= 0.003). The model was compared with the null model where a linear regression was fit only with age and sex (r-square =0.012) and the two models did not differ (ANOVA P=0.29).

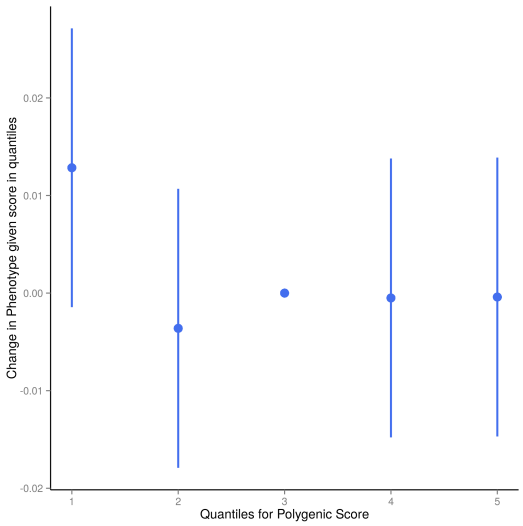


Figure 6.13: Quantile plot showing the effect of increasing score in RNFL thickness: the plot shows the quantiles of the polygenic score created with 37 SNPs (x-axis) plotted against the change in the phenotype

I also investigated whether the 50 AD SNPs identified as significant in RNFL GWAS meta-analysis had a cumulative effect on RNFL thickness variation. 16 of these 50 were independent and selected as shown in Table 6.11. The polygenic score created was still not significantly associated with RNFL thinning (β=12.08, SE=10.46, P=0.24). The overall R-squared was 0.014 and still, this model did not significantly differ from the null model

(ANOVA p=0.24).

SNP	CHR	BP	GENE
rs17056877	5	159083691	LOC285627
rs9480861	6	108858459	LACE1
rs1332027	6	108798497	LACE1
rs2984	6	166826020	RPS6KA2
rs11767557	7	143109138	EPHA1/LOC285965
rs1532278	8	27466314	CLU
rs2814022	10	63360810	C10orf107
rs17444979	10	10994158	LOC254312
rs725206	10	6394994	PRKCQ
rs567075	11	85830156	PICALM
rs17125944	14	53400628	FERMT2
rs7250924	19	45585705	GEMIN7
rs377702	19	45362666	PVRL2
rs7359852	19	45336034	BCAM
rs439401	19	45414450	APOE
rs846848	19	45074176	CEACAM22P

Table 6.11: List of 16 independent SNP associated with both RNFL thickness and AD: the table shows the list of 16 independent AD (non in linkage disequilibrium) SNPs that are also significant in RNFL thickness meta-analysis and were used to create the second polygenic score. CHR: chromosome; BP: base-position; GENE: closest gene/transcript

Given the small sample size in the BMES cohort (960 individuals), a multi-SNP score analysis, which allows summary results from both RNFL thickness meta-analysis and AD meta-analysis, was used. The weighted sum of the individual SNP coefficients leads also to an estimate of the

explained variance. The risk score was weighted using the AD effects for the 16 independent SNPs, previously selected to be associated with both AD and RNFL thickness. This multi-SNP risk score explained 7% of the variance of RNFL thickness. Then the association of this risk score with RNFL thickness variation was tested. The multi-SNP genotypic risk score was associated with a decrease of RNFL thickness (β=-0.0047, SE=0.001, P=0.005), where β is the average effect of AD risk alleles on the β of RNFL thickness (see Figure 6.14)

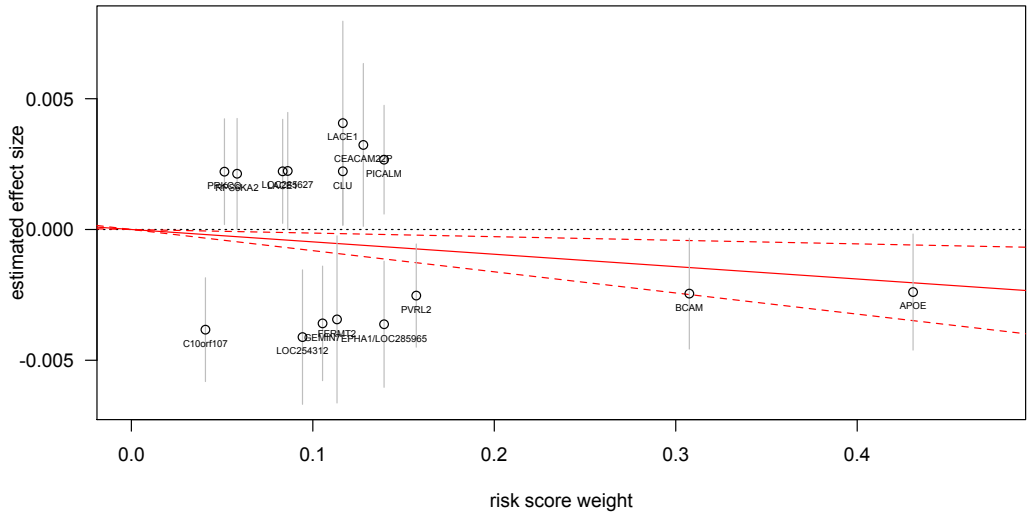


Figure 6.14: Coefficients in the AD risk score versus estimated effect size for RNFL thickness: this figure shows each SNP plotted by coefficient in the AD risk score (x-axis) versus estimated effect size for RNFL thickness in the meta-analysis (y-axis). Vertical grey lines show 95% confidence interval (CI) for each individual SNP. The solid red line shows the effect size estimate for the AD risk score on RNFL thickness, with red dashed lines representing the 95% interval.

6.4 Discussion

In this chapter the relationship between RNFL and POAG endophenotypes, the genetic architecture of RNFL thickness, the genetic factors underlying its variation and shared genetic effects between RNFL thickness and

Alzheimer's disease were investigated.

The relationship between RNFL thickness and POAG endophenotypes was investigated in 960 individuals in the BMES cohort. Findings showed that RNFL thickness was correlated with optic measurements. In addition, using regression models, RNFL thickness was also linked with IOP and refraction. RNFL has already been identified as a good indicator of POAG and it has been used in the diagnosis and for checking the progression (Jonas & Dichtl 1996). Also, increasing IOP has been shown to lead to disrupted axonal transport and a decrease of the mean RNFL thickness in rats (Abbott *et al.* 2014).

As other POAG quantitative traits, RNFL thickness is likely to be affected by environmental and genetic factors. The latter were investigated in the general population, first in the BMES and then in an extensive meta-analysis including more than 10,000 individual of European ancestry. Several common variants have been identified. The most significant signal lies in proximity of two genes: *C1orf100* and *ZBTB18/ZNF238*. *ZBTB18/ZNF238* is highly involved in neuronal activity, for example as a key regulator of neural migration (Ohtaka-Maruyama *et al.* 2013). It is also essential for the growth and organisation of the cerebellum and regulates the development of both GABAergic and glutamarergic neurons (Baubet *et al.* 2012). Finally, it is also been recognised as brain tumour suppressor (Tatard *et al.* 2010).

Another gene identified playing a role in neurogenesis was *BRINP3* which inhibits neuronal cell proliferation and it is involved since the early phases of neural development during the mouse embryogenesis.

BRINP3-mRNA is expressed in the adult brain in the olfactory bulb, cerebellum and neuronal layers in the hyppocampus (Kawano *et al.* 2004).

Another signal was identified close to *DLK1* transcript. The protein is involved in the differentiation of several cell types, including adipocytes; it is also thought to be a tumor suppressor. Muller *et al.* found that it is expressed by 30% of motor neurons. Dlk1 suppresses Notch signalling and activates expression of the K(+) channel subunit Kcng4 to modulate delayed-rectifier currents (Müller *et al.* 2014).

Another genetic signal was found within *CACNB2* gene. The protein encoded increases calcium current density by facilitating channel opening and increasing the number of channels in the plasma membrane (Stölting *et al.* 2015). In addition, *CACNB2* is a well-known gene for neuropsychiatric disorders, in particular ADHD (attention deficit hyperactive disorder), anxiety, bipolar disorder, major depressive disorder and schizophrenia. Voltage-gated calcium signalling (calcium channel activity) could be an important biological process in psychiatric disorders (of the Psychiatric Genomics Consortium *et al.* 2013) and seems to be implicated in postsynaptic density (Lotan *et al.* 2014). In addition, three rare missense mutations of *CACNB2* were found in autism spectrum disorder where the calcium channel dysfunction is likely to play a role in the disease (Breitenkamp *et al.* 2014). Xueying *et al.* identified a modest risk for AD in a gene-gene interaction study investigating a broad region on chromosome 10 in a Caucasian cohort of 506 cases and 558 controls (Liang *et al.* 2009). The authors concluded arguing that calcium level (controlled by calcium channels) may cause mitochondrial damage and induce apoptosis. Also, the protein may affect prefrontal-hippocampal

functional coupling in the human brain (Dixson *et al.* 2014). The gene has been also associated with hypertension in several studies (Hong *et al.* 2013; Lin *et al.* 2011; Takeuchi *et al.* 2010; Levy *et al.* 2009).

PITPNM3 is a well-known gene in ophthalmology as associated with retinal dystrophy. The gene encodes for calcium-binding protein and two Swedish families have a mutation in this gene which causes visual impairment from childhood, progressive decline in visual acuity, decreased macular thickness and enlarged fovea (Reinis *et al.* 2013).

BEST3 gene belongs to the bestrophin family of anion channels, which includes *BEST1*, the gene mutant in vitelliform macular dystrophy. It is expressed in a variety of tissue (cardiac, smooth muscle and renal) and it prevents apoptotic cell death induced by endoplasmatic reticulum stress (Jiang *et al.* 2013).

Interestingly these genes are not shared with other POAG endophenotypes. Only the variant in *BEST3* gene reached nominal significance for VCDR, suggesting that the genetic factors underlying RNFL thickness might be different to those associated with other POAG phenotypes thus underlying different processes of the disease.

To extrapolate biological meaning from the meta-analysis of GWASs and thus elucidate possible mechanisms underlying RNFL thinning, a gene-set pathway analysis was performed. Biological processes such as biological and cell adhesion in plasma membrane have been identified. In addition voltage-dependent ion channel, such as calcium, have been found to be enriched in the result. Using KEGG database, I identified

ECM-receptor interaction and focal adhesion to be overrepresented in RNFL thickness genes.

Alzheimer's disease has commonalities with POAG and RNFL thickness has been proposed as a bio marker of the disease. In this chapter, the shared genetics between RNFL thickness and AD was investigated. 50 variants out of 435 identified through AD meta-analysis were also significant for RNFL thickness in the GWAS meta-analysis. This cluster of genes includes *CLU, PICALM, APOE* and *PVRL2*, which are well-known genes for AD (Naj *et al.* 2011). These genes are likely to be involved in relevant mechanisms underlying RNFL thinning and POAG. *CLU* gene, for example, play many roles in a variety of extracellular processes, including apoptosis, lipid transport, cell-cell and cell-matrix interaction and prevents protein misfolding (Wilson & Easterbrook-Smith 2000; Poon *et al.* 2002). Interestingly, an increase of Clusterin protein has been found in the aqueous humour of individuals with pseudo-exfoliative glaucoma (PXG) (Padhy *et al.* 2014). It is also involved in the TGF-beta signalling (Lee *et al.* 2008). Epigenetic factors may regulate Clusterin expression of retinal pigment epithelial cells, affecting the pathogenesis of AMD via the inhibition of angiogenesis and inflammation (Suuronen *et al.* 2007). Finally, Clusterin produced and released by Muller cell may play a role in the pathogenesis of ischemic injury in the rat retina, following increase of IOP-muller cell die by apoptosis (Gwon *et al.* 2004). In addition, *PVRL2* gene is highly involved in cell adhesion, it encodes Nectin, which is a protein essential for adherens junction formation (Indra *et al.* 2013) and cell-cell adhesion (Samanta *et al.* 2012). *APOE* gene is surely known for AD and might be involved in RNFL thinning. Apolipoprotein E plays a role in the human body as a carrier of cholesterol (Mahley 1988). In the nervous

system, the main source of apoe synthesis are the astrocytes and microglia. Secreted apoe binds to cholesterol and this complex is internalised into the RGCs through the low density lipoprotein (LDL) receptors before being transferred to the optic nerve (Amaratunga *et al.* 1996). Apoe-E4 is higher in AD and also in glaucoma patients (Tamura *et al.* 2006). Several SNPs in *APOE* locus have been associated with both POAG and NTG (normal tension glaucoma) (Lam *et al.* 2006; Mabuchi *et al.* 2005; Ressiniotis T *et al.* 2004; Zetterberg *et al.* 2007; Guo *et al.* 2015). A study in mice showed that apoe has a critical role in axonal damage-induced retinal ganglion cell death and it is likely to involve cytotoxic KA pathway (Omodaka *et al.* 2014).

There was an inverse linear relationship between the effect sizes for RNFL thickness and AD. For example, the variant in *APOE* gene had a protective effect for AD and increase RNFL thickness. Vice versa the one within *PRVL2* gene increased the risk for AD and decrease RNFL thickness.

Two different polygenic risk scores were used to investigate the polygenic effect of AD loci. The first score included 37 independent SNPs associated with AD (Naj *et al.* 2011) and the second one included 16 independent SNPs associated with AD and also associated with RNFL thickness variation in the meta-analysis. Both scores were not significantly associated with RNFL thickness in the BMES.

However, the BMES cohort only has 960 observations for RNFL thickness leading to a lack of power. For this reason, summary statistics, such as effect sizes and SE, from RNFL thickness and AD meta-analyses were used to calculate a multi-SNP score including 16 variants associated with both

phenotypes. Multi-SNP risk score analysis identified a weak but statistical significant protective effect of thicker RNFL on AD risk.

There are a number of limitations to consider. Firstly, genetic associations results for RNFL thickness need to be replicated and validated or assessed for clinical importance in POAG case-controls cohort. Second, results are generated in HapMap2 which covers common genetic variants at the expense of rare variants. Third, results for polygenic risk analysis need replication and to be validated in future studies.

