

This thesis is submitted for the degree of Doctor of Philosophy

Understanding the genetics and function of complex human retinal phenotypes

Hannah Currant

November 2020

University of Cambridge
Murray Edwards College
European Bioinformatics Institute

Declaration of Originality

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee

Summary

Understanding the genetics and function of complex human retinal phenotypes

Hannah Currant

The human retina is the tissue at the back of the eye responsible for converting light stimulus into neuronal signal that can be interpreted by the brain. To perform this integral role within the central nervous system, the retina has a complex and layered structure, with each layer performing a vital step in the signal transformation process.

Changes in the morphology of this structure are often a consequence of disease, which can affect the function of the eye. Better understanding of the genetics influencing this structure may teach us about the biological processes underlying these diseases as well as general eye development.

The retina is imaged routinely in the clinic using optical coherence tomography (OCT), providing a non-invasive imaging technique that produces high-resolution, three-dimensional representations of the retina from which measures describing retinal morphology can be extracted.

This thesis summarises my research into the genetic variation underlying retinal morphology.

Firstly, I explored the morphology of the inner retina, whose thickness is used as a biomarker of glaucoma, using quantitative phenotypes extracted from OCT. I conducted genome-wide association studies (GWAS) of the thickness of the retinal nerve fibre layer and the ganglion cell inner plexiform layer to understand the genetic variation driving inner retinal morphology. I further explored the causal relationship between the inner retina and glaucoma using Mendelian randomisation analysis.

I next performed GWAS of the thickness of the outer retinal layers, including both the component photoreceptor cell layers (the outer nuclear layer, inner segment, and outer segment), and the retinal pigment epithelium layer. I explored how genetic variation was affecting the outer retinal morphology at a higher dimension by looking for genetic

variants that were differentially affecting the outer retinal thickness at the central macula compared to the peripheral macula.

To further explore the rich dimensionality of OCT data, I developed several image analysis techniques to gain more granular information about the morphological variation being affected by the discovered genetic variants. In doing so I established a novel population level trait and examined its effect on visual acuity.

In summary, this thesis provides a well-rounded and detailed look into the genetic variation underlying morphological variation of the retinal layers. As the largest study of retinal layer genetics of its kind, it offers insight into clinical ophthalmology and retinal development, and furthermore opens new avenues for clinical research.

Acknowledgements

I would like to take the time to thank the wonderful group of people that have supported me throughout my career so far, particularly during the last four years whilst I complete my PhD.

Firstly, I would like to thank my supervisor, Ewan Birney, for the opportunity to work in his group and the support that he has provided throughout. He is consistently enthusiastic, encouraging and generous as a supervisor and mentor and has helped shape me into the scientist I am today. I would also like to thank the members of the Birney research group, past and present, that have made for a friendly and enriching working environment: Hannah Meyer, Tomas Fitzgerald, Carl Barton, Maria Herrero-Zazo, Anat Melamed, Ian Brettell, Victoria Keevil, Adrien Leger and Jack Monahan. Each of you have never hesitated to offer me academic advice and a chat over a cup of tea. The wider support from Debbie Howe, Christina Kirikades and the rest of the Director's office has been always helpful throughout my time as a graduate student.

My research throughout the PhD was enabled and enriched through collaboration with amazing clinicians who were generous with their time and knowledge of ophthalmology. Particularly Anthony Khawaja, Praveen Patel and Paul Foster have each taught me lots about the clinical space, provided new insight into my research and have always shown excitement for my work. I am also grateful for the other members of my scientific support network. My thesis advisory committee members, Judith Zaugg, Oliver Stegle and Ed Bullmore, have all offered advice that has helped me improve my research. I would also like to express my gratitude to Graeme Black and John Danesh for the time they invested in examining my thesis, for giving thoughtful and considerate feedback, and for an enjoyable discussion about my work.

I am very grateful for the care and thought of those who took time to proofread my thesis: Anthony Khawaja, Praveen Patel, Susan McGill, Hannah Meyer, Tomas Fitzgerald, Ian Brettell, Ewan Birney and Cameron Lockhart.

I would like to thank hot water bottles for being excellent writing company, Taylor Swift for her perfectly timed release of the ideal writing album, and chocolate fingers, for sustaining me throughout this process.

I am forever grateful to the community of people that throughout this PhD have provided me with encouragement, moral support and friendship. Thank you to all of my fellow EBI and Sanger PhD students, your friendship has made this PhD a fun and memorable experience. Particular thanks to Anna, Elsa, Sushmita and Harald, for afternoon walks and shared baked goods. Thank you to Hannah and Maria, who have been, and continue to be, amazing colleagues, mentors and friends. Thank you to the people I have met during my time in Cambridge and the support you have given me: My CoderDojo family, for reminding me why I love to code. Murray Edwards College, for giving me support when I needed it and for a welcoming community. Elizabeth, for chats over tea about books and plants. To my fellow St Andrews transplants whose continued friendship is a joy. To Léonie for welcoming me to Heidelberg and continued London adventures. And to Harry for pastries at the allotment and much needed breaks from thesis writing. To Ellie, Susan and Amy, thank you being amazing cheerleaders, indulging in rom-coms with me and all of your friendship. Thank you Courtney for all of your support and the best care packages. Thank you to Charlotte, Natalie, Conner, Nick, Jo and Tom for supporting me from near and far. Thank you to the Lockharts for the encouragement and kind words.

Finally thank you to my family for always being supportive of my work, even if they don't entirely understand it, for cheering me on in all my pursuits and for giving me the loving support to help me through any challenges. And thank you to Cameron, for helping me be my best self, for unlimited kindness and cups of tea, and for always keeping me smiling.

Contents

List of Figures

List of Tables

List of Abbreviations

3D	three-dimensional. 140, 142, 143
ALSPAC	Avon Longitudinal Study of Parents and Children. 55
AMD	age-related macular degeneration. 29, 33, 35, 50, 60, 97, 99, 102, 104, 105, 109, 111, 121, 125, 126, 150, 152, 156
ANOVA	Analysis of Variance. 42
BiLEVE	UK Biobank Lung Exome Variant Evaluation. 62
BM	Bruch's Membrane. 28
BMI	body mass index. 50, 139
CNS	central nervous system. 54, 104, 109, 118, 126, 137
COJO	Conditional Joint Analysis. 81, 101, 114
DNA	Deoxyribonucleic Acid. 45–47
ETDRS	Early Treatment Diabetic Retinopathy Study. 58, 60, 61, 69, 99, 101, 113, 126, 128, 129, 131, 134
GA4GH	Global Alliance for Genomic Healthcare. 37
GCC	Ganglion Cell Complex. 67
GCIPL	Ganglion Cell Inner Plexiform Layer. 58, 67, 69, 72, 74, 76, 78, 81, 90, 95, 96, 129, 130, 134, 135, 155
GCL	Ganglion Cell Layer. 28, 58
GIANT	Genetic Investigation of Anthropometric Traits. 50

GWAS genome-wide association studies. 21, 33, 49–51, 54, 67, 69, 70, 72, 74, 76, 78, 81, 85, 88, 90, 95–97, 99, 101, 102, 104, 105, 109, 111, 113–115, 125, 127–131, 134, 135, 144, 148, 155, 156

HMGCR 3-hydroxy-3-methylglutaryl–coenzyme A reductase. 53

IAMDGC International AMD Genomic Consortium. 105, 109
ICD International Classification of Diseases. 57
IGGC International Glaucoma Genetics Consortium. 33, 85, 88
ILM inner limiting membrane. 60
INL Inner Nuclear Layer. 28, 58
IOP Intraocular Pressure. 9, 32, 58, 61, 85, 87, 88, 90, 96
IPL Inner Plexiform Layer. 28, 58
IQR Interquartile Range. 61
IS inner segment. 28, 58, 97, 99, 101, 102, 109, 114, 118, 125, 129, 131, 134, 135, 137, 156

laser light amplification by stimulated emission of radiation. 30
LD Linkage disequilibrium. 51, 63, 74, 78, 81, 88, 101, 111, 114, 148
LDL-C low density lipoprotein cholesterol. 53
LDSC Linkage Disequilibrium Score. 51, 81, 101–103
LOD Logarithm of Odds. 46
logMAR logarithm of the minimum angle of resolution. 58

MAF Minor Allele Frequency. 51, 63
MR Mendelian Randomisation. 52, 53, 88, 96, 109, 111, 126, 155
MRI Magnetic Resonance Imaging. 127
MTAG Multi-Trait Analysis of GWAS. 81, 101

NHS National Health Service. 56

OCA Oculocutaneous albinism. 34, 85, 96, 107, 109, 124, 144, 147, 152, 157

OCT Optical coherence tomography. 29, 30, 54, 58, 60, 63, 67, 69, 70, 96, 99, 127, 129, 134, 135, 140, 150, 152, 155–158

ONL outer nuclear layer. 28, 58, 97, 99, 101, 102, 109, 114–116, 125, 129, 131, 134, 135, 137, 156

ONT Oxford Nanopore Technologies. 47

OPL Outer Plexiform Layer. 28

OS outer segment. 28, 58, 97, 99, 101, 102, 109, 114, 121, 122, 124, 125, 129, 131, 134, 135, 137, 156

PC principal component. 61, 63, 64, 72, 76, 78, 95, 128–131, 133, 134, 144, 147, 148, 152

PCA principal component analysis. 40, 63, 69, 76, 95, 129, 131, 133, 152

POAG Primary Open Angle Glaucoma. 9, 32, 33, 67, 70, 85, 88, 90, 96, 111

PRC photoreceptor cells. 28, 97, 99, 101, 102, 104, 105, 109, 115, 118, 124, 125, 131, 134, 135, 137, 144, 148, 152, 156

PRL photoreceptor layer. 25, 148

qq quantile-quantile. 51

RFLPs restriction fragment length polymorphisms. 46

RNFL Retinal Nerve Fibre Layer. 28, 58, 60, 67, 69, 72, 74, 76, 78, 81, 85, 90, 95, 96, 129, 130, 134, 135, 155

RPE retinal pigment epithelium. 28, 33, 58, 97, 99, 101, 102, 107, 109, 111, 114, 121, 122, 124–126, 129, 131, 134, 139, 144, 150, 156

SDOCT Spectral Domain Optical Coherence Tomography. 58

SNP single nucleotide polymorphisms. 49–51, 53, 62–64, 74, 78, 85–88, 90, 95, 96, 102, 105, 109, 111, 114, 115, 124, 127, 128, 134, 135, 139, 143, 144, 147, 148, 152, 155

TABS	Topcon Advanced Boundary Segmentation. 58, 69, 99, 128, 140
tRNA	transfer ribonucleic acid. 45
UK	United Kingdom. 56
USA	United States of America. 63, 64
WES	whole exome sequencing. 62
WGS	whole genome sequencing. 62

1. Introduction

Academics and scholars have been studying the human eye and the mechanism of sight for millennia. The eye is a key part of the central nervous system, and its dysfunction leads to notable changes in quality of life [Moschos, 2014]. The eye has a complex structure and the anatomy of the eye has therefore been of great scientific interest, with disruption to the normal morphology often leading to pathology. In particular, the retina, the ocular tissue responsible for conversion of light stimuli into neurological signal that is interpreted by the brain, has an intricate structure that is acutely specialised to its function and disruptions can lead to a diverse catalogue of diseases. Advances in imaging technologies have enabled non-invasive imaging of live retinal tissue. This has allowed for a deeper understanding of retinal structure and pathology as well as aiding diagnosis and disease management. This advancement of imaging and its use in the clinical field, coupled with the increased use of genomics, has helped elucidate the underlying genetic cause of ocular diseases and their link to morphology. New large scale databases that contain optical imaging and genetic data are also allowing for study of normal variation in retinal morphology and their consequences on health and disease. There is great potential for the further integration of genomics into ophthalmic care [Black & al., 2020], and an understanding of the genetic variation underlying normal morphological variation may aid this.

In this thesis I used the data-rich, large-scale UK Biobank database, which contains both retinal imaging and genotype data, to study the genetic variation underlying morphological variation of the retina. Throughout this thesis the term morphology is used to encompass a measure of thickness and topography of retinal layers. I performed the first large-scale genome-wide association studies (GWAS) of morphology of individual retinal layers. I looked at the relationship between the genetic drivers of retinal morphology and a number of retinal diseases. I also developed image processing methodologies in order to study the spatial variation in retinal morphology across the field of the retina.

For the remainder of this chapter, I will describe some of the key aspects of ocular and retinal morphology. I will briefly describe the history of the study of the eye. I will further explain some of the methods for imaging the retina and some of the diseases that

can affect the retina. In the second part of the introduction, I will describe the journey to our current understanding and utility of genetics that has enabled the study of human variation using quantitative genetics.

1.1. The Eye

Human curiosity with the eye, both its function and dysfunction, can be documented back to ancient history. The eye has long held a place in mythology and religion alongside medicine and biology. Consider the Egyptian eye of Horus, a symbol of wealth and prosperity; the third eye, a symbol seen in both Buddhism and Hinduism associated with spiritual insight and awareness; or the evil eye in Greek legend, a curse cast by an evil stare that can be protected against using the iconic blue eye-shaped talisman [Potts, 1982; Abbasi, 2017]. Through history, the eyes were seen to have great power, providing spiritual insight and knowledge as well as the less philosophical power of physical sight. As academics began to study the anatomy, function and dysfunction of the eye further, it became clear that this organ could provide scientific insight into many aspects of biology, development and health.

1.2. History of Ocular Study

Study and speculation of the mechanism of sight and the biology of the eye can be dated back to ancient Egyptian times. Both the ancient Egyptians and Babylonians recorded descriptions of eye diseases, most of which were thought to be caused by demons and thus cured by spiritual practises and application of various ointments [Finger, 1994]. In ~1500BC, Suśruta, an Indian physician, provided further descriptions of eye anatomy and diseases including glaucoma. He is also one of the earliest known physicians to have performed a form of cataract surgery known as couching, in which he physically pushed the cataract out of the field of vision [Dutt, 1938], a technique copied by many visiting Greek scholars.

Various contributions were made by the Ancient Greeks to both eye anatomy and theories of visual mechanism. In terms of anatomical advancements, Alcmaeon is thought to be the first to identify the optic nerves via dissection in the 5th century BC. Hippocrates and his disciples split the eye into three areas, the sclera, the uveal tunic and the retina, a model furthered by Herophilus through the inclusion of the vitreous humor and the optic nerves. Proposed theories of visual mechanism included that of Alcmaeon who hypothesised that beams of visual fire emanated from the eye and captured the im-

age of objects. This concept, of the eye emitting something that allowed for sight, was reiterated by many and given the name *extramission* [Finger, 1994].

The Roman scholar Galen was also a supporter of extramission believing that pneuma, a spirit, was released from the eye through hollow optic nerves to capture images. Despite his incorrect beliefs on the mechanism of vision, Galen did provide detailed descriptions of eye anatomy, outlining the cornea, iris, lens, choroid, sclera and humors [Ajita, 2015]. Galen and the majority of his contemporaries believed the lens to be the structure within the eye responsible for vision, with the retina providing only nutrition to the eye. He also noted that whilst the optic nerves of the two eyes appeared to meet at the chiasma, he believed they did not cross but merely communicated with one another to aid binocular vision.

Arabic academics and physicians are often credited with many important ophthalmological advancements. Hunayn ibn Is-Hâq, a physician born in 809AD in Iraq, is thought to have written the earliest Middle Eastern text about the eye, and went on to be a prolific writer on ophthalmology. Most famously he wrote *The Book of Hunayn ibn Is-Hâq on the Structures of the Eye, Its Diseases, and Their Treatment, According to the Teachings of Hippocrates and Galen, in Ten Treatises* [1928], in which the earliest anatomical diagrams of the eye can be found.

One of the most famous Arab scholars was Hasan Ibn al-Haytham, born in Iraq in the early 11th century. Ibn al-Haytham is one of the first scholars to show understanding of the inversion of images within the optic pathway. He also was clear in his opposition to the theory of extramission and instead proposed light and colour travelled in straight lines from an object and was converged at the eye [Al-Khalili, 2015]. His theory was seen to disprove extramission. A major advance was the suggestion by Ibn Sina and Abū Bakr Muhammad ibn Zakariyyā al-Rāzī that the optic nerves crossed at the chiasma [Finger, 1994].

Moving to mid-16th century Europe, Andreas Vesalius was the first to depict the optic nerves as non-hollow, a fact later confirmed in the late 17th century by Antony van Leeuwenhoek using microscopy. Vesalius is also credited with suggesting the retina, not the lens, is the receptive tissue within the eye, although his originality is disputed by those who claim the Arab scholar Averroes suggested this in the early 12th century. The importance of the retina was further supported by Felix Platter, who in the late 17th and early 18th century performed experimental procedures to evidence his theory [Lindberg, 1981].

Other notable advances include the discovery of the blind spot in 1668 by l'Abbé Edme Mariotte [Mariotte & Pecquet, 1668], the identification of the macula and the fovea by

Samuel Thomas von Sömmerring in 1791 [Finger, 1994], and the description of the crossing over of nerve fibres at the optic chiasma by Isaac Newton in 1704 [Newton, 1704].

Johannes Müller was the first to identify rods and cones, a series of cells that he described as rod like and cylindrical in shape, although he and several others incorrectly described them as facing inwards to the eye [Müller, 1837]. Heinrich Müller corrected this, describing the receptors as facing outwards from the eye. He also identified a number of the other key retinal layers. Max Schultze would later describe a distinction between rod and cone cells, as well as theorise about their unique roles based on observations of their ratios within animals of different chronotypic tendencies [Finger, 1994].

Advances in our understanding of retinal structure were enabled by the development of new fixation and staining techniques in the late 19th and early 20th century. Santiago Ramón y Cajal employed these techniques to identify the separation of the retinal nerve fibres and the optic nerve, the independent photoreceptors, bipolar cells, horizontal cells and amacrine cells [Ramón y Cajal, 1893].

The study of the eye has developed through history to a point of detailed understanding. We now possess an in-depth knowledge of the eye, both anatomically and mechanically. This has been enabled in part by novel imaging techniques discussed later.

1.3. Anatomy of the Eye

The eye has a complex structure, acutely tuned to its function, so much so that its design caused Charles Darwin to doubt his theory of natural selection [Darwin, 1859]. The eye is an organ made of numerous layers, each with a unique function (figure 1.1A). The outermost layer of the eye consists of the sclera and the cornea. The sclera makes up the majority of the outer layer and is a tough protective layer with a white colouring. The cornea constitutes the outermost layer at the anterior portion of the eye. It is a clear, pain sensitive connective tissue that allows light into the eye. The next layer is the choroid, the vascular layer that provides blood supply to all layers of the eye. The inner most layer of the eye is the retina, which absorbs light energy and converts it to neural signal that can be interpreted by the brain [James & al., 2003].

The journey of light to the retina starts at the cornea, which acts as a window for light into the eye. The cornea's curvature is responsible for the first stage of light refraction and focusing. It then passes through the pupil, an opening at the front of the eye whose size is controlled by the iris. The area between the cornea and the iris is the anterior chamber. The iris is the coloured muscle tissue in the eye whose contraction and dilation

control the amount of light allowed into the eye. The light then passes through the biconvex crystalline lens which provides further refraction. The space between the iris and the lens is known as the posterior chamber. Collectively the anterior and posterior chambers of the eye are called the anterior segment and are filled with aqueous humor. The lens is held in place in part by the ciliary bodies, which regulate the shape of the lens and are also responsible for production of aqueous humor which flows from the posterior chamber, into the anterior chamber [Fautsch & al., 2006]. The lens focuses the light onto the retina, which is the flat surface at the back of the eye that captures light and converts it to neural signal. The large cavernous space between the lens and the retina is known as the posterior segment and is filled with vitreous humor [Freddo, 2018].

1.3.1. Optics of the Eye

The human eye is required to detect light at a broad range of intensities and distances whilst maintaining high visual acuity and colour vision. Both the cornea and the lens contribute to focusing light onto the retina to maintain this visual acuity. The difference in the index of refraction between the air (1.0) and the cornea (1.38), and the subsequent change in the speed of light as it passes through the cornea, means that the cornea is accountable for about two-thirds of the refractive power of the eye. The lens performs the remaining focusing of the light onto the retina, aided by its convex shape. Due to this shape, the image is projected onto the retina inverted, and the brain is responsible for flipping it to the correct orientation (figure 1.1B). As the distance between the cornea to the retina is fixed, variation in the focal length is controlled by altering the shape of the lens. This allows focusing of light from objects at varying distances. The ciliary bodies control thickening and thinning of the lens to focus light onto the retina. Individuals with uncalibrated distance between the cornea and the retina, for example those who are myopic or hypermetropic, may require glasses that alter the focal length [Urone, 1998].

1.4. The Retina

The retina itself is a highly structured layered tissue (figure 1.2A). As previously stated, it is responsible for conversion of light energy into neurological signal. It can be broadly separated into the inner retina - the neural section nearest the front of the eye - and the outer retina - nearest the posterior of the eye. Light enters from the anterior of the eye,

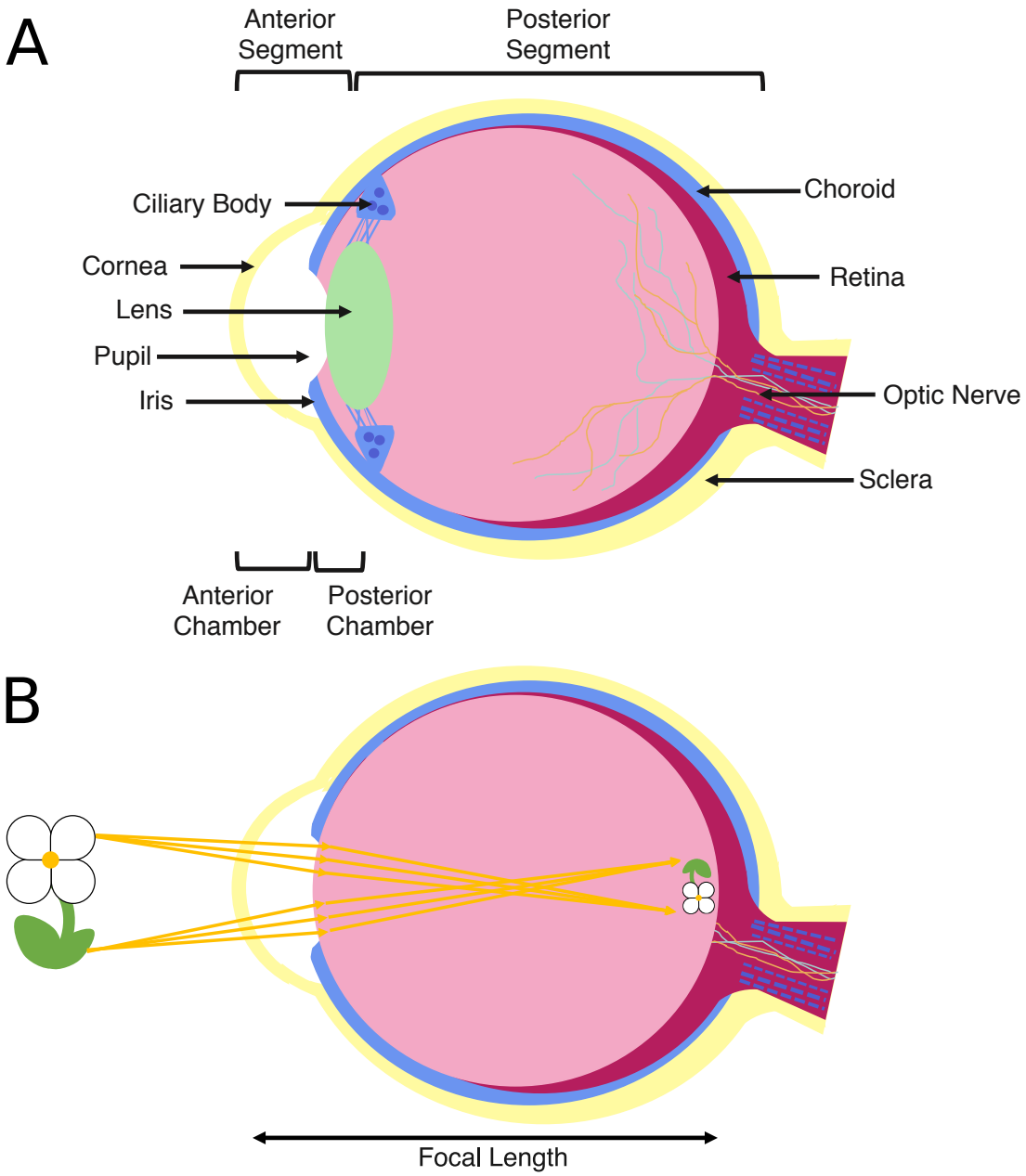

Figure 1.1: Anatomy and optics of the eye. A) A cross-section of the human eye highlighting different tissues and substructures. B) A diagram of the optics of vision, demonstrating the focal length and projection of an inverted image onto the retina.

travels through the layers of the retina until it is absorbed in the posterior of the retina by the photoreceptors in the photoreceptor layer (PRL).

Photoreceptors take light energy as an input and convert it to chemical energy in the form of neurotransmitters. Two types of photoreceptors can be found in the human retina: rods and cones. Rods are more populous in the peripheral areas of the retina, but are in the minority in the central regions - the macula and fovea (figure 1.2B). They are responsible for vision in dim light but lack colour definition, and cannot provide detailed vision. Conversely, cones are present mainly in the central retina and are responsible for high clarity and colour vision, but only function in higher light levels. Cones come in three forms: one that absorbs red light, one absorbing green light and one blue. By the activation of multiple cone types at one time, humans are able to see a spectrum of colour. The layer of photoreceptor cells is backed by the retinal pigment epithelium (RPE) and Bruch's Membrane (BM), the outer most layers of the retina. Both absorb any scattered light which increases definition in vision. They also both supply the other layers with nutrients and a blood supply, and remove waste from all the layers [James & al., 2003].

The component layers of the photoreceptor cells (PRC) are the outer segment (OS), inner segment (IS) and outer nuclear layer (ONL). The OS stores the photopigments within stacked membranes. The IS contains mitochondria and ribosomes that aid the assembly of the photopigments. The ONL contains both the cell body of the PRC, including the nucleus and other organelles, and the synaptic terminal [Purves & al., 2001]. Signal from the photoreceptors is integrated by horizontal cells in the Outer Plexiform Layer (OPL), which helps increase contrast in vision [Masland, 2012]. Signal from the photoreceptors and horizontal cells is passed to the ganglion cells in the Ganglion Cell Layer (GCL) by the bipolar cells in the Inner Nuclear Layer (INL). Once again, information from multiple bipolar cells is combined in the Inner Plexiform Layer (IPL), this time by amacrine cells. The ganglion cells, present in the GCL, then transfer signal to the nerve cells in the Retinal Nerve Fibre Layer (RNFL) and exit along the optic nerve to the brain via the optic disc. The optic disc contains no photoreceptors and is therefore a blindspot, an area in each eye's visual field that cannot see [Freddo, 2018].

Retinal layer structure differs slightly at the macula. The macula is an area at the centre of the retina responsible for high acuity vision. It has a characteristic valley-shaped morphology (figure 1.2B). At the base of the macula dip is the fovea. The macula has a much higher ratio of cones to rods than the rest of the retina. At the fovea only cones are present, hence this area being responsible for highest acuity vision. At the macula, the inner retinal layers - the neural layers - taper to non-existence, allowing a direct path

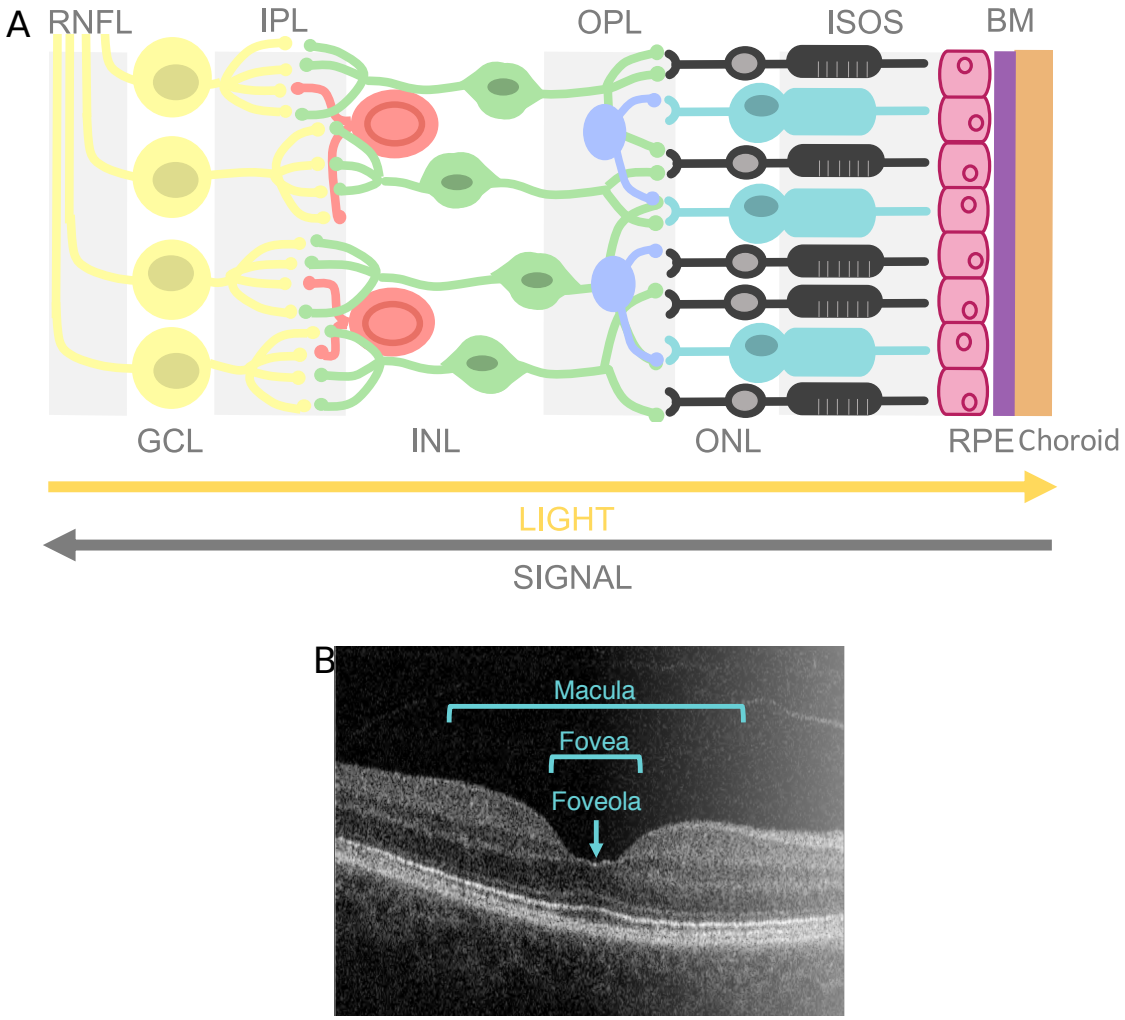

Figure 1.2: Anatomy of the retina and macula. A) A schematic of the layers of the retina. The light enters via the inner retina and is absorbed by the photorecpetors in the outer retina. The direction of travel of both light and neuronal signal is labelled. The light energy is converted into chemical signal which is then propagated back towards the inner retina before exiting the eye via the optic nerve. The layers of the retina are labelled: retinal nerve fibre layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), inner segment/outer segment (ISOS) and retinal pigment epithelium (RPE). Additionally the Bruch's membrane (BM) and choroid are labelled. B) A cross-section of the human macula, the central area of the retina, imaged with optical coherence tomography (OCT) and labelled with substructures. This shows a characteristic valley-shaped morphology that is responsible for highest acuity vision.

for light to the photoreceptors.

In many ways the direction of travel of both light and signal within the human eye seems counter-intuitive or back to front. The light must traverse all layers of the retina before it is absorbed, and the signal must then travel back from whence the light came. This is in contrast to other organisms such as cephalopods, whose photoreceptors are found inverted compared to humans, with the photoreceptor cells in the inner retina directed towards the light [Hanke & Kelber, 2020]. Indeed, there is much variation in eye structure across vertebrates which originally led to the initial conclusion that eyes had independently evolved more than 40 times [Lamb & al., 2009]. However, the discovery of the *PAX6* gene, which is highly conserved across the animal kingdom, supported a shared evolutionary history of eye development [Wawersik & Maas, 2000; Treisman, 2004; Bhatia & al., 2014] and further comparison of common structures suggest there are more shared aspects than originally thought [Lamb & al., 2009].

A structure that has developed differently in different species is the fovea. Fish, such as the teleosts, and reptiles all possess a fovea of differing depths with convex walls [Bringmann & al., 2018]. Comparatively the majority of mammals lack a fovea, although some maintain an area of increased photoreceptor density such as the visual streak in rabbits [Levick, 1967], or the area centralis in cats [Beltran & al., 2014]. The fovea appears to have "re-emerged" in higher primates, where it is present as a valley with concave sides that sees thinning of inner retinal layers at the centre, as in humans [Bringmann & al., 2018].

1.5. Imaging the Retina

Our current understanding of the retina has been largely informed by our ability to visualise it [Li & al., 2018]. There are multiple forms of ocular imaging that allow for visualisation of the tissue at different resolutions in a non-invasive manner. Here I will describe two of these methods in greater detail.

1.5.1. Fundus Photography

Fundus photography is a technique for obtaining a two dimensional view of the retinal surface using normal optical wavelengths [Abramoff & al., 2010]. It is non-invasive and allows ophthalmologists to observe the optic disc, macula and retinal vasculature [Bernardes & al., 2011]. It is a standard technique used in ophthalmological examinations [Yannuzzi & al., 2004] and a primary tool in the diagnosis pathway of several

diseases including age-related macular degeneration (AMD) [Fleckenstein & al., 2018] and diabetic retinopathy [Goh & al., 2016].

1.5.2. Optical Coherence Tomography

Optical coherence tomography (OCT) is a non-invasive imaging method used medically in dermatology, cardiology and ophthalmology [Kume & Uemura, 2018; Ulrich & al., 2016; Tsang & Sharma, 2018]. It is able to image optically scattering biological tissues, building three dimensional representations of the tissue by repeatedly imaging their cross-section [Gabriele & al., 2010]. The technique takes advantage of the coherent nature of light produced by a light amplification by stimulated emission of radiation (laser). Coherent light has a constant phase difference [Ng & al., 2016] and thus interference, the combination of two light waves, can be calculated. In OCT light is emitted by a coherent light source, a laser, and split by a mirror. One of the split beams is sent to a fixed mirror, and the other is sent to the tissue sample. Light is reflected by both the reflecting mirror and the tissue, and combined again at the splitting mirror. This combined signal is then measured by a detector or spectrometer (figure 1.3A). The return distance from the beam splitter to the reflecting mirror is l_1 and the return distance between the beam splitter and the sample is l_2. During OCT, l_1 remains fixed. The detector uses the interference of the beams coming from the reflecting mirror and the tissue to select the single scattered photons and calculate the relative location of the sample (figure 1.3B). The location of the reference mirror can be altered (axial scanning) to obtain information on different axial locations. To limit the area in which scanning is conducted, the light source contains a mix of different but very similar wavelength lights. The difference in wavelengths is proportional to the coherence length, the distance in which interference can be calculated, thus limiting the scanning field (figure 1.3C) [Huang & al., 1991]. OCT is performed routinely in hospitals and by primary care opticians [Fujimoto & Swanson, 2016], and assists in diagnosis and monitoring of a broad range of ocular diseases due to its visualisation of all retinal layers [Jaffe & Caprioli, 2004].

1.6. Dysfunction of the Retina

There are a wide range of retinal pathologies, each with different symptoms and underlying biology. Due to the interconnectivity of the different retinal layers, and the specific role each cell type has in the transmission of signal, disruption to any of these layers can lead to changes in sight. Here I highlight a few notable conditions that affect the retina.

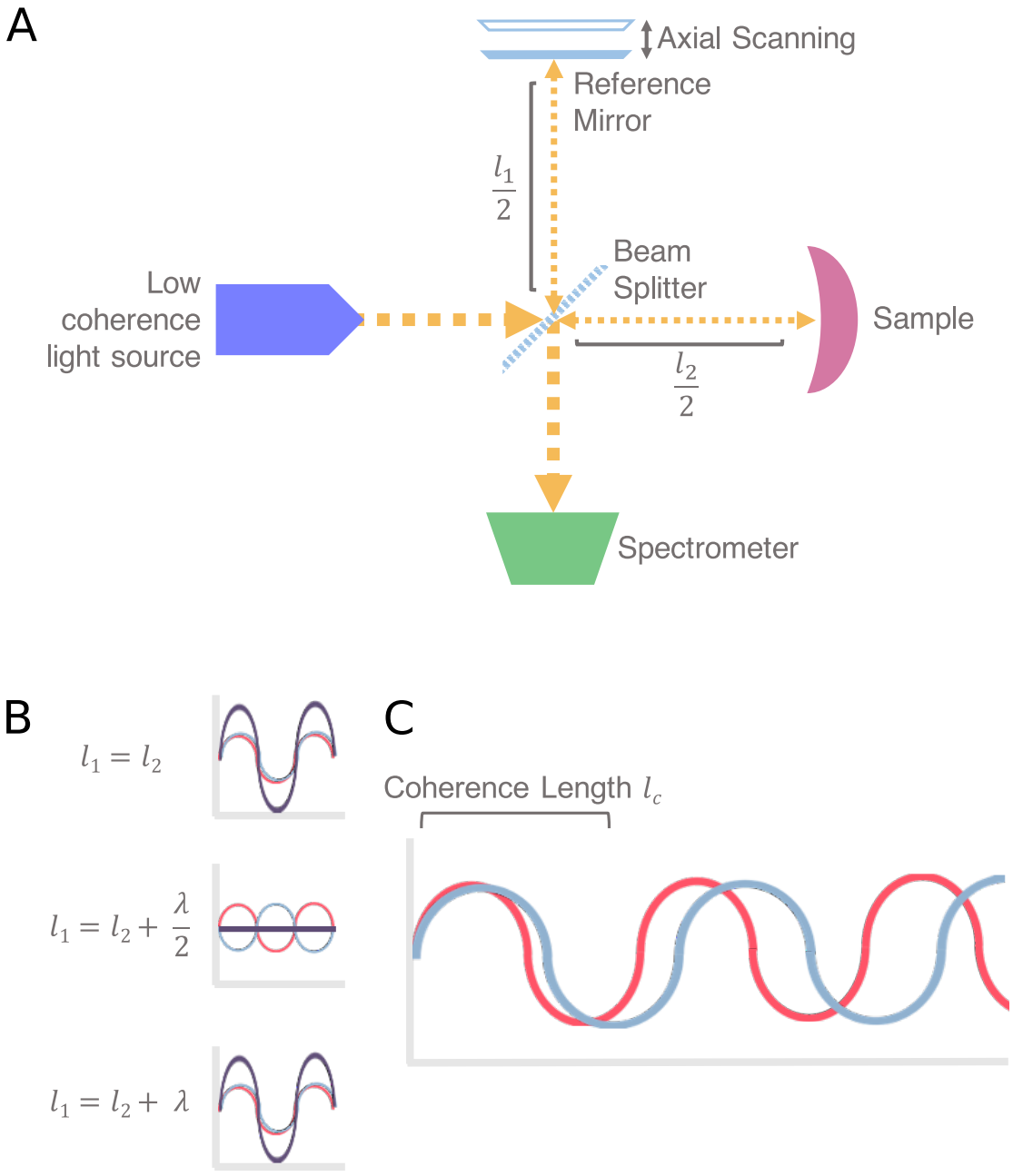

Figure 1.3: Optical Coherence Tomography (OCT). A) A simplistic representation of the mechanism underlying OCT. A low coherence light source, often a laser, is emitted and split by a mirror. The split beams are sent to a a reference mirror, a distance of $l_1/2$ away, and the sample, a distance of $l_2/2$ away. The light is reflected by the reference mirror and sample and combined at the beam splitter where it is reflected to a spectrometer. B) The difference between l_1 and l_2 affects the interference. C) If a light source contains a mix of light with slightly different wavelength, the length for which the wavelength is similar enough that interference can be calculated is known as the coherence length l_c.

1.6.1. Glaucoma

Glaucoma is the leading cause of irreversible blindness globally [King & al., 2013]. Broadly, it is a disease that causes damage of the optic nerve (figure 1.4) leading to a characteristic cupped optic disc [Chan & al., 2017]. There are four main types of glaucoma [Kwom & al., 2009]. Firstly, developmental glaucoma, which presents in infancy or childhood and is largely genetic [Taylor & al., 1999], or results from corrective surgery for other ocular diseases [Gothwal & al., 2020]. Secondly, secondary glaucoma refers to forms of glaucoma that are caused by other diseases, trauma or medication [Weinreb & al., 2014]. Thirdly, primary angle closure glaucoma is caused by a rapid increase in intraocular pressure (IOP) due to the passage way for the drainage of aqueous humor being too narrow [Sun & al., 2017]. And the fourth, primary open angle glaucoma (POAG) is the most common form of glaucoma [Weinreb & Tee Khaw, 2004].

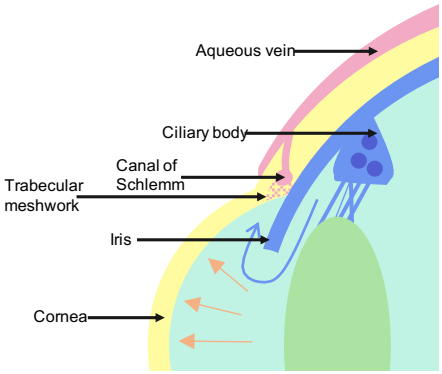

Figure 1.4: Aqueous humor production and drainage The anterior segment of the eye is filled with aqueous humor which is produced by the ciliary bodies. The humor then passes through the passage between the iris and lens to reach the anterior chamber (shown by the blue arrow). Aqueous humor is drained from the anterior chamber via the trabecular meshwork and canal of Schlemm, into the aqueous vein. Narrowing of the passage between the iris and cornea causes primary angle closure glaucoma (PACG). The gradual blockage of the trabecular meshwork leads to primary open angle glaucoma (POAG).

Primary Open Angle Glaucoma

POAG is the most common form of glaucoma. It is a complex disease with multiple causes that leads to degradation of retinal ganglion cells. POAG first presents as loss of peripheral vision due to damage of the outer layers of the optic nerve [Weinreb & al., 2014]. POAG can in extreme cases lead to blindness as optic nerve damage continues to the central field [King & al., 2013]. Glaucoma commonly co-exists with raised IOP.

One cause of this is the gradual blockage of the canal of Schlemm and trabecular meshwork (Fig1.4) [Andrés-Guerrero & al., 2017]. This leads to disruption in the drainage of aqueous humor from the anterior segment and often subsequent increased IOP and cumulative damage to the retinal ganglion cells and optic nerve [Hood & al., 2013]. Main treatment methods of POAG involve decreasing IOP [Beidoe & Mousa, 2012] whether that be through medication or surgical intervention [Weinreb & al., 2014]. However, some individuals who are diagnosed with POAG present with low IOP. There are several risk factors associated with POAG, mainly age and familial history of the disease [Quigley, 2011].

There have also been extensive studies on the genetic predisposition of individuals to develop POAG [Bailey & al., 2016]. The most recent meta GWAS by the International Glaucoma Genetics Consortium (IGGC) identified 127 independent loci associated with POAG [Gharahkhani & al., 2020]. There are a number of well characterised genes that are associated with POAG. *SIX6* has been found associated with POAG and is involved in regulating early retinal progenitor cell proliferation [Li & al., 2002]. This is also supported by *SIX6* affecting the thickness of the retinal nerve fibre layer [Cheng & al., 2014; Ulmer Carnes & al., 2014]. *CAV1* has been associated with POAG and encodes a protein which has been shown to act by protecting the canal of Schlemm from elevated IOP [Elliott & al., 2016]. *CDKN2B-AS* has also been shown to be associated with POAG and acts synergistically with *CAV1*, and *TMCO1*, which has been linked to apoptosis [Burdon & al., 2011].

1.6.2. Age-Related Macular Degeneration

Age-related macular degeneration (AMD) is a condition that affects the centre of the retina, the macula, and causes progressive loss of vision [van Lookeren Campagne & al., 2014]. It is generally categorised into early-stage and late-stage AMD [Ferris & al., 2013]. Late stage AMD can be further categorised into non-neovascular, or dry AMD, and neovascular, or wet AMD [Jager & al., 2008]. The early stages of AMD are characterised by the build up of lipid deposits known as drusen below the RPE. Late stage non-neovascular AMD is characterised by the accumulation of these drusen in both size and number [Mitchell & al., 2018]. This leads to degradation of the RPE [Al-Zamil & Yassin, 2017], the retinal layer that provides nutrients and blood supply to the photoreceptors. This will present as increasing loss of vision as the photoreceptors are progressively damaged [Cunningham, 2017]. Late stage neovascular AMD is characterised by the additional abnormal growth of blood vessels in the choroid. If left untreated these blood vessels may burst, and the resulting scar tissue leads to irreversible loss of sight

[Lim & al., 2012]. Age, smoking and ethnicity have been identified as risk factors for AMD [Cunningham, 2017]. The very first GWAS conducted was of AMD [Klein & al., 2005] and since then further work has led to well established genetics underlying the disease. A recent GWAS identified 52 variants associated with AMD including rare coding variants in *CFH*, *CFI* and *TIMP3* [Fritsche & al., 2016].

1.6.3. Retinal Dystrophy

Retinal Dystrophy refers to a group of conditions, mostly rare genetic conditions, that cause deterioration of vision associated with the retina [Ziccardi & al., 2019]. These conditions are all incurable [Sahel & Dalkara, 2019] but have a wide range of management therapies [Dias & al., 2018; Gill & al., 2019; Sahel & Léveillard, 2018]. A recent study suggests that, worldwide, approximately 5.5 million individuals are affected by an autosomal recessive inherited retinal disease or dystrophy [Hanany & al., 2020]. Retinal dystrophies are broadly categorised into those that are isolated to just the eye, isolated retinal dystrophies, and those that affect the retina in addition to other areas of the body, syndromic retinal dystrophies [Verbakel & al., 2018; Hamel, 2006]. The isolated retinal dystrophies can be further categorised into: early onset retinal dystrophies; rod-cone dystrophies - in which rods are affected prior to cones [Hamel, 2007]; cone-rod dystrophies - in which cones are affected prior to rods [Gill & al., 2019]; and macular dystrophies, which affect central vision [Rahman & al., 2020].

Macular dystrophies include retinitis pigmentosa [Mrejen & al., 2017], the most common form of retinal dystrophy [Hanany & al., 2020] which affects 1 in 3000-7000 individuals and causes progressive loss of peripheral vision [Ratnapriya & Swaroop, 2013]. More than 40 genes have been found associated with one of the forms of retinitis pigemtosa [Parmeggiani & al., 2011]. Genes of particular importance in the disease include *RHO*, *USH2A* and *RPGR*, which when considered together account for ~30% of retinitis pigmentosa cases [Hartong & al., 2006]. Another example is fundus albipunctatus, characterised by impaired night vision and lesions across the retina excluding the fovea [Skorczyk-Werner & al., 2015]. This condition is known to be caused by a mutation in *RDH5*, the gene encoding the 11-*cis* retinol dehydrogenase 5 [Yamamoto & al., 1999].

1.6.4. Oculocutaneous Albinism

Oculocutaneous albinism (OCA) is a rare genetic condition that affects levels of pigmentation in the skin, hair and eyes [Kamaraj & Purohit, 2014]. A key clinical feature

of OCA is foveal hypoplasia, the flattening of the foveal dip due to lack of excavation of central inner retinal layers [McAllister & al., 2010; McCafferty & al., 2015]. Both lack of retinal pigment and foveal hypoplasia leads to reduced visual acuity. Affected individuals also sometimes experience nystagmus, rapid repetitive movements of the eye [Grønskov & al., 2007].

Mutations in several genes have been previously identified as underlying causes of OCA. Such genes include *TYR*, *OCA2*, *TYRP1*, *SLC45A2* and *SLC24A5* [Montoliu & al., 2014; Simeonov & al., 2013]. Defects in the different genes lead to distinct sub-categories of the disease, with different types of the disease having higher frequency in different populations and causing slightly different symptoms [Norman & al., 2017; Kirkwood, 2009].

1.7. Understanding Retinal Morphology

As has been outlined, changes in retinal morphology are a key feature of retinal disease. Due to the nature of the eye, the structure is integral to its function, and even slight changes to its anatomy can cause changes in its capacity for accurate vision. Therefore there is a body of research into the morphology of the retina, its links to pathology, and its underlying genetic drivers. This has been enhanced by the advances in retinal imaging technologies. Studies have shown changes in retinal layer thickness in diseases including glaucoma [Khawaja & al., 2020; Kim & Park, 2018], AMD [Wood & al., 2011; Nivison-Smith & al., 2018] and, more recently several neurodegenerative diseases including Alzheimer's disease [Kim & Kang, 2019; Chan & al., 2019; Mutlu & al., 2018] and Parkinson's disease [Ma & al., 2018]. Studies have been conducted to look at the effect of lifestyle variables such as smoking [Duman & al., 2017], and biological variables such as ethnicity [Kashani & al., 2010] on retinal thickness. Most recently, work has been conducted looking for genetic variation affecting the overall thickness of the retina, finding several novel loci, and some loci previously associated with ocular diseases [Gao & al., 2019].

Work is still needed to understand the genetic contribution to the morphology of the individual retinal layers. This is of value because, as previously described, the different retinal layers are responsible for different biological processes and steps in visual signal processing. For this reason, each of the layers is also affected by different pathologies. Therefore, insight on the common genetic variation underlying the individual layers may help understand visual mechanism and pathology at a more detailed layer, providing clinical value. The creation of large datasets containing both retinal imaging

data and genetic data, as well as the development of methods for analysing such data, has allowed for my further exploration of this topic.

2. Genetics

Genetics, the study of heredity and its molecular basis, has a long and rich history. To reach our current understanding it was necessary to elucidate the fundamental theories underlying inheritance, the mechanisms by which it occurs, and the molecular reagents that enable it.

Scientists and philosophers began theorising about the transmission of traits between generations over 2000 years ago [Mukherjee, 2016]. Farmers have been utilising mechanisms of inheritance for even longer [Burian & Zallen, 2009]. However, our understanding of the underlying biological mechanisms behind inheritance are relatively recent. Technological advances now allow for rich and detailed studies that help us understand the fine-grain effect of genetics on human biology.

Scientists have long been interested in recording the variation in human traits. The creation of databases containing large scale phenotypic data has enabled novel studies of human variation. More recently the ever-falling costs of genetic sequencing have led to increasing masses of genetic data. This genetic data has been aggregated with phenotypic data into 'biobanks'. These vaults of data allow large-scale analysis of genetic variation underlying a broad array of traits.

The access to this coupled phenotypic and genotypic data is important for clinically-relevant research. Not only is the boom in data accumulation being found in the research arena, it is massing in the clinical space. The Global Alliance for Genomic Healthcare (GA4GH) predicts that by 2025, 60 million patients will have had their genome sequenced within the healthcare system [Birney & al., 2017] . More and more, healthcare systems are turning to coupled genetic and phenotypic data for patients as an aid to targeted clinical diagnosis and treatment. This translation of basic research on phenotype-genotype relationships into the clinical space is increasingly important, and requires large-scale datasets and novel techniques that utilise clinically interpretable phenotypes.

Here I provide a brief history of genetics, highlighting the discoveries that made the research within this thesis possible, and detailing some of the particular methods used in my analysis.

2.1. Early Theories of Inheritance

2.1.1. Ancient Ideas of Inheritance

For millennia, people have noted that the offspring of plants, animals and humans bear resemblance to their parents, and have utilised this in the way they selectively breed crops [Burian & Zallen, 2009]. Although they recognised these patterns, they had yet to understand them. Early theories of inheritance can be dated back to Pythagoras, who suggested that semen traversed the body collecting information about various traits that was then nurtured in the womb into an offspring [Mukherjee, 2016]. Aristotle rejected this, noting that traits from both parents are present in the offspring, and most obviously, how could male sperm learn information to produce female genitalia? Consequently he postulated that whilst the father provided the instructions to build an offspring, the mother provided the materials for construction [Harper, 2008]. Various other mechanisms were suggested in the following centuries, including the idea of the homunculus, a miniature fully formed human that resides in the sperm and is nurtured to full size in the womb [Mukherjee, 2016].

Hippocrates put forward another theory of inheritance which was supported by multiple other philosophers including Galen, previously discussed for his contributions to eye anatomy (section 1.2). This theory proposed that all parts of both parents were involved in the process of trait inheritance [Zirkle, 1935]. Charles Darwin later championed and formalised this theory, christening it *Pangenesis*. His interpretation stated that each part of the body secreted a molecule, known as a *gemmule*, that possessed information about the production of that cell. This information passes to offspring through the gametes, activated during fertilisation to create a fully formed human [Darwin, 1868]. An advantage of this theory was that it allowed for traits acquired throughout one's lifetime to be passed on to offspring.

2.1.2. Evolution and Inheritance

Charles Darwin's more famous theory is that of natural selection, originally published in tandem papers at the Linnean Society with Alfred Wallace, titled *On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection* [1858], and later in his book *On the Origin of Species* [1859]. During his travels around Central and South America, Darwin made observational studies of trait variation between species and concluded that speciation occurs when random advantageous specialisations to niche environments are sexually selected for, and persist in, sub-

sequent offspring. This theory revolutionised the field of natural history, suggesting gradual change similar to that proposed by Charles Lyell [1832] in relation to formation of geological formations. This directly contradicted the religiously-underpinned notions of both Lyell and Darwin's time.

Darwin also realised that his theory required some integral assumptions about the nature of hereditary. Firstly, that variation amongst a species can exist in order for advantageous mutants to arise. Secondly, that this variation must be fixed in some manner for the newly acquired traits to be passed on to new generations and selected for. The two assumptions seemed at odds: there needed to be both rigidity of the system to enable transfer between generations, but flexibility to allow for creation of new traits. In *The Variation of Animals and Plants under Domestication* [1868], Darwin adopted Hippocrates' theory of Pangenesis in an attempt to meld and satisfy these two assumptions, however it was widely criticised.

2.2. Mendelian Laws of Inheritance

In parallel with Darwin's famous work, Gregor Mendel was conducting research on the mechanisms of inheritance that would also revolutionise the understanding of genetics.

Mendel was a monk at St Thomas' Abbey in Brno, now the Czech Republic, where he dedicated his time to understanding patterns of inheritance in the monastic gardens [Mukherjee, 2016]. Using pea plants as his experimental reagent, Mendel identified a number of true-breeding traits including pea texture, height and flower colour. By systematically cross-breeding both pure-bred plants, and subsequently their offspring at a large scale, Mendel noted distinct patterns in which traits were inherited.

His main observations were that each trait had two distinct forms, which others later would term alleles, and that these traits did not blend but were inherited in distinct patterns. He condensed these patterns and his subsequent inferences about inheritance into three laws, presented in *Versuche uber Pflanzen-Hybriden* (*Experiments on Plant Hybridization*) [1866]. Firstly, the law of dominance: of the two forms of a trait, one is recessive and one is dominant. The dominant form will be expressed as an observable trait, overruling and masking the recessive form. Secondly, the law of segregation: the two alleles are separated in gamete formation with a single allele present in each gamete. Finally, the law of independent assortment: this segregation process is random, with different traits non-reliant on one another.

These three laws bind together to form a solution to the problem that Darwin aimed to solve with Pangenesis. Considering the centrality of his findings to modern genetics,

Mendel's work remained unrecognised for an astonishingly long period.

2.3. Biometrics

Meanwhile an alternative theory of inheritance was being cultivated, that of the biometricians. Francis Galton, Darwin's half cousin, was born in the same year as Mendel. Galton latched on to many of Darwin's proposed theories of evolution and began to draw parallels and apply the theories to human beings. Inspired by Adolphe Quetelet's extensive work measuring various traits in a cohort of soldiers [Quetelet, 1835], Galton set out to perform a similar study of his own.

Using sometimes unscientific methodology, Galton collected large-scale data on multiple facets of human variation including height, intelligence and beauty. Similar to the findings of Quetelet, Galton noticed that the distributions of continuous traits were bell-shaped: the majority of data points sit around the mean, with a minority of data points at the extreme values. Galton consequently constructed the concept of regression towards the mean, presented in *Regression towards mediocrity in heredity stature* [Galton, 1886]. Karl Pearson further developed these initial concepts into statistical analyses that remain commonly used in modern genetics. These include correlation, the chi-square (χ^2) test and the use of p-values, and principal component analysis (PCA) [Pearson, 1900; Pearson, 1901]. The joint school of study advocated by Galton, Pearson, and Walter Weldon, who applied Galton's statistical technique in his zoological studies [Weldon, 1890; Weldon, 1892], was termed *Biometrics*.

Having made several inferences about human variation, Galton wished to expand his work and understand how such variation was inherited. To do so, Galton studied family pedigree and concluded that an ephemeral mix of social class, intelligence and success, a concept he coined as "eminence", was a biologically heritable trait [Galton, 1869]. However, his theories about normal distribution of anthropometric traits did not easily mesh with his concepts of "eminence". Whereas height tended towards an average, with only exceptional cases being strongly influenced by parents, eminence (a concept we might now know as nepotism) appeared to be much more strongly influenced by family structures.

As a recipient of a powerfully nepotistic upbringing, and in an age of industrial and social revolution, Galton felt existentially threatened by the rise of the working masses and the "undesirable" qualities they possessed. These traits were grounded in his knowledge of class, yet piqued with underlying racism and sexism: the upper-class white man was the ideal. Misinterpreting his own results, and fuelled by this social backdrop,

Galton advocated for selective breeding of those with desirable traits. He termed this practise *eugenics*, meaning "well born" [Galton, 1883], and thus birthed a movement that would cast a shadow over modern genetics. In spite of both the long-term negative impacts of eugenics and biometrics ultimately being eclipsed by Mendelian theories, the statistical techniques developed by the founders of biometrics are still central to current statistical genetics.

2.4. Marrying Evolution and Mendelian Laws of Inheritance

2.4.1. Rediscovering Mendel

Mendel's work was simultaneously rediscovered by multiple scientists in 1900, following many years of being overlooked. Hugo de Vries replicated many of Mendel's studies in pea-plants. He supported Mendel's concept of discontinuous variation which he proposed worked through sudden mutations causing large shifts in offspring traits as compared to their parents. He summarised that traits have distinct particles that carry this hereditary information, which he called *pangenes* [De Vries, 1889]. The botanist Carl Correns also reiterated Mendel's findings [Correns, 1900], as did Erich von Tschermak [von Tschermak, 1900].

One of Mendel's largest champions was William Bateson who dedicated a portion of his career to spreading the Mendelian perspective of inheritance and translated his works into English [Bateson, 1902]. Bateson is credited for coining the word *genetics* to describe the study of inheritance, as well as the term *allelomorph*, now shortened to allele, to describe the two variants of a trait [Wilks, 1907]. His colleague Reginald Punnett was also important in bridging the gap between Mendelian concepts of inheritance and statistics, establishing the Punnett square to illustrate the combinations of alleles that can arise [Punnett, 1907]. Alongside Punnett and Bateson's independent research successes, they collaborated with Edith Rebecca Saunders to recreate Mendel's experiments.

It is interesting to note that many with a broad knowledge of the history of genetics may be unfamiliar with Saunders. Prior to her collaborative work with Bateson, she had independently studied inheritance in *Biscutella laevigata*, observing that the texture of the plant's leaves was inherited discontinuously and not blended [Saunders, 1897]. In her time, she was a reputable and successful scientist. She published many papers as an individual, was elected the first female member of the Linnean society, and served as the president of the British Genetical Society in 1938 [Bronstein & Bolnick, 2018]. Despite her ordainment by J B S Haldane as the "mother of British plant genetics" [Haldane, 1945], Saunders contributions to the field have been largely minimised by scientific his-

torians. This speaks to a wider culture of reducing the role of women and other minority groups in the field of genetics and wider science. One of her many important contributions, which is often attributed to only Bateson and Punnett, is the first account of gene linkage deduced from proportions of flower colour and pollen grain shape [Bateson & al., 1905].

With this increased prominence, Mendelian patterns of inheritance became a powerful rubric for the explanation of human disease in the early 20th century. Archibald Garrod, a physician working in the 1920s, recognised Mendelian patterns of recessive inheritance in alkaptonuria, a disorder affecting urine. He later published this and his similar observations in cystinuria, pentosuria and albinism in *Inborn Errors of Metabolism* [1923].

Having established the use of the term genetics, more necessary vocabulary was created by various scientists to help describe the mechanisms of inheritance explained by Mendelian theories. Wilhelm Johannsen defined a number of integral genetics terms including *gene*, used to refer to a unit of inheritance. He also established the terms *phenotype*, the external trait expressed by an individual, and *genotype*, the underlying genetic trait of an individual [Johannsen, 1909]. Other important lexicographical inventions of the time include *polygenetic* (or polygenic), referring to a trait that is influenced by multiple genes [East, 1910], and *pleiotropic* referring to genes that influence multiple, seemingly unrelated traits [Stearns, 2010].

2.4.2. The Birth of Quantitative Genetics

With the re-emergence of Mendelian theory, and its undeniably astute observations about inheritance, there was a gap to be bridged between these laws and the theories of evolution posed by Darwin. In 1918, Ronald A Fisher published *The Correlation Between Relatives on the Supposition of Mendelian Inheritance* [1919]. In this paper, he established that through polygenicity, the effect of multiple genes on a single trait, continuous variation can occur as a result of Mendelian inheritance. This paper also features the first use of the term *variance*, defined as the square of the standard deviation.

Throughout his career, Fisher established a number of additional statistical concepts that are integral to modern quantitative genetics. He initially studied experimental design [Fisher, 1935], with other landmark contributions including his work on maximum likelihood, establishing a number of statistical distributions, and development of a number of statistical tests. These included Analysis of Variance (ANOVA) [Fisher, 1925] and Fisher's exact test [Fisher, 1922]. Later in his book *The Genetical Theory of Natural Selection* [Fisher, 1930], he outlined a number of concepts that further linked theor-

ies of evolution with statistical interpretations of genetics, now known as quantitative genetics [Fisher, 1919].

Unfortunately, Fisher extrapolated his work on natural selection to human fitness, claiming that societal influence was enabling the survival of "less well adapted" individuals, as detailed in the latter portion of *The Genetical Theory of Natural Selection* [1930]. This speaks to a growing popularity of eugenics at the time, and a further shift from positive eugenics, in which individuals with reportedly desirable traits were encouraged to breed, towards negative eugenics, in which individuals with supposedly undesirable traits were forcibly prevented from breeding [Mukherjee, 2016].

In eugenics, Fisher and others abused Darwin and Mendel's works to fit pre-existing racism and classism. Fisher was a British advocate for sterilisation [Blacker, 1931], working in close parallel to eugenics researchers in Nazi Germany who provided the cloak of science to a brutal programme of genocide.

Eugenics has since been widely discredited. Morally, its use in justifying the atrocities perpetrated in Nazi Germany exposed eugenics as a profoundly reprehensible practice. In parallel, the field of science has come to recognise the large flaws in the scientific justification for eugenics. The importance of genetic diversity, exemplified by the negative effects of persistent selective breeding in agriculture and domestic animals, is diametrically opposed to the goals of eugenics. In spite of being discredited, Fisher remained unfaltering in his support of the eugenics movement until his death [Weiss, 2010].

In the modern field of genetics, and throughout this thesis, statistical methods developed by Fisher and his contemporaries who supported eugenics are utilised and considered central. Meanwhile, a growing minority of society still holds moral sentiment for the eugenics movement. It is a nuanced line to walk in which we must admit the importance of Fisher's work to quantitative genetics, whilst not providing a tacit endorsement for the re-emerging support of his beliefs. More broadly it is vital that scientists remain aware that increased knowledge of the scientific mechanisms of life is not sufficient justification for actions and policy. Self-criticism, scientific rigour and regulation by outside authorities must all contribute to ensure science remains both advancing and ethical.

2.5. Deciphering the Medium of Inheritance

As the mechanism of inheritance was gradually deciphered, work was concurrently being conducted into the underlying molecular medium. In 1871, Friedrich Miescher first published his work identifying *nuclein*, an acidic molecule that is located in the cells

nuclei [Miescher, 1871]. Over a decade later, Walter Flemming published *Zell-substanz, Kern und Zelltheilung* (*Cytoplasm, Nucleus and Cell Division*) [1882] in which he presented his discovery of chromatin, the thread-like structures in the nucleus that stained using basophilic dyes. Flemming further described the chromatin's movements during *mitosis*, the term he coined for the division of cells.

Later, both Walter Sutton and Theodor Boveri advanced the research with a description of *meiosis*, the formation of gametes. Sutton, in his initial paper on the subject, *On the morphology of the chromosome group in Brachystola magna* [1902], established that chromosomes occur in pairs when studying grasshoppers. This led to his hypothesis that chromosomes are the medium of inheritance, and that division of chromosomes during meiosis is in-keeping with Mendel's laws. His later paper *The Chromosomes in Heredity* [Sutton, 1903] reinforced the connection between Mendel's Laws of Inheritance and chromosomes as the medium of this process. Boveri's investigation into the reproduction of sea urchins established that all chromosomes were essential to normal embryo development [Laubichler & Davidson, 2008]. Sutton and Boveri's combined work is the foundation of the Boveri-Sutton chromosome theory [Wilson, 1925]. This theory stated that chromosomes occur in pairs, with each parent contributing one chromosome to each pair. Further, it proposed that all chromosomes are needed for an organism's normal development. Therefore each chromosome must be performing its own role. Nettie Maria Stevens contextualised the Boveri-Sutton theory when she proposed sex determination was controlled by chromosomes based on her observations of mealworm reproduction [Stevens, 1905b; Stevens, 1905a].

In 1904, Thomas Hunt Morgan joined Colombia University where he was interested in studying inheritance and species variation. He chose *Drosophila melanogaster*, the fruit fly, as his animal model which he studied in his famous fly room [Allen, 1968]. There he worked on identifying inheritable artificial mutations. This took some time, but eventually he identified a white-eyed mutant amongst red-eyed wild-type flies. Cross-breeding of mutants and wild-types produced only red-eyed flies. However interbreeding of these red-eyed offspring produced a mixture of red-eyed flies and white-eyed flies, with all the white-eyed mutants being male. Although linkage had been previously observed, Morgan synthesised this observation with the knowledge of sex-determination being chromosome bound. Therefore Morgan concluded that eye colour was likely carried on the sex-chromosome as the traits of sex and eye colour are linked [Morgan, 1910]. He extrapolated this conclusion to hypothesise that other genes may reside on particular chromosomes [Morgan, 1911]. Linkage was further noted in both maize by Guy N Collins [Collins, 1912] and later in mice by John B S Haldane [Haldane & al., 1915].

Thomas Hunt Morgan continued to introduce mutations in his flies that caused phenotypic variants. By doing so he identified variants that were inherited as four distinct "linkage groups", equivalent to the number of *Drosophila* chromosomes. This provided additional support for chromosome theory. One of the identified trait variants produced miniature wings. Although the gene responsible for this trait also appeared to be located on the sex chromosome it was not always inherited alongside eye colour. Morgan suggested that he was observing a phenomena of *chiasmatypie* previously described by Frans Alfons Janssens in 1909 [Koszul & al., 2012]. Thomas renamed the phenomena chromosomal crossover, and updated the concept to suggest that the different genes may cross over with differing frequencies, representative of their physical distance in relation to one another. Barbara McClintock and Harriet Creighton observed chromosomal crossover in maize, and described the physical process by which it occurred, having observed it using microscopy [Creighton & McClintock, 1931].

