THE UNIVERSITY *of* EDINBURGH

Genetic analysis of retinal traits

Mirna Kirin

Presented for the Degree of Doctor of Philosophy

University of Edinburgh

August, 2013

Declaration

I declare that this thesis was composed by myself and that the work presented here is my own, unless otherwise stated.

The work has not been submitted for any other degree or professional qualification.

Signed:

Date:

Acknowledgements

The work presented in this thesis was carried out over a period of four years, while studying as a postgraduate student at the University of Edinburgh. I would like to thank the University of Edinburgh, as well as the Overseas Research Students Awards Scheme, for providing funding, without which this work would not come to fruition.

Furthermore, I would like to thank the members of the MRC Human Genetics Unit and Department of Population Health Sciences, who provided support and assistance during this period of research.

My special thanks go to my supervisors, Dr Veronique Vitart and Dr Jim Wilson, for their consistent guidance; patient sharing of knowledge; enthusiastic readiness for discussion and kind support.

I am also grateful to my fellow PhD students, especially to Jennifer Huffman for lively and constructive conversations that led to creative solutions on numerous occasions during this time.

A great deal of thanks is owed to my husband, Nasser, for his calm and patient presence, as well as endless encouragement and support throughout every step of the way.

Lastly, I would like to thank my family and friends for being there for me and for their unconditional understanding.

Abstract

Retina is a unique site in the human body where the microcirculation can be imaged directly and non-invasively, allowing us to study in vivo the structure and pathology of the human microcirculation. Retinal images can be quantitatively assessed with computerized imaging techniques, enabling us to measure several different quantitative traits derived from the retinal vasculature. Arterial and venular calibres are the most extensively studied traits of the retinal microvasculature and numerous epidemiological studies demonstrated promising associations with systemic and ocular diseases as well as with disease markers. However, there has been a lack of research into pathophysiological processes leading to retinal vascular signs, and how they link retinal microcirculation with coronary and cerebral microvasculature change. Information about genetic determinants underlying retinal vascular structure is therefore important for understanding the processes leading to microvascular pathophysiology. Two genome wide association studies have been published so far revealing four loci associated with retinal venular calibre and one locus with arteriolar calibre. Here the results from the genome-wide association analysis of 10 different retinal vessel traits in two population based cohorts are presented. Retinal images were measured in non-mydriatic fundus images from 808 subjects in the Orkney Complex Disease Study (ORCADES) and 390 subjects from the Croatian island of Korcula, using the semi-automated retinal vasculature measurement programme SIVA and VAMPIRE. Using pairwise estimates of kinship based on genomic sharing, heritability was calculated for each trait. Estimates of tortuosity measure and fractal dimensions present first published reports of heritability estimates for those traits. In addition correlation analysis with systemic risk factor was also completed, confirming already published results as well as revealing some new associations. A genome wide association analysis of retinal arteriolar width revealed a genome wide significant hit ($1.8x10^{-7}$) in a region of chromosome 2q32 (within *TTN* gene). Replication was sought in a further independent Scottish population (LBC) and additional 400 retinal images were graded. The result did not replicate, however the

direction of the effect was consistent and a larger sample size is required. Analysis of the remaining traits did not yield genome wide significant result,s and will also require larger sample sizes.

Genetic analysis of a binary retinal trait was also explored in a case control study of retinal detachment, which is an important cause of vision loss. A two-stage genetic association discovery phase followed by a replication phase in a combined total of 2,833 RRD cases and 7,871 controls was carried out. None of the SNPs tested in the discovery phase reached the threshold for association. Further testing was carried out in independent case-control series from London (846 cases) and Croatia (120 cases). The combined meta-analysis identified one association reaching genome-wide significance for rs267738 (OR=1.29, $p=2.11x10^{-8}$), a missense coding SNP and eQTL for *CERS2* encoding the protein ceramide synthase 2. Additional genetic risk score, pathway analysis and genetic liability analysis were also carried out.

Publication

The following publication has resulted from the work involved in genetic analysis of rhegmatogenous retinal detachment and is attached as Appendix VIII:

Kirin, M., Chandra, A., Charteris, D.G., Hayward, C., Campbell, S., Celap, I., Bencic, G., Vatavuk, Z., Kirac, I., Richards, A.J., Tenesa, A., Snead, M.P., Fleck, B.W., Singh, J., Harsum, S., Maclaren, R.E., den Hollander, A.I., Dunlop, M.G., Hoyng, C.B., Wright, A.F., Campbell, H., Vitart, V., Mitry, D. (2013). Genome-wide association study identifies genetic risk underlying primary rhegmatogenous retinal detachment. Hum Mol Genet 22, 3174-3185.

Table of contents Page

Chapter 1. Introduction

The purpose of this thesis was to examine and possibly decipher the genetic basis of quantitative retinal vasculature parameters using unselected population samples and of a binary trait, retinal detachment in a case-control study. Various analysis processes were included in this study. Analysis of quantitative vascular traits involved phenotypic data collection using semi-automated software for analysis of retinal microvasculature. In total 1593 retinal vessel images were graded and seventeen retinal microvascular phenotypes were derived and analysed for correlation with cardiovascular biomarkers, estimation of heritability, genetic correlation and genome wide association studies (GWAS). Analysis of a binary trait (rhegmatogenous retinal detachment) was a continuation of an on-going population-based epidemiology study of rhegmatogenous retinal detachment in Scotland. This part of the thesis involved different aspects of genetic analysis, from the selection and genotyping of genetic markers, data imputation, to genetic association analysis in a case control design and downstream pathway analysis. This chapter presents the background and the overview on the retina and retinal microvasculature, and its abnormalities. The pathology of rhegmatogenous retinal detachment is also presented, as well as the background on the genetic architecture of common disease. Finally the statement of aims and objectives of this project is presented.

1.1. Overview of the retinal anatomy and physiology

1.1.2. Anatomy and physiology of the eye

1.1.2.1 The eye (*Bulbus oculi*)

The eye of an adult is spherical organ and is approximately 25 mm in diameter. The interior of the eye is separated by the lens into two main cavities (anterior and posterior cavity). The anterior cavity is subdivided by the iris in two chambers (anterior and posterior chamber). Those two chambers connect through the pupil and are filled with a watery fluid (aqueous humour, the daily production of which maintains an internal eye pressure). Behind the lens is the posterior cavity filled with a transparent jellylike vitreous humour (which also contributes to the internal eye pressure to maintain the shape of the eyeball). Unlike aqueous humour, vitreous humour is not produced daily, but formed prenatally. The wall of the eye consists of three basic layers (fibrous tunic, vascular tunic, internal tunic-retina). The fibrous tunic is the outer layer of the eye and is divided into sclera (the white of the eye) and cornea (convex shaped to refract incoming light). The sclera is a protective outer layer of the eye and it maintains the shape of the eye. The cornea is transparent and allows light to enter the eye. The vascular tunic consists of the choroid, the ciliary body and the iris. The choroid contains the blood vessels that provide nourishment to the outer layers of the retina. The iris is the part of the eye that gives its colour. It consists of muscular tissue that responds to surrounding light. Light enters the pupil, falls onto the lens of the eye where it is altered before passing through to the retina. The lens is a transparent, biconvex structure and its function is to refract and focus incoming light onto the retina, where the light reaches photoreceptors and is transformed into neural impulses. The internal tunic of the eye or retina is the inner most layer of the eye (Zigmond et al. 1999, van de Graaff et al. 1997).

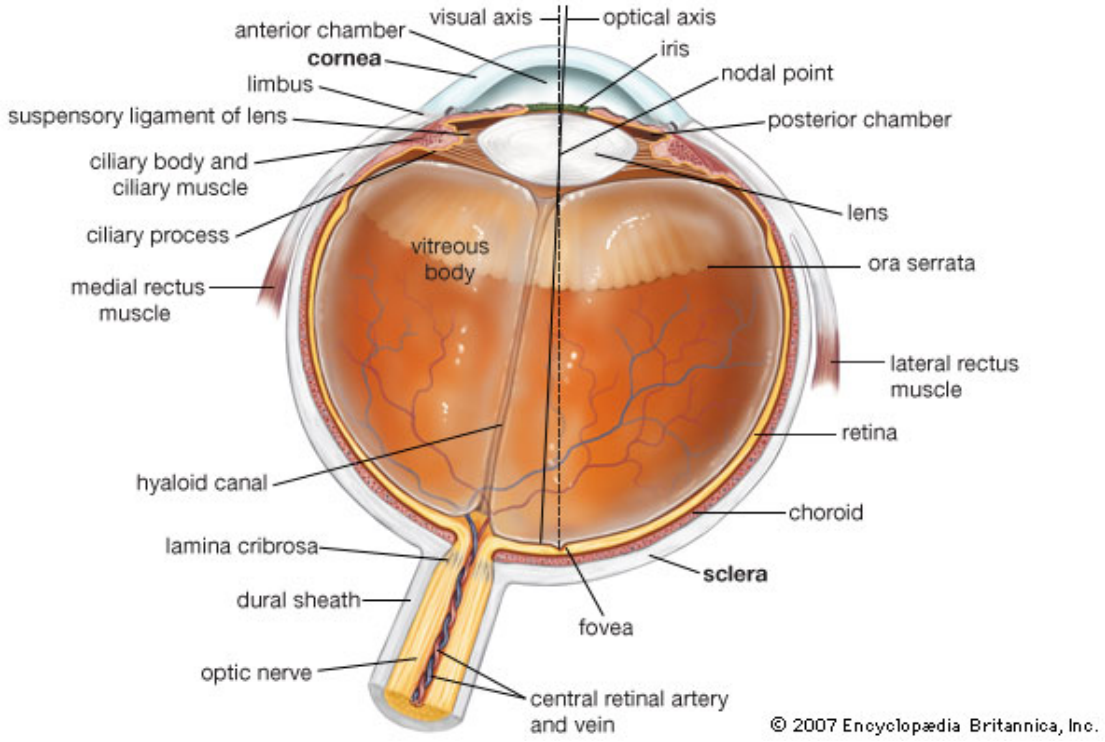
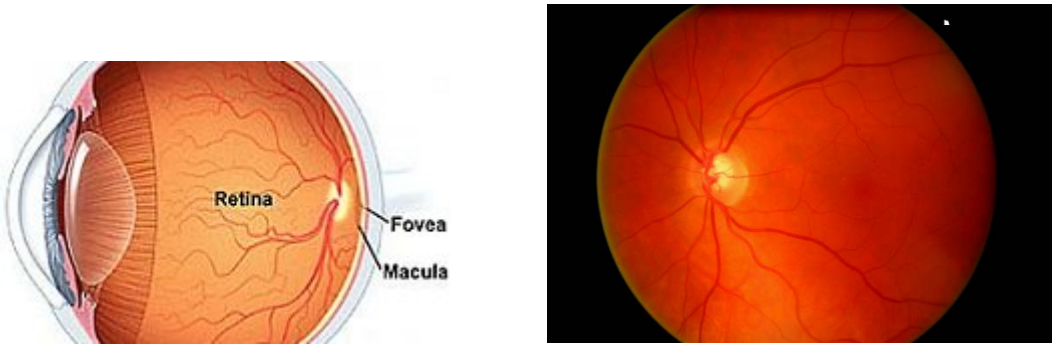


Figure 1.1. Anatomical diagram of human eye (from Encyclopaedia Britannica)

1.1.2.2. Retina

The retina consists of two layers, the outer pigmented layer and inner nervous layer or visual portion. The nervous layer is composed of three layers of neurons; rod and cone cells, bipolar neurons and ganglion neurons. Light first passes the layer of ganglion cells, then the bipolar cells and then stimulates the photoreceptor cells (rods and cones). Rods are longer and thinner and they respond to dim light to provide black and white vision. Cones are less sensitive to light and they provide colour vision and greater visual acuity. Cones are concentrated in a depression near the centre of the retina called the fovea centralis (the area of keenest vision). The macula lutea surrounds the fovea centralis and also contains an abundance of cones. There are no photoreceptors where

the optic nerve is attached to the eye. This area is called the blind spot and is usually referred to as the optic disc. The retina is considered part of the central nervous system (CNS). Embryologically, it is formed from the optic cup, an outpouching of the neural tube. It is connected with the brain through a bundle of fibres (the optic nerve) and is the only part of the CNS that can be visualised non-invasively (Zigmond et al. 1999, Taarnhoj 2008, van de Graaff et al. 1997).



a) **b)**

Figure1. 2 a) Cross-sectional view of the human eye (form science.nasa.gov); b) Retinal fundus image from ORCADES

1.1.2.3. Anatomy and physiology of the retinal vascular system

There are two different arterial systems that supply, abundantly, the retina with blood. The choroidal circulation supplies the outer third of the retina and the branches from the central retinal vessels that supply the inner two thirds of the retina. The central retinal artery and vein branch from the ophthalmic artery and vein and they pierce the optic nerve and come to the retina through the gap in the lamina cribrosa. Both the retinal artery and vein branch into four major vessels. Each of the four major retinal artery

branches supplies one quadrant of the retina. The only part of the retina that is not supplied by the retinal artery is the fovea, which is a capillary free zone in order to allow light to pass through. The fovea is supplied by choriocapillaries. The perfusion pressure of the ocular circulation is related to both systemic blood pressure and intraocular pressure. For the retinal circulation to regulate blood flow over a range of different systemic blood pressures, its vascular resistance has to adapt to it. This alteration of the vascular resistance is regulated by vascular smooth muscles and pericytes (connective tissue cells that occur in small blood vessels) (Taarnhoj et al. 2008, Zhang, 1994, Henkind et al. 1979, Patton et al. 2005).

Retinal blood pressure is regulated by the combination of myogenic (smooth muscle reaction) and metabolic mechanisms. Metabolic mechanisms act in way to increase blood pressure in areas of increased metabolic activity, especially under the conditions of light exposure (Yu et al. 2007). The vascular endothelium cells and the neural and glial cells are mainly involved in the regulation of retinal blood pressure. Nitric oxide (NO) play a vital role in dilation of retinal vessels when flickering light is projected onto the retina. Myogenic mechanisms are defined as the contraction of vascular smooth muscle cells and pericytes in response to transmural pressure (difference in pressure between two sides of a wall) (Ursino, 1991). However, a big increase in systemic blood pressure and intraocular pressure can overcome these auto regulatory mechanisms causing the retinal blood flow to increase (Robinson et al. 1986).

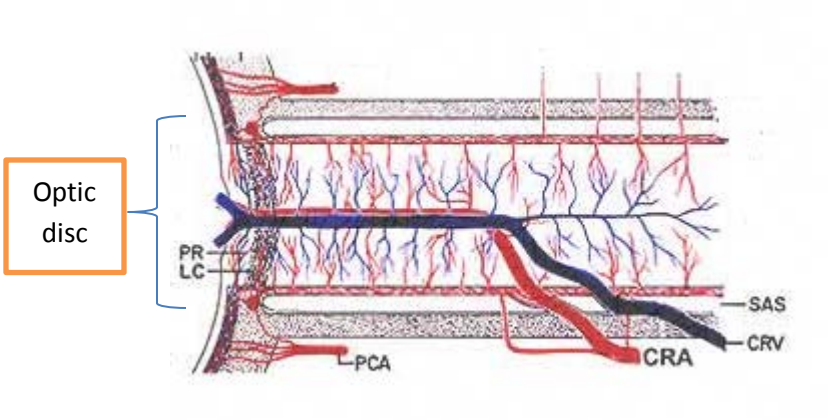


Figure 1.3. Vascular system supplying the retina. CRA –central retinal artery; CRV – central retinal vein; SAS –subarachnoid space; PCA- posterior ciliary artery; PR –prelaminar region; LC-lamina cribrosa. Modified from Hayreh, (1974).

1.1.2.4. Anatomy of the optic disc and cup

The optic disc is the location where ganglion cell axons exit the eye to form the optic nerve. The diameter of the optic nerve is 1.62mm, made up of about 1 million nerve fibres that form the optic disc. In this area of the eye there are no light sensitive rods or cones and therefore no response to light stimuli. It is place 3 to 4 mm to nasal side of the fovea. The optic cup is the white, cup-like area in the centre of the optic disc. The cup is really an empty space in the middle of the optic nerve surrounded by optic nerve fibres. The physiologic cup size varies from less that 0.1 to 0.9 of the total disc diameter and the average cup to disc ratio is about 0.3. The larger the cup, the larger the scleral opening at the back of the globe. The space between the cup and the edge of the disc is the neuro-retinal rim. The area around the disc is called the peri-papillary disc area. (Digre and Corbet, 2002).

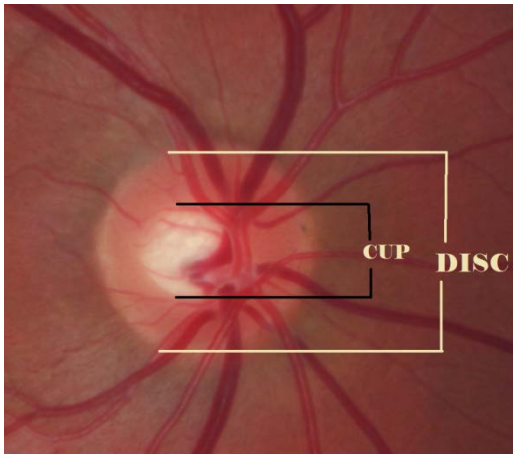


Figure 1.4. Retinal fundus image displays of optic disc and cup (Orkney image).

1.2. Retinal microvascular changes in response to stress and disease

Most changes to the retinal microvasculature are caused by hemodynamic changes, oxygen saturation and metabolic changes associated with elevated blood pressure, diabetes mellitus and old age. Persistent elevated blood pressure results in increased peripheral vascular resistance, which is a result of protective vasoconstriction of arterioles. Increased arteriolar intraluminal pressure leads to narrowing of the vessel through stretch-activated calcium channels (Muraki et al. 2003). This narrowing of the arterioles is seen best in vessels where sclerotic changes did not start and the vessel most affected are the precapillary arterioles (Garner et al. 1975). Persistent hypertension (and diabetes mellitus) can lead to arteriosclerotic changes to the small vessels, that include intimal thickening, hyperplasia of the tunica media and hyaline degeneration of the arteriolar wall (the general architecture of blood vessels is similar across the human vasculature and is given in Figure 1.5.). Arteriosclerotic changes also occur in a normal aging population. Long term hypertension can lead to degeneration of vascular smooth muscle and endothelial cell necrosis, along with narrowing of the arteriolar lumen and problematic blood flow (Wong et al. 2001).

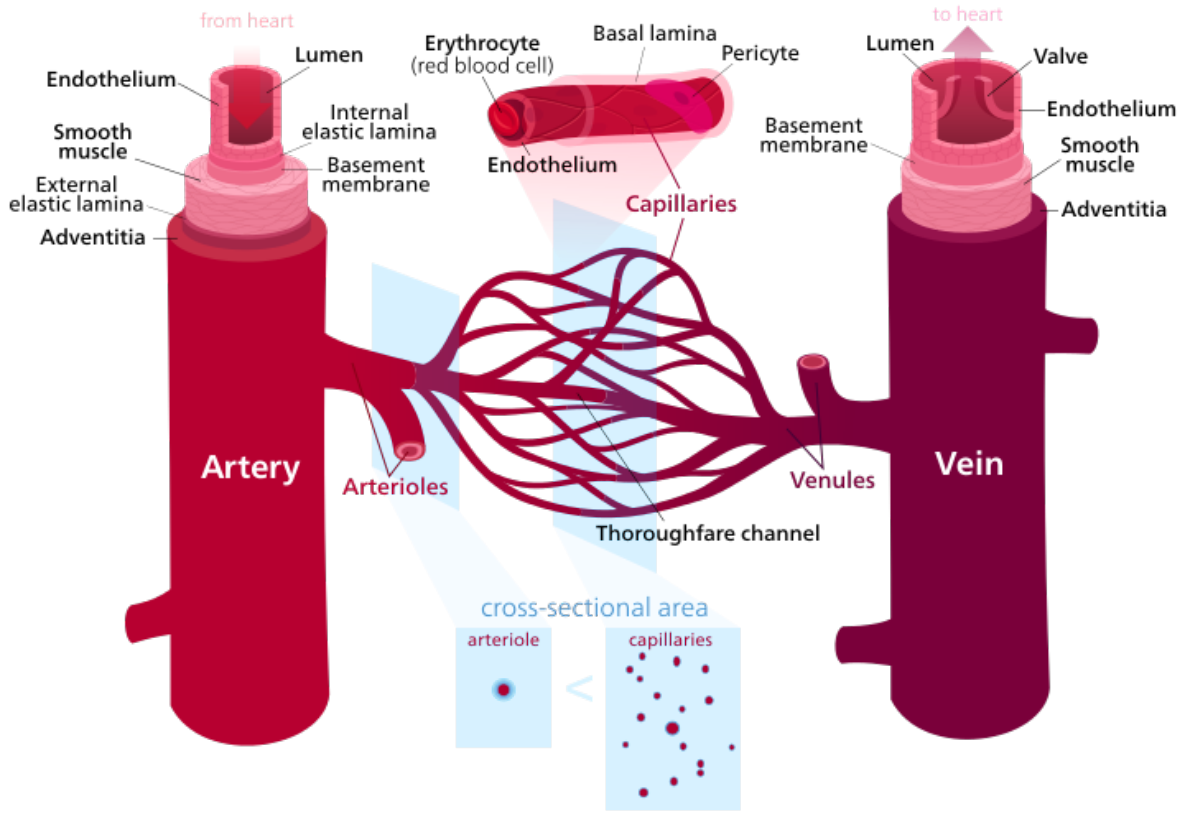


Figure 1.5. Schematic image of different layers of blood vessels (taken from Wikipedia). Both arteries and venes have three different layers. Tunic intima (layer of endothelial cells surrounded by the connective tissue), tunica media (may be rich in vascular smooth muscle (arteries), which control the caliber of the vessel) and tunica adventitia (outer layer made of connective tissue).

The retina adapts to variations in oxygen saturation, especially to oxygen requirements of photoreceptors. Hypoxia has a profound effect on retinal microvasculature as its chronic form leads to an increase in Hypoxia-inducible factor (HIF)-1alpha, which in turn induces the up-regulation of molecules such as endothelin-1 (ET-1) and vascular endothelial growth factor (VEGF). These molecules can induce neovascularisation, weakening of the blood–retina barrier, and vasoconstriction of veins.

Type 1 and Type 2 diabetes mellitus patients present retinal microvascular changes, as a result of several metabolic stresses (elevated blood glucose levels, dyslipidaemia) and hypertension. Increased blood flow and lack of auto-regulation (myogenic and

metabolic) are the main features of diabetic retinopathy. In addition to these haemodynamic abnormalities, increased capillary and venular pressure and diameters, thickening of the basement membrane and loss of perycites are observed. A combination of these abnormalities leads to chronic hypoxia.

Retinal arterial and venous narrowing have also been reported to be associated with open-angle glaucoma in the Blue Mountain Eye Study (Mitchell et al. 2005), although this finding was not seen in another study, the Beaver Dam Eye study (Klein et al. 2004).

1.2.1. Retinal vessel microvasculature and cardiovascular disease

Since retinal vasculature is thought to characterise the vascular system elsewhere in the body and it is the only part of the human vasculature that can be visualised non-invasively in vivo in humans, retinal imaging has developed into a very commonly used and accessible way of studying vascular traits. Retinal vasculature is thought to share the same physiological and anatomical characteristics as cerebral and coronary vascular systems (Goto et al.1972). The main hypothesis in retinal microvascular research is that retinal vessels parallel pathological processes occurring in coronary microvessels and can represent processes occurring there, since they are approximately the same scale as coronary microvasculature (~100-250μm in diameter). There is some evidence that retinal microcirculation changes do parallel pathophysiological changes in the coronary micro and macro circulation. Narrowing of the retinal arterioles have been reported to be associated with coronary artery disease diagnosed by angiography (coronary artery occlusion) (Michelson et al. 1979, Tedeschi-Reiner et al. 2005) although, these studies did not use a quantitative measure of vessel width, and included a very small number of subjects (respectively N=70and, N=109.). In addition a study using MESA (multi ethnic study of atherosclerosis) cohort found association between retinal arteriolar narrowing

and retinopathy with MRI findings suggestive of left ventricular remodelling (Cheung et al. 2007). Recent studies using large population-based cohorts and quantitative vessel calibre measurements have found significant associations between retinal microvascular changes and incident coronary disease or cardiovascular death (Wong et al. 2002, Witt et al. 2006, Wong et al 2006, Wang et al. 2006, Cheung et al. 2007). The ARIC study (atherosclerosis risk in community study) found a two-fold increased risk of incident heart failure in all persons with signs of retinopathy, however no significant relationship between retinal arteriolar narrowing was found (Wong et al. 2005). Similarly no association was found between increased coronary artery calcium score and vessel calibre (Wong et al. 2008). There is also evidence of alterations in retinal microvasculature in patients that suffered lacunar stroke (Schneinder et al. 1993). Despite that some evidence exists supporting the hypothesis that retinal microvascular changes give insight into coronary microcirculation, the extent to which it represents microvasculature changes elsewhere in the body still remains unclear. Direct post mortem histopathological evidence is still needed in order to confirm this homology.

The important question in microvascular research is whether these changes are the consequences of the systemic vascular disease or if they play a part in its development, or both. It has been proposed that changes in renal microvasculature (decrease in arteriolar calibre) occur in the early stages of essential hypertension (Norrelund et al. 1994). A few prospective studies have presented evidence that narrowing of the arteriolar calibre precedes the development of hypertension (Wong et al. 2004, Wong et al. 2004, Ikram et al. 2006), supporting the hypothesis that it may be an integral part in the development of cardiovascular disease.

Vascular architecture develops in a way that optimises the efficacy of blood flow, as theoretical and experimental evidence suggest, and deviation from this optimality occurs in pathological processes (Zamir, 1976). Retinal microvascular parameter that can

contribute to this blood flow optimality (tortuosity, branching angles, fractal dimensions - a measure of the general complexity of microvasculature) should also help in understanding the changes that occur in microvasculature during cardiovascular disease. However these measurements have not been studied as extensively as vessel calibres, due to the unavailability of more automated measurements. Progress in the field of computerised imaging software enabled more accurate and reproducible measurements of these quantitative retinal vessel traits. New softwares are now available, and will allow inclusion of these geometrical parameters of retinal vasculature in the explaining the pathological changes of the microvasculature.

Quantitative retinal vessel characteristics have been measured in several epidemiological studies so far and demonstrated promising associations with systemic and ocular diseases as well as with disease markers (more detailed description given in Chapter 5). All the associations found so far account for a small proportion of variation in retinal vessel morphology. There is therefore scope for studying more traits describing retinal architecture. In particular, the extent of genetic research in the area of retinal microcirculation has been limited. Epidemiological studies have focused primarily on the associations of retinal vessel characteristics and systemic risk factors, and only a few of them explored the heritability of some of the possible retinal vessel traits. Determining the genetic basis of these traits could give us more insight into the pathogenesis of the associated diseases.

1.3. Genetic analysis of retinal microvasculature

The identification of genetic factors influencing common diseases such as cardiovascular and metabolic diseases is a major goal in medicine today. Identification of susceptibility variants often gives new insights into disease pathways and can help in identifying factors contributing to the earlier and more rapid development of the disease. It can also help in the development of genetic risk profiling either alongside or in place of traditional risk prediction using conventional risk factors. Revolutions in genetic technology, deriving originally from the human genome project have driven the discovery of over 2000 gene-disease or disease-trait associations (Hindorff et al. 2009). An understudied, but important area of inquiry is the retina, since its vasculature is thought to share the same physiological and anatomical characteristics as the cerebral and coronary vascular systems (Goto et al. 1975).

Only a few studies so far have investigated the contribution of genetic factors to the morphology of retinal blood vessels. Lee et al. (2004) reported the first study about familial aggregation of quantitative retinal vessel traits. They reported a high correlation of retinal vessel calibres (CRAE –central retinal arterial equivalent, CRVE - central retinal arterial equivalent, AVR –arteriolar to venular diameter ratio) between relatives (more highly correlated than between unrelated individuals). Taarnhoj et al. (2006) reported heritabilities of 70% for CRAE, 61% for CRVE and 82% for tortuosity in a very small numbers - 55 monozygotic and 50 dizygotic same sex healthy individuals. Two studies have reported results of the linkage analysis of quantitative traits of retinal vessel (Xing et al. 2006, Sun et al.2009). Xing et al. reported both specific and common signals for CRAE and CRVE (common signals in 3 regions and 7 specific regions for both CRAE and CRVE). The Australian Twins Eye study reported two multipoint peaks on chromosome 3p12.3 and 8p23.1 for CRAE, and two suggestive logarithm odds score for CRVE (on 2p14 and 9q21.13). Finally, two large scale genome wide meta-analyses have reported genetic association between retinal arteriolar and venular calibres. Sim et

al. (2013) reported one marker significantly associated with CRAE on chromosome 5. There was no overlap between linkage studies and GWAS results, however both studies reported regions common to both retinal arteriolar and venular calibres (Ikram et. al 2010, Sim et al. 2013). Four markers were reported to be associated with venular calibre in a large meta-analysis involving 7 different cohorts (Ikram et a.l 2010).

1.4. Retinal detachment

Retinal detachment is a condition in which the neurosensory retina separates from the underlying retinal pigment epithelium causing accumulation of the fluid in this space. The pigmented layer is attached to the choroid (highly vascular layer) and therefore retinal detachment leaves the neurosensory retina deprived of oxygen. The longer the separation lasts, the greater the risk of permanent vision loss in the affected eye. Retinal detachment can be separated into four types depending on the underlying causes (Ghazi and Green, 2002), but all types have in common the accumulation of subretinal fluid. Rhegmatogenous retinal detachment occurs due to a hole, tear, or break in the retina that allows fluid to pass from the vitreous space into the subretinal space between the sensory retina and the retinal pigment epithelium. Exudative retinal detachment occurs due to inflammation, injury or vascular abnormalities that results in fluid accumulating underneath the retina without the presence of a hole, tear, or break. Tractional retinal detachment occurs when fibrovascular tissue, caused by an injury, inflammation or neovascularization, pulls the sensory retina from the retinal pigment epithelium. The fourth type is combined traction-rhegmatogenous. Predisposing factors for retinal detachment are vitreoretinal adhesions in association with posterior vitreous detachment, local ocular diseases such as retinoschisis and myopia, cataract surgery, trauma, and most of the hyaloideoretinopathies. Most of these risk factors are predisposing to RRD

because of the high incidence of associated vitreous detachment and retinal breaks (Ghazi and Green, 2002).

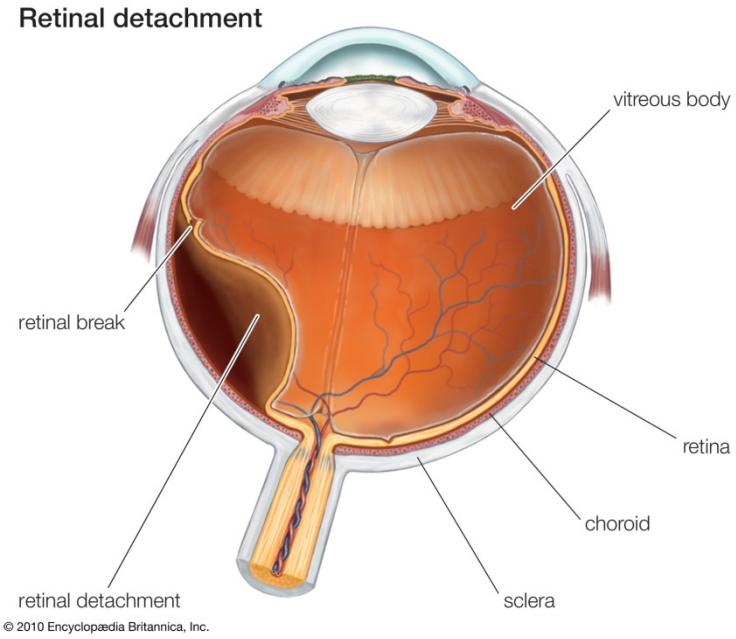


Figure 1.5. Schematic image of retinal detachment (from Encyclopaedia Britannica, 2010).

The vitreous is attached to the retina through a vitreous base, and the layer between vitreous and neural retinal is called the inner limiting membrane (ILM). ILM is a typical basement membrane and has a role in the adhesion of the vitreo-retinal border to the Muller cell (glial cell that provide support to the retinal neurons; they also channell light to the photoreceptor cell in the retina). The vitreous is comprised of collagen fibrils and glycoaminoglycans, and is supported by hyaluronic acid molecules. With aging the vitreous may collapse and liquify, due to the reduction in hyaluronic acid which causes loss of support to the collagen, resulting in detachment. As vitreous detachment occurs peripherally, areas of vitreoretinal adhesion can result in a tear in the sensory retina which can allow liquid vitreous to enter the potential subretinal space, resulting in a rhegmatogenous retinal detachment. However not all PVD lead to retinal tears. It occurs more often in myopic eyes, which have more pronounced vitreous liquefaction. Retinal break is important in formation of RRD, however most eyes with breaks do not develop

RRD as vitreoretinal traction is also needed to keep the break open. As described above, the separation of the neural retina from the underlying retinal pigment following a retinal tear, leaves the neurosensory retina deprived of oxygen and nutrients. The adhesion between the NSR and retinal pigment epithelium is relatively weak, since no junctional attachments exist between these two layers (Steinberg et al. 1974). The adhesion is maintained by the physical interconnection of the RPE cells' apical microvilli with the outer segments of NSR, mechanical support of the interphotoreceptor matrix, and through the net fluid transport from retina to choroid. Once these forces are overwhelmed, the interphotoreceptor space expands allowing for the fluid to accumulate and retinal detachment to occur. The exact mechanisms involved in this complex and synergistic adhesion system are not well understood.

1.4.1. Rhegmatogenous retinal detachment

Rhegmatogenous retinal detachment (RRD) is a major cause of severe vision loss. Despite surgical advances it accounts for up to 4% of the blind and partial sight registrations in Ireland (Munier et al. 1998). On average the annual incidence of RRD is 10-20 cases in 100,000 individulas (Mitry et al. 2010), with big differences between races (Wong et al. 1999, Mowat et al. 2003, Peters et al. 1995). The rate of RRD is reported to be higher in elderly patients (Haimann et al. 1982, Wong et al. 1999, Sasaki et al. 1995) and in men compared to women (Algvere et al. 1999, Ivanisevic et al. 2000, Polkinghorne et al. 2004, Mowat et al. 2003). Other risk factors for rhegmatogenous retinal detachment include high myopia, lattice degeneration, trauma and previous cataract surgery. Up to 10% of RRD cases will also develop retinal detachment or horseshoe tear (precursor of retinal detachment) in the other eye (Richards et al. 2005).

1.4.1.1. Genetic predisposition to RRD

Evidence of the genetic predisposition to RRD comes from a number of inherited disorders (with known causal genetic variants) which are associated with RRD and with familial aggregation studies reporting familial incidence of non-syndromic RRD (Mitry et al. 2011). Familial aggregation studies reported increased risk in the relatives of RRD patients. Siblings of patients with RRD presented two to three time higher frequency of the disease than the general population (Go et al. 2005, Mitry et al. 2011). However these studies used a relatively small number of sibships (N=501, Mitry et al 2011, N=181 Go et al. 2005). RRD has previously been reported to occur as a dominantly inherited trait (Richards et al. 2005, Go et al. 2003), but some of the cases had minimal systemic characteristic of Stickler syndrome and linkage analysis mapped the disease region to the COL2A1 gene (linked to Stickler's syndrome) leaving the distinction between the two conditions uncertain (given the high variability of the phenotype of the syndrome). Very little is known about the genes underlying non-syndromic RRD in general population and no previous study has investigated common genetic variants associated with non-syndromic RRD.

1.5. Genetic architecture of complex disease

Explaining the heritable component of many complex diseases is one major avenue in biomedical research today to better understand their etiologies. Complex traits and disease are believed to result from the interplay of multiple susceptibility genetic variants and environmental factors. Genetic basis for common and complex diseases has been supported by evidence of familial clustering and family history in the diagnosis of such diseases. Finding genes influencing common late-onset human disease is important

for a number of reasons. The ultimate goal is to translate these findings into public health benefits and improvements in clinical decision making and disease management.

1.5.1. Genetic analyses of complex traits

Associations between complex disease and genotypes can be studied by genotyping candidate genes with a prior hypothesis of being in disease pathways, or agnostically by testing associations with genetic variants capturing a large part of the genome variations as facilitated by the amount of linkage disequilibrium in the human genome (linkage or association studies). Candidate gene studies presented very little success in identifying genes underlying complex diseases, mainly because of the incomplete knowledge of disease pathways and assumption involved in the design of these types of studies.

Linkage and genome wide association studies have the advantage of no a priori hypothesis regarding biology of the disease. Linkage studies necessitate large pedigrees of individuals who are measured for the trait or disease, and are aimed at identifying regions of the genome that segregate with disease. The regions identified through linkage analysis tend to be very large (millions of base pairs). This type of analysis has not been very successful for complex traits since its power is in detecting variants that , singly or as a group of closely linked markers, account for large proportion of the phenotypic variance (large effect) and high penetrance, a non-common scenario for complex trait (Risch and Merikangas 1996). There were simply very few such loci to be found, linkage studies did not have the power to identify loci of smaller effect.

Association analysis that test association between trait variation and genotypes of large numbers of genetic markers in a large sample of individuals, under the assumption that

individuals with similar phenotypes share an excess of the same risk variants are better suited for the dissection of complex traits. It represents a population-based alternative mapping approach to linkage analysis and is well powered for mapping of common variants (minor allele frequency, MAF > 0.05) of small effect size if large sample size can be assembled (Risch and Merikangas 1996). The principle is based on the sharing of a small genomic region identical by descent between distantly related individuals. These blocks descend from the same ancestral chromosome and over successive generations recombination reduces the portions of DNA passed on the successive generation which are identical to the original ancestral segment of DNA. These shared DNA regions are broken down to small blocks of a few tens to hundreds of kilobases and genetic loci within DNA blocks are described as being in linkage disequilibrium (LD) within a population. LD occurs due to recombination and is a description of the non-random association of alleles from two or more loci on the same chromosome. The extent of LD between two loci is calculated by checking how often the same alleles at these loci occur together compared to what would be expected by chance, given their allele frequencies. Haplotypes are blocks of alleles at neighboring loci that are inherited together, and therefore a smaller set of alleles can represent the entire haplotype block. This is the principle of tagging markers, which represent haplotype blocks, and other alleles within the block, throughout human genome. By carefully choosing tagging markers number of these markers needed to represent the whole genome is radically reduced, at least for the common variants coverage. Single nucleotide polymorphisms (SNPs) are the markers used in the whole genome association studies. A SNP is a DNA sequence variation and occurs when a single nucleotide may take the form of either of two or more bases (A, T, C, G). SNPs are usually bi-allelic and there are thought to be about 30-50 million common SNPs in the human genome. Mutation rate of single base substitution is approximately 1.3×10^{-8} per site per generation (Manolio et al. 2008), and the chance of reversal of mutation is very low, which makes them heritable. However, their frequencies can be the result of ancestral founder effects and can be very different between different populations.

Regions identified through linkage studies are typically 5–10 Mb in size (containing tens to hundreds of genes), while association studies are able to define quantitative trait or risk loci to roughly 10–100 kb (containing a few genes).

1.5.1. Genome wide association analysis (GWAS)

The basic principle behind genome wide association studies is that a marker located within the same haplotype as the causal variant should also be associated with a trait of interest. And by testing numerous tagging SNPs throughout the genome, haplotypes containing the causal variants will be identified. Identification of the tagging SNPs became possible after the release of the international HapMap Project (HapMap consortium 2005) which revealed a haplotype map of the human genome. Technology which enabled GWAS is a high-throughput method of hybridization of DNA in order to identify polymorphisms at specific nucleotides over the entire genome. This gene-chip technology allows for simultaneous genotyping of hundreds of thousands to millions of SNPs. Gene chips have oligonucleotide probes attached to a solid base and are used to hybridize a target sequences containing the SNP of interest. DNA is firstly fragmented and fluorescently labelled in order to generate a signal that depends on the hybridization conditions. The strength of the fluorescent signal is determined by the sequence/polymorphism at that site. Fluorescent signal intensities are then analysed and converted into a genotype.

As mentioned before GWAS test the correlation between allele counts throughout the genome and phenotypes. One of the most important steps in the analysis is accurate phenotype definition, disease or trait, since phenotypic heterogeneity can reduce power

(Ioannidis et al. 2009). The statistical power of GWAS depends on sample size, effect size of the causal allele, its allele frequency, and how well it is captured by the panel of markers tested i.e. the correlation between marker and causal variants (Stranger et al. 2011). Therefore large sample sizes are needed to detect moderate to small effect sizes, and joint meta-analysis of multiple populations have been proven the best way to detect variants of moderate effect sizes in GWAS. One of the major issues in GWAS is the multiple testing, given the vast number of tests performed and resulting Type I error. A conservative Bonferroni correction is used; for genotyped SNPs the most used threshold has been the value set by the Wellcome Trust Case Control Consortium of 5×10^{-7} (WTCCC, 2007). For analysis of imputed data the general consensus is to correct for the number of independent SNPs in non-African population (estimated at a million independent variants); $p=5x10^{-8}$ (Hoggart et al. 2008, Pe'er et al. 2008). However, by controlling for Type 1 errors we may inflate Type 2 errors, particularly in the analyses with low power (small sample sizes, low frequency alleles, genetic factors with small effect sizes). This is usually overcome by inclusion of large sample, however due to various restrictions in practice this is not always possible. These issues are major considerations both for designing GWAS and interpreting GWAS results.

There are two main GWAS study designs, case-control and quantitative trait analyses. Quantitative trait studies are easier to set-up (as aiming at explaining the natural variation in the trait it requires sampling cross-population) and allow to study a deep array of phenotypes in the same study. As some quantitative traits are strong predictors of disease and are considered to be intermediate phenotypes (endophenotypes) of more complex disease (e.g. HDL levels for cardiovascular disease), they can often be used to identify genes underlying the correlated disease. A big advantage of a quantitative trait over complex disease end-points is that they are less complex in nature. Since endophenotypes have fewer sources of variation than complex diseases, a process of identifying the variation underlying them will be simpler and have more statistical power. In addition many quantitative phenotypes are easier to define and measure, which is not the case with complex disease which often encompasses heterogeneous

conditions. By identifying risk variants underlying intermediate traits, new insights into the pathogenesis of complex diseases may be exposed. However some diseases do not have well characterized endophenotypes and testing is performed on a qualitative trait (affected, not affected). In this scenario it is of great importance to correctly identify affected individuals, as well as selecting matched unaffected controls.

Once genome wide significant results are detected they should be validated though replication in an independent set of subjects. The replication sample should be of sufficient sample size in order to detect the effect of the allele. However, since initial GWAS usually suffer from the effect of "winners curse", the sample sizes should generally be larger in order to account for the over-estimation of the effect sizes. It is also of great importance to use the same phenotype inclusion criteria as for the original GWAS population. In order for a marker to be replicated besides the significant p-value, it should also report the same direction and ideally size of the effect across all populations tested.

SNPs identified in GWAS are unlikely to be the causal variants influencing a particular phenotype. Due to the LD pattern and selection of tagging SNPs, it is in most cases in high LD with the variant influencing a trait (Figure 1.6). In this case further work involving sequencing and possible functional work is needed in order to elucidate the real causal variant. The final goal of GWAS is to identify this causal variant and to characterize the functional effects. This involves fine mapping of an associated locus, followed by deep resequencing of the associated region which reveals possible functional variant. This is then followed by different bioinformatic and genomic methods in order to select variants for experimental studies to reveal the functional consequences of underlying variants

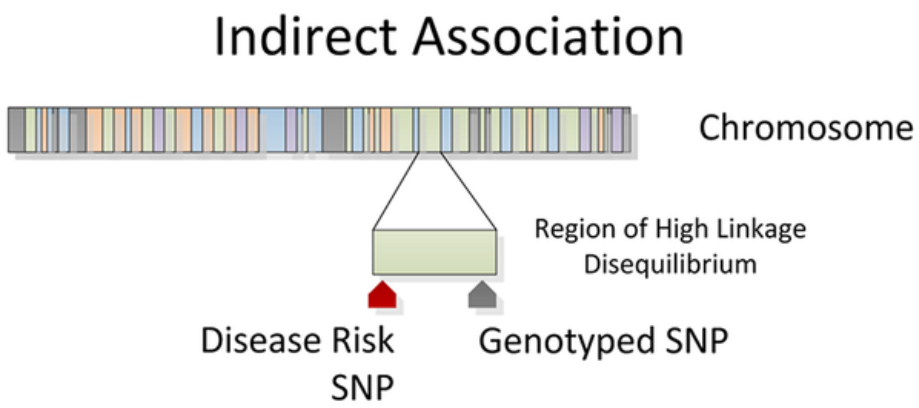


Figure 1.6. SNPs identified in GWAS are usually not the causal variants underlying a particular phenotype. Due to the LD pattern and selection of tagging SNPs, it is in most cases in high LD with the variant influencing a trait (Bush and Moore 2012).

1.5.2. GWAS design achievements

Genome wide association studies in recent years, in which hundreds of thousands to millions of genetic markers (SNPs- single nucleotide polymorphisms) are genotyped in typically tens of thousands of individuals have helped to find over a thousand of common genetic variants associated with such diseases and endophenotypes. These studies helped to gain new insights into the pathogenesis of complex human diseases and the underlying genetic architecture, and have identified around 2000 genotype-phenotype associations with 300 different complex diseases and traits in around 1600 publications since 2005 (Manolio et al. 2013). The main driving force behind GWAS was the common disease/common variant (CD/CV) hypothesis (Lander, 1996), which suggested that small number of genetic variants, which occur commonly in the population (>5% of the population; Reich and Lander, 2001) can account for the genetic

variance of many common diseases. Another reason why GWAS become the focal point of genetic research in the last few years was the development and drop in price of off-the-shelf commercial SNP arrays represent most of the common variation in the genome. Although GWAS helped to access some of the genetic architecture of common disease, most of the variants discovered explained only a small portion of the heritability of these diseases, even when as many as 40-60 variants affecting a trait have been identified (Voight et al. 2010, Morris et al. 2012, Scott et al. 2012), and raised a question of the so-called "missing herritability" problem. Identifying the sources of missing herritability still remains unresolved and the contribution of rare (<1% minor allele frequency; MAF), and low frequency (1-5% MAF) variants not known. Many reasons explaining missing heritability has been proposed (Manolio et al. 2009), and since then three major causes contributing to genetic component have been recognised; a large number of small effects common variants across the allele frequency spectrum (the infinitesimal model), a large number of large effects rare variants (the rare allele model) and combination of genotypic, environmental and epigenetic interactions (broad sense heritability model) (Gibson, 2012).

Empirically there is evidence that both common and rare variants play a role in the genetic architecture of complex diseases. As it was previously mentioned, around 1500 significant associations with around 300 traits have been listed in the US National Human Genome Resources Institute GWAS catalogue. Also re-sequencing efforts of regions identified through GWAS have presented coding mutations conferring increases in risk for disease (Nejentsev et al. 2009).

1.5.3. Isolated populations and GWAS

The use of isolated population in genetic research has a long history in genetic mapping studies. They represent small, geographically (or culturally) isolated populations, removed from the migration processes, which in turn results in endogamy. Such populations have higher levels of background inbreeding as a result of cryptic relatedness. This results in differences in allele frequencies from the more cosmopolitan populations due to both random genetic drift and the increased levels of homozygosity (McQuillan et al. 2009). Increased levels of homozygosity amplify the probability that a rare allele will be homozygous, increasing somewhat the chances of finding them.

Easier detection of variants underlying complex traits also come from the founding effects and population bottlenecks that are typical for isolated populations. These effects accentuate the effect of genetic drift or natural selection and in turn lead to reduced genetic heterogeneity (multiple alleles influencing a trait). Some allele will be removed and some polymorphic alleles will be fixed (Pardo et al. 2005), resulting in smaller number of risk variants when compared to more heterogeneous populations.

As a result of population bottlenecks and low effective population size over time, which is characteristic for isolated populations, a higher degree of linkage disequilibrium is observed. Consequently, coverage achieved with a tagging SNP will be greater in isolated populations and more reliable than in heterogeneous populations (where there is a greater allelic diversity of associated alleles).

Finally, isolated populations have the advantage of reduced environmental variance, which should result in the increase proportion of genetic variance in the total phenotypic variation of a trait in question.

Two isolated populations were used in this thesis for the analysis of quantitative retinal vessel traits; ORCADES and CROATIA-Korcula studies. Orkney is an archipelago in

northern Scotland and The Orkney Complex Disease Study (ORCADES) is an on-going family-based study which aims to identify common and rare genetic variants influencing the risk factors for common complex disease important in Scotland. In total 889 genotyped individuals were available for the genetic analysis. The island of Korcula is located in the southern Croatian coastal region of Dalmatia. A total of 944 subjects were included in this study and most underwent complete eye examination (898 subjects available for genetic analysis). Family history data and pedigrees were available for ORCADES subjects, but not for Korcula. Both of these cohorts have been densely phenotyped and used for many GWAS and GWAMA (genome wide association meta-analysis) studies. Largely, the common variants found by published GWAS were replicated across both populations, without large difference in effect size.

1.6. Aims and objectives

1.6.1. Aims

The aims of this thesis were:

1) to generate retinal microvasculature quantitative measures derived from fundus photographies available from the ORCADES and CROATIA-Korcula studies, test the strength of their correlation with systemic cardiovascular measures available, and analyse the genetic component of their variations using appropriate methods

2) to perform genetic analysis on a binary retinal trait, rhegmatogenous retinal detachment, using a case control study

This second aim was added after my first year report as analysis of fundus photographs turned out to be a laborious and slow process and statistical analyses would not be possible before reaching sufficient number of reliable measures. This thus provided an opportunity to learn genetic association analysis, in a case control design. This was a continuation of an existing epidemiological and genetic study, and represented the first genome wide association study of this syndrome. The aim was to quantify the extent of genetic component and identify variants underlying disease pathology.

1.6.2. Objectives

Objectives of this thesis were as follows:

For the retinal detachment in a case control study:

a. Select markers for the second stage (i.e. follow up stage) of the genome-wide association study
b. Test for genetic association for these markers in independent case-control sets and in all of the discovery sets combined
c. Select best markers for replication (or follow-up if no significant results in discovery step revealed) and test for genetic association
d. Quantify how much of the polygenic component underlying RRD disease liability can be accounted for by the combined common variants tested globally, i.e. whether reaching individual level of significance or not

For the genetic analysis of quantitative retinal vasculature parameters; retinal vascular traits describing vessel calibres, vessel curvature and general complexity of the microvascular network (fractal dimensions).

a. Assess methodologies available-Measure retinal microvasculature traits and assess quality of measures
b. Basic statistics on traits collected- and correlations with available cardiovascular traits
c. Calculate heritabilities and genetic correlations for all retinal quantitative traits derived
d. Perform a genome-wide association studies on individual studies, meta-analyse the results, possibly perform replication in another population sample

Chapter 2. Genetic analysis of rhegmatogenous retinal detachment

2.1. Introduction

This chapter describes the genetic analysis of rhegmatogenous retinal detachment using a case-control design. Rhegmatogenous retinal detachment (RRD) is a major cause of severe vision loss with the average annual incidence of 10-20 cases in 100 000 individuals (Mitry et al. 2010) and with the prevalence of 0.8% (Mitry et al. 2011). RRD occurs following a retinal tear that separates neural retina from the underlying retinal pigment epithelium (RPE) due to the passing of the fluid from the vitreous cavity into a subretinal space. This leaves the neurosensory retina (NSR) deprived of nutrients and oxygen, potentially severely reducing visual function especially when the macula is affected.

Due to financial constrain, the analysis consisted of a two-staged discovery phase, a first stage using a genome-wide genotyping array, a second stage probing ~6000 variants, and a replication/follow-up phase which was limited to testing the seven best candidate SNPs from the discovery phase. A combined total of 2833 RRD cases and 7871 controls were analysed (Figure 2.1). The genome wide association study with close to 300 000 single nucleotide polymorphisms (SNPs) was performed prior to my starting PhD, using 867 affected Scottish individuals and 1953 Scottish controls, and is briefly described here. The main focus of this chapter is the follow up of the most significant SNPs from the original GWAS (second step of the discovery phase), the joint meta-analysis and the replication phase which I have carried on. Firstly the most significant GWAS results were tested in independent sets of cases and controls of Dutch (252 cases; 320 controls) and British origin (748 cases; 2592 controls), after which replication and follow–up was sought through the use of further samples from populations of English (846 cases; 2737 controls) and Croatian descent (120 cases; 269 controls). Finally the

overall meta-analysis including all stages (discovery and replication) was performed. The analysis was expanded with the investigation of the genetic risk score of the most significant results in order to improve the empirical evidence of the genetic basis of RRD. Furthermore the proportion of risk of RRD associated with the collective additive effects of the genotyped SNPs was also estimated from the discovery genome-wide scan. Potential biological interactions between the most significant results were sought out using pathway analysis.

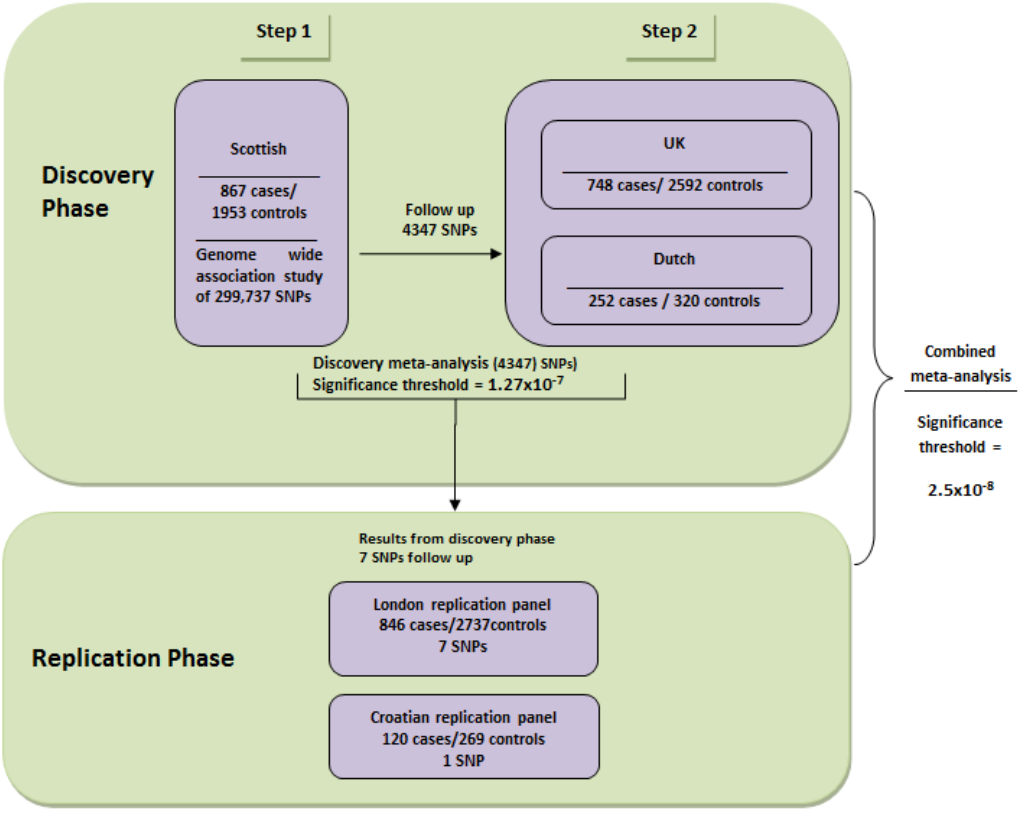


Figure 2.1. The study included a two-stage discovery phase followed by a replication phase where the seven most promising findings were brought forward for testing. GWAS was conducted in a Scottish study in the first stage of the discovery phase. The best ranked SNPs were selected for follow-up and complemented with candidate gene markers. 4347 SNPs were thus genotyped and analysed in the UK and Dutch studies in the second stage of the discovery phase, and results from both discovery stages were combined in a meta-analysis. An overall combined meta-analysis of discovery (2 stages) and replication phases was completed for the seven SNPs carried all along.

Work in this chapter was mainly done by me, unless stated otherwise.

2.2. Materials and methods

2.2.1. Participants

All participating studies received approval from their local ethics committees and followed the tenets of the Declaration of Helsinki.

2.2.1.1. Retinal detachment cases

All incident Scottish cases, aged 9-91 (mean=58.91), were recruited as a part of a population-based epidemiology study of rhegmatogenous retinal detachment recently carried out in Scotland (Mitry et al. 2009). 912 were included in the first step and 121 as a part of the second step of the discovery phase. All incident cases of rhegmatogenous retinal detachment reported in any of the six vireo-retinal surgical sites (Ayr, Glasgow, Edinburgh, Dundee, Aberdeen and Inverness) were invited to participate in the study. The diagnosis of RRD was based on a case definition of "a full thickness break in the neurosensory retina with a surrounding area of sub-retinal fluid extending greater than 2 disc diameters" (Zou et al. 2002, Li et al. 2003, Polkinghorne et al. 2004, Schepens, 1951). All other types of retinal detachment (exudative, tractional and combined) and re-detachment cases regardless of duration of attachment post-operatively were excluded. Also excluded were participants with a cataract surgery within 2 years of to the detachment diagnosis or with known syndromic features. 6.9 percent of Scottish RRD subjects evaluated reported an affected 1st degree relative (Mitry et al. 2011).

Cases from London, UK (N=470 in discovery phase; aged 9-92, mean=57.87, N=846 in replication phase (aged 14-90, mean=58.58), Cambridge, UK (N=217; aged 15-91,

mean=61.38), Nijmegen, Netherlands (N=256; aged12-95, mean=59.92) and Zagreb, Croatia (N=120 aged 20-83, mean=59.96) were collected using the same phenotypic inclusion criteria as for the Scottish samples. Positive family history was recorded in 11.3% of cases from the London recruitment, 7.1% of Croatian cases, and 5.6% of a subset of 180 Dutch cases had an affected first degree relative (Go et al. 2005). Family history information was not available for cases collected in Cambridge.

2.2.1.2. Controls

Control groups were readily available population-matched samples unselected with regards to RRD; i.e. unselected controls, a small fraction of which would have RRD (prevalence of RRD is 0.8%, Mitry et al. 2011). The Scottish Colorectal Cancer Study (SOCCS) was used as a control group for the Scottish samples in the first stage of the discovery phase. This is a prospective population-based study in Scotland (1999–2006), with colorectal cancer cases drawn from all over Scotland, matched to healthy controls by age, sex and area of residence (Tenesa et al. 2008). 1967 genotyped individuals were used in the analysis, aged 18-62 (mean=50.25).

Two sets of control groups from the Wellcome Trust Case Control Consortium (WTCCC) were also used in subsequent stages: 2595 samples from the 1958 UK birth cohort were used in stage 2 of the discovery phase and 2737 individuals from National Blood Service (aged 17-69, mean=43.5) in the replication phase. The Wellcome Trust Case Control Consortium (WTCCC) is a group of 50 research groups located across the UK (established in 2005) with the goal of studying association of common genetic variants to disease (Welcome Trust Case Control Consortium 2007). Access to the WTCCC data was obtained by application to the Wellcome Trust Case Control Consortium Data Access Committee. Control samples were selected based on the

genotyping platform used; samples genotyped on the Illumina arrays were chosen in order to gain the optimal overlap with the markers genotyped in RRD cases.

Controls for the Dutch samples were 326 control individuals for an independent schizophrenia study that were locally recruited, aged 36-74 (mean=50.22). Controls used in the final analysis represent only a subset of the original control sample (N=653), as the age distribution of the original controls sample were not matched to the cases'. In order to match the cases and controls regarding age, individuals older than 35 years were selected (326 out of the 653 controls).

Finally, Croatian participants to an ongoing colorectal cancer case and control study were used as a control dataset for the Croatian RRD cases (245 individuals; aged 24-92, mean= 66.1).

2.2.2. Genotyping and quality control

2.2.2.1. Genome-wide scan

912 Scottish RRD cases were genotyped using the Illumina CNV370v3-Quad array, and compared to population matched Scottish controls from the SOCCS study previously genotyped on the Illumina-300 and 240S HumanHap arrays. Raw data analysis was done using the Illumina genotype analysis software GenomeStudio with a GenCall scoring cut-off of 0.15 and solely on the new data itself (not using external references). A final clustering was done using only the 892 samples with an overall genotype call above 90%. Further quality control was performed separately for cases and controls using the check.marker function of the "GenABEL" R package (Aulchenko et al. 2007). This step

removed SNPs with a low call rate (<97%), those in high Hardy-Weinberg disequilibrium ($p<1x10^{-6}$), as well as samples identified as duplicates or highly related, mixed (high heterozygosity), individuals with a low call rate (<97%), sex discrepancies, and outliers based on whole-genome identity-by-state sharing. In total, 870 cases, 1,968 controls and 299,789 common SNPs (minor allele frequency >0.02) remained for analysis. When cases and controls were analysed jointly, three additional samples failed quality control. Further to these standard QC steps, SNP showing strong signals in association analysis were examined more closely. Several were removed due to poor clustering in either cases or controls or both after visualization of individual clusters (e.g. presence of 4 rather than 3 clusters) and two SNPs, rs4862110 and rs4957798, were removed due to discordance between SNP arrays. In the final analysis presented here, 299 737 SNPs, genotyped in both cases and controls and QCed, were used. This part of the quality control was done previously by Veronique Vitart.

A deeper relatedness check based on genomic kinship coefficients derived from the whole-genome identity-by-state sharing identified 15 pairs of individuals related to first-cousins or greater level (π-hat> 0.12) in PLINKv1.0 (Purcell et. al 2007), Φ >0.0625 in "GenABEL". Removing one individual of each related pair left 867 cases and 1,953 controls for the final analysis. Plots using principal components derived from the identity-by-state sharing measures showed good overlap of ancestry between the Scottish RRD cases and controls (Figure 2.2).

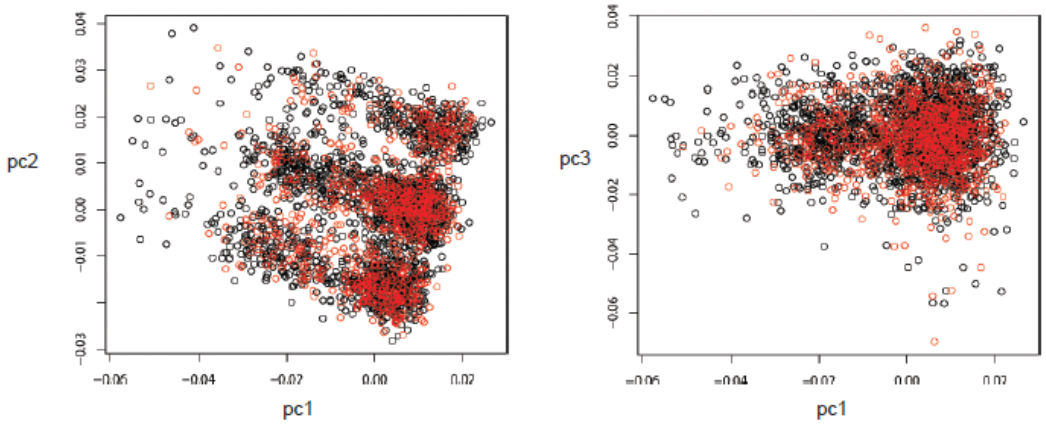


Figure2.2. Ancestry clustering for cases and controls used in the Discovery phase one. Clustering was achieved by multi-dimentional scaling of the matrix of pairwise genetic distances derived from the identity by state sharing statistics obtained using the ibs function in the GenABEL R package. N=858 Scottish RRD cases represented by red dots. N=1958 SOCCS controls represented by black dots (figure was done by V.V).

2.2.2.2. Discovery stage step 2 genotyping

2.2.2.2.1. Selection of the genotyping platform

My work on the project started here, with the selection of genotyping options for the discovery step two, carried out with Dutch and UK newly collected cases and using population-matched controls already genotyped using genome-wide arrays. As the funding was limited, only whole genome genotyping platform considered was the Illumina HumanCytoSNP-12 chip, however there were a few issues related to this chip; a risk of imperfect coverage compared to controls because it centers primarily on markers targeting all regions of known cytogenetic importance. I tested the proportion of

overlapping SNPs between HumanCytoSNP-12 chip and SNPs genotyped in Orkney (combination of Illuminas HumanHap 300v2 and 370CNV-Quad) and Korcula (370CNV-Quad chip). Overlap was low (83054 in Orkney and 86898 in Kocula). The overlap was greater when the SNPs present in the imputed genotypes of both ORCADES and Split individuals (a Croatian cohort available and genotyped on a genome array similar to the one used for the Scottish RRD cases) when compared to HumanCytoSNP-12 chip. The overlap was good (77% of SNPs on HumanCytoSNP-12 chip are represented in the imputed genotypes) indicating that if cases genotyped by Cyto-12 were used 77% of the SNPs would be present in imputed controls and used in the analysis. However this would require imputation of RRD cases which proved difficult (see paragraph 2.3.1.1.1.1. Analysis of the imputed data). In addition, when the overlap was done between the SNPs present on the chip and the most significant results from the GWAS carried out on genotyped data only 27% of the most significant SNPs were represented on the HumanCytoSNP-12 chip. Therefore this option was considered too risky, and custom SNP array of the most significant SNPs was considered to be a better solution for discovery step two analysis.

2.2.2.2.2. iSelect platform

5981 SNPs were selected for the second stage of the discovery phase to be assayed on an iSelect Illumina platform, 4706 selected from the genome-wide scan results complemented by 1275 from 18 candidates gene regions.

The top 4076 genome-wide scan SNPs were selected from merging the 1500 top results from 5 different primary GWAS analyses: all RRD Scottish cases versus all SOCCS Scottish controls, RRD cases with known high myopes (Spherical equivalent refraction $\leq$ -5) removed, SOCCS controls with cancer cases removed and gender

separate. Haplotype tagging SNPs in the candidate gene regions were selected with the Tagger program within the Haploview v4.0 software (Barret et al. 2005) using Centre d'Etude du Polymorphisme Humain from Utah (CEU) population of the HapMap project (International HapMap consortium 2005). A threshold for r^2 was set at 0.8. Region of 100kb upstream and downstream from each gene were used in the selection process. The candidate genes for this study were selected based on their function in some of the known inherited diseases associated with RRD and chosen by D.M and A.W..

Clustering was done using Illumina Genome Studio software and QC steps (call rate, HWE, sex discrepancies) were carried as described for the genome-scan analysis. Gencall score cutoff was at the default genome studio's of 0.15. Clustering of the allele-probe intensities was done using our own dataset, which contained samples with call rates above 80%. Cluster plots for each SNP were inspected manually in order to assess its genotyping quality. Poorly clustered SNPs were excluded from further analysis. Difference in genotype missingness rates between cases and control was also tested separately for each new study using the function –test-missing in PLINKv1.07 (Purcell et al. 2007). No SNPs displayed significant missingness difference ($p<0.00001$) between cases and controls. Summary statistics of the QC procedures are given in Table 2.1.

Study	SNPs before QC	Samples before QC	SNPs with MAF<2%	SNPs call rates < 97%	SNPs out of HWE (p<1e-06)	Samples with call rate <97%	Samples with sex discrepancies	Related individuals
Dutch cases	5168	256	179	199	1	2	2	0
Dutch controls	533620	653	5178	8047	0	3	0	22
UK cases (English and Scottish)	5168	764	161	190	7	6	3	0
English controls	4456	2595	41	35	109	1	0	0

Table 2.1. Quality control summary statistics for Dutch and UK populations genotyped in the discovery stage step two analysis. 4456 SNPs were extracted from WTCCC control data from the user-unfriendly format they were available in. Number of SNPs and individuals prior to the QC procedure is given, along with the subsequent number of exclude SNPs and individuals.

813 (13.6%) assays failed manufacture, comprising 183 of the candidate genes set and 630 top ranked GWAS SNPs (in agreement with failure rate to be expected from the Illumina Infinium technology). These markers failed to be synthesised on the final iSelect platform (the manufacturer guarantees the conversion rate of 80%). Each manufacturing step can potentially lead to loss of a SNP (synthesising the oligo probes, attaching probes to beads, loading beads on to the array etc.). As we were not informed about this prior to the ordering of our array, several potentially important SNPs from the first stage of the analysis were omitted in the replication stage. Therefore the list of potentially important omitted SNPs was created to be genotyped additionally by the Taqman assay. The decision of which SNPs to genotype by TaqMan was based on the results of the original GWAS. The top 200 most significant SNPs from six different analysis were cross checked with the missing SNPs. This left 92 missing iSelect SNPs that were potentially important. These SNPs were investigated for the representativity by proxy (LD $r^2 > 0.2$) of those which failed on the iSelect chip or QC. SNPs located in the candidate gene regions were not included in this extra genotyping step, due to the financial restrictions of the project. The final number of SNPs carried on for the second step of the discovery phase was 4347.

2.2.2.2.3. Single SNP typing

Taqman technology was used for the additional genotyping of eight SNPs missed by iSelect technology, with initial p-values $< 10^{-4}$ and not represented by proxy (LD $r^2 > 0.2$) or in plausible functional gene, *FRMD4A* and *GRM7*, and for the genotyping of seven SNPs to be followed up in the replication phase. The call rates of all of the SNPs were above 90%; with the highest being 96.6% (rs1485179) and the lowest 91.8%

(rs2663014). As a result all of the genotyped SNPs were included in the subsequent analysis.