6.5 Conclusion

This is the first study investigating RNFL thickness genetics in an extensive European population and results suggest a role of several common variants in RNFL variation. The genetic factors identified are mainly involved in glutamate, calcium channels and cell adhesion, which were also identified mechanisms through the gene-set enrichment analysis. The study also showed a possible shared genetics between RNFL thickness and AD underlying the importance of using RNFL measurements as a bio marker for AD, but also the role of neuronal protection in RNFL thickness and thus POAG. These findings might help to shed a light on POAG mechanisms since a complete understanding of glaucoma pathogenesis has not been achieved. At the moment the only therapy aims to lower IOP, however, therapeutic approaches based on neuroprotection might be used. Neuroprotection aims to protect retinal ganglion cell and their surrounding microglia by targeting pathways involved in neurodegeneration. Emerging

lines of investigation are examining neuroprotection therapies aimed to halt retinal ganglion cell apoptosis. For example, glutamate excitotoxicity, which leads to elevated intracellular calcium levels and thus triggering RGC death, appear to be a promising target. Memantine, which is used for AD, has shown protective properties in a rat glaucoma model by acting on NMDA receptor (Danesh-Meyer 2011; Lagreze *et al.* 1998), blocking NMDA channels normally opened by glutamate in glaucoma models (Osborne 2009).

CHAPTER 7

General discussion and conclusions

In this chapter I will summarise and discuss the research of this thesis.

The main purpose of the PhD course is to gain insight in the genetic factors underlying POAG using GWAS techniques. My work shows that GWAS methodology may still have the potential to unravel the genetics of complex traits and disease.

My work includes several interests. In Chapter 3, I described GWAS on POAG quantitative traits in the BMES cohort, such as VCDR and IOP, including a candidate gene study on *ATOH7* gene in optic disc measurements, and consequent meta-analyses as part of the IGGC consortium. I also investigated common mechanisms underlying IOP variation using the GWAS results from the IGG meta-analysis through gene-set enrichment analysis. In Chapter 4 I focused on genetic analysis that can combine information from different POAG endophenotypes, such as principal factor analysis, genetic correlation and multivariate regression models. Then, in Chapter 5 I investigated epistatic interactions in association with VCDR and IOP. Finally, in Chapter 6 I described

association analyses on RNFL thickness, a POAG trait whose genetics has not been fully investigated, and its genetic relationship with AD.

Here I will start with a summary and discussion of the main findings for each chapter and then I will discuss what these results mean globally and why these are important for understanding the genetic architecture of POAG. As Griffith mentioned ”There is no hope of doing perfect research” (Levi 2013) and I will describe how my work is not perfect, the main flaws and limitations of my PhD project. Following this, I will consider a few points for future research and studies.

7.1 Discussion

7.1.1 Which are the genetic variants underlying POAG endophenotypes in the general population and the biological mechanisms underlying them?

In chapter 3 the genetic architecture of POAG endophenotypes (VCDR and IOP) was investigated using GWAS. Specifically, I performed GWAS in the BMES cohort on VCDR and IOP and then meta-analysed the results with other European and Asian populations from the IGGC. The findings suggest a role of cellular stress response in VCDR variation involving DUSP1 and HSF2 genes (Keyse & Emslie 1992; Kwong *et al.* 2006). In addition, cell growth and differentiation may have an effect on VCDR (*PLCE1* gene) (Wang *et al.* 2010; Wang *et al.* 2014; Abnet *et al.* 2010). Results also suggest that neuronal degeneration may be partially responsible of the loss

of cells in the optic disc nerve (EXOC2 gene) (Lipschutz & Mostov 2002; Coleman & Yao 2003). These results also confirm the association between VCDR and variants within *ATOH7* gene, which is a well-known gene for disc size (Ramdas *et al.* 2010; Macgregor *et al.* 2010; Khor *et al.* 2011). The findings corroborate the idea that *ATOH7* gene affects VCDR variation mainly through its effect on disc size. The role of disc size for POAG is controversial (Crowston *et al.* 2004) and therefore whether *ATOH7* gene affects POAG susceptibility varying VCDR or merely the disc size has clinical implication. In addition, findings in this thesis support the role of cellular adhesion in the plasma membrane and cell junctions in IOP variation, suggested also by loci near to *CAV1/2* genes. In addition, calcium ion binding may be partially responsible for IOP regulation.

Taken together, these results suggest that GWAS, especially supported by extensive meta-analysis, were still able to identify novel variants for IOP and VCDR and more important to investigate the biological and molecular mechanisms that may be responsible for these traits, thus affecting POAG susceptibility.

7.1.2 Are POAG endophenotypes correlated with each other and if so is it possible to combine POAG endophenotypes in order to unravel novel shared genetic loci?

Chapter 4 describes my work on pleiotropy within POAG endophenotypes and to identify cross-phenotypes loci within POAG endophenotypes. Firstly, epidemiological correlations between POAG quantitative traits was investigated. The link between IOP and CCT measurements (Doughty &

Zaman 2000) and between optic disc measurements, including disc size and VCDR (Quigley HA *et al.* 1990) was confirmed. More interesting is the link between IOP and VCDR, as the relationship between IOP and the glaucomatous damage is not clear, however it is critical to understand and treat the disease. To clarify whether phenotypic correlation was due to pleiotropy genetic correlation was investigated. Interestingly, a genetic correlation of 64% was identified for IOP and VCDR, showing a shared genetic determination between the two traits and thus suggesting the existence of pleitropic factors underlying both phenotypes.

To reduce the dimensions of POAG traits, principal factor analysis was carried out and two factors were used for further analysis. GWAS using the factors as outcomes were performed in the BMES and TwinsUK cohorts and results were then meta-analysed. The findings suggested several loci involved, for example nearby GLRX3 gene, which is involved in apoptosis and cellular growth (Cheng *et al.* 2011) and close to *RGS7* gene. The latter plays a role in intracellular signalling (Rose *et al.* 2000) and it is expressed in the eye and the retina (Cabrera *et al.* 1998).

Multivariate regression models that can accommodate multiple phenotypes in one model were also investigated. Suggestive signals, such as near *SEC16B* gene and and *SYT14*, were identified through MANOVA method. These genes are involved in organisation of transitional endoplasmatic reticulum sites and protein export (Watson *et al.* 2006) and synaptic transmission, however these data must be interpreted with caution because the correlation between IOP and VCDR is not taken into account. For this purpose, another model using GEE was carried out and a signal within *CACNAD3* gene was identified. This gene modulates the expression and

the function of the voltage-gated calcium channel complex and it may be involved in pain sensitivity (Neely *et al.* 2010). Also, a locus in *SNED1* gene was identified. GO annotations for this gene include calcium ion binding and its paralog is *DLK2* gene, which is an epidermal growth factor and involved as modulator of adipogenesis in vitro (Rivero *et al.* 2012).

The findings of applying multivariate models on POAG traits are far from exhaustive. This work may suggest that, despite a few suggestive genetic loci, mechanisms underlying IOP and optic nerve head are different. The existence of normal tension glaucoma, where optic nerve damage is present without an increase of IOP, provides evidence to this hypothesis. Mechanisms that not involve IOP may play a role. For example, vascular factors has been proven crucial where vascular dysregulation impacts the optic disc without an increase of IOP (Mi *et al.* 2014). Other mechanisms that contribute independently to death of retinal cells and optic disc fibres, without high IOP, may be excessive stimulation of the glutamatergic system (Dreyer *et al.* 1996), poorly functioning cellular pumps and glutamate transporters, oxidative stress and aberrant immunity (Schwartz 2003).

7.1.3 Are genetic variants for POAG endophenotypes interacting with other loci or between each other to determine IOP and VCDR?

In Chapter 5 I described the work on epistatic interactions between genetic variation associated with POAG endophenotypes in order to identify novel variants that interact with known ones to alter CDR and IOP. Several suggestive genetic variants were identified for VCDR. For example, on

chromosome 7 one genetic variation lying close to *PDE1C* gene interacted with a variant in *PLCE1*, identified in the VCDR IGGC meta-analysis. *PDE1C* gene plays a role in calcium-calmodulin complex (Repaske *et al.* 1992), it regulates soluble adenylyl cyclase/ cAMP signalling and lysosome-mediated collagen 1 protein degradation, in particular regulating collagen homeostasis during pathological vascular remodelling (Cai *et al.* 2011). I identified also an interaction affecting VCDR variation between variants located in DUSP1 transcript and *PITX2* gene, which is involved in early development, especially of the anterior segment of the eye (Gage *et al.* 2005). Another interesting interaction was found between variants located in *CDC7/TGFBR3* and *EFHB* genes. The latter plays a role in calcium-ion binding and it is associated with type 1 diabetes (Howson *et al.* 2009). Another gene for VCDR *SALL1* had a significant interaction with a variant within *PCDH7* gene, which encodes an integral membrane protein, involved in cell-cell recognition and adhesion, calcium-ion binding and axon guidance (Leung *et al.* 2013). Finally, I identified another novel variant lying within *C2CD5* gene, which had a significant interaction with *SIX1/6* gene. *C2CD5* gene is an insulin-stimulated glucose transporter. I also investigated the epistatic interactions between the 17 VCDR-associated SNPs to study whether these variants interacted together to affect VCDR variation. The locus within *DUSP1* gene appeared to interact with genes located in *PLCE1*, *SSSCA1*, *RPAP3* and *TMTC2* genes. *TMTC2* gene is also connected with *CHEK2* gene. These interactions underline the role of cell growth and cell differentiation in VCDR variation. Another group of genes included *RERE* gene, which had significant interactions with variants lying on *SALL1*, *CDKN2BAS* and *CDC7/TGFBR3* genes, suggesting an influence during embrionic development on VCDR variation. Indeed, *RERE* gene is involved in eye and cerebellum development (Kim *et al.* 2013; Kim & Scott

2014), *SALL1* gene in neural tube development and *CDC7/TGFBR3* gene in cell growth.

Similarly, suggestive interactions were identified for IOP. Variants located within *PTPRJ* gene had suggestive interactions with loci in *SCAF4/SOD1* and *SORBS2* genes. *SCAF4* gene has been associated with diabetic nephropathy (Alwohhaib *et al.* 2014). *SOD1* is a well known gene for amyotrophic lateral sclerosis (Rotunno *et al.* 2014) and the lack of the protein increases the vulnerability to stress of motor neurons leading to cell death and axon injury (Saccon *et al.* 2013). Interestingly, knock-out mice exhibit a contractile vascular phenotype with impairment of endothelial function during ageing (Didion *et al.* 2006). *SORBS2* gene encodes a adapter protein which is released from damaged cardiac tissue into the bloodstream upon lethal acute myocardial infarction (Kakimoto *et al.* 2013) and it has been also associated with diabetic nephropathy (Nakatani *et al.* 2011). I also identified an interaction between a locus within *ABO* transcript with SLC3A1 gene, which is renal amino acid transporter and involved in cystinuria (Miyamoto *et al.* 1996), and *DOCK7* gene, crucial for axon formation (Pinheiro & Gertler 2006), neural polarisation (Watabe-Uchida *et al.* 2006) and encepholopathy and cortical blindness (Perrault *et al.* 2014). Another suggestive interaction was between loci located in *TMCO1* and *GNG2* genes. *GNG2* gene may be important for signalling mechanisms across membranes and loss of function of this protein inhibits the ability of VEGF (vascular endothelium growth factor) to promote the angiogenic sprouting of blood vessels in zebra-fish models (Leung *et al.* 2006). Finally, I identified a suggestive interaction between variants in *CAV1* and *TOX2* and *TBC1D1* genes. *TOX2* gene regulates human natural killer cell development by controlling T-BET expression (Vong *et al.*

2014) and it has been associated with diastolic blood pressure(Basson *et al.* 2014). *TBC1D1* gene is crucial for cell growth regulation and differentiation. I also checked whether IOP associated variants interacted with each other. *PTPRJ*'s variants had significant interaction with loci within *GAS7* and *FNDC3B* genes, suggesting a role of cell growth and neuronal differentiation in IOP.

The evidence of these findings suggest that genetic factors interact between each other to affect POAG endophenotypes. In particular, mechanisms such as cell growth, neuronal differentiation, calcium channels activity and axon guidance may affect VCDR variation and be the product of interaction between genes. The role of neural development was confirmed also in IOP variation, providing evidence for a link between increased IOP and optic nerve defects in POAG. IOP is also affected by vascular and systemic processes and conditions, including vascular endothelium mechanism, blood pressure and diabetes.

7.1.4 Is RNFL thickness correlated with other POAG endophenotypes? Are there genetic variants affecting RNFL thickness? Do RNFL thickness and Alzheimer's disease have a shared genetics?

Finally, Chapter 6 describes my work on retinal nerve fibre layer thickness, its relationship with POAG endophentoypes, common genetic variants and pathways underlying its variation and its relationship with Alzheimer's disease. RNFL thickness is used in clinic for diagnosis and evaluation of progression of POAG (Jonas & Dichtl 1996) and epidemiological findings in

the BMES cohort confirmed its relationship with other POAG quantitative traits. Therefore RNFL thickness may be used as a quantitative traits for POAG in genetic association studies.