The specification of hereditary material was further narrowed by Oswald Avery. He disproved the more common hypothesis that protein was the genetic material and established Deoxyribonucleic Acid (DNA) as the genetic material instead [Avery & al., 1944]. This led to a cascade of discovery. Firstly Erwin Chargraff established that DNA is composed of four bases - adenine, cytosine, guanine and thymine - and determined the ratios in which these bases occur [Chargaff, 1950]. This knowledge aided James Watson, Francis Crick, Rosalind Franklin and Maurice Wilkins in the determination of DNA structure: two complementary chains of nucleic acids arranged in a double helix [Watson & Crick, 1953]. Resolution of the structure allowed inference of mechanisms of replication. Subsequently Crick proposed the central dogma, the idea that DNA sequence encodes protein sequence, but this flow of information is one-way and may be mediated by an intermediate molecule [Crick, 1958]. Mahlon Hoagland and Paul Zamecnik identified transfer ribonucleic acid (tRNA) as this intermediate molecule [Hoagland & al., 1958]. This was confirmed by Sydnney Brenner and Francois Jacob, whilst studying *Escherichia coli* [Brenner & al., 1961]. Subsequently Brenner, Crick and Leslie Barnett established that the genetic code that translates into protein is read in triplets [Crick & al., 1961] and by 1966 work by Marshall Nirenberg, Heinrich Matthei [Nirenberg & Matthaei, 1961], and Har Gobind Khorana [Khorana & al., 1966] had "cracked" this genetic code.

The research into both the mechanism and reagents of inheritance is broad, rich and deserving of more detailed exploration. However for the purposes of this chapter, I will now focus on several areas that are practically important to the research presented in this thesis.

2.6. Mapping Genomes

2.6.1. Creating the First Genetic Maps

Alfred Sturtevant extended his supervisor Thomas Hunt Morgan's concept of *crossing over* to theorise that the frequency by which traits were jointly inherited was inversely proportional to the distance of their corresponding genes along chromosomes. Using this logic, he was able to construct maps about the relative locations of genes. He published the first genetic map in 1913, detailing the relative location of genes along the X chromosome of *D. melanogaster* [Sturtevant, 1913].

In 1921, the first example of chromosome crossing-over in vertebrates was observed between *Aplocheilus latipes* (Japanese Medaka fish) sex chromosomes [Aida, 1921]. In 1937, Haldane and Julia Bell published the first example of linkage in humans, identifying linkage between haemophilia and colour blindness when studying family pedigrees [Bell & Haldane, 1937]. The potential for human gene maps was furthered by James Renwick and Sylvia Lawler, who utilised the novel concept of Logarithm of Odds (LOD) to analyse large scale family pedigrees and identify linked conditions, including the ABO blood group and nail-patella syndrome [Renwick & Lawler, 1955]. Renwick and Jane Schulze later computerised this process [Renwick & Schulze, 1961].

In 1970 Hamilton Smith discovered restriction enzymes that could break and recombine DNA in a targeted manner. This finding enabled the creation of *restriction maps* [Smith & Welcox, 1970]. These restriction maps could identify genetic variation as differences in the characteristic sizes of DNA fragments, known as restriction fragment length polymorphisms (RFLPs). David Botstein proposed taking advantage of this by identifying RFLPs throughout the genome that would act as landmarks when studying linkage [Botstein & al., 1980]. This was initially applied in a number of model organisms before allowing the mapping of the Huntington's gene in 1983 by Nancy Wexler [Gusella & al., 1983]. One of the genetic variants causing retinitis pigmentosa, the retinal dystrophy, was also mapped using this method [Bhattacharya & al., 1984].

2.6.2. Reading DNA Sequence

These initial genetic maps still lacked fine grain nucleotide-by-nucleotide information, and reading direct genetic code remained an elusive challenge. The first sequence deciphered was a twelve nucleotide section of bacteriophage lambda DNA by Ray Wu and Ellen Taylor [1971] using primer extension.

In the early 1970s, two groups were competing to establish a novel method of sequen-

cing: Walter Gilbert and Allan Maxam at Harvard, and Fredrick Sanger in Cambridge. Sanger's methodology recruited DNA polymerase, the enzyme that copies DNA. By stalling the replication process via the incorporation of modified nucleotides he could map out the sequence of DNA bases [Sanger & al., 1973; Sanger & Coulson, 1975]. This method eclipsed that of Gilbert and Maxwell [Maxam & Gilbert, 1977], and in 1977 Sanger published the 5,386 nucleotide sequence of ϕX174 virus [Sanger & al., 1977]. He would later go on to sequence human DNA for the first time, in the form of mitochondrial DNA, using shotgun sequencing [Anderson & al., 1981], an expansion of his previous method [Messing & al., 1981]. In this method, the DNA is broken up into random fragments, cloned and separately sequenced. These fragments are later reconstructed into a single sequence by utilising the overlapping nature of the fragments. This allowed for sequencing of longer sections of DNA.

Following the establishment of this technique, the 1990s saw a rapid succession of organismal genomes being sequenced: firstly the bacteria *Haemophilus influenzae* [Fleischmann & al., 1995]; followed by the first eukaryotic organism, yeast *Saccharomyces cerevisiae* [Goffeau & al., 1996]; then *Caenorhabditis elegans* [The C. elegans Sequencing Consortium, 1998]; and *D. melanogaster* [Adams & al., 2000]. In 2000 the first draft of the human genome was announced [Lander & al., 2001; Craig Venter & al., 2001]. The sequence was complete in 2003 [Abdellah & al., 2004] and consisted of 3 billion base pairs.

Since then, huge advancements have been made in the speed and volume in which sequencing can be performed. Next-generation (or second-generation) techniques allow for parallelisation of the process [Shendure & al., 2017]. The advent of genotyping using microarray chips allowed for the simultaneous identification of genetic variants at a large scale [Wang & al., 1998]. We have now moved into what is known as the third-generation of DNA sequencing marked by the development of technologies such as Oxford Nanopore Technologies (ONT) sequencing [Heather & Chain, 2016].

2.7. Genotype-Phenotype Association Analysis

The study of the relationship between genotypes and phenotypes in humans was enabled by the various methods of genome mapping. The relationship was initially explored using genetic linkage studies. The advent of second-generation sequencing and subsequent increase in genetic datasets brought with it the use of association studies. Both of these methods are detailed in brief below.

2.7.1. Genetic Linkage

Linkage studies have been used to explore the relationship between human genotypes and phenotypes since the 1930s. They utilised the concepts articulated by Morgan and Sturevant and applied them to humans to identify the location of genes of interest, often those involved in disease. By studying the transmission of disease or other traits within family structures, and the joint inheritance of parallel traits, the relative location of a gene can be described. This allowed for the discovery of a number of disease loci including that of Huntington's disease [Gusella & al., 1983] and cystic fibrosis [Tsui & al., 1986]. Despite its utility in the early study of genetic disease, it is limited to use in rare and monogenic disease.

2.7.2. The "Common Disease-Common Variant" Hypothesis

As previously described, linkage analysis had great success in mapping the location of rare and monogenic diseases. This was possible as the genetic variation underlying such diseases has a near binary effect on the presence of phenotype. The variation is rare, but the size of effect of the variant on the phenotype is large. Conversely, it was observed that many common diseases, such as coronary heart disease, did not follow the same pattern. These diseases were termed polygenic, meaning they were affected by a number of genetic variants. Each variant would contribute a small effect to the outcome phenotype, but their occurrence in the population was much more frequent. The cumulative effect of these genetic variants produces the disease phenotype. This is the underlying principle of the "common disease-common variant" hypothesis, stating that genetic variants involved in complex disease are likely to be common, but affect disease risk in smaller increments. For these diseases, linkage is less likely to be an effective method for studying the genetic-phenotypic relationship, as people with the same phenotype are not always representative of the same locus variant. Therefore new methods were needed that could take account of variant effects at the population level.

2.7.3. Association Studies

Association studies move away from the study of family structures, and instead look at the effect of genetic variants at the population level. Conducted in a case-control manner, a chi-squared test, originally outlined by Pearson in the 1900s, is commonly used to test if there is a significant difference in the frequency of a genetic variant between those who possess a trait, and those who don't. This technique was popular throughout the 1990s and allowed for the identification of variants affecting Alzheimer's disease

[Strittmatter & Roses, 1996] amongst others. Noting the success of these studies when applied in a targeted manner, alongside the growth of genetic sequencing capabilities, a clear need emerged for linked genotypic-phenotypic databases.

2.7.4. Databases of Genetic and Phenotypic Variation

Scientists have understood the benefits of collecting information on human phenotypic variation at scale for a long time, evidenced by Galton's records as above. Sequencing technologies now enable the collection of genetic variation information at the genome-wide level as well. A number of databases have been constructed since the turn of the 21st century, some genetic, some phenotypic, and some with coupled data.

In 2002, the international HapMap Project was launched with the goal of mapping variation in individuals from distinct ethnic lineages. It aimed to study inter-human variation at the base pair level by measuring frequencies of single nucleotide polymorphisms (SNP), with the long-term goal of better understanding disease. Taking advantage of haplotypes, groups of SNPs inherited together, allowed for genotyping of relatively few SNPs whilst retaining a rich data resource [Frazer & al., 2007].

Multiple databases were established for studying disease. The Wellcome Trust Case Control Consortium was a collaborative project established in 2005 to study seven core diseases: type 1 diabetes, type 2 diabetes, coronary heart disease, hypertension, bipolar disorder, rheumatoid arthritis and Crohn's disease. For each of these seven diseases, 2,000 individuals were genotyped in addition to a 3,000-participant control group, and the subsequent data used in GWAS to find disease associated variants [Burton & al., 2007]. Closely paralleled, the Cancer Genome Atlas was designed as a database of genetic sequences of 20,000 cancer samples coupled to control samples. The project began in 2006 and looked to discover novel genetic variants influencing cancer [McLendon & al., 2008]. The UK10K was launched in 2010 and aimed to sequence the genomes of 10,000 individuals with rare diseases to provide a resource for studying the genetic impact on these disorders [Walter & al., 2015].

The UK Biobank was developed to combine broad phenotypic and genotypic data allowing for the study of general and pathological biology. It is a database containing data collected from 2006 onwards from ~500,000 participants. The project aimed to collect broad data at a national scale covering lifestyle, anthropometrics, and basic clinical measures. For each of the participants, biological samples were taken and genotyping was conducted, with exome and whole genome sequencing to be completed in the near future [Sudlow & al., 2015]. For more information see section 3.1.

2.7.5. Genome-wide Association Studies

The growth of these databases documenting common human variation allowed for associations between common variation and complex disease. The statistical framework behind GWAS is relatively simple, with many GWAS relying on an additive linear model. Statistical significance is evaluated with a chi-square test. The first GWAS was published in 2005 [Klein & al., 2005] and identified two loci associated with AMD using a cohort of 146 individuals.

Since the first GWAS, the technique has been central to the discovery of genetic loci involved in a large number of common and complex diseases, including type 2 diabetes [Xue & al., 2018], schizophrenia [Ripke & al., 2014] and Crohn's disease [Franke & al., 2010]. The use of GWAS has also shown great success in eye pathology, with discovery of SNPs associated with diabetic retinopathy [Huang & al., 2011; Sheu & al., 2013; Awata & al., 2014; Burdon & al., 2015; Liu & al., 2019], and glaucoma [Choquet & al., 2018].

GWAS is also used to look for genetic variation underlying general biological variation. The Genetic Investigation of Anthropometric Traits (GIANT) consortium investigated the genetic loci that affect human anthropometrics including height, body mass index (BMI) and body shape [Speliotes & al., 2010; Allen & al., 2010; Heid & al., 2010].

The combined success in pathology and general biology means that GWAS remains one of the key tools used to progress the study of genetics.

2.7.6. Population Structure in Genome-wide Association Studies

Variation in phenotypes is affected by a number of confounders in addition to genetic variation. Increased use of association studies in human populations identified new confounders that had to be controlled for. One of the the largest confounders is population structure, which represents shared common-ancestry as well as shared environment and culture. Whereas in association studies in animals, such as those in agriculture, where you can leverage population structure to fit complex models with the knowledge that the environment has been randomised, the same assumptions can not be made in human studies.

Therefore when conducting GWAS of a trait, if population structure is not controlled for it will confound the results. Whereas other confounders can be controlled for through inclusion in the GWAS model as covariates, population structure is complex to disentangle. Several methods are used to regulate this confounder, including conducting GWAS on distinct well-mixed populations (see section 3.5.1), where it can be assumed that individuals share common ancestry and environment, or linear mixed

models can be used that account for this complex confounder.

Even with these precautions, it remains important to test that this population bias has been removed. Commonly, a quantile-quantile (qq)-plot is generated that plots the negative log of ranked p-values from the GWAS against the expected distribution. Traditionally, qq-plots are deemed "healthy" if the GWAS p-values largely correlate with the expected values, aside from a small proportion that have a higher than average negative log p-value. These present as an upward tick in the latter quartile of the data. However, this analysis is not calibrated to study complex polygenic traits at increasingly large-scales, for example using the UK Biobank. Here inflation in the observed negative log p-values may be caused by increased numbers of SNPs with low frequency in the population (Minor Allele Frequency (MAF)), as well as confounding population structure. Given this, it is necessary to be able to distinguish the underlying cause.

Linkage Disequilibrium Score (LDSC) regression aims to partition the traditional measure of inflation due to population stratification, *genomic control* [Devlin & Roeder, 1999]. The term Linkage disequilibrium (LD) was first described in Lewontin & Kojima, 1960 and is defined as a nonrandom association of alleles at two or more loci [Slatkin, 2008]. LDSC regression partitions genomic control into two parts: that caused by polygenicity and that caused by residual confounding population structure [Bulik-Sullivan & al., 2015]. To complete this, it utilises the concept of shared LD within population factions. Statistics commonly reported from LDSC regression include the *LDSC intercept*, which subtracted from 1 represents the confounding bias scaled to the cohort size. This means an intercept close to 1 is representative of limited residual population structure. Additionally the *ratio* is also reported, which represents the proportion of inflation caused by factors other than polygeneicity.

2.8. Mendelian Randomisation

Observational studies are used in many fields, from econometrics to epidemiology, to try and discover relationships between different variables. In order to identify such relationships, many variables are measured. However, inference of a relationship between variables is not simple. As well as knowing if a relationship exists, we want to know why it is so: correlation does not imply causation. Variables exist in complex and interconnected systems. Because of this complexity, variables can be confounded by one another. The confounding measures are not always easily identifiable, and even when identifiable, not always measurable. Therefore techniques have been developed to try and deal with this problem.

Instrumental variable analysis has been developed to try and mediate the issue of confounders. This technique aims to identify a causal relationship between two variables, known as the *exposure variable* and *outcome variable*. The *exposure variable* is that which we predict will be the cause of an outcome. The *outcome variable* is the observed variable we wish to explain.

In order to identify this causal relationship, instrumental variable analysis utilises an external variable believed to be unconfounded. Here unconfounded refers to a variable that is directly associated with the exposure variable, and singularly associated with the outcome variable through the exposure variable. This external variable is known as the *instrumental variable*.

This method has been applied in multiple fields including econometrics. An example of this is the use of the Vietnam war draft as an instrumental variable in analysing the effect of military service on lifetime earnings. In many circumstances, membership of the military is voluntary and thus confounded by many variables including geographic location and socioeconomic status. However, the draft lottery acts a instrumental variable as it randomised recruitment thus randomising these confounders [Angrist, 1990].

In the biomedical field, researchers often wish to identify causative relationships between exposure and outcome variables that relate to disease. Evidence of causality is necessary before pursuing further avenues of disease treatment or prevention. One way to do this is to use randomised controlled trials in place of observational studies. Here an exposure variable is randomly assigned to different cohorts, thus randomising their environment and minimising the effect of confounders. In most clinical trails, a drug and placebo is randomly assigned to trial participants. Thus, the confounders are randomised with respect to the treatment.

However the use of randomised controlled trials is not always feasible or ethical, and observational studies must be utilised instead. Mendelian Randomisation (MR) borrows from instrumental variable analysis to look for evidence of a causal relationship between an exposure variable and an outcome variable in epidemiological observational studies. First proposed in 1986 [Katan, 1986; Gray & Wheatley, 1991], this method utilises instrumental variable analysis to look for the relationship between two variables, often with one being disease. It utilises genotypes as the instrumental variables. It is named for its grounding in the concept developed by Mendel that germline genetics do not alter from conception and thus cannot be influenced during the disease process or by other environmental factors.

MR can be analogised through clinical trials. Whereas in clinical trials, the exposure variable is randomised between participants, MR uses genotypes as a proxy meas-

ure for the exposure variable. Due to the random distribution of genotypes across the population, groups with the same genotype should be randomised with respect to confounders. Therefore *genetic instruments*, genetic variants associated with the exposure variable, can be used to test for a causal relationship. It should be noted that despite the naming of the variables, MR is inherently non-directional, and a positive MR analysis with specified exposure and outcome variables does not denote a directional effect. This can be further tested for with the Steiger test [Hemani & al., 2017].

A classic example of the utility of MR is its use to verify the causative relationship between low density lipoprotein cholesterol (LDL-C) and the risk of coronary heart disease. Evidence for a causative relationship between elevated LDL-C and increased risk of coronary heart disease has been shown numerous times using MR [Burgess & Harshfield, 2016]. The genetic effect of SNPs associated with levels of LDL-C correlate with the genetic effect of the same SNPs on coronary heart disease. Using MR bypasses many confounders of the variables including lifestyle. The use of MR has been further applied to help interrogate potential therapeutic targets of coronary heart disease through the LDL-C pathway. Randomised controlled trials are expensive and timely and so using MR to establish support for a relationship prior to testing it in a clinical trial is incredibly useful.

In Ference & al., 2019, this technique was employed to look at the relationship between a proposed therapeutic intervention and LDL-C levels. The therapeutic intervention involved the inhibition using bempedoic acid of an enzyme within the cholesterol biosynthesis pathway, ATP citrate lyase. Genetic variants that mimic the inhibition of the enzyme were used as genetic instruments to find evidence supporting its causal relationship with LDL-C levels. This was compared to the effect of 3-hydroxy-3-methylglutaryl–coenzyme A reductase (HMGCR) inhibition using statins on LDL-C. Statins are widely used to reduce the risk of coronary heart disease by lowering LDL-C levels. Again genetic variants were identified that mimic the inhibition of HMGCR and used as genetic instruments in MR. This study concluded that MR showed similar size of effect of ATP citrate lyase inhibition and HMGCR inhibition on reduction of LDL-C levels and subsequent coronary events. This affirms the enzyme's potential as a therapeutic target and gives support to pursuit of bempedoic acid as a therapeutic agent. This example highlights the power of MR in biomedical research in directing further clinical research.

2.9. Thesis Outline

This thesis aims to explore the effects of genetics on the morphology of the human retina. The retina, as previously described, is an integral instrument in sight, and also a key part of the central nervous system (CNS). In understanding more about the genetics underlying its structure, I aim to further our understanding of its function and, conversely, dysfunction.

In chapter 3 I describe the data used for the analysis within this thesis. I detail both the genotypic and phenotypic data, including OCT images, available within the UK Biobank and the measures derived from these OCT images. I further outline the quality control metrics applied to this data, and describe the resulting dataset. The quality control methods described in this chapter are also presented in [Currant & al., 2020]. These methods were developed and implemented with input from Anthony Khawaja, a clinical colleague. I am an author on an epidemiology paper with clinical colleagues that further describes some of this data [Khawaja & al., 2020].

In chapter 4 I present the largest GWAS of inner retinal thickness to date. I explain the choice of phenotype used to describe inner retinal morphology and the choice of model. I then present the loci discovered to be significantly associated with inner retinal morphology. I further perform Mendelian randomisation analysis to explore the causative relationship between inner retinal disease and glaucoma. A manuscript detailing these results is currently under review and available as a pre-print [Currant & al., 2020].

Chapter 5 presents the largest GWAS of outer retinal morphology to date and the discovered genetic loci associated with the outer retinal layers. Here, I extend the analysis from the clinically-used phenotype of mean retinal layer thickness across the macula, to retinal thickness in the peripheral, intermediate and foveal fields of the macula. I further analyse the causal relationship between outer retinal thickness and age-related macular degeneration. I intend to compile the work in this chapter into another first author paper with my clinical colleagues upon submission of this thesis.

Finally, in chapter 6 I expand the dimensionality of the phenotype used to describe retinal morphology. I first explore morphological variation of the macula, and the genetic variation underlying it, at a higher dimensionality using the novel visualisation technique, maculagrams. I then capitalise on the rich pixel-wise data available in the raw OCT images, developing image analysis techniques that are able to extract retinal morphology measures at a finer grain. I then look at the effect of genetic variants on this granular representation of retinal morphology. The image analysis techniques presented in this chapter were developed in collaboration with Tomas Fitzgerald, and the ana-

lysis applied on the inner retina is also presented in a pre-print [Currant & al., 2020].

3. Data

The ever decreasing cost of genetic sequencing has allowed the use of genomics within medical practice to increase across recent years. This growth, alongside increased digitisation of health records, means the benefits of personalised medicine have already shown to enhance clinical treatment with remaining unexplored potential [Abul-Husn & Kenny, 2019; Vogenberg & al., 2010].

As this clinical usage grows, further research on common genetic variation in populations and complex modelling of genetic variation underlying disease is required to maximise applicability in a clinical setting. Furthermore, to ensure that newly available clinical technologies are benefiting all facets of the population, research must be done on a representative sample of the population.

In order to facilitate such research, several databases have been, or are being, developed that provide rich, large-scale genetic and phenotypic data on cross-sections of the population (See section 2.7.4). These datasets can largely be divided into either clinical cohorts, where participants are selected for disease, or population cohorts that aim to be representative of the cross-section of the general population. The category of population cohorts can be further split into birth cohorts and adult recruitment cohorts. Birth cohorts recruit individuals from birth, thus removing some of the bias of self selected inclusion, however there is a long time period between initiation of the study and the time at which some of the data can be utilised. There may also be difficulty in follow-up measures with participant drop-out and logistical challenges of remaining in contact over long time periods. A number of birth cohorts, including the two Lothian birth-cohort studies started in 1921 and 1936 respectively [Deary & al., 2011], and the Avon Longitudinal Study of Parents and Children (ALSPAC) study created in the 1990s [Boyd & al., 2013; Northstone & al., 2019], have been utilised in many biomedical research projects. On the other hand, adult recruitment population cohorts have the advantage of a much shorter turn around time. However due to the nature of self-selection for inclusion, bias is introduced into the population. Furthermore, a portion of data collection is often retrospective, relying upon accuracy of self-reporting or the completeness of records. Examples of this include the Danish effort to genotype and

phenotype a large portion of their population [Thygesen & al., 2011] and the FinnGen study (https://www.finngen.fi/en) that combines multiple Finnish biobanks. The birth cohort and adult recruitment cohort approach can be considered complementary to one another in many ways, with both providing huge utility in academic research. Another approach is that seen in the INTERVAL study [Moore & al., 2014], in which a cohort initially recruited for a randomised control trial, is utilised to study broader variation. The INTERVAL study originally aimed to assess the safety of different blood donation intervals. Since then, the cohort has been genotyped, and the dataset used to study common variation affecting human blood cell traits [Astle & al., 2016]. Additionally exome sequencing of the cohort has allowed for its use as a control dataset in studies of the genetics of schizophrenia [Singh & al., 2016]. The expanded use of the INTERVAL dataset following collection of genetic data, exemplifies an efficient method to collect a population cohort.

The UK Biobank is an adult recruitment population cohort of unprecedented size. Initiated in 2006, it provides coupled phenotypic and genotypic data for a large-scale cohort that enables novel research into human variation and clinical traits [UK Biobank, 2007]. Its design goal was to study late onset disease by collecting longitudinal data on a middle-late aged population. The enormous size of the UK Biobank provides huge statistical power when performing analysis of the epidemiological and genetic data, and the richness of the phenotypic data opens up new avenues of research and allows for complex modelling.

In this chapter, I will give an overview of the UK Biobank, the cohort it comprises, and the different data types that are available within it. I will give a more detailed description of the data used within my analysis and how said data was generated. I will then describe the sources of variation in this data and which quality control methods are applied to create a dataset that is used throughout this thesis.

3.1. The UK Biobank

The UK Biobank is a database of genotypic and phenotypic information for ∼500,000 individuals from across the United Kingdom (UK). Participants were between 40 and 69 years of age at enrolment and were recruited via the National Health Service (NHS) registry. The database contains genetic, biological, medical, demographic and lifestyle data for each of the individuals, as well as imaging data for a subset of the cohort. Data were collected at one of 20 assessment centres across the UK, with some participants making repeat visits. The participants within the UK Biobank were not selected on the

basis of having disease [UK Biobank, 2007].

3.1.1. Baseline Data

Upon their first visit to the assessment centre, participants had a number of baseline measurements taken. This included anthropometric measures, for example height, weight and hip and waist measurements. Other indicators of general health and biology were taken, such as bio-impedance measurement and hand-grip strength. A heel bone ultrasound was performed to record density, spirometry was conducted to measure lung capacity and blood pressure was measured. All participants also completed a questionnaire and several cognitive function tests. Finally several biological samples, including blood and urine, were taken [UK Biobank, 2007].

3.1.2. Lifestyle Data

All UK Biobank participants were required to complete an extensive questionnaire. This included information on exercise, alcohol consumption, smoking and diet. Additionally the database contains information on socio-economic status including income, education, vocation and identity. Participants were given the option to skip questions that they did not wish to answer.

3.1.3. Medical Data

In addition to providing self-reported medical history within the questionnaire, participants within the UK Biobank consented to the linkage of their medical records with the database. This enables access to the cancer and death registries, as well as medication history and disease diagnoses using the International Classification of Diseases (ICD)10 codes. There are future plans to link more extensive medical records with the database including primary care records.

3.1.4. Caveats

Due to the method of recruitment, it must be acknowledged that there is bias within the cohort. Several studies have shown that participants in the UK Biobank are on average of better overall baseline health than those who have not participated [Fry & al., 2017; Batty & al., 2020]. As with any study performed on the UK Biobank cohort, this must be considered when trying to generalise to a general population.

3.1.5. Ocular Examination

As part of the final round of UK Biobank participant recruitment, a subset of 132,041 participants also underwent a number of ocular examinations at one of six study centres around the UK (Sheffield, Liverpool, Birmingham, Croydon, Hounslow and Swansea). Three different measures were taken: best corrected visual acuity, using logarithm of the minimum angle of resolution (logMAR); refractive error, using a Tomey RC-5000 auto refkeratometer [Cumberland & Rahi, 2016]; and IOP using an Reichert Ocular Response Analyzer [Keane & al., 2016], from which corneal compensated IOP was calculated that accounts for rigidity of the cornea [Khawaja & al., 2018]. The three measurements, as well as OCT and retinal fundus photography, were completed within an ∼11 minute time span [Keane & al., 2016] so were performed on undilated eyes as not all tests are compatible with dilation [Khawaja & al., 2018].

3.2. Optical Coherence Tomography

A further subset of the UK Biobank cohort, consisting of 67,321 individuals, underwent Spectral Domain Optical Coherence Tomography (SDOCT) imaging. The imaging was conducted using the Topcon 3D OCT1000 Mark II machine using the 3-dimensional 6x6-mm² macular volume scan mode (512 A scans per B scan; 128 horizontal B scans). The imaging was conducted in a dark room, following other optical measurements, on undilated eyes. The right eye was consistently imaged first, and the majority of individuals had imaging repeated in their left eye. The images are stored in the common medical image format, dicom as well as .fda and .png files [Patel & al., 2016].

3.2.1. Topcon Image Processing

The manufacturers of the OCT device, Topcon, provide an industry standard segmentation software, Topcon Advanced Boundary Segmentation (TABS). TABS automatically segments either five or eight retinal layers, dependent on confidence requirements. It can segment RNFL, Ganglion Cell Inner Plexiform Layer (GCIPL) (GCL and IPL in combination), ONL, OS and RPE at a high confidence (figure 3.1A), and RNFL, GCL, IPL, INL, ONL, IS, OS and RPE with a less high confidence level. From this segmentation, the thickness of the different retinal layers is calculated, either as a mean taken across the entire macular field, or per segment of one of two segmentation grids. The two grids are known as the Early Treatment Diabetic Retinopathy Study (ETDRS) grid and the Macula 6 grid. The ETDRS grid is widely used when analysing the outer retina,

and consists of a bullseye shaped grid with 9 sub-segments (figure 3.1B) [Ooto & al., 2015]. The Macula 6 grid is conversely used to analyse the inner retina. It is made up of six equally sized segments arranged radially around the fovea, with the centre of the grid excluded (figure 3.1C) [Chaglasian & al., 2018]. Glaucoma is the main disease that affects the inner retina, and for this, the line segregating the superior (upper) segment from the inferior (lower) segment of the retina is important. This is because visual defects symptomatic of glaucoma usually affect the superior and inferior field of the retina differentially [Gupta & Asrani, 2016].

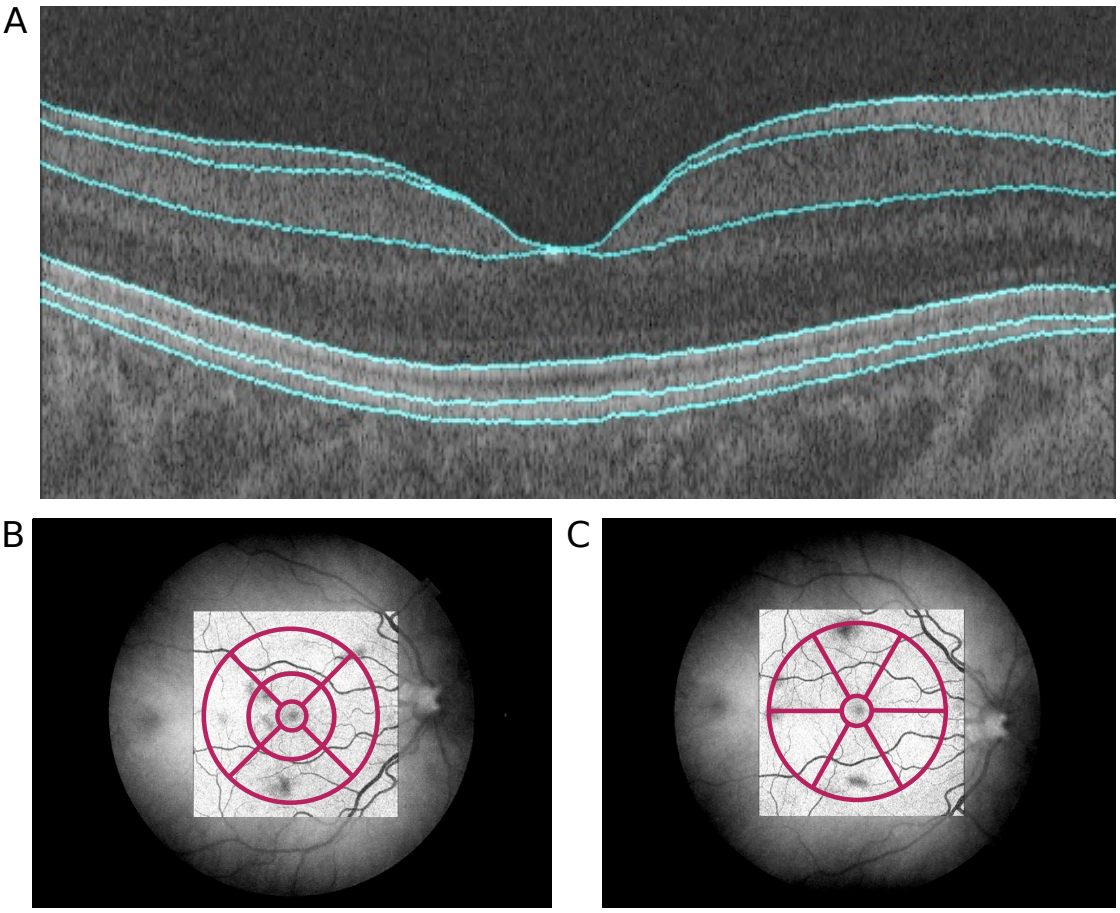

Figure 3.1: Optical coherence tomography images and Topcon advanced boundary segmentation processing. Topcon's propriety software, TABS, segments the different retinal layers. These segmentations are used to calculate the thickness of each layer as a mean across the whole macular field, or as means of sub-fields defined by one of two segmentation grids. A) An example of an OCT cross-section with segmented boundaries (higher confidence level) shown in blue. B and C show a fundus image, with the Early Treatment Diabetic Retinopathy Study (B) and Macula 6 (C) grids superimposed, showing the boundaries of each segmentation scheme.

3.3. Image Quality Control

There are several sources of potential variation and error in segmented thickness values derived from the OCT images. Some of these sources of error can be attributed to technical and operator-related error occurring during image acquisition or processing. Other artefacts can be attributed to biological factors such as disease.

Examples of potential sources of technical artefacts include poor centration of the eye within the segmentation grid, especially important when looking at the separate segments of the ETDRS or Macula 6 grids. Movement of the eye or blinking may also introduce artefact into the image leading to blurred images or missing portions of the scan. Occasionally, the image is clipped at the edge, or shadows may be present in the image causing low signal. All of these technical issues may interfere with the automated segmentation of the image, leading to incorrect thickness values [Liu & al., 2015].

Equally, there are a number of biological factors that may affect the quality of the scan. Presence of fluid, drusen (associated with AMD, see section 1.6.2), or significant disruption to the retinal morphology may cause disruption to the scan or issues with automated segmentation again causing inaccurate retinal thickness values. Dry or cataract eyes may cause low signal due to disruption of the OCT beam reaching the nerve tissue. Equally, blood vessels can sometimes present as artefacts as they traverse the choroid.

For these reasons, a quality control scheme is applied to the image data. The quality control method is adapted from the methods used in [Patel & al., 2016]. The quality control protocol is based on a number of parameters outputted from the OCT device during the scan, as well as the segmented layer thickness's themselves. Initially, all images which received an overall image quality score lower than 45 were removed. Following that the poorest 20% of images in several image segmentation indicators were removed from the dataset. Those image segmentation indicators were: valid count, inner limiting membrane (ILM) indicator, minimum motion correlation, maximum motion delta and maximum motion factor. Valid count identifies scans with a substantial degree of clipping in the OCT scan's z-axis. ILM indicator is a measure of the minimum localised edge strength around the ILM boundary across all B scans. Minimum motion correlation, maximum motion delta and maximum motion factor are a group of motion indicators. Minimum motion correlation is the minimum value of Pearson's r when correlation is performed between the RNFL thickness and the the full retinal thickness of consecutive B-scans. Max motion delta is the maximum absolute difference in RNFL thickness and overall retinal thickness between consecutive B-scans. Both the ILM in-

dicator and the motion indicators are indicative of blinks, motion in the scan, severe signal fading and segmentation errors. Participants with outlier values of refractive error, defined as outside ±1.5 ×the Interquartile Range (IQR) from the median of the data, were removed from the dataset. The demographics of those that pass and fail the image quality control are comparable and are detailed in table 3.1. To assess the image quality control methods, I examined the principal component (PC)s of the component ETDRS retinal thickness measurements for the cohort before and after quality control was applied (figure 3.2). As can be seen in figure 3.2A, prior to quality control filtering there is some structure to the data, with a distinct group seen on the right-hand side of the plot that is removed by the quality control process. These participants could be representative of a group with image artefacts, a group in which the image segmentation had error, or a group with a certain pathology causing extreme values. It was decided that for the analysis presented in this thesis in which I wanted to explore general biology and variation in a healthy population, that this group should be removed via the image quality control. Following image quality control, the data largely groups into a single blob as seen in figure 3.2B, however a few outliers remain, with several in the upper right section showing subtle signs of structure. It was decided that for now, the removal of structure by the image quality metrics described was sufficient, and that any outliers identified in the PC analysis could warrant further exploration in the future.

Table 3.1: Demographics of quality control subsets Comparison of biological, particularly optical, characteristics between all UK Biobank participants with optical coherence tomography data, the group that passes quality control and the group that fails. Results are presented as mean ±standard deviation.

	Total Topcon Population	**Pass Filter**	**Fail Filter**
N	67,310	31,434	35,876
Height (m)	168.68 ±9.25	169.43 ±9.15	168.03 ±9.29
Age (years)	57 ±8	57 ±8	57 ±8
Weight	78.12 ±16.02	78.48 ±15.83	77.81 ±16.18
Sex (m/f)	29,713/34,929	14,837/16,597	14,876/18,332
Refractive error left	-0.32 ±2.73	0.23 ±1.49	-0.81 ±3.42
Refractive error right	-0.38 ±2.73	0.19 ±1.48	-0.88 ±3.41
logMAR left	0.02 ±0.21	-0.00 ±0.18	0.05 ±0.23
logMAR right	0.03 ±0.21	0.00 ±0.18	0.05 ±0.22
IOP left	15.87 ±4.30	15.84 ±4.14	15.90 ±4.44
IOP right	15.99 ±4.26	15.97 ±4.12	16.02 ±4.38

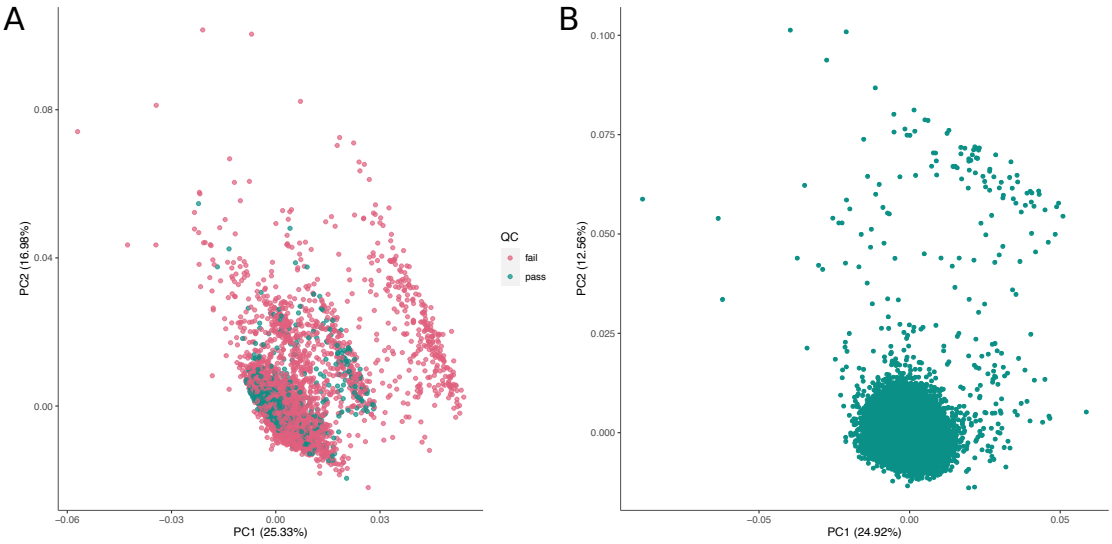

Figure 3.2: Principal component analysis of retinal thickness in filtered and unfiltered populations. The first two principal components from a principal component analysis (PCA) of the ETDRS grid, high confidence segmentation layer thickness performed on A) The full cohort with optical coherence tomography segmentation data, coloured by whether an individuals scan passes or fail our quality control. B) Only individuals that pass our quality control.

3.4. Genetic Data

There are three types of genetic data currently available, or soon to be available, within the UK Biobank: genome-wide genotyping data, the first genetic data type to be released; whole exome sequencing (WES), the second release; and whole genome sequencing (WGS), the final phase of genetic sequencing. Here I will describe the genotype data which is used within the analysis presented in this thesis.

3.4.1. Genotyping Data

Genotyping of the UK Biobank was conducted over a period of 18 months using genetic material extracted from blood samples provided at the initial assessment centre visit. The genotype data within the UK Biobank were generated using two closely-related custom arrays. The first array, the Applied Biosystems™ UK BiLEVE Axiom™ Array by Affymetrix, included 807,411 markers and was used to genotype a subset of 49,950 individuals who were part of the UK Biobank Lung Exome Variant Evaluation (BiLEVE). The remaining 438,427 participants were genotyped for 825,927 markers using the Applied Biosystems™ UK Biobank Axiom™ Array. In total, 95% of markers were shared across the two subset arrays and included both SNPs and indels, some of which had

been selected for their known associations with particular diseases or phenotypes. Affymetrix performed various quality control methods resulting in the return of genotyping data for 489,212 samples across 812,428 markers. Further quality control and imputation were performed as described in detail in Bycroft & al., 2018, and resulting in genotype information for $\sim$ 96 million genetic variants which were returned to the UK Biobank.

3.5. Preparing Genotype Data

3.5.1. Identifying a Well-Mixed Population

Population structure based on mixing of population groups, which often falls along similar boundaries to ethnic background, can confound genetic analyses. One method to control for this is to perform the genetic analysis on a single, well-mixed population. Within the UK Biobank, due to the fact that 94% of the cohort self-reports a white ethnic background [Bycroft & al., 2018], this dense well-mixed sub-population is largely representative of those that self report as white. We select for said well-mixed population based on PCs of the genotype data, comparing them to the PCs of the HapMap dataset whose participants have known ethnic backgrounds.

Methods

The original genotype files, provided in the bgen format by UK Biobank, were converted to PLINK format (bim, bed and fam files) using PLINK [Chang & al., 2015], and the subset (n=48,711) which have quality controlled processed OCT data available selected. The genotypes for this subset were then filtered to select for markers with a MAF > 0.001. The genotypes are further pruned based on LD (window size = 50kb, step size = 1 variant count, r^2 threshold = 0.8). Duplicate SNPs based on either rsid or position were removed and the resulting pruned chromosome files merged into a single genome-wide genotype file. This genotype file was further pruned to MAF > 0.1. The cross-section of SNPs available in our pruned UK Biobank genotype dataset, and the HapMap dataset was identified and selected for in each dataset. Duplicate markers were removed, and the two datasets merged. A PCA was performed on the merged genotype dataset. The HapMap data contains individual level data of ethnic background, dividing the cohort into sub populations representing: African ancestry from the South West United States of America (USA) (ASW); Luhya in Webuye, Kenya(LWK); Yoruba in Ibadan, Nigeria (YRI); Maasai in Kinyawa, Kenya (MKK); Caucasians from North

and western Europe (CEU); Toscani in Italy (TSI); Gujarati Indians in Houston, USA (GIH); Mexican ancestry in Los Angeles, USA (MXL); Han Chinese in Beijing, China (CHB); Chinese in metropolitan Denver, USA (CHD); and Japanese in Tokyo, Japan (JPT). These populations are further grouped to represent four broader groups of ethnic background. The four groups are: ASW, LWK, YRI and MKK (African ancestry); GIH and MXL (Mexican/Native American ancestry); CHB, CHD and JPT (Asian ancestry); and CEU and TSI (European ancestry). CEU and TSI jointly are representative of a European ancestry. The mean value of PC1 and PC2 for the combined TSI and CEU populations within the Hapmap data were calculated. The Euclidean distance of each TSI or CEU individual within the Hapmap data to the mean in PC space was calculated and the maximum value identified. The Euclidean distance of each of the UK Biobank participants in PC space to the mean of the HapMap European population is calculated. UK Biobank participants are selected for inclusion if their Euclidean distance to the mean PC1/PC2 point is $< 1.5 \times$the maximum Euclidean distance of HapMap Europeans to the mean (figure 3.3).

3.5.2. Identifying an Unrelated Population

Another source of confounding within the genetic data is family structure. As both genetic variation and environmental variation are often inherited within families, it is important that only single members from genetically related family units are included in the study. Within the UK Biobank there are several family structures present that it is necessary to remove.

Methods

Pairwise relatedness estimates for all individuals within the UK Biobank are available for download from the UK Biobank database, generated by Bycroft & al., 2018. The resulting table contains: an IBS0 score, for paired genotypes the proportion of SNPs where one sample is homozygous for the reference allele, and the other sample is homozygous for the alternative allele; the kinship coefficient for each pair of samples; and the HetHet value, the proportion of markers which are heterozygous in both samples of the pair. A filter adapted from Meyer, 2020 was used to remove individuals related to third-degree or more whilst maintaining the maximum number of samples in the cohort.

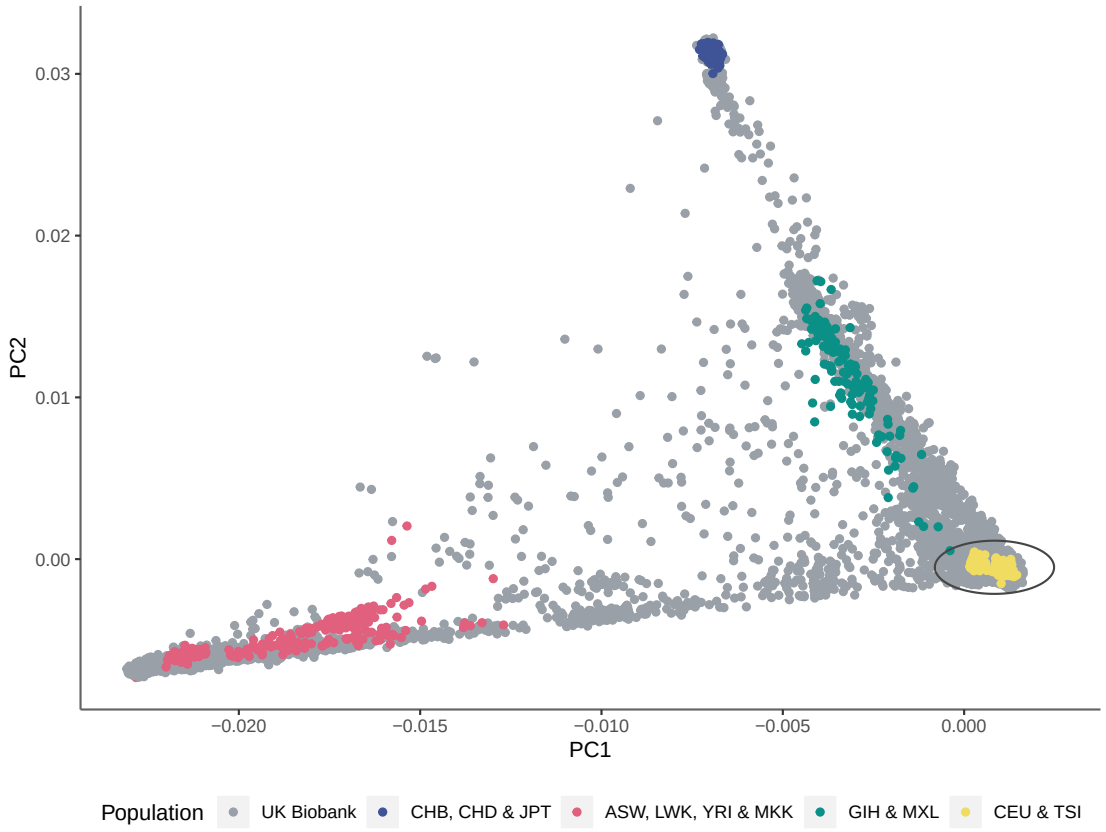

Figure 3.3: Population selection using principal component analysis. PC1 and PC2 of a PCA performed on the UK Biobank and HapMap genotype data. HapMap participants are coloured by the population to which they belong. An ellipse encloses the well-mixed population selected for inclusion in my study. This ellipse has a radius of 1.5 ×the maximum Euclidean distance of HapMap Europeans to the mean of the European population in PC space (shown in yellow).

3.5.3. Final Dataset

Following both phenotypic and genotypic quality control, we are left with an unrelated, genetically well-mixed population, with which I feel confident in proceeding to further analyses. There may still be some regional population structure, which I aim to address using covariates within the model.

4. Genome-wide Association Study of Inner Retinal Morphology

The inner retina is made up of ganglion cells and is responsible for the final stages of transmission before the neuronal signal exits the eye towards the brain via the optic nerve [Purves & al., 2001]. As previously mentioned, the distinction of the inner retina from the outer retina is common in the clinic. Each of the layers is affected by different diseases, as well as being differentially affected by a set of common diseases, with some pathologies affecting the two retinal areas similarly [Müller & al., 2019]. Here I analyse the genetic variation underlying inner retina morphology.

The inner retina is located closest to the front of the eye (figure 4.1A). It is comprised of the RNFL and the GCIPL. In combination these layers are known as the Ganglion Cell Complex (GCC). One well-characterised aspect of inner retinal morphology is that both inner retinal layers taper to non-existence at the fovea, the centre of the macula (figure 4.1B) [Khurana, 2007]. This tapered structure reduces interference of these layers, leading to reduced light scattering. This morphological characteristic, alongside the increased density of cone cells, leads to the fovea being the area of highest visual acuity [James & al., 2003].

It is well established that glaucoma causes damage to the inner retina. This is commonly seen as thinning of the inner retinal layers in clinical settings [Tatham & Medeiros, 2017; Bhagat & al., 2014]. Monitoring changes in the thickness of the inner retina is a key part of glaucoma diagnosis and management. More recently, inner retinal thickness has been recognised as a biomarker for glaucoma, particularly POAG [Hood, 2017; Shin & al., 2017]. The genetic variation underlying inner retinal morphology has not been studied at large-scale previously, owing to a lack of data availability, especially in non-disease based cohorts.

In this chapter, utilising measures derived from the OCT images now available in the UK Biobank, I present the first large-scale GWAS of quantitative inner retinal thickness. Before this, I will present an outline of the analysis performed to choose both the model and the phenotype used in the final GWAS.

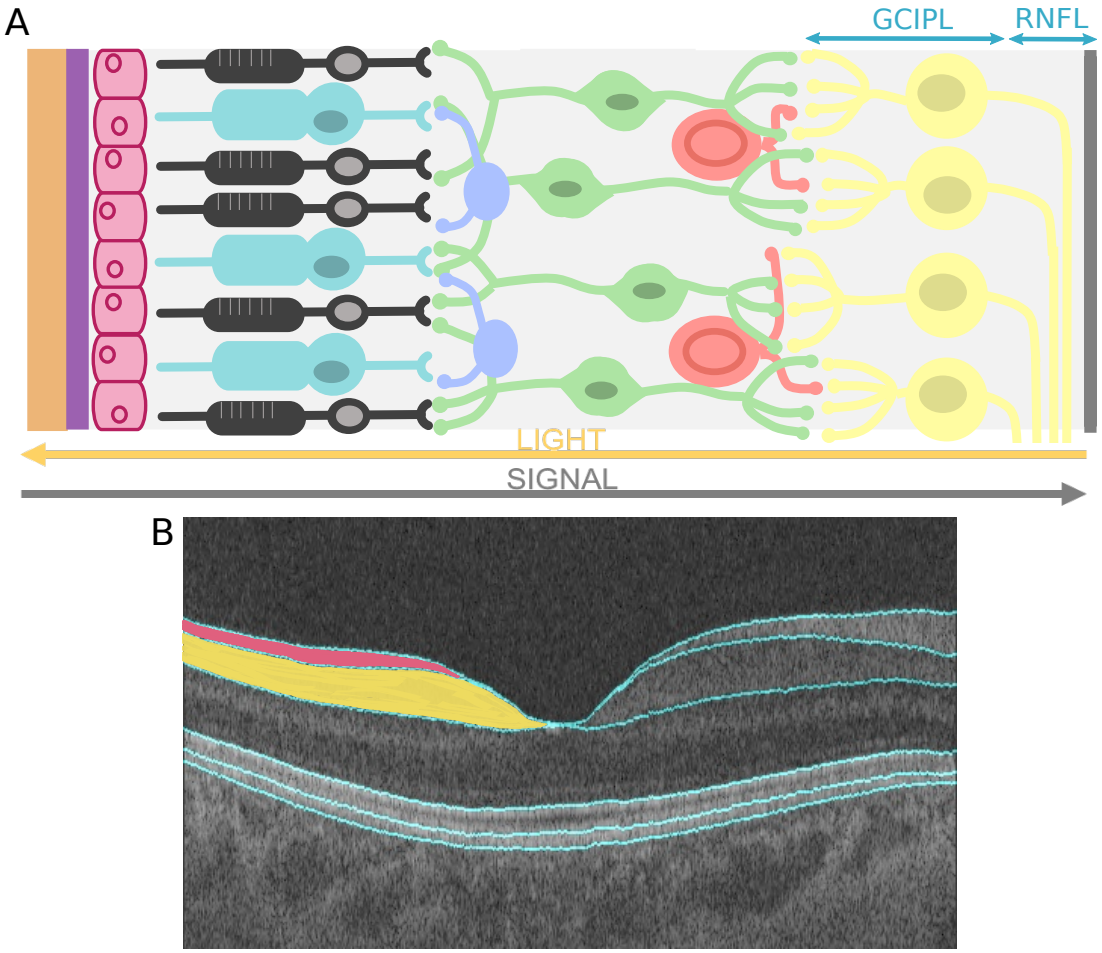

Figure 4.1: Inner retinal morphology and extracted measures. A) A diagram of the retinal layers. The inner retina is comprised of the retinal nerve fibre layer (RNFL) and ganglion cell inner plexiform layer (GCIPL).This is the first layer that light passes through before being absorbed by the outer retina, but is the final layer of cells involved in neural signal transmission before signal exits towards the brain via the optic nerve. B) An optical coherence tomography (OCT) image that has been segmented by the Topcon Advanced Boundary Segmentation (TABS). On the left half of the image the RNFL is shaded pink, and the GCIPL is shaded yellow. Taken together these retinal layers are known as the ganglion cell complex (GCC).

4.1. Inner Retina Data

The UK Biobank contains OCT images from which measures of RNFL and GCIPL thickness have been extracted using the industry standard software TABS. Measurements are provided for right and left eyes separately. As well as the mean thickness of the inner retinal layers across the macular field, thickness of subfields, as defined by the ETDRS [Ooto & al., 2015] and Macula 6 grids [Chaglasian & al., 2018], are available. Of the two, the Macula 6 grid is the one more commonly used to study the inner retina (figure 4.2A). It is particularly used when studying glaucoma. It differs from the ETDRS grid in that it excludes the central area over the fovea, as the inner retina does not extend into this area, and also divides the macular field across the horizontal midline. This is important as it is well established that glaucoma affects the superior and inferior segments of the retina differentially, and this feature is used in diagnosis [Gupta & Asrani, 2016]. Due to its clinical utility when studying the inner retina, measures derived from the Macula 6 grid are used throughout this chapter. Using the Macula 6 grid there are a total of 28 available measures to describe inner retinal morphology (figure 4.2B). These are the thicknesses of the RNFL and GCIPL at each of the six Macula 6 grid segments for both eyes, alongside the mean thickness of each layer for each eye.

4.1.1. Quality Control

To select the cohort of individuals for use in the inner retinal analysis, the quality control methods, both genetic and phenotypic, described in chapter 3 are applied. When applied specifically to the inner retinal data, this produces a final cohort of 31,434 participants that are used in the inner retinal GWAS (figure 4.2C). As previously mentioned in section 3.3, the quality control methodology removes structured groups as identified using PCA of the retinal thickness measures. To reassure ourselves that a group of participants had not been removed in the quality control process that were representative of notable biological variation, I performed a GWAS on the status of whether a participant passed or failed the quality control criteria. The GWAS was implemented using SNPTEST [Marchini & Howie, 2010] and used both biological and technical covariates as described below (section 4.2). This GWAS yielded no significantly associated loci ($P < 5 \times 10^{-8}$, figure 4.3). This reassured us that variation seen in future GWAS were not an artefact of the quality control process. A structured group with shared common variation and biological importance had not been removed. Therefore, the genomic studies of the inner retina going forward use the quality controlled cohort consisting of 31,434 participants.

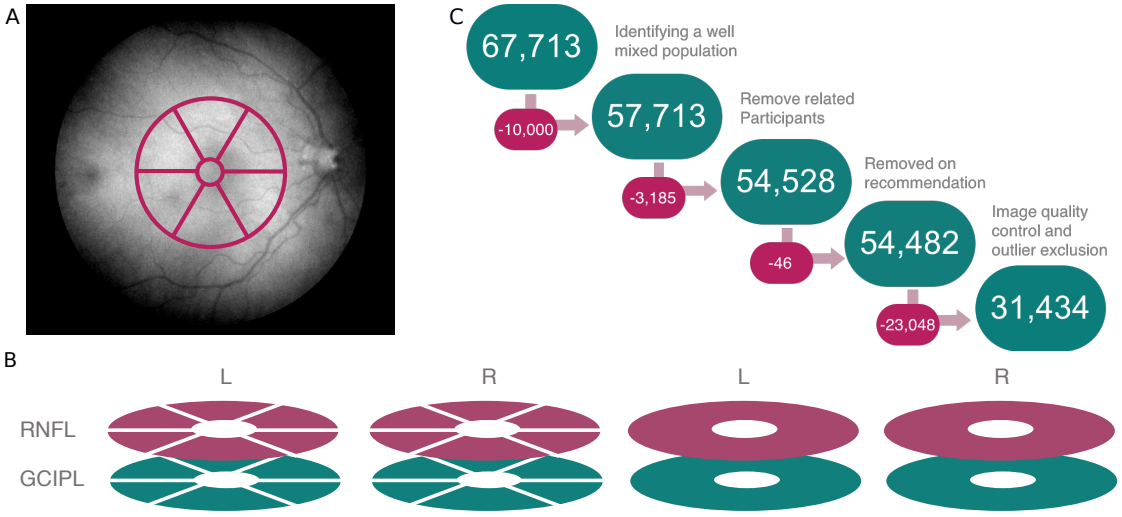

Figure 4.2: Inner retinal data dimensions and quality control. A) An example fundus image of a retina, with the Macula 6 grid, a commonly used partition matrix, superimposed atop in pink. The matrix is comprised of six sections, with the central field being excluded from analysis. B) A schematic of the 28 available measures that describe the inner retinal morphology. This includes the thickness of the RNFL and GCIPL at each of the six segments of the Macula 6 grid, as well as the thickness averaged across the Macula 6 grid, for left and right eyes separately. C) A schematic of quality control stages in selecting a cohort of individuals to include in inner retinal analysis based on criteria detailed in chapter 3.

4.2. Covariates

Before performing any GWAS, it was important to thoroughly consider necessary covariates to include in the models of genetic variation. These were likely a combination of both biological and technical variables. A number of technical variables are outputted from the OCT machine, described in detail in chapter 3. The OCT machine ID was also considered as this would account for any batch-like effect caused by different machines, examination set-ups and regional differences.

In terms of biological covariates, there are several phenotypes with known association to inner retinal thickness. Both sex and age [Song & al., 2010; Demirkaya & al., 2013; Nieves-Moreno & al., 2018] have been previously linked to inner retinal thickness, with increased age being a major risk factor for thinning of the inner retina, and development of POAG [Leske & al., 2003; Choquet & al., 2018]. The size of the eye was not directly measured as part of UK Biobank phenotype acquisition, so height was considered as a proxy measure. Weight was also considered given its correlation with health and lifestyle variables. Finally, differences in refractive error may alter the magnification of the OCT images, and so may affect the thickness measures of the inner retina [Hirasawa

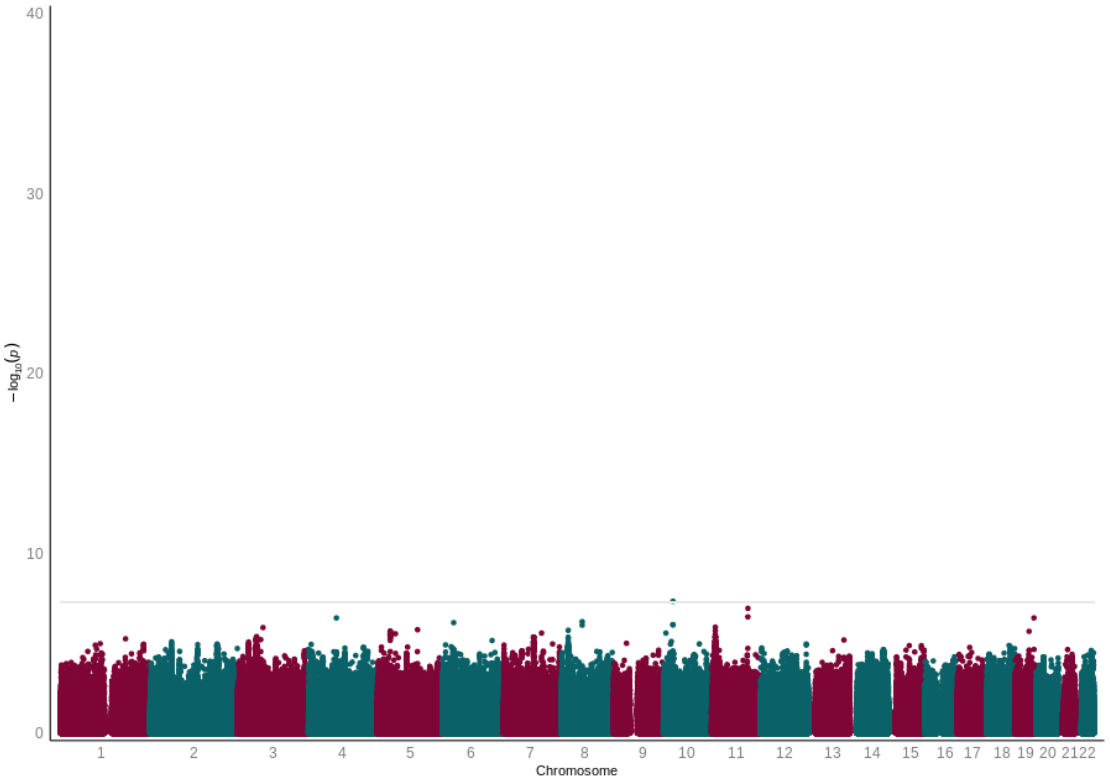

Figure 4.3: Quality control pass-fail status genome-wide association study. Manhattan plot for a genome-wide association study on the status of whether an individual passes or fails our quality control metrics. Variants were considered significant if they reached genome-wide significance ($P < 5 \times 10^{-8}$).

& al., 2014]. Refractive error comprises several conditions, the most common of which is myopia (near-sightedness) [Hysi & al., 2020]. Although not directly measured within the UK Biobank, it can be calculated from spherical and cylindrical power which were recorded [Taylor & al., 2011].

$$refractive\,error = spherical\,error + (0.5 \times cylinderical\,error)$$

The inner retinal phenotypes, and their derivative PCs, were used in correlations with the technical and biological factors mentioned (figure 4.4). Correlation was seen between both the raw and dimensionality transformed phenotypes and many of the covariates. In particular, several PCs appear to be correlated with some of the covariates (figure 4.4B). This suggests these covariates are considerably affecting variation in the inner retinal thickness measurements. Notably, PC4 has a strong correlation with macula center frame, the measure of how central the macula is within the image frame. PC2 is also associated to a milder extent with macula center frame, whilst PC3 has some correlation with mean refractive error and PC4 is additionally correlated with macula center aline. This is in-keeping with the expected effect of these cofactors, and as such, all of the described variables are used as covariates in future models of genetic variation on inner retinal morphology.

4.3. Eye-Specific Analysis

As previously stated, there are a number of different variables available to describe the morphology of the inner retina (figure 4.2B). Therefore we initially explored some of these phenotypes, aiming for a final phenotype that balanced genetic power and interpretability. A complexity of studying eye biology is that humans possess two eyes which, although often sharing in considerable biological symmetry, should not be assumed to be identical. This can be evidenced by the two eyes often having different visual acuity, and the occurrence of ocular dominance, the preference of visual input from one eye over the other [Carey & Hutchinson, 2013]. Equally, while glaucoma generally affects both eyes, it often presents asymmetrically [Gupta & Asrani, 2016]. Therefore, I have initially explored the left and right eyes separately.

4.3.1. Methods

GWAS were performed for the mean thickness across the Macula 6 grid of the RNFL and GCIPL for the right and left eyes separately. The GWAS were implemented in

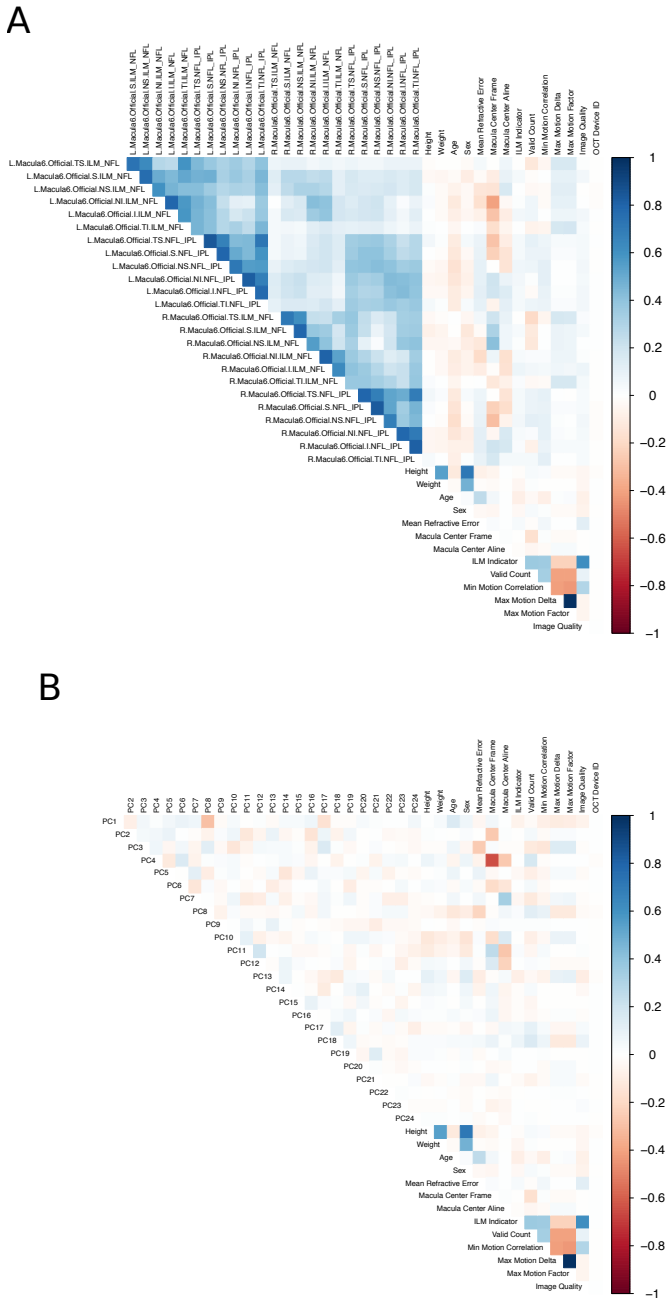

Figure 4.4: Correlation plots of inner retinal phenotypes with covariates. Correlation plots of inner retinal phenotypes with a number of technical and biological factors. A) Correlation between the thickness of both inner retinal layers at the Macula 6 subfields in the left and right eye, and covariates. B) Correlation between PCs derived from the phenotypes in A and covariates. The strength of correlation is represented by saturation of colour.

BGENIE [Bycroft & al., 2018] using an additive linear model. The covariates described above were included in the model. Loci were considered significantly associated if they reached genome-wide significance ($P < 5 \times 10^{-8}$).

4.3.2. Results

The four GWAS, one for the thickness of each of the two retinal layers in each eye, yielded a number of significantly associated genetic variants (figure 4.5). GWAS of the RNFL thickness in the left and right eyes discovered 1,927 and 1,715 significantly associated genetic variants respectively. GWAS of the GCIPL thickness in the left and right eyes discovered 1,024 and 1,090 significantly associated genetic variants respectively. In both cases, filtering of SNPs based on LD to identify lead variants in loci was not performed because at this stage the goal is still phenotype exploration. In both layers, many of the significant loci were shared across left and right eyes.