2.2.2.3. Replication phase genotyping

Seven SNPs in total were genotyped in the replication phase using the Taqman technology. The six top associated SNPs which displayed consistent direction of effect across all cohorts and one additional SNP (rs913444) based on the pathway analysis results. Quality control checks were carried out as described above. The highest call rate in the London replication samples was 99.8% (rs8132771) and the lowest 93.5% (rs955943). Only one marker (rs12960119; call rate=93.6%) passed the 90% threshold in the Croatian samples.

2.2.3. Candidate gene SNPs selection

The majority of the candidate genes for this study were selected based on their function in some of the known inherited diseases associated with RRD. Vitreoretinal degenerations are a group of hereditary retinal dystrophies that often lead to retinal detachment. The most common one is Stickler's syndrome. Different forms of this disease are associated with mutations in *COL2A1* gene (type II collagen), *COL11A1* gene (type XI collagen) and *COL9A1* gene (type IX collagen). Other retinal dystrophies include Wagner syndrome (*CSPG* gene), Knobloch syndrome (*COL18A1* gene), familial exudative vitreoretinopathy (*FZD4* gene) and snowflake vitreoretinal degeneration

(*KCNJ13* gene). Another genetic condition associated with greater risk of RRD is Marfan syndrome thus the fibrilin 1 gene (*FBN1*) was also included in our study. Given the proposed mechanism in which changes in the vitreoretinal adhesion strength play an important role in RRD pathogenesis, molecules involved in the vitreoretinal adhesion represent plausible candidates (Mitry et al. 2010, Ghazi and Green 2002, Foos and Wheeler 1982). Laminin, fibronectin and collagens type 4 (*COL4A4*) and 6 (*COL6A1*) are present in the inner limiting membrane of the retina and are thought to play a role in the vitreoretinal adhesion (Ponsioen et al. 2008). In addition COL6A1 has been proposed to surround vitreous cortex and mediate adhesion between vitreous cortex and ILM. Opticin (*OPTC* which binds to collagen fibrils of the vitreous humour), hyaluron synthase (produced during tissue repair) and hyaluronan and proteoglycan-binding link protein were also included in the final candidate gene list. Number of selected and tested candidate gene SNPs are given in Table 2.2. (* selected that passed iSelect design, ** that passed QC and were present in controls).

Candidate gene	*Number of selected tagging SNPs in the gene (Number successfully tested)*	*Candidate gene*	*Number of selected tagging SNPs in the gene (Number successfully tested)*
COL11A1	34 (14)	*LAMA1*	161 (67)
COL18A1	58 (29)	*LAMB1*	54 (24)
COL2A1	54 (21)	*LAMC1*	41 (13)
COL4A4	62 (42)	*OPTC*	33 (7)
COL6A1	39 (16)	*VERSICAN*	116 (42)
COL9A1	89 (45)	*FBN1*	56 (20)
HAS1	95 (50)	*FN1*	46 (12)
KCNJ13	16 (4)	*FZD4*	62 (24)
HAPLN	28 (13)	*COL11A2*	48 (15)

Table 2.2.Number of selected* and tested** tag SNPs for candidate genes selected for follow-up in the step 2 of the Discovery Phase to complement the GWAS SNPs taken forward.

2.2.4. Imputation of the genotyped SNPs

2.2.4.1. RRD data preparation

870 of the 912 genotyped RRD individuals on the Illumina CNV370v3-Quad array were imputed against CEU reference panel using the MACH software. After the QC with the following parameters (MAF<0.01, HWE<10-6, call rate >0.98) 870 individuals and 324256 markers remained. SNPs with bad genotype clustering (identified in Illumina's Genome Studio during genotype call check of the candidates following the association analysis using genotyped data) were removed (only 3 SNPs from the whole list were actually present in the QCed data) which left 324253 SNPs. Additionally all the SNPs with AT, TA, GC, CG genotypes were removed (due to possible flips between Illlumina strand designation and MACH references haplotypes) which left 323180 SNPs for imputation. The population-matched controls used for this analysis are from the Scottish Colorectal Cancer Study (SOCCS). Imputed genotypes for the SOCCS individuals were available for this analysis and had used the same reference. Identical individuals between case and control groups and close relatives within or between case and control groups were removed as already described (2.2.2.1.).

2.2.4.2. RRD data imputation

Imputation was done for over 2 million SNPs (on 22 autosomal chromosomes) with reference to the CEU population (HapMap2 build 36 release 22). Imputation is run in two stages. The first step is to compare our samples to the phased haplotype data in the reference panel (HapMap CEU) to generate recombination and error maps. Error map includes estimates of the *"error"* rate for each marker (it captures both genotyping error,

discrepancies between our platform and the reference panel, and recurrent mutation), while recombination map includes *"crossover"* rates for each interval (a parameter that describes breakpoints in haplotype spans shared between our samples and the reference panel). The second step uses the parameters estimated in the first step to impute all SNPs represented in the reference panel (from phased HapMap data) in our sample of individuals. Approximately 2.5 million SNPs can be imputed to sufficient quality with this method. Imputation data for the Scottish controls was performed previously. Imputed data from cases and controls were merged together and used for association tests.

2.2.5. Statistical analysis

2.2.5.1. Association testing

The genome-wide scan was performed using a regression model with age and sex as covariates and fitting an additive SNP effect, using the egscore function of the "GenABEL" R package (Aulechenko et al. 2007) which takes into account possible population stratification (Price at al. 2006). QQplots of observed versus expected ordered test statistics were drawn using the "snpMatrix" R package (Clayton 2007). The grey area represents the area between the 2.5% and 97.5% probability bounds of non-departure to the null hypothesis. Five analyses were performed, all RRD Scottish cases versus all SOCCS Scottish controls, RRD cases with known high myopes (Spherical equivalent refraction $\leq$ -5) removed , SOCCS controls with cancer cases removed and gender separate analysis. This part of the analysis was done by V.V. The top ranking

associated SNP from each analysis (p-value< 10^{-3}) were listed to be taken forward in the second stage of the discovery.

In the second stage of the discovery phase, association between RRD with SNP genotypes was tested using logistic regression including sex and age as covariates and using the PLINKv1.07 software (Purcell et al. 2007). Both additive and dominant models were tested. For additive model depending on the number of copies of the minor allele genotypes were coded 0 (minor allele homozygote), 1 (heterozygote) or 2 (major allele homozygote). The direction of the regression coefficient (β) represents the addition of one extra minor allele. Odds ratios (OR) greater than 1 indicate that minor allele increases risk of the disease relative to major allele. For dominant model genotypes are coded 0 (major allele homozygote) and 1 (heterozygote and minor allele homozygote; combines these two genotypes into 1 category), while the regression coefficient compares the effect of the heterozygote and minor allele homozygote genotype to the major allele homozygote one. Since the genetic model coding is done relative to the minor allele and the recessive model tests minor allele homozygote against heterozygote and major allele homozygote, the recessive model gives many low counts for low MAF which often leads to inflated test results; or no counts and analysis cannot be performed for a substantial number of markers. For these reasons it was decided not to pursue testing of the recessive model.

Association analyses were done for each cohort separately and the results from all cohorts (with the Scottish discovery step 1 rerun using the PLINKv1.07 implementation of the logistic regression adjusting for age sex and 3 ancestry principal components) combined in an inverse variance weighted fixed effect meta-analysis using the statistical package METAL (Willer et al. 2010). The fixed effect model assumes that all studies are estimating the same effect size, and the weights (for all these studies) are assigned based on the amount of information captured by each study (a study with greater number of samples is given a bigger weight compared to a smaller study). METAL assigns a

reference allele to each SNP (all studies are aligned to the same reference allele) and its z -statistic is calculated. Based on this an overall z-statistic and p-value are calculated for each SNP from a weighted sum of the individual statistics. Standard errors were used to calculate the effect size weights (inverse variance weights). Forrest plots and heterogeneity measures for the most significant SNPs were obtained using the "meta" R package (Swarzer, 2012).

β_i- effect size estimate for study i
se_i- standard error for study i

$w_i = 1/se^2_i$ (weight assigned to each study)
$se = \sqrt{1}/\sum_i wi$ (the standard error of the combined effect)

weighted effect size $\beta = \sum_i \beta i wi / \sum_i wi$
z-score $Z = \beta/se$
Overall P-value ; $P = 2\Phi(|-Z|)$
(Φ is the standard normal cumulative distribution function)

Given the low sample size of our replication cohorts and that inclusion of covariates in case-control association studies has been shown to reduce their power in this scenario and with low disease prevalence (Pirinen et al. 2012) we carried out analysis in the replication phase without the inclusion of covariates. Results from both discovery stages and the replication phase were also combined in a meta-analysis (unadjusted for covariates).

2.2.5.2. P value threshold

In order to correct for multiple testing in the initial genome wide scan, we used a multiple testing correction method that takes into account the correlation structure among the GWAS genotyped SNPs due to linkage disequilibrium using "SimpleM" R library (Gao et al. 2010). It uses composite linkage disequilibrium correlation, which is

calculated directly from SNP genotypes to create the correlation matrix of SNPs. This correlation matrix is then used to compute the eigenvalues for the principal components analysis which then generates an effective number of independent (non-correlated) SNPs. Imputation of the missing genotypes was carried out using k Nearest Neighbors using the "scrime" R package (Schwender 2012), since SimpleM does not allow for missing values. The calculated effective number of independent SNPs, n=197628, was then used to correct for multiple testing using the standard Bonferroni procedure. The discovery stage significance threshold was additionally corrected for the testing of two genetic models (additive and dominant). The genome wide significance threshold was therefore set at 1.27×10^{-7} (=0.05/(197628x2)). Despite following up only a subset of SNPs used in the first step of the discovery phase into the second step, the same genome wide significance threshold was used for the meta-analysis of the two steps to be conservative. The overall meta-analysis (combining discovery stage and replication stage data) used the conventional 5×10^{-8} significance threshold, additionally corrected for the use of two genetic models, hence set at 2.5×10^{-8}.

2.2.5.3. Permutation procedures

We calculated empirical significance levels for the association analysis in all of the cohorts, and used these empirical significance levels for further meta-analysis. Adaptive permutation testing was carried out in PLINKv1.07 (Purcell et al. 2007). Firstly labels between cases and controls are swapped and a new dataset (which is then tested for association) that is sampled under the null hypothesis is created. This dataset is then permuted using the adaptive permutation approach which permutes non-significant SNPs more quickly (smaller number of performed permutations) than the possibly significant ones.

2.2.6. Genetic risk score

Multilocus genetic risk scores were calculated for each individual and included 5 risk variants (most significant results from meta-analysis of discovery steps 1 and 2 that had significant associations in all of the cohorts; $p<0.05$; rs12960119, rs267738, rs2045084, rs8132771, rs955943). Simple allele counting was used to calculate the genetic risk score. The simple count uses the total sum of the risk allele present in an individual. Genetic risk scores were calculated using PLINK's v1.07 "score" option. Individuals with one or more missing genotypes were not included in the analysis. Odds ratios were calculated relative to the median number of risk alleles among controls in SPSS (subjects with 5 or more risk alleles were combined due to the low number of individuals with high risk allele numbers). The risk score therefore runs from 0-10 in theory.

2.2.7. Disease liability's genetic variance explained

GCTA (Genome Wide Complex Trait Analysis; Yang et al. 2010) software was used to estimate the proportion of variation in disease liability capture by the genotyped SNPs in the study population of discovery stage 1 (Scottish RRD cases and SOCCS controls). GCTA enables this estimation in the samples of unrelated individuals (Lee et al. 2011). Transformation of a disease trait (0-1) to an underlying scale of liability is performed using a probit transformation. Heritability is estimated by fitting a linear mixed model and using restricted maximum likelihood and with a random SNP effect on the transformed trait. This estimation is then corrected for the incomplete linkage disequilibrium (LD) between the tag SNP and the causal variant. The program uses LD between genotyped SNPs as a guide to LD between causal variants and genotyped

SNPs. In case-control studies individuals are selected from the tails of the distribution of liability, thus the proportion of cases is usually larger than the prevalence in the general populations. This ascertainment can be corrected for when calculating the heritability estimates in case–controls study with GCTA. Very stringent QC procedures (Lee et al. 2011) were implemented using plink on the genotyped data to avoid experimental biases causing false case-control differences in relatedness. SNPs with MAF <0.03 and missing rates >0.05, and individuals with missing rates >0.01 were excluded from the analysis. Additionally, very stringent threshold for Hardy-Weinberg equilibrium and differences between the missing genotype rates between cases and control were applied (SNPs with $p<0.05$ were excluded). Closely related individuals were removed to make sure only distant relationships are taken into account (kinship coefficient threshold was set at 0.05 as implemented within the GCTA software). As further check, the control cohort (SOCCS without colorectal cancer cases) was split into two groups and the procedure repeated in order to confirm that the detected explained variance was not a population structure artefact.

2.2.8. Pathway analysis

The top 6 SNPs from the discovery stage 1 and 2 meta-analysis were selected for a first pathway analysis. Identification of all possible genes underlying these SNPs was done using "seed" output from GRAIL (Gene Relationships Among Implicated Loci). For each input SNP GRAIL finds SNPs in LD ($r^2>0.5$, CEU HapMap) with it in the 3′ and 5′ direction, and locates the nearest recombination hotspot to the furthest SNPs in both directions. The interval between these two hotspots is then chosen to locate the genes associated with the SNP of interest (because our top SNPs could be tagging a stronger SNP signal from another variant in the region, particularly because we genotyped only a subset of SNPs from the initial GWAS in the majority of cases). When there are no genes within this region, the interval is extended for 250 kb in each direction. All genes

located in this interval were considered in the subsequent pathway analysis using IPA software (Ingenuity Pathway Analysis; Ingenuity ®Systems, www.ingenuity.com). IPA uses an internal manually maintained database to identify networks that maximize specific connectivity of the entered genes. It searches through networks (in its database) in which the genes from the entered list are included and searches how they associated with other genes (looking for experimentally demonstrated relationships among genes that are associated with distinct known pathways). It then creates a novel network including some or/all genes of interest together with other interacting genes which maximizes its connectivity. When running the analysis in IPA, endogenous chemicals were excluded from the search and only the direct interactions were taken into account. After the analysis of the 6 most significant hits from the meta-analysis, additional analysis using all SNPs with association p-value $<10^{-4}$ (22 SNPs in total, Table 2.8) was performed in order to test whether possible direct biological interactions exist between our top SNPs. Parameters were set identically to the first analysis. Based on this extended pathway analysis we identified three direct interactions and further literature search confirmed their biological grounds for one identified interaction (*TIAM1*-*NTRK2*). SNP (rs913444) associated with *NTRK2* gene (using the GRAIL criteria) was therefore included in the stage 3 of our analysis (SNP associated with *TIAM1* was already among top 6 SNPs from the meta-analysis).

MAGENTA (Meta-Analysis Gene-set Enrichment of variaNT Associations) software was also used to explore the enrichment of our genetic associations in predefined biological pathways. Since MAGENTA only takes genome-wide data, only the results from the first step of the discovery stage data was used in this analysis. The input data comprises SNP association p-values and the location of the variants. Firstly, SNPs are mapped onto genes after which each gene is assigned a gene association score that is a function of its regional SNP association *p*-values (is assigned a set of SNPs that lie within 110 kb upstream and 40 kb downstream of the gene's most extreme transcript boundaries is assigned to each gene in the genome). Then the confounding effects on

gene association scores are identified and corrected for (linear regression-based approach that adjusts for the effects of multiple confounders on the gene score is used; number of SNPs per kb, number of independent SNPs per kb, recombination spots per kb, linkage disequilibrium units per kb, genetic distance across the gene). Each mapped gene in the genome is ranked by this corrected gene score. Finally, a Gene Set Enrichment Analysis (GSEA)-like statistical test is used in order to check whether any gene sets (predefined biologically relevant gene sets from several public databases) are enriched for significant gene association scores (compared to randomly selected gene sets). The observed number of gene scores in a given pathway, above a chosen significance threshold (95th or 75th percentiles of all gene scores) is calculated and compared with 1 000 000 randomly permuted pathways of identical size giving the empirical GSEA *P*-value for each pathway. 2572 pathways from Gene Ontology, PANTHER, KEGG and Ingenuity were tested. A pathway is considered significant if reaching false discovery rate of <0.05 in either analysis.

2.3. Results

2.3.1. Association analysis –discovery stage step 1 –analysis of genotyped data

A genome-wide association study (GWAS) was carried out with 867 RRD cases and 1953 controls from Scotland, using 299 737 genotyped SNPs, after applying stringent quality controls. GWAS results show a very slight excess of low p-values compared to the null hypothesis of no association (Figure 2.3; A,B). One SNP (rs10510663) reached genome wide significance after correcting for multiple testing (significance threshold set at 1.27x10-7 as described in the methods; $p=1.88x10^{-9}$). This marker is located on chromosome 3p22.3, upstream of *ARPP21* gene (analysis done previously by V.V.).

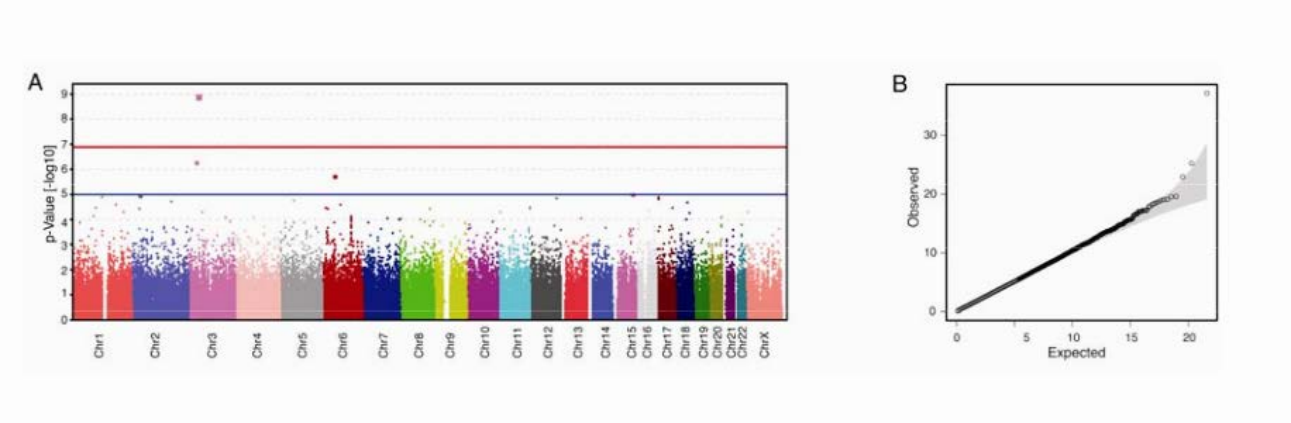


Figure 2.3. A. Manhattan plot (A) and Quantile-quantile (B) plot for the discovery phaseI genome wide scan. Statistics from a logistic regression adjusting for age sex and co-ancestry performed on 867 Scottish RRD cases and 1953 SOCCS participants using 299,737 genotyped SNPs passing quality control checks are plotted (-log10 (P-value) on the x axis, chromosomal location on the y axis for the Manhattan plot.- Observed and expected chi-square values were used for the QQ –plot).

2.3.2. Association analysis –discovery stage step 1 - analysis of the imputed data

The ProbABEL software was used to carry out the association analysis. The analysis resulted in numerous genome-wide significant hits located all around the genome when standard post-analysis QC criteria were applied (markers with imputation quality score Rsq less than 0.3 and low minor allele frequency variants, MAF of <0.02, removed) too significant and too ubiquitous to be true positives. Therefore various more stringent QC criteria were applied, but the distribution of the significant SNPs did not change dramatically. Finally a very stringent QC criteria was applied, removing SNPs with Rsq of <0.98 and MAF of <0.05 in both cases and controls, leaving 38 SNPs with p-value $\leq 10^{-5}$ in total (Table 2.3). Excess of significant result is observed on the QQ plot (λ=1.09) (Figure 2.4).

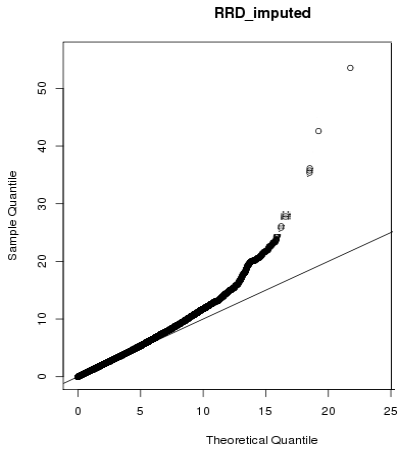


Figure 2.4. Observed and expected chi-square values were used for the QQ –plot. Excess of significant test statistic is observed after SNPs with Rsq<0.98 and MAF<0.05 are removed, for cases and controls separately.

name	CHR	position	effective allele freq RRD	effective allele freq SOCCS	CEU allele freq (LBC freq)	Rsq_soccs	Rsq_RRD	OR	Lower 95% CI	Upper 95% CI	p
rs7816179	8	26832051	0.86	0.93	0.94 (0.91)	0.9998	0.9955	0.46	0.37	0.57	4.78E-13
rs16846471	2	1.42E+08	0.04	0.09	0.036 (0.033)	0.9926	0.9839	0.38	0.28	0.51	1.77E-10
rs7816187	8	26832063	0.83	0.90	0.94 (0.89)	0.9993	0.9961	0.57	0.47	0.68	1.7E-09
rs12677733	8	26833426	0.17	0.10	0.08 (0.11)	0.9992	0.9936	1.76	1.46	2.12	1.99E-09
rs10124811	9	7609582	0.69	0.75	0.75 (0.75)	0.9885	0.9843	0.69	0.59	0.79	2.32E-07
rs4670165	2	36795770	0.64	0.57	0.55 (0.54)	0.9849	0.9971	1.41	1.24	1.61	2.9E-07
rs978591	4	10227767	0.84	0.77	0.88 (0.76)	0.9969	0.9904	1.53	1.30	1.79	3.59E-07
rs2531196	4	10228045	0.16	0.23	0.13 (0.24)	0.9964	0.9905	0.66	0.56	0.77	3.67E-07
rs7672373	**4**	**60856165**	**0.04**	**0.09**	**0.05 (0.09)**	**0.9990**	**0.9862**	**0.48**	**0.36**	**0.64**	**3.84E-07**
rs2531197	4	10228287	0.84	0.77	0.85 (0.76)	0.9952	0.9911	1.52	1.30	1.79	3.86E-07
rs13140301	4	60850026	0.04	0.09	0.0 (0.09)	0.9989	0.9867	0.48	0.36	0.64	3.87E-07
rs11727316	4	60856708	0.96	0.91	0.97 (0.91)	0.9989	0.9860	2.07	1.56	2.74	3.88E-07
rs13113866	**4**	**60857530**	**0.96**	**0.91**	**0.96 (0.91)**	**0.9988**	**0.9859**	**2.07**	**1.56**	**2.74**	**3.95E-07**
rs2531200	4	10230190	0.84	0.77	0.9 (0.76)	0.9944	0.9909	1.52	1.29	1.79	3.97E-07
rs2531192	4	10226798	0.84	0.77	0.88 (0.76)	0.9967	0.9896	1.52	1.29	1.79	4.07E-07
rs2720359	4	10232389	0.16	0.23	0.125 (0.24)	0.9929	0.9904	0.66	0.56	0.77	4.08E-07
rs2720372	4	10226723	0.16	0.23	0.125 (0.24)	0.9969	0.9895	0.66	0.56	0.77	4.2E-07
rs2720361	4	10234723	0.16	0.23	0.123 (0.24)	0.9847	0.9849	0.66	0.56	0.78	6.73E-07
rs2720362	4	10234995	0.16	0.23	0.106 (0.24)	0.9846	0.9849	0.66	0.56	0.78	6.76E-07
rs17383444	4	10230614	0.84	0.77	0.88 (0.77)	0.9894	0.9917	1.51	1.28	1.78	8.01E-07
rs2720365	4	10237359	0.16	0.23	0.11 (0.24)	0.9821	0.9916	0.67	0.57	0.78	9.81E-07
rs2531205	4	10239857	0.16	0.23	0.10 (0.24)	0.9822	0.9917	0.67	0.57	0.78	9.92E-07
rs7615092	3	1.7E+08	0.33	0.27	0.275 (0.29)	0.9999	0.9873	1.39	1.21	1.59	1.97E-06
rs12638698	3	1.7E+08	0.33	0.27	0.275 (0.29)	0.9999	0.9870	1.39	1.21	1.59	1.97E-06
rs1362521	3	1.7E+08	0.67	0.73	0.725 (0.71)	1.0000	0.9856	0.72	0.63	0.82	1.97E-06
rs13061789	3	1.7E+08	0.67	0.73	0.725 (0.71)	0.9996	0.9881	0.72	0.63	0.83	2.01E-06
rs12633440	3	1.7E+08	0.67	0.73	0.725(0.71)	0.9983	0.9847	0.72	0.63	0.82	2.03E-06
rs13094430	3	1.7E+08	0.67	0.73	0.725 (0.71)	0.9974	0.9843	0.72	0.63	0.83	2.07E-06
rs7602707	2	1.42E+08	0.15	0.22	0.13 (0.15)	0.9974	0.9939	0.67	0.56	0.79	3.24E-06
rs11789902	**9**	**7608288**	**0.28**	**0.23**	**0.23 (0.23)**	**0.9946**	**0.9841**	**1.41**	**1.22**	**1.63**	**4.11E-06**
rs10815648	**9**	**7607978**	**0.28**	**0.23**	**0.24 (0.23)**	**0.9987**	**0.9864**	**1.41**	**1.22**	**1.63**	**4.32E-06**
rs9518902	13	1.02E+08	0.09	0.13	0.067 (0.13)	0.9975	0.9877	0.61	0.50	0.76	6.26E-06
rs9518901	13	1.02E+08	0.09	0.13	0.075 (0.13)	0.9985	0.9885	0.61	0.50	0.76	6.28E-06
rs10520735	**15**	**91874783**	**0.82**	**0.75**	**0.77 (0.77)**	**0.9945**	**0.9857**	**1.44**	**1.23**	**1.69**	**6.68E-06**
rs283552	**6**	**51071956**	**0.48**	**0.46**	**0.43 (0. 49)**	**0.9985**	**0.9984**	**0.75**	**0.66**	**0.85**	**7.12E-06**
rs7209554	17	5487877	0.41	0.34	0.40 (0.38)	0.9992	0.9873	1.35	1.18	1.54	9.09E-06
rs6458751	**6**	**51079672**	**0.48**	**0.46**	**0.45 (0.49)**	**0.9950**	**0.9975**	**0.75**	**0.66**	**0.85**	**9.64E-06**
rs7224241	17	5487689	0.41	0.34	0.42 (0.38)	0.9979	0.9874	1.35	1.18	1.53	9.68E-06

Table 2.3 Most significant GWAS results of the imputed data after very stringent QC; for each SNP MAF for case and controls is given, imputed MAF frequencies for RRD cases and controls. P-value, OR with 95% confidence intervals are also reported. Imputation quality score measurement Rsq is also given for each SNP. SNPs having MAF difference when compared to CEU population in both cases and controls are highlighted in bold.

Deviation from CEU's allele frequencies in both cases (RRD) and controls (SOCCS) were calculated for top 38 hits. 37% SNPs in cases and 53% in controls had MAF difference between CEU and RRD greater than 5 %. These inconsistencies suggest that association test is detecting MAF difference between cases and controls that are most

likely the result of imputation errors, considering such big discrepancies in allele frequencies with the reference population for the imputation. This scenario is more likely when allele frequency discrepancies are observed for the controls. For RRD cases it is more difficult to make this assumption, since their allele frequencies may genuinely differ from CEU if they present a true positive result. For the RRD samples, these large discrepancies flagged checking of the genotype clusters of the original data (regions on chromosomes 3, 8 and 9). However genotyping clusters did not reveal issues in the genotyping clustering. In addition, frequencies of the alleles in SOCCS were compared to another Scottish sample (Lothian Birth Cohort) to confirm the possibility of imputation errors. SNPs showing large allele frequency differences between SOCCS and CEU also presented a considerable difference in frequencies when LBC data was compared. The exception was the region on chromosome 4p16.1 within the *CLNK* gene (cytokine-dependent hematopoietic cell linker), with in role in the regulation of immunoreceptor signalling.

In addition when MAF difference between directly genotyped and imputed SNPs were checked, only two markers (rs4670165 and rs10520735) were represented in the set with the QC criteria of Rsq>0.98. Therefore these discrepancies were checked for markers having Rsq values greater than 0.9 in both RRD cases and SOCCS controls (in total 11 SNP frequencies were compared). Differences greater than 2.5% were detected among 4 out of those 11 SNPs (36%). The proportion of SNPs with MAF difference between the genotyped and imputed sets was greater after more relaxed QC criteria were applied. This observation in addition to the observed SOCCS MAF difference (when compared to CEU's and LBC's) indicates imputation errors that partly contribute to the abundance of highly significant results when SNPs with lower quality scores were included. Marker with the highest p-value when cut-off of R_{sq}<0.9 was applied (rs13095640, $p=1.76x10^{-12}$) presented a difference of 5.4% between the imputed and directly genotyped MAF. Unfortunately, marker on chr8 showing highest p-value (rs7816179; R_{sq}<0.98) has not been genotyped and it was no possible to check for these discrepancy To see whether

there is a greater difference in MAF between directly genotyped and imputed SNPs in the RRD cases sample than in another imputed dataset (ORCADES) where GWAS have been routinely carried without report of excess type 1 error rate, the difference in MAF between genotyped and imputed SNPs was inspected for all relevant markers on one chromosome (chromosome 22). 1.13% of SNPs had MAF difference greater than 2.5 % in the RRD individuals, which is similar to the number of SNPs on Chr22 with the same difference in MAF in Orkney individuals (1.5% of SNPs have difference in MAF >2.5% between directly genotyped and imputed SNPs; tested on the similar number of SNPs, 4940 for Orkney and 5303 for RRD cases). Suggesting the quality of imputation of the RRD dataset was in line with the standards for imputation data, and that the excess of positive results problem probably lies in the comparison of the imputed data from two datasets genotyped on different genotyping platforms (and imputed separately). Many of the checks described here were done after the time of selecting markers for the iSelect platform follow-up at which time it was decided too risky to follow up results from the imputation analysis. The issue of extra care when convenience controls are used in case-controls analysis of imputed data has also been documented since I started this analysis. (Sinnot and Kraft, 2011). And although in a standard setting (case and controls genotype together), imputation methodology has been proven to be very accurate and reliable, working with cases and controls genotyped independently requires extra care and stringent QC.

2.3.3. Association analysis –discovery stage Step 2

The most significant SNPs associated with RRD by GWAS of genotyped data, plus some selected candidate gene markers were then genotyped on a custom-made Illumina iSelect array in an additional 1,000 RRD cases (Discovery stage 2; 252 Dutch and 748 UK individuals). The meta-analysis of the SNP association data from these two steps

constituted the discovery phase and involved analyses of three distinct case-control cohorts (of Scottish, English and Dutch ancestry). A stringent genome wide significance threshold based on the number of SNPs tested in the initial whole genome scan ($1.27x10^{-7}$; see Methods section) was applied in this meta-analysis. Index SNPs with association p-value less than 10^{-4} are displayed in Table 2.4.A for an additive model of allelic effect and in Table 2.4.B for a dominant model. The marker reaching genome wide significance in the GWAS of Scottish cases (rs10510663) did not remain significant in the meta-analysis (non-significant associations in both UK (p=0.86 additive; p=0.92 dominant model) and Dutch cohorts (p=0.24 additive; p=0.27 dominant model) and its association p-value after meta-analysis of discovery studies increased to greater than 10^{-4}.

A. Additive model

SNP	Minor allele	CHR	Position (build36)	CEU minor allele freq	OR(95%CI)	Association p-value	Direction	Heterogeneity p-value
rs12960119	G	18	21868701	0.075	1.46(1.26-1.67)	1.58E-07	+++	0.67
rs267738	C	1	149207249	0.275	0.79(0.71-0.87)	6.7E-06	---	0.65
rs955943	A	4	16123506	0.067	1.54(1.27-1.87)	9.9E-06	+++	0.67
rs7097067	A	10	71676508	0.075	1.65(1.32-2.07)	1.28E-05	-++	0.99
rs1074463	A	5	22671126	0.092	1.31(1.16-1.48)	1.28E-05	+++	0.73
rs2045084	G	8	144764982	0.425	1.21(1.11-1.31)	1.54E-05	+++	0.25
rs8132771	A	21	31921294	0.075	1.43(1.21-1.69)	1.96E-05	+++	0.21
rs11259960	A	15	81379707	0.200	1.35(1.17-1.54)	2.36E-05	+-+	0.41
rs2368106	A	2	180815911	0.125	1.35(1.17-1.54)	2.58E-05	-++	0.76
rs7234959	A	18	21816517	0.083	1.34(1.17-1.53)	2.68E-05	+++	0.77
rs6070015	A	20	55192062	0.067	1.67(1.31-2.13)	2.91E-05	+++	0.91
rs913444	A	9	87051470	0.250	1.22(1.11-1.34)	3.00E-05	+++	0.94
rs4893905	A	2	180867194	0.100	1.33(1.16-1.52)	3.32E-05	+++	0.59
rs1477441	A	5	22569230	0.124	1.31(1.15-1.49)	4.03E-05	+++	0.87
rs12193473	G	6	121301788	0.375	1.20(1.10-1.31)	4.25E-05	+++	0.32
rs218843	A	6	121388196	0.375	1.20(1.10-1.31)	4.40E-05	+++	0.30
rs4715056	G	6	47977103	0.275	0.81(0.73-0.90)	4.94E-05	+--	0.59
rs2817896	G	1	22988636	0.242	1.24(1.12-1.37)	5.35E-05	-++	0.82
rs10515162	C	5	73090012	0.050	0.72(0.61-0.85)	6.48E-05	---	0.95
rs11181447	A	12	41026298	0.150	1.28(1.13-1.45)	6.68E-05	+++	0.85
rs12202993	A	6	121330001	0.375	1.20(1.10-1.31)	7.48E-05	+++	0.30
rs564351	A	6	47927523	0.208	0.79(0.70-0.89)	7.49E-05	+--	0.91
rs6035211	C	20	19008513	0.117	1.30(1.14-1.47)	8.66E-05	+++	0.99

B dominant model

SNP	Minor allele	CEU allele frequency	OR (95%CI)	p-value	Direction	HetPVal
rs12960119	G	0.075	1.47(1.27-1.71)	4.45E-07	+++	0.531
rs267738	C	0.275	0.74(0.65-0.84)	1.55E-06	---	0.4144
rs955943	A	0.067	1.54(1.26-1.88)	2.40E-05	+++	0.5753
rs7097067	A	0.075	1.55(1.26-1.91)	3.18E-05	-++	0.1682
rs1074463	A	0.092	1.32(1.16-1.50)	3.24E-05	+++	0.8182
rs8132771	A	0.075	1.43(1.20-1.69)	4.04E-05	+++	0.4098
rs6035211	C	0.117	1.32(1.16-1.51)	4.45E-05	-++	0.4106
rs9956700	A	0.267	0.71(0.60-0.84)	4.94E-05	+?-	0.03752
rs7234959	A	0.083	1.36(1.17-1.58)	5.09E-05	+++	0.7646
rs267733	G	0.133	0.76(0.67-0.87)	5.29E-05	---	0.9093
rs2368106	A	0.125	1.33(1.16-1.53)	5.86E-05	+++	0.382
rs11637235	G	0.225	1.40(1.19-1.65)	6.73E-05	++?	0.01017
rs11181447	A	0.150	1.31(1.14-1.49)	8.73E-05	+++	0.6992

Table 2.4.Highest ranking RRD association signals in the meta-analysis of the 3 discovery studies. A additive model; B dominant model .Odds ratios with 95% confidence intervals, association p-values, direction of the effect for respectively the populations Netherlands, UK and Scotland and heterogeneity metrics for all markers showing suggestive p-values ($< 10^{-4}$).

The most significantly associated marker after this stage was rs12960119, with an association p-value of $1.58x10^{-7}$(additive model), just below the genome wide significance threshold. This marker is located on chromosome 18q11.2 within an intron of the *SS18* gene (OR=1.46; 1.26-1.67 95%CI). The magnitude and direction of the signals were consistent across all three populations, translating into a low heterogeneity score across studies (I^2=0; p=0.67). Cases showed an excess of the rs12960119 minor allele G compared to controls in all three cohorts (Table 2.5.).

Population	Allele frequency (cases/controls)	OR(95%CI) additive model	OR(95%CI) dominant model	p-value (additive/dominant)
Scottish	0.115/0.086	1.42 (1.16-1.75)	1.40(1.30 -1.73)	0.0009/0.003
UK	0.104/0.077	1.41(1.13-1.76)	1.46(1.17-1.82)	0.002/0.0008
Dutch	0.117/0.075	1.76(1.13-2.76)	1.90(1.17-3.1)	0.013/0.009

Table 3.4. rs12960119 minor allele (G) effect across the three discovery cohorts .rs12960119 allele frequencies, odds ratios and association p-values for the minor G allele by discovery population.

Another marker (rs7234959), located upstream of the *SS18* gene, displayed a suggestive p-value in the meta-analysis ($p=2.68x10^{-5}$), strengthening the validity of an association signal in this region. A regional plot if the SNPs in the region of 500kb of the rs12960119 is presented in Figure 2.5.

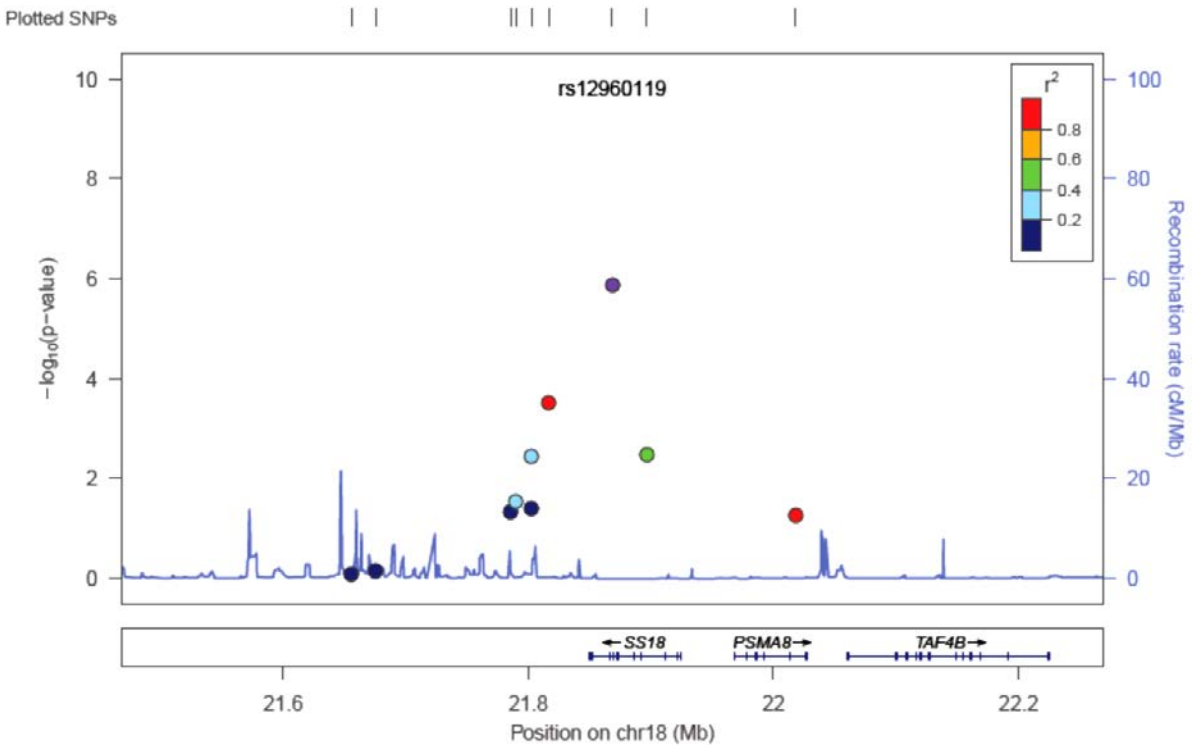


Figure 2.5. Regional SNP association plots for SNP rs12960119 by the combined population meta-analysis of association signals in discover stage populations (Scottish, UK and Dutch). A region of 500 kb around the association signals is displayed. The degree of LD between the top genotyped SNP and other 9 SNP tested in the meta-analysis is indicated by colours (red indicating high LD, purple indicated no LD). Recombination rate is displayed by a blue line with scale on the right-hand axis. Genes are displayed in the bottom part of the figure, with the arrow showing the direction of transcription.

rs12960119 is in high LD with an evolutionarily highly conserved SNP, usually suggesting functional importance (rs765529; Figure 2.6.). This SNP was not successfully typed in the step two of the discovery stage and therefore not included in the previous regional plot. This was observed once the results from the genome scan were plotted in a regional association plot around *SS18* gene.

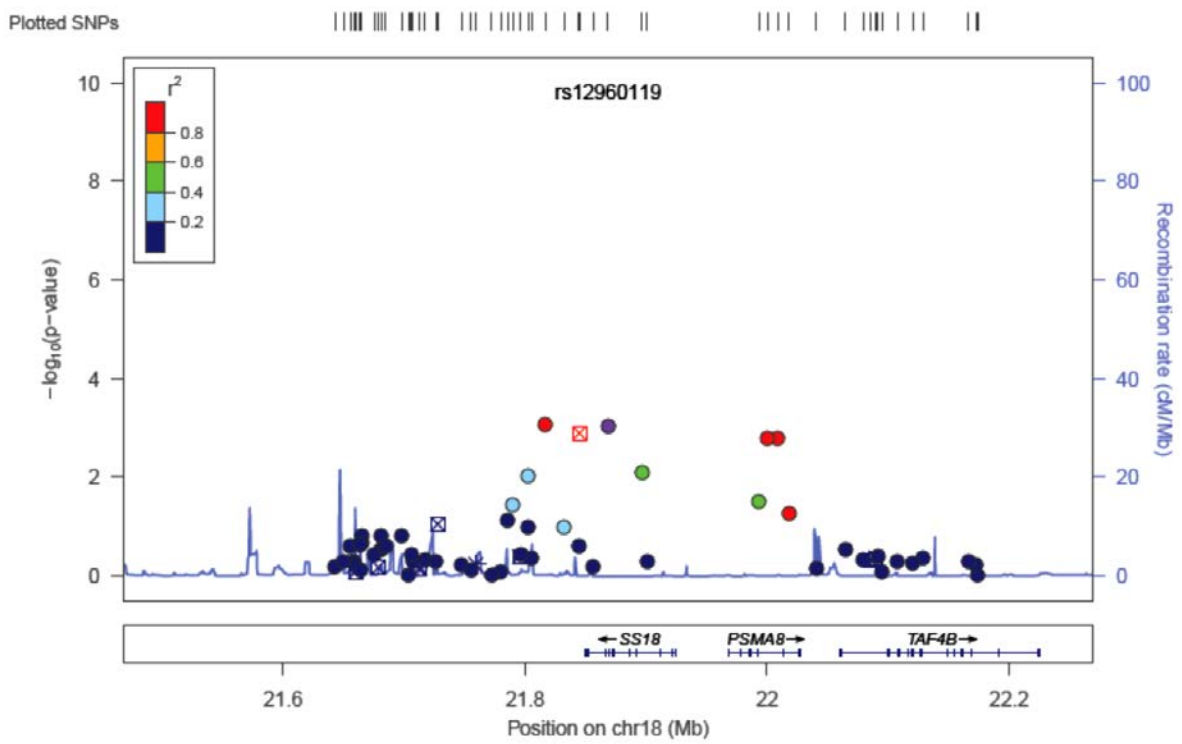


Figure 2.6. Regional plot of genome scan results (discovery step one) showing rs12960119 is in high LD with a SNP in an evolutionary conserved region (rs765529), annotated with an ⊠ symbol in Locus Zoom (tool to plot regional association results).

This region represents the highly conserved 3' UTR region of the *SS18* gene (Figure 2.7.).

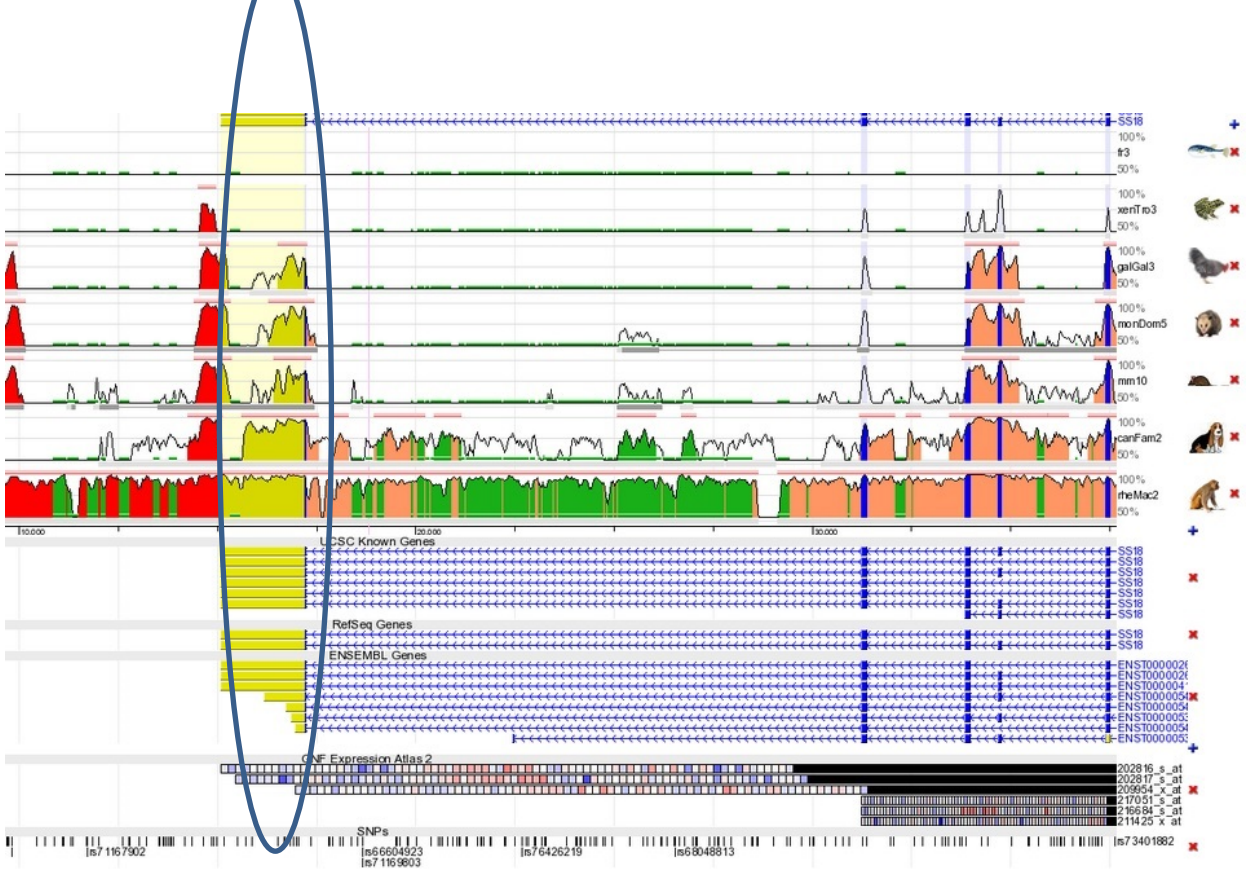


Figure 2.7. Graphical representation of the evolutionary conserved regions (ECR) within the *SS18* gene using ECR browser (Ovcharenko et al. 2004). ECRs are detected by aligning the sequence of the base genome (human in this case) with others in a pairwise fashion. The human sequence (the base genome) is represented on the bottom axis. The height of the conservation peaks represents the number of nucleotides conserved in a windows of 100 nucleotides centred on that position. Blue color correspond to positions of coding exons, yellow to UTRs. Peaks within the conservation profile that do not correspond to transcribed sequences are in red (intergenic) or salmon (within an intron). Green regions relate to transposable elements and simple repeats. Vertical axis represents genomes of different species the base genome is aligned/compared to. Conserved UTR region is highlighted.

Most suggestive SNPs ($p<10^{-4}$) also showed comparable directions of effect across all three populations (Table 2.5.). Sex separate analysis was also performed but yielded no genome wide significant results (results in Appendix I). Since some of the study populations had very low number of cases (Dutch), splitting them based on sex produced study groups containing very small number of cases and significantly reduced power to detect association signal.

2.3.4. Association analysis – replication stage

In the replication stage, we followed up the six most significant markers with low heterogeneity across studies (rs12960119, rs267738, rs955943, rs8132771, rs2045084, rs1074463) from the discovery stage results and an additional marker (rs913444) identified in the pathway analysis, with direct experimental evidence supporting protein-protein interactions between *NTRK2* (assigned to rs913444) and *TIAM1* (assigned to rs8132771) (Miyamoto et al. 2006). A brief description of the genes included in or flanking the corresponding association signals is presented in Table 2.6.

SNP	Gene in LD r^2>0.8 with index SNP
rs12960119	***PSMA8*** - encodes for proteasome subunit, alpha type, 8. The proteasome is a multicatalytic proteinase complex, which cleaves peptides with Arg, Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or slightly basic pH. It is expresses in the mouse retinal tissue (Freeman et al. 2011). ***SS18*** - encodes for the human synovial sarcoma translocated to chromosome X protein. This gene plays a role in cell adhesion processes (Kim et al. 2009). It is expressed in many adult tissues and, in mouse, the protein was detected in all cell types of the retina (Kim et al. 2009).
rs267738	***CERS2*** encodes the protein ceramide synthase 2 (CerS2), the most abundantly and ubiquitously expressed member of the ceramide synthase family. Ceramide-enriched membranes domains amplify diverse signals and are critically involved in inducing apoptosis (Schenck et al. 2007).It is expressed in mouse retina (Freeman et al. 2011).
rs955943	***LDB2*** encodes LIM domain binding transcription cofactor. LIM domains are present in many proteins that have diverse cellular roles as regulators of gene expression, cytoarchitecture, cell adhesion, cell motility and signal transduction(Kadrmas et al. 2004). Expressed in rat retinal ganglion cells (Wang et al. 2009). LDB2 has been shown to interact with SLK which plays a role in actin dependent apoptosis and adhesion dynamics (Storbeck et al. 2009)
rs1074463	***CDH12*** encodes cadherin 12, a member of a family of cadherin molecules. Cadherins are integral membrane proteins that mediate calcium-dependent cell-cell adhesion. Cadherin 12 appears to be expressed specifically in the brain. Expressed in the developing mouse retina (Faulkner-Jones et al. 2009).
rs2045084	***TSTA3*** encodes the enzyme GDP-L-fucose-synthetase, which catalyses the conversion of GDP-4-dehydro-6-deoxy-D-mannose to GDP-fucose in the *de-novo*-synthesis of GDP-fucose. Expressed in human retinal pigment epithelium (Bonaldo et al 1996).
rs8132771	***SOD1*** encodes superoxide dismutase 1 , enzyme that catalyzes th dismutation ofsuperoxide into oxygen and hydrogen peroxide. SOD1-/- mice present oxidative damage of retinal pigment epithelial cells. Expressed in mouse retina (Imamura et al. 2006). ***SCAF4*** encodes for an RNA binding protein. It is the human ortholog of rat rA4 protein that interacts with the C-terminal domain of RNA polymerase II. May act to physically and functionally link transcription and pre-mRNA processing (Yurjev et al. 1996).
rs913444	***TRKB*** encodes neurotrophic tyrosine receptor kinase type B. It is located within the membrane, and signalling pathway involving this kinase leads to cell differentiation, proliferation and survival (Numakawa et al. 2010). TrkB protein is expressed in the inner plexiform layer (IPL) and retinal pigment epithelium in mouse retina (Grishanin et al. 2008). It plays a role in the development of synaptic networks in the inner retina.

Table 2.6 Function of the genes within or flanking association signals followed-up in the replication phase. A brief description of the genes included in or flanking the corresponding association signals are described.

The seven selected markers were typed in 846 cases from London and 120 cases from Croatia. Only one SNP (rs12960119) was successfully typed in the Croatian cohort. All markers displayed comparable direction of effect to the discovery analysis, and one reached the Bonferroni corrected significance threshold of 0.0035 (7 SNPs; two genetic models). Results of the replication stage analysis are given in tables 2.7.A and 2.7.B.

A	**London**				**Croatia**			
SNP	p-value	OR	L95% CI	U95% CI	p-value	OR	L95% CI	U95% CI
rs2045084	0.0009	1.31	1.12	1.54	NA	NA	NA	NA
rs267738	0.0075	0.81	0.69	0.94	NA	NA	NA	NA
rs8132771	0.0552	1.22	1.00	1.49	NA	NA	NA	NA
rs955943	0.2027	1.19	0.91	1.57	NA	NA	NA	NA
rs12960119	0.4167	1.07	0.91	1.27	0.45	1.23	0.72	2.10
rs1074463	0.7500	0.96	0.73	1.25	NA	NA	NA	NA
rs913444	0.6640	1.03	0.88	1.21	NA	NA	NA	NA

B	**London**				**Croatia**			
SNP	p-value	OR	L95% CI	U95% CI	p-value	OR	L95% CI	U95% CI
rs267738	0.0071	0.83	0.72	0.95	NA	NA	NA	NA
rs2045084	0.0278	1.13	1.01	1.26	NA	NA	NA	NA
rs8132771	0.1050	1.20	0.96	1.49	NA	NA	NA	NA
rs955943	0.2919	1.15	0.88	1.50	NA	NA	NA	NA
rs1074463	0.6203	0.96	0.83	1.12	NA	NA	NA	NA
rs12960119	0.3704	1.07	0.92	1.25	0.36	1.23	0.79	1.91
rs913444	0.1080	1.10	0.97	1.25	NA	NA	NA	NA

Table 2.7. Replication phase association results using a dominant (A) and additive (B) effect model for the SNP minor allele Odds ratios with 95% confidence intervals are given for each SNP with respect to the minor allele.*NA-Too many assays failed

This marker, rs2045084, is located on chromosome 8q24.3, 949bp 3' of *TSTA3* and 7.8kb 5' of *PYCRL* gene and displayed a statistically significant association under a dominant model ($p=9x10^{-4}$, OR=1.31) and nominally significant association under an additive model (p=0.028, OR=1.13). The next best associated marker, rs267738, is located on chromosome 1q21.3 within an exon of the *CERS2* gene and displayed a nominally significant association under both additive and dominant models (additive $p=7.1x10^{-3}$, OR=0.83). rs12960119, which almost reached genome wide significance in the discovery phase, did not replicate in the London cohort (p=0.37 additive effect). The minor allele of this marker (allele G, frequency=0.092) was however associated with increased risk, as in the previous analysis, but the effect was smaller in size (OR additive model=1.07, OR dominant model=1.07). This reflected the frequency of the G allele of rs12960119 in the WTCCC NBS controls used in this replication analysis being higher (MAF=0.087) than that of the CEU HapMap population (MAF=0.075) or the WTCCC 1958BC control sample used in the discovery stage (MAF=0.078), whereas the minor allele frequencies were comparable in cases. This marker was successfully genotyped in 109 cases and 265 controls from Croatia. It did not reach statistical significance (p=0.36) in this sample due to the small sample size but the direction and size of the effect (OR=1.23; 0.79-1.91 95%CI) were comparable to those in the discovery cohorts, with the minor allele frequency of the risk allele higher in cases (MAF=0.14) compared to the controls (MAF=0.11).

Genotyping was unsuccessful for the remaining 6 SNPs in the samples from Croatia. Less than 90% of the samples were successfully typed in the cases (45% rs2045084, 51% 267738, 68% rs913444, 51% rs8132771, 45% rs955943) and therefore not included in the replication analysis.

A few genotyping steps were involved in typing rs12960119. First attempt of Taqman genotyping resulted in missing results for 53 individuals (71 subjects were successfully

genotyped). Genotyping of these individuals were repeated, which resulted in additional genotyping of 24 individuals. Sequencing was attempted for the remaining subjects. PCR did not work for 11 of those individual so sequencing could not be attempted. 18 samples were sequenced, and results obtained for only 2 subjects. Additionally, 13 subjects were re-recruited and blood samples were drawn in a new set of vacutainer blood tubes. DNA extraction and genotyping were successful for all of the samples, and it was concluded that that blood tubes used in the original group of cases must have past their expiry dates, since effectively when few samples were redrawn no issue were observed. Given this long and costly process of genotyping rs12960119, the same was not attempted for the rest of the SNPs included in the replication stage.

2.3.5. Overall meta-analysis

Results for the seven SNPs analysed through both discovery and replication steps were combined into an overall meta-analysis (Table 2.8.) and this yielded one variant (rs267738) reaching the conventional genome-wide significance threshold and the next most significant SNP (rs2045084) falling just below our initial GWAS threshold (dominant model p-value=2.59×10^{-7}; threshold set at $1.27x10^{-7}$). SNP rs267738, within the *CERS2* gene, showed the strongest association signal under a dominant effect model ($p=2.11x10^{-8}$). The SNP effect size was consistent in both direction and magnitude across all four populations (Figure 2.8.), with cases showing an excess of the major allele homozygote genotype AA (OR=0.78; 0.71-0.85 95%CI).

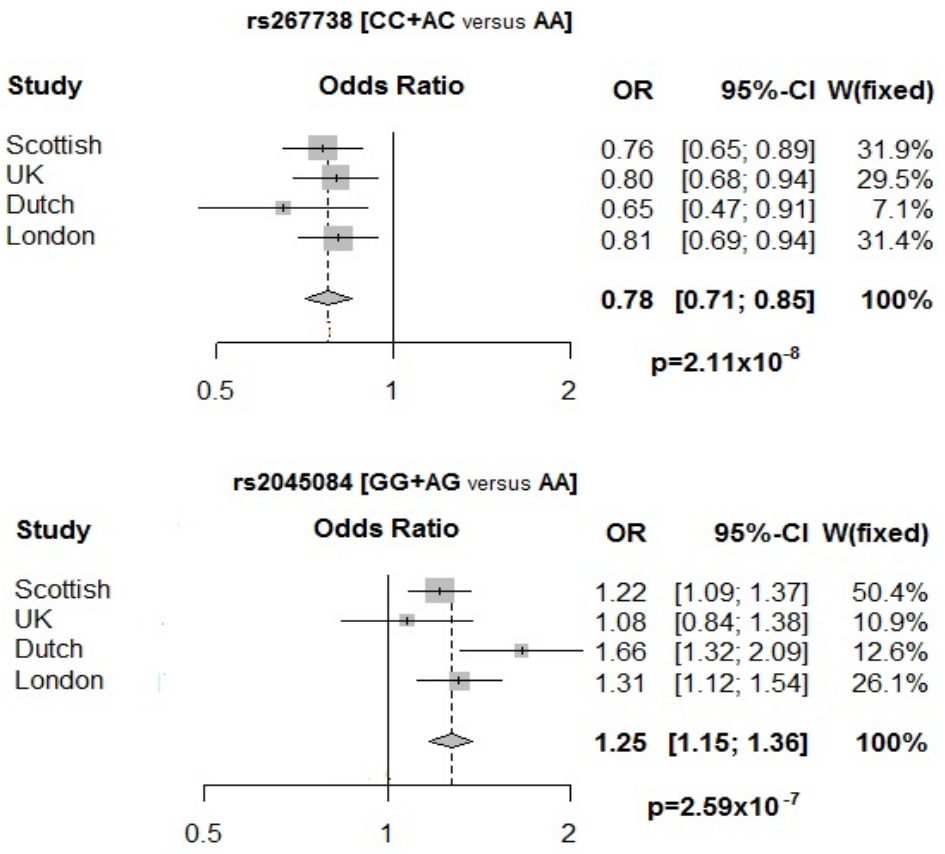


Figure 2.8. Forrest plots for the SNPs with the strongest association signals in the combined analysis under a dominant model for the minor allele effect. Population odds ratio with respect to the minor allele and its 95% CI is represented by a grey box and horizontal line. The bottom rhombus shows the weighted estimate of the odds ratio overall studies.

The signal in this region was further supported by an additional SNP (rs267733) showing suggestive significance in the discovery phase meta-analysis ($p=5.29x10^{-5}$). This marker is located within the gene *ANXA9*, which is 5' of *CERS2*.

The next most significant SNP association in the combined meta-analysis (rs2045084) is located 3' of the *TSTA3* gene, and showed some heterogeneity between studies ($I^2=32.8$, p=0.22). The association was not significant in the UK discovery cohort (p=0.55, OR=1.09 (0.85-1.40 95%CI). The rs12960119 SNP, located in the *SS18* gene, which reached genome wide significance at the discovery phase, did not replicate in the London cohort and did not reach genome wide significance in the overall meta-analysis

($p=2.04x10^{-6}$, dominant model; $p=2.49x10^{-6}$ additive model). Heterogeneity in the overall meta-analysis was high under both models ($I^2=50.2$, p=0.09, under a dominant model; $I^2=49.6$, p=0.09, under an additive model) and contrasted with the homogeneity of association signal in the discovery phase, where three diverse study populations had been used.

Additive model						**Dominant model**				
SNP	OR	L95% CI	U95 %CI	P-value	Direction	OR	L95% CI	U96 %CI	P-value	Direction
rs267738	0.81	0.75	0.88	1.43E-07	----?	0.78	0.71	0.85	**2.11E-08**	----?
rs2045084	1.16	1.09	1.24	2.85E-06	++++?	1.25	1.15	1.36	2.59E-07	++++?
rs12960119	1.26	1.14	1.39	2.49E-06	+++++	1.27	1.15	1.41	2.04E-06	++++ +
rs955943	1.41	1.21	1.64	5.45E-06	++++?	1.42	1.21	1.66	1.51E-05	-+++?
rs8132771	1.31	1.16	1.48	1.65E-05	++++?	1.31	1.16	1.49	1.64E-05	++++?
rs913444	1.17	1.09	1.26	2.09E-05	++++?	1.17	1.08	1.29	3.38E-04	++++?
rs1074463	1.15	1.05	1.26	2.59E-03	+++-?	1.21	1.09	1.35	2.90E-04	+++-?

Table 2.8. Combined discovery and replication phases meta-analysis results Odds ratios with 95% confidence intervals and direction of effect with respect to the SNP minor allele. Directions of the effect are given for respectively the Netherlands, UK, Scotland, London and Croatia studies. Genome wide significant result is highlighted in bold.

2.3.6. Genetic risk score

Each of these markers contributed only a modest amount to individual risk of RRD. In order to assess the combined effect of our top ranked markers, we prioritized five of the SNPs that were followed up at the replication phase and combined them in a genetic risk score by using a simple allele counting approach. The difference in mean number of risk

alleles between cases and controls was statistically significant in all studies using the Mann-Whitney test (discovery Dutch $p=2.05x10^{-5}$; discovery Scottish $p=2.19x10^{-11}$; discovery UK $p=4.1x10^{-6}$; replication London $p=2.4x10^{-3}$). Figure 2.9. illustrates the distribution of the risk alleles in our four independent studies, displaying, in each study, a higher proportion of cases compared to controls amongst individuals carrying four or more risk alleles.

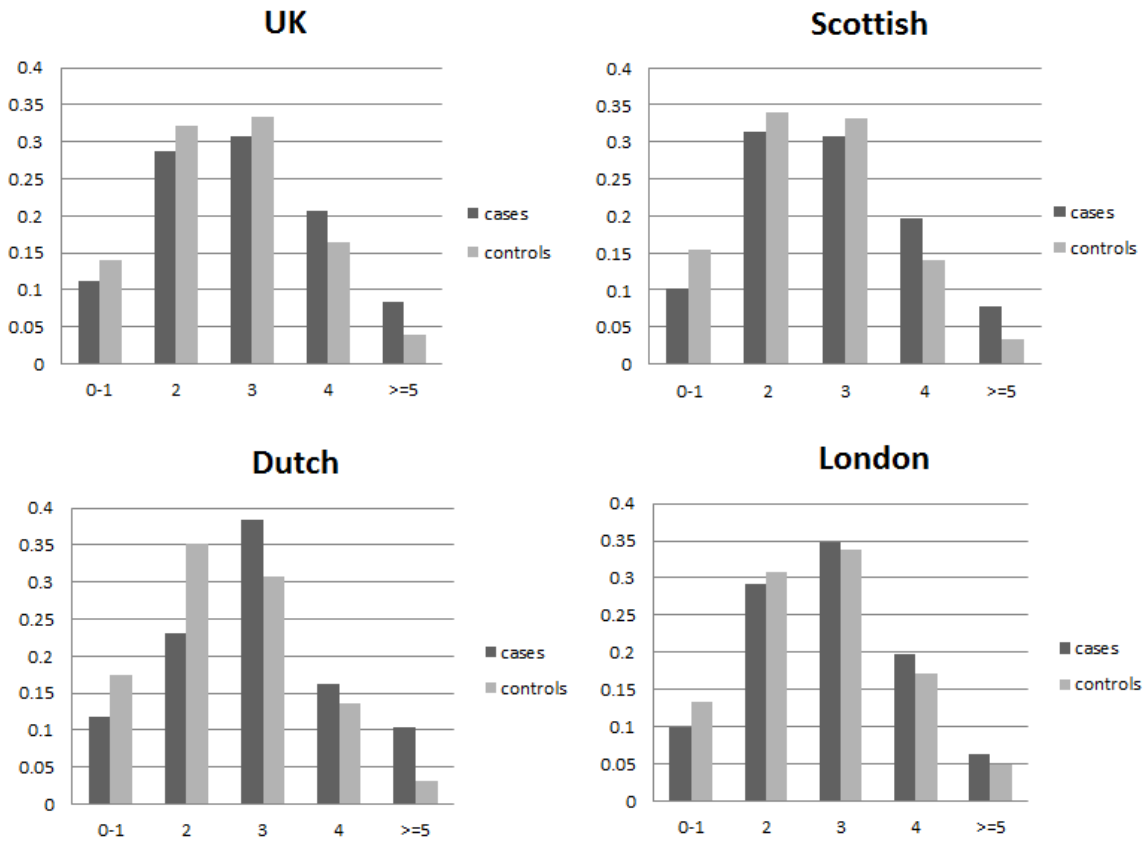


Figure 2.9. Distribution of the number of identified risk alleles in cases (dark grey) and controls (light grey) in four different populations. The horizontal axis represents the number of RRD risk alleles and the vertical axis represents the percentage of individuals with a given number of risk alleles in cases and controls.

The risk for RRD was estimated for each number of risk alleles, relative to the median number of risk alleles in the control populations (3 risk alleles). Individuals with no and one allele were grouped together, as well as individuals with 5 or more alleles. ORs were calculated for each population separately. ORs ranged from 0.70 (95% CI 0.53-0.93) in Scottish population for 0-1 risk alleles to an odds ratio of 6.05 (95% CI 3.31-11.02) for 5 or more risk alleles in Dutch study group. A significant increase of RRD risk (OR>1) for the individuals with five or more risk alleles was detected in all four populations: with odds ratios of at least two in the discovery cohorts and >1.22 in the London replication cohort (Figure 2.10.). Confidence intervals around risk estimates for 4 or more risk alleles were very wide due to small numbers of individuals.

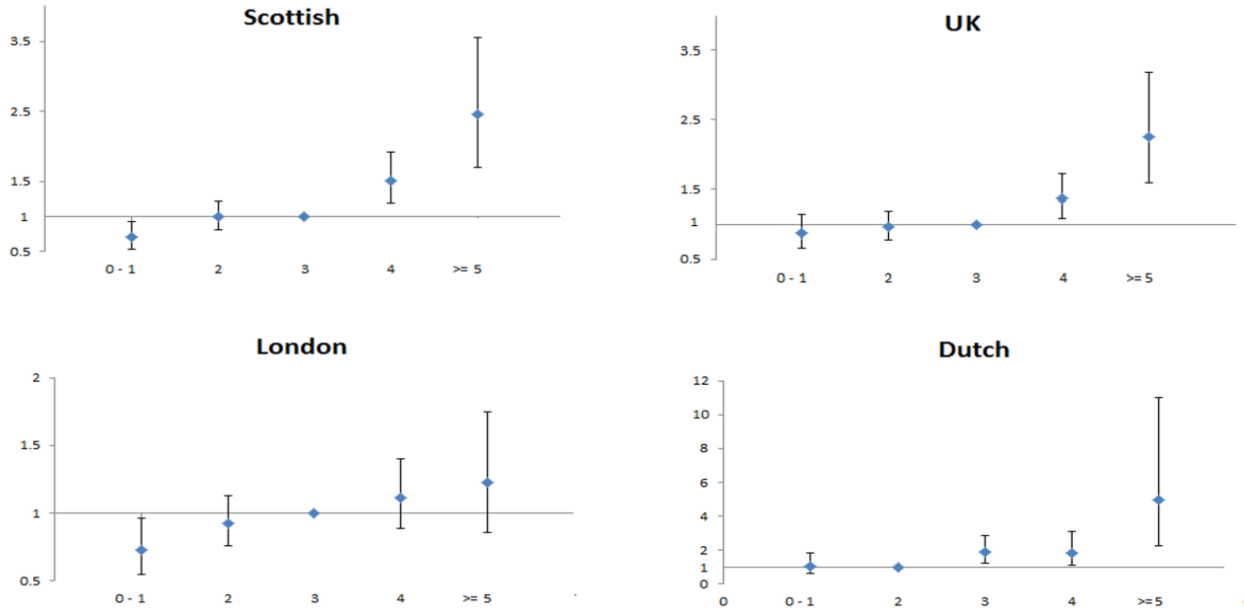


Figure 2.10. ORs and 95% CI for the 5 risk alleles used to build a genetic risk score. The vertical axis represents the OR and the horizontal axis represents number of RRD risk alleles. ORs are calculated relative to the carriers of median number of risk allele in the control population for each of the four populations separately. Individuals with five of more risk alleles were grouped together due to the small number of individuals with high number of risk alleles.