RNFL thickness, as POAG endophenotypes, is a complex trait and it is affected by genetic and environmental factors. Five novel genetic regions associated with RNFL thickness variation were identified in this thesis. The first signal lies on chromosome 1 nearby *C1orf100* and *ZBTB18/ZNF238* genes. The latter is highly involved in neuronal activity, such as neural migration (Ohtaka-Maruyama *et al.* 2013), development of GABAergic and glutamarergic neurons and cerebellum's growth and organisation (Baubet *et al.* 2012). Still on chromosome 1 another genetic variants was found located near to *BRINP3* transcript, which is involved in neuronal cell proliferation inhibition, neural development, and it is expressed in the adult brain (Kawano *et al.* 2004). Another signal was found on chromosome 14 nearby *DLK1* gene that plays a role in cellular differentiation and it is expressed in motor neurons (Müller *et al.* 2014). On chromosome 10 two signals were identified within *CACNAB2* gene whose protein increases calcium current density in the plasma membrane (Stölting *et al.* 2015). This gene is known to be associated with psychiatric disorders (ADHD, schizophrenia, depression, anxiety) ("Identification of risk loci with shared effects on five major psychiatric disorders" 2013) as well as autism (Breitenkamp *et al.* 2014) and may affect prefrontal-hippocampal functional coupling in the human brain and induce apoptosis. Also, the protein affects prefrontal-hippocampal functional coupling in the human brain (Dixson *et al.* 2014). It has also been associated with hypertension (Hong *et al.* 2013; Lin *et al.* 2011; Takeuchi *et al.* 2010; Levy *et al.* 2009). On chromosome 17, I identified a suggestive signal within *PITPNM3* gene,

which is well known for its association with retinal dystrophy (Reinis *et al.* 2013). Even though it did not cross genome-wide significance level, on chromosome 12, I pinpointed an interesting genetic signal lying on *BEST3* gene whose protein prevents apoptotic cell death induced by endoplasmatic reticulum stress (Jiang *et al.* 2013). This gene is also interesting because it is the only one, among the most significance hits for RNFL thickness, to reach significance also for another POAG trait, VCDR.

In Chapter 6, the shared genetics between RNFL thickness and AD was also investigated. Findings showed that 50 AD-associated variants (Naj *et al.* 2011) reached nominal significance in RNFL meta-analysis results. For instance, a locus within *CLU* gene, which is involved in apoptosis, cellular interaction (Wilson & Easterbrook-Smith 2000; Poon *et al.* 2002) and interestingly in pseudo-exfoliative glaucoma (Padhy *et al.* 2014). Another signal associated with both AD and RNFL thickness was identified within *PVRL2* gene that plays a role in cell adhesion (Indra *et al.* 2013; Samanta *et al.* 2012). Interestingly, also a locus located within *APOE* gene was associated with both phenotypes. *APOE* gene is a cholesterol carrier and might play a role in axonal damage. Apoe-E4 is higher in AD and interestingly in glaucoma patients (Tamura *et al.* 2006). There was an inverse relationship between RNFL thickness' effect sizes and AD ones, showing that, for example, the APOE allele investigated had a protective effect on AD and also increase the thickness of RNFL. Following this, different polygenic score were built with AD risk loci in the BMES cohort. I then investigated whether the scores were associated with RNFL thickness. No score was associated with RNFL thickness variation. However, given that these results may reflect the lack of power in the BMES cohort due to a small sample size, a multi-SNP score was created using effect estimates

of 16 SNPs from RNFL meta-analysis and AD study (Naj *et al.* 2011). Multi-SNP risk score analysis identified a weak but statistical significant protective effect of thicker RNFL on AD risk.

Taken together these findings provide evidence that the genetics underlying RNFL thickness may involve glutamate, calcium channels and cell adhesion, also confirmed by the gene-set enrichment analysis. The relationship with AD suggests a possible role of neural protection in RNFL thickness and thus POAG.

7.1.5 General discussion

A complete understanding on POAG pathogenesis has not been yet achieved, however several hypotheses have been proposed. Genetic studies such as GWAS can help to clarify POAG endophenotypes variation and may provide an insight of mechanisms underlying the disease.

The first mechanism my work suggests is cellular adhesion, in particular for IOP variation. Disruptions in cell adhesion mechanisms may lead to an accumulation of extracellular protein within the trabecular meshwork, restricting aqueous humour flow. The mechanism is in line with the mechanical stress model where an increase in IOP creates a compression of the optic disc causing death of the retinal ganglion cells. A pathway thought to be involved in mechanical stress mechanisms involved in POAG is TGB-β super-family of cell-signalling molecules. An increase of TGB-β can cause an increase of tissue stiffness through expression of of other extracellular matrix proteins such as fibronectin, collagen, and elastin (Wordinger *et al.*

2014). TGB-β may affect the capacity of cell to withstand mechanical pressure and/or restricting the flexibility of the trabecular meshwork cells. In combination with IOP fluctuation, cell stiffness may increase cell death over time, leading to the presentation of POAG (Wordinger *et al.* 2014).

Findings on VCDR have underlined a possible role of cellular stress in the pathogenesis of POAG. Cellular stress might involve several cell responses such as endoplasmatic reticulum stress response. Continuous exposure of the trabecular meshowork to oxidative stress seems to be a risk factor for POAG as the stress in the trabecular meshwork may induce death of cells and increase of IOP (Babizhayev & Yegorov 2011).

The findings on RNFL thickness, but also on VCDR and IOP, have potentially highlight the importance of calcium channels mechanisms for POAG in support to excitotoxicity theory. Excitotoxicity is caused by disruptions in glutamate transport that lead in abnormally high levels of glutamate in the synapse. This increased level of glutamate can lead to neurotoxicity through the N-methyl-D-aspartate (NMDA) receptor that causes influx of calcium into the cell and activates apoptotic pathways (Almasieh *et al.* 2012; Li *et al.* 1997; Masliah *et al.* 1996).

These findings have significant implications for the understanding and developing of POAG therapies. Current therapies focus on lowering IOP, however with limited effects as some patients have glaucoma despite normal IOP and others continue to suffer retinal ganglion cell death after medication (Group 1998; Shiose *et al.* 1990). It is also clear that there are additional pathogenic mechanisms that are not related to IOP variation. The current findings on cellular stress and calcium mechanisms

add to a growing body of literature on another therapeutic approach called neuroprotection (Tamm *et al.* 2013). Neuroprotection aims to protect retinal ganglion cells and the surrounding microglia and help their regeneration. Pathways targeted in this approach are oxidative stress and excitotoxicity. For example, memantine acts on excitotoxicity that leads to elevated levels of calcium as a non-competitive NMDA receptor antagonist. Interestingly, memantine is a drug currently used in Alzheimer's disease, supporting the idea of shared mechanisms between the two diseases as suggested by my work on RNFL and Alzheimer's disease (Danesh-Meyer 2011; Lagreze *et al.* 1998).

The generalisability of these results is subject to certain limitations. Firstly, POAG is a complex disease and it is a combination of genetic and environmental factors. Several environmental factors are generally considered to affect glaucomatous mechanisms and might interact with genetic factors. For instance, Kang *et al.* found an interaction between a a nitric oxide synthase gene variant and smoking (Kang *et al.* 2011). Secondly, the potential of false-positive results is an important limitation of GWAS. The balance between sensitivity and specificity, where false positive rates are minimised while maintaining the power to identify the true associations, is difficult to achieve. Due to the multiple tests performed, Bonferroni correction was generally used, however it can be very conservative if the independence assumption does not hold as for SNPs in GWAS. On the other hand, the same efforts to reduce false-positive results may results in overlooking a true association. Another limit of GWAS designs is that most the variants identified have very modest effects on the POAG endophenotypes, explaining only a small fraction of the total estimated heritability. SNPs used in GWAS would not probably be the causal

variants and might not be in LD with a causal variant rare allele. Selection will drive one susceptibility allele to low frequency, thus the increase of interest in the role of rare variants in complex diseases (Yang *et al.* 2010). However GWAS data in this work were imputed to HapMap 2 which does not guarantee an adequate covert of allele with a frequency less than 3-4%. Another potential drawback of studying quantitative traits of POAG may relate to their clinical relevance. In particular for RNFL thickness results and the epistatic study, investigation on the importance of the results in POAG onset and progression is needed. More precise phenotypes that can also give information about the stage of the disease would provide valuable insight in genetic studies. For example, longitudinal data could be implemented in GWAS and would inform about the progression of POAG. In addition, these studies are clearly far from actual clinical application. GWAS results need functional confirmation to be translated into new therapeutic targets. Gene-set enrichment analyses are surely useful in identifying biological pathways, however more research is needed to clarify the functional role of these variants and related genes.

7.2 Future studies and conclusion

The etiology of complex diseases, such as POAG, is still far from understood and much more has to be elucidated. Future directions following from the studies described in this thesis will need to clarify the functional role of the genes identified and validate their clinical relevance for POAG. Firstly, more research is needed to better understand whether combining multiple phenotypes might be a useful tool to improve our understanding on POAG

mechanisms. My first results are not encouraging, but larger sample size and replication are needed. This research has thrown up many questions on the link between POAG and Alzheimer's disease. My results on RNFL thickness and its implication on Alzheimer's disease need replication and validation. Further studies need to be carried out in order to validate epistatic interactions identified in this thesis.

More broadly, research is also needed to determine the functional aspects of the genes identified through my research and in general with association genetic studies. These investigations include fine mapping, including also rare variants investigation, genomic analysis of gene expression in human tissues, possibly including ocular tissue and epigenetic analysis of human tissue using the GWAS SNPs and their LD regions (Axton 2010). Once a gene or genes have been securely implicated, other models could be used. For instance, mouse knockouts, fish or cell knockdowns of a nearby gene and in vitro expression studies (Axton 2010).

This research extends our knowledge of genetic architecture of POAG endophenotypes and enhance our understanding of POAG onset and progression mechanisms. A key strength of the present studies was to show how GWAS can still help us to understand more on the genetics of complex disease. This thesis will serve as a base for future studies on mechanisms and future therapies on POAG.

APPENDIX A

Supporting publications

Venturini C, Nag A, Hysi PG, Wang JJ, Wong TY, Healey PR, Mitchell P, Hammond CJ, Viswanathan AC; Wellcome Trust Case Control Consortium 2, BMES GWAS Group. *Clarifying the role of ATOH7 in glaucoma endophenotypes*. Br J Ophthalmol. 2014 Apr;98(4):562-6. doi: 10.1136/bjophthalmol-2013-304080. Epub 2014 Jan 23. PubMed PMID: 24457358.

Springelkamp H, Hohn R, Mishra A, Hysi PG, Khor CC, Loomis SJ, Bailey JN, Gibson J, Thorleifsson G, Janssen SF, Luo X, Ramdas WD, Vithana E, Nongpiur ME, Montgomery GW, Xu L, Mountain JE, Gharahkhani P, Lu Y, Amin N, Karssen LC, Sim KS, van Leeuwen EM, Iglesias AI, Verhoeven VJ, Hauser MA, Loon SC, Despriet DD, Nag A, **Venturini C** et. al. *Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process*. Nat Commun. 2014 Sep 22;5:4883. doi: 10.1038/ncomms5883. PubMed PMID: 25241763.

Hysi PG, Cheng CY, Springelkamp H, Macgregor S, Bailey JN, Wojciechowski R, Vitart V, Nag A, Hewitt AW, Hohn R, **Venturini C**, et al. *Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma.* Nat Genet. 2014 Aug 31. doi: 10.1038/ng.3087. [Epub ahead of print] PubMed PMID: 25173106.

Book's chapters

Genetic Epidemiology. **C Venturini**, L. van Koolwijk, C. Bunce and A. C. Viswanathan. In: Glaucoma. Elsevier. Eds. J. Crowston, T. Sharaawy, R. Hitchings, 2013, Vol.1.

Oral presentations

Cristina Venturini *et al. Common mechanisms underlying intraocular pressure identified in functional analysis of gene lists from genome-wide association study results in IGGC cohorts*. The Association for Research in Vision and Opthalmology (ARVO) meeting, Orlando, 2014. Selected by the Annual Meeting Program Committee as a "hot topic".

Cristina Venturini,*et al. Genetic factors and pathways affecting retinal nerve fiber layer thickness*.The Association for Research in Vision and Opthalmology (ARVO) meeting, Seattle, 2013.

Poster presentations

Cristina Venturini *et al. Novel genetic loci and Alzheimer' s disease genes influence retinal nerve fibre layer thinning: an European ancestry meta-analysis of 10,502 individuals*. American Society Human Genetics (ASHG) meeting 2015, Baltimore. Abstract accepted as a poster.

Cristina Venturini *et al. Common mechanisms underlying intraocular pressure identified in functional analysis of gene lists from genome-wide association study results in IGGC cohorts*. Quantitative Genomics meeting 2014 (QG14), London.

References

1. Abbott, C. J. *et al.* “Comparison of retinal nerve fiber layer thickness in vivo and axonal transport after chronic intraocular pressure elevation in young versus older rats”. *PloS One* **9,** e114546 (2014).
2. Abelson, G. & Smith. “Sustained reduction of intraocular pressure in humans with the calcium channel blocker verapamil.” *American journal of ophthalmology* **105,** 155–159 (Feb. 1988).
3. Abnet, C. C. *et al.* “A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma.” [**inlang**eng]. *Nat Genet* **42,** 764–767 (Sept. 2010).
4. Abreu, M. M., Kim, Y. Y., Shin, D. H. & Netland, P. A. “Topical verapamil and episcleral venous pressure”. *Ophthalmology* **105,** 2251–2255 (Dec. 1, 1998).
5. Airaksinen, P. J., Drance, S. M., Douglas, G. R., Mawson, D. K. & Nieminen, H. “Diffuse and localized nerve fiber loss in glaucoma”. *American Journal of Ophthalmology* **98,** 566–571 (Nov. 1984).
6. Alexander, C. *et al.* “OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28”. *Nature Genetics* **26,** 211–215 (Oct. 2000).

7. Allingham, R. R., Liu, Y. & Rhee, D. J. "The genetics of primary open-angle glaucoma: A review". *Experimental Eye Research. Current aspects of aqueous humor dynamics and glaucoma A Tribute to Douglas H. Johnson* **88,** 837–844 (Apr. 30, 2009).
8. Almasieh, M., Wilson, A. M., Morquette, B., Vargas, J. L. C. & Di Polo, A. "The molecular basis of retinal ganglion cell death in glaucoma". *Progress in retinal and eye research* **31,** 152–181 (2012).
9. Alwohhaib, M. *et al.* "Single nucleotide polymorphisms at erythropoietin, superoxide dismutase 1, splicing factor, arginine/serin-rich 15 and plasmacytoma variant translocation genes association with diabetic nephropathy". *Saudi Journal of Kidney Diseases and Transplantation: An Official Publication of the Saudi Center for Organ Transplantation, Saudi Arabia* **25,** 577–581 (May 2014).
10. Amaratunga, A. *et al.* "Apolipoprotein E is synthesized in the retina by Müller glial cells, secreted into the vitreous, and rapidly transported into the optic nerve by retinal ganglion cells". *The Journal of Biological Chemistry* **271,** 5628–5632 (Mar. 8, 1996).
11. Armstrong, R. A. "Visual field defects in Alzheimer's disease patients may reflect differential pathology in the primary visual cortex". *Optometry and Vision Science: Official Publication of the American Academy of Optometry* **73,** 677–682 (Nov. 1996).
12. Aulchenko, Y. S. *et al.* "Linkage disequilibrium in young genetically isolated Dutch population". *European Journal of Human Genetics* **12,** 527–534 (Mar. 31, 2004).
13. Aung, T. *et al.* "Investigating the association between OPA1 polymorphisms and glaucoma: comparison between normal tension

and high tension primary open angle glaucoma". *Hum Genet* **110** (May 2002).