Upon observation that many of the significantly associated loci were common to both eyes, I examined the correlation of their effect sizes between the left and right eye (figure 4.6). The effect size of SNPs significantly associated with the mean thickness of the GCIPL across the Macula 6 grid was highly correlated between left and right eyes (Pearson's $r = 0.956$, $P < 2.2 \times 10^{-16}$). The effect size of SNPs significantly associated with RNFL thickness is correlated between left and right eyes, but to a milder extent (Pearson's $r = 0.743$, $P < 2.2 \times 10^{-16}$). It is apparent in the RNFL (figure 4.6A) that there is a group of SNPs that are less correlated between left and right eyes. As mentioned in section 3.2, the scan order of left and right eyes was not randomised, with the right eye consistently imaged first. It is therefore difficult to decipher if the discrepancy in genetic effect between the two eyes is representative of underlying biology specific to the RNFL, or if it can be accounted for by the lack of randomisation in the data acquisition protocol. This discrepancy warrants further investigation, and can hopefully inform future protocol design. However, for the purposes of the analysis within this thesis, a mean across the two eyes was taken to generate the final selected phenotype.

4.4. Dimensionality Reduction Analysis

The data available that describes inner retinal morphology using the clinical segmentation grids is relatively high dimensional. In total 28 measurements are available to describe the morphology of the inner retinal layers (figure 4.2B). Analysing the genetics underlying retinal morphology using the variation in each of these individual measurements is unlikely to be easy to interpret. These dimensions also may not best represent

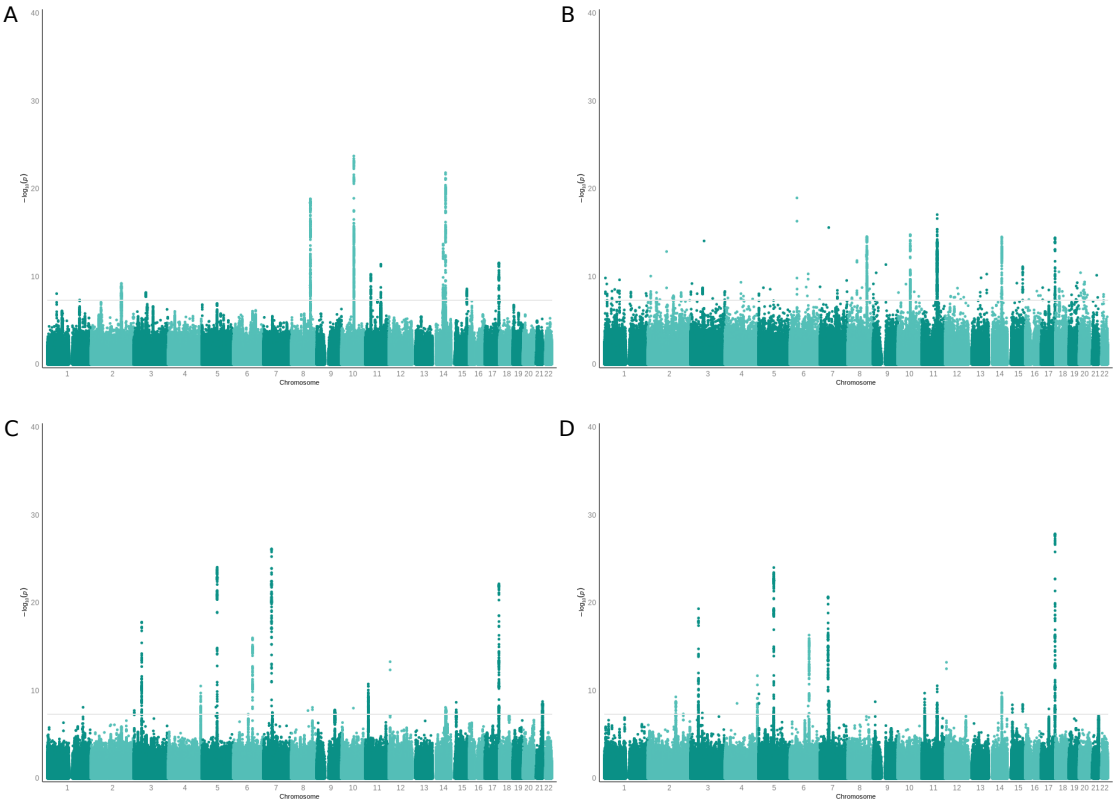

Figure 4.5: Genome-wide association study of inner retinal thickness phenotypes in left and right eyes. Manhattan plots of the GWAS results of mean macular inner retinal thickness phenotypes: A) retinal nerve fibre layer (RNFL) thickness in the right eye. B) RNFL thickness in the left eye. C) ganglion cell inner plexiform layer (GCIPL) thickness in the right eye. D) GCIPL thickness in the left eye. Variants were considered significant if they reached genome-wide significance ($P < 5 \times 10^{-8}$).

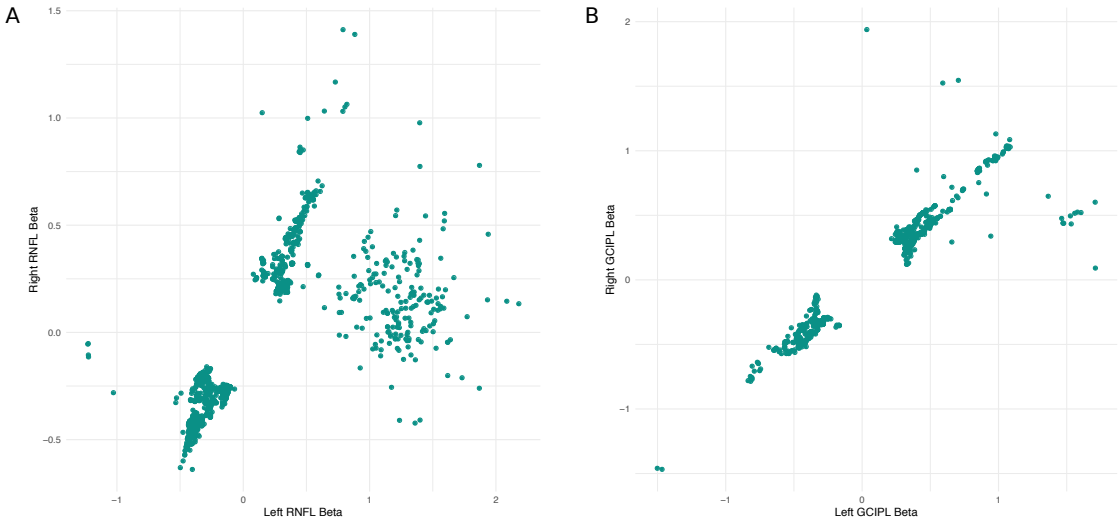

Figure 4.6: Comparison of effect size from GWAS of inner retinal thickness in left and right eyes separately A correlation plot of effect size of SNPs found significantly associated with the inner retinal thickness in either the left or right eye for either: A) Retinal nerve fibre layer (RNFL), Pearson's r = 0.743, P <2.2 $\times 10^{-16}$. B) Ganglion cell complex (GCIPL), Pearson's r = 0.956, P <2.2 $\times 10^{-16}$.

the continuous morphological variation occurring within the retina. However, I wanted to explore the high dimensional data in order to take advantage of all data available. To do this I performed PCA on the high dimensional data available within the clinical segmentation grids. These data comprises of the thickness of the RNFL and GCIPL at each of the segments of the Macula 6 grid, considering the left and right eye separately (figure 4.2B). This resulted in 24 PCs, as the mean values across the Macula 6 grid were excluded, with the first 10 PCs explaining 90% of variation (figure 4.7).

4.4.1. Principal Component GWAS

Observing the variance explained by the PCs, a GWAS was performed for each of the first 10 PCs. An additive linear model was used, implemented in BGENIE [Bycroft & al., 2018]. The same biological and technical covariates as detailed above (section 4.2) were used in the model. The resulting summary statistics were meta-analysed using Multi-Trait Analysis of GWAS (MTAG) [Turley & al., 2018] to create a single discovery list. Variants were considered significant if they reached genome-wide significance (P <5 $\times 10^{-8}$).

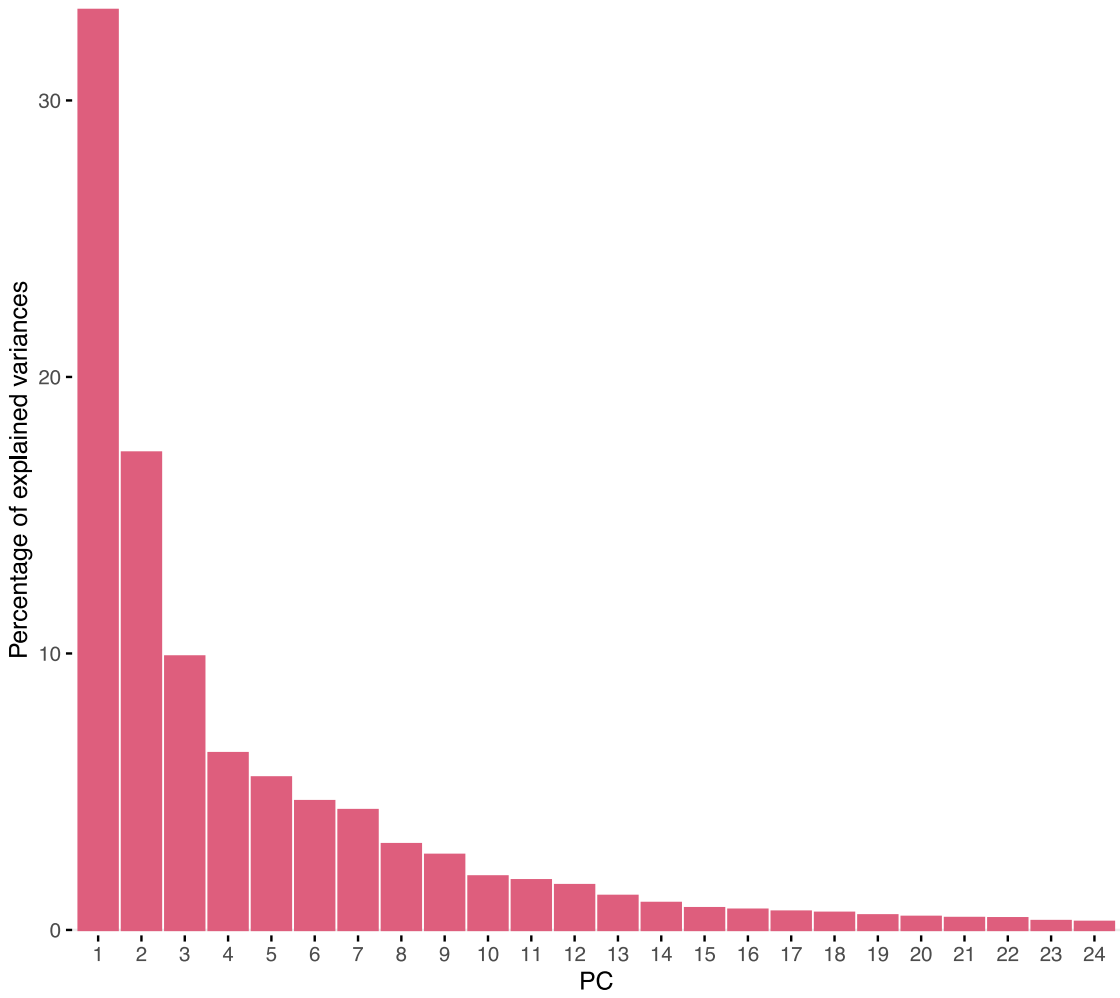

Figure 4.7: Scree plot from principal component analysis of inner retinal thickness measures. A scree plot showing the variance explained by each of the PCs resulting from PCA of the 24 thickness phenotypes describing the inner retinal morphology. These phenotypes describe the thickness of the two inner retinal layers, in left and right eyes separately, at each of the six segments of the Macula 6 grid.

4.4.2. Results

In total, 3,533 variants were found to be significantly associated with one of the first ten PCs (figure 4.8). In comparison, the GWAS of the final selected simple phenotype, as described below (section 4.5), discovered 3,434 significantly associated genetic variants. Filtering of SNPs based on LD to identify lead variants in loci was not performed here because the aim is still phenotype exploration. Many of the genetic variants found in the PC GWAS and final GWAS overlapped. Correlation of the effect size of SNPs discovered in the GWAS of the PCs, and the effect size of the same SNPs in the final GWAS, results in Pearson's r=-0.528, P $<2.2 \times 10^{-16}$, figure 4.9. The correlation pattern of the effect size on each phenotype is not linear. Whilst a subset of the SNPs appears to be positively correlated with one another, another subset appears to be negatively correlated. This suggests that some of the SNPs have opposite effects on the simple inner retinal thickness and the inner thickness derived PCs.

Considering the results of this analysis and the broader characteristics of the two phenotypes, there are advantages and disadvantages to using both the traditional clinical measures and the dimensionality transformed measures. The two phenotypes appear to identify a group of jointly associated genetic variants in addition to each identifying a unique set of significantly associated SNPs. The use of the higher dimensional phenotypes does not appear to yield more statistical power, and subsequent increased number of overall discovered loci, than when using the simple phenotype. Following discussion with clinical colleagues, considering that minimal statistical power is lost, it was decided to use the simple measure of retinal layer thickness going forward. One of the main advantages of using this phenotype is the easier interpretation of the results and subsequent translation into the clinical space.

4.5. Final GWAS

Having pursued the use of other possible phenotypes to represent inner retinal morphology from the available data within the Macula 6 grid, I decided to use a simple representation: the mean thickness across the Macula 6 grid of the two inner retinal layers, the RNFL and GCIPL, averaged across left and right eyes. This value is clinically valuable, easily interpretable and as shown below provides rich results likely informed by the size of the dataset providing large power to the statistical analysis.

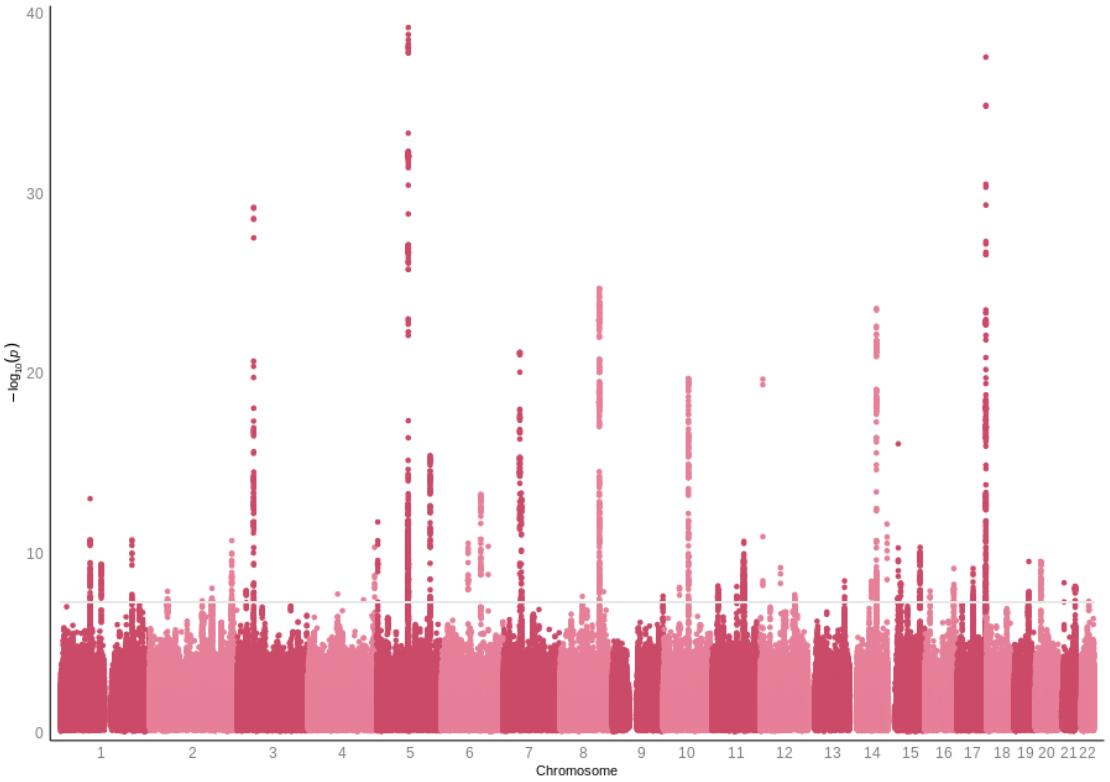

Figure 4.8: Genome-wide association study of first ten principal components of inner retinal phenotypes. Manhattan plot of meta-analysed results of genome-wide association studies of first ten principal components (PCs) of inner retinal phenotypes. The PCs are derived from the thicknesses of the retinal nerve fibre layer (RNFL) and ganglion cell inner plexiform layer (GCIPL) at the six segments of the Macula 6 grid in the left and right eyes. Variants are considered significantly associated if they meet genome wide significance ($P < 5 \times 10^{-8}$).

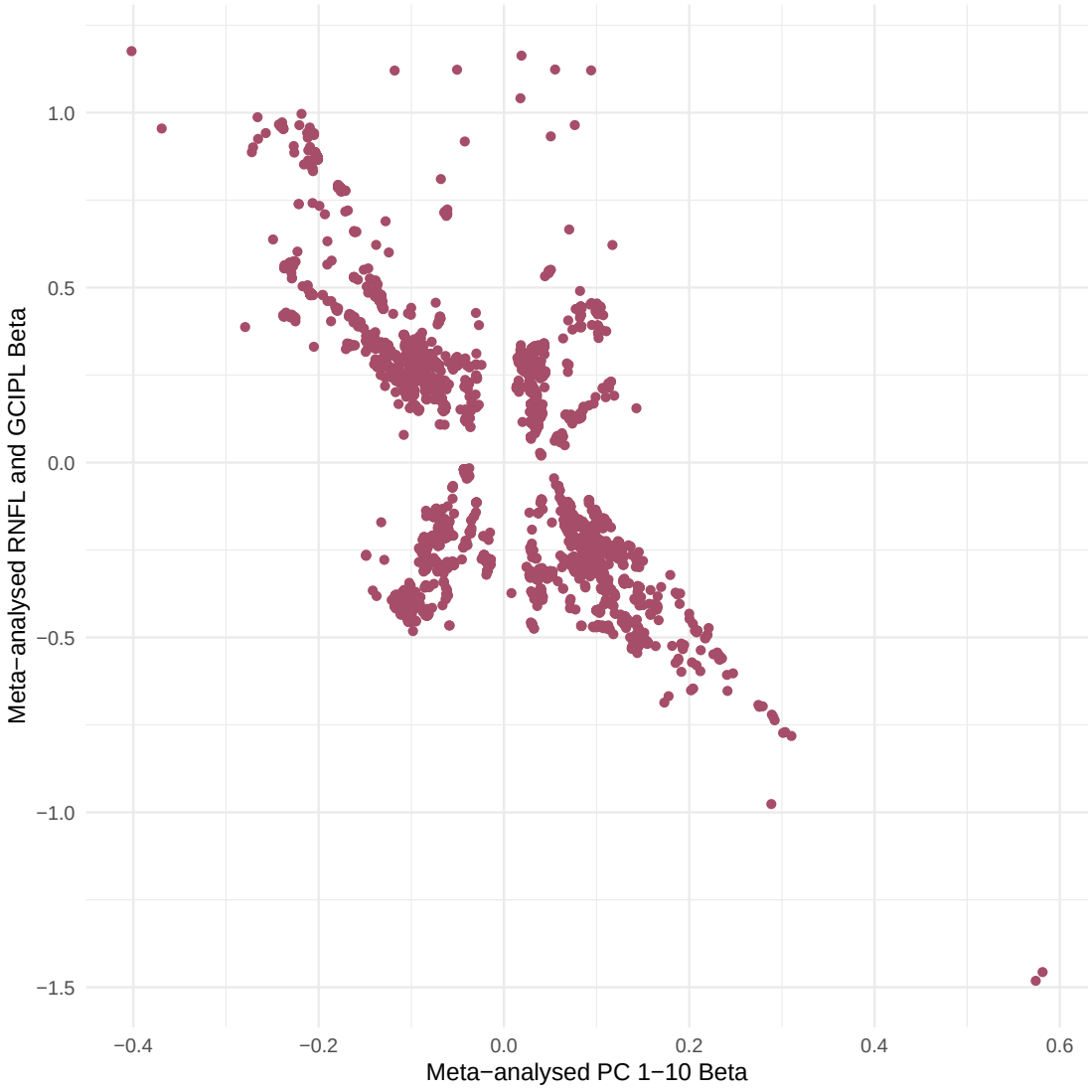

Figure 4.9: Comparison of effect size and direction of SNPs on simple inner retinal phenotypes, and inner retinal principal components. A correlation plot of the effect sizes of the cross section of SNPs significantly associated with the first 10 PCs of the inner retinal phenotypes, and the final simple inner retinal phenotypes representing meta analysed mean retinal nerve fibre layer thickness and mean ganglion cell inner plexiform layer thickness. Pearson's r=-0.528 ($P < 2.2 \times 10^{-16}$).

4.5.1. Methods

Two GWAS were performed of the mean thickness across the Macula 6 grid of each of the inner retinal layers, the RNFL and GCIPL, averaged across left and right eyes. The GWAS were conducted using a linear additive model implemented in BGENIE [Bycroft & al., 2018]. Eye specific covariates including technical covariates associated with scan acquisition and refractive error were regressed out of the individual eye measurements before a mean was taken of the thickness across the left and right eyes. The remaining covariates were included in the model, as detailed in section 4.2. Variants were considered significantly associated if they met the genome-wide significance level ($P < 5 \times 10^{-8}$). LDSC regression was implemented using LDSCore v1.0.1 [Bulik-Sullivan & al., 2015]. As there was a crossover of loci significantly associated with the two inner retinal layers, meta-analysis implemented in MTAG [Turley & al., 2018] was applied, and the associated variants are considered as a combined discovery set. Following application of MTAG, Conditional Joint Analysis (COJO) [Yang & al., 2012] was applied to perform conditional analysis and step wise model selection to filter for independent loci. The resulting 46 loci were further filtered and loci within 1.5Mb of one another were labelled as being in the same locus, as indicated in table 4.1. These 46 variants were also manually annotated with nearest gene, and any prior association with other phenotypes using ENSEMBL [Cunningham & al., 2019] and the Open Targets Genetics [Carvalho-Silva & al., 2019].

4.5.2. Results

In total, 46 loci were identified as significantly associated with one or both of the inner retinal layers following meta-analysis of the two layers and conditional joint analysis to filter those within LD (figure 4.10, table 4.1, table A.1). Both GWAS showed limited evidence of inflation due to residual population structure, as assessed by LD-scores (RNFL: λ GC = 1.11, Linkage Disequilibrium Score regression (ldsc) intercept=1.01, Ratio=0.05; GCIPL: λ GC = 1.12, ldsc intercept=1.01, Ratio=0.05, where ratio = (ldsc intercept-1)/(mean(χ^2)-1)).

Replication of our results was sought in two cohorts, the Rotterdam study [Hofman & al., 2015], and the Raine study [McKnight & al., 2012]. The Rotterdam study contains participants of comparable age to those in the UK Biobank and, despite the smaller cohort size, replication in this dataset saw correlation with the effect sizes of meta analysed loci associated with RNFL and GCIPL (Pearson's r = 0.74, P = 1.25×10^{-6}, table A.2). The Raine study is a cohort of considerably younger individuals. Despite this, replica-

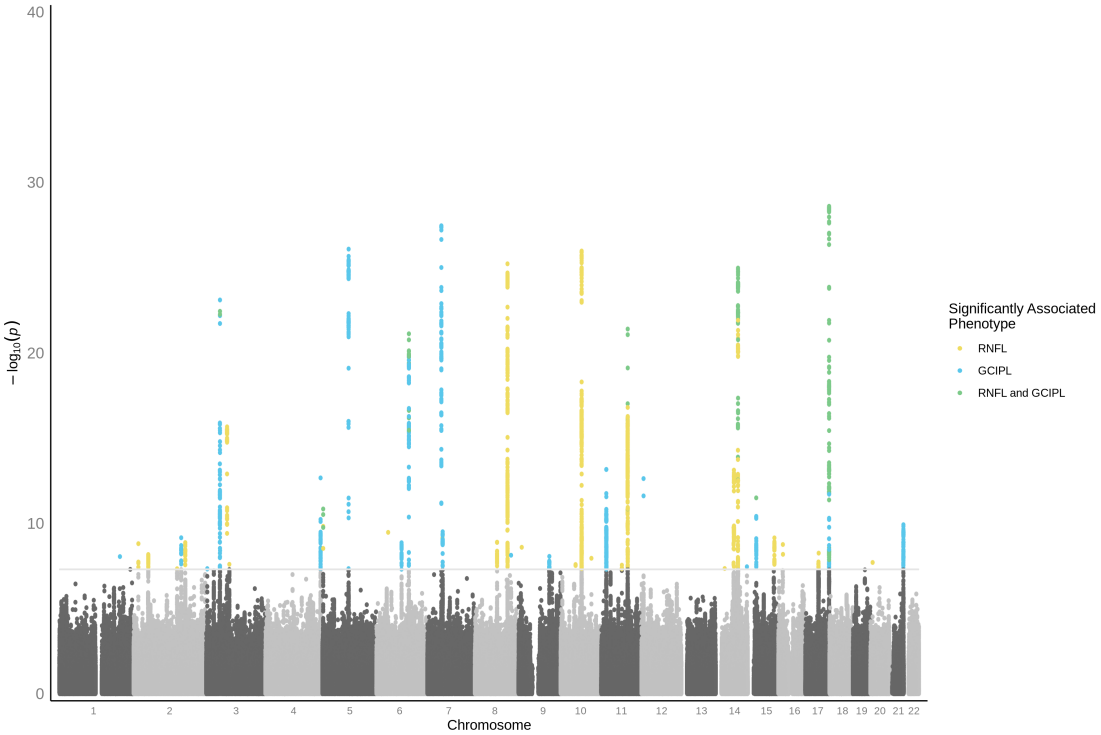

Figure 4.10: Genome-wide association study of mean macular inner retinal thickness phenotypes. Manhattan plot of p-values, resulting from meta-analysis across RNFL and GCIPL mean macular thickness phenotype GWAS. Variants significantly associated ($P < 5 \times 10^{-8}$) with only RNFL are highlighted in yellow, those significantly associated with only GCIPL are highlighted in blue, and those significantly associated with both inner retinal layers are highlighted green.

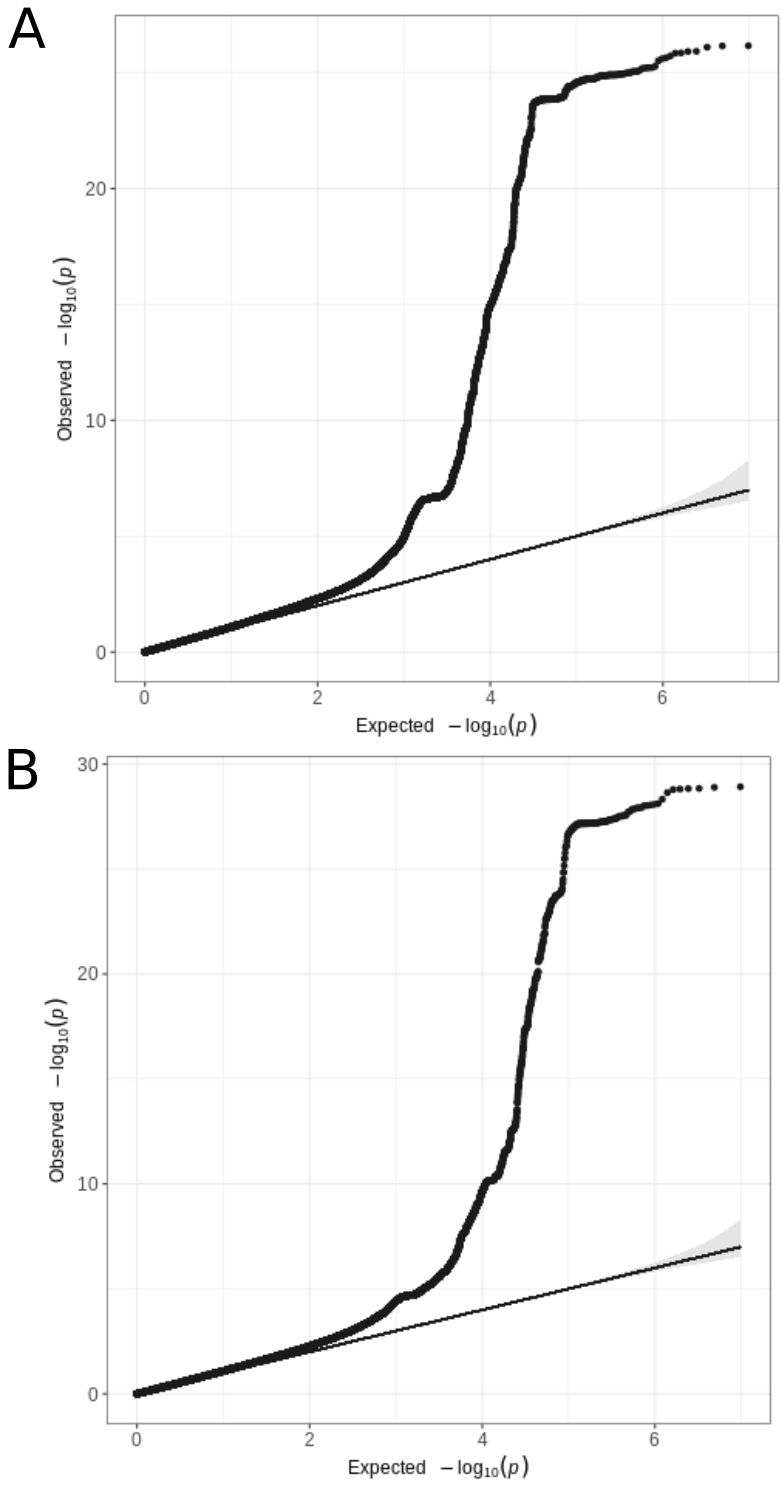

Figure 4.11: Quantile-quantile plots of inner retinal thickness GWAS. The quanitile-quantile plot (qq-plot) for A) the GWAS of RNFL thickness prior to meta-analysis (Lambda GC = 1.11, Intercept = 1.01, Ratio = 0.05). B) the GWAS of GCIPL thickness prior to meta analysis (Lambda GC = 1.12, Intercept = 1.01, Ratio = 0.05).

Table 4.1: 46 SNPs associated with one or more of the meta-analysed inner retinal layers (GCIPL or RNFL) thicknesses. The variants have been annotated with ocular and general biology phenotypes. The p-values presented are resultant from meta-analysis. Variants considered to be representative of a single locus, examples of allelic heterogeneity, are highlighted in the same colour alternating white and grey. For full results including effect size, effect allele specification and standard error please see supplementary Table A.1.

SNP	Chr	P-value	Nearest gene	Ocular Phenotypes	Non-ocular Phenotypes
rs72739513	1	8.88E-09	*ADORA1*		
rs79833181	2	1.55E-09	*NBAS*		
rs13010692	2	6.72E-09	*STON1-GTF2A1L*		
rs980772	2	4.62E-08	*TEX41*		Eosinophil count, Lymphocytes, Neutrophil count, Smoking, White blood cell count
rs12998032	2	6.97E-10	*CCDC148*		
rs2271758	2	1.34E-09	SLC25A12		Age completed full time education, Anthropometric traits, Reaction time
rs13083522	3	4.51E-08	*CRBN*		
rs17279437	3	7.81E-24	*SLC6A20*	Macular thickness	Blood metabolite levels, BMI, Hyperglycinuria, Iminoglycinuria, Overall health rating, Urinary metabolites
rs62252355	3	2.17E-16	*FRMD4B*		
rs149831820	3	2.53E-08	*ROBO2*		
rs66511946	4	2.15E-13	*STOX2*		
rs2004187	5	1.43E-11	*IRX2*	Macular thickness	Acute renal failure
rs17421627	5	8.09E-27	*LINC00461*	Retinal vascular calibre, Macular thickness	Seen doctor for nerves/anxiety
rs527871768	6	3.37E-10	*KIF6*		
rs13215351	6	1.36E-09	*SNAP91*	Spherical power	Bronchiectasis, Menarche, Napping, Standing height
rs9398171	6	7.51E-22	*FOXO3*	Macular thickness	Anthropometric traits, BMI, Coffee intake, Fat-free mass, Intelligence, Lung function, Menarche, Schizophrenia
rs11762530	7	3.45E-28	*IGFBP3*		Body Mass
rs73348111	7	7.15E-10	*IKZF1*		Mean reticulocyte volume, Monocyte percentage
rs35001871	7	1.17E-09	*GRB10*		
rs12719025	7	3.09E-10	*COBL*	Macular thickness, Spherical power, Strong/weak meridian	
rs6989495	8	1.27E-09	*RDH10*		
rs115520750	8	2.54E-10	*ANGPT1*		
rs13271359	8	5.89E-26	*RSPO2*		
rs376067714	8	1.28E-18	*RSPO2*		
rs4871827	8	7.41E-09	*DEPTOR*		Asthma, Height, Heel bone mineral density, Platelet count
rs118031671	9	2.53E-09	*PTPRD*		
rs2787394	9	8.64E-09	*INVS*	Macular thickness	BMI, Body Mass, Weight
rs1947075	10	2.60E-08	*ARHGAP22*	Macular thickness	
rs10762201	10	1.05E-26	*ATOH7*		Anthropometric traits
rs181211282	10	1.12E-08	*MRPL43*		
rs2008905	11	6.81E-14	*PIK3C2A*		Platelet count, Schizophrenia, Standing height
rs12574166	11	2.82E-08	*LINC02747*		Breast cancer
rs1042602	11	3.96E-22	*TYR*	Eye colour, IOP, Macular thickness, Oculocutaneous albinism	Depression, Hair colour, Heel bone mineral density, Nerves, Skin pigmentation, Tanning
rs5442	12	2.36E-13	*GNB3*	Hypermetropia, Macular thickness, Myopia, Spherical power	
rs146652416	14	4.46E-08	*FOXG1*		
rs17095953	14	1.76E-10	*DAAM1*		
rs1254276	14	7.52E-14	*SIX6*	Age started wearing glasses, Primary open angle glaucoma (POAG)	Anthropometric traits, Heel bone mineral density, Menarche
rs10140252	14	1.05E-25	*BBOF1*		Mean corpuscular haemoglobin, Mean corpuscular volume, Red blood cell count, Red cell distribution width
rs35337422	14	3.50E-08	*TDRD9*	Myopia, Spherical equivalent	
rs1800407	15	3.19E-12	*OCA2*	Age started wearing glasses, Cataract, Eye colour, Oculocutaneous albinism	Hair colour, Skin colour, Tanning response
rs1470108	15	7.09E-10	*AEN*		Standing height
rs117304899	16	1.74E-09			
rs117300236	17	5.57E-09	*NSF*		Balding, Forced expiratory volume, Haemoglobin concentration, Height, Mean corpuscular volume, Mean sphered cell volume, Neuroticism, Neutrophil percentage, Red blood cell count, Sensitivity
rs7503894	17	2.49E-29	*TSPAN10*	Age started wearing glasses, Astigmatism, Cataract, Cylindrical power, Spherical power	Hair colour, Tanning
rs143330165	20	1.97E-08	*LINC01428*		
rs7277632	21	1.20E-10	*PCBP3*		

tion in the Raine dataset saw correlation of effect sizes for loci associated with the GCIPL (Pearson's r = 0.84, P = 5.95 x 10-6), however not with RNFL (Pearson's r = -0.03, P = 0.88). It is interesting to note that the RNFL showed more discrepancy of genetic effect between left and right eyes than the GCIPL in the analysis in section 4.3. Full replication results are available in table A.3.

The absence of a significant replication result in the Raine cohort may be attributed to a number of its population's characteristics. Notably the Raine cohort comprises of considerably younger individuals than the UK Biobank - whilst the mean age of our filtered UK Biobank population is 57, participants of the Raine dataset were phenotyped at age 20. As previously mentioned the thickness of retinal layers has been shown to change with age [Khawaja & al., 2020], so this significant age disparity may in part account for the lack of replication. Another potential cause for the lack of replication may be the small size of the replication cohort. The Raine cohort contained 1,014 participants compared to the 31,434 individuals used in the UK Biobank inner retinal GWAS. The smaller sample size of the replication dataset may have meant that many loci found in the inner retinal GWAS would not pass genome-wide significance in these replication cohorts. Considering these factors, and the robust replication in the Rotterdam cohort, the lack of replication in the Raine cohort is unlikely a cause for concern.

A large portion of the significant variants lie within or near genes that have previously been linked to eye phenotypes. Only one well-known glaucoma locus was found to be significantly associated with inner retinal thickness near the *SIX6* gene (rs1254276). *SIX6* has previously been associated with a thinner RNFL and overall development of the inner retinal neural layers [Ulmer Carnes & al., 2014]. Other well known POAG loci were not significantly associated including *CAV1* (P= 0.92 for rs10258482), *CDKN2B-AS1* (P= 1.70×10^{-3}, 1.20×10^{-2} and 1.20×10^{-2} for rs4977756, rs2157719 and rs1333037 respectively) and *TMCO1* (P= 0.20 and 0.28 for rs4656461 and rs7555523 respectively) [Liu & al., 2017]. This is not entirely unsurprising as the method of action for increased susceptibility to the disease that *CAV1* and *CDKN2B-AS1* take is through protection of the canal of Schlemm, not susceptibility of the neural cells [Elliott & al., 2016; Burdon & al., 2011]. Furthermore, very few of the loci discovered in the inner retinal GWAS reached genome wide significance in the IGGC GWAS of POAG (table A.4). In spite of little crossover with previously known glaucoma genetics, several loci, including *TYR*, *PIK3C2A*, *NSF*, *TSPAN10* and *STOX2* have previously been associated with IOP, an endophenotype of glaucoma (Further exploration below).

Two loci significantly associated with the inner retinal phenotypes are in or near genes with well-documented involvement in . These loci are *TYR* (rs1042602) and

OCA2 (rs1800407). Oculocutaneous albinism (OCA) is a rare condition in which there is lowered pigmentation in the retina, contributing to lower visual acuity [Kamaraj & Purohit, 2014] (previously discussed in chapter 1). Individuals with OCA also display foveal hypoplasia. Foveal hypoplasia is the underdevelopment of the foveal dip leading to a flatter fovea with unexcavated inner retinal layers, which also contributes to visual impairment [McAllister & al., 2010]. OCA is very under-represented in the UK Biobank, with only 19 participants identifying as having the condition. This is likely due to under-diagnosis of the condition, in combination with its low incidence in the population. The SNP at *TYR* has previously been discovered as part of a haplotype that causes Ocular Albinism, a partial form of OCA in which the ocular traits of OCA persist but the pigmentation traits are much milder [Campbell & al., 2019]. This result suggests a continuum between rare disease causing genetic variation, and common variation that produces much milder non-pathological forms of the same trait. This will be explored further in chapter 6.

Several of the discovered loci have also previously been associated with refractive error. These included *TSPAN10*, *GNB3*, *SNAP91* and *COBL*. This is in spite of us controlling for refractive error by using it as a covariate in the model. However, the best established refractive error locus, at *LAMA2* (rs12193446) [Hysi & al., 2020], was not significantly associated with either of our inner retinal phenotypes (P=0.32 in the meta-analysis). This suggests that any association with loci previously identified as being involved in refractive error is due to shared genetic process, not underlying magnification bias. Magnification bias is understood as discrepancies in accuracy of individuals' retinal thickness measurements resulting from refractive error magnifying the OCT image [Nowroozizadeh & al., 2014].

There were many other loci significantly associated with inner retinal thickness that are in or near genes with well established roles in eye biology. The *ATOH7* (rs10762201) and *IRX2* (rs2004187) loci have both previously been associated with eye development [Brown & al., 2001; Zagozewski & al., 2014]. Studies have shown the particular importance of *ATOH7* in the development of retinal ganglion cells [Gao & al., 2014; Zhang & al., 2018]. Additionally variants at *IKZF1* (rs73348111) and *RSPO2* (rs13271359 and rs376067714) have been shown to be involved in differentiation and retinal cell definition [Mattar & Cayouette, 2015; Takata & al., 2017]. These findings highlight that retinal thickness at the individual level is affected by a number of fundamental ocular developmental genes.

There are additional loci associated with other ocular pathologies including a SNP near *NBAS* (rs79833181) which has been previously associated with optic atrophy [Mak-

simova & al., 2010]. *TSPAN10* and *OCA2* have also previously been associated with cataracts. These findings illustrate the complex relationship between non-pathological population level variation and disease biology.

In addition to the loci that had prior associations to both ocular and general phenotypes, some of the variants found significantly associated with inner retinal thickness had no prior associations. As an example, *PCBP3* (rs7277632) encodes a protein that binds to single stranded nucleotides. The gene has also shown an enriched expression in the inner nuclear layer (INL) of mice [McKee & al., 2005]. For each of the SNPs with no prior annotations, further work is needed to better understand the role they play in retinal morphology.

4.5.3. Enrichment Analysis

To gain further insight into the underlying biological pathways and mechanisms underlying the inner retinal phenotypes, enrichment analysis was conducted, implemented using GARFIELD [Iotchkova & al., 2016] (figure 4.12). This looks for associations between a trait, in this case inner retinal thickness, and regulatory features from different cell or tissue types. A fold enrichment of more than 10 was seen in eye tissue for loci associated with RNFL or GCIPL thickness ($P < 0.05$). Similar fold enrichment (>15×) was seen in brain tissue. Additional fold enrichment was seen in the pancreas (>15×), blood (>15×) and kidney (>10×), all consistent with broad pleiotropy of many of the loci, which have also been associated with anthropometric, developmental, and blood cell traits. This is likely due to their general role in development, as well as specific roles in retinal development. Interestingly, the enrichment in kidney tissue is also supported by prior associations between kidney disease and retinal disease [Deva & al., 2011; Bodaghi & al., 2014; Savige & al., 2011]. This is supported by the discovery of loci including *IRX2* (rs2004187) which has previously been associated with renal failure.

4.5.4. Mendelian Randomisation

To further interrogate the unexpected discovery of few variants associated with inner retinal thickness that are also associated with glaucoma, or its endophenotype IOP, I applied MR. As detailed in the section 2.8, MR analysis allows us to look for support for a causal relationship between two phenotypes, an exposure variable and an outcome variable. There are well-powered GWAS of both POAG and IOP available that enable this analysis.

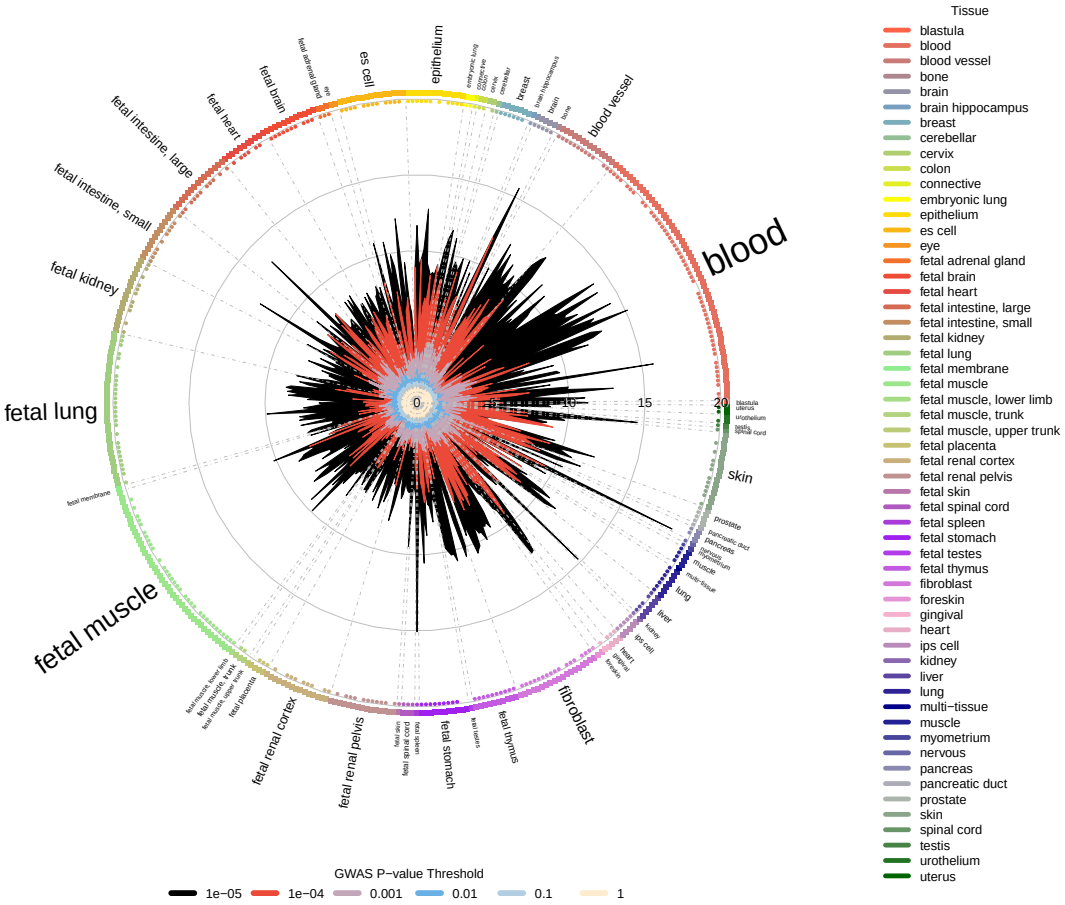

Figure 4.12: Regulatory feature association using GARFIELD. Wheel plot of enrichment analysis on meta-analysed GCIPL and RNFL GWAS results across a number of cell types, as performed in GARFIELD. Associations at different GWAS P-value thresholds are represented in different colours.

4.5.5. IOP and POAG

There is a plethora of research supporting a relationship between IOP and POAG. Clinically, lowering IOP via medication or surgery is one of the main clinical interventions to slow the progression of, or prevent POAG [Beidoe & Mousa, 2012; Weinreb & al., 2014]. Therefore I used MR to look for support of a causal relationship between the two variables. Due to the well established relationship between these two phenotypes, MR of the two variables acts as a positive control. IOP was used as the exposure variable, with genetic instruments selected using genome-wide associated ($P < 5 \times 10^{-8}$) variants from summary statistics of Khawaja & al., 2018. Summary statistics for POAG were taken from the IGGC [Gharahkhani & al., 2020]. Two sample MR analysis was conducted using TwoSampleMR [Hemani & al., 2018] and the in-package LD-pruning function was applied to the exposure variable to obtain a final list of genetic instruments. POAG was used as the outcome variable and, considering the cross-section of variants within the two studies, a final total of 97 instruments were used within the analysis. Five methods of meta-analysis were initially used and presented in the MR figures (inverse variance weighted, MR Egger, simple mode, weighted median and weighted mode). Two meta-analysis methods were used to definitively analyse the bidirectional results: MR-Egger and Inverse Variance Weighting. The Steiger test [Hemani & al., 2017] was also used to test for the directionality of the relationship. The results show strong support for a causal relationship between the two variables using both meta-analysis methods (figure 4.13, table 4.2). The Steiger test confirms that IOP affects POAG. This is congruent with current clinical practice and scientific literature. As well as establishing support for this relationship, this analysis acts as an appropriate control for our further analyses of phenotypic relationships that have been less well studied thus far. It also demonstrates that the two datasets used are well-powered to identify causal relationships using MR.

4.5.6. Inner Retinal Thickness, POAG and IOP

I used summary statistics from my GWAS of inner retinal thickness to look for a causal relationship between inner retinal thickness and either POAG or IOP. The null hypothesis stated that there was no causal relationship between the thickness of the inner retinal layers and either POAG or its endophenotype IOP. The genetic instruments were selected as the cross-section between the list of conditionally filtered SNPs significantly associated with RNFL or GCIPL thickness following meta-anlysis, and the SNPs with available information from the appropriate consortia GWAS summary statistics. Summary statistics for POAG and IOP were taken from Gharahkhani & al., 2020 and Khawaja

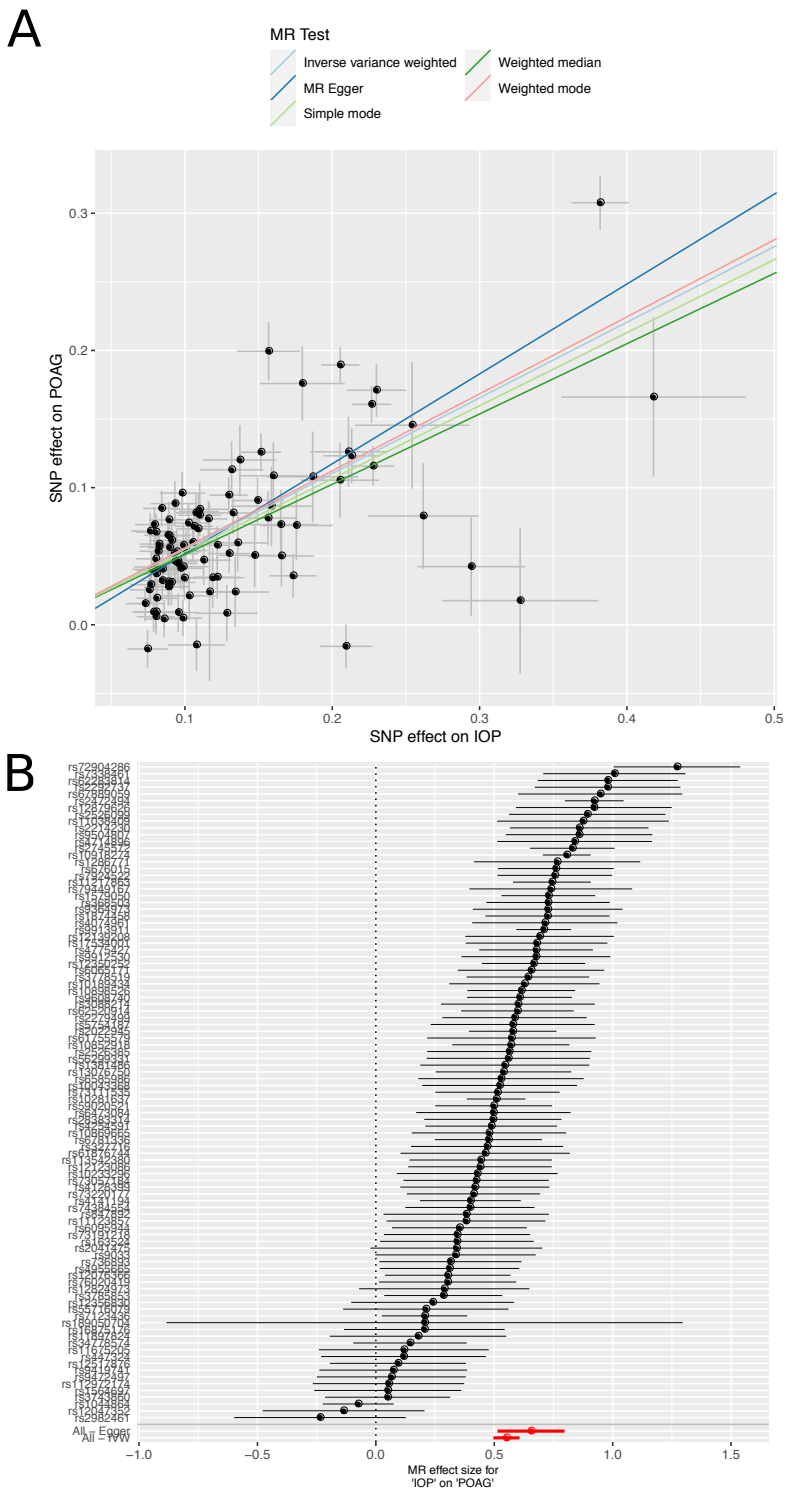

Figure 4.13: Mendelian randomisation analysis of relationship between IOP and POAG. A) Scatter plot showing the relationship between the effect size of SNPs significantly associated with intraocular pressure (IOP), and the effect of the same SNPs on primary open angle glaucoma (POAG). B) Forest plot showing the effect size and direction of SNPs significantly associated with IOP on POAG.

Table 4.2: Results of bidirectional two-sample Mendelian randomisation analysis between IOP and POAG, IOP and inner retinal layer thickness, and POAG and inner retinal layer thickness. Values reported for two meta analysis methods: MR Egger and Inverse Variance Weighted. POAG summary statistics were taken from the POAG International Glaucoma Genetics Consortium (IGGC) meta analysis [Gharahkhani & al., 2020]. Summary statistics for genetic association studies of IOP, were taken from [Khawaja & al., 2018].

Exposure	Outcome	MR Egger			Inverse Variance Weighted		
		Effect size	P-value	SE	Effect size	P-value	SE
IOP	**POAG**	**0.66**	**1.68E-14**	**0.07**	**0.55**	**2.80E-85**	**0.03**
POAG	GCIPL	-0.53	0.02	0.22	-0.16	0.06	0.09
POAG	RNFL	-0.41	0.03	0.18	-0.12	0.09	0.07
POAG	GCIPL and RNFL meta	-0.72	3.62E-03	0.24	-0.22	0.02	0.1
IOP	GCIPL	-0.08	0.65	0.18	-0.06	0.4	0.07
IOP	RNFL	-0.12	0.38	0.13	0.05	0.3	0.05
IOP	GCIPL and RNFL meta	-0.25	0.19	0.19	0.01	0.91	0.75
GCIPL	POAG	-4.39E-03	0.92	0.05	0.01	0.5	0.02
GCIPL	IOP	4.70E-03	0.94	0.07	-7.93E-03	0.73	0.02
RNFL	POAG	-0.01	0.91	0.09	-0.04	0.18	0.03
RNFL	IOP	-0.08	0.26	0.07	-7.26E-03	0.77	0.02
GCIPL and RNFL meta	POAG	0.02	0.66	0.05	-0.01	0.58	0.02
GCIPL and RNFL meta	IOP	-0.03	0.56	0.04	0.01	0.76	0.02

& al., 2018 respectively. Initially, for the analysis using both POAG and IOP, the inner retinal thickness was used as the exposure variable, and the disease phenotype was used as the outcome variable. Following Mendelian randomisation analysis, there was no support for a causal relationships between either RNFL or GCIPL and POAG or IOP (figure 4.14, figure 4.15 and table 4.2). When conducted with exposure and outcome variables reversed, meaning POAG or IOP is used as the exposure, and inner retinal thickness is used as the outcome, there was only minor support for a causal relationship between POAG and the inner retinal layers via one meta-analysis method (table 4.2). The Steiger test was also conducted bidirectionally, and the results of the relationships directionality were inconsistent. This provides additional evidence that there is no causal relationship between inner retinal thickness and POAG or IOP.

These results contrast the effective use of inner retinal thickness as a biomarker for POAG. A reasonable explanation is that clinical monitoring of POAG relies upon repeat measurements of inner retinal thickness. It is a change in the thickness over time that is more indicative of POAG rather than an absolute value. These results suggest that one's genetically determined inner retinal thickness is not a risk factor for the development of glaucoma and clinically it is a change in inner retinal thickness that is symptomatic of the disease. These SNPs and the resulting normal variation they cause are affecting the inner retinal measurement, and therefore may confound its use as a biomarker of glaucoma. In the future, it may be possible to control for such normal variation with

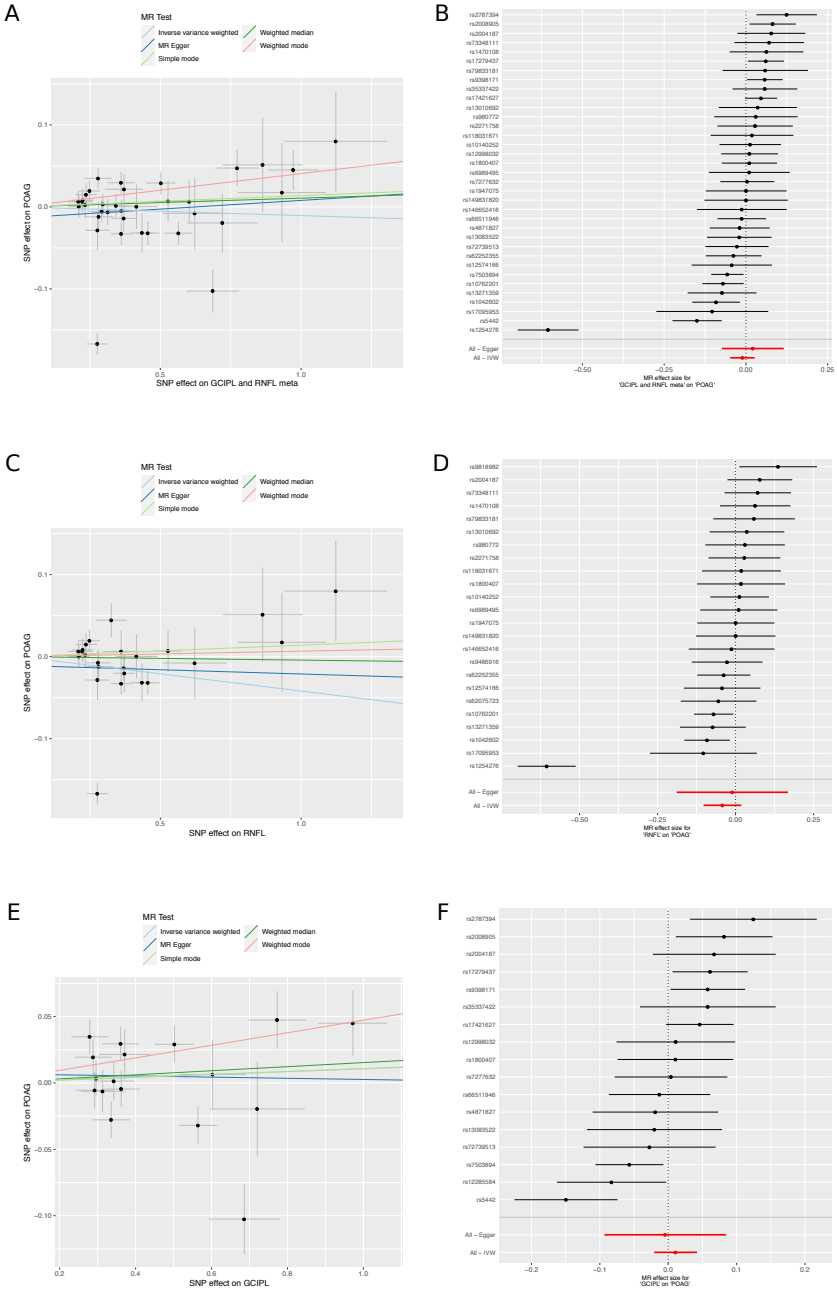

Figure 4.14: Mendelian randomisation analysis of the relationship between inner retinal thickness and POAG. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with ganglion cell inner plexiform layer (GCIPL) thickness, and the effect size of the same SNPs on POAG. B) Forest plot showing the effect size and direction of SNPs associated with GCIPL thickness on POAG. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with retinal nerve fibre layer (RNFL) thickness, and the effect size of the same SNPs on POAGs. D) Forest plot showing the effect size and direction of SNPs associated with RNFL thickness on POAG. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with meta-analysed GCIPL and RNFL thicknesses, and the effect size of the same SNPs on POAG. F) Forest plot showing the effect size and direction of SNPs associated with meta-analysed GCIPL and RNFL thickness on POAG.

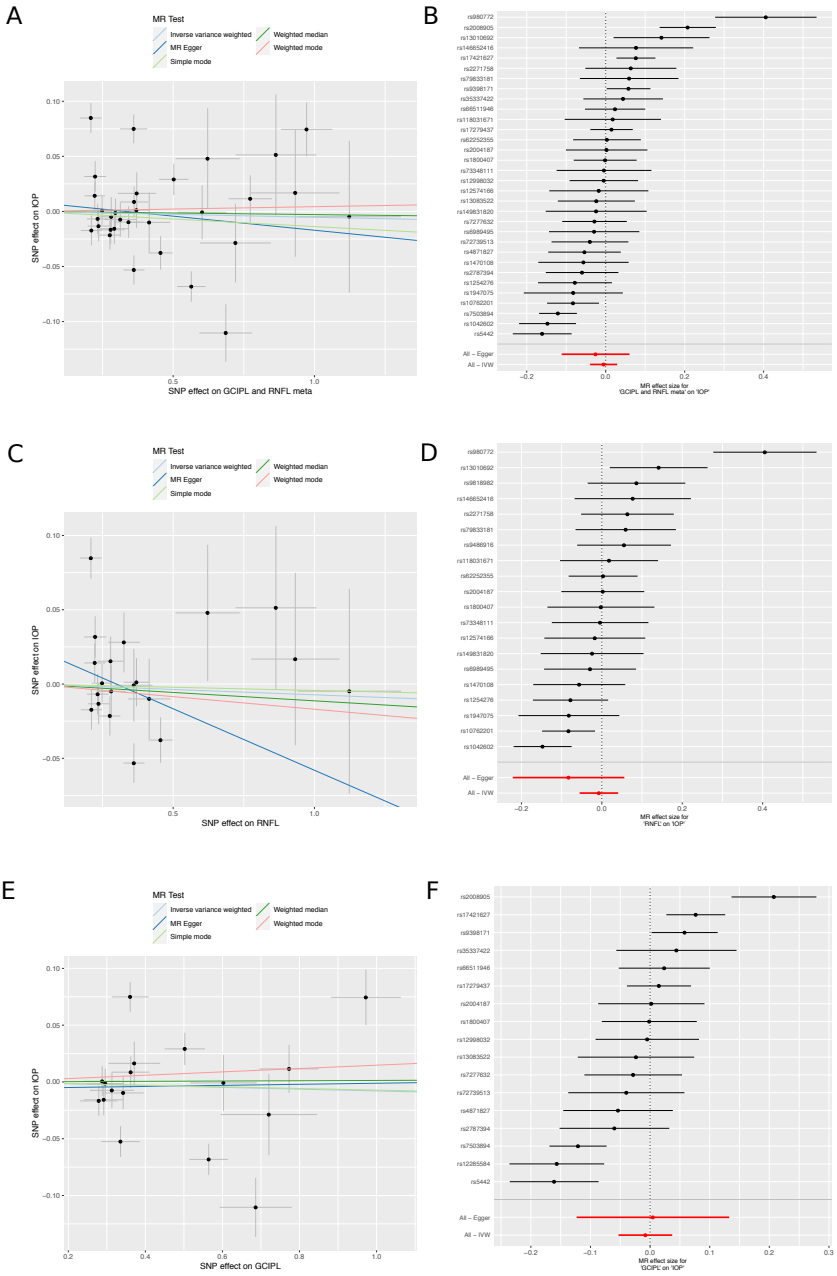

Figure 4.15: Mendelian randomisation analysis of the relationship between inner retinal thickness and IOP. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with ganglion cell inner plexiform layer (GCIPL) thickness, and the effect size of the same SNPs on IOP. B) Forest plot showing the effect size and direction of SNPs associated with GCIPL thickness on IOP. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with retinal nerve fibre layer (RNFL) thickness, and the effect size of the same SNPs on IOP. D) Forest plot showing the effect size and direction of SNPs associated with RNFL thickness on IOP. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with meta-analysed GCIPL and RNFL thickness, and the effect size of the same SNPs on IOP. F) Forest plot showing the effect size and direction of SNPs associated with meta-analysed GCIPL and RNFL thickness on IOP.

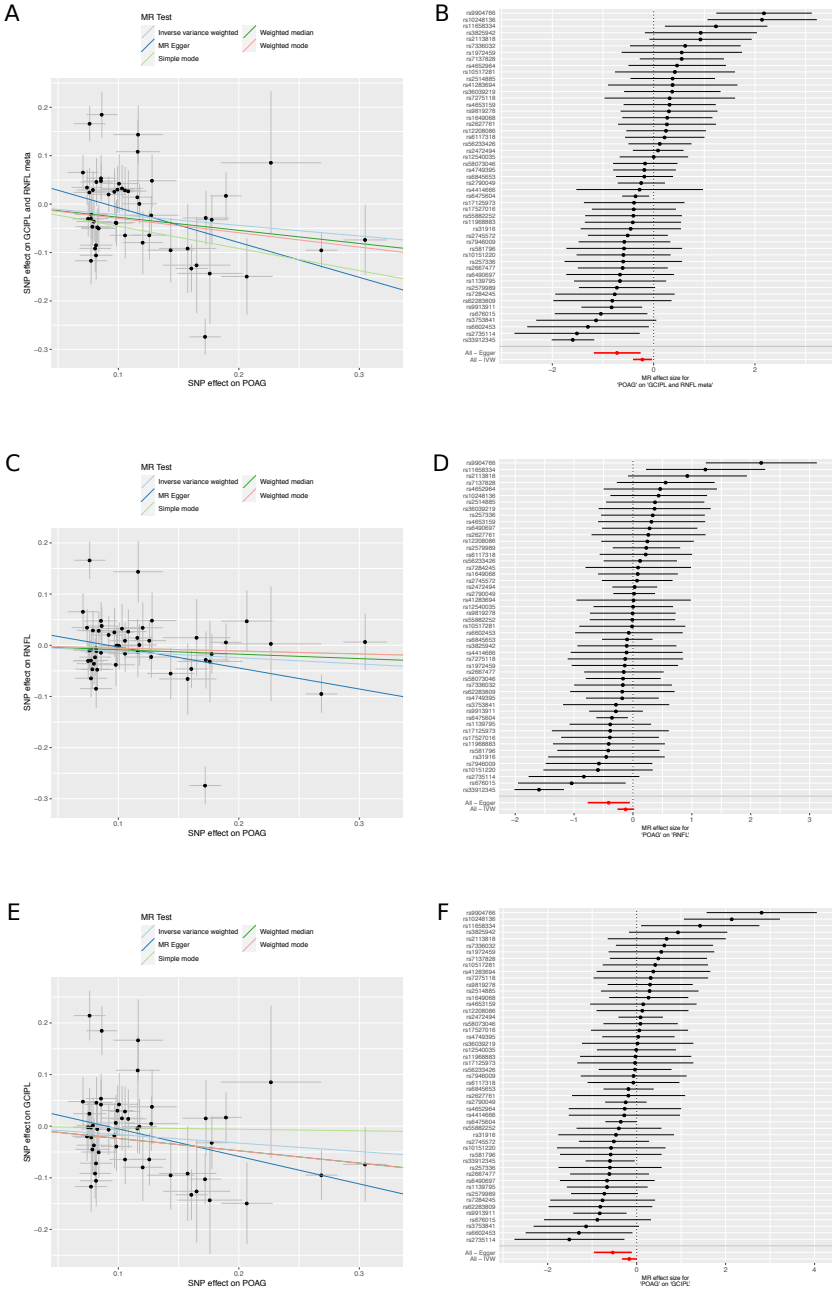

Figure 4.16: Mendelian randomisation analysis of the relationship between POAG and inner retinal thickness. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with POAG, and the effect size of the same SNPs on ganglion cell inner plexiform layer (GCIPL) thickness. B) Forest plot showing the effect size and direction of SNPs associated with POAG on GCIPL thickness. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with POAG, and the effect size of the same SNPs on retinal nerve fibre layer (RNFL) thickness). D) Forest plot showing the effect size and direction of SNPs associated with POAG on RNFL thickness. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with POAG, and the effect size of the same SNPs in meta analyses of GCIPL and RNFL thickness. F) Forest plot showing the effect size and direction of SNPs associated with POAG on meta-analysed GCIPL and RNFL thickness.