2.3.7. Variation in disease liability explained

Following stringent QC (Lee et al. 2011) and using the method of Yang et al. (Yang et al. 2010) implemented in the GCTA software, 27.4% of the variance in underlying disease liability could be assigned to the combined effect of the markers tagged by the primary genome wide scan SNPs. The proportion of the variance explained remained comparable when different minor allele frequencies and missing genotype cut-offs were used (Table 2.9.). The signal was not due to population structure within Scottish cases and controls since it was not found when the Scottish (SOCCS) control group was artificially split into cases and controls, (liability transformed estimate=0.000001, SE=0.23). These GCTA estimates further supported the evidence of a modest polygenic contribution by common genetic variants in RRD.

Threshold	Number of SNPs	Estimate (SE)	Adjusted estimate (SE)	Liability transformed estimate(SE)
MAF >0.01				
SNP missingness 0.05	257 095	0.3587 (0.1206)	0.4479(0.1498)	0.275 (0.092)
SNP missingness 0.005	255 704	0.3573 (0.1207)	0.4467 (0.1502)	0.274 (0.092)
MAF>0.05				
SNP missingness 0.05	250 817	0.3701 (0.1192)	0.4612 (0.1480)	0.283 (0.091)
SNP missingness 0.005	249 439	0.3687 (0.1194)	0.4601 (0.1484)	0.282 (0.091)

Table 2.8.Variation in RRD liability explained by the combined GWAS SNPs using different parameters. Estimated proportion of the variance in disease liability on the observed scale (Estimate), corrected for imperfect LD between SNPs and causal variants (Adjusted estimate) and liability scale (Liability transformed estimate), explained by the consensus panel of SNPs typed on both the Illumina CNV370v3-Quad array (Scottish RRD cases) and Illumina-300 and 240S HumanHap array (SOCCS controls)

2.3.8. Pathway analysis

2.3.8.1. Ingenuity Pathway Analysis (IPA)

The genes putatively underlying the association signals were identified using a protocol implemented in GRAIL (Raychaudhuri et al.2009) (several genes were assigned to one association signal - Table 3.9). Genes underlie top 6 SNPs from the discovery stage meta-analysis were selected for the first step of pathway analysis. This analysis revealed one highly significant network ($p=10^{-62}$) involving 35 molecules in total, 21 of which were present in the top six signals (Figure 2.11.). This network showed enrichment for molecules involved in cell death, DNA replication, recombination and repair, and haematological system development and function.

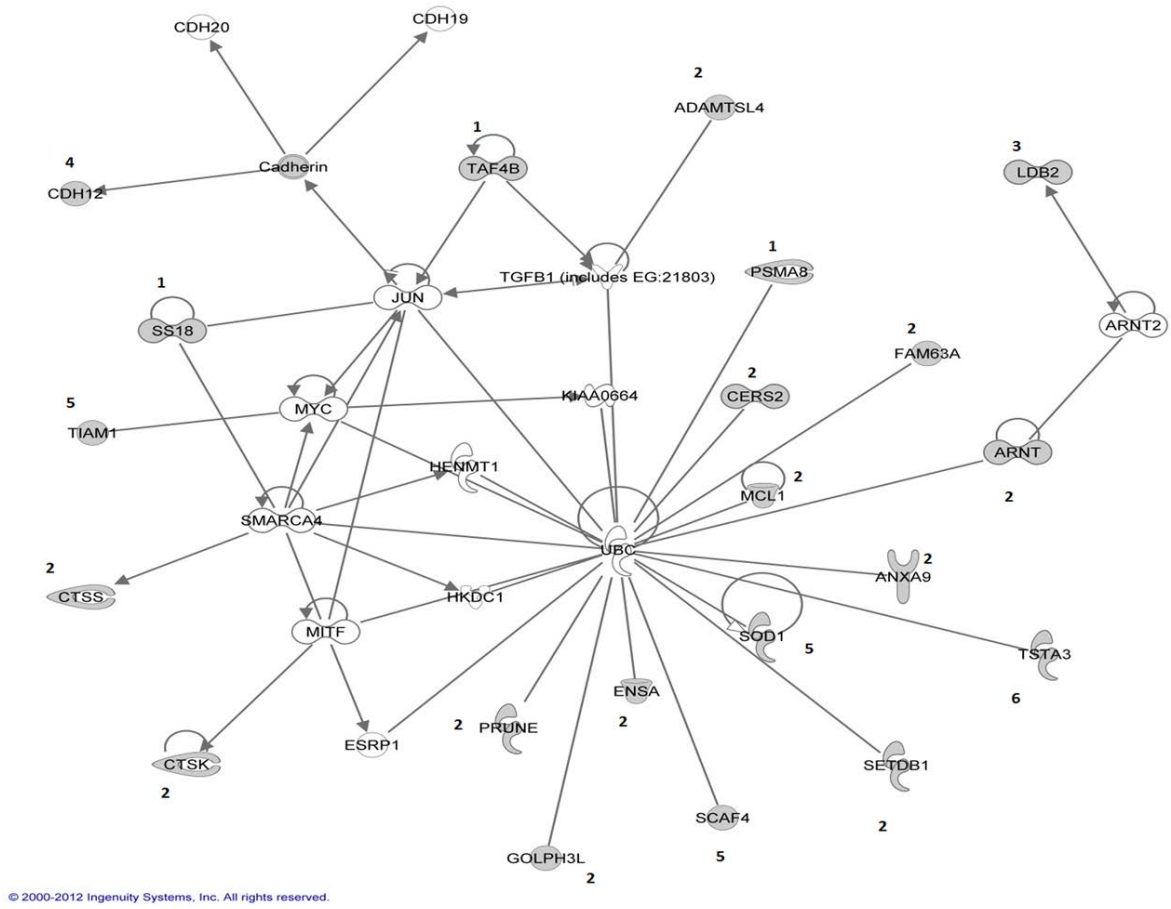


Figure 2.11. IPA network obtained with all genes possibly involved in the six top association signals from the discovery phase. Molecules colored in grey represent the genes products input. The number assigned to such entry represents one of the six SNPs signals (rs12960119-1, rs267738-2, rs955943-3, rs1074463-4, rs8132771-5, rs2045084-6) they are attached to.

Additional pathway analysis was carried out using the 23 top results ($p < 10^{-4}$) from the discovery meta-analysis additive model (Table 2.10). Four direct interactions were identified in this extended pathway analysis (*MCL1-RAE1*, rs267738-rs6070015*; BMP7-SETDB1*, rs6070015-rs267738; *SETDB1-PPA1*, rs267738-rs7097067: *TIAM1-TrkB*, rs81232771-rs913444) (Figure 2.12.). The latter is strongly supported by experimental evidence of a protein-protein interaction (Myamoto et al. 2006). The SNP rs913444 was therefore included for testing in the replication stage of our analysis.

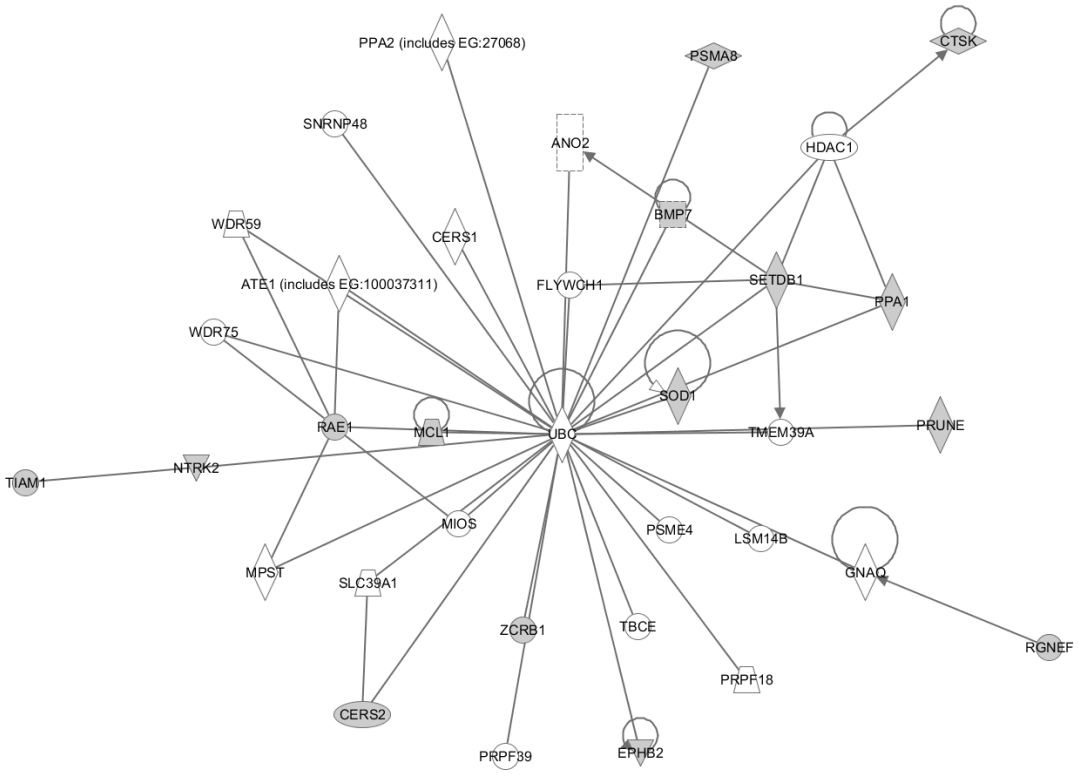


Figure 2.12. IPA network obtained with all genes possibly involved in the top 23 association signals from the discovery phase. Molecules colored in grey represent the genes products input.

SNP	GRAIL input genes	Genes in LD block r^2>0.8 with index SNP
rs12960119	PSMA8, TAF4B, SS18	SS18, PSAM8
rs267738	GOLPH3L, CTSK, FAM63A, SETBD1, CTSS, MCL1, CERS2, ANXA9, PRUNE, ARNT, HORMAD1, ADAMTSL4, ENSA	CERS2
rs955943	LDB2	LDB2
rs7097067	PPA1, NPFFR1,TYSND1, AIMF2, SAR1A	*PPA1,NPFFR1
rs1074463	CDH12	CDH12
rs2045084	TSTA3	TSTA3
rs8132771	TIAM1, SOD1, SFRS15	SOD1, SFRS15
rs11259960	HOMER2	HOMER2
rs2368106	CW22	*CW22,UB2E3
rs7234959	TAF4B,PSMA8, SS18	SS18, PSMA8
rs6070015	RAE1, SPO11, BMP7, RBM38	BMP7
rs913444	TRKB	*TRKB,AGTBPBP1
rs4893905	CW22	*CW22,UB2E3
rs1477441	CDH12	CDH12
rs12193473	C6orf170	*MAN1A,C6orf170
rs218843	C6orf170	*MAN1A,C6orf170
rs4715056	GPR115, OPN5,CD2AP,GPR111	C6orf38
rs2817896	EPHB2	EPHB2
rs10515162	RGNEF	RGNEF
rs11181447	GLT8D3, YAF2, ZCRB1, PPHLN1, PRICKLE1	ZCRB1, PPHLN1, PRICKLE1
rs12202993	C6orf170	*MAN1A,C6orf170
rs564351	GPR115, OPN5,CD2AP,GPR111	*OPN5, PTCHD4
rs6035211	SLC24A3	*C20orf79,SLC24A3

Table 2.10. List of genes underlying most significant ($p<10^{-4}$) SNPs from discovery stage meta-analysis (additive model). List identified by GRAIL is given as well as the list of genes in the LD block with index SNP ($r^2>0.8$)

2.3.8.2. MAGENTA analysis

In order to test for enrichment of GWAS results with a different pathway analysis tool, MAGENTA pathways analysis was performed. However the analysis revealed no significant enrichment of any biological pathways in the GWAS results from discovery stage step 1 analysis. No pathway reached the set significance threshold (FDR <0.05). Top ranking pathways were pathways involved in nuclear transport, mRNA transcription and phagocytosis. Top 50 ranked networks of the analysis are given in Appendix II.

2.4. Discussion

To my knowledge, this is the first genome wide association study ever performed for rhegmatogenous retinal detachment, a condition which continues to result in significant vision loss. Genetic component to RRD has been suggested by the familial aggregation studies reporting two to three fold increased frequency of RRD in siblings of affected subjects compared with siblings of unaffected subjects in Scottish (λ=2.1) and Dutch (λ=2.6) populations (Mitry et al. 2011, Go et al. 2005). This result is similar to the type 2 diabetes sibling risk (λ~3; Lyssenko et al. 2005) and much lower than the sibling risk in type 1 diabetes (λ =15; Davies et al. 1994). Here the results supporting a polygenic component for this condition is confirmed, and several contributing common risk variants, of small effect, were identified.

2.4.1. Functional relevance of best associated SNPs

A genome-wide significant signal of association with RRD was found for the marker rs267738. This is a missense (Glu to Ala) coding SNP located within the *CERS2* gene, in a gene rich region. The index SNP has been associated with the expression level of *CERS2* in lymphoblastoid cell lines (exon eQTL for *CERS2*) (Montogomery et al. 2010) making the regulation of *CERS2* expression a strong functional candidate for the RRD association. *CERS2* is expressed in mouse retina (Freeman et al. 2011) and encodes the protein ceramide synthase 2 (CerS2), the most abundantly and ubiquitously expressed member of the ceramide synthase family (Laviad et al. 2008). Ceramide-enriched membranes domains amplify diverse signals and are critically involved in inducing apoptosis (Schenck et al. 2007). Ablation of CerS2 in mice causes changes in biophysical properties of the cell membrane, due to the lack of long chain ceramides (Sylva et al. 2012). In vitro, ceramides have been shown to mediate apoptosis of mammalian photoreceptors (German et al. 2006, Sanvicens et al. 2006) and retinal pigment epithelium cell lines (Kannan et al. 2004, Tomita et al. 2000, Sreekumar et al. 2009). In a mouse model of retinitis pigmentosa, inhibition of ceramide synthesis slowed disease progression (Strettoi et al. 2010) and mutation in the ceramide kinase like gene *CERKL* has been implicated in a recessive form of retinitis pigmentosa in humans (Tuson et al. 2004). Increased cellular ceramide levels were found in brains of patients with Batten's disease (Puranam et al. 1997) and in retinas of patients with Farber's diseas (Zarbin et al. 1988). A feature of both diseases is blindness, although this is due to retinal degeneration rather than detachment. An association signal with RRD in this region is supported by an additional association of SNP rs267733 ($p=5.3x10^{-5}$, dominant model) located 11kb from *CERS2*. A variant in high LD (rs267734; $r^2=1$) with this SNP has previously been reported to be associated with chronic kidney disease (Kottgen et al. 2010).

The second strongest association signal is located on chromosome 8q24.3 between two genes (*PYCRL* and *TSTA3*) but the block of high LD with the index SNP comprises only the *TSTA3* (tissue specific transplantation antigen P35B) gene (Figure 2.13.). *TSTA3* encodes the enzyme GDP-L-fucose-synthetase, which catalyses the conversion of GDP-4-dehydro-6-deoxy-D-mannose to GDP-fucose in the *de-novo*-synthesis of GDP-fucose. GDP-fucose is a substrate in the posttranslational fucosilation reactions. Mice with induced null mutation in the TSTA3 gene show a complete deficiency in cellular fucosylation (N-fucosylation, O-fucosylation and E-/P-/Lselectin activities) (Smith et al. 2002). This enzyme has been shown to play an important role in the post-translational fucosylation of selectin ligands (Smith et al. 2002). Selectins are cell adhesion molecules which bind to specific carbohydrate determinants of molecules on opposing cells, and play an important part in the inflammation processes (Ley, 2003). It has been demonstrated that GDP-L-fucose-synthetase controls adhesion of colorectal cancer cell to endothelial cells (Zipin et al. 2004). Furthermore, the de novo synthesis of GDP-fucose plays an important role in the flagellar adhesion in Trypanosomabrucei (Turnock et al. 2007). Given this evidence, GDP-L-fucosesynthetase could potentially play a role in the inflammation and cell adhesion processes involved in the pathogenesis or RRD.

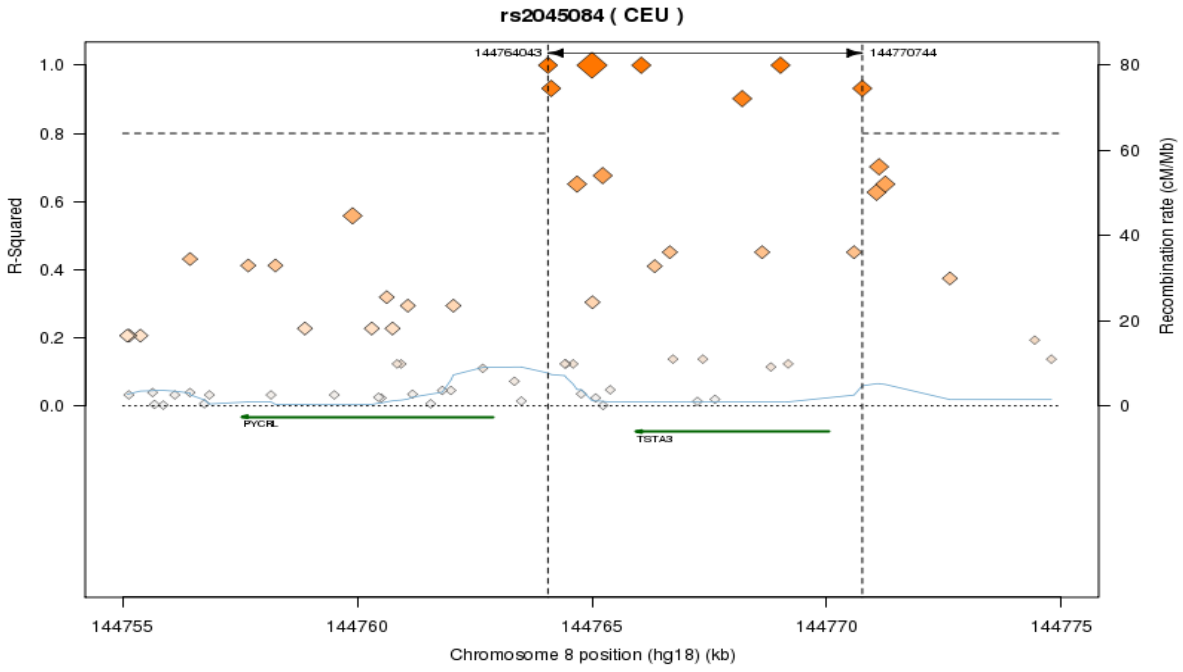


Figure 2.13. Regional LD plot around rs2045084 is presented. A region of 500 kb around the SNP is displayed. Recombination rate is displayed by a blue line with scale on the right-hand axis and the R-squared scale on the left-hand side. Genes are displayed in the bottom part of the figure, with the arrow showing the direction of transcription

Finally, marker rs12960119 which reached genome wide significance in the discovery step, but failed in the replication step, is located within the *SS18* gene. This remains, however, a particularly strong candidate gene for association with RRD because of the role of SS18 in cell adhesion. SS18 is expressed in many adult tissues and, in mouse, the protein was detected in all cell types of the retina (Kim et al. 2009). Integrins are transmembrane heterodimer glycoproteins that play an important role in cell signalling. They are the main receptor proteins for both binding and responding to the extracellular matrix and various matrix proteins are recognized by multiple integrins. Integrins bind the extracellular matrix to the actin cytoskeleton, which activates signaling cascade involved in cytoskeletal reorganization and cell adhesion processes. Cytosolic SS18 has been shown to bind F-actin and to play an essential role in the formation of stress fibers and focal adhesion (Eid et al. 2003). Ablation of SS18 in cultured human cells led to

compromised adhesion on specific extracellular matrix substrates (fibronectin and laminin-111 but not on collagen type I and IV) (Kim et al. 2009) prompting the authors to propose cytosolic SS18 as a sensor of extra cellular signals. In chicks and rats retina, β1 integrin and fibronectin had been localized exclusively at the basal membrane of RPE, chroriocapillaries and at the vitreo-retinal border (Philip et al. 2007). A type of integrin molecule ($\alpha_v\beta_5$-Integrins) that may play a role in the binding of RPE to the interphotoreceptor matrix (Fineman et al. 1997, Finneman et al. 2003) have been found at the apical part of the retinal pigment epithelium (RPE-photoreceptor interface) in human and rats (Finneman et al. 2007). However the role of SS18 in activating this specific class of integrin is unknown, and retinal detachment occurs on the basal part of RPE.

The region of association with RRD is in high LD with an evolutionarily highly conserved sequence in the 3' UTR region of the *SS18* gene, suggesting a functional role. While the association signal in this region was not replicated in a large UK sample using NBSC controls, the minor allele frequency difference for the index marker was high at 3% between cases and controls in the Croatian replication sample. However it did not reach statistical significance due to the small sample size.

2.4.2. Pathway analysis

The non-randomness of the top association signals in our discovery phase analysis was tested through their connectivity in "*in-silico*" pathway networks occurring more often than by chance. A highly significant network was created based on the six most significant association signals and included 21 of our 22 entries, in a network of 35 molecules, which is an unusual result (Figure 2.11.). However when a different approach was used that tests for enrichment of the top signals in the known pathways, and does

not create a new "in silico" networks, no significant results were revealed. IPA analysis was expanded to include more signals (index SNP with association p-value $<10^{-4}$). In the resulting network, direct protein-protein interaction annotations were detected. For example, an interaction between TrkB (encoded by *NTRK2* gene) and TIAM1. *TIAM1* encodes a guanine nucleotide exchange factor which activates a small GTPase of the rho family and modulates changes in cytoskeleton downstream of membrane receptor stimulation. These receptors include TrkB (Miyamoto et al. 2006) a receptor for brain derived neurotrophic factors, but also adhesion molecules such as β1 integrins (Hamelers et al. 2005, O'Toole et al. 2011). The former, TrkB, and β1 integrin links functionally with the *SS18* product. TIAM1 signalling is also required in E-cadherin mediated cell-cell adhesion (Malliri et al. 2004). In mouse retina, TrkB protein is detected in the inner plexiform layer (IPL) and RPE (Grishanin et al. 2008) and while it plays a role in the development of synaptic networks in the inner retina, a role in the RPE has not been established. The TrkB-TIAM1 interactions, and the interaction of SS18 to actin filaments and activation of integrin, suggest a significant role for these players in the regulation of actin cytoskeletal organisation in the pathogenesis of RRD. They may also participate in the strength of the apical RPE filopodia – photoreceptor outer segments' inter-digitation and adhesion. Our pathway analysis did not demonstrate an enrichment of highly significant results in any of the known pathways (MAGENTA analysis), however this analysis only utilized results from the genome-wide scan and therefore omitting some significant results obtained from the later meta-analysis of the discovery stage populations. This problem is also present in the IPA analysis as only top signals were chosen.

2.4.3. Risk conferred by combined variants

While single marker associations were weak, association with RRD jointly using the five markers showing the most significant and consistent association was significant in all four studies (Scottish discovery, Dutch discovery, UK discovery, London replication) so

that the combined genetic score was a potentially useful predictor of disease. Using all the genotyped SNP additive effects from the discovery genome-wide scan jointly, explained 27.4% of the underlying liability to RRD, supporting a moderate polygenic contribution to disease status (solely due to the common variants tested in this scan). This method have been successful in explaining the proportion of genetic variance in other complex diseases and traits with previously identified risk loci (Chron's disease 23%, bipolar disorder 39%, type I diabetes 30%, (Lee et al. 2011)), as well as in traits where no single risk variants have been identified in genetic association analysis (cognitive ability change from childhood to old age; account for 24% of the variation (Deary et al. 2012)). It has been suggested that estimating the proportion of variance explained that is due to all SNPs jointly may be inflated due to the confounding effect of population structure (Browning and Browning 2011). However, I believe that this should not be a strong contributory factor in our discovery cohort as the variance explained in the controls alone, randomly labeled as cases and controls, was negligible. The ancestry clusters drawn in RRD cases and Scottish controls were overlapping (Figure 2.2.) and care was taken to remove systematic differences between RRD cases and controls due to the different genotyping platforms used (Mitry et al. 2011).

2.4.4. Issues regarding imputation of cases-control data

Analysis of the imputed data was not successful. An overabundance of highly significant associations when standard QC restrictions were applied highlighted the known issue of Type 1 error inflation in the analysis of case-control data imputed from different genotyping platforms. After more stringent QC criteria were applied, the number of highly significant SNPs dropped. However, even with this rigorous criteria 37-53% of highly significant SNPs ($p<10^{-5}$) in cases and controls presented allele frequency difference greater than 5% when compared to CEU's frequencies. In addition, when highly significant marker's frequencies were compared between the control set (SOCCS) and another Scottish cohort (LBC), similar inconsistencies were observed

except for one region on chromosome 4. Since the data are imputed against the HapMap CEU population, when MAF in cases (or controls) differs considerably from the CEU, the end result is a very different MAF for that particular SNP between cases and controls, which drives highly significant associations. Based on these allele frequency difference it can be suggested that association signals could come as a result of the imputation process and are artifacts of data imputation errors (and not as a result of true associations). This differential measurement error introduced when case-control data are imputed from different genotyping platforms has now been well known and documented (Sinott and Kraft, 2011). This problem is improved but not completely resolved if cases and controls from different platforms are imputed together as one population (Sinott and Kraft, 2011), and therefore re-imputing the merged case-control data was not attempted. It has been demonstrated that for SNPs with low MAF, even a small error in genotyped calling produces an inflation of Type 1 error (Moskvina et al. 2006). Finally, re-genotyping a number of controls alongside cases in order to exclude SNPs with problematic gene calling was proposed, and it was shown to reduce Type 1 error (Sinott and Kraft, 2011), however it also requires that a substantial number of SNPs be excluded from the final analysis. This was not attempted in this analysis, due to the financial restrictions of the study. Finally it is important to point out that studies looking into issue of case-control imputation suggest genotyping cases and controls at the same time as the best possible strategy, which was not possible for this particular study. Imputation analysis results were not included in the subsequent steps, but will be worth further investigations in the future.

2.4.5. Study limitations

In addition to the data imputation issues, there were a number of other limitations to this study. As we followed up only a subset of the variants tested in the GWAS discovery step, variants which displayed a non-statistically significant association at this first stage may have been missed and yet could have been statistically significant in the larger

sample size of the combined datasets (false negatives due to lack of power). Furthermore, there was a substantial failure rate (14%) of the SNPs tested on the customized iSelect Illumina platform in the follow-up discovery step. We tried to minimize the effect of this by (re)genotyping the most significant markers not represented by a proxy SNP, however true associations could still have been missed. A further limitation was the use of controls from studies or repositories that were genotyped independently using different genotyping platforms. This can create confounding but true associations should remain consistent across the case-control sets assembled. The controls were unselected with regard to phenotype therefore a number of RRD cases might be expected within the control series given the relatively high prevalence of RRD (0.8%) (Mitry et al. 2011). Since the number of affected RRDs may have fluctuated across control series by chance, this may partly explain the heterogeneity across studies. Furthermore, significant population structure has been reported in two of the WTCCC control cohorts, 1958 BC and NBS (Browning and Browning 2011), that we used in the discovery and replication stages respectively, making it a possible explanation for the non-replication of the *SS18* locus. NBS controls had a 1.2% higher allele frequency for rs12960119 compared to that in the HapMap CEU population, and a 0.9 % higher allele frequency than that in the other control group used in the discovery step (1958 Birth Cohort). The fact that only a subset of SNPs had been typed in the RRD cases used after the stage 1 discovery prevented us from correcting for population structure for the matched RRD cases and controls in the usual manner (by adjusting for ancestry principal components). Heterogeneity of the RRD phenotype is an additional concern as other underlying pathologies usually precede its development, e.g. posterior vitreous detachment, high myopia, lattice degeneration. Larger studies will be useful to tease apart the possible differences between these two pathologies. Results from the published GWAS on high myopia (Kieferet al. 2013, Verhoven et al. 2013) and lattice degeneration (Meguro et al. 2012) did not show any overlap with our most significant associations and neither the *SS18*, *CERS2*, or *TIAM1* signal strengths were affected by removing high myopes from the RRD cases in the primary analysis (Appendix III) although the *TSTA3*, *TrkB* and *CDH12* associations may be influenced by myopia. This

study was not designed to explore the association of rare variants with RRD, and given the fraction of underlying disease liability not accounted for by the common variants tested in this study, the search for rare variants influencing RRD is warranted. Finally, given our sample size, we had limited power to detect and replicate variants of moderate effect sizes.

2.4.6. Summary

Despite these limitations, this first genetic association study of RRD supports a polygenic component underlying RRD risk, and identified several common variants of moderate effect contributing to risk. Additional studies will be required to confirm the individual variants highlighted in this study and additional genotyping and increased sample size will be needed to extend this first exploration of the genetic basis of RRD. Since the pathogenesis of RRD is not fully understood, the identification of genes influencing it should make an important contribution to furthering our understanding of the disorder and in time lead to better prevention and treatment.

Chapter 3. Retinal fundus image analysis and phenotype selection

3.1. Introduction

This chapter aims to provide detailed information about retinal fundus image analysis, the various traits derived from the images and the trait selection procedures used in this study. Semi-automatic computerized methods for the analysis of quantitative traits derived from retinal photographs are detailed, and the most commonly used traits in the analysis of the retinal vessel architecture are also examined. The VAMPIRE software used in this work had been recently updated at the start of this project, hence this chapter provides information on the validation procedures essential for any newly developed software. Finally the selection process for the traits used in the subsequent genetic analyses is explained.

3.2. Retinal fundus image

The medical term “fundus” refers to the base of an organ. For the eye, this term is used for the portion of the inner eye visible though the ophthalmoscope or the fundus camera. It represents the retinal layer of the eye captured by the external camera. The retina is photographed directly though the pupil, which is used as both an entrance and an exit for the fundus camera's illuminating and imaging light rays. An external source of light illuminates the retinal surface, which is then reflected in order to capture a 2-dimensional image of the 3-dimensional retinal tissue surface. This 2-dimensional image

of the retinal surface is termed a retinal fundus image. A typical fundus camera views between 30 and 50 degrees of the retinal surface, and magnifies it 2.5 times.

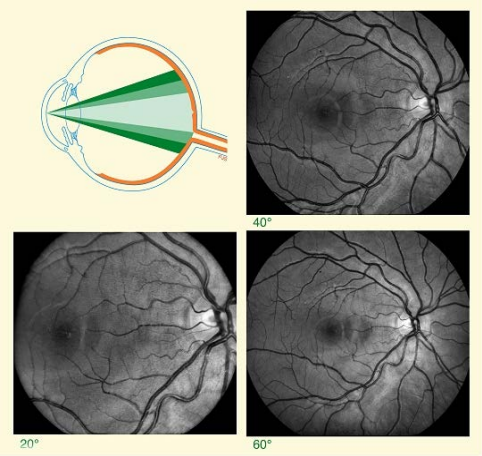


Figure 3.1. Top left panel represents a schematic picture of the creation of fundus image with different angles of the camera. Remaining panels represent fundus images with the 40, 20 and 60 degree angles of the camera. (reproduced form Saine and Tyler, 2002)

The main physiological features of the retinal fundus image include the optic disc, the optic cup, retinal vessels, fovea and macula. Fundus images can be centred on different retinal features, depending on the purpose. The most commonly used ones (due to the medical purpose of each) are: optic disc cantered images (glaucoma), macula cantered images (macular degeneration) and images cantered between the optic disc and the macula(diabetic retinopathy). Most large scale retinal vessel analysis studies use images cantered around the optic disc or between optic disc and the macula. These two image fields provide the most details of the retinal microvasculature, as they capture the area where the circulation enters and leaves the retina. Specialised computer software has been deployed to analyse fundus features.

3.3. Fundus photography and digital imaging

A Canon digital fundus camera which digital back was used to capture non-mydriatic retinal photographs centred between the macula and optic disc for the ORCADES and the Croatia-Korcula study participants. The first step in digital imaging is image capture. The image is composed of an array of pixels of varying intensity across the image. This intensity corresponds to the level of "greyness" (or red, green and blue for a colour image) at a specific point on the image. Modern digital imaging cameras give high-resolution images (camera used in this study is a 12Mpixel) that can later be digitally analysed. Image processing operations transform the "greyness" (or RGB) values of pixels.

Image quality is a major issue in retinal image capture, and image enhancement can develop or improve an image to a point more suitable for subsequent use (image is improved for viewing, processing or analysis). Image restoration helps to reverse damage generated by known causes. The major challenge in the processing and analysis of digital image is the effect of the magnification, due to differences in refraction, axial length, and distance and angle between the camera and the eye (Patton et al. 2006). This challenge can be overcome by taking into the account the optics of the eye and the position of the camera, however, in studies that collect large numbers of individuals this can be difficult to achieve. Therefore many studies have turned to dimensionless or magnification-independent measures to nullify any magnification effect. Such dimensionless traits are arteriovenus ratio (AVR), bifurcation angles, vascular tortuosity and fractal dimensions. These measures overcome the magnifying effect, but can also mean a loss of pertinent information. For example, AVR is a ratio between arterial and venular widths, and many studies showed that different conditions relate specifically to arterioles or venules. Many studies chose to ignore magnifying effect in order to capture

more information. Researchers in the field of retinal image analysis have acknowledged magnification bias, but it still remains an issue in the field of retinal vessel research.

3.4. Computer assisted methods for retinal vessel analysis

Measuring retinal vessel topography by hand is time consuming and laborious activity that effectively precludes the assessment of large number of individuals and hinders the detection of small differences. Hence, computerised approaches have been developed to increase the number of images that can be assessed. Further, the process itself is objective compared to the more subjective manual measurements. Two different types of software were used for the quantitative retinal vessel analysis: VAMPIRE and SIVA.

VAMPIRE

VAMPIRE was developed by Dr Tom MacGillivray at the University of Edinburgh (Wellcome Trust Clinical Research Facility) in collaboration with School of Computing, University of Dundee (Prof Emanuel Trucco and Adria Perez). The software employs a high-performance supervised classification Gabor wavelet algorithm (Soares et al. 2006) as a pre-processing step to segment the retinal vasculature in fundus images (and to divide pixels into vessel and non vessel classes). The image is filtered at different scales to highlight vessels of different sizes. Processing is performed on the green channel of the colour fundus images as this typically exhibits the greatest contrast between vessel and background. All algorithms were implemented in Matlab (The Mathworks Inc., USA). The green channel is first inverted so that vessels appear brighter than the background, and then filtered. The pixels of a fundus image are considered to be objects represented by feature vectors which allow the application of statistical classifiers to

segment the image into two classes, vessel and non-vessel. Training sets for the classifier were constructed from manual segmentations of the images by trained researchers.

All vessels captured by the fundus image are automatically tracked, with branching points automatically selected by the program (Figure 3.2.). Optic disc grid placement is done manually by the grader, and is required for the initialization of the tracking procedures.

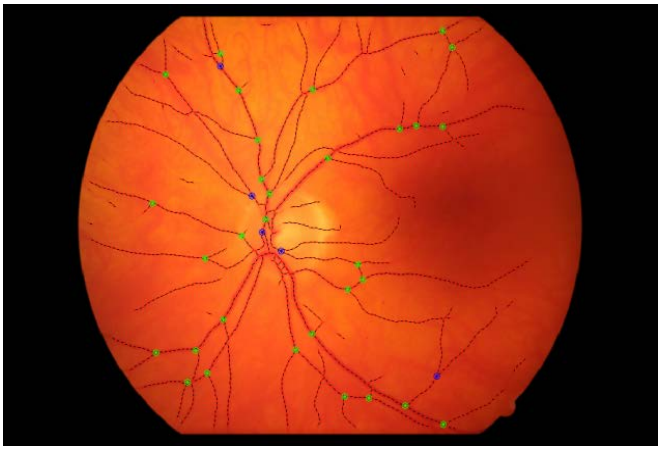


Figure 3.2. A model image analysed automatically by the VAMPIRE software (from the ORCADES study). Fractal geometrical analysis (analysis of the complexity of the microvasculature) and detection of vascular bifurcation points are highlighted on the fundus image.

The program uses information from all tracked vessel from each image to calculate fractal dimensions of a given vascular tree. The user can decide whether to use all or to select specific tracked branching angles in the given output of the image analysis. Separate protocols for analysing branching angles was devised and followed during the image analysis (Appendix IV).

SIVA

Semi-automated computer-assisted image program SIVA (Singapore I Vessel Assessment) was the second software used in this project (Cheung et al. 2010). It was developed at the School of Computing, National University of Singapore. Their approach includes "techniques from wavelet analysis, texture analysis, and curvature ridge/trench analysis, to attain the desired clinical sensitivity" (SIVA Brochure, Exploit technologies). The software supports automatic optic disc grid placement with an additional manual placing option (when the automated assignment needs to be corrected). Vessels located 0.5 - 2 disc diameters away from the optic disc are tracked automatically, if wider than 40µm. Vessel type (venules or arterioles) is also automatically detected by the software. Tracking of the vessels and vessel typing (venules/arterioles) is subsequently improved manually by the grader, as the software does not detect or track all the vessels correctly (Figure 3.3.).

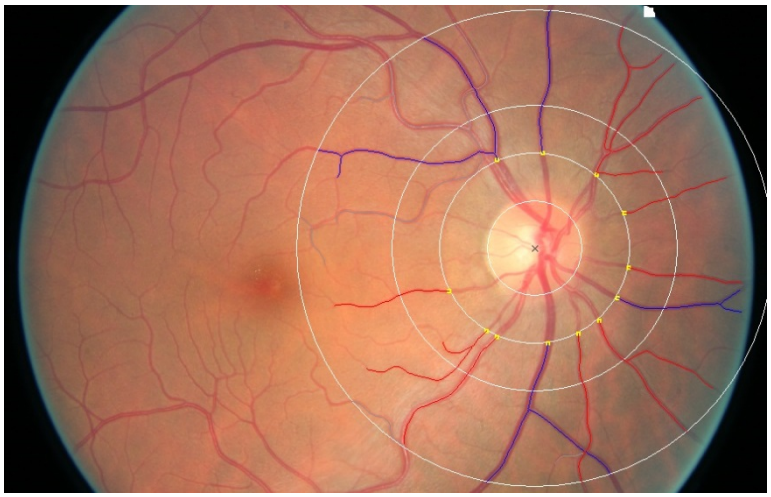


Figure 3.3. Example of the automated tracking of the retinal vessel completed by the SIVA software. Blue vessels represent automatically detected venules and the red ones automatically detected arterioles. The majority of the vessels are not tracked or typed in the correct way, and manual correction is needed for each image analyzed. (Example of an ORCADES images)

The software automatically generates series of transverse stripes (referred to as 'covers' in SIVA) measuring the distances between pixels at the edges of the vessel, later used to generate vessel width measurements (described later in the text). These stripes are perpendicular to a tracked vessel wall and appear at regular intervals down the length of

the vessel. In the final step of the image analysis by SIVA, these lines have to be checked and manually corrected by the grader (Figure 3.4.).

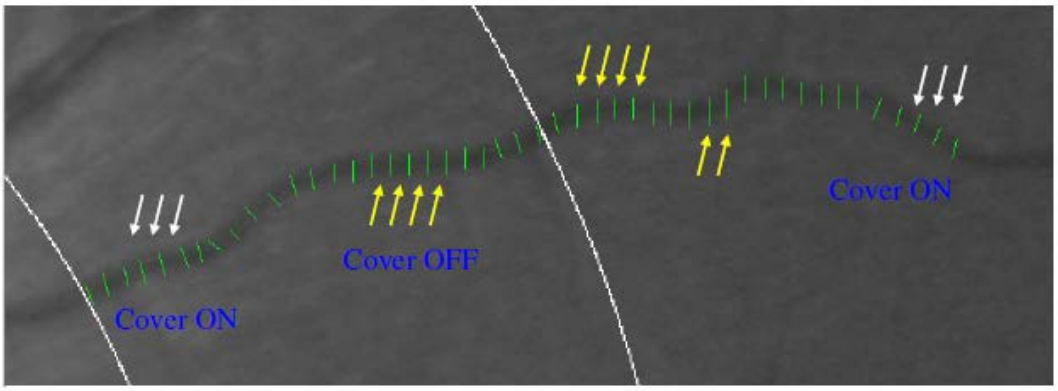


Figure 3.4. Lines perpendicular to the vessel are generated by the software, and later manually corrected by the grader. Lines under/over estimating the width of a vessel (in yellow) are removed and are not used in the subsequent automatic generation of the vessel width measurements (Image from - SERI training manual).

Two different areas of the fundus image are used to derive vessel traits; areas are between 0.5 and 1 disc diameter, and between 0.5 and 2 disc diameters (Figure 3.5.). Once the tracking process is completed, measurements are automatically exported. Quantitative vessel traits derived by the SIVA software are: vessel caliber measurements, tortuosity measurements, branching parameters, fractal analysis parameters.

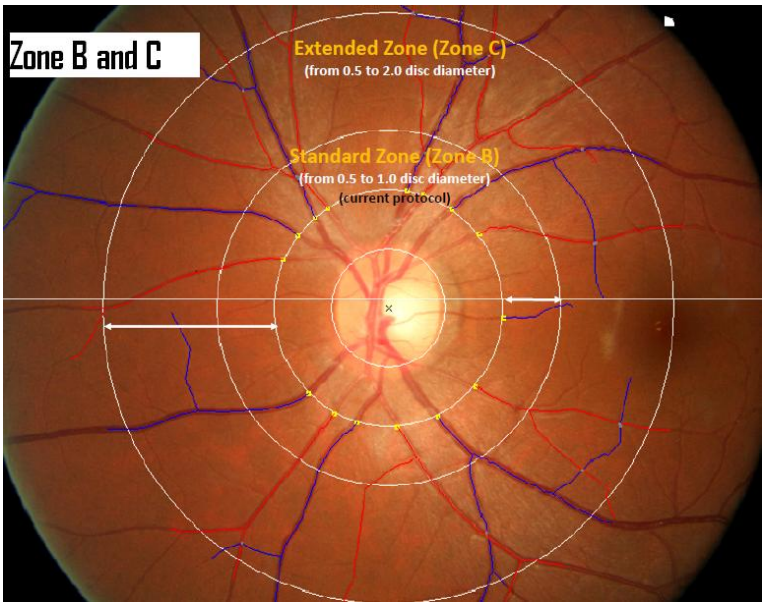


Table 3.5. Example of a retinal fundus images analyzed with SIVA software. SIVA is semi-automated retinal vasculature measurement program- with automatic optic disc detection placement - measures peripheral vessels up to 2 disc diameters away from optic disc. Zones B and C are highlighted in the image. (Image from - SERI training manual).

The grading protocol developed in the Singapore Eye Research Centre was followed during the grading process (Appendix V) and required relatively extensive manual editing as it had been set for scoring of Asian eyes.

3.5. Quantitative retinal vessel traits

3.5.1. Vessel widths

The retinal vessel calibre measured from the colour fundus images was believed to measure the width of the streaming column of reflective erythrocytes (central red blood cell column), and does not include the surrounding plasma zone and the vessel wall, which are transparent to the fundus photography. This is based on the notion that erythrocytes flow in the centre of the retinal vessel, just like in other vessels, while the outer segment, that flows at a slower pace, consist mainly of plasma. However, one recent study found only very slight differences between width measurements from fundus images and fluorescein angiographs (Pakter et al. 2011). As fluorescein angiography is the method of choice for blood column visualisation, the authors of this study have concluded that the overall lumen (erythrocytes and plasma) is also measured by the fundus photography. Assessment of the retinal vessel's diameters from a digital colour fundus applies a micro-densitometric approach which uses intensity profiles of a grey-scale image. Grey scale digital images consist of many pixels, and each one has a spatial co-ordinate and an intensity value, also known as its grey value. Each pixel point in a line perpendicular to the retinal vessel will have various intensity levels, subsequently producing a distinct double Gaussian distribution curve against the background intensity of the surrounding retina (Patton et al.2005; Figure 3.6). This

double Gaussian curve is then used in the image analysis for the estimation of the vessel width, most commonly by estimating it at half height of the peak of the intensity profile of the Gaussian curve (half-height method). Image analysis software automatically generates a series of perpendicular lines to a retinal vessel in a predefined zone, information gathered by the tracking algorithm is then summarized using formulae to obtain values representing the average arteriolar and venular calibre of that particular eye.

a. b.

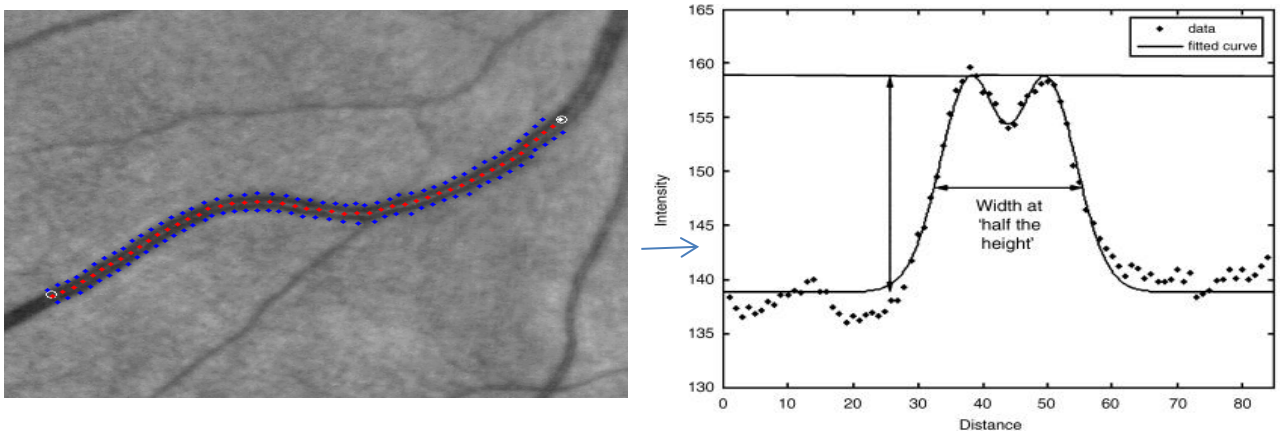


Figure 3.6. a. Example of the tracking procedure used by the software used for tracking the vessels (Orkney image) , b. Information gathered by the tracking algorithm is then used to calculate vessel widths (Patton et al. 2005).

The most commonly used measures used for describing arteriole and venule widths are CRAE (central retinal artery equivalent) and CRVE (central retinal vein equivalent). These measures were developed to standardise the way in which retinal vessel calibres are reported, and to summarize the average widths of vessels surrounding the optic disc. The CRAE measurement was devised by Parr and Spears (1974) and improved by Hubbard et al. (1999). The CRVE was developed by Hubbard et al. 1992 based on the Parr and Spears CRAE formula. Both of these formulae were derived by examining numerous vascular junctions in retinal images of young healthy normotensive

individuals. The relationships between the trunk vessel and its respective branches, for numerous vessels in a predefined zone around the optic disc, were calculated using a least root mean square deviation model that best fit the observed data. Hubbard et al. further developed the measure by combining the largest vessel with the smallest one, the second largest with the second smallest one, until all vessels are combined. Knudtson et al. (2003) developed a revised measure based on the six largest arterioles and venules passing through the predefined zone, making this measurement independent of the number of vessel measured in each image. They used the same iterative process of pairing the largest and the smallest vessels (there are only 5 iterations because only six vessels are being measured). In this method a "branching coefficient" is measured which gives the ratio of the branch and trunk vessels:

Branching coefficient= $(w^2_1+w^2_2)/W^2$

Where W is the width of the trunk vessel and w_1 and w_2 are the widths of the two branch vessels. They measured 187 arterial and 151 venular junctions in a sample of 44 young healthy individuals. By placing the values calculated from these images in the above formula, they calculated that:

CRAE: $Wa=0.88\sqrt{(w_1^2+w_2^2)}$ - arterioles

CRVE: $Wv=0.95\sqrt{(w_1^2+w_2^2)}$ - venules

By using the pairing method (combining smallest and largest vessel in each pairing), CRAE and CRVE are then calculated. In SIVA these measurements are done for the standard (0.5-1 OD diameter) and the extended (0.5-2 OD diameter) zones (the measures of vessel widths were not available in the VAMPIRE software). The computation of the retinal vascular calibre in the extended zone additionally includes the width of the first branch (if it is larger than 50% of the trunk width).

The most commonly used dimensionless measure is AVR, the arteriolar-venular ratio. AVR=CRAE/CRVE (smaller AVR signals relatively smaller arterial diameter or wider venous diameter). A limitation of the AVR measures is that arterioles and venules can have different responses to different pathological mechanisms, and therefore the independent use of CRAE and CRVE can possibly provide more information.

3.5.2. Bifurcation parameters

Branching angles

The bifurcation angle refers to the angle between the two daughter vessels, branching from the parent vessel. It provides information about the architecture of the vessel network, which determines the efficiency of the circulation. In a healthy state, the vascular network tends to conform to some fluid dynamic principals of the optimal structure in order to minimise the shear stress and work and in turn maximises efficiency. Murray (1926) calculated that if the blood flow across the vessel network is proportional to the cubed power of the vessel radius it achieves the most efficient circulation (Murray's law). In addition to this, an optimal angle between the two daughter vessels at a vascular junction was calculated to be 75˚ in order to minimize the principle work across the vascular network (Murray 1926).

When measuring branching angles, the software generates a skeleton from a binary (i.e. segmented) map of the vasculature and it locates the intersecting junction points of 3 vessels. A set number of points are extracted from each branch and a straight line fitted. For each junction, the 3 straight lines are used to find an improved junction point in a least squares manner. The improved junction point in conjunction with the previously

extracted vessel branch points are fitted with another straight line this time anchored at the new bifurcation point. These fits are used to calculate the branching angle.

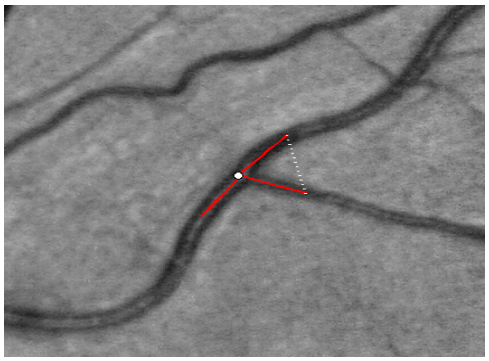


Figure 3.7. Bifurcations (red lines) are found by the vessel tracking algorithm.

Branching coefficient

The measurement of the branching coefficient is based on vessel widths between the parent and two daughter vessels. In addition to the branching angle, it is also an important factor determining retinal vascular architecture. And changes in these measurements have implications in the distribution of shear stress (defined as a stress which is applied parallel or tangential to a surface of a material) across the network. Theoretical predictions estimated its optimal value to be 1.26. Branching coefficients are used for the calculation of "revised" CRAE and CRVE measurement, a method developed by Knudtson et al.(2003)

3.5.3. Tortuosity

The tortuosity is a measure of the degree of curvature of a vessel. Degree of tortuosity of a vessel can be defined as the ratio between the distance a vessel travels from A to B and the shortest distance between the points A and B drawn by a straight line. It is a diagnostic parameter assessed by ophthalmologists on the basis of examples and experience; however no quantitative model is specified in clinical practice. Image analysis techniques implemented in the SIVA software measure two types of tortuosity: simple and curvature. Simple tortuosity is estimated as the difference between the lengths of the vessel segment measured by vessel tracking and the straight line length of the same segment divided by the straight line length. This is a simple measure and it does not distinguish the curve bowing from multiple points of inflection (Figure 8). The second measurement, curvature tortuosity, is defined as the integral of the curvature square along the path of the vessel, normalized by the total path length and provides a more accurate measure of the vessel's curvature (Figure 3.8).

a b

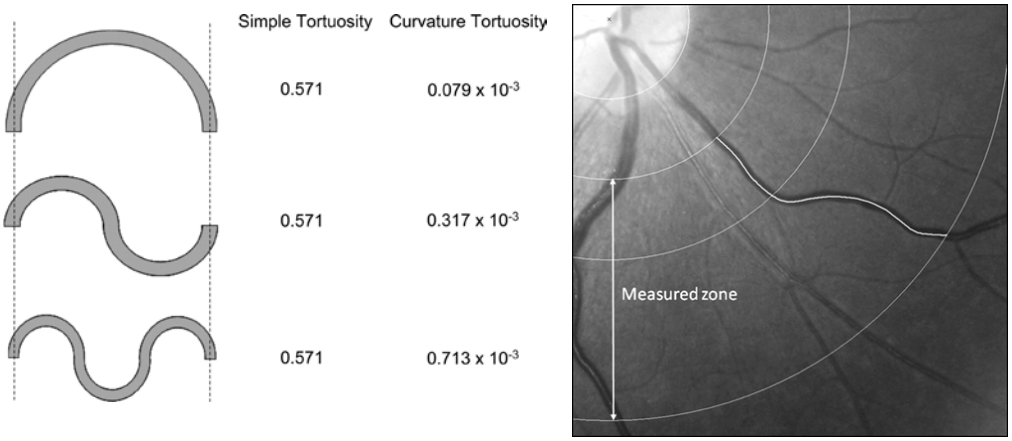


Figure 3.8. a) Difference between simple and curvature tortuosity measurements with respect to curving and inflection points (image from Witt et al. 2006). b) Tortuosity measurements are obtained from the extended zone (image from Koh et al. 2010).

3.5.4. Fractal analysis

Retinal vasculature is a complex branching structure and difficult to describe using only the simple Euclidian geometry, thus the individual measures described above do not provide a global measure of the vasculature as a whole. Fractal analysis denotes one type of analysis used to describe geometrical patterns that defy Euclidian concepts. Any type of fractal analysis is based on fractal dimensions, which are a measure of the geometric complexity of a spatial object (e.g. the vascular tree). Fractal dimensions describe how well a pattern fills two-dimensional space, described as a form of symmetry (Master, 2004). In retinal image analysis it attempts to describe the branching patterns of the vascular system, and is based on the concept of self-similarity of spatial geometrical patterns (fractals are self similar objects).

The first step in a fractal analysis is the determination of its fractal dimensions (D). There are several methods for determining fractal dimension; box counting, mass-radius relation and pair correlation function method. The most commonly used in retinal image analysis is the box counting method. In this method a binary image is covered with square boxes of the length L. The number of boxes covering the images is counted (N). The size of the box is then increased and the process repeated. A log-log plot of numbers (N) vs box length (L) is plotted and the slope of the plot calculated by least square regression. The slope of the plot is –D, and D is the box-counting fractal dimension. This is a monofractal analysis and it yields a single fractal dimension, which is a dimensionless quantity that measures the degree of branching complexity for the structure. However, retinal blood vessels may have different properties in different regions of the retina, and so different characteristics can be found depending on the location or scale of the measurement considered. A multifractal approach calculates multiple fractal dimensions from numerous (~1000) randomly chosen starting points on the vessel tree. Based on the fact that this measure takes into account the retinal vascular

system as a whole and at different scales, it could be expected that it will prove to be the best quantitative measure of overall vascular architecture. (MacGillivray et al. 2007).

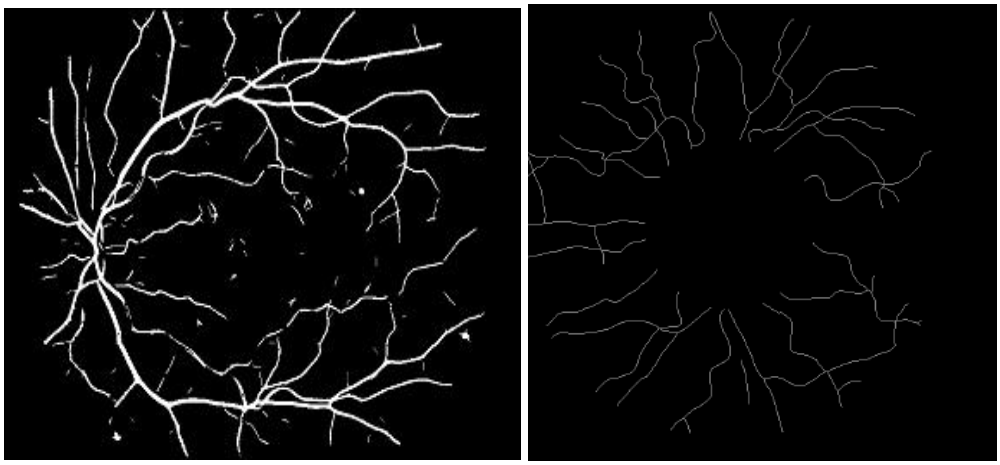


Figure 3.9. Fractal analysis areas measure by VAMPIRE(a) and SIVA (b) (using two different images)

The VAMPIRE software utilises both mono and multiracial analysis and captures the vasculature from the entire retinal fundus image (Figure 3.9. a). In contrast to SIVA, the fractal dimension is only measured in the extended zone (0.5-2 OD diameters; Figure 9b) and returns a monofractal value, using a curvature based segmentation (Cossato et al. 2010). On the other hand, it offers separate fractal measurement for arteriolar and venular networks separately, in addition to the measurement of the complete vasculature.

3.6. Quantitative traits derived from the optic disc and their assessment

Alongside retinal vessel traits, traits derived from the optic nerve or optic disc are most commonly analysed from the retinal fundus images. Three major traits were typically analysed from the fundus images are optic disc area, an optic cup area and a cup to disc ratio (CDR). These traits play a very important role in the diagnosis of glaucoma.

Images from the ORCADES participants were sent to Singapore to be automatically analysed by the AGLAYA software. This system automatically detects glaucoma from multiple image features of retinal images. As the optic disc region is very small compared to the rest of the fundus image, the first step is to extract the region of interest via pixel intensity analysis. This is followed by the segmentation of the optic disc using a variational level-set algorithm and boundary smoothing of the detected disc. The same procedure is repeated for the optic cup. Due to the decreased visibility of the boundary between the optic cup and the surrounding optic disc region and the vessels, this is more challenging. Therefore, a multi modal approach is used combining several different algorithms (level-set algorithm and colour intensity method). After extracting the disc and the cup are, CDR is calculated from the vertical extents of the cup and disc. Complete analysis of the ORCADES images was done conducted in Singapore and the results were sent in the Excel spreadsheet.

3.7. Measuring retinal vessel vasculature parameters

Retinal fundus images were obtained at a 45° angle using a high resolution digital camera (Canon CR-DGi non mydriatic retinal fundus camera). All images were stored in JPEG or TIFF format (high resolution digital images), depending on the study. Retinal photographs were centred between optic disc and macula. Images centred overly towards the macula, with optic disc on the side of the field, were not used in the further image analysis as they would not allow grading of the vascular network in the extended zone (if more than half of one quadrant of the extended zone was missing, the image was considered ungradable). Furthermore, only retinal image with at least four major arteries and venules present in the zone of grading were followed up. Images of very poor quality (e.g out of focus, overexposing of the lens) or with known pathologies

influencing image quality (e.g cataract, asteroid hyalosis) were not used in the further image analysis.

a. b. c.

Figure 3.10. Examples of images removed before the image analysis.an a- cataract; b-asteroid hyalosis; c-more than half quadrant missing

Measurements were obtained on one eye for each subject. If both images were of the same quality, the right eye was chosen, otherwise the eye with the better photographic quality was chosen. The analysis of the ORCADES and CROATIA-Korcula images was conducted by the me. The Lothian Birth Cohort (LBC) images were graded by the me and two other graders. Retinal images from ORCADES, CROATIA-Korcula and Lothian Birth Cohort included in the study were analysed by semi-automated computer image analysis software SIVA and VAMPIRE. VAMPIRE was used for branching parameter (ORCADES) and fractal analysis measurements (ORCADES, Korcula, LBC). Images graded by SIVA include measurements for vessel calibres, tortuosity, branching parameters and fractal dimensions. A detailed list is given in Table 3.1.

Retinal image analysis software	Vascular parameters measured
VAMPIRE	Branching angles Branching coefficient Monofractal analysis Multifractal analysis
SIVA	Vessel calibre (standard and extended zone) Simple tortuosity Curvature tortuosity Monofractal analysis (complete vasculature, arterial vasculature and venular vasculature in extended zone C) Branching angles Branching coefficient Length to width parameter

Table 3.1. Retinal vascular parameters measure by VAMPIRE and SIVA softwares.

In total 820 images from Orkney, 420 from Korcula and 397 LBC images were analysed by SIVA and VAMPIRE.

3.8. Validation of the branching angles analysis with VAMPIRE

Prior to the use of a computerized image software analysis in a large study, the method/sytem requires to be tested and validated. The SIVA licence was purchased from the Singapore Eye Research Institute and tested and validated by them, with a very detailed grading protocol. Extensive onsite training of the author was also required, which was necessary for comprehensive and accurate usage of the program. VAMPIRE was developed by the University of Edinburgh in collaboration with University of Dundee and, at the start of this PhD study, was still under development. These developments to improve software available in Edinburgh meant that measurements of bifurcation branching angles had to be tested before use in the ORCADES study. The agreement between human and computer measurements was assessed using 20 retinal images. On each image we aimed to measure 5 venule and 5 arteriole branching points. Due to the different retinal patterns and varying image quality, it was not always possible to identify these and so the number of measures for each image varied between 2 and 5. Images were manually analyzed by two different graders using the angle measurement tool in ImageJ (http://rsbweb.nih.gov/ij/) and compared. Subsequently, human measurements were compared to the results obtained from a semi-automated computer method. Results are shown in Figure 3.11.

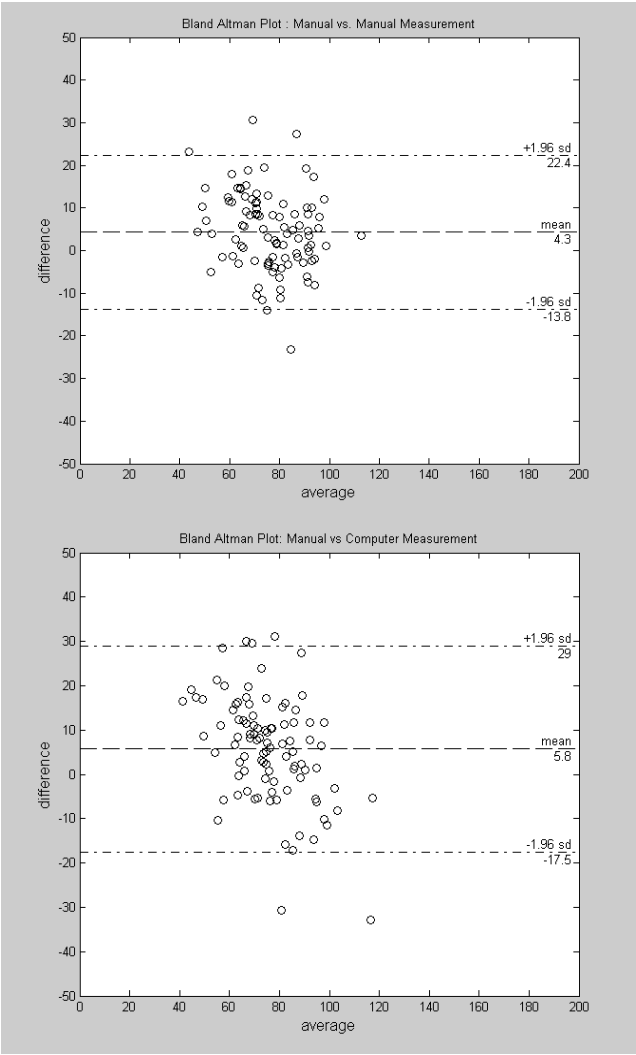


Figure 3.11. Bland-Altman plots of human vs human and human vs computer measurements

A Bland-Altman plot describes the agreement between two quantitative sets of measurements. In the plot the difference between the paired sets is plotted against the average of the two measurements. The agreement between two quantitative methods is considered to be good if 95% confidence intervals include 95% of differences between the two measurement methods. The first plot in Figure 11 shows the mean difference between two human raters was 4.3° with 95% Confidence Interval (CI) of -13.9° to 22.4° while the second shows the mean difference between human and computer was 5.8° with 95% CI of -17.5° to 29°. Thus we were confident that the computational

technique is returning valid branching angle measurements, and it could be used on the ORCADES images.

The same procedure was repeated for branching coefficient measurements, on the same branching points. Here, only human versus computer measurements were possible because the second grader was unavailable to repeat the process. Results are shown in Figure 12.

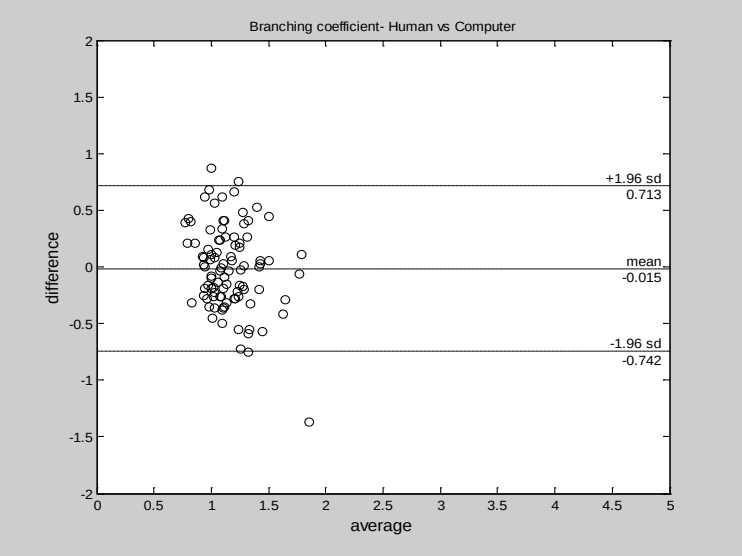


Figure 3.12. Bland-Altman plots human vs computer measurements for branching coefficients

After the validation of the technique for measuring branching angles and coefficients, VAMPIRE was used to analyse branching points of the ORCADES images. Before starting the measurements, a measurement protocol was established and followed when each image was analysed. The protocol is given in the appendix (Appendix IV). Fractal analysis of the retinal vascular network with VAMPIRE software was validated prior to the start of this study.

3.9. Selection of the retinal fundus image phenotype for the genetic analysis

Both left and right eye fundus images were obtained for each subject included in the study. However due to the laborious process involved (average of 20 minutes spent on one image) in the image analysis with semi-automated software, only one eye could be measured for each study subject. A good correlation between left and right eye phenotypes is needed in order to ensure that a randomly selected eye represent an overall eye phenotype. Correlation between the two eyes is expected to be high due to the sharing of the same environment and genes. This has been reported for human limb bones (Choi et al. 1997), retinal vessel diameters (Leung et al. 2003, Ikram et al. 2006) and retinal nerve fibre layer measurements (Budenz 2008). In addition, in their study of small vessel disease and retinal vessel diameters Ikram et al. (2006) did not find statistical significant differences between measured diameters in left and right eyes in a random subsample of 100 individuals. Investigating the correlations between the left and the right eye measures was therefore the first step in the analysis of the traits available.

First, branching parameters estimated by VAMPIRE program were tested. Both bifurcation angles and coefficient were tested in a set of 40 eyes, and both median and mean values of the total measurements for the eye (between 3 and 6) with the results given below Table 3.2.

Vessel type	**arterioles**		**venules**	
	mean	median	mean	median
branching angles	0.31	0.26	0.44	0.40
bifurcation coefficients	0.04	-0.16	-0.04	0.08

Table 3.2. Pearson's Correlation coefficients between left and right eye for mean and median values of branching angles and branching coefficients measured by VAMPIRE software (of 3 to 6 angles, depending on the availability).

Correlation between the two eyes of the same individuals were low for branching angles, and very poor for branching coefficients, signalling a problem either in the measurement procedure or suggesting a poor correlation of these traits between both eyes. One likely reason is the inability of the software to detect all the branching points on the image, with a consequent bias in the selection of the angles chosen for the derivation of the final trait (mean or median of maximally 6 bifurcation points). This is introducing bias due to the high variance in the number of measured bifurcation angles within each eye. Further, the consensus is that angles of the first, second and third branching point from the optic disc are used for the measurements. They are not always randomly distributed across the image and are often masked by the neighbouring vessels.

Monofractal and multifractal dimensions of the retinal vasculature were also tested for left and right eye correlation. In total, 130 images were tested. Correlation values were higher compared to the branching parameters measured by VAMPIRE. Left-right eye correlations were 0.78 for monofractal dimensions and 0.50 for multifractal dimensions. These results are more in line with data published on retinal vessel calibres (Leung et al. 2003, Joachim et al. 2012) increasing confidence in using these measures for further analyses.

Next, correlations were assessed for the parameters obtained from the retinal image analysis with SIVA. 30 ORCADES subjects and their left and right retinal images were used. Subjects were chosen if they had good quality images for both eyes. Results are given in Table 3.4.

