Investigation of the genetic epidemiology of age-related macular degeneration, primary open-angle glaucoma, and diabetic retinopathy in diverse populations

By

Nicole Ann Restrepo

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Human Genetics

May, 2015

Nashville, Tennessee

Approved:

Dana C. Crawford, Ph.D.

David J. Calkins, Ph.D.

Jonathan L. Haines, Ph.D.

Milam A. Brantley, MD

Bingshan Li, Ph.D.

To the family and friends who believed in me,

provided support,

and gave comfort

where comfort was needed.

ACKNOWLEDGEMENTS

This work presented in this dissertation was supported by in part by the Vanderbilt University Center for Human Genetics Research and the Computational Genomics Core that provided computational and/or analytical support for this work. Additionally support was provided by funding from the NHGRI NIH Quantitative Ocular Genomics Training Grant 1T32EY021453 and NIH U01 HG004798 and its ARRA supplements. My thanks go out to the Population Architecture Using Genomics and Epidemiology (PAGE) consortium staff and participants of all PAGE studies and our collaborators from the SP2 and SiMES cohorts for their important contributions. The PAGE program is funded by the National Human Genome Research Institute (NHGRI), supported by U01HG004803 (CALiCo), U01HG004798 (EAGLE), U01HG004802 (MEC), U01HG004790 (WHI), and U01HG004801 (Coordinating Center), and their respective NHGRI ARRA supplements. None of this would have been possible without funding and support from the "Epidemiologic Architecture for Genes Linked to Environment (EAGLE)" which is funded through the NHGRI PAGE program (U01HG004798 and its NHGRI ARRA supplements). The study participants derive from the National Health and Nutrition Examination Surveys (NHANES), and these studies are supported by the Centers for Disease Control and Prevention.

Over the past years, I have received support and encouragement from a number of individuals to whom my thanks cannot be adequately expressed.

To my advisor, Dr. Dana Crawford, my sincerest gratitude for the tremendous opportunity and experience of pursuing my PhD. Dana was an exceptional mentor that supported me on each step in my graduate career from beginning to end. Her guidance taught me how to approach scientific problems, how to ask-the right scientific questions, and most importantly how not to present scientific work. Dana had a magical talent to encourage my own scientific explorations and provide invaluable guidance. I would like to thank Dr. Crawford for making the past five years such a memorable one.

To the members of my dissertation committee, Dr. Jonathan L. Haines, Dr. David J. Calkins, Dr. Milam A. Brantley, and Dr. Bingshan Li, my thanks for being a great source of support and guidance for my graduate work. Particularly when informing me that my endeavors were at times miscalculated.

And always my love and appreciation for the many friends and colleagues in the Crawford lab and the Human Genetics PhD program that provided great help and support during the past years. From cupcakes, tears, and the occasional Saturday aperitif.

Finally, I want to thank my companion and part time computational tool belt Mariusz Butkiewicz. His love and support were vital in buckling me in for one of life's rockiest roller-coaster rides and for ensuring that this work made it out, mostly on time.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AGE	-	advanced glycation end products
AGER	-	the advanced glycation end product receptor
Akt	-	activation of protein kinase B
AMD	-	age-related macular degeneration
ANDRE	-	Analytic Data Research by Email portal
AREDS	-	Age-Related Eye Disease Study group
ARIC	-	Atherosclerosis Risk in Communities study
ASW	-	African Americans with ancestry in Southwest United States
ATP	-	adenosine triphosphate
BMI	-	body mass index
CAF	-	coded allele frequencies
CALiCo	-	Causal Variants Across the Life Course program
CCT	-	thinner central corneal thickness
CDC	-	Centers for Disease Control and Prevention
CDR	-	cup-to-disc ratio
CETP	-	cholesteryl ester transfer protein
CEU	-	Europeans CEU
CHB	-	Han Chinese in Beijing
CHD	-	coronary heart disease
CHS	-	Cardiovascular Health Study
CI	-	confidence interval
CPT	-	Current Procedural Terminology billing codes
CTSI	-	National Center of Advancing Translational Technologies
DCCT	-	Diabetes Control and Complication Trial
DNA	-	Deoxyribonucleic Acid
DR	-	Diabetic retinopathy
EAGLE	-	Epidemiologic Architecture for Genes Linked to Environment study
eMERGE	-	Electronic Medical Records and Genomics Network
EMR	-	electronic medical records
FN	-	false negative
FP	-	false positive
GWAS	-	genome-wide association studies
GxE	-	gene environment interactions
HbA1c	-	glycated hemoglobin levels
HDL-C	-	high density lipoprotein cholesterol
HF	-	heart failure
HPLC	-	high performance liquid chromatography
HTR	-	hormone therapy replacement
IBD	-	identity-by-descent
ICAARE	-	International Consortium of African Ancestry Research in Glaucoma
ICD-9	-	International Classification of Diseases

IOP	-	intraocular pressure
IRB	-	Institutional Review Board
LALES	-	Los Angeles Latino Eye study
LCV	-	last clinic visit
LD	-	linkage disequilibrium
LOD	-	logarithm of the odds
MA	-	Mexican American
MAF	-	minor allele frequency
MEX	-	Mexican Americans
MRI	-	magnetic-resonance imaging
mtRNA	-	mitochondrial DNA
MYOC	-	myocilin
NCHS	-	National Center for Health Statistics
NHANES	-	National Health and Nutrition Examination Survey
NHB	-	non-Hispanic Blacks
NHGRI	-	National Human Genome Research Institute
NHGRI	-	the National Human Genome Research Institute
NHW	-	non-Hispanic Whites
NIH	-	National Institutes of Health
NPDR	-	non-proliferative diabetic retinopathy
NPV	-	negative predictive value
NTG	-	normal tension glaucoma
OCT	-	optical coherence tomography
OPP	-	ocular perfusion pressure
OPTN	-	optineurin
OR	-	odds ratio
PAGE	-	Population Architecture using Genomics and Epidemiology
PC	-	principal components
PDGFR	-	platelet-derived growth factor
PDR	-	proliferative diabetic retinopathy
POAG	-	Primary open-angle glaucoma
POS	-	photoreceptor outer segment
PPV	-	positive predictive value
QC	-	quality control
RE	-	refractive errors
RGC	-	retinal ganglion cells
ROS	-	cellular reactive oxygen species
RPE	-	the retinal pigment epithelium
SD	-	Synthetic Derivative
sDR	-	severe Diabetic retinopathy
SiMES	-	the Singapore Malay Eye Study
SNP	-	single nucleotide polymorphism
SNPxtrait	-	pairwise SNP x quantitative trait interactions were modeled using an interaction term
SP2	-	Singapore Prospective Study Programme
T1D	-	type 1 diabetes

T2D	-	type 2 diabetes
TIA	-	transient ischemic attack
TN	-	true negative
TP	-	true positive
VEGF	-	vascular endothelial growth factor
VEI	-	Vanderbilt Eye Institute
VUMC	-	Vanderbilt University Medical Center
WHO	-	The World Health Organization
WHR	-	waist to hip ratio

CHAPTER I

INTRODUCTION

Global and national impact of vision loss and blindness

Studies of the impact of vision impairment and blindness in various countries are rapidly accruing, resulting in a clearer picture of how vision loss will burden future economies and quality of life for global citizens. The World Health Organization (WHO) annually collects health data from among 193 of the Member States. As of 2010, WHO estimates that 285 million people (i.e. ,4.25% of global population) are affected by some form of visual impairment(Pascolini and Mariotti, 2012) and that 80% of visual impairment is correctable or curable. Of those afflicted by visual impairment, fourteen percent are blind. Age-related eye diseases (e.g., cataracts, age-related macular degeneration, glaucoma, and diabetic retinopathy) account for the majority of vision loss/blindness. Populations in developing countries are rapidly expanding, but much of the developed economies are experiencing a general aging of their populations. With this trend, 27% of the expected 7.5 billion individuals from the global population are expected to be over the age of fifty years by 2019. As the population ages, the prevalence of vision loss/blindness will accordingly increase.

In the United States, nearly 38 million Americans over the age of 40 are visually impaired or blind(National Alliance for Eye and Vision Research, 2006). Visual disability can drastically reduce an individual's quality of life and increase the risk of mortality. Quality of life factors encompass wealth, employment, education, recreation, and physical and mental health. The effects on quality of life can be seen on a larger scale in the economic burden of lost worker productivity and medical treatment, approximately $35.4 billion in 2004(Rein et al., 2006). In a survey conducted by the National Eye Institute, adults were asked to think about conditions that would affect their day-to-day life, on a scale of 1-10 with a 10 indicating the greatest impact. Seventy-one percent of surveyed adults answered that loss of vision would rank as a 10(Lions Clubs International Foundation, 2007). Some vision loss is not uncommon as part of the process of normal aging. Aging eyes take longer to adapt to changes in light and dark, glare, and distortion in distance and depth. These types of impairments have been found to be independent risk factors for falls among the

elderly(Freeman et al., 2007; Ivers et al., 1998), where a fall is an event defined in which an individual inadvertently comes to rest on the ground. Individuals over seventy years are at greatest risk of a fall-related mortality(Ivers et al., 1998).

Physiology of the human eye

The human eye is a sensory organ that allows for vision via the transmission of energy from photons to visual pigments contained in specialized photoreceptor cells. This process, called phototransduction, generates an electrical signal that is carried to the brain via the optic nerve where the image is interpreted. Still an active area of research, the complexity of the mechanical and biochemical processes of vision was noted early by scientists such as Charles Darwin, who noted in *The Origin of Species*, how "staggering" that "an organ so perfect as the eye could have been formed by natural selection."(Darwin)

The eye (Figure 1) is composed of two primary structural units with the smaller frontal unit containing the components necessary for the entry and focusing of light rays into the larger structural globe. Light rays initially enter the eye via the cornea, a clear, curved lens responsible for nearly two-thirds of the eye's refractive power(Cassin and Solomon) . The cornea is a fixed structure that covers the lens, iris, and pupil and is composed of a matrix of collagen fibers(Nejad et al., 2014) that maintains its clarity due to a lack of blood vessels. Oxygen is supplied to the cells via two mechanisms: dissolved oxygen from the environment that comes in contact with the cornea via the tear film layer and from the aqueous humor behind the cornea. As light enters the cornea, it is focused by changes in the geometry of the corneal walls.

Light focused by the cornea then passes through the pupil, a hole located in the center of the iris. Pupil size determines the quantity of light that can enter the inner globe of the eye. Size is controlled by dilator and sphincter muscles of the iris (i.e., colored structure of the eye) which react to increase pupil size in the dark or to decrease in response to bright light. After light has passed through the pupil it subsequently must pass through the crystalline lens, a biconvex structure that provides roughly a third of the eye's refractive power, and separates the two chambers of the eye. Similar to the cornea, the crystalline lens is flexible and

facilitates the focusing of light. Ciliary muscles bend and contract through a process called accommodation that alters the shape and thickness of the lens and brings into focus images at varying distances(Schachar, 2006).

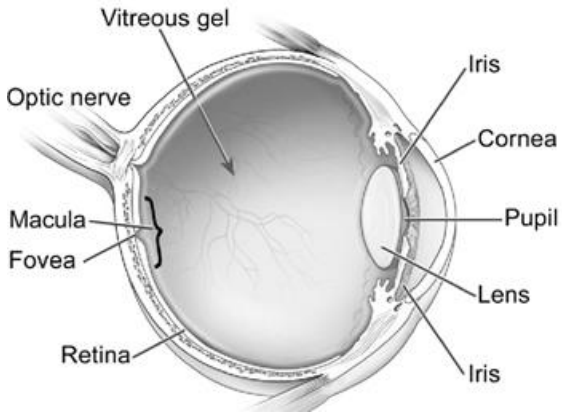


Figure 1: Diagram of the major structural features of the human eye.
Taken from the National Eye Institute, National Institutes of Health. Website accessed December 17, 2014. http://www.flickr.com/photos/nationaleyeinstitute/7544457228/in/photostream/

The globe of the eye maintains its shape in large part by the pressurization of the vitreous humor, a clear gelatinous substance composed of ~98% water, sugar, inorganic salts, collagen, and proteins(Angi et al., 2012; Ulrich et al., 2008). The vitreous is mainly stagnant and as such does not refresh itself over the course of an individual's lifetime. In healthy eyes it maintains contact with retina but is physiologically only attached to the head of the optic nerve. The retina is the light-sensing tissue that lines the inner surface of the eye responsible for transmitting incoming images to the brain through a series of chemical and electrical signals. It is composed of photoreceptor cells (e.g. rods and cones). Rods and cones are highly specialized neurons responsible for triggering phototransduction. Rods are efficient at absorbing photons allowing them to function in low light. A rod contains only one type of light-sensitive pigment and as such contributes to black and white vision. Cones require a higher concentration of photons to trigger transduction but respond to changing images more readily and work best in bright light. Cones are composed of three light-sensitive pigments that respond to short, medium, and long wavelengths of light respectively (i.e., 1) 564-580 nm,

2) 534-545 nm, and 3) 420-440 nm). Due to variation in the sensitivity and pigment composition of cones in an individual eye, the brain is able to interpret a spectrum of colors. Cone density is greatest in the center of the retina, a region called the macula, which generates sharp central vision. The periphery of the retina is composed predominately of rods which correspond to grainier/peripheral vision.

Lastly, the photoreceptor cells transmit electrical information through retina bipolar cell that converge below the retinal tissue to be transmitted to retinal ganglion cells (RGC). These RGC form the optic nerve which exit from the back of the eye and continue to the image centers of the brain.

Major contributors to vision loss and blindness

Vision loss and blindness are multifactorial conditions whose definitions can range between institutions, organizations, and countries. Complete blindness, a condition in which an individual sees no light, occurs rarely. In the United States," legal blindness" is defined by the Social Security Administration as a visual impairment of 20/200 or worse in an individual's better eye after correction with refractive lenses or a visual field limitation in the better eye that subtends at less than a 20 degree angle at the widest diameter of the visual field(OCOMM.OCPT). Causes of vision loss/blindness encompass a wide range of diseases, environmental factors, accidents, and simple access to health care.

Perhaps the greatest contributors to vision impairment are refractive errors (RE). A RE occurs when an eye's ability to focus light is reduced or impeded. A typical RE occurs when light that enters the eye is focused to a point in front of the retina, instead of directly on the retina, resulting in myopia (i.e., nearsightedness). In this state an individual is capable of clearly seeing objects up close while those at a distance are blurred. Myopia can be caused by one of three conditions. Either the cornea or crystalline lens is too curved or the length of the eye globe exceeds 26.5mm(Friedman and Kaiser, 2007). All of these conditions can result in the focused beams of light falling short of the retina. In contrast, hyperopia occurs when the cornea or lens is too flat or else the globe of the eye is too short, thus resulting in light rays being focused 'behind' the retina. Individuals with hyperopia see distant objects clearly while those up close are

blurred. A third form of RE is called astigmatism. Astigmatism results from variation in the curvature of either the lens or cornea in one location in comparison to the rest of the structure. This optical defect leads to blurring of the vision at any distance as light is focused differently through the various reflective planes in the cornea/lens. Together RE are the greatest form of correctable vision impairment, constituting 43% of total global visual impairment cases as of 2010(Pascolini and Mariotti, 2012). In the United States it is

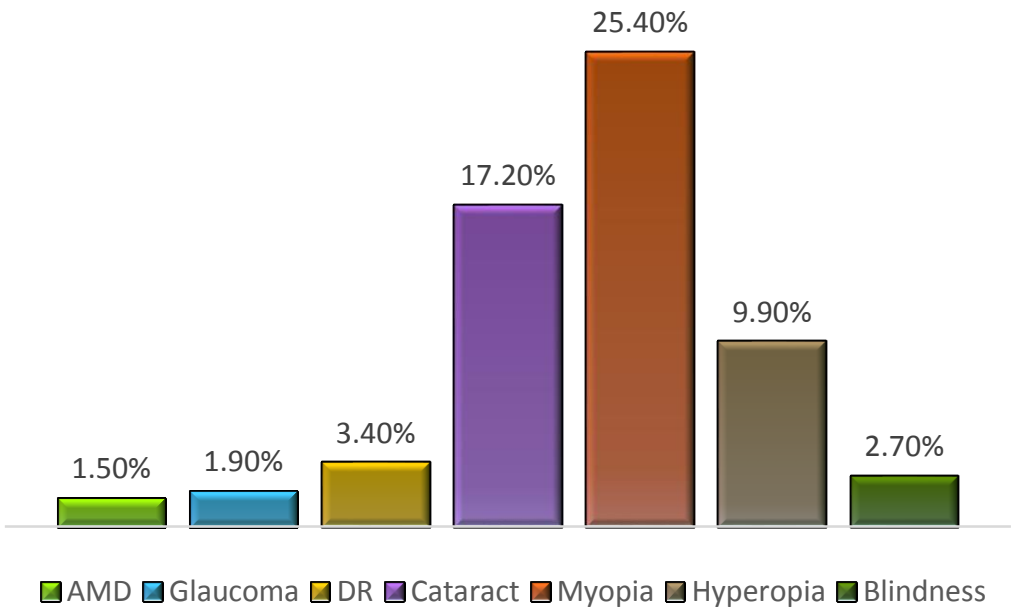


Figure 2: Prevalence rates of major ocular diseases in United States adults over the age of 40 years. Estimates were taken from the publication *The Vision Problems in the U.S.*

estimated that myopia and astigmatism affect 33.1% and 36.2% of the general population over twenty years of age, with hyperopia occurring less frequently at 3.6%(Vitale S, 2008).

Four primary causes of vision impairment and blindness in the elderly are cataracts, age-related macular degeneration (AMD), glaucoma, and diabetic retinopathy (Figure 2). Collectively these conditions contribute to nearly 75% of total cases of blindness and 81% of total cases of low vision in the United States(Congdon et al., 2004). Cataracts form through the gradual opacification of the eye's crystalline lens and are the leading cause of reversible blindness(Asbell et al., 2005) accounting for 51% of global cases of

blindness(Brian and Taylor, 2001). Cataracts commonly forms due to normal aging with cataractogenesis taking place in the nucleus of the lens (i.e., nuclear cataract), at the back of the lens (i.e., subcapsular cataract), or at the periphery of the lens' outer layer (i.e., cortical cataract). Absolute clouding of the lens prevents light from entering the eye. Currently the only available treatment option is for surgical extraction of the deteriorated lens and implantation of a new artificial lens.

AMD is the third leading cause of visual impairment worldwide,(Resnikoff et al., 2004) commonly affecting seniors more than any other form of blindness. AMD is often a bilateral condition involving the destruction of photoreceptors in the macula that leads to loss of central vision. Typically AMD presents on a disease continuum from mild, to intermediate, and then severe. Intermediate AMD is defined as the presence of either many medium-sized drusen or at least one large drusen in one or both eyes. Severe AMD is referred to as "late" AMD and presents as either atrophic (dry) or neovascular (wet) AMD. Symptoms of dry AMD include the formation of drusen deposits, pigment disruption, and geographic atrophy. Wet AMD is a severe condition resulting from the neovascularization of the choriorcapillaris. These aberrant blood vessels are fragile and can rupture leading to blood and/or fluids entering the extracellular space between the retina and the retinal pigment epithelium (RPE). AMD affects over 1.8 million adults over the age of 40 years in the U.S. As the baby boomer generation enters retirement and life expectancy continues to increase, the incident number of AMD cases is expected to grow from 11 million today to approximately 22 million by the year 2050(David S. Friedman (last) et al., 2012).

As the second leading cause of blindness in the United States(Resnikoff et al., 2004), glaucoma is a driving force behind vision disability. Glaucoma is a heterogeneous group of eye diseases that are characterized by chronic degeneration of the optic nerve. Glaucoma is distinguishable from other conditions of optic neuropathy via the presentation of the optic-nerve tissue (pink versus loss of color) and formation of optic-nerve cup (present versus not present)(Kwon et al., 2009). Cupping occurs due to loss of retinal ganglion cells axons and the support vasculature. Individuals with glaucoma gradually lose peripheral vision and, if

left untreated, the disease can result in blindness. As a whole, glaucoma is responsible for approximately 130,000 cases of blindness in the United States as of 2000(Quigley and Vitale, 1997).

Diabetic retinopathy (DR) is the leading cause of blindness in working age adults (i.e., 20-74 years old) in the United States(Klein and Klein, 1995). DR is caused by abnormalities in the microvasculature of the retina and is present in 82% of type 1 diabetes (T1D)(Roy et al., 2004) and 40% of type 2 diabetes (T2D)(The Eye Diseases Prevalence Research Group, 2004) patients in the United States. DR is traditionally classified according to the absence or presence of new blood vessel growth within the retina: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR), respectively. Early stages of DR are characterized by little vision loss, which are the most effective times at which to treat. Early intervention of glycemic control has been shown to delay progression of retinopathy. Late stages of DR are often accompanied by irreversible, severe vision loss or blindness.

Racial/ethnic differences in risk of common ocular diseases

Over time trends in vision disease and impairment have fluctuated with rates for AMD declining in recent decades from 9.4% in U.S. adults in the Third (III) National Health and Nutrition Examination Survey (NHANES) conducted between 1988 and 1994 to 6.5% in the NHANES 2005-2008 cohorts(Klein R, 2011). Simultaneously rates for myopia have surged in developed and Asian countries affecting 20-40% of

Table 1: Prevalence rates of common ocular diseases in adults within the United States by race/ethnicity

Populations	AMD >40 yrs	DR* >40 yrs	Glaucoma >40 yrs	Myopia >20 yrs
European American	7.3%	24.8%	5.6%	35.2%
African American	2.4%	36.7%	12%	28.6%
Hispanics	5.1%	37.4%	6.5%	25.1%
Asian Americans	6.8%	25.7%	6.5%	-

Estimates were obtained from published NHANES estimates.
*Percentage rates are based on the proportion of the population over 40 years with diabetes.

surveyed populations(Rosman et al., 2012; Schache and Baird, 2013; Vitale S, 2009). Hypotheses concerning the causality of the these changes in incidence vary from access to better health care, changes in clinical ascertainment(Klein R, 2011; Vitale S, 2009), changes in life-style factors (i.e., reduction in smoking rates), and even variations in the racial/ethnic make-up of nations. Understanding the role of race/ethnicity in disease susceptibility may provide a better understanding of the risk factors and etiology disease as well as affected biological pathways of disease states.

Prevalence rates of AMD vary by race and ethnicity with non-Hispanic whites (7.3%) experiencing a higher burden than non-Hispanic blacks (2.4%), and Mexican Americans (5.1%) in a study of the U.S. population over the age of 40(Klein R, 2011). The prevalence of AMD in Asian populations has been found to be similar to that seen in European-descent populations at approximately 6.8%(Kawasaki et al., 2010). Although prevalence rates are similar between Asians and European populations, there are differences in the rate and risk of AMD sub-types by specific Asian ethnicity. For example, in one study, Chinese Americans and Pakistani Americans had an increased risk, as calculated by hazards ratio, for dry AMD compared to European-descent populations but Japanese Americans had comparatively less risk for developing wet AMD(Stein et al., 2011a).

The racial/ethnic disparities in prevalence and progression of DR have been noted by multiple studies but little has been done to better understand these differences. Surveys have shown that the prevalence of DR is higher in African Americans (36.7%) and Hispanics (37.4%) compared to European Americans (24.8%) with diabetes(WONG et al., 2006) (Table 1). This difference may be explained, in part, by the higher rates of diabetes in these two racial/ethnic groups compared with European Americans, but the overall trend is higher even after accounting for differences in risk factors, suggesting that other population-specific variables are at play. In studies comparing phenotypic differences between African Americans and Latinos, it was found that Latinos on average suffer from a greater number of intraretinal hemorrhages and typically experience a faster and more severe progression of the disease(Chen et al., 2009). Other differences in DR initiation and progression have been observed in a cross-sectional study utilizing the NHANES 2005-2006

database where differences in the prevalence of DR between European Americans and Africans Americans were observed dependent on glycated hemoglobin levels (HbA1c)(Cheng et al., 2009). The International Expert Committee proposed in 2010 to establish the diabetes HbA1c diagnostic criteria threshold at 6.5% (HbA1c/HbA). This threshold was based on the prevalence of DR, an accepted early indicator of diabetes complication, in a European-descent cohort. Screening based on this criterion identified African Americans as having a crude prevalence of DR at 13.1% versus 6.3% in European Americans(Tsugawa et al., 2012). These differences suggest that ancestry-specific guidelines may need to be developed in order to accurately screen and diagnose individuals of non-European descent.

Glaucoma is the leading cause of blindness in African Americans(Congdon et al., 2004) with prevalence rates approximately double that observed in European-descent populations(Congdon et al., 2004; Friedman et al., 2004a; Stein et al., 2011b) (Table 1). The rates of glaucoma have been found to be similar between European, Japanese, and Indian populations with rates approaching those observed in African descent populations in China in the oldest age categories(Quigley and Broman, 2006). Although African Americans comprise the group of highest risk of developing glaucoma-related vision problems, many cases remain undiagnosed until later stages of disease. Previous studies have suggested that nation-wide implementation of screening middle aged African Americans could decrease the rate of undiagnosed glaucoma from 50% to 27%(Ladapo et al., 2012). Earlier screening and diagnosis enables patients to more effectively leverage current treatment options to reduce the risk of bilateral blindness later in life(Ladapo et al., 2012).

Globally, myopia is the most common eye problem afflicting upwards of 80 million children(Siatkowski et al., 2008). Previous epidemiological studies have suggested that it affects more than 70% of inhabitants in some Asian populations(Lin et al., 2004; Wong et al., 2000; Woo et al., 2004). In the United States, 25-30% of the population experiences some level of myopia. In children, myopia afflicts more Hispanic and Asian children than their European-descent peers (13.2% and 18.5% versus 4.4%)(Kleinstein et al., 2003). This difference is reversed in adults with Hispanics experiencing the lowest levels of myopia while Europeans and Asians experience the highest rates (14.2% vs 31.0% and 37.2%, respectively)(Pan et al.).

Evidence for the role of genetics in common, complex ocular disease

Genetics plays an integral role in the evolution of diverse human populations. As a discipline, human genetics is relatively young compared with other disciplines in the natural sciences. Gregor Mendel was the first to publish on the concept of discrete traits in 1865[1], which is the passing on of a heritable unit from parent to offspring. Mendel's work was all but unknown and forgotten until the 1890s. Spurred in part by the re-discovery of Mendel's work, quantitative genetics germinated and began to flourish in the early 20th century despite few or no natural datasets available apart from phenotypic collections. In this period, methods were developed to describe and measure heritability(Visscher et al., 2008) (the variability in phenotype attributable to a genetic component), epistasis(Norton and Pearson, 1976) (departure from additivity), and linkage(Bateson et al.). The advancement of genetic knowledge would gain speed with the elucidation of the DNA molecule in 1944[2], to the discovery of the DNA structure by Watson, Crick, and Franklin in 1953[3]. Irrefutably the most publicized advancement was the race to complete the first draft of the human genome sequence in 2000(Lander et al., 2001; Venter et al., 2001). This milestone coupled with the International HapMap Project(International HapMap Consortium, 2005) is the foundation on which current genome-wide association studies (GWAS) are based

GWAS is a study design that interrogates common genetic variation across the genome for associations with a disease or trait of interest and is currently the study design of choice for common, complex diseases(Hirschhorn and Daly, 2005). In the past decade, both candidate gene studies and GWAS have identified over 14,000 genomic loci associated with upwards of 5,000 quantitative traits and diseases(Welter et al., 2014). As detailed below, candidate gene studies and GWAS have been performed for diseases of vision loss with varying success.

Diabetic retinopathy

It is well established that diabetes, necessary for the development of diabetic retinopathy, is heritable. Multiple populations have been utilized in the study of the heritability of diabetes(WONG et al., 2006)·(Leslie and Pyke, 1982). As estimated in a Finnish twin study(Hyttinen et al., 2003), the broad sense

heritability for T1D was high as 88%. Other studies including twin and familial aggregation studies have calculated heritability estimates for T2D in the range of 0.31-0.69(Das and Elbein, 2006) dependent on race/ethnicity and age-of-onset.

It is well understood that risk and development of diabetes has a genetic component in conjunction with strong environmental influences, but other evidence suggests DR has additional independent genetic components unrelated to exposure to diabetes. The Diabetes Control and Complication Trial (DCCT)(1987), a multicenter, clinical study of the effects of intensive blood glucose control on progression of diabetes, identified novel risk factors for DR as well as validated important known risk factors. Major risk factors include male sex, duration of diabetes, glycemic control, hypertension, hyperlipidemia, type of diabetes (T1D versus T2D), age, and race/ethnicity. Data from the DCCT suggest that by maintaining tight glycemic control, it is possible to both reduce the risk of developing DR (by 76%) and slow progression of pre-existing DR (by 54%). Other studies determined that glycemic control and duration of diabetes only account for approximately 11% of the variation(Lachin et al., 2008) in retinopathy risk, which suggests that other factors play a role. Family studies of the relative risk of severe DR (sDR), as defined by the presence of diabetic macular edema or proliferative DR, have shown that family members of an individual with sDR were three times more likely to develop the condition(Arar et al., 2008). Other studies calculated heritability (h^2) for proliferative diabetic retinopathy at $h^2=0.52\pm0.31$(Hietala et al., 2008). Collectively, these family studies suggest that while diabetes is a major risk factor for the development of DR, it is not deterministic. Indeed, in some clinical cohorts diabetic patients have suffered from diabetes and other diabetic complications for upwards of twenty years without presentation of retinal disease(Kullberg and Arnqvist, 1995).

Age-related macular degeneration

Genetic factors may contribute anywhere from 46% to 71% of the variability see in AMD pathology(Seddon et al., 2005), exceeding the contribution of environmental risk factors to variation in AMD susceptibility and clinical severity. Heritability estimates of AMD have been determined from male

twins registered in the World War II Veterans Twin database with an overall AMD heritability of 46%, 67% in intermediate AMD cases, and 71% in advanced AMD(Seddon et al., 2005). Recently, the heritability of advanced AMD subtypes (geographic atrophy or choroidal neovascularization) was assessed in siblings recruited from a multitude of clinical practices across the United States and found to be concordant between sibs and clinical subtypes(Sobrin et al., 2012). Quantitative traits associated with clinical manifestations of the disease, such as the size and coverage of drusen deposits, in particular contain a strong genetic component that influences the biological variability of these traits. The variation in the presence of small hard drusen, a precursor of early AMD, was 63% heritable in a young adult twin pairs study(Munch et al., 2007).

The genetics of AMD has been extensively explored in numerous cohorts over the last few decades, but before the successful identification of genetic factors, studies of epidemiological risk variables for AMD consistently found that specific groups demarcated by demographics and/or environmental exposures experience a greater burden of disease. Those at an increased risk are women(Smith et al., 1997), older individuals, smokers, and individuals of European descent(Chakravarthy et al., 2010; 2000a). Several studies have identified hypertension, increased body-mass-index, and poor lipid profiles as environmental modifiers of AMD risk. A substantial body of research has implicated lipid levels as a major risk factor in AMD, particularly low high density lipoprotein (HDL)-cholesterol, though the relationship between serum HDL-C levels and AMD risk is inconsistent(Klein et al., 2003a; van Leeuwen et al., 2004; Reynolds et al., 2010; Tomany et al., 2004). Cigarette smoking is perhaps the greatest contributor to modifiable environmental risk factors. Individuals with bilateral AMD were found to be more likely to have been heavy smokers (odds ratio 5.1) than with those who presented with unilateral AMD(Chakravarthy et al., 2007).

Glaucoma

Studies of glaucoma have been complicated by the heterogeneous nature of the disease with family history inconsistently being identified as a risk factor(Budde, 2000)(Tielsch et al., 1994)(Leske et al., 1995). Given the difficulties of phenotypic heterogeneity in glaucoma, most heritability studies available are based on

more easily-measured glaucoma endophenotypes. The following are the heritability estimates for quantitative traits reproducibly found to be associated with glaucoma: central corneal thickness ranges from 0.35-0.72%(Charlesworth et al., 2010; Freeman et al., 2013; van Koolwijk et al., 2007), intraocular pressure (IOP) 0.35– 0.94%(Charlesworth et al., 2010; Freeman et al., 2013), and cup-to-disc ratio 0.56-0.66%(Chang et al., 2005; Freeman et al., 2013). Pulsatility of choroidal blood flow and velocity are additional quantitative traits whose variation from normal parameters has been seen in individuals with glaucoma(Findl et al., 2000; Fontana et al., 1998), yet heritability studies have found no significant genetic contribution to their variability(Freeman et al., 2013).

Evidence suggests that specific non-modifiable and environmental factors drive the emergence and progression of the disease, such as African ancestry(Tielsch et al., 1994), age(Leske et al., 1995), myopia(Pan et al., 2013), and high intraocular pressure(Chandrasekaran et al., 2006; Jiang et al., 2012; Leske et al., 1995).

Primary open-angle glaucoma (POAG) is the most prevalent clinical subtype of glaucoma in the United States. Early linkage and family-based genetic association studies identified the *MYOC* (myocilin), *OPTN* (optineurin), and *WDR36* (WD repeat domain 36)(Monemi et al., 2005; Rezaie et al., 2002; Stone et al., 1997) genes as susceptibility loci for POAG. Mutations in *MYOC* are known to cause hereditary early-onset POAG in multiple populations(Adam et al., 1997; Stone et al., 1997; Suzuki et al., 1997). *MYOC* is expressed in several tissues including the trabecular meshwork(Polansky et al., 1997) where it is hypothesized that mutated versions of the protein are not being adequately secreted into the aqueous humor(Jacobson et al., 2001). In a study of African Americans, the frequency of myocilin mutations was comparably lower (~1.4%) than in other populations (~2-4%)(Liu et al., 2012). This suggests that other genetic loci are driving risk in this group.

Summary

In summary, common ocular diseases are multifaceted conditions driven in part by both environmental influences and genetics. Separately age-related macular degeneration, primary open-angle glaucoma, and diabetic retinopathy afflict unique structural attributes of the vision pathway, and together they are the driving cause of vision loss and blindness across the globe. Additional research is necessary to understand the genetic and epidemiologic architecture of these disease states in diverse populations; knowledge of which may lead to a better understanding of ancestry-specific etiologies and development of targeted screening programs. Genetic association studies, such as the ones outlined here in subsequent chapters, may provide new insights into the genetic epidemiology of common, ocular diseases and further our understanding of the biological processes inherent to disease risk and prevention. Briefly, here is an outline of the work contained in the following chapters.

In Chapter II, I take a candidate gene approach to determine if variants previously identified as risk loci for AMD in European-descent populations also contribute to risk in non-European populations. To-date, most genetic association studies have been performed in European-descent and Asian populations. A meta-analysis was performed on samples with targeted genotyping data for known AMD and lipid trait-associated SNPs from the Population Architecture using Genomics and Epidemiology (PAGE) study. Populations included in this study were European Americans, African Americans, Mexican Americans, and Singaporeans. Furthermore, I explore whether mitochondrial genetic variants affect risk of AMD in a subset of NHANES III and NHANES 2007-2008 samples. Additionally, I explore the role of gene x environment interactions in risk of AMD as single SNP associations are known to account for only a portion of the risk.

In Chapter III, I pursue strategies to extract ocular phenotypes and traits from the Vanderbilt University Medical Center DNA repository linked to de-identified medical records (i.e., the Synthetic Derivative (SD)). The SD contains over 1.4 million records with the average patient's record containing 6.5 years of medical history and an average of 8 prescriptions. These records consist of both inpatient and outpatient records. I developed phenotyping algorithms to identify DR and POAG cases and controls via a

combination of International Classification of Disease (ICD-9) billing codes, Current Procedural Terminology (CPT) procedural codes, and free text searches.

The work in Chapter IV and V explores the role of common genetic variation in risk of POAG and DR, respectively, in African Americans that were extracted from the SD in Chapter III. In chapter IV, I utilized both a candidate gene approach and a hypothesis-free analyses to jointly identify whether previously identified risk loci contribute to risk of POAG in this population and whether novel loci could be discovered. Similarly in Chapter V, I explored the role of common genetic variation in risk of DR in African Americans from the Vanderbilt University Medical Center SD. Limited progress has been made in identifying susceptibility variants for DR and given that mitochondria are known to play a pathological role in DR, I additionally performed a meta-analysis of mitochondrial variants genotyped in a diverse set of NHANES populations to ascertain their potential contribution to DR risk.

Finally in Chapter VI, I will summarize the work present in Chapters II through V for each ocular phenotype presented. I will also discuss future directions for the field and a potential new project.

CHAPTER II

INVESTIGATION OF THE GENETIC ARCHITECTURE OF AGE-RELATED MACULAR DEGENERATION IN DIVERSE POPULATIONS

Introduction

Age-related macular degeneration (AMD) is a bilateral, multifactorial condition of the retina and the retinal support system leading to gradual deterioration of photoreceptor cells and subsequent loss of central vision. The fundamental mechanics of vision loss of AMD is the painless destruction of highly specialized photoreceptor "cone" cells packed within the macula which is located in the central region of the retina (Figure 3). Cones are comprised of three visual pigments termed opsins, which are light-sensitive G protein-coupled receptors. These G protein-coupled receptors absorb light at various wavelengths and give the perception of trichromatic vision. L-cones absorb wavelengths in the range of 564-580 nm (i.e. long), which correspond with the color red in the electromagnetic spectrum. M-cones (i.e. medium) respond to wavelengths in the range of 534-545 nm for green, and S-cones (i.e. short) absorb wavelengths in the blue region at 420-440 nm. Together these three types of cones, when stimulated by light to varying degrees, create the perception of a color spectrum in the brain. The fovea, a small pit in the center of the macula, contains the greatest density of cones.

(Restrepo et al., 2014, 2015) Adapted from: Restrepo NA, Spencer KL, Goodloe R, Garrett TA, Heiss G, Buzkova P, Jorgensen N, Jensen RA, Matise TC, HIndorff LA, Klein R, Wong TY, Cheng CY, Cornes BK, Tai ES, Ritchie MD, Haines JL, Crawford DC. Genetic determinants of age-related macular degeneration in diverse populations from the PAGE Study. *Investigative Ophthalmology & Visual Sciences* 10, 6839-6850 (2014).

Restrepo, NA, Mitchell SL, Goodloe RJ, Murdock DG, Haines JL, Crawford DC. Mitochondrial variation and the risk of age-related macular degeneration across diverse populations. *Pacific Symposium Biocomputing. Pac.* **20,** 243–254 (2015).

The fovea, which is located at the very center of the macula encompassing a diameter of approximately 1.5 mm, is responsible for high-acuity vision. Although the fovea comprises but a small portion of the retina, it contains roughly half of the nerve fiber cells contained in the optic nerve bundle(Curcio and Allen, 1990). Photoreceptors each contain an outer segment structure (i.e. photopigment), a nucleus, synaptic terminal, an inner fiber, and an inner segment composed of the endoplasmic reticulum, ribosomes, and mitochondria(Hildebrand and Fielder, 2003). The outer and inner segments are connected by series of plasma membrane discs that contain rhodopsin, photopigment essential in phototransduction, and these discs are shed and replenished every ten days(Chuang et al., 2007).

In healthy eyes, the retina is supported by a unique vascular network made up in large part by the choroid which is a layer of connective tissue positioned between the sclera, the hard outer layer of the eyeball, and

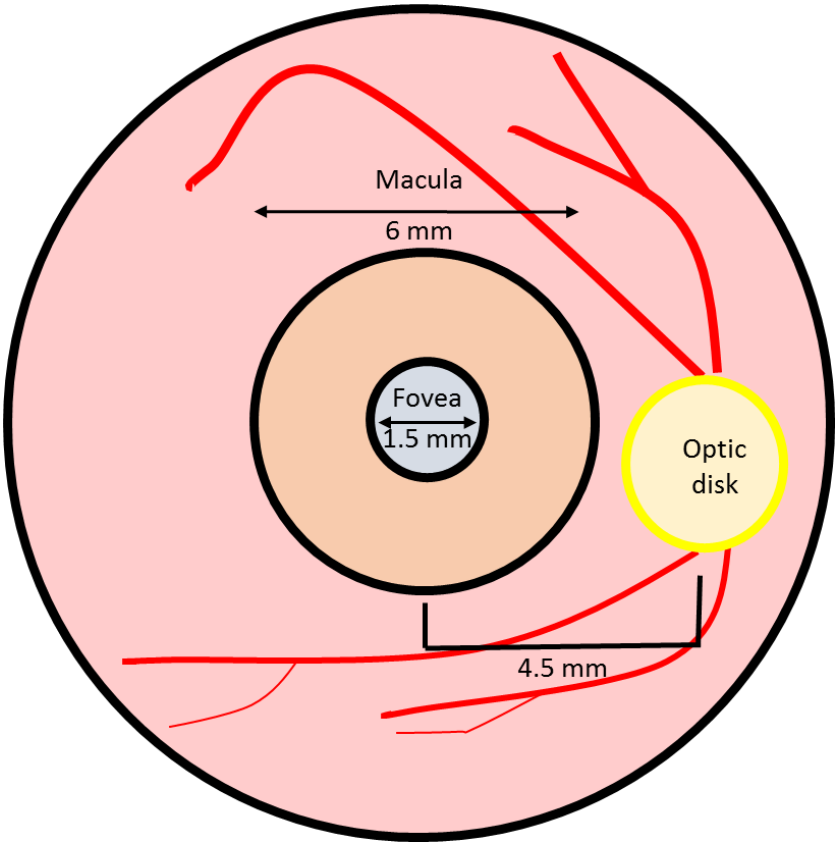


Figure 3: Simple diagram of the human retina;
Visualizing the location of the macula and fovea as would be seen from a funduscopic image.

the retina. Retinal oxygen demand exceeds almost all other tissues in the body(Yu and Cringle, 2001), and this demand is met in large part by the ophthalmic artery. The ophthalmic artery provides approximately 80% of total ocular blood flow to the choroid(Hildebrand and Fielder, 2003). The choroidal network in turn supplies nutrients, oxygen, and glucose to the retinal pigment epithelium (RPE) and photoreceptors and acts as a conduit for disposal of cellular waste. RPE is the pigmented cell layer situated between the retina and the choroid and acts as the outer layer of the blood-retina barrier, preventing molecules from leaky choriocapillaries from crossing into the retinal tissue. RPE cells are tightly compacted in the fovea (i.e. 7,500 cells/mm^2) versus the peripheral macula (5,000 cells/mm^2)(Ach et al., 2014). The RPE plays an integral role in the phagocytosis of shed photoreceptor outer-segment discs. Perturbations in the macula/fovea region can have devastating consequences for visual integrity and health.