14. Axenovich, T. *et al.* "Linkage and association analyses of glaucoma related traits in a large pedigree from a Dutch genetically isolated population". *Journal of Medical Genetics* **48,** 802–809 (Dec. 2011).
15. Axton, M. "Editorial: On beyond GWAS". *Nat Genet* **42,** 551 (2010).
16. Babizhayev, M. & Yegorov, Y. "Senescent phenotype of trabecular meshwork cells displays biomarkers in primary open-angle glaucoma". *Current molecular medicine* **11,** 528–552 (2011).
17. Bamji, S. X. "Cadherins: Actin with the Cytoskeleton to Form Synapses". *Neuron* **47,** 175–178 (July 21, 2005).
18. Basson, J. *et al.* "Gene-education interactions identify novel blood pressure loci in the Framingham Heart Study". *American Journal of Hypertension* **27,** 431–444 (Mar. 2014).
19. Bateson, W. *Mendel's Principles of Heredity*. 462 pp. (Cosimo, Inc., Nov. 1, 2007).
20. Baubet, V. *et al.* "Rp58 is essential for the growth and patterning of the cerebellum and for glutamatergic and GABAergic neuron development". *Development (Cambridge, England)* **139,** 1903–1909 (June 2012).
21. Bayer, A. U. & Ferrari, F. "Severe progression of glaucomatous optic neuropathy in patients with Alzheimer's disease". *Eye (London, England)* **16,** 209–212 (Mar. 2002).
22. Bayer, A., Ferrari, F. & Erb, C. "High Occurrence Rate of Glaucoma among Patients with Alzheimer’s Disease". *European Neurology* **47,** 165–168 (2002).
23. Bellezza, A. J., Hart, R. T. & Burgoyne, C. F. "The Optic Nerve Head as a Biomechanical Structure: Initial Finite Element Modeling".

Investigative Ophthalmology & Visual Science **41,** 2991–3000 (Sept. 1, 2000).

24. Bellezza, A. J. *et al.* "Deformation of the lamina cribrosa and anterior scleral canal wall in early experimental glaucoma". *Investigative Ophthalmology & Visual Science* **44,** 623–637 (Feb. 2003).
25. Belyea, D. A. *et al.* "Utility of Heidelberg retinal tomography as a screening tool for analyzing retinal nerve fiber layer defects". *Clinical Ophthalmology (Auckland, N.Z.)* **8,** 2409–2414 (Nov. 28, 2014).
26. Bill, A. "Blood circulation and fluid dynamics in the eye". *Physiological Reviews* **55,** 383–417 (July 1, 1975).
27. Bill, A. "The role of ciliary blood flow and ultrafiltration in aqueous humor formation". *Experimental Eye Research. Symposium on Transport Processes in the Eye Part III* **16,** 287–298 (Aug. 10, 1973).
28. Blennow, K., de Leon, M. J. & Zetterberg, H. "Alzheimer's disease". *Lancet* **368,** 387–403 (July 29, 2006).
29. Böhm, J. *et al.* "Sall1, sall2, and sall4 are required for neural tube closure in mice". *The American Journal of Pathology* **173,** 1455–1463 (Nov. 2008).
30. Bohnsack, B. L., Kasprick, D. S., Kish, P. E., Goldman, D. & Kahana, A. "A zebrafish model of axenfeld-rieger syndrome reveals that pitx2 regulation by retinoic acid is essential for ocular and craniofacial development". *Investigative Ophthalmology & Visual Science* **53,** 7–22 (Jan. 2012).
31. Bonomi, L. *et al.* "Vascular risk factors for primary open angle glaucoma: the Egna-Neumarkt Study". *Ophthalmology* **107,** 1287–1293 (July 2000).

32. Borrás, T. “Gene expression in the trabecular meshwork and the influence of intraocular pressure”. *Progress in Retinal and Eye Research* **22,** 435–463 (July 2003).
33. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. A. “Epistasis as the primary factor in molecular evolution.” [**inlang**eng]. *Nature* **490,** 535–538 (Oct. 2012).
34. Breitenkamp, A. F. S. *et al.* “Rare mutations of CACNB2 found in autism spectrum disease-affected families alter calcium channel function”. *PloS One* **9,** e95579 (2014).
35. Brown, A. A. *et al.* “Genetic interactions affecting human gene expression identified by variance association mapping”. *eLife* **3,** e01381 (2014).
36. Bunce, C., Hitchings, R. A., Duijn, C. M. V., Jong, P. T. V. M. D. & Vingerling, J. R. “Associations between the deletion polymorphism of the angiotensin 1-converting enzyme gene and ocular signs of primary open-angle glaucoma”. *Graefe’s Archive for Clinical and Experimental Ophthalmology* **243,** 294–299 (Apr. 1, 2005).
37. Burdon, K. P. *et al.* “Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1”. *Nature Genetics* **43,** 574–578 (June 2011).
38. Cabrera, J. L., de Freitas, F., Satpaev, D. K. & Slepak, V. Z. “Identification of the G03b25-RGS7 Protein Complex in the Retina”. *Biochemical and Biophysical Research Communications* **249,** 898–902 (Aug. 28, 1998).
39. Cáceres, A. *et al.* “Involvement of the cellular phosphatase DUSP1 in vaccinia virus infection”. *PLoS pathogens* **9,** e1003719 (2013).
40. Cai, Y. *et al.* “Cyclic nucleotide phosphodiesterase 1 regulates lysosome-dependent type I collagen protein degradation in vascular

smooth muscle cells". *Arteriosclerosis, Thrombosis, and Vascular Biology* **31,** 616–623 (Mar. 2011).

41. Cannon, T. D. & Keller, M. C. "Endophenotypes in the genetic analyses of mental disorders". *Annu. Rev. Clin. Psychol.* **2,** 267–290 (2006).
42. Carlborg, Ö. & Haley, C. S. "Epistasis: too often neglected in complex trait studies?" *Nature Reviews Genetics* **5,** 618–625 (Aug. 2004).
43. Cartee, G. D. "Of mice and men: filling gaps in the TBC1D1 story". *The Journal of Physiology* **588,** 4331–4332 (Pt 22 Nov. 15, 2010).
44. Challa, P. *et al.* "Analysis of LOXL1 polymorphisms in a United States population with pseudoexfoliation glaucoma". *Mol Vis* **14,** 146–9 (2008).
45. Chang, C. C. *et al.* "Second-generation PLINK: rising to the challenge of larger and richer datasets." [**inlang**eng]. *Gigascience* **4,** 7 (2015).
46. Chang, T. C. *et al.* "Determinants and Heritability of Intraocular Pressure and Cup-to-Disc Ratio in a Defined Older Population". *Ophthalmology* **112,** 1186–1191 (July 2005).
47. Charlesworth, J. *et al.* "The Path to Open-Angle Glaucoma Gene Discovery: Endophenotypic Status of Intraocular Pressure, Cup-to-Disc Ratio, and Central Corneal Thickness". *Investigative Ophthalmology & Visual Science* **51,** 3509–3514 (July 1, 2010).
48. Chen, L. *et al.* "Evaluation of LOXL1 polymorphisms in exfoliation syndrome in a Chinese population". *Mol Vis* **15,** 2349–57 (2009).
49. Cheng, C.-Y. *et al.* "Association of common SIX6 polymorphisms with peripapillary retinal nerve fiber layer thickness: the Singapore Chinese Eye Study". *Investigative Ophthalmology & Visual Science* **56,** 478–483 (Jan. 2015).

50. Cheng, N.-H. *et al.* "A mammalian monothiol glutaredoxin, Grx3, is critical for cell cycle progression during embryogenesis". *FEBS Journal* **278,** 2525–2539 (July 1, 2011).
51. Chi, Z.-L. *et al.* "Mutant WDR36 directly affects axon growth of retinal ganglion cells leading to progressive retinal degeneration in mice". *Human molecular genetics,* ddq299 (2010).
52. Cirulli, E. T. & Goldstein, D. B. "Uncovering the roles of rare variants in common disease through whole-genome sequencing". *Nat Rev Genet* **11,** 415–25 (June 2010).
53. Clapham, D. E. "Calcium Signaling". *Cell* **131,** 1047–1058 (Dec. 14, 2007).
54. Coleman, P. D. & Yao, P. J. "Synaptic slaughter in Alzheimer′s disease". *Neurobiology of Aging. Molecular and Cellular Basis of Synaptic Loss and Dysfunction in Alzheimer's Disease* **24,** 1023–1027 (Dec. 2003).
55. Congdon, N. *et al.* "Causes and prevalence of visual impairment among adults in the United States". *Archives of Ophthalmology* **122,** 477–485 (Apr. 2004).
56. Conte, I. *et al.* "Proper differentiation of photoreceptors and amacrine cells depends on a regulatory loop between NeuroD and Six6". *Development (Cambridge, England)* **137,** 2307–2317 (July 2010).
57. Copin, B. *et al.* "Apolipoprotein E–Promoter Single-Nucleotide Polymorphisms Affect the Phenotype of Primary Open-Angle Glaucoma and Demonstrate Interaction with the Myocilin Gene". *The American Journal of Human Genetics* **70,** 1575–1581 (June 2002).
58. Cordell, H. J. "Detecting gene-gene interactions that underlie human diseases". *Nature Reviews. Genetics* **10,** 392–404 (June 2009).

59. Cordell, H. J. “Epistasis: what it means, what it doesn’t mean, and statistical methods to detect it in humans”. *Human Molecular Genetics* **11,** 2463–2468 (Oct. 1, 2002).

60. Costanzo, M. *et al.* “The Genetic Landscape of a Cell”. *Science* **327,** 425–431 (Jan. 22, 2010).

61. Crawford, M. L., Harwerth, R. S., Smith, E. L., Mills, S. & Ewing, B. “Experimental glaucoma in primates: changes in cytochrome oxidase blobs in V1 cortex”. *Investigative Ophthalmology & Visual Science* **42,** 358–364 (Feb. 2001).

62. Cronin-Golomb, A., Rizzo, J. F., Corkin, S. & Growdon, J. H. “Visual function in Alzheimer’s disease and normal aging”. *Annals of the New York Academy of Sciences* **640,** 28–35 (1991).

63. Cronin-Golomb, A., Sugiura, R., Corkin, S. & Growdon, J. H. “Incomplete achromatopsia in Alzheimer’s disease”. *Neurobiology of Aging* **14,** 471–477 (Oct. 1993).

64. Crowston, J. G., Hopley, C. R., Healey, P. R., Lee, A. & Mitchell, P. “The effect of optic disc diameter on vertical cup to disc ratio percentiles in a population based cohort: the Blue Mountains Eye Study”. *British Journal of Ophthalmology* **88,** 766–770 (June 1, 2004).

65. Czudowska, M. A. *et al.* “Incidence of Glaucomatous Visual Field Loss: A Ten-Year Follow-up from the Rotterdam Study”. *Ophthalmology* **117,** 1705–1712 (Sept. 2010).

66. Danesh-Meyer, H. V. “Neuroprotection in glaucoma: recent and future directions”. *Current Opinion in Ophthalmology* **22,** 78–86 (Mar. 2011).

67. Dastani, Z. *et al.* "Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals". *PLoS genetics* **8,** e1002607 (2012).

68. De Vries, L., Zheng, B., Fischer, T., Elenko, E. & Farquhar, M. G. "The regulator of G protein signaling family". *Annual Review of Pharmacology and Toxicology* **40,** 235–271 (2000).

69. Delettre, C. *et al.* "Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy". *Nature Genetics* **26,** 207–210 (Oct. 2000).

70. Denis, M. *et al.* "Expression, regulation, and activity of ABCA1 in human cell lines". *Molecular Genetics and Metabolism* **78,** 265–274 (Apr. 2003).

71. Didion, S. P., Kinzenbaw, D. A., Schrader, L. I. & Faraci, F. M. "Heterozygous CuZn superoxide dismutase deficiency produces a vascular phenotype with aging". *Hypertension* **48,** 1072–1079 (Dec. 2006).

72. Dimasi, D. P., Hewitt, A. W., Green, C. M., Mackey, D. A. & Craig, J. E. "Lack of association of p53 polymorphisms and haplotypes in high and normal tension open angle glaucoma". *Journal of Medical Genetics* **42,** e55–e55 (Sept. 1, 2005).

73. Dixon, A. L. *et al.* "A genome-wide association study of global gene expression". *Nature Genetics* **39,** 1202–1207 (Oct. 2007).

74. Dixson, L. *et al.* "Identification of gene ontologies linked to prefrontal-hippocampal functional coupling in the human brain". *Proceedings of the National Academy of Sciences of the United States of America* **111,** 9657–9662 (July 1, 2014).

75. Do, R. *et al.* “Common variants associated with plasma triglycerides and risk for coronary artery disease”. *Nature Genetics* **45,** 1345–1352 (Nov. 2013).

76. Donegan, R. K., Hill, S. E., Turnage, K. C., Orwig, S. D. & Lieberman, R. L. “The Glaucoma-associated Olfactomedin Domain of Myocilin Is a Novel Calcium Binding Protein”. *Journal of Biological Chemistry* **287,** 43370–43377 (Dec. 21, 2012).

77. Doughty, M. J. & Zaman, M. L. “Human Corneal Thickness and Its Impact on Intraocular Pressure Measures: A Review and Meta-analysis Approach”. *Survey of Ophthalmology* **44,** 367–408 (Mar. 2000).