Vascular trait	Correlation coefficient(p-value)
CRAE zoneB	**0.68 ($5x10^{-5}$)**
CRVE zoneB	**0.78($7.5x10^{-7}$)**
AVR zoneB	**0.56(0.002)**
CRAE zoneC	**0.86($2.4x10^{-9}$)**
CRVE zone C	**0.78($4.6x10^{-7}$)**
AVR zoneC	**0.78($6.6x10^{-7}$)**
Fractal Dimensions zone C –complete vasculature	**0.66($1x10^{-4}$)**
Fractal Dimensions zone C –arteriol vasculature	0.24(0.139)
Fractal Dimensions zone C –venular vasculature	**0.59(0.001)**
Simple tortuosity arteries	**0.69($3x10^{-5}$)**
Curvature tortuosity arteries	**0.90($1.8x10^{-11}$)**
Branching coefficient – arteries	-0.48(0.001)
Branching angles – arteries	0.34(0.66)
Angle of first daughter vessel – arteries	0.08(0.653)
Angle of second daughter vessel – arteries	0.16(0.41)
LDR- length to diameter ratio –arteries	0.055(0.778)
Simple tortuosity venules	**0.59(0.001)**
Curvature tortuosity venules	**0.66($1x10^{-4}$)**
Branching coefficient – venules	-0.03(0.885)
Branching angles – venules	-0.11(0.563)
Angle of first daughter vessel – venules	0.06(0.76)
Angle of second daughter vessel – venules	0.17(0.37)
LDR- length to diameter ratio – venules	0.04(0.842)
Simple tortuosity – complete vasculature	**0.70($2.5x10^{-5}$)**
Curvature tortuosity – complete vasculature	**0.82($4.5x10^{-8}$)**
Branching coefficient - complete vasculature	-0.43(0.18)
Branching angles - complete vasculature	0.14(0.476)
Angle of first daughter vessel - complete vasculature	0.20(0.292)
Angle of second daughter vessel - complete vasculature	0.22(0.253)
LDR- length to diameter ratio - complete vasculature	0.21(0.286)

Table 3.4. Pearson's correlation and the significance p-values are given for parameters measure by SIVA. Measure with $r^2>0.5$ are highlighted in bold.

Retinal vascular network branching parameters presented generally low correlation values with both programmes for retinal vessel image analysis. Therefore it was decided not to include them in the subsequent analysis as there was not enough evidence that these phenotypes measure in one eye would accurately represent the phenotype. Remaining traits measure by VAMPIRE and SIVA showed an inter-eye correlation of 50% and more, considered a cut-off for phenotype inclusion.

Finally, traits pertaining to the optic nerve were assessed. Left and right retinal images (of good quality for both eyes) of 100 subjects were sent for the analysis to Singapore. Correlation values were low, compared to published data of the inter-eye correlation of the retinal nerve fiber layer (Budenz 2008); 0.22 for the optic disc are, 0.26 for the optic cup are and 0.18 for cup to disc ratio. A proportion of ORCADES subjects also underwent Heidelberg retinal tomography (HRT) measurements of the optic nerve layer. 135 subjects had CDR results from both of these measurements, however the correlation between these results was also extremely poor (r^2=-0.0089), suggesting a larger issue with software measurements of optic disc parameters from retinal fundus images. It was therefore decided not to pursue these parameters in the subsequent analysis.

In summary, quantitative traits derived from the retinal fundus images selected for the subsequent genetic and epidemiological analysis were: vessel calibre traits, tortuosity traits and fractal analysis traits.

3.10. Grader variability of the retinal vessel measurements

Estimation of the reliability of a quantitative measurement is essential for its use in future research. Reliability enables us to assess the consistency (repeatability) of our measurement technique. In the case of semi-automated computer software where human input accounts for a substantial part of the measurement procedure, evaluation of reliability is essential. Principal measure of reliability is the intraclass correlation coefficient (ICC). It represents correlation when variables are different measurements of the same construct (agreements between two or more individuals measuring the same set of traits). The general purpose of ICC is to determine the proportion of variation in a

measure that is due to being a member of a particular group. The benefit of ICC is that it is adjusted for the effects of the scale of measurements as opposed to the correlation coefficient, and that it can represent agreements between more than two raters. First, a two-way analysis of variance is carried out, and if the between column mean sum squares (between raters) is not significantly greater than the residual error, then concordance can be concluded. The intraclass correlation coefficient provides a measure of agreement or concordance between all the methods (0-0.2 indicates poor agreement; 0.3-0.4 indicates fair agreement; 0.5-0.6 indicates moderate agreement; 0.7-0.8 indicates strong agreement; and >0.8 indicates almost perfect agreement).

Several large epidemiological studies reported moderate to strong reliability for the retinal vessel calibre measures for both inter and intra-grader reliability measurements (Hubbard et al. 1999, Liew et al. 2008, Wang et al. 2004, Ikram et al 2004). Despite the different reliability measurements used by these studies, (Pearsons's correlation coeficient, ICC, kappa), high repeatability values were consistently reported. High reliability was also reported for the measurements of fractal dimensions (Doubal et al. 2010, Lim et al. 2009, Sng et al. 2010) and tortuosity (Koch et al. 2010, Sasongko et al. 2010).

Intra-grader and inter-grader reliability for all the selected traits were evaluated. A set of 60 images from a Singapore Research Eye Study (SERI) were used to calculate inter-grader reliability (i.e two different graders analysing the same sets of images), and for intra-observer reliability assessment (i.e. the same grader at different times) 60 images from the ORCADES study were used. SERI images were previously analysed by an experience grader (L.J.). Results obtained by these replicated grading exercises are presented in Table 3.5. They confirm previously published results of high reliability of quantitative retinal vessel traits. Due to the higher level of automatisation in the grading process, reliability was higher compared to the previously published results (older retinal image software required more human input).

Traits	Intraclass correlation coefficient	
	Intra-grader repeatability	Inter-grader repeatability
CRAE zone B	0.95	0.90
CRVE zone B	0.97	0.95
AVR zone B	0.98	0.90
CRAE zone C	0.99	0.95
CRVE zone C	0.98	0.98
AVR zone C	0.98	0.93
Fractal dimension zone C	0.97	0.92
Arterial curvature tortuosity	0.99	0.88
Venular curvature tortuosity	0.97	0.94

Table 3.5. Reliability estimates for retinal vascular traits analysed by SIVA software. Both inter (right column) and intra grader(left column) ICC measurements are reported.

3.11. Discussion

This chapter has shown that while the recent improvements in retinal image analysis software has led to a quicker and more reliable analysis of the traits derived for the retinal fundus images, limitations remain. Traits available from the two softwares, VAMPIRE and SIVA, were therefore evaluated for further use. One selection process was the assessment of the phenotypic stability of each trait, by estimating the correlation between the two eyes in a subset of our study subjects. This was essential since only one eye was selected to represent a retinal vessel phenotype for each study subject, and even if both eyes were available traits with the high variation between left and right eyes are not likely to have significant genetic basis. Some of the traits (branching angles and derived group of traits) showed particularly low left to right eye correlation with both image analysis programs used. A more in depth analysis of the reasons for this was beyond the scope of this thesis, and the causes are therefore not possible to deduce.

However, since the area of the retina analyzed by the software is dependent on the area of the retina photographed by the camera (this area often differs between two eyes), number and type of the angles analyzed in the left and the right eye may not be the same.

The standard practice is to measure first, second and third degree angles (numbering is given depending on the distance of the bifurcation from the main retinal artery or vein), and the total number of each type measured is often different between the two eyes (left/right). Furthermore the total number of the measured bifurcation points is often different between the two eyes, adding to the inconsistency. Another potential reason could be the inability of the algorithm to correctly and consistently detect the branching angles. The inability of the VAMPIRE software to automatically detect all of the branching points, and the limitation of the SIVA grading protocol which only appoints tracking of the vessels that the program automatically detects (with a few exceptions) are most likely the main reasons for the lack of phenotypic consistency of the branching angle measurements.

It is crucial that any trait used in the following epidemiological or genetic analysis is accurate and reproducible. Estimating the reliability of the computer assisted methods used in the previous step was therefore the next phase of the selection process. As mentioned previously, the estimation of reliability is essential when human input is substantial, which was the case in our image analysis. Both intra and inter grader reliability of the measurements chosen in the previous step was very high, which was not surprising due to the high level of automatisation of the SIVA software and the removal of the traits with poor left eye-right eye correlations. However it is important to acknowledge several factors that can influence the accuracy of the retinal vessel measurements despite good correlations, since in some cases it is not possible to fully account or correct for them. All images taken by the fundus camera go through the image magnification process, which in turn affects the precision of the absolute value of the vessel caliber measurements (CRAE, CRVE). The scale depends on the distance from the camera to the eye and on the eye's refractive error. To partially correct for these magnification differences a calibration factor was used (ratio of micrometers per pixel). The calibration factor is obtained by measuring the average vertical height of the optic disc (in pixels) in 10% of the retinal images in a given population, and dividing the assumed average of the disc diameter (1800μm, average disc diameter measured for the

Caucasian subjects by Varma et al. 1994) with the population's average value. The relationship between refractive error and the axial length of the eye with retinal vascular caliber has been explored in several studies. Associations between smaller retinal vessel diameters and longer axial length and more myopic refraction were reported (Patton et al. 2005, Wong et al. 2004, Lim et al. 2011).

However, these associations have been explained as both the artifacts of the magnification effect and as the result of a true biological process. Axial length and the refractive error were found not be associated with vascular caliber after the correction for the refractive error using the Bengstsson formula (Bengstsson et al. 1992) in the Singapore myopia study (Cheung et al. 2007) and in the Blue Mountain Eye Study (Wong et al 2004). On the other hand, underlying biological relationship is supported by the studies of decreased ocular pulse and retinal blood flow with myopia (Perkins 1981, Benavente et al. 2010). Tortuosity (a dimensionless retinal vessel trait) was also found to be correlated with axial length and spherical equivalent (Lim et al. 2011). However it is important to point out that the correction for refraction had no impact on the association between blood pressure and retinal vessel diameters (Wong et al. 2004, Tan et al. 2004), suggesting that the overall impact of axial length and refraction error on the retinal vessel morphology is not large enough to skew associations with systemic disease. Another issue concerning the retinal vessel width measurements is its variability with regards to the cardiac cycle (Knudtson et al. 2004), degree of nervous stimulation (Lanigan et al. 1988) and the pigmentation of the fundus. The variation caused by the cardiac pulse has been shown to be small and random, and it appears not to bias known associations (Sun et al. 2009). Several studies reported difference in the vessel calibre between ethnicities (Rotchinina et al. 2008, Cheung et al. 2007) suggesting that retinal pigmentation could introduce a source of bias in studies with subjects with different ethnic backgrounds and should be accounted for. Finally, any study that attempts to quantify retinal vascular morphology should take into account factors directly affecting retinal microcirculation such as age, sex, BMI, blood pressure, diabetes, etc. However

the difficulty of accounting for all possible systemic, ocular and environmental confounding effects has to be acknowledged.

In summary, this part of the study provided evidence of good correlation between eyes for the measurements of several retinal vessel traits (diameters, tortuosity and fractal dimensions), and excellent reliability for all measurement selected for further analysis. The evidence from the published data and those obtained in this study allowed selection of retinal vascular morphological traits best suited for further genetic and epidemiological investigations as less subject to grader variability and measurement errors.

Chapter 4. Heritability and genetic correlation of retinal vessel traits

4.1. Introduction

This chapter evaluates the genetic component of the retinal vessel traits. Estimation of the extent to which genetic variation contributes to these quantitative retinal vessel traits' variances and the amount of genetic correlation between these traits is the main focus. A few studies so far estimated the heritability of retinal vessel width parameters, and the majority of them were derived from twin studies, which tend to overestimate the actual heritability (due to difficulty of teasing out the shared environmental component).

All retinal vessel traits used in the study are essentially describing the architecture of the retinal vasculature in different ways (each quantifying one of the aspects of the vessel morphology). In order to investigate the degree to which these various retinal vessel traits are influenced by the same genes, genetic correlations between the traits were tested. Some reports of correlation between calibre measurements are published, but no study systematically estimated genetic correlation between various retinal vascular traits. Thus, estimation of heritabilities and correlation between retinal vascular is the first step in elucidating the influence of genetic components on retinal vasculature.

4.2. Heritability estimates of the retinal vessel traits

4.2.1. Definition and measurements of heritability

The variance of a quantitative trait can be decomposed into several components. In its simple form, phenotypic variance is decomposed into a genetic and environmental components.

$$\mathrm{V(phenotypic)} = V\mathrm{(genetic)} + V\mathrm{(environmental)}$$

The broad sense heritability is then defined as the ratio of genetic variance over the overall phenotypic variance; i.e it measures the proportion of the total variance of the trait accounted by genetic effects.

$$H_2 = \frac{\mathrm{V(genetic)}}{\mathrm{V(phenotypic)}}$$

However, the genetic variation contains contribution from additive and dominant effects, and those can be further modulated by epistasis, gene by environment interactions, maternal and paternal effects. Only the additive effects are transmitted to descendants and the fraction of the phenotypic variance attributable to the additive genetic variance is defined as the narrow sense heritability. It is the one usually referred to when measuring heritabilities of quantitative traits measured in human populations.

$$h_2 = \frac{\mathrm{V(additive\ genetic\ effect)}}{\mathrm{V(phenotypic)}}$$

Estimates of heritability in human populations are based on correlation between phenotypic measures of related individuals and their expected sharing of genetic material inherited by descent, and therefore a particular trait is considered to be heritable

if the similarity between relatives arises from shared genotypes. The most commonly used study designs have been the twin study and parent-offspring regression, which use data on the observed and expected resemblance between relatives to estimate heritabilities. However these estimates rely on assumptions that are often violated (e.g shared environment component assumed to be the same in mono and dizygotic twins, or environmental component similar in parent and offspring generations) and estimating heritabilities from a mix of relationships has been advocated (Visscher et al. 2008).

In the case of population-based studies, phenotypic measures are available from individuals with a mixture of relationships and multiple generations, estimates of additive genetic variance are calculated by fitting the so-called "animal model": a linear mixed-effects model (one containing both fixed (e.g. SNP) and random (polygenic) effects). This polygenic model of inheritance is considered the model of choice in studying heritability of quantitative traits, in settings where phenotypic values can be correlated between pairs of individuals of multiple different types of relationships (Almasy and Blangero 1998). The model describes the vector of an observed quantitative trait as:

$$Y = \mu+\Sigma\beta jCj+A+e$$

where μ is the intercept, Cj the design vector for the jth covariate, Bj its associated effect (coefficient of regression), *A* is the matrix of polygenic additive contribution, and *e* is the matrix of random residual contribution (Lynch and Walsh, 1998). A, the additive polygenic component, describes the contribution from multiple independently segregating genes all having a small additive effect onto the trait (infinitesimal model) and its distribution in a pedigree can be described as multivariate normal with variance-covariance matrix proportional to the relationship matrix Φ and the variance of additive effects (σ_A^2) : $A \sim MVN(0,\Phi\sigma_A^2)$. It is assumed that residuals, from each of the individuals in a pedigree, are independent between the pedigree members, and the joint distribution of residuals in the pedigree can be modelled using multivariate normal

distribution with variance-covariance matrix proportional to the identity matrix I and residual variance (σ_e^2); $e \sim MVN(0, I\sigma_e^2)$.

Covariances between individuals are used to estimate the additive genetic component given the coefficients of coancestry based on the family pedigree or genetic markers:

$$Cov(yi,yj) \sim 2\ \phi ij \sigma_A^2$$

ϕij expected kinship (co-ancestry) coefficient between individual yi and yj

Maximum likelihood and restricted maximum likelihood methods are used to estimate fixed effects and variance components in the mixed model and for significance testing. Likelihood of obtaining the observed data given the family structure and the scale of parameters to be estimated (the genetic and environmental variance components) is calculated. The model then finds the values of these unknown parameters which maximises the likelihood function. Estimated parameters are used to calculate heritability. Heritability estimated in this manner takes into account all the informative relationships within the population (e.g. parent offspring, full-sib, half-sib, avuncular, first cousin and the more distant ones), which in turn improves the power of the estimates as all the data from all informative relationships are used. In addition this model allows the estimation of different components of environmental variance. Isolated populations represent populations with large complex pedigrees and with relatively stable and uniform diet, climate and living conditions. However, reduction of the genetic diversity and more uniform shared environmental influences may give estimates that are not representative of other populations. Two such isolated populations (Orkney and Korcula) were used in this study.

4.2.2. Reported heritability measurements from the literature

Estimates of the heritability of retinal vessel traits have been investigated in only a few studies to date. All of the studies reported substantial genetic effect on these traits, however most of them applied either a twin or parent-offspring study design known for possible biased estimates. No studies to my knowledge have reported data from large population-based samples.

Retinal vessel widths were the most extensively studied trait. Three twin studies (Taarnhoj et al. 2006, Sun et al. 2009, Fahy et al. 2011) reported heritabilities between 0.56-0.70 for CRAE (central retinal arteriolar equivalent) and 0.62-0.83 for CRVE (central retinal venular equivalent). One familial aggregation study (Lee et al. 2004) estimated heritability by doubling the parent–child correlations and reported lower but still substantial heritability estimates; 0.48 for CRAE and 0.54 for CRVE. Finally, one smaller population based study (413 participants from 70 families from a Flemish population; Liu et al. 2013) reported lower heritability estimates of 0.21 for CRAE and 0.34 for CRVE.

A Danish study (Taarnhooj et al. 2008) estimated heritability for retinal arterial tortuosity at 0.82 in a sample of 57 monozygotic and 52 dizygotic same-sex healthy twin pairs. Estimates of other parameters of retinal vessel architecture have not been published to my knowledge.

4.2.3. Material and Methods - heritability estimates of retinal vessel traits in the population isolates of Orkney and Korcula

4.2.3.1. Study populations

The retinal vascular traits described previously were chosen to estimate heritability in two isolated populations from Scotland (Orkney islands- ORCADES study) and Croatia (Korcula island).

The Orkney islands -ORCADES

Orkney is an archipelago in northern Scotland of around seventy islands (of which 17 are inhabited) situated on average ~30 kilometres north of the coast of Caithness. The total populationis ~20,000 people. The advantage of the Orkney islands is their geographic isolation and the fact that family trees can be tracked back as many as eight generations. The North Isles of Orkney, the focus of this study, consist of ten inhabited islands: Westray, Papa Westray, North Ronaldsay, Rousay, Egilsay, Wyre, Stronsay, Sanday, Eday and Shapinsay.

The Orkney Complex Disease Study (ORCADES) is an ongoing family-based study which aims to identify common and rare genetic variants influencing the risk factors for common complex disease which are important in Scotland. The ORCADES study collected data for 6 years and has a resource of ~2100 richly phenotyped and densely genotyped subjects (46% male; mean age at recruitment 53 years, range 18 - 100), including data on numerous quantitative traits associated with cardiovascular, metabolic and other diseases, detailed family history data and 6 generation pedigrees. Illumina genome-wide scans were available for 889 participants (after the merge of two sets of individulas genotyped on different platforms described in Chapter 6) at the time of this study.

Good quality retinal images were available for 811 genotyped individuals. All of the images from Orkney were graded by one grader (M.K.).

KORCULA

The island of Korcula is located in the southern Croatian coastal region of Dalmatia. According to the 2001 Census, the total island population was 16,182 inhabitants, distributed in five larger districts: Korcula, Lumbarda, Smokvica, Blatoand Vela luka. The target population comprised inhabitants of the Eastern part of the island (6,544 registered inhabitants). The study aimed to recruit healthy volunteers from the town of Korčula and villages of Lumbarda, Žrnovo and Račišće. A total of 944 subjects were included in this study and most underwent complete eye examination. Family history data and pedigrees were not available for the recruited subjects. Data on numerous quantitative traits associated with cardiovascular, metabolic and other diseases was also collected as a part of the study.

The eye examination included fundus image photography. However, only 421 images were graded by the SIVA and the VAMPIRE software. Most of the ungradable images were rejected as they were excessively centered towards the macula, with optic disc on the side of the field. These images do not allow grading of the vascular network in the extended zone C. All the images were grade by one grader (M.K).

4.2.3.2. Genotyping and quality control

DNA was extracted from peripheral blood. Extraction was done by Nucleon kits by WTCRF (Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh) for ORCADES and at the Medical School, University of Split for Korcula participants. Participants were genotyped using a dense SNP arrays; a mix of HumanHap 300v2 and 370CNV-Quad for ORCADES (data integration is described in chapter 6),

370CNV-Quad for Korčula, following the manufacturer's standard recommendations. Genotypes were determined using Illumina BeadStudio software.

Genotyping rates for both SNPs and individuals are checked using the GenABEL package implemented in R. Individuals with less than 97% genotyping rate were removed from further analysis (to ensure poor quality data are not used). SNPs with call rate below 98%, minor allele frequency (MAF) below 2% and *P*-value for Fisher's exact test of Hardy–Weinberg equilibrium (HWE) below 10^{-10} are also removed to remove low-quality SNPs. In addition, samples with excess autosomal heterozygosity or gender inconsistency (based on the sex chromosomes genotypes) are removed due to possible sample mix-up. Ethnic outliers are removed based on principal components analysis of genotypic data. After this quality control step, the number of individuals available with ocular measures and genotypes was 811 for Orcades and 387 for Korcula

4.2.3.3. Statistical analysis – heritability estimates

Heritability estimates were calculated using the GenABEL package in R (Aulchenko et al. 2007) which implements a maximum likelihood method for the animal model parameters' estimates. The first step of the analysis is the calculation of a realised kinship matrix from the genome-wide data to account for covariances between individual genomes, using the gkin function of the statistical package GenABEL. This way of estimating the relationship between individuals is superior to the one derived from classic pedigrees as this solution does not rely on the completeness and quality of pedigrees. Furthermore pedigree-based kinships provide only the expected average amount of material shared IBD between pairs of individual genomes, whereas genomic-based kinship measures the realised sharing and is expected to lead to better estimates of polygenic model parameters. Estimator of the kinship matrix can be obtained by

calculating the kinship coefficients between individuals i and j (Astle and Balding 2009):

$$\hat{f}_{ij} = \frac{1}{L}\sum_{l=1}^{L}\frac{(g_{l,i}-p_l)(g_{l,j}-p_l)}{p_l(1-p_l)}$$

L - the number of loci,

p_l - the allelic frequency at l-th locus

$g_{l,j}$ - the genotype of j-th person at the l-th locus, coded as 0, 1/2, and 1, corresponding to the homozygous, heterozygous, and other type of homozygous genotype

The kinship matrix was calculated with the following GenABEL argument:

gkin= ibs(dfq[,autosomal(dfq)],w="freq")

The next step is fitting a polygenic model using a maximum likelihood function. GenABEL implements the method developed by E. A. Thompson group (Thompson and Shaw 1992).

polygenic(trait~covariates, kinship.matrix, data, trait.type="gaussian")

Additive genetic and residual variance components that maximise the polygenic model are estimated and used to calculate narrow sense heritability. Maximal likelihood estimates (MLEs) are given in the GenABEL output (mean, estimates of effect size for covariates, heritability, and polygenic + residual variance estimates) as well as the value of twice negative maximum log-likelihood.

The significance of the explanatory covariates was tested via a likelihood ratio test (LRT). The LRT is used to compare the fit of two models; it compares obtained likelihood for the full model (model with the covariates) to the likelihood of the nested model (the covariates are constrained to zero). The test statistic is twice the difference in these log-likelihoods:

$$D = -2\ln\left(\frac{likelihood\ for\ null\ model}{likelihood\ for\ alternative\ model}\right)$$

The probability distribution of the test statistic is approximately a chi-squared distribution with degrees of freedom equal to the difference in the number for parameters included in two models. Significance of the heritability estimates was tested by comparing the full and the null model of heritability using the following option in GenABEL:

herritability<-polygenic(trait~covariates, kinship.matrix, data, trait.type="gaussian")

heritability_nullmodel<-polygenic(trait~covariates, kinship.matrix, data, trait.type="gaussian", fixh2=0)

and by comparing the likelihood estimates of those two model, as described above.

Goodness of fit of a polygenic model was tested using the Akaike information criterion (AIC);

$$2k-2\ln(L)$$

where k is the numbers of parameters in the model and L the maximum likelihood of the model (Akaike 1974) . This criterion introduces a penalty for extra covariates used. Each extra covariate included in the model was tested this way and the optimal model for each trait was chosen as the one with the lowest AIC value.

Normality of the distribution for each trait was tested using the Kolmogorov-Smirnov normality test (lillie.test) implemented in nortest package in R (Gross 2012), in addition to the graphical representation of the distribution of each traits. Traits for which the distribution deviated from normality were transformed using quantile normalisation with

the rank transformation function of GenABEL (Aulchenko et al. 2009). Interquartile range was used to identify and remove outliers. An outlier was defined as any value that is bellow below $Q_1 - 1.5 \times IQR$ or above $Q_3 + 1.5 \times IQR$. IQR is interquartile range and is defined as the difference between upper (Q3) and lower (Q1) quartiles.

4.2.4. Heritability estimates - Results

4.2.4.1. Descriptive statistics

The descriptive statistics for the 17 retinal vessel traits measured in the participants of Orkney and Korcula (for which both quality-controlled phenotypic and genotypic data were available) are displayed in Table 4.1. In total 811 (751 – for fractal dimension measures by VAMPIRE) individuals from Orkney and 387 (328 - fractal dimension measures by VAMPIRE) from Korcula were used in the analysis. 31.4% tested subjects were male in Korcula, and 46.1% in Orkney. Measures were comparable between the two populations, however measurements of vascular calibres in zone B were lower to those previously published (CRAE mean=136.1-168.8, CRVE mean=196.5-253.0). This can be explained by the average disc diameter used in the calculation of the conversion factor from pixels to μm. In the present study I used 1800 microns as the standard disc diameter (following the SIVA protocol), while most of the published study use 1850 microns. Several studies reported monofractal dimension of the retinal vessel network, and results from Orkney and Korcula are in line with the published data (mean=1.42, SD=0.02 - Doubal et al. 2010, mean=1.41, SD=0.046 - Cheung et al. 2012, mean=1.44, SD=0.024 - Liew et al. 2011). Doubal et al. 2010 reported values for multifractal dimensions (mean=1.67, SD=0.03), which are again in line with the results reported in this study. Published means for retinal vessel tortuosity exist, however the methodology

of tortuosity measurement in these studies is different and the scales are not comparable to the results obtained from SIVA.

Orkney	N	Minimum	Maximum	Mean	Std. DeviationStd. Error	Kolmogorov-Smirnov test p-value
Big6craeB	811	79.90	166.92	127.60	0.50	0.88
Big6crveB	811	125.29	236.54	178.22	15.44	0.10
Big6avrB	811	0.55	0.98	0.72	11.61	0.009
Big6crae	811	79.41	168.20	128.33	16.13	0.64
Big6crve	811	131.05	232.55	180.37	0.06	0.91
Big6avr	811	0.54	0.912	0.71	11.04	0.03
FractalDimC	811	1.19	1.51	1.40	15.37	0.001
FractalDimCv	811	0.99	1.30	1.17	0.05	0.04
sTORTa	811	1.05	1.23	1.10	0.045	$<2.2x10^{-16}$
cTORTa	811	$3x10^{-5}$	0.00015	$6.63E^{-05}$	0.05	$<2.2x10^{-16}$
sTORTv	811	1.06	1.24	1.10	0.05	$1.1x10^{-14}$
cTORTv	811	$3.8x10^{-5}$	0.0002	$7.33x10^{-5}$	0.03	$<2.2x10^{-16}$
sTORTt	811	1.06	1.20	1.1	0.00002	$5.6x10^{-11}$
cTORTt	811	$3.9x10^{-5}$	0.00015	$6.97x10^{-5}$	0.02	$<2.2x10^{-16}$
Dbox_scel	751	1.30	1.46	1.40	0.00002	$2.88x10^{-6}$
D0_scel	751	1.54	1.7	1.64	0.02	$4.23x10^{-7}$

Korcula	N	Minimum	Maximum	Mean	Std. Deviation	Kologorov-Smirnov test p-value
Big6craeB	387	90.15	188.93	126.34	12.51	0.11
Big6crveB	387	121.8	234.00	177.47	0.47	0.72
Big6avrB	387	0.52	0.93	0.71	12.94	0.56
Big6crae	387	92.91	186.57	126.60	17.35	0.054
Big6crve	387	125.99	231.58	179.73	0.06	0.83
Big6avr	387	0.56	0.93	0.71	12.50	0.05
FractalDimC	387	1.22	1.47	1.37	16.97	0.36
FractalDimCv	387	0.99	1.26	1.15	0.05	0.25
sTORTa	387	1.06	1.24	1.10	0.06	$<2.2x10^{-16}$
cTORTa	387	0.000036	0.00019	$7.4x10^{-05}$	0.05	$<2.2x10^{-16}$
sTORTv	387	1.06	1.37	1.098	0.03	$<2.2x10^{-16}$
cTORTv	387	0.00004	0.00032	$8.1x10^{-05}$	0.00002	$<2.2x10^{-16}$
sTORTt	387	1.06	1.23	1.10	0.03	$1.65x10^{-9}$
cTORTt	387	0.000045	0.00019	7.7E-05	0.00002	$3.44x10^{-10}$
Dbox_scel	328	1.26	1.41	1.33	0.03	0.05
D0_scel	328	1.51	1.76	1.64	0.04	0.60

Table 4.1. Descriptive statistics of retinal vessel traits in two isolated populations. Departure from normality of the traits distribution is given by the p-value obtained from the Kologorov-Smirnov test

4.2.4.2. Effect of the covariates

The majority of the studies published so far included only sex and age as covariates. In addition, we tested the following covariates: the 3 first principal components of ancestry, the average systolic blood pressure (mean of two measurements), the average diastolic blood pressure, the average difference between systolic and diastolic blood pressure (pulse pressure) and the body mass index (BMI). Numerous epidemiological studies showed significant correlations of retinal vessel traits with past, present and future blood pressure and obesity (detailed literature review in chapter 5), and therefore these covariates were tested for the size of effects in our model.

			Single covariate effect (% trait variance explained)								
trait	Transformation	Population	age	sex	PC1	PC2	PC3	avsbp	avdbp	diffbp	Bmi
CRAE zone B		Orkney Korcula	**7.97%**** **3.18%****	0.42% **3.07%**	0.23% 0.41%	0.45% 1.92%	0.001% 0.16%s	**11.69%**** **8.64%****	**8.58%**** **9.46%****	**6.27%**** **4.16%****	**0.59%**** **0.66%****
CRVE zone B		Orkney Korcula	**5.6%**** **3.12%****	0.03% 0.09%	0.6% 0.002%	0.36% 0.63%	0.24% 1.36%	**3.59%** **2.18%****	**0.57%*** **1.08%****	**4.06%**** **1.26%****	**0.11%**** 0.07%
AVR zone B	rank transformed (Orkney)	Orkney Korcula	0.01% 0.003%	0.02% **2.83%****	0.01% 0.74%	0.002% 0.43%	0.03% 0.74%	**2.72%**** **1.99%****	**5.66%**** **4.05%****	0.22% **0.59%***	**0.87%**** **1.03%****
CRAE zone C		Orkney Korcula	**10.2%**** **4.4%****	0.43% **3.82%**	0.07% 0.16%	0.61% 2.13%	0.01% 0.28%	**15.48%**** **9.95%****	**11.84%**** **9.53%****	**8.1%**** **5.27%****	**1.38%**** 0.27%
CRVE zone C		Orkney Korcula	**5.57%**** **2.81%**	0.04% 0.21%	0.44% 0.002%	0.48% 0.64%	0.18% 1.34%	**3.23%**** **2.20%**	**0.78%*** **1.42%**	**3.1%**** **1.24%**	**0.18%**** **0.57%**
AVR zone C	rank transformed	Orkney Korcula	0.52% 0.16%	0.66% **3.00%****	0.21% 0.35%	0.02% 0.62%	0.38% 0.43%	**5.88%** **4.35%****	**8.92%** **5.83%****	**1.2%** **2.33%****	**1.54%**** **1.53%***
Arterial curvature tortuosity	rank transformed	Orkney Korcula	**4.17%**** **1.79%***	0.02% 0.81%	0.21% 0.54%	0.001% 1.32%	0.15% 0.001%	**3.2%**** **0.48%**	**1.51%**** **0.43%**	**2.1%**** **0.45%**	**0.82%*** **0.54%****
Venular curvature tortuosity	rank transformed	Orkney Korcula	0.007% 0.02%	0.02% 0.15%	0.46% 0.07%	0.002% 0.23%	0.05% 0.19%	0.02% 1.71%	0.007% 1.74%	0.07% 1.67%	0.81% 0.07%
Total curvature tortuosity	rank transformed	Orkney Korcula	**2.0%**** **1.14%***	0.008% 0.79%	0.48% 0.13%	0.002% 1.00%	0.11% 0.18%	**0.9%**** **1.1%***	**0.58%*** **1.19%***	**0.51%*** **1.07%***	**0.24%*** 0.36%
Arterial simple tortuosity	rank transformed	Orkney Korcula	**3.2%**** **1.11%***	0.001% 0.24%	0.01% 0.52%	0.00 2% 0.004%	0.07% 0.32%	**1.8%**** **0.20%**	**0.74%*** **0.43****	**1.3%**** **0.16%***	**0.71%*** **3.39%****
Venular simple tortuosity	rank transformed	Orkney Korcula	0.4% 1.68%	0.07% 0.37%	0.47% 0.02%	0.02% 0.73%	0.1% 0.45%	0.003% **0.30%***	0.02% **0.37%***	0.02% **2.11%***	0.29% **2.07%***
Total simple tortuosity	rank transformed	Orkney Korcula	**2.2%**** **0.60%***	0.01% 0.15%	0.31% 0.26%	0.003% 0.07%	0.11% 0.29%	**0.77%*** **0.63%***	0.15% **0.37%***	**0.74%*** **0.49%***	0.66% **1.24%***
Fractal Dimension zone C	rank transformed (Orkney)	Orkney Korcula	**12.37%**** **8.9%****	0.16% 0.90%	0.27% 0.35%	0.01% 2.47%	0.06% 0.50%	**8.8%**** **5.68%****	**2.9%**** **4.10%****	**6.9%**** **4.61%****	**1.3%**** **2.24%****
Fractal dimension Zone c - veins	rank transformed (Orkney)	Orkney Korcula	**4.42%**** **1.94%****	0.01% 0.86%	0.31% 0.74%	0.14% 1.20%	0.001% 0.67%	**1.4%**** **1.85%****	0.04% **1.74%****	**1.9%**** **1.73%****	0.47% 0.02%
Monofractal dimension Dbox	rank transformed	Orkney Korcula	**25.2%**** **19.39%****	0.04% 1.25%	0.11% 0.21%	0.08% 1.23%	0.003% 2.02%	**15.1%**** **13.95%**	**7.8%**** **8.44%**	**10.39%**** **10.46%**	**1.73%**** 1.18%
Multifractal dimesion – Do	rank transformed	Orkney Korcula	**16.47%**** **3.16%***	0.03% 0.76%	0.35% 0.15%	0.19% 0.11%	0.004% 0.46%	**8.4%**** **2.95%**	**4.8%**** **1.79%**	**5.51%**** **3.26%**	**1.23%**** 0.18%

Table 4.2. Variance of retinal vessel traits explained by 9 different covariates (age, sex, three first principal component, average systolic blood pressure, average diastolic blood pressure, difference between the systolic and diastolic blood pressure-average, body mass index). Effect of a single covariate was tested in a general linear model taking family structure into account. Significant results are highlighted, and significance level marked with ** for p-val < 0.01 and * for P-val < 0.05.

The effect of 9 different covariates on retinal vessel traits was tested singly in the both populations separately and expressed as % of the trait variance explained (Table 4.2). Age and average systolic blood pressure were covariates with greatest effects on retinal vessel traits. Average systolic blood pressure (avdbp) had the greatest effect on central arterial width in both zones B (8.64-11.7%) and C (9.96-15.48%), while age showed the strongest effect on several measurements of fractal dimensions (SIVA 8.9 -12.37%; VAMPIRE (16.47-25.4%). Overall age was the covariate with biggest effect on most retinal vessel traits. Average diastolic blood pressure (avdbp) also presented a strong effect on several traits; arterial diameter (zone B 8.58- 8.64%; zone C 9.95-11.84%) and AVR (zone B 4.05-5.66%; zoneC 5.82-8.91%). Avdp was the covariate with the biggest effect on AVR in both zones. Other covariates had much smaller effects. All three ancestry principal components had very small effects on the traits, which was anticipated since the polygenic effect is already fitted in the model. Body mass index (BMI) showed little effect on the traits (explaining up to a maximum of 1.73% of variance in monofractal dimensions in Orkney, but not significant in Korcula). Likewise sex had a very small effect on all of the traits tested; at most explaining 3.82% of variance in CRAE in zone C in Korcula (but not significant in Orkney). The traits most affected by any of the covariates were fractal dimensions measured by the VAMPIRE software; monofractal dimension measured in the Orkney sample showed a very strong effect of age (25.2%). The least affected trait by any of the covariates in both populations were venular tortuosity measurements (simple and curvature), withvery small and non-significant effects.

The effect of most covariates was stronger and more significant in the Orkney population than in Korcula, most likely due to the large difference in the subjects tested (811 Orkney, 387 Korcula).

It has been reported that axial length and intra ocular pressure affect the retinal vessel measurements explored in this analysis, and would be obvious candidates for potential

covariates. However only 529 individuals available for heritability analysis in Orkney have these measurements taken, and it was therefore decided not to explore their effects due to the inevitable loss of power. The effects of combined multiple covariates was explored further in the next section with testing of the best fitting models.

4.2.4.3. Heritability estimates

Several different covariates presented substantial and significant effect on retinal vessel traits, and therefore the subsequent step was to establish the best fitting model for the heritability analysis. The effects of various single and multiple explanatory covariates were tested, and the best possible model based on the Akaike criterion was selected for further analysis. Covariates with the individually strongest effects were considered in a multiple covariate model (age, BMI, average/systolic/diastolic/difference blood pressure, and age^2 in addition to age when significant in case the effect of age on the trait is non-linear). The following protocol was followed in the selection of the best fitting model: the variable with the greatest and most significant effect was first tested for each trait as the base model (if there were no significant variables, the no covariate model was kept as the best model); then the next best covariate was fitted and tested based on the Akaike criterion to check if it gives a better fitting model, if not the previous model was used otherwise the next best covariate was fitted and testing repeated. As the effect of age squared in not usually tested alone but in addition to age, for traits in which it had significant effect, it was tested in addition to age in the model. The best fitting models (based on the Akaike criterion) were selected to calculate the heritability estimates for each trait. Results are given in Table 4.3.

A Orkney

trait	Best model covariates	Effect size	% trait variance explained by covariates	h2best model (p-value)
CRAE zone B	Age Avsbp	-0.098 -0.167	13.11%	$0.40(2.74x10^{-9})$
CRVE zone B	Age bmi	-0.268 0.296	6.21%	$0.43(2.12x10^{-10})$
AVR zone B	Avdb	-0.0619	5.66%	$0.26(2.7x10^{-4})$
CRAE zone C	Age Avsbp	-0.1051 -0.1818	17.12%	$0.42(2.44x10^{-9})$
CRVE zone C	Age	-0.2553	5.57%	$0.39(1.99x10^{-8})$
AVR zone C	Avdbp	-0.0289	0.78%	$0.28(2x10^{-8})$
Arterial curvature tortuosity cTORTa	Age Avsbp	-0.009 -0.0049	4.47%	$0.52(2.9x10^{-12})$
Venular curvature tortuosity	Bmi	0.0172	0.81%	0.22(0.14)
Total curvature tortuosity	Age	-0.009	2.0%	0.44(0.001)
Arterial simple tortuosity	Age	-0.011	3.22%	$0.55(9.16x10^{-12})$
Venular simple tortuosity	Age	-0.003	0.4%	0.20(0.006)
Total simple tortuosity	Age	-0.009	2.86%	$0.43(4.1x10^{-9})$
Fractal Dimension zone C	Age Avsbp	-0.0176 -0.008	15%	$0.28(1.84x10^{-4})$
Fractal dimension Zone c - veins	Age avsbp	-0.014 -0.006	5.15%	0.20(0.005)
Monofractal dimension Dbox	Age Avsbp	-0.0282 -0.0081	26.79%	$0.24(9.6x10^{-4})$
Multifractal dimesion - Do	Age Avsbp	-0.0239 -0.0053	17.13%	0.11(0.12)

trait	Best model covariates	Effect size	% trait variance explained by covariates	h2 best model (p-value)
CRAE zone B	Avsbp Avdbp	-0.079 -0.28	10.59%	0.35 (0.07)
CRVE zone B	avsbp	-0.14	2.18%	0.59(0.005)
AVR zone B	avdbp	-0.001	4.05%	0.275(0.04)
CRAE zone C	Avsb Avdbp	-0.097 -0.24	11.4%	0.40(0.06)
CRVE zone C	Avsbp	-0.13	2.2%	0.54(0.02)
AVR zone C	Avdbp	-0.023	5.83%	0.41 (0.07)
Arterial curvature tortuosity - cTORTa	Age avsbp	-0.009 -0.004	0.88%	0.58(0.14)
Venular curvature tortuosity	Bmi	0.021 -0.041	0.07%	0.64(0.03)
Total curvature tortuosity	avsbp	0.002	1.1%	0.58(0.08)
Arterial simple tortuosity	avsbp bmi	-0.002 -0.033	3.2%	0.17(0.6)
Venular simple tortuosity	diffbp	-0.003	2.11%	$1.08x10^{-9}$(0.39)
Total simple tortuosity	avsbp	-0.004	0.63%	0.08 (0.38)
Fractal Dimension zone C	Age Avsbp	-0.00 -0.0001	11.2%	0.22(0.37)
Fractal dimension Zone c - veins	avsbp	-0.0001	1.85%	0.05(0.79)
Monofractal dimension Dbox	Age Avsbp	-0.0025 -0.0079	17.81%	0.79(0.009)
Multifractal dimesion - Do	Age Avsbp	-0.0006 -0.00002	2.3%	0.013(0.14)

Table 4.3 Heritability estimates and covariates effects for retinal vessel traits in Orkney and Korcula islands. Best models represent the best fitting models (obtained after following the described protocol) with significant covariates. Fixed effects are reported as regression coefficients expressed in unit of the covariate fitted. Rank transformed traits to normality are highlighted.

A combination of age and average systolic blood pressure resulted in the best fitting models for several traits in Orkney (CRAE zone B and C, arterial curvature tortuosity, SIVA total fractal dimensions and venular fractal dimensions, VAMPIRE monofractal dimensions and multifractal dimensions) and Korcula (arterial curvature tortuosity,

monofractal dimensions (by VAMPIRE and SIVA) , and VAMPIRE multifractal dimensions Vampire).

Similar to the single covariates model, the trait with the largest effect of the combined covariates in both populations was monofractal dimension (Dbox); age and average systolic blood pressure explained 26.79% of Dbox variance in Orkney and 17.81% in Korcula. Heritability estimates for Dbox were 0.24 in Orkney and 0.79 in Korcula. Fractal dimension estimated by SIVA was also highly affected by the same covariates in both populations (avsbp and age explained 15% of variance in Orkney and 11.2% in Korcula). Heritability estimates for traits adjusted for covariates were more comparable for this trait between the two populations (0.28 for Orkney and 0.22 for Korcula). All tortuosity measurements were least affected by the combination of any explanatory variables. In both populations, venular simple tortuosity was the trait least affected by the covariates (age explaining 0.4% in Orkney and diffbp 2.11% of the traits variance in Korcula).

The majority of the tested traits displayed moderate heritability values. Heritability estimates ranged from 0.11 for multifractal dimension to 0.55 for simple arterial tortuosity in Orkney. More extreme heritability estimates were recorded for several traits in Korcula (ranging from $1.08x10^{-8}$ for venular simple tortuosity to 0.79 for monofractal dimensions estimated by Vampire). Such a wide range of heritability estimates could be explained by the low number of subjects tested (N=387/N=328). Furthermore, heritability estimates for the majority of the traits in Korcula were not significant, again most likely due to the low number of tested subjects.

4.3. Genetic correlation of retinal vessel traits

4.3.1. Definition and measurements of genetic correlation

Phenotypic values of different traits are often correlated. The source of these correlations can be environmental and genetic, and similar to partitioning the traits variance into genetic and environmental, the same can be done for phenotypic correlation (r_p). It can be decomposed into genetic (r_a) and environmental correlation (r_e).

$$r_p = r_a + r_e$$

The genetic correlation quantifies the amount of the genetic influence shared between two traits. The magnitude and direction of genetic correlation depends on the pleiotropic effects (alleles in one locus affecting more than one trait) of genes and linkage disequilibrium (non-random association of alleles at two or more loci). The genetic correlation can be negative or positive, with a magnitude from -1 to +1.
Genetic covariance can be partitioned in the same manner as the genetic variance; into components of additive, dominant, epistatic and maternal and paternal effects. Therefore in order to estimate genetic correlation, genetic covariance and variance estimates are required. Additive genetic correlation is calculated as a ratio of the covariance between the additive genetic effects of two traits and the square root of the product of the variances for the additive genetic effects of the traits.

Estimation of genetic correlation requires genetically informative samples, e.g twins, large pedigrees, isolated populations. The software used for the estimation of genetic correlation requires pedigree information and this was available only for the Orkney population. The genetic correlation estimates were therefore calculated only for the Orkney population. In addition, the small population size for Korcula was an additional reason in the decision not to pursue the analysis in Korcula.

Genetic correlations were calculated for traits that were observed to have significant additive genetic variance because r_A is theoretically undefined when one trait has heritability equal to zero. Heritability analysis revealed substantial additive genetic variance for all tested traits in Orkney (therefore all traits were considered for the genetic correlation analysis). Best fitting model covariates identified in the heritability analysis were used.

4.3.2. Methodology

Genetic correlation estimates were assessed using the ASReml package (Gilmour et al. 2002). ASReml is a statistical package that fits linear mixed models using Restricted Maximum Likelihood (REML). ASReml requires a known genetic pedigree to estimate a range of quantitative genetic parameters (only available for Orkney population). Since the heritability estimates are also given in the ASReml output, the two heritability estimates for each trait were compared.

A multivariate animal model was used to estimate genetic correlations between selected traits. A multivariate model is used for correlation estimates as it allows for observation and analysis of more than one outcome variable at a time. Selection of fixed effects was based on the previous analysis of best fitting model. Multivariate models also allow for estimation of covariance components between traits. These include the (additive) genetic covariance COV_A which is then rescaled to give the genetic correlation r_G.

Input files include standard pedigree information for each individual (individual's ID along with maternal and paternal ID), and a second file consisting of phenotype information as well as the covariates used in the subsequent analysis. All the individuals

with the phenotypic data have to be included as offspring in the pedigree file. Following code was used:

First the multivariate model is specified:

```
Trait1   Trait2  ~  Trait   Trait.covariate   !r   Trait.ID
```

#Trait 1 and Trait 2 -represent traits tested for correlation

Trait -represents the mean of each trait

Trait.covariates- covariates used in the model

!r - qualifier tells ASReml to fit the terms that appear after this qualifier as random effects (in this case additive genetic effect).

Triat.ID –additive genetic effect

Then the model of the variance structure is specified:

```
1 2 1                    #Variance header: 1 residual structure with two dimensions, and 1 additional random effects (additive genetic effect)
```

Residual (R) structure defined as an unstructured covariance matrix:

```
0
Trait 0 US !S2== 1       # Starting values supplied as lower triangle
0.1                      # VR1,
0.1 0.1                  # COVR, VR2
```

Random effect (G) structure defined as an unstructured covariance matrix:

```
Trait.ID 2
Trait 0 US
1                        # Starting values supplied as lower triangle
0.1 1                    # VA1
ID                       # COVA, VA2
```

After this model is run, the output results file are generated containing six variance components (residual and additive variances of both traits along with their covariances). Finally a post-processing file is used to estimates heritability and genetic correlation. Total phenotypic variance for each trait is computed and heritability calculated a ratio of phenotypic and additive variance. Genetic correlation for the two traits is calculated a ratio of the covariance between the additive genetic effects of two traits and the square root of the product of the variances for the additive genetic effects of those traits.

4.3.3. Results : genetic and phenotypic correlations between the traits

Rank transformation of non-normally distributed traits was done prior to the correlation analysis. Covariate selection was based on the previous best fitting model analysis. In cases where two tested traits had different covariates, all covariates were fitted. Pairs of retinal vessel traits were selected based on the potential biological association between traits. In addition correlation between similar traits derived with different methodology (fractal dimensions) was also included in the analysis. In order to avoid redundant measurements, width measurements from only zone B and curvature tortuosity measurements were included in the analysis. Both phenotypic and genetic correlations were estimated for all tested pairs of traits. Strong positive genetic and phenotypic correlations were found for several pairs of retinal vessel traits (Table 4.4).

Trait	Covariates	Genetic correlation (SE)	Phenotypic correlation (SE)
CRAE and CRVE	Age, avsbp	0.65 (0.11)**	0.55 (0.03)**
CRAE and AVR	Age, avsbp	0.27 (0.18)	0.45 (0.03)**
CRVE and AVR	Age, bmi	-0.51 (0.14)**	-0.50 (0.03)**
CRAE and fractal dimension (SIVA)	Age, avsbp	0.69(0.20)**	0.13 (0.03)**
CRVE and fractal dimension (SIVA)	Age, avsbp	0.41 (0.18)*	0.16(0.04)**
CRVE and fractal dimension veins (SIVA)	Age, bmi	0.90 (0.28)**	0.19 (0.03)**
CRAE and curvature arterial tortuosity	Age, avsbp	0.003 (0.14)	0.17 (0.04)**
CRAE and curvature total tortuosity	Age, avsbp	0.13 (0.17)	0.14 (0.04)**
CRVE and curvature venular tortuosity	Age, bmi	0.54(0.28)	0.07 (0.04)
CRVE and curvature total tortuosity	Age, bmi	0. 22(0.16)	0.13(0.04)**
curvature arterial and curvature venular tortuosity	Age, avsbp	0.69 (0.21)**	0.32 (0.03)**
curvature arterial and curvature total tortuosity	Age, avsbp	0.97 (0.03)**	0.84 (0.01)**
Curvature vein and curvature total tortuosity	Age, avsbp	0.84 (0.11)**	0.75 (0.02)**
curvature arterial tortuosity and fractal dimension (SIVA)	Age, avsbp	0. 43(0.15)**	0.29(0.03)**
curvature venular tortuosity and fractal dimension (SIVA)	Age, avsbp	0.27 (0.27)	0.25 (0.03)**
curvature venular tortuosity and fractal dimension veins (SIVA)	Age, avsbp	0.27 (0.36)	0.26 (0.03)**
curvature total tortuosity and fractal dimension (SIVA)	Age, avsbp	0.37 (0.17)	0.32(0.03)**
CRAE and monofractal dimesion (VAMPIRE)	Age, avsbp	0.69 (0.09)**	0.42 (0.03)**
CRAE and multifractal dimesion (VAMPIRE)	Age, avsbp	0. 45(0.14)**	0.25(0.03)**
CRVE and monofractal dimesion (VAMPIRE)	Age, avsbp	0.25 (0.17)	0.32 (0.04)**
CRVE and multifractal dimesion (VAMPIRE)	Age, avsbp	0.41 (0.18)*	0.34 (0.03)**
curvature arterial tortuosity and monofractal dimesion (VAMPIRE)	Age, avsbp	0.40 (0.11)**	0.23(0.04)**
curvature total tortuosity and monofractal dimesion (VAMPIRE)	Age,avsbp	0.39(0.12)**	0.25(0.04)**
curvature arterial tortuosity and multifractal dimesion (VAMPIRE)	Age, avsbp	0. 46(0.13)**	0.21(0.04)**
curvature total tortuosity and multifractal dimesion (VAMPIRE)	Age, avsbp	0.48 (0.15)**	0.25 (0.04)**
monofractal dimesion (VAMPIRE)and fractal dimension (SIVA)	Age, avsbp	0.95 (0.09)**	0.51(0.03)**
multifractal dimesion (VAMPIRE) and fractal dimension (SIVA)	Age, avsbp	0. 84(0.14)**	0.43(0.03)**
multifractal dimesion (VAMPIRE) and monofractal dimension (VAMPIRE)	Age, avsbp	0.89 (0.11)**	0.78 (0.02)**

Table 4.4 Genetic and phenotypic correlations between retinal vessel traits estimated by the bivariate genetic analysis. Correlation values with standard error are given for pairs of traits. Significance levels are marked with ** for p-val < 0.01 and * for P-val < 0.05.

The strongest correlations (both phenotypic and genetic) were found for pairs of fractal dimension traits, as expected. High phenotypic correlation (78%) between mono and multifractal dimensions measured by VAMPIRE demonstrates close relationship between these two measurements. When fractal measurements are compared between SIVA and VAMPIRE software, moderate phenotypic and very high genetic correlation is detected. Very high phenotypic and genetic correlations were also found between several tortuosity measurements. As expected, high correlation between total tortuosity and its components (arterial and venular tortuosity) is confirmed. However, more interestingly high genetic correlation is found between arterial and venular tortuosity measurements (r_a=0.69, SE=0.21), as for the arterial and venular width measurements (r_a=0.65, se=0.11) suggesting a common genetic factors underlying these retinal vessel structures. Surprisingly genetic correlations between arterial widths and tortuosity measurements were extremely low (r_a=0.003, SE=0.14), which was not the case for the analogous venular measurements (r_a=0.54, SE=0.28). A strong genetic relationship was also found between measurements of arterial widths and fractal dimensions (SIVA- r_a=0.69, SE=0.20; VAMPIRE - r_a=0.69, SE=0.09).

In addition herrtiabilites obtained from ASReml and GenABEL were compared. Results are given in Table 4.5., and displaying consistencies between the two methods; one using genomic kinship matrix and the other constructed pedigree information in order to calculate relatedness between subjects.

Trait	ASReml herritability estimates	GenABEL herritability estimates
CRAE zone B	0.37**	0.40**
CRVE zone B	0.45**	0.43**
AVR zone B	0.30**	0.26**
Arterial curvature tortuosity - cTORTa	0.53**	0.52**
Venular curvature tortuosity	0.17	0.22
Total curvature tortuosity	0.43**	0.44**
Fractal Dimension zone C	0.31**	0.28**
Fractal dimension Zone c – veins	0.16	0.20**
Monofractal dimension Dbox	0.24**	0.24*
Multifractal dimesion – Do	0.18*	0.11

Table 4.5. Herritability estimates using two methods are presented. Significance of the herritability estimates are marked with ** for p-val < 0.01 and * for P-val < 0.05.

4.4. Discussion

Heritability estimates of retinal vessel traits found in the two isolated population of Croatia and Scotland showed substantial genetic contribution for the majority of the traits inter-individual variabilities. The results reported in this study were in agreement with published literature for vessel width traits. CRAE and CRVE reports from twin study were higher (CRAE 0.56-0.70; CRVE 0.62-0.83) than the ones reported here (CRAE Orkney 0.40, Korcula 0.35; CRVE Orkney 0.43, Korcula 0.59) as twin-based studies provide an upper limit for the heritability due to shared environments. However, when differences in heritabilities between the published measurements and the ones obtained in this study were tested (given their standard errors and tested by two sided test on z-scores (Jensen et al. 2003)), no significant differences were found. However our results are in agreement with the only published familial aggregation study (CRAE 0.48, CRVE 0.54; Lee et al. 2004). A Belgian population based study (413 participants from 70 families from a Flemish population) reported lower heritability estimates (CRAE 0.21, CRVE 0.34), but the general trend is similar across various studies, displaying higher heritabilites for venular diameters, compared to arterial. Interestingly heritability estimates for AVR (ratio of arterial and vein diameters) demonstrates very similar estimates across studies (Orkney 0.26, Korcula 0.28, Beaver Dam familial aggregation study 0.32, Flemish population based study 0.27). This perhaps is not surprising, since AVR gives a ratio of the two widths, and therefore lowers the total variance of the traits (it combines information from both arms of the circulatory system).

Only one study so far reported heritability estimates for arterial tortuosity measurements (Taarnhoj et al. 2008). No other estimates of tortuosity traits have been published to my knowledge. Thus this study represents the first comprehensive study of additive genetic contribution to retinal vessel tortuosity, as well as phenotypic and genetic correlations between these and other retinal vessel traits. Taarnhoj et al. reported very high heritability of arteriolar tortuosity arterioles (h_2=0.82), however as this was a twin study

including only 218 subjects the result is very likely to be inflated. Heritability of arterial curvature tortuosity was high and similar in both populations (h2=0.52 Orkney, h2=0.58 Korcula), and is the trait with highest reported heritability across both populations. Venular tortuosity gave lower values in Orkney (curvature h2=0.22, simple h2=0.20). This is not surprising because of the anatomical difference between arterioles and venules. The arteriolar wall is thicker (because of the tunica media; a layer made up of smooth muscle cells and elastic tissue) and less compliant to the influences of different mediators of vasodilatation and vasoconstriction, and thus is subject to less contortion than is the venular wall. Venular tortuosity is therefore more likely to be shaped by environmental factors (hypertension, vasodilatation and vasoconstriction), and less of its variance can be attributable to the genetic factors. Very inconsistent values of venular tortuosity were documented in Korcula (curvature h_2=0.64, simple h_2=1.08x10^{-8}).

No reported data on the genetic contribution to the variance of fractal dimensions exist, and this study represents the first reports of the traits heritability. Heritability reports for fractal dimensions measured by SIVA were consistent between the two populations (0.24-0.29), however other fractal traits provide very different estimates between the two isolates. In Orkney, monofractal measures from VAMPIRE and SIVA report very consistent values (0.24-0.28), confirming a good agreement between the two methods. Multifractal measurements from VAMPIRE report lower heritability (0.11). Values in Korcula vary from high (0.79) for monofractals (VAMPIRE) to very low (0.01) for multifractals (VAMPIRE). Overall the heritability measures were consistent between the two populations, with a few exceptions. Good agreement between the two isolates is reported for arterial tortuosity, the majority of the calibre measurements and the fractal dimensions measured by SIVA. The most probable source of inconsistencies is the low sample size of Korcula population (N=387), which is less than half the number of Orkney subjects (N=809). In addition, lower number of informative familial relationships is available as a result of the lower number of individuals phenotyped in Korcula. Since the precision of the heritability measures depend on the sample size and

pedigree structure (among other sources of bias), unbalanced estimates of heritability obtained from Korcula sample are not surprising.

Here we also report high phenotypic and genetic correlation between the calibre measurements CRAE and CRVE (r_a=0.68, r_p=0.52), in line with twin data estimates (r_a=0.65-0.77, r_p=0.55; Fachy et al. 2011) and higher than results reported in the Flemish study (r_a=0.36). Flemish study only included 413 participants randomly selected from 70 families, presenting a less genetically informative sample than the isolated population of Orkney. Phenotypic correlations were in line with reported data (r_p=0.55, Fachy et al. 2011). Correlations between other retinal vessel traits have not been published. Genetic correlations, representing a measure of the overlap in genetic effects between traits, conveyed moderate to high values, with the exception of the correlation between arterial widths and arterial tortuosity (r_a=0.003, SE=0.14). A novel finding was of high genetic correlation between both arterial and venular tortuosity measurements (r_a=0.69, SE=0.21)) and the arterial and venular width measurements (r_a=0.65 se=0.11) suggesting a common genetic factors underlying these structures in both arterioles and veins. Moderate to high genetic correlations of most pairs of tested traits suggest the possibility that common genetic variants are influencing general retinal vascular morphology, and together with documented substantial heritabilites of those retinal vessel traits propose suggest this to be a fruitful area for further gene mapping analysis. Phenotypic correlations between traits were also moderate to high in most pairs tested, with the exception of the pairs of vessel calibres - vessel tortuosity and vessel calibres - fractal dimension traits measured by SIVA (r_a=0.07-0.19). These results represent the first reports of phenotypic correlation between the retinal vessel traits other than vessel widths.

Large effects of age, blood pressure values (systolic; diastolic in some studies) and a moderate BMI effect have been reported in numerous studies for retinal vessel widths (Leung et al 2003, Wong et al 2003, Wang et al. 2006, Kaushik et al. 2007, Taylor et al. 2007, Gopinath et al. 2013, Klein et al. 2012, Li et al. 2012) tortuosity (Taarnhoj et al.

2008, Cheung et a.2011, Owen et al. 2011) and fractal dimensions (Liew et al. 2008, Kurniawan et al. 2012, Cheung et. al. 2012). In the two isolated populations examined in this study these covariates demonstrated moderate to large effect on retinal vessel traits. Age and average systolic blood pressure had the largest effects on fractal dimension measurements (age explaining 25.4% of monofractal dimensions variance in Orkney, and 19.39% in Korcula; avsbp explaining 15.1 % of variance in Orkney and 13.95% in Korcula of monofractal dimensions). Systolic blood pressure increase of 10mmHg was associated with 1.7 µm narrowing of retinal arterioles in Orkney and 1.5µm in Korcula, and 0.2 µm in Orkney and 0.9µm in Korcula narrowing of venules (adjusted for age and sex for comparison with published data), which was in agreement with reported effect sizes of systolic blood pressure on CRAE and CRVE in adult populations (1.1-1.9µm for CRAE , 0.2-0.5 µm for CRVE; lkram et al. 2004, Wong et al. 2006, Leung et al. 2003, Klein et al. 2012). Fewer studies reported effects of diastolic blood pressure, and data from Korcula (3.3 µm CRAE, 1.7 µm CRVE decrease per 10mmHg) and Orkney (2.7 µm CRAE, 0.8 µm CRVE decrease per 10mmHg) were also in agreement with published reports (2.1-4.3 µm CRAE, 0.7-0.8 µm CRVE; Ikram et al. 2004, Leung et al. 2003). Effect of bmi on retinal vessel traits was smaller compared to age and blood pressure traits (explaining at most 1.85% of the variance in simple venular tortuosity), however when compared to published results effect sizes were comparable in both Orkney and Korcula samples. Reduction in retinal arteriolar width was 0.003µm in Orkney and 0.08 µm in Korcula per kg/m^2 , while venular calibres were 0.29µm wider in Orkney and 0.45µm in Korcula per kg/m^2 (published reports; CRAE narrowing 0.07-0.72, CRVE widening 0.03-0.44 ;Wang et al. 2006, Ikram et al. 2004, Hughes et al. 2009). Identifying the sources of the phenotypic variation is a very important step in the heritability analysis, since it lowers the overall phenotypic variance and therefore allows for more accurate heritability estimates.

In summary, the present study provides evidence of substantial additive genetic contribution to retinal vessel morphology and provides novel data on the heritability of various retinal vessel traits, as well as the genetic and phenotypic correlation between them. Furthermore, heritability estimates reported here are in agreement with the published reports (for calibre traits). Estimates reported here together with the evidence of common genetic variants influencing general retinal vascular morphology provide good basis for the following gene mapping analysis, even though heritability estimates are not always a good indicator of the success of GWAS. The choice of covariates selected will be taken into account in the follow-up gene mapping studies.

Chapter 5. Retinal vessel traits: Association with systemic risk factors

5.1. Introduction

The main focus of this chapter was to explore the relationship of retinal microvasculature with subclinical cerebrovascular, cardiovascular and metabolic outcomes (including some inflammatory biomarkers) measured in individuals from Orkney and Korcula (representing a cross-section of general population from those two islands). Since the retina presents the lone site in the human body where microcirculation can be assessed non-invasively, exploring associations between traits derived from its vasculature and subclinical markers of systemic disease developed into a standard way of studying manifestation of the human microvasculature. The associations of systemic risk factors with retinal vessel calibers have been explored in several epidemiological studies already, providing a comparison set when similar traits are tested. Additional systemic risk factors are tested here for the first time (e.g pulse wave analysis traits) as the ORCADES and CROATIA-Korcula studies included a deep array of phenotypes. All retinal vessel measures will be tested which include retinal vessel calibers, tortuosity and fractal dimension measurements. The analysis described in this chapter is exploratory, and the main purpose was to validate traits measured and aid future research into retinal vessel traits; to identify possible risk factors or confounding factors and guide the future epidemiological and genetic analysis of retinal vessel traits. It does not attempt to explore deeper analysis of joint effects of the systemic factors tested, and is purely designed to detect new and so far unknown associations of retinal vessel traits and systemic factors. Simple correlation between ten different retinal vessel traits and medically relevant traits was examined in this chapter.

5.2. Literature review

There are a great number of epidemiological studies focusing on quantitative traits related to retinal vessels diameters, and their association with systemic diseases (type 2 diabetes, cerebrovascular disease, coronary heart disease). Nine major studies account for the majority of research in this field. Those studies are: Atherosclerotic Risk in Communities Study (ARIC), Blue Mountain Eye Study (BMES), The Beaver Dam Eye Study (BDES), Cardiovascular Health Study (CHS), The Wisconsin epidemiological study of diabetic retinopathy (WESDR), the Rotterdam study, Danish Twin Register study (DTRS), Singapore-Malay Eye Study (SMES) and Multi-Ethnic study of atherosclerosis (MESA). Majority of these studies have primarily focused on arterial and venous calibres and their association with systemic conditions and diseases. Details and cohort references to these studies are given in Table 5.1.

Study name	Sample size	Type of study	Reference
Atherosclerotic Risk in Communities Study (ARIC)	~17000	Prospective	http://www2.cscc.unc.edu/aric/
Blue Mountain Eye Study	~3500	Prospective, cross-sectional	http://www.cvr.org.au/bmes.htm
The Beaver Dam Eye Study	~2500	Prospective, cross-sectional	http://www.bdeyestudy.org/
Cardiovascular Health Study (CHS)	~6000	Prospective, cross-sectional	Fried et al. 1991
The Wisconsin epidemiological study of diabetic retinopathy (WESDR)	~1000	Prospective	http://www.epi.ophth.wisc.edu/content/wesdr
Rotterdam study	~8000	Prospective, cross-sectional	http://www.epib.nl/research/ergo.htm
Singapore-Malay Eye Study	~3300	Prospective, cross-sectional	Rossman et al. 2012,
Multi-Ethnic study of atherosclerosis (MESA)	~5000	Prospective, cross-sectional	http://www.mesa-nhlbi.org/
Danish Twin Register study (DTRS)	~100	Twin study	Taarnhoj et al. 2006

Table 5.1 Details of nine main epidemiological study included in the literature review. Sample size, type of study and the study reference is given.

Generalised retinal arterial and venous narrowing have been associated with increasing age in four study populations (ARIC, DTRS, BMES, BDES). The main features of hypertension are increased peripheral vascular resistance and increased vascular stiffness. Even though it is still unknown whether arterial narrowing precedes and contributes to the development of hypertension, or is just a consequence of it, testing the association between the blood pressure and retinal microvascular features is an obvious line of inquiry in microvascular epidemiological research. In addition, pathological studies have shown that cerebral and retinal vasculatures change with increasing blood pressure (Hill 1970, Goto et al. 1975, Tso and Jampol, 1982). Retinal arterial narrowing have been shown to be associated with current high blood pressure (ARIC, DTRS,

BMES, BDES, Rotterdam study, CHS, SMES, MESA), past high blood pressure (ARIC, CHS, BMES) and future blood pressure (ARIC, BMES), whilst retinal venous dilation was associated with high current and future blood pressure (BMES).

Since inflammation is considered to be one of the major underlying mechanisms of small and large vessel disease, associations with these factors and retinal calibres have also been tested. Several studies found association between venular dilation and markers of inflammation (C reactive protein-CRP, white blood cell count, interleukin-6; BDES, Rotterdam study, MESA). In addition, the Rotterdam study reported association of higher white blood cell counts and erythrocyte sedimentation rate with wider retinal venules.

Microvascular damage is one of the main pathological features of diabetes mellitus, and many studies explored associations between hyperglycaemia and retinal vessel calibres. Higher levels of serum glucose were reported to be associated with wider arterioles (MESA). In diabetic patients higher levels of HbA1c (glycated haemoglobin) were associated with both wider arterioles and venules (WESDR), and reported wider venular calibers than in a control group (BMES). Venous dilation alone has been shown to be associated with increased risk of diabetes mellitus (Rotterdam study, WESDR, ARIC).

Several studies reported association of retinal vascular calibre and dyslipidaemia (MESA, ARIC, SMES) and obesity (BMES, Rotterdam study, MESA, SMES). It is reported that dyslipidaemia is associated with wider venules (higher triglycerides – ARIC,MESA; lower HDL –BMES,MESA, Rotterdam studies). Higher LDL levels were associated with venular dilation in MESA, but not in BMES. Obesity was associated with retinal arterial narrowing and wider venules (Rotterdam study, ARES, MESA).

In the ARIC study, arterial narrowing was found to share some risk factors with atherosclerosis (carotid arterial plaque and carotid arterial stiffness, but not with carotid intima-media thickness (CIMT)), but wasn't associated with clinical indicators of atherosclerosis. This was not confirmed in Rotterdam study, where lower AVR was associated with higher CIMT and carotid plaque score, nor in the CHS study which found no associations between AVR and measures of atherosclerosis. ARIC and Rotterdam studies found that both wider retinal venules and narrower arterioles are related to increased risk of stroke, but the Rotterdam study reported that venous dilation rather than arterial narrowing explained this association. The CHS study reported association between larger venous diameter and smaller arterial width increased 5-year risk of coronary heart disease (CHD), whilst the ARIC study reported association of lower AVR with increased risk of CHD and acute myocardial infarction in women but not in men. Both CHD and stroke mortality were found to be associated with larger venous and smaller arterial calibres (BDES, BMES). In addition, SMES reported association of retinal arterial narrowing with chronic kidney disease. The Sydney childhood eye study found association between narrower retinal arterioles and lower birth weight (Gopinath et al. 2013).