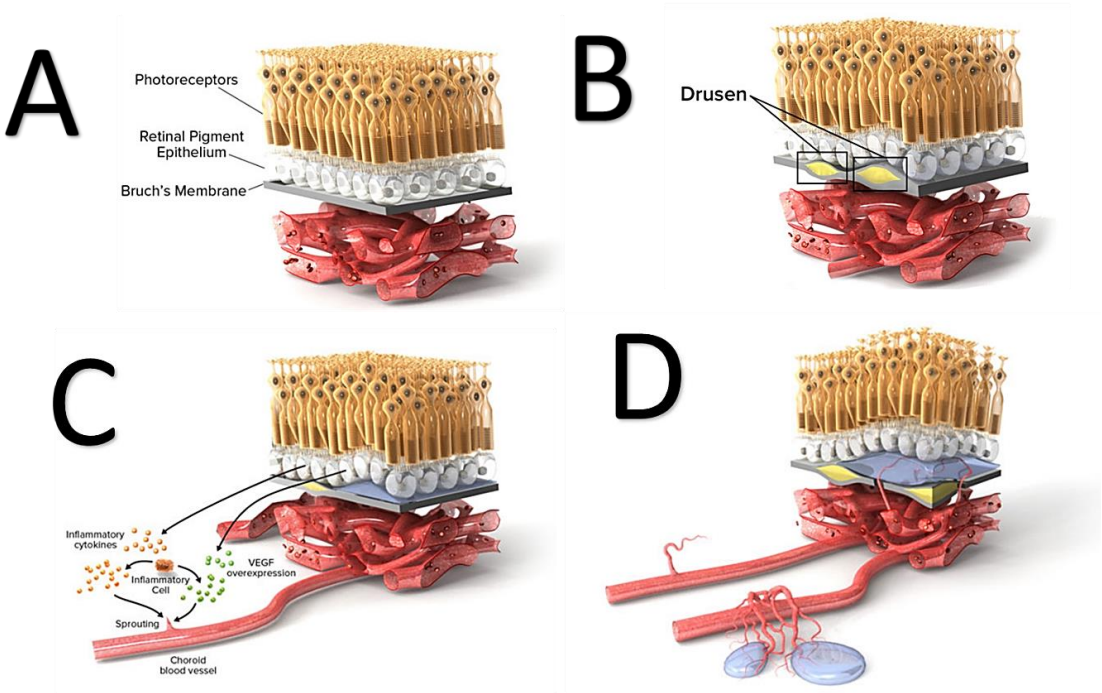
Clinical features and diagnosis of age-related macular degeneration cases and controls

In studies that follow, controls were at least 60 years old with gradable retinal photographs showing an absence of hallmark AMD features. AMD cases were classified by two clinical subtypes termed early and late AMD according to a modified version of the Wisconsin Age-related Maculopathy Scale.(Klein et al., 1991) Early AMD cases were at least 60 years of age and included participants with: 1) soft drusen, 2) depigmentation of the retinal pigment epithelium in the presence of soft and/or hard drusen, or 3) hyperpigmentation in the presence of soft and/or hard drusen (Figure 4).

Drusen (Figure 4.B) are extracellular deposits that occur between the basement membrane of the RPE and the inner layer of Bruch's membrane. Drusen occur as part of the normal aging process but are also important early signs of AMD pathology. The presence of small hard drusen is observed in elderly individuals over the age of 60 years. In comparison, soft drusen, which mark early pathological changes in the retina, are larger and have soft indistinct edges.

Pigmentary abnormalities of the retina, such as hypopigmentation, result as a consequence of RPE cell loss. The loss of RPE increases the visualization of the underlying choroidal vasculature. Hyperpigmentation

inversely is the result of over proliferation of RPE cells(de Jong, 2006). Patients that present with only soft drusen are less likely to progress to late-stage AMD over the course of 5 years(van Leeuwen et al., 2003) versus those that contain soft/hard drusen in conjunction with pigmentary abnormalities(Klein et al., 1997).



This figure was adapted from The Angiogenesis Foundation: The Science of AMD presentation on January 15th, 2015 at http://www.scienceofamd.org/learn/.

Figure 4: ocular tissue overview

A) Diagram of the primary ocular tissues which make up the retina and the retinal support systems in a human eye. B) The 3D structure and placement of a drusen deposit located between the Bruch's membrane and the retinal pigment epithelium layer. Drusen appear as a yellow/white deposit of extracellular waste products containing apoE, CFH, and lipids. C) Accumulation of drusen deposits in the subretinal space can trigger an inflammatory response and recruitment of inflammatory cytokines which in turn can trigger over expression of VEGF. D) Over expression of VEGF protein triggers angiogenesis and the formation and growth of new blood vessels into Bruch's membrane. These fragile new vessels can easily break and leak blood and plasma fluids into the subretinal space.

Late AMD cases included individuals 60 years of age and older with geographic atrophy, sub-retinal hemorrhage, sub-retinal fibrous scarring, or sensory serous sub-retinal detachments. Geographic atrophy is

the hallmark of advanced dry AMD, which begins with demarcated areas of hypopigmentation situated around severe regions of atrophy whereby large choroidal blood vessels are visible. Neovascular AMD occurs, in part, as the result of overexpression of the vascular endothelial growth factor (VEGF) protein, which signals a cascade response for the growth and proliferation of new blood vessel growth (Figure 4. C)(Rudolf et al., 2008). Fragile growth of new blood vessels in the choroid can lead to inflammation and disposition of the RPE and damage to the macula.

Epidemiology of age-related macular degeneration

The epidemiology of AMD risk involves several modifiable and non-modifiable factors. One recently described modifiable risk factor is HDL-C. The relationship between serum HDL levels and AMD risk, however, is inconsistent in the literature.(van Leeuwen et al., 2004; Reynolds et al., 2010) Reynolds and colleagues found a protective effect of higher HDL-C levels and AMD. Conversely, the Rotterdam Study found that higher levels of HDL-C increased the risk (OR=1.20 per standard deviation increase) of incident AMD.(van Leeuwen et al., 2004) Other lipid trait studies have observed a greater level of lipoproteins in AMD patients along with an increase in serum C-reactive protein(Colak et al., 2011) compared to age-matched controls. The exact role that lipids play in AMD pathology is still unknown; however, genetic association studies have found evidence of a correlation between lipid-trait genes and AMD risk. These studies have led to increased efforts to study the effect of lipid-lowering drugs (e.g. statins) on the progression and treatment of AMD with conflicting results(Gehlbach et al., 2009; Shalev et al., 2011). While there is currently a lack of strong evidence to suggest that statins can prevent AMD, statin use in one study appears to protect against soft drusen, which is a precursor to late AMD(Tan et al., 2007a).

Genetics of age-related macular degeneration

Substantial progress has been made in identifying susceptibility variants for AMD (Table 2). The most widely replicated loci are Complement Factor H (*CFH*) and the Age-related Maculopathy Susceptibility-2 (*ARMS2*)/*HTRA1* complex. (Edwards et al., 2005; Haines et al., 2005; Jakobsdottir et al., 2005; Klein et al., 2005; Rivera et al., 2005) Multiple polymorphisms within the chromosome 10q26 region have been

proposed as the functional variation including *ARMS2* A69S, the *HTRA1* promoter variant rs11200638 adjacent to *ARMS2*,(Dewan et al., 2006; Yang et al., 2006) and a complex insertion/deletion variant in the untranslated region (UTR) of *ARMS2*.(Fritsche et al., 2008; Wang et al., 2010)

Lipid levels and lipid metabolism have been associated with susceptibility and progression of AMD in various populations(Gemmy Cheung et al., 2012; Klein et al., 2007a; van Leeuwen et al., 2004; Tomany et al., 2004). As a result, numerous studies have been conducted to determine if genetic variants associated with lipid levels and lipid metabolism also impact risk of AMD. Indeed, the results of a targeted genotyping study suggested that the minor (T) allele for *LIPC* rs10468017 is associated with an increase of HDL and a subsequent lower risk of developing AMD(Reynolds et al., 2010). Recently, a meta-analysis(Fritsche et al., 2013) of over 17,000 advanced AMD cases confirmed the findings of genetic variants previously associated with HDL cholesterol on susceptibility to AMD that had been observed in two previous GWAS(Chen et al., 2010; Neale et al., 2010).

The association between lipid levels (and the genetic variants associated with these levels) makes biological sense. For example, genetic variants in the cholesteryl ester transfer protein (*CETP*) have been consistently associated with HDL-C levels in multiple populations(Dumitrescu et al., 2011; Kathiresan et al., 2009; Willer et al., 2008a). The product of *CETP* is a plasma glycoprotein involved in the reverse cholesterol transport pathway known to transport and remove lipoproteins from blood circulation. CETP has been found to cause an increase or decrease in blood lipid levels with corresponding high or low blood levels of CETP(Chang et al., 2011; Ridker et al., 2009). A study utilizing monkey retinas has shown that CETP localizes to the lipid rich, photoreceptor outer segment (POS)(Tserentsoodol et al., 2006). POS are recycled by the RPE several thousands of times a day to replenish vital components for the creation of new photoreceptor cells in a process that maintains visual integrity(Ebrahimi and Handa, 2011). Disruption in this process is hypothesized to leads to a buildup of lipid peroxidation products that ultimately leads to apoptosis of RPE and release of inflammatory factors into the Bruch's membrane where inflammation leads to the formation of drusen.

Studies of the genetic architecture of AMD have progressed rapidly in European-descent and Asian populations with much success in identifying GWAS-significant associations (Table 2). However, accounting for intra- and inter-population determinants of AMD are necessary for better understanding the pathophysiology of the condition and for targeted treatment options. Furthermore, the modifying effect of environmental factors and the role that gene x environment interactions play has not been fully realized. Genetic association studies, such as the ones outlined in this chapter can help to identify these differences and could further our understanding of the complex interplay between genes and the environment.

Table 2: Variants published in the scientific literature that have been associated with AMD.

SNP	CHR	Closest Gene	Population	PMID
rs10033900	4	*CFI*	European	20385826
rs10468017	15	*LIPC*	European	20385826
rs10490924	10	*ARMS2*	European Asian	20385826 23455636
rs1061170	1	*CFH*	European	20385826
rs10737680	1	*CFH*	European	20385819
rs11200638	10	*HTRA1*	Asian	17053108
rs11755724	6	*RREB1*	European	20385826
rs13081855	3	*COL8A1*	European Asian	23455636
rs13095226	3	*COL8A1*	European	20385826
rs13278062	8	*TNFRSF10A*	European Asian	21909106 23455636
rs1329424	1	*KCNT2*	European Asian	20385819 23326517
rs1410996	1	*CFH*	European	20385826
rs1864163	16	*CETP*	European Asian	23455636
rs2230199	19	*C3*	European Asian	20385819 20385826
rs2285714	4	*CFI*	European	20385819
rs3130783	6	*TRNAI25*	European Asian	23455636
rs334353	9	*TGFBR1*	European Asian	23455636
rs3764261	16	*CETP*	European	20385819
rs3793917	10	*ARMS2*	European Asian	20385819 23326517
rs380390	1	*CFH*	European	15761122
rs429608	6	*CFB*	European Asian	20385819 23455636

rs4420638	19	*APOE*	European Asian	23455636
rs4698775	4	*CFI*	European Asian	23455636
rs493258	15	*RPL28P4*	European	20385819
rs5749482	22	*TIMP3*	European Asian	23455636
rs641153	6	*CFB*	European	20385826
rs8135665	22	*SLC16AB*	European Asian	23455636
rs920915	15	*LIPC*	European Asian	23455636
rs9380272	6	*C2*	European	20385819
rs943080	6	*VEGFA*	European Asian	23455636
rs8017304	14	*RAD51B*	European Asian	23455636
rs6795735	3	*ADAMTS9*	European Asian	23455636
rs3812111	6	*COL10A1*	European Asian	23455636
rs9542236	13	*B3GALTL*	European Asian	23455636

Variants are listed for SNPs found to be associated with AMD and deposited in the NHGRI GWAS catalog as of January 1st, 2015.

Meta-analysis of known and novel age-related macular degeneration loci in diverse /racial ethnic populations

The prevalence of AMD varies across ancestral populations. I hypothesize that this is due to differences in genetic architecture and environmental exposures across these diverse populations. I will statistically ascertain the genetic architecture of AMD in non-European populations beginning with replicating and/or generalizing previously identified GWAS variants discovered primarily in European Americans to the minority populations ascertained by EAGLE. Results between European Americans and the diverse populations will be compared to determine if trends/differences in risk and allele frequencies exist between the various populations.

The Population Architecture using Genomics and Epidemiology (PAGE) I study is a consortium of four projects representing eight population-based studies in different regions of the United States. It was established by the National Human Genome Research Institute (NHGRI) to address a deficiency in genetic research in ethnically diverse populations. The PAGE initiative is to study the genetic architecture of common, complex diseases across diverse populations, identify genetic and environmental modifiers of disease, and to study associations within novel phenotypes. With more than 121,000 DNA samples from African American, Hispanic, Native American, European American, and other populations for use in large scale meta-analyses of several phenotypes, PAGE I is uniquely poised to address the questions of replication and generalization of genotype-phenotype associations in ethnic minorities(Matise et al., 2011a).

Methods

Study Populations

Participants from three PAGE study sites are included in this study: the Atherosclerosis Risk in Communities (ARIC) study, the Cardiovascular Health Study (CHS), and Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study accessing the Third National Health and Nutrition Examination Survey (NHANES III). In addition to data from the PAGE study sites, additional data are presented from the Singapore Prospective Study Programme (SP2) and the Singapore Malay Eye Study (SiMES). Brief descriptions of each study are given below in Table 3. The original phenotypic focus of each study varies from cardiovascular traits (ARIC and CHS), type 2 diabetes and cardiovascular diseases (SP2), ocular traits (SiMES), and a representative sample of the United States regardless of health status (EAGLE).

The Atherosclerosis Risk in Communities Study (ARIC) is a population-based cohort study that included 15,792 women and men between 45 and 64 years of age at recruitment in 1987 through 1989.(1989) The participants were selected by probability sampling from four U.S. communities: suburbs of Minneapolis, Minnesota; Washington County, Maryland; Jackson, Mississippi; and Forsyth County, North Carolina. Retinal photographs were taken at the third visit, allowing inclusion of a subset of participants in this study of AMD.(Klein R et al., 1999) A digitized, 45° color fundus photograph was taken of one eye from participants between 48-72 years of age. AMD was graded according to a modified version of the Wisconsin Age-Related Maculopathy Grading System.(Klein et al., 1991) Current smoking was defined by "Do you now smoke cigarettes?" After the institutional review board at every participating university approved the ARIC Study protocol, written informed consent was obtained from each participant.

The Cardiovascular Health Study (CHS) is a population-based longitudinal study of risk factors for cardiovascular disease in adults 65 years of age or older, recruited at four field centers (Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; Pittsburgh, Pennsylvania).(Fried et al., 1991) Overall, 5,201 predominantly white individuals were recruited in 1989-

1990 from random samples of Medicare eligibility lists, followed by an additional 687 African Americans recruited in 1992-1993. Starting in 1989 participants underwent standardized clinical exams at enrollment, which included blood pressure, lipid profiles, echocardiography of the heart, carotid ultrasound, cranial magnetic-resonance imaging (MRI), and fundus photography. Retinal photographs were taken with a 45-degree nonmydriatic camera of one randomly selected eye from participants.(Klein et al., 2003b; Klein R et al., 1999) AMD status was graded according to the modified Wisconsin Age-relate Maculopathy classification scheme.(Klein et al., 1991) The main outcomes are coronary heart disease (CHD), angina, heart failure (HF), stroke, transient ischemic attack (TIA), claudication, and mortality. Current smoking status was self-reported at baseline.

The National Health and Nutrition Examination Surveys (NHANES), conducted by the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC), are U.S. population-based, cross-sectional surveys collected without regard to health status and include detailed demographic, health, lifestyle, laboratory, clinical, and physical examination data for study participants. NHANES III was conducted in two phases between 1988-1994, and DNA samples were collected in the second phase (1991-1994).(2004, 1996) Genetic NHANES III consists of 7,159 DNA samples, and the method of collection has been previously described.(Chang et al., 2009; Steinberg et al., 1997) We used study participant data from NHANES III, of which 3,131 had available fundus photographs and laboratory measurements of serum cotinine (ng/mL). Participants over the age of 40 were selected to have a non-stereoscopic, 45° color fundus photograph taken of one randomly selected eye. AMD was graded according to the Wisconsin Age-Related Maculopathy Grading System.(Klein et al., 1991) Current smokers were defined using "yes" to the question "do you smoke cigarettes now?" or cotinine levels > 15ng/ml. All procedures were approved by the CDC Ethics Review Board and written informed consent was obtained from all participants. Because no identifying information is available to the investigators, Vanderbilt University's Institutional Review Board determined that this study met the criteria of "non-human subjects."

The Singapore Prospective Study Programme (SP2) is a population-based cohort consisting of Singaporean Chinese participants aged between 40 and 80 years.(Nang et al., 2009) Initially, 10,747 subjects were invited from four population-based cross-sectional surveys conducted in Singapore (1982–1998) to participate in a repeat examination from 2004 to 2007. (Cutter et al., 2001; Hughes et al., 1990, 1997; Tan et al., 1999) Participants who were successfully re-contacted and completed a questionnaire (n=7,744; 76.8% response rate) were then invited to attend a clinic health examination that included physicals and ocular assessment, retinal photography, and collection of biologic specimens, of which 5,163 (66.7% of those who completed the questionnaire or 51.2% of all eligible subjects) attended. Retinal photographs were taken of both eyes with a 45° digital retinal camera and are available for 4,110 participants.(Jeganathan et al., 2009) AMD was graded according to the Wisconsin Age-Related Maculopathy Grading System.(Cheung et al., 2012) A structured interviewer-administered questionnaire was used to collect information about smoking status. Current smoking status was self-reported.

The Singapore Malay Eye Study (SiMES) is a population-based cross-sectional study of urban Singaporean Malay adults, conducted to assess prevalence, risk factors, and the public health impact of common age related eye diseases.(Foong et al., 2007) An age-stratified (by 10-year age group) random sample of the Malay population residing in 15 residential districts in Southwestern Singapore age 40 to 80 years was drawn from the computer-generated random list of 16,069 Malay names provided by the Ministry of Home Affairs. Of 4,168 eligible participants, 3,280 (overall response rate 78.7%) participated in the study, conducted from August 2004 through June 2006. Retinal photographs were taken of both eyes in participants with a digital retinal camera and AMD was graded according to the Wisconsin Age-Related Maculopathy Grading System.(Kawasaki et al., 2008) A questionnaire was used to collect information about smoking status, with participants self-reporting current smoking status.

Phenotyping

At all study sites, fundus photographs were graded according to a modified version of the Wisconsin Age-related Maculopathy Scale.(Klein et al., 1991) AMD cases and controls were at least 60 years of age with gradable retinal photographs.

SNP selection and genotyping

As part of the PAGE study, SNPs were selected for genotyping based on previous GWAS and candidate gene studies for a variety of common human diseases and traits including AMD, lipid-traits (HDL-C, LDL-C, and triglycerides), body mass index/obesity, type 2 diabetes, hypertension, and inflammation (C-reactive protein), to name a few. For this study, we included both AMD and lipid-trait associated SNPs reported in genome-wide association studies or the National Human Genome Research Institute's GWAS Catalog as of 2009(Welter et al., 2014). Previously, a total of 57 SNPs were selected for analysis and included the two primary AMD variants (*CFH* rs1061170 and *ARMS2* rs10490924), other variants involved in the complement system pathway (*CFH* rs800292, rs1065489, rs3753394, rs3766404, rs6677604, rs800292; *CFI* rs10033900, rs11726949; *C2* rs547154), and variants associated with lipid-related traits (Supplementary Table 1). As described in Matise et al, genotyping in the PAGE Study was performed at each PAGE study site independently using a variety of genotyping assays and platforms.(Matise et al., 2011a) Not all 57 SNPs were genotyped in each study site (Supplementary Table 2). For quality control, all PAGE study sites genotyped 360 DNA samples (CEU, YRI, CHB, JPN, and MEX) from the International HapMap Project, including 77 parent-child trios.(2003)

In ARIC, the *CFH* Y402H variant (rs1061170) was previously genotyped in a candidate gene study using the TaqMan assay and polymerase chain reaction amplification (Applied Biosystems, Foster City, CA).(Volcik et al., 2008) Genotypes were called using the ABI 7900HT and the Sequence Detection System software (Foster City, CA). DNA genotyping for the *ARMS2* A69S variant (rs10490924) was part of a GWAS using the Affymetrix GeneChip SNP Array 6.0 (Santa Clara, CA).(Psaty et al., 2009) The HDL-associated SNPs used in this study were genotyped as previously described.(Dumitrescu et al., 2011)

In CHS, *CFH* Y402H and *ARMS2* were not genotyped. Genotyping data for other variants were obtained from two sources. First, SNP genotyping was conducted in the Houston central lab using TaqMan (see details above for ARIC). The second source of genotyping data from CHS was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system. Genotypes were called using the Illumina BeadStudio software.(Psaty et al., 2009)

In EAGLE, *CFH* Y402H (rs1061170) and *ARMS2* A69S (rs10490924) were genotyped by the Center for Human Genetics Research DNA Resources Core in NHANES III using Sequenom iPLEX® Gold assay (San Diego, California) according to the manufacturer's instructions. Blinded duplicates were genotyped as required by CDC, and both SNPs passed quality control metrics required by CDC. The following lipid trait-associated SNP data were accessed from existing genotype data in Genetic NHANES, a collection of DNA samples collected from the NHANES III and Continuous NHANES (1999+) surveys: rs3890182 (*ABCA1*), rs3135506 (*APOA5*), rs1800775 (*CETP*), rs1323432 (*GRIN3A*), rs1800588 (*LIPC*), and rs328 (*LPL*).(Keebler et al., 2009) The remaining lipid trait-associated SNPs were genotyped using either Sequenom or the Illumina BeadXpress. All genotype data reported here were deposited into the NHANES III Genetic database and are available for secondary analysis through CDC.

In SP2, Chinese samples were previously genotyped as controls for a type 2 diabetes case control study which used either the Illumina HumanHap 610 Quad or 1MDuo-v3 BeadChips and as part of a psoriasis case control study which used the Illumina HumanHap 550 BeadChip. Genotyping details have been published elsewhere.(Sim et al., 2011)

SiMES genotyping methods followed those of the SP2 cohort as previously described.(Sim et al., 2011)

Genotyping for any given SNP, in a particular population, was not always available across all study sites. Therefore we provide the following maximum and minimum ranges which include cases and controls for

genotyping data by population: European Americans (N = 4546 – 6540), African Americans (N = 796 – 1267), Mexican Americans (N = 275 – 317), and Asians (N = 972 – 1197).

Statistical methods

Each study site performed tests of association locally using a common analysis protocol prior to meta-analysis. Early and late cases of AMD were combined for analyses to increase power. Each genetic variant was tested for association with AMD using logistic regression assuming an additive genetic model stratified by self-described race/ethnicity (e.g. European American, African American, Mexican American, and Asian). SNPs available for analysis by study site are given in Supplementary Table 2. All models were adjusted for site of ascertainment (Model 1, minimally adjusted). Models 2 and 3 were adjusted for age, sex, body mass index (BMI), smoking status (current versus ever/never), and high density lipoprotein (HDL) cholesterol, fasting (≥8 hours) (Model 2) or regardless of fasting status (Model 3). Participants on lipid-lowering medications were excluded in both Models 2 and 3. Asians were represented by the SP2 cohort only in Model 2 given the lack of fasting HDL-C in the SiMES cohort.

Local analyses were conducted using Stata 9.0 (ARIC), SAS v9.3 (CHS), SAS v9.2 (by EAGLE using the Analytic Data Research by Email (ANDRE) portal of the CDC Research Data Center in Hyattsville, MD), and PLINK v1.06 (SP2/SiMES).

After analyses were conducted locally at each site, meta-analyses using summary statistics were carried out in the Crawford lab by Robert Goodloe using a fixed-effects, inverse-variance weighted approach implemented in METAL.(Willer et al., 2010) Between study heterogeneity was tested for in METAL with the Cochran's *Q*-test. Genomic control was applied on the combined meta-analysis and not each specific cohort analysis to avoid overcorrection of population stratification. Meta-analysis results were plotted using Synthesis-View.(Pendergrass et al., 2010)

All p-values listed in the manuscript and tables are uncorrected for multiple hypothesis testing. The following Bonferroni corrected thresholds were calculated by population based on the number of available

SNPs in that population across all study sites regardless of whether the SNP was available at each study site: European Americans (0.05/49 SNPs) = 0.0010; African Americans (0.05/47 SNPs) = 0.0010; Mexican Americans (0.05/31 SNPs) = 0.0016; and Asians (0.05/43 SNPs) = 0.0012.

Ethics statement

This research adhered to the tenets of the Declaration of Helsinki. Approval for the study was obtained from the appropriate institutional review boards at all participating institutions, and all study participants gave informed consent where appropriate. All work for studies conducted in the United States was Health Insurance Portability and Accountability Act - compliant.

Results

Population characteristics

The meta-analysis is composed of multiple studies within the PAGE study (ARIC, CHS, and EAGLE) and Asian cohorts comprised of SP2 and SiMES. Combined, this meta-analysis includes various racial/ethnic groups: European Americans (830 cases and 5,710 controls), African Americans (95 cases and 1,172 controls), Mexican Americans (47 cases and 270 controls), Singaporean Chinese (21 cases and 206 controls), and Singaporean Malays (107 cases and 863 controls). All study sites ascertained both men and women. Smoking status was defined as current vs. ever/never smoker and varied across studies (Table 3). The average BMI was greater than 25 kg/m^2 across all studies regardless of case status, except in the SP2 dataset.

Table 3: Description of study participants, by study site

	ARIC				CHS				EAGLE						SP2		SiMES	
	European American		African American		European American		African American		European American		African American		Mexican American		Asian		Asian	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
N	289	3128	34	582	351	1918	31	381	190	664	30	209	47	270	21	206	107	863
Age	65.1 (3.1)	64.3 (3.0)	63.9 (3.1)	64.0 (3.1)	72.4 (4.6)	70.7 (4.2)	72.4 (4.4)	71.5 (4.7)	76.4 (8.1)	71.2 (7.5)	70.2 (7.5)	68.5 (6.9)	69.3 (7.0)	67.2 (6.3)	69.3 (5.3)	67.1 (4.9)	71.3 (5.4)	69.0 (5.6)
Body Mass Index	27.8 (4.6)	27.7 (4.8)	29.1 (3.9)	29.8 (5.9)	26.1 (4.2)	26.5 (4.2)	28.9 (6.0)	28.7 (4.8)	26.8 (5.7)	27.1 (4.7)	31.0 (6.3)	28.2 (6.2)	28.6 (5.3)	28.4 (5.3)	21.3 (3.3)	23.2 (3.1)	25.5 (5.8)	26.2 (5.0)
Total Fasting Cholesterol	209.6 (35.0); n=249	208.8 (36.7); n=2647	210.9 (31.6); n=30	209.3 (38.4); n=501	210.0 (36.7)	213 (39.0)	198.9 (29.2)	210.1 (35.9)	230.0 (79.5)	228.6 (76.0)	235.7 (56.2)	229.9 (104.5)	256.6 (145.9)	222.1 (70.9)	197.2 (30.9)	208.8 (34.8)	220.4 (42.5)*	224.2 (46.4)*
Fasting HDL Cholesterol	50.0 (16.5)	51.5 (18.4)	53.1 (20.0)	55.4 (18.1)	56.0 (16.6)	54.3 (15.4)	59.4 (14.4)	58.8 (15.0)	52.7 (15.5)	66.7 (113.9)	63.2 (19.5)	71.2 (116.8)	84.6 (180.0)	53.9 (71.6)	54.1 (11.6)	54.1 (11.6)	54.1 (11.6)*	50.2 (11.6)*
% female	49.1	51.3	44.1	63.6	57.0	60.0	74.2	65.1	66.8	54.5	60.0	47.4	44.7	44.1	19.05	46.60	29.91	51.68
% current smokers	61.6	60.1	58.8	50.5	9.6	8.7	12.9	13.3	44.2	52.0	50.0	58.9	53.2	54.8	38.10	22.82	48.57	36.74

Atherosclerosis Risk in Communities (ARIC) Study; Cardiovascular Health Study (CHS); Epidemiologic Architecture for Genes Linked to Environment (EAGLE); Singapore Prospective Study Programme (SP2); Singapore Malay Eye Study (SiMES)

Means and standard deviations (SD) are given unless otherwise noted.

* Study site did not collect fasting data; statistics are shown for non-fasting data.

Replication and generalization of known age-related macular degeneration risk variants

We meta-analyzed tests of associations for up to 19 SNPs previously associated with AMD. Among European Americans, 13 AMD SNPs were tested, and 7/13 (54%) were significant at an uncorrected p< 0.05 in Model 1 (Table 4 and Supplementary Table 1; Figure 6). As expected, both *CFH* rs1061170 (OR=1.55; $p=3.05x10^{-8}$) and *ARMS2* rs10490924 (OR=1.55; $p=6.36x10^{-6}$) were strongly associated with AMD risk (Table 4) at a Bonferroni corrected p-value (p< 0.001). The genetic effect sizes estimated here were smaller for these variants compared with previous reports with other studies estimating risk between 2.7 and 4.6 for heterozygotes (Figure 5).(Despriet et al., 2006; Haines et al., 2005; Klein et al., 2005; Yu et al., 2011) Three additional *CFH* SNPs were associated with AMD risk at p<0.05 in European Americans: missense rs800292 (OR=0.58; $p=3.80x10^{-5}$), intergenic rs3753394 (OR=1.25; p=0.03), and intronic rs6677604 (OR=0.77; p=0.04). The remainder of the Complement Factor SNPs that were tested failed to replicate at this liberal significance threshold. Two lipid-related variants previously associated with AMD(Chen et al., 2010; Neale et al., 2010; Yu et al., 2011) were nominally associated with AMD in European Americans: *CETP* rs3764261 (OR=1.14; p=0.04) and *ABCA1* rs1883025 (OR=0.82; p=0.03). After adjustment for age, sex, BMI, smoking status, and HDL-C (Model 3), only *CFH* rs1061170, *ARMS2*

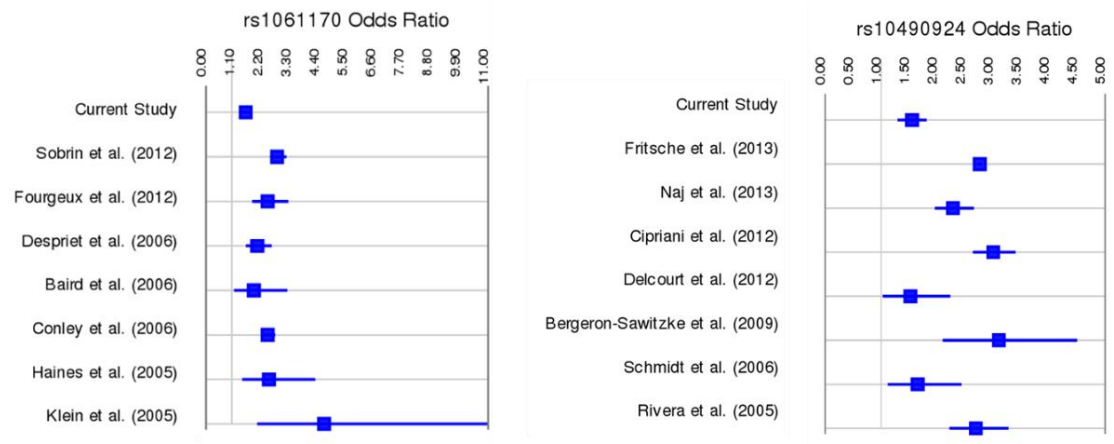


Figure 5: Forest plots with odds ratios and confidence intervals for previously reported associations. *CFH* rs1061170 and *ARMS2* rs10490924 are compared to results for current study for European-descent populations.

rs10490924, and *CFH* rs800292 remained significant at a Bonferroni corrected p-value (Supplementary Table 3; Supplementary Figure 1).

Table 4: Significant AMD meta-analysis association results.

Each study site performed tests of association using logistic regression assuming an additive genetic model. Data were meta-analyzed using a fixed-effects inverse-variance weighted approach. Results are shown for nominally significant tests at an uncorrected threshold (p<0.05) adjusted for site of ascertainment (Model 1).

Rsid	Gene	Chr	OR (95% CI)	Direction of Effect *	P-value	Coded Allele	CAF	Race/Ethnicity
rs1061170	*CFH*	1	1.55 (1.34 – 1.78)	+,.,+	**3.05×10^{-8}**	C	0.37	European American
			1.36 (0.91 – 2.04)	+,.,+	0.13		0.38	African American
			1.22 (0.70 – 2.13)	~~A~~	0.47		0.20	Mexican American
			4.43 (0.35 – 55.29)	~~A~~	0.24		0.04	Asian
rs800292	*CFH*	1	0.59 (0.47 – 0.74)	~~A~~	**3.80×10^{-5}**	A	0.24	European American
			0.55 (0.29 – 1.03)	~~A~~	0.06		0.69	African American
			0.89 (0.68 – 1.15)	-,-,-,+	0.36		0.42	Asian
rs3753394	*CFH*	1	1.25 (1.03 – 1.51)	~~A~~	0.03	T	0.29	European American
			1.45 (0.60 – 3.53)	~~A~~	0.41		0.08	African American
			0.97 (0.75 – 1.25)	+,-,+,-	0.80		0.54	Asian
rs6677604	*CFH*	1	0.77 (0.61 – 0.97)	~~A~~	0.04	A	0.22	European American
			0.63 (0.31 – 1.28)	~~A~~	0.20		0.37	African American
			0.97 (0.54 – 1.75)	-,-,+,+	0.92		0.05	Asian
rs328	*LPL*	8	0.95 (0.79 – 1.14)	-,-,+	0.60	G	0.10	European American
			1.75 (1.06 – 2.91)	+,+,+	0.03		0.07	African American
			1.30 (0.55 – 3.08)	~~A~~	0.53		<0.01	Mexican American
rs6987702	*TRIB1*	8	1.01 (0.87 – 1.17)	-,+,.	0.93	T	0.73	European American
			1.61 (1.03 – 2.52)	+,+,.	0.04		0.29	African American
			1.20 (0.90 – 1.58	+,+,+,+	0.21		0.43	Asian
rs1883025	*ABCA1*	9	0.82 (0.69 – 0.96)	-,.,-	0.03	A	0.26	European American
			0.87 (0.57 – 1.33)	-,.,-	0.52		0.35	African American
			0.93 (0.57 – 1.54)	~~A~~	0.78		<0.01	Mexican American
			0.40 (0.08 – 1.91)	.,-,.,.	0.25		0.23	Asian
rs10490924	*ARMS2*	10	1.55 (1.29 – 1.81)	+,.,+	**6.36×10^{-6}**	T	0.22	European American
			0.83 (0.51 – 1.33)	+,.,-	0.43		0.24	African American

			1.63 (1.02 – 2.60)	~~A~~	0.04		0.26	Mexican American
rs2338104	*KCTD10*	12	1.05 (0.92 – 1.20)	.,-,+	0.50	G	0.53	European American
			0.86 (0.61 – 1.22)	-,-,+	0.40		0.74	African American
			1.69 (1.08 – 2.64)	~~A~~	0.02		0.45	Mexican American
rs261332	*LIPC*	15	0.77 (0.60 – 0.98)	~~A~~	0.05	A	0.20	European American
			1.11 (0.57 – 2.18)	~~A~~	0.75		0.25	African American
			0.39 (0.05 – 3.45)	.,-,.,.	0.40		0.12	Asian
rs1800775	*CETP*	16	1.00 (0.87 – 1.16)	-,.,+	0.98	C	0.52	European American
			1.57 (1.03 – 2.38)	+,.,+	0.04		0.42	African American
			1.13 (0.70 – 1.83)	~~A~~	0.59		0.46	Mexican American
			1.03 (0.79 – 1.34)	+,+,+,-	0.84		0.51	Asian
rs3764261	*CETP*	16	1.14 (1.01 – 1.28)	+,+,-	0.04	T	0.33	European American
			0.99 (0.71 – 1.40)	-,+,-	0.96		0.33	African American
			0.78 (0.46 – 1.32)	~~A~~	0.37		0.31	Mexican American
			1.07 (0.74 – 1.53)	+,-,+,+	0.73		0.17	Asian

* Direction of effect is given for ARIC, CHS, and EAGLE for EA, AA, and MA if data are available. Direction of effect is given for Asians for SiMES and SP2 1M, 550, and 610 platforms. Otherwise, the study site is set to missing (".").
~~A~~ Only a single study site is represented.
Bold p-values are those that met strict Bonferroni correction

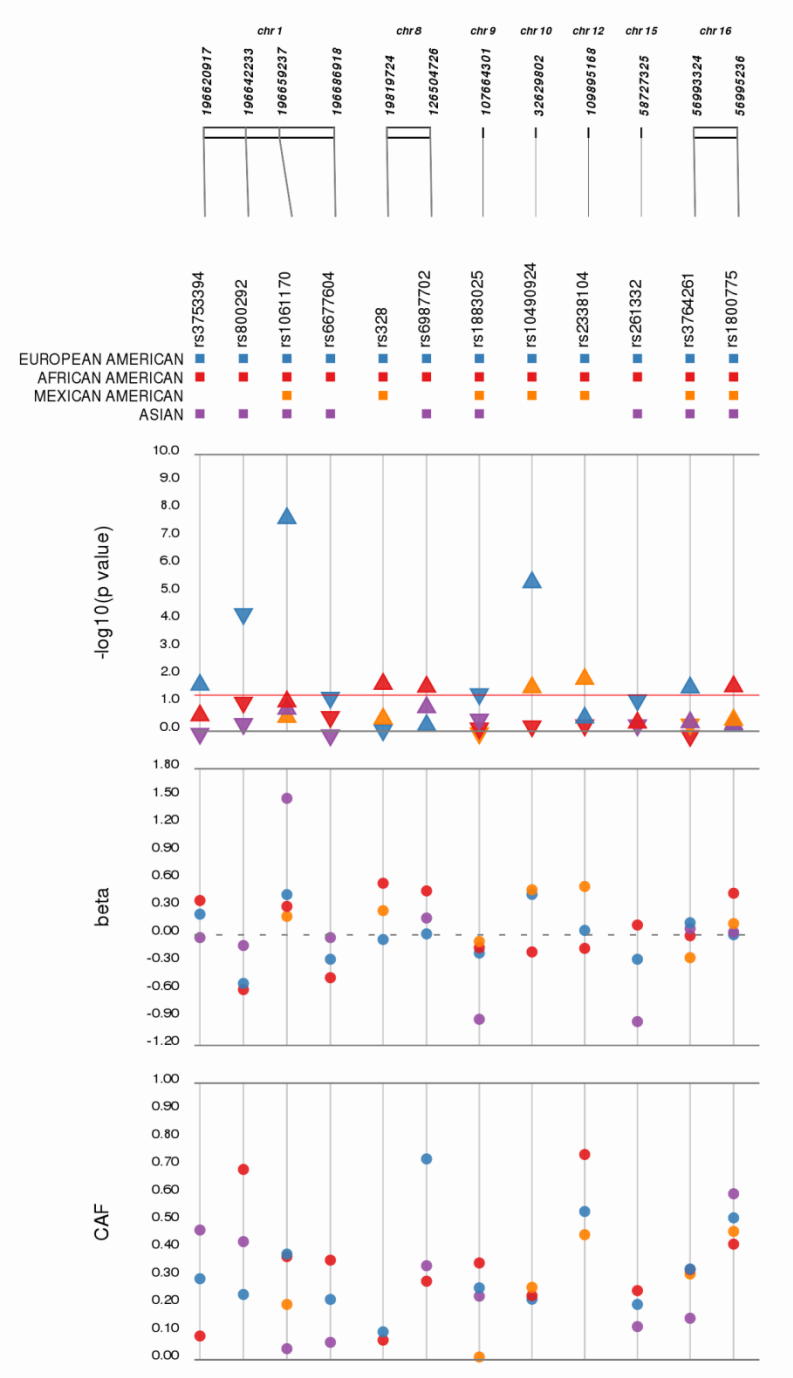


Figure 6: Meta-analysis association results

Synthesis view plot of nominally significant (p < 0.05) meta-analysis association results for Model 1 which is minimally adjusted only for site of ascertainment, for all race/ethnicities. P- values are represented by the colored arrows and are transformed by the –log10, with the threshold of p = 0.05 marked by the red line. Colored arrows also show the direction of effect (beta). P-values, beta's, and coded allele frequencies (CAF) are plotted by race/ethnicity.