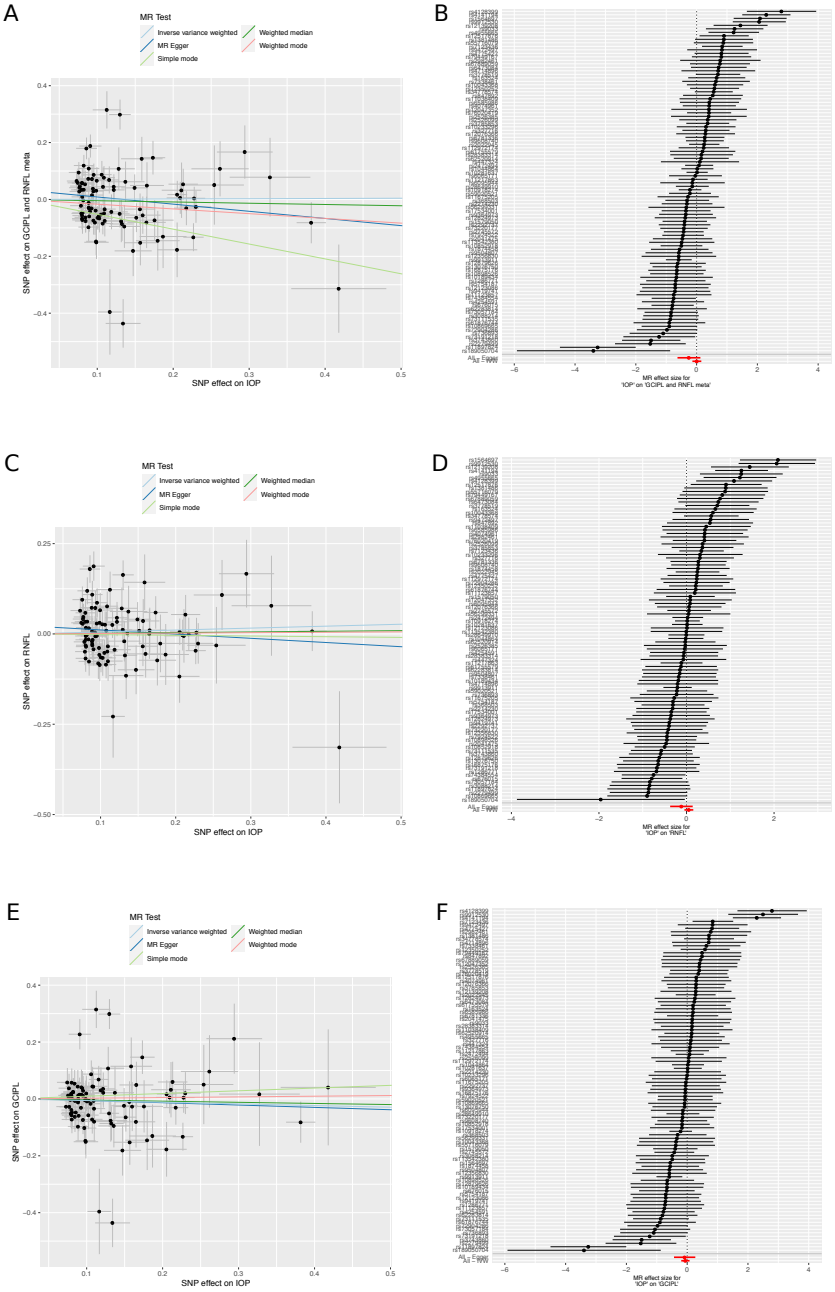

Figure 4.17: Mendelian randomisation analysis of the relationship between IOP and inner retinal thickness. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with IOP, and the effect size of the same SNPs on ganglion cell inner plexiform layer (GCIPL) thickness. B) Forest plot showing the effect size and direction of SNPs associated with IOP on GCIPL thickness. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with IOP, and the effect size of the same SNPs on retinal nerve fibre layer (RNFL) thickness. D) Forest plot showing the effect size and direction of SNPs associated with IOP on RNFL thickness. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with IOP, and the effect size of the same SNPs in meta analyses of GCIPL and RNFL thickness. F) Forest plot showing the effect size and direction of SNPs associated with IOP on meta-analysed GCIPL and RNFL.

these SNPs, using them as an approximation for the baseline genetically determined inner retinal thickness. This may help improve the speed and accuracy of glaucoma diagnosis.

4.6. Discussion

In this chapter I have presented the first large-scale GWAS of inner retinal layer thickness. Prior to conducting the final GWAS, I applied quality control metrics to the imaging data and explored the validity of these metrics. A GWAS of the status of whether an individual passed or failed said metrics discovered no significantly associated genetic variants. This suggests that the quality control is not removing a genetically interesting group that would likely represent biological interest. Thus this quality control system was accepted and used in the rest of the analysis.

I also explored a number of alternative phenotypes representing the morphology of the inner retina. This included examining the left and right eyes separately. The GWAS results of the mean thickness of both the RNFL and GCIPL were largely congruent between left and right eyes, with a small subset of SNPs differentially affecting the measures in the two eyes. There is a possibility that this discrepancy is grounded in biological variation, however the confounding nature of scan acquisition - where order of the left and right eye scans were not randomised - makes it impossible to disentangle any biological and confounding signal of the order of measurement. This should be considered when future protocols of ocular study are designed, as it would be interesting to be able to study the biology and genetics of any asymmetry.

To consider the full scope of dimensional data available to describe the morphology of the inner retina using the clinical segmentation grids, PCA was conducted on the thickness of the individual retinal layers at each of the Macula 6 segments and considering the left and right eyes separately. A GWAS was conducted for each of the first 10 resultant PCs and the results of the ten GWAS were meta-analysed. A comparable number of genetic variants were found significantly associated with both the dimensionality transformed phenotypes, and the simpler clinical phenotypes used in the final GWAS. This similarity in discovery power is likely due to the large size of the dataset, meaning the increased power usually afforded by using dimensionally transformed phenotypes is unnecessary. As the magnitude of the results is similar, the simpler phenotype was selected, enabling easier interpretation of results and translation into the clinical space. The dimensionality of the data available to describe retinal morphology is further explored in chapter 6.

The final GWAS of mean RNFL and GCIPL thickness across the macular field were meta-analysed due to the overlap of discovered genetic variants associated with each phenotype. In total, 46 loci were found to be significantly associated with one or both of the inner retinal layers. These loci included several that have previously been associated with ocular disease such as OCA and optic atrophy as well as with general ocular development and function. Many of the loci identified by this GWAS highlight the link between common population level variation, optical disease and ocular biology. There were also several genetic variants associated with inner retinal thickness that had no prior annotations. These loci may require further investigation to better understand their role in ocular biology.

There was only a single SNP found to be significantly associated with the thickness of one or both of the inner retinal layers with prior associations to glaucoma. This is in spite of the use of inner retinal thickness as a biomarker of glaucoma. Using MR analysis, no support was found for a causal relationship between the inner retinal phenotypes and glaucoma, or its endophenotype IOP. From this, it can be concluded that baseline thickness of the inner retinal layers is not indicative of glaucoma development. It is rather the change in inner retinal thickness that is symptomatic of glaucoma. Considering this, it is possible that the normal variation caused by the SNPs discovered in the GWAS is confounding the measure of inner retinal thickness when it is used to diagnose glaucoma. Therefore, it may be possible to utilise the genetic markers of inner retinal thickness during diagnosis to help control for baseline inner retinal thickness. This may increase accuracy and efficiency of POAG diagnosis. Unfortunately an appropriate dataset in which to test this theory does not currently exist, as the glaucoma incidence within the UK Biobank is too low at this time. If an appropriate dataset becomes available, this would be an interesting avenue of research, providing further translational utility to the GWAS results.

OCT is fairly unique in its ability to image live tissue, in particular a component of the central nervous system, in an easy and non-invasive manner. Here we have demonstrated the potential of OCT derived quantitative measures to explore eye biology and the impacts of common variation on specific eye pathologies. This endorses OCT derived quantitative measures as an exciting path of exploration to identify novel biological findings and candidate disease biomarkers.

5. Genome-wide Association Study of Outer Retinal Morphology

The outer retina is found furthest from the front of the eye, and is responsible for receiving light stimulus and the initial stages of conversion to neural signal (figure 5.1A). Compared to the inner retina which is made up of neural cells, the outer retinal layers are made up of photoreceptor cells (PRC) and epithelial cells. This cell type difference leads to the inner and outer retina being affected by a set of different diseases, in addition to being differentially affected by diseases common to both areas [Müller & al., 2019]. Here, we perform GWAS of outer retinal layer thicknesses to uncover the genetic variation underlying outer retinal morphology.

The outermost layer of the outer retina is the RPE, the pigmented layer of cells that plays two main roles in the retina. Firstly it is responsible for absorbing any scattered light to improve clarity of vision. Secondly it provides nourishment to the rest of the retinal layers and additionally aids disposal of waste [James & al., 2003]. The rest of the outer retina is made up of photoreceptor cells, both rods and cones, responsible for absorbing light stimulus and transmitting it through the retina. Rods are responsible for vision in low light levels but lack colour definition. Conversely, cones are responsible for high acuity colour vision, but only in high light levels. The cones are found at highest number at the macula and fovea, with limited numbers of rods, and the inverse pattern is seen at the peripheral retina. Component layers of the PRC are the outer segment (OS), inner segment (IS) and outer nuclear layer (ONL), working inwards from the RPE (figure 5.1B). The OS contains the photopigments stored within stacked membranes. The IS contains the mitochondria of the PRC in addition to ribosomes that aid the assembly of the photopigments. The ONL contains both the cell body of the photoreceptor, including the nucleus and other organelles, and the synaptic terminal of the photoreceptor [Purves & al., 2001].

There are a number of diseases associated with outer retinal layers, most notably AMD. As previously detailed in section 1.6.2, AMD is defined by the build up of lipid deposits known as drusen under the RPE. This accumulation of deposits leads to the degradation of the RPE and progressive damage to the photoreceptors, causing loss of

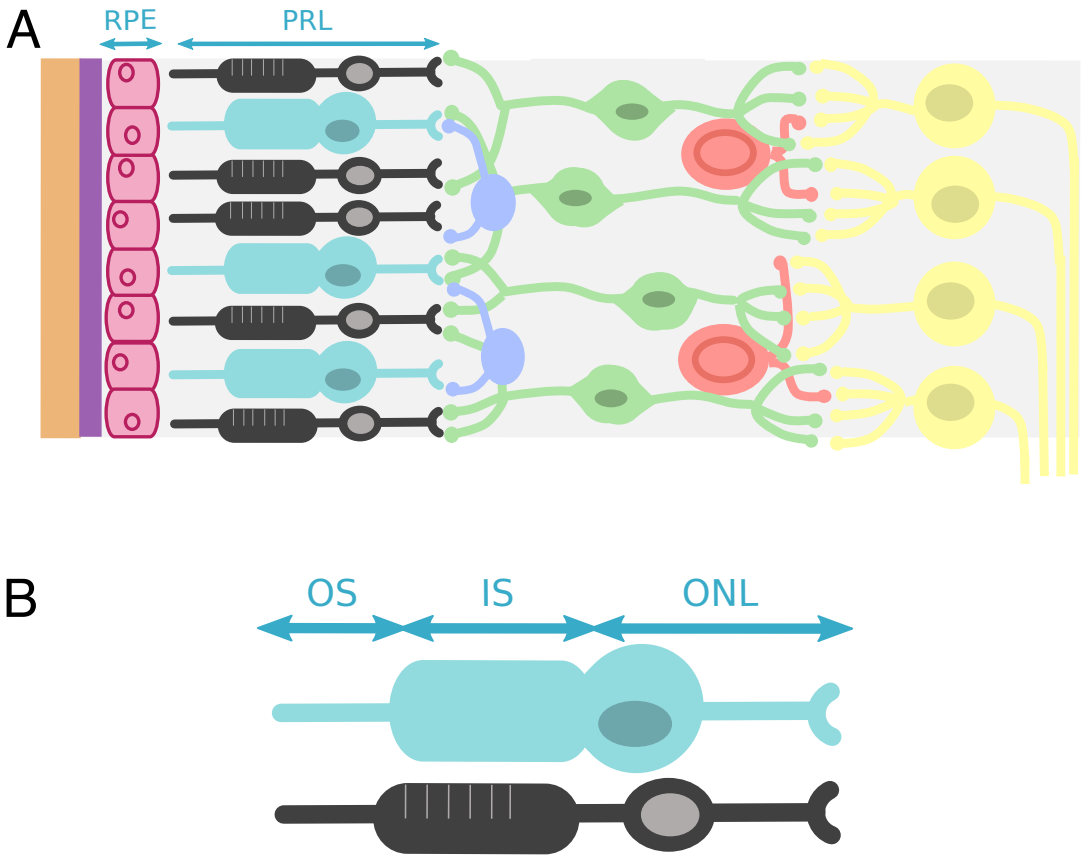

Figure 5.1: Layers of the outer retina A) A diagram of the retinal layers. The outer retina is comprised of the photoreceptor cells (PRC) and the retinal pigment epithelium (RPE). The light is first absorbed by the photoreceptor cells (PRC) in the PRL. The RPE absorbs scattered light and provides nutrients and waste removal to the rest of the retina. B) A schematic of the PRC and their component layers. The outer nuclear layer (ONL) contains the cell body of the PRC and the synaptic terminal. The inner segment (IS) contains the mitochondria of the PRC and the ribosomes that aid assembly of the photopigments. The outer segment (OS) contains the photopigments stored within stacked membranes.

vision [James & al., 2003]. It has been shown that the development of drusen cause thinning of the photoreceptor layers, as identifiable from OCT images [Schuman & al., 2009]. Thinning of the outer retinal layers in AMD has also been linked to reduced visual acuity [Pappuru & al., 2011].

In this chapter, I perform the first large-scale GWAS of outer retinal layer thickness, and analyse subsequent links to outer retinal diseases. I further explore the genetic variation underlying morphological variation at an increased dimensionality.

5.1. Outer Retina Data

The OCT image processing software TABS outputs a number of measurements that describe the morphology of the outer retina. Information is available on the thickness of the ONL, IS, OS and RPE (figure 5.2A).

In contrast to the inner retinal analysis in which I used the Macula 6 grid, throughout the outer retinal analysis I used data derived from the ETDRS grid (figure 5.2 B). The ETDRS grid is made up of concentric rings, and also split into superior and inferior, and nasal and temporal fields. This segmentation grid is more commonly used to look at the outer retina clinically. Unlike the Macula 6 grid, it includes the foveal field. This is because outer retinal layers, unlike inner retinal layers, extend into the fovea.

Therefore the available measures to describe the morphology of the outer retina includes the thickness of the ONL, IS, OS and RPE at each of the nine fields of the ETDRS grid, as well as the thickness averaged across the ETDRS grid, for left and right eyes (figure 5.2C). In-keeping with the analysis of the inner retina, I utilised simple phenotypes, measurements of thickness of each outer retinal layer averaged across both eyes and across the ETDRS grid, to describe the outer retinal morphology. This led to four measures that describe the morphology of the retina being used as phenotypes in GWAS: the mean thickness across the ETDRS grid of the ONL, IS, OS and RPE averaged across the left and right eye. Throughout the rest of this chapter, the component layers of the PRC (the ONL, IS and OS) are considered separately to the RPE due to the different cell type.

5.1.1. Quality Control

The quality control metrics described in chapter 3 were again applied to obtain the analysis cohort used in the outer retinal GWAS. In brief, following removal of related individuals, selection to the densest, well-mixed population, and removal of individuals who do not pass the quality control metrics based on technical imaging variables and other criteria, 31,237 participants remained.

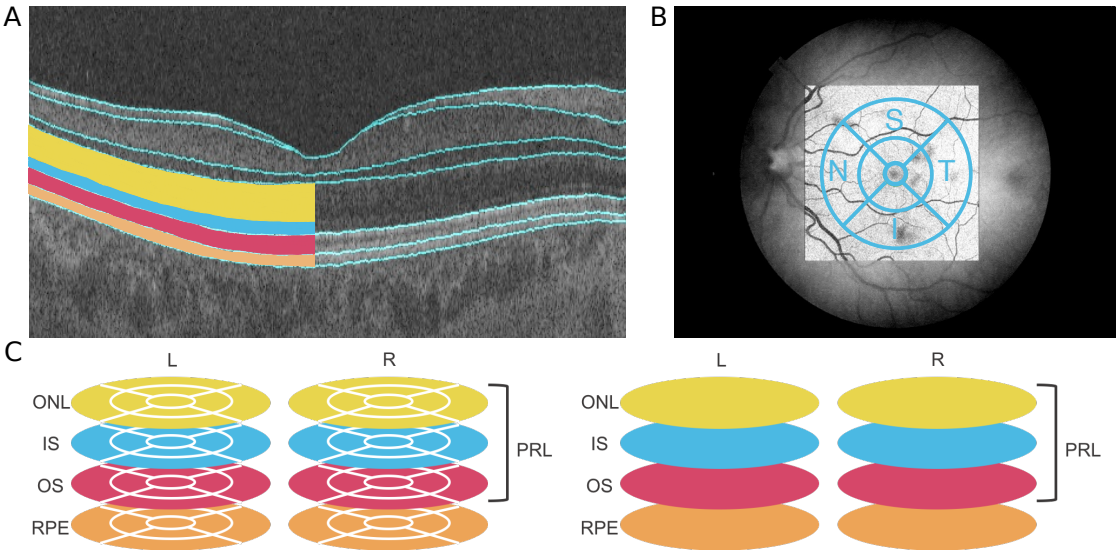

Figure 5.2: Data describing the outer retinal morphology. A) An optical coherence tomography (OCT) image that has had the Topcon Advanced Boundary Segmentation (TABS) software applied to it. On the left half of the image the outer nuclear layer (ONL) is shaded yellow, the inner segment (IS) is shaded blue, the outer segment (OS) is shaded pink, and the retinal pigment epithelium (RPE) is shaded orange. The ONL, IS and OS are components of the photoreceptor cell (PRC). B) An example fundus image of a retina with the Early Treatment Diabetic Retinopathy Study (ETDRS) grid, a commonly used partition matrix, superimposed atop in blue. The four quadrants of the grid are labelled as superior (S), temporal (T), inferior (I), and nasal (N). C) A schematic of the available measures that describe the outer retinal morphology. This includes the thickness of the ONL, IS, OS and RPE at each of the nine fields of the ETDRS grid, as well as the thickness averaged across the ETDRS grid, for left and right eyes separately.

5.2. GWAS of Outer Retinal Layers

5.2.1. Methods

A total of four GWAS were performed, one for each of the ONL, IS, OS, and RPE thicknesses taken as a mean across the ETDRS grid, and across left and right eyes. The GWAS were conducted using a linear additive model implemented in BGENIE [Bycroft & al., 2018]. As described in chapter 4, eye specific covariates including technical covariates outputted during scan acquisition and refractive error were regressed out of the individual eye measurements before a mean thickness was calculated across the left and right eyes. The remaining covariates, as detailed in section 4.2, were included in the model. Variants were considered significantly associated if they reached genome-wide significance ($P < 5 \times 10^{-8}$). LD score regression was applied using LDSCore V1.01 [Bulik-Sullivan & al., 2015]. The results of the GWAS on the component layers of the PRC, namely the ONL, IS and OS GWAS, were meta-analysed using MTAG [Turley & al., 2018] and combined into a single discovery list. COJO [Yang & al., 2012] was applied to the meta-analysed PRC component layer results and the RPE results to perform stepwise model selection and conditional analysis to select for independent loci. Further filtering was applied and loci within 1.5Mb of one another were labelled as being in the same locus as indicated in the results tables. The resulting genetic variants were annotated with nearest gene, and any prior association with ocular or general phenotypes using ENSEMBL [Cunningham & al., 2019] and Open Targets Genetics [Carvalho-Silva & al., 2019].

5.3. Results

For the four GWAS, calculation of LDSC regression yielded the values shown in table 5.1 where ratio = (LDSC intercept-1)/(mean(χ^2) (figure 5.3). As described in section 2.7.6, the LDSC intercept subtracted from 1 represents the confounding bias scaled to cohort size. Therefore an intercept close to 1 represents limited residual population structure. Additionally the ratio represents the proportion of p-value inflation caused by factors other than polygenicity. Considering this, the values obtained from the outer retinal GWAS indicate limited inflation suggesting minimal effect of residual population structure on the results. The ratio value for RPE is slightly higher than expected, but considering the LDSC intercept of the RPE is low, and the same population is used throughout the analysis, I concluded this may be due to the small number of significantly associated SNPs and proceed with the results.

Table 5.1: Results from linkage disequilibrium score (LDSC) regression on the GWAS results of the outer nuclear layer, inner segment, outer segment and retinal pigment epithelium thickness. For each GWAS λ GC is reported alongside the LDSC intercept and ratio. The LDSC intercept subtracted from 1 represents the confounding bias scaled to the cohort size. The LDSC ratio represents the proportion of inflation caused by factors other than polygeneicity.

GWAS trait	λ GC	LDSC intercept	LDSC ratio
ONL	1.15	1.04	0.13
IS	1.07	1.00	<0
OS	1.09	1.02	0.15
RPE	1.03	1.01	0.41

For each of the GWAS, many of the discovered variants lie in or near genes that have previously been associated with ocular phenotypes, including ocular development and ocular disease. When analysing the results, I have largely considered the component layers of the PRC and the RPE separately due to their different cell type.

5.3.1. Photoreceptor Layer

The three GWAS of the layers that comprise the PRC, the ONL, IS, and OS, yielded 74, 15 and 38 significantly associated ($P < 5 \times 10^{-8}$) loci respectively following conditional analysis (table A.5, table A.6, table A.7). As previously stated, due to the crossover of loci associated with the component layers of the PRC, we meta-analysed the genome-wide set of the three PRC GWAS summary statistics together using MTAG [Turley & al., 2018] (figure 5.4). Following conditional analysis, the meta-analysed discovery list consists of 59 genetic variants associated with one or more of the PRC layers (table 5.2 and table A.8).

Many of the discovered loci are in or near genes associated with other ocular phenotypes and conditions. A considerable number of variants have been previously associated with the age at which one starts wearing glasses or contact lenses. These include *LINC00461* (rs12653396), *RORB* (rs10781177), *CDHR1* (rs11200922), *SLC6A13* (rs2080402), *RDH5* (rs3138144), *FAAP100* (rs56737642) and *TSPAN10* (rs62075724 and rs9910935). *LINC00461*, a long non-coding RNA, has previously been associated with AMD [Han & al., 2020] and macular telangiectasia [Scerri & al., 2017]. *RORB* has previously been shown to be important in rod cell development in mice [Jia & al., 2009]. *CDHR1* has been associated with several retinal dystrophies, and is known to be key to the maintenance of photoreceptor structure [Stingl & al., 2017]. *RDH5* deficient mice have significantly thinner outer nuclear layer and rod/cone layers. The *RDH5* gene is

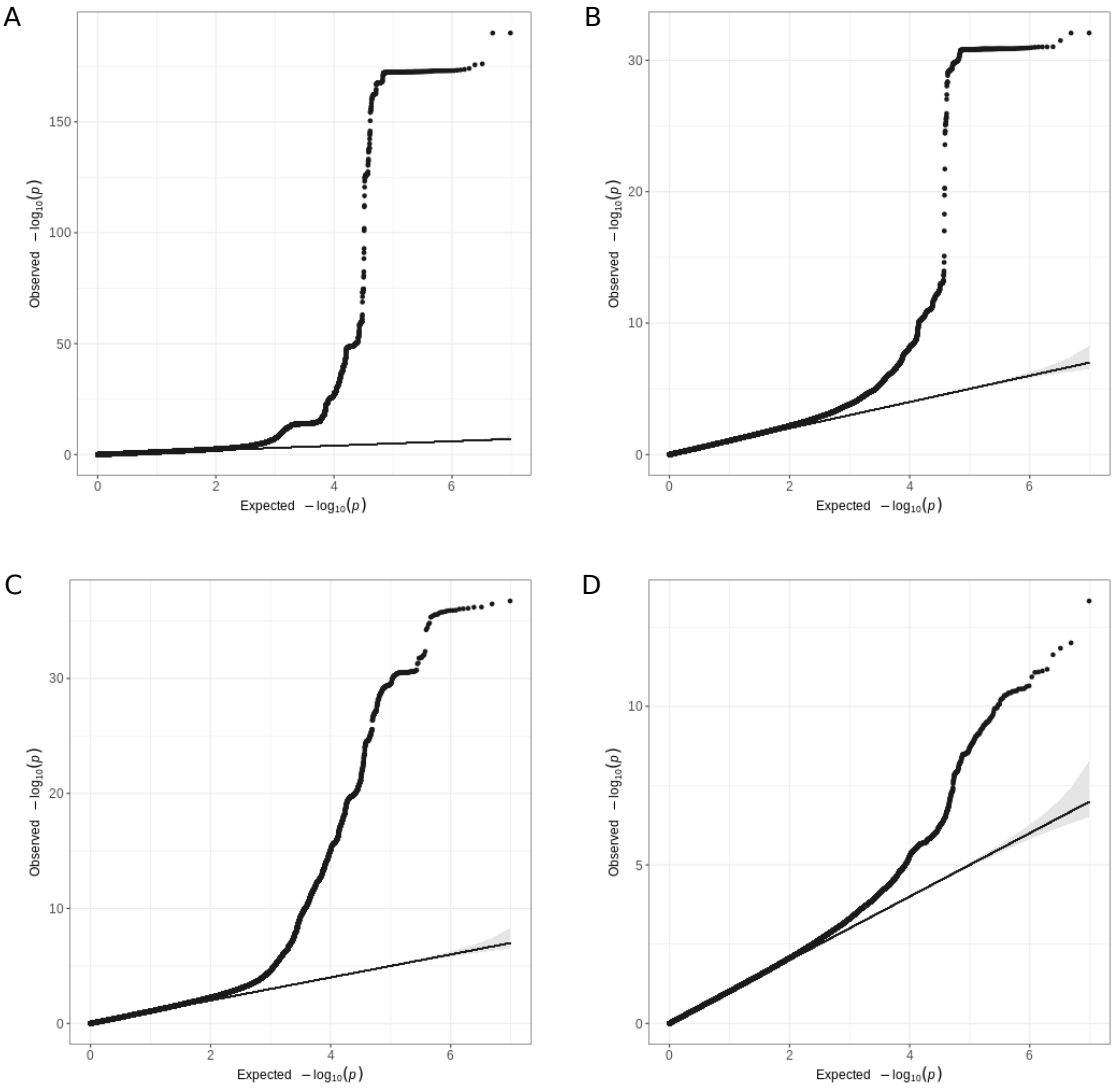

Figure 5.3: Quantile-quantile plots of outer retinal thickness GWAS The quanitile-quantile plot (qq-plot) for A) the GWAS of ONL thickness (λ GC = 1.15, LDSC intercept= 1.04, Ratio= 0.13); B) the GWAS of IS thickness (λ GC = 1.07, LDSC intercept= 1.00, Ratio= < 0); C) the GWAS of OS thickness (λ GC = 1.09, LDSC intercept= 1.02, Ratio= 0.15); D) the GWAS of RPE thickness (λ GC = 1.03, LDSC intercept= 1.01, Ratio= 0.41).

also known to cause fundus albipunctatus, a form of night blindness [Xie & al., 2020]. The *TSPAN10* locus has been associated with a number of ocular conditions including strabismus, the misalignment of the direction of the eyes commonly known as a squint [Plotnikov & al., 2019], and AMD [Orozco & al., 2020]. The association of these loci to age one started wearing glasses and the various types of photoreceptor-related ocular phenotypes supports the general importance of outer retinal structure in visual acuity.

Many of the loci found significantly associated with one or more of the meta-analysed PRC component layers have also been previously associated with either refractive error or spherical power. However, the best established refractive error locus, at *LAMA2* (rs12193446) [Hysi & al., 2020], was not significantly associated with the meta-analysed PRC layers (P= 0.04). Therefore, given refractive error is also accounted for as a covariate in the GWAS model, we conclude that any associations seen with refractive error are due to underlying shared biological pathways, and not due to residual confounding. *RORB* (rs1078177), *CDHR1* (rs11200922), *RDH5* (rs3138144) and *TSPAN10* (rs62075724 and rs9910935) are all loci associated with both the thickness of the PRC layers and refractive error. Several additional loci were found to also have prior associations with refractive error. These include *MOV10* (rs2932531), a locus that has known involvement in CNS development [Skariah & al., 2017], and when knocked out in *Xenopus laevis*, causes reduced differentiation of retinal cells [Skariah & al., 2018]. Additionally *ZNF281* (rs6673000) has prior associations with refractive error and has been found to be involved in neuronal differentiation in neuroblastoma cells [Pieraccioli & al., 2018]. This suggests a role of common variation in ocular function, specifically, visual acuity.

One locus, *MYO7A*, was found to be significantly associated with the PRC component layers and has also previously been associated with Usher syndrome [Well & al., 1995; Jacobson & al., 2011; Cheng & al., 2018]. Usher syndrome is characterised by deafness and retinitis pigmentosa. Retinitis pigmentosa refers to a group of incurable conditions that cause progressive loss of vision as rod cells are damaged [James & al., 2003].

The *CFHR2* locus (rs410895) which we found to be significantly associated with thickness of the component PRC layers is a locus known to be involved in AMD. *CFHR2* encodes a complement factor H (CFH)-related protein. *CFH* was one of two loci identified as significantly associated with AMD in the landmark study by Klein & al., 2005. CFH is an important part of the complement cascade, a system of immune response that promotes inflammatory response [Janeway & al., 2001]. Mutations in *CFH* are thought to be a factor in the formation of drusen [Hageman & al., 2005].

In total, four loci found significantly associated with the meta analysed PRC layers reached genome-wide significance ($P < 5 \times 10^{-8}$) in the International AMD Genomic

Consortium (IAMDGC) GWAS of AMD [Fritsche & al., 2016]. These loci are *RORB* (rs10781177, P=1.71 $\times 10^{-9}$), *RLBP1* (rs3825991, P=3.08 $\times 10^{-8}$), *FAAP100* (rs56737642, P=2.15 $\times 10^{-10}$) and *TSPAN10* (rs9910935, P=1.53 $\times^{-10}$)

In addition to loci with prior associations to ocular and general phenotypes, some of the variants found associated with thickness of the component layers of the PRC had no prior annotations. These include *KDM5B* (rs11577827) which is a histone demethylase that helps regulate embryonic stem cell differentiation [Kidder & al., 2014]. This mirrors the involvement of some of the known ocular loci in development. *DUOX1* (rs1648303) encodes Dual oxidase 1, a transmembrane protein that generates hydrogen peroxide by transporting electrons across membranes [de Faria & Fortunato, 2020]. It is involved in the thyroid hormone synthesis pathway, host defence and cancer. *LMAN1* (rs4940460) encodes Protein ERGIC-53, a lectin that has a role in endoplasmic rectticulum to Golgi apparatus transport [Zhang & al., 2003]. A more recent study has suggested that through this function, the protein may play a role in photoreceptor gene transport and homeostasis in mice [Hao & al., 2014]. For each of these SNPs, more research is needed to understand their role in retinal structure.

Table 5.2: 59 SNPs significantly associated with one or more of the meta-analysed photoreceptor cell component layer (ONL, IS or OS) thicknesses. The variants have been annotated with ocular and general biology phenotypes. The p-values presented are resultant from meta-analysis. Variants considered to be representative of a single locus, examples of allelic heterogeneity, are highlighted in the same colour alternating white and grey. For full results including effect sizes, effect allele specification and standard error please see supplementary Table A.8.

SNP	Chr	P-value	Nearest Gene	Ocular Phenotypes	Non-ocular Phenotypes
rs11587687	1	5.02E-09	*GNAT2*		BMI, Frequency of tiredness/lethargy
rs2932531	1	2.79E-09	*MOV10*	IOP, Myopia	Blood pressure, Hayfever, allergic rhinitis or eczema, Hypertension, Platelet distribution width
rs410895	1	8.19E-22	*CFHR2*	Macular degeneration	
rs6673000	1	2.73E-13	*ZNF281*	Spherical power	
rs11577827	1	2.77E-13	*KDM5B*		
rs6665290	1	8.24E-11	*COQ8A*	Cataract	Impedance of limbs, Lymphocyte traits, Neutrophil traits, White blood cell traits
rs2113422	2	2.89E-10	*ATAD2B*		Anthropometric traits, Basal metabolic rate, Forced expiratory volume, Guilt, Intelligence, Systolic blood pressure, White blood cell traits
rs13382582	2	5.99E-09	*STK39*		
rs838718	2	2.05E-08	*DGKD*		Crohn's disease, Disorders of bilirubin excretion, Haematocrit percentage, Haemoglobin concentration, Ibuprofen, Inflammatory bowel disease, Platelet count, Urinary calculus
rs34234056	3	2.37E-08	*SLC6A6*		

Table 5.2: continued

rs348866	3	1.96E-13	*IMPG2*	Corneal hysteresis	Body fat percentage, Crohn's disease, Eosinophill traits, Heel bone mineral density, Inflammatory bowel disease, Platelet traits
rs1376026	3	5.86E-09	*TSC22D2*		Basal metabolic rate, Body mass, High blood pressure, Hypertension, Number of self-reported non-cancer illnesses
rs2929724	5	4.21E-08	*BASP1-AS1*		Number of days per week walked >10 minutes,Open wounds on head neck and trunk
rs12653396	5	4.67E-11	*LINC00461*	Age started wearing glasses	ADHD, Age first had sexual intercourse, Alcohol intake frequency, Anxiety, Basal metabolic rate, Body mass, Educational attainment, Frequency of travelling from home to work place, Full cream consumption, Hot drink temperature, Impedance of limbs, Mean corpuscular volume, Sodium in urine
rs1109114	5	2.65E-13	*AFAP1L1*		BMI, Forced expiratory volume, Hearing difficulties, High blood pressure, Hypertension
rs1438692	5	1.32E-11	*AFAP1L1*		Forced expiratory volume
rs450611	5	5.82E-10	*BOD1*		Height at age 10
rs375435	6	3.56E-16	*PRPH2*		Neutrophil percentage
rs1926098	6	3.35E-08	*IMPG1*		
rs12719025	7	6.39E-22	*COBL*	Macular thickness, Spherical power, Strong/weak meridian	
rs2437002	8	5.77E-13	*RSPO2*		
rs2514842	8	3.84E-16	*RSPO2*		Ankle spacing width, Balding pattern, Contracture of palmar fascia, Disorders of muscle, ligament and fascia, Fascitis, Hair colour, Heel bone mineral density
rs10781177	9	1.49E-08	*RORB*	Age started wearing glasses, Spherical power	
rs717299	9	3.80E-13	*RORB*		Impedance of arms
rs11200922	10	1.62E-26	*CDHR1*	Age started wearing glasses, Hypermetropia, Spherical power	
rs618838	11	1.79E-09	*CTSF*		Anthropometric traits, Impedance of limbs, Overall health rating, Red blood cell traits
rs7128814	11	4.56E-08	*SMIM38*		Hair colour, Platelet crit, Skin colour, Tanning
rs631695	11	1.71E-09	*CCND1*		Breast cancer, Leg mass
rs948962	11	4.55E-11	*MYO7A*	Retinitis pigmentosa	Deafness, Non-syndromic hearing loss
rs28620862	11	7.57E-11	*TYR*		Hearing difficulties, Home location, Malignant melanoma, Seen doctors for nerves, anxiety or tension, Skin colour, Sunburn, Tanning
rs7930541	11	2.65E-09	*SESN3*		
rs2080402	12	2.21E-12	*SLC6A13*	Age started wearing glasses	Life time number of sexual partners, Size at age 10
rs3138144	12	6.54E-28	*RDH5*	Age started wearing glasses, Myopia, Spherical power	
rs6538677	12	7.36E-10	*SNRPF*		Mean platelet volume, Platelet count
rs9796234	13	4.69E-18	*GRK1*		

Table 5.2: continued

rs7151797	14	5.46E-14	*ACTN1*		Alternative exercises, Forced vital capacity, Frequency of tiredness/lethargy, Mean platelet volume, Sensitivity/hurt feelings
rs11626048	14	1.08E-45	*ELMSAN1*		Corpuscular traits, Mean sphered cell volume, Red blood cell traits
rs10147158	14	9.62E-16	*BBOF1*		Red blood cell distribution width
rs11159058	14	1.30E-24	*BBOF1*		
rs11159063	14	2.71E-50	*LIN52*		Chronotype, Mean corpuscular volume, Red blood cell distribution width
rs1972564	14	2.93E-74	*LIN52*		
rs28488340	14	9.02E-31	*VSX2*		
rs888413	14	1.92E-14	*YLPM1*		Anthropometric traits, Depression, Mood swings, Nervous feelings, Neuroticism, Platelet crit
rs1648303	15	2.72E-08	*DUOX1*		
rs10083695	15	8.00E-13	*WDR72*		Haematocrit, Red blood cell count
rs3825991	15	2.70E-31	*RLBP1*		Menarche
rs7206532	16	1.86E-09	*DYNLRB2*	Spherical power	
rs6564059	16	1.16E-08	*MEAK7*		Monocyte count, White blood cell count
rs11640338	16	2.25E-08	*GINS2*		Monocyte traits, Neutrophil traits, Platelet traits
rs1852269	17	1.88E-09			
rs11079866	17	1.58E-08	*PHOSPHO1*	Spherical power	Allergic disease, Asthma, Coronary atherosclerosis, Coronary artery disease, Eosinophill count, Forced expiratory volume, Height, High blood pressure, Hypercholesteromua, Impedance of limbs, Long-standing illness, disabilities or infirmary, Overall health rating
rs56737642	17	2.87E-11	*FAAP100*	Age started wearing glasses, Cataract, Logmar, Spherical power	Ankle spacing width, Hair colour, Skin colour, Tanning
rs62075724	17	2.91E-12	*TSPAN10*	Age started wearing glasses, Amblyopia, Cylindrical power, Logmar, Myopia, Spherical power	Hair colour, Tanning
rs9910935	17	2.75E-28	*TSPAN10*	Age started wearing glasses, Amblyopia, Cataract, Spherical power	Ankle width spacing, Hair colour, Skin colour, Tanning
rs4940460	18	2.86E-12	*LMAN1*		
rs6077977	20	4.98E-19	*JAG1*		Heel bone mineral density, Height
rs8132685	21	3.58E-09	*C21orf62*		
rs5752975	22	1.80E-09	*MTMR3*		Ankle spacing width, Basal metabolic rate, Body mass, Depression, Educational attainment, Forced vital capacity, Impedance of body, Lymphocyte traits, Mean sphered cell volume, Mood swings, Neutrophil percentage
rs8138196	22	2.38E-09	*OSM*		Red blood cell count

5.3.2. Retinal Pigment Epithelium

Seven loci were found significantly associated with the thickness of the RPE (figure 5.5, table 5.3 and table A.9). This included two well-known OCA loci, *TYR* (rs1126809) and

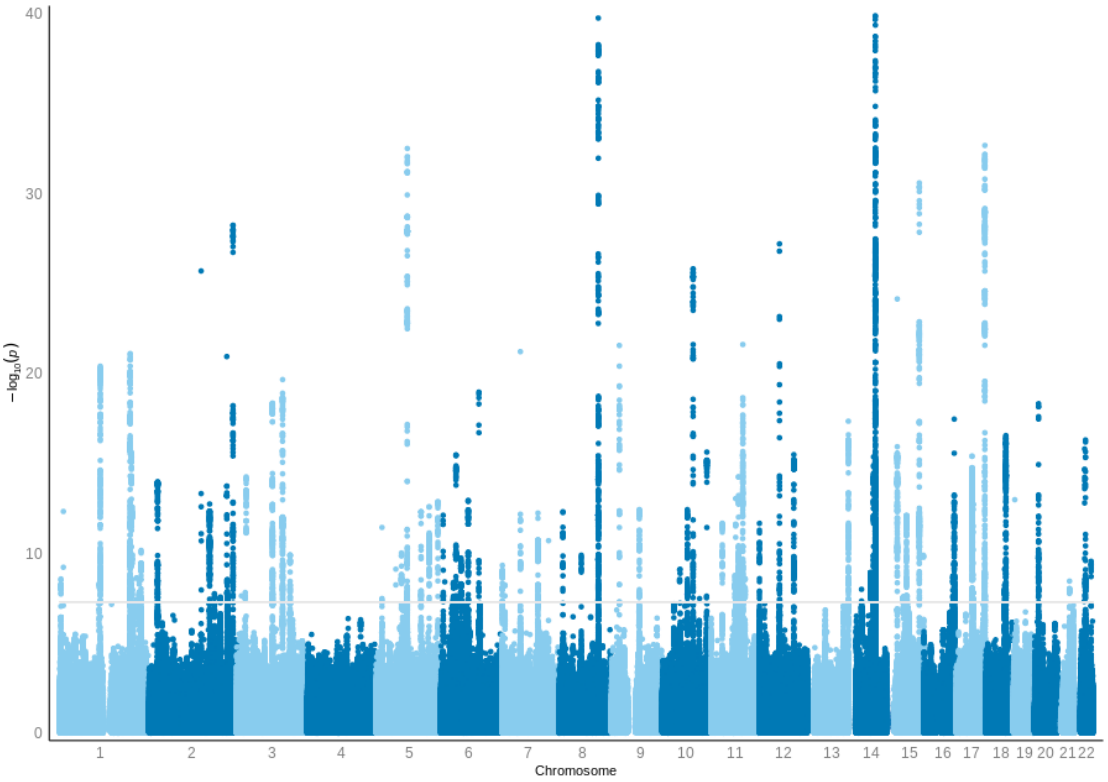

Figure 5.4: Genome-wide association study of the mean component photoreceptor cell layer thickness phenotypes. Manhattan plot of the p-values resulting from meta-analysis across IS, OS and ONL mean macular thickness phenotype GWAS. These three layers are the component PRC layers. Variants are considered significantly associated if they reach genome wide significance ($P < 5 \times 10^{-8}$).

OCA2 (rs1800407), which are also associated with inner retinal thickness. In the context of the RPE, OCA is characterised by the underproduction of melanin in tissues such as the skin, hair and, within the eye, the iris and RPE. OCA is well known to disrupt the morphology of the fovea through lack of foveal excavation and subsequently less cone maturation and lengthening [Kamaraj & Purohit, 2014]. However research has also suggested that cone maturation in OCA can happen on a continuum, which is potentially reflected in a mild form here [McAllister & al., 2010; Wilk & al., 2014]. Additionally the SNP at the *TYR* locus is part of a haplotype (with one of the SNPs found associated with inner retinal layer thickness) that has been shown to be associated with ocular albinism, a partial form of OCA [Campbell & al., 2019].

One of the loci identified, *CFHR1* (rs61822227), has previously been associated with AMD. *CFHR1* is a *CFH*-related gene. As noted above, *CFH* was one of two loci found in the very first GWAS of AMD [Klein & al., 2005]. In the IAMDGC GWAS of AMD [Fritsche & al., 2016], this SNP reached genome-wide significance ($P=1.62 \times 10^{-344}$). *CFHR1* encodes Complement factor H-related protein 1 which is involved in complement regulation. As stated above, the complement system is a key part of the immune response. It involves a network of many proteins that function in a cascade to induce inflammatory responses [Janeway & al., 2001]. It has been shown that *CFHR1* acts as a competitive antagonist to *CFH*, thus regulating the complement pathway[Goicoechea De Jorge & al., 2013].

Furthermore, three of the loci discovered, *GPR12* (rs77001109), *NETO1* (rs56194068) and *KIF16B* (rs77561053), have prior associations with neurological development. *GPR12* has been found to be expressed in areas of neuronal differentiation in the CNS during embryonic development of mice [Ignatov & al., 2003]. *NETO1*, which encodes a transmembrane protein, is known to be involved in the regulation of off-centre cone bipolar cells of mice [Chow & al., 2004]. *KIF16B* has been shown to be integral to localisation of endosomes in mouse neurons , and is also involved in the degradation pathway of neurons [Hoepfner & al., 2005; Farkhondeh & al., 2015] The RPE is integral in the disposal of redundant photoreceptors and associated waste [James & al., 2003].

The link of several retinal pigment thickness associated loci to both ocular-specific, but also general neural development, again reinforce the importance of the eye as a window into the tissue of the CNS.

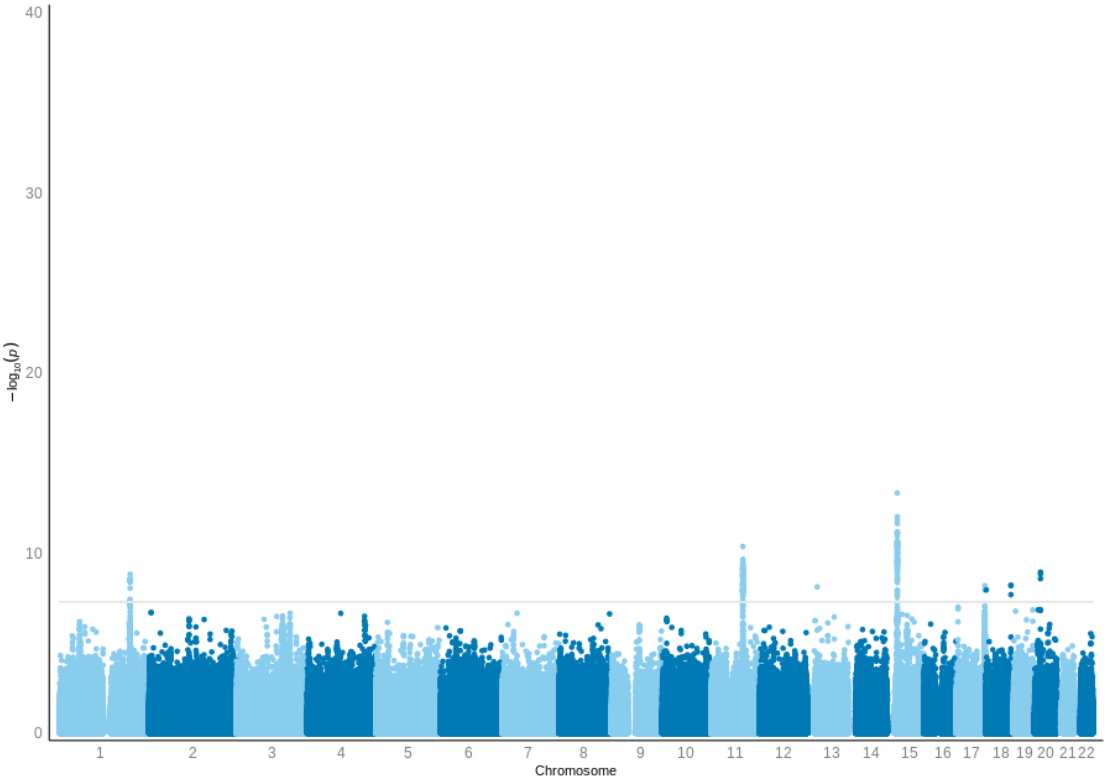

Figure 5.5: Genome-wide association study of mean retinal pigment epithelium layer thickness. Manhattan plot of mean macular retinal pigment epithelium layer thickness GWAS p-values. Variants are considered significantly associated if they reach genome wide significance ($P < 5 \times 10^{-8}$).

Table 5.3: Seven SNPs associated with RPE layer thickness and annotations of ocular and general biology phenotypes. For full results including effect size, effect allele specification and standard error please see supplementary table A.9

SNP	Chr	P-value	Nearest Gene	Ocular Phenotypes	Non-ocular Phenotypes
rs61822227	1	2.85E-09	*CFHR1*	Macular degeneration	
rs1126809	11	4.47E-11	*TYR*	Eye colour, Hypoplasia of fovea, Oculocutaneous albinism, Optic disc size	Basal cell carcinoma, Hair colour, Hearing difficulties, IOP, Malignant melanoma, Skin pigmentation, Squamous cell carcinoma, Sunburns, Tanning, Vitiligo
rs77001109	13	7.74E-09	*GPR12*		
rs1800407	15	4.78E-14	*OCA2*	Age started wearing glasses, Cataract, Central retinal arteriolar equivalent, Cylindrical power, Eye colour, Macular thickness, Oculocutaneous albinism	Hair colour, Relative age of first facial hair, Skin cancer, Skin colour, Sunburn, Tanning
rs7239443	18	1.14E-08	*METTL4*		
rs56194068	18	2.09E-08	*NETO1*		
rs77561053	20	1.19E-09	*KIF16B*		

5.4. Mendelian Randomisation

Due to the known clinical relationship between AMD and the outer retinal layers, I conducted MR analysis to look for support for a causal relationship between the two phenotypes (MR is discussed in detail in section 2.8). This analysis is enabled by robust GWAS studies of AMD within the scientific literature.

Bidirectional MR was conducted to interrogate the relationship between AMD and the thickness of the outer retinal layers: the ONL, the IS, the OS, the meta-anlysed PRC component layers, and the RPE. GWAS summary statistics for AMD were taken from Fritsche & al., 2016. The MR analysis was implemented using the TwoSampleMR package [Hemani & al., 2018]. Initially, thickness of the outer retinal layers was used as the exposure variable, and AMD was used as the outcome variable. The cross-section of the conditionally filtered SNPs found significantly associated with each outer retinal thickness phenotype, and the SNPs found in the AMD summary statistics were used as the genetic instruments. The in-package LD-pruning was further applied to the list of genetic instruments. MR analysis was also conducted in the reverse direction, with AMD used as the exposure variable and outer retinal thickness used as the outcome variable. In this instance, the cross-section of SNPs found to be significantly associated with AMD in the [2016] study, defined as $P < 5 \times 10^{-8}$, and SNPs found in the GWAS of the thickness of each outer retinal layer were used as genetic instruments. Again the in-package LD-pruning was applied to the list of genetic instruments. Five methods of meta-analysis are presented in the MR figures (Inverse variance weighted, MR Egger, simple mode, weighted median and weighted mode). Two meta-analysis methods were used to analyse the bidirectional results, MR-Egger and Inverse Variance Weighting. Sufficient support for a causal relationship was considered significance in at least one of two meta-analysis methods, with congruent effect size between the two meta-analysis methods. The Steiger test for directionality was also applied.

The results of this bi-directional MR analysis were mixed (table 5.4, figure B.5, figure B.6, figure B.7, figure B.8). There was support for causal relationships between some of the outer retinal layers and AMD (figure 5.6). However the Steiger test showed incongruent results when applied bidirectionally for all relationships apart from the relationship of AMD on RPE thickness. When AMD is used as the exposure variable and RPE is used as the outcome variable, there is statistically significant evidence for a causal relationship (Effect size = -0.19, $P = 4.31 \times 10^{-3}$ using Inverse Variance Weighting, figure 5.6A and B). AMD is defined by the development of drusen, abnormal lipid deposits, under the RPE. It is therefore logical that AMD has a causal relationship with

RPE thickness.

Little support is seen for a causal relationship when the thickness of the outer retinal layers is used as the exposure variable and AMD is used as the outcome variable (figure B.5, figure B.6). Overall, these results suggest that whilst AMD does damage the outer retina, seen clinically as changes in the outer retinal layer thickness, one's genetically determined outer retinal thickness does not alter the risk of AMD development. This relationship is similar to the pattern of relationship seen between the inner retina and POAG.

Table 5.4: Results of bidirectional two-sample Mendelian randomisation analysis between AMD and outer retinal layer thickness. Values are reported for two meta-analysis methods, MR Egger and Inverse Variance Weighted. AMD summary statistics were taken from the International AMD Genomic Consortium GWAS [Fritsche & al., 2016]. Results considered significant are bold

Exposure	Outcome	MR Egger			Inverse Variance Weighted		
		Effect size	P-value	SE	Effect size	P-value	SE
ONL	AMD	0.02	0.63	0.04	0.03	0.04	0.01
IS	AMD	0.14	0.67	0.31	0.11	0.22	0.09
OS	AMD	0.05	0.90	0.39	0.04	0.75	0.13
PRL meta	AMD	0.08	0.20	0.06	0.04	0.05	0.02
RPE	AMD	0.26	0.70	0.62	-0.22	0.42	0.27
AMD	ONL	-0.17	0.50	0.24	0.08	0.56	0.14
AMD	IS	0.02	0.23	0.02	0.02	0.05	0.01
AMD	OS	-0.44	2.51E-03	0.14	-0.15	0.07	0.09
AMD	PRL meta	-0.41	0.02	0.17	-0.11	0.26	0.10
AMD	**RPE**	**-0.20**	**0.09**	**0.11**	**-0.19**	**4.31E-03**	**0.07**

5.5. Exploring the ETDRS Grid

There are a large number of individual phenotypes available using the clinical segmentation grids that describe the morphology of the outer retina. As outlined above, performing GWAS of the thickness of outer retinal layers at each of the ETDRS grid subcompartments may be hard to interpret. However I was keen to explore the biological variation occurring at a slightly higher dimensionality than that utilised in my initial GWAS. An interesting feature of the ETDRS grid, especially compared to the Macula 6 grid used in the inner retinal analysis, is its division into concentric circles. As the macula dip varies radially in a similar fashion to the segmentation of the ETDRS grid, I decided to investigate whether genetic variation was affecting the foveal area, the intermediate ring, and the peripheral ring differentially (figure 5.7). Therefore I calculated the mean thickness for each of these areas for each of the outer retinal layers and performed GWAS of the resulting phenotypes.

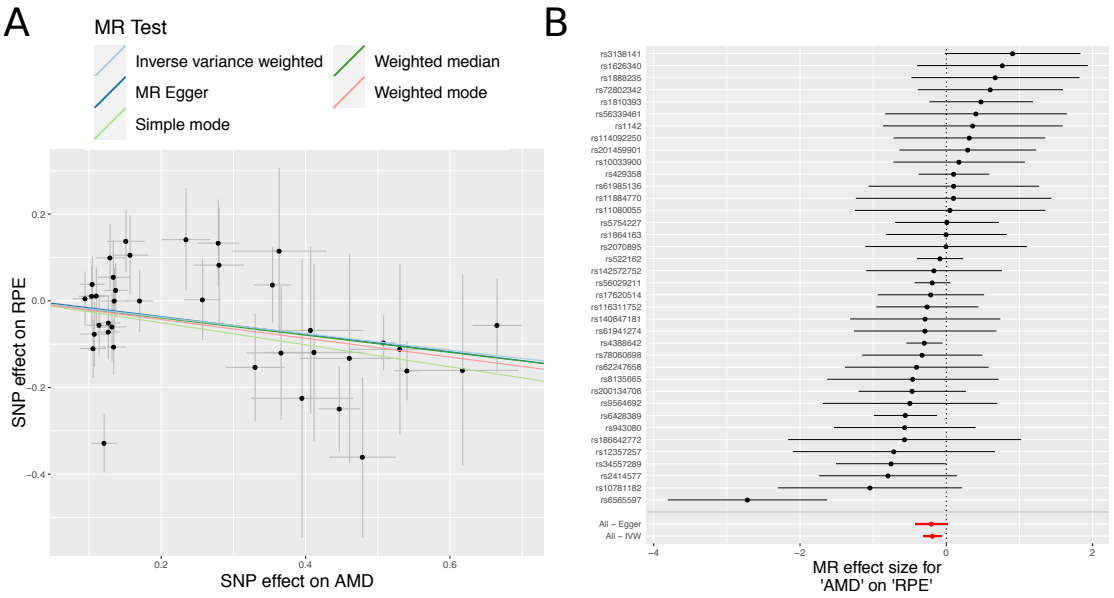

Figure 5.6: Mendelian randomisation analysis of the relationship between retinal pigment epithelium thickness and AMD. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on RPE thickness. B) Forest plot showing the effect size and direction of SNPs associated with AMD on RPE thickness. The GWAS summary statistics for AMD come from Fritsche & al., 2016

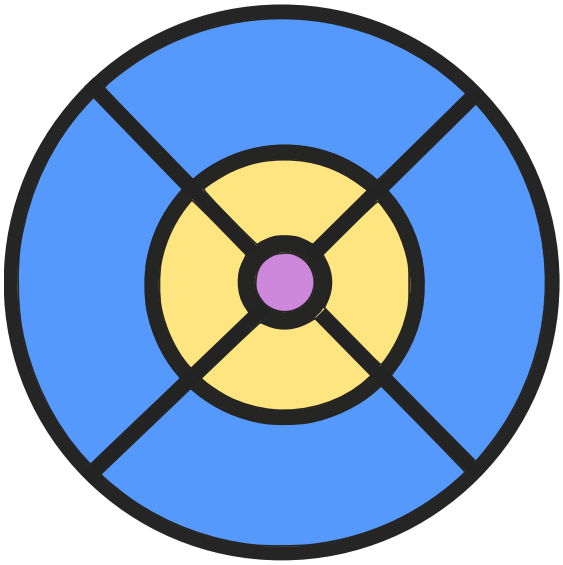

Figure 5.7: Division of Early Treatment Diabetic Retinopathy Study grid into concentric fields. A diagram of the Early Treatment Diabetic Retinopathy Study grid divided into concentric fields as denoted by shading. The foveal field is shaded purple, the intermediate field is shaded yellow, and the peripheral field is shaded blue.

5.5.1. Methods

A mean thickness across the four component areas of the intermediate ring, alongside the four component areas of the peripheral ring, was calculated. GWAS of the thickness of the ONL, IS, OS and RPE across the foveal area, intermediate ring and peripheral ring averaged across left and right eyes were performed, making a total of 12 GWAS. For each quantitative phenotype, an additive linear model was used implemented in BGENIE [Bycroft & al., 2018], with eye specific covariates regressed out prior to a mean across left and right eyes being calculated and the remaining covariates included in the model (as detailed above and in chapter 4).

To examine the results of the GWAS, comparisons were made between the effect sizes obtained from the GWAS of different concentric fields across the same retinal layer. For each SNP a z-score was calculated to compare its effect size on three areas: the foveal field compared to the intermediate field, the foveal field compared to the peripheral field, and the intermediate field compared to the peripheral field. The z-score was defined as:

$$Z = \frac{|\beta_0 - \beta_1|}{\sqrt{SE_0^2 + SE_1^2}}$$

A p-value is calculated for each z-score to identify SNPs with significantly different effects on the retinal fields. These loci were conditionally filtered for LD using COJO [Yang & al., 2012]. P-values were then Bonferroni adjusted and SNPs with $P < 0.05$ considered significant. The lists of significant SNPs were manually annotated using ENSEMBL [Cunningham & al., 2019] and Open Targets Genetics [Carvalho-Silva & al., 2019].

5.5.2. Results

Upon first inspection of the GWAS results across the different concentric fields, several observations stand out (figure 5.8). Firstly, in the ONL, there is a locus on chromosome 8 that does not appear to be significantly associated with the layer's thickness at the fovea, but is increasingly significantly associated with the layer's thickness as we move towards the peripheral ring. The opposite effect can be seen at a locus on chromosome 6 where the significance of the association with ONL thickness is much higher at the fovea than it is in the peripheral ring (figure B.1). In the IS GWAS, loci on chromosome 3, chromosome 7, chromosome 17 and chromosome 21 are seen to be significantly associated with the layers thickness at both the fovea and in the intermediate ring, but not the

peripheral ring (figure B.2). This pattern is seen in reverse for a locus on chromosome 13 in the OS GWAS, with the loci being significantly associated with the intermediate and peripheral rings, but not at the fovea (figure B.3). Finally in the RPE, a locus on chromosome 17 is significantly associated with the thickness of the layer at the fovea and the intermediate ring, but not in the peripheral ring (figure B.4). These initial observations suggest genetic variation is affecting the areas of the retina differentially, most probably owing to the complex organisation of the PRC about the fovea.

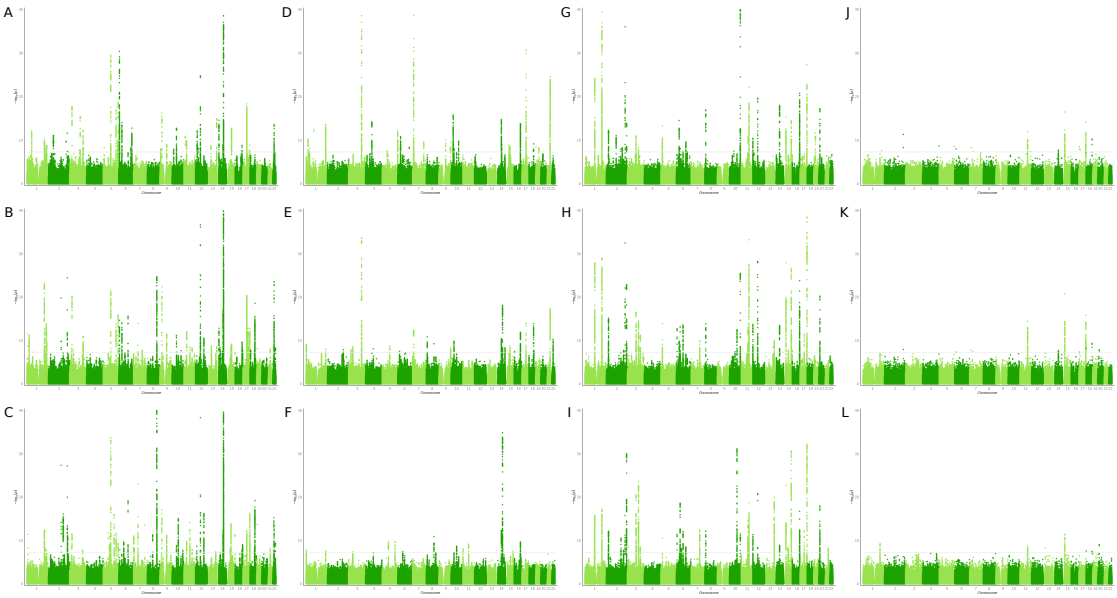

Figure 5.8: GWAS of outer thickness at concentric ETDRS fields A) ONL thickness across foveal field, B) ONL thickness across intermediate ring, C) ONL thickness across peripheral ring, D) IS thickness across foveal field, E) IS thickness across intermediate ring, F) IS thickness across peripheral ring, G) OS thickness across foveal field, H) OS thickness across intermediate ring, I) OS thickness across peripheral ring, J) RPE thickness across foveal field, K) RPE thickness across intermediate ring, L) RPE thickness across peripheral ring. The results are also shown in figure B.1, figure B.2, figure B.3 and figure B.4.

To further analyse the differential effect of SNPs on the sub-fields of the retina, we looked at which SNPs had notable differences in the effect size values obtained from their respective GWAS of the different macular fields. For each of the PRC layers there were numerous loci that appeared to be affecting the fields differentially. The loci discovered specific to each layer are detailed below.

Outer Nuclear Layer

Many loci were found to be differentially affecting the thickness of the ONL across the macular field (table 5.5). The ONL contains both the cell body and synaptic terminal of the PRC. Some interesting loci to highlight include *MIR29B2CHG* (rs6701735), *PLEKHA7* (rs10766368) and *LINC00880* (rs9860481). All three loci have the strongest effect on the thickness of the ONL at the fovea, and a reduced effect in the peripheral field. They also all have prior association with corneal hysteresis. Corneal hysteresis is a measure of the cornea's resistance to deformation [Agarwal & al., 2012] and has been shown to correlate with glaucoma progression [Medeiros & al., 2013]. As described in chapter 4, it is the inner retina that is usually affected in glaucoma patients. This suggests a possible link between the outer retina and glaucoma, however the association with corneal hysteresis could also be through a different pathway.

Another interesting spatial effect to note is that of *PLEKHA5* (rs10841116). *PLEKHA5* had nearly no effect on the thickness of the ONL in the peripheral fields of the macula (Effect size = 4.43×10^{-3}), with a slight effect in the intermediate field (Effect size = 0.14) and a larger effect at the fovea (Effect size = 0.57). *PLEKHA5* has previously been associated with localisation of melanoma metastasis to the brain with an apparent role in regulating the blood brain barrier [Jilaveanu & al., 2015]. The considerable difference in effect size on the different macula fields of the ONL could warrant further investigation into its role in the eye.

Table 5.5: List of SNPs differentially affecting the thickness of the ONL at the three concentric fields of the macula. List of SNPs with a significant z-score (P <0.05 following Bonferroni correction) describing the differential effect on the ONL thickness at the foveal (F), intermediate (I) and peripheral (P) fields. The field (F1 or F2) and corresponding effect size from GWAS of thickness in each field are listed alongside the Bonferroni adjusted p-value of the comparative z-score. Each genetic variant is also annotated with nearest gene and any ocular and non-ocular phenotypes previously associated with it.

SNP	Chr	F1	F2	F1 effect size	F2 effect size	Adjusted p-value	Nearest gene	Ocular phenotypes	Non-ocular phenotypes
rs36052911	2	F	I	0.56	0.21	0.05	*CCDC88A*		
rs11720108	3	F	I	0.71	0.12	8.37E-06	*ADCY5*		Birth weight, Forced expiratory volume, Heel bone mineral density, Height, Type 2 diabetes
rs55914544	5	F	I	0.65	0.09	1.26E-03	*WWC1*		Body fat percentage
rs75757892	6	F	I	-1.12	-0.55	2.85E-04	*RREB2*		Balding pattern, Heel bone mineral density, Height, Impedance of limbs, Red blood cell traits, Self reported math ability

Table 5.5: continued

rs13263941	8	F	I	0.22	0.69	1.05E-03	*RSPO2*	Strong/weak meridian	Coffee intake, Hair colour
rs1512380	9	F	I	-0.51	-0.12	0.04	*PTPRD*		
rs10841116	12	F	I	0.57	0.14	1.31E-03	*PLEKHA5*		Time spent outside in winter
rs1493529	13	F	I	-0.48	-0.08	6.18E-03	*ENOX1*		
rs12147951	14	F	I	-1.57	-2.48	2.91E-04	*VSX2*		
rs887595	14	F	I	1.28	1.94	5.51E-06	*VSX2*	Macular thickness	Red blood cell traits
rs3766775	1	F	P	-0.60	-0.06	1.19E-05	*SLC1A7*		Duration of fitness test
rs6701735	1	F	P	0.67	-0.07	7.20E-05	*MIR29B2CHG*	Corneal hysteresis	Pulse rate
rs36052911	2	F	P	0.56	0.10	3.86E-04	*CCDC88A*		
rs1494836	3	F	P	0.91	0.46	0.04	*EOMES*	Macular thickness	
rs73857644	3	F	P	-0.82	-0.18	0.04	*ZBTB20*		
rs11720108	3	F	P	0.71	-0.04	1.35E-10	*ADCY5*		Birth weight, Forced expiratory volume, Heel bone mineral density, Height, Type 2 diabetes
rs7623382	3	F	P	0.54	0.18	8.48E-03	*LINC00880*	Corneal hysteresis	BMI, Coronary artery disease, Height, Impedance of limbs
rs12515565	5	F	P	-0.07	-0.51	0.01E-03	*LMNB1-DT*		Height, Height at age 10, Multiple myeloma
rs55914544	5	F	P	0.65	-0.04	3.54E-06	*WWC1*		Body fat percentage
rs258873	5	F	P	-0.69	-0.31	6.23E-03	*BOD1*		
rs6875105	5	F	P	-0.69	-0.32	8.48E-03	*BOD1*		
rs75757892	6	F	P	-1.12	-0.20	2.57E-13	*RREB2*		Balding pattern, Heel bone mineral density, Height, Impednce of limbs, Red blood cell traits, Self reported math ability
rs4711420	6	F	P	0.72	0.29	0.01	*TULP1*		Height, Forced expiratory volume, Lymphocyte counts, Monocyte percenatge, Sum basophil neutrophil counts, Trunk mass, White blood cell count
rs4896369	6	F	P	0.44	0.05	2.29E-03	*HEBP2*		
rs4870049	6	F	P	0.55	0.11	5.81E-04	*LRP11*	Strong/weak meridian	
rs13263941	8	F	P	0.22	0.93	6.81E-10	*RSPO2*	Strong/weak meridian	
rs376067714	8	F	P	0.15	0.85	3.62E-05	*RSPO2*		
rs1512380	9	F	P	-0.51	0.02	1.08E-04	*PTPRD*		
rs7893507	10	F	P	0.45	0.03	1.56E-03	*CUBN*		
rs3906617	10	F	P	0.62	0.05	2.38E-06	*FRMPD2*	Age started wearing glasses, Spherical power	
rs10766368	11	F	P	0.56	0.14	3.27E-03	*PLEKHA7*	Corneal hysteresis, Glaucoma, IOP, Strong/weak meridian	Height, Impedance of limbs
rs1122316	11	F	P	0.17	-0.45	5.98E-05	*CCND1*		Breast cancer, Number of non-cancer illnesses
rs10841116	12	F	P	0.57	4.43E-03	4.54E-07	*PLEKHA5*		Time spent outside in winter

Table 5.5: continued

rs3138142	12	F	P	1.37	0.81	1.44E-05	*RDH5*	Age started wearing glasses, Cataract, Disorders of refraction and accommodation, Hypermetropia, Macular thickness, Myopia, Pigmentary retinal dystrophy, Spherical power	Atrial fibrillation
rs1493529	13	F	P	-0.48	0.07	1.47E-06	*ENOX1*		
rs12428305	13	F	P	-0.68	-0.21	6.68E-04	*IRS2*		Peripheral nerve disorders
rs3748358	14	F	P	-0.60	-0.20	1.84E-03	*RPGRIP1*		
rs12147951	14	F	P	-1.57	-2.35	3.71E-03	*VSX2*		
rs887595	14	F	P	1.28	1.94	1.29E-06	*VSX2*	Macular thickness	Red blood cell traits
rs1972565	14	F	P	1.27	1.94	1.20E-06	*VSX2*		
rs7183832	15	F	P	0.56	0.20	0.01	*BNIP2*	Myopia	
rs12515565	5	I	P	-0.18	-0.51	0.02	*LMNB1-DT*		Height, Height at age 10, Multiple myeloma
rs3138142	12	I	P	1.28	0.81	7.05E-06	*RDH5*	Age started wearing glasses, Cataract, Disorders of refraction and accomodation, Hypermetropia, Macular thickness, Myopia, Pigmentary retinal dystrophy, Spherical power	Atrial fibrilation

Inner Segment

As a reminder, the IS contains the mitochondria and ribsomomes of the PRC. Several of the genetic variants differentially affecting the thickness of the IS across the retinal fields had prior associations to neurodevelopmental traits (table 5.6). This includes *SLC44A4* (rs9267659) which has previously been strongly associated with multiple sclerosis and Parkinson's disease. *SLC44A4* had a very minor effect on inner segment thickness in the peripheral field (Effect size = 3.76×10^{-4}) and a stronger effect at the fovea (Effect size = -0.13). *SLC44A4* encodes a choline transporter and mutations in the gene have also been linked to hearing loss [Ma & al., 2017]. Additionally, *TMEM106B* (rs7791726 and rs1990622) has a stronger effect on IS thickness at the fovea (Effect size = -0.23) compared to in the peripheral field (Effect size = -0.02). This locus has prior associations to both dementia and depression. *TMEM160B* encodes a transmembrane protein which a study in mice reports has an apparent effect on myelination [Zhou & al., 2020]. There appears to be an overarching link between spatial variation in IS thickness and traits of CNS development.