Vessel changes	Association	Study
Narrower arterioles	increasing age	ARIC, DTRS, BMES, BDES
Narrower arterioles	current high blood pressure	ARIC, DTRS, BMES, BDES, Rotterdam study, CHS, SMES
Narrower arterioles	past high blood pressure	ARIC, CHS, BMES
Narrower arterioles	future high blood pressure	ARIC, BMES
Narrower arterioles	increased risk of stroke	ARIC
Narrower arterioles	increased risk of coronary heart disease and acute myocardial infarction in women	ARIC
Narrower arterioles	increased incidence of coronary heart disease	CHS
Narrower arterioles	stroke mortality	BDES, BMES
Narrower arterioles	coronary heart disease mortality	BDES, BMES
Narrower arterioles	chronic kidney disease	SMES
Narrower arterioles	lower birth weight, shorter birth length and smaller head circumference	Sydney childhood eye study
Narrow venules	increasing age	ARIC, DTRS, BMES, BDES
Venular dilation	high blood pressure (present and future)	BMES
Venular dilation	markers of inflammation (CRP, fibrinogen) (CRP, fibrinogen, interleukin 6)	Rotterdam study BDES, MESA
Venular dilation	hyperglycemia	BDES, Rotterdam study, MESA, SMES
Venular dilation	obesity	BMES, Rotterdam study, MESA, SMES
Venular dilation	increased risk of stroke	ARIC, Rotterdam study
Venular dilation	coronary heart disease mortality	BDES, BMES
Venular dilation	stroke mortality	BDES, BMES
Venular dilation	increased risk of diabetes mellitus	ARIC, Rotterdam study, WESDR
Venular dilation	dyslipidemia	MESA, ARIC, SMES

Table 5.2. Association of retinal vessel diameters with different disease outcomes, markers and risk factors.

Two major studies have focused on association of quantitative vessel traits and cognition (ARIC, and Lothian Birth Cohort (LBC) study). The ARIC study found that the presence of retinal abnormalities (retinopathy, micro aneurysms, retinal haemorrhages and exudates) was associated with small decrease in cognitive function, at the same time they reported no association between retinal arterial narrowing and cognitive function. In addition, LBC found no association of vascular width measurement (CRAE, CRVE, AVR) and cognition (Patton et al. 2007). However they found association of suboptimality of the retinal bifurcation junctional geometry (the root mean square of the difference of the median angle/branching coefficient from 75°/1.26) and cognitive decline. They reported associations of suboptimal angles with logical memory (GLM, p=0.03), and suboptimal branching coefficients with verbal fluency (GLM, p=0.01) and with cognitive ability score (GLM, p=0.02). However, they did not apply adjustments for multiple testing and so the effect, if real, is of questionable significance.

Several studies reported associations of monofractal dimensions with systemic diseases. Grauslund et al. (2010) investigated association of retinal monofractal dimensions with vascular complications in long-term surviving type 1 diabetes patients, and found that patients with lower fractal dimension were more likely to have proliferative retinopathy and neuropathy. In addition they found that patients with lower fractal dimensions were older, which was confirmed in several additional studies (Liew et al. 2008, Cheung et al. 2012, Azemin et al. 2012). SMES (Cheung et al. 2012) also reported association between lower monofractal dimensions and higher mean arterial blood pressure, which was confirmed in further studies (Sng et. al 2010, Kurniawan et al. 2012).Association between lower fractal monodimensions and diastolic blood pressure were also reported (Liew et al. 2008, Kurniawan et al 2012). Decreased fractal dimensions have also been associated with the incidence of stroke (Doubal et al. 2010, Kawasaki et al. 2011).

Patients with chronic kidney disease were associated with suboptimal fractal dimension (Sng et al. 2010). Reduced monofractal dimensions in children were reported to be associated with smaller head circumference (Gropinath et al. 2013).

Vessel changes	Association	Study
lower fractal dimensions	proliferative retinopathy and neuropathy in long term surviving type 1 diabetes patients	Grauslund et al. 2010
lower fractal dimensions	increasing age	Grauslund et al. 2011, Cheung et al. 2012, Azemin et. al 2012, liew et al. 2008
lower fractal dimensions	increased mean arterial blood pressure	Cenung et al. 2011, Sng et al. 2013, Kurniawan et al. 2011
lower fractal dimensions	elevated diastolic blood pressure	Kurniawan et al. 2011, Liew et al. 2008
lower fractal dimensions	stroke	Doubal et al. 2010, Kawasaki et al. 2011
suboptimal fractal dimensions	chronic kidney disease	Sng et al. 2010
reduced arterioral tortuosity	older age	Cheung et al. 2011,
Reduced arterial tortuosity	higher blood pressure	Cheung et al. 2011, Hasan et al. 2013, Taarnhoj et al 2008
decreased tortuosity	increased risk of death from ischemic heart disease	Witt et al. 2006
Decreased arterial tortuosity	higher body mass index	Taarnhoj et al 2009, Cheung et al. 2008
Decreased arterial tort	increased Hb1c	Sasongko et al. 20130
Increased venular tortuoasity	younger age, higher blood pressure, lower HDL	Cheung et al. 2008

Table 5.3. Association of fractal dimensions and tortuosity with different disease outcomes, markers and risk factors

SMES study reported associations between reduced arteriolar tortuosity and older age, higher blood pressure and higher body mass index (BMI) (Cheung et al. 2011). The same study also reported associations between greater venular tortuosity and younger

age, higher blood pressure and lower high-density lipoprotein (HDL) cholesterol level. Association between reduced arterial tortuosity and increased mean arterial blood pressure was reported in several other studies (Hasan et al. 2013, Taarnhoj et al 2008). In contrast, a study of 10-year old children reported positive associations between arteriolar tortuosity and systolic and diastolic blood pressure, levels of triglyceride, total and LDL cholesterol. Higher HbA1c was associated with increased arteriolar tortuosity (Sasongko et al. 2010), whilst less tortuous ones have been associated with increased risk of death from ischemic heart disease (Witt et al. 2006). Taarnhoj et al. (2008) have also reported in their twin study that less tortuous vessels are associated with increased body mass index.

To summarise, out of all the retinal vessel traits retinal vessel width measurements have been studied to the greatest extent. There have been independent sources of evidence that retinal arterial narrowing and increased retinal vein diameter are associated with several cardiovascular disease outcomes and markers. Other quantitative vessel measurements haven't been studied in such extent as vessel diameters. Research into fractal analysis and systemic disease have only started recently and the first results are looking very promising, and further research into this field is of great interest. Retinal vessel tortuosity still remains fairly unexplored area of research of retinal vasculature, and more research into associations with risk factors is essential.

5.3. Material and methods

5.3.1 Selection of cardiovascular risk factors

The selection of cardiovascular risk factors used in the correlation analysis was based on the following criteria; well established factors used in previously published retinal vessel research, and availability of the phenotypes in the two populations tested. Established guidelines from prior epidemiological studies on the associations of the systemic risk factors with retinal vessel calibers was complemented with one novel trait (from pulse wave analysis). Risk factors identified though literature could be divided into six distinct groups: blood pressure traits (systolic and asystolic blood pressure measurement), markers associated with hyperglycemia (glucose, Hb1c), markers of dyslipidemia (cholesterol, triglycerides, LDL, HDL), markers of obesity (bmi, waist to hip ratio), markers of atherosclerosis (CIMT) and markers of inflammation (CRP, fibrinogen). One novel traits tested has been recognized as a marker of systemic atherosclerosis (pulse wave analysis). Selected risk factors were used for the correlation analysis of all of retinal vessel traits available.

5.3.2. Measurement of cardiovascular risk factors

Blood was drawn directly into monovette tubes. Total and HDL cholesterol, triglycerides, HDL, CRP and glucose were measured EDTA plasma, fibrinogen was measured in citrated plasma, and HbA1c was measured in whole blood in EDTA. LDL levels were calculated using the Friedewald equation. All analytes were determined by enzyme-based technologies using auto-analysers in accredited laboratories.

Ankle brachial pressure index is the ratio of the blood pressure in the lower legs to the blood pressure in the arms. The measurements used in the analysis were calculated as the lowest posterior tibial (ankle) systolic pressure divided by the highest brachial systolic pressure.

Augmentation index, was one trait selected from the pulse wave analysis (PWA). PWA is used to estimate the degree of arterial and was performed after 1 hour of horizontal rest. Radial arterial tonometry was performed using the SphygmoCor Px system (AtCor Medical, NSW, Australia) with a Miller tonometer (Miller instruments, TX, USA), continuously for 30 seconds in a resting state and an average waveform was generated. Four reading were performed in total (two successful readings followed by 5 minutes rest followed by two further readings). Different PWA measurements were derived from each waveform and the average value was obtained.

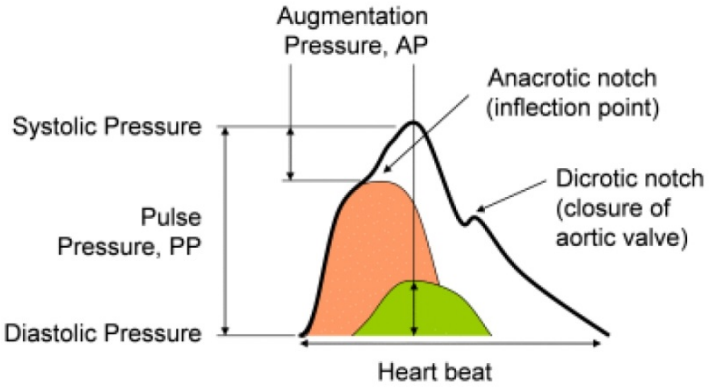


Figure 5.1. Example pulse wave describing key elements form which the PWA traits are derived from (http://www.pulsecor.com/augmentation-index.html)

The augmentation index is a ratio calculated from the blood pressure waveform and one of the most commonly used non-invasive measures of arterial stiffness (a risk factor for cardiovascular events). It is defined as the ratio of augmentation pressure (AP) to pulse pressure (PP), or by how much the reflected pressure wave augments or raises the overall pulse pressure. The stiffer the arteries the more quickly the reflected pressure wave travels and so is superposed on the forward wave thus augmenting the pressure.

CIMT (carotid intima media thickness) was assessed by ultrasound examination of the common carotid artery 2 cm below the carotid bifurcation. The mean of the three measurements was used as a final trait in the analysis. This measure was only available in the ORCADES study.

Systolic and diastolic blood pressures were measured twice using OMRON digital sphygmomanometers, and the mean of the two measurements was used. Smoking information was collected (categorized as current, former or never smoker) through the questionnaire. Percentage of body fat was measured using bioimpedance with Tanita scales. BMI and waist to hip ratio was calculated accordingly to the following formulae:

BMI = Weight in Kilograms / (Height in Meters x Height in Meters)

Waist to hip ratio= The Circumference of the Waist/Circumference of the Hips

5.3.3. Statistical analysis

5.3.3.1. Testing distributional properties of the traits and adjusting for non-normal distribution

Normality of the distribution for each trait was tested using the Kolmogorov-Smirnov normality test (lillie.test) implemented in nortest package in R (J. Gross 2012). Traits for which the distribution deviated from normality were transformed using quantile normalisation with the rank transformation function (*trait* ~ age + sex), of GenABEL (Aulchenko et al. 2007). It ranks the Z-scored residuals of the regression of the traits on the relevant covariates and performs quantile normalization.

5.3.3.2. Correction for the relatedness of the individuals

Since both Orkney and Korcula represent genetically isolated island populations and therefore contain substantial proportion of relatively closely related individuals, correcting for this relatedness was required. As a result of many close relationships between individuals, similarity between certain heritable traits could arise from these familial relationships and not from the true correlations between the retinal vessel traits and systemic risk factors. In order to correct for this, polygenic model of inheritance (explained in chapter 4) was used to obtain residuals, that have the estimated random polygenic effect and the effects of covariates (age, sex) factored out. Age and sex were the only covariates considered in this analysis, as a result of exploratory nature of the analysis. More complex selection of the covariates for each of the 180 pairs of trait tested was beyond the scope of this type of analysis. All of the cardiovascular risk traits, but waist to hip ratio, of interest were rank-transformed prior to the analysis, because of their departure from normal distribution.

5.3.3.3. Correlation analysis

Linear correlation between retinal vessel traits and cardiovascular risk factors were tested using bivariate Pearson's correlation. Pearson's correlation coefficient between two variables is defined as the ratio of the covariance of the two variables divided and the product of their standard deviations. It ranges from -1 to 1 (correlation coefficient of 1 and -1 implies a perfect linear relationship between two variable, and 0 implies no linear correlation). For all pairs of traits tested (both cardiovascular and retinal vessel traits), residuals obtained from the polygenic model were used to test the correlation. Zone B diameter measures and curvature tortuosity measurement were used (leaving out measurements in zone C and simple tortuosity) in order to minimize redundancy and

keep the analysis more clear. Statistical analyses were performed using statistical software SPSS and R.

5.4. Results

6.4.1. Descriptive statistics for cardiovascular risk factors

Descriptive statistics for 18 different cardiovascular risk factors are displayed in Table 5.4. Subjects form Korcula did not have the CIMT measurements taken, so the values for CIMT are missing from the table. Values for majority of the traits were comparable between the two independent populations, however glycated hemoglobin levels and the pulse wave augmentation index presented slightly different mean values. This is probably due to the lower numbers of individuals having these measures taken in Korcula (HbA1c and augmentation index) and Orkney (augmentation index). The normality of the distribution of all traits were tested, and rank transformed accordingly. In addition 22.5% in Korcula and 33.7% in Orkney are current smokers, while 48.2% in Korcula are ever-smokers, compared to 42.1% in Orkney.

Orkney	N	Minimum	Maximum	Mean	Std. Deviation	Kologorov-Smirnov test
ABPI	871	0.55	1.31	1.03	0.10	$3.8x10^{-4}$
Avdbp	873	0	164	75.99	10.55	$<2.2x10^{-16}$
Avsbp	873	58.5	212	130.88	19.12	$5.07x10^{-8}$
Bmi	867	16.97	47.63	27.67	4.84	$4.94x10^{-13}$
BodyFat	866	8.7	53.8	31.81	8.64	0.007
CAGPHHR75mean	587	-28	47.5	15.60	14.41	4.88x10-$^{-15}$
Chol	889	2.7	12	5.45	1.15	$4.1x10^{-6}$
CIMT	866	0.02	0.35	0.05	0.02	$<2.2x10^{-16}$
CRP	882	0.15	90.1	2.55	5.06	$<2.2x10^{-16}$
Fibrinogen	882	1.18	9.79	3.4	0.86	$1.31x10^{-11}$
Glucose	889	3.9	17.1	5.43	0.99	$<2.2x10^{-16}$
HbA1cIFCC	873	2.1	9.7	3.64	0.78	$<2.2x10^{-16}$
HDL	889	0.7	4.9	1.55	0.43	$3x10^{-10}$
LDL	889	0	9	3.36	1.07	$<2.2x10^{-16}$
Triglyc	889	0.2	11.6	1.18	0.76	$<2.2x10^{-16}$
waist_hip ratio	872	0.47	1.23	0.91	0.09	0.036

Korcula	N	Minimum	Maximum	Mean	Std. Deviation	Kologorov - Smirnov test
Abpi	860	0	2	0.99	0.15	$6.22x10^{-11}$
Avdbp	878	58	113	81.29	9.31	$1.19x10^{-8}$
Avsbp	879	93	215	138.95	21.02	$1.1x10^{-4}$
Bmi	881	16.59	53.84	27.96	4.14	0.008
Body_fat	856	14.47	50.33	36.23	6.51	$4.75x10^{-16}$
Cholesterol	895	2.7	11.3	5.95	1.23	$5.6x10^{-7}$
Crp	870	0.012	14.12	3.19	3.05	$<2.2X10^{-16}$
Fibrinogen	801	2	11	4.56	1.51	$<2.2X10^{-16}$
Glucose	895	3.8	17.5	5.81	1.55	$<2.2X10^{-16}$
hba1c	456	5	13	5.79	0.765	$<2.2X10^{-16}$
Hdl	892	0.76	2.7	1.46	0.35	$8.16X10^{-10}$
Ldl	889	1.3	8	3.82	1.04	$5.8X10^{-8}$
Pulse-wave	486	-21.5	54.5	22.01	11.31	$6.92x10^{-16}$
triglycerides	895	0.3	21.84	1.49	1.26	$<2.2X10^{-16}$
Waisthip	847	0.64	1.35	0.91	0.09	0.67

Table 5.4. Descriptive statistics of selected cardiovascular risk traits in two isolated populations. Departure from normality of the traits distribution is reflected by the p-value obtained from the Kolmogorov-Smirnov test (the lower, the greater)

5.4.2 Correlation between retinal vessel traits and cardiovascular disease risk factors

Correlation matrices representing correlation coefficients for all pairs of tested traits are given in table 5.5. Correlation coefficients between most pairs of traits were very low and non-significant in both populations. However some of the tested traits displayed consistent results in both of the populations tested. Correlation between arterial caliber in zone B and blood pressure measurements (both systolic and diastolic) showed the highest significant correlation values in both Orkney and Korcula (Orkney – 0.24 CRAE-avsbp $p=4.3x10^{-12}$, 0.238 CRAE-avdbp $p=7.6x10^{-12}$; Korcula 0.242 - CRAE-avsbp $p=1.9x10^{-6}$, 0.214 CRAE-avdbp $p=2.8x10^{-5}$). Correlation between arterial diameter and blood pressure measures were negative in both population, confirming the widely previously reported associations between high blood pressure and arterial narrowing. AVR showed similar consistent values in both population for blood pressure measurements, as well as percentage of body fat (-0.14 in Orkney and -0.15 in Korcula). Fractal dimension measurements showed significant correlation values with blood pressure measurements in Orkney, but not in Korcula. Three different measurements of fractal dimension of the whole vasculature presented inverse correlation coefficients with both systolic and diastolic blood pressure (ranging from -0.09 to -0.15). In Korcula, measurement of multifractal dimensions presented the only significant negative correlation with diastolic blood pressure (-0.13). Pulse wave augmentation showed significant inverse correlation coefficient with AVR, of the similar size, in both populations (Orkney 0.10, Korcula 0.18). Puls wave augmentation index represents the only marker of atherosclerosis displaying consistent correlation values across both population for the same trait. CIMT was measured only in the subject form Orkney and showed significant correlation values with arterial curvature tortuosity (0.08). ABPI did not display any significant correlation values with any of the retinal vascular trait in both of the tested populations. Traits showing significant correlation values with the most retinal vessel traits in Orkney were the total cholesterol level and LDL levels; displaying significant values with all of the tortuosity and fractal dimension measurements (ranging

from 0.07 to 0.09 from cholesterol, 0.08 to 0.11 for LDL). Current smoking displayed significant correlations with venular dilatation in both populations.

Correlations between AVR adjusted for body fat, and average systolic blood pressure (and average diastolic blood pressure) also adjusted for body fat were calculated in order to see if the effects were independent or not. The same procedure was repeated to adjust for pulse wave augmentation. Lower correlation coefficients were observed between both blood pressure measurements in Orkney (AVR-avsbp r=-0.152, AVR-avdbp r=-0.211 (adjusted for pulsewave augmentation); AVR-avsbp r=-0.160, AVR-avdbp r=-0.23 adjusted for body fat percentage)) when traits were corrected for pulse wave augmentation and body fat percentage. This drop in the correlation was not nominally big, but given the high significance of the correlation coefficients it could be concluded that the associations between AVR and blood pressure measurements are not in large part due to confounding with body fat or pulse wave. Results for Korcula did not show a similar trend (AVR-avsbp r=-0.172, AVR-avdbp r=-0.149 (adjusted for pulsewave augmentation); AVR-avsbp r=-0.210, AVR-avdbp r=-0.157 adjusted for body fat percentage)). This observation should be tested in a further independent cohort, in order to confirm the results observed in Orkney.

	BMI	avsbp	avdbp	bodyfat	ABPI	CIMT	CRP	Fibrinogen	Glucose	HbA1c	Choll	HDL	LDL	Triglycerides	Waist to hip ratio	Puls wave	regular smokers	Ever never smoked
CRAE	-0.006	**-0.240****	**-.238****	-0.036	-0.022	0.016	-0.002	0.048	-0.035	0.051	0.003	-0.015	0.013	-0.022	-0.023	-.114**	-0.021	0.027
CRVE	0.080*	-0.071*	-0.012	.093**	0.027	0.059	0.067	0.048	0.018	0.038	0.045	-0.051	0.028	.071*	0.06	-0.022	0.041	.097**
AVR	-0.090*	**-0.178****	**-.241****	**-.136****	-0.054	-0.045	-.073*	0.001	-0.053	0.019	-0.042	0.036	-0.014	-.095**	-.089*	-.103*	-0.064	-.072*
cTORTa	-0.011	-0.056	-0.038	-0.03	0.01	-.082*	0.005	0.037	-0.014	0.023	.071*	0.03	.078*	0.048	-0.003	-0.003	-.069*	-0.029
cTORTv	0.080*	0.019	0.003	.082*	-0.033	0.026	0.034	.079*	.077*	0.061	.091**	0.033	.091**	0.03	-0.012	0.011	-0.016	-0.007
cTORTt	0.034	-0.014	-0.012	0.024	-0.019	-0.028	0.017	0.064	0.029	0.043	.093**	0.046	.093**	0.048	-0.007	0.02	-0.055	-0.02
FractlaDimc	0.019	-0.142**	-.089*	0.033	0.057	-0.053	0.042	.083*	0.055	-0.01	.086*	-0.001	.080*	0.038	0.034	-0.032	0.002	0.014
FractalDimv	0.007	-0.023	0.026	0.048	0.048	-0.003	0.021	0.056	0.046	-0.008	.071*	0.019	0.059	0.037	0.037	0.027	0.001	0.002
Dbox	0.037	-0.147**	-.132**	0.043	-0.033	-0.04	0.014	0.04	0.049	0.014	.072*	-0.057	.091*	0.042	0.043	-0.067	0.037	.088*
D0	0.031	-0.092*	-.093*	0.03	-0.021	0.016	-0.016	0.042	0.069	0.023	.081*	-0.045	.110**	0.034	0.047	-0.022	-0.017	0.065

	BMI	avsbp	avdbp	Bodyfat	ABPI	CRP	Fibrinogen	glucose	HbA1c	chol	HDL	LDL	Triglycerides	waist-hip ratio	Puls wave	Regular smokers	Ever never smokers
CRAE	-0.07	**-.242****	**-.214****	-0.06	0.08	0.09	-0.03	-0.002	-0.01	-0.06	-0.05	-0.05	0.01	.147**	-0.10	0.07	.113*
CRVE	-.125*	-0.10	-0.09	0.05	0.01	0.09	0.02	0.07	-0.04	0.09	0.06	0.06	0.05	.144**	0.08	0.10	.121*
AVR	0.08	**-.175****	**-.152****	**-.139****	0.08	-0.001	-0.06	-0.07	0.04	-.170**	-.132**	-.131**	-0.05	0.01	-.185**	-0.03	0.01
cTORTa	0.01	0.001	0.03	-.160**	-0.06	-0.06	-0.07	-0.03	0.08	-0.02	-0.002	-0.002	-0.04	0.03	-0.10	-0.03	0.04
cTORTv	-0.04	0.03	0.004	0.04	0.05	0.01	0.01	0.04	0.05	0.02	-0.03	0.02	0.03	0.06	-0.01	0.01	-0.002
cTORTt	-0.03	0.03	0.04	-0.08	0.01	-0.02	-0.04	-0.0004	0.09	-0.002	-0.02	-0.0004	-0.01	0.08	-0.03	-0.01	0.03
FractalDimC	0.00	-0.05	-0.04	0.08	0.03	-.102*	-0.03	0.04	-0.003	0.06	-0.01	0.07	0.05	-0.03	0.11	0.06	-0.04
FractlaDimv	-0.02	0.01	0.03	0.04	-0.003	-0.04	-0.03	0.06	-0.05	0.06	0.01	0.05	0.06	-0.01	.173*	0.10	-0.001
D0	-0.05	0.05	0.004	.151**	-0.03	0.04	.139*	0.05	0.15	0.05	-0.06	0.06	0.00	0.08	0.06	0.03	.131*
Dbox	-0.10	-0.10	-.134*	.129*	0.00	0.02	0.12	0.02	.165*	0.03	-0.05	0.05	0.02	0.04	0.02	0.07	0.08

Table 5.5 Correlation matices representing person's correlation coefficient between retinal vessel traits and selected cardiovascular risk factors. *. Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level, (tailed).

5.5. Discussion

The results of the correlation analysis have confirmed previously known associations between retinal arterial diameters and blood pressure measurements. Several novel associations were detected with traits that have been tested in this analysis for the first time (pulse wave augmentation index). This study further extends previous work on associations of retinal vessel traits and systemic risk factors, which previously reported that the retinal vascular calibre and fractal dimensions of the retinal vascular network is influenced by blood pressure in adult populations (Liew et al. 2008, Sun et al. 2009, Kurniawa.n et al. 2012). However we did not find significant correlations between retinal vessel tortuosity and blood pressure measurements as previously reported (Taarnhoj et al. 2009, Cheung et al. 2011, Hasan et al. 2013).

Highest correlation coefficient values were reported for CRAE and both systolic and diastolic blood pressure, and these values were consistent across both populations, even though Korcula included only 389 subject with both phenotypes measured (compared to 809 from Orkney), confirming the robustness of this association that has been reported in numerous retinal vascular research studies. Correlation coefficient between systolic blood pressure and venular diameter was significant in Orkney, but of a smaller size when compared to arterial association. Positive correlation between venular dialation and obesity has also been reported (Klein et al. 2006, Ikram et al. 2004, Wong et al. 2004, Wong et al. 2006), and confirmed here in Orkney (r(CRVE-bmi)=0.08), but not in Korcula where it displayed inverse correlation with bmi; opposite to the published reports. This is most likely the result of low number of subjects tested in Korcula. In addition, association of venular dilation with dyslipidaemia, hyperglycaemia and inflammation markers reported in literature has not been replicated here. CRVE presented positive but non-significant associations with serum levels of glucose,

triglycerides, total cholesterol in both Orkney and Korcula and inverse correlation with HDL in Orkney.

Lower fractal dimesion measurements have been shown to correlate with increased blood pressure measurements, which was also confirmed in the analysis of subjects from Orkney. Results were consistent between different monofractal and multifractal measurements, while in Korcula significant correlations were only reported between multifractal measurement and diastolic blood pressure. Blue Mountains Eye Study (Liew et al. 2008) reported correlation coefficient of -0.29 (after adjustment for age and sex) between systolic blood pressure and monofractal measurement of retinal vasculature which is sizably higher than reported here (0.147/0.142), however their study included only 300 participants compared to 809 included in the analysis of subjects from Orkney.

Relationship between retinal vessel traits and markers of atherosclerosis has not been studied extensively and a few published reports showed no significant relationship between CRAE, except for the CIMT (Ikram et al. 2004, Torres et al. 2013). The same study reported association of CRVE with several markers of atherosclerosis. CIMT was reported to be associated with arterial narrowing (Ikram et al 2004, Torres et al. 2013) and venular diameter (Torres et al. 2013). Second atherosclerotic marker tested here (ABPI) showed no significant relationships between retinal vessel traits tested in both populations. Previous reports suggested a linear relationship of lower ABPI and larger venular diameters (Ikram et al. 2004), but it was not detected here. A novel finding was the significant correlation between arterial curvature tortuosity and CIMT (r=-0.08), suggesting a role of the atherosclerotic plaque in the tortuosity of arterioles and should be tested further in larger studies. A novel atherosclerotic marker (marker of arterial stiffness) tested here, pulse wave augmentation index, displayed significant correlations between CRAE (Orkney; -0.114) and AVR (Orkney -0.103; Korcula -0.185). This is the first report of this type of associations and warrants a further research into the

relationship between this and other traits derived from pulse wave analysis and retinal vessel diameters. Augmentation index and central blood pressure are thought to measure different aspects of arterial stiffness. The low drop in the correlation coefficients between AVR and blood pressure measurements in Orkney (after adjusting both AVR and blood pressure measurements for pulse wave augmentation index) suggested they are acting independently on AVR. However the same drop in correlation coefficient was not observed in Korcula, and therefore it would be of interest to see whether this result replicates in other larger studies.

Results obtained from this correlation analysis largely confirmed previously reported associations of some cardiovascular risk factors with retinal vessel diameters and fractal dimensions, validating our measurements. In addition some new significant correlations were reported that should guide the future epidemiological and genetic research, especially in terms of vessel diameters and pulse wave analysis traits. Data also showed the lack of significant correlation for a large number of pairs of traits tested, suggesting a need for more detailed and careful selection of the covariates in the future studies.

Chapter 6. Genome wide association studies of retinal vessel traits

6.1. Introduction

This chapter describes the process of genome wide association studies of 10 traits derived from the retinal vessel vasculature. Each trait has been tested in the two populations from Orkney and Korcula and joint meta-analysis was performed. Two studies so far reported genome-wide association scans for retinal vessel calibers (CRVE (Ikram et al. 2010) and CRAE (Sim et al. 2013)) in 8 populations of European ancestry totaling 22010 individuals for CRVE and 7 populations totaling 22661 for CRAE analysis. These uncovered four genome-wide significant loci associated with CRVE and one with CRAE, allowing to look-up if these results replicate in our populations. Since two genetically isolated populations were used here, it is also possible that the allele frequency changes within these populations would enable the detection of genetic variants that have lesser allele frequencies in the larger population (used in the previously published GWAS). In addition, here we use very uniform phenotypes, acquired in identical manner in both populations, which has not been the case in the larger published GWAS where slightly different measurements of retinal widths were used between the many populations included. Finding potential novel associations not identified in the larger meta-analysis studies and finding potential association for traits never before tested for genetic association was the primary purpose of the work reported in this chapter. Directly genotyped data as well as imputed data were analysed for each trait, and replication was sought in an independent population form Scotland (Lothian Birth Cohort) for any potential genome-wide significant findings. Pathway analysis was also performed for the whole group of retinal traits results, selecting each GWAS SNPs with p values $<5\times10^{-5}$ and testing for non-randomness of these signals.

6.2. Methods

6.2.1. ORCADES data preparation and quality control

Prior to the start of this project 758 individuals from Orkney have been genotyped using an Illumina HumanHap300 v2 array and converted to a format for genome wide data analysis. A second batch of 162 individuals was genotyped just before the start of this project using the Illumina Human 370CNV Quad array and was available to be merged with the previously genotyped individuals. Genotypes clustering and calling was done using Illumina Genome Studio software by C.H.. Genotyping data for these 162 individuals were then merged with the 758 individuals in plink using *"–merge"* option and converted to GenABEL format, using available GenABEL package option (*convert.SNP.illumina*). Before the quality control (QC) procedures there were 894 individuals and 351251 SNPs (the combination of typed subjects and SNPs in the first and second genotyping stage).

Appropriate quality control of the genotyped data is essential when executing association analysis of large-scale SNP data, primarily to avoid an abundance of false positive and false negative results arising from genotyping errors or population stratification (Cardon and Palmer 2003; Anderson et al. 2010). In addition a strict QC protocol is essential to allow for the most accurate data imputation of the genotyped data, which was done subsequently.

293687 SNPs passed QC criteria (SNPs with call rate≤0.98, overall call rate per individual (perid.call) ≤ 0.97, minor allele frequency (maf) ≤ 0.02, Hardy Weinberg equilibrium (HWE) departure p.level ≤ 1.e-10, heterozygosity fdr ≤ 0.01 were removed). Before the QC there were 351251 markers in total. 10908 (3.105472%) markers

excluded as having low (<2%) minor allele frequency. 56299 (16.02814%) were markers excluded due to low (<98%) call rate and 258 (0.07345175%) markers were excluded because they were out of HWE (P <1e-10).

Before QC in GenABEL, 2 individuals (ORCA1033, ORCA1142) were removed to a known sample mix-up. Six individuals in total failed QC. Three were removed due to the mismatch between reported gender and chromosome X-linked markers genotypes. Even though X chromosome aneuploidy or gender reassignment could explain this mismatch, it is more likely that mislabeling of a DNA sample or another pre-analytical error has occurred. Genotyped call rate and average heterozygosity of an individual are used to detect individuals with poor quality genotype data. Two individuals were removed because of the low call rate, which can signify a low concentration or degraded DNA sample. One individual failed the heterozygosity test (HET >= 0.3641862) indicating possible sample contamination sample or a population admixture which also indicates the subject does not represent the population accurately. Relationships between the samples were tested by calculating the Identity by state (IBS) sharing levels, which enables the detection of twin or duplicated samples.. 1 individual was removed due to the very high IBS-sharing with another sample (IBS>=0.95), indicating a possible duplicate.

We did not exclude any individuals after inspection of ancestry principal components plots based on IBS sharing distances (Figure 6.1.) as there were no significant outliers and inspection of pedigrees of multiple individuals from the extremes of principal components 1 and 2 show them to be fully Orcadian, but from different isles.

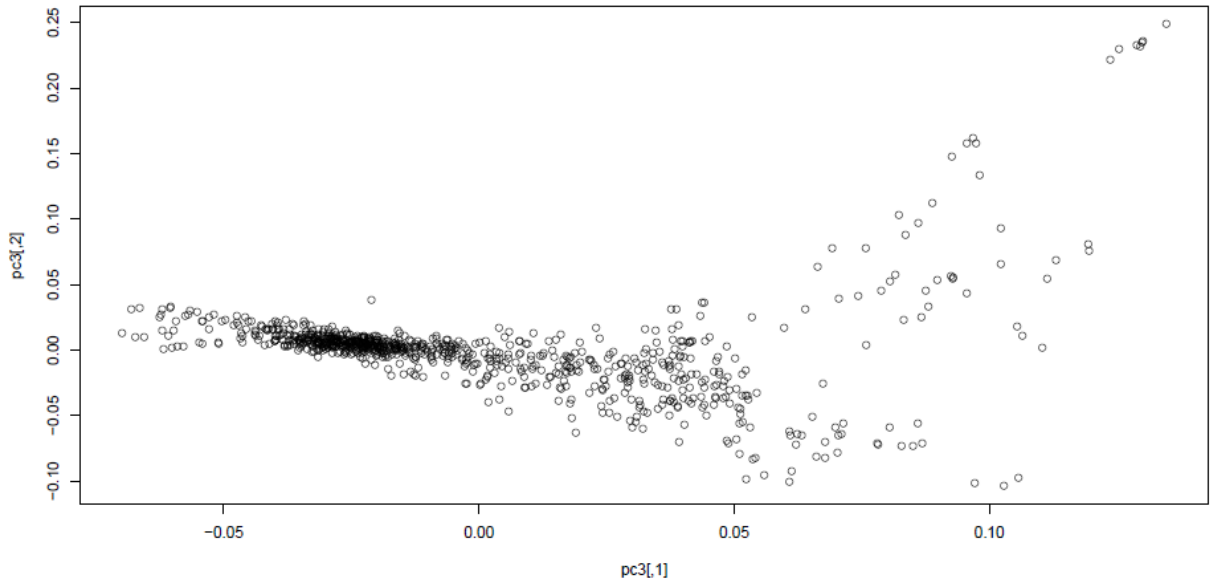


Figure 6.1. Ancestry clustering for Orkney subjects passing QC measures. Clustering was achieved by multi-dimentional scaling of the matrix of pairwise genetic distances derived from the identity by state sharing statistics obtained using the ibs function in the GenABEL R package.

6.2.1.1. Data imputation

889 genotyped individuals and 293,687 SNPS were then used for imputation to the CEU HapMap2 reference build 36 release22. Additionally all the SNPs with AT, TA, GC, CG genotypes were removed (due to possible discrepancies between Illlumina strand designation and reference haplotypes) which left 293,600 SNPs. Imputation of the genotyped data was done using MACH software. This allowed 2.74 million genotypes to be estimated. Details on the imputation procedures are given in Chapter 2. The data was therefore ready for testing for association of common SNPs with quantitative retinal vessel traits.

6.2.3. Korcula genotyping and quality control

Details on the genotyping and subsequent QC procedure of the Korcula data are given in Chapter 4. Identical QC procedures as for Orkney were applied, which left 345555 SNPs and 898 individuals to use in the association analysis, and subsequent data imputation. Data management and imputations was done previous to this project by C.H., V.V and J.H.

6.2.4. Genome wide association analysis

For genome wide association analysis of genotyped data, the GenABEL (Aulchenko et al. 2007) library implemented in the statistical program R was used. It provides facilities for storage and manipulation of large amounts of data, very fast tests for GWA analysis in particular tests accounting for relatedness, and special functions to analyse and graphically present the results of GWA analysis. As mentioned previously, individuals from Orkney and Korcula represent isolated population groups and as many participants of the study are related this source of positive confounding (extended pedigree works as a confounder) needs to be taken into account in the analysis. We expect an inflation of significant results (pedigree inflates the resulting null distribution of chi-square test statistic by a constant λ) due to increased sharing of alleles between related individuals and familial covariance in the tested traits. In a genetically homogenous population, like the isolated populations, it can be shown that the constant λ is a function of a trait's heritability and pedigree structure. A mixed model accounting for a random polygenic inheritance is a standard approach to correct for this non-independence (as described in

more detail Chapter 4 p120) when heritability estimates procedures were explained). Various covariates can be included in the polygenic model, including genetic markers that need to be tested for association with a specific phenotype. However when large number of markers are tested, maximum likelihood(ML)/restricted maximum likelihood (REML) solutions (used for estimating the model parameters) may take very long time to run . Therefore fast approximate tests were developed for the purposes of GWA association analysis in samples of relatives. Two of fast approximations are available to use in GenABEL. FAmily-based Score Test for Association (FASTA, Chen & Abecasis, 2007) and Genome-wide Rapid Analysis using Mixed Models And Score test (GRAMMAS, Amin et al. 2007). Both tests are based on the calculation of the polygenic model of inheritance and are run in two stages.

The first step is that used for heritability estimation (Chapter 4). Briefly a mixed model is fitted that uses the genomic kinship matrix from genome-wide data and ML/REML method is used to estimate the random effects' variances (the polygenic effect and error terms) together with relevant covariates effects fitted as fixed effects . The vector of residuals and the inverse of the variance-covariance matrix calculated in this stage are then used in the second step of the analysis, in which the FASTA score statistic is used to test for association between a trait and genetic marker added as a fixed effect.

$$T_F^2 = \frac{\left((g - E[g])^T \cdot (\Phi \cdot \hat{\sigma}_G^2 + I \cdot \hat{\sigma}_e^2)^{-1} \cdot (Y - \hat{\mu})\right)^2}{(g - E[g])^T \cdot (\Phi \cdot \hat{\sigma}_G^2 + I \cdot \hat{\sigma}_e^2)^{-1} \cdot (g - E[g])}$$

Where (g - E[g]) is the additively coded observed genotype minus the expected average genotype based on allele frequency. Φ is the kinship matrix, σ_g^2 the estimated polygenic variance, σ_e^2 the residual environmental variance and I the identity matrix. *Y* is the phenotype value and μ is the population mean. It can be shown that T_F^2 follows χ_1^2 when

pedigree structure is correct. In an additive coding, effect for the homozygous for the effect allele is twice that of the effect of the heterozygous genotype.

Imputed data were analysed with ProbABEL software (Aulchenko et al. 2010) is used. When dealing with imputed or high-throughput sequencing data, for many of loci there is uncertainty about the genotypic status of the person (we deal with probability distribution, based on the observed information). The degree of confidence about the real status of a genotype is measured with this probability distribution. To generate unbiased estimates of association parameters and to take into account the probability distributions, ProbABEL carries out association analysis by means of regression of the outcome of interest onto estimated allelic doses, which take a continuous value between 0 and 2 . It is designed to carry out the regression analysis in fast and memory efficient ways, since it deals with millions of SNPs.

Genome wide association analyses were done for each population separately, followed by a meta-analysis for each of the traits. Inverse variance weighted meta-analysis was applied (explained in detail in Chapter 2).

The standard Bonferroni correction for multiple testing was applied. Results were considered significant if te p-value was less than $5x10^{-7}$ for genotyped data, and $5x10^{-8}$ for imputed data. Haploview v4.0 software was used for the graphical representation of the genome-wide analysis results (Manhattan plots), and SNAP software version 2.2 (Broad institute) was used for the creation of regional association plots. Using genotyped data allows us to usea greater p-value to declare genome-wide significance as less markers are tested, also insures better genotype call certainty (genotype clustering could be checked back to insure quality of the call) and in our case allows to test association with X-linked markers as those were not imputed.

Genome wide significant results were tested for replication in an independent population from Scotland (LBC). Lothian Birth Cohost study is a follow up cohort study totaling 1091 participants, 396 of which had retinal images grade by SIVA.

6.3. Results

Here the genome wide analyses of 10 different quantitative retinal vessel traits are presented. The meta-analysis of genotyped data revealed one genome wide significant result, which however did not replicate in a small replication sample from Scotland. Several suggestive results are also reported and indicate the need for further studies involving larger number of subjects. Replication analysis of the already published trait-locus associations are also presented. All the suggestive associations reported here represent novel findings.

6.3.1. GWAS of retinal vessel caliber measurements

A genome wide association study was performed for each of the retinal vessel caliber traits in zone B. Covariates used in this analysis were based on the best polygenic model identified in the heritability analysis, with the exception of AVR (Table 6.1). Covariates identified through correlations analysis were also included in two additional analysis of AVR (GWAS of genotyped data) in order to explore their effect on AVR and to

compare it to the effect of the standard covariates. Covariates identified in the subjects from Orkney were used in GWAS of both populations, in order to use the results of the same models in the subsequent meta-analysis.

Trait-Orkney	**Best model covariates**	**Trait-Korcula**	**Best model covariates**
CRAE zone B	Age avsbp	CRAE zone B	avsbp avdbp
CRVE zone B	Age BMI	CRVE zone B	avsbp
AVR zone B	avdb (bodyfat contenet , pulsewave augmentation index)	AVR zone B	avdbp (bodyfat content , pulsewave augmentation index)

Table 6.1. List of covariates selected based on the best fitting polygenic model, for each caliber trait tested.

6.3.1.1. Central retinal arterial equivalent

Genetic association analysis of retinal width was performed for the zone B, because of the high correlation with the width measurements in the zone C. Results of the population specific and joint analyses of this trait in Orkney and Korcula populations are given bellow.

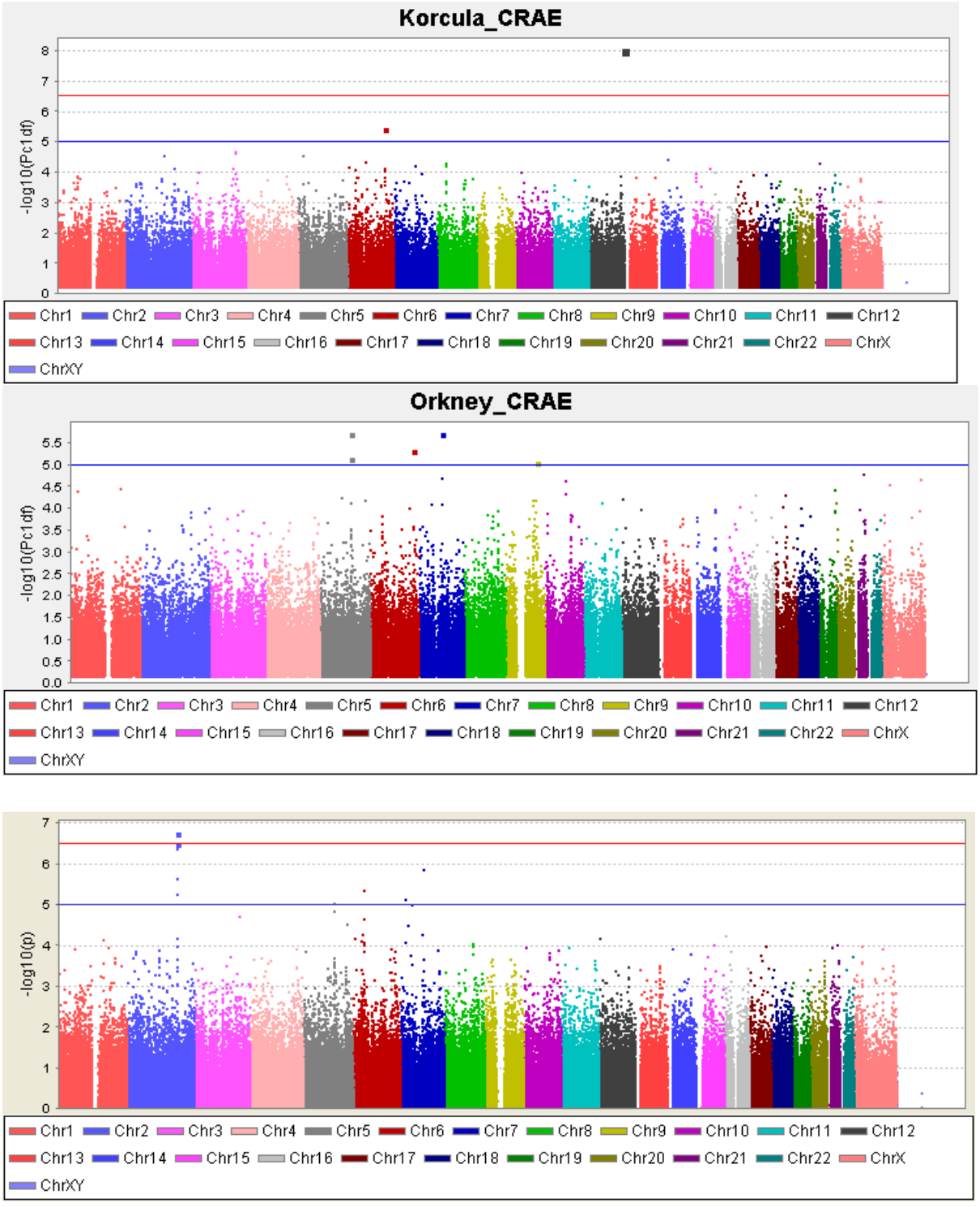


Figure 6.1. Manhattan plot of *P*-values for the association of genotyped SNPs with CRAE in the meta-analysis of ORCADES and CROATIA-Korcula studies. Horizontal lines represent the threshold of suggestive significance (in blue) and genome-wide significance (in red).

name	A1	A2	CHR	position	beta	se beta	frequency (A2)	Beta Orkney	Beta Korcula	se beta Orkney	se beta Korcula	p	gene
rs2288569	A	G	2	179166837	-3.790	0.726	0.870	-3.507	-4.276	0.913	1.197	1.80E-07	TTN
rs3829747	A	G	2	179105807	-3.714	0.726	0.870	-3.507	-4.070	0.913	1.197	3.15E-07	TTN
rs744426	A	G	2	179144266	-3.693	0.727	0.871	-3.507	-4.012	0.913	1.199	3.73E-07	TTN
rs3731749	A	G	2	179114437	-3.696	0.730	0.871	-3.474	-4.070	0.922	1.197	4.19E-07	TTN
rs12707538	A	G	7	82497792	3.304	0.685	0.854	3.830	1.943	0.807	1.298	1.42E-06	PCLO
rs2303838	A	G	2	179153185	-3.079	0.651	0.837	-2.815	-3.611	0.796	1.131	2.26E-06	TTN
rs846523	A	G	6	40673918	2.290	0.498	0.634	2.230	2.443	0.588	0.942	4.36E-06	LRFN2
rs3731746	A	G	2	179139243	-3.025	0.667	0.847	-2.743	-3.573	0.821	1.144	5.75E-06	TTN
rs1011559	A	C	7	20670473	2.975	0.664	0.150	2.658	4.097	0.752	1.416	7.55E-06	ABCB5
rs10061888	A	C	5	111652801	2.561	0.578	0.215	3.189	0.772	0.672	1.134	9.40E-06	EPB41L4A
rs6962367	A	C	7	42518912	-2.542	0.576	0.757	-2.653	-2.238	0.673	1.113	1.01E-05	GLI3
rs1554624	A	G	5	111683281	3.176	0.732	0.124	3.830	1.354	0.854	1.425	1.45E-05	EPB41L4A
rs1033482	A	G	6	40650051	2.113	0.499	0.636	2.051	2.275	0.587	0.949	2.31E-05	LRFN2
rs31694	A	G	5	159225281	2.461	0.590	0.787	2.822	1.638	0.707	1.067	2.99E-05	PWWP2A, ADRA1B, TTC1
rs160375	A	G	7	28621226	-2.145	0.517	0.291	-2.000	-2.553	0.602	1.009	3.30E-05	CREB5

Table 6.2. Markers displaying the most significant association p-values ($p<5x10^{-5}$) in genotyped data meta-analysis of CRAE adjusted for systolic blood pressure and age. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

Analysis of the genotyped data revealed one genome wide significant result when Korcula was analyzed separately, while the GWAS of the Orkney data did not reveal any SNPs reaching the genome wide threshold. The most significant result was reported for rs12366899 on chromosome 12 ($p=8.82x10^{-9}$). However this SNPs was not genotyped in the Orkney samples, nor was it present in the imputed data. It was also detected as significant in the GWAS of Hb1c (C. Franklin thesis, 2010) as well as in several different traits analysed in Korcula, and has been suggested to be a false positive artifact. It is important to point out that in both analysis highly significant results were detected for the same SNP, with small number of subjects tests (CRAE analysis N=374, Hb1c analysis N=411). These two traits did not show a significant correlation in Korcula, although wider arteriolar and venular calibers were associated with higher glycosylated hemoglobin in WESDR (Klein et al. 2003), and with wider venular calibers in Singapore Malay Study (Sun et al. 2009).

When the population specific results were analyzed in a joint meta-analysis a region on chromosome 2 was significantly associated with CRAE. Top ranking SNP (rs2288569) reported a p-value of 1.8×10^{-7}, followed by three additional SNPs passing the Bonferoni corrected significance threshold ($p < 5 \times 10^{-7}$). These markers are located in the region of chromosome 2q32 and fall within the *TTN* (titin) gene, which codes for a large abundant protein of striated muscle. This signal was supported by additional SNPs showing low p-values in the same region of chromosome 2 (Figure 6.2), and showing comparable effect sizes in both populations (Table 6.2)

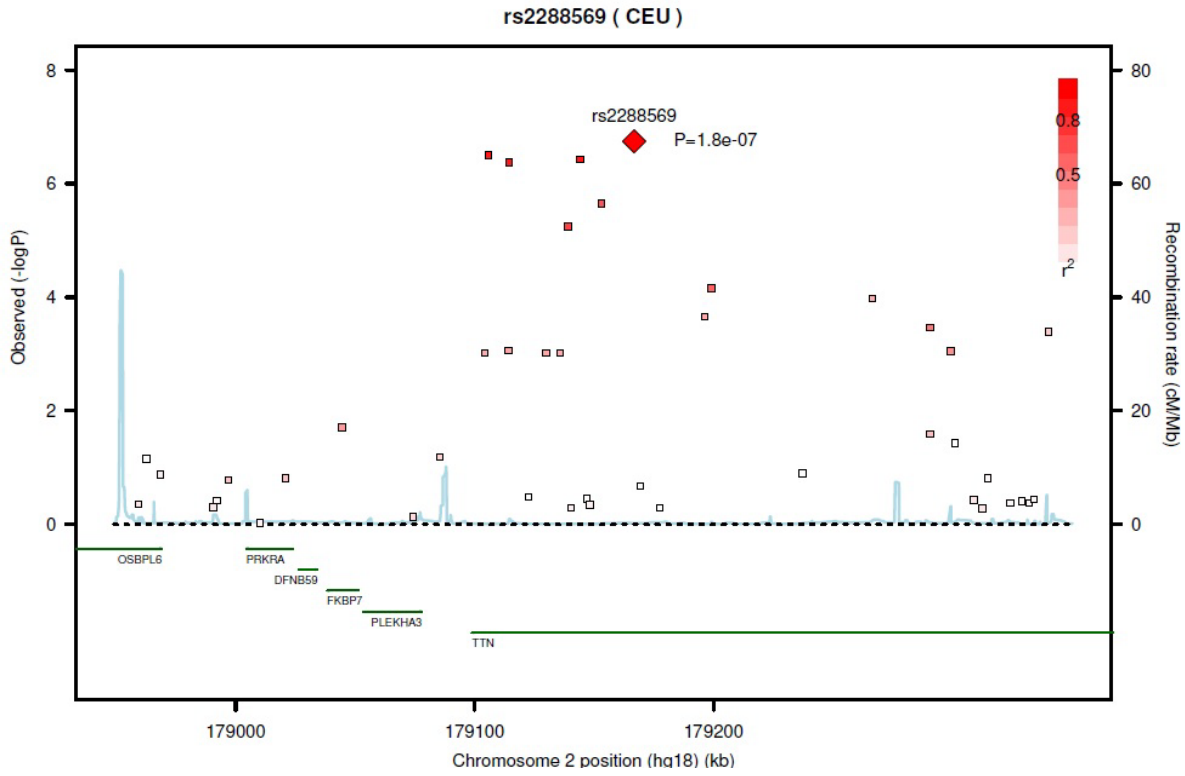


Figure 6.2. Association plot of the region on chromosome 2 associated with retinal arteriolar width (CRAE), by the combined population meta-analysis of Orkney and Korcula. A region of 500 kb around the association signals is displayed. The degree of LD between the top genotyped SNP (rs2288569) and other 42 SNP tested in the meta-analysis of this region is indicated by shades of red (bright red indicating high LD). Recombination rate is displayed by a blue line with scale on the right-hand axis. Genes are displayed in the bottom part of the figure, with the arrow showing the direction of transcription.

The LBC cohort was used for the replication of these significant results. The top 4 markers were not significantly associated with CRAE in the LBC (p=0.18-0.19), however the direction of the effect were comparable to the results in Orkney and Korcula (rs2288569 β=-1.73 (1.32), rs3829747 (-1.73)(1.32)), but the magnitude smaller.

Analysis of the imputed data did not display any new significant association signals (Figure 6.3), The association signals reported in the analysis of the genotyped data were confirmed, however none of the markers reached genome wide significance threshold for imputed data ($p<5x10^{-8}$). Analysis of the imputed data also suggested potential new association between a different region on chromosome 2 (2q37.1) and also a region on chromosome 7 (7q21.11) within the *PCLO* gene (piccolo presynaptic cytomatrix protein). Modest genomic control inflation was observed for each population (genotyped: λ_{orkney}=1.001, $\lambda_{korucula}$=1.006, $\lambda_{meta\text{-}analysis}$=1.014; imputed: λ_{orkney}= 0.99 ,$\lambda_{korcula}$=1.003, λ_{meta}=1.004), and resulting qq plots illustrated minimal genomic controls inflation (Figure 6.4) and excess of positive results likely to be true positives.

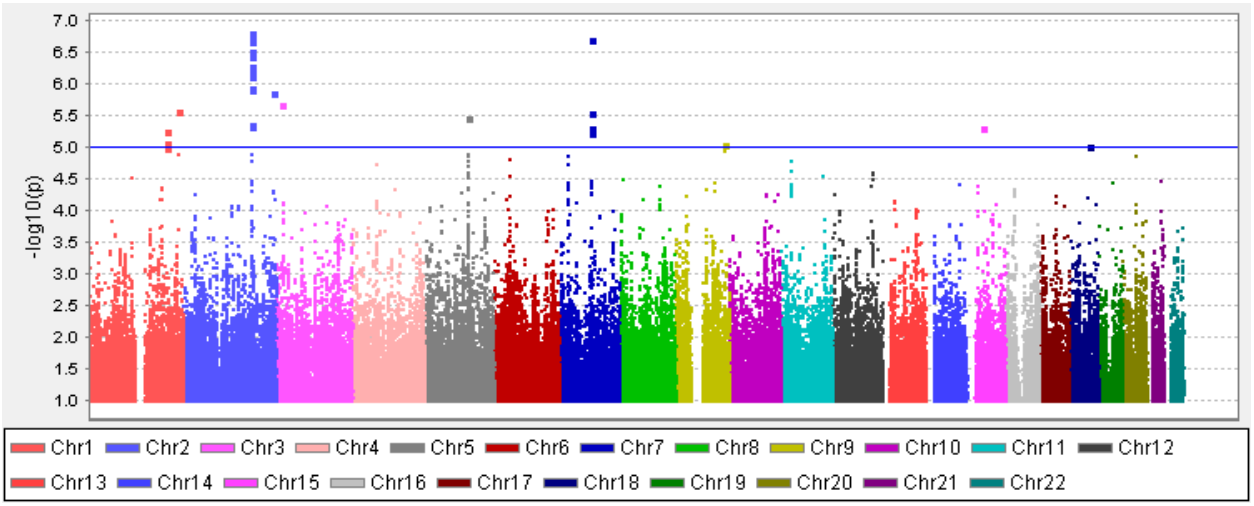


Figure 6.3. Manhattan plot of *P*-values for the association of imputed SNPs with CRAE in the meta-analysis of Orkney and Korcula islands. Horizontal lines represent the threshold of suggestive significance (in blue).

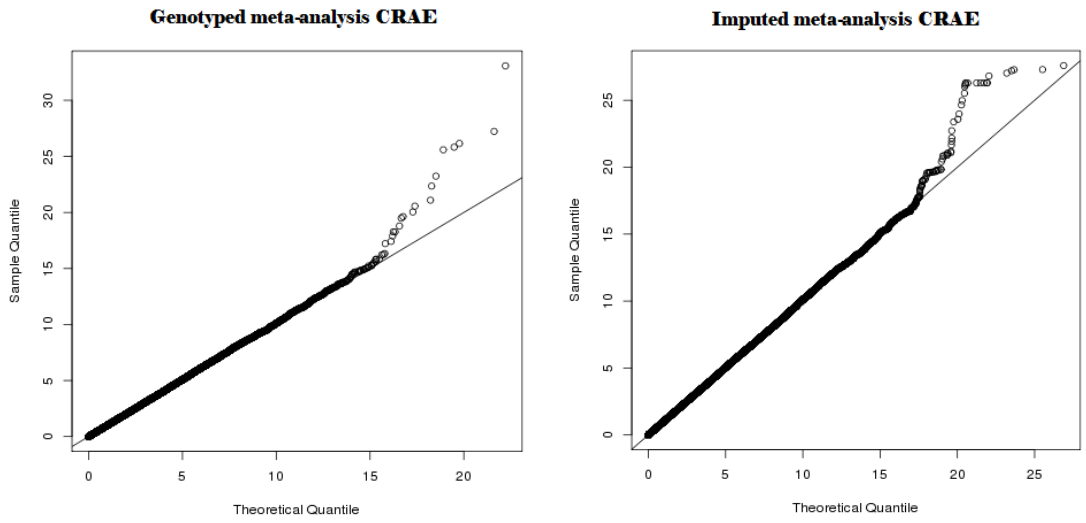


Figure.6.4 Statistics from a genome wide meta-analysis on genotyped and imputed data, adjusting for age and systolic blood pressure, performed on 809 Orkney and 389 Korcula participants using SNPs passing quality control checks are plotted (-log10 (P-value) on the x axis, chromosomal location on the y axis for the Manhattan plot-Observed and expected chi-square values were used for the QQ –plot).

Sim et al. 2013 reported genome-wide significant association between CRAE and marker rs2194025 on chromosome 5q14.3 (overall meta-analysis $p=2.11x10^{-12}$, β=-1.6(0.23SE)) located between (*TMEM161B* and *MEF2C* genes) . The marker was imputed in our datastes, and presented the same direction of the effect in the two Scottish populations (Orkney p=0.6, β=-0.64(1.22SE) and LBC (p=0.99, β=-0.72(1.43SE)), with smaller effect sizes. Effect size of the same SNP was much smaller and in the opposite direction in Korcula (p=0.84, β=0.025(0.13SE). The same study reported additional two markers showing suggestive associations with CRAE (rs3744061, rs2281827). The marker on chromosome 17 (rs3744061) showed comparable effect size for Orkney (combined published meta-analysis β=-0.86(0.13SE), Orkney β=-0.8(0.59SE)), and lower for LBC (β=-0.35(0.89SE). Marker on chromosome 13 (rs2281827) presented same direction of the effect in Orkney (combined published meta-analysis β=-0.70(0.16SE), Orkney β=-0.81(0.72SE)) and Korcula (β=-0.10(0.92SE)), but not in LBC (β=0.06(1.08SE)).

6.3.1.2. Central retinal venular equivalent –CRVE

name	Eff allele	CHR	position	beta	sebeta	Eff allele freq	beta Orkney	beta Korcula	se beta Orkney	se beta Korcula	p	GENE
rs7548537	G	1	40384956	4.04	0.93	0.16	4.35	3.03	1.06	1.93	1.35E-05	CAP1, RLF, ZMPSTE24, PPT1
rs4823569	G	22	45376808	-3.44	0.80	0.75	-3.07	-4.32	0.95	1.47	1.70E-05	GRAMD4, CERK, GRAMD4
rs7542378	C	1	14817628	3.33	0.78	0.74	4.00	1.89	0.94	1.37	1.80E-05	RP1-21O18.1, RP1-21O18.1
rs11786927	G	8	94744788	7.45	1.75	0.96	8.02	6.16	2.10	3.14	2.02E-05	FAM92A1, RBM12B, TMEM67
rs10927473	C	1	14832760	-3.24	0.77	0.27	-3.95	-1.68	0.93	1.38	2.72E-05	RP1-21O18.1, RP1-21O18.1
rs9371850	G	6	1.55E+08	2.77	0.66	0.50	3.43	1.32	0.80	1.18	2.78E-05	TIAM2, RBM16, CLDN20, TFB1M
rs5769010	G	22	45371498	-3.11	0.74	0.67	-2.62	-4.24	0.89	1.35	2.89E-05	GRAMD4, CERK
rs9724754	C	1	1.77E+08	-5.22	1.26	0.92	-6.33	-3.06	1.55	2.16	3.45E-05	TOR3A, ABL2, SOAT1, FAM20B, C1orf125, TDRD5, NPHS2
rs1835661	G	20	12973560	-3.07	0.74	0.72	-3.56	-2.06	0.91	1.30	3.80E-05	SPTLC3
rs10484851	G	6	39086749	-5.51	1.34	0.93	-5.15	-6.23	1.64	2.34	3.95E-05	GLP1R, DNAH8
rs4950887	G	1	2.01E+08	3.54	0.86	0.21	3.05	4.70	1.03	1.58	4.01E-05	RABIF ,CYB5R1, KLHL12, ADIPOR1
rs8179338	G	1	2.01E+08	-3.53	0.87	0.79	-3.05	-4.71	1.03	1.60	4.42E-05	RABIF ,CYB5R1, KLHL12, ADIPOR1
rs7204135	G	16	49183696	-2.85	0.70	0.51	-2.42	-3.89	0.83	1.29	4.64E-05	NKD1
rs9919219	G	1	2.01E+08	3.52	0.86	0.21	3.03	4.71	1.03	1.60	4.68E-05	RABIF, CYB5R1, KLHL12, ADIPOR1
rs10494544	G	1	1.8E+08	5.30	1.30	0.93	5.75	4.02	1.51	2.56	4.78E-05	CACNA1E

Table 6.5 Markers displaying suggestive association p-values ($p<5x10^{-5}$) of genotyped data meta-analysis p-values of CRVE adjusted for systolic age and BMI. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported .List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

Genetic association analysis was carried out for CRVE in the same manner as described for CRAE. No genome-wide significant results were reported; the Manhattan plot of the meta-analysis of the imputed data is displayed in Figure 6.4. and suggestive SNPs are given in Table 6.3.

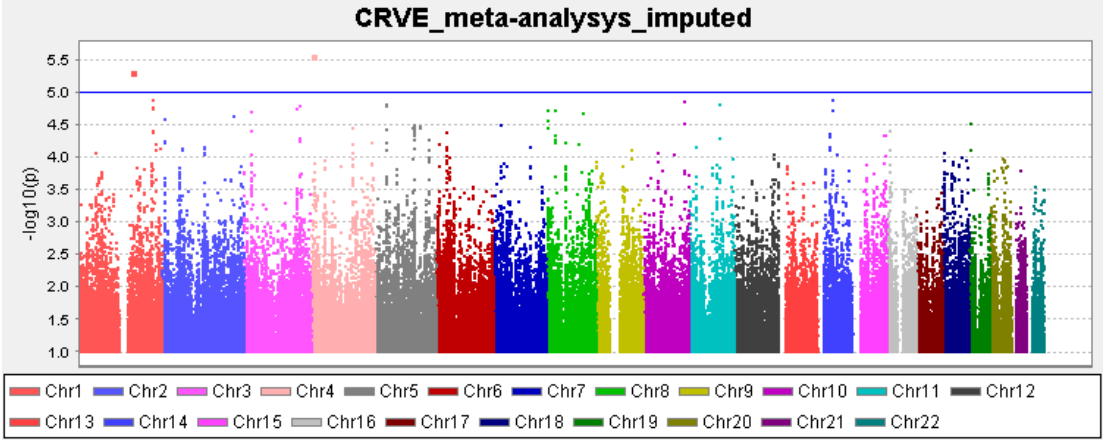


Figure 6.4. Minus log-transformed P-values are shown in a Manhattan plot (signal intensity plot) relative to their genomic position. Horizontal lines represent the threshold of suggestive significance (in blue).

rs6849076 displayed the highest association p-value in the analysis of the imputed data ($2.72x10^{-6}$). It is located on chromosome 4 within *ABLIM2* gene, coding for actin binding LIM protein, but failed to reach genome-wide significance. Modest genomic control inflation was observed for each population (genotyped: λ_{orkney}=1.01, $\lambda_{korucula}$=1.008, $\lambda_{meta\text{-}analysis}$=1.0; imputed: λ_{orkney}= 1.009 ,$\lambda_{korcula}$=1.0002, λ_{meta}=1.0).

Additionally, results for the association of CRVE and four loci previously reported to be associated with CRVE (Ikram et al 2010.) are reported here. Only one of the SNPs was genotyped in all three populations (rs225717), and displayed nominally significant association with CRVE in Orkney (p=0.0092), but not in Korcula (p=0.81) nor LBC (p=0.17). The direction and the size of the effect were comparable in all three population to the published meta-analysis (Table 6.6) (Ikram et al. 2010). Nominally significant p-value was obtained after the results from all three of our populations were meta-analyzed (p=0.048, β=-1.99(1.01SE)). In addition rs10774625 displayed nominally significant association signal in Korcula (p=0.035), but not in Orkney (p=0.44) or LBC (p=0.36). However directions and magnitudes of the effect of this allele were not homogenous between our three populations (Table 6.5). rs2287921 displayed homogeneous effect size and direction in Orkney and LBC, and was comparable to the meta-analysis result.

SNP	Orkney	Korcula	LBC	Published meta-analysis (Ikram et al. 2009)
rs225717 (genotyped)	**P=0.009 B=-2.59(1.69) EAF=0.19**	P=0.81 B=-0.77(2.66) EAF=0.26	P=0.17 B=-1.91(1.42) EAF=0.25	$P=2.11x10^{-16}$ B=-1.9(0.23) EAF=0.23
rs2287921 (imputed)	P=0.11 B=-1.31 (0.84) EAF=0.47	P=0.86 B=0.01(0.07) EAF=0.48	P=0.22 B=-1.53(1.27) EAF=0.46	$P=1.61x10^{-25}$ B=-2.1(0.20) EAF=0.47
rs10774625 (imputed)	P=0.44 B=-0.64(0.83) EAF=0.49	**P=0.034 B=-0.15(0.07) EAF=0.48**	P=0.36 B=1.12(1.23) EAF=0.48	$P=2.15x10^{-13}$ B=1.5(0.22) EAF=0.48
rs17421627 (imputed)	P=0.69 B=0.74(1.87) EAF=0.05	P=0.79 B=0.03(0.13) EAF=0.09	P=0.65 B=0.93(2.11) EAF=0.09	$P=7.32x10^{-16}$ B=3.0(0.73) EAF=0.08
rs1035387 (imputed)	P=0.72 B=0.32(0.89) EAF=0.39	P=0.37 B=-0.07(075) EAF=0.36	P=0.22 B=1.52(1.27) EAF=0.39	$P=3.8x10^{-7}$ B=1.0(0.02) EAF=0.39

Table 6.5. Look-up for SNPs associated with CRVE in a published meta-analysis of 8 different populations. Association p-value, effect and its standard errors, along with the frequency of the effect allele are given for three populations analysed (Orkney, Korcula, LBC).

6.3.1.2. Arteriolar to Venular diameter Ratio – AVR

Analysis of the AVR was performed with two sets of covariates. The first set was the one determined in the heritability analysis using covariates that most studies could reproduce and used throughout this study. Since AVR presented significant correlations with both pulse-wave augmentation index and percentage of body fat in both Orkney and Korcula, additional analysis was done to see whether it impacts the heritability estimates and the results of the genetic association analysis. The main goal was to identify possible novel associations.

Firstly, the analysis of genotyped and imputed data including standard covariates was performed. No significant associations were detected in any population or meta-analysis of genotyped or imputed data (Figure 6.5.). Very modest genomic control inflation was

observed for each population (genotyped: λ_{orkney}=1.004, $\lambda_{korucula}$=1.02, $\lambda_{meta\text{-}analysis}$=1.002; imputed: λ_{orkney}= 0.99 ,$\lambda_{korcula}$=1.003, λ_{meta}=1.0).