Generalization of previously associated AMD variants to diverse populations

We consider a previously associated AMD variant to have generalized if the same variant was associated in a different population with the same direction of effect as observed in the original population (in this case, European-descent populations). For Mexican Americans, only one study contributed data toward generalization: EAGLE accessing NHANES III, which included 47 cases and 270 controls. Among tests of association for variants previously associated with AMD, 1/8 (12%) in Model 1 was significant at an uncorrected $p<0.05$ in Mexican Americans. The association between *ARMS2* rs10490924 (OR=1.63; p=0.04) and AMD in Mexican Americans has been previously reported for EAGLE accessing NHANES III.(Spencer et al., 2012a) The association between *ARMS2* rs10490924 and AMD in Mexican Americans is still significant (p=0.05) in this study after adjustment for Model 3 covariates (Supplementary Table 3). However, after strict correction for multiple testing, none of the SNPs tested in Mexican Americans was associated with AMD.

Among African Americans, none of the 13 previously associated SNPs were associated with AMD in Model 1 in this population at the liberal significance threshold of $p<0.05$. Of note is the test of association between AMD and *ARMS2* rs10490924. A previous report accessing only NHANES III data consisting of 30 cases and 209 controls suggested this variant was marginally associated with AMD in the opposite direction compared with European Americans.(Spencer et al., 2012a) In this meta-analysis, the test of association was expanded to include an additional 34 cases and 582 controls from ARIC. The resulting point estimate of the genetic effect size (OR=0.83; 95% CI: 0.51-1.33) was consistent with the original report by Spencer *et al*(Spencer et al., 2012a), but the test of association was no longer significant (p=0.43).

Similar to African Americans, in Asians none of the previously associated AMD variants were associated with AMD in Model 1 in this population at a liberal significance of $p<0.05$. While *CFH* rs1061170 failed to generalize in terms of statistical significance (p=0.24), the point estimate of the genetic effect (OR=4.43;

95% CI: 0.35-55.29) was in the same direction as the effect sizes observed for European-descent populations (Table 4; Figure 6).

Lipid-associated SNPs and age-related macular degeneration

Given that recent GWAS have highlighted the association of genes traditionally involved in lipid pathways as mediators of AMD risk, we tested an additional 44 lipid-trait variants in these diverse populations for association with AMD. (Fritsche et al., 2013; Neale et al., 2010; Yu et al., 2011) Among the variants associated with lipid traits that were not previously associated with AMD in European-descent populations, only *LIPC* rs261332 (Model 1) was marginally associated with AMD in European Americans at p=0.052 (Table 3; Figure 4). *LIPC* rs261332 (OR=0.77) was only tested in ARIC with a total of 289 cases.

In non-European-descent populations, several lipid-associated SNPs were associated with AMD. Among 95 and 1,172 African American cases and controls, respectively, 3/41 (7%) SNPs that were previously associated with a lipid trait were associated with AMD at the uncorrected threshold of $p<0.05$ in Model 1. These SNPs include *LPL* rs328 (OR=1.75; p=0.03), *TRIB1* rs6987702 (OR=1.61; p=0.04), and *CETP* rs1800775 (OR=1.57; p=0.04) (Table 4 and Supplementary Table 1; Figure 6). In Mexican Americans, 1/29 (3%) lipid-related SNPs tested reached significance at an uncorrected $p<0.05$ (Model 1). This variant was rs2338104 (*KCTD10/MVK)*, previously associated with HDL cholesterol, associated here with AMD (OR=1.69; p=0.02).(Kathiresan et al., 2009; Willer et al., 2008a) The associations observed for *LPL* rs328 and *TRIB1* rs6987702 in African Americans and *KCTD10/MVK* rs2338104 in Mexican Americans remained significant after adjusting for covariates including fasting HDL-C (Supplementary Table 3). Of the 35 lipid-associated SNPs tested in the SP2 and SiMES meta-analysis of Asians, none of the association tests reached significance at an uncorrected $p<0.05$. After strict correction for multiple testing none of the lipid-associated SNPs were associated with AMD in any population.

Discussion

We tested up to 57 SNPs representing 19 previously associated AMD variants and 38 previously associated lipid trait variants for association with AMD in European Americans, African Americans, Mexican Americans, and Asians. At an uncorrected p-value for multiple testing ($p<0.05$), we replicated up to 54% (7/13) of the previously reported associations for AMD tested here in European Americans. Only one previously reported AMD association (*ARMS2* rs10490924) generalized to Mexican Americans (Model 1; Table 4) whereas none generalized to African Americans and Asians. In contrast, several associations were observed between lipid trait-associated variants and AMD in African Americans and Mexican Americans.

Factors that impact generalization of AMD-associated variants

The lack of replication and/or generalization could be due to power. Indeed, sample sizes for non-European descent populations were limited. For the two most strongly associated variants observed in European Americans (rs1061170 *CFH* and rs10490924 *ARMS2*), we had greater than 90% power to detect published effect sizes of 2.41 and 2.94 in African Americans and Mexican Americans, assuming an additive genetic model at a p-value of 0.05.(Yu et al., 2011) Therefore, compared to the effect sizes described in the literature for European Americans (Figure 5), we were generally well powered to detect previously reported associations with these two SNPs in African Americans and Mexican Americans. In the present study, the direction of effect for *CFH* rs1061170 (OR=1.36) in African Americans was the same as that previously reported in European Americans (OR=1.80-4.60; Figure 6); however, the direction of effect for *ARMS2* rs10490924 (OR=0.83) was opposite that of published studies of European Americans. In Mexican Americans, the direction of effect was the same for both variants as that observed in European Americans. The consistent direction of effect observed in Mexican Americans may represent European admixture at or surrounding this genomic region. However, due to the limited genetic data available in the present study, we could not explicitly test this hypothesis. In the Asian population, we were underpowered to detect an association for *CFH* rs1061170 due to the limited number of cases (n=8) genotyped for this particular variant. *ARMS2* rs10490924 was not genotyped in this Asian population. Although our study was powered

to detect published effect sizes in African Americans and Mexican Americans, these previous estimates were based on studies of European-descent individuals and may not be representative of the risk of AMD in diverse populations explored here.

Differences in linkage disequilibrium could also have adversely affected our ability to generalize associations originally identified in European-descent populations. In this study, we only tested the index variants reported in the literature. The index variants are often not the causal or functional variant; rather, the index variant is in linkage disequilibrium (LD) or "tags" the unknown functional variant. Failure to genotype or target the functional or causal variant can potentially reduce power to detect the association if the genotyped variant is not in perfect LD with the true functional or causal variant.

GWAS or fine-mapping data are not available for this study; therefore, LD patterns cannot be directly compared across the study populations described here. However, examination of linkage disequilibrium (LD) patterns in HapMap III for Europeans (CEU), African Americans (ASW), Mexican Americans (MEX), and Han Chinese (CHB) suggests that LD differs in AMD-associated genomic regions across populations as expected.(Gabriel et al., 2002; Jakobsson et al., 2008) For example, in the region of *ARMS2* (rs10490924), Haploview(Barrett et al., 2005) plots of Europeans (CEU) (Figure 7) show rs10490924 in perfect LD (r^2=1) or near perfect LD (r^2>0.90) with four intronic SNPs (rs3750848, rs3750847, rs2284665, rs932275) in an approximately 17 kb region. These SNPs are also in perfect/near perfect LD with rs10490924 in HapMap Mexicans (MEX) (Figure 9). However in HapMap African ancestry in Southwest United States (ASW) (Figure 8), neither rs2284665 nor rs932275 are in perfect LD although they are still correlated to rs10490924 (r^2=0.47 and r^2=0.70, respectively). In HapMap Han Chinese in Beijing (CHB) (Figure 10), rs10490924 is again in perfect LD with rs3750848 and rs3750847 and high LD with rs2284665 and rs932275 (r^2=0.86 and r^2=0.63, respectively). Many other variants upstream of rs10490924 are in moderate LD (r^2 >0.25) in CEU and MEX. There is less evidence of LD present in ASW and CHB. Generalization of rs10490924 in Mexican Americans (Table 4) but not in African Americans in our study

coupled with LD patterns suggest that rs10490924 may not be a major factor behind AMD susceptibility in African Americans.

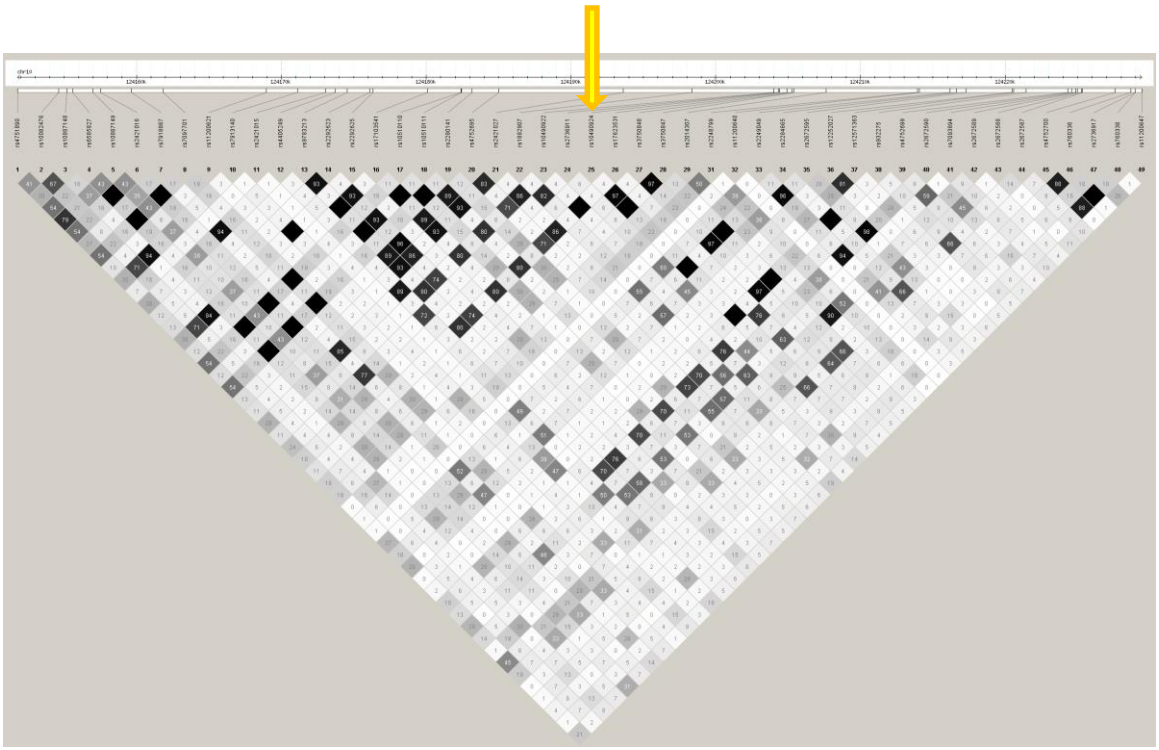


Figure 7: Haploview (v. 4.2) LD plot of Chromosome 10q26 in HapMap III CEU population.
Plot generated for an approximately 80kb window around rs10490924 (positions 124150 through 124230). Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$. The arrow denotes the location of rs10490924.

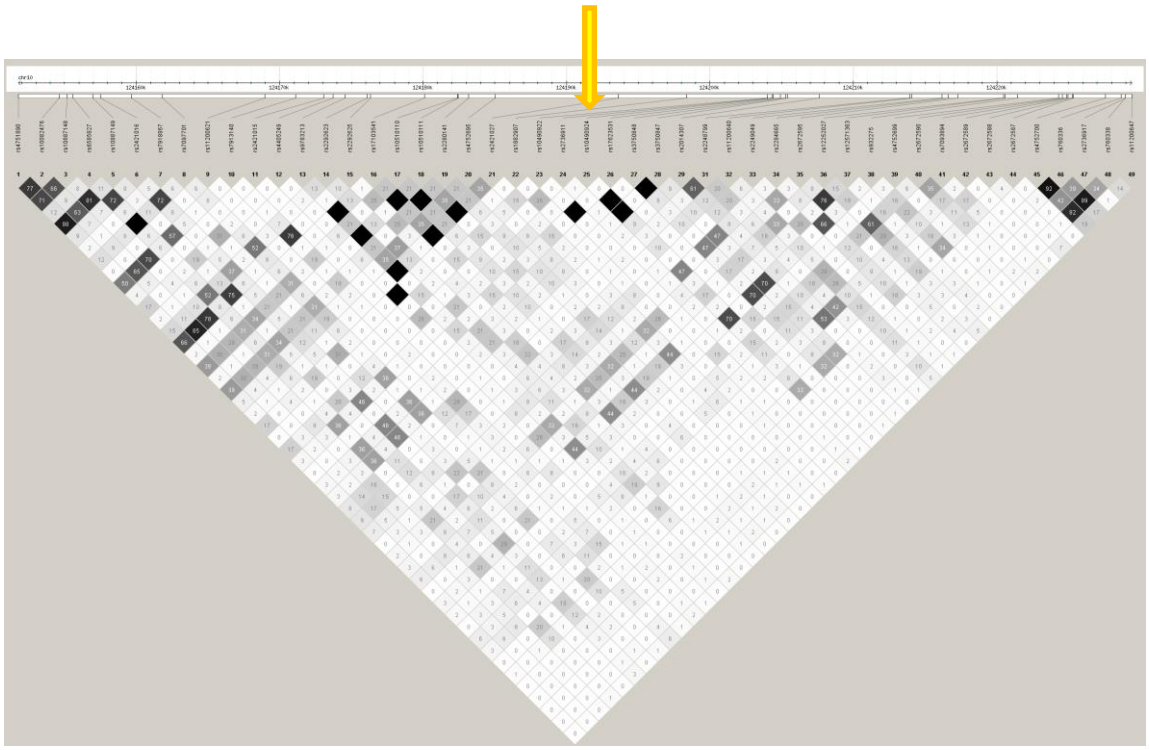


Figure 8: Haploview (v. 4.2) LD plot of Chromosome 10q26 in HapMap III ASW population.

Plot generated for an approximately 80kb window around rs10490924 (positions 124150 through 124230). Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$. The arrow denotes the location of rs10490924.

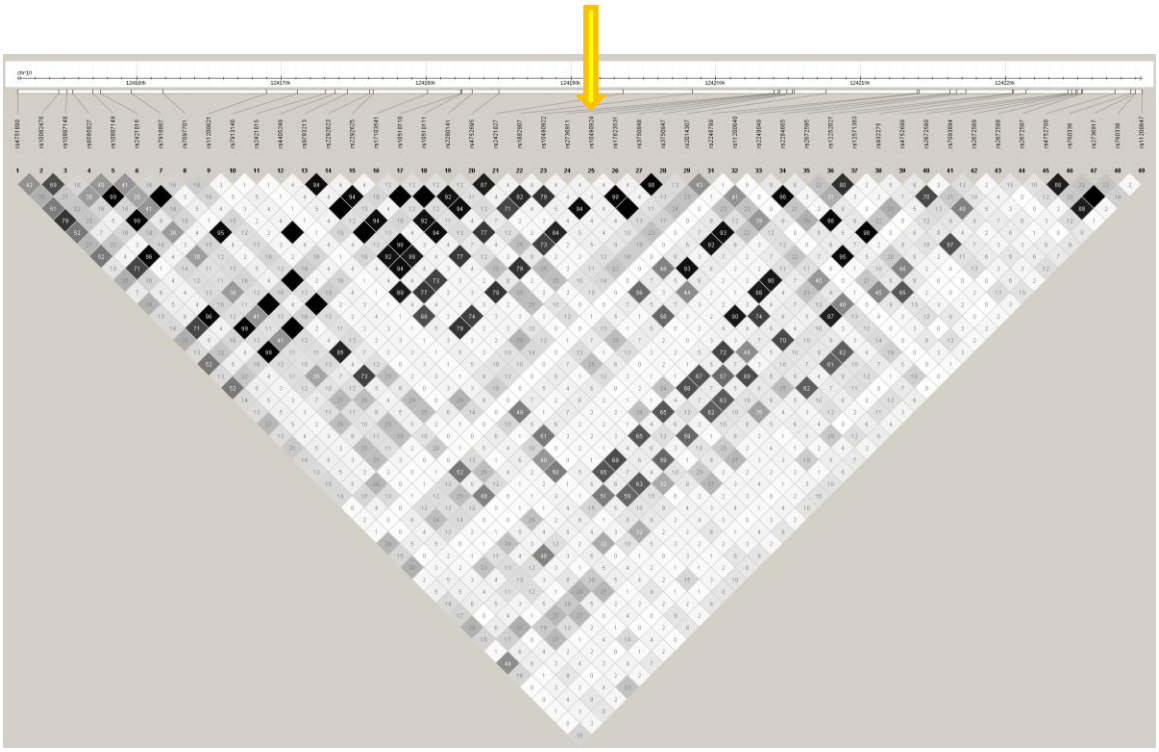


Figure 9: Haploview (v. 4.2) LD plot of Chromosome 10q26 in HapMap III MEX population.

Plot generated for an approximately 80kb window around rs10490924 (positions 124150 through 124230). Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$. The arrow denotes the location of rs10490924.

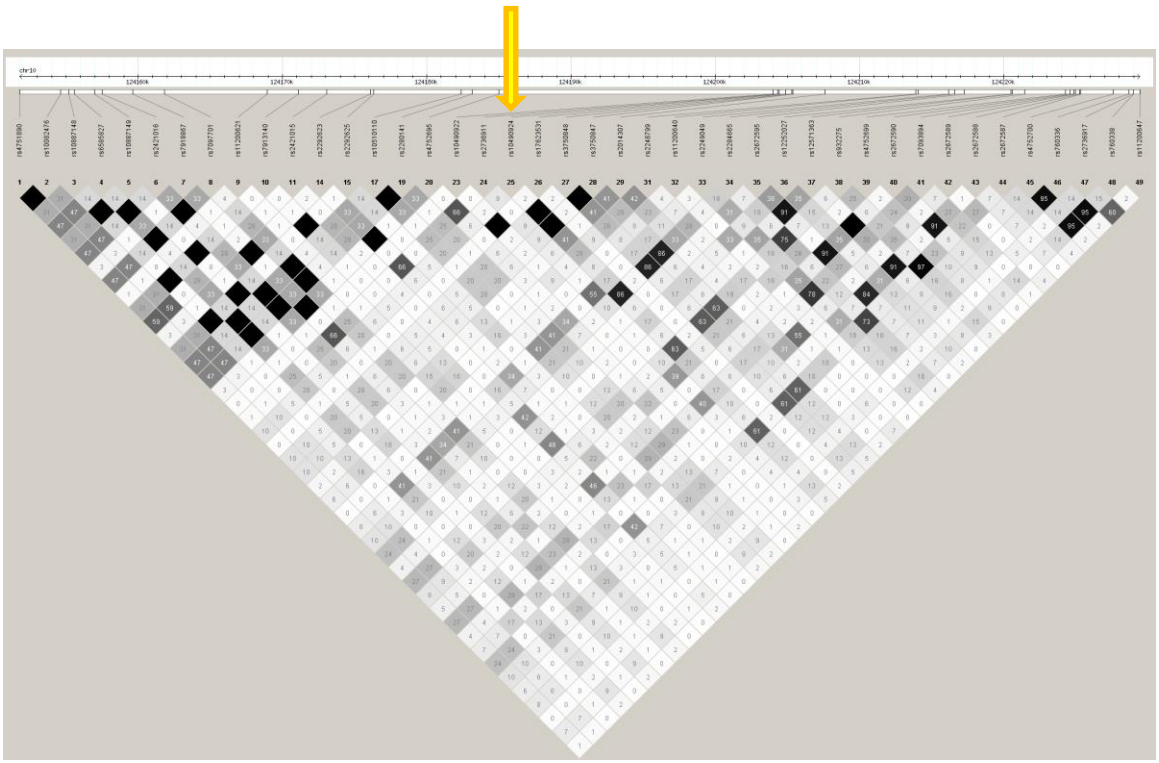


Figure 10: Haploview (v. 4.2) LD plot of Chromosome 10q26 in HapMap III CHB population.
Plot generated for an approximately 80kb window around rs10490924 (positions 124150 through 124230). Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$. The arrow denotes the location of rs10490924.

Power in this study could have also been affected by differences in allele frequencies between populations. In a comprehensive review of estimates of *CFH* Y402H frequency, four studies of *CFH* Y402H in African-descent populations were identified.(Grassi et al., 2006; Volcik et al., 2008; Ziskind et al., 2008) The allele frequency estimates from these four studies range from 0.34 to 0.43, in line with our estimate of 0.38. It is intriguing that despite a high frequency of the risk allele similar to European-descent populations, the prevalence of AMD in African and African-descent populations is much lower.(Klein et al., 2008) This suggests that other genetic and/or environmental factors are at play. Of the four race-ethnicities represented, Mexican-Americans and Asians had the lowest frequency of Y402H (0.20 and 0.04, respectively). The

lower frequency of this variant may account for at least some of the difference in prevalence of AMD in Hispanics compared with European-descent populations.(Klein et al., 2008)

CETP variant rs3764261, which has been associated with HDL levels and AMD in European-descent populations, was replicated in our European American analyses (uncorrected $p<0.05$).(Chen et al., 2010; Sabatti et al., 2009; Willer et al., 2008b; Yu et al., 2011) The allele frequency for this SNP was 0.33 in our European Americans, similar to the 0.32 and 0.31 in the African Americans and Mexican Americans, respectively. A study by Chen *et al* found that rs3764261 trended toward significance in a Japanese population with a frequency of 0.22[25]. In the present study the Singaporean population had a similar frequency of 0.17 but was not associated with AMD. As with *CFH* Y402H, the frequency of the rs3764261 allele is similar across diverse populations but does not appear to contribute to AMD risk in all populations suggesting that this variant may only be "tagging" the casual variant.

AMD risk and lipid trait-associated variants

Lipid levels and lipid metabolism have been associated with susceptibility and progression of AMD in various populations.(Gemmy Cheung et al., 2012; van Leeuwen et al., 2004; Tomany et al., 2004) The cholesteryl ester transfer protein (*CETP*) is a plasma glycoprotein involved in the transport and removal of lipoproteins from blood circulation and as a component of the reverse cholesterol transport pathway. Variants in *CETP* have been found to cause an increase or decrease in blood lipid levels with corresponding high or low blood levels of CETP.(Chang et al., 2011; Ridker et al., 2009) This gene has been identified in various studies as a modifier of AMD susceptibility.(Chen et al., 2010; Yu et al., 2011) Our study found nominally significant associations with *CETP* variants, rs1800775 and rs3764261, in European Americans and African Americans. These *CETP* variants, along with other lipid-related genes in our study, were found to be nominally significant [rs328 (*LPL*) and rs6987702 (*TRIB1*)] in African Americans; and rs2338104 (*KCTD10/MVK*) in Mexican Americans. We observed more of these lipid-related associations in diverse populations than the European American population supporting the role of lipids in disease susceptibility

across diverse populations. Still, more studies are needed to elucidate the role they play in risk of AMD due to limited sample size.

Overall limitations and strengths

The major limitation for the PAGE study of age-related macular degeneration was small sample size. As expected, the European-descent samples (~6,500) constitute the largest population. Sample sizes for African Americans, Mexican Americans, and Asians were smaller, which adversely impacted power to detect associations. Sample size was further impacted by the fact that not all SNPs were genotyped or available in all study populations. Also, as already emphasized, only the index variant was genotyped in PAGE. The lack of GWAS-level or fine-mapping data abolished our ability to examine differences in LD in these study populations.

Despite the small sample sizes, a major strength of the present study is the diversity of populations included. Several large-scale studies have examined the genetic risk variants associated with AMD in European-descent,(Cipriani et al., 2012; Fritsche et al., 2013; Holliday et al., 2013; Naj et al., 2013; Yang et al., 2006; Yu et al., 2011) Japanese,(Arakawa et al., 2011; Chen et al., 2010; Goto et al., 2009; Tanaka et al., 2011) and Chinese populations.(Liu et al., 2013a; Ng et al., 2008; Tian et al., 2012; Wu et al., 2013) Indeed, some of these previous GWAS for AMD have included many of the same European American and Singaporeans examined here. In contrast, only a handful of limited genotyping studies have examined various Hispanic and African American groups in association with AMD and complement factor genes, *APOE*, and *ARMS2/HTRA1*.(Klein et al., 2008; Nonyane et al., 2010; Spencer et al., 2012a; Tedeschi-Blok et al., 2007; Tikellis et al., 2007) Compared to previous publications, this study contributes new data beyond *CFH* and *ARMS2/HTRA1*in African Americans and Mexican Americans. These new data, coupled with the cumulative data collected in European American and Asian populations, highlight shared as well as unique associations for AMD across diverse populations.

Acknowledgements

The PAGE consortium thanks the staff and participants of all PAGE studies and our collaborators from the SP2 and SiMES cohorts for their important contributions.

Funding

(a) The Population Architecture Using Genomics and Epidemiology (PAGE) program is funded by the National Human Genome Research Institute (NHGRI), supported by U01HG004803 (CALiCo), U01HG004798 (EAGLE), U01HG004802 (MEC), U01HG004790 (WHI), and U01HG004801 (Coordinating Center), and their respective NHGRI ARRA supplements. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The complete list of PAGE members can be found at http://www.pagestudy.org.

(b) The data and materials included in this report result from a collaboration between the following studies:

The "Epidemiologic Architecture for Genes Linked to Environment (EAGLE)" is funded through the NHGRI PAGE program (U01HG004798 and its NHGRI ARRA supplements). The study participants derive from the National Health and Nutrition Examination Surveys (NHANES), and these studies are supported by the Centers for Disease Control and Prevention. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. We would like to thank Dr. Geraldine McQuillan and Jody McLean for their help in accessing the Genetic NHANES III data. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core provided computational and/or analytical support for this work.

Funding support for the Genetic Epidemiology of Causal Variants Across the Life Course (CALiCo) program was provided through the NHGRI PAGE program (U01HG004803 and its NHGRI ARRA supplement). The following studies contributed to this manuscript and are funded by the following

agencies: The Atherosclerosis Risk in Communities (ARIC) Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022. The Cardiovascular Health Study (CHS) is supported by contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01-HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, N01-HC-85239, HHSN268201200036C, grants U01HL080295 and R01 HL087652, HL105756 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. See also http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center of Advancing Translational Technologies CTSI grant UL1TR000124 and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center and Cedars-Sinai Board of Governors' Chair in Medical Genetics (JIR).

Assistance with phenotype harmonization, SNP selection and annotation, data cleaning, data management, integration and dissemination, and general study coordination was provided by the PAGE Coordinating Center (U01HG004801-01 and its NHGRI ARRA supplement). The National Institutes of Mental Health also contributes to the support for the Coordinating Center.

SiMES is funded by National Medical Research Council (grants 0796/2003, IRG07nov013, IRG09nov014, STaR/0003/2008 and CG/SERI/2010) and Biomedical Research Council (grants 09/1/35/19/616), Singapore.

SP2 is funded by a grant from the Biomedical Research Council of Singapore (grant nos. BMRC Grant No 03/1/27/18/216 and BMRC 05/1/36/19/413). Ching-Yu Cheng is supported by an award from NMRC (CSA/033/2012). The Singapore Tissue Network and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services for tissue archival and genotyping, respectively.

Association of mitochondrial variants and haplogroups in diverse populations

In European Americans over the age of 80, nearly 1 out of 10 is likely to be diagnosed with some form of AMD.(Friedman et al., 2004b) Major genetic risk loci include *CFH*,(Edwards et al., 2005; Haines et al., 2005; Klein et al., 2005) *ARMS2*,(Jakobsdottir et al., 2005; Rivera et al., 2005) and *C2/C3*(Gold et al., 2006) which account for most of the known heritable risk of AMD. In all, 20 risk loci have been associated with risk of AMD accounting for upwards of 60% of the heritable risk.(Fritsche et al., 2014) Nearly all of these loci are located within the nuclear genome as there have been few studies investigating the potential role that mitochondrial genetic variation plays in the development of AMD.

(Restrepo et al., 2015)Adapted from: Restrepo, NA, Mitchell SL, Goodloe RJ, Murdock DG, Haines JL, Crawford DC.. Mitochondrial variation and the risk of age-related macular degeneration across diverse populations. *Pacific Symposium Biocomputing. Pac.* **20,** 243–254 (2015).

In vitro studies have found that mitochondrial DNA (mtDNA) variants can affect the replication rate of the mitochondrial genome and thus mtDNA copy number.(Kenney et al., 2014) Mitochondria are both particularly sensitive to and a major contributor of cellular reactive oxygen species (ROS), which are a byproduct of oxidative phosphorylation. These free radicals play a large role in chronic inflammation, the complement system pathways, and cardiovascular disease.(Cristina Kenney et al., 2014) Exposure to excessive oxidative stress can lead to mitochondrial dysfunction in the RPE layer,(Cano et al., 2014; Liang and Godley, 2003) a decrease in cellular bioenergetics imperative to photoreceptor initiation/maintenance,(Kenney et al., 2013, 2014; Sheu et al., 2013a) and susceptibility to apoptosis. Additionally, mitochondrial genetic variation has been associated with AMD risk in European Americans. MtDNA variants on mitochondrial haplogroup H, the most common European haplogroup, have been associated with decreased risk of AMD,(Jones et al., 2007; SanGiovanni et al., 2009; Udar et al., 2009)

while mitochondrial haplogroups J(Mueller et al., 2012) and T(Canter et al., 2008) are associated with increased risk. Collectively, these data suggest that the health of ocular mitochondria may play a role in AMD pathology.

Determining the role that mitochondrial genetic variation plays in AMD risk across populations may provide new insights into the underlying disease pathology. This study explores the contribution of mitochondrial genetic variation to AMD risk in European Americans, African Americans, and Mexican Americans.

Methods

Study populations

We accessed study participant data from NHANES III and NHANES 2007-2008 with the exclusion of NHANES 1999-2002 which did not collect ophthalmologic data. NHANES III, of which 3,131 participants had available fundus photographs and laboratory measurements of serum cotinine (ng/mL), oversampled non-Hispanic blacks and Mexican-Americans. NHANES 2007-2008 oversampled Mexican Americans and other-Hispanic blacks and had a total of 3,172 participants who completed the fundus exam. Current smokers were defined as those responding "yes" to the question "do you smoke cigarettes now?" or those with serum cotinine levels > 15ng/ml.

SNP selection and genotyping

The method of collection for NHANES III and Genetic NHANES has been previously described.138,139 We targeted a total of 63 mitochondrial SNPs for genotyping using Sequenom iPLEX® Gold MassArray as previously described.(Mitchell et al., 2014a)

Phenotyping

Participants over the age of 40 were selected to have a non-stereoscopic, 45° color fundus photograph taken of one randomly selected eye in NHANES III and a 45° non-mydriatic digital photo taken of both eyes in NHANES 2007-2008. Fundus photographs were graded according to a modified version of the Wisconsin

Age-related Maculopathy Scale.(Klein et al., 1991) AMD cases and controls were at least 60 years of age with gradable retinal photographs. Controls were not excluded if they presented with another retinal disease.

Statistical methods

We tested for an association between each individual mtSNP and mitochondrial haplogroups J, T, and U with AMD. Given that AMD largely occurs on a disease spectrum, data for early and late AMD were pooled for analyses to increase power to detect an association. Each mitochondrial variant was tested for an association with AMD using logistic regression assuming a dominant genetic model stratified by self-described race/ethnicity (e.g. non-Hispanic white, non-Hispanic black, and Mexican American). Of the SNPs that passed quality control (QC) standards (call rate >95%), a total of 55 SNPs were included in analyses for NHANES III, and 60 SNPs were analyzed in NHANES 2007-2008. A total of 50 SNPs were available for meta-analysis. Haplogroups were assigned to each NHANES participant as previously described.(Mitchell et al., 2014a) Haplogroup analyses were conducted in the same manner as the individual mtSNPs but with participants identified as having either haplogroup J, T, and U each being compared to participants in all other haplogroups. All models were adjusted for age, sex, BMI, and smoking status (current versus ever/never). Analyses were conducted using SAS v9.2 via the Analytic Data Research by Email (ANDRE) portal of the CDC Research Data Center in Hyattsville, MD. All p-values presented are uncorrected for multiple testing.

Ethics statement

All procedures were approved by the CDC Ethics Review Board and written informed consent was obtained from all participants. Because no identifying information is available to the investigators, Vanderbilt University's Institutional Review Board determined that this study met the criteria of "non-human subjects."

Results

Population characteristics

The study population consisted of a total of 416 AMD cases (312 non-Hispanic whites, 37 non-Hispanic blacks, and 67 Mexican Americans) and 2,200 controls (1,349 non-Hispanic whites, 430 non-Hispanic blacks, and 421 Mexican Americans) 60 years or older at the time of examination (Table 5). In the combined NHANES III/2007-2008 dataset, cases were generally female, except in Mexican Americans (49% female), and overweight defined as BMI >25 kg/m^2. Non-Hispanic black cases were nearly twice as likely to be smokers (62% smokers) compared to non-Hispanic white (29% smokers) and Mexican American (36%) cases. On average, controls were younger compared to cases across all race/ethnicities and were nearly as likely to be smokers compared to cases with the exception of non-Hispanic black cases (62% smokers) versus controls (50% smokers).

Table 5: Combined NHANES III and 2007-2008 study population demographics listed by case/control status and race/ethnicity

	non-Hispanic whites		non-Hispanic blacks		Mexican Americans	
	Case	Control	Case	Control	Case	Control
N	312	1349	37	430	67	421
Age (years)	76.0	71.0	69.2	67.8	69.0	67.1
% Female	60	51	57	49	49	48
% Smoker	29	30	62	50	36	31
BMI (kg/m^2)	27.2	27.8	31.1	29.0	29.2	28.9

Means are presented unless otherwise noted.

Mitochondrial SNP associations with age-related macular degeneration

A total of 50 mitochondrial SNPs passed QC and were tested across the three race/ethnicities in NHANES III and NHANES 2007-2008. Not all SNPs were available across each population as some SNPs were monomorphic in one population or did not pass QC. Of these 50 SNPs, 41 were available for analysis in non-Hispanic whites, 44 in non-Hispanic blacks, and 42 Mexican Americans (Figure 11).

In Mexican Americans, five mtSNPs were associated with AMD at $p < 0.05$ (Table 6). Of these, three are located in the mitochondrial control region, the non-coding region responsible for the initiation of transcription of the MT-genome: mt16111 ($p = 0.005$; OR = 2.90; 95% CI 1.38 – 6.11), mt16362 ($p =$

0.007; OR = 2.80; 95% CI 1.32 – 5.95), and mt16319 (p = 0.01; OR = 2.59; 95% CI 1.24 – 5.42). A synonymous variant, mt12007, located within the NADH dehydrogenase subunit 4 (MT-*ND4*) gene, was also found to increase risk of AMD in Mexican Americans (p = 0.018; OR = 2.47; CI 1.18 – 5.18). Lastly, mt1736 located in the mitochondrial 16S ribosomal RNA (*MT-RNR2*) gene, was found to be protective (p = 0.01; OR = 0.40; CI 0.19 – 0.83) in this population.

In non-Hispanic blacks and non-Hispanic whites, no test of association was significant at $p < 0.05$ in the adjusted model following inclusion of individual study results into the meta-analysis.

Table 6: Mitochondrial genetic variants associated with AMD risk in Mexican Americans

SNPID	Gene	OR	lower CI	upper CI	p-value	CA	CAF (%)	Race/Ethnicity
mt16111	control region	2.90	1.38	6.11	0.005	A	0.31	Mexican American
		—	—	—	0.98	A	0.01	non-Hispanic white
		—	—	—	0.98	A	0.02	non-Hispanic black
mt16362	control region	2.80	1.32	5.95	0.01	C	0.42	Mexican American
		1.70	0.93	3.10	0.08	C	0.08	non-Hispanic white
		2.29	0.84	6.27	0.10	C	0.13	non-Hispanic black
mt16319	control region	2.59	1.24	5.42	0.01	A	0.34	Mexican American
		0.42	0.05	3.52	0.43	A	0.01	non-Hispanic white
		3.23	0.33	31.60	0.31	A	0.02	non-Hispanic black
mt1736	*MT-RNR2*	0.40	0.19	0.83	0.01	A	0.65	Mexican American
		—	—	—	0.98	A	0.99	non-Hispanic white
		—	—	—	0.99	A	—	non-Hispanic black
mt12007	*MT-ND4*	2.47	1.18	5.18	0.01	A	0.34	Mexican American
		1.18	0.31	4.50	0.81	A	0.02	non-Hispanic white
		0.91	0.11	7.50	0.93	A	0.06	non-Hispanic black

Most significant meta-analysis results for the model adjusted by age, sex, body mass index, and smoking status. Results are listed for tests with the smallest p-value
Abbreviations: Odds Ratio (OR), confidence interval (CI), coded allele (CA), coded allele frequency (CAF)
"—" denotes genetic association tests with uninterpretable results due to very few case counts or monomorphic allele

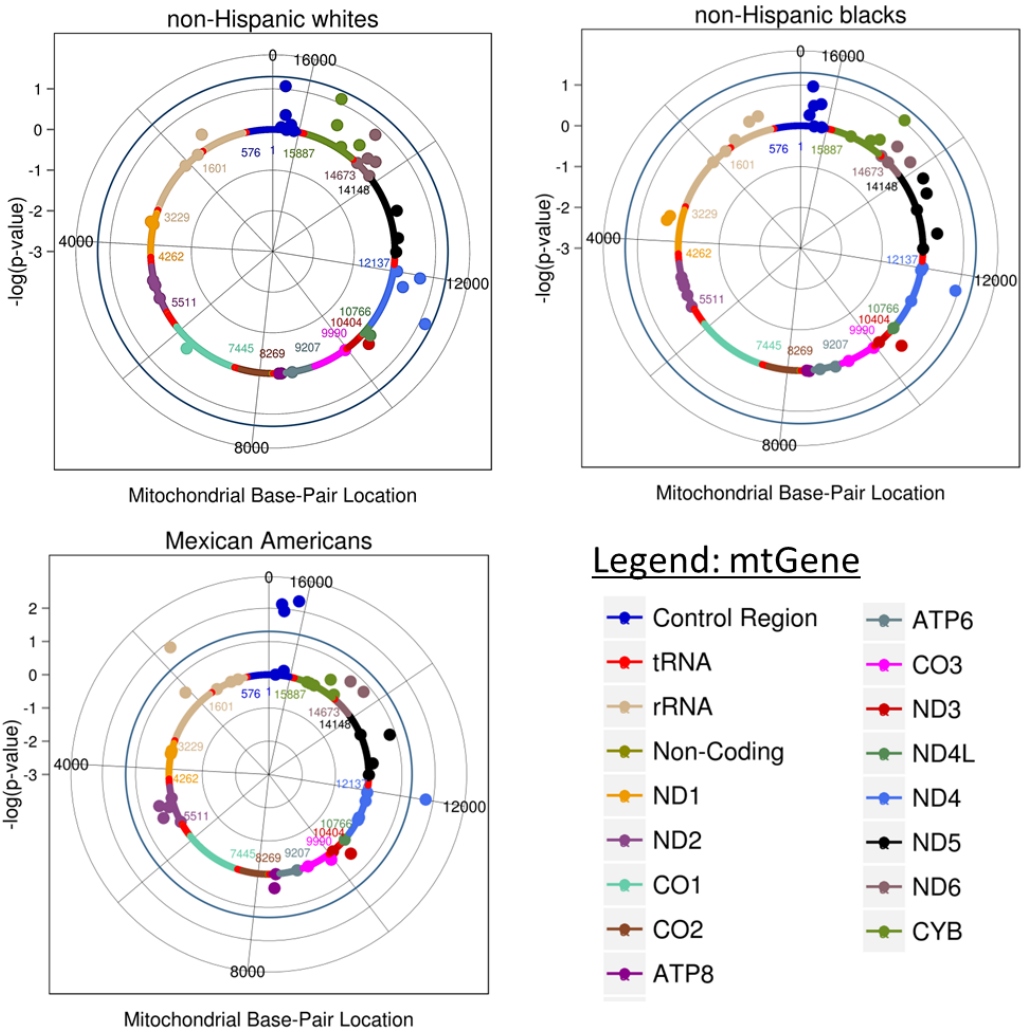


Figure 11: Meta-analysis single SNP association results by race/ethnicity.

Log(p) values were plotted using R. The outer blue ring represents a significance threshold of p = 0.05. SNPs are color coded by mitochondrial gene/regions as denoted in the legend. Cases/controls for the three populations are as follows: non-Hispanic whites (case = 312, control = 1,349), non-Hispanic blacks (case = 37, control = 430), and Mexican Americans (case = 67, control = 421).

Mitochondrial haplogroup analysis

Previous studies have suggested that mitochondrial haplogroups J, T, and U are associated with AMD.(Canter et al., 2008; Jones et al., 2007; Mueller et al., 2012; SanGiovanni et al., 2009; Udar et al., 2009) In NHANES III, none of the three haplogroups was associated with AMD at $p < 0.05$ although haplogroup J was associated in non-Hispanic whites at $p = 0.057$ (OR = 2.03; 95% CI 0.98 – 4.20). In NHANES 2007-2008, only haplogroup T was significantly associated at $p < 0.05$ in non-Hispanic whites (OR = 2.50; 95% CI 1.17 – 5.33). No haplogroup was found to be associated with AMD at $p < 0.05$ in any of the racial/ethnic groups in the NHANES III/2007-2008 meta-analysis.