RAX2 (rs76076446) is affecting the thickness of the IS layer more at the fovea (Effect size = 0.31) than the periphery (Effect size = -0.09). Mutations in this locus have been

linked to both cone and rod dysfunction [Yang & al., 2015] and retinitis pigmentosa, a rod-cone disease [Van de Sompele & al., 2019]. These two diseases have distinct and opposing patterns in which they affect the rods and cones. As stated in the introduction to this chapter, rods and cones have a specific localisation across the macula field with rods found most densely populated in the periphery and cones found at highest frequency in the foveal field. There appears to be a parallel between the spatial effects of this locus on thickness of the IS and the organisation of PRC affected by the diseases associated with this locus.

Table 5.6: List of SNPs differentially affecting the thickness of the IS at the three concentric fields of the macula. List of SNPs with a significant z-score (P <0.05 following Bonferroni correction) describing the differential effect on the IS thickness at the foveal (F), intermediate (I) and peripheral (P) fields. The field (F1 or F2) and corresponding effect size from GWAS of thickness in each field are listed alongside the Bonferroni adjusted p-value of the comparative z-score. Each genetic variant is also annotated with nearest gene and any ocular and non-ocular phenotypes previously associated with it.

SNP	Chr	F1	F2	F1 effect size	F2 effect size	Adjusted p-value	Nearest gene	Ocular phenotypes	Non-ocular phenotypes
rs4587551	1	F	I	-0.24	-0.10	0.05	*LRRC8D*		
rs78878924	4	F	I	0.38	0.17	0.05	*RUFY3*		
rs10479543	5	F	I	-0.14	-0.05	0.03	*RUFY1*		
rs7791726	7	F	I	-0.23	-0.09	2.47E-09	*TMEM106B*		Coronary artery disease, Depression, Frontotemporal dimentia, Height, Irritability, Lethargy, Loneliness, Mood swings, Neuroticism
rs1990622	7	F	I	-0.23	-0.09	2.62E-09	*TMEM106B*		Coronary artery disease, Depression, Differential ageing of the frontal cortex, Height, Irritability, Loneliness, Mood swings
rs56280639	8	F	I	0.01	0.12	0.03	*PRSS51*	Strong meridian	Anxious feelings/worrier, Forced expiratory volume Heel bone mineral density, Neutrophil count, Platelet distribution width, Red blood cell distribution width, White blood cell count
rs141629142	9	F	I	0.40	0.11	0.01	*RORB*		
rs1409396	10	F	I	0.13	0.06	0.02	*PIP4K2A*		Frequency of walking for pleasure, Platelet disrtibution width, Sitting height
rs11200920	10	F	I	0.12	0.02	4.33E-03	*CDHR1*		
rs55634267	17	F	I	0.20	0.11	2.06E-03	*GNGT2*	Spherical power	Angina, Anthropometric traits, Asthma, Coronary artery disease, Forced expiratory volume, High blood pressure, Impedance of limbs

Table 5.6: continued

rs76076446	19	F	I	0.31	0.05	9.47E-03	*RAX2*	Cone-Rod Dystrophy, Macular degerenation, Macular thickness	
rs11701177	21	F	I	0.15	0.07	0.02	*ZNF295-AS1*		
rs12128178	1	F	P	-0.10	-0.02	3.20E-03	*KDM1A*		Anthropometric traits, Forced expiratory volume, Hayfever, Mean corpuscular volume, Mean reticulocyte volume
rs4587551	1	F	P	-0.24	-0.04	1.38E-04	*LRRC8D*		
rs6763422	3	F	P	-0.26	-0.06	1.05E-16	*TSC22D3*	Age started wearing glasses	
rs78878924	4	F	P	0.38	0.02	5.67E-07	*RUFY3*		
rs72761109	5	F	P	-0.02	0.08	2.26E-04	*MAP1B*	Age started wearing glasses	
rs10479543	5	F	P	-0.14	-0.03	1.72E-03	*RUFY1*		
rs9267659	6	F	P	-0.13	3.76E-04	6.08E-06	*SLC44A4*		Asthma, Diabetes, Height, Hyperthyroidism, Intestinal malabsorption, Forced vital capacity, Insulin, Mean platelet volume, Multiple sclerosis, Overall health rating, Parkinson's disease, Red blood cell traits
rs2749938	6	F	P	0.11	0.02	5.98E-03	*MED23*		
rs7791726	7	F	P	-0.23	-0.02	1.67E-24	*TMEM106B*		Coronary artery disease, Depression, Frontotemporal dimentia, Height, Irritability, Lethargy, Loneliness, Mood swings, Neuroticism
rs4731631	7	F	P	-0.10	-0.02	4.34E-03	*UBE2H*		
rs141629142	9	F	P	0.40	0.06	6.59E-04	*RORB*		
rs3904960	9	F	P	-0.12	-0.02	0.01	*UCK1*		Anthropometric traits, Basal metabolic rate, Impedance of leg, Snoring
rs1409396	10	F	P	0.13	0.02	5.49E-07	*PIP4K2A*		Frequency of walking for pleasure, Platelet disrtibution width, Sitting height
rs11200920	10	F	P	0.12	-0.04	3.69E-10	*CDHR1*		
rs1255910	14	F	P	0.10	0.03	0.04	*SYNE2*		Facial aging, Hair colour
rs887595	14	F	P	0.08	0.20	6.19E-04	*VSX2*	Macular thickness	Mean corpuscular volume/ haemoglobin,Red blood cell width
rs1201689	15	F	P	-0.10	-0.01	1.56E-03	*MAPKBP1*		
rs55634267	17	F	P	0.20	0.05	7.81E-10	*GNGT2*	Spherical power	Angina, Anthopometric traits, Asthma, Coronary artery disease, Forced expiratory volume, High blood pressure, Impedance of limbs
rs35638197	17	F	P	-0.31	-0.12	3.51E-04	*GNGT3*	Spherical power	
rs76076446	19	F	P	0.31	-0.09	1.27E-07	*RAX2*	Cone-Rod Dystrophy, Macular degerenation, Macular thickness	
rs231235	19	F	P	-0.09	-0.01	4.94E-03	*ARHGAP33*		

Table 5.6: continued

rs11701177	21	F	P	0.15	0.02	1.77E-07	*ZNF295-AS1*		
rs2282526	21	F	P	0.17	0.02	2.70E-11	*RRP1B*		Blood pressure, Hayfever, Hypertension, Red blood cell traits
rs1463229	3	I	P	-0.18	-0.06	2.60E-08	*TSC22D2*	Age started wearing glasses	
rs1990622	7	I	P	-0.09	-0.02	3.82E-04	*TMEM106B*		Coronary artery disease, Depression, Differential ageing of the frontal cortex, Height, Irritability, Loneliness, Mood swings
rs2994042	9	I	P	-0.10	-0.03	0.05	*PRRC2B*		Anthropometric traits, Basal metabolic rate, Impedance of leg, Postlaminectomy syndrome
rs969060	21	I	P	0.11	0.02	5.44E-06	*RRP1B*		Blood pressure, Hypertension, Red blood cell traits

Outer Segment

Numerous loci differentially affecting the thickness of the OS layer across the concentric retinal fields were noted to have involvement in AMD (table 5.7). This included the well established AMD locus *CFH* [Klein & al., 2005] which has an increasingly strong effect on the thickness of the OS layer as you move inwards from the peripheral field (Effect size = 0.24) to the fovea (Effect size =0.64). As previously mentioned, *CFH* encodes complement factor H. Complement is an integral part of the innate immune response, aiding removal of bacterial and apoptotic cell debris as well as recruiting immune cells to sites of infection [Landowski & al., 2019]. This is of note as the OS interfaces with the RPE whose responsibilities include the disposal of waste from outer retinal layers. Further loci associated with AMD that were also differentially affecting OS thickness across the macular field include *ARMS2* and *HTRA1*. These two genes are located in the same region on chromosome 10. The two genes have a known effect on AMD [Kortvely & al., 2010]. Both *ARMS2* and *HTRA1* effect the thickness of OS most strongly at the fovea (Effect size = -0.75 and -0.74 respectively), with less of an effect as you move outwards towards the peripheral field (Effect size = -0.24 and -0.24 respectively). *ARMS2* has been shown to be involved in proper function of the extracellular matrix and has a suggested protective role against drusen formation [Kortvely & al., 2010]. *HTRA1* is a heatshock protease [DeWan & al., 2006] that is involved in the processing of extracellular matrix proteins. It is thought that it may help regulate inhibitors of angiogenesis and thus the neovascularisation seen in AMD [Lin & al., 2018]. AMD affects the central retina and spares peripheral vision [Khurana, 2007]. The spatial effect of these loci on

thickness of the OS appears to correlate with the spatial effect of AMD across the retina.

Another interesting locus that is differentially affecting OS thickness across the macular field is *RDH5* (rs3138142). *RDH5* appears to affect the foveal and intermediate fields to the same magnitude (Effect size = 0.76 and 0.78, respectively), and the peripheral field less (Effect size = 0.32). This genetic variant has previously been associated with a number of ocular phenotypes including macular thickness, pigmentary retinal dystrophy and spherical equivalent. *RDH5* has also been shown to be involved in fundus albipunctatus. Fundus albipunctatus is a form of night blindness that affects the rods [Skorczyk-Werner & al., 2015]. *RDH5* encodes 11-*cis*-retinol dehydrogenase which is expressed in the RPE and is an enzyme involved in the processing of the chromophore of mammalian visual pigments [Lidén & al., 2001]. The OS is the storage location of photopigments in the retina, and thus the differential effect of this locus on OS thickness and its association with the photopigment production pathway is of note.

MERTK has close to no effect on the thickness of the OS in the peripheral field (Effect size = 3.98×10^{-4}), but a moderate effect at the intermediate field (Effect size = 0.13), and an increased effect at the fovea (Effect size = 0.30). *MERTK* encodes a receptor localised to the RPE surface, which interfaces with the OS. Mutations in *MERTK* have been found in rod-cone and cone-rod dystrophies [Audo & al., 2018]. Again, the density of rods and cones is different across the retina, with cones being found at high density in the foveal field, and the rods being found most densely in the peripheral field. This distribution of photoreceptors corresponds to the spatial observations of the effect of this locus on thickness of the OS across the macula.

Table 5.7: List of SNPs differentially affecting the thickness of the OS at the three concentric fields of the macula. List of SNPs with a significant z-score (P <0.05 following Bonferroni correction) describing the differential effect on the OS thickness at the foveal (F), intermediate (I) and peripheral (P) fields. The field (F1 or F2) and corresponding effect size from GWAS of thickness in each field are listed alongside the Bonferroni adjusted p-value of the comparative z-score. Each genetic variant is also annotated with nearest gene and any ocular and non-ocular phenotypes previously associated it.

SNP	Chr	F1	F2	F1 effect size	F2 effect size	Adjusted p-value	Nearest gene	Ocular phenotypes	Non-ocular phenotypes
rs1329428	1	F	I	0.64	0.42	5.92E-03	*CFH*	Age related macular degeneration, Chronic central serous retinopathy, Macular degeneration	Forced vital capacity
rs4848901	2	F	I	0.30	0.13	0.04	*MERTK*		
rs3750846	10	F	I	-0.75	-0.48	2.28E-03	ARMS2	Advanced age related macular degeneration	

Table 5.7: continued

rs3793917	10	F	I	-0.74	-0.47	2.70E-03	*HTRA1*	Age related macula degeneration, Spherical power	Anthropometric traits, Basal metabolic rate, Stroke
rs72683442	1	F	P	-0.57	-0.29	1.41E-03	*SLC16A1*	Spherical power	Birth weight
rs1329428	1	F	P	0.64	0.24	3.35E-11	*CFH*	Age related macular degeneration, Chronic central serous retinopathy, Macular degeneration	Forced vital capacity
rs410895	1	F	P	0.58	0.25	2.36E-07	*CFHR2*	Macular degeneration	
rs4848901	2	F	P	0.30	3.98E-04	2.07E-06	*MERTK*		
rs288282	2	F	P	-0.28	-0.08	0.03	*DNAJC10*		
rs1864251	2	F	P	-0.74	-0.34	1.33E-03	*MREG*		
rs148388367	2	F	P	1.25	0.51	3.70E-05	*MREG*		
rs556679	6	F	P	0.56	0.18	5.68E-04	C2		Ankle spacing width Anthropometric traits, Hyperthyroidism, Platelet count, Psoriasis
rs3750846	10	F	P	-0.75	-0.24	1.10E-12	*ARMS2*	Advanced age related macular degeneration	
rs60401382	10	F	P	-0.70	-0.24	1.11E-10	*HTRA1*	Macular degeneration	Anthropometric traits, Basal metabolic rate, Intelligence, Stroke
rs1126809	11	F	P	-0.49	-0.29	0.04	*TYR*	Eye colour, Foveal hypoplasia Oculocutaneous albinism, Optic disc size	Bipolar disorder, Hair colour, Hearing difficulties, IOP, Neuroticism, Skin cancer, Skin colour, Tanning, Vitiligo
rs3138141	12	F	P	0.76	0.32	3.14E-11	*RDH5*	Age related macular degeneration, Cataract, Disorders of refraction and accomodation, Myopia, Spherical power	Atrial fibrilation
rs1254257	14	F	P	-0.44	-0.19	2.09E-03	*PPM1A*	Age started wearing glasses, Glaucoma, Spherical power	Anthropometric traits, Basal metabolic rate, Forced expiratory volume, Heel bone mineral density, Impedance of limbs, Menarche, Sodium in urine
rs1254294	14	F	P	-0.44	-0.19	2.21E-03	*SIX6*	Age started wearing glasses, Glaucoma, Spherical power	Anthropometric traits, Basal metabolic rate, Forced expiratory volume, Impedance of legs Menarche, Sodium in urine
rs142963458	16	F	P	-1.21	-0.69	0.04	*MEAK7*	Macular thickness	
rs62125245	19	F	P	0.44	0.11	1.95E-044	*NXNL1*		
rs72683442	1	I	P	-0.49	-0.29	0.02	*SLC16A1*	Spherical power	
rs1329428	1	I	P	0.42	0.24	4.11E-03	*CFH*	Age related macular degeneration, Chronic central serous retinopathy, Macular degeneration	
rs148388367	2	I	P	0.98	0.51	0.01	*MREG*		
rs3793917	10	I	P	-0.47	-0.24	1.70E-03	*HTRA1*	Age related macular degeneration, Spherical power	

Table 5.7: continued

rs60401382	10	I	P	-0.46	-0.24	0.00376442	*HTRA1*	Macaulr degeneration
rs1126809	11	I	P	-0.49	-0.29	3.95E-03	*TYR*	Eye colour, Foveal hypoplasia, Oculocutaneous albinism, Optic disc size
rs3138142	12	I	P	0.78	0.32	1.35E-15	*RDH5*	Age started wearing glasses, Cataract, Disorders of refraction and accomodation, Hypermetropia, Macular thickness, Pigmentary retinal dystrophy, Myopia, Spherical power
rs3138141	12	I	P	0.76	0.32	2.37E-15	*RDH5*	Age related macular degeneration, Cataract, Disorders of refraction and accomodation, Myopia, Spherical power
rs7405453	17	I	P	-0.55	-0.37	0.01	*TSPAN10*	Age started wearing glasses, Amblyopia, Astigmatism, Cataract, Cylinderical power, IOP, Macular thickness, Myopia, Sphierical power, Visual acuity (Logmar)

Retinal Pigment Epithelium

Many of the loci with significantly different effects on the thickness of the RPE layer across the fields of the macula had prior associations with ocular traits (table 5.8). These include *OCA2* (rs1800407) which has a similar magnitude of effect on the foveal and intermediate fields of the retina (Effect size = -1.20), but less of an effect on the periphery field (Effect size = -0.75). *OCA2* is a locus with well-known involvement in OCA. As mentioned previously, OCA causes decreased amounts of pigment in the RPE, leading to impaired visual acuity. This is in addition to foveal hypoplasia, the flattening of the macular dip that also causes impaired vision. It is interesting to note that the increased effect of the SNP corresponds to the shape of the macula dip. A similar pattern can be seen for *TSPAN10* (rs7405453), which has similar magnitude of effect on the foveal and intermediate fields (Effect size =0.58 and 0.57 respectively), and less of an effect on the peripheral field (Effect size = 0.27). *TSPAN10*, similarly to *OCA2*, is a locus involved in pigmentation. These results suggest an interesting relationship between spatial effect of pigmentation and retinal morphology.

Another locus that affects the RPE layer thickness differentially at the fovea compared to the peripheral field is *MREG*. *MREG* has double the effect on the RPE thickness at the fovea (Effect size = -1.24) compared to the peripheral field (Effect size = -0.45).

MREG has been shown to play a role in retinal health, regulating the homoeostasis of OS degradation [Frost & al., 2013]. One of the main roles of the RPE is the disposal of redundant PRC components. Therefore it is interesting to see a locus with a known role in this pathway differentially affecting the thickness of the RPE across the macula field.

Table 5.8: List of SNPs differentially affecting the thickness of the retinal pigment epithelium at the three concentric fields of the macula. List of SNPs with a significant z-score ($P < 0.05$ following Bonferroni correction) describing the differential effect on the RPE thickness at the foveal (F), intermediate (I) and peripheral (P) field. The field (F1 or F2) and corresponding effect size from GWAS of thickness in each field are listed alongside the Bonferroni adjusted p-value of the comparative z-score. Each genetic variant is also annotated with nearest gene and any ocular and non-ocular phenotypes previously associated it.

SNP	Chr	F1	F2	F1 effect size	F2 effect size	Adjusted p-value	Nearest gene	Ocular phenotypes	Non-ocular phenotypes
rs148388367	2	F	P	-1.24	-0.45	0.04	*MREG*		
rs7405453	17	F	P	0.58	0.27	0.04	*TSPAN10*	Age started wearing glasses, Amblyopia, Astigmatism, Cataract, Corneal resistance, IOP, Myopia, Refractive error	Ankle spacing width, Hair colour, Impedance of leg, Macular thickness, Tanning, Snoring
rs1800407	15	I	P	-1.20	-0.75	0.04	*OCA2*	Age started wearing glasses, Cataracts, Eye colour, Macular thickness, Oculocutaneous albinism, Refractive error, Retinal vasculature	Age of first facial hair, Hair colour, Skin cancer, Skin colour, Sunburn, Tanning
rs9895741	17	I	P	0.57	0.27	0.01	*TSPAN10*	Age started wearing glasses, Amblyopia, Astigmatism, Catract, Corneal resistance, Spherical power, Strabismus	Ankle spacing width, Hair colour, Impedance of leg, Tanning

5.6. Discussion

Here, I have conducted the first large-scale GWAS of outer retinal layer thickness. In-keeping with our analysis of the inner retina, the main GWAS were conducted on the simple phenotype of mean thickness across the macula of the outer retinal layers - the ONL, IS, OS and RPE - averaged across left and right eyes. Due to shared associations between the three component PRC layer - the ONL, IS and OS - were meta-analysed and considered together in a single robust discovery list. The results of the RPE thickness GWAS were considered separately due to the different cell type found in this layer. 59 loci were found robustly associated to one or more of the PRC component layers. Many of these loci have prior associations with ocular phenotypes including refractive error and age one started wearing glasses, in addition to ocular pathologies such as AMD

and retinitis pigmentosa. There were also a number of genetic variants significantly associated with the component PRC layer thicknesses that had no prior annotations. These loci may warrant further investigation as to their role in eye biology. Similarly, seven loci were found to be significantly associated with RPE thickness. These loci also had prior associations with ocular pathologies including oculocutaneous albinism and AMD.

MR analysis revealed support for a causal relationship between AMD and the thickness of the RPE. A Steiger test confirmed the directionality to be AMD affecting RPE thickness. The RPE is the layer under which drusen develop during AMD, causing damage to the outer retina and subsequent visual impairment. Support was not found for a causal relationship between outer retinal thickness as an exposure variable and AMD as an outcome variable. These results mirror the relationship seen between inner retinal thickness and glaucoma. It suggests that, although AMD causes thinning of outer retinal layers as the disease progresses, the baseline thickness of the outer retinal layers as determined by genetics is not indicative of risk for AMD development.

Enabled by the use of the ETDRS grid, I explored spatial differences of the effect of genetic variants on outer retinal thickness. For each of the retinal layers, there were notable differences in discovery of loci across each macular subfield, representative of concentric circles outwards from the fovea. Further interrogation produced a comprehensive list of loci with differential effect on the retinal sub-fields for each outer retinal layer. It must be acknowledged that this is highly exploratory research. In line with the nature of such comparative statistical analyses, there is the risk that probabilistic chance may have influenced the discovery of variants. However the loci discovered included those with prior associations with ocular diseases such as inherited retinal dystrophies and AMD. Interestingly, several of the loci also had prior associations to neurodevelopmental diseases such as multiple sclerosis and dementia. This highlights the role of the eye in the CNS and the potential of using ocular imaging as a biomarker for a broad array of diseases.

These results emphasize the additional biology explained by a more granular phenotype, that is enabled by the use of imaging data from the clinical segmentation grids. A finer grain understanding of the spatial genetic effect on retinal morphology may provide further understanding of the biological pathways underpinning ocular development and pathology. Such finer grain representations of retinal morphology are explored in the following chapter.

6. Higher Dimensional Phenotyping

The OCT images used throughout this thesis contain rich high dimensional data describing the morphology of the retina. Thus far in this thesis I have focused on studying simple phenotypes which are already utilised in the clinical environment and which have the advantage of being more easily interpreted when discussing clinical aspects of results. I now explore more of the high-dimensional data afforded by OCT images, harnessing the pixel-wise information to study variation in retinal morphology at a finer grain level. In doing so I aim to capture the continuous variation across the macula field, enabling new and interesting biological findings.

There are several examples of high dimensional phenotypes from medical images being used in conjunction with genetics to interrogate the biology of human organs. Notable examples include the use of Magnetic Resonance Imaging (MRI) derived measures in GWAS that led to novel associations of 148 clusters of SNPs with brain structure and function [Elliott & al., 2018]. This study bears similarity to ours in that we both aim to use complex image derived phenotypes to study both general human biology and disease. Another example includes a study of heart morphology that provided further evidence as to the function of heart trabeculation [Meyer & al., 2020]. In this case, machine learning segmentation of heart MRI were used to identify phenotypes for use in GWAS. Given these contributions, this high-dimensional complex phenotype approach appears to hold value in making new biological discoveries.

The retina, as previously discussed, has a specific and highly organised structure which is integral to its function. As well as being organised into a number of distinct layers, the retina has complex topographical variation. The macula dip, a valley like structure in the centre of the retina, is important to visual acuity. The distribution of photoreceptors also varies across the macular field. At the central area of the macula, the fovea, only cones are present which are responsible for high acuity colour vision in high light levels. These reduce in density moving towards the peripheral retina. Conversely, rods increase in density in the peripheral fields of the retina. The rods are responsible for vision in low light levels, but lack colour definition [James & al., 2003].

Considering this interesting topography, examination of the morphology of the retina

at a higher dimensionality provides new opportunities for biological discovery. This was initially explored by analysing the ETDRS grid in chapter 5, where even increasing the dimensionality of the data from one to three provided new information about spatial biological variation and its underlying genetics.

Therefore, looking at alternative representations of the eye is an interesting path of exploration to understand the genetic variation underlying a finer-grain measure of retinal morphology. In this chapter, I present the techniques used to obtain these novel representations of retinal morphology, and explore the biology and genetics that underlies them.

6.1. Maculagrams

To aid easy clinical interpretations of results within this thesis thus far, we collapse the data describing retinal morphology from pixel-wide magnitude to a single measure per retinal layer. In chapter 5 we utilised more of the data available from the TABS output to expand the dimensionality and compare effects of SNPs on the thickness of outer retinal layers at the foveal, intermediate and peripheral fields.

I now look to extend this in two stages. First I aim look at the spatial morphological variation in the individual sections of the two segmentation grids, the Macula 6 grid and the ETDRS grid. I then will study the underlying genetics of the variations across the segmentation grids. As a reminder, the Macula 6 grid segments the macula into six radial segments, excluding the fovea, and is generally used to study the inner retina. The ETDRS grid is a nine space segmentation grid that resembles a bullseye and includes the foveal field. It is mainly used to study the outer retina.

To enable this analysis, I created a novel visualisation technique for the macular area, coined in the research group as maculagrams. These maculagrams allow the plotting of principal component loadings and genetic effect sizes back into the macular grid space. This enables visualisation of variation patterning across the macula.

6.1.1. Principal Component Maculagrams

To examine how retinal layer thickness is varying across the macular field, I calculated the principal components (PCs) of the measurements that are the composite of the single measures used in the simple clinical GWAS. These measurements comprise the thickness of the individual retinal layers at each segment of either of the segmentation grids - the Macula 6 or the ETDRS grid - across left and right eyes separately. The loadings of these PCs were mapped back into the original space in the plots we have termed

maculagrams. The aim of this is to identify which areas of the macula are contributing the most to morphological variation and interrogate any distinct spatial patterns.

Methods

PCA was conducted on three sets of measures. Firstly the thickness of the GCIPL and RNFL at each of the sections of the Macula 6 grid in both the left and right eye (24 dimensions, figure 4.2B). Secondly the thickness of the ONL, IS and OS at each of the sections of the ETDRS grid in both the left and right eye (54 dimensions, figure 5.2C). And finally the thickness of the RPE at each of the sections of the ETDRS grid in both the left and right eye (18 dimensions, figure 5.2C). In each case the measures were scaled and centred prior to PCA. The loadings of the first 10 PCs for each is returned and plotted back to the original grid space. The grids are oriented so that, for both eyes and both segmentation grids, the easterly segments represent the temporal segments of the eye (those closest to the ear) and the westerly segments represent the nasal sides (those closest to the nose).

Results

There are several patterns of variation that are common to all three of the retinal areas. Firstly, in each of the three retinal layers, one of the PCs corresponds to variation between the left and right eyes. This is likely evidence of confounding caused by scan acquisition order, as the right eye was consistently scanned first [Patel & al., 2016], introducing a batch like effect. Further analysis of the difference between the two eyes can be found in section 4.3. In the GWAS conducted in the previous chapters of this thesis, this is controlled for by regressing out eye specific traits such as technical measures outputted from the OCT machine and by making a mean across the two eyes.

Across the layers, several PCs also appear to occur in pairs. A feature of PCA is that each PC must be orthogonal to one another. In the spatial PC analysis of each of the three retinal areas, one of the first 10 PCs has been representative of variation between left and right eyes. The presence of a PC representing this variation forces all other PCs to honour this asymmetry. Therefore it is possible that some of the PCs represent the same pattern of spatial variation, split across the two eyes.

Inner Retinal Principal Component Spatial Analysis

The first ten PCs of the inner retinal layer component thicknesses account for 90% of variation (figure 6.1).

The first PC appears to affect all components of the inner retinal thickness equally and in the same direction. Comparatively the second PC appears to be driven by variation between the left and right eye, as discussed above.

The third PC appears to be driven by variation between the two inner retinal layers, the RNFL and the GCIPL. This is represented in the GWAS results where, although there is a large crossover of significantly associated loci between the two layers, some loci are uniquely associated with one of the layers.

Many of the other PCs show complex patterning of loadings. The fourth and sixth PCs have an interesting pattern in that they appear to show differential and antagonistic effect of the superior and inferior segments of the eye. Notably, changes in inner retinal morphology across this midline are a key feature used in clinical glaucoma diagnosis [Gupta & Asrani, 2016].

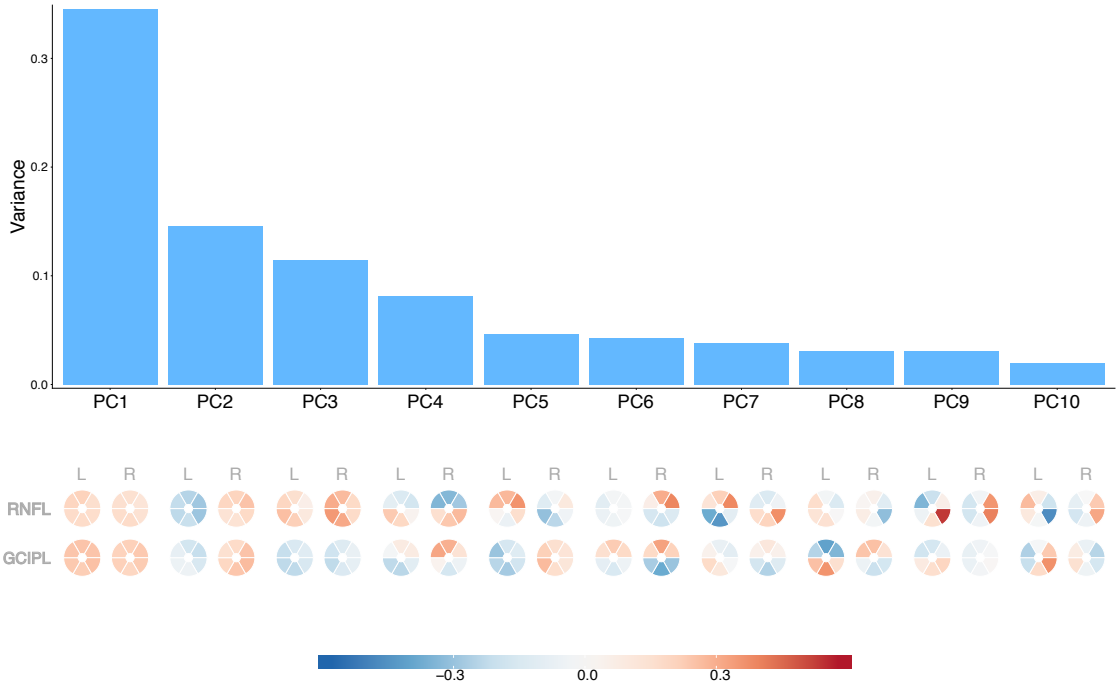

Figure 6.1: Principal component analysis of inner retinal thickness and spatial mapping Principal component analysis (PCA) was conducted on the thickness of the inner retinal layers - the retinal nerve fibre layer (RNFL) and the ganglion cell inner plexiform layer (GCIPL) - at the 6 segments of the Macula 6 grid across left (L) and right (R) eyes. The variance explained by each of the first ten principal components is plotted in the top panel. The loadings of the first ten principal components are plotted back to the original Macula 6 grid space using maculagrams. The saturation of colour indicates the magnitude of the effect a segment is having on a PC, with direction of effect (red or blue) being arbitrary. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

Photoreceptor Cell Principal Component Spatial Analysis

The first 10 PCs of the PRC component layer thicknesses account for 83% of variation (figure 6.2).

The first two PCs appear to be mostly driven by an antagonistic effect on the ONL and IS compared to the OS. Although inverse of one another, the first two PCs share a common pattern of variation. Considering that PCs are inherently orthogonal, the biology driving this variation can be assumed to be complex. Comparatively the third PC appears to be affected by the IS and OS, more so by the IS, in the same direction, and the ONL in the opposite direction. The three component layers of the PRC represent different parts of the same cell. These results may therefore indicate a complex relationship between the different parts of the PRC and their unique functions.

The fourth PC represents variation between the left and right eyes which, as noted in the previous section, is likely evidence of confounding by scan acquisition order.

Other interesting patterning can be seen in PC6, where in the ONL and IS, there appears to be an antagonistic relationship between the peripheral fields of the macula and the central foveal area. This finding may reinforce the results of the spatial GWAS conducted on the PRC component layers, in which several loci were shown to affect the foveal area, intermediate ring and peripheral ring differentially (section 5.5). This also aligns with the spatial differences in PRC density across the macula, with cones being more densely populated at the fovea, and rods being present in higher quantities in the periphery [Purves & al., 2001].

Retinal Pigment Epithelium Principal Component Spatial Analysis

Following a PCA of the component thicknesses of the RPE in the segments of the ETDRS grid across left and right eyes, the first two PCs explain 89% of variation (figure 6.3). This is a stark pattern which may be due to the relatively uniform nature of the RPE across the macula.

Similarly to the inner retina, the first PC is being influenced by all components of the RPE in the same direction and to roughly the same magnitude. Again, the second PC is representative of variation between the left and right eyes.

The third PC appears to be affected differentially by the peripheral field, as compared to the intermediate field and the foveal field. This is supportive of the spatial effects seen in the GWAS of these concentric fields in section 5.5. This aligns with the distribution of rods and cones across the macula as well as the structure of the macula's valley like structure.

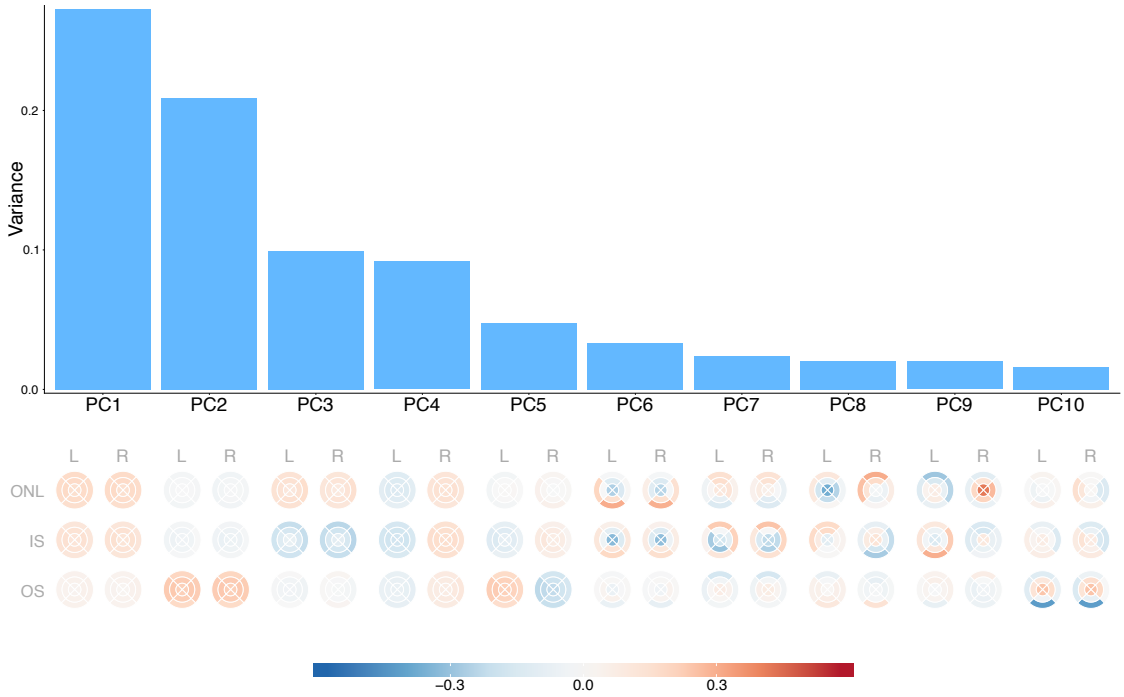

Figure 6.2: Principal component analysis of the component photoreceptor cell layer thicknesses and spatial mapping Principal component analysis was conducted on the thickness of the photoreceptor cell component layers - the outer nuclear layer (ONL), the inner segment (IS) and the outer segment (OS) - at the 9 segments of the ETDRS grid across the left (L) and right (R) eyes. The variance explained by each of the first ten principal components is plotted in the top panel. The loadings of the first ten principal components are plotted back to the original ETDRS grid space using maculagrams. The saturation of colour indicates the magnitude of the effect a segment is having on a PC, with direction of effect (red or blue) being arbitrary. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

There are several obvious examples of variation patterning that appear to have been split into left and right eyes by the orthogonality of PCA as detailed above. For example PC5 and PC6 appear to occur in a pair that represent the same pattern of spatial variation, just separated by eye. In this instance both PC5 and 6 show an opposing influence from the inferior quadrant and the superior quadrant of the grid in one eye.

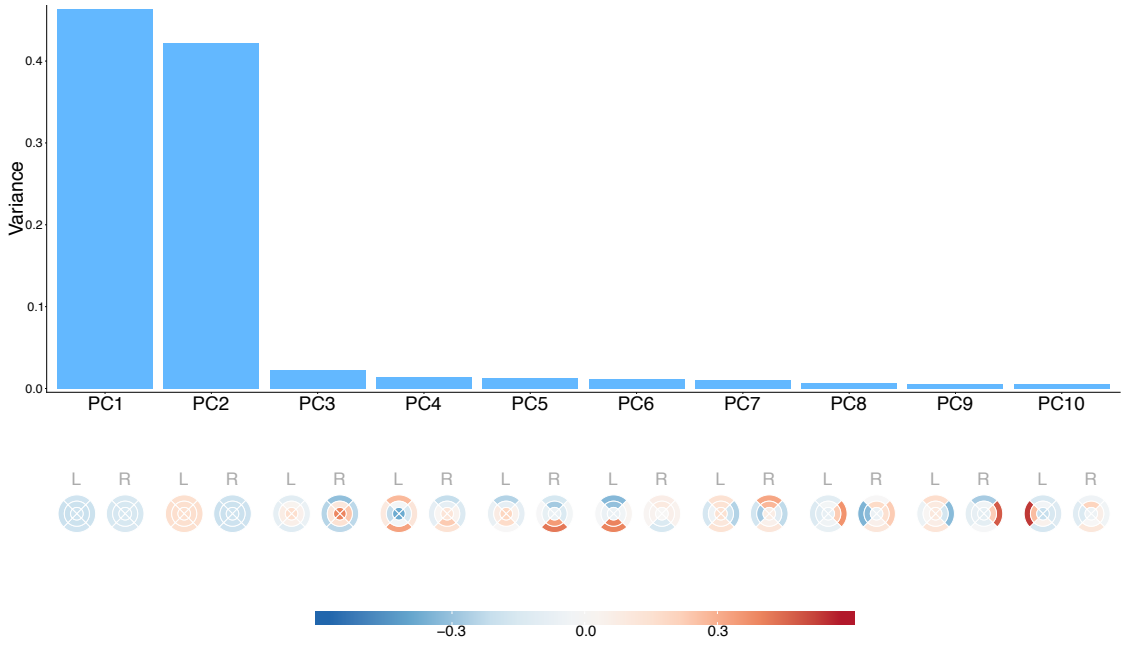

Figure 6.3: Principal component analysis of retinal pigment epithelium thickness and spatial mapping Principal component analysis was conducted on the thickness of the retinal pigment epithelium (RPE) layer at the 9 segments of the ETDRS grid across the left (L) and right (R) eyes. The variance explained by each of the first ten principal components is plotted in the top panel. The loadings of the first ten principal components are plotted back to the original ETDRS grid space using maculagrams.The saturation of colour indicates the magnitude of the effect a segment is having on a PC, with direction of effect (red or blue) being arbitrary. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

6.1.2. Higher Dimensional Genetic Variation

Considering the interesting patterns of spatial variation in morphology observed using the maculagrams in the previous section, it can be hypothesised that the genetic variation underlying retinal morphology may be operating at a similarly high dimension. I previously explored utilising PCA derived phenotypes to represent inner retinal mor-

phology at a higher dimension in section 4.4. In that instance there was not a notable increase in discovery power, most probably due to the cohort size. Therefore I opted to use a simpler clinically utilised measure which aided easier interpretation. However, exploring a higher dimensional space is an interesting route to further understand SNPs discovered in the GWAS of the simpler phenotypes. Therefore I next analysed the way in which the SNPs found to be significantly associated with the three retinal areas - the inner retina , the PRC, and the RPE - are affecting the morphology of the layers component areas across both eyes. This is a direct extension of the work in section 5.5 where I looked at three dimensions of data per layer of the PRC. I now move to study dimensions of 24, 54 and 18 for the inner retina, PRC component layers and RPE respectively.

Methods

The three areas of the retina were considered separately: the inner retina, the PRC and the RPE. Firstly, the thickness of the component inner retinal layers, the GCIPL and RNFL, at each of the sections of the Macula 6 grid across the left and right eyes (24 measures) were linearly modelled as a result of each of the 46 SNPs significantly associated with the thickness of one or both of the meta analysed inner retinal layers. Secondly, the thickness of the component PRC layers - the ONL, IS and OS - at each of the sections of the ETDRS grid across left and right eyes (54 measures) was linearly modelled as a result of each of the 59 SNPs significantly associated with the thickness of one or more of the meta-analysed component PRC layers. Thirdly, the thickness of the RPE at the component sections of the ETDRS grid across left and right eyes (18 measures) was linearly modelled as the outcome of each of the 7 SNPs significantly associated with overall RPE thickness. In all of these models, height, age, weight, sex, ID of the OCT machine used in the scan and the first 20 genotype PCs were used as covariates. The effect sizes of the SNPs were plotted back into the original grid space from which they were calculated as maculagrams. Once again the grids are oriented so that for both eyes and both grids, the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

Spatial Variation of Genetic Effect on the Inner Retina

I observed various interesting patterns of variation in genetic effect on the inner retinal layers across the macula field (figure 6.4). For several of the SNPs, a difference was seen in their effect on the left and right eyes. For example, *KIF6* (rs527871768, figure 6.4A), *IKZF1* (rs73348111, figure 6.4B) and *ANGPT1* (rs11520750, figure 6.4C) all showed no-

ticeably different effect size and direction between left and right eyes. The genetic effect on the two eyes was explored in section section 4.3. The variation underlying this differential genetic effect is likely due to a lack of randomisation in the order that the left and right OCT scans were taken, with the right eye routinely being taken first [Patel & al., 2016], as discussed in section 6.1.1.

Some of the patterning delineates the effect of loci on a specific inner retinal layer. *GNB3* (rs5442, figure 6.4G) and *SLC6A20* (rs17279437, figure 6.4H) have a stronger effect on the GCIPL than the RNFL. This is reflected in their p-values from the GWAS of the individual layer, *GNB3* was found significantly associated with GCIPL ($P=2.36 \times 10^{-13}$) but not in the RNFL ($P=0.27$), and *SLC6A20* was found to have a much stronger association with GCIPL ($P=7.81 \times 10^{-24}$) compared to the RNFL ($P=1.23 \times 10^{-7}$). *LINC00461* (rs17421627, figure 6.4K) shows the inverse pattern, in that it has a stronger effect on the RNFL than the GCIPL. This is once again reflected in the p-values obtained from the GWAS of the individual inner retinal layer. *LINC00461* had a much stronger association with the RNFL ($P = 8.09 \times 10^{-27}$) than the GCIPL ($P=4.87 \times 10^{-2}$).

RSPO2 (rs376067714 and rs13271359, figure 6.4 N and O respectively) shows a slightly stronger effect on the nasal side (western side of the plot) of the RNFL than the temporal side. The nasal side of the retina is the side from which the optic nerve exits the eye [Freddo, 2018]. *RSPO2* encodes R-spondin-2 which acts as an activator of the Wnt signalling pathway [Glinka & al., 2011]. *RSPO2* has previously been associated with retinal cell differentiation and definition [Takata & al., 2017]. This suggests a potential involvement of this locus in the organisation of the neural structures within the retina.

Spatial Variation of Genetic Effect on the Photoreceptor Cell Component Layers

Some distinct patterns of spatial genetic effect on the component PRC layers can be noted (figure 6.5). Firstly, there are several loci that appear to have much stronger effects on the ONL compared to the IS and OS. These include *LIN52* (rs1972564 and rs11159063, figure 6.5 A and B respectively). The ONL comprises the cell body of the PRC, and is also the location of the synaptic terminal [James & al., 2003]. *LIN52* has previously been linked to regulation of the cell cycle [Litovchick & al., 2011]. This speaks to the complex relationship between the different sectors of the PRC and the developmental pathways that underlie them.

RDH5 (rs3138144, figure 6.5C) shows an interesting pattern of effect on the ONL and OS. In these layers the SNP appears to be affecting the foveal and intermediate ring more so than the peripheral ring. A similar finding was made in chapter 5 when comparing

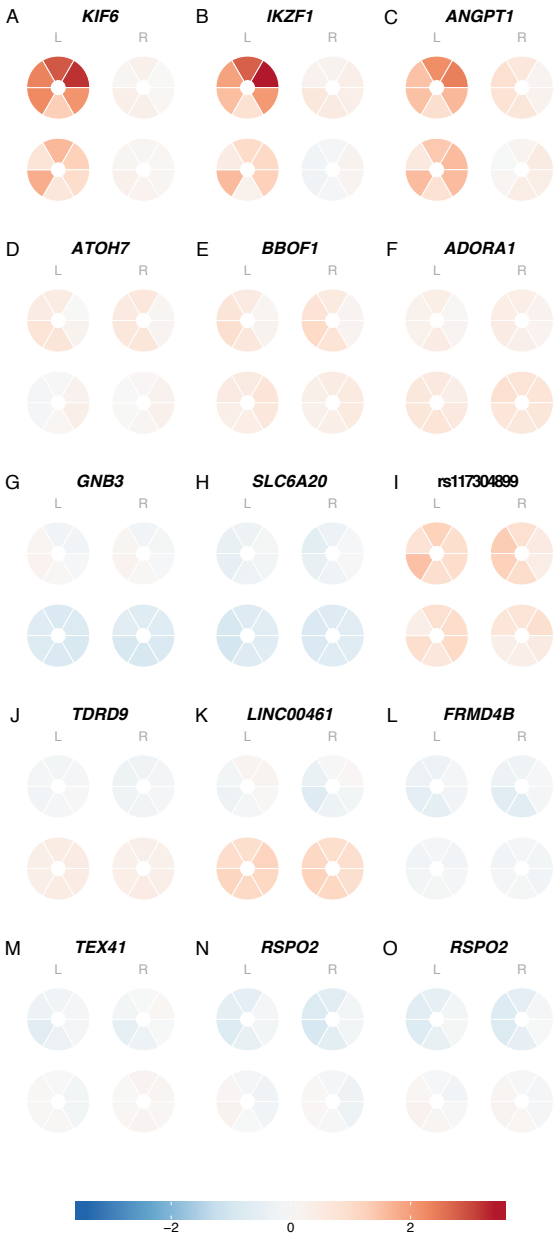

Figure 6.4: Spatial variation of genetic effect on the inner retina across the Macula 6 grid, illustrated as a maculagram The thickness of each inner retinal layer, both the retinal nerve fibre layer (RNFL) and the ganglion cell inner plexiform layer (GCIPL), in each segment of the Macula 6 grid across both left and right eyes was modelled as the outcome of each of the genetic variants found significantly associated with mean inner retinal thickness. The effect sizes from each of these models is plotted back into the Macula 6 grid space using maculagrams. 15 of the 46 SNPs are plotted here, selected based on interesting patterning. For each plot, the RNFL is pictured at the top, and the GCIPL is pictured below. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

the genetic effect of *RDH5* on the foveal, intermediate and peripheral areas separately. *RDH5* encodes a protein involved in the production of the human chromofore of visual pigments[Lidén & al., 2001]. Also *RDH5* is a gene known to be involved in fundus albipunctatus, a form of macular dystrophy [Yamamoto & al., 2003; Xie & al., 2020]. Fundus albipunctatus is characterised by white lesions across the retina, excluding the fovea [Skorczyk-Werner & al., 2015]. This differential spatial pathology aligns with the spatial genetic effect we see in these results.

VSX2 (rs28488340, figure 6.5D), *OSM* (rs8138196, figure 6.5E), *CCND1* (rs631695, figure 6.5F) and *FAAP100* (rs56737642, figure 6.5G) all show a similar pattern of genetic effect across the macula. Each of the four loci have a stronger genetic effect on the ONL compared to the IS and OS. Conversely, *TSPAN10* (rs9910935, figure 6.5M) shows a much stronger effect on the OS compared to the ONL and IS. Although part of the same cell, the ONL, IS and OS each have different roles within the PRC. These patterns of genetic effect are in keeping with layer specific roles within the PRC.

CCND1 also shows a pattern of effect on the ONL whereby it has a stronger effect on the intermediate and peripheral fields of the macula compared to the fovea. This genetic variant has previously been strongly associated with breast cancer. *CCND1* encodes cyclin D1, a regulator of cyclin-dependent kinase that is involved in the control of the cell cycle [Motokura & al., 1991]. This highlights the complex relationship between development and disease and the potential insights that the eyes, as direct windows to tissue of the CNS, could provide.

Additionaly *TSC22D2* (rs1376026, figure 6.5O) shows an opposing effect on the thickness of the ONL in the foveal field compared to the intermediate and peripheral fields. This locus has previously been associated with the presence of tumours using linkage analysis and has since been suggested to have roles in cellular metabolism and cell-cycle regulation [Li & al., 2016]. The variant at this locus has also been previously associated with basal metabolic rate, body mass and high blood pressure. Further investigation is needed to fully elucidate the role of this genetic variant in retinal morphology.

ZNF281 (rs6673000, figure 6.5K) appears to affect the thickness of the intermediate ring of the ONL more than the fovea and peripheral ring. *ZNF281* encodes a zinc finger transcriptional regulator which has been shown to be required for embryonic stem cell development [Pieraccioli & al., 2016]. The locus has also previously been associated with spherical power. This suggests spatial patterns of development that have effects on retinal function.

AFAP1L1 (rs1109114 and rs1438692, figure 6.5I and J respectively) has a notably stronger effect on the nasal quadrant (westerly quadrant) of the ONL than on other

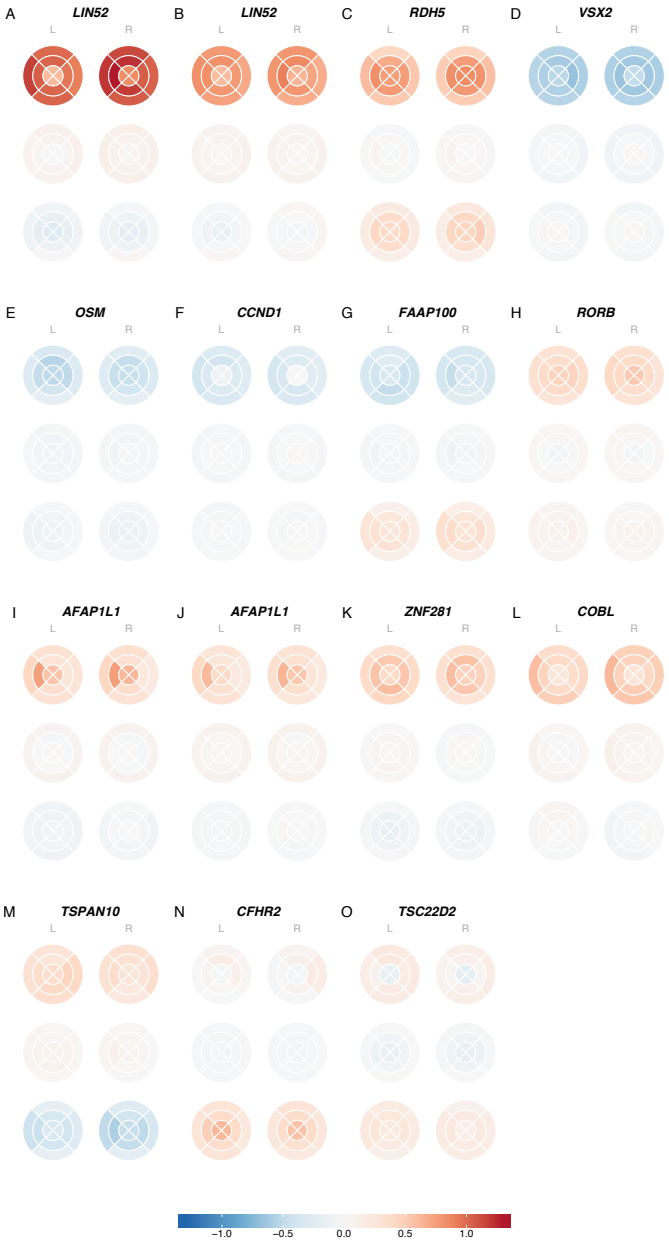

Figure 6.5: Spatial variation of genetic effect on the component photoreceptor cell layers across the ETDRS grid, illustrated as a maculagram The thickness of each component photoreceptor layer, the outer nuclear layer, the inner segment and the outer segment, in each segment of the ETDRS grid across both left and right eyes was modelled as the outcome of each of the genetic variants significantly associated with mean inner retinal thickness. The effect sizes from each of these models is plotted back into the ETDRS grid space using maculagrams. 15 of the 59 SNPs are plotted here, selected based on interesting patterning. For each plot, the layers are organised from top to bottom as follows: the outer nuclear layer, the inner segment, then the outer segment. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

areas of that layer, and this effect is antagonistic to the effect exerted on the other two layers. *COBL* (rs12719025, figure 6.5L) also shows a milder form of this nasally localised genetic effect. As previously mentioned, the optic nerve exits the retina via the nasal side of the retina [Freddo, 2018]. *COBL* has prior associations with spherical power and total macular thickness. *AFAP1L1* has previous associations with BMI, hearing difficulties, hypertension and lung-related traits. The association with hearing difficulties suggests a role of this locus in fine neuronal structure and development that provides an interesting avenue for future research.

Spatial Variation of Genetic Effect on the Retinal Pigment Epithelium

The majority of SNPs associated with RPE layer thickness have a uniform effect across the macular field and the two eyes (figure 6.6). There are two SNPs, *METTL4* (rs7239443, figure 6.6B) and *KIF16B* (rs77561053, figure 6.6C), that show a distinction between the left and right eyes, but once again this is likely an example of confounding due to lack of randomisation in which eyes was scanned first. The simplicity of the topographical genetic effect may be due to the fact that the RPE layer is comprised of a single cell type, is a single cell layer thick and itself is relatively uniform across the macula.

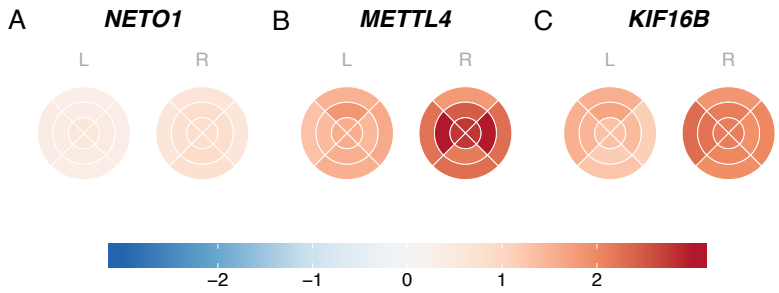

Figure 6.6: Spatial variation of genetic effect on the retinal pigment epithelium across the ETDRS grid, illustrates as a maculagram The thickness of the retinal pigment epithelium in each segment of the ETDRS grid across both left and right eyes was modelled as the outcome of the genetic variants significantly associated with mean retinal pigment epithelium layer thickness. The effect sizes from each of these models is plotted back into the ETDRS grid space using maculagrams. Three of the seven SNPs are plotted here, selected based on interesting patterning. The grids are oriented so that for both left and right eyes the easterly segments represent the temporal segments of the eye, those closest to the ear, and the westerly segments represent the nasal sides, those closest to the nose.

6.2. Extracting Higher Dimensional Phenotypes

Utilising more of the data extracted using the segmentation grids enabled novel discoveries about genetic effect on spatial morphology variation. However, even by analysing the subsections of the TABS segmented grids, much of the rich data available in the OCT images is ignored. Therefore, I developed image analysis methods to extract phenotypes representing higher dimensional variation in retinal morphology from OCT images. I aim to extract several measures of retinal morphology, mainly quantified as retinal thickness, from raw OCT images fully utilising the pixel level information available. Studying a finer grain phenotype enables analysis of the genetic effect on a more continuous and detailed representation of retinal morphology.

6.2.1. Image Analysis Methods

To extract higher dimensional phenotypes, I segmented the two limiting membranes of the retina, and from this calculated the overall retinal thickness. I also extracted an additional phenotype measuring the total retinal thickness at the central fovea.

Input Images

Raw OCT images were available for each of the ophthalmologically imaged participants of the UK Biobank. Whereas the TABS derived phenotypes collapsed the available data into a number of discrete values, information is available in the raw images at a pixel-wise level. Each three-dimensional (3D) image is comprised of 128 B-scans, along the z-dimension (figure 6.7). Each B-scan is comprised of 512 A-scans, along the x-dimension. Each A-scan has 650 values in the y dimension. In total, there are >40 million pixel-wise values per OCT image.

Segmentation of Retinal Limiting Membranes

To generate an initial phenotype, I segmented the limiting retinal membranes from the raw OCT images. First a running median, with window size of three, is applied to the intensity values for each A-scan per B-scan. The median intensity of the A-scan is subtracted from this vector, and the absolute value calculated, resulting in a normalised intensity vector per A-scan. A threshold is set as the median of this normalised intensity vector plus two times the standard deviation of the vector. The maximum and minimum indices of the values that pass the threshold are considered the preliminary limiting membrane positions (figure 6.8 A).

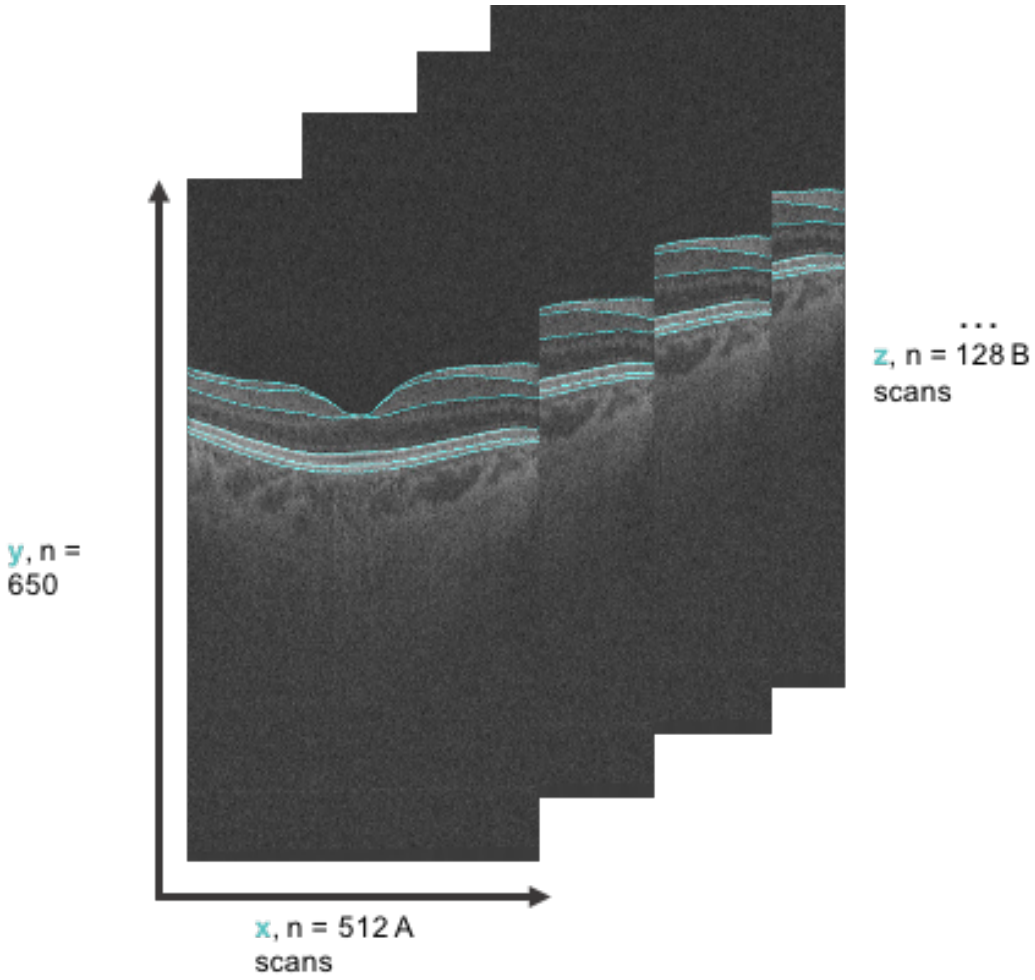

Figure 6.7: Dimensions of raw OCT image data A single slice of an OCT image, or B-scan, has 512 pixel values in the x dimension, and 650 pixel values in the y dimension. y $\times$ z dimension are known a the A scans. There are 128 B-scans per OCT image.

Having identified the preliminary position of the limiting retinal membranes for every A-scan in a B-scan, the vector of the segmented surface for the outer and inner limiting membranes are considered separately. For each surface vector across a B-scan, a running median is calculated with a median window of 51. The derivative log spread of the membrane vector was calculated as a measure of the noise between consecutive data measurements. Where the preliminary value of the limiting membrane was outside of a confidence interval - defined as the running median value $\pm$three times the derivative log spread - the original value was replaced with the interpolated value (figure 6.8 B). This two step process was performed to minimise noise and attempt to model a more consistent gradual change in thickness across the macula. The result of this process is two vectors, representing the inner and outer limiting membrane of the macular retina for a B-scan. This process was repeated for every B-scan. This creates a 3D segmented representation of the retinal limiting membranes (figure 6.8 C). For each limiting membrane, we have 65,536 values describing their locations.

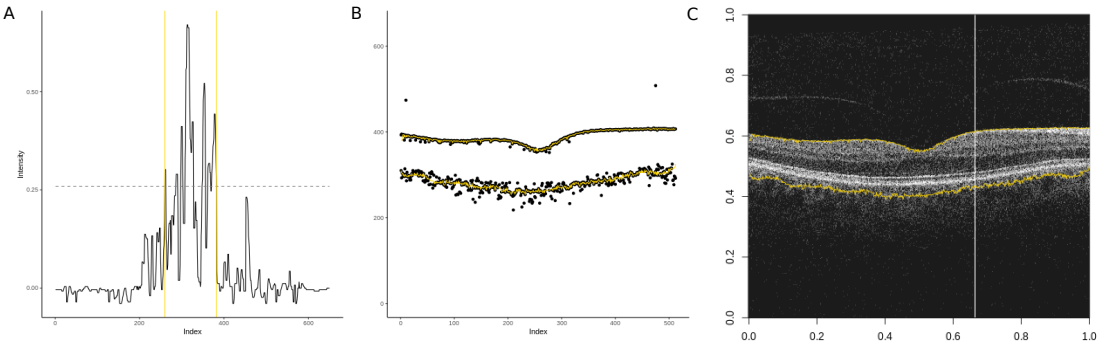

Figure 6.8: Segmentation of retinal boundaries The main stages of identifying the limiting membranes of the retina from raw OCT images. A) A running median of intensity is calculated along each A-scan for every B-scan, and the median intensity across the overall A-scan subtracted, the absolute values of which are plotted as the black line. A threshold is set as the median of this vector plus two times the standard deviation, plotted as the dotted line. The maximum and minimum values that pass this threshold are considered the limiting membranes of the retina, the positions of which are plotted as yellow lines. B) Having selected the limiting membrane for each A-scan across a B-scan, the data outside of band specific noise confidence intervals are interpolated to smooth the line. C) This smoothed line is used as the segmented limiting membranes of the macular retina.

New High Dimensional Measures

As an initial high dimensional phenotype, the total retinal thickness across the macula was calculated for every pixel dimension. This was done by calculating the difference between the two segmented retinal limiting membranes. This resulted in 65,536 measures of total retinal thickness per 3D image.

To look at the genetic effect on this high-dimensional phenotype, participants were stratified by genotype into three groups at each locus of interest: homozygous reference, heterozygous and homozygous alternative (labeled as 0, 1 and 2 respectively). The mean total retinal thickness at each pixel-wise location of the macular area was calculated across each genotype group. I next compared the phenotype between the genotype groups using the homozygous reference as the baseline. It should be noted that the assignment of reference and alternative allele is arbitrary and the alternative allele does not necessarily refer to the derived allele. For each SNP, two phenotype matrices were calculated. The difference between the mean retinal thickness matrix of the heterozygous genotype group was subtracted from that of the homozygous reference, labelled as "0 to 1". The homozygous alternative mean retinal thickness matrix was subtracted from the homozygous reference, labelled as "0 to 2". The resulting matrices are plotted as heatmaps.

Results

On initial examination of the resulting heatmaps, several spatial patterns of morphological variation were notably different between the genotype groups. I observed that for several of the SNPs, there is noise within the heatmap visualisation. This is seen as vertical stripes in the heatmaps. This is likely an artefact of scan acquisition and image processing, which in both cases is done in a slice-wise manner through the retina. Another observation was that, for multiple SNPs, changes in the morphology between genotype groups was localised to the fovea. This suggested that these genetic variants were affecting the overall thickness of the retina at the fovea differentially to the rest of the retina. Several SNPs also appeared to have an effect on overall retinal thickness localised to a ring enclosing the fovea. These results are explored in more detail below.

Central Foveal Thickness Phenotype

Following observation of a foveal localised genetic effect upon visualisation using the heatmaps, I looked to explore the differential genetic effect on total retinal thickness at the fovea as compared to the rest of the macula. To do so a measure of retinal thickness

at the centre of the fovea was calculated. The vectors representing the thickness at the central cross-section of the macula in both the x and z dimensions in the left eye were identified. A spline was applied to each of these vectors. Two values of total retinal thickness at the central fovea were defined as the value at the midpoint of both these vectors. These values were modelled as the outcome of each of the SNPs found to be associated with the three retinal areas - the inner retina, the PRC and the RPE. The first 20 genotype PCs were included as covariates in the models. Loci were considered significantly associated with the central foveal thickness phenotype if P>0.05 for models of both the x and z dimension value following Bonferroni correction for multiple testing.