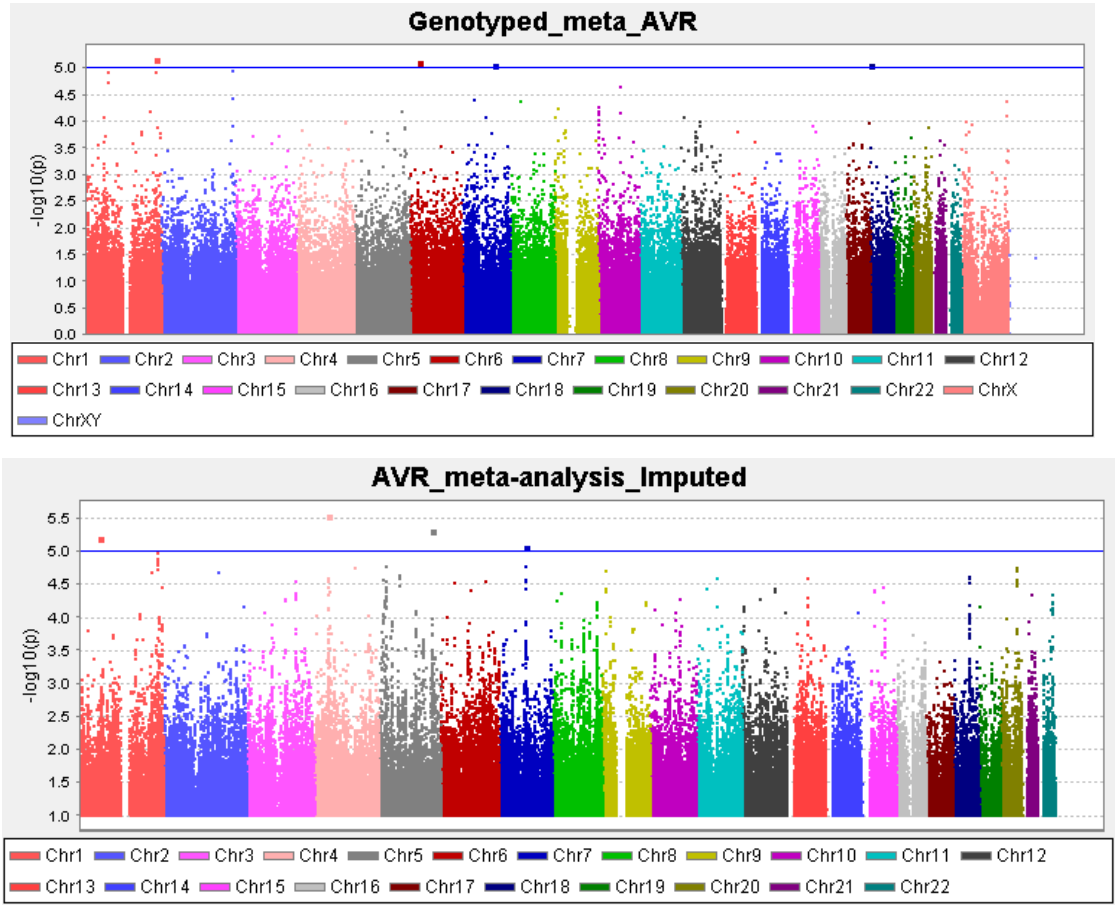


Figure 6.6. Minus log-transformed P-values are shown in a Manhattan plot (signal intensity plot) relative to their genomic position. Horizontal lines represent the threshold of suggestive significance (in blue).

name	effallele	CHR	position	beta	sebeta	Effallele freq	beta Korcula	beta Orkney	se Korcula	se Orkney	p	gene
rs12734075	G	1	2.28E+08	-0.25	0.05	0.22	-0.36	-0.19	0.09	0.06	9.45E-07	NUP133,ACTA1, ABCB10
rs2884064	G	1	2.3E+08	-0.29	0.06	0.87	-0.35	-0.23	0.09	0.09	5.90E-06	SIPA1L2,DISC1
rs2026	G	2	2.34E+08	-0.27	0.06	0.87	-0.42	-0.20	0.11	0.08	1.27E-05	ATG16L,USP40,DGKD
rs7068299	G	10	72992635	0.21	0.05	0.77	0.18	0.23	0.08	0.06	1.93E-05	CDH23
rs7121351	G	11	72086807	0.33	0.08	0.08	0.45	0.26	0.12	0.10	2.06E-05	CENTD2,STARD10
rs277366	G	1	75073505	-0.20	0.05	0.69	-0.19	-0.20	0.08	0.06	2.17E-05	CRYZ, LHX8,TYW3, SLC44A5
rs1563188	G	18	1728361	-0.18	0.04	0.49	-0.20	-0.17	0.07	0.05	2.26E-05	
rs5969812	G	X	1.51E+08	0.15	0.04	0.39	0.19	0.14	0.07	0.04	2.29E-05	PASD1
rs3122713	G	1	2.11E+08	0.27	0.06	0.87	0.33	0.21	0.09	0.08	2.37E-05	SIPA1L2,DISC1
rs9569076	C	13	54125829	-0.23	0.06	0.18	-0.24	-0.23	0.10	0.07	2.69E-05	ATG16L,USP40,DGKD
rs1897139	G	3	1.43E+08	-0.28	0.07	0.88	-0.39	-0.21	0.11	0.08	2.83E-05	ATF3, FAM71AATP1B3
rs29232	G	6	29719410	-0.18	0.04	0.64	-0.21	-0.17	0.08	0.05	3.56E-05	
rs12839220	G	X	1.51E+08	0.18	0.04	0.80	0.22	0.16	0.07	0.05	3.74E-05	PASD1
rs277355	G	1	75063372	0.19	0.05	0.31	0.18	0.19	0.08	0.06	4.35E-05	CRYZ, LHX8,TYW3, SLC44A5
rs10520531	G	4	1.84E+08	-0.36	0.09	0.94	-0.32	-0.36	0.20	0.10	4.38E-05	ODZ3
rs9905895	G	17	22609947	0.18	0.04	0.61	0.15	0.19	0.08	0.05	4.76E-05	WSB1
rs2306409	G	10	1036712	0.17	0.04	0.51	0.16	0.18	0.07	0.05	4.97E-05	GTPBP4

Table6.6. Markers displaying suggestive association p-values in genotyped data analysis with diastolic blood pressure as a covariate. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

A second set of analysis was done using body fat percentage and pulse wave augmentation index as covariates. Heritability estimates were also evaluated to see whether the inclusion of these novel covariates would modify results. Analysis was done for the Orkney and Korcula populations on directly genotyped data. Percentage of body fat was a significant covariate in both populations (Orkney $p=6x10^{-5}$, Korcula $p=0.025$), but explaining only a small proportion of variance (Orkney 1.6%, Korcula 1.5%). Heritabilty estimates were significant and very similar to the estimates with standard covariates (Orkney h2=0.27, Korcula h2=0.41). Results for pulse wave augmentation index should be taken with caution, since only 343 individuals form Orkney and 207 individuals from Korcula had both AVR and pulse wave augmentation phenotypes

measured. Heritabilty estimates in Orkney when pulse wave was included in the analysis as a covariate did change, h_2=0.35, however this is probably due to the lower number of individuals included in the calculation. Heritability estimates in Korcula was very low and incomparable (h_2=1.01x10^{-8}) to the values determined in the analysis using standard covariates, again most likely to the very low number of subjects included (N=207).

GWAS was also performed using body fat percentage as a covariate, since the pulse wave analysis would include only half of the subjects tested in the original GWAS, and results would not be comparable. No significant associations were detected, nor did the results differ from the original GWAS (Results in Appendix VI).

6.3.2. GWAS of retinal vessel tortuosity measurements

Genome wide association study was performed for each of the retinal vessel curvature tortuosity traits in zone C. Each trait has been rank transformed, because of the departure from normality (Chapter 4). Covariates used in this analysis were based on the best polygenic model identified in the heritability analysis. Analysis of the simple tortuosity measurements were performed, but yielded almost identical results to the curvature analysis, and is therefore not included in this thesis.

6.3.2.1. Arteriolar curvature tortuosity

Firstly genome-wide association analysis was performed using genotyped data for each of the populations separately. This analysis in Orkney yielded one biologically

interesting, yet not genome-wide significant result. The most significant SNP in the analysis of Orkney genotyped data (only reaching suggestive significance; rs2296845, $p=2x10^{-6}$, $\beta=0.25(0.05\ SE)$) is located within the COL4A2 gene, a very good biological candidate gene since it codes for the type 4 collagen subunit; the major structural component of basement membrane (a thin sheet of fibres which lie under the endothelium). This marker however, did not present even nominal significance in Korcula ($p=0.18$, $\beta=0.076(0.072SE)$), nor LBC cohorts ($p=0.19$, $\beta=0.094(0.072SE)$). The direction of the effect was consistent in all three populations, suggesting the need for greater number of subjects to be tested in order to confirm this result found in Orkney. Only very modest genomic control inflation was observed for each population (inflation factor, λ = 1.0 for and 1.005 for Orkney; meta-analysis λ=1.009), suggesting that the observed results were not due to population stratification. Meta-analysis of the imputed data confirmed the association between the same region on chromosome 13 (13q34, *COL4A2*), and indicated a new region on chromosome 12 (12q24) as a possible region of association with cTORTa (Figure 6.6). The genomic inflation factor lambda was 1.003 for Orkney, 1.004 for Korcula, and 1.008 for the joint meta-analysis. The resulting quantile–quantile plot was consistent with a clear excess of true genetic associations .

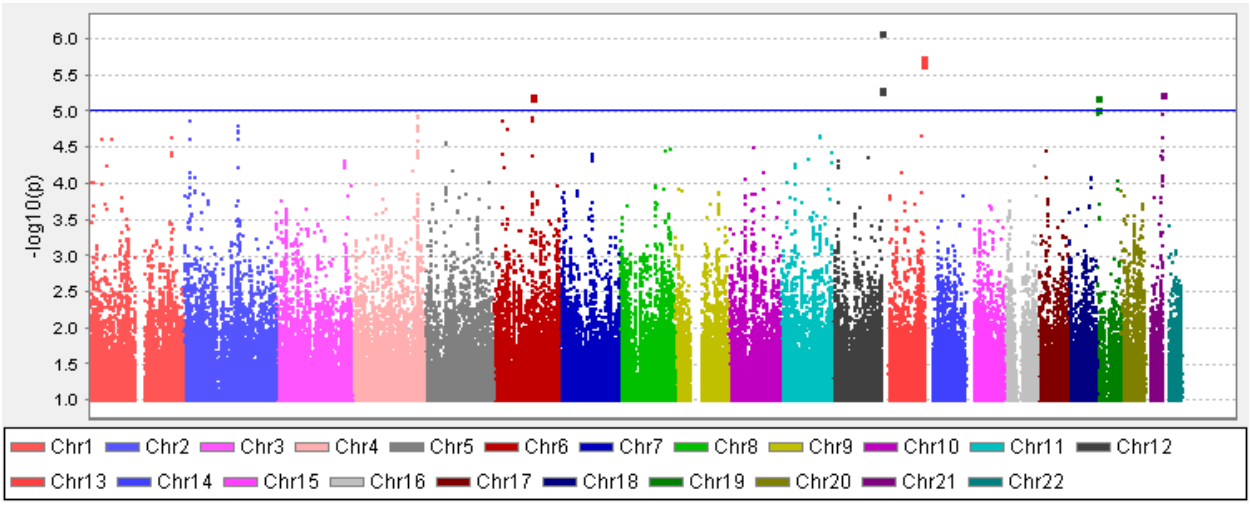


Figure 6.7. Minus log-transformed P-values are shown in a Manhattan plot (signal intensity plot) relative to their genomic position. Results of meta-analysis of imputed data from Orkney and Korcula are given.

SNP	Eff allele	CHR	POS	beta	sebeta	eff allele freq	beta Orkney	beta Korcula	Se Orkney	se Korcula	p	Gene
rs2450546	C	12	129813641	0.06	0.01	0.46	0.06	0.05	0.01	0.08	8.00E-07	STX2,RAN, GPR133,RIMBP2
rs9555699	G	13	109898019	0.06	0.01	0.47	0.06	0.08	0.01	0.07	1.78E-06	COL4A2
rs2296845	G	13	109898948	0.06	0.01	0.47	0.06	0.07	0.01	0.07	1.78E-06	COL4A2
rs4103	C	13	109896227	-0.06	0.01	0.46	-0.06	-0.08	0.01	0.07	2.24E-06	COL4A2
rs2632601	C	12	129834650	-0.06	0.01	0.60	-0.06	-0.03	0.01	0.08	5.04E-06	STX2,RAN, GPR133,RIMBP2
rs4412792	G	12	129834522	-0.06	0.01	0.60	-0.06	-0.03	0.01	0.08	5.04E-06	STX2,RAN, GPR133,RIMBP2
rs4759783	G	12	129831462	-0.06	0.01	0.60	-0.06	-0.03	0.01	0.08	5.04E-06	STX2,RAN, GPR133,RIMBP2
rs7977276	G	12	129835141	-0.06	0.01	0.60	-0.06	-0.03	0.01	0.08	5.14E-06	STX2,RAN, GPR133,RIMBP2
rs221579	C	6	99853258	-0.06	0.01	0.62	-0.06	-0.09	0.01	0.08	6.25E-06	C6orf168
rs734931	G	19	2015016	0.11	0.02	0.10	0.12	0.04	0.03	0.12	6.43E-06	MOBKL2A, C19orf36
rs221582	G	6	99850828	0.06	0.01	0.38	0.06	0.08	0.01	0.08	6.51E-06	C6orf168
rs8105426	G	19	2016196	-0.11	0.03	0.91	-0.12	-0.03	0.03	0.12	9.55E-06	MOBKL2A, C19orf36

Table6.7. Markers displaying suggestive association p-values in genotyped data meta-analysis of curvature arteriolar tortuosity withage and systolic blood pressure as a covariates. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported is given. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

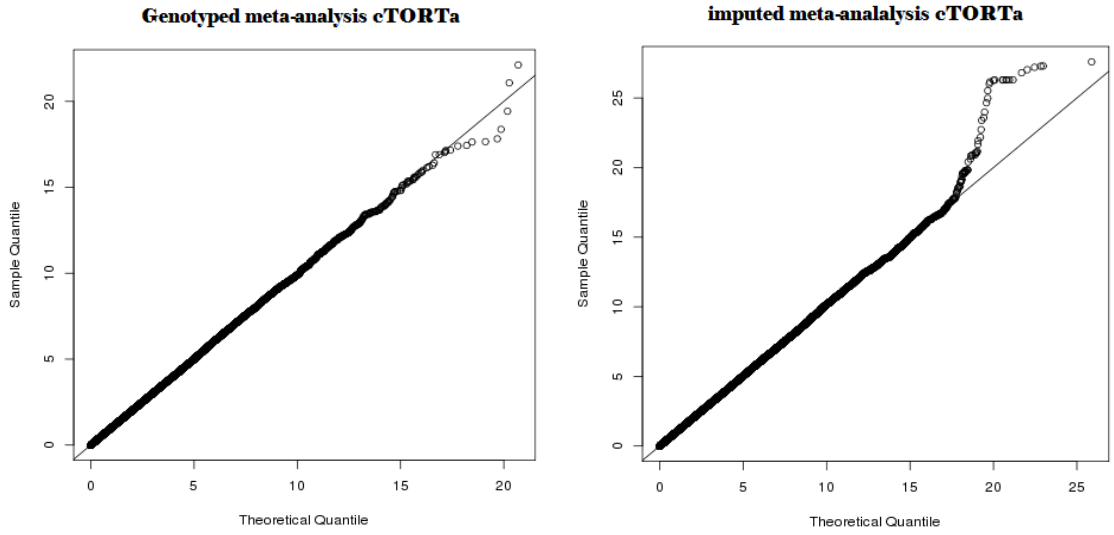


Figure 6.8. Statistics from a genome wide meta-analysis on genotyped and imputed data, adjusting for age and systolic blood pressure, performed on 809 Orkney and 389 Korcula participants using SNPs passing quality control checks are plotted (-log10 (P-value) on the x axis, chromosomal location on the y axis for the Manhattan plot-Observed and expected chi-square values were used for the QQ –plot)..

6.3.2.2. Venular curvature tortuosity

Analysis of venular curvature tortuosity did not reveal any genome-wide significant results in either genotyped or imputed data analysis. The marker with the lowest p-values was rs5952756 located on the X chromosome and reaching significance p-value of 2.88E-05 (genotyped analysis) (Table 6.10). Inflation factor λ (Orkney=1.04/1.006, Korcula=1.08/1.01, meta-analysis= 1.09/1; lambdas are given for genotyped/ imputed data) did not suggest inflation of significant results due to the population stratification.

name	Eff allele	CHR	position	beta	se beta	eff allelefreq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	Gene
rs5952756	G	X	45177745	0.24	0.06	0.11	0.25	0.22	0.07	0.10	2.88E-05	CXorf36, UTX
rs13108458	G	4	32614391	0.21	0.05	0.26	0.23	0.16	0.06	0.08	3.14E-05	
rs2380422	G	2	12266840	0.24	0.06	0.16	0.29	0.11	0.07	0.10	3.97E-05	
rs1982542	G	4	1.26E+08	-0.26	0.06	0.12	-0.20	-0.35	0.08	0.10	4.24E-05	ANKRD50

Table 6.8. Markers displaying suggestive association p-values in genotyped data meta-analysis of venular curvature tortuosity with BMI as a covariate. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported is given. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

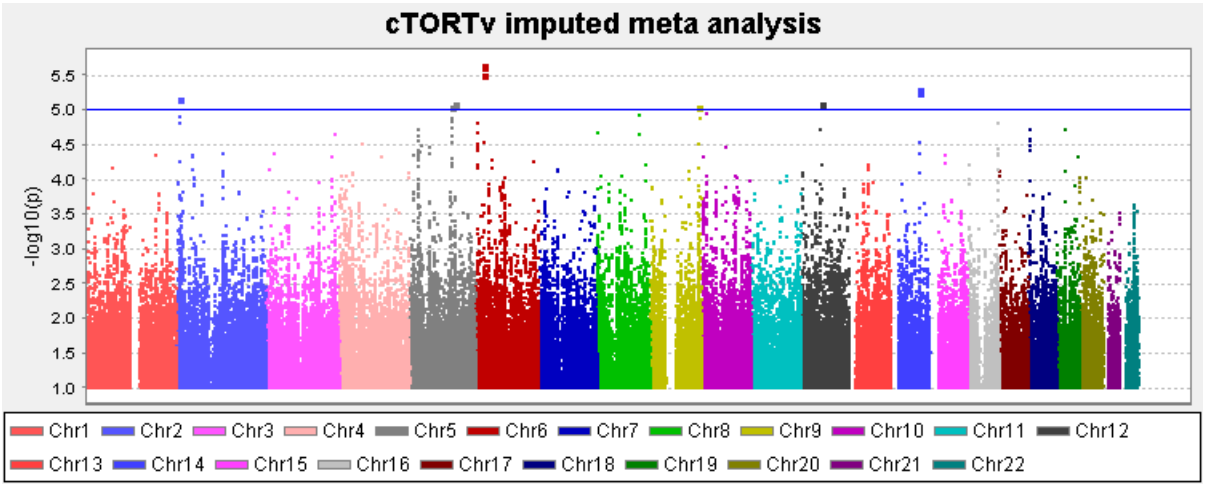


Figure 6.9. Minus log-transformed P-values are shown in a Manhattan plot (signal intensity plot) relative to their genomic position. Results of meta-analysis of imputed data from Orkney and Korcula are given.

6.3.2.3 Total curvature tortuosity

name	eff allele	CHR	position	beta	sebeta	eff allele freq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	gene
rs4434655	G	8	1.39E+08	0.27	0.06	0.12	0.31	0.21	0.08	0.11	1.19E-05	
rs1546331	G	4	1.73E+08	-0.19	0.04	0.55	-0.17	-0.22	0.05	0.08	1.22E-05	GALNT17
rs2220655	G	4	1.73E+08	-0.18	0.04	0.42	-0.21	-0.13	0.05	0.08	2.13E-05	GALNT17
rs200791	G	12	97832413	0.23	0.05	0.18	0.21	0.28	0.07	0.10	2.33E-05	ANKS1B
rs1495324	G	11	40630351	-0.18	0.04	0.59	-0.10	-0.32	0.05	0.07	2.58E-05	LRRC4C
rs2088147	G	11	40625056	0.18	0.04	0.41	0.10	0.32	0.05	0.07	2.77E-05	LRRC4C
rs5908533	G	X	1.42E+08	0.17	0.04	0.25	0.14	0.24	0.05	0.07	3.17E-05	SAPANXN4
rs1505727	G	2	1.56E+08	0.37	0.09	0.06	0.46	0.21	0.11	0.15	3.66E-05	KCNJ3
rs16874849	G	5	77237118	0.23	0.05	0.83	0.21	0.27	0.07	0.10	3.75E-05	AP3B1, OTP, TBCA

Table 6.9. Markers displaying suggestive association p-values in genotyped data meta-analysis of curvature venular tortuosity with age fitted as a covariate. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported is given. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

Results of the genome wide analysis of both genotyped and imputed data did not reveal any significant results, and similar to the analysis of venular tortuosity marker with the lowest p-value in the genotyped analysis only reached $1.19x10^{-5}$.

6.3.3. Fractal dimension measurements

Genome-wide association analyses were carried out for four measurements of retinal vascular fractal dimensions (two measurements obtained by SIVA software and two obtained by VAMPIRE). No genome wide significant results were found in meta-analysis of genotyped and imputed data. Results are given in Tables 6.12.-6.13. Two measurements of monofractal dimensions of the complete retinal microvasculature (obtained by two different softwares) did not show an overlap between top results,

which could be explained by a small number of individuals tested in these analysis, resulting in high variation of the two phenotypes. Also, it has to be noted that only a subset of subjects from Korcula had fractal dimensions measured with VAMPIRE software (N=296; 390 individuals had SIVA measurements). When additional 396 LBC individuals were included in the analysis, a lack of overlap between the top results for the two traits was still seen (Appendix VI). Inflation factors λ was low for all population included in the analysis (monofractal dimension measurements; D_{box} $\lambda_{meta\text{-}analysis}$=1.003 ; FractalDimC $\lambda_{meta\text{-}analysis}$=0.99; multifractal dimension measurement $Do_{\lambda meta\text{-}analysis}$=1.002; monofractal dimensions of venules $\lambda_{meta\text{-}analysis}$=0.99). Measurements obtained by SIVA for monofractal dimension of the whole retinal microvasculature and for the venular microvasculature did show a substantial overlap between the top results as expected (Tables 6.12-6.13), as did the results for mono and multifractal dimensions obtained by VAMPIRE.

name	Eff allele	CHR	position	beta	se beta	eff allele freq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	gene
rs2193849	G	3	1.77E+08	-0.23	0.05	0.21	-0.18	-0.32	0.06	0.08	1.74E-06	NAALADL2
rs2681384	G	19	21517965	-0.23	0.05	0.82	-0.19	-0.34	0.06	0.09	3.76E-06	ZNF493, ZNF429, ZNF708
rs6875935	G	5	1.73E+08	0.19	0.04	0.34	0.21	0.14	0.05	0.07	7.79E-06	CPEB4
rs34118	G	5	80394308	-0.20	0.05	0.78	-0.17	-0.27	0.06	0.08	1.66E-05	RASGRF2
rs17124619	G	20	31397839	-0.19	0.05	0.76	-0.14	-0.32	0.05	0.08	2.08E-05	ZNF341, CBFA2T2, PXMP4, E2F1, NECAB3, CDK5RAP1, SNTA1
rs9294390	G	6	88556380	-0.17	0.04	0.52	-0.19	-0.12	0.05	0.07	2.12E-05	C6orf166, ORC3L, RARS2
rs4877107	G	9	90573421	0.18	0.04	0.40	0.17	0.20	0.05	0.08	2.16E-05	C9orf47, SHC3, SPIN1, EDG3, NXNL2
rs520865	G	5	80373213	0.17	0.04	0.54	0.14	0.22	0.05	0.07	2.37E-05	RASGRF2
rs527506	G	5	80344906	0.17	0.04	0.48	0.14	0.24	0.05	0.07	2.71E-05	RASGRF2
rs9315991	C	13	42629541	-0.23	0.05	0.84	-0.25	-0.17	0.06	0.10	2.74E-05	ENOX1, DNAJC15, EPSTI1
rs9386455	G	6	1.06E+08	0.18	0.04	0.70	0.18	0.18	0.05	0.07	3.21E-05	PREP
rs1423726	G	16	72025417	-0.20	0.05	0.78	-0.21	-0.17	0.06	0.08	3.51E-05	
rs7837028	G	8	1.16E+08	-0.17	0.04	0.34	-0.19	-0.13	0.05	0.07	3.85E-05	
rs2677872	G	18	2498671	0.21	0.05	0.82	0.22	0.17	0.06	0.09	3.96E-05	METTL4, NDC80
rs9388176	G	6	1.23E+08	0.17	0.04	0.69	0.12	0.28	0.05	0.07	4.02E-05	RLBP1L2, RLBP1L2
rs7153045	G	14	55911652	-0.17	0.04	0.39	-0.12	-0.25	0.05	0.07	4.46E-05	PELI2, C14orf101
rs2604296	G	11	1.14E+08	-0.16	0.04	0.59	-0.11	-0.28	0.05	0.07	4.55E-05	RBM7, REXO2
rs7982593	G	13	42645747	-0.22	0.05	0.84	-0.25	-0.14	0.06	0.10	4.74E-05	ENOX1, DNAJC15, EPSTI1
rs1012040	G	19	21464254	0.20	0.05	0.20	0.17	0.28	0.06	0.09	4.83E-05	ZNF493, ZNF429, ZNF708
rs9878975	C	3	43143226	0.22	0.05	0.17	0.19	0.27	0.07	0.09	4.90E-05	C3orf39

A

name	eff allele	CHR	position	beta	se beta	eff allele freq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	gene
rs4970420	G	1	1096336	-0.26	0.06	0.84	-0.19	-0.36	0.07	0.09	5.20E-06	TTLL10, C1orf159
rs17124619	G	20	31397839	-0.22	0.05	0.76	-0.20	-0.25	0.06	0.09	5.20E-06	ZNF341, CBFA2T2, PXMP4, E2F1, NECAB3, CDK5RAP1, SNTA1
rs7007158	G	8	1.44E+08	0.19	0.04	0.41	0.14	0.31	0.05	0.08	1.02E-05	C8orf31
rs4768689	C	12	44845032	-0.21	0.05	0.24	-0.18	-0.25	0.06	0.08	1.19E-05	SFRS2IP, SLC38A2, SLC38A1
rs2177043	G	15	94918992	-0.19	0.04	0.69	-0.16	-0.26	0.05	0.08	1.40E-05	NR2F2
rs6057924	G	20	31938232	0.17	0.04	0.52	0.16	0.19	0.05	0.08	2.77E-05	CHMP4B, ZNF341
rs9315991	C	13	42629541	-0.23	0.06	0.84	-0.25	-0.20	0.07	0.10	3.02E-05	ENOX1, DNAJC15, EPSTI1
rs2193849	G	3	1.77E+08	-0.21	0.05	0.21	-0.13	-0.37	0.06	0.09	3.47E-05	NAALADL2
rs9386455	G	6	1.06E+08	0.19	0.04	0.70	0.17	0.21	0.06	0.08	3.85E-05	PREP
rs7982593	G	13	42645747	-0.23	0.06	0.84	-0.25	-0.17	0.07	0.10	4.05E-05	ENOX1, DNAJC15, EPSTI1
rs13282211	G	8	8797548	-0.20	0.05	0.74	-0.17	-0.25	0.06	0.08	4.10E-05	MFHAS1, CLDN23
rs3762988	G	5	32745410	0.18	0.04	0.67	0.21	0.13	0.05	0.08	4.58E-05	NPR3
rs9294390	G	6	88556380	-0.17	0.04	0.52	-0.17	-0.16	0.05	0.07	4.65E-05	C6orf166, ORC3L, RARS2

B

Table 6.10. Markers displaying suggestive association p-values in the genotyped data analysis with diastolic blood pressure and age as a covariate. Table A represent results of association analysis for fractal dimensions of the compete retinal microvasculature analysed by SIVA, while table B represents results for venular vasculatue fractal dimension. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported is given. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

.name	eff allele	CHR	position	beta	se beta	eff allele freq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	gene
rs2279475	**C**	**2**	**28398613**	**-0.24**	**0.05**	**0.81**	**-0.22**	**-0.25**	**0.08**	**0.06**	**1.02E-06**	BRE
rs833605	G	4	1.67E+08	-0.18	0.04	0.37	-0.06	-0.22	0.08	0.05	8.02E-06	TLL1
rs7830335	G	8	75024038	-0.18	0.04	0.63	-0.02	-0.23	0.08	0.05	1.29E-05	LY96, TMEM70, TCEB1
rs1559797	G	8	75033201	0.18	0.04	0.36	0.02	0.23	0.08	0.05	1.55E-05	LY96, TMEM70, TCEB1
rs9428780	G	1	2.37E+08	-0.27	0.06	0.90	-0.20	-0.31	0.12	0.08	2.02E-05	
rs1440371	G	15	64728138	0.18	0.04	0.71	0.16	0.19	0.08	0.05	2.26E-05	SMAD6, ZWILCH, RPL4, MAP2K1, LCTL, SNAPC5
rs1563517	C	5	60099268	-0.24	0.06	0.87	-0.24	-0.24	0.11	0.07	2.50E-05	DEPDC1B, NDUFAF2, ERCC8, ELOVL7
rs9428757	G	1	2.37E+08	0.27	0.06	0.10	0.20	0.30	0.12	0.08	2.57E-05	
rs4738402	G	8	75042873	0.17	0.04	0.37	0.02	0.23	0.08	0.05	2.58E-05	LY96 ,TMEM70, TCEB1
rs7827230	G	8	62117524	-0.31	0.07	0.92	-0.45	-0.25	0.14	0.09	2.97E-05	CHD7, RLBP1L1
rs2176347	C	2	45821737	0.17	0.04	0.40	0.21	0.15	0.08	0.05	3.23E-05	PRKCE
rs2274037	**G**	**10**	**6102114**	**-0.40**	**0.10**	**0.96**	**-0.35**	**-0.42**	**0.16**	**0.12**	**3.34E-05**	IL2RA
rs2250952	C	3	67710973	0.20	0.05	0.81	0.11	0.25	0.09	0.06	3.50E-05	SUCLG2
rs1473092	G	4	46535701	-0.17	0.04	0.73	-0.07	-0.22	0.08	0.05	4.43E-05	COX7B2

.

name	eff allele	CHR	position	beta	se beta	eff allele freq	beta Orkney	beta Korcula	se Orkney	se Korcula	p	gene
rs6703753	G	1	2.17E+08	0.21	0.05	0.78	0.26	0.20	0.09	0.06	1.90E-05	LYPLAL1
rs1806507	G	13	96836578	-0.18	0.04	0.65	-0.12	-0.21	0.08	0.05	2.25E-05	RAP2A, MBNL2
rs1893799	G	18	50868723	0.19	0.04	0.66	0.15	0.20	0.09	0.05	2.26E-05	CCDC68 ,TCF4 RAB27B
rs2274037	**G**	**10**	**6102114**	**-0.42**	**0.10**	**0.96**	**-0.55**	**-0.36**	**0.17**	**0.13**	**2.84E-05**	IL2RA
rs10778641	C	12	1.07E+08	0.18	0.04	0.41	0.18	0.18	0.08	0.05	2.87E-05	SART3, ISCU, CORO1C, WSCD2, TMEM119, CMKLR1, SELPLG, FICD
rs8020821	G	14	63284561	0.35	0.08	0.94	0.44	0.31	0.15	0.10	3.25E-05	WDR89, SGPP1
rs2959232	G	2	45164823	-0.17	0.04	0.52	-0.15	-0.18	0.08	0.05	3.76E-05	SIX2, SIX3
rs2279475	**C**	**2**	**28398613**	**-0.21**	**0.05**	**0.81**	**-0.25**	**-0.19**	**0.09**	**0.06**	**4.66E-05**	BRE

Table 6.11. Markers displaying suggestive association p-values in the genotyped data analysis with diastolic blood pressure and age as a covariate. Table A represent results of association analysis for monofractal dimensions of the complete retinal microvasculature analysed by VAMPIRE, while table B represents results for multifractal dimension analysed by the same software. Markers displaying suggestive association p-values in genotyped data meta-analysis of curvature venular tortuosity with age fitted as a covariate. Meta-analysis effect size estimates along with its standard errors are given for each SNP, as well as population specific values. Allele frequency of the allele for the effect is reported is given. List of genes underlying suggestive ($p<5x10^{-5}$) SNPs identified by GRAIL is given.

6.4. Pathway analysis

Genes putatively underlying the association signals were identified using a protocol implemented in GRAIL, just as for the retinal detachment data (Raychaudhuri et al.). Association signals with p-value $<5x10^{-5}$ from the meta-analysis of the genotyped data of all 10 traits tested were used to identify genes included in the pathway IPA analysis (detailed description in chapter 2), totalling 185 genes. Networks with a maximum of 35 genes were created using IPA software. The analysis yielded 9 significant networks (Table 6.12). The two most significant networks showed functional associations with Cell Signalling, Neurological Disease, Organismal Injury and Abnormalities in the first network and Auditory and Vestibular System Development and Function, Connective Tissue Development and Function, Embryonic Development in the second highly significant network. In addition, network overlapping was observed for networks 4, 5 and 6, and for the networks 2 and 8. Nine significant "in-silico" pathway networks with a substantial degree of overlapping present evidence for the non-randomness of the top association signals of retinal microvasculature traits.

Molecules in Network	Score	Focus Molecules	Top Functions
ADCY,ADRA1B,ADRB, AKIRIN2,alcohol group acceptor phosphotransferase,Calmodulin, CAP1,CMKLR1,Cyclin E, DISC1, ERK1/2,FAM20B, GLP1R,GNRH,Gpcr, GPR133,Lh,MAP2K1,MAP2K1/2 ,NPR3 ,PDGF BB,PLC,PREP,PRKCE,Proinsulin,RAN, Ras,RASGRF2, RIMBP2,S1PR3, SGPP1, SHC3,SMAD6,SNTA1,TIAM2	$1x10^{-39}$	21	Cell Signaling, Neurological Disease, Organismal Injury and Abnormalities
ACTA1, Actin, AP3B1, ATG16L1 ,BRE, CACNA1E,caspase,CD3,CHD7,E2F1,ERK,FSH,GLI3,GR AMD4 ,GTPBP4, Histone h4, IL12 (complex), Insulin, KCNJ3, KLHL12,LY96,MFHAS1, NDC80, NFkB (complex), NR2F2, P38 MAPK, PCLO, Pka, Pkc(s),Ras homolog, SELPLG, SLC38A2, SOAT1, TFB1M, Vegf	$1x10^{-39}$	21	Auditory and Vestibular System Development and Function, Connective Tissue Development and Function, Embryonic Development
ABL2,ADIPOR1,Akt,Alp,Ap1,ATF3, BCR (complex),Ck2, CLDN23,COL4A2, CREB5, Creb,DGKD,ERCC8, estrogen receptor,F Actin,Hdac,Histone h3,IL1 ,IL2RA, Immunoglobulin, Interferon alpha, NKD1, NPHS2,ORC3, PELI2, PI3K (complex), PPT1, RAP2A,RNA polymerase II,SCAF8,SIX3,TCF4,TMEM119,TTN	$1x10^{-37}$	20	Developmental Disorder, Hereditary Disorder, Lipid Metabolism
ANKRD50,APP,C18orf21,C9orf142,CBFA2T2,CHAC2,C HID1,CLDN20,CPEB4,DCAKD,DESI2,DNAJC15,ELAV L1,EPB41L4A,GDPGP1,GPATCH1,GTDC2,HINT3,ISC U,KIAA2013,MBNL2,MOB3A,NAALADL2,PASD1,RN ASE11,RNF212,SART3,SIX2,SLC48A1,TOB1,UBC,USP 40,WDR89,ZNF529,ZSWIM1	$1x10^{-26}$	16	Developmental Disorder, Hereditary Disorder, Neurological Disease
AP1M2,CEP192,CYB5R1,DEPDC1B,DNAJB9,ERN1,FA M92A1,GEMIN5,KDELR1,LPCAT2,LYPLAL1,METTL4 ,MKRN2,MPC1,MPZL2,NDUFAF2,NECAB3,PARP16,P EX19,PMF1,POF1B,PXMP4,RARS2,RBM12B,RNASEL, SIPA1L2,SPTLC1,SPTLC2,SPTLC3,SYT11,TMEM67,T MEM70,UBC,WFS1,ZNF493	$1x10^{-26}$	15	Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport
BROX,CHMP7,CHMP4B,CRYZ,DNAH6,DNAH8,DNA H10,DNAH11,DNAH14,DYNC2H1,DYNLRB2,DYNLT1 ,ENOX1,ENPP5,MRPL30,NBAS,NUP37,NUP133,RBM7, REXO4,RLF,SCAF11,SCFD2,SPDL1,SPIN1,SUCLG1,S UCLG2,TBCA,TLL1,TROAP,UBC,USP54,ZNF429,ZNF 708,ZWILCH	$1x10^{-26}$	15	Lipid Metabolism, Nucleic Acid Metabolism, Small Molecule Biochemistry
ABCB5,ATG5,COX11,COX15,COX4I2,COX6B2,COX7 B2,COX7B,COX8C,CUL1,cytochrome-c oxidase,EPSTI1,FGF4,FGF14,FGF22,FGFR1,FICD,FOS, GKN1,HSF1,HTT,LCTL,Mapk,MITF,PDSS2,POMC,RA B15,RAB27B,RABIF,REXO2,RPL4,SLC44A5,SNAPC5, SURF1,ZMPSTE24	$1x10^{-18}$	12	Cell Signaling, Post-Translational Modification, Protein Synthesis
Ca2+,CASP1,CD177,CDH23,CERK,CORO1C,dihydrotest osterone,ELOVL7,Erdr1,EXTL2,GPX5,GRIN1,HPCA,HR AS,Jnk,KDM6A,LHX8,LRFN2,MIF,MTCH1,N-arachidonylglycine,NOS2,OTP,PCDH15,Pkc(s),platelet activating factor phosphatidate,PSEN1,RGN,Rps17,SRF,STX2,TCHH,TOR 3A,TTC1,vitamin K1	$1x10^{-17}$	11	Gene Expression, Auditory Disease, Hereditary Disorder
ABCB10,ANKS1B,ASB1,ASB2,ASB9,beta-estradiol,BMI1,CDK5RAP1,CDKL5,CLVS2,DLG4,DSCA M,ELL3,FOLH1,GALNT7,GFM1,HOXC5,HUNK,LAGE 3,LRRC4C,mir-322,miR-16-5p (and other miRNAs w/seed AGCAGCA),PCDH10,PWWP2A,SDK1,SLC38A1,SLC52 A1,SLC6A9,SSTR4,SUSD2,TAF1L,TCEB1,TMF1,Ubiqui tin,WSB1	$1x10^{-13}$	9	Cell Death and Survival, Developmental Disorder, Endocrine System Disorders

Table 6.12. Genes creating nine functional networks obtained by IPA analysis are given, along with the significance score for each network and number of genes included in the output of our GWAS meta-analysis. Top functions represent enrichment of known functional networks in the hypothetical pathways generated by IPA.

6.5. Discussion

Cardiovascular diseases are complex influenced by numerous risk factors and their interactions. Identification of susceptibility variants underlying quantitative intermediate phenotypes (e.g blood pressure, arterial stiffness, BMI) has been used to identify novel disease pathways. The majority of the genetic association studies to date have focused on macrovascular disease phenotypes, and only two GWAS of the microvascular retinal traits have been published so far (Sim et al. 2013, Ikram et al 2009). Here we analysed rarely studied endophenotypes, derived from retinal vessel microvasculature which have been associated with various systemic risk factors and diseases (Chapter 5). None of the analyses led to genome-wide significant replicated findings but the sample sizes were modest. Still, most traits looked promising for future larger-scale GWAS analyses given the shape of the QQ plots obtained (excess of low p-values while genomic inflation was kept low) and the biological relevance of some of the suggestive hits.

.

CRAE

Narrowing of the retinal arterioles is a characteristic of hypertension, due to the increased peripheral vascular resistance (Levy et al. 2001). However it is still unresolved whether it is a consequence of it or it precedes and contributes to the development of hypertension. A few studies have presented evidence that retinal arterial narrowing may be linked to the development of hypertension (Wang et al. 2008, Wong et al 2004), and that retinal arteriolar diameter therefore presents an endophenotype for hypertension. However changes to microvasculature similar to the ones observed in hypertension are also seen in hypertrophic cardiomyopathy. In this study the region on chromosome 2 found to be significantly associated with CRAE is located within the *TTN* gene, which encodes a large and abundant protein of striated muscle. Titin is involved in the passive

and restoring force of the cardiac sarcomere (the basic contractile unit of the myocyte or cardiac muscle cell) and makes a major contribution to the diastolic wall stress of the left ventricle (LeWinter and Granzier 2010). Mutations in this gene have been implicated in numerous myopathies involving cardiac and skeletal muscle (Vogan et al. 2012). A recessive mouse mutation (shrunken head mutation) that disrupts function of the *Ttn* gene has been associated with changes in blood vessel morphology making it a good biological candidate for a CRAE. This mutation results in weak heart contractility (May et al. 2004) causing reduced circulation in these mouse mutants which affects blood vessel morphogenesis (May et al. 2004). The previously identified locus associated with CRAE was imputed in all of our cohorts and could be directly tested in this study. No significant association were found in any of the three cohorts tested, but the direction of the effect was consistent in Orkney and LBC, but not in Korcula. Interestingly, this marker (rs2194025) identified in the large meta-analysis of five cohorts of European ancestry and replicated in two additional independent cohorts of European ancestry, is located between *TMEM161B* and *MEF2C* genes. The closest gene, *MEF2C;* myocyte enhancer factor, is a transcription factor in MEF2 family and is involved in cardiac myogenesis (Lin et al. 1997, Han et al. 2000). Overexpression of *MEF2C* induces dilated cardiomyopathy and lengthening of myocytes in mice (Xu et al. 2006). The same locus was found to be associated with venular calibre in a large meta-analysis involving ~22000 subjects. It did not reached genome-wide significance in our study, but it showed the same direction and smaller size of the effect as in the published data . Interestingly, similar phenotypes were reported for muscles from *mef2c/d* double knockdown zebrafish and mice lacking titin's M-line region (Hinits and Hughes, 2007). Therefore, even though our genome wide significant association result at the titin's locus was not strengthened by the analysis of imputed data or look-up in the small independent Scottish cohort, functional evidence make this locus a strong candidate and requires further testing involving larger numbers of tested subjects.

Curvature tortuosity

GWAS of curvature arterial tortuosity revealed another functionally interesting candidate gene. The most significantly associated marker (rs2296845) in the analysis of the Orkney genotyped data falls with the region of the *COL4A2* gene which encodes one of the six subunits of type IV collagen. This marker did not reveal significant association with arteriolar curvature tortuosity in Korcula nor LBC cohorts, however the direction and size of the effect was consistent across all three populations. Type IV collagens are the major structural component and most abundant protein of basement membranes (BM). Basement membranes provide support to the vessel wall and in retinal vessels and act as matrix controlling the passage and sequestration of vasoactive substances. In addition, type IV collagens play a role in angiogenesis. The role of COL4A2 in vessel structure makes this association a very interesting functional candidate. Mutations in *COL4A2* have been found in sporadic and familial porencephally and small vessel disease in humans (Yoneda et al. 2012, Verbeek et al. 2012). Porencephally is a neurological disorder characterized by cysts in the brain, most likely caused by impaired blood supply. In addition, *Col4a2* mutations in mice are associated with defects in the eye, brain and with vessel stability (Favor et al. 2007). *COL4A2* shares a promoter and is co-expressed in the basement membrane of epithelia and endothelia with the neighbouring paralogue *COL4A1*. Mendelian variants in COL4A1 underlie syndromes which include tortuous retinal vessels (Plaisier et al. 2011). Further evidence that this retinal tortuosity association needs further research comes from the fact that SNPs in this region (COL4A1/COL4A2 locus) show genome-wide significant associations with vascular abnormalities, including coronary artery calcification (O'Donnell et al. 2011), coronary heart disease (Cradiogram; Schunkert et al. 2011) and arterial stiffness (Tarasov et al. 2009). Col4a2 mouse mutants report small vessel disease, intracerebral haemorrhage and retinal changes. The *SLC2A10* gene found to be associated with general arterial tortuosity (Coucke et al. 2006, Cheng et al. 2008) did not show significant associations in analysis of this or other retinal tortuosity measurements. The marker presenting lowest p-value in the meta-analysis of the imputed data is located

between RIMBP2 (encodes RIM-binding protein) and *STX2* gene (encodes a protein that regulates epithelial-mesenchymal interactions and epithelial cell morphogenesis and activation). This SNP is located within the LD block that spans STX2 gene (Figure 6.10). It is expressed in the connective tissue around various epithelial tissues, muscles and vessels; with strong staining present on the adventitia (connective tissue that surrounds the vessel) of blood vessels (Zhang et al. 1998).

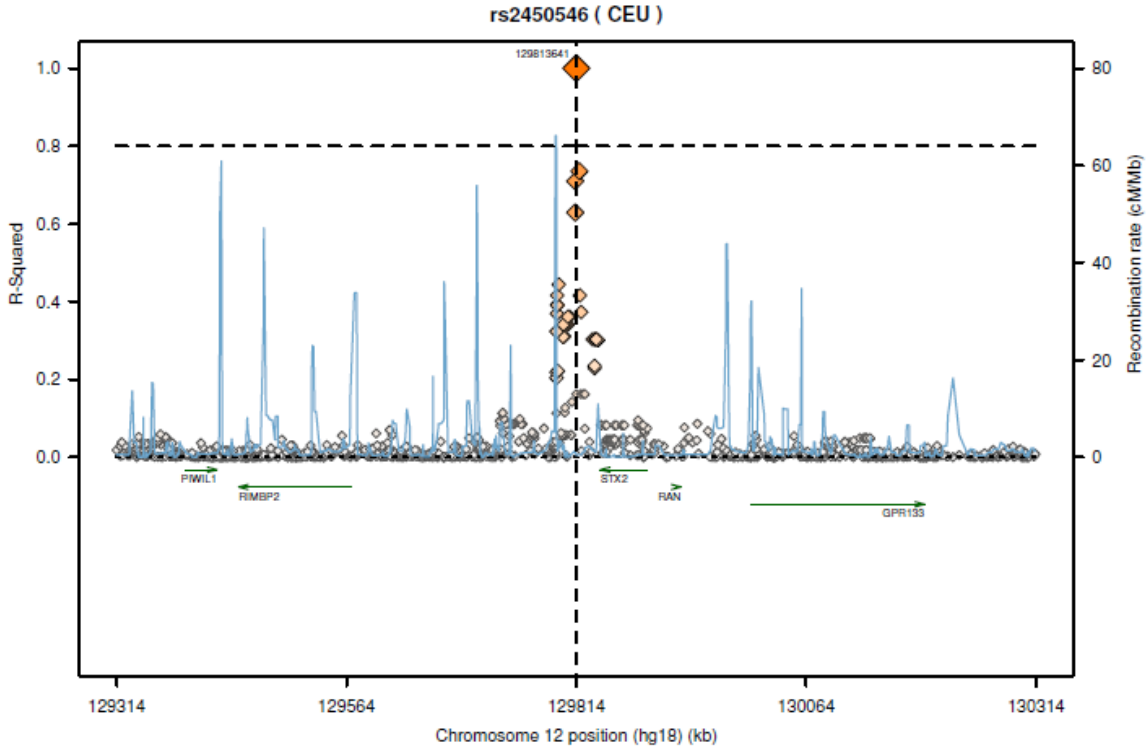


Figure 6.11. Regional LD plot around rs2450546 is presented. A region of 500 kb around the SNP is displayed. Recombination rate is displayed by a blue line with scale on the right-hand axis and the R-squared scale on the left-hand side. Genes are displayed in the bottom part of the figure, with the arrow showing the direction of transcription.

Look-up of published vessel caliber associations

Look-up of the published associations between several loci and retinal vessel calibers was performed. Even though only one SNP identified in the large meta-analysis to be associated with CRVE replicated in Orkney and another one in Korcula, all tested markers presented same direction of the effect. Most likely due to the considerably smaller number of subjects included in our cohorts compared to the published study and winners curse effect, smaller p-values were reported. Each population tested in the meta-analysis included between 2579 and 15,792 subjects and despite larger sample sizes, one of the markers did not present nominally significant association with CRVE in all of the cohorts included in the discovery stage meta-analysis (AGES study, rs2287921 p=0.075; rs17421627 p=0.058). Three out of four replication populations did not provide nominally significant p value for at least one of the markers identified in the discovery stage. A similar observation is true for the meta-analysis of arterial width (Sim et al. 2013), results of which did not replicate in our populations. The smallest population of the replication stage (Australian Twin Study; N=1769) did not present nominally significant association for the identified SNP. Nevertheless our analysis yielded one genome wide significant result not identified in the larger published meta-analysis, but markers identified in both analyses lie in the regions of genes with very similar biological functions (*TTN* and *MEF2C).*

Two large GWAS published on retinal microvasculature involved around 22000 individuals, compared to ~1600 individuals used in this study. This study was underpowered to detect genes with small effects and will require inclusion of further population in order to confirm suggestive results identified, and to identify further variants associated with additional retinal vessel traits.

Chapter 7. Conclusion

This chapter briefly recapitulates the main achievements and the remaining challenges in the two main areas studied in this thesis; genetics of the retinal detachment and quantitative genetics of retinal microvasculature traits.

Genetics of the retinal detachment

The genetic basis of rhegmatogenous retinal detachment has been suggested by familial aggregation studies and a number of inherited disorders which present RRD as one of the symptoms. However no study prior to this explored the underlying genetics of the non-syndromic RRD. The genetic association study presented here was a continuation of the larger epidemiological study carried out in Scotland (Mitry et al. 2010). The genome-wide genetic association analysis was completed prior to this project, and my work started after a multi-center collaboration was formed in order to expand the original genetic association analysis of the Scottish samples. Based on the most promising results of the original GWAS, the most significant SNPs were selected for testing in two additional populations. Candidate gene markers were also included, in order to test the associations between genes with functional importance in genetic syndromes associated with RRD or the ones with the known role in vitreo-retinal adhesion. Finally, the most promising associations were tested in the replication set consisting of cases from two additionally recruited populations. This multi-stage design was applied due to the financial restrictions. One genome wide significant and several suggestive common variants of moderate effect contributing to RRD risk were identified. Interestingly, several of the suggestive signals encompassed genes with a documented role in cell adhesion or migration; in particular these results suggested a major role of actin cytoskeletal organisation in the pathogenesis of RRD. None of the candidate gene SNPs were present among the suggestive results in the discovery stage meta-analysis. Among the selected candidate genes, the majority were molecules that

play a significant role in vitreo-retinal adhesion. Adhesion of the internal limiting membrane and vitreous is mediated thought the extra-cellular matrix (Sebag 1992). It is mainly sustained by collagens (types VI, VII and XVIII) that play a role in binding matrix components and stabilizing ILM (type IV) (Ponsioen et al. 2008), along with laminin, fibronectin and other glycoproteins. However, viteo-retinal adhesion plays a more important role in the development of retinal tears during posterior vitreous detachment, which is one of the risk factors for RRD development. Nevertheless, only a small proportion of the patients presenting with posterior vitreous detachment develop RRD, and the adhesion processes suggested by the genes underlying variants identified in GWAS are more likely to be involved in pathways involved in the adhesion between the neurosensory retina (NSR) and the retinal pigment epithelium (RPE). Since detachment of these two layers is responsible for the development of RRD from a predisposing tear. Very little is known about the pathogenesis behind RRD and results presented here suggest that factors involved in the cell adhesion are potentially crucial in the progression from retinal tear to RRD. This study also highlighted the difficulty of imputation of case-control data genotyped on different genotyping platforms. The abundance of positive association in the analysis of the imputed data emphasised the need for careful planning of the study design and genotyping procedure before the imputation is carried out. The consequence of selecting just a subset of variants tested in the GWAS step, to be tested in the follow up steps resulted in the loss of potential associations that could have been detected given the larger sample size of the meta-analysis. Given the constraints involved in the design of this genetic association study, which resulted from the financial limitations, there is a good prospect for further research. Whole genome genotyping of cases collected in the second discovery step and the replication step would be the obvious extension of the work presented here. This would enable the detection of variants missed due to the SNP selection procedure. Analysis of imputed data using the guidelines developed in this thesis would be also worth applying.

The main limitation of this study was the number of cases and controls included in the analysis. For a SNP with MAF of 10%, under a multiplicative model with heterozygote relative risk of 1.3, at least 1100 cases and controls are required to obtain 80% power at a significance level of 5% with no correction for multiple testing, and at least 6000 if a standard genome wide significance cut-off is applied $5x10^{-8}$ (Purcell et al. 2003). Based on this, the RRD study was underpowered to detect many of the variants of small effect, which are mainly revealed in GWA studies of complex traits (i.e., a median OR of 1.33 (Hiendorf et al 2009) - which would be much lower if calculated today given the larger datasets and consequently smaller effects that may be found). Our results were mostly consistent with common variants individually contributing small effect, with a few of the suggestive markers showing slightly higher OR; probably due to the "winners curse" effect. Studies identifying risk alleles of larger effect can be performed with substantially lower number of individuals, e.g. 1660 European ancestry cases and 7,306 controls in the GWAS of multiple myeloma reporting genome wide significant variants with OR of 1.8-1.95 (Weinhold et al. 2013). One important result was the quantification of the overall contribution of the common variants tested using the newly developed method GCTA (Yang et al. 2010). It suggests a polygenic basis of RRD, with 27.4% of the disease risk attributable to the additive effects of the genotyped SNPs. Rare variant are not tested and detected by regular GWAS platforms and analysis and may contribute to the genetic risk.

In addition to collecting larger number individuals for genome wide association studies, other sources of genetic variation should also be explored in future studies, in order to elucidate the genetic factors influencing RRD and retinal vessel morphology. Detection of the rare variants (MAF<0.05) would be of interest, since common variants are not likely to explain all of the genetic variation underlying complex diseases and traits. An alternative to the common disease-common variant hypothesis (Lander 1996, Reich and Lander 2001) is the common disease/rare variant - CD/RV hypothesis or the rare allele model (Wright et al. 2003). The only technology for detection of such rare SNPs is

sequencing (of either regions of interest or whole genome). Sequencing has been done on regions of repeated associations with common SNPs (identified by GWAS), or in large regions of the genome in people with extreme phenotypes (Manolio et al. 2009). Whole genome sequencing in general has become more affordable and sequencing individuals on the extreme of a quantitative trait (more likely to carry loss-of-function alleles (Romeo et al., 2007)) would be a potential way forward in the analysis of retinal traits.

Empirical evidence supporting both hypotheses (common disease-common variants, common disease-rare variants) exists and therefore both directions should be followed in the future research. The infinitesimal model proposing that genetic component of complex traits can be attributed to a large number of small effects common variants is supported by the results of recently published GWAS including large number of individuals (N>100 000) that have uncovered numerous previously unidentified association signals (Lango–Allen et al. 2010, Kottgen et al. 2013, Morris et al. 2011, Speliotes et al. 2010). Evidence from the expression quantitative trait loci analysis showing that gene expression and splicing are heavily influenced by common variants support the CD-CV hypothesis (Cookson et al 2009, Lalonde et al. 2011). Evidence favouring the rare allele model include population genetics data which gives evidence that deleterious variants are rare (exome-sequence data analysis demonstrated that non-synonymous substitutions are found to have very low frequencies (Kryukov et al. 2007), and data showing that many rare familial genetic disorders are caused by rare alleles of big effect (e.g. breast cancer susceptibility mutation BRCA1 and BRCA2 (Easton et al. 2007).

Finally structural variants (poorly captured by existing arrays used for GWAS), epistatic and epigenetic effects may also be important in explaining the genetic architecture of complex traits (Manolio et al. 2009). Growing number of studies are giving evidence of

epigenetic effects (Feinberg et al. 2007), parent of origin effect (Kong et al. 2009, Small et al. 2011) and DNA methylation inheritance (Jablonka et al. 2009), suggesting these effects play a role in the genetic component of complex disease and are a part of the missing heritability not explained by the current GWAS. Contributions of rare variants and CNVs should also be explore in future studies.

Retinal vessel traits

The same issue of low sample size was observed in the genetic association analysis of the retinal vessel traits. Our sample size consisted of 1196 individuals used in the meta-analysis and further 397 used as a replication cohort, which is very low when compared to the majority of the published GWAMA studies of quantitative traits (more recent studies included close to or even more than 100,000 individuals (Kotgen et al. 2013, Morris et al. 2011, Speliotes et al. 2010, Teslovich et al. 2010)). Both published reports on genome wide associations of retinal vessel calibre used approximately 20,000 individuals and reported one variant for arterial calibre and four for venular widths that were confirmed in the replication cohorts (Sim e al. 2013, Ikram et al. 2010). Even though the sample sizes available in the retinal vessel traits GWAS were not comparable to large scale meta-analysis using numerous cohorts, two biologically very interesting results were reported and warrant further testing in additional cohorts.

In addition to the genetic association studies, the work presented in this thesis revealed the first reported heritability estimates of vascular tortuosity measurements and fractal dimensions, providing evidence of the substantial additive genetic contribution to these traits. This is further supported by the fact that the reported estimates were consistent between the two isolates for the majority of the traits tested. Subsequent association studies did not reveal any significant genetic associations, but some of the suggestive results present a good basis for further work involving more subjects. Reports of

moderate to high genetic correlations of most pairs of tested traits suggest common genetic variants influencing general retinal vascular morphology.

Associations of several known systemic factors with retinal vessel morphology were confirmed, and presented a form of validation for this study. In addition, several novel associations were detected, and some of the systemic risk factors were tested for the first time (pulse wave augmentation index). Results presented confirmed the well-known associations, but also highlighted a need for more careful selection of covariates in future epidemiological and genetic studies involving retinal microvasculature, since some of the systemic factor tested showed very low and insignificant correlations with majority of the tested retinal microcirculation traits (e.g. glucose, HbA1c, fibrinogen). Future work involving associations between systemic risk factors and retinal microvasculature could potentially include a more dynamic measures of vasculare function (e.g flowmetry and dynamic vessel assessment), as they could convey more information on the prediction of cardiovascular disease. These measures are more difficult to obtain and therefore majority of studies have concentrated on more the static structural features of retinal microvasculature.

The next important step in the retinal vessel analysis would be the additional phenotyping of fundus images available in both Orkney and Korcula in order to achieve larger sample sizes needed to increase the power of the genetic analysis. The work presented here should form a basis for such future investigation.

Since SNPs identified in GWAS are unlikely to be the causal variants influencing our phenotype of interest (due to the selection of tagging SNPs) but are in most cases in high LD with the variant influencing the trait, further work is needed in order to elucidate the real causal variant and to characterize its functional effects. First step in this process is fine mapping of associated locus, which involves target re-sequencing of a region of association based on a known LD structure. The size of the region to be sequenced is usually defined by the LD structure, however the strength of the correlation between the associated SNP and causal variant is not known and therefore even the use of low r^2 values may be needed. One approach to shorten the region of re-sequencing is to use non-European populations due to the different LD pattern. Populations with overall shorter LD (populations of African descent (HapMap consortium, 2005)) blocks can potentially reduce the target region. Once the region of association is narrowed down, further steps in identifying causal variants can be performed. Firstly the computational methods should be explored, since they are quick to perform and can simplify further experimental work. Existing databases are explored to determine known information about our variants of interest (association with other disease or phenotypes) and to determine possible associations with levels of gene expression though expression quantitative trait loci (eQTLs; genetic variants affecting transcript levels). Potential genes underlying our variant of interest can also be selected by using text mining techniques of research papers (GRAIL; Raychaundhuri et al. 2009). Further approaches, especially in the large non-coding intergenic regions include characterization of the regulatory elements of the regions of associations. Regulatory elements include enhancers, promoters, insulators and silencers. While promoters are usually found at the transcription start sites of the genes, other regulatory elements could be found far from the genes they are influencing. Analysis of histone modification or DNase sensitivity sites has proven to be a good method of annotating regulatory elements (Visel et al. 2009). However, these distal regulatory sequences are usually tissue specific and it should be taken into account in these types of analysis. Once there is enough evidence for one or several genes as potential candidates underlying our region of association, gene expression analysis could be carried out. Empirical or experimental analyses of

gene expression in tissues of interest are usually performed. Regulatory regions analysis is usually done using ChIP-sequencing methodologies (used for identifying binding sites of DNA-associated proteins). Resequencing of the identified regulatory regions is used to capture all the variation within them. Following these steps, multiple experimental manipulations in model systems to functionally validate our implicated genes can be carried out. Homologous recombination uses modification of cell's or organism's genotype to explore allele specific effects on phenotypes using recombination vectors, gene replacements and knockouts in animal models. Silencing of gene expression (RNA interference) is a method used to monitor the effect of different levels of gene expression. In addition, the use of morpholino (artificially synthesised molecule) is another method for modification of gene expression. Gene targeting by zinc finger nuclease can be also used for gene editing, and artificial ectopic gene expression for the modulation of expression levels using constructs with strong promoters. In summary, a good candidate for causal variant can be confirmed with significant associations with transcription levels (overlap between eQTL and GWAS signal), tissue specificity and effect on phenotypes in question on model systems (Freedman et al. 2011).

Genetic analysis of rhegmatogenous retinal detachment revealed one good candidate SNP (rs267738) already associated with the expression levels of *CERS2* gene (Montgomery et al. 2010). Literature search also confirmed expression of this gene in mouse retina (Freeman et al. 2011), and expression analysis in human retinal tissue would be the next step. Particularly expression in different layers of retinal tissue would be of interest, since the detachment is occurring in the basal part of the retinal pigment epithelium. Similar analysis will be considered for all variants showing associations close to genome wide significance. Further step of silencing genes found to be expressed in human retinal tissue using morpholino in a model organism (zebrafish) could be used to see whether it has an effect on the development of the retinal tissue. This would be of particular interest for rs12960119 considering the evolutionary highly conserved region of the *SS18* gene it is in high LD with.

Similar functional follow up will be considered for the genes identified in the analysis of retinal vessel traits. Following the phenotyping of larger numbers of subjects, and subsequent genetic association analysis, functional analysis will follow. *COL4A4* gene was already identified as a potential functional candidate gene associated with retinal arterial tortuosity considering the known role of this gene in the angiogenesis and confirmed association with a disease phenotype in humans which affects the vessel stability. It would be of interest to see the effects on retinal vessels in particular. Knockout mouse models using homologous recombination could help in identifying and confirming the role of *COL4A4*, and the role of specific variants within the gene, in the development and function of retinal arteries. Gene expression analysis of the *TTN* gene, and other markers reaching, or close to, genome wide significance will be considered following a genetic association analysis including larger number of individuals.

References

Akaike, H. (1974). A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19, 716-723.

Algvere, P.V., Jahnberg, P., and Textorius, O. (1999). The Swedish Retinal Detachment Register. I. A database for epidemiological and clinical studies. Graefes Arch Clin Exp Ophthalmol 237, 137-144.

Almasy, L., and Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62, 1198-1211.

Amin, N., van Duijn, C.M., Aulchenko, Y.S. (2007). A genomic background based method for association analysis in related individuals. PLoS One. 2007 Dec 5;2(12):e1274.

Anderson, C.A., Pettersson, F.H., Clarke, G.M., Cardon, L.R., Morris, A.P., and Zondervan, K.T. (2010). Data quality control in genetic case-control association studies. Nat Protocols 5, 1564-1573.

Anderson, C.A., Soranzo, N., Zeggini, E., and Barrett, J.C. (2011). Synthetic associations are unlikely to account for many common disease genome-wide association signals. PLoS Biol 9, 1000580.

Aulchenko, Y.S., Ripke, S., Isaacs, A., and van Duijn, C.M. (2007). GenABEL: an R library for genome-wide association analysis. Bioinformatics 23, 1294-1296.

Aulchenko, Y., Struchalin, M., van Duijn, C. (2010). ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 11: 134.

Azemin, M.Z., Kumar, D.K., Wong, T.Y., Wang, J.J., Mitchell, P., Kawasaki, R., and Wu, H. (2012). Age-related rarefaction in the fractal dimension of retinal vessel. Neurobiol Aging 33, 15.

Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263-265.

Benavente-Perez, A., Hosking, S.L., Logan, N.S., Broadway, D.C. (2010.) Ocular blood flow measurements in healthy human myopic eyes. Graefes Arch Clin Exp Ophthalmol. 248: 1587– 1594

Bengtsson, B., and Krakau, C.E. (1992). Correction of optic disc measurements on fundus photographs. Graefes Arch Clin Exp Ophthalmol 230, 24-28.

Bonaldo, M.F., Lennon, G., and Soares, M.B. (1996). Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res 6, 791-806.

Browning, S.R., and Browning, B.L. (2011). Population structure can inflate SNP-based heritability estimates. Am J Hum Genet 89, 191-193; author reply 193-195.

Budenz, D.L., Fredette, M.J., Feuer, W.J., and Anderson, D.R. (2008). Reproducibility of peripapillary retinal nerve fiber thickness measurements with stratus OCT in glaucomatous eyes. Ophthalmology 115, 661-666

Bush, W.S., and Moore, J.H. (2012). Chapter 11: Genome-wide association studies. PLoS Comput Biol 8, 27.

Cardon, L.R., and Palmer, L.J. (2003). Population stratification and spurious allelic association. Lancet 361, 598-604.

Chen, W.M., and Abecasis, G.R. (2007). Family-based association tests for genomewide association scans. Am J Hum Genet 81, 913-926.

Cheung, C.Y., Thomas, G.N., Tay, W., Ikram, M.K., Hsu, W., Lee, M.L., Lau, Q.P., and Wong, T.Y. (2012). Retinal vascular fractal dimension and its relationship with cardiovascular and ocular risk factors. Am J Ophthalmol 154, 663-674.

Cheung, C.Y., Zheng, Y., Hsu, W., Lee, M.L., Lau, Q.P., Mitchell, P., Wang, J.J., Klein, R., and Wong, T.Y. (2011). Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. Ophthalmology 118, 812-818.

Cheung, C.Y., Hsu, W., Lee, M.L., Wang, J.J., Mitchell, P., Lau, Q.P., Hamzah, .H, Ho, M., Wong, T.Y. (2010). A new method to measure peripheral retinal vascular caliber over an extended area. Microcirculation. 17,495-503.

Cheung, N., Islam, F.M., Saw, S.M., Shankar, A., de Haseth, K., Mitchell, P., and Wong, T.Y. (2007). Distribution and associations of retinal vascular caliber with ethnicity, gender, and birth parameters in young children. Invest Ophthalmol Vis Sci 48, 1018-1024.

Cheung, N., Tikellis, G., Saw, S.M., Amirul Islam, F.M., Mitchell, P., Wang, J.J., and Wong, T.Y. (2007). Relationship of axial length and retinal vascular caliber in children. Am J Ophthalmol 144, 658-662.

Cheng, C.H., Kikuchi, T., Chen, Y.-H., Sabbagha, N.G.A.-A.-A., Lee, Y.-C., Pan, H.-J., Chang, C., and Chen, Y.-T. (2008). Mutations in the SLC2A10 gene cause arterial abnormalities in mice. Cardiovascular Research.

Choi, B.Y., Chae, Y.M., Chung, I.H., Kang, H.S. (1997) Correlation between the postmortem stature and the dried limb-bone lengths of Korean adult males. Yonsei Med J 38:, 79-85.

Clayton, D., and Leung, H.T. (2007). An R package for analysis of whole-genome association studies. Human heredity 64, 45-51.

Consortium, W.T.C.C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-678.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. Nat Rev Genet 10, 184-194.

Cosatto, V.F., Liew, G., Rochtchina, E., Wainwright, A., Zhang, Y., Hsu, W., Lee, M.L., Lau, Q.P., Hamzah, H.H., Mitchell, P., et al. (2010). Retinal vascular fractal dimension measurement and its influence from imaging variation: results of two segmentation methods. Curr Eye Res 35, 850-856.

Coucke, P.J., Willaert, A., Wessels, M.W., Callewaert, B., Zoppi, N., De Backer, J., Fox, J.E., Mancini, G.M., Kambouris, M., Gardella, R., et al. (2006). Mutations in the facilitative glucose transporter GLUT10 alter angiogenesis and cause arterial tortuosity syndrome. Nat Genet 38, 452-457.

Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.C., Jenkins, S.C., Palmer, S.M., et al. (1994). A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371, 130-136.

de, M.J., and Boyd, T.A. (1961). A statistical investigation of the correlation of retinal arteriolar tortuosity with blood pressure and age. Trans Can Opthalmolog Soc 24, 6-17.

Deary, I.J., Yang, J., Davies, G., Harris, S.E., Tenesa, A., Liewald, D., Luciano, M., Lopez, L.M., Gow, A.J., Corley, J., et al. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. Nature 482, 212-215.

Digre, K.B., Corbett, J.J. (2003) Practical Viewing of the Optic Disc. Burlington: Butterworth Heineman

Doubal, F.N., MacGillivray, T.J., Patton, N., Dhillon, B., Dennis, M.S., and Wardlaw, J.M. (2010). Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke. Neurology 74, 1102-1107

Easton, D.F., Deffenbaugh, A.M., Pruss, D., Frye, C., Wenstrup, R.J., Allen-Brady, K., Tavtigian, S.V., Monteiro, A.N., Iversen, E.S., Couch, F.J., et al. (2007). A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am J Hum Genet 81, 873-883.