Table 7: Haplogroup frequencies in the combined NHANES study populations
NHANES III, NHANES 1999-2002, and NHANES 2007-2008 as previously published(Mitchell et al., 2014a)

Haplogroup	non-Hispanic whites	non-Hispanic blacks	Mexican Americans
J	9.2 %	0.4%	1.4%
T	9.6%	0.4%	0.9%
U	13.6%	1.4%	1.6%

Discussion

In this study, haplogroup analyses did not replicate previous associations of the European haplogroups J, T, and U with risk of AMD in non-Hispanic whites.(Canter et al., 2008; Mueller et al., 2012; Udar et al., 2009) Other studies have not always replicated these associations,(Jones et al., 2007; Tilleul et al., 2013) which may be due in part to heterogeneity across these studies or else suggesting a weak role of mitochondrial variation in the risk of AMD. However, we did observe that individual variants on the T haplogroup in the NHANES 2007-2008 non-Hispanic whites were associated with AMD risk in this population. Unsurprisingly, neither individual variants nor haplogroups that were previously associated with AMD in European-descent populations generalized to non-Hispanic blacks. African-descent populations suffer from a lesser burden of AMD, and previous studies have suggested that African-descent populations may have a different genetic architecture contributing to AMD etiology(Friedman et al., 1999; Klein et al., 2007b; Sadigh et al., 2014; Spencer et al., 2012b) compared with other populations. We

observed that mitochondrial variants mt16111, mt16362, mt16319, mt1736, and mt12007 were associated with AMD risk within the meta-analyzed Mexican American population after adjusting for well-known non-modifiable factors and environmental modifiers. The direction of the genetic effect was the same across the individual NHANES analyses for these SNPs in the Mexican American populations as follows: mt16111 OR = 2.90 and 2.33; mt16362 OR = 2.49 and 3.35; mt16319 OR = 2.17 and 3.72; mt1736 OR = 0.55 and 0.22; mt12007 OR = 1.83 and 4.26 in NHANES III and NHANES 2007-2008 respectively.

Limited studies have been performed to assess the genetic factors of AMD in Mexican or Latino populations. A handful of studies have examined whether *CFH*, *ARMS2*, and *C2/C3*, strong risk loci in European populations, contribute to risk of AMD in Mexican-descent populations.(Buentello-Volante et al., 2012; Contreras et al., 2014; Spencer et al., 2012b) These studies, although limited in case size, did find a correlation between these European-derived variants and risk of advanced AMD in Mexicans and Latinos, suggesting that risk of AMD is being driven in part by European risk variants in these admixed populations. All five of the mtSNPs associated with AMD in Mexican Americans in this study are located on the A-A2 haplogroup background. Haplogroup A developed in Asia over 30,000 years ago and occurs most frequently in the Indigenous peoples of the Americas, with its subgroup A2 found to be the most common haplogroup in many of the indigenous ethnic groups of Central and North America.(Fagundes et al., 2008) In the combined NHANES III, 1999-2002, and 2007-2008 populations, haplogroup A is the most prevalent among Mexican Americans(Mitchell et al., 2014a) with a frequency of 34.2% while composing less than 1% in non-Hispanic whites and non-Hispanic blacks. This observation is interesting given that Mexican Americans, who experience similar rates of AMD as that observed in European-descent groups (5.1% vs 7.3%),(Klein et al., 2011) may contain a set of genetic risk factors on this haplogroup that are driving AMD risk in addition to or in combination with the already known European-derived variants.

Three of the significant mtSNPs (mt1611, mt16362, and mt16319) are located within the control region of the mitochondrial genome containing the origin of replication and the origin of transcription. These SNPs have not previously been associated with AMD but have been identified as contributors to various forms

of cancer. A high load of somatic mtSNPs in the mitochondrial control region (i.e. mt16111) was found in patients suffering from prostate cancer.(Chen et al., 2002) In a study examining the effect of mtSNPs located more specifically within the D-loop of the control region on risk for renal cell carcinoma, mt16319 was specifically found to be reduced in cases of clear cell renal cell carcinoma.(Zhang et al., 2013) Lastly, mt16362 was found to be a risk factor associated with familial breast cancer.(Cheng et al., 2014)

Strengths of this study include the systematic fashion in which all participants over the age of 40 years were included in ophthalmologic exams to ascertain eye health and AMD status. This ensures a strong degree of homogeneity in case and control status across the various NHANES cohorts and minimizes between study heterogeneity. Over sampling of minority groups also likely increased the number of cases available for study in these underrepresented groups. Limitations include differences in data collection between NHANES III and NHANES 2007-2008, as NHANES III only performed fundus photography on one randomly selected eye whereas NHANES 2007-2008 performed fundus photography on both eyes. Other limitations include low statistical power and lack of correction for multiple hypothesis testing. When considering a significance threshold for mitochondrial variation analyses it should be noted that independence of each test is questionable as all variation is inherited as a whole. Statistical power was a limitation for generalizing previously reported variants to the NHANES African Americans and Mexican Americans. A number of variants in the T1 and T2 haplogroups were identified by *SanGiovanni et al* in a Caucasian population as being associated with advanced AMD(SanGiovanni et al., 2009) with strong ORs ranging from 3.0 – 10. Our study did not genotype the exact variants reported in this study, but a number of variants were genotyped nearby. Variants such as mt11812 had an allele frequency of ~23% and an OR of 10 in the *SanGiovanni* study. In our study, if we assume a MAF of 25%, case:control of 1:3, a dominant genetic mode of inheritance, and a two-sided t-test then we had 80% power to detect an association with an OR = 1.50 in NHANES European Americans, OR = 3.0 in African Americans, and an OR = 2.30 in Mexican Americans. We likely did not find similar associations for a number of reasons, one of the least of which is differences between inclusion criteria of cases between our study and *SanGiovanni* study. Our study

included both early and late/advanced AMD cases. Late AMD cases make up only a fraction of total AMD cases (i.e., ~8-12%). Also, the frequency of the T-haplogroup in NHANES African Americans and Mexican Americans is a scarce 0.4% and 0.9%, respectively. Lastly, we relied on self-reported race/ethnicity as opposed to genetic ancestry of the mitochondrial genome which may lead to false positive associations that are in actuality identifying differences in haplogroup ancestry. Future studies are needed to validate the results of the present study.

Despite limited sample sizes available for AMD analyses, NHANES is one of only a few surveys to include ophthalmologic exams of minority populations. As large scale epidemiology surveys become more cost prohibitive, a stronger emphasis on the utilization of electronic medical records (EMR) to identify cases and controls for inclusion in future studies will become more pronounced. Given that many of these EMR systems are still predominately composed of European-descent patients, a concerted effort must be made to increase the number of minorities with access to routine, continuous health care.

Acknowledgements

This work was supported by NIH U01 HG004798 and its ARRA supplements. The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core provided computational and/or analytical support for this work. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Gene x Environmental interaction effects on risk of age-related macular degeneration in diverse populations

Environmental factors such as smoking and BMI are well known to influence AMD risk. While these environmental factors may increase the odds of developing AMD, other environmental factors such as vitamin supplementation may indeed protect against disease onset, slow progression, or limit severity of AMD. Previous studies from the NIH-sponsored Age-Related Eye Disease Study (AREDS) group have suggested that vitamin supplementation could provide a potential treatment option for AMD. AREDS enrolled participants from 1992-1998 for a prospective study of AMD and cataracts(Age-Related Eye Disease Study Research Group, 1999) which included a double-masked clinical control study of high dose supplementation with varying formulations of vitamin C, vitamin E, beta-carotene, and zinc(Age-Related Eye Disease Study Research Group, 2001). A collaboration with Bausch and Lomb Pharmaceuticals later paved the way for additional clinical trials to begin to rigorously test the hypothesis that antioxidants could treat AMD. Eventually, AREDS would develop their patented formula consisting of zinc, Vitamin E, Lutein, Zeaxanthin, and copper which was found to slow progression of AMD in patients with intermediate disease severity. AREDS intermediate AMD was defined as the presence of several medium-sized drusen or one or more large druse. AREDS trials along with other studies have set a strong precedent that antioxidants play a strong role in retinal health(Age-Related Eye Disease Study Research Group et al., 2007; Chong et al., 2007; Tan et al., 2008; 2000a).

As previously established in this chapter, specific genetic variants contribute to susceptibility to AMD; however, neither the collective known genetic factors nor the collective known environmental factors explain all of the risk for AMD(Buitendijk et al., 2013; Fritsche et al., 2014). Indeed, this observation is a common one for complex diseases where multiple common genetic variants have been identified and replicated by GWAS and subsequent studies but collectively the variants do not account for the majority of the genetic component estimated in heritability studies(Manolio et al., 2009). It has been postulated that the "missing heritability(Maher, 2008)" can be found, in part, within gene-environment interactions not tested in GWAS(Eichler et al., 2010).

Gene-environment studies in AMD have been mostly limited to candidate gene studies. Previous studies have described an interaction effect between smoking and a rare haplotype variant in *CFH*(Biswas et al., 2014). Other studies have identified potential interactions between smoking and the *APOE E2* allele(Eichler et al., 2010), variants in *ARMS2*(Schmidt et al., 2006), and variants in *NOS2A*(Ayala-Haedo et al., 2010). However, these potential interactions have either failed to replicate(Conley et al., 2006; DeAngelis et al., 2007; Hughes et al., 2007) or have yet to be replicated in subsequent studies.

Beyond smoking, very few studies have tested the potential modifying effects of other environmental risk factors. In the following study, the hypothesis that carotenoids, organic plant-based pigments, interact with strong AMD risk variants to further influence risk of disease was tested. Additionally, BMI which has been inconsistently associated with AMD(Chakravarthy et al., 2010; Goldberg et al., 1988; Tan et al., 2007b) was tested as a surrogate for visceral fat. Increasing evidence suggests that the pro-inflammatory effects of visceral fat may play a role in AMD pathogenesis(Haas et al., 2015; Howard et al., 2014). Gene x environment interaction analyses were conducted for three of the preeminent AMD risk variants (i.e. rs10490924 *ARMS2*, rs1061170 *CFH*, and rs547154 *C2*) in combination with serum carotenoid levels and BMI across the NHANES III populations.

Methods

Study populations and SNP selection and genotyping

We accessed study participant data from NHANES III which was described in greater detail in the previous methods section for the Association of mitochondrial variants and haplogroups in diverse populations. Details of the genotyping are also available in the previous section.

Collection of quantitative traits

Blood samples were obtained for all NHANES III participants over the age of one year at the mobile examination center. Carotenoids (e.g. lutein/zeaxanthin, alpha carotene, beta carotene) and vitamin A were analyzed via isocratic high performance liquid chromatography (HPLC) at three different wavelengths

within designated sites(Gunter et al., 1996). Test detection limits are as follows: lutein (0.43 µg/dL); vitamin A (0.5 µg/dL); alpha carotene (0 µg/dL); beta carotene (0.67 µg/dL). Testing sites set the following reference ranges as normal serum concentration levels of the following: lutein/zeaxanthin (5-65 µg/dL); vitamin A (25-115 µg/dL); alpha carotene (1-15 µg/dL); beta carotene (2-80 µg/dL).

Statistical methods

We tested for an association with AMD and each quantitative trait using linear regression adjusting for age, sex, BMI, and smoking status (Supplementary Table 5). Each genetic variant was tested for association with AMD using logistic regression assuming an additive genetic model adjusting for age, sex, BMI, and smoking status. Quantitative traits without a normal distribution were log transformed (i.e., vitamin A and alpha carotene). Exhaustive pairwise SNP x quantitative trait interactions were modeled using an interaction term (SNPxtrait). Gene–environment interactions were modeled using a multiplicative interaction term between the environmental variable and the additively encoded SNP. All models were adjusted for the main effect of the SNP and the environmental variable, along with age, sex, BMI, and smoking status.

Results

Population characteristics

The study population consisted of a total of 190 non-Hispanic Whites (NHW), 30 non-Hispanic Blacks (NHB), and 47 Mexican American (MA) AMD cases from the NHANES III (Table 8). As expected, cases were older than controls in all three populations. However, NHB and MA cases were on average younger than NHW cases (70.2 and 69.3 years vs 76.4 years). NHB cases were more likely to be obese (BMI > 30 kg/m^2) and smokers compared to NHW and MA cases. Serum concentration of carotenoids for all populations, regardless of case status, were within normal limits as described below.

Table 8: EAGLE AMD study population demographics and mean values for NHANES III

	Non-Hispanic White		Non-Hispanic Black		Mexican American	
	case	control	case	control	case	control
N	190	664	30	209	47	270
% females	66.8	54.5	60.0	47.4	44.7	44.1
Age	76.4	71.2	70.2	68.5	69.3	67.2
BMI (kg/m^2)	26.8	27.1	31.0	28.2	28.6	28.4
% Current Smokers	46.3	53.5	70.0	65.1	53.2	55.9
Serum vitamin A (ug/dL)	65.4	63.9	59.9	61.8	54.3	56.3
Serum lutein/zeaxanthin (ug/dL)	23.2	23.4	28.6	30.4	25.3	25.6
Serum alpha carotene (ug/dL)	6.8	5.7	4.2	4.5	4.9	4.7
Serum beta carotene (ug/dL)	32.2	26.3	30.5	28.9	19.8	18.2

Overall, *CFH* rs1061170 (C) occurred less frequently in Mexican American controls (0.19) compared with non-Hispanic whites (0.34) and non-Hispanic blacks (0.37). Conversely, the total *ARMS2* rs10490924 (T) frequency was similar across the three groups (not shown): 0.22 (non-Hispanic white), 0.24 (non-Hispanic black), and 0.26 (Mexican American). *C2* was more frequent in non-Hispanic blacks 0.21 compared with non-Hispanic whites 0.11 and Mexican Americans 0.11.

Table 9: Allele frequencies of AMD SNPs in NHANES III, stratified by race/ethnicity and AMD case/control status.

			Non-Hispanic White		Non-Hispanic Black		Mexican American	
SNP	Gene	CAF	case	control	case	Control	case	Control
rs10490924	*ARMS2*	T	0.30	0.19	0.13	0.26	0.35	0.25
rs1061170	*CFH*	C	0.41	0.34	0.48	0.37	-	0.19
rs547154	*C2*	T	0.03	0.11	-	0.20	0.11	-

"-"denotes cells in which the number of individuals included would be less than five. These numbers were excluded to prevent potential re-identification of individuals.

Replication and Generalization of AMD variants

As previously reported in the PAGE meta-analysis, *CFH* rs1061170 (OR =1.32; p=0.02) and *ARMS2* rs10490924 (OR = 1.88; p=0.0001) were associated with AMD in non-Hispanic whites. As previously reported by *Spencer et al*(Spencer et al., 2012b), *ARMS2* rs10490924 was also associated with AMD in Mexican Americans (OR =1.64; p=0.03) and non-Hispanic blacks (OR = 0.43; p=0.04) at $p < 0.05$ but for non-Hispanic blacks the association is in the opposite direction. It should also be noted that the (T) allele is also noticeably less frequent in non-Hispanic black cases (0.13) vs. non-Hispanic white cases (0.30) and Mexican American cases (0.35). We found no association between *C2* rs547154 and risk of AMD in any of the populations.

Table 10: Replication and generalization results of AMD variants across racial/ethnic populations in NHANES III

SNP	Gene	Coded Allele	Ethnicity	OR	CI	p-value
rs10490924	*ARMS2*	T	Non-Hispanic White	1.88	1.42-2.50	0.0001
			Non-Hispanic Black	0.43	0.19-0.97	0.04
			Mexican American	1.64	1.03-2.60	0.03
rs1061170	*CFH*	C	Non-Hispanic White	1.32	1.03-1.69	0.02
			Non-Hispanic Black	1.61	0.89-2.88	0.11
			Mexican American	1.26	0.72-2.21	0.42
rs547154	*C2*	T	Non-Hispanic White	0.82	0.54-1.23	0.34
			Non-Hispanic Black	1.17	0.62-2.24	0.60
			Mexican American	1.07	0.52-2.20	0.85

Gene x Environment associations with age-related macular degeneration

We found that AMD risk was modified by an interaction between *C2* rs547154 and BMI ($beta_{interaction}$ = 0.85; p = 0.0004) in non-Hispanic whites. Results suggest that vitamin A modifies *CFH* in Mexican Americans ($beta_{interaction}$ = 9.48; p = 0.02) and *ARMS2* in non-Hispanic blacks ($beta_{interaction}$ = 11.35; p = 0.04). Carotenoids (i.e., alpha-carotene, beta-carotene, and lutein/zeaxanthin) did not appear to modify the effect of AMD SNPS in any population. Lastly, we did not find a significant interaction ($p < 0.05$) that generalized across all three populations.

Table 11. Environmental modifiers of AMD-associated genetic variants in NHANES III populations with p < 0.05. Model adjusted for the main effect of the SNP, the environmental variable, and covariates: age, sex, BMI, and smoking status.

SNP	Gene	Coded Allele	Ethnicity	Interaction	OR	SE	p - value
rs547154	*C2*	T	Non-Hispanic White	SNP x BMI	0.85	0.04	0.0004
rs1061170	*CFH*	C	Mexican American	SNP x VAP	9.48	0.96	0.02
rs10490924	*ARMS2*	T	Non-Hispanic Black	SNP x VAP	11.35	1.19	0.04

BMI = body mass index kg/m^2
VAP = serum Vitamin A (ug/dL)

Discussion

We set out to determine whether carotenoids, vitamin A, or BMI modify the effect of major AMD index SNPs across diverse populations. In the interaction model adjusted for age, sex, BMI, and smoking status, BMI modified *C2* rs547154 to reduce the risk of AMD in non-Hispanic whites. In addition, we find a significant interaction effect between serum Vitamin A levels in conjunction with *CFH* and *ARMS2* variants for AMD risk in Mexican Americans and non-Hispanic Blacks, respectively. We did not identify an interaction between lutein/zeaxanthin or alpha- and beta- carotene with AMD index SNPs in any of the three NHANES III populations.

Lack of evidence for modification by antioxidant exposures is corroborated by the larger AREDS study. In the 2nd AREDS trial addition of lutein/zeaxanthin did not appreciably alter reduction in the risk of advanced

AMD, although the trial results were suggestive that lutein/zeaxanthin could prove more beneficial than beta-carotene(Age-Related Eye Disease Study 2 (AREDS2) Research Group et al., 2013). Similarly, an AREDS study testing whether rs1061170 and rs10490924 interacted with the AREDS supplement (containing vitamin C, vitamin E, beta-carotene, zinc, and copper) to alter progression from intermediate to late AMD found no association(Chew et al., 2014). A joint Blue Mountains Eye Study and Rotterdam Study investigation examined dietary intake of lutein/zeaxanthin, beta carotene, and vitamin C from food frequency questions and AMD genetic risk as defined by the number of rs1061170 and rs10490924 alleles. They identified a positive interaction effect between lutein/zeaxanthin intake and genetic risk(Wang et al., 2013). There are a number of differences between our studies which may explain lack of replication. The largest difference is sample size with our study containing 190 non-Hispanic white cases in comparison to 1,327 cases in the joint study. Additionally, the manner in which each study assessed lutein/zeaxanthin (i.e., blood sample measurement versus food frequency questionnaires) may have contributed to differences in analysis outcomes.

It is interesting to note that our study did not find a significant association for a main effect in non-Hispanic Whites with either the *C2* SNP. Previous work has repeatedly identified rs547154 as a protective variant for AMD(Kaur et al., 2010; Richardson et al., 2009; Spencer et al., 2007; Thakkinstian et al., 2012) and polypoidal choroidal vasculopathy(Nakata et al., 2012) (PCV). The minor allele frequency of 11% in non-Hispanic white controls is similar with the HapMap CEU estimates of 6-9%, but given our limited sample sizes and a minor allele frequency of 3% in non-Hispanic white cases, we were underpowered to detect the previously reported association. Lack of replication may also be due to differences in study design, population ascertainment, and heterogeneity.

NHANES III is a cross-sectional survey that measured blood serum vitamin levels at one time point in contrast to other studies which collect food questionnaires. A major strength of the NHANES III design is that it provides a quantitative measurement of vitamin exposure as opposed to food questionnaires which

rely on an individual's recollection of what they ate in the past days or months. Subsequently, a single time point estimate does not allow for studies to look at long-term exposure variables.

Due to limited statistical power, we did not adjust for multiple testing which is a major limitation of this study. Had we taken into account the fifteen tests performed for each population (i.e. 3 SNPs x 5 traits), our Bonferroni threshold would be $p < 0.003$. Excitingly, the interaction for *C2* x BMI in non-Hispanic whites surpasses this threshold ($p = 0.0004$). Replication in other datasets is necessary to determine if this is a true effect. The interaction observed here between vitamin A levels and *CFH* in Mexican Americans and *ARMS2* in non-Hispanic blacks also requires additional testing but provides an attractive opportunity to further assess antioxidants role in AMD pathology.

Although underpowered, the present study suggests quantitative traits known to be associated with AMD risk may further modify genetic risk. Further studies are needed to determine if lack of generalization is due to statistical power or differences in allelic distribution across these diverse populations.

Summary

In summary, we have contributed to the characterization of the genetic architecture of AMD in diverse populations. We performed a meta-analysis of known AMD-related variants and variants in cholesterol pathways in the three major race-ethnicities in the United States, and Chinese and Malay individuals from Singapore. Additionally, we identified potential novel associations between mitochondrial variants and risk of AMD in the NHANES Mexican Americans. All of the associated mitochondrial variants are found on the A-A2 haplogroup which is common among this racial/ethnic group and varies from the haplogroups previously reported to influence risk in European-descent populations. Future studies in larger Mexican-descent datasets may clarify some of these findings. Lastly in a study to assess whether antioxidants or body-mass-index modify the effect of prominent AMD risk variants, we identified a significant interaction effect of body-mass-index on the *C2* SNP rs547154 in NHANES III non-Hispanic whites.

Although limited in scope, the work presented in this chapter has identified potential genetic factors influencing the risk of AMD in diverse populations. The potential exists for elucidating inter- and intra-population causative pathways as highlighted by the presence or lack of an association of *CFH* Y402H and *ARMS2* A69S across populations. These new findings may offer insight into the cellular mechanics underlying the etiology of AMD.

CHAPTER III

UTILIZATION OF ELECTRONIC MEDICAL RECORD SYSTEMS FOR GENETIC ASSOCIATION STUDIES

Introduction

The use of electronic medical records (EMR) in biomedical research has gained traction in recent years, driven in part by the large-scale collection of biospecimens associated with EMRs. When linked to extensive biobanks, these large, data-dense resources can be utilized in large-scale genetic and genomic studies(Kohane, 2011). Compared to traditional epidemiological cohorts, which are expensive and can take years to decades to collect, EMRs are a relatively cost efficient alternative with years of medical information immediately accessible. These medical records can be interrogated for a range of phenotypes often not available or included in epidemiological cohorts. Also, large sample sizes can be acquired from the development of phenotype algorithms standardized for use across multiple institutions(Kho et al., 2012; Newton et al., 2013).

Here we introduce two algorithms to extract cases and controls of primary open-angle glaucoma (POAG) and diabetic retinopathy (DR) from a single institution's EMR among for genetic association studies. POAG is a clinical subtype of glaucoma, a heterogeneous group of eye diseases characterized by chronic degeneration of the optic nerve and gradual vision-loss. As the second leading cause of blindness in the United States(Resnikoff et al., 2004), glaucoma is a leading cause of vision disability. The most common form of glaucoma is POAG which disproportionately afflicts African Americans. The prevalence of POAG in African Americans is nearly double that observed in European-descent populations(Congdon et al., 2004; Friedman et al., 2004a; Stein et al., 2011b) with rates as high as 5.6% vs. 1.7%, respectively, in individuals over the age of 40 years(Vajaranant et al., 2012a). Likewise, DR, another driving cause of blindness and vision loss, occurs more predominately in African Africans with diabetes (36.7%) in comparison to European Americans (24.8%) (Table 1: Chapter 1). Unfortunately few studies focus on African American patients, which may be due in part to the perception that African Americans are less willingly to participate

in research and clinical trials.(Corbie-Smith et al., 1999, 2002; Freimuth et al., 2001) Though this perception persists, epidemiological study surveys find that African Americans are equally interested in participating in research trials as their European American peers(Brown and Topcu, 2003; Katz et al., 2007; Wendler et al., 2006).

In general, participation rates in epidemiology studies have declined in the last 30 years. Participation rates in clinical trials have been harder to ascertain due to a lack of a centralized repository for participation. This decline may be due to many factors, a few of which include feelings of exploitation by medical institutions and pharmaceutical companies and the time commitment needed for a series of assessments and lengthy follow-up appointments(Galea and Tracy, 2007). It is also possible that with the increase in the total number of clinical and research trials being carried out by government agencies, medical/academic institutions, and pharmaceutical companies, people are experiencing volunteer burnout(Galea and Tracy, 2007). The use of EMRs not only limits the commitment burden of an active participant population, assuming permission has already been granted for research use of biomedical data, but may completely eliminate the onus if patient data has been de-identified. With the de-identification of personal information, researchers have the opportunity to quickly accrue data and pursue studies with the approval of an Institutional Review Board (IRB).

Vanderbilt University Medical Center de-identified electronic medical records systems

One such EMR resource available for research is the Vanderbilt University Medical Center (VUMC) Synthetic Derivative (SD). The SD is a de-identified version of Vanderbilt's institutionally developed EMR system StarChart and contains inpatient and outpatient medical records collected at VUMC and affiliated clinics. Patient records consist of both structured (e.g., billing codes, procedure codes, laboratory values) and unstructured (e.g., clinical free text) data. To date, the Vanderbilt EMR contains over 2.2 million records with each record containing on average 6.5 years of medical history and an average of eight prescriptions. The SD is linked with VUMC's DNA repository known as BioVU(Roden et al., 2008). These DNA samples are extracted from discarded blood samples collected from outpatient clinical laboratories.

The SD in conjunction with BioVU has been used recently in several genetic association studies to identify genetic variants associated with a range of phenotypes such as multiple sclerosis(Ritchie et al., 2010), type-2 diabetes(Long et al., 2012), cancer(Cheng et al., 2013; Kocarnik et al., 2014), electrocardiographic traits(Denny et al., 2010; Jeff et al., 2013; Ritchie et al., 2013), pharmacogenomic-related outcomes(Delaney et al., 2012; Oetjens et al., 2014; Ramirez et al., 2012), clinical quantitative traits(Crosslin et al., 2012, 2013; Ding et al., 2013; Gong et al., 2013; Rasmussen-Torvik et al., 2012), and hypothyroidism(Denny et al., 2011). As part of the PAGE I study(Matise et al., 2011a), we as the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study accessed ~15,000 DNA samples of non-European descent individuals in BioVU (EAGLE BioVU) to perform genetic association studies in diverse populations.

A major challenge associated with use of the Vanderbilt EMR for ocular research is that the specialty eye clinic (the Vanderbilt Eye Institute or VEI) currently lacks an interface to digitally upload forms into the broader EMR. Due to this lack of interface, digital photographs are not readily available for research use in the SD. Given these limitations, we sought to develop a phenotype algorithms using searchable and parsable elements available in the SD. With this goal in mind, we have 1) developed and implemented a data-mining algorithm to classify individuals as POAG and DR cases or controls; and 2) manually verified the case/control status of individuals to evaluate the algorithm's performance.

Methods

Ethics Statement

BioVU is an opt-out biorepository. DNA is collected from discarded blood samples remaining after routine clinical testing and is linked to de-identified medical records. According to the Vanderbilt Institutional Review Board (IRB) and the Federal Office of Human Research Protections provisions, the Vanderbilt protocol is considered nonhuman subjects research (The Code of Federal Regulations, 45 CFR 46.102 (f)). The IRB at Vanderbilt University approved this research.

Population

The EAGLE Study, as part of the PAGE I Study(Matise et al., 2011a), accesses clinical and epidemiological collections with racially/ethnically diverse populations. These collections are used to perform and generalize genetic association studies for common human diseases, including common ocular diseases such as POAG. As part of PAGE I, EAGLE genotyped all DNA samples in BioVU from non-European-descent individuals as of 2011 (EAGLE BioVU; n=15,863). All DNA samples were genotyped on the Illumina Metabochip by the Vanderbilt University Center for Human Genetics Research DNA Resources Core. The Metabochip is a custom array designed for replication and fine mapping of genome-wide association study (GWAS)-identified variants for metabolic and cardiovascular traits(Voight et al., 2012). The data described here will be available through the database of Genotypes and Phenotypes (dbGaP).

Calculation of PPV, NPV, Sensitivity, Specificity, and Accuracy

The performance of the case and control algorithms were calculated as follows: PPV was calculated as the ratio of true positives (TP) over TP + false positive (FP) [PPV = TP/(TP + FP)]. A TP was an individual who was identified by the case algorithm as a case and was then confirmed by manual review to be a true case. A FP in turn was an individual identified as a case that was determined not to be a case during manual review. NPV is the ratio of true negative (TN) (i.e. a control who was confirmed as a control) over TN and FN [NPV = TN/(TN + FN)]. Sensitivity was calculated as the ratio of TP over TP and false negatives (FN) [Sensitivity = TP/(TP + FN)]. A FN was an individual identified by the control algorithm as a control who then by manual review was determined not to be a control. Specificity was calculated as the ratio of the TN over TN and FP [Specificity = TN/(TN +FP)]. Lastly, Accuracy was calculated as the ratio of the sum of TP and TN over the sum of all positives and all negatives [Accuracy = (TP + TN)/(Positives (TP + FP) + Negatives (TN + FN))].

Extraction and calculation of individual demographic elements

Demographic data and laboratory measurements were extracted and calculated with an emphasis for use in future case/control studies. For cases, age at POAG diagnosis was determined by the date in the records for

the first mention of a POAG ICD-9 (365.11). For controls, age at last clinic visit (LCV) was taken as the date of the last CPT mentioned in the records. An individual, regardless of case/control status, was classified as a hypertensive if he/she met one of three criteria: 1) systolic blood pressure > 140 mm/Hg, 2) diastolic blood pressure > 90 mm/Hg, or 3) mention of a hypertension medication within a two year window of when an individual was diagnosed with POAG for cases or within a two year window of their LCV date for controls. Median values were calculated for the following laboratory measurements within a two year window of an individual's POAG diagnosis or LCV for controls: blood pressure (systolic and diastolic), lipids (total cholesterol, high-density cholesterol, low-density cholesterol, and triglycerides), and body mass index (height and weight).

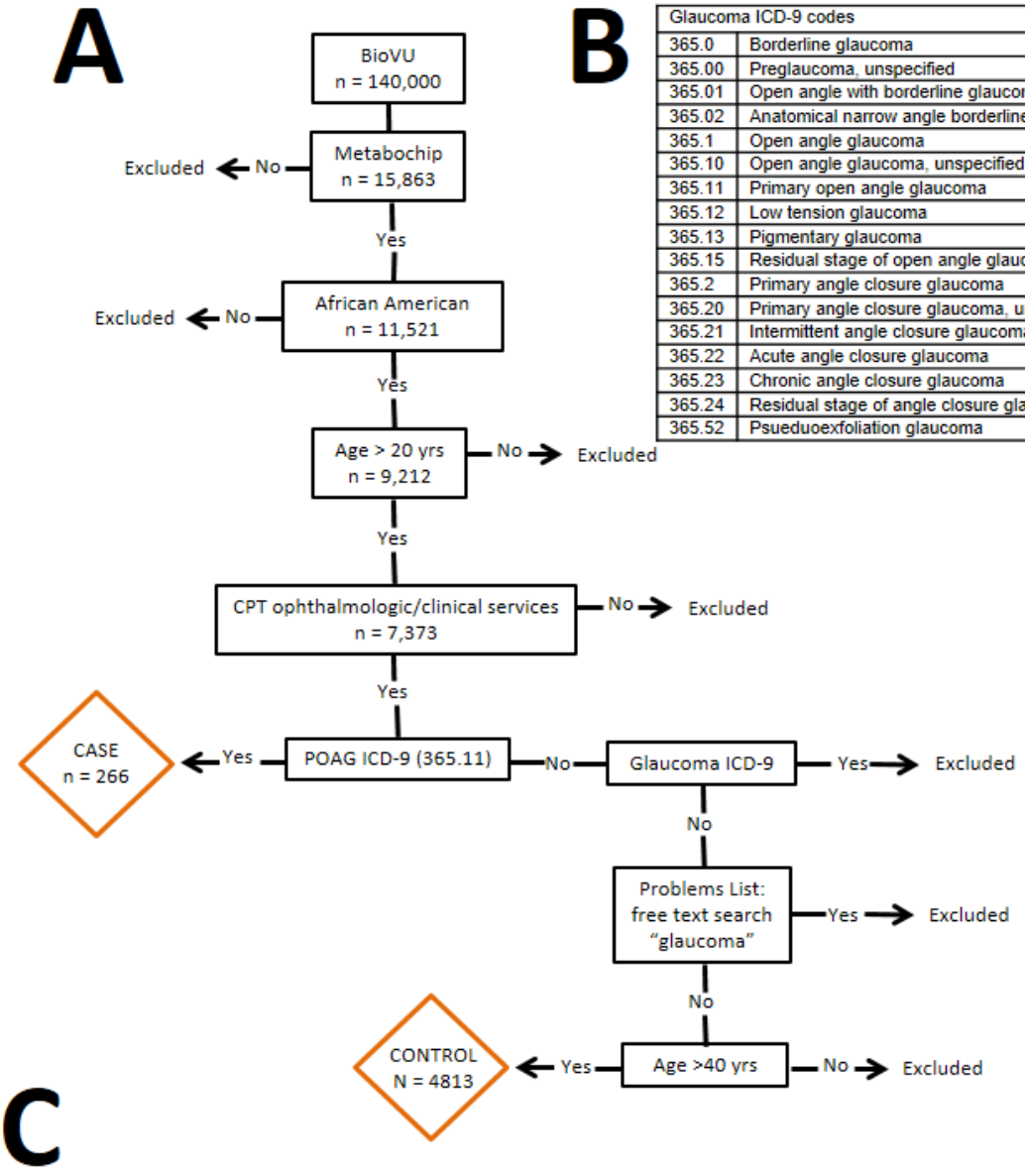
Algorithm development for primary open-angle glaucoma

Briefly, we developed an algorithm designed to identify POAG cases and controls from EAGLE BioVU, a subset of records and DNA samples from the larger Vanderbilt SD and BioVU, using a combination of International Classification of Diseases (ICD-9) diagnostic codes for glaucoma, Current Procedural Terminology (CPT) billing codes for ophthalmology/general clinic, and free text searches (Figure 12 A, B, and C). Digital photographs are the gold standard for diagnosing ocular diseases such as POAG; however these are not readily available in the SD for research as the VEI currently uses a manual/paper process for storing patient data, with most forms being scanned and uploaded into the EMR as portable document format (PDF) forms. Because it is difficult to de-identify and parse PDF forms, this POAG case/control algorithm targeted structured and easily parsed unstructured data and was designed to maximize the number of individuals eligible for manual review to confirm case/control status.

Initial screening criteria for POAG study population

Individuals included for this POAG study were African American adults over the age of 20 years as of March 20, 2013. We excluded pediatric glaucoma cases because it is a separate condition caused by developmental issues prior to birth and/or by a very rare genetic mutation. The genetics of pediatric glaucoma are likely to be fundamentally different from the genetics of glaucoma in older adults, and this

genetic heterogeneity would result in lower power to detect genetic associations for adult POAG. The final inclusion criteria required that an individual's medical records include either one mention of a CPT code for ophthalmology or a CPT code for general clinic procedures (Figure 12 C).



Glaucoma ICD-9 codes	
365.0	Borderline glaucoma
365.00	Preglaucoma, unspecified
365.01	Open angle with borderline glaucoma finding
365.02	Anatomical narrow angle borderline glaucoma
365.1	Open angle glaucoma
365.10	Open angle glaucoma, unspecified
365.11	Primary open angle glaucoma
365.12	Low tension glaucoma
365.13	Pigmentary glaucoma
365.15	Residual stage of open angle glaucoma
365.2	Primary angle closure glaucoma
365.20	Primary angle closure glaucoma, unspecified
365.21	Intermittent angle closure glaucoma
365.22	Acute angle closure glaucoma
365.23	Chronic angle closure glaucoma
365.24	Residual stage of angle closure glaucoma
365.52	Psueduoexfoliation glaucoma

Ophthalmology/Clinical CPT codes	
92002, 92004	Ophthalmological services: medical examination and evaluation with initiation of diagnostic and treatment program
92012, 92014	Ophthalmological services: medical examination and evaluation, with initiation or continuation of diagnostic and treatment program
99201, 99202, 99203, 99204, 99205	Office/outpatient visit for evaluation of new patient, problem focused history, exam, counseling, and/or coordination with other physicians or health care professionals
99212, 99213, 99214, 99215	Office/outpatient visit for evaluation and management of established patient: problem focused history, exam, medical decision making, counseling, and/or coordination with other physicians or health care professionals
99307, 99308, 99309, 99310	Nursing facility care for evaluation and management of patient, problem focused interval history, exam, medical decision making, counseling, and/or coordination with other physicians or health care professionals
99324, 99325, 99326, 99327, 99328	Domiciliary or rest home visit for evaluation and management of new patient, problem focused history, exam, medical decision making, counseling, and/or coordination with other physicians or health care professionals
99334, 99335, 99336, 99337	Domiciliary or rest home visit for evaluation and management of established patient, problem focused history, exam, medical decision making, counseling, and/or coordination with other physicians or health care professionals

Figure 12: Flow diagram of POAG case/control algorithm

Primary open-angle glaucoma cases

POAG cases were identified as individuals with at least one ICD-9 code for POAG (365.11). Cases were not excluded if they contained an additional ICD-9 code for another sub-type of glaucoma (list of glaucoma ICD-9 codes included in Figure 12 B). It is not uncommon for patients to experience different sub-types of glaucoma bilaterally or within the same eye which is also known as mixed- or combined- mechanism glaucoma. Also, over the course of a lifetime a patient may progress from one form to another such as closed-angle glaucoma to open-angle glaucoma.

Primary open-angle glaucoma controls

Controls were defined as individuals whose records were devoid of any glaucoma ICD-9 code. If the algorithm identified "glaucoma", "glaucome", "glocoma", "gloucoma", "gluacoma", "glucoma", or "glycoma" in a free text search of problems lists and clinical notes, the individual was excluded. We excluded individuals under the age of 40 years, as calculated from a given birth-date, to reduce contamination of controls with future cases.

Manual review of primary open-angle glaucoma

The SD records of all EAGLE BioVU POAG cases and a random sample of controls identified by the algorithm were manually reviewed by myself to verify POAG case/control status. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were then calculated to evaluate the algorithm's performance. Of the 267 individuals identified as cases, 138 were determined to be definite cases based on records retrieved from the SD. The records accessed for verification of case status were primarily composed of but not limited to surgical reports, optometry, and ophthalmology clinic notes. Other data that was taken into account were medication lists, general and specialty clinic reports, clinical communications, and problems lists. An individual was classified as a definite case if they met one of two criteria: 1) a written diagnosis by a Vanderbilt ophthalmologist/optometrist as pertained to a patient's exact clinical sub-type of glaucoma (An example of this is seen in Figure 13) and 2) the patient's medical record contained *all* of the following: at least two independent mentions of the POAG ICD-9 code (365.11),

a glaucoma medication (Table 12), and a surgical procedure for treatment of POAG complications as identified by a surgical report. Surgical procedures for treatment of POAG may include argon laser trabeculoplasty, selective laser trabeculoplasty, MicroPulse laser trabeculoplasty, and Ex-Press Mini Glaucoma Shunt. Previous studies which successfully developed phenotype algorithms with high PPV for ocular traits found that CPT codes for surgical procedures were sufficient for positively identifying individuals with cataracts(Peissig et al., 2012). We classified individuals whose records were positive for glaucoma case status but lacked sufficient clinical records to determine the sub-type of glaucoma as "potential cases". The criteria for potential cases includes *all* of the following: at least one mention of POAG ICD-9 (365.11), an ophthalmology/fundus CPT code (i.e. 92012, 92014, or 92250), glaucoma medication, and a text mention of "POAG" or "glaucoma" in a clinic note or problems history. Of the records reviewed, we identified 67 potential cases. Sixty-two individuals were determined to be false positives. Upon examination of the records, 28 of the 62 false positives were found to be diagnosed with another clinical sub-type of glaucoma, such as uveitic glaucoma (n=6), chronic angle closure (n=11), pediatric (n=2), neovascular (n=6), or glaucoma steroid responder (n=3). Reviews of other false positives lacked sufficient records or communications to determine glaucoma/POAG status. Ambiguous notes would state that a patient had "advanced glaucoma",

"ocular hypertension", or was a "glaucoma suspect". These individuals were excluded from case status as we could not reasonably verify a diagnosis.