6.2.2. Higher Dimensional Analysis of the Inner Retina

Each of the SNPs found significantly associated with the meta-analysed inner retinal layers were used as input into the above stratification analysis. Three loci were found to be significantly associated with the central foveal thickness phenotype. These were *OCA2* (rs1800407) *TYR* (rs1042602), and *TSPAN10* (rs7503894). Comparing their effect on retinal thickness against a control genetic variant - the SNP with the lowest p-value in the inner retinal thickness GWAS which was not significantly associated with central foveal thickness - using visualisation by heatmaps highlights the fovea localised effect (figure 6.9).

To better visualise the spatial effect of these SNPs on overall retinal thickness, the cross-section through the central macula was plotted stratified by genotype (figure 6.10). As compared to a control (*IGFBP3*, figure 6.10 A), in which there is no noticeable separation between the genotype groups in terms of overall retinal thickness, the three loci significantly associated with the retinal thickness at the fovea show a noticeably different pattern. The majority of differentiation in total retinal thickness between genotype states appears to occur at the fovea, with smaller or no differences at the macula shoulders and peripheral areas.

This foveal localised pattern of effect noted using the central foveal thickness bears similarity to the rare clinical condition of foveal hypoplasia. Foveal hypoplasia clinically refers to a shallower foveal dip, and a consequently thicker retinal thickness at the fovea. Foveal hypoplasia is the result of underdevelopment of the foveal dip, due to lack of excavation of the inner retinal layers. Foveal hypoplasia is commonly seen in several conditions including OCA [McAllister & al., 2010]. Both *TYR* and *OCA2* are well-studied OCA genes. The SNPs associated with retinal thickness are not those found previously mutated in oculocutaneous albinism but are common variants. As previously mentioned, the SNP at the *TYR* locus, has been found in a haplotype that causes

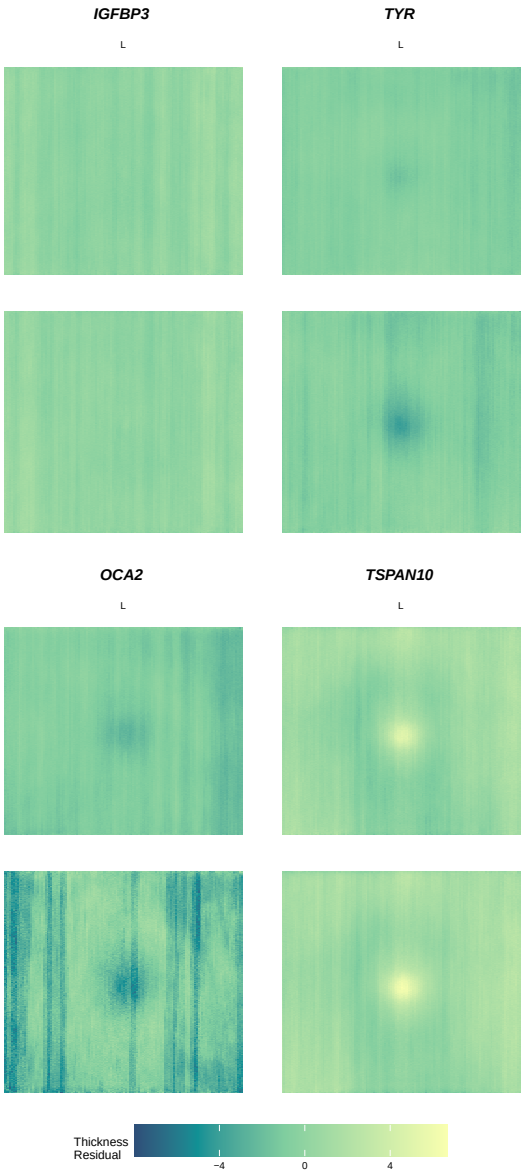

Figure 6.9: Difference in overall retinal thickness of groups stratified by genotype at loci associated with inner retinal thickness The overall retinal thickness is calculated across the macular field for each participant in the left eye. The participants are stratified by genotype, homozygous reference, heterozygous or homozygous alternative (encoded 0, 1, and 2 respectively) at each of the loci found to be significantly associated with inner retinal thickness. The mean retinal thickness of each of these groups is calculated. The difference in thickness between the homozygous reference and heterozygous (0 to 1), and homozygous reference and homozygous alternative (0 to 2) are calculated and plotted as heatmaps. *IGFBP3* (rs11762530), the control, was not significantly associated with total retinal thickness at the foveola. *TYR* (rs1042602), *OCA2* (rs1800407) and *TSPAN10* (rs7503894) were all found to be significantly associated with total retinal thickness at the central fovea.

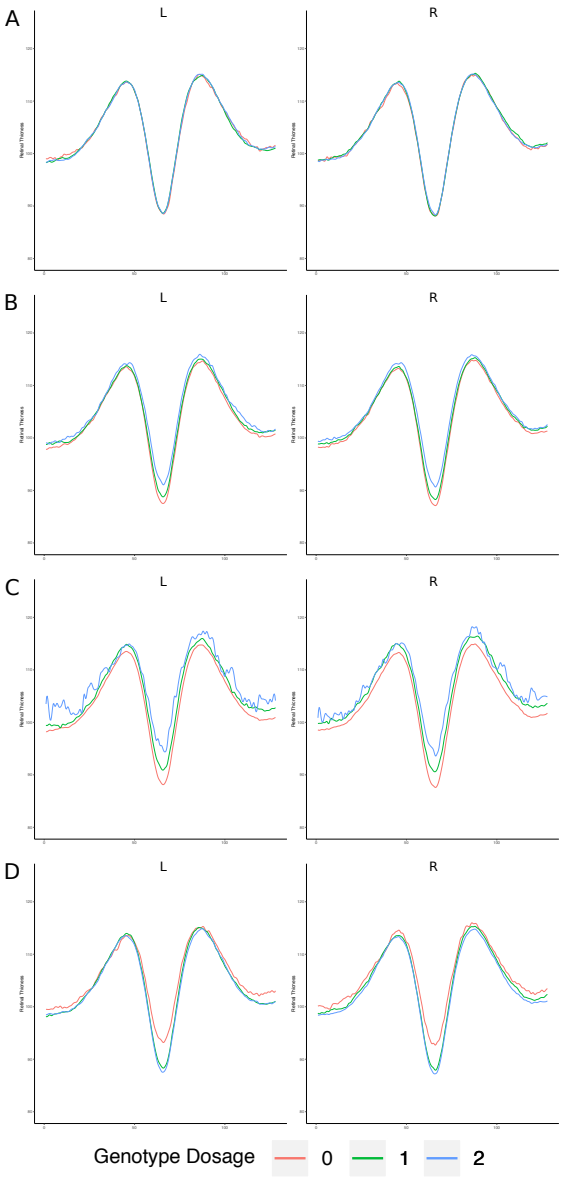

Figure 6.10: Cross-section of mean total retinal thickness stratified by genotype state. The mean total retinal thickness at the central cross-section of the retina in the z-dimension across groups stratified by the genotype state. The three genotype states are homozygous reference (0), heterozygous (1) and homozygous alternative (2). Stratified by genotype state at: A) *IGFBP3* (rs11762530), the variant with the lowest p-value aside from *TSPAN10* in the inner retinal GWAS, used as a control. B) *TYR* (rs1042602), C) *OCA2* (rs1800407) D) *TSPAN10* (rs7503894). The y-axis represents the overall retinal thickness.

ocular albinism [Campbell & al., 2019]. Ocular albinism displays the same ocular phenotypes as OCA, including foveal hypoplasia, but not the pigmentation phenotypes. In spite of this, both *TYR* and *OCA2*, in addition to *TSPAN10* share prior associations to pigmentation.

To test this association between pigmentation and the central foveal thickness trait, the hair colour by percentage of the population stratified by genotype at these three loci was plotted (figure 6.11). To confirm the link of these loci to pigmentation, a model of SNP effect on hair colour was built. Self-reported hair colour within the UK Biobank, excluding red hair, was numerically encoded from light (blonde) to dark (black) and the effect of the three SNPs determined by linearly modelling their effect with the first 20 genotype PCs as covariates. All three loci were significantly associated with hair colour. Interestingly, the direction of effect was not consistent across the three loci. For both *TYR* and *TSPAN10*, the allele associated with increased central foveal thickness was associated with lighter hair colour. The genetic variant at *TSPAN10* in particular appears to follow a classic dominant recessive model. Conversely, in the case of *OCA2*, the allele associated with increased central foveal thickness was associated with darker hair. This result hints at the complex relationship between foveal morphology and pigmentation.

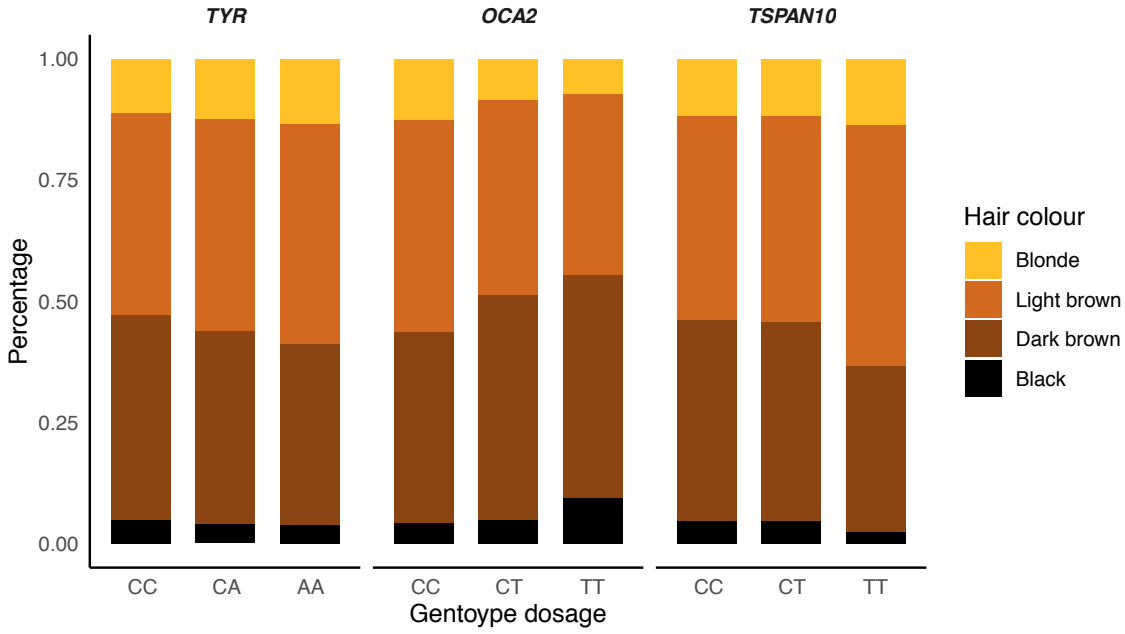

Figure 6.11: Hair colour proportions within the population stratified by allele state. Percentage of self-reported hair colour within each group, defined by their stratification based on genotype state at loci significantly associated with central foveal thickness. From left to right: *TYR* (rs1042602), *OCA2* (rs1800407), *TSPAN10* (rs7503894).

The clinical condition of foveal hypoplasia which shows acute thickening of the central fovea is often associated with poor visual acuity. Therefore I proceeded to explore the functional impact of this milder central foveal thickness phenotype seen above. Therefore the 46 SNPs significantly associated with inner retinal thickness were checked for association with visual acuity. To do this, the effect of each of the 46 SNPs on the logMAR score was linearly modelled with the first 20 genotype PCs included as covariates. Three of the loci were found to be significantly associated with visual acuity following Bonferroni correction for multiple testing. These three loci were the same loci that were significantly associated with the central foveal thickness phenotype. It is of great interest that only the three loci that significantly display the foveal localised effect on thickness appear to also affect a change in function.These results show that common variants in disease loci show milder versions of the disease phenotype, both structurally and functionally.

It should be noted that other patterns of high dimensional spatial variation may also be affected by genetic variants. For example, in addition to the effect of the *TSPAN10* locus localised to the fovea, there appears to be an effect on the intermediate concentric field, presenting as a dark ring around the fovea (figure 6.9). Future research would be needed to find a quantitative measure of this effect and thus test for significant examples of this pattern.

6.2.3. Higher Dimensional Analysis of the Photoreceptor Cell Layer

Visualisation of the spatial genetic effect again led to the observation that several SNPs appear to have a differential effect on total retinal thickness at the fovea compared to the rest of the macula. Therefore, the genetic variants found significantly associated with the thickness of the component PRC layers were each used to model the total retinal thickness at the central fovea. Of the 59 SNPs tested, four were found to be associated with this measure. The effect of these SNPs on the total retinal thickness across the macula is visualised in heatmaps (figure 6.12). A distinct difference in pattern can be seen between SNPs with no association (*LIN52*, rs1972564, the locus with the smallest p-value in the PRL GWAS), and those significantly associated with total retinal thickness at the foveola, *TYR* (rs28620862), *FAAP100* (rs56737642) and *TSPAN10* (rs62075724 and rs9910935).

TYR and *FAAP100*, are both known pigmentation loci, with previous associations to hair and skin colour. *FAAP100* is also in LD with the two SNPs associated with central macula thickness at *TSPAN10*. The three SNPs at this locus are likely an example of allelic heterogeneity and highlight the complex genetic effects caused by this locus.

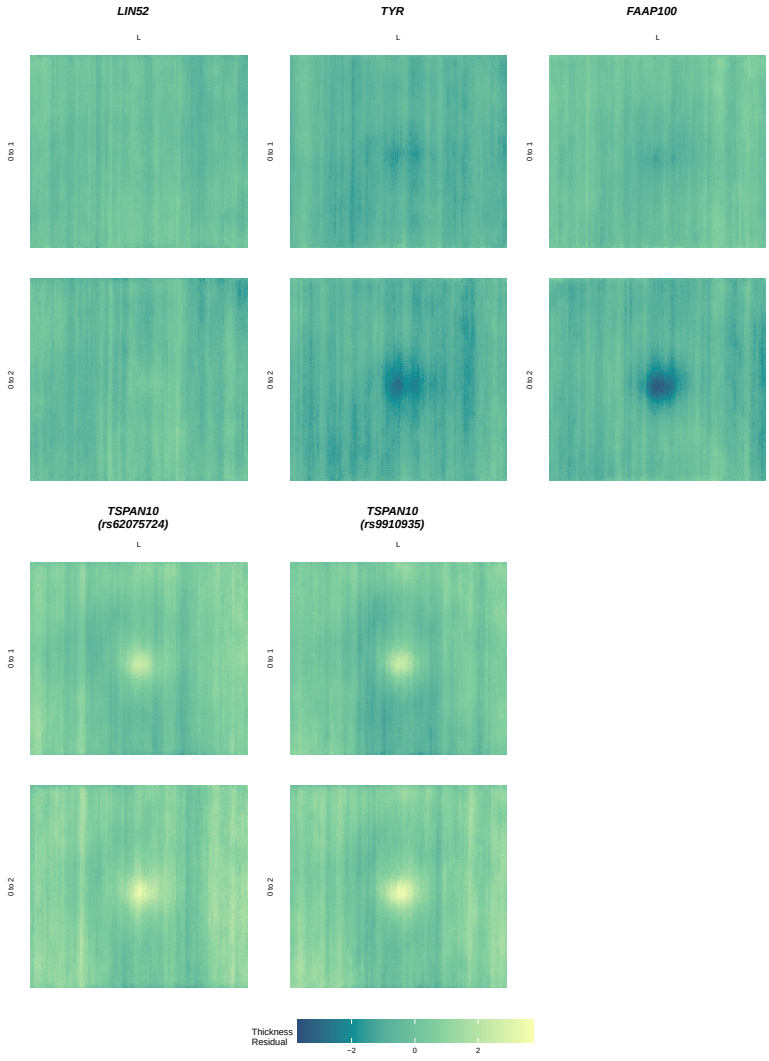

Figure 6.12: Difference in overall retinal thickness of groups stratified by genotype at loci associated with photoreceptor cell layer thickness. The overall retinal thickness is calculated across the macular field for each participant in the left eye. The participants are stratified by genotype, homozygous reference, heterozygous and homozygous alternative (encoded 0, 1, and 2 respectively) at each of the 59 loci found to be significantly associated with photoreceptor layer thickness. The mean retinal thickness of each of these groups is calculated. The difference in mean thickness between the homozygous reference and heterozygous (0 to 1), and homozygous reference and homozygous alternative (0 to 2) are calculated and plotted. *LIN52* (rs1972584), the control, was not significantly associated with total retinal thickness at the foveola. *TYR* (rs28620862), *FAAP100* (rs56737642) and *TSPAN10* (rs62075724 and rs9910935) were all found to be significantly associated with total retinal thickness at the central fovea.

Similarly to the analysis of the SNPs associated with inner retinal thickness, these results highlight the complex relationship between retinal structure in particular central foveal thickness, and pigmentation.

Once again, an additional pattern of effect is localised to a ring surrounding the fovea for the *TSPAN10* variants. This effect requires further investigation in the future.

6.2.4. Higher Dimensional Analysis of the Retinal Pigment Epithelium

Visualisation of the effect of genetic variants associated with RPE thickness on total retinal thickness across macula highlighted a differential effect at the fovea as compared to the rest of the macula. Three of the loci associated with RPE layer thickness were found to be significantly associated with total retinal thickness at the central fovea. The fovea specific changes are starkly noticeable when comparing to those not significantly associated (*KIF16B*, rs77561053) (figure 6.13).These include the *TYR* (rs1126809) and *OCA2* (rs1800407), well-characterised oculocutaneous albinism genes recurrently discovered across the retinal areas. Additionally, a variant at the *CFHR1* (rs61822227) locus was also found to be significantly associated with differential total retinal thickness at the central fovea. *CFHR1* is known to play a role in AMD, a condition that causes progressive loss of vision. As AMD mainly affects central areas of the macula, this spatial effect aligns with genetic effect on the thickness of the retinal epithelium seen here. This again highlights an interesting relationship between common variation causing mild spatial effects and disease phenotypes.

6.3. Discussion

In this chapter, I explored the biological variation in retinal morphology at a higher dimensionality, and looked at the genetics underlying these more granular phenotypes. To do this, I developed a technique, termed maculagrams, for visualising and assessing the morphological variation across the macula field and the genetics underlying it. I then developed image analysis techniques that can extract high dimensional phenotypes from the raw OCT images, taking further advantage of the higher dimensionality available within the data. The combined application of both of these analyses enabled novel findings about retinal morphology and genetics.

The maculagrams allowed for identification of several interesting patterns of morphological variation in retinal layer morphology across the macular field. Common to all of the layers, a considerable source of variation in layer thickness appears to be differences between left and right eyes My main hypothesis is that much of the between-eye

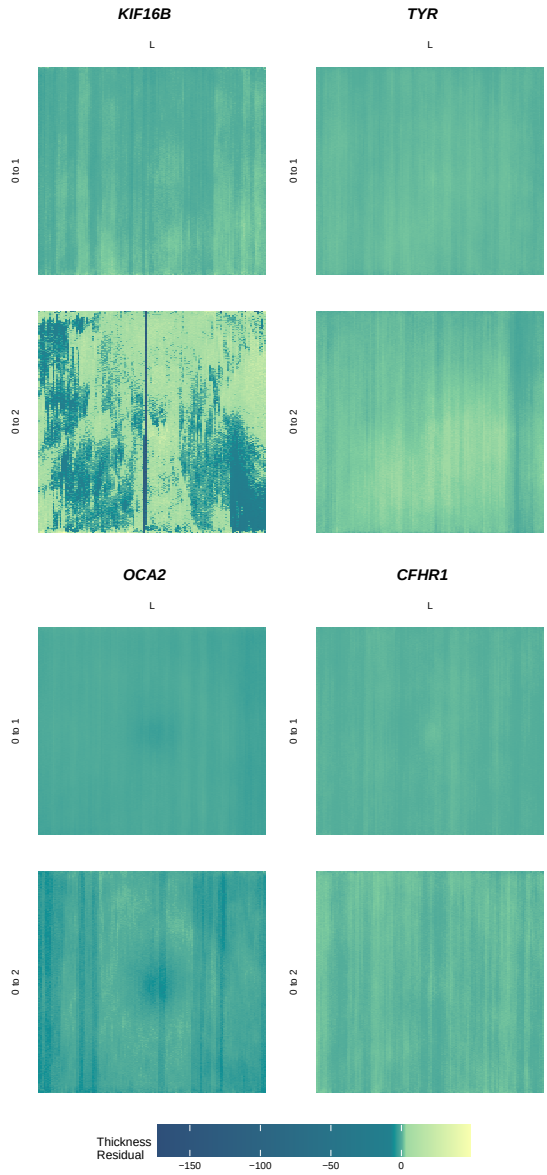

Figure 6.13: Difference in overall retinal thickness of groups stratified by genotype at loci associated with retinal pigment epithelium layer thickness. The overall retinal thickness is calculated across the macular field for each participant in the left eye. The participants are stratified by genotype, homozygous reference, heterozygous and homozygous alternative (encoded 0, 1, and 2 respectively) at each of the 7 loci found to be significantly associated with retinal pigment epithelium layer thickness. The mean retinal thickness of each of these groups is calculated. The difference in mean thickness between the homozygous reference and heterozygous (0 to 1), and homozygous reference and homozygous alternative (0 to 2) are calculated and plotted. *KIF16B* (rs77561053), the control, was not significantly associated with total retinal thickness at the foveola. *TYR* (rs1126809), *OCA2* (rs1800407), and *CFHR1* (rs61822227) were all found to be significantly associated with total retinal thickness at the central fovea.

variation results from the lack of randomisation of order of eye scans. The protocol was designed so that the right eye was consistently scanned first, and the left eye second. It is likely that this has produced a batch-like effect, caused by differential light exposure, eye dryness, blinking behaviour and other eye-specific scanning responses. There may be biological variation underlying these differences, as suggested blinking response or adaptation to reduced light levels. However in this instance, this is not the variation we wish to investigate, so the best option was to control for it. In an ideal situation, the order in which the left and right eyes were scanned would have been randomised, which would allow for us to identify if the left-right eye variation had a biological basis or was a result of measurement order.

Some of the PCs representing variation in retinal thickness across the macula field appeared to cluster into pairs representative of the same spatial patterning, but split by eye. This is due to the nature of PCA in which all resulting PCs must be orthogonal to one another [Pearson, 1901]. Considering that for each retinal area, one of the first PCs represented between eye variation, this forces all other PCs to respect this asymmetry. Therefore it can be assumed that several of the PCs occur in pairs that represent the same pattern of variation, but across the two eyes separately. This makes interpretation of the visualisation technique less intuitive. Therefore in the future it would be interesting to explore other methods of dimensionality transformation that do not assume orthogonality of the resulting dimensions.

The use of maculagrams also allowed for interesting observations on the spatial genetic effects on retinal layer thickness. Some patterns of note include a differential effect of SNPs on the thickness of both the inner retinal layers and the component PRC layers at the fovea compared to the peripheral macula. The structure of the macula dip varies radially. Additionally the distribution of the two forms of PRC, the rods and cones, varies in a concentric pattern from the fovea outwards. Some diseases affect only one type of PRC and loci associated with such diseases - including fundus albipunctatus and AMD - were found to exhibit this distinctive genetic effect across the macula field. The alignment of the pattern of genetic effect of these SNPs on retinal layer thickness with the shape of the macular dip and PRC distribution suggest a relationship between retinal disease, common variants at disease loci, and retinal structure.

Moving to an even higher dimensional representation of the macular space that utilised the pixel-wise information available in OCT images revealed new insights into the complexities of retinal morphology. This is highlighted by the results showing the differential effect of multiple SNPs at the fovea as compared to other areas of the macula. Several of the loci associated with total retinal thickness at the central macula were also

associated with visual acuity. Some of these loci are known to be involved in optical pathologies including OCA or its milder form, ocular albinism [Campbell & al., 2019]. In these conditions, foveal hypoplasia, the underdevelopment of the foveal dip and subsequent thicker central macula, occurs. The foveal localised thickness represented by the central foveal thickness phenotype may exist on a spectrum that also includes the rare and more acute condition of foveal hypoplasia. This analysis appears to have identified common variants at loci associated with rare disease that present with phenotypes related to the disease phenotype. This highlights the continuum between rare and common variation. It is probable that the milder phenotypes are often not recognised due to the lower impact, but that they are still affecting retinal function.

Further, the discovery of central foveal thickness being affected by genetic variants that also affect retinal layer thickness - with links to pigmentation and visual acuity - suggests that the thickness of the central fovea may be confounding measures of retinal thickness. This is of importance in several instances where retinal layer thickness is used as a biomarker of disease. In the future, the genetic variants could be controlled for to gain more accurate retinal layer thickness measures. Alternatively, the way in which the scans are conducted could be altered or the macula field better modelled to exclude more of the central foveal field, minimising the effect of the central foveal thickness on the mean retinal thickness measure.

7. Concluding Remarks

In this thesis, I have conducted the first large-scale GWAS of individual retinal layer thicknesses. This has elucidated a number of genetic variants significantly associated with retinal morphology. A portion of these genetic variants are novel, with many others having prior associations to ocular and general biology. By employing MR on the outcome of my GWAS, I have been able to interrogate commonly observed correlations between clinical traits. I have also demonstrated the discovery power enabled by utilising high-dimensional phenotypes derived from medical imaging.

7.1. Summary of the main findings

The key results of this thesis are:

- I explored the range of available phenotypes extracted from OCT images to describe retinal morphology. I showed that a simple measure of mean thickness across the macula and across both eyes provides comparable genetic discovery power to measures derived from higher dimensional morphological measures. This is likely owing to the large size of the UK Biobank dataset. The use of the simple measure is also beneficial in terms of ease of interpretation as it is a measure more closely aligned with that used in routine clinical examinations.

- In chapter 4 I conducted the largest GWAS of inner retinal thickness as extracted from OCT images. As a result I found 46 genetic variants associated with the thickness of one or both of the inner retinal layers, the RNFL and the GCIPL. Many of these loci had prior associations to ocular phenotypes. There were also loci discovered with prior associations to anthropometric and blood related-traits, asthma and neurodevelopmental conditions. Additionally a number of the SNPs were novel, with no prior associations to ocular or general biological phenotypes.

- Later in chapter 4, using MR, I found evidence for a causative relationship between POAG and IOP, consistent with the treatment of POAG through the lowering of IOP via medication or surgery. In contrast, MR did not provide evidence for a

causal relationship between inner retinal thickness and POAG, suggesting they do not share the same underlying genetic pathway. This is in contrast to the use of inner retinal thickness as a diagnostic biomarker of POAG. These results indicate that a genetically determined thinner inner retina does not predispose one to glaucoma, rather it is the thinning of the inner retina over time that is symptomatic of glaucoma.

- In chapter 5 I conducted the largest GWAS of outer retinal layer thicknesses to date, including the thickness of the ONL, IS, OS and RPE layer. In total this led to the discovery of more than 100 loci associated with one or more of these layers. The ONL, IS and OS were jointly meta-analysed as they constitute subsections of the PRC. This resulted in 59 loci significantly associated with one or more of the layers of the PRC. Seven loci were discovered to be significantly associated with RPE thickness. In both cases, discovered loci included those previously associated with ocular phenotypes and general biology. Each GWAS also found novel genetic variants with no prior associations. Using MR, support was shown for a causal relationship between AMD and RPE thickness but not between AMD and the other outer retinal layers.

- In the latter half of chapter 5, I explored expanding the dimensionality of the data we use to describe outer retinal morphology from one measure to three. Moving from the mean thickness across a retinal layer, to the thickness in the peripheral, intermediate and foveal fields of the retina, provided novel discoveries about genetic variation driving both intra and inter-layer morphological variation. Numerous loci were discovered to be differentially affecting the retinal fields of the PRC component layers and the RPE layer. These included loci associated with retinal disease such as cone-rod dystrophy, already known to have a disease process with spatial variation.

- In chapter 6 I developed a visualisation technique, maculagrams, that allow for visualisation and identification of variation patterns across the macula field as defined by clinical segmentation grids. These maculagrams enabled me to explore variation at two levels. First, the variation in retinal morphology across the macula field. Second, the differential effect of genetic variants on retinal morphology across the macula field. These maculagrams enabled visualisation of the complex patterns in which morphology of the retinal layers was varying. There were also several notable patterns of spatial genetic effect on the thickness of the retinal layers across the macular field. This includes a differential effect of several loci on

the thickness of retinal layers at the fovea compared to the peripheral macula. The macula dip varies radially, as does the distribution of rods and cones across the macula. Some of the loci found to affect retinal thickness in this pattern had prior associations with ocular diseases that affected a particular type of photoreceptor cell.

- Later in chapter 6, I further developed image analysis techniques that extracted phenotypes from OCT images, taking advantage of the pixel-wise data to generate more granular representations of retinal variation. These phenotypes describe overall retinal thickness across the macula. I explored how this measure was influenced by the genetic variants I found associated with individual retinal layers. I discovered a novel population level trait, a fovea-localised change in total retinal thickness that bears resemblance to foveal hypoplasia, a clinical trait previously associated with Mendelian disease. Several loci that affected individual layer retinal thicknesses were associated with the central foveal thickness trait, and the same loci were associated with pigmentation and visual acuity. However interestingly the direction of the effect of hair colour relative to direction of effect of the central foveal thickness was not consistent across loci. Meanwhile the direction of effect of visual acuity and the effect of the central foveal thickness remained consistent across SNPs. Some of these loci are known to be involved in ocular pathologies including, OCA which causes severe foveal hypoplasia, and ocular albinism, a partial form of OCA.

7.2. Conclusions

7.2.1. Implications in understanding of ocular disease

Despite the use of inner retinal thickness as a biomarker for glaucoma, MR analysis did not find evidence to support a causal relationship between the measure and the disease. This does not diminish the value of inner retinal thickness as a biomarker of glaucoma, but rather informs us of the mechanism underlying disease progression. These results suggest that whilst thinning of the inner retina is indicative of glaucoma, a genetically determined thinner retina does not increase susceptibility to the disease. It is a change in thickness over time that is of diagnostic importance. This finding may help efficiently direct therapeutic research and diagnosis (discussed further in section 7.4).

A similar relationship is seen between AMD and the thickness of the outer retinal layers excluding the RPE. Despite the thinning of the PRL during the progression of AMD,

MR found no evidence for them being on the same genetic causal pathway. This suggests that whilst PRC thinning is seen as a result of AMD, genetically determined PRC thickness does not alter the risk of developing the disease. There was however evidence of a causal relationship between RPE thickness and AMD, when AMD is used as the exposure variable and RPE as the outcome. This is in-keeping with the development of drusen, lipid deposits characteristic of AMD, occurring under the RPE.

The analysis presented in this thesis has identified common variants at loci associated with rare disease that present with phenotypes that appear to be on the same spectrum as the disease phenotypes. Specifically, the loci with prior associations to OCA, a disease marked by foveal hypoplasia, were found associated with a mild localised change in thickness at the centre of the fovea in our population. This discovery suggests there is a continuum between rare and common variation. This is coupled with the effect of the same loci on visual function as measured by visual acuity. It is likely that milder phenotypes often go unrecognised due to their lower impact, but that they maintain an effect on retinal function.

Beyond this, variants were found to be associated with both mean thickness of individual retinal layers and the central foveal thickness trait at the loci that cause clinical foveal hypoplasia. This suggests mild central foveal thickening may be affecting the measurement of individual layer thicknesses. This is particularly pertinent in the inner retina, where this mild central foveal thickening may be confounding inner retinal thickness measures that are used as a biomarker for glaucoma. There are several ways in which this could be accounted for; either through controlling for the genetic variants associated with the central foveal thickness trait when using inner retinal thickness to diagnose glaucoma. Alternatively, by excluding a larger field surrounding the fovea, or modelling the macula differently, it may be possible to minimise the effect of this phenotype on inner retinal thickness measures. Both of these theories require further exploration but offer interesting avenues of research to improve the diagnostic efficiency of glaucoma.

7.2.2. Strengths of high-resolution biomedical imaging

The analysis within this thesis discovered loci associated with the thickness of each of the retinal layers. Of these, numerous loci had prior associations to visual function. This highlights the role of morphology in function, the specificity of ocular structure, and the potential for using non-invasive methodologies that can monitor morphology to diagnose and manage disease. In particular the association of several of the SNPs with both inner retinal thickness and neurodevelopmental traits reiterates the eyes' role

within the CNS and its potential role as a source of biomarkers for complex neurological conditions.

Equally the exploration of genetic effect on the morphology of the retinal layers across the macula field has provided evidence that there is both inter- and intra-layer variation that is affected by genetic variation. The ability of OCT to visualise the 3D space provides a rich source of data that allows for modelling of the continuous topographical variation. Further investigation may help us better understand development of the eye, and the intricacies of disease that have spatial effects (discussed further in section 7.4).

Further, analysis of the three layers constituting the PRC has offered a unique opportunity. Although OCT imaging does not offer single cell resolution, the measures extracted from the images are describing the sub-structures of cells. This allowed me to study the micro-structures of the retina. The genetic variation we see underlying the specific sub-structures of this cell type speaks to the incredible specificity of retinal structure and function. This analysis highlights the discoveries enabled by high-resolution biomedical imaging in progressing our understanding of developmental processes, cell biology and function.

7.3. Limitations of the analysis and potential improvements

The scope of this research allowed us to look at the effect of genetic variation on the retinal morphology in a population with large statistical power. This was largely enabled by the richness and size of the UK Biobank dataset. However, the dataset and thus the analysis was not without limitations. One limitation to studying the morphology of the human retina using this dataset is the ability to comprehensively interrogate any asymmetry between left and right eyes. When exploring potential phenotypes that could be used to represent inner retinal morphology, the thickness of the left and right eyes were used as separate input into GWAS (section 4.3). Although the results were largely congruent between the two eyes, there was some discrepancy. This discrepancy was also seen when looking at the high dimensional data describing retinal morphology in chapter 6. However the data acquisition protocol consistently scanned the right eye first, introducing a scan acquisition confounder. Therefore any variation grounded in biology is inseparable from that due to scan acquisition confounding. It would be interesting to be able to interrogate any asymmetry further and investigate if there are some biological groundings to the differences that can be observed in left and right eye morphology. For this to be possible, a dataset would be necessary in which scanning of the left and right eye has been randomised so that the confounding due to scan acquisition

can be minimised.

A strength of the UK Biobank cohort in the analysis presented was that the population was not selected on the basis of having a disease, meaning that we were able to look at the natural variation in morphology of the human retina. However several of the hypotheses that this analysis generated require cohorts of individuals with rare disease, preferably of similar scale and phenotype depth. Unfortunately such cohorts are rare which has limited follow up on some findings including the potential role of these morphological measures as biomarkers (discussed further in section 7.4) . As the UK Biobank population ages, some may present with these currently rare diseases making the proposed hypotheses testable.

7.4. Outlook and future research

The research presented in this thesis has been enabled by the increasing number of biobanks that contain coupled human phenotypic and genotypic information. Amongst this phenotypic information, image data provides a non-invasive but rich method of phenotyping humans, both in the research arena, but also in a clinical setting. OCT images of the eye are routinely collected in the clinic as well as by primary care opticians, and have been included in the UK Biobank. This has provided me with a novel research opportunity to explore basic ocular biology and pathology. The analysis of data that is routinely collected in clinical and primary care settings allows for broad translational potential.

Of the SNPs discovered to be associated with retinal morphology, several had no prior associations to additional phenotypes, ocular or otherwise. Following further inquiry, the genes the loci were labelled with suggested potential pathways through which their involvement in retinal morphology may act. Further investigation of these novel loci, for example using animal models, may help us better understand the intricacies of retinal morphology and function.

The use of MR to interrogate our understanding between the clinically utilised morphological traits and disease has developed our understanding of them as biomarkers. In the case of POAG, no evidence was found to support a causal genetic relationship between glaucoma and one of its clinical biomarkers, inner retinal thickness. This suggests it is a change in thickness over time that is indicative of the disease, and a genetically determined thinner inner retina does not predispose one to glaucoma. This suggests that the variants discovered within this thesis are unlikely to be viable therapeutic targets of the disease. However one opportunity would be to utilise the genetic variants

associated with non-pathological inner retina thickness to help estimate genetically determined inner retinal thickness and thus improve efficiency of glaucoma diagnosis. Exploration of this idea would require a dataset with higher glaucoma incidence. However this finding highlights the importance of reporting negative results and their value in research.

Conversely, evidence was found in support of a causal genetic relationship between AMD and RPE thickness, with RPE thickness being downstream of AMD. This suggests that those with a genetically determined thinner RPE may have a predisposition for AMD. It would be of great interest to further this work by testing the predictive power of RPE thickness for AMD. Unfortunately at this time, the incidence of AMD within the UK Biobank is insufficient to carry out such analysis. This is partly owing to the age of the participants being younger than is common for the development of AMD. Therefore an alternative dataset would be required for this analysis to be completed, or as the population within the UK Biobank ages and follow up measures are collected, this analysis may become feasible.

The work also highlighted the intersection between common and rare variation, and the spectrum of clinical and sub-clinical phenotypes this can affect. In particular the effects of common variants at *TYR* and *OCA2*, loci known to be associated with the rare clinical condition OCA is of particular interest. SNPs at these loci were consistently found associated with the thickness of all retinal layers. This consistency emphasises the importance of this finding and the need for additional exploration of the biological underpinnings of this observation. Additionally the presence of a phenotype associated with these common variants (central foveal thickness) that appears to be on the same spectrum as a clinical phenotype associated with OCA (foveal hypoplasia) is suggestive of a phenotype continuum with complex genetic causation. Previous work has looked at how some of these common variants interact with rare variants to produce the clinical phenotype. It would be of great interest to further expand this work in light of other common variants that may exhibit similar effect. The recent release of exome sequencing data for participants in the UK Biobank may offer one opportunity, however the incidence of OCA is very low, so an alternative clinical dataset may be better suited.

At an overarching level, care must be taken into how the results from this thesis are extrapolated to the general population. The analysis in this thesis was consistently done on the densest well-mixed population within the UK Biobank. This largely aligns with a self-reported white population. Selection of a single well-mixed population was a necessary measure to avoid confounding by population structure. However, the results are not automatically scalable to the entire population. The consistent bias of biomedical

research to focus on white populations influences subsequent medical decisions and can widen the healthcare inequality gap that runs along axes of ethnicity. This signals a requirement for me and the broader biomedical community to consciously evaluate the data we are using.

The work within this thesis demonstrates the utility and value of biomedical imaging derived phenotypes. This has already been established with application to other human organs such as the use of MRI in cardiology [Meyer & al., 2020] and neurology [Elliott & al., 2018]. Image derived phenotyping is advantageous in its non-invasive nature and density of data collected, and the work here provides further evidence for applying similar analysis to other organs. In the future similar work could be expanded to the skeleton using x-ray images and the reproductive organs using ultrasound.

Looking at the broader scope of the work in this thesis, the importance of studying phenotypes at a high dimensionality and the additional discoveries it can enable is clear. The clinical segmentation grids used within this research are medically useful, easily interpretable and have shown to provide interesting insight. However their segmentation boundaries may not best align with biological variation as morphological variation is a continuous trait. Initial expansion of dimensionality to look at intra layer variation in chapter 5 suggested that spatial variation of individual layers is occurring, with underlying causative genetics. Further expansion of phenotype dimensionality in chapter 6 to a more granular space allowed for additional biological discoveries. Considering the evidence for both inter- and intra-layer variation, further discoveries could result from expansion to a higher dimensional space applied to the individual layers. To facilitate this higher dimensional phenotype extraction, employment of cutting edge image analysis techniques, particularly machine learning methods, and collaboration with those with such expertise would be required. High level segmentation of retinal structure from OCT has already been conducted and proven to provide rich information that has diagnostic capabilities [De Fauw & al., 2018]. Harnessing similar technologies to interrogate non-pathological retinal morphology may help us better understand ocular biology.

In summary, this work provides a detailed approach to study the morphology of the human retina using biomedical imaging. It has demonstrated the insights into underlying genetics and function that can be discovered when we harness the high dimensional data available. This work has identified genetic variation underlying morphological variation of the human retina, opening new avenues of research in the fields of clinical ophthalmology and biological development. It also provides support for the power of this type of analysis in application to other areas of human physiology. This will help us

obtain a better understanding of general human biology, whilst giving clinical insights in a world with increasing biomedical phenotyping.

A. Supplementary Tables

A.1. Genome-wide Association Study of Inner Retinal Morphology

Table A.1: Full summary statistics for GWAS of RNFL and GCIPL thickness. Full GWAS summary statistics for 46 significant variants, including effect size, p-value and standard error (SE) for the individual phenotypes, GCIPL and RNFL (labelled accordingly) after meta-analysis, as well as the values selected within the meta analysis (labelled "MTAG"). A1 is the effect allele.

SNP	Chr	BP	A1	A2	AF	MTAG effect size	RNFL effect size	GCIPL effect size	MTAG p-value	RNFL p-value	GCIPL p-value	MTAG SE	RNFL SE	GCIPL SE
rs72739513	1	203080149	A	G	0.04	0.72	0.37	0.72	8.88E-09	1.05E-04	8.88E-09	0.13	0.10	0.13
rs12998032	2	159095496	C	T	0.44	0.30	0.08	0.30	6.97E-10	2.12E-02	6.97E-10	0.05	0.04	0.05
rs13010692	2	48800667	C	T	0.32	0.22	0.22	0.09	6.72E-09	6.72E-09	0.09	0.04	0.04	0.05
rs2271758	2	172701157	G	T	0.59	-0.22	-0.22	-0.20	1.34E-09	1.34E-09	3.63E-05	0.04	0.04	0.05
rs79833181	2	15666802	C	T	0.02	0.86	0.86	0.57	1.55E-09	1.55E-09	2.43E-03	0.14	0.14	0.19
rs980772	2	145442190	T	G	0.67	-0.21	-0.21	0.03	4.62E-08	4.62E-08	0.60	0.04	0.04	0.05
rs13083522	3	3270368	G	A	0.78	0.31	0.08	0.31	4.51E-08	0.08	4.51E-08	0.06	0.04	0.06
rs149831820	3	77192591	C	T	0.06	-0.42	-0.42	-0.31	2.53E-08	2.53E-08	1.48E-03	0.07	0.07	0.10
rs17279437	3	45814094	A	G	0.11	-0.77	-0.31	-0.77	7.81E-24	1.23E-07	7.81E-24	0.08	0.06	0.08
rs62252355	3	69572006	C	T	0.20	-0.37	-0.37	-0.15	2.17E-16	2.17E-16	1.22E-02	0.05	0.05	0.06
rs66511946	4	184932935	G	A	0.40	-0.36	-0.11	-0.36	2.15E-13	4.36E-03	2.15E-13	0.05	0.04	0.05
rs17421627	5	87847586	G	T	0.07	0.97	-0.14	0.97	8.09E-27	4.87E-02	8.09E-27	0.09	0.07	0.09
rs2004187	5	2612747	C	A	0.60	0.25	0.25	0.29	1.43E-11	1.43E-11	2.77E-09	0.04	0.04	0.05
rs13215351	6	84313801	T	A	0.25	-0.33	-0.06	-0.33	1.36E-09	0.12	1.36E-09	0.05	0.04	0.05
rs527871768	6	39419992	A	G	0.01	1.16	1.16	0.71	3.37E-10	3.37E-10	3.43E-03	0.19	0.19	0.24
rs9398171	6	108983527	T	C	0.71	0.50	0.23	0.50	7.51E-22	1.45E-08	7.51E-22	0.05	0.04	0.05
rs11762530	7	46630602	C	G	0.59	0.53	0.07	0.53	3.45E-28	4.77E-02	3.45E-28	0.05	0.04	0.05
rs12719025	7	51100190	G	A	0.46	0.30	0.04	0.30	3.09E-10	0.29	3.09E-10	0.05	0.04	0.05
rs35001871	7	50975439	C	G	0.29	0.32	0.19	0.32	1.17E-09	2.98E-06	1.17E-09	0.05	0.04	0.05
rs73348111	7	50364291	C	T	0.01	1.12	1.12	0.66	7.15E-10	7.15E-10	5.70E-03	0.18	0.18	0.24
rs115520750	8	108739734	T	G	0.01	1.12	1.12	0.82	2.54E-10	2.54E-10	4.51E-04	0.18	0.18	0.23
rs13271359	8	109114426	T	C	0.26	-0.44	-0.44	-0.02	5.89E-26	5.89E-26	0.75	0.04	0.04	0.05
rs376067714	8	109141863	G	A	0.18	-0.48	-0.48	-0.11	1.28E-18	1.28E-18	0.13	0.05	0.05	0.07
rs4871827	8	121061879	A	G	0.33	-0.29	-0.12	-0.29	7.41E-09	2.00E-03	7.41E-09	0.05	0.04	0.05
rs6989495	8	74230223	T	G	0.33	0.23	0.23	0.14	1.27E-09	1.27E-09	4.33E-03	0.04	0.04	0.05
rs118031671	9	10521068	G	T	0.01	0.93	0.93	0.76	2.53E-09	2.53E-09	2.21E-04	0.16	0.16	0.21
rs2787394	9	103007414	T	C	0.41	-0.28	-0.06	-0.28	8.64E-09	0.09	8.64E-09	0.05	0.04	0.05
rs10762201	10	70040111	G	A	0.76	0.46	0.46	0.15	1.05E-26	1.05E-26	6.72E-03	0.04	0.04	0.06
rs181211282	10	102746829	A	G	0.03	0.67	0.67	0.24	1.12E-08	1.12E-08	0.11	0.12	0.12	0.15
rs1947075	10	49741135	T	C	0.64	-0.21	-0.21	-0.19	2.60E-08	2.60E-08	1.14E-04	0.04	0.04	0.05
rs1042602	11	88911696	A	C	0.37	-0.36	-0.36	-0.33	3.96E-22	3.96E-22	2.10E-11	0.04	0.04	0.05
rs12574166	11	69291285	T	C	0.15	0.28	0.28	0.24	2.82E-08	2.82E-08	3.92E-04	0.05	0.05	0.07
rs2008905	11	17184623	T	C	0.42	-0.36	-0.16	-0.36	6.81E-14	1.92E-05	6.81E-14	0.05	0.04	0.05
rs5442	12	6954864	A	G	0.07	-0.69	-0.08	-0.69	2.36E-13	0.27	2.36E-13	0.09	0.07	0.09
rs10140252	14	74528023	T	G	0.16	0.53	0.53	0.49	1.05E-25	1.05E-25	1.39E-13	0.05	0.05	0.07
rs1254276	14	60847001	T	C	0.39	-0.28	-0.28	-0.10	7.52E-14	7.52E-14	3.24E-02	0.04	0.04	0.05
rs146652416	14	29907103	G	A	0.03	0.62	0.62	0.45	4.46E-08	4.46E-08	2.44E-03	0.11	0.11	0.15
rs17095953	14	59719393	A	G	0.24	-0.28	-0.28	-0.08	1.76E-10	1.76E-10	0.14	0.04	0.04	0.06
rs35337422	14	104407243	C	A	0.15	0.37	-0.20	0.37	3.50E-08	1.28E-04	3.50E-08	0.07	0.05	0.07
rs1470108	15	89153744	A	C	0.34	0.24	0.24	0.21	7.09E-10	7.09E-10	3.77E-05	0.04	0.04	0.05
rs1800407	15	28230318	T	C	0.08	-0.60	-0.36	-0.60	3.19E-12	4.08E-08	3.19E-12	0.09	0.07	0.09
rs117304899	16	15055042	G	C	0.02	0.92	0.92	0.75	1.74E-09	1.74E-09	1.82E-04	0.15	0.15	0.20
rs117300236	17	44753350	G	A	0.72	-0.25	-0.25	-0.30	5.57E-09	5.57E-09	1.24E-07	0.04	0.04	0.06
rs7503894	17	79583473	C	T	0.65	0.56	0.37	0.56	2.49E-29	3.58E-22	2.49E-29	0.05	0.04	0.05
rs143330165	20	7154672	T	C	0.01	1.04	1.04	0.41	1.97E-08	1.97E-08	0.09	0.19	0.19	0.24
rs7277632	21	47327542	G	A	0.72	0.34	0.09	0.34	1.20E-10	3.56E-02	1.20E-10	0.05	0.04	0.05

Table A.2: Association of UK Biobank inner retinal thickness-associated variants with ganglion cell complex thickness in the Rotterdam Study. Comparison between effect sizes and p-values from meta-analysed inner retinal GWAS (labelled "MTAG"), and GWAS of the GCC thickness in the Rotterdam study (labelled "Rotterdam").

SNP	MTAG effect size	MTAG p-value	Rotterdam effect size	Rotterdam p-value
rs1042602	-0.36	3.96E-22	-0.77	6.71E-03
rs10762201	0.46	1.05E-26	0.10	0.78
rs117300236	-0.25	5.57E-09	-0.03	0.94
rs11762530	0.53	3.45E-28	0.25	0.39
rs1254276	-0.28	7.52E-14	-0.20	0.48
rs12574166	0.28	2.82E-08	0.21	0.59
rs12719025	0.30	3.09E-10	0.99	5.89E-04
rs12998032	0.30	6.97E-10	0.22	0.44
rs13010692	0.22	6.72E-09	0.54	0.08
rs13083522	0.31	4.51E-08	0.03	0.93
rs13215351	-0.33	1.36E-09	-0.04	0.91
rs1470108	0.24	7.09E-10	0.34	0.27
rs149831820	-0.42	2.53E-08	0.27	0.69
rs17279437	-0.77	7.81E-24	-0.79	0.12
rs17421627	0.97	8.09E-27	1.62	2.21E-03
rs1800407	-0.60	3.19E-12	-1.12	0.16
rs1947075	-0.21	2.60E-08	-0.07	0.80
rs2004187	0.25	1.43E-11	0.31	0.29
rs2008905	-0.36	6.81E-14	0.07	0.81
rs2271758	-0.22	1.34E-09	-0.51	0.07
rs2787394	-0.28	8.64E-09	-0.36	0.21
rs35001871	0.32	1.17E-09	0.36	0.25
rs35337422	0.37	3.50E-08	0.24	0.55
rs4871827	-0.29	7.41E-09	-0.87	4.49E-03
rs5442	-0.69	2.36E-13	0.38	0.46
rs62252355	-0.37	2.17E-16	-0.27	0.43
rs66511946	-0.36	2.15E-13	-0.51	0.11
rs6989495	0.23	1.27E-09	1.00	9.50E-04
rs7277632	0.34	1.20E-10	1.35	2.25E-05
rs7503894	0.56	2.49E-29	0.57	0.06
rs9398171	0.50	7.51E-22	0.69	2.55E-02
rs980772	-0.21	4.62E-08	-0.23	0.43

Table A.3: Association of UK Biobank inner retinal thickness-associated variants with retinal nerve fibre layer (RNFL) and ganglion cell inner plexiform layer (GCIPL) thickness in the Raine Study. Comparison between effect sizes and p-values from meta-analysed inner retinal GWAS (labelled "MTAG"), and GWAS of the RNFL and GCIPL in the Raine Study (labelled "Raine RNFL" and "Raine GCIPL").

SNP	MTAG effect size	MTAG p-value	Raine RNFL effect size	Raine RNFL p-value	Raine GCIPL effect size	Raine GCIPL p-value
rs1042602	-0.36	3.96E-22	-0.20	0.11	-	-
rs10762201	0.46	1.05E-26	0.42	2.95E-03	-	-
rs115520750	1.12	2.54E-10	-0.66	0.39	-	-
rs117300236	-0.25	5.57E-09	-0.15	0.28	-	-
rs11762530	0.53	3.45E-28	-	-	0.78	8.00E-04
rs118031671	0.93	2.53E-09	-0.26	0.66	-	-
rs1254276	-0.28	7.52E-14	-0.24	0.05	-	-
rs12574166	0.28	2.82E-08	0.35	3.32E-02	-	-
rs12719025	0.30	3.09E-10	-	-	0.57	1.66E-02
rs12998032	0.30	6.97E-10	-	-	0.50	2.64E-02
rs13010692	0.22	6.72E-09	0.09	0.49	-	-
rs13083522	0.31	4.51E-08	-	-	-0.28	0.32
rs13215351	-0.33	1.36E-09	-	-	-0.27	0.30
rs143330165	1.04	1.97E-08	-1.81	0.21	-	-
rs146652416	0.62	4.46E-08	0.37	0.38	-	-
rs1470108	0.24	7.09E-10	0.10	0.43	-	-
rs149831820	-0.42	2.53E-08	-0.37	0.13	-	-
rs17279437	-0.77	7.81E-24	-	-	-1.76	1.53E-06
rs17421627	0.97	8.09E-27	-	-	0.58	0.20
rs1800407	-0.60	3.19E-12	-0.28	0.18	-0.49	0.23
rs181211282	0.67	1.12E-08	-0.19	0.65	-	-
rs1947075	-0.21	2.60E-08	0.02	0.86	-	-
rs2004187	0.25	1.43E-11	0.20	0.10	0.45	0.06
rs2008905	-0.36	6.81E-14	-	-	0.03	0.90
rs2271758	-0.22	1.34E-09	-0.41	9.13E-04	-	-
rs2787394	-0.28	8.64E-09	-	-	-0.69	3.71E-03
rs35337422	0.37	3.50E-08	-	-	0.12	0.71
rs4871827	-0.29	7.41E-09	-	-	-0.34	0.16
rs5442	-0.69	2.36E-13	-	-	-1.04	2.18E-02
rs62252355	-0.37	2.17E-16	-0.05	0.72	-	-
rs66511946	-0.36	2.15E-13	-	-	-0.70	1.07E-02
rs6989495	0.23	1.27E-09	-0.01	0.96	-	-
rs72739513	0.72	8.88E-09	-	-	0.26	0.71
rs7277632	0.34	1.20E-10	-	-	0.69	6.14E-03
rs73348111	1.12	7.15E-10	0.88	0.21	-	-
rs7503894	0.56	2.49E-29	-	-	0.38	0.14
rs79833181	0.86	1.55E-09	-0.06	0.92	-	-
rs9398171	0.50	7.51E-22	-	-	0.37	0.14
rs980772	-0.21	4.62E-08	-0.35	5.36E-03	-	-

Table A.4: Association of UK Biobank inner retinal thickness-associated variants with primary open-angle glaucoma in the International Glaucoma Genetics Consortium meta-analysis [Gharahkhani & al., 2020] Each variant is annotated with the nearest gene and the odds ratio (OR) for glaucoma reported alongside the p-value.

SNP	Risk allele	Other allele	Chr	Position	Nearest gene	OR for POAG	P-value
rs1254276	T	C	14	60847001	SIX6	1.18	2.03E-37
rs117300236	G	A	17	44753350	NSF	0.93	1.82E-05
rs5442	A	G	12	6954864	GNB3	1.11	9.97E-05
rs11762530	C	G	7	46630602	IGFBP3	1.04	1.52E-03
rs2787394	T	C	9	103007414	INVS	0.97	8.44E-03
rs1042602	A	C	11	88911696	TYR	1.03	1.47E-02
rs2008905	T	C	11	17184623	PIK3C2A	0.97	2.35E-02
rs7503894	C	T	17	79583473	NPLOC4	0.97	2.56E-02
rs17279437	A	G	3	45814094	SLC6A20	0.95	2.87E-02
rs10762201	G	A	10	70040111	ATOH7	0.97	2.95E-02
rs9398171	T	C	6	108983527	FOXO3	1.03	3.79E-02
rs17421627	G	T	5	87847586	LINC00461	1.05	0.07
rs2004187	C	A	5	2612747	IRX2	1.02	0.14
rs13271359	T	C	8	109114426	RSPO2	1.03	0.18
rs73348111	C	T	7	50364291	IKZF1	1.08	0.19
rs17095953	A	G	14	59719393	DAAM1	1.03	0.24
rs35337422	C	A	14	104407243	TDRD9	1.02	0.25
rs1470108	A	C	15	89153744	AEN	1.02	0.27
rs115520750	T	G	8	108739734	ANGPT1	1.07	0.37
rs79833181	C	T	2	15666802	NBAS	1.05	0.37
rs62252355	C	T	3	69572006	FRMD4B	1.01	0.39
rs12574166	T	C	11	69291285	LINC02747	0.99	0.49
rs13010692	C	T	2	48800667	STON1-GTF2A1L	1.01	0.55
rs143330165	T	C	20	7154672	LINC01428	1.07	0.57
rs72739513	A	G	1	203080149	ADORA1	0.98	0.58
rs117304899	G	C	16	15055042	NA	1.05	0.62
rs12719025	G	A	7	51100190	COBL	1.01	0.63
rs2271758	G	T	2	172701157	SLC25A12	0.99	0.63
rs980772	T	G	2	145442190	TEX41	0.99	0.64
rs181211282	A	G	10	102746829	MRPL43	0.98	0.67
rs4871827	A	G	8	121061879	DEPTOR	1.01	0.69
rs13083522	G	A	3	3270368	CRBN	0.99	0.69
rs66511946	G	A	4	184932935	STOX2	1.00	0.74
rs118031671	G	T	9	10521068	PTPRD	1.01	0.77
rs10140252	T	G	14	74528023	BBOF1	1.01	0.78
rs12998032	C	T	2	159095496	CCDC148	1.00	0.80
rs1800407	T	C	15	28230318	OCA2	0.99	0.81
rs13215351	T	A	6	84313801	SNAP91	1.00	0.86
rs146652416	G	A	14	29907103	FOXG1	0.99	0.86
rs6989495	T	G	8	74230223	RDH10	1.00	0.87
rs7277632	G	A	21	47327542	PCBP3	1.00	0.92
rs1947075	T	C	10	49741135	ARHGAP22	1.00	0.99
rs149831820	C	T	3	77192591	ROBO2	1.00	0.99

A.2. Genome-wide Association Study of Outer Retinal Morphology

Table A.5: Full summary statistics for GWAS of ONL thickness. Full GWAS summary statistics for the 74 variants associated with ONL thickness, including effect size, p-value and standard error (SE). A1 is the effect allele. Allele frequency (AF) refers to the frequency of the effect allele.

SNP	Chr	BP	A1	A2	AF	Effect size	P-value	SE
rs2128416	1	10700448	C	T	0.15	-0.52	1.91E-12	0.07
rs6427827	1	200398387	A	T	0.62	0.42	7.24E-16	0.05
rs4950952	1	202739602	G	A	0.51	0.36	4.43E-13	0.05
rs919655	1	214157972	A	G	0.12	0.47	1.05E-09	0.08
rs12032598	1	222148552	C	A	0.21	0.38	3.05E-09	0.06
rs6426584	1	227374949	T	A	0.34	0.30	1.64E-08	0.05
rs116350483	2	145338686	T	C	0.02	-2.13	1.05E-26	0.20
rs80265589	2	169027979	A	T	0.27	-0.42	1.40E-13	0.06
rs148388367	2	216850944	A	T	0.04	-1.27	1.08E-20	0.14
rs72959706	2	218533177	G	A	0.07	-0.60	1.84E-09	0.10
rs11129176	3	25049310	A	G	0.28	0.31	1.66E-08	0.06
rs6775323	3	27721085	T	G	0.20	0.49	3.40E-15	0.06
rs4681617	3	150122569	C	T	0.10	0.47	1.94E-08	0.08
rs30373	5	55745334	G	C	0.37	0.32	9.60E-10	0.05
rs63338061	5	71486228	T	C	0.36	0.33	4.66E-10	0.05
rs17421627	5	87847586	G	T	0.07	1.15	3.64E-33	0.10
rs62391700	5	126093395	C	T	0.20	-0.43	1.08E-11	0.06
rs1109114	5	148615946	T	C	0.43	0.38	6.11E-14	0.05
rs6875105	5	173054917	C	T	0.38	-0.37	1.50E-12	0.05
rs12192672	6	7229619	A	G	0.30	-0.33	1.23E-09	0.05
rs17507554	6	11394287	A	G	0.05	0.69	3.73E-09	0.12
rs4711420	6	35495811	T	G	0.80	0.36	1.08E-08	0.06
rs6910414	6	56726737	G	A	0.18	0.37	1.12E-08	0.07
rs74526772	6	106515218	A	T	0.04	-1.25	1.35E-19	0.14
rs9639276	7	867033	T	C	0.17	0.43	1.12E-10	0.07
rs929511	7	8003017	T	C	0.12	0.46	2.97E-09	0.08
rs12719025	7	51100190	G	A	0.46	0.49	9.38E-22	0.05
rs34926272	7	129591807	C	G	0.03	-1.07	9.77E-12	0.16
rs61675430	8	61671071	A	G	0.20	-0.41	1.42E-10	0.06
rs13263941	8	109121945	C	T	0.26	0.85	6.15E-50	0.06
rs376067714	8	109141863	G	A	0.18	0.77	4.39E-24	0.08
rs9298817	9	21576591	C	A	0.66	-0.49	2.67E-20	0.05
rs10781177	9	76593011	T	C	0.42	-0.29	1.02E-08	0.05
rs717299	9	77185933	G	A	0.45	0.36	1.28E-12	0.05
rs111245635	10	48389841	T	C	0.02	1.29	3.56E-10	0.21
rs1947075	10	49741135	T	C	0.64	-0.33	5.92E-10	0.05
rs7916697	10	69991853	G	A	0.76	-0.43	3.36E-13	0.06
rs1016934	11	31720621	G	A	0.30	-0.37	1.64E-11	0.06
rs116233906	11	68968271	A	C	0.04	-0.93	2.74E-13	0.13
rs1485995	11	69307707	A	G	0.67	-0.33	1.34E-09	0.05
rs6483429	11	95239787	T	C	0.46	-0.30	4.51E-09	0.05
rs3138142	12	56115585	T	C	0.24	0.93	1.64E-56	0.06
rs76629482	12	96178789	G	C	0.18	-0.53	6.87E-16	0.07
rs2793783	13	100253216	G	A	0.62	-0.29	3.88E-08	0.05
rs28468687	14	36026342	A	G	0.18	0.38	1.30E-08	0.07
rs1956524	14	68800393	A	G	0.64	0.29	3.49E-08	0.05
rs10135971	14	69517494	A	G	0.34	0.41	1.11E-14	0.05
rs112145470	14	74356090	A	G	0.03	1.30	2.87E-15	0.17
rs368205955	14	74497636	G	T	0.02	-3.21	1.85E-46	0.22
rs12147951	14	74642451	C	A	0.07	-2.36	5.13E-113	0.10
rs1972565	14	74666824	G	A	0.82	1.92	8.13E-191	0.06
rs1972564	14	74666944	T	C	0.48	1.03	7.10E-74	0.06
rs118186707	14	74686575	A	G	0.03	1.78	9.73E-33	0.15
rs28488340	14	74693803	G	C	0.40	-0.59	2.62E-30	0.05
rs888413	14	75267396	T	C	0.47	0.39	7.11E-15	0.05
rs12904079	15	45483124	A	G	0.34	0.30	2.64E-08	0.05
rs10083695	15	53987147	G	A	0.49	-0.37	3.48E-13	0.05
rs1372613	15	101204835	T	C	0.30	0.36	6.90E-11	0.05
rs7206532	16	80490131	C	T	0.52	0.29	5.19E-09	0.05

Table A.5: continued

rs4782636	16	84585176	T	A	0.19	0.41	2.51E-10	0.06
rs35444818	16	85718918	A	C	0.29	-0.33	1.76E-09	0.06
rs111825734	17	44019107	G	T	0.23	0.48	6.28E-16	0.06
rs8077430	17	79511135	G	A	0.77	-0.49	2.78E-16	0.06
rs8070929	17	79530993	T	G	0.35	-0.43	3.60E-16	0.05
rs57820851	18	53413566	T	G	0.35	-0.30	1.58E-08	0.05
rs189921298	18	56922264	G	A	0.30	0.46	1.70E-16	0.06
rs17696543	18	56971398	T	C	0.18	-0.56	1.71E-17	0.07
rs76076446	19	3771586	A	G	0.02	-1.23	2.68E-13	0.17
rs8132685	21	34220618	T	C	0.52	-0.30	4.29E-09	0.05
rs2032576	22	27089655	T	G	0.56	-0.28	4.46E-08	0.05
rs5752638	22	28188203	C	T	0.22	-0.51	2.20E-16	0.06
rs5763593	22	30302238	C	T	0.63	0.35	2.57E-11	0.05
rs2073946	22	30619599	A	G	0.31	0.47	2.05E-17	0.06
rs75159625	22	46377008	G	T	0.31	0.34	3.92E-10	0.05

Table A.6: Full summary statistics for GWAS of IS thickness. Full GWAS summary statistics for the 15 variants associated with IS thickness, including effect size, p-value and standard error (SE). A1 is the effect allele. Allele frequency (AF) refers to the frequency of the effect allele.

SNP	Chr	BP	A1	A2	AF	Effect size	P-value	SE
rs112248193	1	3718402	A	G	0.13	-0.11	1.39E-09	0.02
rs6665290	1	227201106	T	C	0.46	0.07	3.28E-09	0.01
rs2351248	3	47829277	T	G	0.70	-0.07	2.75E-08	0.01
rs1463229	3	150167912	G	A	0.76	-0.09	2.84E-11	0.01
rs72761109	5	71506529	T	C	0.31	0.07	2.31E-08	0.01
rs7816990	8	10463944	C	A	0.74	-0.08	7.13E-09	0.01
rs3779791	8	87676710	C	T	0.76	0.10	4.46E-12	0.01
rs2808051	10	32205630	C	T	0.51	0.07	2.47E-09	0.01
rs35314358	10	62647787	A	T	0.22	0.09	2.56E-09	0.01
rs2292962	11	1778321	T	C	0.13	0.10	2.78E-08	0.02
rs618838	11	66328719	C	T	0.55	0.07	2.27E-09	0.01
rs1272131	14	60886150	C	T	0.39	-0.09	1.03E-13	0.01
rs887595	14	74666641	G	A	0.82	0.18	8.38E-33	0.02
rs10220961	16	71789550	G	T	0.48	0.08	4.27E-12	0.01
rs35638197	17	47284735	C	T	0.06	-0.14	2.12E-08	0.02

Table A.8: Full summary statistics for GWAS of PRL component layer thickness. Full GWAS summary statistics for 59 significant variants, including effect size, p-value and standard error (SE) for the individual phenotypes, IS, OS and ONL (labelled accordingly), as well as the values selected within the meta analysis (labelled "MTAG"). A1 is the effect allele.