Eid, J.E., Kung, A.L., Scully, R., and Livingston, D.M. (2000). p300 interacts with the nuclear proto-oncoprotein SYT as part of the active control of cell adhesion. Cell 102, 839-848.

Fahy, S.J., Sun, C., Zhu, G., Healey, P.R., Spector, T.D., Martin, N.G., Mitchell, P., Wong, T.Y., Mackey, D.A., Hammond, C.J., et al. (2011). The relationship between retinal arteriolar and venular calibers is genetically mediated, and each is associated with risk of cardiovascular disease. Invest Ophthalmol Vis Sci 52, 975-981.

Faulkner-Jones, B.E., Godinho, L.N., and Tan, S.S. (1999). Multiple cadherin mRNA expression and developmental regulation of a novel cadherin in the developing mouse eye. Experimental neurology 156, 316-325.

Favor, J., Gloeckner, C.J., Janik, D., Klempt, M., Neuhauser-Klaus, A., Pretsch, W., Schmahl, W., and Quintanilla-Fend, L. (2007). Type IV procollagen missense mutations associated with defects of the eye, vascular stability, the brain, kidney function and embryonic or postnatal viability in the mouse, Mus musculus: an extension of the Col4a1 allelic series and the identification of the first two Col4a2 mutant alleles. Genetics 175, 725-736.

Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. Nature 447, 433-440.

Finnemann, S.C. (2003). Focal adhesion kinase signaling promotes phagocytosis of integrin-bound photoreceptors. The EMBO journal 22, 4143-4154.

Finnemann, S.C., Bonilha, V.L., Marmorstein, A.D., and Rodriguez-Boulan, E. (1997). Phagocytosis of rod outer segments by retinal pigment epithelial cells requires alpha(v)beta5 integrin for binding but not for internalization. Proc Natl Acad Sci U S A 94, 12932-12937.

Foos, R.Y., and Wheeler, N.C. (1982). Vitreoretinal juncture. Synchysis senilis and posterior vitreous detachment. Ophthalmology 89, 1502-1512.

Franklin, C. Linkage and Association Mapping for Quantitative Phenotypes in Isolated Populations, 2010

Freedman, M.L., Monteiro, A.N.A., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D. et al. (2011). Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet, 43, 513-518.

Freeman, N.E., Templeton, J.P., Orr, W.E., Lu, L., Williams, R.W., and Geisert, E.E. (2011). Genetic networks in the mouse retina: growth associated protein 43 and phosphatase tensin homolog network. Mol Vis 17, 1355-1372.

Fried, L.P., Borhani, N.O., Enright, P., Furberg, C.D., Gardin, J.M., Kronmal, R.A., Kuller, L.H., Manolio, T.A., Mittelmark, M.B., Newman, A., et al. (1991). The Cardiovascular Health Study: design and rationale. Ann Epidemiol 1, 263-276.

Gabriel, L.A., Wang, L.W., Bader, H., Ho, J.C., Majors, A.K., Hollyfield, J.G., Traboulsi, E.I., and Apte, S.S. (2012). ADAMTSL4, a secreted glycoprotein widely distributed in the eye, binds fibrillin-1 microfibrils and accelerates microfibril biogenesis. Invest Ophthalmol Vis Sci 53, 461-469.

Gao, X., Starmer, J., and Martin, E.R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol 32, 361-369.

Garner, A., Ashton, N., Tripathi, R., Kohner, E.M., Bulpitt, C.J., and Dollery, C.T. (1975). Pathogenesis of hypertensive retinopathy. An experimental study in the monkey. British Journal of Ophthalmology 59, 3-44.

German, O.L., Miranda, G.E., Abrahan, C.E., and Rotstein, N.P. (2006). Ceramide is a mediator of apoptosis in retina photoreceptors. Investigative ophthalmology & visual science 47, 1658-1668.

Ghazi, N.G., and Green, W.R. (2002). Pathology and pathogenesis of retinal detachment. Eye (London, England) 16, 411-421.

Gibson, G. (2012). Rare and common variants: twenty arguments. Nat Rev Genet 13, 135-145

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. 2002 ASReml User Guide Release 1.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Go, S.L., Hoyng, C.B., and Klaver, C.C. (2005). Genetic risk of rhegmatogenous retinal detachment: a familial aggregation study. Arch Ophthalmol 123, 1237-1241.

Go, S.L., Maugeri, A., Mulder, J.J., van Driel, M.A., Cremers, F.P., and Hoyng, C.B. (2003). Autosomal dominant rhegmatogenous retinal detachment associated with an Arg453Ter mutation in the COL2A1 gene. Invest Ophthalmol Vis Sci 44, 4035-4043.

Gopinath, B., Baur, L.A., Wang, J.J., Teber, E., Liew, G., Cheung, N., Wong, T.Y., and Mitchell, P. (2010). Smaller birth size is associated with narrower retinal arterioles in early adolescence. Microcirculation 17, 660-668.

Gopinath, B., Wang, J.J., Kifley, A., Tan, A.G., Wong, T.Y., and Mitchell, P. (2013). Influence of blood pressure and body mass index on retinal vascular caliber in preschool-aged children. J Hum Hypertens 9, 523-528

Goto, I., Kimoto, K., Katsuki, S., Mimatsu T., Ikui H. (1975) Pathological studies on the intracerebral and retinal arteries in cerebrovascular and noncerebrovascular diseases. Stroke 6, 263–269

Gowen, M., Lazner, F., Dodds, R., Kapadia, R., Feild, J., Tavaria, M., Bertoncello, I., Drake, F., Zavarselk, S., Tellis, I., et al. (1999). Cathepsin K Knockout Mice Develop Osteopetrosis Due to a Deficit in Matrix Degradation but Not Demineralization. Journal of Bone and Mineral Research 14, 1654-1663.

Grauslund J,Green A, Kawasaki R, Hodgson L, Sjølie AK, Wong TY (2010) Retinal Vascular Fractals and Microvascular and Macrovascular Complications in Type 1 Diabetes, Ophthalmology 117, 1400-1405.

Grishanin, R.N., Yang, H., Liu, X., Donohue-Rolfe, K., Nune, G.C., Zang, K., Xu, B., Duncan, J.L., Lavail, M.M., Copenhagen, D.R., et al. (2008). Retinal TrkB receptors regulate neural development in the inner, but not outer, retina. Mol Cell Neurosci 38, 431-443.

Gross, J. (2013). Nortest: R package version 1.0-2.

Haimann M.H., Burton T. C., and Brown C. K. (1982). Epidemiology of retinal detachment. Archives of Ophthalmology 100, 289-292.

Hamelers, I.H., Olivo, C., Mertens, A.E., Pegtel, D.M., van der Kammen, R.A., Sonnenberg, A., and Collard, J.G. (2005). The Rac activator Tiam1 is required for (alpha)3(beta)1-mediated laminin-5 deposition, cell spreading, and cell migration. The Journal of cell biology 171, 871-881.

Han, J., and Molkentin, J.D. (2000). Regulation of MEF2 by p38 MAPK and its implication in cardiomyocyte biology. Trends Cardiovasc Med 10, 19-22.

Hasan, T, Sharma, S., Satyam,K., Shilpi, J. Retinal Arterial Tortuosity In Young Indians: A Cross Sectional Study. The Internet Journal of Health. 2013 Volume 14 ,Number 1.

Hayreh, S.S., (1974) Anatomy and physiology of the optic nerve head. Transactions - American Academy of Ophthalmology and Otolaryngology 78: 240-254

Henkind, P., Hansen, R.I., Szalay J. (1979) Ocular circulation. In "Physiology of the human eye and visual system" Harper & Row, New York

Heron, L., Virsolvy, A., Peyrollier, K., Gribble, F.M., Le Cam, A., Ashcroft, F.M., and Bataille, D. (1998). Human alpha-endosulfine, a possible regulator of sulfonylurea-sensitive KATP channel: molecular cloning, expression and biological properties. Proc Natl Acad Sci U S A 95, 8387-8391.

HILL, G.S. (1970). Studies on the Pathogenesis of Hypertensive Vascular Disease: Effect of High-Pressure Intra-Arterial Injections in Rats. Circulation Research 27, 657-668

Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). 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.

Hinits, Y., and Hughes, S.M. (2007). Mef2s are required for thick filament formation in nascent muscle fibres. Development 134, 2511-2519

Hoggart, C.J., Clark, T.G., De Iorio, M., Whittaker, J.C., and Balding, D.J. (2008). Genome-wide significance for dense SNP and resequencing data. Genet Epidemiol 32, 179-185.

Hu, P., Shi, B., Geng, F., Zhang, C., Wu, W., and Wu, X.Z. (2008). E-cadherin core fucosylation regulates nuclear beta-catenin accumulation in lung cancer cells. Glycoconjugate journal 25, 843-850.

Hubbard, L.D., Brothers, R.J., King, W.N., Clegg, L.X., Klein, R., Cooper, L.S., Sharrett, A.R., Davis, M.D., and Cai, J. (1999). Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the Atherosclerosis Risk in Communities Study. Ophthalmology 106, 2269-2280

Hughes, A.D., Wong, T.Y., Witt, N., Evans, R., Thom, S.A.M., Klein, B.E., Chaturvedi, N., and Klein, R. (2009). Determinants of Retinal Microvascular Architecture in Normal Subjects. Microcirculation 16, 159-166.

Ikram, M.K., De Jong, F.J., Van Dijk, E.J., Prins, N.D., Hofman, A., Breteler, M.M., and De Jong, P.T. (2006). Retinal vessel diameters and cerebral small vessel disease: the Rotterdam Scan Study. Brain 129, 182-188.

Ikram, M.K., de Jong, F.J., Vingerling, J.R., Witteman, J.C.M., Hofman, A., Breteler, M.M.B., and de Jong, P.T.V.M. (2004). Are Retinal Arteriolar or Venular Diameters Associated with Markers for Cardiovascular Disorders? The Rotterdam Study. Investigative Ophthalmology & Visual Science 45, 2129-2134.

Ikram MK, Janssen JA, Roos AM, Rietveld I, Witteman JC, Breteler MM, Hofman A, vanDuijn CM, de Jong PT (2006) Retinal vessel diameters and risk of impaired fasting glucose or diabetes: the Rotterdam study. Diabetes ,55,506-510.

Ikram, M.K., Sim, X., Jensen, R.A., Cotch, M.F., Hewitt, A.W., Ikram, M.A., Wang, J.J., Klein, R., Klein, B.E., Breteler, M.M., et al. (2010). Four novel Loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. PLoS Genet 6, 1001184

Imamura, Y., Noda, S., Hashizume, K., Shinoda, K., Yamaguchi, M., Uchiyama, S., Shimizu, T., Mizushima, Y., Shirasawa, T., and Tsubota, K. (2006). Drusen, choroidal neovascularization, and retinal pigment epithelium dysfunction in SOD1-deficient mice: A model of age-related macular degeneration. Proceedings of the National Academy of Sciences 103, 11282-11287.

Ioannidis, J.P., Thomas, G., and Daly, M.J. (2009). Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 10, 318-329.

Ivanisevic, M., Bojic, L., and Eterovic, D. (2000). Epidemiological study of nontraumatic phakic rhegmatogenous retinal detachment. Ophthalmic research 32, 237-239.

Jablonka, E., and Raz, G. (2009). Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q Rev Biol 84, 131-176.

Jensen, H., Saether, B.E., Ringsby, T.H., Tufto, J., Griffith, S.C., and Ellegren, H. (2003). Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (Passer domesticus). J Evol Biol 16, 1296-1307.

Joachim, N., Rochtchina, E., Tan, A.G., Hong, T., Mitchell, P., and Wang, J.J. (2012). Right and left correlation of retinal vessel caliber measurements in anisometropic children: effect of refractive error. Invest Ophthalmol Vis Sci 53, 5227-5230.

Jurata, L.W., and Gill, G.N. (1997). Functional analysis of the nuclear LIM domain interactor NLI. Mol Cell Biol 17, 5688-5698.

Kadrmas, J.L., and Beckerle, M.C. (2004). The LIM domain: from the cytoskeleton to the nucleus. Nat Rev Mol Cell Biol 5, 920-931.

Kannan, R., Jin, M., Gamulescu, M.A., and Hinton, D.R. (2004). Ceramide-induced apoptosis: role of catalase and hepatocyte growth factor. Free radical biology & medicine 37, 166-175.

Kaushik, S., Kifley, A., Mitchell, P., and Wang, J.J. (2007). Age, Blood Pressure, and Retinal Vessel Diameter: Separate Effects and Interaction of Blood Pressure and Age. Investigative Ophthalmology & Visual Science 48, 557-561.

Kawasaki, R., Che Azemin, M.Z., Kumar, D.K., Tan, A.G., Liew, G., Wong, T.Y., Mitchell, P., and Wang, J.J. (2011). Fractal dimension of the retinal vasculature and risk of stroke: a nested case-control study. Neurology 76, 1766-1767.

Kiefer, A.K., Tung, J.Y., Do, C.B., Hinds, D.A., Mountain, J.L., Francke, U., and Eriksson, N. (2013). Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. Plos Genet in press.

Kim, J., Swee, M., and Parks, W.C. (2009). Cytosolic SYT/SS18 isoforms are actin-associated proteins that function in matrix-specific adhesion. PloS one 4, e6455.

Klein, R., Klein, B.E., Tomany, S.C., and Wong, T.Y. (2004). The relation of retinal microvascular characteristics to age-related eye disease: the Beaver Dam eye study. Am J Ophthalmol 137, 435-444.

Klein R, M.C.E.K.M.D., and et al. (2012). Relationship of blood pressure and other factors to serial retinal arteriolar diameter measurements over time: The beaver dam eye study. Archives of Ophthalmology 130, 1019-1027.

Knudtson, M.D., Lee, K.E., Hubbard, L.D., Wong, T.Y., Klein, R., and Klein, B.E. (2003). Revised formulas for summarizing retinal vessel diameters. Curr Eye Res 27, 143-149.

Koh, V., Cheung, C.Y.-l., Zheng, Y., Wong, T.Y., Wong, W., and Aung, T. (2010). Relationship of Retinal Vascular Tortuosity with the Neuroretinal Rim: The Singapore Malay Eye Study. Investigative Ophthalmology & Visual Science 51, 3736-3741.

Kondo, M., Wang, L., and Bill, A. (1997). The role of nitric oxide in hyperaemic response to flicker in the retina and optic nerve in cats. Acta Ophthalmologica Scandinavica 75, 232-235.

Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Kristinsson, K.T., Frigge, M.L., et al. (2009). Parental origin of sequence variants associated with complex diseases. Nature 462, 868-874.

Kottgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., Pistis, G., Ruggiero, D., O'Seaghdha, C.M., Haller, T., et al. (2013). Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nat Genet 45, 145-154.

Kottgen, A., Pattaro, C., Boger, C.A., Fuchsberger, C., Olden, M., Glazer, N.L., Parsa, A., Gao, X., Yang, Q., Smith, A.V., et al. (2010). New loci associated with kidney function and chronic kidney disease. Nat Genet 42, 376-384.

Kryukov, G.V., Pennacchio, L.A., and Sunyaev, S.R. (2007). Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. Am J Hum Genet 80, 727-739.

Kurniawan, E.D., Cheung, N., Cheung, C.Y., Tay, W.T., Saw, S.M., and Wong, T.Y. (2012). Elevated Blood Pressure is Associated with Rarefaction of the Retinal Vasculature in Children. Investigative Ophthalmology & Visual Science 53, 470-474.

Lalonde, E., Ha, K.C., Wang, Z., Bemmo, A., Kleinman, C.L., Kwan, T., Pastinen, T., and Majewski, J. (2011). RNA sequencing reveals the role of splicing polymorphisms in regulating human gene expression. Genome Res 21, 545-554.

Lander, E.S. (1996). The New Genomics: Global Views of Biology. Science 274, 536-539.

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467, 832-838.

Lanigan, L.P., Clark, C.V., and Hill, D.W. (1988). Retinal circulation responses to systemic autonomic nerve stimulation. Eye 2, 412-417.

Laviad, E.L., Albee, L., Pankova-Kholmyansky, I., Epstein, S., Park, H., Merrill, A.H., and Futerman, A.H. (2008). Characterization of Ceramide Synthase 2. Journal of Biological Chemistry 283, 5677-5684.

Lee, K.E., Klein, B.E.K., Klein, R., and Knudtson, M.D. (2004). Familial Aggregation of Retinal Vessel Caliber in the Beaver Dam Eye Study. Investigative Ophthalmology & Visual Science 45, 3929-3933.

Lee, Sang H., Wray, Naomi R., Goddard, Michael E., and Visscher, Peter M. (2011). Estimating Missing Heritability for Disease from Genome-wide Association Studies. American journal of human genetics 88, 294-305.

Lee, T., Henry, J.D., Trollor, J.N., and Sachdev, P.S. (2010). Genetic influences on cognitive functions in the elderly: a selective review of twin studies. Brain Res Rev 64, 1-13.

Leung, H., Wang, J.J., Rochtchina, E., Tan, A.G., Wong, T.Y., Hubbard, L.D., Klein, R., and Mitchell, P. (2003). Computer-assisted retinal vessel measurement in an older population: correlation between right and left eyes. Clinical & Experimental Ophthalmology 31, 326-330.

Leung, H., Wang, J.J., Rochtchina, E., Tan, A.G., Wong, T.Y., Klein, R., Hubbard, L.D., and Mitchell, P. (2003). Relationships between Age, Blood Pressure, and Retinal Vessel Diameters in an Older Population. Investigative Ophthalmology & Visual Science 44, 2900-2904.

Levy, B.I., Ambrosio, G., Pries, A.R., and Struijker-Boudier, H.A. (2001). Microcirculation in hypertension: a new target for treatment? Circulation 104, 735-740.

LeWinter, M.M., and Granzier, H. (2010). Cardiac titin: a multifunctional giant. Circulation 121, 2137-2145.

Ley, K. (2003). The role of selectins in inflammation and disease. Trends in molecular medicine 9, 263-268.

Li, L.J., Cheung, C.Y., Ikram, M.K., Gluckman, P., Meaney, M.J., Chong, Y.S., Kwek, K., Wong, T.Y., and Saw, S.M. (2012). Blood pressure and retinal microvascular characteristics during pregnancy: Growing Up in Singapore Towards Healthy Outcomes (GUSTO) Study. Hypertension 60, 223-230.

Li, X. (2003). Incidence and epidemiological characteristics of rhegmatogenous retinal detachment in Beijing, China. Ophthalmology 110, 2413-2417.

Liew, G., Wang, J.J., Cheung, N., Zhang, Y.P., Hsu, W., Lee, M.L., Mitchell, P., Tikellis, G., Taylor, B., and Wong, T.Y. (2008). The Retinal Vasculature as a Fractal: Methodology, Reliability, and Relationship to Blood Pressure. Ophthalmology 115, 1951-1956.e1951.

Liew, G., Wang, J.J., Mitchell, P., and Wong, T.Y. (2008). Retinal vascular imaging: a new tool in microvascular disease research. Circ Cardiovasc Imaging 1, 156-161.

Lim, L.S., Cheung, C.Y.-l., Lin, X., Mitchell, P., Wong, T.Y., and Mei-Saw, S. (2011). Influence of Refractive Error and Axial Length on Retinal Vessel Geometric Characteristics. Investigative Ophthalmology & Visual Science 52, 669-678.

Lim, S.W., Cheung, N., Wang, J.J., Donaghue, K.C., Liew, G., Islam, F.M.A., Jenkins, A.J., and Wong, T.Y. (2009). Retinal Vascular Fractal Dimension and Risk of Early Diabetic Retinopathy: A prospective study of children and adolescents with type 1 diabetes. Diabetes Care 32, 2081-2083.

Lin, Q., Schwarz, J., Bucana, C., and Olson, E.N. (1997). Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. Science 276, 1404-1407.

Liu, Y.-P., Kuznetsova, T., Jin, Y., Thijs, L., Asayama, K., Gu, Y.-M., Bochud, M., Verhamme, P., Struijker-Boudier, H.A.J., and Staessen, J.A. (2013). Heritability of The Retinal Microcirculation in Flemish Families. American Journal of Hypertension.

Lynch, M. and Walsh, B. (1998). Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland, MA.

Lyssenko, V., Almgren, P., Anevski, D., Perfekt, R., Lahti, K., Nissén, M., Isomaa, B., Forsen, B., Homström, N., Saloranta, C., et al. (2005). Predictors of and longitudinal changes in insulin sensitivity and secretion preceding onset of type 2 diabetes. Diabetes 54, 166-174.

MacGillivray TJ, Patton N, Doubal FN, Graham C, Wardlaw JM (2007) Fractal analysis of the retinal vascular network in fundus images. In: Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE. pp 6455-6458.

Malliri, A., van Es, S., Huveneers, S., and Collard, J.G. (2004). The Rac exchange factor Tiam1 is required for the establishment and maintenance of cadherin-based adhesions. J Biol Chem 279, 30092-30098.

Manolio, T.A., Brooks, L.D., and Collins, F.S. (2008). A HapMap harvest of insights into the genetics of common disease. J Clin Invest 118, 1590-1605.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R. et al. (2009)Finding the missing heritability of complex diseases. Nature 8;461,747-753

Manolio, T.A. (2013). Bringing genome-wide association findings into clinical use. Nat Rev Genet 14, 549-558.

Mao, Y., and Finnemann, S.C. (2012). Essential diurnal Rac1 activation during retinal phagocytosis requires alphavbeta5 integrin but not tyrosine kinases focal adhesion kinase or Mer tyrosine kinase. Mol Biol Cell 23, 1104-1114.

Masters BR (2004) FRACTAL ANALYSIS OF THE VASCULAR TREE IN THE HUMAN RETINA. Annual Review of Biomedical Engineering 6: 427-452.

May, S.R., Stewart, N.J., Chang, W., and Peterson, A.S. (2004). A Titin mutation defines roles for circulation in endothelial morphogenesis. Developmental Biology 270, 31-46.

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., et al. (2008). Runs of homozygosity in European populations. Am J Hum Genet 83, 359-372.

Meguro, A., Ideta, H., Ota, M., Ito, N., Ideta, R., Yonemoto, J., Takeuchi, M., Uemoto, R., Nishide, T., Iijima, Y., et al. (2012). Common variants in the COL4A4 gene confer susceptibility to lattice degeneration of the retina. PloS one 7, e39300.

Michelson, E.L., Morganroth, J., Nichols, C.W., MacVaugh, H., (1979) III Retinal Arteriolar Changes as an Indicator of Coronary Artery Disease. Arch Intern Med.139:1139–1141

Mitchell, P., Leung, H., Wang, J.J., Rochtchina, E., Lee, A.J., Wong, T.Y., and Klein, R. (2005). Retinal vessel diameter and open-angle glaucoma: the Blue Mountains Eye Study. Ophthalmology 112, 245-250.

Mitry, D., Campbell, H., Charteris, D.G., Fleck, B.W., Tenesa, A., Dunlop, M.G., Hayward, C., Wright, A.F., and Vitart, V. (2011). SNP mistyping in genotyping arrays--an important cause of spurious association in case-control studies. Genet Epidemiol 35, 423-426.

Mitry, D., Charteris, D.G., Yorston, D., Siddiqui, M.A., Campbell, H., Murphy, A.L., Fleck, B.W., Wright, A.F., and Singh, J. (2010). The epidemiology and socioeconomic associations of retinal detachment in Scotland: a two-year prospective population-based study. Invest Ophthalmol Vis Sci 51, 4963-4968.

Mitry, D., Fleck, B.W., Wright, A.F., Campbell, H., and Charteris, D.G. (2010). Pathogenesis of rhegmatogenous retinal detachment: predisposing anatomy and cell biology. Retina (Philadelphia, Pa) 30, 1561-1572.

Mitry, D., Williams, L., Charteris, D.G., Fleck, B.W., Wright, A.F., and Campbell, H. (2011). Population-based estimate of the sibling recurrence risk ratio for rhegmatogenous retinal detachment. Invest Ophthalmol Vis Sci 52, 2551-2555.

Miyamoto, Y., Yamauchi, J., Tanoue, A., Wu, C., and Mobley, W.C. (2006). TrkB binds and tyrosine-phosphorylates Tiam1, leading to activation of Rac1 and induction of changes in cellular morphology. Proc Natl Acad Sci U S A 103, 10444-10449.

Montgomery, S.B., Sammeth, M., Gutierrez-Arcelus, M., Lach, R.P., Ingle, C., Nisbett, J., Guigo, R., and Dermitzakis, E.T. (2010). Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 464, 773-777.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segre, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 44, 981-990.

Moskvina, V., and Schmidt, K.M. (2008). On multiple-testing correction in genome-wide association studies. Genet Epidemiol 32, 567-573

Mowatt, L., Adrien, G., and Price, N. (2003). Ethnic differences in the demand incidence of retinal detachments in two districts in the West Midlands. Eye 17, 63-70.

Munier, A., Gunning, T., Kenny, D., and O'Keefe, M. (1998). Causes of blindness in the adult population of the Republic of Ireland. British Journal of Ophthalmology 82, 630-633.

Muraki, K., Iwata, Y., Katanosaka, Y., Ito, T., Ohya, S., Shigekawa, M., and Imaizumi, Y. (2003). TRPV2 is a component of osmotically sensitive cation channels in murine aortic myocytes. Circ Res 93, 829-838.

Murray C (1926) The physiological principle of minimum work 1. The vascular system and the cost of blood volume. Proc. Natl. Acad. Sci. USA,12,207–214

Nandrot, E.F., Kim, Y., Brodie, S.E., Huang, X., Sheppard, D., and Finnemann, S.C. (2004). Loss of synchronized retinal phagocytosis and age-related blindness in mice lacking alphavbeta5 integrin. The Journal of experimental medicine 200, 1539-1545.

Nejentsev, S., Walker, N., Riches, D., Egholm, M., and Todd, J.A. (2009). Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science 324, 387-389.

Norrelund, H., Christensen, K.L., Samani, N.J., Kimber, P., Mulvany, M.J., and Korsgaard, N. (1994). Early narrowed afferent arteriole is a contributor to the development of hypertension. Hypertension 24, 301-308.

Numakawa, T., Suzuki, S., Kumamaru, E., Adachi, N., Richards, M., and Kunugi, H. (2010). BDNF function and intracellular signaling in neurons. Histol Histopathol 25, 237-258.

O'Donnell, C.J., Kavousi, M., Smith, A.V., Kardia, S.L., Feitosa, M.F., Hwang, S.J., Sun, Y.V., Province, M.A., Aspelund, T., Dehghan, A., et al. (2011). Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. Circulation 124, 2855-2864.

O'Toole, T.E., Bialkowska, K., Li, X., and Fox, J.E. (2011). Tiam1 is recruited to beta1-integrin complexes by 14-3-3zeta where it mediates integrin-induced Rac1 activation and motility. Journal of cellular physiology 226, 2965-2978.

Owen, C.G., Rudnicka, A.R., Nightingale, C.M., Mullen, R., Barman, S.A., Sattar, N., Cook, D.G., and Whincup, P.H. (2011). Retinal Arteriolar Tortuosity and Cardiovascular Risk Factors in a Multi-Ethnic Population Study of 10-Year-Old Children; the Child Heart and Health Study in England (CHASE). Arteriosclerosis, Thrombosis, and Vascular Biology 31, 1933-1938.

Parr, J.C., and Spears, G.F. (1974). General caliber of the retinal arteries expressed as the equivalent width of the central retinal artery. Am J Ophthalmol 77, 472-477

Pakter, H.M., Fuchs, S.C., Maestri, M.K., Moreira, L.B., Dei Ricardi, L.M., Pamplona, V.F., Oliveira, M.M., and Fuchs, F.D. (2011). Computer-Assisted Methods to Evaluate Retinal Vascular Caliber: What Are They Measuring? Investigative Ophthalmology & Visual Science 52, 810-815.

Pardo, L.M., MacKay, I., Oostra, B., van Duijn, C.M., and Aulchenko, Y.S. (2005). The effect of genetic drift in a young genetically isolated population. Ann Hum Genet 69, 288-295.

Patton, N., Maini, R., MacGillivary, T., Aslam, T.M., Deary, I.J., and Dhillon, B. (2005). Effect of axial length on retinal vascular network geometry. Am J Ophthalmol 140, 648-653.

Patton N, Aslam TM, MacGillivray T, Deary IJ, Dhillon B, Eikelboom RH, Yogesan K, Constable IJ (2006) Retinal image analysis: Concepts, applications and potential. Progress in Retinal and Eye Research 25, 99-127.

Pe'er, I., Yelensky, R., Altshuler, D., and Daly, M.J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 32, 381-385.

Perez-Rovira, A., MacGillivray, T., Trucco, E., Chin, K.S., Zutis, K., Lupascu, C., Tegolo, D., Giachetti, A., Wilson, P.J., Doney, A., et al. (2011). VAMPIRE: Vessel assessment and measurement platform for images of the REtina. Conf Proc IEEE Eng Med Biol Soc 4, 3391-3394.

Perkins, E.S. (1981) The ocular pulse. Curr Eye Res. 1: 19– 23

Peters, A.L. (1995). Retinal detachment in black South Africans. South African medical journal 85, 158-159.

Philp, N.J., and Nachmias, V.T. (1987). Polarized distribution of integrin and fibronectin in retinal pigment epithelium. Invest Ophthalmol Vis Sci 28, 1275-1280.

Pirinen, M., Donnelly, P., and Spencer, C.C. (2012). Including known covariates can reduce power to detect genetic effects in case-control studies. Nat Genet 44, 848-851.

Plaisier, E., Chen, Z., Gekeler, F., Benhassine, S., Dahan, K., Marro, B., Alamowitch, S., Paques, M., and Ronco, P. (2010). Novel COL4A1 mutations associated with HANAC syndrome: a role for the triple helical CB3[IV] domain. Am J Med Genet A 10, 33659.

Polkinghorne, P.J., and Craig, J.P. (2004). Northern New Zealand Rhegmatogenous Retinal Detachment Study: epidemiology and risk factors. Clin Experiment Ophthalmol 32, 159-163.

Ponsioen, T.L., van Luyn, M.J., van der Worp, R.J., van Meurs, J.C., Hooymans, J.M., and Los, L.I. (2008). Collagen distribution in the human vitreoretinal interface. Invest Ophthalmol Vis Sci 49, 4089-4095.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38, 904-909.

Puranam, K., Qian, W.H., Nikbakht, K., Venable, M., Obeid, L., Hannun, Y., and Boustany, R.M. (1997). Upregulation of Bcl-2 and elevation of ceramide in Batten disease. Neuropediatrics 28, 37-41.

Purcell, S., Cherny, S.S., and Sham, P.C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19, 149-150.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American journal of human genetics 81, 559-575.

Raychaudhuri, S., Plenge, R.M., Rossin, E.J., Ng, A.C.Y., Purcell, S.M., Sklar, P., Scolnick, E.M., Xavier, R.J., Altshuler, D., Daly, M.J., et al. (2009). Identifying Relationships among Genomic Disease Regions: Predicting Genes at Pathogenic SNP Associations and Rare Deletions. PLoS Genet 5, e1000534.

Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. Trends Genet 17, 502-510.

Richards, A.J., Meredith, S., Poulson, A., Bearcroft, P., Crossland, G., Baguley, D.M., Scott, J.D., and Snead, M.P. (2005). A novel mutation of COL2A1 resulting in dominantly inherited rhegmatogenous retinal detachment. Invest Ophthalmol Vis Sci 46, 663-668.

Risch, N., and Merikangas, K. (1996). The Future of Genetic Studies of Complex Human Diseases. Science 273, 1516-1517.

Robinson, F, Riva, C.E., Grunwald, J.E., Petrig, B.L., Sinclair SH (1986) Retinal blood flow autoregulation in response to an acute increase in blood pressure, Invest. Ophthalmol. Vis.Sci. 27 ;722-726.

Rochtchina, E., Wang, J.J., Taylor, B., Wong, T.Y., and Mitchell, P. (2008). Ethnic Variability in Retinal Vessel Caliber: A Potential Source of Measurement Error from Ocular Pigmentation?—The Sydney Childhood Eye Study. Investigative Ophthalmology & Visual Science 49, 1362-1366.

Romeo, S., Yin, W., Kozlitina, J., Pennacchio, L.A., Boerwinkle, E., Hobbs, H.H., and Cohen, J.C. (2009). Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. J Clin Invest 119, 70-79.

Rosman, M., Wong, T.Y., Ong, S.G., and Ang, C.L. (2001). Retinal detachment in Chinese, Malay and Indian residents in Singapore: a comparative study on risk factors, clinical presentation and surgical outcomes. International ophthalmology 24, 101-106.

Rosman, M., Zheng, Y., Lamoureux, E., Saw, S.M., Aung, T., Tay, W.T., Wang, J.J., Mitchell, P., Tai, E.S., and Wong, T.Y. (2012). Review of key findings from the Singapore Malay Eye Study (SiMES-1). Singapore Med J 53, 82-87.

Rosman, M., Zheng, Y., Wong, W., Lamoureux, E., Saw, S.M., Tay, W.T., Wang, J.J., Mitchell, P., Tai, E.S., and Wong, T.Y. (2012). Singapore Malay Eye Study: rationale and methodology of 6-year follow-up study (SiMES-2). Clin Experiment Ophthalmol 40, 557-568.

Royal College of Ophthalmologists. (2010). Management of retinal detachment. In. (

Sanvicens, N., and Cotter, T.G. (2006). Ceramide is the key mediator of oxidative stress-induced apoptosis in retinal photoreceptor cells. Journal of neurochemistry 98, 1432-1444.

Sasaki, K., Ideta, H., Yonemoto, J., Tanaka, S., Hirose, A., and Oka, C. (1995). Epidemiologic characteristics of rhegmatogenous retinal detachment in Kumamoto, Japan. Graefes Arch Clin Exp Ophthalmol 233, 772-776.

Sasongko, M.B., Wang, J.J., Donaghue, K.C., Cheung, N., Benitez-Aguirre, P., Jenkins, A., Hsu, W., Lee, M.-L., and Wong, T.Y. (2010). Alterations in Retinal Microvascular Geometry in Young Type 1 Diabetes. Diabetes Care 33, 1331-1336.

Schenck, M., Carpinteiro, A., Grassme, H., Lang, F., and Gulbins, E. (2007). Ceramide: physiological and pathophysiological aspects. Archives of biochemistry and biophysics 462, 171-175.

Schepens, C.L. (1951). REtinal detachment and aphakia. Archives of Ophthalmology 45, 1-7.

Schneider, R., Rademacher, M., and Wolf, S. (1993). Lacunar infarcts and white matter attenuation. Ophthalmologic and microcirculatory aspects of the pathophysiology. Stroke 24, 1874-1879

Schunkert, H., Konig, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F., Barbalic, M., Gieger, C., et al. (2011). Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 43, 333-338.

Schwarzer, G. (2012). meta: Meta-Analysis with R. R package version 2.1-2.

Schwender, H., and Fritsch, A. (2012). scrime: Analysis of High-Dimensional Categorical Data such as SNP Data. R package version 1.2.9.

Scott, R.A., Lagou, V., Welch, R.P., Wheeler, E., Montasser, M.E., Luan, J.a., Magi, R., Strawbridge, R.J., Rehnberg, E., Gustafsson, S., et al. (2012). Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 44, 991-1005.

Sebag, J. (1992). Anatomy and pathology of the vitreo-retinal interface. Eye 6, 541-552.

Silva, L.C., Ben David, O., Pewzner-Jung, Y., Laviad, E.L., Stiban, J., Bandyopadhyay, S., Merrill, A.H., Jr., Prieto, M., and Futerman, A.H. (2012). Ablation of ceramide synthase 2 strongly affects biophysical properties of membranes. Journal of lipid research 53, 430-436.

Singapore Eye Research Intitute. (2011). Singapore "I" Vessel Assessment (SIVA):Computer-aided Integrated Platform for Large-scale non-invasive Observation of Cardiovascular Disorders Using Retina Image Analysis.

Sim, X., Jensen, R.A., Ikram, M.K., Cotch, M.F., Li, X., MacGregor, S., Xie, J., Smith, A.V., Boerwinkle, E., Mitchell, P., et al. (2013). Genetic loci for retinal arteriolar microcirculation. PloS one 8.

Sinnott, J.A., and Kraft, P. (2012). Artifact due to differential error when cases and controls are imputed from different platforms. Human genetics 131, 111-119.

Small, K.S., Hedman, A.K., Grundberg, E., Nica, A.C., Thorleifsson, G., Kong, A., Thorsteindottir, U., Shin, S.Y., Richards, H.B., Soranzo, N., et al. (2011). Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet 43, 561-564.

Smith, P.L., Myers, J.T., Rogers, C.E., Zhou, L., Petryniak, B., Becker, D.J., Homeister, J.W., and Lowe, J.B. (2002). Conditional control of selectin ligand expression and global fucosylation events in mice with a targeted mutation at the FX locus. The Journal of cell biology 158, 801-815.

Sng, C.C.A., Sabanayagam, C., Lamoureux, E.L., Liu, E., Lim, S.C., Hamzah, H., Lee, J., Tai, E.S., and Wong, T.Y. (2010). Fractal analysis of the retinal vasculature and chronic kidney disease. Nephrology Dialysis Transplantation.

Soares, J.V., Leandro, J.J., Cesar Junior, R.M., Jelinek, H.F., and Cree, M.J. (2006). Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. IEEE Trans Med Imaging 25, 1214-1222.

Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Lango Allen, H., Lindgren, C.M., Luan, J., Magi, R., et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 42, 937-948

SPSS Inc. (2007). SPSS for Windows. In Version 16 Chicago. (SPSS Inc.

Sreekumar, P.G., Ding, Y., Ryan, S.J., Kannan, R., and Hinton, D.R. (2009). Regulation of thioredoxin by ceramide in retinal pigment epithelial cells. Experimental eye research 88, 410-417.

Steinberg, R.H., and Wood, I. (1974). Pigment epithelial cell ensheathment of cone outer segments in the retina of the domestic cat. Proceedings of the Royal Society of London Series B, Containing papers of a Biological character Royal Society (Great Britain) 187, 461-478.

Storbeck, C.J., Wagner, S., O'Reilly, P., McKay, M., Parks, R.J., Westphal, H., and Sabourin, L.A. (2009). The Ldb1 and Ldb2 transcriptional cofactors interact with the Ste20-like kinase SLK and regulate cell migration. Mol Biol Cell 20, 4174-4182.

Stranger, B.E., Stahl, E.A., and Raj, T. (2011). Progress and Promise of Genome-Wide Association Studies for Human Complex Trait Genetics. Genetics 187, 367-383.

Strettoi, E., Gargini, C., Novelli, E., Sala, G., Piano, I., Gasco, P., and Ghidoni, R. (2010). Inhibition of ceramide biosynthesis preserves photoreceptor structure and function in a mouse model of retinitis pigmentosa. Proceedings of the National Academy of Sciences.

Sun, C., Zhu, G., Wong, T.Y., Hewitt, A.W., Ruddle, J.B., Hodgson, L., Montgomery, G.W., Young, T.L., Hammond, C.J., Craig, J.E., et al. (2009). Quantitative Genetic Analysis of the Retinal Vascular Caliber: The Australian Twins Eye Study. Hypertension 54, 788-795.

Tan, A., Wong, T.Y., Wang, J.J., Rochtchina, E., Klein, R., Mitchell, P., and Blue Mountains Eye Study. (2004). Does Refractive Error Influence the Association of Blood Pressure and Retinal Vessel Diameters? The Blue Mountains Eye Study. Invest Ophthalmol Vis Sci 45, 5257-.

Taarnhøj, N.C.B.B., Larsen, M., Sander, B., Kyvik, K.O., Kessel, L., Hougaard, J.L., and Sørensen, T.I.A. (2006). Heritability of Retinal Vessel Diameters and Blood Pressure: A Twin Study. Investigative Ophthalmology & Visual Science 47, 3539-3544.

Taarnhoj, N.C., Munch, I.C., Sander, B., Kessel, L., Hougaard, J.L., Kyvik, K., Sorensen, T.I., and Larsen, M. (2008). Straight versus tortuous retinal arteries in relation to blood pressure and genetics. Br J Ophthalmol 92, 1055-1060.

Tarasov, K.V., Sanna, S., Scuteri, A., Strait, J.B., Orru, M., Parsa, A., Lin, P.I., Maschio, A., Lai, S., Piras, M.G., et al. (2009). COL4A1 is associated with arterial stiffness by genome-wide association scan. Circ Cardiovasc Genet 2, 151-158

Taylor, B., Rochtchina, E., Wang, J.J., Wong, T.Y., Heikal, S., Saw, S.M., and Mitchell, P. (2007). Body mass index and its effects on retinal vessel diameter in 6-year-old children. Int J Obes 31, 1527-1533.

Tedeschi-Reiner, E., Strozzi, M., Skoric, B., Reiner, Z. (2005) Relation of Atherosclerotic Changes in Retinal Arteries to the Extent of Coronary Artery Disease. Am J Cardiol. 96,1107–1109.

Tenesa, A., Farrington, S.M., Prendergast, J.G., Porteous, M.E., Walker, M., Haq, N., Barnetson, R.A., Theodoratou, E., Cetnarskyj, R., Cartwright, N., et al. (2008). Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 40, 631-637.

The International HapMap Consortium. (2005). A haplotype map of the human genome. Nature 437, 1299-1320.

Thompson, E.A., and Shaw, R.G. (1992). Estimating polygenic models for multivariate data on large pedigrees. Genetics 131, 971-978.

Tomita, H., Abe, T., and Tamai, M. (2000). Ceramide-induced cell death in cultured rat retinal pigment epithelial cells. The Tohoku journal of experimental medicine 190, 223-229.

Torres, F.S., Fuchs, S.C., Maestri, M.K., Fuchs, F.D., Oliveira, M.M., Moreira, L.B., and Gus, M. (2013). Association between carotid intima-media thickness and retinal arteriolar and venular diameter in patients with hypertension: A cross-sectional study. Atherosclerosis 229, 134-138.

Tso, M.O., and Jampol, L.M. (1982). Pathophysiology of hypertensive retinopathy. Ophthalmology 89, 1132-1145

Turner, A., Lau, P., Soubeyrand, S., Jarinova, O., and McPherson, R. (2012). 421 Functional Relationship of the COL4A1/COL4A2 Locus on Chromosome 13q34 to Coronary Artery Disease (CAD). The Canadian journal of cardiology 28, S257-S258.

Turnock, D.C., Izquierdo, L., and Ferguson, M.A. (2007). The de novo synthesis of GDP-fucose is essential for flagellar adhesion and cell growth in Trypanosoma brucei. J Biol Chem 282, 28853-28863.

Tuson, M., Marfany, G., and Gonzalez-Duarte, R. (2004). Mutation of CERKL, a novel human ceramide kinase gene, causes autosomal recessive retinitis pigmentosa (RP26). Am J Hum Genet 74, 128-138.

Ursino, M. (1991) Mechanisms of cerebral blood flow regulation. Crit. Rev. Biomed. Eng. 18, 255–288.

Van der Graaf., K.M., Fox., S.I., Lafleur., K.M. (1997) Synopsis of human natomy and physiology. WCB McGraw-Hill, 307-317.

Varma, R., Tielsch, J.M., Quigley, H.A., Hilton, S.C., Katz, J., Spaeth, G.L., and Sommer, A. (1994). Race-, age-, gender-, and refractive error-related differences in the normal optic disc. Arch Ophthalmol 112, 1068-1076.

Verhoeven, V.J., Hysi, P.G., Wojciechowski, R., Fan, Q., Guggenheim, J.A., Hohn, R., Macgregor, S., Hewitt, A.W., Nag, A., Cheng, C.Y., et al. (2013). Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. Nat Genet 10.

Verbeek, E., Meuwissen, M.E., Verheijen, F.W., Govaert, P.P., Licht, D.J., Kuo, D.S., Poulton, C.J., Schot, R., Lequin, M.H., Dudink, J., et al. (2012). COL4A2 mutation associated with familial porencephaly and small-vessel disease. Eur J Hum Genet 20, 844-851

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F. et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457:854-858.

Visscher, P.M., Hill, W.G., and Wray, N.R. (2008). Heritability in the genomics era--concepts and misconceptions. Nat Rev Genet 9, 255-266.

Vogan, K. (2012). TTN mutations in cardiomyopathy. Nat Genet 44, 368-368.

Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 42, 579-589.

Wang, D.Y., Ray, A., Rodgers, K., Ergorul, C., Hyman, B.T., Huang, W., and Grosskreutz, C.L. (2010). Global Gene Expression Changes in Rat Retinal Ganglion Cells in Experimental Glaucoma. Invest Ophthalmol Vis Sci 51, 4084-4095.

Wang, J.J., Mitchell, P., Rochtchina, E., Tan, A.G., Wong, T.Y., and Klein, R. (2004). Retinal vessel wall signs and the 5 year incidence of age related maculopathy: the Blue Mountains Eye Study. British Journal of Ophthalmology 88, 104-109.

Wang, J.J., Rochtchina, E., Liew, G., Tan, A.G., Wong, T.Y., Leeder, S.R., Smith, W., Shankar, A., and Mitchell, P. (2008). The Long-term Relation among Retinal Arteriolar Narrowing, Blood Pressure, and Incident Severe Hypertension. American Journal of Epidemiology 168, 80-88.

Wang, J.J., Taylor, B., Wong, T.Y., Chua, B., Rochtchina, E., Klein, R., and Mitchell, P. (2006). Retinal vessel diameters and obesity: a population-based study in older persons. Obesity 14, 206-214.

Weinhold, N., Johnson, D.C., Chubb, D., Chen, B., Forsti, A., Hosking, F.J., Broderick, P., Ma, Y.P., Dobbins, S.E., Hose, D., et al. (2013). The CCND1 c.870G>A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma. Nat Genet 45, 522-525

Welcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-678.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190-2191.

Witt, N., Wong, T.Y., Hughes, A.D., Chaturvedi, N., Klein, B.E., Evans, R., McNamara, M., Thom SAM, Klein, R. (2006) Abnormalities of Retinal Microvascular Structure and Risk of Mortality From Ischemic Heart Disease and Stroke. Hypertension 47, 975-981.

Wong, T.Y., Tielsch, J.M., and Schein, O. (1999). Racial difference in the incidence of retinal detachment in Singapore. Archives of Ophthalmology 117, 379-383.

Wong, T.Y., Klein, R., Klein, B.E.K., Tielsch, J.M., Hubbard, L., and Nieto, F.J. (2001). Retinal Microvascular Abnormalities and their Relationship with Hypertension, Cardiovascular Disease, and Mortality. Survey of Ophthalmology 46, 59-80.

Wong, T.Y., Hubbard, L.D., Klein, R., Marino, E.K., Kronmal, R., Sharrett, A.R., Siscovick, D.S., Burke, G., Tielsch, J.M. (2002) Retinal microvascular abnormalities and blood pressure in older people: the Cardiovascular Health Study. Br J Ophthalmol 86, 1007-1013

Wong, T.Y., Knudtson, M.D., Klein, R., Klein, B.E., Meuer, S.M., and Hubbard, L.D. (2004). Computer-assisted measurement of retinal vessel diameters in the Beaver Dam Eye Study: methodology, correlation between eyes, and effect of refractive errors. Ophthalmology 111, 1183-1190.

Wong, T.Y., Shankar, A., Klein, R., Klein, B.E., and Hubbard, L.D. (2004). Prospective cohort study of retinal vessel diameters and risk of hypertension. Bmj 329, 2.

Wong, T.Y., Wang, J.J., Rochtchina, E., Klein, R., and Mitchell, P. (2004). Does refractive error influence the association of blood pressure and retinal vessel diameters? the blue mountains eye study. American journal of ophthalmology 137, 1050-1055.

Wong, T.Y., Rosamond, W., Chang, P.P., Couper, D.J., Sharrett, A.R., Hubbard, L.D., Folsom, A.R., Klein, R. (2005) Retinopathy and Risk of Congestive Heart Failure. JAMA. 293:63–69

Wong, T.Y., and McIntosh, R. (2005). Systemic associations of retinal microvascular signs: a review of recent population-based studies. Ophthalmic and Physiological Optics 25, 195-204

Wong, T.Y., Islam, F.M.A., Klein, R., Klein, B.E.K., Cotch, M.F., Castro, C., Sharrett, A.R., and Shahar, E. (2006). Retinal Vascular Caliber, Cardiovascular Risk Factors, and Inflammation: The Multi-Ethnic Study of Atherosclerosis (MESA). Investigative Ophthalmology & Visual Science 47, 2341-2350.

Wong, T.Y., Cheung, N., Islam, F.M.A., Klein, R., Criqui, M.H., Cotch, M.F., Carr, J.J., Klein, B.E.K., and Sharrett, A.R. (2008). Relation of Retinopathy to Coronary Artery Calcification: The Multi-Ethnic Study of Atherosclerosis. American Journal of Epidemiology 167, 51-58.

Wright, A., Charlesworth, B., Rudan, I., Carothers, A., and Campbell, H. (2003). A polygenic basis for late-onset disease. Trends Genet 19, 97-106.

Xing C,. Klein BEK, Klein R,. Jun g, Lee KE, Iyengar SK (2006) Genome-Wide Linkage Study of Retinal Vessel Diameters in the Beaver Dam Eye Study, Hypertension 47, 797-802.

Xu, J., Gong, N.L., Bodi, I., Aronow, B.J., Backx, P.H., and Molkentin, J.D. (2006). Myocyte enhancer factors 2A and 2C induce dilated cardiomyopathy in transgenic mice. J Biol Chem 281, 9152-9162.

Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42, 565-569.

Yoneda, Y., Haginoya, K., Arai, H., Yamaoka, S., Tsurusaki, Y., Doi, H., Miyake, N., Yokochi, K., Osaka, H., Kato, M., et al. (2012). De novo and inherited mutations in COL4A2, encoding the type IV collagen alpha2 chain cause porencephaly. Am J Hum Genet 90, 86-90.

Yu, T.-Y., Acosta, M.L., Ready, S., Cheong, Y.-L., and Kalloniatis, M. (2007). Light exposure causes functional changes in the retina: increased photoreceptor cation channel permeability, photoreceptor apoptosis, and altered retinal metabolic function. Journal of neurochemistry 103, 714-724

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M., and Corden, J.L. (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc Natl Acad Sci U S A 93, 6975-6980.

Zamir, M. (1976). Optimality principles in arterial branching. J Theor Biol 62, 227-251.

Zarbin, M.A., Green, W., Moser, A.B., and Tiffany, C. (1988). Increased levels of ceramide in the retina of a patient with Farber's disease. Archives of Ophthalmology 106, 1163-1163.

Zhang, L., Ishikawa, O., Takeuchi, Y., and Miyachi, Y. (1998). Immunohistochemical distribution of epimorphin in human and mouse tissues. Histochem J 30, 903-908.

Zhang HR (1994) Scanning electron-microscopic study of corrosion casts on retinal and choroidal angioarchitecture in man and animals. Prog. Ret. Eye Res. 13, 243-270.

Zhao, Y., Itoh, S., Wang, X., Isaji, T., Miyoshi, E., Kariya, Y., Miyazaki, K., Kawasaki, N., Taniguchi, N., and Gu, J. (2006). Deletion of core fucosylation on alpha3beta1 integrin down-regulates its functions. J Biol Chem 281, 38343-38350.

Zigmond , M., Bloom F.,Landis S., Roberts J. , Squire L. (1999) Fundamental Neuroscience,Academic Press, San Diego; 727-743.

Zipin, A., Israeli-Amit, M., Meshel, T., Sagi-Assif, O., Yron, I., Lifshitz, V., Bacharach, E., Smorodinsky, N.I., Many, A., Czernilofsky, P.A., et al. (2004). Tumor-microenvironment interactions: the fucose-generating FX enzyme controls adhesive properties of colorectal cancer cells. Cancer research 64, 6571-6578.

Zou, H., Zhang, X., Xu, X., Wang, X., Liu, K., and Ho, P.C. (2002). Epidemiology survey of rhegmatogenous retinal detachment in Beixinjing District, Shanghai, China. Retina (Philadelphia, Pa) 22, 294-299.

APPENDICES

Appendix I

Sex separated results RRD

Female only meta-analysis; Dutch, Englsih and Scottish original cohort

chr	position	MarkerName	Allele1	Allele2	OR	P-value	Direction	HetISq	GENE
6	121168181	rs2789074	a	g	1.58	2.10E-06	+++	0	intergenic; c6orf170
9	93321320	rs7856591	a	c	1.52	2.87E-06	+++	35.4	intergenic;ROR2/NFIL3
6	121301788	rs12193473	a	g	0.74	9.22E-06	---	19.4	intergenic; c6orf170
6	121410867	rs6914477	a	g	1.37	1.57E-05	+++	0	intergenic; c6orf171
6	121193027	rs2789019	a	g	0.68	1.73E-05	---	52.6	intergenic; c6orf172
6	121481259	rs12206641	a	c	1.37	1.83E-05	+++	41.2	intergenic; c6orf173
11	132275836	rs2725451	a	g	0.74	1.90E-05	---	0	OPCML
6	121388196	rs218843	a	g	1.37	1.92E-05	+++	16.7	intergenic; c6orf173
6	121477769	rs13199297	a	g	1.35	3.21E-05	+++	51.4	intergenic; c6orf174
6	121330001	rs12202993	a	g	1.36	3.49E-05	+++	9.8	intergenic; c6orf175

Male only meta-analysis; Dutch, English and Scottish original cohort

chr	pos	MarkerName	Allele1	Allele2	OR	P-value	Direction	HetISq	GENE
2	41814174	rs1920480	a	g	1.29	2.01E-05	+++	0	
20	59890586	rs8118581	a	g	1.99	2.13E-05	+++	22.6	CDH4
3	23900401	rs6779573	a	c	1.26	3.50E-05	+++	7.3	UBE2E1
2	60068526	rs6713648	a	g	1.65	3.57E-05	+++	0	intergenic
13	30014373	rs7337496	a	g	0.77	4.23E-05	---	25.4	HPCAL4
19	24045852	rs4932927	a	g	1.78	5.16E-05	+++	39.4	intergenic; RPSA P58/ZNF7 26
2	41835431	rs1560260	a	g	1.26	5.51E-05	+++	0	
19	367313	rs585487	a	g	0.79	6.05E-05	---	0	THEG
17	52843452	rs9914327	a	g	1.33	6.11E-05	-++	21.7	intergenic; TOM1L1
18	56165682	rs8086573	a	c	0.73	6.90E-05	+?-	60.2	ALPK2
19	366451	rs689448	a	c	0.79	8.07E-05	---	0	THEG
18	24494875	rs1143735	a	g	1.72	8.64E-05	+-+	0	CHST9
10	74090186	rs11000266	a	g	0.67	8.93E-05	---	0	intergenic; DNAJB12
16	24802325	rs3803716	a	g	0.73	0.000113	+--	0	TNRC6A

Appendix II
Magenta top 50 pathway results

DB	GS	FDR 95PERC CUTOFF	GENES FLAGGED	FLAGGED GENE_NAMES
PANTHER_BIOLOGICAL_PROCESS	Nuclear_transport	1.20E-01	0	
PANTHER_BIOLOGICAL_PROCESS	mRNA_transcription	2.85E-01	0	
PANTHER_BIOLOGICAL_PROCESS	Phagocytosis	3.23E-01	1	RGNEF
PANTHER_BIOLOGICAL_PROCESS	Signal_transduction	6.97E-01	1	RGNEF
PANTHER_BIOLOGICAL_PROCESS	DNA_degradation	7.10E-01	0	
GOTERM	negative regulation of gene-specific transcription	7.52E-01	1	BMP7
PANTHER_BIOLOGICAL_PROCESS	Cell_structure_and_motility	7.65E-01	2	TIAM1, RGNEF
GOTERM	superoxide metabolic process	7.85E-01	1	SOD1
GOTERM	purinergic nucleotide receptor activity, G-protein coupled	7.85E-01	0	
PANTHER_BIOLOGICAL_PROCESS	Regulation_of_lipid,_fatty_acid_and_steroid_metabolism	7.89E-01	0	
GOTERM	mismatch repair	7.95E-01	0	
GOTERM	protein tyrosine kinase activity	7.95E-01	1	EPHB2
GOTERM	cilium membrane	7.96E-01	0	
GOTERM	mitochondrial membrane	8.05E-01	0	
GOTERM	quinone binding	8.08E-01	0	
GOTERM	hippocampus development	8.12E-01	0	
GOTERM	microvillus	8.13E-01	0	
GOTERM	phagocytosis	8.18E-01	0	
GOTERM	manganese ion binding	8.19E-01	1	PRUNE
GOTERM	SH3 domain binding	8.23E-01	1	CD2AP
GOTERM	extracellular matrix	8.26E-01	3	BMP7, SOD1, ADAMT
GOTERM	trophectodermal cell differentiation	8.27E-01	0	
PANTHER_BIOLOGICAL_PROCESS	Other_amino_acid_metabolism	8.29E-01	0	
GOTERM	negative regulation of MAP kinase activity	8.33E-01	1	BMP7
GOTERM	negative regulation of microtubule depolymerization	8.35E-01	0	
GOTERM	proteolysis involved in cellular protein catabolic process	8.35E-01	1	PSMA8
GOTERM	peripheral nervous system development	8.38E-01	0	
GOTERM	oligodendrocyte development	8.38E-01	0	
GOTERM	transcription initiation factor activity	8.46E-01	1	TAF4B
GOTERM	release of cytochrome c from mitochondria	8.47E-01	0	
GOTERM	septin complex	8.50E-01	0	
GOTERM	translation factor activity, nucleic acid binding	8.51E-01	0	
GOTERM	germ cell development	8.51E-01	0	
GOTERM	protein amino acid deacetylation	8.51E-01	0	
GOTERM	embryonic pattern specification	8.52E-01	1	BMP7
GOTERM	protein self-association	8.53E-01	0	
GOTERM	specific transcriptional repressor activity	8.54E-01	0	
GOTERM	integral to Golgi membrane	8.54E-01	0	
GOTERM	negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle	8.55E-01	0	
GOTERM	protein polyubiquitination	8.56E-01	0	
GOTERM	spermatogenesis	8.57E-01	2	SOD1, SPO11
GOTERM	neuropeptide receptor activity	8.57E-01	1	NPFFR1
PANTHER_MOLECULAR_FUNCTION	DNA_methyltransferase	8.58E-01	0	
GOTERM	exonuclease activity	8.61E-01	0	
GOTERM	transcription factor TFIID complex	8.62E-01	1	TAF4B
GOTERM	metal ion transmembrane transporter activity	8.65E-01	0	
GOTERM	regulation of Rho protein signal transduction	8.65E-01	2	TIAM1, RGNEF
GOTERM	snRNA processing	8.65E-01	0	
GOTERM	basolateral plasma membrane	8.67E-01	0	

Appendix III

Primary GWAS results for the top 23 most significant SNPs in the discovery phase

No high myopes in RRD cases									
SNP	chr	position	build	effallele	effallelefreq	n	effect	se(effect)	p-value
rs12960119	18	21868701	36	G	0.09	2525	0.0178	0.0050	0.00038
rs267738	1	149207249	36	C	0.23	2524	-0.0091	0.0026	0.00065
rs955943	4	16123506	36	A	0.04	2525	0.0147	0.00797	0.0654
rs7097067	10	71676508	36	A	0.03	2525	0.0238	0.0089	0.0076
rs1074463	5	22671126	36	A	0.13	2524	0.0117	0.00408	0.0041
rs2045084	8	144764982	36	G	0.34	2500	0.0057	0.00195	0.0034
rs8132771	21	31921294	36	A	0.07	2525	0.0213	0.00620	0.0006
rs11259960	15	81379707	36	A	0.20	2525	0.0104	0.00301	0.00054
rs2368106	2	180815911	36	A	0.10	2524	0.0107	0.0047	0.02354
rs7234959	18	21816517	36	A	0.11	2523	0.0169	0.0046	0.00027
rs6070015	20	55192062	36	A	0.05	2524	0.0164	0.0074	0.027
rs913444	9	87051470	36	A	0.26	2523	0.0059	0.0024	0.016
rs4893905	2	180867194	36	A	0.10	2523	0.0083	0.0049	0.0867
rs1477441	5	22569230	36	A	0.13	2525	0.0115	0.0041	0.0053
rs12193473	6	121301788	36	G	0.39	2524	0.0058	0.0017	0.0005
rs218843	6	121388196	36	A	0.40	2525	0.0054	0.0016	0.0010
rs4715056	6	47977103	36	G	0.28	2524	-0.0065	0.0023	0.0042
rs2817896	1	22988636	36	G	0.26	2524	0.00964	0.0025	0.000098
rs10515162	5	73090012	36	C	0.09	2524	-0.0082	0.0051	0.109
rs11181447	12	41026298	36	A	0.13	2525	0.0076	0.0040	0.057
rs12202993	6	121330001	36	A	0.39	2525	0.0056	0.0017	0.0008
rs564351	6	47927523	36	A	0.19	2525	-0.0083	0.0032	0.0089
rs6035211	20	19008513	36	C	0.13	2525	0.0067	0.0041	0.105
All RRD cases									
rs12960119	18	21868701	36	G	0.10	2819	0.019	0.0054	0.00039
rs267738	1	14920729	36	C	0.23	2817	-0.0095	0.0030	0.00123
rs955943	4	16123506	36	A	0.04	2819	0.0238	0.0085	0.00534
rs7097067	10	71676508	36	A	0.04	2819	0.0283	0.0096	0.00308
rs1074463	5	22671126	36	A	0.13	2818	0.0154	0.0044	0.00047
rs2045084	8	144764982	36	G	0.35	2794	0.0081	0.0021	0.00012
rs8132771	21	31921294	36	A	0.07	2819	0.0183	0.0068	0.00701
rs11259960	15	81379707	36	A	0.20	2819	0.0144	0.0032	0.000009
rs2368106	2	180815911	36	A	0.11	2817	0.0188	0.0051	0.00021
rs7234959	18	21816517	36	A	0.11	2817	0.0181	0.0051	0.00035
rs6070015	20	55192062	36	A	0.05	2818	0.0291	0.0078	0.00018
rs913444	9	87051470	36	A	0.27	2817	0.00799	0.0026	0.0024
rs4893905	2	180867194	36	A	0.10	2817	0.0179	0.0052	0.00063
rs1477441	5	22569230	36	A	0.13	2819	0.0157	0.0044	0.00040
rs12193473	6	121301788	36	G	0.40	2817	0.0070	0.0018	0.00012
rs218843	6	121388196	36	A	0.40	2819	0.00698	0.0017	0.000095
rs4715056	6	47977103	36	G	0.28	2818	-0.0074	0.0025	0.0032
rs2817896	1	22988636	36	G	0.26	2818	0.0067	0.0027	0.0144
rs10515162	5	73090012	36	C	0.09	2818	-0.0173	0.0057	0.0024
rs11181447	12	41026298	36	A	0.14	2819	0.013	0.0043	0.00187
rs12202993	6	121330001	36	A	0.40	2819	0.0068	0.00180	0.00015
rs564351	6	47927523	36	A	0.19	2819	-0.0099	0.00349	0.0046
rs6035211	20	19008513	36	C	0.13	2819	0.0129	0.0044	0.0032

Appendix IV

Protocol for bifurcation angles and branching coefficient measurements

Parameters are measured using semi-automated software VAMPIRE (by TM and AP, written in Matlab) and are operated by one individual (MK)

1. Open VAMPIRE interface and load the images for analysis.
2. Identify optic disc location manually (press "Locate OD" button on the interface) for every image using the greyscale image (press "Grey image" button). The boundary is set on the inner circle. In some cases there may be lighter (fuzzy) area outside the optic disc, you should ignore this.
3. Start processing images (press "Process images" button)
4. After the processing is done tick the "processed junctions" box (all the junctions detected by the program appear on the screen).
5. Choose 3-6 junction points for both venules and arteriols
    - if images of both eyes are available for an individual choose the one with more actual branching points detected
    - images that are out of focus and ones with only few measurements are not used for further analysis
    - for every image check for the presence of faint arteries and interceptions of the vessel paths by ticking the "binary segmentation" box, such images are not used in the further analysis (they are put in a separate folder until the new segmentation algorithm is completed)
    - start with the most proximal junction points to the optic disc and go to up to the 3rd order junctions until at least 3 and maximum of 6 branching points are selected for both venules and arterioles
6. Mark the branching points of venules and arterioles separately (arterioles-white, venules-black) . Venule branching points are marked black by clicking on the point and holding the "shift" key. Arteriole points are left white (default after clicking on a branching point).
7. Save the chosen values (angle and branching coefficient) by pressing the "Save Results" button

Appendix V

SIVA grading protocol

1. Open image into SIVA
2. Process the images
3. Check if the image is gradable

    Criteria for gradable image:

    - Clear view of optic dsic
    - At least 4 venules and arterioles
    - Ensure zone C is not blocked

4. Optic disc placement
5. Detection of types of vessel
    - For hidden vessel start where the area becomes visible-closest to zone B
    - Arterioles- straighter, less tortuous, thinner than veins, paler in colour
    - Venules- more tortuous , thicker than arterioles, darker in colour
6. Trace the full length of vessels
    - Let SIVA trace as much as possible
    - Avoid manual tracing as much as possible
    - Covers traced all the way till the end of the Zone B; up to u appoint where grader is confident
    - Use black and white or Red-free filter for better view
    - Amend incorrect tracing
    - Take only vessel SIVA automatically traced
    - Ignore other small vessel not picked up by SIVA
    - Use red-free function to locate the end point of small vessels
7. Connect broken segments
8. “Cover” vessels
    - use covers to switch on and off only
    - a minimum of 5 covers are needed, no maximum number
    - use edit segment only if all the covers are wrong
    - use standard deviation of the mean width as a guide only
    - need to accept that SIVA will sometimes take shadows, but not when the covers are twice the width size
9. Check all vessels
10. Save data
11. Export data to .csv file

Appendix VI

AVR genome wide association analysis results with body fat percentage and pulse wave augmentation index used as covariates

Meta-analysis of AVR (Orkney and Korcula genotyped data) with pulse wave augmentation index as a covariate

name	effallele	CHR	position	beta	Se beta	eff allele freq	beta Korcula	beta Orkney	se Korcula	se Orkney	p
rs10904588	G	10	981737	0.26	0.05	0.43	0.27	0.25	0.10	0.06	1.85E-06
rs2884064	G	1	230386210	-0.37	0.08	0.88	-0.34	-0.41	0.12	0.11	2.75E-06
rs12734075	G	1	227759755	-0.29	0.06	0.23	-0.30	-0.29	0.11	0.08	4.15E-06
rs4610449	C	5	152139377	-0.70	0.16	0.97	-0.58	-0.72	0.39	0.17	6.48E-06
rs7088825	G	10	17544941	0.33	0.08	0.85	0.32	0.33	0.14	0.09	1.43E-05
rs7856322	C	9	9821142	-0.30	0.07	0.79	-0.37	-0.26	0.12	0.08	1.43E-05
rs11135327	G	5	164191345	0.27	0.06	0.76	0.22	0.30	0.11	0.08	1.55E-05
rs7735637	G	5	164191750	-0.27	0.06	0.24	-0.22	-0.29	0.11	0.08	1.59E-05
rs1993023	G	4	163484911	-0.28	0.07	0.79	-0.12	-0.35	0.12	0.08	1.78E-05
rs1460060	C	4	167034879	-0.26	0.06	0.31	-0.20	-0.28	0.12	0.07	1.92E-05
rs2167190	G	4	163495091	0.28	0.06	0.21	0.12	0.34	0.12	0.08	1.98E-05
rs8055236	C	16	81769899	-0.31	0.07	0.84	-0.35	-0.30	0.14	0.09	2.40E-05
rs2761748	C	9	9807514	0.23	0.06	0.43	0.21	0.24	0.10	0.07	3.23E-05
rs2570171	G	8	76000717	0.32	0.08	0.14	0.19	0.39	0.12	0.10	3.32E-05
rs5969812	G	X	150577130	0.19	0.05	0.40	0.33	0.14	0.09	0.05	3.48E-05
rs7208001	G	17	5095000	0.33	0.08	0.87	0.32	0.33	0.16	0.09	4.13E-05
rs2761752	G	9	9809028	0.23	0.06	0.42	0.22	0.24	0.10	0.07	4.14E-05
rs9984592	G	21	14449488	0.27	0.07	0.78	0.32	0.25	0.12	0.08	4.45E-05
rs10518706	G	15	51240801	-0.28	0.07	0.82	-0.16	-0.33	0.14	0.08	4.81E-05

Meta-analysis of AVR (Orkney and Korcula genotyped data) with percentage of body fat as a covariate

name	effallele	CHR	position	beta	se beta	eff allele freq	beta Korcula	beta Orkney	se Korcula	se Orkney	p
rs12734075	G	1	227759755	-0.25	0.05	0.22	-0.36	-0.19	0.09	0.06	9.45E-07
rs2884064	G	1	230386210	-0.29	0.06	0.87	-0.35	-0.23	0.09	0.09	5.90E-06
rs2026	G	2	234045405	-0.27	0.06	0.87	-0.42	-0.20	0.11	0.08	1.27E-05
rs7068299	G	10	72992635	0.21	0.05	0.77	0.18	0.23	0.08	0.06	1.93E-05
rs7121351	G	11	72086807	0.33	0.08	0.08	0.45	0.26	0.12	0.10	2.06E-05
rs2689790	G	7	79176147	0.74	0.17	0.95	0.74	NA	0.17	NA	2.16E-05
rs277366	G	1	75073505	-0.20	0.05	0.69	-0.19	-0.20	0.08	0.06	2.17E-05
rs1563188	G	18	1728361	-0.18	0.04	0.49	-0.20	-0.17	0.07	0.05	2.26E-05
rs5969812	G	X	150577130	0.15	0.04	0.39	0.19	0.14	0.07	0.04	2.29E-05
rs3122713	G	1	210866493	0.27	0.06	0.87	0.33	0.21	0.09	0.08	2.37E-05
rs9569076	C	13	54125829	-0.23	0.06	0.18	-0.24	-0.23	0.10	0.07	2.69E-05
rs1897139	G	3	143088651	-0.28	0.07	0.88	-0.39	-0.21	0.11	0.08	2.83E-05
rs29232	G	6	29719410	-0.18	0.04	0.64	-0.21	-0.17	0.08	0.05	3.56E-05
rs12839220	G	X	150549192	0.18	0.04	0.80	0.22	0.16	0.07	0.05	3.74E-05
rs277355	G	1	75063372	0.19	0.05	0.31	0.18	0.19	0.08	0.06	4.35E-05
rs10520531	G	4	183572185	-0.36	0.09	0.94	-0.32	-0.36	0.20	0.10	4.38E-05
rs9905895	G	17	22609947	0.18	0.04	0.61	0.15	0.19	0.08	0.05	4.76E-05
rs710652	C	12	69079783	0.30	0.07	0.50	0.30	NA	0.07	NA	4.89E-05
rs2306409	G	10	1036712	0.17	0.04	0.51	0.16	0.18	0.07	0.05	4.97E-05

Appendix VII

Genome wide meta-analysis of genotyped data form Orkney, Korcula and LBC for fractal dimensions measurements

Dbox meta-analysis of three cohorts

name	effallele	chr	position	beta	sebeta	eff allele freq	beta Korcula	beta Orkney	beta LBC	se Korcula	se Orkney	se LBC	p
rs7759428	G	6	85428117	-0.01	0.00	0.72	0.03	-0.02	-0.01	0.08	0.05	0.00	1.03E-05
rs7943162	G	11	91195221	0.01	0.00	0.94	0.11	-0.07	0.01	0.13	0.10	0.00	1.19E-05
rs748344	C	15	54177908	0.01	0.00	0.43	-0.02	-0.09	0.01	0.08	0.05	0.00	1.90E-05
rs625204	G	6	36077746	-0.01	0.00	0.41	-0.03	0.02	-0.01	0.07	0.05	0.00	2.23E-05
rs4339430	C	6	1.62E+08	0.01	0.00	0.10	-0.09	-0.05	0.01	0.11	0.08	0.00	3.33E-05
rs1254643	G	2	2.21E+08	-0.01	0.00	0.66	0.01	0.04	-0.01	0.08	0.05	0.00	4.23E-05
rs4411488	C	16	81693004	-0.01	0.00	0.10	0.18	0.02	-0.01	0.13	0.08	0.00	4.29E-05

D0

name	effallele	chr	position	beta	sebeta	eff allele freq	beta Korcula	beta Orkney	beta LBC	se Korcula	se Orkney	se LBC	p
rs2159042	G	17	66920924	-0.021	0.004	0.96	0.05	-0.23	-0.02	0.20	0.12	0.004	2.18E-07
rs917344	G	17	66910868	-0.021	0.004	0.96	0.01	-0.23	-0.02	0.23	0.13	0.004	2.21E-07
rs10488477	G	7	53518990	0.014	0.003	0.07	0.30	0.05	0.01	0.20	0.09	0.003	4.02E-06
rs10231525	G	7	1.44E+08	-0.012	0.003	0.89	0.09	0.06	-0.01	0.13	0.07	0.003	5.82E-06
rs2124216	G	11	1.02E+08	-0.007	0.002	0.73	-0.04	0.02	-0.01	0.10	0.05	0.002	1.26E-05
rs10487665	G	7	1.39E+08	0.011	0.003	0.10	0.15	-0.04	0.01	0.16	0.08	0.003	2.23E-05
rs3779258	G	7	1.23E+08	-0.011	0.003	0.90	-0.11	0.01	-0.01	0.12	0.08	0.003	2.24E-05
rs2361285	G	14	75956281	-0.008	0.002	0.83	-0.04	-0.01	-0.01	0.11	0.06	0.002	2.44E-05
rs2145641	C	14	96309884	0.011	0.003	0.09	0.08	0.03	0.01	0.12	0.09	0.003	2.61E-05
rs1991549	G	17	49927285	-0.012	0.003	0.92	-0.09	0.01	-0.01	0.12	0.10	0.003	3.25E-05
rs748344	C	15	54177908	0.006	0.002	0.43	-0.05	-0.10	0.01	0.08	0.05	0.002	3.56E-05
rs4533289	G	16	25840965	-0.010	0.002	0.88	-0.04	-0.04	-0.01	0.12	0.08	0.002	4.08E-05
rs2107785	G	7	80802211	0.010	0.002	0.10	-0.01	0.04	0.01	0.12	0.09	0.002	4.48E-05
rs12864063	G	13	92432395	-0.008	0.002	0.16	0.02	-0.07	-0.01	0.10	0.07	0.002	4.52E-05

Appendix VIII

Published paper

Human Molecular Genetics, 2013, Vol. 22, No. 15 ***3174–3185***
doi:10.1093/hmg/ddt169
Advance Access published on April 11, 2013

Genome-wide association study identifies genetic risk underlying primary rhegmatogenous retinal detachment

Mirna Kirin[1,†], Aman Chandra[2,3], David G. Charteris[2], Caroline Hayward[4], Susan Campbell[4], Ivana Celap[5], Goran Bencic[6], Zoran Vatavuk[6], Iva Kirac[7], Allan J. Richards[8,9], Albert Tenesa[4,10], Martin P. Snead[11], Brian W. Fleck[12], Jaswinder Singh[12], Steven Harsum[2], Robert E. MacLaren[2], Anneke I. den Hollander[13,14], Malcolm G. Dunlop[4], Carel B. Hoyng[13], Alan F. Wright[4], Harry Campbell[1], Veronique Vitart[4,†,*] and Danny Mitry[1,†]

[1]Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9AG, UK, [2]Moorfields Eye Hospital, City Road, London EC1 V2PD, UK, [3]UCL Institute of Ophthalmology, Bath Street, London EC1 V9EL, UK, [4]MRC Human Genetics Unit, IGMM, University of Edinburgh, Edinburgh EH4 2XU, UK, [5]Clinical Institute of Chemistry, [6]Department of Ophthalmology and, [7]Surgical Oncology Department, University Hospital for Tumours, University Clinical Hospital 'Sestre Milosrdnice', 10000 Zagreb, Croatia, [8]Molecular Genetics and Medical Genetics, Cambridge University NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK, [9]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK, [10]The Roslin Institute, The University of Edinburgh, Edinburgh EH25 9RG, UK, [11]Vitreoretinal Service, BOX 41, Cambridge University NHS Foundation Trust, Hills Road, Cambridge, UK, [12]Princess Alexandra Eye Pavilion, Edinburgh, UK, [13]Department of Ophthalmology and [14]Department of Human Genetics, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, the Netherlands

Received November 26, 2012; Revised and Accepted April 5, 2013

Rhegmatogenous retinal detachment (RRD) is an important cause of vision loss and can potentially lead to blindness. The underlying pathogenesis is complex and incompletely understood. We applied a two-stage genetic association discovery phase followed by a replication phase in a combined total of 2833 RRD cases and 7871 controls. The discovery phase involved a genome-wide association scan of 867 affected individuals and 1953 controls from Scotland, followed by genotyping and testing 4347 highest ranking or candidate single nucleotide polymorphisms (SNPs) in independent sets of cases (1000) and controls (2912) of Dutch and British origin. None of the SNPs selected reached a Bonferroni-corrected threshold for significance ($P < 1.27 \times 10^{-7}$). The strongest association, for rs12960119 ($P = 1.58 \times 10^{-7}$) located within an intron of the *SS18* gene. Further testing was carried out in independent case–control series from London (846 cases) and Croatia (120 cases). The combined meta-analysis identified one association reaching genome-wide significance for rs267738 (OR = 1.29, $P = 2.11 \times 10^{-8}$), a missense coding SNP and eQTL for *CERS2* encoding the protein ceramide synthase 2. Several of the top signals showing suggestive significance in the combined meta-analysis encompassed genes with a documented role in cell adhesion or migration, including *SS18*, *TIAM1*, *TSTA3* and *LDB2*, which warrant further investigation. This first genetic association study of RRD supports a polygenic component underlying RRD risk since 27.4% of the underlying RRD liability could be explained by the collective additive effects of the genotyped SNP from the discovery genome-wide scan.