Table 12: List of glaucoma medications used in validation of primary open-angle glaucoma cases

Drug name(s)	Generic drug name/active ingredient	Indications*
Diamox	Acetazolamide	Some types of glaucoma, epilepsy, and cardiac edema
Alphagan	Brimonidine	Prevention of elevated intraocular pressure (IOP)post operatively in individuals undergoing argon laser trabeculoplasty (ALT)
Iopidine	Apraclonidine	Prevention of elevated IOP post operatively in individuals undergoing ALT, argon laser iridotomy, or Nd:YAG posterior capsulotomy
Azopt	Brinzolamide	Treatment of elevated IOP for individuals with open angle glaucoma and ocular hypertension
Betoptic	Betaxolol	Treatment of elevated IOP for individuals with chronic open angle glaucoma and ocular hypertension
Cosopt	Dorzolamide and Timolol	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension
Trusopt	Dorzolamide	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension
Latanoprost/ Xalatan	Latanoprost	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension
Lumigan	Bimatoprost	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension
Timoptic(-xe)/ Betimol/ Istalol/ Blocadren	Timolol maleate	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension
Travatan	Travoprost	Treatment for elevated IOP in individuals with open angle glaucoma or ocular hypertension

Indications is a limited list of FDA approved uses as stated on the Drugs@FDA website as of October 4th, 2014: http://www.accessdata.fda.gov/scripts/cder/drugsatfda/

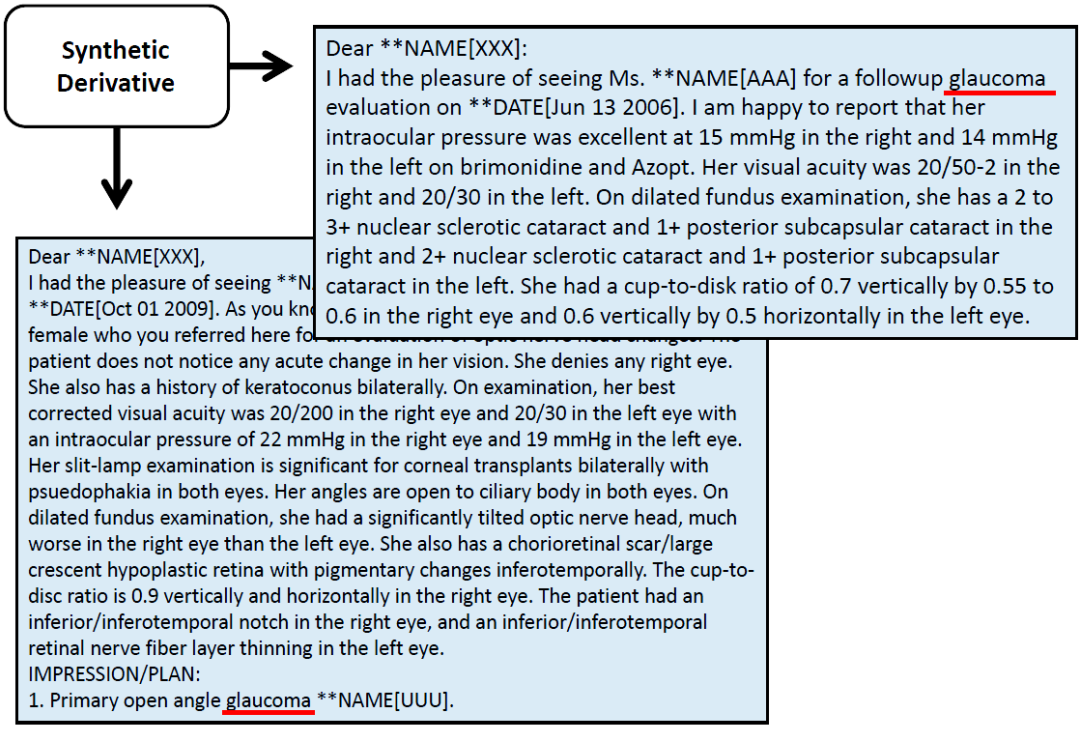


Figure 13: Screen shots of clinic notes from the Vanderbilt Synthetic Derivative as pertains to patients' glaucoma status.

Cup-to-disc ratio (CDR)

As mentioned, ocular phenotypes are difficult to extract from the VUMC EMR due to the lack of a structured electronic interface with VEI. Extracting quantitative phenotypes such as cup-to-disc ratios, the ratio of optic cup diameter to optic disc diameter, proved particularly challenging as there is not a uniform field or manner in which ophthalmologists report these data. We have found, however, that these data can be extracted via expression matching from referral letters and clinic notes sent between ophthalmologists and clinicians in the process of patient care (Figure 13). The following key word phrases were included in the search in reference to the right eye (O.D.) and left eye (O.S.): CDR, cup-to-disc ratio, cup-to-disk, and cup-to-disk ratio. Upon manual review of the results it became apparent that a more sophisticated approach was necessary to differentiate whether the numbers recorded were for the horizontal optic cup ratio, the vertical optic cup ratio, or the quotient of the horizontal-to-vertical CDR which ophthalmologists report interchangeably. When multiple measures of CDR were available for a patient, the most recent measurement was taken.

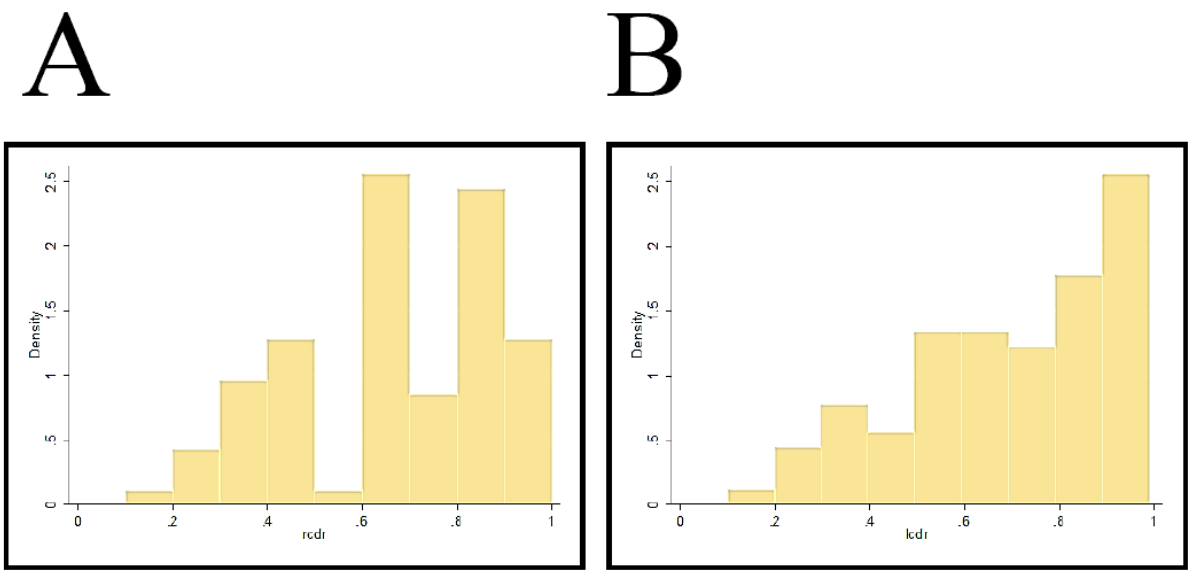


Figure 14: Histogram plots of the distribution of CDR (A – left eye, B – right eye) in EAGLE BioVU African American POAG cohort

A total of 132 POAG case records were positive for a CDR key word phrase, and all were manually reviewed for CDR. Only 9 of the 132 records (7 %) were missing a measure for CDR. The median values for CDR in this study are 0.7 (SD 0.22) in the right eyes and 0.7 (SD 0.23) in the left eyes (Figure 14 A and B).

Algorithm development for diabetic retinopathy

Similar to the development of the POAG case/control algorithms, the DR algorithms relied on the use of ICD-9, CPT, and free text searches to discriminate between cases and controls. The exception between the two is that we applied a previously vetted algorithm to first identify type-2 diabetics from among the African Americans included in the EAGLE Metabochip disease. The type 2 diabetes algorithm was developed as part of the Electronic Medical Records and Genomics (eMERGE) Network(Kho et al., 2012). The eMERGE Network utilized the clinical diagnostic criteria set forth by the American Diabetes Association and categorized individuals based on data extracted from an EMR. In brief, the eMERGE Network excluded individuals with ICD-9 codes for type 1 diabetes. For individuals with ICD-9 codes for type 2 diabetes, cases were required to have 1) a prescription for insulin or 2) a prescription for a type 2 diabetes medication and then in conjunction with either insulin/T2D medications the individual must have EITHER 1) more than two clinic visits with a recorded T2D diagnosis or 2) a prescription of type 2 diabetes medication prior to the insulin prescription. Cases were also identified among individuals without a prescription for insulin but with a prescription for type 2 diabetes medications, among individuals without prescriptions for either insulin or type 2 diabetes medications but who have abnormal glucose or glycated hemoglobin levels, and among those individuals without an ICD-9 code for type 2 diabetes but who are taking medication for type 2 diabetes and have abnormal glucose or glycated hemoglobin levels. Pregnant women were excluded as these patients may have developed gestational diabetes a condition separate from T2D (Supplementary Figure 2)

After the initial evaluation of African Americans in EAGLE Metabochip, 630 cases of T2D were identified. These individuals were then included in the following study for the identification of DR cases and controls.

Initial screening criteria for DR study population

As part of the eMERGE Network algorithm for T2D, individuals were excluded if medical records contained an ICD-9 for T2D (250.xx) before the age of twenty years as it is more likely this individual was a T1D.

Diabetic retinopathy cases

DR cases were individuals identified as having T2D as defined by the eMERGE Network and at least one mention of a DR ICD-9 code, excluding ICD-9 362.01, in conjunction with at least one mention of a clinic or ophthalmology CPT code (Figure 15 A and B). ICD-9 362.01 was excluded from the case definition given that background diabetic retinopathy, the presence of microaneurisms and hemorrhages, resolves on its own and does not impede vision.

Table 13: International Classification of Disease (ICD-9) Codes for diabetic retinopathy

DR ICD-9 codes	
362.0	Diabetic retinopathy
362.02	Proliferative diabetic retinopathy
362.03	Nonproliferative diabetic retinopathy
362.04	Mild nonproliferative diabetic retinopathy NOS
362.05	Moderate nonproliferative DR
362.06	Severe nonproliferative DR

ICD-9 362.01 was excluded from the case definition, as patients with background retinopathy typically see it resolve without treatment

Diabetic retinopathy controls

Initially, controls were cases of T2D whose medical records included a CPT code for clinic/ophthalmology visit (Figure 12 C) but did not include any of the three following: any ICD-9 for DR (Table 13), any CPT code for treatment procedures commonly used in the treatment of DR (Table 14), and any text mention of "diabetic retinopathy" or "retinopathy" in the problems lists. Lastly, controls with T2D duration of less than two years were removed as potential future incident cases. It should be noted that in the case of ophthalmology visit CPT codes it cannot be presumed that patients had a dilated fundus examination nor that physicians looked for or made note of the presence of any retinopathy.

After the initial review of controls, it was noted that many were being excluded. Part of the T2D routine medical plan at Vanderbilt includes an annual eye exam to screen for potential development of DR. As such, compliant T2D cases will have DR ICD-9 codes in their records marking these annual eye exams regardless of DR diagnosis. After additional review, we adjusted the inclusion/exclusion criteria for controls to allow for the inclusion of ICD-9 codes for background diabetic retinopathy (362.01) and mild nonproliferative DR (362.04). These two codes were selected based on initial case chart reviews that were determined to be false positives. DR controls with ICD-9 codes (362.01 and 362.04) were flagged for additional screening. Inclusion of ICD-9 codes 362.01 and 362.04 result in an additional 15 individuals meeting control criteria.

Table 14: Current procedural terminology codes for common treatment options for diabetic retinopathy

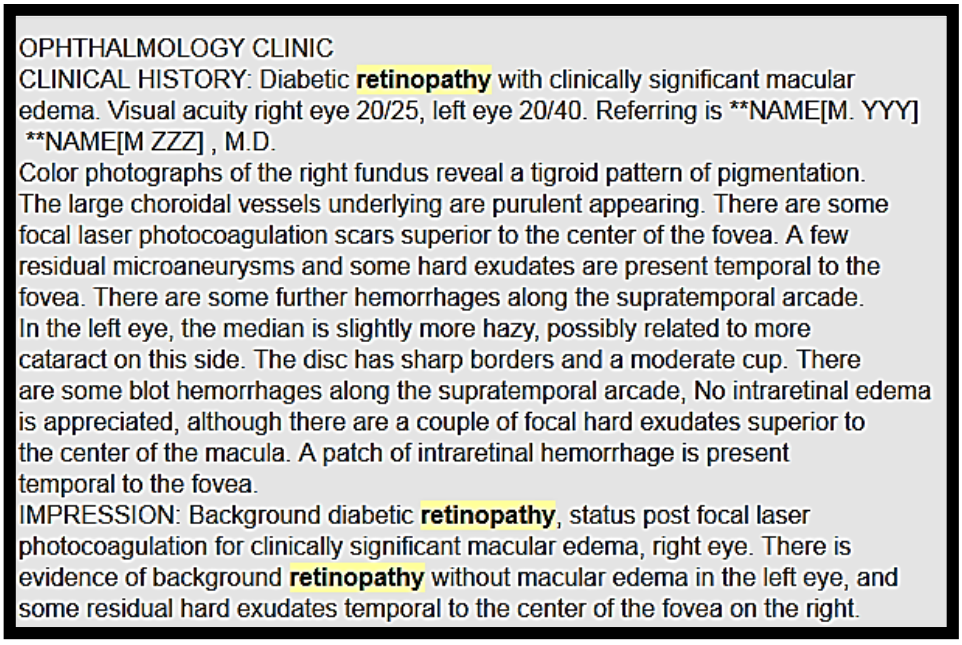
Procedures	**CPT Codes**
Impltj Intravitreal Drug Dlvr Sys Rmvl Vts Implantation of intravitreal drug delivery system (eg, ganciclovir implant), includes concomitant removal of vitreous	67027
Vtrc Mchnl Pars Plna Vtrc Mchnl Pars Plna Focal Endolaser Pc Vtrc Mchnl Pars Plna Endolaser Panrta Pc Vitrectomy Pars Plana Remove Preretinal Membrane Vitrectomy, mechanical, pars plana approach; Vitrectomy, mechanical, pars plana approach; with epiretinal membrane stripping Vitrectomy, mechanical, pars plana approach; with focal endolaser photocoagulation Vitrectomy, mechanical, pars plana approach; with endolaser panretinal photocoagulation Vitrectomy, mechanical, pars plana approach; with removal of preretinal cellular membrane (eg, macular pucker) Repair of retinal detachment; with vitrectomy, any method, with or without air or gas tamponade, focal endolaser photocoagulation, cryotherapy, drainage of subretinal fluid, scleral buckling, and/or removal of lens by same technique Repair of complex retinal detachment (eg, proliferative vitreoretinopathy, stage C-1 or greater, diabetic traction retinal detachment, retinopathy of prematurity, retinal tear of greater than 90 degrees), with vitrectomy and membrane peeling, may include air, gas, or silicone oil tamponade, cryotherapy, endolaser photocoagulation, drainage of subretinal fluid, scleral buckling, and/or removal of lens	67039 67040 67041 67113
Dstrj Loclzd Les Retina 1+ Sess Crtx Dthrm Destruction of localized lesion of retina (eg, macular edema, tumors), 1 or more sessions; cryotherapy, diathermy	67208

Manual review of diabetic retinopathy cohort

Review procedures followed those set forth in the manual review of the primary open-angle glaucoma cohort. Verification of case/control status was carried out utilizing a combination surgical reports, optometry, and ophthalmology clinic notes. Other data that was taken into account were medication lists, general and specialty clinic reports, clinical communications, and problems lists. DR cases were classified

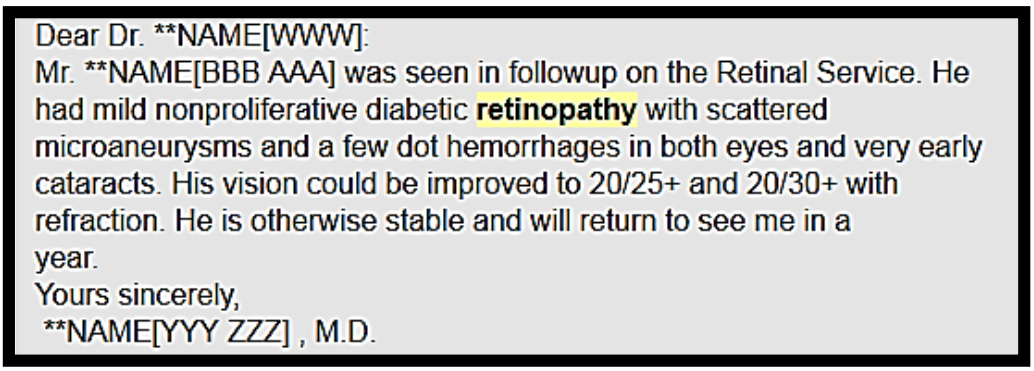
as either a definite or potential case. Definite DR cases met one of two criteria: 1) a written diagnosis by a Vanderbilt ophthalmologist/optometrist as pertained to a patient's exact diabetic retinopathy status (An example of this is seen in Figure 15 A and B) or 2) the patient's medical record contained *both* of the following: at least three independent mentions of a DR ICD-9 code and a surgical procedure for treatment of DR complications as identified by a surgical report. Surgical procedures for treatment of DR complications may include membrane peel, pars plana vitrectomy, scleral buckle, laser/photo coagulation, silicone oil, and intraocular/expanding gas. We classified individuals whose records were positive for text mention of DR case status but lacked surgical notes or ophthalmology clinic notes as "potential cases." The criteria for potential cases include *at least one* of the following: at least three mentions of DR ICD-9 with text mention of "diabetic retinopathy", text mention(s) for "surgery for diabetic retinopathy" in a clinic note or problems history, or thorough ophthalmology/optometry notes with diagnosis of background DR. Of the records reviewed, we identified 119 definite cases and 26 potential cases. Thirteen individuals were determined to be false positives. False positives tended to be T2D cases with thorough ophthalmology/optometry notes that explicitly state an individual was clear of signs of DR at the time of last visit (n = 7), records which lacked sufficient data to determine a diagnosis (n = 2), T1D (n = 1), or individuals diagnosed with other clinical forms of retinopathy such as "hypertensive" or "herpes retinitis" (n = 3).

Four-hundred and seventy-three individuals met DR control criteria. Of these, 100 controls were randomly selected for manual chart review. Each record was searched for text mention of "diabetic retinopathy" in all clinic notes and problems lists. Of these 100, all were clear of either a text mention of DR or else contained an ophthalmology/optometry report with negative findings for DR.



OPHTHALMOLOGY CLINIC
CLINICAL HISTORY: Diabetic **retinopathy** with clinically significant macular edema. Visual acuity right eye 20/25, left eye 20/40. Referring is **NAME[M. YYY] **NAME[M ZZZ] , M.D.
Color photographs of the right fundus reveal a tigroid pattern of pigmentation. The large choroidal vessels underlying are purulent appearing. There are some focal laser photocoagulation scars superior to the center of the fovea. A few residual microaneurysms and some hard exudates are present temporal to the fovea. There are some further hemorrhages along the supratemporal arcade. In the left eye, the median is slightly more hazy, possibly related to more cataract on this side. The disc has sharp borders and a moderate cup. There are some blot hemorrhages along the supratemporal arcade, No intraretinal edema is appreciated, although there are a couple of focal hard exudates superior to the center of the macula. A patch of intraretinal hemorrhage is present temporal to the fovea.
IMPRESSION: Background diabetic **retinopathy**, status post focal laser photocoagulation for clinically significant macular edema, right eye. There is evidence of background **retinopathy** without macular edema in the left eye, and some residual hard exudates temporal to the center of the fovea on the right.

Figure 15 A: Screen shot of a clinic note from the Ophthalmology Clinic at Vanderbilt University Medical Center, as seen in the Synthetic Derivative. Notes pertain to a patient's retinal eye exam and diagnosis.



Dear Dr. **NAME[WWW]:
Mr. **NAME[BBB AAA] was seen in followup on the Retinal Service. He had mild nonproliferative diabetic **retinopathy** with scattered microaneurysms and a few dot hemorrhages in both eyes and very early cataracts. His vision could be improved to 20/25+ and 20/30+ with refraction. He is otherwise stable and will return to see me in a year.
Yours sincerely,
**NAME[YYY ZZZ] , M.D.

Figure 15 B: Screen shot of a brief clinic note as seen in the Synthetic Derivative and pertains to a patient's retinal eye exam.

Results

Performance of primary open-angle glaucoma algorithms

For the POAG case algorithm, we determined that the PPV for definite cases was 51.6% with a sensitivity of 96.5% and accuracy of 76.3% (Table 15). When including potential cases, PPV was 76.7% with a sensitivity of 97.6% and an accuracy of 83.1% (Table 15).We manually reviewed the SD medical records of 300 randomly selected controls identified by the algorithm to calculate NPV, specificity, and accuracy. Of the 300 individuals identified as controls with this algorithm, five showed evidence of glaucoma at the time of review or else had mention of glaucoma in their records. As an additional review function, we performed a free-text search in all available documents for mention of the following words and abbreviations: glaucoma, fundus, opth, ophth, and vision. The shorthand identifiers such as fundus, opth, ophth [i.e. funduscopic eye exam, ophthalmology clinic, and opth (common abbreviation for ophthalmology)] are used throughout the clinical records especially within communications across clinics. This function was performed to ensure that algorithm-identified controls without qualifying ICD-9 or CPT codes also did not have evidence of POAG in the free clinical text. Performance of the POAG control algorithm was found to have a NPV of 98.3% with a specificity of 69.6% (Table 15).

Table 15: Evaluation of primary open-angle glaucoma phenotype algorithm in African Americans from EAGLE BioVU

	Sample Size	Manually reviewed	PPV	NPV	Sensitivity	Specificity	Accuracy
Cases	267	267	-	-	-	-	-
-Definite		138	51.6%	-	96.5%	-	76.3%
-Potential		67	76.7%	-	97.6%	-	83.1%
Controls	4813	300	-	98.3%	-	69.6%	-

Definite cases were individuals whose POAG status could be determined with high likelihood.

Potential cases were individuals whose medical records lacked sufficient information to make a definitive decision. Potential case results were calculated by including both potential and definite case numbers.

POAG study characteristics for African Americans

Of the 11,521 African Americans, 9,441 were over the age of 20 years and considered for downstream genetic association studies of POAG. Of these adults, 267 were identified as POAG cases (2.82%) and 4,813 as POAG controls. As might be expected for an age-related ocular disease, the median age of cases was older than controls (62 versus 54 years; Table 16). More than half of the cases and controls were female; and approximately half were hypertensive. On average, both cases and controls were obese (median body mass index > 30.0 kg/m^2).

Table 16: Study population characteristics of POAG definite cases and controls among African Americans in EAGLE BioVU

	Definite Cases > 20 yrs (SD)	Controls >40 yrs (SD)
N	138	4813
Age at Diagnosis (years)	62.0 (12.0)	--
Age at Last Clinic (years)	--	54 (11.7)
Sex (% female)	63.7	60
Hypertensive (%)	55.1	46.6
BMI (kg/m^2)	30.1 (6.7)	30.1 (8.0)
Diastolic (mm/Hg)	74.5 (8.1)	80 (33.6)
Systolic (mm/Hg)	134.5 (14.1)	124 (26.2)
Cholesterol (mg/dL)	183 (40.6)	161 (65.2)
HDL (mg/dL)	52.5 (25.0)	53 (38.6)
LDL (mg/dL)	103 (42.9)	99 (50.7)
Triglycerides (mg/dL)	125 (76.3)	98 (67.8)

Median values were calculated for the following: Age at POAG diagnosis was determined by the date of when POAG ICD-9 (365.11) was first mentioned in the records. Age at last clinic visit (LCV) was taken as the date of the last CPT mentioned in the records for controls. An individual was classified as hypertensive if he/she met one of three criteria: systolic blood pressure > 140 mm/Hg, diastolic blood pressure > 90 mm/Hg, or on hypertension medications all within a two year window of when they were diagnosed with POAG in cases and a two year window of their LCV date for controls. Blood pressure (systolic and diastolic), lipids (total cholesterol, high-density cholesterol, low-density cholesterol, and triglycerides), and body mass index (height and weight) were calculated from labs or measurements within two years of POAG diagnosis or LCV. Abbreviations: standard deviation (SD)

Performance of diabetic retinopathy algorithms

The PPV for definite DR cases was 75.3% with a sensitivity of 1.0% and accuracy of 84.8% (Table 17). With the inclusion of potential cases, PPV increased to 91.7% with a sensitivity of 1.0% and an accuracy of 94.9% (Table 17).We manually reviewed the SD medical records of 100 randomly selected controls identified by the algorithm to calculate NPV, specificity, and accuracy. Performance of the DR control algorithm was found to have a NPV of 1.0% with a specificity of 71.9% (Table 17).

Table 17: Evaluation of diabetic retinopathy phenotype algorithm in African Americans from EAGLE BioVU

	Sample Size	Manually reviewed	PPV	NPV	Sensitivity	Specificity	Accuracy
Cases	158	158	-	-	-	-	-
-Definite		119	75.3%	-	1.0%	-	84.8%
-Potential		26	91.7%	-	1.0%	-	94.9%
Controls	473	100	-	1.0%	-	71.9%	-

DR study characteristics for African Americans

Of the 1,672 African Americans screened for T2D, 145 were identified as DR cases (i.e. definite and potential) (30.6%) and 473 as DR controls. The median age of cases was older than controls (62.2 versus 49.8 years; Table 18). Controls were predominately female and tended to have better control of their diabetes (HbA1c = 6.7%) compared with cases. On average, both cases and controls were obese (median body mass index > 30.0 kg/m^2).

Table 18: Median demographics of EAGLE BioVU Metabochip African Americans with diabetic retinopathy

	Diabetic retinopathy	
	cases	controls
N	145	473
Age (yrs)	62.2	49.8
% female	57%	66%
BMI	32.7	30.4
Systolic (mmHg)	140.1	127.9
Diastolic (mmHg)	77.1	77.4
Cholesterol (mg/dL)	221.1	208.2
Glucose (mg/dL)	187	100.2
LDL (mg/dL)	95	91.3
HbA1c	9.5	6.7

For Controls, anyone under 40 years of age were excluded. Cases include individuals ascertained as definite and potential

Discussion and Summary

Overall, the POAG algorithm identified 267 cases and 4813 controls among African Americans from EAGLE BioVU. After manual review we determined that the algorithm only had a positive predictive value of 51.6% to identify definite POAG cases with a sensitivity of 96.5% and an accuracy of 76.3%. The control algorithm demonstrated a high performance with a negative predictive value of 98.3% and a specificity of 69.6%. Despite the limited positive predictive value of the algorithm in defining POAG case status, our strategy overall provides a starting point in the absence of digital photographs to identify definitive cases of POAG from amongst over 10,000 African American adults. Manual review of these records would be prohibitively time- and resource-consuming. An alternative strategy to triaging records for manual review is the requirement of a single ICD-9 code for POAG. Based on these relaxed criteria, we identified 309 African Americans in EAGLE BioVU whose EMR contained an ICD-9 code for POAG. After review of these additional individuals for definite POAG case status, we determined that all measures of performance were lower (PPV = 47.2%, specificity = 64.4%, and accuracy = 72.4%) compared with the algorithm developed here.

The age and gender composition of the EAGLE BioVU definite POAG cases (Table 16) and controls differ when compared with other clinical and epidemiologic cohorts. More than half of EAGLE BioVU definite POAG cases (63.7%) and controls (60%) are female. In contrast, the International Consortium of African Ancestry Research in Glaucoma (ICAARE-Glaucoma; n=2,150) study identified fewer female cases and controls(Liu et al., 2013b): African American cases (49.5%), African American controls (54.6%), Ghanaian cases (43.9%), and Ghanaian controls (57.3%). Also, the mean age at diagnosis for ICAARE-Glaucoma African American cases (57.0 yrs) is younger than EAGLE BioVU POAG cases (62.0 years) while the ICAARE-Glaucoma control group (59.4 yrs) is older compared with EAGLE BioVU POAG controls (54 yrs). These demographic differences are likely the result of differences in recruitment or ascertainment of African Americans within their respective communities, which can potentially introduce heterogeneity in downstream studies.

As evident by this study, broad phenotyping of a disease cohort is well within the capabilities of an EMR, even one with limited access to clinical data, through the use of simple text-mining techniques incorporating pattern matching and structured data from the EMR. The algorithm designed here for POAG is stringent yet the PPV (51.6%) of this POAG algorithm for definite case status is well below the threshold of 95% adopted by consortia such as the eMERGE network(McCarty et al., 2011) for use in large-scale genetic association studies(Kho et al., 2012; Peissig et al., 2012). However, the addition of potential cases substantially increases the PPV to 76.7%. If we had designed the algorithm to merely detect an individual with a "general" glaucoma classification, the PPV increased to 87.2%. Furthermore, we have developed a highly discriminatory algorithm (NPV 98.3%) that can identify ocular controls. In the development of phenotype algorithms there is the constant tradeoff between overly strict or vague criteria resulting in loss of cases or misclassification of subjects, both of which will lead to a loss in statistical power in downstream genetic association studies.

The DR case/control algorithms notably exceeded the performance of the POAG case/control algorithms. Under definite case status, the DR case algorithm had a PPV = 75.3%, accuracy = 84.8 %, and sensitivity

= 1.0. The DR control algorithm had a NPV on par with the POAG control algorithm (1.0% vs 98.3%) and a specificity of 71.9% versus 69.65, respectively. The marked increase in performance between the two can reasonably be attributed to the disparity in health care and counseling provided to diabetics in comparison to non-diabetics. As mentioned previously, when a patient is diagnosed with diabetes in the Vanderbilt health care system they are advised to seek the attention of health care providers at the diabetes clinic and subsequently to seek annual eye exams for the assessment of diabetes-related eye changes (e.g. diabetic retinopathy and diabetic macular edema). Irrespective of compliance, diabetics are closely monitored and followed for the duration of a patient's life. In contrast, fifty percent of African American glaucoma patients go undiagnosed in the United States(Ladapo et al., 2012).

Although most clinics within a health care organization that maintain an EMR adopt the digital system, occasionally a clinic may be excluded from converting to an all-digital interface or choose to opt out. The reasons for these exclusions may vary, but the results of missing data can impact a study in many ways. Evaluation of the algorithms was limited by the lack of available funduscopic images (i.e. the "gold standard") for validation, a limitation of this study. Missing data limited our ability to identify definite cases of POAG and DR. Several cases identified here could only be classified as potential cases given the lack of sufficient clinical records to determine the sub-type of glaucoma. This limitation underscores the need for better implementation of EMRs across healthcare organizations for use in biomedical research.

Missing data can also introduce misclassification bias into studies. The control algorithms developed here were designed around the concept that an individual is free of DR/POAG. However, without a complete medical work-up, there is the potential that a control is an undiagnosed case. Misclassification and potential ascertainment bias is also possible in case identification where cases are only those individuals who have been evaluated by a specialist and are therefore potentially extreme or overly symptomatic cases. A well-known barrier for individuals seeking medical attention is low socioeconomic status which disproportionally affects African Americans(Anderson et al., 2004). This may in part explain the limited number of African American POAG cases (n=267) representing only 2.82% of African Americans in

EAGLE BioVU. Given the expected prevalence of POAG at 4-5% among African Americans (Friedman et al., 1999, 2004a), presumably the cases in EAGLE BioVU are only being diagnosed once vision loss becomes severe. The same can be said for DR cases (n=119) which at a rate of 21.6% of the EAGLE BioVU T2D population over 40 years (n=431) is notably lower than would be expected in the general U.S. African American population of diabetics over 40 years (36.7%).

Limitations and Conclusions

Our study has a number of limitations. The algorithms were developed at VUMC under the restrictions of a de-identified EMR, which limits access to certain data types, and the SD, which does not currently contain all pertinent medical records such as digital fundus photographs. As mentioned before, the VEI does not utilize an electronic interface with structured fields. Lack of these structured data makes it impossible to search the SD for ophthalmology exam results. Without the actual exam results it is not possible for researchers to definitively ascertain an individual's ocular disease status without a clinician's explicit note detailing a patient's results in writing. Due to these limitations it is unclear if our algorithms can be exported for use in the EMRs of other medical institutions. Also, only one investigator performed the manual review making it impossible to assess intra- or inter-grader variability. This investigator was aware of the algorithm's determination of case and control status, which may have introduced bias. However, we modeled our algorithm development process after the electronic Medical Records and Genomics (eMERGE) Phenotype Working Group workflow(Kho et al., 2011; Newton et al., 2013). In the eMERGE Network, algorithm development and assessment are an iterative process. That is, content experts design the initial algorithm and deploy it. After a round of manual reviews and performance calculations, the algorithm is altered, re-deployed, and re-evaluated. There is potential bias in reviewing and assessing the performance of the algorithm when adopting the eMERGE workflow.

Additional complications in utilizing EMRs is that a scarcity of notes or large time gaps between exams for the identification of cases and controls may lead to incident cases or false negatives. In the instance of DR cases, patients with mild to moderate non-proliferative DR have been shown to see their DR resolve on its

own with little to no medical intervention. So depending on time between exams, incident cases may develop.

Utilizing clinic notes to extract CDR measurements introduces a number of biases, most notably human error. In ophthalmology clinics, current clinical practice involves a manual examination of the optic nerve head and the screener's subsequent assessment or perception of the CDR. The accuracy of the CDR will rely on the screener's experience and examination conditions (e.g., monoscopic versus stereoscopic conditions). The assessment of CDR also varies across examiners. Intra- and inter- reader variation can reach as much as 0.2 disc diameters(Arthur et al., 2006; Varma et al., 1992). Therefore, the interpretation of CDR measurements extracted from clinical notes in an EMR system should be approached with a bit of skepticism.

Despite the limitations in portability and data access, we were able to define primary open-angle glaucoma and diabetic retinopathy from the VUMC's SD. We have a diverse population connected to a depth of medical data, even if not all of it is easily searchable. Our study has made available more case counts for African Americans with ocular disease. And, BioVU continues to accrue samples as well as update the medical records associated with samples already collected; therefore, the accrual of additional cases is anticipated. The ability to extract ocular phenotypes from EMRs will provide researchers with previously unaccessed datasets to further advances in ocular genetic research and vision-loss prevention.

Acknowledgements

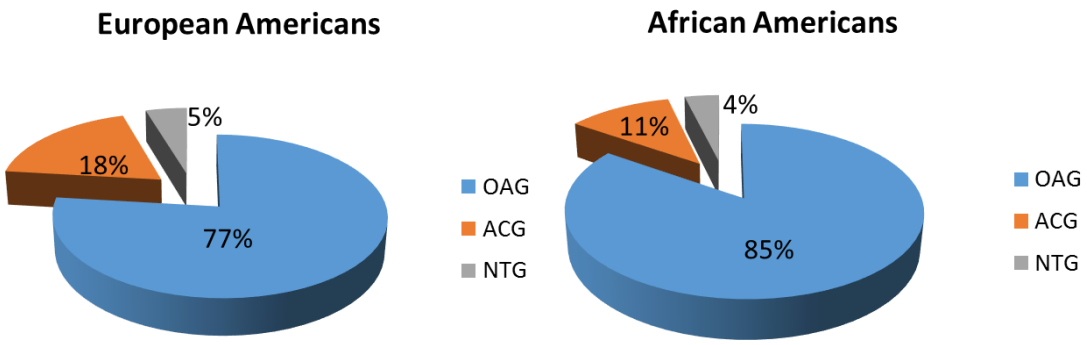
The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core provided computational and/or analytical support for this work.

CHAPTER IV

THE ASSOCIATION OF COMMON GENETIC VARIATION IN PRIMARY OPEN-ANGLE GLAUCOMA IN AFRICAN AMERICANS

Introduction

Primary open-angle glaucoma (POAG) is the most common form of glaucoma and a major driver of irreversible vision loss (Figure 16). Understandably the loss of vision impacts an individual's quality of life through loss of independence, income, and mobility. Less tangible indicators of life quality include mental health with depression occurring more frequently in the elderly with glaucoma(Wang et al., 2012). Research by the Salisbury Eye Evaluation group has even suggested that visual field loss, a defining component of glaucomatous disease, is the fundamental vision component responsible for falls in older adults(Freeman et al., 2007). This is in comparison to visual acuity, contrast sensitivity or stereoacuity. The visual field is the spatial array of visual sensations available for observation(Smythies, 1996) which in humans includes a field 60° nasally, 60° superiorly, 70° inferiorly, and 100° temporally(Anderson and Patella, 1998). The



Abbreviations: open angle glaucoma (OAG), angle closure glaucoma (ACG), and normal tension glaucoma (NTG)

Figure 16: Pie graphs of the approximate breakdown of major glaucoma types in European Americans and African American within the United States[5–7].

visual field can be altered by disease onset in the vision pathway (i.e. retina, optic nerve, brain) or physiologic trauma. Typical forms of field loss are altitudinal field defects affecting vision above or below the horizons, bitemporal hemianopia involving loss of peripheral vision, or scotomas which are regions of vision loss affecting the central vision.

Clinical features and diagnosis of primary open-angle glaucoma

Glaucomatous eye changes may be diagnosed in multiple ways, one of which includes testing for loss of visual fields. Visual field testing can be carried out manually (i.e. confrontation visual field) or by automated or frequency doubling perimetry, whereby the extent of a patient's visual field can be measured and potential blind-spots identified. Blind spots in the visual field can be indicative of eye disease. In the case of glaucoma, visual field deficits occur in a particular pattern and location(Asman and Heijl, 1992). Most glaucoma patients suffer some visual field loss within the central 24-30°(Ballon et al., 1992).

Additionally, an optometrist or ophthalmologist may perform a visual examination of the optic cup-to-disc ratio (CDR), a measurement used to assess the thickness of the nerve fiber layer. In a healthy eye (Figure 17, far left), the optic disc is the physiological location where the retinal ganglion cell axons, forming the optic nerve bundle, and ocular blood vessels exit the retina. The nerve fibers are visualized as the pink rim of the disc, while the central white cup is an area lacking any nerve fibers. Healthy eyes maintain a CDR

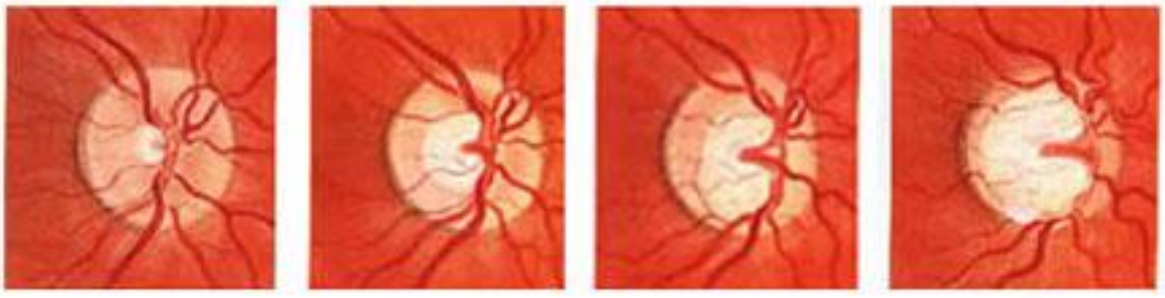


Figure 17: Funduscopic image of the optic nerve head in a healthy eye (far left) as it progresses to glaucoma (far right) as shown by signs of physiological cupping. This image was adapted from the website of the Glaucoma Associates (http://www.glaucomaassociates.com) accessed February 4, 2015.

of 0.3. As nerves in the optic bundle begin to atrophy and die, the pink rim decreases in size. The subsequent increase in the central cup is a typical sign of glaucomatous changes. CDRs greater than 0.7 (the third and fourth pane of Figure 17) are clinically significant indications of glaucoma.

Lastly, elevated IOP or ocular hypertension (i.e. 21 mmHG ≥ IOP) may be used as part of a glaucoma diagnosis. Its utilization in diagnosing POAG has waned in part because it is not an independent predictor of POAG(Dielemans et al., 1994; Levine et al., 2006; Shah et al., 1999; Wolfs et al., 2000). Individuals with ocular hypertension do not definitively go on to develop glaucoma; inversely, individuals with normal IOP are still at risk of developing POAG. The presence of OAG in the absence of elevated IOP is commonly referred to as normal tension glaucoma (NTG). Previously NTG and POAG were considered two separate clinical subtypes of glaucoma, but the current consensus wavers as to whether they are separate entities or else IOP dependent/independent ends of a spectrum. IOP is the only known modifiable factor for POAG with viable treatment options shown to slow disease progression(Heijl et al., 2002; Kass et al., 2002; 2000b).

Epidemiology of primary open-angle glaucoma

The risk factors and mechanisms underlying POAG risk and progression have been studied extensively across populations. Yet for each new risk variable that these studies identify, the question emerges as to how these factors are biologically impacting disease. To recapitulate, African ancestry(Tielsch et al., 1994), age(Leske et al., 1995), myopia(Pan et al., 2013), and IOP(Chandrasekaran et al., 2006; Jiang et al., 2012; Leske et al., 1995) are well established risk factors for POAG. IOP in particular is such a strong modifier of POAG risk that some clinicians still use it as a diagnostic criterion. Mechanistically high IOP can damage the optic nerve and retina by inflicting structural damage at the optic nerve head, reducing retinal blood flow, and increasing the expression of cytokines(Morrison et al., 2005). Part of this mechanical theory of POAG involves the lamina cribrosa, a mesh-like network of collagen fibers located at the optic nerve head. It is through the lamina cribrosa that the retinal ganglion cell axons pass before forming the optic nerve bundle. The lamina cribrosa forms a barrier between the inside of the eye and the surrounding

tissue(Morgan-Davies et al., 2004). Changes in IOP have been hypothesized to lead to structural changes in the lamina cribrosa that in turn pinch and damage the optic nerve fiber tendrils and blood vessels. Correlations between myopia and POAG risk observed in epidemiologic studies may in fact represent the mechanical properties of myopia, such as an increase in axial length of the ocular globe. Axial length has been independently associated with an increase in risk of OAG, which is postulated to occur as a result of increased stress on the sclera. As IOP rises the mechanical stress is greater in axially longer eyes(Kuzin et al., 2010). In the prospective Los Angeles Latino Eye study (LALES), individuals with axial myopia had a higher incidence of developing OAG particularly with higher baseline levels of IOP(Jiang et al., 2012).