78. Dreyer, E. B., Zurakowski, D., Schumer, R. A., Podos, S. M. & Lipton, S. A. “Elevated glutamate levels in the vitreous body of humans and monkeys with glaucoma”. *Archives of Ophthalmology* **114,** 299–305 (Mar. 1996).

79. Dudbridge, F. & Gusnanto, A. “Estimation of significance thresholds for genomewide association scans”. *Genetic Epidemiology* **32,** 227–234 (2008).

80. Eichler, E. E. *et al.* “Missing heritability and strategies for finding the underlying causes of complex disease”. *Nature Reviews Genetics* **11,** 446–450 (June 2010).

81. El Fatimy, R. *et al.* “Heat shock factor 2 is a stress-responsive mediator of neuronal migration defects in models of fetal alcohol syndrome”. *EMBO molecular medicine* **6,** 1043–1061 (2014).

82. Erickson, K. A., Schroeder, A. & Netland, P. A. “Verapamil increases outflow facility in the human eye”. *Experimental Eye Research* **61,** 565–567 (Nov. 1995).

83. Estermann, S. *et al.* “Effect of oral donepezil on intraocular pressure in normotensive Alzheimer patients”. *Journal of Ocular Pharmacology and Therapeutics: The Official Journal of the Association for Ocular Pharmacology and Therapeutics* **22,** 62–67 (Feb. 2006).

84. Euesden, J., Lewis, C. M. & O’Reilly, P. F. “PRSice: Polygenic Risk Score software.” [**inlang**eng]. *Bioinformatics* **31,** 1466–1468 (May 2015).

85. Evans, D. M. *et al.* “Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility”. *Nature Genetics* **43,** 761–767 (Aug. 2011).

86. Fan, B. J., Wang, D. Y., Pasquale, L. R., Haines, J. L. & Wiggs, J. L. “Genetic variants associated with optic nerve vertical cup-to-disc ratio are risk factors for primary open angle glaucoma in a US Caucasian population”. *Invest Ophthalmol Vis Sci* **52,** 1788–92 (Mar. 2011).

87. Fingert, J. H. *et al.* “Analysis of myocilin mutations in 1703 glaucoma patients from five different populations”. *Human molecular genetics* **8,** 899–905 (1999).

88. Fisher, R. A. in *Breakthroughs in Statistics* (eds Kotz, S. & Johnson, N. L.) 66–70 (Springer New York, Jan. 1, 1992).

89. Flammer, J. *et al.* “The impact of ocular blood flow in glaucoma”. *Progress in Retinal and Eye Research* **21,** 359–393 (July 2002).

90. Foong, A. W. P. *et al.* “Rationale and Methodology for a Population-Based Study of Eye Diseases in Malay People: The Singapore Malay Eye Study (SiMES)”. *Ophthalmic Epidemiology* **14,** 25–35 (Jan. 1, 2007).

91. Foster, P. J. *et al.* "The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district". *Arch Ophthalmol* **118** (Aug. 2000).

92. Foster, P. J., Buhrmann, R., Quigley, H. A. & Johnson, G. J. "The definition and classification of glaucoma in prevalence surveys". *British Journal of Ophthalmology* **86,** 238–242 (Feb. 1, 2002).

93. Frazer, K. A., Murray, S. S., Schork, N. J. & Topol, E. J. "Human genetic variation and its contribution to complex traits". *Nature Reviews Genetics* **10,** 241–251 (Apr. 2009).

94. Funayama, T. *et al.* "SNPs and Interaction Analyses of Noelin 2, Myocilin, and Optineurin Genes in Japanese Patients with Open-Angle Glaucoma". *Investigative Ophthalmology & Visual Science* **47,** 5368–5375 (Dec. 1, 2006).

95. Furney, S. *et al.* "Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease". *Molecular psychiatry* **16,** 1130–1138 (2011).

96. Gage, P. J., Rhoades, W., Prucka, S. K. & Hjalt, T. "Fate maps of neural crest and mesoderm in the mammalian eye". *Investigative Ophthalmology & Visual Science* **46,** 4200–4208 (Nov. 2005).

97. Gibson, G. "Epistasis and Pleiotropy as Natural Properties of Transcriptional Regulation". *Theoretical Population Biology* **49,** 58–89 (Feb. 1996).

98. Gilmore, G. C., Wenk, H. E., Naylor, L. A. & Koss, E. "Motion perception and Alzheimer's disease". *Journal of Gerontology* **49,** P52–57 (Mar. 1994).

99. Gordon, M. O. *et al.* "The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma". *Arch Ophthalmol* **120** (June 2002).

100. Gottesman, I. I. & Gould, T. D. “The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions”. *American Journal of Psychiatry* **160,** 636–645 (Apr. 1, 2003).

101. Gould, D. B. *et al.* “Mutant Myocilin Nonsecretion In Vivo Is Not Sufficient To Cause Glaucoma”. *Molecular and Cellular Biology* **26,** 8427–8436 (Nov. 15, 2006).

102. Gratton, J.-P., Bernatchez, P. & Sessa, W. C. “Caveolae and Caveolins in the Cardiovascular System”. *Circulation Research* **94,** 1408–1417 (June 11, 2004).

103. Greenland, S. “Interactions in Epidemiology: Relevance, Identification, and Estimation:” *Epidemiology* **20,** 14–17 (Jan. 2009).

104. Grieshaber, M. C. & Flammer, J. “Blood flow in glaucoma”. *Current Opinion in Ophthalmology* **16,** 79–83 (Apr. 2005).

105. Group, C. N.-T. G. S. “The effectiveness of intraocular pressure reduction in the treatment of normal-tension glaucoma”. *Am J Ophthalmol* **126,** 498–504 (1998).

106. Grundberg, E. *et al.* “Mapping cis- and trans-regulatory effects across multiple tissues in twins”. *Nature Genetics* **44,** 1084–1089 (Oct. 2012).

107. Guo, H., Li, M., Wang, Z., Liu, Q. & Wu, X. “Association of MYOC and APOE promoter polymorphisms and primary open-angle glaucoma: a meta-analysis”. *International Journal of Clinical and Experimental Medicine* **8,** 2052–2064 (2015).

108. Gwon, J.-S., Kim, I.-B., Lee, M.-Y., Oh, S.-J. & Chun, M.-H. “Expression of clusterin in Müller cells of the rat retina after pressure-induced ischemia”. *Glia* **47,** 35–45 (July 2004).

109. Haines, J. L. *et al.* “Complement factor H variant increases the risk of age-related macular degeneration”. *Science* **308** (Apr. 15, 2005).

110. Hashizume, K. *et al.* “Genetic Polymorphisms in the Angiotensin II Receptor Gene and Their Association with Open-Angle Glaucoma in a Japanese Population”. *Investigative Ophthalmology & Visual Science* **46,** 1993–2001 (June 1, 2005).

111. Hauser, M. A. *et al.* “Distribution of WDR36 DNA Sequence Variants in Patients with Primary Open-Angle Glaucoma”. *Investigative Ophthalmology & Visual Science* **47,** 2542–2546 (June 1, 2006).

112. Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A. & Evans, D. A. “Alzheimer disease in the US population: prevalence estimates using the 2000 census”. *Archives of Neurology* **60,** 1119–1122 (Aug. 2003).

113. Hemani, G., Knott, S. & Haley, C. “An Evolutionary Perspective on Epistasis and the Missing Heritability”. *PLoS Genet* **9,** e1003295 (Feb. 28, 2013).

114. Hemani, G. *et al.* “Detection and replication of epistasis influencing transcription in humans”. *Nature* **508,** 249–253 (Apr. 10, 2014).

115. Hertel, N., Redies, C. & Medina, L. “Cadherin expression delineates the divisions of the postnatal and adult mouse amygdala”. *The Journal of Comparative Neurology* **520,** 3982–4012 (Dec. 1, 2012).

116. Higgins, J. P. T. & Thompson, S. G. “Quantifying heterogeneity in a meta-analysis”. *Statistics in Medicine* **21,** 1539–1558 (2002).

117. Hildebrand, G. D. & Fielder, A. R. in *Pediatric Retina* (eds Reynolds, J. & Olitsky, S.) 39–65 (Springer Berlin Heidelberg, 2011).

118. Hindorff, L. A. *et al.* “Potential etiologic and functional implications of genome-wide association loci for human diseases and traits”. *Proceedings of the National Academy of Sciences* **106,** 9362–9367 (June 9, 2009).

119. Hm, L. *et al.* "The Framingham Eye Study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975." *Survey of ophthalmology* **24,** 335–610 (Suppl Dec. 1979).

120. Hoffmann, E. M., Zangwill, L. M., Crowston, J. G. & Weinreb, R. N. "Optic Disk Size and Glaucoma". *Survey of Ophthalmology* **52,** 32–49 (Jan. 2007).

121. Hofman, A. *et al.* "The Rotterdam Study: 2012 objectives and design update". *European Journal of Epidemiology* **26,** 657–686 (Aug. 1, 2011).

122. Hong, G. L. *et al.* "Genetic variations in MOV10 and CACNB2 are associated with hypertension in a Chinese Han population". *Genetics and molecular research: GMR* **12,** 6220–6227 (2013).

123. Hotta, K. *et al.* "Association between obesity and polymorphisms in SEC16B, TMEM18, GNPDA2, BDNF, FAIM2 and MC4R in a Japanese population". *Journal of Human Genetics* **54,** 727–731 (Dec. 2009).

124. Hougaard, J. L. *et al.* "Evaluation of heredity as a determinant of retinal nerve fiber layer thickness as measured by optical coherence tomography". *Invest Ophthalmol Vis Sci* **44** (July 2003).

125. Howie, B. N., Donnelly, P. & Marchini, J. "A flexible and accurate genotype imputation method for the next generation of genome-wide association studies". *PLoS genetics* **5,** e1000529 (June 2009).

126. Howson, J. M. M., Walker, N. M., Smyth, D. J., Todd, J. A. & Type I Diabetes Genetics Consortium. "Analysis of 19 genes for association with type I diabetes in the Type I Diabetes Genetics

Consortium families". *Genes and Immunity* **10 Suppl 1,** S74–84 (Dec. 2009).

127. Hsu, Y.-H. *et al.* "An Integration of Genome-Wide Association Study and Gene Expression Profiling to Prioritize the Discovery of Novel Susceptibility Loci for Osteoporosis-Related Traits". *PLoS Genet* **6,** e1000977 (June 10, 2010).

128. Hu, T., Darabos, C., Cricco, M. E., Kong, E. & Moore, J. H. "Genome-wide genetic interaction analysis of glaucoma using expert knowledge derived from human phenotype networks." [**inlang**eng]. *Pac Symp Biocomput,* 207–218 (2015).

129. Huang, D. W., Sherman, B. T. & Lempicki, R. A. "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources". *Nature Protocols* **4,** 44–57 (Dec. 2008).

130. Huang, W. *et al.* "Epistasis dominates the genetic architecture of Drosophila quantitative traits". *Proceedings of the National Academy of Sciences* **109,** 15553–15559 (Sept. 25, 2012).

131. Hunt, R. A., Edris, W., Chanda, P. K., Nieuwenhuijsen, B. & Young, K. H. "Snapin interacts with the N-terminus of regulator of G protein signaling 7". *Biochemical and Biophysical Research Communications* **303,** 594–599 (Apr. 4, 2003).

132. Hyman, L. *et al.* "Prevalence and causes of visual impairment in The Barbados Eye Study". *Ophthalmology* **108,** 1751–1756 (Oct. 2001).

133. Hysi, P. G. *et al.* "Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma". *Nature Genetics* (Aug. 31, 2014).

134. "Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis". *The Lancet* **381,** 1371–1379 (Apr. 26, 2013).

135. Iglesias, A. I. *et al.* “Exome sequencing and functional analyses suggest that SIX6 is a gene involved in an altered proliferation-differentiation balance early in life and optic nerve degeneration at old age”. *Human Molecular Genetics* **23,** 1320–1332 (Mar. 1, 2014).

136. Indra, I., Hong, S., Troyanovsky, R., Kormos, B. & Troyanovsky, S. “The adherens junction: a mosaic of cadherin and nectin clusters bundled by actin filaments”. *The Journal of Investigative Dermatology* **133,** 2546–2554 (Nov. 2013).

137. Iseri, P. K., Altinaş, O., Tokay, T. & Yüksel, N. “Relationship between cognitive impairment and retinal morphological and visual functional abnormalities in Alzheimer disease”. *Journal of Neuro-Ophthalmology: The Official Journal of the North American Neuro-Ophthalmology Society* **26,** 18–24 (Mar. 2006).

138. Ishikawa, K. *et al.* “Association between glaucoma and gene polymorphism of endothelin type A receptor”. *Molecular vision* **11,** 431–437 (2005).

139. Jacobson, N. *et al.* “Non-secretion of mutant proteins of the glaucoma gene myocilin in cultured trabecular meshwork cells and in aqueous humor”. *Human Molecular Genetics* **10,** 117–125 (Jan. 15, 2001).

140. Jang, Y., Lincoff, A., Plow, E. F. & Topol, E. J. “Cell adhesion molecules in coronary artery disease”. *Journal of the American College of Cardiology* **24,** 1591–1601 (Dec. 1, 1994).

141. Janssen, S. F. *et al.* “Gene Expression and Functional Annotation of the Human Ciliary Body Epithelia”. *PLoS ONE* **7** (ed Libby, R.) e44973 (Sept. 18, 2012).

142. Jiang, C. & Zeng, Z. B. “Multiple trait analysis of genetic mapping for quantitative trait loci”. *Genetics* **140,** 1111–1127 (July 1995).

143. Jiang, L. *et al.* “Mitochondria dependent pathway is involved in the protective effect of bestrophin-3 on hydrogen peroxide-induced apoptosis in basilar artery smooth muscle cells”. *Apoptosis: An International Journal on Programmed Cell Death* **18,** 556–565 (May 2013).

144. Jonas, J. B. & Dichtl, A. “Evaluation of the retinal nerve fiber layer”. *Survey of Ophthalmology* **40,** 369–378 (Apr. 1996).