*To whom correspondence should be addressed. Tel: +44 1316511071; Fax: +44 1314678456; Email: veronique.vitart@igmm.ed.ac.uk
†These authors contributed equally to this work.

INTRODUCTION

Despite advances in surgical treatment, rhegmatogenous retinal detachment (RRD) remains an important cause of visual loss. It accounts for up to 4% of the blind and partial sight registrations in Ireland (1). Following the formation of a break in the retina, the neural retina separates from the underlying retinal pigment epithelium (RPE) due to the egress of fluid from the vitreous cavity into the subretinal space. This leaves the neurosensory retina (NSR) deprived of nutrients and oxygen, severely reducing visual function, especially when the macula is affected. Pathological changes at the vitreoretinal interface are critical both to the development of a retinal break and to the progression of retinal detachment, since liquefied vitreous passes through the open break and the NSR is separated from the underlying RPE. The adhesion between the NSR and RPE is relatively weak since no junctional attachments exist between these two layers (2). The adhesion is maintained by the physical interconnection of the RPE cells' apical microvilli with the outer segments of NSR photoreceptors, mechanical support from the interphotoreceptor matrix and the net fluid transport from the retina to the choroid. Once these forces are overwhelmed, the subretinal space expands, allowing the fluid to accumulate and retinal detachment to occur. The exact mechanisms involved in this complex and synergistic adhesion system are incompletely understood.

The annual incidence of RRD is ~10–15 cases per 100 000 individuals (3), and prevalence ~1%, (4) with notable differences between ethnic groups (5–7). The rate of RRD has been reported to be higher in elderly patient (5,8,9) and in men compared with women (6,10–13). Other risk factors for RRD include high myopia, lattice retinal degeneration, trauma and previous cataract surgery.

A genetic predisposition to RRD is suggested by a number of inherited disorders (with known causal genetic variants) presenting RRD as an associated feature and by familial aggregation studies (14,15). Cases of non-syndromic familial RRD, inherited in a dominant fashion, have also been ascribed to mutations in the *COL2A1* gene (a known Stickler's syndrome gene [MIM 609508]) (16,17). Very little is known about the genes underlying non-syndromic RRD in the general population and no previous study has investigated the role of common genetic variants.

We conducted a genome-wide association study of RRD using a two-stage discovery phase followed by a replication phase (Fig. 1). A full genome-wide scan was first performed with close to 300 000 single nucleotide polymorphisms (SNPs) in a large collection of Scottish RRD cases ($n = 867$) (18) and ethnically matched controls ($n = 1953$) (19). The most significant

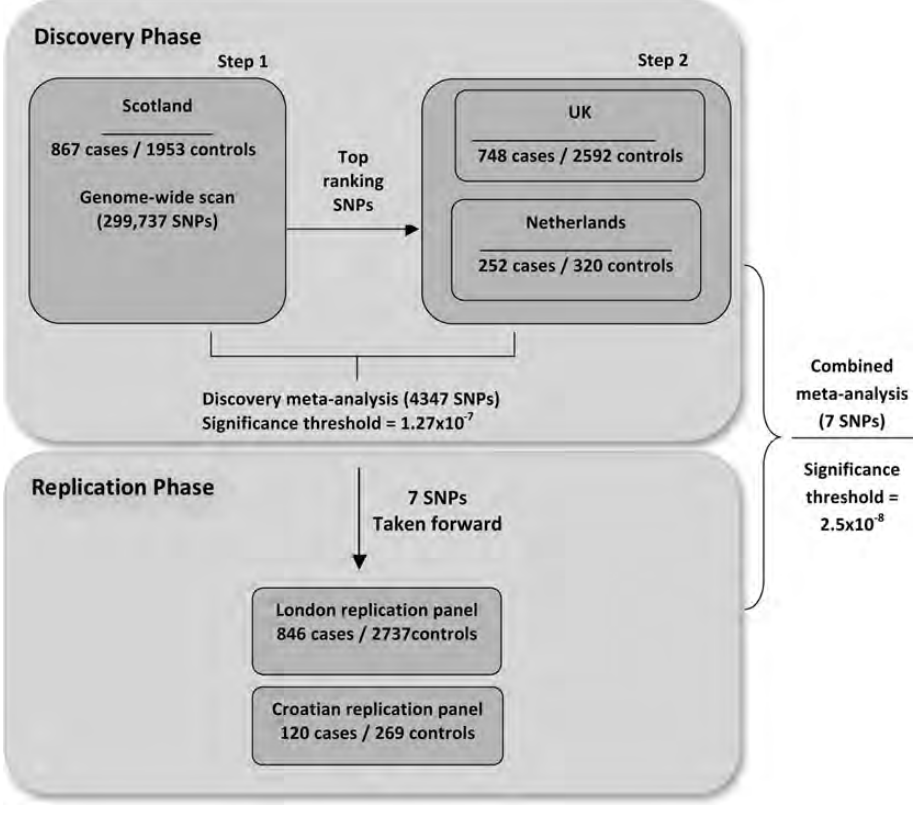


Figure 1. Our study included a two-stage discovery phase followed by a replication phase where the seven most promising findings were brought forward for testing. GWAS was conducted in a Scottish study in the first stage of the discovery phase. The best ranked SNPs were selected for follow-up and complemented with candidate gene markers. Four thousand and three hundred and forty-seven SNPs were successfully genotyped and analyzed in the UK and Dutch studies in the second stage of the discovery phase, and results from both discovery stages were combined in a meta-analysis (significance threshold for discovery phase meta-analysis was set to 1.27×10^{-7}). An overall combined meta-analysis of discovery (two stages) and replication phases was completed for the seven SNPs carried all along (significance threshold set at 2.5×10^{-8}). Studies are labeled according to the location of the center(s) where the RRD cases were collected.

results were then tested in independent sets of cases and controls of Dutch (252 cases; 320 controls) and British origin (748 cases; 2592 controls). Replication and follow-up was sought through the use of further samples from populations of English (846 cases; 2737 controls) and Croatian descent (120 cases; 269 controls). In total, 2833 cases and 7871 controls were analyzed in this study.

RESULTS

Association analysis

Discovery phase

A genome-wide association study (GWAS) was carried on 867 RRD cases and 1953 controls from Scotland, using 299 737 genotyped SNPs, after applying stringent quality control criteria. GWAS results showed a very slight excess of low *P*-values compared with those expected under the null hypothesis of no association (Fig. 2 and Supplementary Material, Fig. S2). One SNP, rs10510663, achieved genome-wide significance after correction for multiple testing ($P = 1.88 \times 10^{-9}$). This marker is located on chromosome 3p22.3, upstream of the *ARPP21* gene. The full GWAS results corresponding to Figure 2 are available Supplementary Material, Table S7.

The most significant SNPs associated with RRD by GWAS plus some selected candidate gene markers (see section 'Methods' and Supplementary Material, Table S1) were then genotyped on a custom-made Illumina iSelect array in an additional 1000 RRD cases (Discovery stage 2; 252 Dutch and 748 UK individuals). The meta-analysis of the SNP association data from these two steps constituted the discovery phase and involved analyses of three distinct case–control cohorts (of Scottish, British and Dutch ancestry). A stringent genome-wide significance threshold based on the number of SNPs tested in the initial whole-genome scan (1.27×10^{-7}; see section 'Methods') was applied in this meta-analysis. Index SNPs with association *P*-value $<10^{-4}$ are displayed in Table 1 for an additive model of allelic effect (for a dominant model, see Supplementary Material, Table S2). The highest ranking marker in the initial Discovery GWAS in Scottish cases (rs10510663) was not significantly associated with RRD in either the UK ($P = 0.86$) or Dutch cohorts ($P = 0.24$) and its association *P*-value after meta-analysis of discovery studies increased to be $>10^{-4}$. The most significantly associated marker after this stage was rs12960119, with an association *P*-value of 1.58×10^{-7}, just below the genome-wide significance threshold. This marker is located on chromosome 18q11.2 within an intron of the *SS18* gene (OR = 1.46; 1.26–1.67 95% CI). The magnitude and direction of the signals were consistent across all three populations, translating into a low heterogeneity score across studies ($I^2 = 0$; $P = 0.67$). Cases showed an excess of the rs12960119 minor allele G compared with controls in all three cohorts (Table 2). Another marker (rs7234959), located upstream of the *SS18* gene, displayed a suggestive *P*-value in the meta-analysis ($P = 2.68 \times 10^{-5}$), strengthening the validity of an association signal in this region. Most suggestive SNPs ($P < 10^{-4}$) also showed comparable directions of effect across all three populations (Table 1).

The genes putatively underlying the association signals were identified using a protocol implemented in GRAIL (20) (several genes were assigned to one association signal—Table 1) and inter loci functional networks sought out using the pathway analysis tool implemented in IPA (Ingenuity® Systems, www.ingenuity.com, last accessed 04/09/2012). Analysis using only the six most significant SNPs (additive model) as index SNPs revealed one highly significant network ($P = 10^{-62}$) involving 35 molecules in total, 21 of which were present in the top six signals (Supplementary Material, Fig. S3). This network showed enrichment for molecules involved in cell death, DNA replication, recombination and repair, and haematological system development and function. Additional pathway analysis was carried out using the 23 top results ($P < 10^{-4}$) from the discovery meta-analysis (Table 1). Four direct interactions were identified in this extended pathway analysis (*MCL1-RAE1*, rs267738-rs6070015; *BMP7-SETDB1*, rs6070015-rs267738; *SETDB1-PPA1*, rs267738-rs7097067; *TIAM1-TrkB*, rs81232771-rs913444). The latter is strongly supported by experimental evidence of a protein–protein interaction (21).

Replication phase

In the replication phase, we followed up the six most significantly associated markers displaying low heterogeneity across

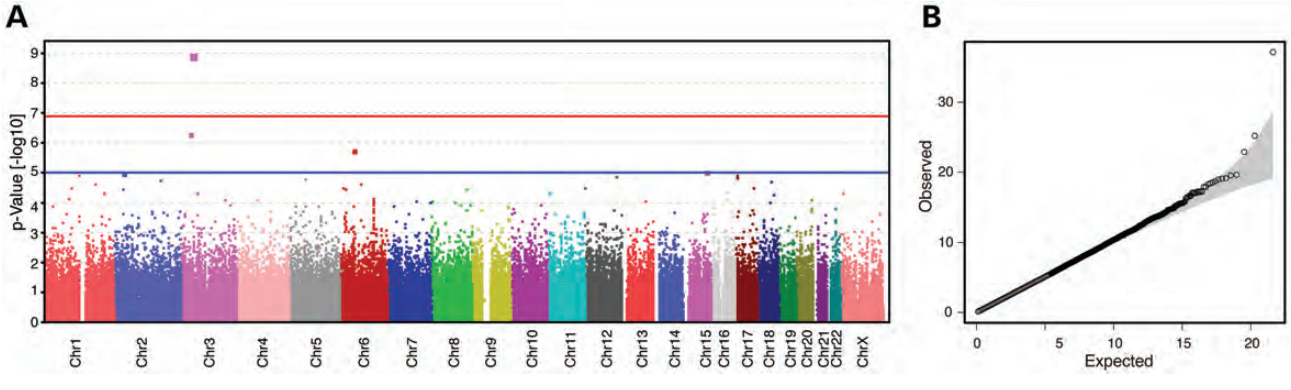


Figure 2. Manhattan plot (**A**) and Quantile–quantile (**B**) plot for the discovery stage1 genome-wide scan results. Statistics from a logistic regression adjusting for age sex and co-ancestry, performed on 867 Scottish RRD cases and 1953controls (SOCCS participants) using 299 737 genotyped SNPs which passed quality control, are displayed. In the Manhattan plot (A) −log10 (*P*-value) is plotted on the *x*-axis, and chromosomal location on the *y*-axis. In the QQ plot (B), distributions of observed and expected Chi-square values are compared.

Table 1. Highest ranking RRD association signals in the meta-analysis of the three discovery studies

SNP	Minor allele	CHR	Position (build36)	CEU minor allele freq	OR(95% CI)	Association *P*-value	Direction	Heterogeneity *P*-value	GRAIL input genes	Genes in LD block $r^2 > 0.8$ with index SNP
rs12960119	G	18	21 868 701	0.075	1.46 (1.26–1.67)	1.58E-07	+++	0.67	*PSMA8, TAF4B, SS18*	*SS18, PSAM8*
rs267738	C	1	149 207 249	0.275	0.79 (0.71–0.87)	6.7E-06	−−−	0.65	*GOLPH3L, CTSK, FAM63A, SETBD1, CTSS, MCL1, CERS2, ANXA9, PRUNE, ARNT, HORMAD1, ADAMTSL4, ENSA*	*CERS2*
rs955943	A	4	16 123 506	0.067	1.54 (1.27–1.87)	9.9E-06	+++	0.67	*LDB2*	*LDB2*
rs7097067	A	10	71 676 508	0.075	1.65 (1.32–2.07)	1.28E-05	−++	0.99	*PPA1, NPFFR1, TYSND1, AIMF2, SAR1A*	*PPA1, NPFFR1*[a]
rs1074463	A	5	22 671 126	0.092	1.31 (1.16–1.48)	1.28E-05	+++	0.73	*CDH12*	*CDH12*
rs2045084	G	8	144 764 982	0.425	1.21 (1.11–1.31)	1.54E-05	+++	0.25	*TSTA3*	*TSTA3*
rs8132771	A	21	31 921 294	0.075	1.43 (1.21–1.69)	1.96E-05	+++	0.21	*TIAM1, SOD1, SCAF4*	*SOD1, SCAF4*
rs11259960	A	15	81 379 707	0.200	1.35 (1.17–1.54)	2.36E-05	+−+	0.41	*HOMER2*	*HOMER2*
rs2368106	A	2	180 815 911	0.125	1.35 (1.17–1.54)	2.58E-05	−++	0.76	*CW22*	*CW22, UB2E3*[a]
rs7234959	A	18	21 816 517	0.083	1.34 (1.17–1.53)	2.68E-05	+++	0.77	*TAF4B, PSMA8, SS18*	*SS18, PSMA8*
rs6070015	A	20	55 192 062	0.067	1.67 (1.31–2.13)	2.91E-05	+++	0.9084	*RAE1, SPO11, BMP7, RBM38*	*BMP7*
rs913444	A	9	87 051 470	0.250	1.22 (1.11–1.34)	3.00E-05	+++	0.94	*TRKB*	*TRKB, AGTPBP1*[a]
rs4893905	A	2	180 867 194	0.100	1.33 (1.16–1.52)	3.32E-05	+++	0.59	*CW22*	*CW22, UB2E3*[a]
rs1477441	A	5	22 569 230	0.124	1.31 (1.15–1.49)	4.03E-05	+++	0.87	*CDH12*	*CDH12*
rs12193473	G	6	121 301 788	0.375	1.20 (1.10–1.31)	4.25E-05	+++	0.32	*C6orf170*	*MAN1A, C6orf170*[a]
rs218843	A	6	121 388 196	0.375	1.20 (1.10–1.31)	4.40E-05	+++	0.30	*C6orf170*	*MAN1A, C6orf170*[a]
rs4715056	G	6	47 977 103	0.275	0.81 (0.73–0.90)	4.94E-05	+−−	0.59	*GPR115, OPN5, CD2AP, GPR111*	*C6orf38*
rs2817896	G	1	22 988 636	0.242	1.24 (1.12–1.37)	5.35E-05	−++	0.82	*EPHB2*	*EPHB2*
rs10515162	C	5	73 090 012	0.050	0.72 (0.61–0.85)	6.48E-05	−−−	0.95	*RGNEF*	*RGNEF*
rs11181447	A	12	41 026 298	0.150	1.28 (1.13–1.45)	6.68E-05	+++	0.85	*GLT8D3, YAF2, ZCRB1, PPHLN1, PRICKLE1*	*ZCRB1, PPHLN1, PRICKLE1*
rs12202993	A	6	121 330 001	0.375	1.20 (1.10–1.31)	7.48E-05	+++	0.30	*C6orf170*	*MAN1A, C6orf170*[a]
rs564351	A	6	47 927 523	0.208	0.79 (0.70–0.89)	7.49E-05	+−−	0.91	*GPR115, OPN5, CD2AP, GPR111*	*OPN5, PTCHD4*[a]
rs6035211	C	20	19 008 513	0.117	1.30 (1.14–1.47)	8.66E-05	+++	0.99	*SLC24A3*	*C20orf79, SLC24A3*[a]

Odds ratios with 95% confidence intervals are given with respect to the minor allele for an additive model of allelic effect, association *P*-values, direction of the minor effect in, respectively, the Netherlands, UK and Scotland study (+ increasing risk, − decreasing) and the heterogeneity test *P*-values are also displayed. Genes potentially underlying each SNP signal as identified by a protocol within the GRAIL software and genes falling within the higher LD block tagged by the top associated SNP (defined by markers in high LD with it, $r^2 > 0.8$, based on based on CEU haplotypes HapMap2 build36 using the SNAP webtool http://www.broadinstitute.org/mpg/snap/ldplot.php, date last accessed 14/04/2013) are listed.

[a]Intergenic association signal, flanking genes are indicated.

studies (rs12960119, rs267738, rs955943, rs8132771, rs2045084, rs1074463—Table 1) and in addition, rs913444, the next highest ranking with low heterogeneity across studies marker with the highest minor allele frequency (to increase replication power). Furthermore, the product of the gene associated with this additional marker, *TrkB*, is known to bind and phosphorylate the product of *TIAM1* (21), a gene close to the rs8132771 association signal. A brief description of the genes included in or flanking the corresponding association signals is reported in Supplementary Material, Table S3. The seven selected markers were typed in 846 cases from London and 120 cases from Croatia (Fig. 1). All but one marker (rs1074463) displayed same direction of effect in the replication analysis compared with those in the discovery analysis, and one of them reached the Bonferroni-corrected significance threshold of 0.0035 (Supplementary Material, Tables S4 and S5). This marker, rs2045084, is located on chromosome 8q24.3, 949 bp 3′ of *TSTA3* and 7.8 kb 5′ of *PYCRL* gene and displayed a statistically significant association under a dominant model ($P = 9 \times 10^{-4}$, OR = 1.31) and nominally significant association under an additive model ($P = 0.028$, OR = 1.13). The next best associated marker, rs267738, is located on chromosome 1q21.3 within an exon of the *ceramide synthase 2* (*CERS2*) gene and displayed a nominally significant association under both additive and dominant models (additive $P = 7.1 \times 10^{-3}$, OR = 0.83). rs12960119, which almost reached genome-wide significance in the discovery phase, did not replicate in the London cohort ($P = 0.37$ additive effect). The minor allele of this marker (allele G, frequency = 0.092) was, however, associated with increased risk, as in the previous analysis, but the effect was smaller in size (OR additive model = 1.07, OR dominant model = 1.07). This reflected the frequency of the G allele of rs12960119 in the Wellcome Trust Case Control Consortium (WTCCC) NBS controls used in this replication analysis being higher (MAF = 0.087) than that of the CEU HapMap population (MAF = 0.075) or the WTCCC 1958BC control sample used in the discovery stage (MAF = 0.078), whereas the minor allele frequencies were comparable in cases. This marker was successfully genotyped in 109 cases and 265 controls from Croatia. It did not reach statistical significance ($P = 0.36$) in this sample due to the small sample size, but the direction and size of the effect (OR = 1.23; 0.79–1.91 95% CI) were comparable with those in the discovery cohorts, with the minor allele frequency of the risk allele higher in cases (MAF = 0.14) compared with the controls (MAF = 0.11).

Table 2. rs12960119 minor allele (G) effect across the three discovery cohorts

Population	Allele frequency (cases/controls)	OR (95% CI) additive model	OR (95% CI) dominant model	*P*-value (additive/dominant)
Scottish	0.115/0.086	1.42 (1.16–1.75)	1.40 (1.30–1.73)	0.0009/0.003
UK	0.104/0.077	1.41 (1.13–1.76)	1.46 (1.17–1.82)	0.002/0.0008
Dutch	0.117/0.075	1.76 (1.13–2.76)	1.90 (1.17–3.1)	0.013/0.009

rs12960119 allele frequencies, odds ratios and association *P*-values for the minor G allele by discovery population.

Combined analysis

Results for the seven SNPs analyzed through both discovery and replication steps were combined into an overall meta-analysis (Table 3) and this yielded one variant (rs267738) reaching the conventional genome-wide significance threshold and the next most significant SNP (rs2045084) falling just below our initial GWAS threshold (dominant model *P*-value = 2.59×10^{-7}; threshold set at 1.27×10^{-7}). SNP rs267738, within the *CERS2* gene, showed the strongest association signal under a dominant effect model ($P = 2.11 \times 10^{-8}$). The SNP effect size was consistent in both direction and magnitude across all four populations (Fig. 3), with cases showing an excess of the major allele homozygote genotype AA (OR = 0.78; 0.71–0.85 95% CI). The signal in this region was further supported by an additional SNP (rs267733) showing suggestive significance in the discovery phase meta-analysis ($P = 5.29 \times 10^{-5}$). This marker is located within the gene *ANXA9*, which is 5′ of *CERS2*. The next most significant SNP association in the combined meta-analysis (rs2045084) is located 3′ of the *TSTA3* gene, and showed some heterogeneity between studies ($I^2 = 32.8$, $P = 0.22$, Fig. 3). The association was not significant in the UK discovery cohort [$P = 0.55$, OR = 1.09 (0.85–1.40 95% CI)].

The rs12960119 SNP, located in the *SS18* gene, which almost reached genome-wide significance at the discovery phase, did not replicate in the London cohort and did not reach genome-wide significance in the overall meta-analysis ($P = 2.04 \times 10^{-6}$, dominant model; $P = 2.49 \times 10^{-6}$ additive model). Heterogeneity in the overall meta-analysis was high under both models ($I^2 = 50.2$, $P = 0.09$, under a dominant model; $I^2 = 49.6$, $P = 0.09$, under an additive model) and contrasted with

Table 3. Combined discovery and replication phases meta-analysis results

SNP	Additive model					Dominant model				
	OR	L95% CI	U95% CI	*P*-value	Direction	OR	L95% CI	U96%CI	*P*-value	Direction
rs267738	0.81	0.75	0.88	1.43E-07	−−−−?	0.78	0.71	0.85	2.11E-08	−−−−?
rs2045084	1.16	1.09	1.24	2.85E-06	++++?	1.25	1.15	1.36	2.59E-07	++++?
rs12960119	1.26	1.14	1.39	2.49E-06	+++++	1.27	1.15	1.41	2.04E-06	+++++
rs955943	1.41	1.21	1.64	5.45E-06	++++?	1.42	1.21	1.66	1.51E-05	−+++?
rs8132771	1.31	1.16	1.48	1.65E-05	++++?	1.31	1.16	1.49	1.64E-05	++++?
rs913444	1.17	1.09	1.26	2.09E-05	++++?	1.17	1.08	1.29	3.38E-04	++++?
rs1074463	1.15	1.05	1.26	2.59E-03	+++−?	1.21	1.09	1.35	2.90E-04	+++−?

Odds ratios with 95% confidence intervals and direction of effect with respect to the SNP minor allele. Direction of the effect is given for, respectively, the Netherlands, UK, Scotland, London and Croatia studies.

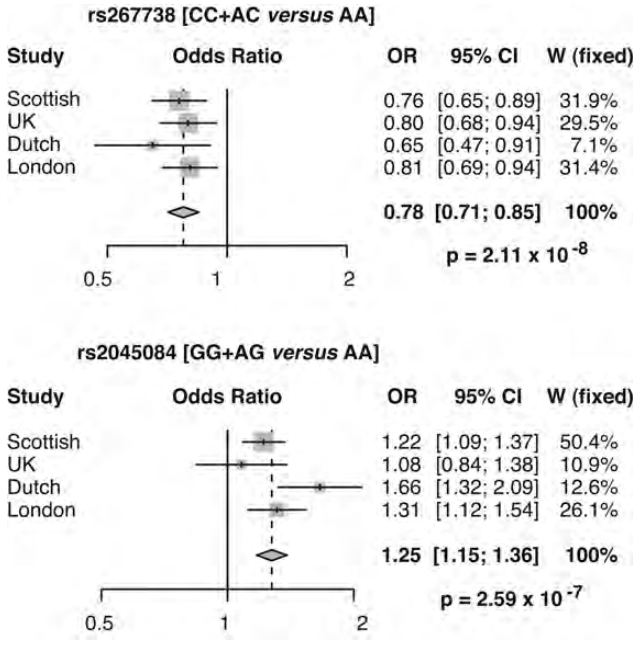


Figure 3. Forest plots for the SNPs with the strongest association signals in the combined analysis under a dominant model for the minor allele effect. Population odds ratio with respect to the minor allele and its 95% CI is represented by a grey box and horizontal line. The bottom rhombus shows the weighted estimate of the odds ratio overall studies.

the homogeneity of association signal in the discovery phase, where three diverse study populations had been used.

Genetic risk score

Each of these markers contributed only a modest amount to individual risk of RRD. To assess the combined effect of our top-ranked markers, we prioritized five of the SNPs that were followed up at the replication phase (see section 'Methods') and combined them in a genetic risk score using a simple allele counting approach. The difference in a mean number of risk alleles between cases and controls was statistically significant in all studies using the Mann–Whitney test (discovery Dutch $P = 2.05 \times 10^{-5}$; discovery Scottish $P = 2.19 \times 10^{-11}$; discovery UK $P = 4.1 \times 10^{-6}$; replication London $P = 2.4 \times 10^{-3}$). Figure 4 illustrates the distribution of the risk alleles in our four-independent studies, displaying, in each study, a higher proportion of cases compared with controls among individuals carrying four or more risk alleles. The odds ratios calculated relative to the median number of risk alleles in the controls for each population show a significantly increased RRD risk (OR >1) for the individuals with five or more risk alleles: with odds ratios of at least two in the discovery cohorts and >1.22 in the London replication cohort (Supplementary Material, Fig. S4).

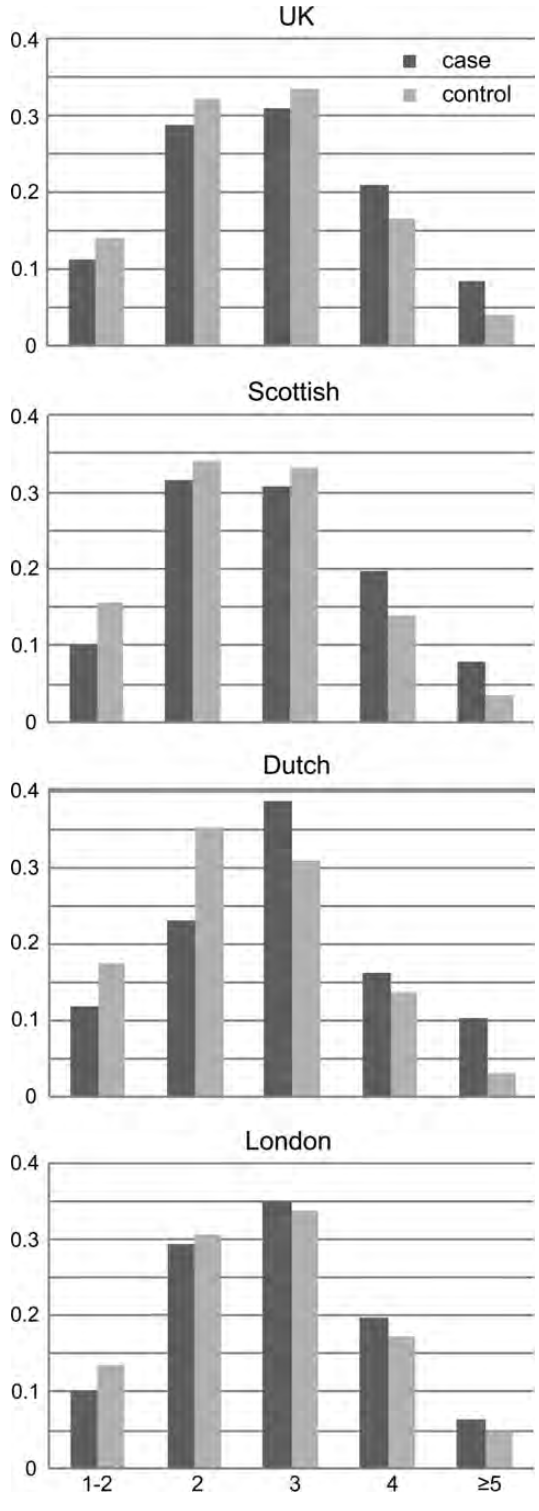


Figure 4. Distribution of the number of identified risk alleles in cases (dark grey) and controls (light grey) in four different populations. The horizontal axis represents the number of RRD risk alleles and the vertical axis represents the percentage of individuals with a given number of risk alleles in cases and controls.

Variation in disease liability explained

Following stringent QC procedures and using the method of Yang *et al*., implemented in the Genome Wide Complex Trait Analysis (GCTA) software, 27.4% of the variance in underlying disease liability could be assigned to the combined effect of the markers tagged by the primary genome-wide scan SNPs. The proportion of the variance explained remained comparable when different minor allele frequencies and missing genotype cut-offs were used (Table 4). The signal was not due to

Table 4. Variation in RRD liability explained by the combined GWAS SNPs using different parameters

Threshold	Number of SNPs	Estimate (SE)	Adjusted estimate (SE)	Liability transformed estimate (SE)
MAF >0.01				
SNP missingness 0.05	257 095	0.3587 (0.1206)	0.4479 (0.1498)	0.275 (0.092)
SNP missingness 0.005	255 704	0.3573 (0.1207)	0.4467 (0.1502)	0.274 (0.092)
MAF >0.05				
SNP missingness 0.05	250 817	0.3701 (0.1192)	0.4612 (0.1480)	0.283 (0.091)
SNP missingness 0.005	249 439	0.3687 (0.1194)	0.4601 (0.1484)	0.282 (0.091)

Estimated proportion of the variance in disease liability on the observed scale (estimate), corrected for imperfect LD between SNPs and causal variants (adjusted estimate) and liability scale (Liability transformed estimate), explained by the consensus panel of SNPs typed on both the Illumina CNV370v3-Quad array (Scottish RRD cases) and Illumina-300 and 240S HumanHap array (SOCCS controls).

population structure within Scottish cases and controls since it was not found when the Scottish Colorectal Cancer Study (SOCCS) control group was artificially split into cases and controls (liability transformed estimate = 0.000001, SE = 0.23). These GCTA estimates further supported the evidence of a modest polygenic contribution by common genetic variants in RRD.

Discussion

This is the first large-scale genome-wide association study ever performed for RRD, a condition which continues to result in significant vision loss. Our results support a polygenic etiology for this condition and identified several common contributory risk variants of small effect.

A genome-wide significant signal of association with RRD was found for the marker rs267738. This is a missense (Glu to Ala) coding SNP located within the *CERS2* gene, in a gene-rich region. The index SNP has been associated with the expression level of *CERS2* in lymphoblastoid cell lines (exon eQTL for *CERS2*) making the regulation of *CERS2* expression a strong functional candidate for the RRD association. *CERS2* encodes the protein CerS2, the most abundantly and ubiquitously expressed member of the ceramide synthase family (22). Ceramide-enriched membrane domains amplify diverse signals and are critically involved in inducing apoptosis (23). *In vitro*, ceramides have been shown to mediate apoptosis of mammalian photoreceptors (24,25) and RPE cell lines (26–28). In a mouse model of retinitis pigmentosa, inhibition of ceramide synthesis slowed disease progression and mutation in the ceramide kinase-like gene *CERKL* has been implicated in a recessive form of retinitis pigmentosa in humans (29). Increased cellular ceramide levels were found in brains of patients with Batten's disease (30) and in retinas of patients with Farber's disease (31). A feature of both diseases is blindness, although this is due to retinal degeneration rather than detachment. An association signal with RRD in this region is supported by an additional association with SNP rs267733 ($P = 5.3 \times 10^{-5}$, dominant model), located 11 kb from *CERS2*. A variant in high LD (rs267734; $r^2 = 1$) with this SNP has previously been reported to be associated with chronic kidney disease (32).

Within or near each of the five suggestive signals of the combined analysis ($2.5 \times 10^{-8} < P$ combined $< 10^{-4}$) was a gene expressed in the retina, several of which with a documented role in cell adhesion or migration (Supplementary Material, Table S3 for genes within suggestive signals). Both TSTA3 and SS18 can modify integrins. Integrins are transmembrane glycoproteins that bind extracellular matrix components to the actin cytoskeleton and activate a signaling cascade leading to cytoskeletal reorganization and cell adhesion. The products of *LDB2* and *TIAM1*, the later close to- not within- a suggestive signal (rs8132771), have a role in cytoskeletal reorganization. *TIAM1* encodes a guanine nucleotide exchange factor which activates a small GTPase of the rho family and modulates changes in cytoskeleton downstream of membrane receptor stimulation. These receptors include TrkB (21), encoded by a gene within another suggestive signal, but also adhesion molecules such as β1 integrins (33,34).

While single marker associations were weak, association with RRD jointly using the five markers showing the most significant and consistent association was significant in all four studies (Scottish discovery, Dutch discovery, UK discovery, London replication) so that the combined genetic score was a potentially useful predictor of disease. Using all the genotyped SNP additive effects from the discovery genome-wide scan jointly, explained 27.4% of the underlying liability to RRD, supporting a moderate polygenic contribution to disease status (solely due to the common variants tested in this scan). It has been suggested that estimating the proportion of variance explained that is due to all SNPs jointly may be inflated due to the confounding effect of population structure (35). However, we believe that this should not be a strong contributory factor in our discovery cohort as the variance explained in the controls alone, randomly labeled as cases and controls, was negligible. The ancestry clusters drawn in RRD cases and Scottish controls were overlapping (Supplementary Material, Fig. S1) and care was taken to remove systematic differences between RRD cases and controls due to the different genotyping platforms used (36).

There were a number of limitations to this study. As we followed up only a subset of the variants tested in the GWAS discovery step, variants which displayed a non-statistically significant association at this first stage may have been missed and yet could have been statistically significant in the larger sample size of the combined data sets (false negatives due to lack of power). Furthermore, there was a substantial failure rate (14%) of the SNPs tested on the customized iSelect Illumina platform in the follow-up discovery step. We tried to minimize the effect of this by (re)genotyping the most significant markers not represented by a proxy SNP; however, true associations could still have been missed. A further limitation was the

use of controls from studies or repositories that were genotyped independently using different genotyping platforms. This can create confounding but true associations should remain consistent across the case–control sets assembled. The controls were unselected with regard to phenotype therefore a number of RRD cases might be expected within the control series given the relatively high prevalence of RRD (0.8%) (15). Since the number of affected RRDs may have fluctuated across control series by chance, this may partly explain the heterogeneity across studies. Furthermore, significant population structure has been reported in two of the WTCCC control cohorts, 1958 BC and NBS (35), that we used in the discovery and replication stages, respectively, making it a possible explanation for the non-replication of the *SS18* locus. NBS controls had a 1.2% higher allele frequency for rs12960119 compared with that in the HapMap CEU population, and a 0.9% higher allele frequency than that in the other control group used in the discovery step (1958 Birth Cohort). The fact that only a subset of SNPs had been typed in the RRD cases used after the stage 1 discovery prevented us from correcting for population structure for the matched RRD cases and controls in the usual manner (by adjusting for ancestry principal components). Heterogeneity of the RRD phenotype is an additional concern as other underlying pathologies usually precede its development, e.g. posterior vitreous detachment, high myopia, lattice degeneration. Larger studies will be useful to tease apart the possible overlap between these pathologies. Nevertheless, results from the published GWAS on myopia (37,38) and lattice degeneration (39) did not show any overlap with our most significant associations and neither the *SS18*, *CERS2* or *TIAM1* signal strengths were affected by removing high myopes from the RRD cases in the primary analysis (Supplementary Material, Table S6) although the *TSTA3*, *TrkB* and *LDB2* associations may be influenced by myopia. This study was not designed to explore the association of rare variants with RRD, and given the fraction of underlying disease liability not accounted for by the common variants tested in this study, the search for rare variants influencing RRD is warranted. Finally, given our sample size, we had limited power to detect and replicate variants of moderate effect sizes.

Despite these limitations, this first genetic association study of RRD supports a polygenic component underlying RRD risk, and identified several common variants of moderate effect contributing to risk. Additional studies will be required to confirm the individual variants highlighted in this study and additional genotyping and increased sample size will be needed to extend this first exploration of the genetic basis of RRD. Since the pathogenesis of RRD is not fully understood, the identification of genes influencing it should make an important contribution to furthering our understanding of the disorder and in time lead to better prevention and treatment.

MATERIALS AND METHODS

Study subjects

All participating studies received approval from their local ethics committees and followed the tenets of the Declaration of Helsinki.

Retinal detachment cases

All Scottish cases [aged 9–91 (mean = 58.91)] were recruited as a part of a population-based epidemiology study of RRD recently carried out in Scotland (18). All incident cases of RRD reported in any of the six vitreoretinal surgical sites (Ayr, Glasgow, Edinburgh, Dundee, Aberdeen and Inverness) were invited to participate in the study. The diagnosis of RRD was based on a case definition of 'a full thickness break in the NSR with a surrounding area of sub-retinal fluid extending greater than 2 disc diameters' (12,40–42). All other types of retinal detachment (exudative, tractional and combined) and re-detachments cases regardless of the duration of attachment post-operatively were excluded. Also excluded were participants with a cataract surgery within 2 years of the detachment diagnostic or with known syndromic features. 6.9% of RRD Scottish subjects evaluated reported an affected first degree relative (15). Care was taken that cases from London, UK $n = 457$ in the discovery phase (aged 9–92, mean = 57.87); $n = 846$ in the replication phase (aged 14–90, mean = 58.58), Cambridge, UK ($n = 173$, aged 15–91, mean = 61.38), Nijmegen, the Netherlands ($n = 252$; aged 12–95, mean = 59.92) and Zagreb, Croatia ($n = 120$, aged 20–83, mean = 59.96) were collected using the same phenotypic inclusion criteria as for the Scottish sample. Positive family history was recorded in 11.3% of cases from the London recruitment, 7.1% of Croatian cases and 5.6% of a subset of 180 Dutch cases had an affected first degree relative (14). Family history information for cases collected in Cambridge was not available.

Controls

Control groups were readily available population-matched samples unselected with regard to RRD;, i.e. unselected controls, a small fraction of which would have RRD. The SOCCS was used as a control group for the Scottish samples in the first stage of the discovery. This is a prospective population-based study in Scotland (1999–2006), with colorectal cancer cases drawn from all over Scotland, matched to healthy controls by age, sex and area of residence (19). One thousand and nine hundred and sixty-seven genotyped individuals were used in the analysis, aged 18–62 (mean = 50.25). Two sets of control groups from the WTCCC were also used in subsequent stages: 2592 samples from the 1958 UK birth cohort (1958BC) was used in stage two of the discovery phase and 2737 individuals from National Blood Service (aged 17–69, mean = 43.5) in the replication phase (43). Controls for the Dutch samples were 320 control individuals from an independent schizophrenia study that were locally recruited, aged 36–74 (mean = 50.22). Finally, Croatian participants to a colorectal cases and controls on-going study were used as a control data set for the Croatian RRD cases (269 individuals; aged 24–92, mean = 66.1).

Genotyping

Genome-wide scan

Nine hundred and twelve Scottish RRD cases were genotyped using the Illumina CNV370v3-Quad array, and compared with population-matched Scottish controls from the SOCCS study previously genotyped on the Illumina-300 and 240S HumanHap arrays. Raw data analysis was done using the Illumina genotype analysis software GenomeStudio with a GenCall scoring cut-off

of 0.15 and solely on the new data itself (not using external references). A final clustering was done using only the 892 samples with an overall genotype call >90%. Further quality control was performed separately for cases and controls using the check marker function of the 'GenABEL' R package. This step removed SNPs with a low call rate (<97%), those in a high Hardy–Weinberg disequilibrium ($P < 1 \times 10^{-6}$), as well as samples identified as duplicates or highly related, mixed (high heterozygosity), those with a low call rate (<97%), sex discrepancies, and outliers based on whole-genome identity-by-state sharing. In total, 870 cases, 1968 controls and 299 789 common SNPs (minor allele frequency >0.02) remained for analysis. When cases and controls were analyzed jointly, three additional samples failed quality control. Deeper relatedness check based on genomic kinship coefficients derived from the whole-genome identity-by-state sharing identified 15 pairs of individuals related to the first-cousins or greater level (π-hat > 0.125 in PLINKv1.07, $\Phi > 0.0625$ in 'GenABEL' (44)). Removing one individual of each related pair left 867 cases and 1953 controls for the final analysis. Plots using principal components derived from the identity-by-state sharing measures showed good overlap of ancestry between the Scottish RRD cases and controls (Supplementary Material, Fig. S1).

Further to these standard QC steps, SNP showing strong signals in association analysis were examined more closely. Several were removed due to poor clustering in either cases or controls or both after visualization of individual clusters (e.g. presence of four rather than three clusters) and two SNPs, rs4862110 and rs4957798, were removed due to discordance between SNP arrays (36). In the final analysis presented here, 299 737 SNPs, genotyped in both cases and controls and QCed, were used.

Imputation of the genotyped SNPs and the subsequent analysis was not performed since different genotyping platforms had been used in cases and controls. It is well documented that in this scenario imputation is tricky and often leads to type I error inflation (45).

iSelect platform

Five thousand and nine hundred and eighty-one SNPs were selected for the second stage of the discovery phase to be assayed on an iSelect Illumina platform, 4706 selected from the genome-wide scan results complemented by 1275 from 18 candidate gene regions (listed in Supplementary Material, Table S1). Tagging SNPs in the candidate gene regions were selected with Tagger program within the Haploview v4.0 software using Centre d'Etude du Polymorphisme Humain from Utah (CEU) population of the HapMap project (International HapMap consortium) (46). A threshold for r^2 was set at 0.8. Region of 100 kb upstream and downstream from each gene was used in the selection process. Eight hundred and thirteen assays failed design, including 183 of the candidate genes set (in agreement with failure rate to be expected from the Illumina Infinium technology). The list of the candidate genes tagging SNPs that were tested are listed Supplementary Material, Table S1. Clustering was done using Illumina Genome Studio software and QC steps (call rate, HWE, sex discrepancies) were carried as described for the genome scan analysis. Difference in genotype missingness rates between cases and control was also tested separately for each new study using the function—test-missing in PLINKv1.07 (47). No SNPs displayed significant missingness difference ($P < 0.00001$) between cases and controls. To ensure that signals from all of the SNPs with best ranks in the discovery stage 1 were taken forward in stage 2 (i.e. among the best 200 of any stage 1 analysis), we investigated the representativity by proxy (LD $r^2 > 0.2$) of those which failed on the iSelect chip or QC. Eight SNPs with initial P-values $< 10^{-4}$ or in possible functional gene, *FRMD4A* and *GRM7* were thus additionally typed using the Taqman technology. The final number of SNPs carried on for the second stage of the discovery phase was 4347.

Single SNP typing

Seven SNPs in total were genotyped in the replication phase using the Taqman technology. The six top associated SNPs, which displayed consistent direction of effect across all cohorts and one additional SNP (rs913444) based on the pathway analysis results. Quality control tests were carried out as described above. Call rates for SNPs genotyped using the TaqMan technology were set to 90%. The highest in the London replication samples was 99.8% (rs8132771) and the lowest 93.5% (rs955943). Only one marker (rs12960119; call rate = 93.6%) passed this threshold in the Croatian samples.

Statistical analysis

Association testing

The genome-wide scan was performed using a regression model with age and sex as covariates and fitting an additive SNP effect, using the egscore function of the 'GenABEL' R package (44) which takes into account possible population stratification (48). QQ plots of observed versus expected ordered test statistics were drawn using the 'snpMatrix' R package (49). The grey area represents the area between the 2.5 and 97.5% probability bounds of non-departure to the null hypothesis. Five analyses were performed, all RRD Scottish cases versus all SOCCS Scottish controls, RRD cases with known high myopes (spherical equivalent refraction ≤-5) removed, SOCCS controls with cancer cases removed and gender separate analysis (QQ plots presented in Fig. 2 and Supplementary Material, Fig. S2). The top ranking associated SNPs from each analysis (P-value < 10^{-3}) were listed to be taken forward in the second stage of the discovery.

In the second stage of the discovery phase, association between RRD with SNP genotypes was tested using logistic regression including sex and age as covariates and using the PLINKv1.07 software (47). Both additive and dominant models were tested. Association analyses were done for each cohort separately and the results from all cohorts (with the Scottish discovery step 1 rerun using the PLINKv1.07 implementation of the logistic regression adjusting for age sex and three ancestry principal components) combined in an inverse variance-weighted fixed effect meta-analysis using METAL (50). We calculated empirical significance levels for the association in all of the cohorts by permutation as implemented in PLINKv1.07, and used these empirical significance levels for further meta-analysis.

Given the low sample size of our replication cohorts and that inclusion of covariates in case–control association studies has been shown to reduce their power in this scenario and with low

disease prevalence, we carried out analysis in the replication phase without the inclusion of covariates. Results from both discovery stages and the replication phase were also combined in a meta-analysis (unadjusted for covariates). Forrest plots and heterogeneity measures for the most significant SNPs were obtained using the 'meta' R package (51).

P-value threshold

To correct for multiple testing in the initial genome-wide scan, we used a multiple testing correction method that takes into account the correlation structure due to linkage disequilibrium (LD) among the GWAS genotyped SNPs tested, using the 'SimpleM' R library (52). Imputation of the missing genotypes was carried out using k Nearest Neighbors using the 'scrime' R package (53). The calculated effective number of independent SNPs, $n = 197\,628$, was then used to correct for multiple testing using the standard Bonferroni procedure. Discovery stage significance threshold was additionally corrected for the testing of two genetic models (additive and dominant). The genome-wide significance threshold was therefore set at 1.27×10^{-7} ($=0.05/(197\,628 \times 2)$). Despite following up only a subset of SNPs used in the first step of the discovery phase into the second step, the same genome wide significance threshold was used for the meta-analysis of the two steps to be conservative. The overall meta-analysis (combining discovery stage and replication stage data) used the conventional 5×10^{-8} significance threshold, additionally corrected for the use of two genetic models, hence set at 2.5×10^{-8}.

Genetic risk score

The multilocus genetic risk score was calculated for each individual using five risk variants (the most significant homogeneous results from the discovery phase). A simple allele count was used to calculate the genetic risk score using PLINK's v1.07 'score' option. Individuals with one or more missing genotypes were not included in the analysis. Odds ratios were calculated relative to the median number of risk alleles among controls using SPSS v16.0 (54) (subjects with five or more risk alleles were collapsed together due to the low number of individuals with high-risk allele numbers).

Genetic variance of disease liability explained

The GCTA software (55) was used to estimate the proportion of variation in disease liability captured by the genome-wide SNP array (discovery stage 1). The combined-SNP effect was fitted as a random effect in a mixed linear model predicting the transformed trait (probit transformation) and its variance estimated by restricted maximum likelihood. This estimation is then corrected for the incomplete LD between the tag SNP and the causal variant and for the ascertainment of cases (compared with the disease prevalence in general population). Stringent QC procedures (PLINK v1.07) (47) were implemented on the genotyped data to avoid experimental biases causing false case–control differences in relatedness (56). Closely related individuals were removed to make sure only distant relationships were taken into account (kinship coefficient threshold was set at 0.05 and implemented within the GCTA software). As a population structure control, the control cohort (SOCCS without colorectal cases) was split into two groups and the procedure repeated in order to check that the detected explained variance there was negligible.

Pathway analysis

The top six SNPs from the discovery phase were selected in a first pathway analysis. Identification of all possible genes underlying the SNPs' association signals was done using the 'seed' output from the GRAIL (20) (Gene Relationships Among Implicated Loci) software. All genes located in the interval were considered in the subsequent pathway analysis using the IPA software (Ingenuity® Systems, www.ingenuity.com). IPA uses an internal manually maintained database to identify networks that maximize connectivity. When running the analysis, endogenous chemicals were excluded from the search and only the direct interactions were taken into account. An extended analysis, including all SNPs from the discovery phase with associated *P*-value $< 10^{-4}$ (23 SNPs in total, Table 1), was also performed using the same parameters.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank all subjects enrolled for their participation. We are indebted to R. Ophoff for providing the Dutch control data and the WTCCC for the access to the UK control data. DNA extraction and genotyping of the Scottish cases was performed at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health.

Conflict of Interest statement. None declared.

FUNDING

The Scottish RRD cases collection and genotyping were supported by a grant from the W.H. Ross Foundation for the Prevention of Blindness, The Royal College of Surgeons Edinburgh/Royal Blind Asylum/Scottish War Blinded, and the Chief Scientist Office Scotland (grant number CZB/4/705). The SOCCS study (Scottish controls) was funded by grants from Cancer Research UK (grant numbers C348/A3758 and A8896, C48/A6361), Medical Research Council (grant number G0000657-53203) and Scottish Executive Chief Scientist's Office (grant numbers K/OPR/2/2/D333, CZB/4/449). The London cases collection was supported by Fight for Sight UK, the Special Trustee of Moorfields Eye Hospital and the Department of Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology for a Specialist Biomedical Research Centre for Ophthalmology. The overall genetic analysis was supported by the Medical Research Council, the University of Edinburgh (College of

Medicine and Veterinary Medicine) and an Edinburgh Global Research Scholarship to M.K.

REFERENCES

1. Munier, A., Gunning, T., Kenny, D. and O'Keefe, M. (1998) Causes of blindness in the adult population of the Republic of Ireland. *Br. J. Ophthalmol.*, **82**, 630–633.
2. Steinberg, R.H. and Wood, I. (1974) Pigment epithelial cell ensheathment of cone outer segments in the retina of the domestic cat. *Proc. R. Soc. Lond. B Biol. Sci.*, **187**, 461–478.
3. Royal College of Ophthalmologists (2010) *Management of Retinal Detachment*.
4. Brinton, D.A. and Wilkinson, C.P. (2009) *Retinal Detachment: Principles and Practice*. Oxford University Press, Oxford, UK.
5. Wong, T.Y., Tielsch, J.M. and Schein, O.D. (1999) Racial difference in the incidence of retinal detachment in Singapore. *Arch. Ophthalmol.*, **117**, 379–383.
6. Mowatt, L., Adrien, G. and Price, N. (2003) Ethnic differences in the demand incidence of retinal detachments in two districts in the West Midlands. *Eye*, **17**, 63–70.
7. Peters, A.L. (1995) Retinal detachment in black South Africans. *S. Afric. Med. J.*, **85**, 158–159.
8. Haimann, M.H., Burton, T.C. and Brown, C.K. (1982) Epidemiology of retinal detachment. *Arch. Ophthalmol.*, **100**, 289–292.
9. Sasaki, K., Ideta, H., Yonemoto, J., Tanaka, S., Hirose, A. and Oka, C. (1995) Epidemiologic characteristics of rhegmatogenous retinal detachment in Kumamoto, Japan. Graefe's Arch. Clin. Exp. Ophthalmol., **233**, 772–776.
10. Algvere, P.V., Jahnberg, P. and Textorius, O. (1999) The Swedish Retinal Detachment Register. I. A database for epidemiological and clinical studies. Graefe's Arch. Clin. Exp. Ophthalmol., **237**, 137–144.
11. Ivanisevic, M., Bojic, L. and Eterovic, D. (2000) Epidemiological study of nontraumatic phakic rhegmatogenous retinal detachment. *Ophthalm. Res.*, **32**, 237–239.
12. Polkinghorne, P.J. and Craig, J.P. (2004) Northern New Zealand Rhegmatogenous Retinal Detachment Study: epidemiology and risk factors. *Clin. Exp. Ophthalmol.*, **32**, 159–163.
13. Rosman, M., Wong, T.Y., Ong, S.G. and Ang, C.L. (2001) Retinal detachment in Chinese, Malay and Indian residents in Singapore: a comparative study on risk factors, clinical presentation and surgical outcomes. *Int. Ophthalmol.*, **24**, 101–106.
14. Go, S.L., Hoyng, C.B. and Klaver, C.C. (2005) Genetic risk of rhegmatogenous retinal detachment: a familial aggregation study. *Arch. Ophthalmol.*, **123**, 1237–1241.
15. Mitry, D., Williams, L., Charteris, D.G., Fleck, B.W., Wright, A.F. and Campbell, H. (2011) Population-based estimate of the sibling recurrence risk ratio for rhegmatogenous retinal detachment. *Invest. Ophthalmol. Vis. Sci.*, **52**, 2551–2555.
16. Richards, A.J., Meredith, S., Poulson, A., Bearcroft, P., Crossland, G., Baguley, D.M., Scott, J.D. and Snead, M.P. (2005) A novel mutation of COL2A1 resulting in dominantly inherited rhegmatogenous retinal detachment. *Invest. Ophthalmol. Vis. Sci.*, **46**, 663–668.
17. Go, S.L., Maugeri, A., Mulder, J.J., van Driel, M.A., Cremers, F.P. and Hoyng, C.B. (2003) Autosomal dominant rhegmatogenous retinal detachment associated with an Arg453Ter mutation in the COL2A1 gene. *Invest. Ophthalmol. Vis. Sci.*, **44**, 4035–4043.
18. Mitry, D., Charteris, D.G., Yorston, D., Siddiqui, M.A., Campbell, H., Murphy, A.L., Fleck, B.W., Wright, A.F. and Singh, J. (2010) The epidemiology and socioeconomic associations of retinal detachment in Scotland: a two-year prospective population-based study. *Invest. Ophthalmol. Vis. Sci.*, **51**, 4963–4968.
19. Tenesa, A., Farrington, S.M., Prendergast, J.G., Porteous, M.E., Walker, M., Haq, N., Barnetson, R.A., Theodoratou, E., Cetnarskyj, R., Cartwright, N. *et al.* (2008) Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.*, **40**, 631–637.
20. Raychaudhuri, S., Plenge, R.M., Rossin, E.J., Ng, A.C.Y., Purcell, S.M., Sklar, P., Scolnick, E.M., Xavier, R.J., Altshuler, D., Daly, M.J. *et al.* (2009) Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.*, **5**, e1000534.
21. Miyamoto, Y., Yamauchi, J., Tanoue, A., Wu, C. and Mobley, W.C. (2006) TrkB binds and tyrosine-phosphorylates Tiam1, leading to activation of Rac1 and induction of changes in cellular morphology. *Proc. Natl. Acad. Sci. USA*, **103**, 10444–10449.
22. Laviad, E.L., Albee, L., Pankova-Kholmyansky, I., Epstein, S., Park, H., Merrill, A.H. and Futerman, A.H. (2008) Characterization of ceramide synthase 2. *J. Biol. Chem.*, **283**, 5677–5684.
23. Schenck, M., Carpinteiro, A., Grassme, H., Lang, F. and Gulbins, E. (2007) Ceramide: physiological and pathophysiological aspects. *Arch. Biochem. Biophys.*, **462**, 171–175.
24. Sanvicens, N. and Cotter, T.G. (2006) Ceramide is the key mediator of oxidative stress-induced apoptosis in retinal photoreceptor cells. *J. Neurochem.*, **98**, 1432–1444.
25. German, O.L., Miranda, G.E., Abrahan, C.E. and Rotstein, N.P. (2006) Ceramide is a mediator of apoptosis in retina photoreceptors. *Invest. Ophthalmol. Vis. Sci.*, **47**, 1658–1668.
26. Kannan, R., Jin, M., Gamulescu, M.A. and Hinton, D.R. (2004) Ceramide-induced apoptosis: role of catalase and hepatocyte growth factor. *Free Radic. Biol. Med.*, **37**, 166–175.
27. Sreekumar, P.G., Ding, Y., Ryan, S.J., Kannan, R. and Hinton, D.R. (2009) Regulation of thioredoxin by ceramide in retinal pigment epithelial cells. *Exp. Eye Res.*, **88**, 410–417.
28. Tomita, H., Abe, T. and Tamai, M. (2000) Ceramide-induced cell death in cultured rat retinal pigment epithelial cells. *Tohoku J. Exp. Med.*, **190**, 223–229.
29. Tuson, M., Marfany, G. and Gonzalez-Duarte, R. (2004) Mutation of CERKL, a novel human ceramide kinase gene, causes autosomal recessive retinitis pigmentosa (RP26). *Am. J. Hum. Genet.*, **74**, 128–138.
30. Puranam, K., Qian, W.H., Nikbakht, K., Venable, M., Obeid, L., Hannun, Y. and Boustany, R.M. (1997) Upregulation of Bcl-2 and elevation of ceramide in Batten disease. *Neuropediat.*, **28**, 37–41.
31. Zarbin, M.A., Green, W., Moser, A.B. and Tiffany, C. (1988) Increased levels of ceramide in the retina of a patient with Farber's disease. *Arch. Ophthalmol.*, **106**, 1163–1163.
32. Kottgen, A., Pattaro, C., Boger, C.A., Fuchsberger, C., Olden, M., Glazer, N.L., Parsa, A., Gao, X., Yang, Q., Smith, A.V. *et al.* (2010) New loci associated with kidney function and chronic kidney disease. *Nat. Genet.*, **42**, 376–384.
33. Hamelers, I.H., Olivo, C., Mertens, A.E., Pegtel, D.M., van der Kammen, R.A., Sonnenberg, A. and Collard, J.G. (2005) The Rac activator Tiam1 is required for (alpha)3(beta)1-mediated laminin-5 deposition, cell spreading, and cell migration. *J. Cell Biol.*, **171**, 871–881.
34. O'Toole, T.E., Bialkowska, K., Li, X. and Fox, J.E. (2011) Tiam1 is recruited to beta1-integrin complexes by 14–3-3zeta where it mediates integrin-induced Rac1 activation and motility. *J. Cell Physiol.*, **226**, 2965–2978.
35. Browning, S.R. and Browning, B.L. (2011) Population structure can inflate SNP-based heritability estimates. *Am. J. Hum. Genet.*, **89**, 191–193. author reply 193–195.
36. Mitry, D., Campbell, H., Charteris, D.G., Fleck, B.W., Tenesa, A., Dunlop, M.G., Hayward, C., Wright, A.F. and Vitart, V. (2011) SNP mistyping in genotyping arrays—an important cause of spurious association in case-control studies. *Genet. Epidemiol.*, **35**, 423–426.
37. Kiefer, A.K., Tung, J.Y., Do, C.B., Hinds, D.A., Mountain, J.L., Francke, U. and Eriksson, N. (2013) Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PloS Genet*, **9**, e1003299
38. Verhoeven, V.J., Hysi, P.G., Wojciechowski, R., Fan, Q., Guggenheim, J.A., Hohn, R., Macgregor, S., Hewitt, A.W., Nag, A., Cheng, C.Y. *et al.* (2013) Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat. Genet.*, **45**, 314–318.
39. Meguro, A., Ideta, H., Ota, M., Ito, N., Ideta, R., Yonemoto, J., Takeuchi, M., Uemoto, R., Nishide, T., Iijima, Y. *et al.* (2012) Common variants in the COL4A4 gene confer susceptibility to lattice degeneration of the retina. *PLoS ONE*, **7**, e39300.
40. Zou, H., Zhang, X., Xu, X., Wang, X., Liu, K. and Ho, P.C. (2002) Epidemiology survey of rhegmatogenous retinal detachment in Beixinjing District, Shanghai, China. *Retina*, **22**, 294–299.
41. Li, X. (2003) Incidence and epidemiological characteristics of rhegmatogenous retinal detachment in Beijing, China. *Ophthalmology*, **110**, 2413–2417.
42. Schepens, C.L. (1951) REtinal detachment and aphakia. *Arch. Ophthalmol.*, **45**, 1–7.

43. Welcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
44. Aulchenko, Y.S., Ripke, S., Isaacs, A. and van Duijn, C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, **23**, 1294–1296.
45. Sinnott, J.A. and Kraft, P. (2012) Artifact due to differential error when cases and controls are imputed from different platforms. *Hum. Genet.*, **131**, 111–119.
46. The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature*, **437**, 1299–1320.
47. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
48. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
49. Clayton, D. and Leung, H.T. (2007) An R package for analysis of whole-genome association studies. *Hum. Hered.*, **64**, 45–51.
50. Willer, C.J., Li, Y. and Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, **26**, 2190–2191.
51. Schwarzer, G. (2012) meta: Meta-Analysis with R. R package version 2.1–2.
52. Gao, X., Starmer, J. and Martin, E.R. (2008) A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.*, **32**, 361–369.
53. Schwender, H. and Fritsch, A. (2012) scrime: Analysis of High-Dimensional Categorical Data such as SNP Data. R package version 1.2.9.
54. SPSS Inc. (2007), In *Version 16 Chicago*. SPSS Inc.
55. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex tait analysis. *Am. J. Hum. Genet.*, **88**, 76–82.
56. Lee, S.H., Wray, N.R., Goddard, M.E. and Visscher, P.M. (2011) Estimating Missing Heritability for Disease from Genome-wide Association Studies. *Am. J. Hum. Genet.*, **88**, 294–305.