Additional predictors of POAG include lack of vision insurance, waist to hip ratio (WHR), female gender, and thinner central corneal thickness (CCT)(Gordon MO et al., 2002; Jiang et al., 2012). The role of general obesity or BMI in POAG is questionable(Jiang et al., 2012; Leske et al., 1995) but evidence supports that abdominal obesity or WHR may be a more substantial predictor of POAG risk in women compared with men(Jiang et al., 2012). The total number of women affected by POAG exceeds that of men in the U.S.(Vajaranant et al., 2012b). Yet as POAG is an age-related disease, these gender trends can be explained somewhat by women's longer life spans compared with men. The life expectancy in the U.S. for women is currently 81.0 years vs. men at 76.2 years. POAG rates by gender vary from study to study(Dielemans et al., 1994; Leibowitz et al., 1980; Leske et al., 1994; Reidy et al., 1998) yet some find that male sex increases risk of POAG after adjusting for age(Rudnicka et al., 2006). Female sex hormones may be protective against the onset of POAG. Oestrogens regulate signaling pathways involved with neuronal differentiation, synaptic plasticity of neurons and cell migration/survival and death(Arevalo et al., 2011). Estrogen and/or progesterone receptor mRNAs are found in numerous ocular tissues such as the bulbar conjunctivae, cornea, RPE, lens, retina, and iris(Gupta et al., 2005; Wickham et al., 2000). Current research is delving into the use of hormone therapy replacement (HTR) in elderly women to ascertain efficacy in treating and preventing eye disease. Use of estrogen/progestin in postmenopausal women before onset of visual field loss reduced the risk of POAG in the Nurse's Health Study(Pasquale et al., 2007).

Pathogenicity of glaucoma has also been postulated to occur, not as a consequence of the accumulation of specific mechanical or vascular stressors, but as a result of axonopathy in RGC (i.e., the disruption of normal axon function)(Calkins, 2012). Classically, glaucoma associated vision loss is attributed to degeneration and apoptosis of RGCs and their axons, which form the optic nerve bundle, due to IOP-dependent or IOP-independent stressors on a background of age-related factors. Currently, mouse and rat models of glaucoma provide evidential support that glaucomatous optic neuropathy occurs primarily as a result of axonal transport deficits in the presence of varying levels of IOP "sensitivity"(Buckingham et al., 2008; Danias et al., 2003; Vidal-Sanz et al., 2012). These transportation deficits affect the movement of mitochondria to and away from areas of the axon requiring greater availability of ATP for hydrolyzation(Hollenbeck and Saxton, 2005). With the accumulation of mitochondria in key points of the axon, reactive oxygen species may build up and incur damage to the axonal milieu triggering Wallerian-like degeneration at distal points(Calkins, 2012, 2013).

Genetics of primary open-angle glaucoma

The genetics of POAG is a complicated domain, lacking a clear Mendelian mode of inheritance, and further muddled by clinical terminology and diagnostic criteria. It can be argued that this optic neuropathy is a singular disease of complex, multifactorial etiology of which science and medicine have yet to completely tease apart the particulars or else it is the result of multiple diseases being classified under one heading. Either way, genetic studies to-date have only had limited success, the majority of which stemmed from the study of the quantitative traits associated with general glaucomas as discussed in Chapter 1. Still, early genetic studies identified *MYOC ,OPTN,* and *WDR36*(Monemi et al., 2005; Rezaie et al., 2002; Stone et al., 1997) as genes linked to susceptibility of POAG. The myocilin gene (*MYOC*), which was originally named the Trabecular Meshwork-Inducible Glucocorticoid Response Protein gene (*TIGR*), was discovered in 1997(Polansky et al., 1997). *MYOC* is expressed in several ocular tissues including the sclera, choroid, cornea, and the trabecular meshwork(Adam et al., 1997; Ortego et al., 1997; Tamm et al., 1999) where it is hypothesized that mutated versions of the protein are not being adequately secreted into the aqueous

humor(Jacobson et al., 2001). This buildup of mutated MYOC proteins can prevent the flow of fluids through the trabecular meshwork and increase IOP resulting in damage to the optic nerve. The pathological severity of *MYOC* mutations likely differ from one to another. Some *MYOC* mutations are found to segregate with juvenile glaucoma under a Mendelian mode of inheritance in families of varying ethnic and racial origins(Braghini et al., 2013; Geyer et al., 2011; Mimivati et al., 2014; Waryah et al., 2013) suggesting these variants evoke a greater loss in protein stability. Mutations of a lesser deleterious nature likely account for the population-based studies that have found that *MYOC* mutations contribute to risk of POAG in a small subset of patients (3-4%) whose conditions are not solely explained by *MYOC* variants(Alward et al., 2002; Fingert et al., 2002). In the last decade, large-scale GWAS studies exploring associations with smaller effect sizes have reproducibly found that variants in the *CAV1/CAV2*, *CDKN2B-AS1* and *SIX1/SIX6I* genes influence POAG risk in European-descent and Japanese populations(Nakano et al., 2012; Osman et al., 2012; Thorleifsson et al., 2010; Wiggs et al., 2012).

This evidence supports that genetic variation drives, in part, POAG pathology in an additive fashion but much remains to be determined. Additional genetic factors that have yet to be discovered are hypothesized to drive POAG risk. And, the factors that drive POAG incidence differences observed across racial/ethnic groups have yet to be determined. One hypothesis is that there are population-specific genetic associations in addition to trans-population association that contribute to POAG risk. To identify these population-specific and trans-population genetic factors, we conducted hypothesis-testing and hypothesis-generating genetic association studies in African Americans with and without POAG drawn from a clinical cohort in Nashville, Tennessee.

Replication and generalization of published POAG risk variants in primary open-angle glaucoma risk in an African American population

African Americans experience a greater burden of risk for development of POAG compared to European Americans. We hypothesize that this due, in part, to a combination of greater prevalence of common and rare interpopulation risk variants and population-specific variation. To test this hypothesis, we will first test directly highly replicated nuclear variants for an association with POAG in the EAGLE BioVU dataset(Crawford et al.). It is expected that trans-population risk SNPs discovered in Europeans or Asians may generalize or be associated with POAG in African Americans. However, because GWAS variants often are tags or surrogates as opposed to the true causal or functional variant, I will also test gene regions fine-mapped by the Metabochip to identify the most strongly associated SNPs for POAG in African Americans. African Americans are an ideal population for fine-mapping given the lower levels of linkage disequilibrium (statistical correlation between SNPs) compared with European-descent populations(Teo et al., 2010). Finally, I will also perform a hypothesis-generating experiment where I test all genotyped SNPs for an association with POAG.

The EAGLE BioVU dataset includes African Americans that have been genotyped on the Metabochip, a custom array of ~200,000 SNPs commonly used for genetic studies of metabolic and cardiovascular traits(Buyske et al., 2012a; Crawford et al., 2013). Metabochip contains variants known to be associated with optic traits, such as vertical cup-to-disc ratio (*MTAP* rs1063192) and optic cup size (*ATOH7* rs3858145). There are ~55 variants known to be associated with various ocular diseases and ocular quantitative traits that are directly genotyped on the chip or amicable to imputation. Case-control tests of genetic association will be performed on select SNPs and gene regions on Metabochip that have been previously identified as risk modifiers in POAG.

CDKN2B-AS1 is a gene famously associated with cardiovascular disease which has also shown pleiotropy with other conditions such as cancer(Chen et al., 2013, 2014), T2D(Scott et al., 2007; Zeggini et al., 2007), endometriosis(Buggio et al., 2014), and notably glaucoma(Burdon et al., 2011a; Liu et al., 2013b;

Nakano et al., 2012). Primary analyses for common variants in POAG will focus on the *CDKN2B-AS1* gene region which multiple studies have found harbors SNPs that reproducibly contribute to risk of POAG, cup-to-disk ratios, and NTG. The Metabochip array targeted the *CDKN2B-AS1* region for fine-mapping with 459 SNPs that were selected based on the first iteration of the 1000 Genomes dataset which included African-descent reference samples. Additional analyses will include the *SIX6* gene variant (i.e. rs10483727) that has also been repeatedly implicated in POAG risk.

Methods

Metabochip Genotyping of BioVU African American Samples

As part of the PAGE I study(Buyske et al., 2012a; Matise et al., 2011b), EAGLE selected all non-European Americans from BioVU as of 2011 for genotyping on the Metabochip. A total of 11,521 African Americans, 1,714 Hispanics, and 1,122 Asian samples in BioVU were genotyped(Crawford et al.). The Metabochip is a custom array from Illumina designed for replication and fine mapping of metabolic and cardiovascular traits. Many of these traits are related to T2D, and genotypes are available for variants that were found to be of genome-wide significance of any phenotype deposited in the National Human Genome Research Institute (NHGRI) GWAS Catalog as of 2009. This includes several ocular related SNPs (Supplemental Table 1). Metabochip also contains common and rare variants selected from HapMap and the 1000 Genomes project. Fine mapping regions cover 257 loci selected from SNPs that reached genome-wide significance from select consortium meta-analyses. This chip has proven effective in identifying causal variants of lipid traits in African Americans(Buyske et al., 2012b) that were previously missed in GWAS studies that used genotyping chips designed for European Americans. Advantages of the Metabochip over other GWAS chips are the areas of fine mapping mentioned above (Figure 18). These areas include variants that are uncommon or not present in Europeans and get around part of the problem of differences in linkage disequilibrium architecture between populations. Fine-mapping in diverse populations allows researchers to narrow the risk interval in association studies and to discover population specific putative, causal variants.

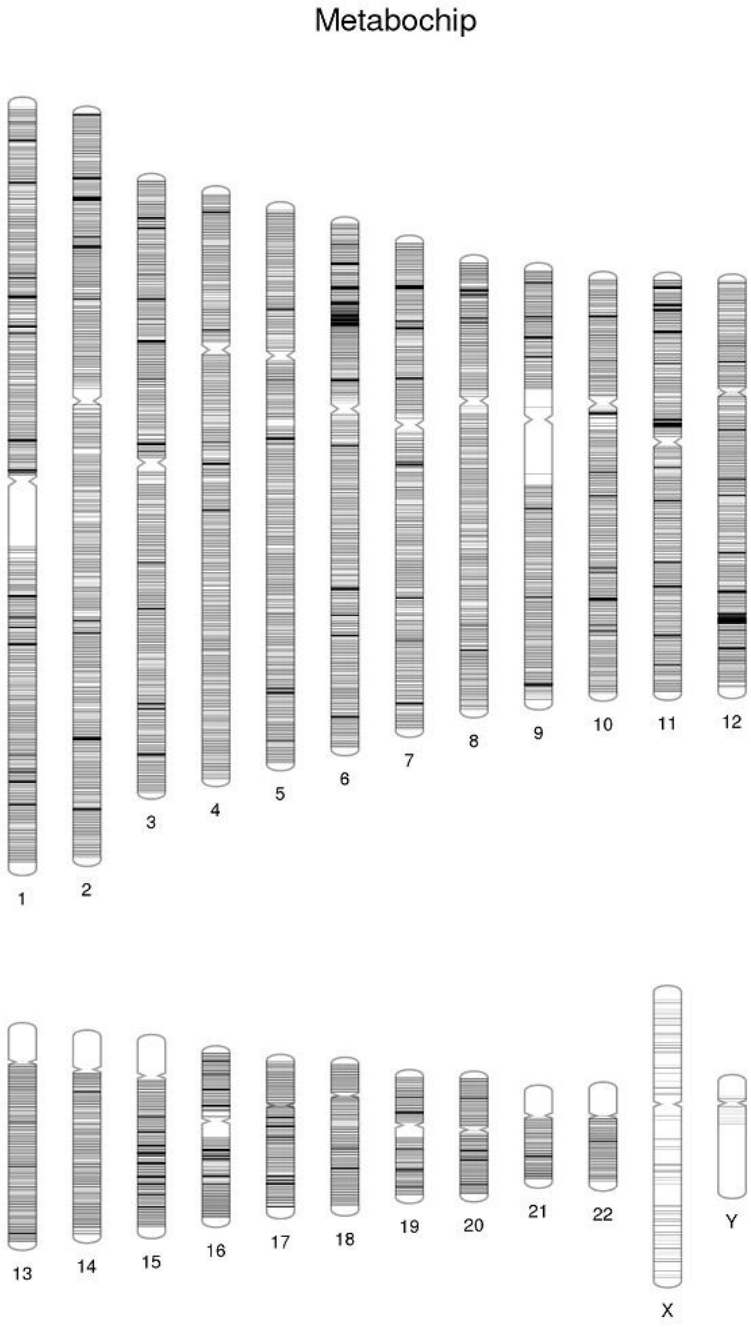


Figure 18: Phenogram of the genomic location of the Metabochip SNPS that passed quality control. Each SNP is represented by a line with fine mapped regions appearing as heavier bands of black. Not shown are the mitochondrial variants.

Samples were genotyped using the Metabochip following the manufacturer's protocol (Illumina, Inc; San Diego, CA.), and 360 HapMap samples, including YRI samples, were genotyped for PAGE-wide cross-study QC standards(Crawford et al., 2013). A description of the genotyping protocols and quality control measures has been previously published (Buyske 2012). In brief, genetic variants were evaluated for deviations from Hardy Weinberg Equilibrium, which may be a result of poor genotyping. Variants with a genotyping call rate < 95% were removed from further analysis. Principal components (PC) were calculated using EIGENSOFT(Patterson et al., 2006). At the sample level, DNA samples with poor sample call rate (<95%), sex discordance, or evidence of cryptic relatedness (based on identify-by-descent; Figure 19) were removed from analyses. For cryptic relatedness we drop the individual with the lowest genotyping rate. Before quality control, 459 SNPs and 138 case samples with DNA were available for study. Post quality control, 286 SNPs (with HWE $p > 0.001$ and minor allele frequency >5%) and 130 case samples remained for analysis.

Study Population and Phenotyping

The study population and phenotyping of POAG cases and controls have been described in detail in Chapter 3: Utilization of Electronic Medical Records systems for ascertainment of case/control cohorts in genetic association studies. Briefly, 138 African Americans POAG cases and 4,813 African American POAG controls were identified in the Vanderbilt SD that were genotyped on the Metabochip. These individuals were further evaluated for inclusion in the following genetic association analysis.

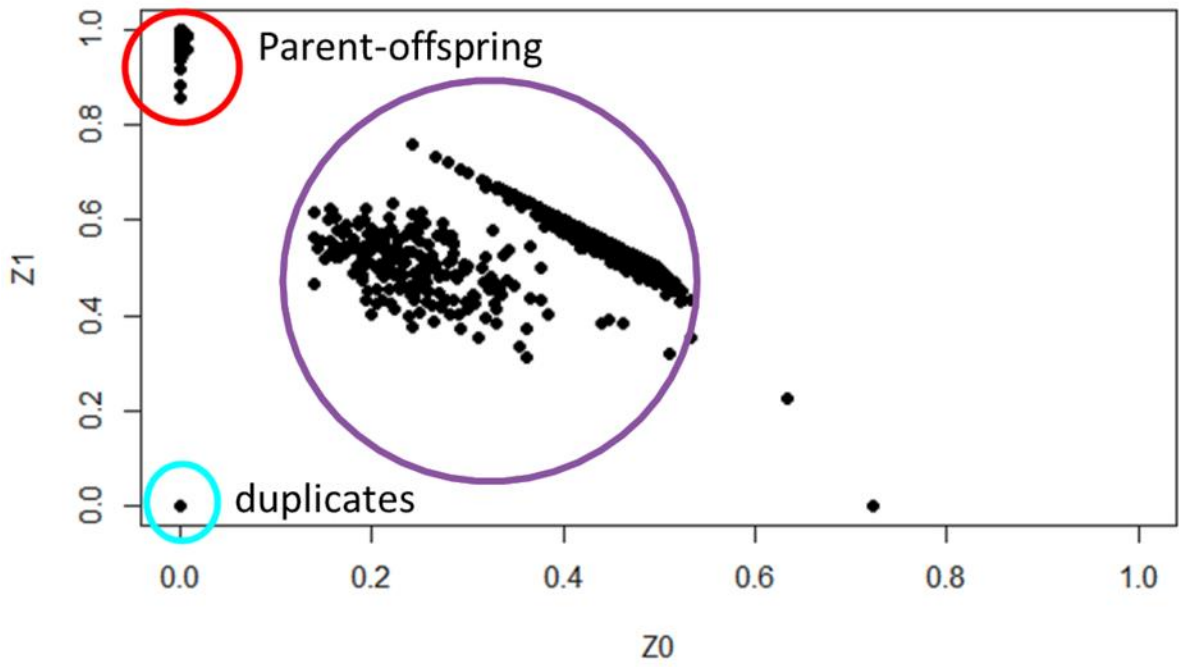


Figure 19: Rplot of the identity-by-descent (IBD) estimation results for pairwise interactions with a phi-hat greater than 0.25 in EAGLE BioVU African Americans (N=11,014). Each dot represents an individual. We identified 17 pairs of monozygotic twins (duplicates in the teal circle), 385 pairs of parent-offspring marked by the red circle, 209 pairs of full siblings, and 289 pairs of half siblings. Full sibling, half sibling, and others fall within the purple circle.

Statistical methods

We tested for an association between *POAG* and 286 SNPs in the *CDKN2B-AS1* region of chromosome 14 in African Americans and then additionally for all common variants (MAF > 0.05) on the Metabochip that passed quality control. Individuals included in this analysis were those identified as "definite" POAG cases over the age of 20 years and POAG controls over the age of 60 years. Age was defined as age at diagnosis in cases and age at last clinical exam in controls.

Each SNP was tested for an association using logistic regression assuming a log-additive genetic model 1) adjusted by age, sex, and the first three principal components (PC) (Supplementary Figure 3) and 2) age, sex, first three PCs, and median diastolic blood pressure (Supplementary Figure 5). Analyses were conducted using PLINKv1.90(Purcell et al., 2007).

Ethics statement

BioVU follows an opt-out model for DNA sample acrrual(Roden et al., 2008). That is, DNA is collected from discarded blood samples remaining after routine clinical testing and is linked to de-identified medical records. According to the Vanderbilt IRB and the Federal Office of Human Research Protections provisions, the Vanderbilt protocol is considered nonhuman subjects research (The Code of Federal Regulations, 45 CFR 46.102 (f)). The IRB at Vanderbilt University approved this research.

Results

Population characteristics

A total of 138 African American POAG cases and 1,376 controls EAGLE BioVU for analysis. In general, cases were more likely to be female and tended to have a higher BMI and elevated cholesterol levels in comparison to controls. The cases also tended to be younger on average (62 years) compared with controls (67.3 years), which was not unexpected given that controls were limited to individuals >60 years of age in EAGLE BioVU (Table 19).

Table 19: Study population characteristics of POAG definite cases and controls over 60 years among African Americans in EAGLE BioVU

	Definite Cases > 20 yrs (SD)	Controls >60 yrs (SD)
N	138	1376
Age at Diagnosis (years)	62.0 (12.0)	--
Age at Last Clinic (years)	--	67.3 (7.8)
Sex (% female)	63.7	56.5
Hypertensive (%)	55.1	52.5
BMI (kg/m^2)	30.1 (6.7)	28.8 (7.35)
Diastolic (mm/Hg)	74.5 (8.1)	76.0 (8.8)
Systolic (mm/Hg)	134.5 (14.1)	135 (14.6)
Cholesterol (mg/dL)	183 (40.6)	169 (46.7)
HDL (mg/dL)	52.5 (25.0)	49 (17.8)
LDL (mg/dL)	103 (42.9)	93 (37.4)
Triglycerides (mg/dL)	125 (76.3)	97 (68.1)

Median values were calculated for the following: Age at POAG diagnosis was determined by the date of when POAG ICD-9 (365.11) was first mentioned in the records. Age at last clinic visit (LCV) was taken as the date of the last CPT mentioned in the records for controls. An individual was classified as hypertensive if he/she met one of three criteria: systolic blood pressure > 140 mm/Hg, diastolic blood pressure > 90 mm/Hg, or on hypertension medications all within a two year window of when they were diagnosed with POAG in cases and a two year window of their LCV date for controls. Blood pressure (systolic and diastolic), lipids (total cholesterol, high-density cholesterol, low-density cholesterol, and triglycerides), and body mass index (height and weight) were calculated from labs or measurements within two years of POAG diagnosis or LCV. Abbreviations: standard deviation (SD)

Generalization of known primary open-angle glaucoma risk variants

We consider a previously associated POAG variant to have generalized if the same variant was associated in a different population with the same direction of effect as observed in the original population (in this case, European-descent and Japanese populations).

In this analysis of African American adults from EAGLE BioVU, POAG cases and controls were tested for an association with 286 common variants (MAF > 5%) in the *CDKN2B-AS1* region using a logistic regression model 1) adjusted for age, sex, and PCs (Supplementary Table 9) and 2) age, sex, PCs, and median diastolic blood pressure (Table 2). Although some SNPs were associated at $p < 0.05$ (Figure 5),

none passed a strict Bonferroni correction ($p < 0.0001$). Nine SNPs were marginally significant at a $p < 0.05$: rs77728904, rs80166549, rs1333049, rs79985856, rs79182326, rs77920300, rs77284052, rs10757277, and rs10757279 (Table 2). In the next study, we aimed to explore whether potential novel variants, available on the Metabochip, modify POAG risk in this population.

Table 20: Results for genetic association analysis of *CDKN2B-AS1* region in African Americans cases (n=138) and controls (n=1,376) adjusted for age, sex, PC, and median diastolic blood pressure. Shown are the ten most significant tests.

SNP	Allele	MAF	Function Class	OR	CI	p-value
rs77728904	C	0.09	Intron	0.40	0.18-0.89	0.03
rs80166549	G	0.10	Intron	0.43	0.20-0.92	0.03
rs1333049	G	0.24	upstream intergenic	0.62	0.40-0.96	0.03
rs79985856	A	0.10	Intron	0.44	0.21-0.94	0.03
rs79182326	A	0.10	Intron	0.44	0.21-0.94	0.03
rs77920300	A	0.10	Intron	0.45	0.21-0.95	0.04
rs77284052	A	0.10	Intron	0.45	0.21-0.95	0.04
rs10757277	G	0.20	upstream intergenic	0.63	0.40-1.00	0.05
rs10757279	G	0.20	upstream intergenic	0.63	0.40-1.00	0.05
rs10738609	G	0.21	Intron	0.65	0.41-1.01	0.06
rs2383206	G	0.41	Intron	0.72	0.51-1.01	0.06
rs17694493	G	0.11	Intron	0.52	0.27-1.02	0.06
rs2069422	C	0.10	Intron	0.50	0.24-1.03	0.06
rs10738610	C	0.21	upstream intergenic	0.65	0.41-1.02	0.06
rs10965234	A	0.47	Intron	1.37	0.98-1.91	0.06
rs77563194	A	0.05	Intron	1.77	0.96-3.27	0.07
rs10965235	A	0.47	Intron	1.37	0.98-1.91	0.07
rs1333046	T	0.25	upstream intergenic	0.68	0.45-1.03	0.07
rs10217426	C	0.47	Intron	1.36	0.98-1.90	0.07
rs12347950	G	0.10	Intron	1.55	0.96-2.50	0.07

MAF = minor allele frequency
OR = odds ratio
CI = confidence interval

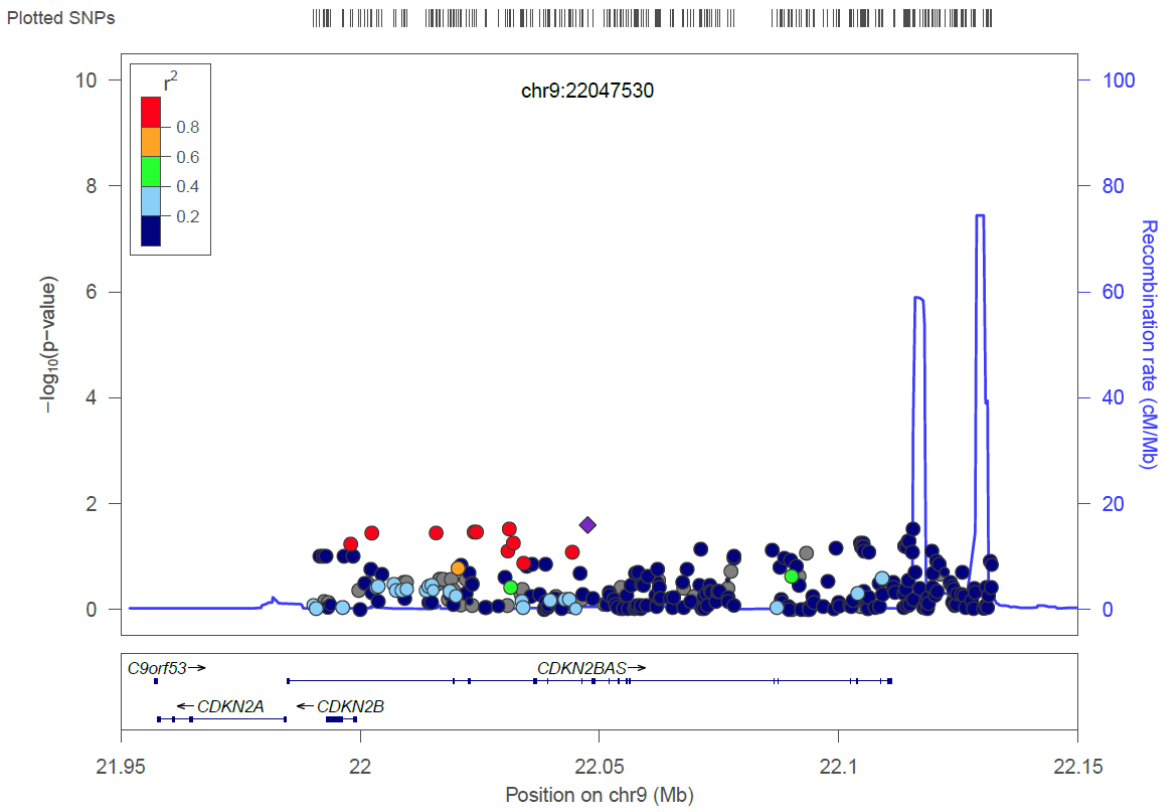


Figure 20 Locus Zoom regional association plot for POAG in African Americans for *CDKN2B-AS1*.
Vertical axis is $-\log_{10}$ of the p-value, the horizontal axis is the chromosomal position. Each dot represents a SNP tested for association with POAG in 138 cases and 1,376 controls. Approximate linkage disequilibrium between the most significant SNP and the other SNPs in the plot is shown by the r^2 legend with LD calculations from 1000 Genomes YRI.

Genetic discovery for primary open-angle glaucoma in African Americans

To reiterate, much of the population-based genetic association studies in POAG have been carried out in European and Asian-descent populations. While these studies have had some success in identifying genetic modifiers of POAG risk, the studies generalizing these associations in African Americans, the population most at risk for POAG, have been limited. The studies that have been published thus far sought to determine whether European index variants generalized to African American and African population. In a study of African American women from the Women's Health Initiative (WHI)(Hoffmann et al., 2014) there was no association between European-index variants and their population, while additional work by Liu *et al* found an association with single variants in the *CDKN2B-AS1* and *SIX1/SIX6* gene regions in subgroup analyses of normal tension and high pressure POAG(Liu et al., 2013b). These conflicting results do not yet provide strong evidence to suggest that European index variants are driving risk in African Americans. By only focusing on genomic locations identified in other populations the scientific community leaves open a huge gap in the knowledge of POAG genetics.

The Metabochip provides an interesting opportunity to explore known cardiovascular and metabolic – associated regions in the genome. The vascular theory of glaucoma hypothesizes that systemic irregularities in the cardiovascular pathways lead to pathogenesis of glaucomatous optic neuropathy via a mechanism of reduced blood supply to the eye. This can result in starving retinal ganglion cells of prerequisite oxygen and nutritents(Flammer, 1994; Flammer et al., 2002). Whether or not systemic blood pressure (bp) is involved in glaucoma pathology has been up for debate(Bonomi et al., 2000; Leske et al., 2007, 2008; Quigley et al., 2001). The correlation between bp and glaucoma found in some studies can perhaps be better explained by ocular perfusion pressure (OPP) and its effect on optic nerve health. OPP is calculated by taking the difference between the arterial and venous blood pressure. Low OPP has been repeatedly associated with glaucoma and glaucoma progression(McGlynn et al., 2013; de Oliveira and Kasahara, 2014). At a molecular level it is hypothesized that low OPP or fluctuations in OPP lead to ischemia in the optic nerve(Cherecheanu et al., 2013).

We under took to investigate whether gene regions associated with cardiovascular and metabolic traits might be driving POAG risk in African Americans by performing a genetic association analysis of all SNPs available on the Metabochip. Although it should be noted that many of these SNPs were associated with other traits not related to cardiovascular or metabolic phenotypes.

Results

We tested all SNPs genotyped on the Metabochip for an association with POAG 1) adjusted for age, sex, first three PCs, and 2) adjust for age, sex, PCs, and median diastolic blood pressure. The 100 most significant results for each model can be found in the appendix (Supplemental Table 3 and Supplemental Table 4, respectively). No SNP was significantly associated with POAG after adjusting for a strict Bonferroni correction ($p < 4.04 \times 10^{-7}$) in either model (Table 21 and Supplementary Table 12). The two most significant associations in the model adjusted for age, sex, PC, and median diastolic blood pressure (chr1:228347779 and chr1:228354829) are located within the protein coding gene for iron-sulfur cluster assembly homolog (*IBA57*), which is a component of the biosynthesis pathway for mitochondrial 4Fe-4S proteins and clinical implications in severe myopathy(Ajit Bolar et al., 2013). A small number of SNPs met our suggestive threshold of $p < 10^{-4}$ (Figure 21).

Table 21: Ten most significant results for POAG African American Metabochip genetic association. Logistic regression assuming an additive genetic model was performed for 138 cases and 1,376 controls adjusted by age, sex, principal components, and median diastolic blood pressure.

CHR	SNP	Gene	Allele	MAF	OR	95% CI	p-value
1	chr1:228347779	*IBA57*	A	0.11	2.37	1.56-3.60	5.00×10^{-5}
1	chr1:228354829	*IBA57*	C	0.15	2.09	1.44-3.02	8.73×10^{-5}
21	rs9982695	*C21orf33*	A	0.24	2.09	1.44-3.02	8.74×10^{-5}
4	rs3775202	*VEGFC*	G	0.43	1.92	1.38-2.66	9.70×10^{-5}
2	rs13423742	*FN1*	C	0.06	3.04	1.73-5.36	1.14×10^{-4}
6	rs7454156	*BMP6*	G	0.18	2.08	1.42-3.02	1.37×10^{-4}
6	rs9479726	*RGS17-OPRM1*	A	0.24	0.41	0.25-0.64	1.54×10^{-4}
19	rs1671152	*GP6*	A	0.32	1.91	1.36-2.68	1.60×10^{-4}
10	rs286489	LOC101929727	A	0.28	1.90	1.35-2.66	1.80×10^{-4}
5	rs4336354	*HTR4*	G	0.09	2.51	1.54-4.07	1.86×10^{-4}

MAF = minor allele frequency
OR = odds ratio
CI = confidence interval

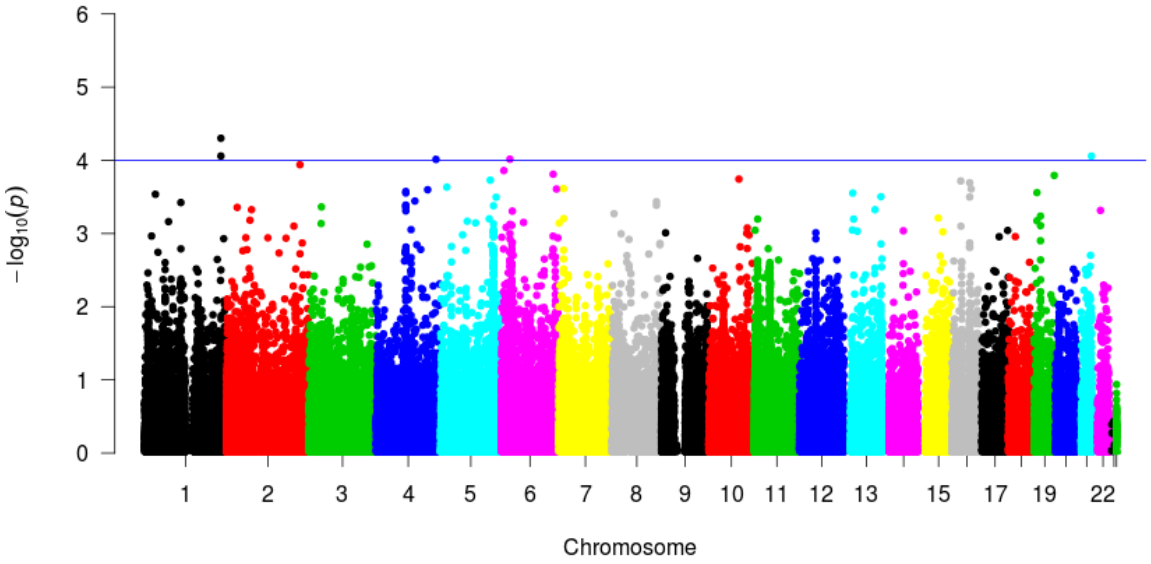


Figure 21: Manhattan plot of EAGLE BioVU African American POAG genetic association results. Logistic regression assuming an additive genetic model was performed for 138 cases and 1,376 controls adjusted by age, sex, principal components, and median diastolic blood pressure. P-values ($-\log_{10}$) on the y-axis) for each test of association are plotted by chromosome (x-axis). The blue line depicts a suggestive significance threshold of $p = 10^{-4}$.

Discussion and Summary

Generalization of previously identified CDKN2B-AS1 and SIX6 variant

This study did not find a significant association in the *CDKN2B*-AS1 region for POAG in African Americans after adjusting for multiple testing. These results are not entirely surprising given the differences in susceptibility and linkage disequilibrium between European and African-descent populations (Supplementary Figure 7 and Supplementary Figure 8). Most of the significant associations between this gene and POAG were discovered in European and Japanese populations (Table 22). Of these SNPs, five were available on the Metabochip but were not found to be associated with POAG in this study. Our results are somewhat consistent with Liu et al(Liu et al., 2013b) who investigated known European POAG risk

loci in both an African American and Ghanaian population. Liu et al failed to replicate many of the European risk loci in their African populations, but they did identify one SNP (rs10120688) in *CDKN2B-AS1* significantly associated with POAG with an OR of 1.21. This SNP is available on the Metabochip but the association was not replicated in our dataset (OR = 1.01 p = 0.93) nor did we have the power to detect this association (Supplemental Figure 5). *CDKN2B-AS1* rs1333049 was available for analysis in this study, and interestingly, our results (OR = 0.72; p = 0.06) were similar to that reported Liu et al for the same SNP (OR = 0.89; p = 0.07). Increasing our power with additional samples may resolve this test.

Table 22: Published index variants for the *CDKN2B-AS1 region* associated with POAG or POAG associated trait and availability of these variants on the Metabochip.

rs#	Gene	population	OR	p-value	Discovery study	Current Study OR	Current Study p-value
rs7865618	*CDKN2B-AS1*	Japanese	1.78	9x10-11	Nakano et al(Nakano et al., 2012)	1.01	0.96
rs1063192	*CDKN2B*	Japanese	1.33	5x10-10	Osman et al(Osman et al., 2012)	0.92	0.75
rs2157719	*CDKN2B-AS1*	European American	1.45	2x10-18	Wiggs et al(Wiggs et al., 2012)	0.97	0.92
rs4977756	*CDKN2B-AS1*	European	1.50	4.7x10-9	Burdon et al(Burdon et al., 2011b)	1.05	0.72
rs10120688	*CDKN2B-AS1*	African American	1.21	0.002	Liu et al(Liu et al., 2013b)	1.01	0.93

Shown are significant index variants which are listed on the NHGRI GWAS catalog and within PubMed as of 2014. Included is the availability of the index variants on the Metabochip and summary results for the current studies association analysis of African Americans with POAG in the *CDKN2B-AS1* region.

A previously identified SNP (rs10483727), located upstream of *SIX6,* associated with POAG and quantitative glaucoma traits(Macgregor et al., 2010; Osman et al., 2012; Ramdas et al., 2010; Wiggs et al., 2012) in separate European and Japanese population studies did not generalize in our study. In EAGLE BioVU African American POAG cases, rs10483727 (p = 0.42) had a genetic effect size (OR) of 1.25 which is on par with the published effect size (1.27-1.32). The A-allele frequency in HapMap CEU

and JPT is approximately 0.38 and 0.74, respectively, while in the EAGLE BioVU African Americans the A-allele was present with a frequency of 0.86. Our study was underpowered to detect a significant association with a genetic effect size smaller than 1.50. Within HapMap ASW, the rs10483727 variant does not appear to be in LD with other SNPs in the region (Supplemental Figure 6) while even in HapMap CEU there is little LD ($r^2 < 50$) to suggest tagging of another SNP (Supplementary Figure 9).

Discovery

In the discovery phase of our study, no SNP reached Bonferroni-corrected significance. However, the POAG analysis identified several interesting SNPs (n = 80) at a $p < 10^{-4}$ to be followed up on in future studies (Supplemental Table 4). The most significant of these associations involve SNPs located within genes involved in the mitochondria, angiogenesis, and serotonin and collagen receptors. Of particular interest for future studies is rs1671152 (OR=1.96; $p = 1.60x0^{-4}$), a known missense variant in the *glycoprotein VI* (*GP6*) gene. *GP6*, a collagen receptor, is involved in platelet aggregation but is expressed in the eye and brain. Potential implications may include scleral collagen organization and integrity of the blood-retinal barrier in glaucoma susceptibility.

Abnormalities in the vascularization of ocular tissues and subsequently the deprivation of oxygen and nutrients to these tissues are known to lead to eye disease such as age-related macular degeneration, diabetic retinopathy, and neovascular glaucoma. Hypoxia-induced ischemia can lead to over expression of pro-angiogenic factors such as the proteins in the VEGF family. Vascular endothelial growth factor C (*VEGFC*) and serotonin receptor 4 (*HTR4*) genes promote angiogenesis in endothelial cells(Profirovic et al., 2013) and are expressed in retinal and central nervous system tissues(Zarkada et al., 2015). Our top results highlight a potential role for angiogenesis (i.e. VEGFC and HTRA) in POAG risk for African Americans.

Strengths and Limitations

This study has a number of strengths and limitations. Perhaps its greatest strength is the pursuit of knowledge in African Americans, a population far too often underrepresented in biomedical research. As

the population at greatest risk of POAG, the work performed here has the potential to offer the greatest impact to those at risk. Studies like ours can help to determine to what extent allele frequency differences are contributing to variation in disease risk and subsequently to ascertain whether population specific factors are at work. Additional strengths involve the utilization of an EMR as a cost efficient and data-dense resource for studies.

A major limitation of our study is the limited number of cases for inclusion in the genetic association studies, which reduces our statistical power to identify moderate to smaller genetic effects (Supplemental Figure 5). Even at a moderate allele frequency of 25%, our study had only 80% power to detect an association with an OR of 1.60. Power is further exasperated by incomplete data in patient's medical charts in the Vanderbilt de-identified medical records (BioVU). We lost twenty-one cases for analysis when incorporating diastolic blood pressure into our model resulting in a total case reduction of 18.4%. Despite limited samples sizes, these samples could be included in a meta-analysis of POAG in African Americans which would increase power.

Because of these strengths, and despite these limitations, we have contributed to the sparse knowledge of the genetic architecture of POAG in a diverse population. One of the findings in our study is that metabolic and cardiovascular risk loci may be contributing to POAG risk in African Americans while known loci discovered in European and Japanese populations do not explain genetic risk in this population.

CHAPTER V

GENETIC DETERMINANTS OF DIABETIC RETINOPATHY IN AFRICAN AMERICANS

Introduction

The American Diabetes Association defines "diabetes" as a group of metabolic diseases typically characterized by high blood sugar and defects in the body's ability to produce and/or use insulin (Figure 22). As of 2014, 9.3% of the United States population was affected by diabetes with estimates predicting that nearly one third of the population will be afflicted by 2050(Antonetti et al., 2012) driven by an increase in incident T2D cases.

Type-1 diabetes (T1D) is an insulin-dependent diabetes that is caused by a lack of insulin production in the body. T1D occurs in youths before the age of 25, with an average diagnosis around12 years of age. Prevalence rates are predicted to be 0.26 to 1.5% in youths < 20 yrs. This insulin deficiency results from the destruction of beta cells in the pancreas due to autoimmune assault. Approximately 80-90% of new diagnoses have anti-islet cell antibodies(Steck et al., 2015) which are thought to be triggered by viral infection or other environmental insult. Classic symptoms include: polyuria, polydipsia, increased hunger, fatigue, and weight loss.

Our study focuses on individuals who suffer from the non-insulin dependent form or type 2 (T2D) which results from complications of insulin resistance on top of an insulin deficiency background. Insulin resistance is the result of the body's inability to efficiently use insulin for the uptake of glucose into tissues such as muscle and the liver. This inability results in high glucose levels that can trigger an increase in glycogen synthesis in liver cells and a failure to suppress glucose production. Chronic elevated basal levels of insulin down regulate insulin receptors and thereby increase insulin resistance. Currently it is estimated that people living with T2D will lose up to 10 years of life expectancy and up to 20 years for T1D(Livingstone et al., 2015; Writing Group for the DCCT/EDIC Research Group et al., 2015).