145. Kakimoto, Y. *et al.* “Sorbin and SH3 domain-containing protein 2 is released from infarcted heart in the very early phase: proteomic analysis of cardiac tissues from patients”. *Journal of the American Heart Association* **2,** e000565 (2013).

146. Kang, J. H. *et al.* “Endothelial nitric oxide synthase gene variants and primary open-angle glaucoma: interactions with hypertension, alcohol intake, and cigarette smoking”. *Archives of ophthalmology* **129,** 773–780 (2011).

147. Kass, M. A. *et al.* “The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma”. *Archives of Ophthalmology (Chicago, Ill.: 1960)* **120,** 701–713, 701–713 (June 2002).

148. Kawano, H. *et al.* “Identification and characterization of novel developmentally regulated neural-specific proteins, BRINP family”. *Brain Research. Molecular Brain Research* **125,** 60–75 (June 18, 2004).

149. Keltner, J. L. *et al.* "The Association between Glaucomatous Visual Fields and Optic Nerve Head Features in the Ocular Hypertension Treatment Study". *Ophthalmology* **113,** 1603–1612 (Sept. 2006).

150. Keyse, S. M. & Emslie, E. A. "Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase". *Nature* **359,** 644–647 (Oct. 15, 1992).

151. Khor, C. C. *et al.* "Genome-wide association studies in Asians confirm the involvement of ATOH7 and TGFBR3, and further identify CARD10 as a novel locus influencing optic disc area". *Human Molecular Genetics* **20,** 1864–1872 (May 1, 2011).

152. Kim, B. J. & Scott, D. A. "Mouse model reveals the role of RERE in cerebellar foliation and the migration and maturation of Purkinje cells". *PloS One* **9,** e87518 (2014).

153. Kim, B. J. *et al.* "An allelic series of mice reveals a role for RERE in the development of multiple organs affected in chromosome 1p36 deletions". *PloS One* **8,** e57460 (2013).

154. Klein, B. E. K., Klein, R. & Knudtson, M. D. "Intraocular pressure and systemic blood pressure: longitudinal perspective: the Beaver Dam Eye Study". *British Journal of Ophthalmology* **89,** 284–287 (Mar. 1, 2005).

155. Klein, B. E. & Klein, R. "Intraocular pressure and cardiovascular risk variables". *Archives of ophthalmology* **99,** 837–839 (May 1981).

156. Koolwijk, L. M. E. v. *et al.* "Association of Cognitive Functioning with Retinal Nerve Fiber Layer Thickness". *Investigative Ophthalmology & Visual Science* **50,** 4576–4580 (Oct. 1, 2009).

157. Koolwijk, L. M. E. v. *et al.* "Genetic Contributions to Glaucoma: Heritability of Intraocular Pressure, Retinal Nerve Fiber Layer

Thickness, and Optic Disc Morphology". *Investigative Ophthalmology & Visual Science* **48,** 3669–3676 (Aug. 1, 2007).

158. Kwon, Y. H., Fingert, J. H., Kuehn, M. H. & Alward, W. L. "Primary Open-Angle Glaucoma". *New England Journal of Medicine* **360,** 1113–1124 (Mar. 12, 2009).

159. Kwong, J. M. *et al.* "Co-expression of heat shock transcription factors 1 and 2 in rat retinal ganglion cells". *Neuroscience letters* **405,** 191–195 (2006).

160. Lagreze, W. A., Knörle, R., Bach, M. & Feuerstein, T. J. "Memantine is neuroprotective in a rat model of pressure-induced retinal ischemia." *Investigative ophthalmology & visual science* **39,** 1063–1066 (1998).

161. Lam, C. Y. *et al.* "Association of Apolipoprotein E Polymorphisms With Normal Tension Glaucoma in a Chinese Population:" *Journal of Glaucoma* **15,** 218–222 (June 2006).

162. Lander, E. S. "The New Genomics: Global Views of Biology". *Science* **274,** 536–539 (Oct. 25, 1996).

163. Larrosa, J. M. *et al.* "Potential new diagnostic tool for Alzheimer's disease using a linear discriminant function for Fourier domain optical coherence tomography". *Investigative Ophthalmology & Visual Science* **55,** 3043–3051 (May 2014).

164. Lee, K.-B. *et al.* "Clusterin, a novel modulator of TGF-beta signaling, is involved in Smad2/3 stability". *Biochemical and Biophysical Research Communications* **366,** 905–909 (Feb. 22, 2008).

165. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. "Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood". *Bioinformatics* **28,** 2540–2542 (Oct. 1, 2012).

166. Leske, M. C., Connell, A. M., Schachat, A. P. & Hyman, L. "The Barbados Eye Study. Prevalence of open angle glaucoma". *Arch Ophthalmol* **112** (June 1994).

167. Leske, M. C. "Open-Angle Glaucoma—An Epidemiologic Overview". *Ophthalmic Epidemiology* **14,** 166–172 (Jan. 2007).

168. Leske, M. C. "The Epidemiology of Open-Angle Glaucoma: A Review". *American Journal of Epidemiology* **118,** 166–191 (Aug. 1, 1983).

169. Leske, M. C., Wu, S.-Y., Hennis, A., Honkanen, R. & Nemesure, B. "Risk Factors for Incident Open-angle Glaucoma: The Barbados Eye Studies". *Ophthalmology* **115,** 85–93 (Jan. 2008).

170. Leske, M. C. *et al.* "Nine-year incidence of open-angle glaucoma in the Barbados Eye Studies". *Ophthalmology* **114,** 1058–1064 (June 2007).

171. Leung, L. C. *et al.* "Coupling of NF-protocadherin signaling to axon guidance by cue-induced translation". *Nature Neuroscience* **16,** 166–173 (Feb. 2013).

172. Leung, T. *et al.* "Zebrafish G protein gamma2 is required for VEGF signaling during angiogenesis". *Blood* **108,** 160–166 (July 1, 2006).

173. Levene, R. Z., Workman, P. L., Broder, S. W. & Hirschhorn, K. "Heritability of ocular pressure in normal and suspect ranges". *Arch Ophthalmol* **84** (Dec. 1970).

174. Levi, C. "There is no hope of doing perfect research". *Publications Oboulo. com* (2013).

175. Levy, D. *et al.* "Genome-wide association study of blood pressure and hypertension". *Nature Genetics* **41,** 677–687 (June 2009).

176. Li, S., Mallory, M., Alford, M., Tanaka, S. & Masliah, E. "Glutamate transporter alterations in Alzheimer disease are possibly associated

with abnormal APP expression." *Journal of Neuropathology & Experimental Neurology* **56,** 901–911 (1997).

177. Liang, X. *et al.* "Genomic convergence to identify candidate genes for Alzheimer disease on chromosome 10". *Human Mutation* **30,** 463–471 (Mar. 2009).

178. Lin, H.-J., Chen, W.-C., Tsai, F.-J. & Tsai, S.-W. "Distributions of p53 codon 72 polymorphism in primary open angle glaucoma". *British Journal of Ophthalmology* **86,** 767–770 (July 1, 2002).

179. Lin, H.-J. *et al.* "Association of interleukin 1beta and receptor antagonist gene polymorphisms with primary open-angle glaucoma". *Ophthalmologica* **217** (2003).

180. Lin, Y. *et al.* "Genetic variations in CYP17A1, CACNB2 and PLEKHA7 are associated with blood pressure and/or hypertension in She ethnic minority of China". *Atherosclerosis* **219,** 709–714 (Dec. 2011).

181. Lipschutz, J. H. & Mostov, K. E. "Exocytosis: The Many Masters of the Exocyst". *Current Biology* **12,** R212–R214 (Mar. 19, 2002).

182. Liton, P. B., Luna, C., Challa, P., Epstein, D. L. & Gonzalez, P. "Genome-wide expression profile of human trabecular meshwork cultured cells, nonglaucomatous and primary open angle glaucoma tissue". *Molecular vision* **12,** 774–790 (July 12, 2006).

183. Liu, Y. & Allingham, R. R. "Molecular genetics in glaucoma". *Experimental Eye Research* **93,** 331–339 (Oct. 2011).

184. Logan, J. F. J. *et al.* "Evidence for Association of Endothelial Nitric Oxide Synthase Gene in Subjects with Glaucoma and a History of Migraine". *Investigative Ophthalmology & Visual Science* **46,** 3221–3226 (Sept. 1, 2005).

185. Lotan, A. *et al.* "Neuroinformatic analyses of common and distinct genetic components associated with major neuropsychiatric disorders". *Neurogenomics* **8,** 331 (2014).

186. Lu, Y. *et al.* "Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus". *Nature Genetics* **45,** 155–163 (Feb. 2013).

187. Mabuchi, F. *et al.* "The apolipoprotein E gene polymorphism is associated with open angle glaucoma in the Japanese population". *Molecular vision* **11,** 609–612 (2005).

188. Macgregor, S. *et al.* "Genome-wide association identifies ATOH7 as a major gene determining human optic disc size". *Human Molecular Genetics* **19,** 2716–2724 (July 1, 2010).

189. Mackey, D. A. *et al.* "Twins Eye Study in Tasmania (TEST): Rationale and Methodology to Recruit and Examine Twins". *Twin research and human genetics : the official journal of the International Society for Twin Studies* **12** (Oct. 2009).

190. Maher, B. "Personal genomes: The case of the missing heritability". *Nature News* **456,** 18–21 (Nov. 5, 2008).

191. Mahley, R. W. "Apolipoprotein E: cholesterol transport protein with expanding role in cell biology". *Science (New York, N.Y.)* **240,** 622–630 (Apr. 29, 1988).

192. Maier, P. C., Funk, J., Schwarzer, G., Antes, G. & Falck-Ytter, Y. T. "Treatment of ocular hypertension and open angle glaucoma: meta-analysis of randomised controlled trials". *BMJ : British Medical Journal* **331,** 134 (July 16, 2005).

193. Manolio, T. A. *et al.* "Finding the missing heritability of complex diseases". *Nature* **461,** 747–753 (Oct. 8, 2009).

194. Mao, C.-A. *et al.* "Neuronal transcriptional repressor REST suppresses an Atoh7-independent program for initiating retinal ganglion cell development". *Developmental biology* **349,** 90–99 (2011).

195. Mapstone, M., Dickerson, K. & Duffy, C. J. "Distinct mechanisms of impairment in cognitive ageing and Alzheimer's disease". *Brain: A Journal of Neurology* **131,** 1618–1629 (Pt 6 June 2008).

196. Martin, R. I. R. *et al.* "Genetic variants associated with risk of atrial fibrillation regulate expression of PITX2, CAV1, MYOZ1, C9orf3 and FANCC". *Journal of Molecular and Cellular Cardiology* **85,** 207–214 (June 11, 2015).

197. Masliah, E., Hansen, L., Alford, M., Deteresa, R. & Mallory, M. "Deficient glutamate tranport is associated with neurodegeneration in Alzheimer's disease". *Annals of neurology* **40,** 759–766 (1996).

198. Mason, R. P. *et al.* "National survey of the prevalence and risk factors of glaucoma in St. Lucia, West Indies. Part I. Prevalence findings". *Ophthalmology* **96,** 1363–1368 (Sept. 1989).

199. Mayama, C. "Calcium channels and their blockers in intraocular pressure and glaucoma". *European Journal of Pharmacology. Special Issue on Calcium Channels* **739,** 96–105 (Sept. 15, 2014).

200. McKee, A. C. *et al.* "Visual association pathology in preclinical Alzheimer disease". *Journal of Neuropathology and Experimental Neurology* **65,** 621–630 (June 2006).

201. Mi, X.-S., Yuan, T.-F. & So, K.-F. "The current research status of normal tension glaucoma". *Clinical Interventions in Aging* **9,** 1563–1571 (2014).

202. Miki, A., Otori, Y., Morimoto, T., Okada, M. & Tano, Y. "Protective effect of donepezil on retinal ganglion cells in vitro and in vivo". *Current Eye Research* **31,** 69–77 (Jan. 2006).

203. Mitchell, P., Hourihan, F., Sandbach, J. & Jin Wang, J. "The relationship between glaucoma and myopia: The blue mountains eye study". *Ophthalmology* **106,** 2010–2015 (Oct. 1, 1999).

204. Mitchell, P., Smith, W., Attebo, K. & Healey, P. R. "Prevalence of Open-angle Glaucoma in Australia: The Blue Mountains Eye Study". *Ophthalmology* **103,** 1661–1669 (Oct. 1996).

205. Mito, T., Delamere, N. A. & Coca-Prados, M. "Calcium-dependent regulation of cation transport in cultured human nonpigmented ciliary epithelial cells". *The American Journal of Physiology* **264,** C519–526 (Mar. 1993).

206. Miyamoto, K. *et al.* "Role of rBAT gene products in cystinuria". *International Journal of Urology: Official Journal of the Japanese Urological Association* **3,** S92–94 (Jan. 1996).

207. Moayyeri, A., Hammond, C. J., Hart, D. J. & Spector, T. D. "The UK Adult Twin Registry (TwinsUK Resource)". *Twin research and human genetics: the official journal of the International Society for Twin Studies* **16,** 144–149 (Feb. 2013).

208. Monemi, S. *et al.* "Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1". *Human Molecular Genetics* **14,** 725–733 (Mar. 15, 2005).

209. Moore, J. H. *et al.* "A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility". *Journal of Theoretical Biology* **241,** 252–261 (July 21, 2006).

210. Morales, J. *et al.* "Homozygous mutations in ADAMTS10 and ADAMTS17 cause lenticular myopia, ectopia lentis, glaucoma, spherophakia, and short stature". *American Journal of Human Genetics* **85,** 558–568 (Nov. 2009).

211. Morris, D. W. *et al.* "Confirming RGS4 as a susceptibility gene for schizophrenia". *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **125B,** 50–53 (Feb. 15, 2004).

212. Morton, D. B., Clemens-Grisham, R., Hazelett, D. J. & Vermehren-Schmaedick, A. "Infertility and male mating behavior deficits associated with Pde1c in Drosophila melanogaster". *Genetics* **186,** 159–165 (Sept. 2010).