SNP	Chr	BP	A1	A2	AF	MTAG effect size	IS effect size	OS effect size	ONL effect size	MTAG p-value	IS p-value	OS p-value	ONL p-value	MTAG SE	IS SE	OS SE	ONL SE
rs11587687	1	110147801	A	G	0.44	0.18	0.01	0.18	-0.08	5.02E-09	0.30	5.02E-09	0.12	0.03	0.01	0.03	0.05
rs2932531	1	113236309	A	T	0.55	0.20	0.01	0.20	0.02	2.79E-09	0.43	2.79E-09	0.70	0.03	0.01	0.03	0.05
rs410895	1	196898226	C	T	0.55	0.30	-0.01	0.30	0.06	8.19E-22	0.17	8.19E-22	0.23	0.03	0.01	0.03	0.05

Table A.8: continued

rs6673000	1	200451943	G	A	0.56	0.37	0.04	-0.03	0.37	2.73E-13	8.15E-05	0.33	2.73E-13	0.05	0.01	0.03	0.05
rs11577827	1	202798831	T	C	0.53	0.37	0.05	0.07	0.37	2.77E-13	6.56E-06	0.03	2.77E-13	0.05	0.01	0.03	0.05
rs6665290	1	227201106	T	C	0.46	0.07	0.07	-0.04	0.24	8.24E-11	8.24E-11	0.15	2.12E-06	0.01	0.01	0.03	0.05
rs2113422	2	23939007	G	A	0.58	0.20	0.02	0.20	0.17	2.89E-10	0.14	2.89E-10	1.08E-03	0.03	0.01	0.03	0.05
rs13382582	2	169079662	G	C	0.42	0.18	0.04	0.18	0.18	5.99E-09	5.09E-04	5.99E-09	4.05E-04	0.03	0.01	0.03	0.05
rs838718	2	234296650	A	G	0.48	-0.17	-2.13E-03	-0.17	-0.08	2.05E-08	0.77	2.05E-08	0.12	0.03	0.01	0.03	0.05
rs34234056	3	14418444	G	A	0.51	0.17	-1.87E-03	0.17	0.01	2.37E-08	0.86	2.37E-08	0.78	0.03	0.01	0.03	0.05
rs348866	3	100992385	T	C	0.49	0.22	0.01	0.22	-0.10	1.96E-13	0.19	1.96E-13	0.05	0.03	0.01	0.03	0.05
rs1376026	3	150182592	G	A	0.51	0.18	-0.02	0.18	0.14	5.86E-09	0.12	5.86E-09	6.11E-03	0.03	0.01	0.03	0.05
rs2929724	5	17200427	G	C	0.52	0.17	-2.71E-03	0.17	2.96E-03	4.21E-08	0.80	4.21E-08	0.95	0.03	0.01	0.0315	0.05
rs12653396	5	87847273	A	T	0.57	0.34	0.05	0.06	0.34	4.67E-11	1.54E-06	0.07	4.67E-11	0.05	0.01	0.03	0.05
rs1109114	5	148615946	T	C	0.43	0.37	0.07	-0.04	0.37	2.65E-13	1.28E-10	0.23	2.65E-13	0.05	0.01	0.03	0.05
rs1438692	5	148659664	A	G	0.42	0.07	0.07	-0.01	0.32	1.32E-11	1.32E-11	0.75	3.59E-10	0.01	0.01	0.03	0.05
rs450611	5	173038618	C	A	0.59	-0.32	-0.03	-0.05	-0.32	5.82E-10	1.37E-03	0.13	5.82E-10	0.05	0.01	0.03	0.05
rs375435	6	42661404	C	T	0.57	-0.25	-0.01	-0.25	-0.06	3.56E-16	0.62	3.56E-16	0.20	0.03	0.01	0.03	0.05
rs1926098	6	76715926	A	G	0.56	-0.17	0.02	-0.17	-0.06	3.35E-08	0.10	3.35E-08	0.27	0.03	0.01	0.03	0.05
rs12719025	7	51100190	G	A	0.46	0.49	0.07	-3.73E-03	0.49	6.39E-22	1.33E-10	0.90	6.39E-22	0.05	0.01	0.03	0.05
rs2437002	8	109034419	T	G	0.58	0.37	0.05	0.06	0.37	5.77E-13	1.06E-06	0.07	5.77E-13	0.05	0.01	0.03	0.05
rs2514842	8	109062046	G	T	0.47	0.41	0.07	0.12	0.41	3.84E-16	2.59E-10	1.01E-04	3.84E-16	0.05	0.01	0.03	0.05
rs10781177	9	76593011	T	C	0.42	-0.29	-0.05	-0.01	-0.29	1.49E-08	1.12E-05	0.64	1.49E-08	0.05	0.01	0.03	0.05
rs717299	9	77185933	G	A	0.45	0.37	0.04	0.06	0.37	3.80E-13	9.57E-05	0.04	3.80E-13	0.05	0.01	0.03	0.05
rs11200922	10	85961758	G	A	0.46	0.32	-0.03	0.33	0.08	1.62E-26	0.01	1.62E-26	0.12	0.03	0.01	0.03	0.05
rs618838	11	66328719	C	T	0.55	0.06	0.06	0.03	0.19	1.79E-09	1.79E-09	0.28	1.49E-04	0.01	0.01	0.03	0.05
rs7128814	11	68953054	A	G	0.58	0.28	0.04	0.07	0.28	4.56E-08	9.68E-05	0.02	4.56E-08	0.05	0.01	0.03	0.05
rs631695	11	69283303	G	T	0.59	-0.31	-0.04	-0.02	-0.31	1.71E-09	9.01E-04	0.46	1.71E-09	0.05	0.01	0.03	0.05
rs948962	11	76919478	A	C	0.47	-0.20	0.01	-0.20	0.01	4.55E-11	0.40	4.55E-11	0.87	0.03	0.01	0.038	0.05
rs28620862	11	88979097	C	A	0.47	0.20	-0.01	0.20	-0.02	7.57E-11	0.45	7.57E-11	0.74	0.03	0.01	0.03	0.05
rs7930541	11	95238880	C	A	0.42	-0.30	-0.04	-0.01	-0.30	2.65E-09	5.64E-04	0.83	2.65E-09	0.05	0.01	0.03	0.05
rs2080402	12	345175	C	T	0.56	0.22	-0.03	0.22	-0.13	2.21E-12	2.18E-03	2.21E-12	0.01	0.03	0.01	0.03	0.05
rs3138144	12	56114769	C	G	0.49	0.56	0.06	0.26	0.56	6.54E-28	1.89E-09	1.17E-16	6.54E-28	0.05	0.01	0.03	0.05
rs6538677	12	96254528	G	A	0.60	0.32	0.04	0.01	0.32	7.36E-10	4.44E-05	0.64	7.36E-10	0.05	0.01	0.03	0.05
rs9796234	13	114323997	C	T	0.52	0.27	0.03	0.27	0.19	4.69E-18	3.13E-03	4.69E-18	1.57E-04	0.03	0.01	0.03	0.05
rs7151797	14	69499509	C	T	0.42	0.39	0.03	0.02	0.39	5.46E-14	2.82E-03	0.58	5.46E-14	0.05	0.01	0.03	0.05
rs11626048	14	74282511	C	T	0.54	0.72	0.09	0.04	0.72	1.08E-45	1.59E-16	0.21	1.08E-45	0.05	0.01	0.03	0.05
rs10147158	14	74465741	T	A	0.54	-0.41	-0.08	-0.02	-0.41	9.62E-16	8.15E-13	0.54	9.62E-16	0.05	0.01	0.03	0.05
rs11159058	14	74497642	G	T	0.56	-0.56	-0.09	-0.02	-0.56	1.30E-24	9.28E-16	0.52	1.30E-24	0.05	0.01	0.03	0.05
rs11159063	14	74614566	C	T	0.53	0.75	0.11	0.02	0.75	2.71E-50	9.84E-25	0.51	2.71E-50	0.05	0.01	0.03	0.05
rs1972564	14	74666944	T	C	0.48	1.03	0.15	-0.04	1.03	2.93E-74	1.40E-37	0.22	2.93E-74	0.06	0.01	0.03	0.06
rs28488340	14	74693803	G	C	0.40	-0.59	-0.08	-0.05	-0.59	9.02E-31	1.07E-13	0.14	9.02E-31	0.05	0.01	0.03	0.05
rs888413	14	75267396	T	C	0.47	0.39	0.07	0.07	0.39	1.92E-14	3.64E-10	0.02	1.92E-14	0.05	0.01	0.03	0.05
rs1648303	15	45445692	G	A	0.57	-0.28	-0.04	-0.01	-0.28	2.72E-08	8.32E-04	0.66	2.72E-08	0.05	0.01	0.03	0.05
rs10083695	15	53987147	G	A	0.49	-0.36	-0.06	-0.06	-0.36	8.00E-13	5.67E-09	0.03	8.00E-13	0.05	0.01	0.03	0.05
rs3825991	15	89761664	A	C	0.47	0.36	-0.04	0.36	0.05	2.70E-31	1/10E-04	2.70E-31	0.29	0.03	0.01	0.03	0.05
rs7206532	16	80490131	C	T	0.52	0.30	0.03	0.02	0.30	1.86E-09	3.06E-03	0.56	1.86E-09	0.05	0.01	0.03	0.05
rs6564059	16	84587249	T	G	0.43	-0.18	0.03	-0.18	0.15	1.16E-08	2.30E-03	1.16E-08	2.87E-03	0.03	0.01	0.03	0.05
rs11640338	16	85689653	G	A	0.46	0.29	0.03	3.28E-03	0.29	2.25E-08	3.93E-03	0.94	2.25E-08	0.05	0.01	0.03	0.05
rs1852269	17	43655273	G	A	0.45	0.38	0.05	-0.01	0.38	1.88E-09	8.63E-05	0.85	1.88E-09	0.06	0.01	0.04	0.06
rs11079866	17	47324485	C	T	0.53	0.17	-0.03	0.17	0.05	1.58E-08	0.00509477	1.58E-08	0.32	0.03	0.01	0.03	0.05
rs56737642	17	79515509	G	A	0.48	-0.34	-0.06	0.21	-0.34	2.87E-11	3.90E-08	4.64E-11	2.87E-11	0.05	0.01	0.03	0.05
rs62075724	17	79611410	C	T	0.46	0.38	0.07	-0.21	0.38	2.91E-12	1.94E-09	1.12E-10	2.91E-12	0.05	0.01	0.03	0.05
rs9910935	17	79613949	C	T	0.53	-0.34	0.05	-0.34	0.30	2.75E-28	8.88E-07	2.75E-28	3.96E-09	0.03	0.01	0.03	0.05
rs4940460	18	57045134	A	G	0.50	-0.35	-0.06	0.03	-0.35	2.86E-12	1.74E-07	0.34	2.86E-12	0.05	0.01	0.03	0.05
rs6077977	20	10930708	G	A	0.49	0.27	-0.01	0.27	-0.09	4.98E-19	0.28	4.98E-19	0.08	0.03	0.01	0.03	0.05
rs8132685	21	34220618	T	C	0.52	-0.30	-0.04	-0.02	-0.30	3.58E-09	1.01E-04	0.54	3.58E-09	0.05	0.01	0.03	0.05
rs5752975	22	30274335	G	C	0.40	0.31	0.06	4.89E-047	0.31	1.80E-09	9.13E-08	0.98	1.80E-09	0.05	0.01	0.03	0.05
rs8138196	22	30656635	G	A	0.53	-0.30	-0.05	-0.07	-0.30	2.38E-09	8.70E-07	0.03	2.38E-09	0.05	0.01	0.03	0.05

Table A.7: Full summary statistics for GWAS of OS thickness. Full GWAS summary statistics for 38 variants associated with OS thickness, including effect size, p-value and standard error (SE). A1 is the effect allele. Allele frequency (AF) refers to the frequency of the effect allele.

SNP	Chr	BP	A1	A2	AF	Effect size	P-value	SE
rs11587687	1	110147801	A	G	0.44	0.19	7.53E-10	0.03
rs3790612	1	113084146	A	G	0.26	-0.27	1.87E-15	0.03
rs72683442	1	113468825	C	T	0.22	-0.34	9.79E-21	0.04
rs410895	1	196898226	C	T	0.55	0.30	7.01E-22	0.03
rs7594221	2	24112724	A	T	0.76	0.27	3.30E-14	0.04
rs6737602	2	169062731	C	T	0.39	0.19	5.14E-10	0.03
rs1866665	2	198925980	G	A	0.17	0.23	1.24E-08	0.04
rs148388367	2	216850944	A	T	0.04	0.64	1.03E-14	0.08
rs7564805	2	234228946	G	A	0.05	-0.79	1.48E-29	0.07
rs34234056	3	14418444	G	A	0.51	0.17	1.53E-08	0.03
rs62282867	3	100972078	A	G	0.79	0.35	9.46E-21	0.04
rs111163508	3	129235423	C	G	0.16	-0.40	3.65E-22	0.04
rs7430585	3	150112041	A	G	0.76	0.22	4.46E-10	0.04
rs115237855	5	17186336	A	G	0.03	-0.69	8.57E-13	0.10
rs78303234	5	17218432	G	C	0.12	-0.26	4.08E-08	0.05
rs2326838	6	6901663	A	G	0.36	-0.23	7.21E-13	0.03
rs9470055	6	35521961	A	G	0.82	0.22	1.97E-08	0.04
rs375435	6	42661404	C	T	0.57	-0.25	2.36E-16	0.03
rs947340	6	76747944	C	A	0.70	-0.25	1.50E-13	0.03
rs111963714	7	99948655	G	T	0.21	0.27	1.27E-12	0.04
rs62490856	8	10469030	A	G	0.13	0.35	3.54E-14	0.05
rs11200922	10	85961758	G	A	0.46	0.32	2.68E-26	0.03
rs60401382	10	124227624	T	C	0.23	-0.30	5.32E-17	0.04
rs12574286	11	76937602	C	G	0.21	-0.30	2.01E-15	0.04
rs1126809	11	89017961	A	G	0.30	-0.34	3.32E-24	0.03
rs2080402	12	345175	C	T	0.56	0.22	2.77E-13	0.03
rs3138141	12	56115778	A	C	0.25	0.43	4.00E-35	0.03
rs9796234	13	114323997	C	T	0.52	0.26	1.15E-17	0.03
rs1254260	14	60835737	A	G	0.29	-0.23	1.52E-11	0.03
rs1800407	15	28230318	T	C	0.08	0.58	9.33E-26	0.06
rs183626410	15	55738274	T	C	0.01	-0.73	1.59E-08	0.13
rs3825991	15	89761664	A	C	0.47	0.36	3.52E-31	0.03
rs142963458	16	84561361	T	C	0.04	-0.78	2.18E-20	0.08
rs8059590	16	84572816	G	C	0.38	-0.25	2.08E-14	0.03
rs4794029	17	47280301	C	T	0.68	0.24	4.22E-13	0.03
rs7405453	17	79615572	G	A	0.65	-0.41	1.84E-37	0.03
rs1232603	20	10612963	T	C	0.33	-0.23	2.25E-12	0.03
rs6077977	20	10930708	G	A	0.49	0.28	3.03E-20	0.03

Table A.9: Full summary statistics for GWAS of RPE thickness. Full GWAS summary statistics for seven variants associated with RPE thickness, including effect size, p-value and standard error (SE). A1 is the effect allele. Allele frequency (AF) refers to the frequency of the effect allele.

SNP	Chr	BP	A1	A2	AF	Effect size	P-value	SE
rs61822227	1	196767705	C	A	0.26	0.45	2.85E-09	0.08
rs1126809	11	89017961	A	G	0.30	0.45	4.47E-11	0.07
rs77001109	13	27469579	T	C	0.01	1.66	7.74E-09	0.29
rs1800407	15	28230318	T	C	0.08	-0.86	4.78E-14	0.11
rs7239443	18	2460759	T	C	0.01	1.78	1.14E-08	0.31
rs56194068	18	71061618	T	G	0.11	0.57	2.09E-08	0.10
rs77561053	20	16403226	C	T	0.01	1.69	1.19E-09	0.28

B. Supplementary Figures

B.1. Genome-wide Association Study of Outer Retinal Morphology

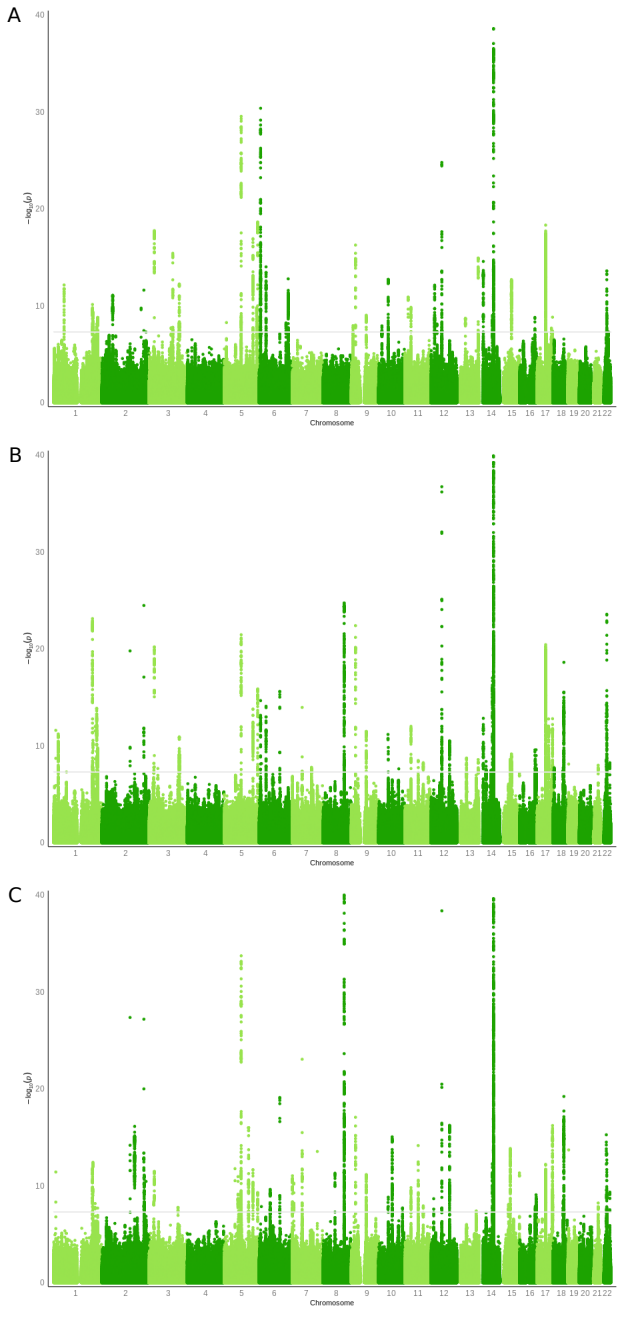

Figure B.1: Genome-wide association study of ONL thickness in different fields of the ETDRS grid. Manhattan plots of ONL thickness averaged across the A)Fovea, B)Intermediate ring, and C)Peripheral ring. Variants are considered significantly associated if they reach consensus genome wide significance ($P < 5 \times 10^{-8}$).

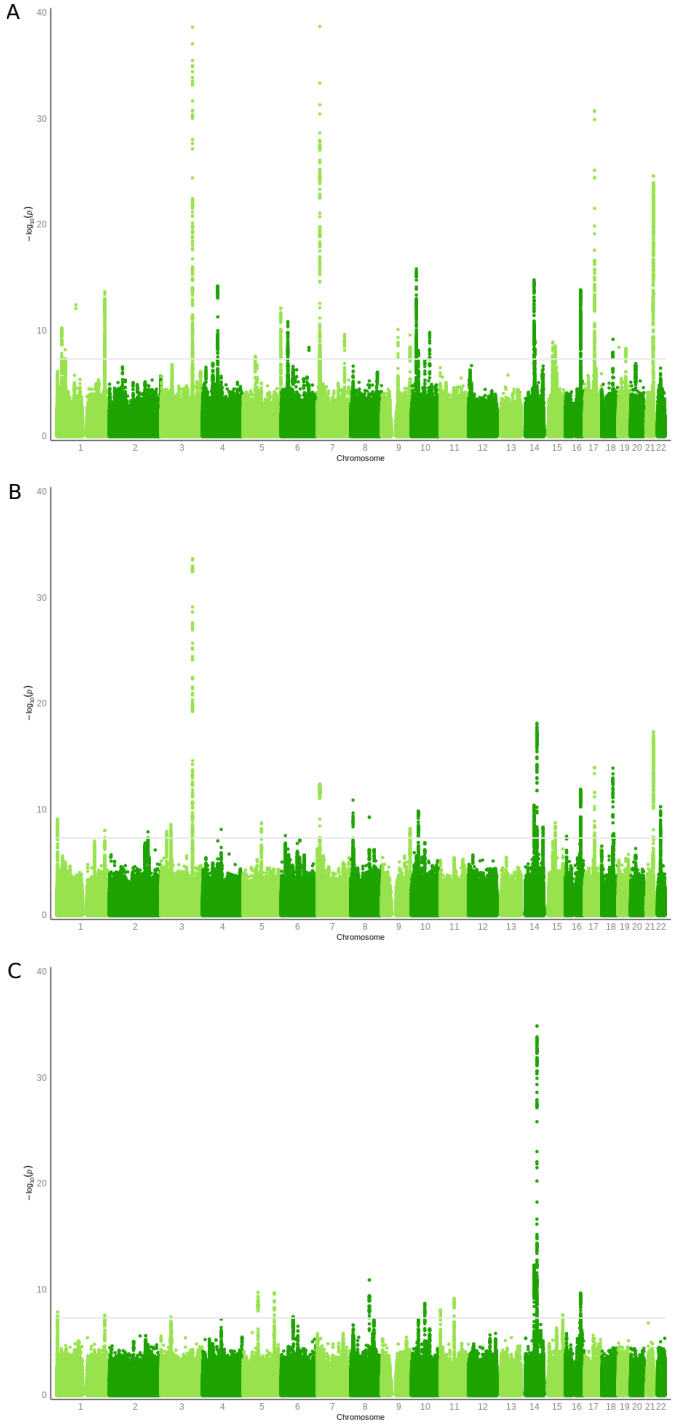

Figure B.2: Genome-wide association study of IS thickness in different fields of the ETDRS grid. Manhattan plots of IS thickness averaged across the A)Fovea, B)Intermediate ring, and C)Peripheral ring. Variants are considered significantly associated if they reach consensus genome wide significance ($P < 5 \times 10^{-8}$).

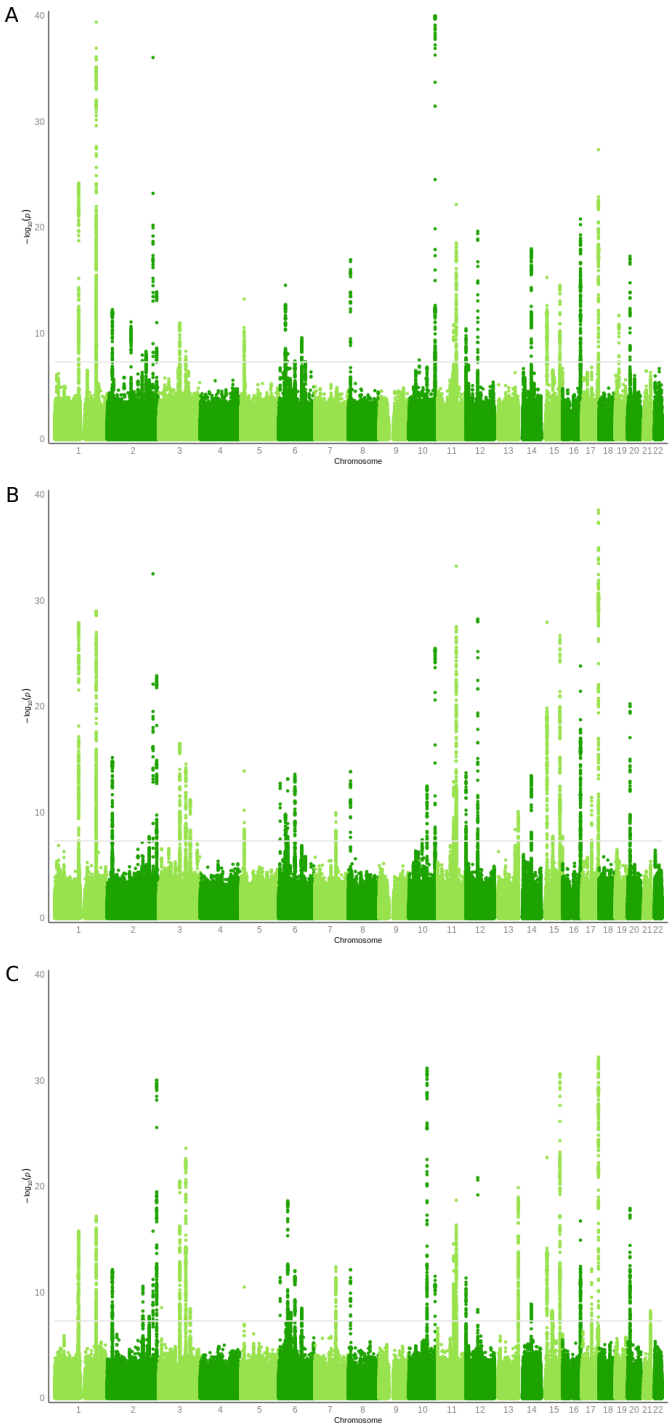

Figure B.3: Genome-wide association study of OS thickness in different fields of the ETDRS grid. Manhattan plots of OS thickness averaged across the A)Fovea, B)Intermediate ring, and C)Peripheral ring. Variants are considered significantly associated if they reach consensus genome wide significance ($P < 5 \times 10^{-8}$).

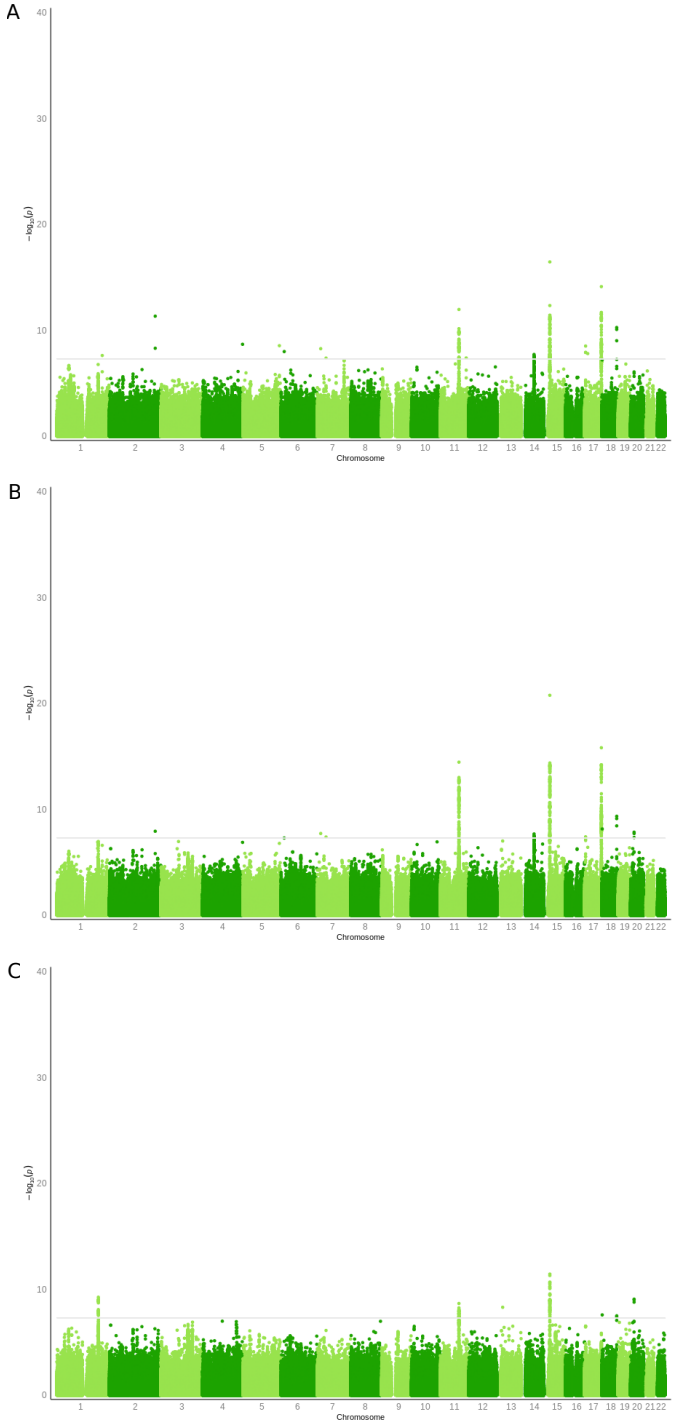

Figure B.4: Genome-wide association study of RPE thickness in different fields of the ETDRS grid. Manhattan plots of RPE thickness averages across the A)Fovea, B)Intermediate ring, C)Peripheral ring. Variants are considered significantly associated if they reach consensus genome wide significance ($P < 5 \times 10^{-8}$).

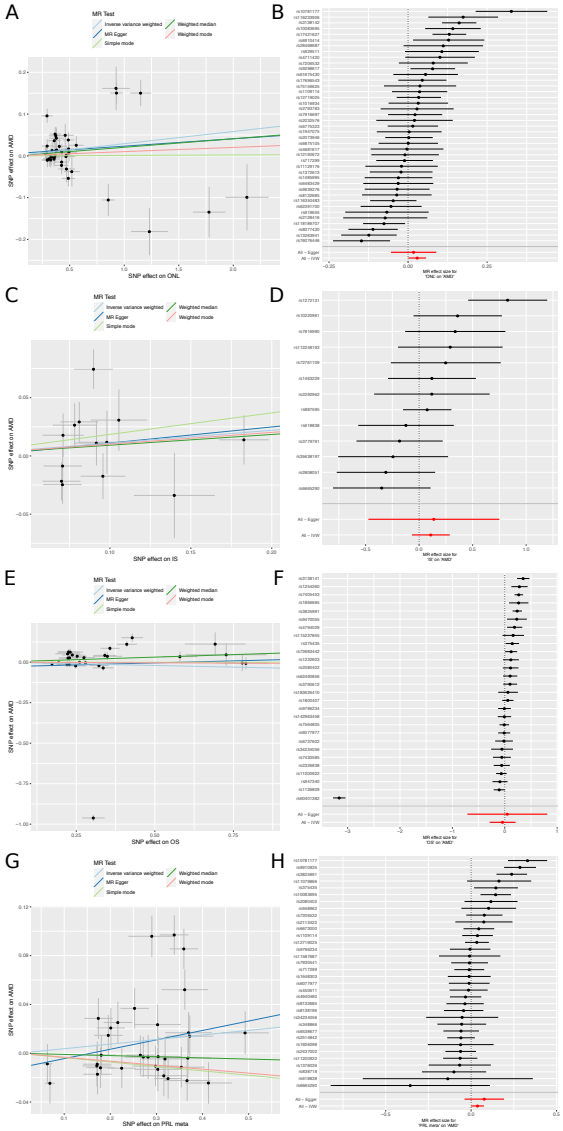

Figure B.5: Mendelian randomisation analysis of the relationship between photoreceptor layer thickness and AMD. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with ONL thickness, and the effect size of the same SNPs on AMD. B) Forest plot showing the effect size and direction of SNPs associated with ONL thickness on AMD. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with IS thickness, and the effect size of the same SNPs on AMD. D) Forest plot showing the effect size and direction of SNPs associated with IS thickness on AMD. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with OS thickness, and the effect size of the same SNPs on AMD. F) Forest plot showing the effect size and direction of SNPs associated with OS thickness on AMD. G) Scatter plot of the relationship between the effect size of SNPs significantly associated with meta-analysed PRL component layer thickness, and the effect size of the same SNPs on AMD. H) Forest plot showing the effect size and direction of SNPs associated with meta-analysed PRL component layer thickness on AMD.

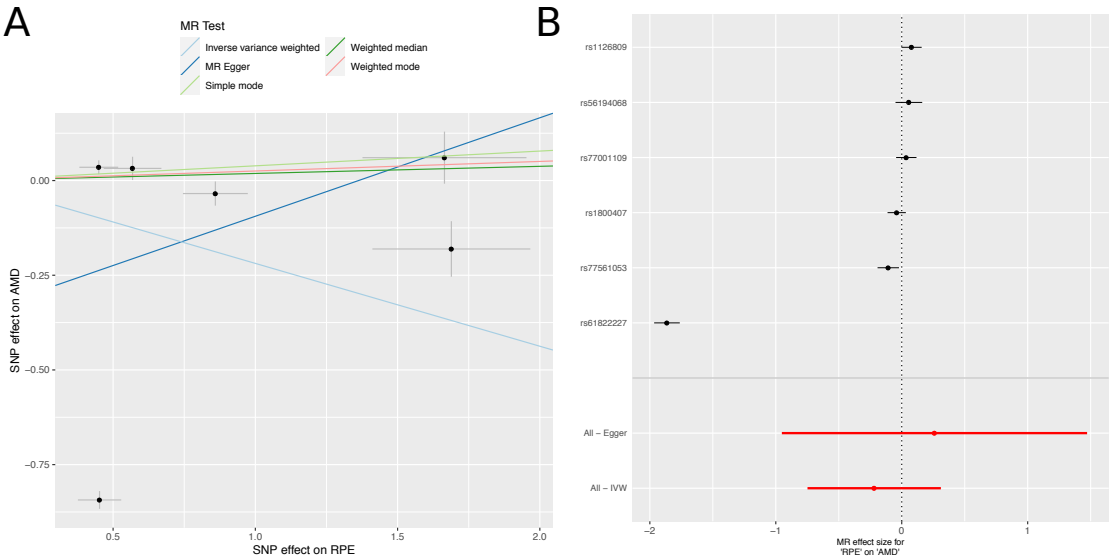

Figure B.6: Mendelian randomisation analysis of the relationship between retinal pigment epithelium layer thickness and AMD. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with RPE thickness, and the effect size of the same SNPs on AMD. B) Forest plot showing the effect size and direction of SNPs associated with RPE thickness on AMD.

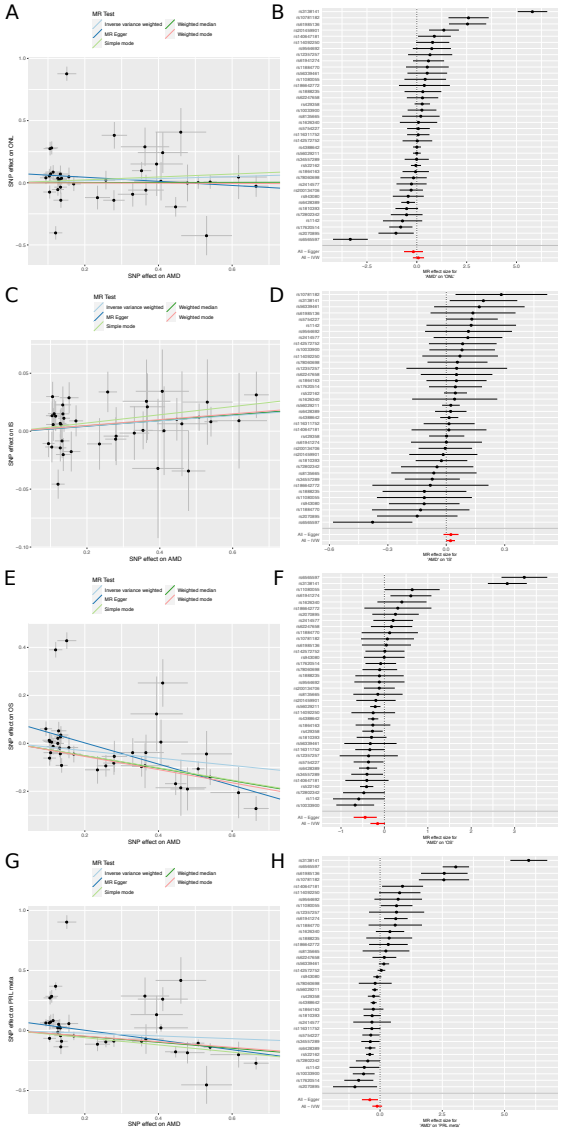

Figure B.7: Mendelian randomisation analysis of the relationship between AMD and photoreceptor layer thickness. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on ONL thickness. B) Forest plot showing the effect size and direction of SNPs associated with AMD on ONL thickness. C) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on IS thickness. D) Forest plot showing the effect size and direction of SNPs associated with AMD on IS thickness. E) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on OS thickness. F) Forest plot showing the effect size and direction of SNPs associated with AMD on OS thickness. G) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on meta-analysed PRL component layer thickness. H) Forest plot showing the effect size and direction of SNPs associated with AMD on meta-analysed PRL component layer thickness.

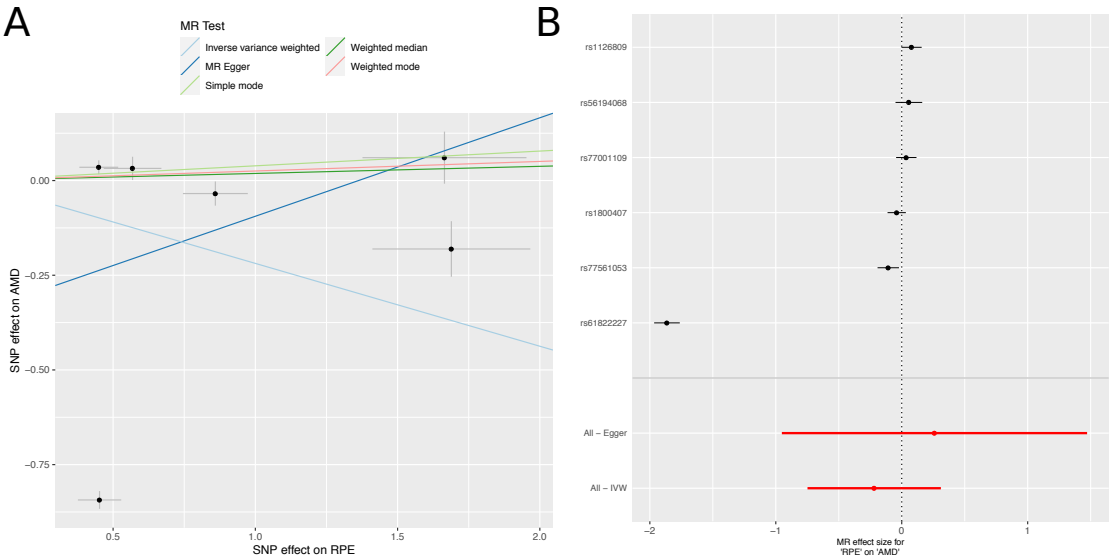

Figure B.8: Mendelian randomisation analysis of the relationship between AMD and retinal pigment epithelium layer thickness. A) Scatter plot of the relationship between the effect size of SNPs significantly associated with AMD, and the effect size of the same SNPs on RPE thickness. B) Forest plot showing the effect size and direction of SNPs associated with AMD thickness on RPE thickness.

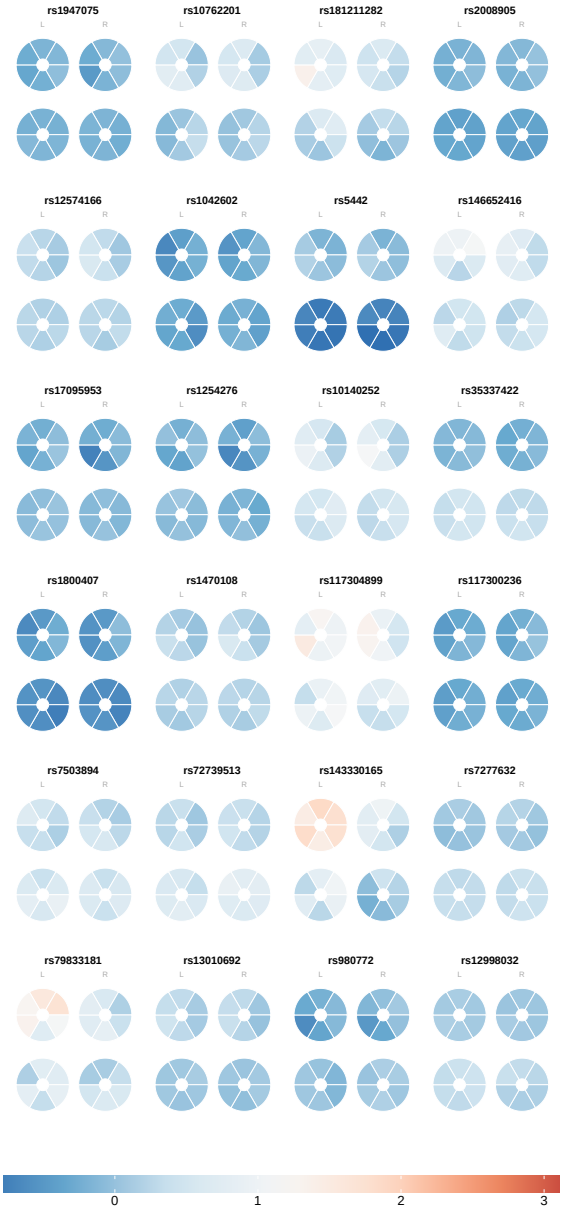

Figure B.9: Full results of spatial variation analysis of genetic effect on the inner retina across the Macula 6 grid, illustrated as maculagrams The thickness of the retinal nerve fibre layer (RNFL) and ganglion cell inner plexiform layer (GCIPL) at each of the segments of the Macula 6 grid across both left and right eyes was modelled as the outcome of each of the genetic variants found significantly associated with mean inner retinal thickness. The effect sizes from each of these models is plotted back into the Macula 6 grid using maculagrams. For each plot, the RNFL is pictured top, and the GCIPL is pictured below. The grids are oriented so that for both the left and right eyes, the easterly segments represent the temporal segments of the eye, those closest tot he ear, and the westerly represent the nasal sides, those closest to the nose.

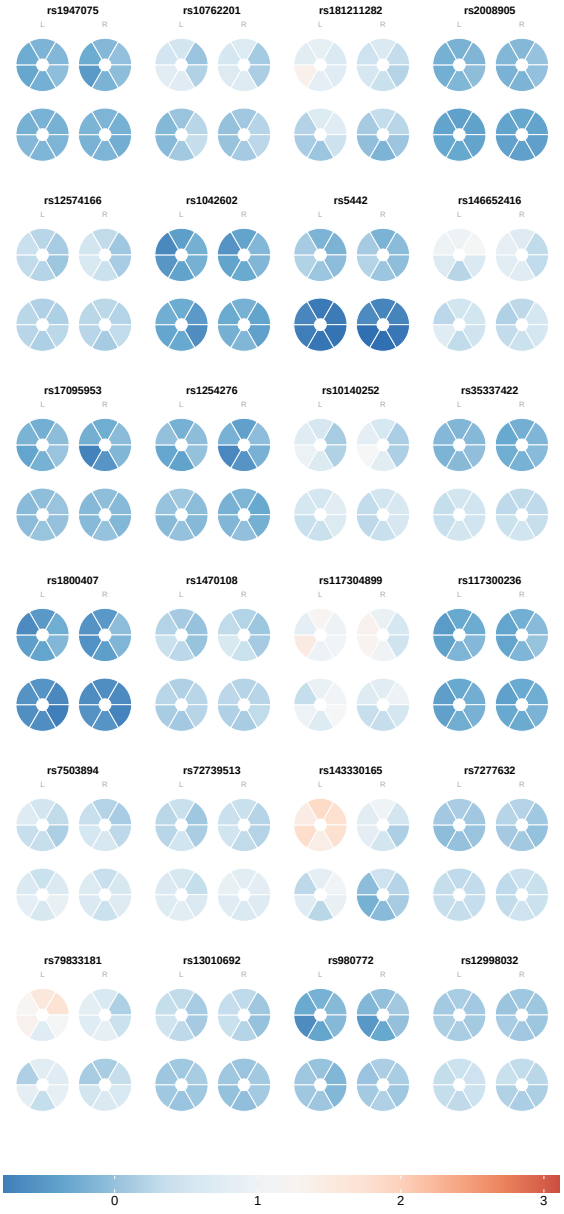

Figure B.10: figure B.9 **continued**

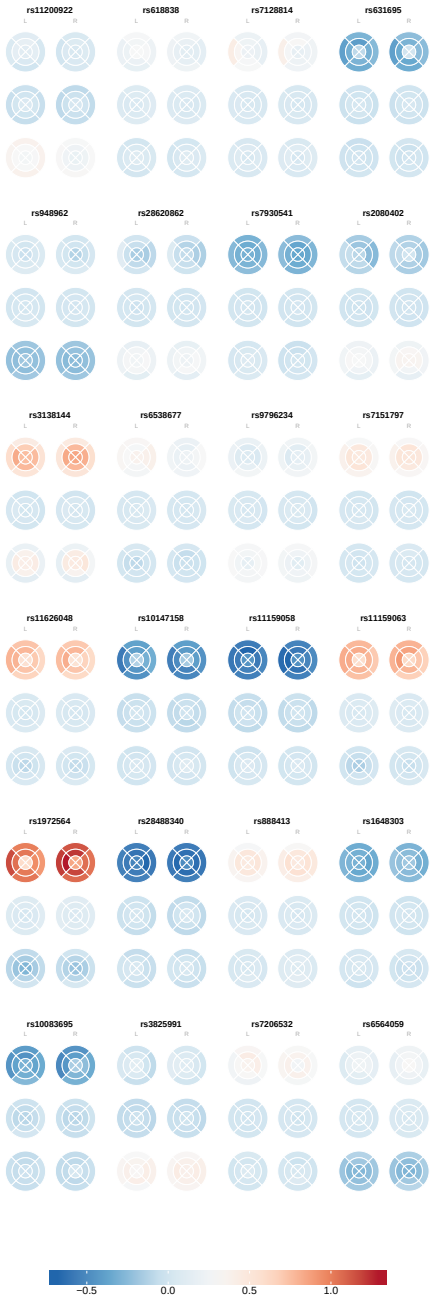

Figure B.11: Full results of spatial variation analysis of genetic effect on the photoreceptor cells across the ETDRS grid, illustrated as maculagrams The thickness of the outer nuclear layer (ONL), Inner segment (IS) and outer segment (OS) at each of the segments of the ETDRS grid across both left and right eyes was modelled as the outcome of each of the genetic variants found significantly associated with mean inner retinal thickness. The effect sizes from each of these models is plotted back into the ETDRS grid using maculagrams. For each plot, the ONL is pictured top, the IS in the middle and the OS is pictured below. The grids are oriented so that for both the left and right eyes, the easterly segments represent the temporal segments of the eye, those closest tot he ear, and the westerly represent the nasal sides, those closest to the nose.

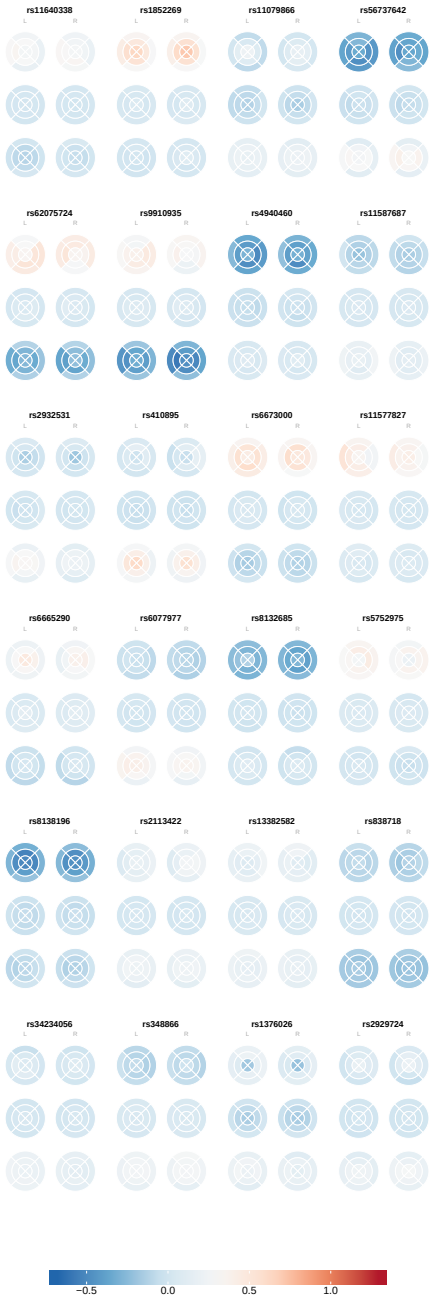

Figure B.12: figure B.9 **continued**

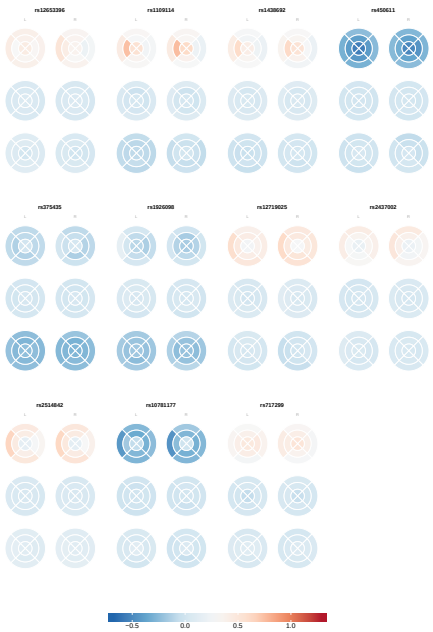

Figure B.13: figure B.9 **continued**

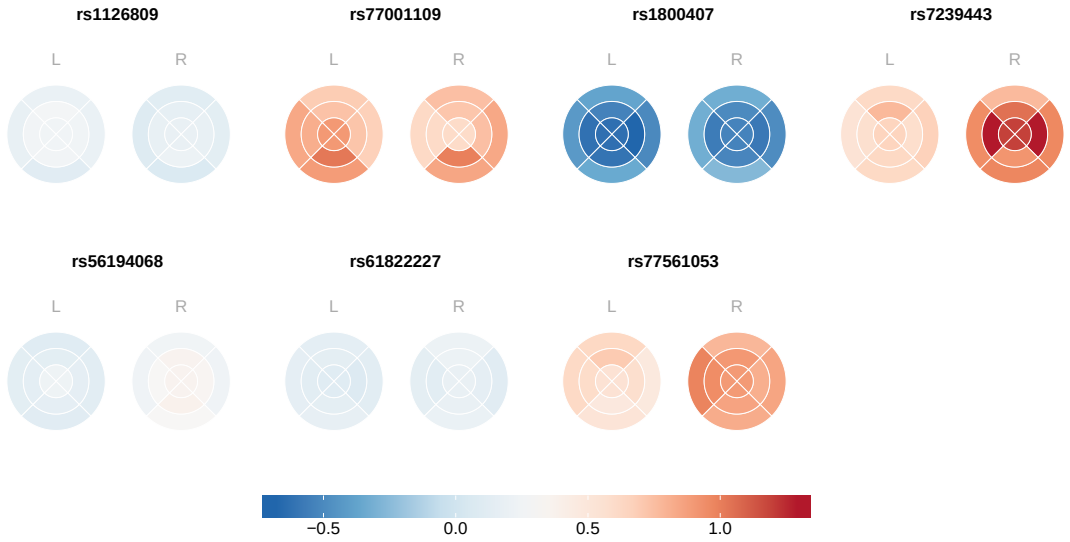

Figure B.14: Full results of spatial variation analysis of genetic effect on the retinal pigment epithelium across the ETDRS grid, illustrated as maculagrams The thickness of the retinal pigment epithelium (RPE) at each of the segments of the ETDRS grid across both left and right eyes was modelled as the outcome of each of the genetic variants found significantly associated with mean inner retinal thickness. The effect sizes from each of these models is plotted back into the ETDRS grid using maculagrams. The grids are oriented so that for both the left and right eyes, the easterly segments represent the temporal segments of the eye, those closest tot he ear, and the westerly represent the nasal sides, those closest to the nose.

References

Abbasi, S. M. (2017) A Study of the Evil Eye Phenomenon and How it is Translated into Modern Fashion, Textiles, and Accessories. *Indian J.Sci.Res* 13(1): 137–143 (cit. on p. 24).

Abdellah, Z., Ahmadi, A., Ahmed, S., Aimable, M., Ainscough, R., Almeida, J., Almond, C., Ambler, A., Ambrose, K., Ambrose, K., Andrew, R., Andrews, D., Andrews, N., Andrews, D., Apweiler, E., Arbery, H., Archer, B., Ash, G., Ashcroft, K., Ashurst, J. & al. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431(7011): 931–945 (cit. on p. 49).

Abramoff, M. D., Garvin, M. K. & Sonka, M. (2010) Retinal imaging and image analysis. *IEEE Reviews in Biomedical Engineering* 3: 169–208 (cit. on p. 31).

Abul-Husn, N. S. & Kenny, E. E. (2019) Personalized Medicine and the Power of Electronic Health Records. *Cell* 177(1): 58–69 (cit. on p. 59).

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X. & al. (2000) The genome sequence of Drosophila melanogaster. *Science* 287(5461): 2185–2195 (cit. on p. 49).

Agarwal, D. R., Ehrlich, J. R., Shimmyo, M. & Radcliffe, N. M. (2012) The relationship between corneal hysteresis and the magnitude of intraocular pressure reduction with topical prostaglandin therapy. *British Journal of Ophthalmology* 96(2): 254–257 (cit. on p. 122).

Aida, T. (1921) On the Inheritance of Color in a Fresh-Water Fish, APLOCHEILUS LATIPES Temmick and Schlegel, with Special Reference to Sex-Linked Inheritance. *Genetics* 6(6): 554–73 (cit. on p. 48).

Ajita, R. (2015) Galen and His Contribution To Anatomy: a Review. *Journal of Evolution of Medical and Dental Sciences* 4(26): 4509–4516 (cit. on p. 25).

Allen, G. E. (1968) Thomas Hunt Morgan and the Problem of Natural Selection. *Journal of the History of Biology* 1(1): 113–139 (cit. on p. 46).

Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., Willer, C. J., Jackson, A. U., Vedantam, S., Raychaudhuri, S., Ferreira, T., Wood, A. R., Weyant, R. J., Segrè, A. V., Speliotes, E. K., Wheeler, E., Soranzo, N., Park, J. H., Yang, J., Gudbjartsson, D. & al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467(7317): 832–838 (cit. on p. 52).

Anderson, S., Bankier, A. T., Barrell, B. G., De Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R. & Young, I. G. (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290(5806): 457–465 (cit. on p. 49).

Andrés-Guerrero, V., García-Feijoo, J. & Konstas, A. G. (2017) Targeting Schlemm's Canal in the Medical Therapy of Glaucoma: Current and Future Considerations. *Advances in Therapy* 34(5): 1049–1069 (cit. on p. 35).

Angrist, J. D. (1990) Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records. *American Economic Review* 80(3): 313–336 (cit. on p. 54).

Astle, W. J., Elding, H., Jiang, T. & Ouwehand, W. H. (2016) The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* 167: 1415–1429 (cit. on p. 60).

Audo, I., Mohand-Said, S., Boulanger-Scemama, E., Zanlonghi, X., Condroyer, C., Démontant, V., Boyard, F., Antonio, A., Méjécase, C., El Shamieh, S., Sahel, J.-A. & Zeitz, C. (2018) MERTK mutation update in inherited retinal diseases. *Human mutation* 39(7): 887–913 (cit. on p. 128).

Avery, O. T., Macleod, C. M. & McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type iii. *Journal of Experimental Medicine* 79(2): 137–158 (cit. on p. 47).

Awata, T., Yamashita, H., Kurihara, S., Morita-Ohkubo, T., Miyashita, Y., Katayama, S., Mori, K., Yoneya, S., Kohda, M., Okazaki, Y., Maruyama, T., Shimada, A., Yasuda, K., Nishida, N., Tokunaga, K. & Koike, A. (2014) A Genome-Wide Association Study for Diabetic Retinopathy in a Japanese Population: Potential Association with a Long Intergenic Non-Coding RNA. *PLoS ONE* 9(11): e111715 (cit. on p. 52).

Bailey, J. N., Loomis, S. J., Kang, J. H., Allingham, R. R., Gharahkhani, P., Khor, C. C., Burdon, K. P., Aschard, H., Chasman, D. I., Igo, R. P., Hysi, P. G., Glastonbury, C. A., Ashley-Koch, A., Brilliant, M., Brown, A. A., Budenz, D. L., Buil, A., Cheng, C. Y.,

Choi, H., Christen, W. G. & al. (2016) Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. *Nature Genetics* 48(2): 189–194 (cit. on p. 35).

Bateson, W. (1902) *Mendel's principles of heredity: a defence, with a translation of Mendel's original papers on hybridisation.* (Cit. on p. 43).

Bateson, W., Saunders, E. R. & Punnett, R. C. (1905) Experimental studies in the physiology of heredity. *Reports to the Evolution Committee of the Royal Society* (cit. on p. 44).

Batty, G. D., Gale, C. R., Kivimäki, M., Deary, I. J. & Bell, S. (2020) Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 368 (cit. on p. 61).

Beidoe, G. & Mousa, S. A. (2012) Current primary open-angle glaucoma treatments and future directions. *Clinical Ophthalmology* 6(1): 1699–1707 (cit. on pp. 35, 93).

Bell, J. & Haldane, J. B. S. (1937) The linkage between the genes for colour-blindness and haemophilia in man. *Proceedings of the Royal Society of London. Series B - Biological Sciences* 123(831): 119–150 (cit. on p. 48).

Beltran, W. A., Cideciyan, A. V., Guziewicz, K. E., Iwabe, S., Swider, M., Scott, E. M., Savina, S. V., Ruthel, G., Stefano, F., Zhang, L., Zorger, R., Sumaroka, A., Jacobson, S. G. & Aguirre, G. D. (2014) Canine retina has a primate fovea-like bouquet of cone photoreceptors which is affected by inherited macular degenerations. *PLoS ONE* 9(3): 90390 (cit. on p. 31).

Bernardes, R., Serranho, P. & Lobo, C. (2011) Digital Ocular Fundus Imaging: A Review. *Ophthalmologica* 226(4): 161–181 (cit. on p. 31).

Bhagat, P. R., Deshpande, K. V. & Natu, B. (2014) Utility of ganglion cell complex analysis in early diagnosis and monitoring of glaucoma using a different spectral domain optical coherence tomography. *Journal of Current Glaucoma Practice* 8(3): 101–106 (cit. on p. 71).

Bhatia, S., Monahan, J., Ravi, V., Gautier, P., Murdoch, E., Brenner, S., Van Heyningen, V., Venkatesh, B. & Kleinjan, D. A. (2014) A survey of ancient conserved non-coding elements in the PAX6 locus reveals a landscape of interdigitated cis-regulatory archipelagos. *Developmental Biology* 387(2): 214–228 (cit. on p. 31).

Bhattacharya, S. S., Wright, A. F., Clayton, J. F., Price, W. H., Phillips, C. I., McKeown, C. M., Jay, M., Bird, A. C., Pearson, P. L., Southern, E. M. & Evans, H. J. (1984) Close ge-

netic linkage between X-linked retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. *Nature* 309(5965): 253–255 (cit. on p. 48).

Birney, E., Vamathevan, J. & Goodhand, P. (2017) Genomics in healthcare: GA4GH looks to 2022. *bioRxiv*: 203554 (cit. on p. 39).

Black, G. C., MacEwen, C. & Lotery, A. J. (2020) The integration of genomics into clinical ophthalmic services in the UK. *Eye* 34(6): 993–996 (cit. on p. 23).

Blacker, C. P. (1931) The sterilization proposals: A history of their development. *The Eugenics review* 22(4): 239–47 (cit. on p. 45).

Bodaghi, B., Massamba, N. & Izzedine, H. (2014) The eye: A window on kidney diseases. *Clinical Kidney Journal* 7(4): 337–338 (cit. on p. 91).

Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32(3): 314–331 (cit. on p. 48).

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S. & Davey Smith, G. (2013) Cohort Profile: the 'children of the 90s'–the index offspring of the Avon Longitudinal Study of Parents and Children. eng. *International journal of epidemiology* 42(1): 111–127 (cit. on p. 59).

Brenner, S., Jacob, F. & Meselson, M. (1961) An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190(4776): 576–581 (cit. on p. 47).

Bringmann, A., Syrbe, S., Görner, K., Kacza, J., Francke, M., Wiedemann, P. & Reichenbach, A. (2018) The primate fovea: Structure, function and development. *Progress in Retinal and Eye Research* 66: 49–84 (cit. on p. 31).

Bronstein, J. L. & Bolnick, D. I. (2018) "Her Joyous Enthusiasm for Her Life-Work ...": Early Women Authors in The American Naturalist. *The American Naturalist* 192(6): 655–663 (cit. on p. 43).

Brown, N. L., Patel, S., Brzezinski, J. & Glaser, T. (2001) Math5 is required for retinal ganglion cell and optic nerve formation. *Development* 128(13): 2497–2508 (cit. on p. 90).

Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., Daly, M. J., Price, A. L., Neale, B. M., Corvin, A., Walters, J. T., Farh, K. H., Holmans, P. A., Lee, P., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M. & al. (2015) LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* 47(3): 291–295 (cit. on pp. 53, 85, 107).

Burdon, K. P., Fogarty, R. D., Shen, W., Abhary, S., Kaidonis, G., Appukuttan, B., Hewitt, A. W., Sharma, S., Daniell, M., Essex, R. W., Chang, J. H., Klebe, S., Lake, S. R., Pal, B., Jenkins, A., Govindarjan, G., Sundaresan, P., Lamoureux, E. L., Ramasamy, K., Pefkianaki, M. & al. (2015) Genome-wide association study for sight-threatening diabetic retinopathy reveals association with genetic variation near the GRB2 gene. *Diabetologia* 58(10): 2288–2297 (cit. on p. 52).

Burdon, K. P., MacGregor, S., Hewitt, A. W., Sharma, S., Chidlow, G., Mills, R. A., Danoy, P., Casson, R., Viswanathan, A. C., Liu, J. Z., Landers, J., Henders, A. K., Wood, J., Souzeau, E., Crawford, A., Leo, P., Wang, J. J., Rochtchina, E., Nyholt, D. R., Martin, N. G. & al. (2011) Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. *Nature Genetics* 43(6): 574–578 (cit. on pp. 35, 89).

Burgess, S. & Harshfield, E. (2016) Mendelian randomization to assess causal effects of blood lipids on coronary heart disease: Lessons from the past and applications to the future. *Current Opinion in Endocrinology, Diabetes and Obesity* 23(2): 124–130 (cit. on p. 55).

Burian, R. & Zallen, D. (2009) Genes. *The Cambridge History of Science*. Ed. by P. J. Bowler & J. V. Pickstone. Vol. 6. The Cambridge History of Science. Cambridge University Press: 432–450 (cit. on pp. 39, 40).

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., Todd, J. A., Donnelly, P., Barrett, J. C., Davison, D., Easton, D., Evans, D., Leung, H. T., Marchini, J. L., Morris, A. P., Spencer, C. C. & al. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447(7145): 661–678 (cit. on p. 51).

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P. & Marchini, J. (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562(7726): 203–209 (cit. on pp. 67, 68, 78, 80, 85, 107, 120).

Campbell, P., Ellingford, J. M., Parry, N. R., Fletcher, T., Ramsden, S. C., Gale, T., Hall, G., Smith, K., Kasperaviciute, D., Thomas, E., Lloyd, I. C., Douzgou, S., Clayton-Smith, J., Biswas, S., Ashworth, J. L., Black, G. C. & Sergouniotis, P. I. (2019) Clinical and genetic variability in children with partial albinism. *Scientific Reports* 9(1): 1–10 (cit. on pp. 90, 115, 153, 159).

Carey, D. P. & Hutchinson, C. V. (2013) Looking at eye dominance from a different angle: is sighting strength related to hand preference? *Cortex* 49(9): 2542–2552 (cit. on p. 76).

Carvalho-Silva, D., Pierleoni, A., Pignatelli, M., Ong, C., Fumis, L., Karamanis, N., Carmona, M., Faulconbridge, A., Hercules, A., McAuley, E., Miranda, A., Peat, G., Spitzer, M., Barrett, J., Hulcoop, D. G., Papa, E., Koscielny, G. & Dunham, I. (2019) Open Targets Platform: new developments and updates two years on. *Nucleic Acids Research* 47(D1): D1056–D1065 (cit. on pp. 85, 107, 120).

Chaglasian, M., Fingeret, M., Davey, P. G., Huang, W. C., Leung, D., Ng, E. & Reisman, C. A. (2018) The development of a reference database with the Topcon 3D OCT-1 Maestro. *Clinical Ophthalmology* 12: 849–857 (cit. on pp. 63, 73).

Chan, M. P., Broadway, D. C., Khawaja, A. P., Yip, J. L., Garway-Heath, D. F., Burr, J. M., Luben, R., Hayat, S., Dalzell, N., Khaw, K. T. & Foster, P. J. (2017) Glaucoma and intraocular pressure in EPIC-Norfolk Eye Study: cross sectional study. *BMJ* (*Clinical research ed.*) 358: j3889 (cit. on p. 34).

Chan, V. T., Sun, Z., Tang, S., Chen, L. J., Wong, A., Tham, C. C., Wong, T. Y., Chen, C., Ikram, M. K., Whitson, H. E., Lad, E. M., Mok, V. C. & Cheung, C. Y. (2019) Spectral-Domain OCT Measurements in Alzheimer's Disease: A Systematic Review and Meta-analysis. *Ophthalmology* 126(4): 497–510 (cit. on p. 37).

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M. & Lee, J. J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4(1) (cit. on p. 67).

Chargaff, E. (1950) Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 6(6): 201–209 (cit. on p. 47).

Cheng, C. Y., Rand Allingham, R., Aung, T., Tham, Y. C., Hauser, M. A., Vithana, E. N., Khor, C. C. & Wong, T. Y. (2014) Association of common SIX6 polymorphisms with peripapillary retinal nerve fiber layer thickness: The Singapore chinese eye study. *Investigative Ophthalmology and Visual Science* 56(1): 478–483 (cit. on p. 35).

Cheng, L., Yu, H., Jiang, Y., He, J., Pu, S., Li, X. & Zhang, L. (2018) Identification of a novel MYO7A mutation in Usher syndrome type 1. *Oncotarget* 9(2): 2295–2303 (cit. on p. 110).

Choquet, H., Paylakhi, S., Kneeland, S. C., Thai, K. K., Hoffmann, T. J., Yin, J., Kvale, M. N., Banda, Y., Tolman, N. G., Williams, P. A., Schaefer, C., Melles, R. B., Risch, N., John, S. W., Nair, K. S. & Jorgenson, E. (2018) A multiethnic genome-wide association

study of primary open-angle glaucoma identifies novel risk loci. *Nature Communications* 9(1): 1–14 (cit. on pp. 52, 74).

Chow, R. L., Volgyi, B., Szilard, R. K., Ng, D., McKerlie, C., Bloomfield, S. A., Birch, D. G. & McInnes, R. R. (2004) Control of late off-center cone bipolar cell differentiation and visual signaling by the homeobox gene Vsx1. *Proceedings of the National Academy of Sciences of the United States of America* 101(6): 1754–1759 (cit. on p. 115).

Collins, G. N. (1912) Gametic Coupling as a Cause of Correlations. *The American Naturalist* 46(550): 569–590 (cit. on p. 46).

Correns, C. E. (1900) G. Mendel's Regel über das Verhalten der Nachkommenschaft der Rassenbastarde. *Berichte der Deutschen Botanischen Gesellschaft* 18(4): 158–168 (cit. on p. 43).

Craig Venter, J., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M. & al. (2001) The sequence of the human genome. *Science* 291(5507): 1304–1351 (cit. on p. 49).

Creighton, H. B. & McClintock, B. (1931) A Correlation of Cytological and Genetical Crossing-Over in Zea Mays. *Proceedings of the National Academy of Sciences* 17(8): 492–497 (cit. on p. 47).

Crick, F. H. (1958) On protein synthesis. *Symposia of the Society for Experimental Biology* 12: 138–163 (cit. on p. 47).

Crick, F. H., Barnett, L., Brenner, S. & Watts-Tobin, R. J. (1961) General nature of the genetic code for proteins. *Nature* 192: 1227–1232 (cit. on p. 47).

Cumberland, P. M. & Rahi, J. S. (2016) Visual function, social position, and health and life chances the UK Biobank study. *JAMA Ophthalmology* 134(9): 959–966 (cit. on p. 62).

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K. J., Gall, A., Girón, C. G., Gil, L., Grego, T., Haggerty, L., Haskell, E., Hourlier, T. & al. (2019) Ensembl 2019. *Nucleic Acids Research* 47(D1): D745–D751 (cit. on pp. 85, 107, 120).

Cunningham, J. (2017) Recognizing age-related macular degeneration in primary care. *Journal of the American Academy of Physician Assistants* 30(3): 18–22 (cit. on pp. 35, 36).