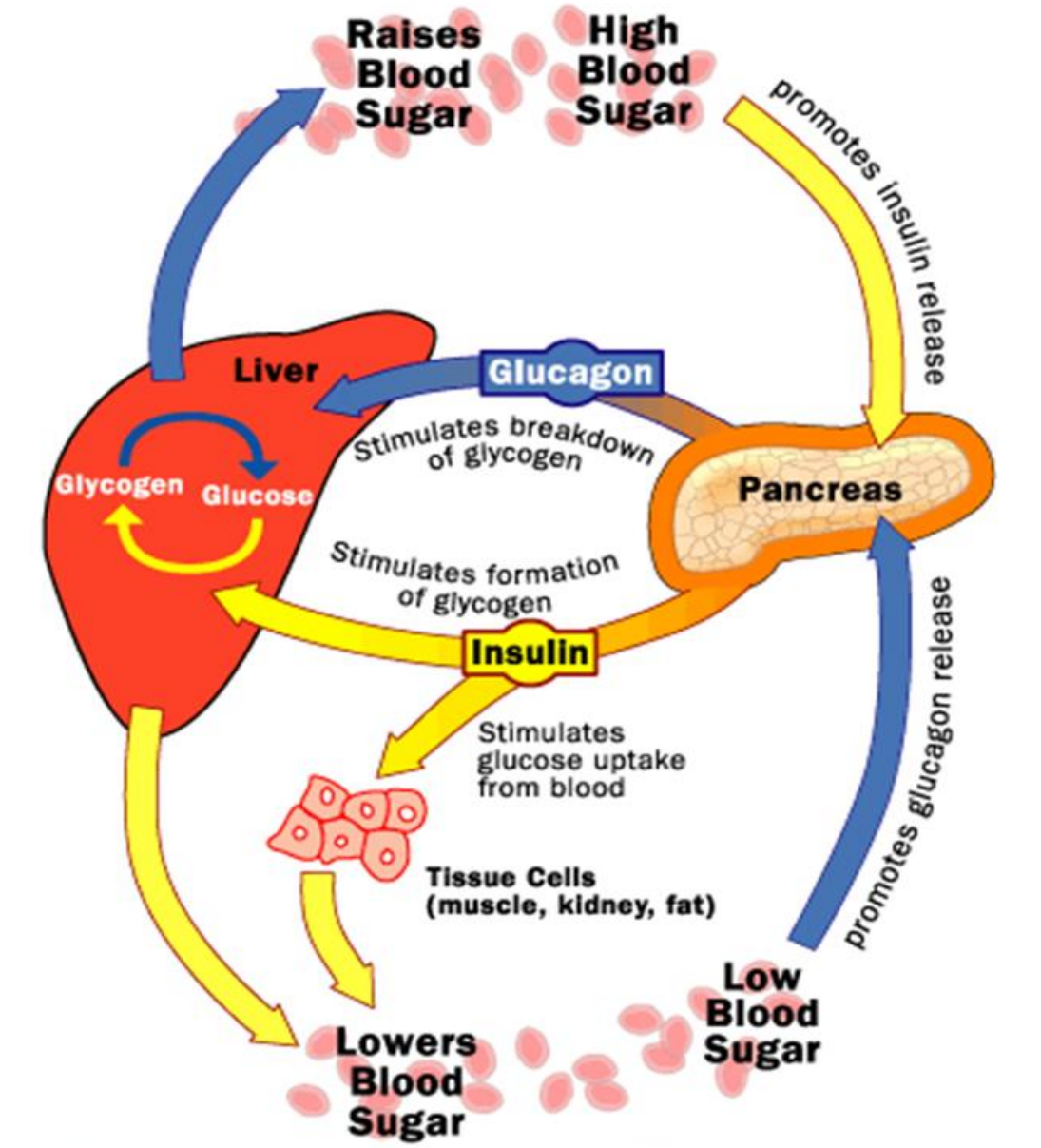


Figure 22: Diagram of the biological pathways involved in the control of blood glucose levels. This figure is adapted from the work of Dr. Craig Fruedenrich at http://science.howstuffworks.com/environmental/life/human-biology/diabetes1.htm.

From the original coining of the term “diabetes” by the Greek Aretaeus, to the discovery of insulin in 1921 by the Canadian surgeon Frederick Banting, the clinical and genetic knowledge of diabetes and diabetic complications has exploded in recent decades. The advent of medical technology has opened up the ability of clinicians to monitor, regulate, and treat the many metabolic dysfunctions of the disease. This has led to a drastic increase in the longevity of patients from an original mortality prognosis of one year before the discovery of insulin, to the ability of most diabetics to live full lives. Greater longevity of today’s diabetic patients has also led to the presentation and study of complications that arise from long term experience with the disease. One such complication is the initiation and progression of diabetic retinopathy (DR). Individuals with DR present with a variety of morphological lesions in the retina. This abnormality in the microvasculature condition of the retina is present in 82% of type 1 diabetes (T1D)(Roy et al., 2004) and 40% in type 2 diabetes (T2D)(The Eye Diseases Prevalence Research Group, 2004) patients in the United States with prevalence rates increasing drastically 15 years after diagnosis to nearly 100% in T1D and 80% in T2D patients(Fong et al., 2004; Klein et al., 1984). In 2007 approximately 4.4% of individuals with diabetes had advance DR putting them at risk for severe vision loss and blindness.

Clinical features and diagnosis of diabetic retinopathy

Diabetic retinopathy is traditionally viewed as one disease that occurs on a spectrum of severity. It is classified as two clinical subtypes according to the absence or presence of new blood vessel growth within the retina: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR), respectively. Each sub-type is further broken down into severity levels (i.e. mild, moderate, severe) based upon the presence and type of microvascular damage, such as, microaneurisms, cotton wool spots, venous beading, or neovascularization (Figure 23). Early stages of DR are characterized by little vision loss and are the most effective time at which to treat. Early intervention of glycemic control has been shown to delay progression of retinopathy. Late stages of DR are often accompanied by irreversible, severe vision loss or blindness and accounts for the number one cause of blindness in working age Americans (i.e. 20-74 years of age) (Figure 24).

Screening for DR is typically carried out through the use of a dilated eye exam with fundus photography for visualization of the retina. Exams look for the presence and/or absence of cataract, abnormal blood vessels, blood or swelling in retina, scar tissue, neovascularization, blood in vitreous humor, and retinal detachment(s). Diagnosis is determined by the Airlie House Classification Scheme. Additional tests include the fluorescein angiogram and optical coherence tomography (OCT).

An angiogram is a picture of the inside or lumen of blood vessels to determine if blood vessels in the retina are leaking (i.e. aneurysms and hemorrhages). During the procedure for a fluorescein angiogram, the patient is injected with sodium fluorescein and a picture of the retina is captured, via fluorescence emitted after illumination of the retina with blue light at a wavelength of 490 nm, to obtain an angiogram. OCT is a noninvasive optical signal acquisition and procession method that captures 3D images of slices of the retina. An OCT can measure the thickness of the retinal tissue, determine if fluid is accumulating within the retina, oxygenation levels of retinal tissue, retinal swelling, and retinal detachment. OCT is more sensitive than fundus photography, which requires skilled personnel for both the imaging and is a qualitative versus a quantitative analysis.

Figure 23 Representation of the vision loss associated with diabetic retinopathy in comparison to normal vision. Images are adapted from the National Eye Institute's webpage:https://www.nei.nih.gov/photo/eye-diseases-and-vision-disorders.

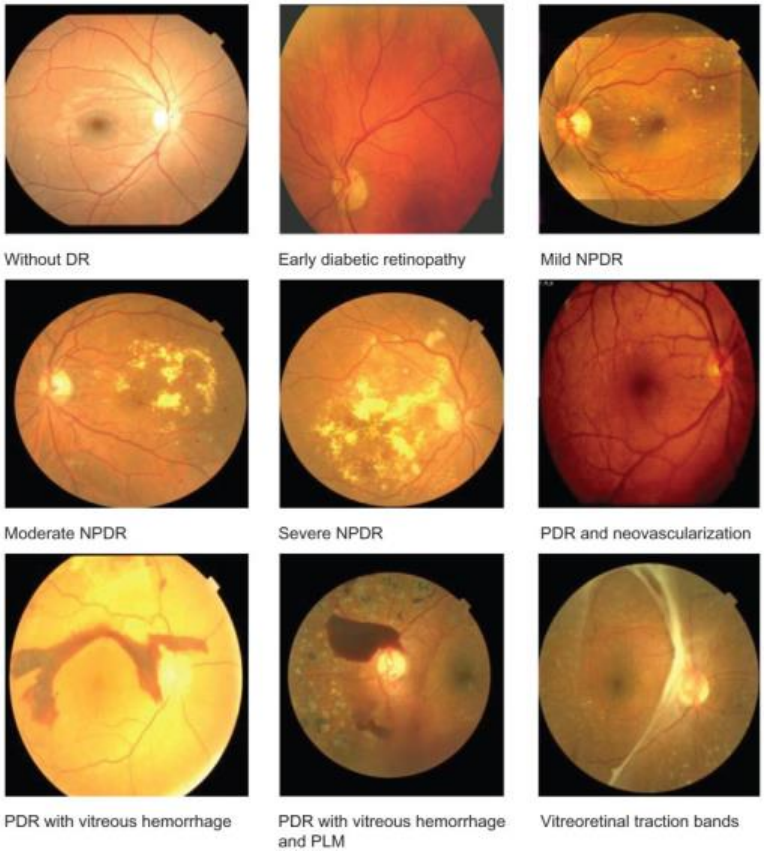


Figure 24: Funduscopic image of the stages of diabetic retinopathy severity start from the publication of El-Bab et al[8].

Etiology and pathogenesis of diabetic retinopathy

The retina is a unique organ within the body that is protected by the blood-ocular barrier which isolates the retina from the body and most other parts of the eye. This barrier acts as protection from immunological attacks and prevents substances such as viruses and drugs from readily crossing over. Retinal tissue also exceeds most other tissues in the body for metabolic energy needs but, interestingly, the inner retina possesses few mitochondria. It requires glycolysis to meet these needs instead of the more common oxidative phosphorylation mode of ATP production(Antonetti et al., 2006). Mitochondria-enriched Muller cells are found in the outer retina following along an inherent oxygen tension gradient that is highest in the outer regions and decreases to an almost hypoxic state in the inner regions. The cycle of energy acquisition (Figure 25) (Antonetti et al., 2006) illustrates that cells in the outer retina (i.e. Muller cells and astrocytes) and neurons can take up glucose from diffusion at the capillaries, but only the neuron can oxidize glucose. Muller cells and astrocytes convert glucose to lactate that is then taken up by neurons for oxidation. The constant production of ATP is vital for cell signaling. It is believed that this high demand for energy but minimal vascular supply of nutrients predisposes the retina to diabetes-induced damage via chronic inflammation.

Current hypotheses suggest that the inflammation is triggered by constant levels of high blood glucose which directly or indirectly leads to cellular apoptosis in retinal pericytes. The pericytes make up a component of the retinal capillary walls that helps regulate blood flow and removal of cellular debris. High levels of glucose are believed to lead to increased levels of advanced glycation end products (AGE) which are presumed to accumulate in the pericytes(Hammes, 2005). The build-up of these toxic components then leads to apoptosis and the formation of microaneurisms and hemorrhages(Geraldes et al., 2009), early signs of DR.

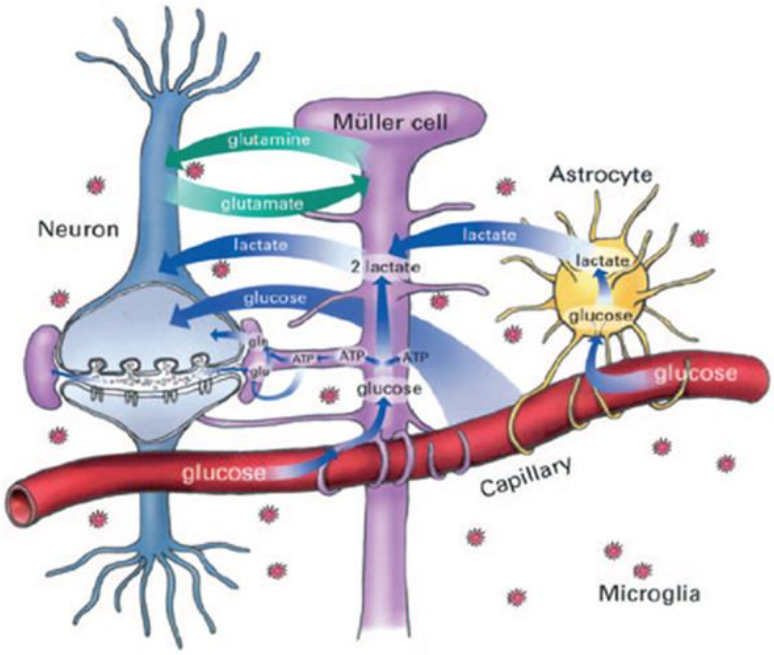


Figure 25: Adapted from Antonetti, et al (2006): Diagram of the flow of glucose utilization by retinal tissue cells

Pericytes interact with endothelial cells and contribute to the homeostasis of the blood-retinal barrier by a complex interplay of cell signaling (Figure 26). Endothelial cells secrete PDGF-B to recruit and maintain pericytes through the activation of protein kinase B (Akt). Akt activity is central to cell survival pathways by inhibiting apoptotic events. Additionally, Pericytes produce molecules that contribute to the endothelial barrier and junctional complexes such as the platelet-derived growth factor (PDGFR) and the PDGF receptor. In diabetes, hyperglycemia destabilizes the interaction between pericytes and the endothelium and can lead to pericytes cell death. An increase in the expression of VEGF and a decrease in pericytes leads to permeability in the vascular system and contributes to angiogenesis. These fragile new blood vessels are prone to breakage and leaking of blood and cellular fluids under the retinal tissue or into the vitreous humour.

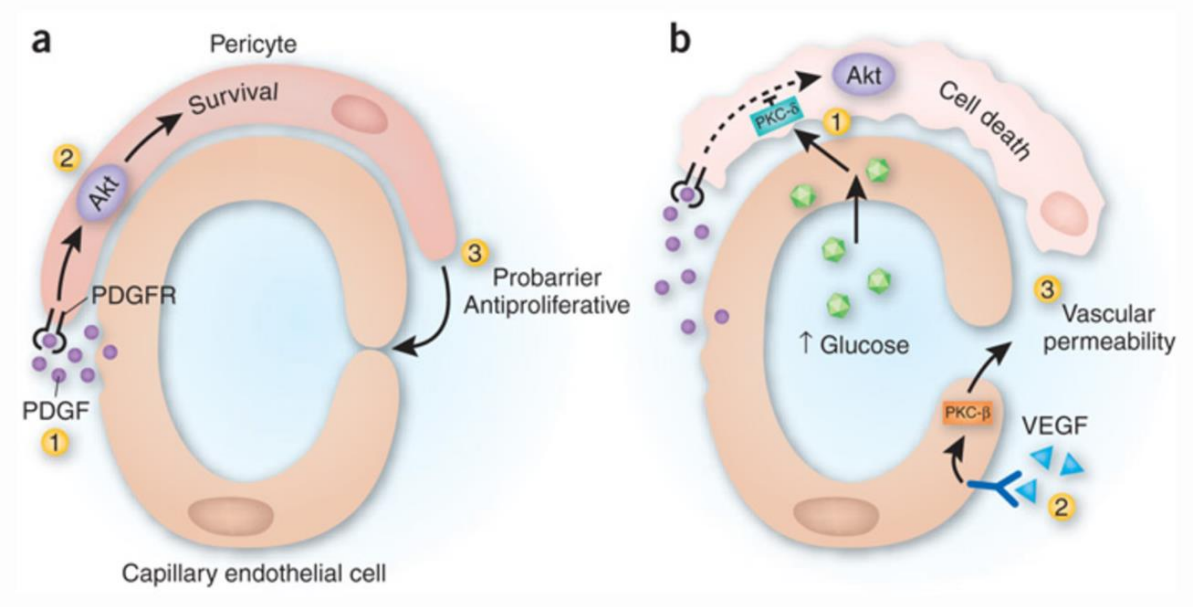


Figure 26*:* Diagram of pericyte cell death as mediated by diabetic complications. Adapted from the publication of Dr. David Antonetti[9].

Genetics and epidemiology of diabetic retinopathy

The genetics of DR are complicated by conflicting results in association studies, but epidemiological studies have been fairly consistent in regards to the environmental risks that play a role in disease susceptibility and intermediate phenotypes. Briefly, primary risk factors include male sex, duration of diabetes, glycemic control, hypertension, hyperlipidemia, type of diabetes (T1D versus T2D), age, and race. The prevalence of DR varies in T2D differs across populations (Figure 27) with African Americans and Hispanics experiencing the greatest level of burden. Evidence from heritability studies supports that genetic factors play a role in this disease. The predisposition to diabetes is the first necessary factor for development of DR. Multiple heritability studies of diabetes have been carried out in different populations(WONG et al., 2006)·(Leslie and Pyke, 1982). It is well understood that diabetes has a genetic component but other studies have also shown evidence for a genetic component to DR beyond the risk of disease due to exposure of diabetes.

Beyond heritability studies, the identification of genes that play a direct or indirect role in DR has been limited. This may be explained partially in that family based linkage studies of common, complex diseases such as diabetes are difficult for a number of reasons including identification of multiplex family pedigrees and genetic heterogeneity across families. Regardless, some studies have been successful in identifying genomic loci and variants correlated with DR. In a modified sibpair study of Pima Indians, a weak linkage peak (LOD = 1.36 - 1.46) with DR was found on chromosome 3 and chromosome 9(Imperatore et al., 1998). Larger GWAS studies have had limited success. One GWAS of severe diabetic retinopathy (sDR) was performed in 973 cases of European-descent individuals with T1D(Grassi et al., 2011). Though this study did not find genome-wide significant ($p < 10^{-8}$) associations for DR in a T1D cohort(Grassi et al., 2011), it did identify 19 variants with suggestive significance ($p < 1x10^{-7}$). One nominally associated variant (rs476141) was located near the *AKT3* gene which is a member of the serine/threonine protein kinase family. This kinase and others in the family are known regulators in cell signaling, insulin, and glucose uptake and suggest a biological link to DR. A study in a Han Chinese population found associations ($p < 10^{-7}$) in intergenic regions of *TBC1D4-COMMD6-UCHL3* and *ARL4C-SH3BP4*(Sheu et al., 2013b). Interestingly, a Taiwanese T2D GWAS(Huang et al., 2011) found a significant association in the heparin sulfate 6-O-sulfotransferase 3 gene (*HS6ST3*) which is known to play a role in lipid metabolism and inflammation, two factors which are thought to contribute to the etiology of DR.

To recapitulate, the etiology of DR is complex and complicated further by shared genetic and epidemiological risk factors of diabetes, such as African ancestry. For nearly every study that identifies ancestry as a risk variable another study does not(Ojaimi et al., 2011; WONG et al., 2006). After accounting for traditional risk factors(Emanuele et al., 2005), African Americans from the Veterans Affairs Diabetes Trial had a higher frequency of severe DR a greater risk (OR = 2.30) of macular edema(Emanuele et al., 2009) compared to their European American counterparts . Still, the question remains whether race/ethnicity plays a significant role in DR and more specifically whether this differentiation is due to unique genetic determinants or disparities in health care and additional socioeconomic factors.

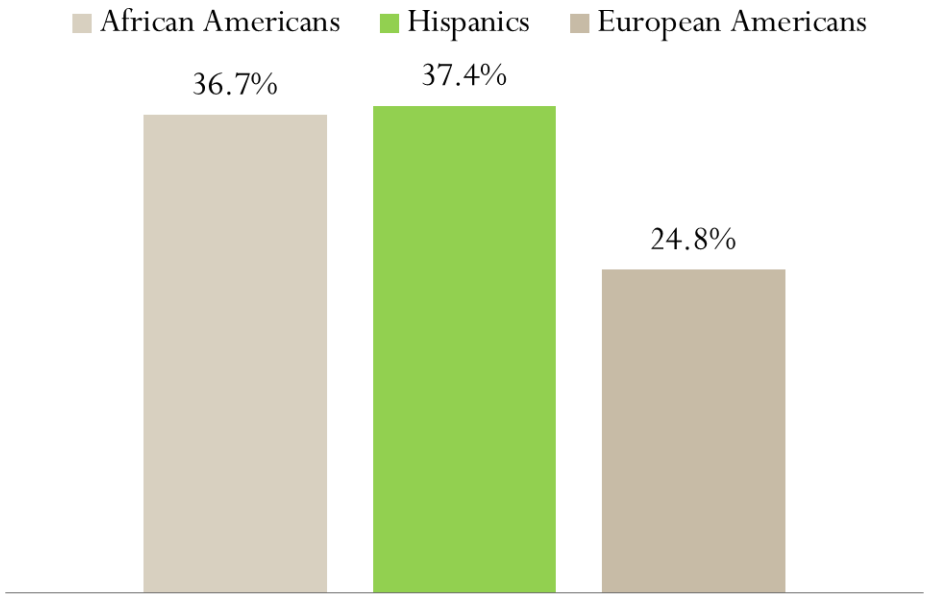


Figure 27: Prevalence rates of diabetic retinopathy in U.S. adults with diabetes as determined by published NHANES data.

Generalization of published diabetic retinopathy risk variant(s) in an African American population

The impact of the obesity epidemic sweeping across developed countries is driving development of T2D and the suite of diabetic comorbidities, such as extreme fatigue, frequent infections, impaired healing ability, and ocular complications that can hinder and reduce the quality of life of individuals in the diabetes community(Narayan K, 2003). Furthermore, this burden is heaviest on individuals of low socioeconomic status who are disproportionately affected by this disease. As the incidence of T2D increases it is expected that the incidence of DR will increase proportionally without appropriate interventions. Therefore, it is important to determine the genetic and environmental factors that play a role in disease onset and progression. Our study proposes to explore the genetic architecture of diabetic retinopathy in African Americans, prior studies of which are altogether missing or uninformative. This knowledge is important for determining potential screening programs and may lead to a better understanding of the underlying etiology specific to African Americans and to identifying potential pathways for pharmacology targets in a population vastly underrepresented in the medical research field.

Here I will present my primary analyses for DR common variants in the advanced glycation end product receptor (*AGER*) gene and the transcription factor 7-like 2 (TCF7l2) gene. *AGER*, which is expressed in the retina, is known to interact with molecules involved in homeostasis and inflammation. Nonsynonymous *AGER* rs2070600 (Gly82Ser) is one of a very few polymorphisms that have been found to correlate with DR risk(Balasubbu et al., 2010), although inconsistently(Kang et al., 2012). The association studies for AGER rs2070600 and DR thus far have been limited to populations of Asian-descent. AGER rs2070600 has also been associated with T2D risk, but this association has also been inconsistent across studies(Goulart et al., 2008; Kang et al., 2012) and the association has yet to be identified at genome-wide levels. The associations between AGER rs2070600 and T2D have been conducted in mostly in populations of European and Asian-descent(Kang et al., 2012).

Unlike *AGER* rs2070600, polymorphisms within and around *TCF7L2* (such as rs7903146) have been consistently associated with type 2 diabetes in both candidate gene and GWAS conducted primarily in European-descent populations(Morris et al., 2012; Scott et al., 2007; Voight et al., 2010; Zeggini et al., 2007). *TCF7L2* SNPs have also recently been associated with T2D in African Americans(DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium et al., 2014; Haiman et al., 2012; Kho et al., 2012; Ng et al., 2014; Palmer et al., 2012), suggesting that this association for T2D generalizes to multiple ancestral populations. Like AGER rs207600, TCF7L2 rs7903146 has been inconsistently associated with diabetic retinopathy(Sudchada and Scarpace, 2014) and has yet to be associated with DR at genome-wide significance.

As part of the PAGE I study, EAGLE genotyped nearly 16,000 DNA samples from non-European Americans in BioVU using the Illumina Metabochip, an array designed to replicate and fine-map previously associated genomic regions that impact type 2 diabetes and cardiovascular disease risk. The Metabochip and DR phenotype algorithm have been previously described in Chapter 3. With 119 DR cases and 473 diabetes controls drawn from a clinical population, I test here whether or not *AGER* rs2070600 and *TCF7l2* rs7903146 are risk factors for DR in African Americans. I also test for the best associated SNPs in these gene regions in African Americans. Finally, I test all SNPs targeted on the Metabochip for an association with DR among African Americans as a hypothesis generating experiment. Overall, I did not identify a significant association with DR in this moderately-sized clinically-derived dataset, highlighting the challenges of performing genetic association studies for a complex disease in minority populations.

Methods

Study Population and Phenotyping

The study population and phenotyping of DR cases and controls have been described in detail in Chapter 3: Utilization of Electronic Medical Records systems for genetic association studies. Briefly, 119 African Americans DR cases and 473 African American T2D controls were identified in the Vanderbilt SD that

were genotyped on the Metabochip. These individuals were further evaluated for inclusion in the following genetic association analysis.

Statistical methods

SNP rs2070600 was genotyped on Metabochip with an additional 27 variants genotyped in a 30 kbp region. After filtering for MAF > 0.05,we tested for an association between DR and 13 SNPs in the *AGER* region of chromosome 6 and 148 SNPs in the *TCF7L2* region of chromosome 10 in EAGLE BioVU African Americans. We also performed tests of association for all variants on the Metobochip as a hypothesis-generating exercise. Individuals included in this analysis were those identified as "definite" DR cases over the age of 20 years and DR controls over the age of 60 years

Each SNP was tested for an association with DR using logistic regression assuming a log-additive genetic model. We performed tests of association adjusted by 1) age and sex and 2) age, sex, and mean serum glucose levels. For definite DR cases (n=119), this study has 80% power to detect an OR of 2.0 at an allele frequency of 0.20 or an OR of 1.80 at MAF 0.40. This assumes a Bonferroni corrected p-value of 0.0018 and a log-additive genetic model. For *TCF7L2* rs7903146, I assumed an allele frequency of 0.30, a published odds ratio of 1.58, sample size of 100 cases, and a Bonferroni corrected p-value of 0.0003. Figure 28 shows this association is adequately powered in our study at a significance threshold of p=0.05. Analyses were conducted using PLINKv1.90(Purcell et al., 2007) and results were plotted using Synthesis View(Pendergrass et al., 2010). Power calculations were performed in Quanto®(Wang and Li, 2004).

Within our diabetic retinopathy study, we have the power to replicate associations found previously in European descent population studies, such as rs7903146 (*TCF7L2*) which is located with a fine mapped region of Metabochip. Power calculations were run in Quanto®(Wang and Li, 2004) and Figure 28 shows an example of variant rs7903146 with an allele frequency of 0.30, a published odds ratio of 1.58, and a sample size of 100, which is adequately powered in our study at a significance threshold of p=0.05.

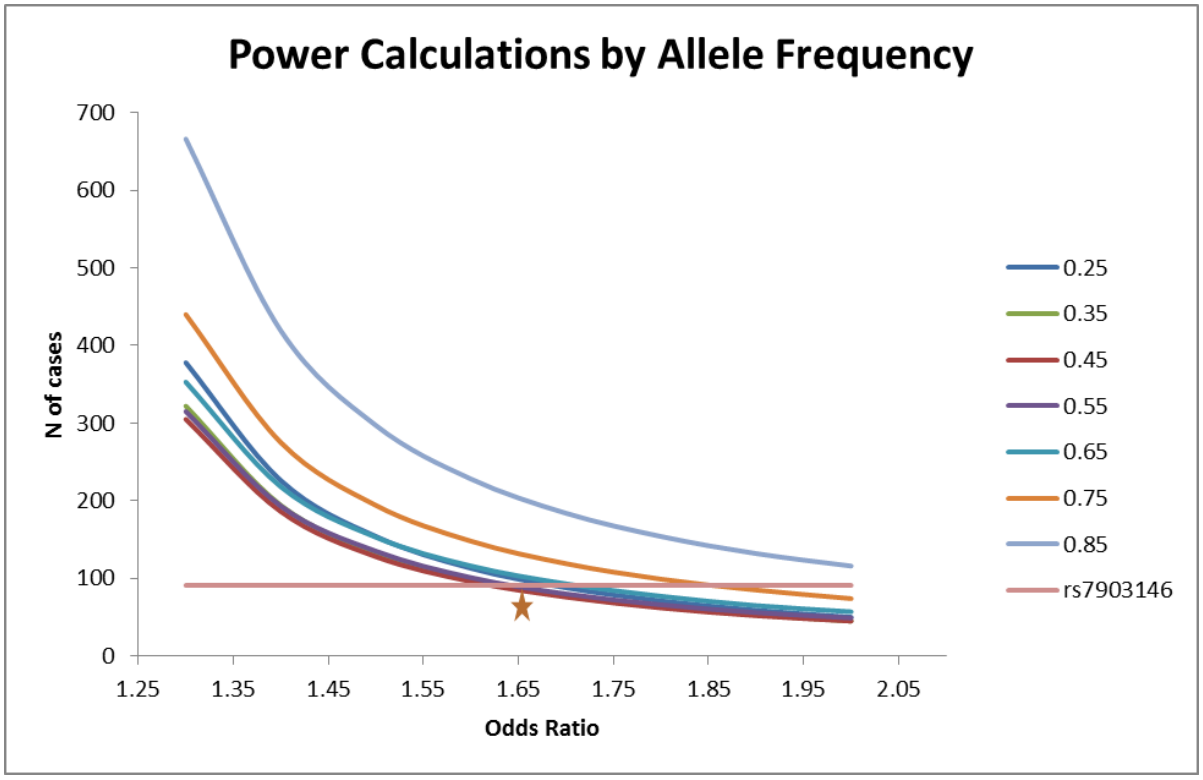


Figure 28: Power curve of the number of cases necessary to identify an association with SNP rs7903146 at a specified odds ratios with 80% power.
Power Calculations determined in Quanto for the EAGLE BioVU African American diabetic retinopathy cases with a case: control ratio of 1:3, log-additive model, and a 2-sided t-test. The color coded lines represented the varying allele frequencies tested for in these power calculations.

Results

Population characteristics

After final quality control, we calculated the demographics for the EAGLE BioVU African American DR cases and T2D controls. In general cases were older, more likely to be male, and had been diagnosed with

T2D earlier compared to controls (Table 23). As expected, controls had better control of their diabetes as determined by their blood glucose and glycated hemoglobin levels. Controls were also more likely to have healthier blood cholesterol levels than their case counterparts.

Table 23: Demographics of EAGLE BioVU Metabochip African Americans with diabetic retinopathy

	Diabetic retinopathy	
	cases	controls
N	119	434
Age (years)	65	63.5
% female	39%	61%
BMI	31.1	32.4
Age T2D diagnosis	52.9	59.9
T2D duration (years)	9	5
Age DR diagnosis	56	-
Systolic (mmHg)	139.1	132.7
Diastolic (mmHg)	76.3	77.9
Cholesterol (mg/dL)	194.5	165
Glucose (mg/dL)	222	117
Triglycerides (mg/dL)	140	101
LDL (mg/dL)	111	91
HbA1c (%)	10.6	6.7

Values shown are for the medians

Generalization of known diabetic retinopathy risk variants

After quality control and MAF frequency cutoffs, we tested a total of 13 SNPs in the *AGER* gene and a total of 148 SNPs in *TCF7L2.* We did not identify an association between DR and the EAGLE BioVU African American cases and controls for either the *AGER* or the *TCF7L2* gene regions. As can be seen in the Synthesis View plots (Figure 29and Figure 30) no SNP was significantly associated with DR at a significance threshold of $p < 0.05$.

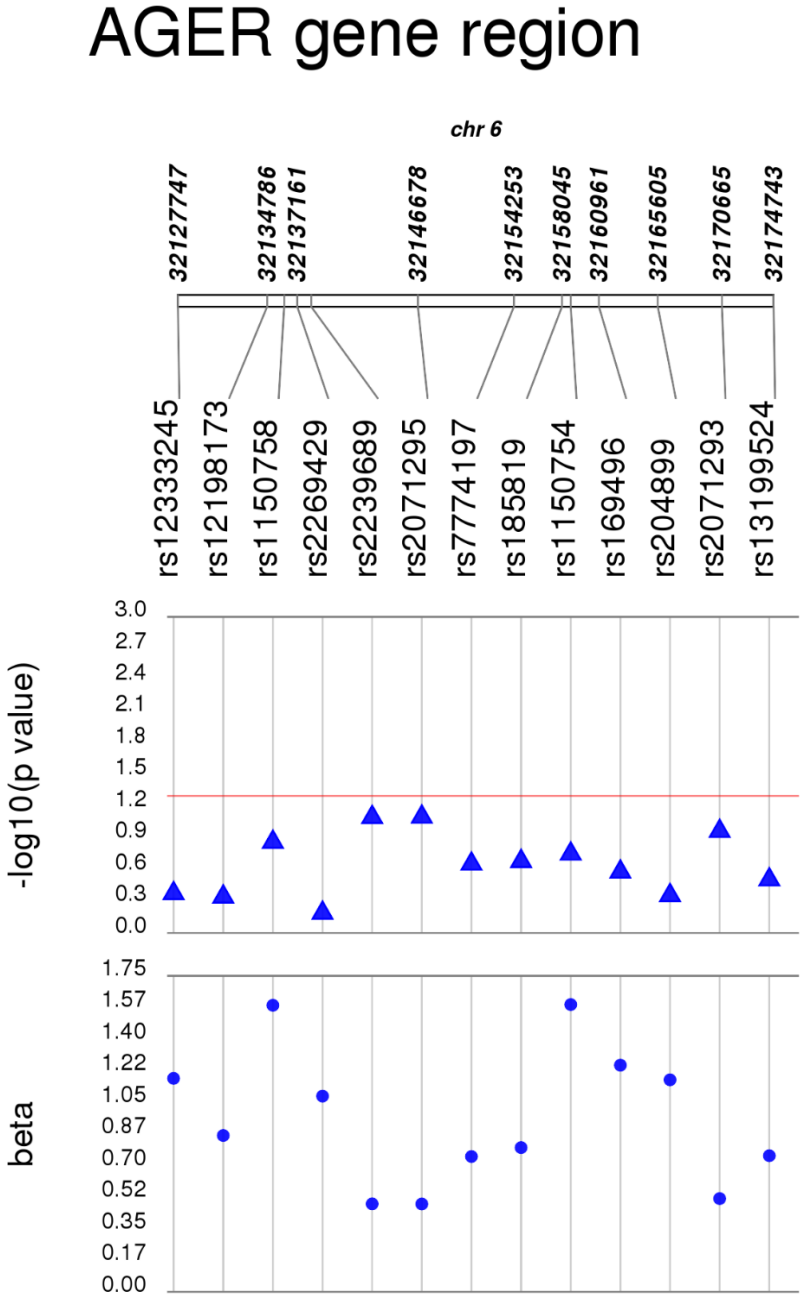


Figure 29: Synthesis view plot of association results for 13 common variants(MAF > 0.05) in the *AGER* gene region.

The model was adjusted for age, sex, and mean blood serum glucose levels. P-values are represented by the blue arrows and are transformed by the –log10, with the threshold of p = 0.05 marked by the red line. Colored arrows also show the direction of effect (beta).

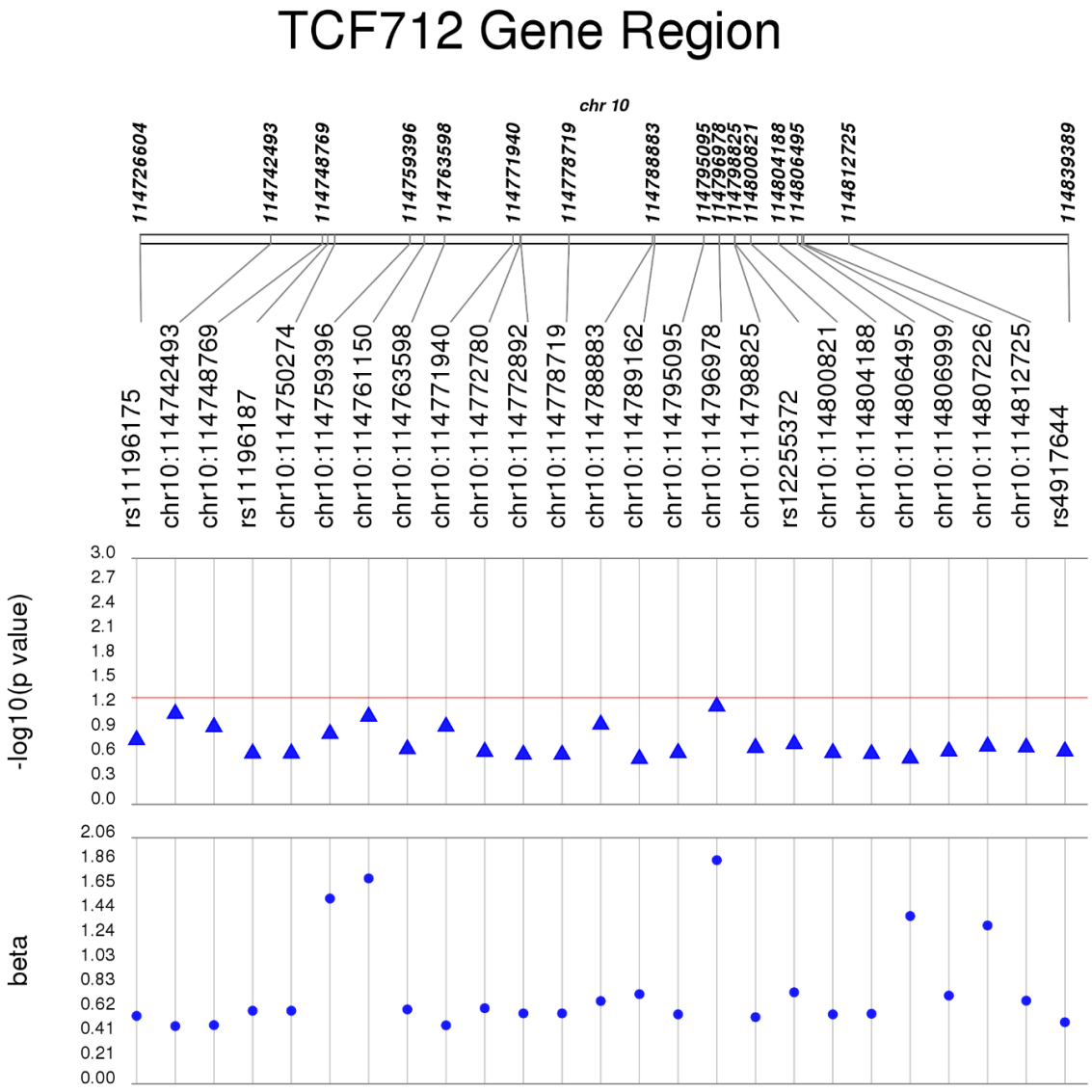


Figure 30: Synthesis view plot of association results for the top 25 common variants (MAF > 0.05) in the *TCF7L2* gene region

The model was adjusted for age, sex, and mean blood serum glucose levels. P-values are represented by the blue arrows and are transformed by the –log10, with the threshold of p = 0.05 marked by the red line. Colored arrows also show the direction of effect (beta).

Discovery study of novel risk loci correlated with diabetic retinopathy in African Americans

Results

We tested for an association between DR and all SNPs available on the Metabochip with a MAF > 0.05 under two models (Figure 31 and Figure 32). In the model adjusted for age and sex, there were few variants that were nominally associated at p $<10^{-4}$ (Supplementary Table 13).

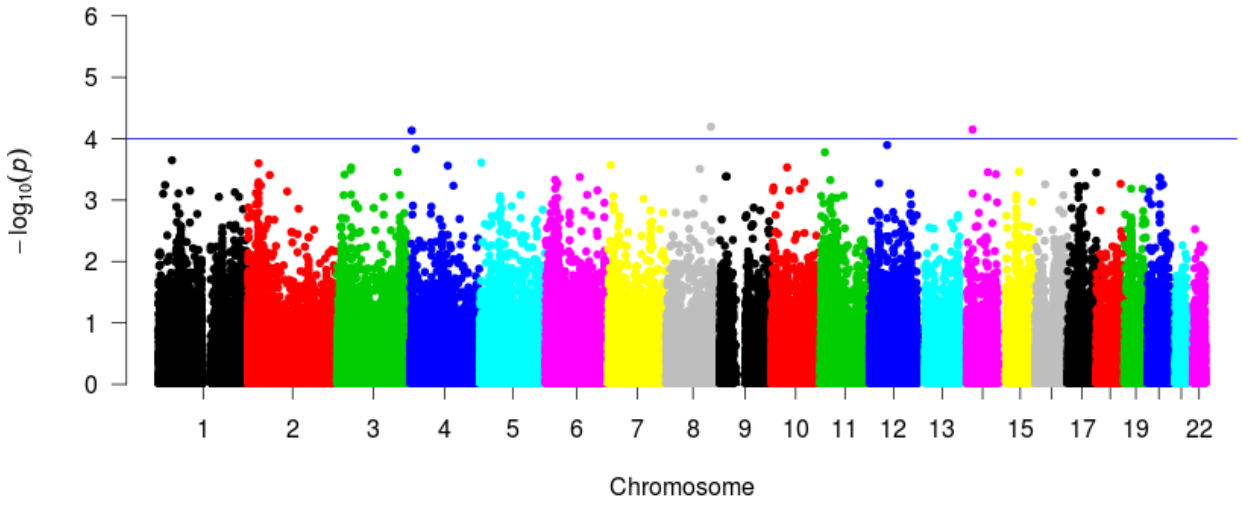


Figure 31: Manhattan plot of EAGLE BioVU African American DR genetic association results. Logistic regression assuming an additive genetic model was performed for 119 cases and 434 controls adjusted by age and sex. P-values ($-\log_{10}$) on the y-axis) for each test of association are plotted by chromosome (x-axis). The blue line depicts a suggestive significance threshold of p =10^{-4}.