213. Motallebipour, M., Rada-Iglesias, A., Jansson, M. & Wadelius, C. "The promoter of inducible nitric oxide synthase implicated in glaucoma based on genetic analysis and nuclear factor binding". *Molecular vision* **11,** 950–957 (2005).

214. Mukhopadhyay, A., Talukdar, S., Bhattacharjee, A. & Ray, K. "Bioinformatic approaches for identification and characterization of olfactomedin related genes with a potential role in pathogenesis of ocular disorders". *Molecular vision* **10,** 304–314 (Apr. 20, 2004).

215. Müller, D. *et al.* "Dlk1 promotes a fast motor neuron biophysical signature required for peak force execution". *Science (New York, N.Y.)* **343,** 1264–1266 (Mar. 14, 2014).

216. Murakami, T. *et al.* "Trophoblast glycoprotein: possible candidate mediating podocyte injuries in glomerulonephritis". *American journal of nephrology* **32,** 505–521 (2010).

217. Murdoch, I. E. *et al.* "Glaucoma prevalence may not be uniformly high in all 'black' populations". *African Journal of Medicine and Medical Sciences* **30,** 337–339 (Dec. 2001).

218. Naj, A. C. *et al.* "Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease". *Nature Genetics* **43,** 436–441 (May 2011).

219. Nakatani, S. *et al.* "Targeted proteomics of isolated glomeruli from the kidneys of diabetic rats: sorbin and SH3 domain containing 2 is a novel protein associated with diabetic nephropathy". *Experimental Diabetes Research* **2011,** 979354 (2011).

220. Neely, G. G. *et al.* "A genome-wide Drosophila screen for heat nociception identifies 03b1203b43 as an evolutionarily conserved pain gene". *Cell* **143,** 628–638 (Nov. 12, 2010).

221. Nemesure, B. *et al.* "Incident open-angle glaucoma and intraocular pressure". *Ophthalmology* **114,** 1810–1815 (2007).

222. Ntim-Amponsah, C. T. *et al.* "Prevalence of glaucoma in an African population". *Eye (London, England)* **18,** 491–497 (May 2004).

223. Of the Psychiatric Genomics Consortium, C.-D. G. *et al.* "Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis". *The Lancet* **381,** 1371–1379 (2013).

224. Ohtaka-Maruyama, C. *et al.* "RP58 regulates the multipolar-bipolar transition of newborn neurons in the developing cerebral cortex". *Cell Reports* **3,** 458–471 (Feb. 21, 2013).

225. Oikkonen, J. *et al.* "A genome-wide linkage and association study of musical aptitude identifies loci containing genes related to inner ear development and neurocognitive functions". *Molecular Psychiatry* **20,** 275–282 (Feb. 2015).

226. Ollila, H. M. *et al.* "Genome-wide association study of sleep duration in the Finnish population". *Journal of Sleep Research* **23,** 609–618 (Dec. 2014).

227. Omodaka, K. *et al.* “Neuroprotective effect against axonal damage-induced retinal ganglion cell death in apolipoprotein E-deficient mice through the suppression of kainate receptor signaling”. *Brain Research* **1586,** 203–212 (Oct. 24, 2014).

228. On Complex Epilepsies, I. L. A. E. C. “Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies.” [**inlang**eng]. *Lancet Neurol* **13,** 893–903 (Sept. 2014).

229. O’Neill, E. C. *et al.* “The optic nerve head in hereditary optic neuropathies”. *Nature Reviews Neurology* **5,** 277–287 (May 2009).

230. Ortuso, F. *et al.* “Discovery of PTPRJ agonist peptides that effectively inhibit in vitro cancer cell proliferation and tube formation”. *ACS chemical biology* **8,** 1497–1506 (July 19, 2013).

231. Osborne, N. N. “Recent clinical findings with memantine should not mean that the idea of neuroprotection in glaucoma is abandoned”. *Acta Ophthalmologica* **87,** 450–454 (June 1, 2009).

232. Padhy, B. *et al.* “Role of an extracellular chaperone, Clusterin in the pathogenesis of Pseudoexfoliation Syndrome and Pseudoexfoliation Glaucoma”. *Experimental Eye Research* **127,** 69–76 (Oct. 2014).

233. Pan, Y. *et al.* “Dok5 is involved in the signaling pathway of neurotrophin-3 against TrkC-induced apoptosis”. *Neuroscience letters* **553,** 46–51 (2013).

234. Paquet, C. *et al.* “Abnormal retinal thickness in patients with mild cognitive impairment and Alzheimer’s disease”. *Neuroscience Letters* **420,** 97–99 (June 13, 2007).

235. Pardo, L. M., MacKay, I., Oostra, B., van Duijn, C. M. & Aulchenko, Y. S. “The Effect of Genetic Drift in a Young Genetically Isolated Population”. *Annals of Human Genetics* **69,** 288–295 (2005).

236. Parisi, V. *et al.* “Morphological and functional retinal impairment in Alzheimer’s disease patients”. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* **112,** 1860–1867 (Oct. 2001).

237. Parton, R. G. & Simons, K. “The multiple faces of caveolae”. *Nature Reviews Molecular Cell Biology* **8,** 185–194 (Mar. 2007).

238. Pasutto, F. *et al.* “Heterozygous NTF4 Mutations Impairing Neurotrophin-4 Signaling in Patients with Primary Open-Angle Glaucoma”. *The American Journal of Human Genetics* **85,** 447–456 (Oct. 9, 2009).

239. Pena, J. D. *et al.* “Increased elastin expression in astrocytes of the lamina cribrosa in response to elevated intraocular pressure”. *Investigative Ophthalmology & Visual Science* **42,** 2303–2314 (Sept. 2001).

240. Pennisi, E. “Disease risk links to gene regulation”. *Science* **332,** 1031–1031 (2011).

241. Perrault, I. *et al.* “Mutations in DOCK7 in individuals with epileptic encephalopathy and cortical blindness”. *American Journal of Human Genetics* **94,** 891–897 (June 5, 2014).

242. Phillips, P. C. “Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems”. *Nature Reviews Genetics* **9,** 855–867 (Nov. 2008).

243. Pinheiro, E. M. & Gertler, F. B. “Nervous Rac: DOCK7 regulation of axon formation”. *Neuron* **51,** 674–676 (Sept. 21, 2006).

244. Pirone, A. *et al.* “03b1203b43 is essential for normal structure and function of auditory nerve synapses and is a novel candidate for auditory processing disorders”. *The Journal of Neuroscience: The*

Official Journal of the Society for Neuroscience **34,** 434–445 (Jan. 8, 2014).

245. Poon, S., Treweek, T. M., Wilson, M. R., Easterbrook-Smith, S. B. & Carver, J. A. "Clusterin is an extracellular chaperone that specifically interacts with slowly aggregating proteins on their off-folding pathway". *FEBS letters* **513,** 259–266 (Feb. 27, 2002).

246. Powell, B. L. *et al.* "Polymorphisms in OPA1 are associated with normal tension glaucoma". *Mol Vis* **9** (Sept. 22, 2003).

247. Prange, K. H., Singh, A. A. & Martens, J. H. "The genome-wide molecular signature of transcription factors in leukemia". *Experimental hematology* **42,** 637–650 (2014).

248. Quigley HA, Brown AE, Morrison JD & Drance SM. "THe size and shape of the optic disc in normal human eyes". *Archives of Ophthalmology* **108,** 51–57 (Jan. 1, 1990).

249. Quigley, H. A. "Number of people with glaucoma worldwide." *British Journal of Ophthalmology* **80,** 389–393 (May 1, 1996).

250. Quigley, H. A., Addicks, E. M. & Green, W. R. "Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy". *Archives of Ophthalmology (Chicago, Ill.: 1960)* **100,** 135–146 (Jan. 1982).

251. Quigley, H. A. & Broman, A. T. "The number of people with glaucoma worldwide in 2010 and 2020". *British Journal of Ophthalmology* **90,** 262–267 (Mar. 1, 2006).

252. Ramdas, W. D. *et al.* "A Genome-Wide Association Study of Optic Disc Parameters". *PLoS Genet* **6,** e1000978 (June 10, 2010).

253. Ramdas, W. D. *et al.* "Common genetic variants associated with open-angle glaucoma". *Human Molecular Genetics* **20,** 2464–2471 (June 15, 2011).

254. Rebolleda, G. *et al.* "OCT: New perspectives in neuro-ophthalmology". *Saudi Journal of Ophthalmology: Official Journal of the Saudi Ophthalmological Society* **29,** 9–25 (Mar. 2015).

255. Reinis, A., Golovleva, I., Köhn, L. & Sandgren, O. "Ocular phenotype of CORD5, an autosomal dominant retinal dystrophy associated with PITPNM3 p.Q626H mutation". *Acta Ophthalmologica* **91,** 259–266 (May 2013).

256. Repaske, D. R., Swinnen, J. V., Jin, S. L., Van Wyk, J. J. & Conti, M. "A polymerase chain reaction strategy to identify and clone cyclic nucleotide phosphodiesterase cDNAs. Molecular cloning of the cDNA encoding the 63-kDa calmodulin-dependent phosphodiesterase." [**inlang**eng]. *J Biol Chem* **267,** 18683–18688 (Sept. 1992).

257. Resnikoff, S. *et al.* "Global data on visual impairment in the year 2002." *Bulletin of the World Health Organization* **82,** 844–851 (Nov. 2004).

258. Ressiniotis T, Griffiths PG, Birch M, Keers S & Chinnery PF. "THe role of apolipoprotein e gene polymorphisms in primary open-angleglaucoma". *Archives of Ophthalmology* **122,** 258–261 (Feb. 1, 2004).

259. Rezaie, T. *et al.* "Adult-Onset Primary Open-Angle Glaucoma Caused by Mutations in Optineurin". *Science* **295,** 1077–1079 (Feb. 8, 2002).

260. Rivero, S., Díaz-Guerra, M. J. M., Monsalve, E. M., Laborda, J. & García-Ramírez, J. J. "DLK2 is a transcriptional target of KLF4 in the early stages of adipogenesis". *Journal of Molecular Biology* **417,** 36–50 (Mar. 16, 2012).

261. Rose, J. J. *et al.* "RGS7 Is Palmitoylated and Exists as Biochemically Distinct Forms". *Journal of Neurochemistry* **75,** 2103–2112 (Nov. 1, 2000).

262. Rotunno, M. S. *et al.* "Identification of a misfolded region in superoxide dismutase 1 that is exposed in amyotrophic lateral sclerosis". *The Journal of Biological Chemistry* **289,** 28527–28538 (Oct. 10, 2014).

263. Rougier, M.-B. *et al.* "Retinal nerve fibre layer thickness measured with SD-OCT in a population-based study of French elderly subjects: the Alienor study". *Acta Ophthalmologica* (Jan. 13, 2015).

264. Rudnicka, A. R., Mt-Isa, S., Owen, C. G., Cook, D. G. & Ashby, D. "Variations in Primary Open-Angle Glaucoma Prevalence by Age, Gender, and Race: A Bayesian Meta-Analysis". *Investigative Ophthalmology & Visual Science* **47,** 4254–4261 (Oct. 1, 2006).

265. Saccon, R. A., Bunton-Stasyshyn, R. K. A., Fisher, E. M. & Fratta, P. "Is SOD1 loss of function involved in amyotrophic lateral sclerosis?" *Brain* **136,** 2342–2358 (Aug. 2013).

266. Sailer, A. & Houlden, H. "Recent advances in the genetics of cerebellar ataxias". *Current Neurology and Neuroscience Reports* **12,** 227–236 (June 2012).

267. Saitoh, O. *et al.* "RGS7 and RGS8 Differentially Accelerate G Protein-mediated Modulation of K+ Currents". *Journal of Biological Chemistry* **274,** 9899–9904 (Apr. 2, 1999).

268. Samanta, D. *et al.* "Structure of Nectin-2 reveals determinants of homophilic and heterophilic interactions that control cell-cell adhesion". *Proceedings of the National Academy of Sciences of the United States of America* **109,** 14836–14840 (Sept. 11, 2012).

269. Sanfilippo, P. G., Hewitt, A. W., Hammond, C. J. & Mackey, D. A. "The Heritability of Ocular Traits". *Survey of Ophthalmology* **55,** 561–583 (Nov. 2010).

270. Schifano, E. D., Li, L., Christiani, D. C. & Lin, X. "Genome-wide Association Analysis for Multiple Continuous Secondary Phenotypes". *The American Journal of Human Genetics* **92,** 744–759 (Feb. 5, 2013).

271. Schram, M. T. & Stehouwer, C. D. "Endothelial Dysfunction, Cellular Adhesion Molecules and the Metabolic Syndrome". *Hormone and Metabolic Research* **37,** 49–55 (Apr. 2005).

272. Schwartz, M. "Neurodegeneration and neuroprotection in glaucoma: development of a therapeutic neuroprotective vaccine: the Friedenwald lecture". *Investigative Ophthalmology & Visual Science* **44,** 1407–1411 (Apr. 2003).

273. Senatorov, V. *et al.* "Expression of Mutated Mouse Myocilin Induces Open-Angle Glaucoma in Transgenic Mice". *The Journal of Neuroscience* **26,** 11903–11914 (Nov. 15, 2006).

274. Shabana, N., Cornilleau Pérès, V., Carkeet, A. & Chew, P. T. K. "Motion perception in glaucoma patients: a review". *Survey of Ophthalmology* **48,** 92–106 (Feb. 2003).

275. Sheffield, V. C. *et al.* "Genetic linkage of familial open angle glaucoma to chromosome 1q21–q31". *Nature Genetics* **4,** 47–50 (May 1993).

276. Shiose, Y. *et al.* "Epidemiology of glaucoma in Japan–a nationwide glaucoma survey." *Japanese journal of ophthalmology* **35,** 133–155 (1990).

277. Smith, A. M. *et al.* “The transcription factor PU. 1 is critical for viability and function of human brain microglia”. *Glia* **61,** 929–942 (2013).

278. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. “Pleiotropy in complex traits: challenges and strategies”. *Nature Reviews. Genetics* **14,** 483–495 (July 2013).