Currant, H., Hysi, P., Fitzgerald, T. W., Gharahkhani, P., M Bonnemaijer, P. W., Biobank, U., Craig, J., Hewitt, A. W., Khaw, P. T., W Klaver, C. C., Kubo, M., Ong, J.-S., Pasquale,

L. R., Reisman, C. A., Simcoe, M. J., H J Thiadens, A. A., van Duijn, C. M., Yazar, S., Jorgenson, E., MacGregor, S. & al. (2020) Genetic variation affects morphological retinal phenotypes extracted from UK Biobank Optical Coherence Tomography images. *bioRxiv* (cit. on pp. 56, 57).

Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life.* (Cit. on pp. 26, 40).

Darwin, C. (1868) *The variation of animals and plants under domestication.* London: J. Murray (cit. on pp. 40, 41).

Darwin, C. & Wallace, A. (1858) On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *Journal of the Proceedings of the Linnean Society of London. Zoology* 3(9): 45–62 (cit. on p. 40).

De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., Visentin, D., van den Driessche, G., Lakshminarayanan, B., Meyer, C., Mackinder, F., Bouton, S., Ayoub, K., Chopra, R., King, D., Karthikesalingam, A., Hughes, C. O. & al. (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine* 24(9): 1342–1350 (cit. on p. 168).

De Vries, H. (1889) *Intracellular pangenesis.* Jena (cit. on p. 43).

Deary, I. J., Gow, A. J., Pattie, A. & Starr, J. M. (2011) Cohort Profile: The Lothian Birth Cohorts of 1921 and 1936. *International Journal of Epidemiology* 41(6): 1576–1584 (cit. on p. 59).

De Faria, C. C. & Fortunato, R. S. (2020) The role of dual oxidases in physiology and cancer. *Genetics and Molecular Biology* 43(1): 1–9 (cit. on p. 111).

Demirkaya, N., van Dijk, H. W., van Schuppen, S. M., Abràmoff, M. D., Garvin, M. K., Sonka, M., Schlingemann, R. O. & Verbraak, F. D. (2013) Effect of age on individual retinal layer thickness in normal eyes as measured with spectral-domain optical coherence tomography. *Investigative Ophthalmology and Visual Science* 54(7): 4934–4940 (cit. on p. 74).

Deva, R., Alias, M. A., Colville, D., Tow, F. K. N. F. H., Ooi, Q. L., Chew, S., Mohamad, N., Hutchinson, A., Koukouras, I., Power, D. A. & Savige, J. (2011) Vision-threatening retinal abnormalities in chronic kidney disease stages 3 to 5. *Clinical Journal of the American Society of Nephrology* 6(8): 1866–1871 (cit. on p. 91).

Devlin, B. & Roeder, K. (1999) Genomic control for association studies. *Biometrics* 55(4): 997–1004 (cit. on p. 53).

DeWan, A., Liu, M., Hartman, S., Zhang, S. S. M., Liu, D. T., Zhao, C., Tam, P. O., Chan, W. M., Lam, D. S., Snyder, M., Barnstable, C., Pang, C. P. & Hoh, J. (2006) HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science* 314(5801): 989–992 (cit. on p. 127).

Dias, M. F., Joo, K., Kemp, J. A., Fialho, S. L., da Silva Cunha, A., Woo, S. J. & Kwon, Y. J. (2018) Molecular genetics and emerging therapies for retinitis pigmentosa: Basic research and clinical perspectives. *Progress in Retinal and Eye Research* 63: 107–131 (cit. on p. 36).

Duman, R., Duman, R., Sabaner, M. C. & Çetinkaya, E. (2017) Effect of smoking on the thickness of retinal layers in healthy smokers. *Cutaneous and Ocular Toxicology* 36(4): 366–369 (cit. on p. 37).

Dutt, K. C. (1938) Cataract Opertaions in the Prehistoric Age. *Archives of Ophthalmology* 20(1): 1–15 (cit. on p. 24).

East, E. M. (1910) A Mendelian Interpretation of Variation that is Apparently Continuous. *The American Naturalist* 44(518): 65–82 (cit. on p. 44).

Elliott, L. T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K. L., Douaud, G., Marchini, J. & Smith, S. M. (2018) Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* 562(7726): 210–216 (cit. on pp. 133, 168).

Elliott, M. H., Ashpole, N. E., Gu, X., Herrnberger, L., McClellan, M. E., Griffith, G. L., Reagan, A. M., Boyce, T. M., Tanito, M., Tamm, E. R. & Stamer, W. D. (2016) Caveolin-1 modulates intraocular pressure: Implications for caveolae mechanoprotection in glaucoma. *Scientific Reports* 6(1): 1–12 (cit. on pp. 35, 89).

Farkhondeh, A., Niwa, S., Takei, Y. & Hirokawa, N. (2015) Characterizing KIF16B in neurons reveals a novel intramolecular "stalk inhibition" mechanism that regulates its capacity to potentiate the selective somatodendritic localization of early endosomes. *Journal of Neuroscience* 35(12): 5067–5086 (cit. on p. 115).

Fautsch, M. P., Johnson, D. H., Acott, T. S., Aihara, M., Bhattacharya, S. K., Borrás, T., Camras, C. B., Civan, M. M., Clark, A. F., Crosson, C. E., Crowston, J. G., Epstein, D., Ethier, C. R., Freddo, T. F., Gong, H., Gonzalez, P., John, S. W., Johnson, M., Kaufman, P. L., Knepper, P. A. & al. (2006) 'Aqueous humor outflow: What do we know? Where will it lead us?' *Investigative Ophthalmology and Visual Science*. Vol. 47. 10. The Association for Research in Vision & Ophthalmology: 4181–4187 (cit. on p. 27).

Ference, B. A., Ray, K. K., Catapano, A. L., Ference, T. B., Burgess, S., Neff, D. R., Oliver-Williams, C., Wood, A. M., Butterworth, A. S., Di Angelantonio, E., Danesh,

J., Kastelein, J. J. & Nicholls, S. J. (2019) Mendelian Randomization Study of ACLY and Cardiovascular Disease. *New England Journal of Medicine* 380(11): 1033–1042 (cit. on p. 55).

Ferris, F. L., Wilkinson, C. P., Bird, A., Chakravarthy, U., Chew, E., Csaky, K. & Sadda, S. R. (2013) Clinical classification of age-related macular degeneration. *Ophthalmology* 120(4): 844–851 (cit. on p. 35).

Finger, S. (1994) *Origins of neuroscience : a history of explorations into brain function*. New York: Oxford University Press (cit. on pp. 24–26).

Fisher, R. A. (1925) *Statistical methods for research workers*. Edinburgh ; London: Oliver & Boyd (cit. on p. 44).

Fisher, R. A. (1919) The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52(2): 399–433 (cit. on pp. 44, 45).

Fisher, R. A. (1922) On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society* 85(1): 87 (cit. on p. 44).

Fisher, R. (1930) *The genetical theory of natural selection*. Oxford: Clarendon (cit. on pp. 44, 45).

Fisher, R. A. (1935) *The design of experiments*. Edinburgh: Oliver & Boyd (cit. on p. 44).

Fleckenstein, M., Mitchell, P., Freund, K. B., Sadda, S. V., Holz, F. G., Brittain, C., Henry, E. C. & Ferrara, D. (2018) The Progression of Geographic Atrophy Secondary to Age-Related Macular Degeneration. *Ophthalmology* 125(3): 369–390 (cit. on p. 32).

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M. & al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* 269(5223): 496–512 (cit. on p. 49).

Flemming, W. (1882) *Zellsubstanz, Kern und Zelltheilung*. Nineteenth Century Collections Online (NCCO): Science, Technology, and Medicine: 1780-1925. Vogel (cit. on p. 46).

Franke, A., McGovern, D. P., Barrett, J. C., Wang, K., Radford-Smith, G. L., Ahmad, T., Lees, C. W., Balschun, T., Lee, J., Roberts, R., Anderson, C. A., Bis, J. C., Bumpstead, S., Ellinghaus, D., Festen, E. M., Georges, M., Green, T., Haritunians, T., Jostins, L., Latiano, A. & al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics* 42(12): 1118–1125 (cit. on p. 52).

Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C. & al. (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449(7164): 851–861 (cit. on p. 51).

Freddo, T. F. (2018) *Anatomy of the eye and orbit : the clinical essentials* (cit. on pp. 27, 29, 141, 145).

Fritsche, L. G., Igl, W., Bailey, J. N., Grassmann, F., Sengupta, S., Bragg-Gresham, J. L., Burdon, K. P., Hebbring, S. J., Wen, C., Gorski, M., Kim, I. K., Cho, D., Zack, D., Souied, E., Scholl, H. P., Bala, E., ELee, K., Hunter, D. J., Sardell, R. J., Mitchell, P. & al. (2016) A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nature Genetics* 48(2): 134–143 (cit. on pp. 36, 111, 115, 117–119).

Frost, L. S., Lopes, V. S., Stefano, F. P., Bragin, A., Williams, D. S., Mitchell, C. H. & Boesze-Battaglia, K. (2013) Loss of melanoregulin (MREG) enhances cathepsin-D secretion by the retinal pigment epithelium. *Visual Neuroscience* 30(3): 55–64 (cit. on p. 131).

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R. & Allen, N. E. (2017) Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American journal of epidemiology* 186(9): 1026–1034 (cit. on p. 61).

Fujimoto, J. & Swanson, E. (2016) *The development, commercialization, and impact of optical coherence tomography* (cit. on p. 32).

Gabriele, M. L., Wollstein, G., Ishikawa, H., Xu, J., Kim, J., Kagemann, L., Folio, L. S. & Schuman, J. S. (2010) *Three dimensional optical coherence tomography imaging: Advantages and advances* (cit. on p. 32).

Galton, F. (1869) *Hereditary genius : an inquiry into its laws and consequences*. London: Macmillan & Co. (cit. on p. 42).

Galton, F. (1883) *Inquiries into human faculty and its development*. London: Macmillan (cit. on p. 43).

Galton, F. (1886) Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246 (cit. on p. 42).

Gao, X. R., Huang, H. & Kim, H. (2019) Genome-wide association analyses identify 139 loci associated with macular thickness in the UK Biobank cohort. *Human Molecular Genetics* 28(7): 1162–1172 (cit. on p. 37).

Gao, Z., Mao, C. A., Pan, P., Mu, X. & Klein, W. H. (2014) Transcriptome of Atoh7 retinal progenitor cells identifies new Atoh7-dependent regulatory genes for retinal ganglion cell formation. *Developmental Neurobiology* 74(11): 1123–1140 (cit. on p. 90).

Garrod, A. E. (1923) *Inborn errors of metabolism.* 2. Oxford: Oxford University Press (cit. on p. 44).

Gharahkhani, P., Jorgenson, E., Hysi, P., Khawaja, A. P., Pendergrass, S., Han, X., Ong, J. S., Hewitt, A. W., Segre, A., Igo, R. P., Choquet, H., Qassim, A., Josyula, N. S., Bailey, J. N. C., Bonnemaijer, P., Iglesias, A., Siggs, O. M., Young, T., Vitart, V., Thiadens, A. A. & al. (2020) A large cross-ancestry meta-analysis of genome-wide association studies identifies 69 novel risk loci for primary open-angle glaucoma and includes a genetic link with Alzheimer's disease. *bioRxiv*: 927822 (cit. on pp. 35, 93, 95, 174).

Gill, J. S., Georgiou, M., Kalitzeos, A., Moore, A. T. & Michaelides, M. (2019) Progressive cone and cone-rod dystrophies: Clinical features, molecular genetics and prospects for therapy. *British Journal of Ophthalmology* 103(5): 711–720 (cit. on p. 36).

Glinka, A., Dolde, C., Kirsch, N., Huang, Y.-L., Kazanskaya, O., Ingelfinger, D., Boutros, M., Cruciat, C.-M. & Niehrs, C. (2011) LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. *EMBO reports* 12(10): 1055–1061 (cit. on p. 141).

Goffeau, A., Barrell, G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S. G. (1996) Life with 6000 genes. *Science* 274(5287): 546–567 (cit. on p. 49).

Goh, J. K. H., Cheung, C. Y., Sim, S. S., Tan, P. C., Tan, G. S. W. & Wong, T. Y. (2016) Retinal Imaging Techniques for Diabetic Retinopathy Screening. *Journal of Diabetes Science and Technology* 10(2): 282–294 (cit. on p. 32).

Goicoechea De Jorge, E., Caesar, J. J., Malik, T. H., Patel, M., Colledge, M., Johnson, S., Hakobyan, S., Morgan, B. P., Harris, C. L., Pickering, M. C. & Lea, S. M. (2013) Dimerization of complement factor H-related proteins modulates complement activation in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 110(12): 4685–4690 (cit. on p. 115).

Gothwal, V. K., Sharma, S. & Mandal, A. K. (2020) 'Beyond Intraocular Pressure: Visual Functioning and Quality of Life in Primary Congenital Glaucoma and Secondary Childhood Glaucoma'. *American Journal of Ophthalmology*. Vol. 209. Elsevier Inc.: 62–70 (cit. on p. 34).

Gray, R. & Wheatley, K. (1991) How to avoid bias when comparing bone marrow transplantation with chemotherapy. *Bone marrow transplantation* 7 Suppl 3: 9–12 (cit. on p. 54).

Grønskov, K., Ek, J. & Brondum-Nielsen, K. (2007) Oculocutaneous albinism. *Orphanet Journal of Rare Diseases* 2(1): 43 (cit. on p. 37).

Gupta, D. & Asrani, S. (2016) Macular thickness analysis for glaucoma diagnosis and management. *Taiwan Journal of Ophthalmology* 6(1): 3–7 (cit. on pp. 63, 73, 76, 136).

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E. & Martin, J. B. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306(5940): 234–238 (cit. on pp. 48, 50).

Hageman, G. S., Anderson, D. H., Johnson, L. V., Hancox, L. S., Taiber, A. J., Hardisty, L. I., Hageman, J. L., Stockman, H. A., Borchardt, J. D., Gehrs, K. M., Smith, R. J., Silvestri, G., Russell, S. R., Klaver, C. C., Barbazetto, I., Chang, S., Yannuzzi, L. A., Barile, G. R., Merriam, J. C., Smith, R. T. & al. (2005) A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proceedings of the National Academy of Sciences of the United States of America* 102(20): 7227–7232 (cit. on p. 110).

Haldane, J. B. S. (1945) Miss E. R. Saunders. *Nature* 156(3961): 385–385 (cit. on p. 43).

Haldane, J. B., Sprunt, A. D. & Haldane, N. M. (1915) Reduplication in mice (Preliminary Communication). *Journal of Genetics* 5(2): 133–135 (cit. on p. 46).

Hamel, C. (2006) Retinitis pigmentosa. *Orphanet Journal of Rare Diseases* 1(1): 40 (cit. on p. 36).

Hamel, C. P. (2007) Cone rod dystrophies. *Orphanet Journal of Rare Diseases* 2(1): 7 (cit. on p. 36).

Han, X., Gharahkhani, P., Mitchell, P., Liew, G., Hewitt, A. W. & MacGregor, S. (2020) Genome-wide meta-analysis identifies novel loci associated with age-related macular degeneration. *Journal of Human Genetics* 65(8): 657–665 (cit. on p. 108).

Hanany, M., Rivolta, C. & Sharon, D. (2020) Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases. *Proceedings of the National Academy of Sciences of the United States of America* 117(5): 2710–2716 (cit. on p. 36).

Hanke, F. D. & Kelber, A. (2020) The Eye of the Common Octopus (Octopus vulgaris). *Frontiers in Physiology* 10 (cit. on p. 31).

Hao, H., Gregorski, J., Qian, H., Li, Y., Gao, C. Y., Idrees, S. & Zhang, B. (2014) In vivo function of the ER-Golgi transport protein LMAN1 in photoreceptor homeostasis. *Advances in Experimental Medicine and Biology* 801: 395–399 (cit. on p. 111).

Harper, P. S. (2008) *A short history of medical genetics*. Oxford monographs on medical genetics ; no. 57. Oxford ; New York: Oxford University Press (cit. on p. 40).

Hartong, D. T., Berson, E. L. & Dryja, T. P. (2006) Retinitis pigmentosa. *Lancet* 368(9549): 1795–1809 (cit. on p. 36).

Heather, J. M. & Chain, B. (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics* 107(1): 1–8 (cit. on p. 49).

Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Ssteinthorsdottir, V., Tthorleifsson, G., Zillikens, C., Sspeliotes, E. K., Mägi, R., Workalemahu, T., White, C., Bouatia-Naji, N., Harris, T. B., Berndt, S. I., Ingelsson, E., Willer, C. J., Weedon, M. N., Luan, J., Vedantam, S. & al. (2010) Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics* 42(11): 949–960 (cit. on p. 52).

Hemani, G., Tilling, K. & Davey Smith, G. (2017) Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genetics* 13(11) (cit. on pp. 55, 93).

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab, H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Davey Smith, G., Gaunt, T. R. & Haycock, P. C. (2018) The MR-base platform supports systematic causal inference across the human phenome. *eLife* 7 (cit. on pp. 93, 117).

Hirasawa, K., Shoji, N., Yoshii, Y. & Haraguchi, S. (2014) Determination of axial length requiring adjustment of measured circumpapillary retinal nerve fiber layer thickness for ocular magnification. *PLoS ONE* 9(9): 107553 (cit. on p. 74).

Hoagland, M. B., L., S. M., Scott, J. F., Hecht, L. I. & Zamecnik, P. C. (1958) A soluble ribonucleic acid intermediate in protein synthesis. *The Journal of biological chemistry* 231(1): 241–257 (cit. on p. 47).

Hoepfner, S., Severin, F., Cabezas, A., Habermann, B., Runge, A., Gillooly, D., Stenmark, H. & Zerial, M. (2005) Modulation of receptor recycling and degradation by the endosomal kinesin KIF16B. *Cell* 121(3): 437–450 (cit. on p. 115).

Hofman, A., Brusselle, G. G., Murad, S. D., van Duijn, C. M., Franco, O. H., Goedegebure, A., Ikram, M. A., Klaver, C. C., Nijsten, T. E., Peeters, R. P., Stricker, B. H., Tiemeier,

H. W., Uitterlinden, A. G. & Vernooij, M. W. (2015) The Rotterdam Study: 2016 objectives and design update. *European Journal of Epidemiology* 30(8): 661–708 (cit. on p. 85).

Hood, D. C. (2017) Improving our understanding, and detection, of glaucomatous damage: An approach based upon optical coherence tomography (OCT). *Progress in Retinal and Eye Research* 57: 46–75 (cit. on p. 71).

Hood, D. C., Raza, A. S., de Moraes, C. G. V., Liebmann, J. M. & Ritch, R. (2013) Glaucomatous damage of the macula. *Progress in Retinal and Eye Research* 32(1): 1–21 (cit. on p. 35).

Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A. & Fujimoto, J. G. (1991) Optical coherence tomography. *Science* 254(5035): 1178–1181 (cit. on p. 32).

Huang, Y. C., Lin, J. M., Lin, H. J., Chen, C. C., Chen, S. Y., Tsai, C. H. & Tsai, F. J. (2011) Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology* 118(4): 642–648 (cit. on p. 52).

Hysi, P. G., Choquet, H., Khawaja, A. P., Wojciechowski, R., Tedja, M. S., Yin, J., Simcoe, M. J., Patasova, K., Mahroo, O. A., Thai, K. K., Cumberland, P. M., Melles, R. B., Verhoeven, V. J., Vitart, V., Segre, A., Stone, R. A., Wareham, N., Hewitt, A. W., Mackey, D. A., Klaver, C. C. & al. (2020) Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. *Nature Genetics* 52(4): 401–407 (cit. on pp. 76, 90, 110).

Ibn Isḥāq al-Ibādī, Ḥ. (1928) *The book of the ten treatises on the eye, ascribed to Hunain ibn Is-hâq (809-877 A. D.) :* Cairo: Government press (cit. on p. 25).

Ignatov, A., Lintzel, J., Hermans-Borgmeyer, I., Kreienkamp, H. J., Joost, P., Thomsen, S., Methner, A. & Schaller, H. C. (2003) Role of the G-protein-coupled receptor GPR12 as high-affinity receptor for sphingosylphosphorylcholine and its expression and function in brain development. *Journal of Neuroscience* 23(3): 907–914 (cit. on p. 115).

Iotchkova, V., Ritchie, G., Geihs, M., Morganella, S., Min, J., Walter, K., Timpson, N., Dunham, I., Birney, E. & Soranzo, N. (2016) GARFIELD - GWAS Analysis of Regulatory or Functional Information Enrichment with LD correction. *bioRxiv*: 7–24 (cit. on p. 91).

Jacobson, S. G., Cideciyan, A. V., Gibbs, D., Sumaroka, A., Roman, A. J., Aleman, T. S., Schwartz, S. B., Olivares, M. B., Russell, R. C., Steinberg, J. D., Kenna, M. A., Kimberling, W. J., Rehm, H. L. & Williams, D. S. (2011) Retinal disease course in Usher

syndrome 1B due to MYO7A mutations. *Investigative Ophthalmology and Visual Science* 52(11): 7924–7936 (cit. on p. 110).

Jaffe, G. J. & Caprioli, J. (2004) Optical coherence tomography to detect and manage retinal disease and glaucoma. *American Journal of Ophthalmology* 137(1): 156–169 (cit. on p. 32).

Jager, R. D., Mieler, W. F. & Miller, J. W. (2008) Age-related macular degeneration. *New England Journal of Medicine* 358(24): 2606 (cit. on p. 35).

James, B., Chew, C. & Bron, A. (2003) *Lecture notes on ophthalmology.* 9th ed. Malden, Mass.: Blackwell Pub. (cit. on pp. 26, 29, 71, 103, 105, 110, 115, 133, 141).

Janeway, C. A. J., Travers, P., Walport, M. & Shlomchik, M. J. (2001) *The complement system and innate immunity*. Garland Science (cit. on pp. 110, 115).

Jia, L., Oh, E. C., Ng, L., Srinivas, M., Brooks, M., Swaroop, A. & Forrest, D. (2009) Retinoid-related orphan nuclear receptor RORβ is an early-acting factor in rod photoreceptor development. *Proceedings of the National Academy of Sciences of the United States of America* 106(41): 17534–17539 (cit. on p. 108).

Jilaveanu, L. B., Parisi, F., Barr, M. L., Zito, C. R., Cruz-Munoz, W., Kerbel, R. S., Rimm, D. L., Bosenberg, M. W., Halaban, R., Kluger, Y. & Kluger, H. M. (2015) PLEKHA5 as a biomarker and potential mediator of melanoma brain metastasis. *Clinical Cancer Research* 21(9): 2138–2147 (cit. on p. 122).

Johannsen, W. (1909) *Elemente der exakten Erblichkeitslehre*. Cambridge Library Collection - Darwin, Evolution and Genetics (cit. on p. 44).

Kamaraj, B. & Purohit, R. (2014) Mutational analysis of oculocutaneous albinism: A compact review. *BioMed Research International* 2014 (cit. on pp. 36, 90, 115).

Kashani, A. H., Zimmer-Galler, I. E., Shah, S. M., Dustin, L., Do, D. V., Eliott, D., Haller, J. A. & Nguyen, Q. D. (2010) Retinal Thickness Analysis by Race, Gender, and Age Using Stratus OCT. *American Journal of Ophthalmology* 149(3): 496–502.e1 (cit. on p. 37).

Katan, M. B. (1986) Apoupoprotein E Isoforms, Serum Cholesterol, and Cancer. *The Lancet* 327(8479): 507–508 (cit. on p. 54).

Keane, P. A., Grossi, C. M., Foster, P. J., Yang, Q., Reisman, C. A., Chan, K., Peto, T., Thomas, D. & Patel, P. J. (2016) Optical Coherence Tomography in the UK Biobank Study – Rapid Automated Analysis of Retinal Thickness for Large Population-Based Studies. *PLOS ONE* 11(10). Ed. by T. L. Sørensen: e0164095 (cit. on p. 62).

Al-Khalili, J. (2015) In retrospect: Book of Optics. *Nature* 518(7538): 164–165 (cit. on p. 25).

Khawaja, A. P., Chua, S., Hysi, P. G., Georgoulas, S., Currant, H., Fitzgerald, T. W., Birney, E., Ko, F., Yang, Q., Reisman, C., Garway-Heath, D. F., Hammond, C. J., Khaw, P. T., Foster, P. J., Patel, P. J., Strouthidis, N., Atan, D., Aslam, T., Barman, S. A., Barrett, J. H. & al. (2020) Comparison of Associations with Different Macular Inner Retinal Thickness Parameters in a Large Cohort. *Ophthalmology* 127(1): 62–71 (cit. on pp. 37, 56, 89).

Khawaja, A. P., Cooke Bailey, J. N., Wareham, N. J., Scott, R. A., Simcoe, M., Igo, R. P., Song, Y. E., Wojciechowski, R., Cheng, C. Y., Khaw, P. T., Pasquale, L. R., Haines, J. L., Foster, P. J., Wiggs, J. L., Hammond, C. J. & Hysi, P. G. (2018) Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. *Nature Genetics* 50(6): 778–782 (cit. on pp. 62, 93, 95).

Khorana, H. G., Büchi, H., Ghosh, H., Gupta, N., Jacob, T. M., Kössel, H., Morgan, R., Narang, S. A., Ohtsuka, E. & Wells, R. D. (1966) Polynucleotide synthesis and the genetic code. *Cold Spring Harbor symposia on quantitative biology* 31: 39–49 (cit. on p. 47).

Khurana, A. K. (2007) *Comprehensive ophthalmology.* 4th ed. New Delhi: New Age International (P) Ltd., Publishers (cit. on pp. 71, 127).

Kidder, B. L., Hu, G. & Zhao, K. (2014) KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biology* 15(2): R32 (cit. on p. 111).

Kim, J. I. & Kang, B. H. (2019) Decreased retinal thickness in patients with Alzheimer's disease is correlated with disease severity. *PLoS ONE* 14(11) (cit. on p. 37).

Kim, K. E. & Park, K. H. (2018) Macular imaging by optical coherence tomography in the diagnosis and management of glaucoma. *British Journal of Ophthalmology* 102(6): 718–724 (cit. on p. 37).

King, A., Azuara-Blanco, A. & Tuulonen, A. (2013) Glaucoma. *BMJ* (*Clinical research ed.*) 346: f3518 (cit. on p. 34).

Kirkwood, B. (2009) Albinism and its implications with vision. *Insight - Journal of the American Society of Ophthalmic Registered Nurses* 34(2): 13–16 (cit. on p. 37).

Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., SanGiovanni, J. P., Mane, S. M., Mayne, S. T., Bracken, M. B., Ferris, F. L., Ott, J., Barnstable, C. & Hoh, J. (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308(5720): 385–389 (cit. on pp. 36, 52, 110, 115, 127).

Kortvely, E., Hauck, S. M., Duetsch, G., Gloeckner, C. J., Kremmer, E., Alge-Priglinger, C. S., Deeg, C. A. & Ueffing, M. (2010) ARMS2 is a constituent of the extracellular matrix providing a link between familial and sporadic age-related macular degenerations. *Investigative Ophthalmology and Visual Science* 51(1): 79–88 (cit. on p. 127).

Koszul, R., Meselson, M., Van oninck, K., Vandenhaute, J. & Zickler, D. (2012) The centenary of Janssens's Chiasmatype theory. *Genetics* 191(2): 309–317 (cit. on p. 47).

Kume, T. & Uemura, S. (2018) Current clinical applications of coronary optical coherence tomography. *Cardiovascular Intervention and Therapeutics* 33(1): 1–10 (cit. on p. 32).

Kwom, Y. H., Fingprt, J. H., Kuehn, M. H. & Alward, W. L. (2009) Primary open-angle glaucoma. *New England Journal of Medicine* 360(11): 1113–1124 (cit. on p. 34).

Lamb, T. D., Arendt, D. & Collin, S. P. (2009) The evolution of phototransduction and eyes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1531): 2791–2793 (cit. on p. 31).

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., McKernan, K. & al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822): 860–921 (cit. on p. 49).

Landowski, M., Kelly, U., Klingeborn, M., Groelle, M., Ding, J. D., Grigsby, D. & Rickman, C. B. (2019) Human complement factor H Y402H polymorphism causes an age-related macular degeneration phenotype and lipoprotein dysregulation in mice. *Proceedings of the National Academy of Sciences of the United States of America* 116(9): 3703–3711 (cit. on p. 127).

Laubichler, M. D. & Davidson, E. H. (2008) Boveri's long experiment: Sea urchin merogones and the establishment of the role of nuclear chromosomes in development. *Developmental Biology* 314(1): 1–11 (cit. on p. 46).

Leske, M. C., Heijl, A., Hussein, M., Bengtsson, B., Hyman, L. & Komaroff, E. (2003) Factors for glaucoma progression and the effect of treatment: The early manifest glaucoma trial. *Archives of Ophthalmology* 121(1): 48–56 (cit. on p. 74).

Levick, W. R. (1967) Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit's retina. *The Journal of Physiology* 188(3): 285–307 (cit. on p. 31).

Lewontin, R. C. & Kojima, K.-i. (1960) THE EVOLUTIONARY DYNAMICS OF COMPLEX POLYMORPHISMS, *Evolution* 14(4): 458–472 (cit. on p. 53).

Li, Q., Chen, P., Zeng, Z., Liang, F., Song, Y., Xiong, F., Li, X., Gong, Z., Zhou, M., Xiang, B., Peng, C., Li, X., Chen, X., Li, G. & Xiong, W. (2016) Yeast two-hybrid screening identified WDR77 as a novel interacting partner of TSC22D2. *Tumor Biology* 37(9): 12503–12512 (cit. on p. 143).

Li, X., Perissi, V., Liu, F., Rose, D. W. & Rosenfeld, M. G. (2002) Tissue-specific regulation of retinal and pituitary precursor cell proliferation. *Science* 297(5584): 1180–1183 (cit. on p. 35).

Li, Y., Xia, X. & Paulus, Y. M. (2018) *Advances in retinal optical imaging* (cit. on p. 31).

Lidén, M., Romert, A., Tryggvason, K., Persson, B. & Eriksson, U. (2001) Biochemical Defects in 11-cis-Retinol Dehydrogenase Mutants Associated with Fundus Albipunctatus. *Journal of Biological Chemistry* 276(52): 49251–49257 (cit. on pp. 128, 143).

Lim, L. S., Mitchell, P., Seddon, J. M., Holz, F. G. & Wong, T. Y. (2012) Age-related macular degeneration. *The Lancet* 379(9827): 1728–1738 (cit. on p. 36).

Lin, M. K., Yang, J., Hsu, C. W., Gore, A., Bassuk, A. G., Brown, L. M., Colligan, R., Sengillo, J. D., Mahajan, V. B. & Tsang, S. H. (2018) HTRA1, an age-related macular degeneration protease, processes extracellular matrix proteins EFEMP1 and TSP1. *Aging Cell* 17(4) (cit. on p. 127).

Lindberg, D. C. (1981) *Theories of vision from al-Kindi to Kepler*. Chicago history of science and medicine. Chicago ; London: University of Chicago Press (cit. on p. 25).

Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P. & Decaprio, J. A. (2011) DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. *Genes and Development* 25(8): 801–813 (cit. on p. 141).

Liu, C., Chen, G., Bentley, A. R., Doumatey, A., Zhou, J., Adeyemo, A., Yang, J. & Rotimi, C. (2019) Genome-wide association study for proliferative diabetic retinopathy in Africans. *Genomic Medicine* 4(1): 1–7 (cit. on p. 52).

Liu, Y., Simavli, H., Que, C. J., Rizzo, J. L., Tsikata, E., Maurer, R. & Chen, T. C. (2015) Patient characteristics associated with artifacts in Spectralis optical coherence tomography imaging of the retinal nerve fiber layer in glaucoma. *American journal of ophthalmology* 159(3): 565–76.e2 (cit. on p. 64).

Liu, Y., Ph, D., Allingham, R. R. & D, M. (2017) Major review : Molecular genetics of primary open-angle glaucoma. *Experimental Eye Research* 160: 62–84 (cit. on p. 89).

Lyell, C. (1832) *Principles of geology.* 2nd ed. London (cit. on p. 41).

Ma, L. J., Xu, L. L., Mao, C. J., Fu, Y. T., Ji, X. Y., Shena, Y., Chen, J., Yang, Y. P. & Liu, C. F. (2018) Progressive changes in the retinal structure of patients with Parkinson's disease. *Journal of Parkinson's Disease* 8(1): 85–92 (cit. on p. 37).

Ma, Z., Xia, W., Liu, F., Ma, J., Sun, S., Zhang, J., Jiang, N., Wang, X., Hu, J. & Ma, D. (2017) SLC44A4 mutation causes autosomal dominant hereditary postlingual non-syndromic mid-frequency hearing loss. *Human Molecular Genetics* 26(2): 383–394 (cit. on p. 124).

Maksimova, N., Hara, K., Nikolaeva, I., Chun-Feng, T., Usui, T., Takagi, M., Nishihira, Y., Miyashita, A., Fujiwara, H., Oyama, T., Nogovicina, A., Sukhomyasova, A., Potapova, S., Kuwano, R., Takahashi, H., Nishizawa, M. & Onodera, O. (2010) Neuroblastoma amplified sequence gene is associated with a novel short stature syndrome characterised by optic nerve atrophy and Pelger-Huët anomaly. *Journal of Medical Genetics* 47(8): 538–548 (cit. on p. 90).

Marchini, J. & Howie, B. (2010) *Genotype imputation for genome-wide association studies* (cit. on p. 73).

Mariotte, E. & Pecquet, J. (1668) A new discovery touching vision. *Philosophical Transactions of the Royal Society of London* 3(35): 668–671 (cit. on p. 25).

Masland, R. H. (2012) The neuronal organization of the retina. *Neuron* 76(2): 266–280 (cit. on p. 29).

Mattar, P. & Cayouette, M. (2015) Mechanisms of temporal identity regulation in mouse retinal progenitor cells. *Neurogenesis* 2(1) (cit. on p. 90).

Maxam, A. M. & Gilbert, W. (1977) A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America* 74(2): 560–564 (cit. on p. 49).

McAllister, J. T., Dubis, A. M., Tait, D. M., Ostler, S., Rha, J., Stepien, K. E., Gail Summers, C. & Carroll, J. (2010) Arrested development: High-resolution imaging of foveal morphology in albinism. *Vision Research* 50(8): 810–817 (cit. on pp. 37, 90, 115, 150).

McCafferty, B. K., Wilk, M. A., McAllister, J. T., Stepien, K. E., Dubis, A. M., Brilliant, M. H., Anderson, J. L., Carroll, J. & Summers, C. G. (2015) Clinical insights into foveal morphology in albinism. *Journal of Pediatric Ophthalmology and Strabismus* 52(3): 167–172 (cit. on p. 37).

McKee, A. E., Minet, E., Stern, C., Riahi, S., Stiles, C. D. & Silver, P. A. (2005) A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restric-

ted expression in the developing mouse brain. *BMC Developmental Biology* 5: 14 (cit. on p. 91).

McKnight, C. M., Newnham, J. P., Stanley, F. J., Mountain, J. A., Landau, L. I., Beilin, L. J., Puddey, I. B., Pennell, C. E. & Mackey, D. A. (2012) Birth of a cohort — The first 20 years of the raine study. *Medical Journal of Australia* 197(11): 608–610 (cit. on p. 85).

McLendon, R., Friedman, A., Bigner, D., Van Meir, E. G., Brat, D. J., Mastrogianakis, G. M., Olson, J. J., Mikkelsen, T., Lehman, N., Aldape, K., Yung, W. K., Bogler, O., Weinstein, J. N., VandenBerg, S., Berger, M., Prados, M., Muzny, D., Morgan, M., Scherer, S., Sabo, A. & al. (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455(7216): 1061–1068 (cit. on p. 51).

Medeiros, F. A., Meira-Freitas, D., Lisboa, R., Kuang, T. M., Zangwill, L. M. & Weinreb, R. N. (2013) Corneal hysteresis as a risk factor for glaucoma progression: A prospective longitudinal study. *Ophthalmology* 120(8): 1533–1540 (cit. on p. 122).

Mendel, G. (1866) Versuche über Plflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr 1865*: 3–47 (cit. on p. 41).

Messing, J., Crea, R. & Seeburg, P. H. (1981) A system for shotgun DNA sequencing. *Nucleic Acids Research* 9(2): 309–321 (cit. on p. 49).

Meyer, H. V. (2020) *meyer-lab-cshl/Genetic-association-studies* (cit. on p. 68).

Meyer, H. V., Dawes, T. J., Serrani, M., Bai, W., Tokarczuk, P., Cai, J., de Marvao, A., Henry, A., Lumbers, R. T., Gierten, J., Thumberger, T., Wittbrodt, J., Ware, J. S., Rueckert, D., Matthews, P. M., Prasad, S. K., Costantino, M. L., Cook, S. A., Birney, E. & O'Regan, D. P. (2020) Genetic and functional insights into the fractal structure of the heart. *Nature* 584(7822): 589–594 (cit. on pp. 133, 168).

Miescher, F. (1871) *Ueber die chemische Zusammensetzung der Eiterzellen* (cit. on p. 46).

Mitchell, P., Liew, G., Gopinath, B. & Wong, T. Y. (2018) Age-related macular degeneration. *The Lancet* 392(10153): 1147–1159 (cit. on p. 35).

Montoliu, L., Grønskov, K., Wei, A.-H., Martínez-García, M., Fernández, A., Arveiler, B., Morice-Picard, F., Riazuddin, S., Suzuki, T., Ahmed, Z. M., Rosenberg, T. & Li, W. (2014) Increasing the complexity: new genes and new types of albinism. *Pigment Cell & Melanoma Research* 27(1): 11–18 (cit. on p. 37).

Moore, C., Sambrook, J., Walker, M., Tolkien, Z., Kaptoge, S., Allen, D., Mehenny, S., Mant, J., Angelantonio, E. D., Thompson, S. G., Ouwehand, W., Roberts, D. J. & Danesh, J. (2014) The INTERVAL trial to determine whether intervals between blood

donations can be safely and acceptably decreased to optimise blood supply: Study protocol for a randomised controlled trial. *Trials* 15(1): 363 (cit. on p. 60).

Morgan, T. H. (1910) Sex limited inheritance in drosophila. *Science* 32(812): 120–122 (cit. on p. 46).

Morgan, T. H. (1911) Random segregation versus coupling in Mendelian inheritance. *Science* 34(873): 384 (cit. on p. 46).

Moschos, M. M. (2014) Physiology and psychology of vision and its disorders: a review. *Medical hypothesis, discovery & innovation ophthalmology journal* 3(3): 83–90 (cit. on p. 23).

Motokura, T., Bloom, T., Kim, H. G., Jüppner, H., Ruderman, J. V., Kronenberg, H. M. & Arnold, A. (1991) A novel cyclin encoded by a bcl1-linked candidate oncogene. *Nature* 350(6318): 512–515 (cit. on p. 143).

Mrejen, S., Audo, I., Bonnel, S. & Sahel, J. A. (2017) Retinitis pigmentosa and other dystrophies. *Developments in Ophthalmology* 58: 191–201 (cit. on p. 36).

Mukherjee, S. (2016) *The gene : an intimate history*. Scribner (cit. on pp. 39–41, 45).

Müller, J. (1837) *Elements of physiology, translated from the German, with notes, by William Baly*. London: Taylor & Walton (cit. on p. 26).

Müller, P. L., Wolf, S., Dolz-Marco, R., Tafreshi, A., Schmitz-Valckenberg, S. & Holz, F. G. (2019) 'Ophthalmic Diagnostic Imaging: Retina'. *High Resolution Imaging in Microscopy and Ophthalmology*. Springer International Publishing: 87–106 (cit. on pp. 71, 103).

Mutlu, U., Colijn, J. M., Ikram, M. A., Bonnemaijer, P. W., Licher, S., Wolters, F. J., Tiemeier, H., Koudstaal, P. J., Klaver, C. C. & Ikram, M. K. (2018) Association of Retinal Neurodegeneration on Optical Coherence Tomography with Dementia: A Population-Based Study. *JAMA Neurology* 75(10): 1256–1263 (cit. on p. 37).

Newton, I. (1704) *Opticks: or, A treatise of the reflexions, refractions, inflexions and colours of light. Also two treatises of the species and magnitude of curvilinear figures.* London: Sam. Smith, & Benj (cit. on p. 26).

Ng, D. S., Cheung, C. Y., Luk, F. O., Mohamed, S., Brelen, M. E., Yam, J. C., Tsang, C. W. & Lai, T. Y. (2016) Advances of optical coherence tomography in myopia and pathologic myopia. *Eye (Basingstoke)* 30(7): 901–916 (cit. on p. 32).

Nieves-Moreno, M., Martínez-de-la-Casa, J. M., Morales-Fernández, L., Sánchez-Jean, R., Sáenz-Francés, F. & García-Feijoó, J. (2018) Impacts of age and sex on retinal layer

thicknesses measured by spectral domain optical coherence tomography with Spectralis. *PLOS ONE* 13(3). Ed. by A. S. Lewin: e0194169 (cit. on p. 74).

Nirenberg, M. W. & Matthaei, J. H. (1961) The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America* 47(2): 1588–1602 (cit. on p. 47).

Nivison-Smith, L., Wang, H., Assaad, N. & Kalloniatis, M. (2018) Retinal Thickness Changes throughout the Natural History of Drusen in Age-related Macular Degeneration. *Optometry and Vision Science* 95(8): 648–655 (cit. on p. 37).

Norman, C. S., O'Gorman, L., Gibson, J., Pengelly, R. J., Baralle, D., Ratnayaka, J. A., Griffiths, H., Rose-Zerilli, M., Ranger, M., Bunyan, D., Lee, H., Page, R., Newall, T., Shawkat, F., Mattocks, C., Ward, D., Ennis, S. & Self, J. E. (2017) Identification of a functionally significant tri-allelic genotype in the Tyrosinase gene (TYR) causing hypomorphic oculocutaneous albinism (OCA1B). *Scientific Reports* 7(1) (cit. on p. 37).

Northstone, K., Lewcock, M., Groom, A., Boyd, A., Macleod, J., Timpson, N. J. & Wells, N. (2019) The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019 [version 1; peer review: 2 approved]. *Wellcome Open Research* 4(51) (cit. on p. 59).

Nowroozizadeh, S., Cirineo, N., Amini, N., Knipping, S., Chang, T., Chou, T., Caprioli, J. & Nouri-Mahdavi, K. (2014) Influence of correction of ocular magnification on spectral-domain OCT retinal nerve fiber layer measurement variability and performance. *Investigative Ophthalmology and Visual Science* 55(6): 3439–3446 (cit. on p. 90).

Ooto, S., Hangai, M. & Yoshimura, N. (2015) Effects of sex and age on the normal retinal and choroidal structures on optical coherence tomography. *Current Eye Research* 40(2): 213–225 (cit. on pp. 63, 73).

Orozco, L. D., Chen, H.-H., Cox, C., Townsend, M. J., Van, M., Campagne, L. & Hackney, J. A. (2020) Integration of eQTL and a Single-Cell Atlas in the Human Eye Identifies Causal Genes for Age-Related Macular Degeneration. *Cell Reports* (cit. on p. 110).

Pappuru, R. R., Ouyang, Y., Nittala, M. G., Hemmati, H. D., Keane, P. A., Walsh, A. C. & Sadda, S. R. (2011) Relationship between outer retinal thickness substructures and visual acuity in eyes with dry age-related macular degeneration. *Investigative Ophthalmology and Visual Science* 52(9): 6743–6748 (cit. on p. 105).

Parmeggiani, F., S. Sorrentino, F., Ponzin, D., Barbaro, V., Ferrari, S. & Di Iorio, E. (2011) Retinitis Pigmentosa: Genes and Disease Mechanisms. *Current Genomics* 12(4): 238–249 (cit. on p. 36).

Patel, P. J., Foster, P. J., Grossi, C. M., Keane, P. A., Ko, F., Lotery, A., Peto, T., Reisman, C. A., Strouthidis, N. G. & Yang, Q. (2016) Spectral-domain optical coherence tomography imaging in 67 321 adults: Associations with macular thickness in the UK biobank study. *Ophthalmology* 123(4): 829–840 (cit. on pp. 62, 64, 135, 141).

Pearson, K. (1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50(302): 157–175 (cit. on p. 42).

Pearson, K. (1901) LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2(11): 559–572 (cit. on pp. 42, 158).

Pieraccioli, M., Nicolai, S., Antonov, A., Somers, J., Malewicz, M., Melino, G. & Raschella, G. (2016) ZNF281 contributes to the DNA damage response by controlling the expression of XRCC2 and XRCC4. *Oncogene* 35(20): 2592–2601 (cit. on p. 143).

Pieraccioli, M., Nicolai, S., Pitolli, C., Agostini, M., Antonov, A., Malewicz, M., Knight, R. A., Raschellá, G. & Melino, G. (2018) ZNF281 inhibits neuronal differentiation and is a prognostic marker for neuroblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 115(28): 7356–7361 (cit. on p. 110).

Plotnikov, D., Shah, R. L., Rodrigues, J. N., Cumberland, P. M., Rahi, J. S., Hysi, P. G., Atan, D., Williams, C. & Guggenheim, J. A. (2019) A commonly occurring genetic variant within the NPLOC4–TSPAN10–PDE6G gene cluster is associated with the risk of strabismus. *Human Genetics* 138(7): 723–737 (cit. on p. 110).

Potts, A. M. (1982) *The World's Eye* (cit. on p. 24).

Punnett, R. C. (1907) *Mendelism.* Cambridge (cit. on p. 43).

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O. & Williams, S. M. (2001) 'The Retina'. *Neuroscience*. Sinauer Associates (cit. on pp. 29, 71, 103, 137).

Quetelet, A. (1835) *Sur l'homme et le développement de ses facultés ou essai de physique sociale : volumes 1 & 2*. Paris: Bachelier (cit. on p. 42).

Quigley, H. A. (2011) Glaucoma. *The Lancet* 377(9774): 1367–1377 (cit. on p. 35).

Rahman, N., Georgiou, M., Khan, K. N. & Michaelides, M. (2020) *Macular dystrophies: Clinical and imaging features, molecular genetics and therapeutic options* (cit. on p. 36).

Ramón y Cajal, S. (1893) La rétine des Vertébrés. *La Cellule* 9: 119–259 (cit. on p. 26).

Ratnapriya, R. & Swaroop, A. (2013) Genetic architecture of retinal and macular degenerative diseases: The promise and challenges of next-generation sequencing. *Genome Medicine* 5(9): 84 (cit. on p. 36).

Renwick, J. H. & Lawler, S. D. (1955) Genetical Linkage Between the ABO and Nail-Patella Loci. *Annals of Human Genetics* 19(4): 312–331 (cit. on p. 48).

Renwick, J. H. & Schulze, J. (1961) A computer program for the processing of linkage data from large pedigrees. *Excerpta Medica International Congress Series* 32(E145) (cit. on p. 48).

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., Bene, J. & al. (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510): 421–427 (cit. on p. 52).

Sahel, J. A. & Dalkara, D. (2019) Gene therapy for retinal dystrophy. *Nature Medicine* 25(2): 198–199 (cit. on p. 36).

Sahel, J. A. & Léveillard, T. (2018) Maintaining cone function in rod-cone dystrophies. *Advances in Experimental Medicine and Biology* 1074: 499–509 (cit. on p. 36).

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M. & Smith, M. (1977) Nucleotide sequence of bacteriophage ϕx174 DNA. *Nature* 265(5596): 687–695 (cit. on p. 49).

Sanger, F. & Coulson, A. R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 94(3): 441–448 (cit. on p. 49).

Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H. & Fischer, D. (1973) Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage f1 DNA. *Proceedings of the National Academy of Sciences of the United States of America* 70(4): 1209–1213 (cit. on p. 49).

Saunders, E. R. (1897) On a Discontinuous Variation Occurring in Biscutella laevigata. *Proceedings of the Royal Society of London* 62: 11–26 (cit. on p. 43).

Savige, J., Ratnaike, S. & Colville, D. (2011) Retinal abnormalities characteristic of inherited renal disease. *Journal of the American Society of Nephrology* 22(8): 1403–1415 (cit. on p. 91).

Scerri, T. S., Quaglieri, A., Cai, C., Zernant, J., Matsunami, N., Baird, L., Scheppke, L., Bonelli, R., Yannuzzi, L. A., Friedlander, M., Egan, C. A., Fruttiger, M., Leppert, M., Allikmets, R. & Bahlo, M. (2017) Genome-wide analyses identify common variants associated with macular telangiectasia type 2. *Nature Genetics* 49(4): 559–567 (cit. on p. 108).

Schuman, S. G., Koreishi, A. F., Farsiu, S., ho Jung, S., Izatt, J. A. & Toth, C. A. (2009) Photoreceptor Layer Thinning over Drusen in Eyes with Age-Related Macular Degeneration Imaged In Vivo with Spectral-Domain Optical Coherence Tomography. *Ophthalmology* 116(3) (cit. on p. 105).

Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A. & Waterston, R. H. (2017) DNA sequencing at 40: Past, present and future. *Nature* 550(7676): 345–353 (cit. on p. 49).

Sheu, W. H., Kuo, J. Z., Lee, I. T., Hung, Y. J., Lee, W. J., Tsai, H. Y., Wang, J. S., Goodarzi, M. O., Klein, R., Klein, B. E., Ipp, E., Lin, S. Y., Guo, X., Hsieh, C. H., Taylor, K. D., Fu, C. P., Rotter, J. I. & Chen, Y. D. I. (2013) Genome-wide association study in a Chinese population with diabetic retinopathy. *Human Molecular Genetics* 22(15): 3165–3173 (cit. on p. 52).

Shin, J. W., Sung, K. R., Lee, G. C., Durbin, M. K. & Cheng, D. (2017) Ganglion Cell–Inner Plexiform Layer Change Detected by Optical Coherence Tomography Indicates Progression in Advanced Glaucoma. *Ophthalmology* 124(10): 1466–1474 (cit. on p. 71).

Simeonov, D. R., Wang, X., Wang, C., Sergeev, Y., Dolinska, M., Bower, M., Fischer, R., Winer, D., Dubrovsky, G., Balog, J. Z., Huizing, M., Hart, R., Zein, W. M., Gahl, W. A., Brooks, B. P. & Adams, D. R. (2013) DNA Variations in Oculocutaneous Albinism: An Updated Mutation List and Current Outstanding Issues in Molecular Diagnostics. *Human Mutation* 34(6): 827–835 (cit. on p. 37).

Singh, T., Kurki, M. I., Curtis, D., Purcell, S. M., Crooks, L., McRae, J., Suvisaari, J., Chheda, H., Blackwood, D., Breen, G., Pietilinen, O., Gerety, S. S., Ayub, M., Blyth, M., Cole, T., Collier, D., Coomber, E. L., Craddock, N., Daly, M. J., Danesh, J. & al. (2016) Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nature Neuroscience* 19(4): 571–577 (cit. on p. 60).

Skariah, G., Perry, K. J., Drnevich, J., Henry, J. J. & Ceman, S. (2018) RNA helicase Mov10 is essential for gastrulation and central nervous system development. *Developmental Dynamics* 247(4): 660–671 (cit. on p. 110).

Skariah, G., Seimetz, J., Norsworthy, M., Lannom, M. C., Kenny, P. J., Elrakhawy, M., Forsthoefel, C., Drnevich, J., Kalsotra, A. & Ceman, S. (2017) Mov10 suppresses retroelements and regulates neuronal development and function in the developing brain. *BMC Biology* 15(1) (cit. on p. 110).

Skorczyk-Werner, A., Pawłowski, P., Michalczuk, M., Warowicka, A., Wawrocka, A., Wicher, K., Bakunowicz-Łazarczyk, A. & Krawczyński, M. R. (2015) Fundus albipunctatus: review of the literature and report of a novel RDH5 gene mutation affecting the invariant tyrosine (p.Tyr175Phe). *Journal of Applied Genetics* 56(3): 317–327 (cit. on pp. 36, 128, 143).

Slatkin, M. (2008) *Linkage disequilibrium - Understanding the evolutionary past and mapping the medical future* (cit. on p. 53).

Smith, H. O. & Welcox, K. W. (1970) A Restriction enzyme from Hemophilus influenzae: I. Purification and general properties. *Journal of Molecular Biology* 51(2): 379–391 (cit. on p. 48).

Song, W. K., Lee, S. C., Lee, E. S., Kim, C. Y. & Kim, S. S. (2010) Macular thickness variations with sex, age, and axial length in healthy subjects: A spectral domain-optical coherence tomography study. *Investigative Ophthalmology and Visual Science* 51(8): 3913–3918 (cit. on p. 74).

Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., Allen, H. L., Lindgren, C. M., Luan, J., Mägi, R., Randall, J. C., Vedantam, S., Winkler, T. W., Qi, L., Workalemahu, T., Heid, I., Steinthorsdottir, V., Stringham, H., Weedon, M. N., Wheeler, E. & al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics* 42(11): 937–948 (cit. on p. 52).

Stearns, F. W. (2010) One hundred years of pleiotropy: A retrospective. *Genetics* 186(3): 767–773 (cit. on p. 44).

Stevens, N. M. (M. (1905a) *Studies in spermatogenesis*. Carnegie Institution of Washington. Publication. no. 36, t. [I]-II. Washington, D.C.: Carnegie Institution of Washington (cit. on p. 46).

Stevens, N. M. (1905b) A study of the germ cells of Aphis rosæ and Aphis œnotheræ. *Journal of Experimental Zoology* 2(3): 313–333 (cit. on p. 46).

Stingl, K., Mayer, A. K., Llavona, P., Mulahasanovic, L., Rudolph, G., Jacobson, S. G., Zrenner, E., Kohl, S., Wissinger, B. & Weisschuh, N. (2017) CDHR1 mutations in retinal dystrophies. *Scientific Reports* 7(1): 1–11 (cit. on p. 108).

Strittmatter, W. J. & Roses, A. D. (1996) Apolipoprotein E and Alzheimer's disease. *Annual Review of Neuroscience* 19: 53–77 (cit. on p. 51).

Sturtevant, A. H. (1913) The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. *Journal of Experimental Zoology* 14(1): 43–59 (cit. on p. 48).

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T. & Collins, R. (2015) UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* 12(3): e1001779 (cit. on p. 51).

Sun, X., Dai, Y., Chen, Y., Yu, D. Y., Cringle, S. J., Chen, J., Kong, X., Wang, X. & Jiang, C. (2017) Primary angle closure glaucoma: What we know and what we don't know. *Progress in Retinal and Eye Research* 57: 26–45 (cit. on p. 34).

Sutton, W. S. (1902) *On the Morphology of the Chromosome Group in Brachystola Magna* (cit. on p. 46).

Sutton, W. S. (1903) The Chromosomes in Heredity. *The Biological Bulletin* 4(5): 231–250 (cit. on p. 46).

Takata, N., Abbey, D., Fiore, L., Acosta, S., Feng, R., Gil, H. J., Lavado, A., Geng, X., Interiano, A., Neale, G., Eiraku, M., Sasai, Y. & Oliver, G. (2017) An Eye Organoid Approach Identifies Six3 Suppression of R-spondin 2 as a Critical Step in Mouse Neuroretina Differentiation. *Cell Reports* 21(6): 1534–1549 (cit. on pp. 90, 141).

Tatham, A. J. & Medeiros, F. A. (2017) Detecting Structural Progression in Glaucoma with Optical Coherence Tomography. *Ophthalmology* 124(12): S57–S65 (cit. on p. 71).

Taylor, R. H., Ainsworth, J. R., Evans, A. R. & Levin, A. V. (1999) The epidemiology of pediatric glaucoma: the Toronto experience. *Journal of AAPOS* 3(5): 308–315 (cit. on p. 34).

Taylor, R. H., Ellingham, R. B., Subramaniam, S. & Wakely, L. A. (2011) Calculating the error in refractive error. *Eye* 25(10): 1333–1336 (cit. on p. 76).

The C. elegans Sequencing Consortium (1998) *Genome sequence of the nematode C. elegans: A platform for investigating biology* (cit. on p. 49).

Thygesen, L. C., Daasnes, C., Thaulow, I. & Brønnum-Hansen, H. (2011) Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. *Scandinavian journal of public health* 39(7 Suppl): 12–16 (cit. on p. 60).

Treisman, J. E. (2004) How to make an eye. *Development* 131(16): 3823–3827 (cit. on p. 31).

Tsang, S. H. & Sharma, T. (2018) Optical coherence tomography. *Advances in Experimental Medicine and Biology* 1085: 11–13 (cit. on p. 32).

Tsui, L. C., Buetow, K. & Buchwald, M. (1986) Genetic analysis of cystic fibrosis using linked DNA markers. *American Journal of Human Genetics* 39(6): 720–728 (cit. on p. 50).

Turley, P., Walters, R. K., Maghzian, O., Okbay, A., Lee, J. J., Fontana, M. A., Nguyen-Viet, T. A., Wedow, R., Zacher, M., Furlotte, N. A., Magnusson, P., Oskarsson, S., Johannesson, M., Visscher, P. M., Laibson, D., Cesarini, D., Neale, B. M., Benjamin, D. J., Agee, M., Alipanahi, B. & al. (2018) Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics* 50(2): 229–237 (cit. on pp. 80, 85, 107, 108).

UK Biobank (2007) *UK Biobank: Protocol for a large-scale prospective epidemiological resource.* Tech. rep. March: 1–112 (cit. on pp. 60, 61).

Ulmer Carnes, M., Liu, Y. P., Allingham, R. R., Whigham, B. T., Havens, S., Garrett, M. E., Qiao, C., Katsanis, N., Wiggs, J. L., Pasquale, L. R., Ashley-Koch, A., Oh, E. C. & Hauser, M. A. (2014) Discovery and Functional Annotation of SIX6 Variants in Primary Open-Angle Glaucoma. *PLoS Genetics* 10(5). Ed. by G. Gibson: e1004372 (cit. on pp. 35, 89).

Ulrich, M., Themstrup, L., De Carvalho, N., Manfredi, M., Grana, C., Ciardo, S., Kästle, R., Holmes, J., Whitehead, R., Jemec, G. B., Pellacani, G. & Welzel, J. (2016) Dynamic Optical Coherence Tomography in Dermatology. *Dermatology* 232(3): 298–311 (cit. on p. 32).

Urone, P. P. (1998) *College physics*. Pacific Grove ; London: Brooks/Cole (cit. on p. 27).

Van de Sompele, S., Smith, C., Karali, M., Corton, M., Van Schil, K., Peelman, F., Cherry, T., Rosseel, T., Verdin, H., Derolez, J., Van Laethem, T., Khan, K. N., McKibbin, M., Toomes, C., Ali, M., Torella, A., Testa, F., Jimenez, B., Simonelli, F., De Zaeytijd, J. & al. (2019) Biallelic sequence and structural variants in RAX2 are a novel cause for auto-

somal recessive inherited retinal disease. *Genetics in Medicine* 21(6): 1319–1329 (cit. on p. 125).

van Lookeren Campagne, M., LeCouter, J., Yaspan, B. L. & Ye, W. (2014) Mechanisms of age-related macular degeneration and therapeutic opportunities. *The Journal of Pathology* 232(2): 151–164 (cit. on p. 35).

Verbakel, S. K., van Huet, R. A., Boon, C. J., den Hollander, A. I., Collin, R. W., Klaver, C. C., Hoyng, C. B., Roepman, R. & Klevering, B. J. (2018) Non-syndromic retinitis pigmentosa. *Progress in Retinal and Eye Research* 66: 157–186 (cit. on p. 36).

Vogenberg, F. R., Barash, C. I. & Pursel, M. (2010) Personalized medicine - Part 1: Evolution and development into theranostics. *Pharmacy and Therapeutics* 35(10): 560 (cit. on p. 59).

Von Tschermak, E. (1900) Ueber künstliche Kreuzung bei Pisum sativum. *Berichte der Deutschen Botanischen Gesellschaft* 18(6): 232–239 (cit. on p. 43).

Walter, K., Min, J. L., Huang, J., Crooks, L., Memari, Y., McCarthy, S., Perry, J. R., Xu, C., Futema, M., Lawson, D., Iotchkova, V., Schiffels, S., Hendricks, A. E., Danecek, P., Li, R., Floyd, J., Wain, L. V., Barroso, I., Humphries, S. E., Hurles, M. E. & al. (2015) The UK10K project identifies rare variants in health and disease. *Nature* 526(7571): 82–89 (cit. on p. 51).

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M. S., Shen, N., Kilburn, D. & al. (1998) Large-scale identification, mapping, and genotyping of single- nucleotide polymorphisms in the human genome. *Science* 280(5366): 1077–1082 (cit. on p. 49).

Watson, J. D. & Crick, F. H. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* 171(4356): 737–738 (cit. on p. 47).

Wawersik, S. & Maas, R. L. (2000) Vertebrate eye development as modeled in Drosophila. *Human Molecular Genetics* 9(6): 917–925 (cit. on p. 31).

Weinreb, R. N., Aung, T. & Medeiros, F. A. (2014) The pathophysiology and treatment of glaucoma: A review. *JAMA - Journal of the American Medical Association* 311(18): 1901–1911 (cit. on pp. 34, 35, 93).

Weinreb, R. N. & Tee Khaw, P. (2004) Primary open-angle glaucoma. *Lancet* 363(9422): 1711–1720 (cit. on p. 34).

Weiss, S. F. (2010) After the fall: Political whitewashing, professional posturing, and personal refashioning in the postwar career of Otmar Freiherr von Verschuer. *ISIS* 101(4): 722–758 (cit. on p. 45).

Weldon, W. F. R. (1890) The variations occurring in certain decapod crustacea. *Proceedings of the Royal Society of London* 47(286-291): 445–453 (cit. on p. 42).

Weldon, W. F. R. (1892) Certain correlated variations in crangon vulgaris. *Proceedings of the Royal Society of London* 51(308-314): 1–21 (cit. on p. 42).

Well, D., Blanchard, S., Kaplan, J., Guilford, P., Gibson, F., Walsh, J., Mburu, P., Varela, A., Levilliers, J., Weston, M. D., Kelley, P. M., Kimberling, W. J., Wagenaar, M., Levi-Acobas, F., Larget-Piet, D., Munnich, A., Steel, K. P., Brown, S. D. & Petit, C. (1995) Defective myosin VIIA gene responsible for Usher syndrome type IB. *Nature* 374(6517): 60–61 (cit. on p. 110).

Wilk, M. A., McAllister, J. T., Cooper, R. F., Dubis, A. M., Patitucci, T. N., Summerfelt, P., Anderson, J. L., Stepien, K. E., Costakos, D. M., Jr, T. B., Wirostko, W. J., Chiang, P. W., Dubra, A., Curcio, C. A., Brilliant, M. H., Gail Summers, C. & Carroll, J. (2014) Relationship between foveal cone specialization and pit morphology in Albinism. *Investigative Ophthalmology and Visual Science* 55(7): 4186–4198 (cit. on p. 115).

Wilks, W. (1907) *Report of the Third International Conference 1906 on Genetics : hybridisation (the cross-breeding of genera or species), the cross-breeding of varieties, and general plant-breeding*. London : Royal Horticultural Society, 1–600 (cit. on p. 43).

Wilson, E. B. (1925) *The cell in development and heredity*. 3rd ed. New York: Macmillan (cit. on p. 46).

Wood, A., Binns, A., Margrain, T., Drexler, W., Povaay, B., Esmaeelpour, M. & Sheen, N. (2011) Retinal and choroidal thickness in early age-related macular degeneration. *American Journal of Ophthalmology* 152(6) (cit. on p. 37).

Wu, R. & Taylor, E. (1971) Nucleotide sequence analysis of DNA: II. Complete nucleotide sequence of the cohesive ends of bacteriophage λ DNA. *Journal of Molecular Biology* 57(3): 491–511 (cit. on p. 48).

Xie, Y., Gonome, T., Yamauchi, K., Maeda-Monai, N., Tanabu, R., Ishiguro, S.-i. & Nakazawa, M. (2020) A spectral-domain optical coherence tomographic analysis of Rdh5-/- mice retina. *PLOS ONE* 15(4). Ed. by K. Stieger: e0231220 (cit. on pp. 110, 143).

Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., Yengo, L., Lloyd-Jones, L. R., Sidorenko, J., Wu, Y., Agbessi, M., Ahsan, H., Alves, I., Andiappan, A., Awadalla, P.,

Battle, A., Beutner, F., Bonder, M. J. J., Boomsma, D., Christiansen, M. & al. (2018) Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nature Communications* 9(1): 1–14 (cit. on p. 52).

Yamamoto, H., Simon, A., Eriksson, U., Harris, E., Berson, E. L. & Dryja, T. P. (1999) Mutations in the gene encoding 11-cis retinol dehydrogenase cause delayed dark adaptation and fundus albipunctatus. *Nature Genetics* 22(2): 188–191 (cit. on p. 36).

Yamamoto, H., Yakushijin, K., Kusuhara, S., Escaño, M. F. T., Nagai, A. & Negi, A. (2003) A novel RDH5 gene mutation in a patient with fundus albipunctatus presenting with macular atrophy and fading white dots. *American Journal of Ophthalmology* 136(3): 572–574 (cit. on p. 143).

Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Weedon, M. N., Loos, R. J., Frayling, T. M., McCarthy, M. I., Hirschhorn, J. N., Goddard, M. E. & Visscher, P. M. (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics* 44(4): 369–375 (cit. on pp. 85, 107, 120).

Yang, P., Chiang, P. W., Weleber, R. G. & Pennesi, M. E. (2015) Autosomal dominant retinal dystrophy with electronegative waveform associated with a novel RAX2 mutation. *JAMA Ophthalmology* 133(6): 653–661 (cit. on p. 125).

Yannuzzi, L. A., Ober, M. D., Slakter, J. S., Spaide, R. F., Fisher, Y. L., Flower, R. W. & Rosen, R. (2004) Ophthalmic fundus imaging: Today and beyond. *American Journal of Ophthalmology* 137(3): 511–524 (cit. on p. 31).

Zagozewski, J. L., Zhang, Q., Pinto, V. I., Wigle, J. T. & Eisenstat, D. D. (2014) The role of homeobox genes in retinal development and disease. *Developmental Biology* 393(2): 195–208 (cit. on p. 90).

Al-Zamil, W. M. & Yassin, S. A. (2017) Recent developments in age-related macular degeneration: A review. *Clinical Interventions in Aging* 12: 1313–1330 (cit. on p. 35).

Zhang, B., Cunningham, M. A., Nichols, W. C., Bernat, J. A., Seligsohn, U., Pipe, S. W., McVey, J. H., Schulte-Overberg, U., De Bosch, N. B., Ruiz-Saez, A., White, G. C., Tuddenham, E. G., Kaufman, R. J. & Ginsburg, D. (2003) Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nature Genetics* 34(2): 220–225 (cit. on p. 111).

Zhang, X. M., Hashimoto, T., Tang, R. & Yang, X. J. (2018) Elevated expression of human bHLH factor ATOH7 accelerates cell cycle progression of progenitors and enhances production of avian retinal ganglion cells. *Scientific Reports* 8(1): 1–13 (cit. on p. 90).

Zhou, X., Nicholson, A. M., Ren, Y., Brooks, M., Jiang, P., Zuberi, A., Phuoc, H. N., Perkerson, R. B., Matchett, B., Parsons, T. M., Finch, N. C. A., Lin, W., Qiao, W., Castanedes-Casey, M., Phillips, V., Librero, A. L., Asmann, Y., Bu, G., Murray, M. E., Lutz, C. & al. (2020) Loss of TMEM106B leads to myelination deficits: Implications for frontotemporal dementia treatment strategies. *Brain* 143(6): 1905–1919 (cit. on p. 124).

Ziccardi, L., Cordeddu, V., Gaddini, L., Matteucci, A., Parravano, M., Malchiodi-Albedi, F. & Varano, M. (2019) Gene therapy in retinal dystrophies. *International Journal of Molecular Sciences* 20(22) (cit. on p. 36).

Zirkle, C. (1935) The Inheritance of Acquired Characters and the Provisional Hypothesis of Pangenesis. *The American Naturalist* 69(724): 417–445 (cit. on p. 40).