279. Sommer, A. *et al.* “Evaluation of nerve fiber layer assessment”. *Archives of Ophthalmology (Chicago, Ill.: 1960)* **102,** 1766–1771 (Dec. 1984).

280. Spector, T. D. & Williams, F. M. K. “The UK Adult Twin Registry (TwinsUK)”. *Twin Research and Human Genetics* **9,** 899–906 (Dec. 2006).

281. Spisak, K. *et al.* “rs2070424 of the SOD1 gene is associated with risk of Alzheimer’s disease”. *Neurologia I Neurochirurgia Polska* **48,** 342–345 (2014).

282. Spring, K., Lapointe, L., Caron, C., Langlois, S. & Royal, I. “Phosphorylation of DEP-1/PTPRJ on threonine 1318 regulates Src activation and endothelial cell permeability induced by vascular endothelial growth factor”. *Cellular Signalling* **26,** 1283–1293 (June 2014).

283. Springelkamp, H. *et al.* “Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process”. *Nature Communications* **5** (Sept. 22, 2014).

284. Steinhausen, K., STUMPFF, F., STRAUß, O., THIEME, H. & WIEDERHOLT, M. “Influence of Muscarinic Agonists and Tyrosine Kinase Inhibitors on L-type Ca2+Channels in Human and Bovine Trabecular Meshwork Cells”. *Experimental Eye Research* **70,** 285–293 (Mar. 2000).

285. Stephens, M. “A Unified Framework for Association Analysis with Multiple Related Phenotypes”. *PLoS ONE* **8,** e65245 (July 5, 2013).

286. Stölting, G. *et al.* “Direct interaction of CaV03b2 with actin up-regulates L-type calcium currents in HL-1 cardiomyocytes”. *The Journal of Biological Chemistry* **290,** 4561–4572 (Feb. 20, 2015).

287. Strange, A. *et al.* “A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1.” [**inlang**eng]. *Nat Genet* **42,** 985–990 (Nov. 2010).

288. Suuronen, T., Nuutinen, T., Ryhänen, T., Kaarniranta, K. & Salminen, A. “Epigenetic regulation of clusterin/apolipoprotein J expression in retinal pigment epithelial cells”. *Biochemical and Biophysical Research Communications* **357,** 397–401 (June 1, 2007).

289. Tabassum, R. *et al.* “Evaluation of DOK5 as a susceptibility gene for type 2 diabetes and obesity in North Indian population”. *BMC medical genetics* **11,** 35 (2010).

290. Takeuchi, F. *et al.* “Blood pressure and hypertension are associated with 7 loci in the Japanese population”. *Circulation* **121,** 2302–2309 (June 1, 2010).

291. Tamm, E. R., Grehn, F. & Pfeiffer, N. “Neuroprotection in glaucoma”. *Cell and tissue research* **353,** 201–203 (2013).

292. Tamura, H. *et al.* “High frequency of open-angle glaucoma in Japanese patients with Alzheimer’s disease”. *Journal of the Neurological Sciences* **246,** 79–83 (July 15, 2006).

293. Tatard, V. M., Xiang, C., Biegel, J. A. & Dahmane, N. “ZNF238 is expressed in postmitotic brain cells and inhibits brain tumor growth”. *Cancer Research* **70,** 1236–1246 (Feb. 1, 2010).

294. Teikari, J. M. “Genetic factors in open-angle (simple and capsular) glaucoma. A population-based twin study”. *Acta Ophthalmologica* **65,** 715–720 (Dec. 1987).

295. Thiyagesh, S. N. *et al.* “The neural basis of visuospatial perception in Alzheimer’s disease and healthy elderly comparison subjects: an fMRI study”. *Psychiatry Research* **172,** 109–116 (May 15, 2009).

296. Thorleifsson, G. *et al.* “Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma”. *Nature Genetics* **42,** 906–909 (Oct. 2010).

297. Thorleifsson, G. *et al.* “Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma”. *Science* **317,** 1397–400 (Sept. 7, 2007).

298. Tian, B., Geiger, B., Epstein, D. L. & Kaufman, P. L. “Cytoskeletal Involvement in the Regulation of Aqueous Humor Outflow”. *Investigative Ophthalmology & Visual Science* **41,** 619–623 (Mar. 1, 2000).

299. Tielsch JM, Katz J, Sommer A, Quigley HA & Javitt JC. “Family history and risk of primary open angle glaucoma: The baltimore eye survey”. *Archives of Ophthalmology* **112,** 69–73 (Jan. 1, 1994).

300. Tielsch, J. M., Katz, J., Sommer, A., Quigley, H. A. & Javitt, J. C. “Hypertension, perfusion pressure, and primary open-angle glaucoma. A population-based assessment”. *Archives of Ophthalmology* **113,** 216–221 (Feb. 1995).

301. Tielsch, J. M., Katz, J., Quigley, H. A., Javitt, J. C. & Sommer, A. “Diabetes, intraocular pressure, and primary open-angle glaucoma in the Baltimore Eye Survey”. *Ophthalmology* **102,** 48–53 (1995).

302. Toh, T. *et al.* “Central corneal thickness is highly heritable: the twin eye studies”. *Invest Ophthalmol Vis Sci* **46** (Oct. 2005).

303. Tonnu, P.-A. *et al.* “The influence of central corneal thickness and age on intraocular pressure measured by pneumotonometry, non-contact tonometry, the Tono-Pen XL, and Goldmann applanation tonometry”. *British Journal of Ophthalmology* **89,** 851–854 (July 1, 2005).

304. Trick, G. L., Trick, L. R., Morris, P. & Wolf, M. “Visual field loss in senile dementia of the Alzheimer’s type”. *Neurology* **45,** 68–74 (Jan. 1995).

305. Tsilis, A. G., Tsilidis, K. K., Pelidou, S.-H. & Kitsos, G. “Systematic review of the association between Alzheimer’s disease and chronic glaucoma”. *Clinical Ophthalmology (Auckland, N.Z.)* **8,** 2095–2104 (2014).

306. Valenti, D. A. “Alzheimer’s disease and glaucoma: imaging the biomarkers of neurodegenerative disease”. *International Journal of Alzheimer’s Disease* **2010,** 793931 (2011).

307. Valenti, D. A. “Alzheimer’s disease: visual system review”. *Optometry (St. Louis, Mo.)* **81,** 12–21 (Jan. 2010).

308. Van Koolwijk, L. M. E. *et al.* “Common Genetic Determinants of Intraocular Pressure and Primary Open-Angle Glaucoma”. *PLoS Genet* **8,** e1002611 (May 3, 2012).

309. VanderWeele, T. J. “Sufficient Cause Interactions and Statistical Interactions:” *Epidemiology* **20,** 6–13 (Jan. 2009).

310. Venturini, C., van Koolwijk, L. M., Bunce, C. & Viswanathan, A. in **by** Crowston, J. G., Sharaawy, T. & Hitchings, R. *Glaucoma.* 2nd (Elsevier, 2014).

311. Venturini, C. *et al.* “Clarifying the role of ATOH7 in glaucoma endophenotypes”. *The British Journal of Ophthalmology* **98,** 562–566 (Apr. 2014).

312. Vickers, J. C. *et al.* “The apolipoprotein epsilon4 gene is associated with elevated risk of normal tension glaucoma”. *Molecular vision* **8,** 389–393 (Oct. 14, 2002).

313. Visscher, P. M., Hill, W. G. & Wray, N. R. “Heritability in the genomics era — concepts and misconceptions”. *Nature Reviews Genetics* **9,** 255–266 (Apr. 2008).

314. Vodopiutz, J. *et al.* “Homozygous SALL1 mutation causes a novel multiple congenital anomaly-mental retardation syndrome.” [**inlang**eng]. *J Pediatr* **162,** 612–617 (Mar. 2013).

315. Vong, Q. P. *et al.* “TOX2 regulates human natural killer cell development by controlling T-BET expression”. *Blood* **124,** 3905–3913 (Dec. 18, 2014).

316. Wang, L.-D. *et al.* “Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54.” [**inlang**eng]. *Nat Genet* **42,** 759–763 (Sept. 2010).

317. Wang, J., Lin, L., Wang, H.-Q. & Chen, N. “PLCE1 rs2274223 polymorphism contributes to risk of esophageal cancer: evidence based on a meta-analysis”. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine* **35,** 6925–6931 (July 2014).

318. Watabe-Uchida, M., John, K. A., Janas, J. A., Newey, S. E. & Van Aelst, L. “The Rac activator DOCK7 regulates neuronal polarity through local phosphorylation of stathmin/Op18”. *Neuron* **51,** 727–739 (Sept. 21, 2006).

319. Watson, P., Townley, A. K., Koka, P., Palmer, K. J. & Stephens, D. J. “Sec16 Defines Endoplasmic Reticulum Exit Sites and is Required for

Secretory Cargo Export in Mammalian Cells". *Traffic* **7,** 1678–1687 (Dec. 1, 2006).

320. Weinreb, R. N. & Khaw, P. T. "Primary open-angle glaucoma". *The Lancet* **363,** 1711–1720 (May 2004).
321. Wen, J. *et al.* "Dok-5 is involved in cardiomyocyte differentiation through PKB/FOXO3a pathway". *Journal of molecular and cellular cardiology* **47,** 761–769 (2009).
322. Wentz-Hunter, K., Kubota, R., Shen, X. & Yue, B. Y. "Extracellular myocilin affects activity of human trabecular meshwork cells". *Journal of Cellular Physiology* **200,** 45–52 (2004).
323. Wiggs, J. L. *et al.* "Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma in Caucasians from the USA". *Human Molecular Genetics* **20,** 4707–4713 (Dec. 1, 2011).
324. Williams, S. E. *et al.* "Major LOXL1 risk allele is reversed in exfoliation glaucoma in a black South African population". *Mol Vis* **16,** 705–12 (2010).
325. Wilson, M. R. & Easterbrook-Smith, S. B. "Clusterin is a secreted mammalian chaperone". *Trends in Biochemical Sciences* **25,** 95–98 (Mar. 2000).
326. Wolfs, R. C. *et al.* "Genetic risk of primary open-angle glaucoma. Population-based familial aggregation study". *Arch Ophthalmol* **116** (Dec. 1998).
327. Woo, S. J. *et al.* "Investigation of the association between OPA1 polymorphisms and normal-tension glaucoma in Korea". *J Glaucoma* **13** (Dec. 2004).
328. Wordinger, R. J., Sharma, T. & Clark, A. F. "The role of TGF-β2 and bone morphogenetic proteins in the trabecular meshwork and

glaucoma". *Journal of Ocular Pharmacology and Therapeutics* **30,** 154–162 (2014).

329. Wright, S. "Evolution in Mendelian Populations". *Genetics* **16,** 97–159 (Mar. 1931).
330. Xu, L., Wang, Y., Wang, S., Wang, Y. & Jonas, J. B. "High Myopia and Glaucoma Susceptibility: The Beijing Eye Study". *Ophthalmology* **114,** 216–220 (Feb. 2007).
331. Xu, L., Zhang, H., Wang, Y. X. & Jonas, J. B. "Central Corneal Thickness and Glaucoma in Adult Chinese: The Beijing Eye Study". *Journal of Glaucoma* **17,** 647–653 (Dec. 2008).
332. Xu, L. *et al.* "Intraocular Pressure in Northern China in an Urban and Rural Population: The Beijing Eye Study". *American Journal of Ophthalmology* **140,** 913–915 (Nov. 2005).
333. Yajima, I., Kumasaka, M. Y., Tamura, H., Ohgami, N. & Kato, M. "Functional analysis of GNG2 in human malignant melanoma cells". *Journal of Dermatological Science* **68,** 172–178 (Dec. 2012).
334. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. "GCTA: A Tool for Genome-wide Complex Trait Analysis". *The American Journal of Human Genetics* **88,** 76–82 (Jan. 7, 2011).
335. Yang, J. *et al.* "Common SNPs explain a large proportion of the heritability for human height". *Nature Genetics* **42,** 565–569 (July 2010).
336. Yang, Q. & Wang, Y. "Methods for Analyzing Multivariate Phenotypes in Genetic Association Studies". *Journal of Probability and Statistics* **2012,** e652569 (July 15, 2012).
337. Yao, W. *et al.* "Evaluation of the association between OPA1 polymorphisms and primary open-angle glaucoma in Barbados families". *Mol Vis* **12** (2006).

338. Yeghiazaryan, K. *et al.* “An enhanced expression of ABC 1 transporter in circulating leukocytes as a potential molecular marker for the diagnostics of glaucoma”. *Amino Acids* **28,** 207–211 (Mar. 1, 2005).

339. Yücel, Y. H. *et al.* “Memantine protects neurons from shrinkage in the lateral geniculate nucleus in experimental glaucoma”. *Archives of Ophthalmology (Chicago, Ill.: 1960)* **124,** 217–225 (Feb. 2006).

340. Zetterberg, M. *et al.* “Apolipoprotein E Polymorphisms in Patients With Primary Open-Angle Glaucoma”. *American Journal of Ophthalmology* **143,** 1059–1060 (June 1, 2007).

341. Zhang, H., Ren, Y., Pang, D. & Liu, C. “Clinical implications of AGBL2 expression and its inhibitor latexin in breast cancer”. *World journal of surgical oncology* **12,** 1–7 (2014).

342. Zhou, Y., Grinchuk, O. & Tomarev, S. I. “Transgenic Mice Expressing the Tyr437His Mutant of Human Myocilin Protein Develop Glaucoma”. *Investigative Ophthalmology & Visual Science* **49,** 1932–1939 (May 1, 2008).

343. Zhu, H. *et al.* “Effects of AGBL2 on cell proliferation and chemotherapy resistance of gastric cancer.” *Hepato-gastroenterology* **62,** 497–502 (2014).

344. Zhu, W. & Zhang, H. “Why do we test multiple traits in genetic association studies?” *Journal of the Korean Statistical Society* **38,** 1–10 (Mar. 2009).

345. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. “The mystery of missing heritability: Genetic interactions create phantom heritability”. *Proceedings of the National Academy of Sciences* **109,** 1193–1198 (Jan. 24, 2012).