In the model adjusted for age, sex, and mean blood serum glucose levels, again no SNP met Bonferroni significance for multiple testing adjustment. One SNP (rs7076968) on chromosome 10, located within the intergenic region of *ZCCHC24-EIF5AL1* had a p - value of 7.60E-06 (Table 8). In the general EAGLE BioVU African American population (n=11,521), this SNP has a MAF of 5.5%. The MAF was lower in T2D controls (MAF = 3.5%) and notably higher in DR cases (MAF = 9.7%).

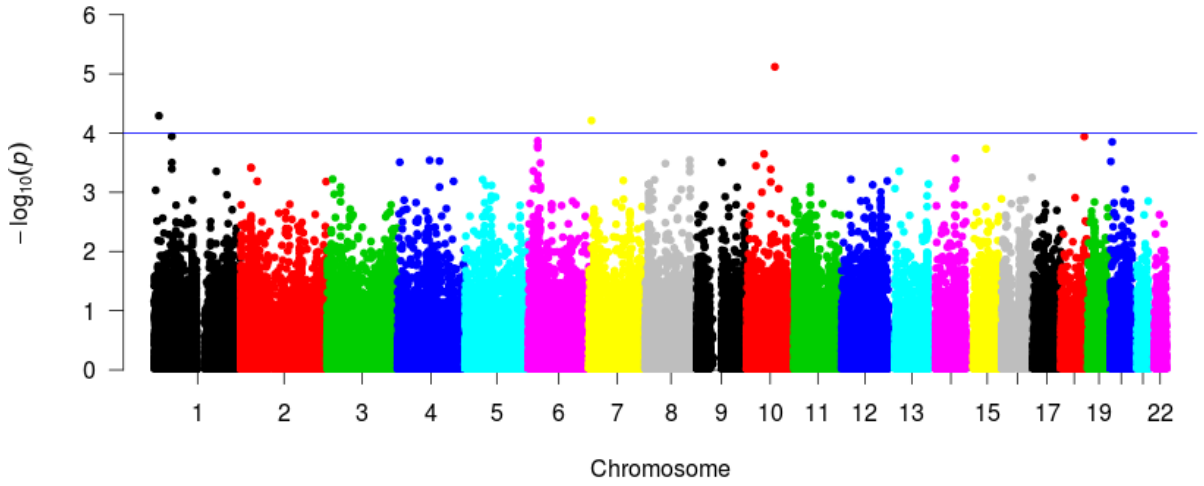


Figure 32: Manhattan plot of EAGLE BioVU African American DR genetic association results.

Logistic regression assuming an additive genetic model was performed for 119 cases and 434 controls adjusted by age, sex, and mean blood glucose levels. P-values ($-\log_{10}$) on the y-axis) for each test of association are plotted by chromosome (x-axis). The blue line depicts a suggestive significance threshold of $p = 10^{-4}$.

Table 24: Twenty most significant results for diabetic retinopathy African American Metabochip genetic association.

Logistic regression assuming an additive genetic model was performed for 119 cases and 434 controls adjusted by age, sex, and mean blood serum glucose levels.

CHR	SNP	Gene	Allele	OR	CI	P
10	rs7076968	ZCCHC24-EIF5AL1	A	5.54	2.62-11.71	7.60E-06
1	rs820626	KAZN	C	2.28	1.53-3.41	5.11E-05
7	rs7790518	C1GALT1	C	2.86	1.71-4.77	6.12E-05
4	rs6849073	upstream-NDST3	A	0.49	0.33-0.72	2.97E-04
20	rs6038207	GPCPD1	A	0.46	0.31-0.68	1.41E-04
10	rs4253162	ERCC6	A	2.13	1.42-3.17	2.25E-04
14	rs3917187	TGFB3	G	0.47	0.31-0.70	2.68E-04
20	rs3830064	PDYN	C	2.16	1.42-3.28	3.01E-04
6	rs2076310	RXRB	G	2.09	1.40-3.13	3.20E-04
9	rs17062682	TJP2	A	2.12	1.41-3.19	3.12E-04
18	rs1529220	downstream-GTSCR1	G	0.44	0.29-0.67	1.14E-04
1	rs12090545	FAF1	G	2.30	1.46-3.62	3.14E-04
1	rs12086130	FAF1	A	2.46	1.56-3.89	1.13E-04
4	rs10516865	CCSER1	A	2.35	1.48-3.74	2.88E-04
4	hg18_4_6916470	-	G	2.99	1.65-5.43	3.11E-04
8	chr8:126581940	-	G	2.48	1.52-4.05	2.85E-04
6	chr6:25835975	-	G	2.14	1.44-3.17	1.61E-04
6	chr6:25826741	-	T	2.12	1.43-3.14	1.78E-04
6	chr6:25820412	-	A	2.17	1.46-3.23	1.35E-04
15	chr15:55827973	-	G	0.39	0.24-0.64	1.84E-04

OR = odds ratio
CI = confidence interval

Role of Mitochondrial variation in risk of diabetic retinopathy in diverse populations ascertained in NHANES

Limited progress has been made in identifying susceptibility variants for diabetic retinopathy (DR). To date, most genetic studies have focused on the nuclear genome and excluded analysis of the mitochondrial genome (mtDNA), regardless that mitochondria are known to play a pathological role in DR. Hyperglycemia-induced oxidative stress has been shown to impair mitochondrial function, damage mtDNA, and increase apoptosis of retinal capillary cells resulting in damage to the retinal microvasculature and development of DR. Here we performed a meta-analysis with mitochondrial variants genotyped in a diverse set of NHANES populations to ascertain their potential contribution to DR risk.

Methods

Study populations

Individuals with T2D were collected from the NHANES III and NHANES 2007-2008. The NHANES surveys were described in greater detail in Chapter 2: Association of mitochondrial variants and haplogroups with AMD in diverse populations. In our study, diabetes was defined by a combination of glycated hemoglobin levels (>6.5%) and an answer of "Yes" to questions of "Have you ever been diagnosed with diabetes" and "Are you currently being treated for diabetes" in NHANES III/2007-2008. Fasting plasma glucose and glucose tolerance tests are only available for morning participants in NHANES 2007-2008. In NHANES III these test results are questionable as a number of participants failed to fast before the test. NHANES III did not differentiate between T1D and T2D. "Age at diagnosis" was utilized to remove any cases/controls that were diagnosed with diabetes before 30 years to avoid potential confounding by T1D patients. DR cases were identified from adults greater than 40 years of age with T2D and gradable fundus photographs. Diagnosis criteria were set according to the Airlie House Classification Scheme at the Wisconsin-Madison University, Ocular Epidemiology Grading Center. Controls are individuals greater than 40 years at the time of survey participation that were diagnosed with T2D years yet showed no sign of diabetic retinopathy upon funduscopic exam. Race was collected by self-identifying surveys in NHANES and verified via Ancestry Informative Markers (AIMs).

Statistical Methods

Individual SNPs and haplogroups were tested for association with DR using logistic regression assuming a dominant genetic model in each NHANES dataset separately. Models were adjusted for age, sex, body mass index, and glycated hemoglobin levels (HbA1c). NHANES analyses were conducted in SAS v9.2 (SAS Institute, Cary, NC) using the Analytic Data Research by Email (ANDRE) portal of the CDC Research Data Center in Hyattsville, MD. Individual NHANES survey results were then meta-analyzed using a fixed-effects inverse-variance weighted approach in METAL(Willer et al., 2010).

Results

Study Population

In NHANES III, African American DR cases were younger on average in comparison to European American and Mexican Americans (Table 3) while in NHANES 2007-2008, the average age of cases was similar for African American and Mexican Americans (Table 4). NHANES 2007-2008 participants were generally obese (BMI > 30 kg/m^2) regardless of race/ethnicity or case/control status and tended to have better control of his/her diabetes. On average, the duration of T2D for African American cases and controls were similar (less than 2 years) across both surveys. Participants within a particular population were less likely to female and had lower HbA1c in the 2007-2008 versus NHANES III.

The NHANES III surveys collected funduscopic images in a randomly selected eye for each participant over the age of 40 years, while the NHANES 2007-2008 ocular examination component performed funduscopic imaging on both eyes. In order to determine the amount of potential discordancy in diabetic retinopathy assessment across the two NHANES collections, we calculated the percent of individuals with varying diabetic retinopathy severity scores in NHANES 2007-2008 populations. Diabetic severity scores were numerated 1-4. The scores were coded as follows: 1 = no retinopathy, 2 = mild non-proliferative retinopathy, 3 = moderate to severe non-proliferative retinopathy, and 4 = proliferative retinopathy. Participants with missing scores were removed from the calculations.

Table 25: NHANES III Type 2 diabetes and diabetic retinopathy population demographics

	European American		African American		Mexican American	
Variable	Case	Control	Case	Control	Case	Control
N	26	68	26	40	31	51
Age at interview	64.2 (12.4)	72.5 (7.0)	57.9 (10.5)	68.1 (6.0)	63.6 (10.3)	67.5 (6.7)
BMI (kg/m2)	29.5 (4.7)	28.9 (5.2)	32.5 (6.1)	30.5 (6.8)	29.5 (5.3)	29.3 (4.8)
Duration T2D	14.8 (9.3)	8.2 (6.5)	12.0 (8.3)	11.3 (9.5)	14.3 (8.3)	8.1 (6.1)
HbA1c (%)	8.3 (1.4)	8.1 (1.4)	8.9 (2.3)	8.2 (1.6)	9.1 (1.8)	8.6 (2.3)
% Female	65.4	48.5	57.7	47.5	51.6	52.9
% Poorly controlled T2D	23.1	2.6	42.3	30	41.9	29.4

Values shown are the means and standard deviations unless otherwise stated
Duration of T2D was calculated as the difference in age at time of interview to the survey question of "age at diagnosis"
An individual with poorly controlled T2D is defined as having a HbA1c greater than 9%

Table 26: NHANES 2007-2008 Type 2 diabetes and diabetic retinopathy population demographics

	European American		African American		Mexican American	
Variable	Case	Control	Case	Control	Case	Control
N	51	122	50	61	29	45
Age at interview	66.1 (11.3)	70.9 (5.9)	62.6 (9.7)	68.4 (6.6)	61.8 (9.5)	69.2 (6.5)
BMI (kg/m2)	31.5 (6.5)	31.1 (5.3)	32.7 (8.3)	31.4 (6.3)	30.8 (4.7)	31.6 (5.6)
Duration T2D	12.3 (7.8)	9.6 (12.2)	13.8 (9.2)	11.9 (11.4)	10.5 (5.5)	7.1 (4.8)
% Female	37.2	34.4	50	55.7	48.3	51.1
HbA1c (%)	7.6 (1.6)	6.7 (1.0)	7.8 (1.6)	7.2 (1.7)	8.4 (1.8)	7.01 (1.4)
% Poorly controlled T2D	23.5	4.9	16.3	11.8	24.1	8.9

Values shown are the means and standard deviations unless otherwise stated
Duration of T2D was calculated as the difference in age at time of interview to the survey question of "age at diagnosis"
An individual with poorly controlled T2D is defined as having a HbA1c greater than 9%

Our results suggest that on average 8 to 10% of participants (Table 5), regardless of race/ethnicity, have some asymmetry in retinopathy severity between both eyes. This difference can occur at any point of the spectrum. For example a participant might present with mild retinopathy in one eye while the other is free of disease or else have disease present in both eyes at varying stages of the disease. Cases in which the severity levels were more than one step removed in a participant's two eyes occurred only nominally, occurring in only one European American, in three African American, two Mexican Americans, and one Hispanic.

Table 27: The percentage of diabetic retinopathy asymmetry in NHANES 2007-2008

NHANES	European American	African American	Mexican American	Hispanics
2007-2008	7.8%	11.4%	8.2%	8.5%

Diabetic retinopathy asymmetry was determined by counting the number of individuals within a racial/ethnic population whose diabetic retinopathy severity score (1-4) differed between their two eyes.

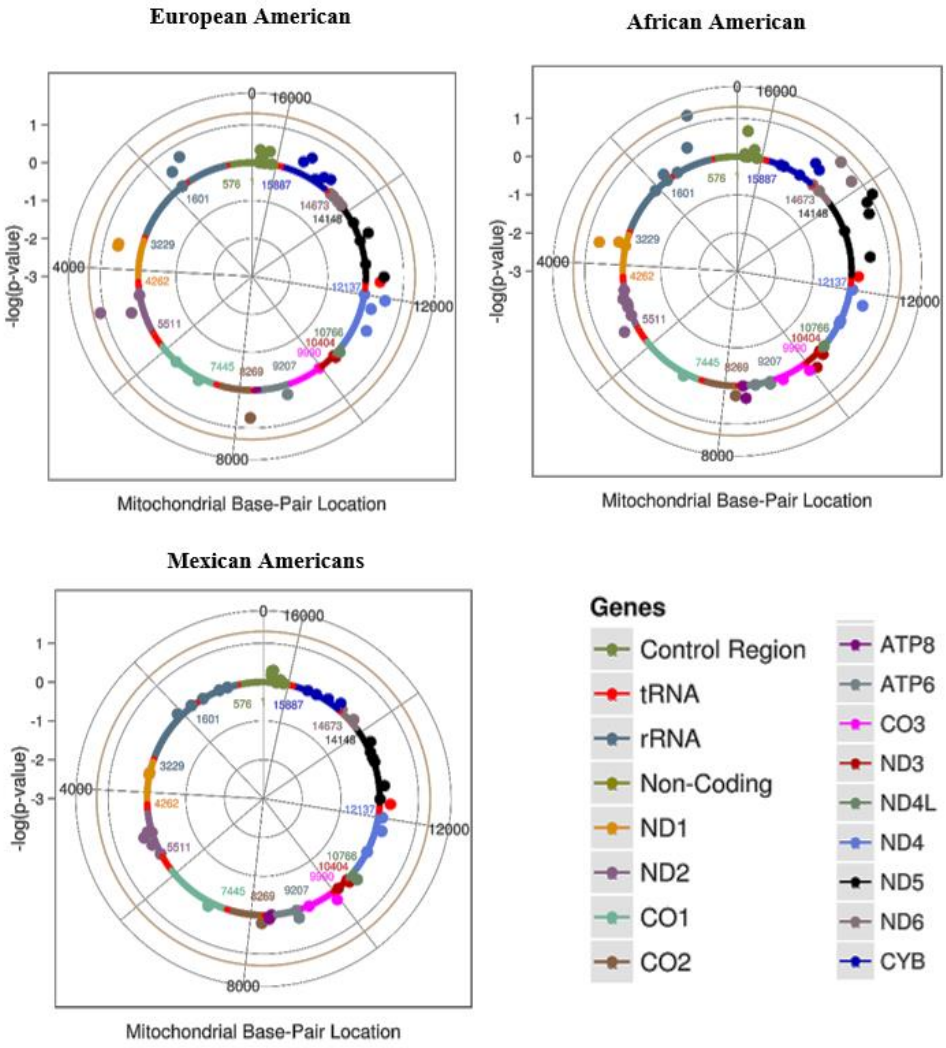


Figure 33*:* Mitochondrial Solar Plots representing single SNP associations by race/ethnicity in adjusted models.

Log(p) values were plotted with the outer brown circle representing a significance threshold of p=0.05. SNPs are color coordinated by mitochondrial gene/region as denoted in the legend.

Mitochondrial genetic association results

We did not find an association with DR in any of the populations tested (Figure 33) at a significance threshold of $p < 0.05$ in single variant tests nor in haplogroup analyses.

Discussion and Summary

In our study of the genetic architecture of diabetic retinopathy in African Americans, we performed directed tests of association for *AGER* rs2070600 and *TCF7L2* rs7903146 as well as tested over 126,000 SNPs on the Metabochip loci previously associated with cardiovascular and metabolic traits. In our hypothesis-driven study of the *AGER* and *TCF7L2* regions, our analyses did not generalize previous associations from European-descent or Indian population studies at a liberal significance threshold of $p < 0.05$. Although, we did not generalize these associations our results did trend in the same direction of effect for both variants. Also, our study did not find an association with DR in our hypothesis-generating study of the variation covered by the Metabochip at a significance threshold corrected for multiple testing ($p < 10^{-7}$). However, we did identify a few variants at a suggestive threshold of $p < 10^{-4}$ that warrant further study in future work. Lastly, in our exploration of potential mitochondrial genetic variants we identified rs2853520 (A) was borderline associated with DR in African Americans at a p=0.052 in NHANES III/2007-2008.

Generalization of previously identified diabetic retinopathy variants

Briefly, we tested 13 common SNPs in the advanced glycosylation end product specific receptor (*AGER*) gene on chromosome six, which spans 3.3 kbp and 148 common SNPs in the Transcription Factor 7-like 2 (*TCF7L2*) gene on chromosome 10 spanning 217.4 kbp. Our study had sufficient case/control numbers to generalize previously associated variants with odds ratios greater than 1.70 or inversely 0.60(Balasubbu et al., 2010; Fu et al., 2010; Grassi et al., 2011) at common allele frequencies. Still, we did not generalize these associations in our study. One reason is that previously identified SNPs such as rs7903146 in *TCF7L2* had allele frequencies below 5%. In our study rs7903146 had an odds ratio of 1.12 with $p = 0.49$. The lack

of generalization could be due to major differences in LD across racial/ethnic populations as observed in the *TCF7L2* region. *AGER* SNP rs2070600 was associated with a decreased risk of DR (OR = 0.49) in an Indian population(Balasubbu et al., 2010). In the Indian population the minor allele frequency (A) ranged from 3.5-5%, while in the EAGLE BioVU African Americans it occurred at only 1%, which dropped it from our final study. When testing without the minor allele frequency cutoff, rs2070600 has an odds ratio of 0.44 and a p = 0.44 in our study. *AGER* is a relatively small gene with few if any variants in strong LD with rs2070600 in either the HapMap III GIH, CEU, or YRI populations. The low minor allele frequency coupled with its presumed function as a nonsynonymous variant suggested that *AGER* rs2070600 may in fact be a casual variant.

Discovery

This study considered how previously GWAS-identified ocular variants and loci associated with cardiovascular and metabolic traits might be associated with DR in African Americans. Not unexpectedly, we failed to identify a significant association after correction for multiple testing ($p < 10^{-7}$). Interestingly, we did identify several variants nominally associated at a $p < 10^{-4}$ which are located in genes known to play a role in epithelial and endothelial tight cellular junctions, wound healing, familial hypercholanemia, and the brain's reward center. At a molecular and biochemical level, these pathways are potentially important in development of diabetes and ocular health. For instance, the SNP rs820626 (OR = 2.28; $p = 5.11 \times 10^{-5}$) and rs17062682 (OR = 2.12; $p = 3.2 \times 10^{-4}$) are located in the kazrin, periplakin interacting protein (*KAZN*) and the tight junction protein 2 (*TJP2*) genes, respectively. Both genes are involved in the assembly and maintenance of the endothelial and epithelial cellular tight junction complexes. The tight junction complexes are integral to forming barriers to protect the body from immunological attacks and mechanical stress. The TJP2 protein is a component of the tight junction complex. Mutations in *TJP2* were found to cause familial Hypercholanemia in the Lancaster County Old Order Amish(Carlton et al., 2003). Hypercholanemia is a liver disease characterized by a decrease in bile acid, resulting in malabsorption of fat and consequently severe vitamin K and D deficiency. This mutation is thought to cause leaks in the tight

junction complexes of the liver and lead to loss of bile acid through drainage into the plasma. Additionally, kazrin encoded by *KAZN* co-localizes with desmoplakin and periplakin in the assembly of desmosomes(Groot et al., 2004; Nachat et al., 2009). Desmosomes form cellular adhesion complexes that link to filaments of the intracellular keratin cytoskeleton. Studies have found that kazrin interacts with BAC, ARC, and Bax where it may be regulating cellular apoptosis(Wang et al., 2009). As discussed in the introduction of this chapter, hyperglycemia can disrupt cell survival pathways that inhibit apoptotic events. Release of cellular debris from apoptosed cells causes inflammation in the ocular environment, which can damage pericytes. Pericytes form part of the blood-retinal barrier and contribute to the junctional complexes with endothelial cells. Permeability in the ocular vasculature can trigger angiogenesis which leads to the hallmark signs of diabetic retinopathy.

Mitochondrial contribution to diabetic retinopathy in diverse populations

Mitochondria are important for the production of cell energy and regulation of cellular apoptosis. Degradation and mutations in the mitochondrial genome can lead to Alzheimer's disease, various cancers, and age-related macular degeneration. Recently, a mitochondrial phenome-wide association study found additional evidence that supports the hypothesis that mitochondria impact risk of cholesterol levels and T2D(Mitchell et al., 2014b).

Given the elevated energy needs of retinal tissue, mitochondria become absolutely vital in the maintenance and functional mechanics of the visual pathway. Therefore, we tried to determine whether mitochondrial variation specifically contributed to risk of DR in a diverse population of NHANES participants. We performed a meta-analysis to test for an association with 60 mitochondrially encoded variants and DR from the NHANES III and NHANES 2007-2008 European American, African American, and Mexican American populations.

We did not find an association with DR in any of the populations tested (Figure 33). The variant rs2853520 (A) found in the mitochondrially encoded 12s RNA gene, was found to trend towards significance in this

African American DR case-control study at p=0.052. In the individual NHANES III analysis, rs2853520 was found to be significant at p=0.047. Notably, the ancestral allele frequency (A) is 14.5 % in the general HapMap Yoruba (YRI) population but was approximately 22% in this NHANES African American populations (23.5% and 21.7%, NHANES III & 2007-2008).

Strengths and Limitations

Beyond limited sample size, a significant limitation of this study is trait heterogeneity and phenotypic variability, known confounders in genetic studies. These confounders may be minimized by the stratification of DR cases into severity levels or subcategories classified by environmental confounders such as income, education, and access to health care. Unfortunately, study power would be detrimentally reduced by the restricted number of subgroup cases. Still, incorporating a variable to model access to health care in future studies may provide us with a clearer image of the underlying genetic architecture of DR given that strict control of a patient's diabetes can drastically delay or prevent progression of DR.

CHAPTER VI

SUMMARY AND FUTURE DIRECTIONS

The impact of vision impairment and blindness on the daily life of individuals will continue to grow as populations age. With aging comes a greater risk of developing age-related ocular diseases such as age-related macular degeneration, glaucoma, and diabetic retinopathy. While a substantial body of research has continued to improve screening regimens and treatment options for these conditions, much of the genetic etiology has yet to be determined. Genetic studies have aided in the identification of biologically actionable targets. Treatment options have been successfully developed for patient's suffering from neovascular age-related macular degeneration by pharmacological targets inhibiting VEGF production, but the neovascular subtype only accounts for 10 to 15% of age-related macular degeneration cases. Similarly, pharmaceuticals are available to protect against vision loss in glaucoma patients who suffer from IOP-dependent forms of the condition while options for normal tension glaucoma remain elusive. Limitations in the current field of medicine can also be seen in patient's suffering from diabetes that despite strict adherence to diabetes treatment regimens are eventually blinded by diabetic retinopathy.

The scientific community strives to elucidate the genetic and molecular etiology of common disease which can lead to new therapies and risk reduction. Indeed this is the objective of personalized medicine, whereby through the understanding of the genetic architecture and molecular constituents of disease in larger populations it is possible to distill this knowledge down into targeted treatment options to meet the needs of unique patients.

One step toward explaining the role of genetic variation in complex ocular phenotypes is the identification of shared and population-specific variants in disease risk. In this body of work, I assessed the role of previously identified variants on disease risk for age-related macular degeneration, primary open-angle glaucoma, and diabetic retinopathy in diverse populations. I began by determining whether variants strongly associated with age-related macular degeneration in European-descent populations, similarly drove risk of

disease in African Americans, Mexican Americans, and Singaporeans from the Population Architecture using Genomics and Epidemiology (PAGE) Study. Differences in disease prevalence and severity range across racial/ethnic populations, with individuals of European ancestry being at greatest risk while those of African ancestry are at the lowest risk for age-related macular degeneration. This study was a meta-analysis of targeted genotyping variants including *CFH* rs1061170, *ARMS2* rs10490924, *C2* rs547154, additional complement system, and lipid-related variants. We replicated up to 54% (7/13) of the previously reported associations for age-related macular degeneration in European Americans including *CFH* rs1061170 and *ARMS2* rs10490924. We additionally replicated an association with *ARMS2* rs10490924 that generalized to Mexican Americans in this study, which had already been reported in NHANES III(Spencer et al., 2012b). None of the complement system or *ARMS2* variants generalized to African Americans or Asians. This study did identify several associations between age-related macular degeneration and lipid trait-associated variants in *LPL* and *CETP* in African Americans and Mexican Americans. The effect sizes for *CFH* rs1061170 and *ARMS2* rs10490924 in European populations were large. If the effect of these variants on disease risk had been equally as strong in African Americans and Singaporeans, we should have been able to detect the associations despite the small sample sizes available for the non-European populations. Lack of generalization in non-European populations suggests that other genetic or environmental variables may be driving risk in these populations.

Nearly all of the age-related macular degeneration risk loci are located within the nuclear genome and account for more than 60% of the heritable risk(Fritsche et al., 2014). A number of large-scale GWAS studies have been carried out which have identified new risk loci with ever decreasing effect sizes that do not account for the remainder of the heritability. I therefore set out to determine whether part of the remaining heritability might be explained in part by mitochondrial variation and interaction effects between environmental factors and strong age-related macular degeneration variants. To do this, we accessed NHANES III and NHANES 2007-2008 participant data and genotyping information from the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study which genotyped 63

mitochondrial variants in European Americans, , and Mexican Americans. We performed tests of association between age-related macular degeneration and previously associated haplogroups J, T, and U. Tests of association were also performed between each mitochondrial variant and age-related macular degeneration. We failed to replicate previous associations of the European haplogroups J, T, and U with risk of age-related macular degeneration in NHANES European Americans nor did we find an association in African Americans. Interestingly we observed several mitochondrial variants that were associated with age-related macular degeneration risk within the meta-analyzed Mexican American population after adjusting for well-known modifiers. Mexican Americans experience prevalence rates of age-related macular degeneration similarly to those of European-descent populations, yet we did not observe a strong association with European-identified risk variants from the specific nuclear genome variants also genotyped in Mexican Americans. The results of the mitochondrial analyses suggest that risk may be driven by variation in the mitochondrial genome within this population.

The study of main effects in genetic association analyses relies on the premise that genetic variants singularly act in an additive or Mendelian fashion. Yet the current state of the field acknowledges that the sum of all known additive variants in non-Mendelian disease do not add up to the whole of the heritability. We therefore further explored whether age-related macular degeneration risk is additionally modified by interactions between environmental factors such as blood serum carotenoid levels and strong European index variants rs10490924 *ARMS2*, rs1061170 *CFH*, and rs547154 *C2*. This study was performed in the NHANES III populations where in a model adjusted for age, sex, BMI, and smoking status the variant rs547154 *C2* interacted with body-mass-index to significantly modify risk of age-related macular degeneration in European Americans.

Assessing the effects of highly impactful risk loci across population can provide important insights into common disease pathways. These pathways may then become relevant for future work targeting expression of these genes or gene products. Still, inconsistencies in the associations of these variants across populations, after taking into account allele frequencies, may suggest that the fundamental etiology varies

based on differences in environmental influences or genetic background. Age-related macular degeneration only occurs in 2.4% of African American adults and yet the primary *CFH* and *ARMS2* variants occur with a MAF of over 10%. These variants are not rare in African Americans, but they do not appear to confer strong risk similar to that observed in European-descent populations. Studies addressing heritability of age-related macular in African Americans and the contribution of age and smoking to risk of disease are also lacking.

Large-scale genetic association studies are primarily being carried out in European and Asian descent populations with little effort being made to elucidate the role of genetics in underrepresented populations. Narrowing this deficit may lead to new insights in diseases that occur predominantly in these underrepresented populations such as diabetes, obesity, and glaucoma. A major hurdle in the study of minority populations is ascertainment of cases and controls which can be a lengthy and expensive process. To rapidly accrue data in a cost-effective manner, I accessed de-identified patient medical information from the Vanderbilt University Medical Center (VUMC) electronic medical records. EMRs are becoming a vital resource for research studies given the dense clinical data potentially available on tens of thousands of patients. My goal was to develop an efficient method for identifying diabetic retinopathy and primary open-angle glaucoma cases and controls for inclusion in genetic association studies, as outlined in Chapter 3. VUMC's EMR available for research does not currently make available the gold-standard clinical exams and records necessary for the definitive diagnosis of either diabetic retinopathy or primary open-angle glaucoma. Lack of access to records from the Vanderbilt Eye Clinic proved to be a challenge. To address this issue, I created a set of phenotype algorithms incorporating International Classification of Diseases codes 9th revision (ICD-9), Current Procedural Terminology (CPT) codes, and free text searches to identify African Americans with and without diabetic retinopathy and primary open-angle glaucoma. These algorithms, when coupled with manual-review follow-up, allowed me to identify high-quality cases and controls for each study despite the lack of gold-standard data available in the EMR for research. Our work further explored common genetic variants in relation to risk of diabetic retinopathy and primary-open angle

glaucoma in African Americans, which is discussed in more detail within Chapters 4 and 5. I performed hypothesis-driven studies targeting SNPs that were previously associated with diabetic retinopathy and primary open-angle glaucoma. Previously identified SNPs did not generalize in our study. Lack of generalization can be due to low statistical power or difference in linkage disequilibrium across populations. These studies highlight many of the challenges faced when working with small sample sizes and multiple populations.

Nominally significant, we identified rs3775202 in *VEGFC* associated with an increased risk of primary open-angle glaucoma (odds ratio = 1.92; $p = 9.70 \times 10^{-5}$) in EAGLE BioVU African Americans. This SNP in conjunction with other potential associations suggests a role for angiogenesis (i.e. *VEGFC* and *HTRA*) and ATP/energy mediation via mitochondria homeostasis (*C21orf33* and *IBA57*).

Future Directions

The work presented here is a small milestone in the larger work needed to uncover the genetic architecture of common, complex ocular disease. While my work has highlighted that African American genetic risk factors for age-related macular degeneration, primary open-angle glaucoma, and diabetic retinopathy may differ from European-descent populations, much work is left to determine whether these differences are due to differences in genetic architecture, environmental factors, or socioeconomic conditions. Additional studies to determine the role of common genetic variants in African Americans would involve ascertaining new cases and genotyping them on a GWAS chip and/or collaborating with other researchers to increase our study size.

Indeed, the PAGE I Study, which included Atherosclerosis Risk in Communities (ARIC), the Cardiovascular Health Study (CHS), and EAGLE NHANES III, only had up to 95 African American cases of age-related macular degeneration. The addition of NHANES 2007-2008 would only increase the case sample size by seven. To date, only one GWAS in the NHGRI GWAS catalog for age-related macular degeneration includes African American cases(Ryu et al., 2010).The lower prevalence of age-related

macular degeneration makes increasing the sample size for GWAS and other genomic studies challenging. The challenge is further compounded by the fact that there are only a limited number of cohorts with African American participants and DNA samples available for research. Thus, for more prevalent diseases such as primary open-angle glaucoma and diabetic retinopathy, sample size and power are still very much real issues for non-European descent populations.

I have attempted to address this issue using electronic medical records linked to DNA samples as a source for new cases and controls for study. Although successful, additional EMRs linked to biobanks will be needed to amass sufficient sample sizes for powerful genomic studies. Therefore, future directions should include collaborations with biobanks such as Northwestern University and Mt. Sinai, both members of the eMERGE Network, to identify additional non-European cases and controls for these common ocular diseases. Previous studies have suggested that algorithms to extract cases and controls are portable(Denny et al., 2011), so it is reasonable to expect that my algorithm can be deployed by these other study sites to identify additional cases and controls. Beyond eMERGE, other sources of new cases could include the Veterans Affairs Million Veteran Project, a research program funded by the Department of Veterans Affairs Office of Research Development, as well as new cohorts resulting from the new Precision Medicine initiative proposed in early 2015 by President Barack Obama(Kaiser and Servick, 2015). Regardless of the source of data, a concerted effort will be needed to ensure there are sufficient sample sizes of non-European-descent participants to conduct properly powered discovery studies for ocular diseases.

Gender disparity is one area of research I would like to explore in future work. In 1993 congress passed the National Institutes of Health Revitalization Act of 1993 mandating that NIH funded studies include women and minorities. Blatant gaps still persist in research and clinical trials that proactively determine gender differences in disease risk and treatment efficacy of pharmaceuticals and surgical intervention outcomes(Institute of Medicine (US) Committee on Women's Health Research, 2010). Cardiovascular disease is the number one killer in women across all race/ethnicities and while women are now included in cardiovascular studies, they are either underrepresented or studies fail to examine gender differences in

treatment outcomes. For example, a clinical trial studied the safety and efficacy of a defibrillator implant in a cohort of men and women(Moss et al., 2002), where women comprised 15% of the cohort. The device was later approved by the FDA but was subsequently found not to offer the same protective effect in women(Ghanbari et al., 2009). Strong gender differences also occur in ocular disease.

In age-related macular degeneration, epidemiology studies have suggested that women are at greater risk than men(Chakravarthy et al., 2010; Cheung et al., 2012; Tomany et al., 2004) for development of disease. The life expectancy of a woman is greater than a man's in developed countries which may explain the difference in risk for age-related diseases. Yet genetic, biochemical, and life style contrasts between the sexes is often ignored. Only within the last few years have age-related macular degeneration researchers pursued the cause of this gender disparity in genetic association studies. Female-specific susceptibility loci and pathways have been identified in the death-associated protein-like 1 (*DAPL1*) on chromosome 2(Grassmann et al., 2015) and the *VEGF* pathway for women with neovascular age-related macular degeneration(Courtenay et al., 2014). Gender studies of ocular blood flow suggest that female sex hormones are protective in diseases that affect the ocular vasculature(Schmidl et al., 2015) providing further incentive to studying the effects of gender on common, complex ocular diseases. Similarly a role for female sex hormones in glaucoma is supported by the expression of estrogen receptors in retinal ganglion cells(Munaut et al., 2001) and reduced risk for women who enter menopause after age fifty-four(Pasquale et al., 2007). The Glaucoma Genes and Environment study and the National Eye Institute Glaucoma Human Genetics Collaboration consortium found an association in estrogen pathways for total and high pressure open-angle glaucoma risk in women, which was notably absent in men(Pasquale et al., 2013).

Future studies could assess gender differences in disease risk by stratifying cases and controls by sex. The decrease in sample size can be mitigated by actively recruiting women from local communities or hospital databases. Assessing long term effects of female sex hormone exposure on risk of AMD or POAG is possible by modeling age at menopause and age at menarche from EMRs(Malinowski et al., 2014) in conjunction with use of hormone therapy and birth control.

Further extensions for studying the effect of gender on genetic models is the incorporation of female specific environmental exposures. Females differ from males, beyond differences in estrogen and androgenic hormone levels, in key behavioral ways. Women significantly spend more and consume more personal care and cosmetic products then men. Spending differences occur by birth cohort and by economic influences, notably the "lip stick effect(Hill et al., 2012)." The infamous "lip stick effect" is the hypothesis that when faced with an economic crisis consumers, notably women, reduce their discretionary spending by purchasing less costly luxury goods. A prime example is the purchase of a high-end tube of lipstick versus a new leather jacket. Spending patterns on personal care products, which include lotions, cosmetics, bathing, and shaving products, varied by age and sex in a United Kingdom study of consumer spending(Twigg and Majima, 2014). In the 1960s young women shopped more frequently than older women over the age of seventy-five. This trend was reversed by 2011. These patterns for consumption of personal care products extend to other European countries. In the Netherlands, women on average used seventeen personal care products a day while men used six(Biesterbos et al., 2013). Of most interest is the use of cosmetics particularly cosmetics applied to the area on or immediately surrounding the eye.

The eyelid performs a number of biological functions that include creating a physical barrier between the anterior surface of the eye globe, tear generation, and light regulation. At less than 1mm, the skin of the eyelid is one of the thinnest in the body. Absorption of chemicals, such as corticosteroids, through the eyelid may lead to increased risk of eye disease such as glaucoma(Cubey, 1976).

I would like to explore the potential consequences to ocular health when individuals are exposed to known toxic chemicals commonly found in cosmetics. Endocrine disruptors bisphenol A(Steinmetz et al., 1997), phthalates(Harris et al., 1997), and parabens(Okubo et al., 2001) and antimicrobial compounds (i.e., Triclosan) are present in many cosmetics. These chemicals are routinely found in urine samples of adults and children in the United States(Calafat et al., 2008). In a Swedish study of mothers and children, paraben metabolites were found in greater concentrations in women who used more personal care products(Larsson et al., 2014). The Scientific Committee on Consumer Safety in Europe initially set daily allowances for

Triclosan exposure at 17.4g/day(Scientific Committee on Consumer Safety Europe, 2010) but due to toxic side effects and over exposure Triclosan was banned from use in products distributed in member countries.

In summary, the work presented here contributes data towards the genetic architecture of age-related macular degeneration, primary open-angle glaucoma, and diabetic retinopathy in non-European descent populations. Larger samples sizes, however, are essential to catalogue the complete genetic architecture of these ocular diseases in African American and Mexican Americans. Future work is needed to identify cases as well as the additional environmental data needed to identify all risk factors that lead to the development and perhaps clues to the prevention of these leading causes of vision loss in the United States.

APPENDIX

Chapter 2 Appendix

Supplementary Table 1: All Model 1 results for AMD meta-analysis association tests. Results are shown for model adjusted for site of ascertainment. Coded Allele frequency is shown for combined case/control.

SNP ID	*Gene*	Chr	OR	lower CI	upper CI	p-value	CA	CAF (%)	Race
rs1061170	*CFH*	1	1.550	1.340	1.780	$3.05X10^{-8}$	C	0.372	European American
			1.360	0.911	2.040	0.135	C	0.382	African American
			1.220	0.702	2.131	0.479	C	0.200	Mexican American
			4.430	0.350	55.291	0.248	C	0.036	Asian
rs203674	*CFH*	1	0.226	0.018	2.819	0.248	T	0.961	Asian
rs3753394	*CFH*	1	1.251	1.034	1.514	0.034	T	0.294	European American
			1.451	0.596	3.528	0.412	T	0.083	African American
			0.968	0.747	1.254	0.805	T	0.542	Asian
rs3753396	*CFH*	1	0.692	0.245	1.955	0.487	A	0.493	Asian
rs3766404	*CFH*	1	1.014	0.586	1.756	0.960	T	0.922	Asian
rs529825	*CFH*	1	0.914	0.328	2.541	0.863	A	0.415	Asian
rs6677604	*CFH*	1	0.767	0.608	0.967	0.039	A	0.217	European American
			0.628	0.308	1.280	0.200	A	0.366	African American
			0.974	0.543	1.747	0.930	A	0.054	Asian
rs800292	*CFH*	1	0.587	0.465	0.741	$3.80X10^{-5}$	A	0.237	European American
			0.552	0.295	1.033	0.063	A	0.692	African American
			0.886	0.680	1.153	0.366	A	0.422	Asian
rs6754295	*APOB*	2	1.107	0.952	1.288	0.225	T	0.751	European American
			0.879	0.450	1.716	0.705	T	0.728	African American
			1.058	0.808	1.386	0.682	T	0.289	Asian
rs11726949	*CFI*	4	1.043	0.764	1.424	0.791	T	0.165	Asian

rs10033900	*CFI*	4	0.969	0.811	1.158	0.753	T	0.475	European American
			1.754	0.939	3.278	0.078	T	0.385	African American
			1.043	0.799	1.362	0.754	T	0.608	Asian
rs547154	*C2*	6	0.688	0.393	1.206	0.192	T	0.058	Asian
rs12678919	*LPL*	8	0.745	0.507	1.095	0.168	A	0.904	European American
			0.733	0.335	1.604	0.437	A	>0.99	African American
			0.583	0.220	1.538	0.276	A	>0.99	Mexican American
			0.990	0.599	1.637	0.970	A	0.893	Asian
rs10503669	*LPL*	8	1.041	0.813	1.333	0.770	A	0.096	European American
			1.503	0.721	3.135	0.277	A	0.06	African American
			1.078	0.406	2.857	0.880	A	>0.99	Mexican American
rs2083637	*LPL*	8	0.927	0.802	1.071	0.346	T	0.717	European American
			0.916	0.447	1.879	0.811	T	0.781	African American
			1.043	0.720	1.512	0.823	T	0.795	Asian
rs2197089	*LPL*	8	0.961	0.840	1.098	0.587	T	0.542	European American
			0.945	0.653	1.368	0.764	T	0.785	African American
			0.906	0.575	1.426	0.669	T	0.498	Mexican American
			1.632	0.568	4.688	0.363	T	0.329	Asian
rs328	*LPL*	8	0.950	0.791	1.140	0.607	G	0.100	European American
			1.750	1.060	2.910	0.030	G	0.071	African American
			1.309	0.556	3.081	0.537	G	<0.01	Mexican American
rs6586891	*LPL*	8	1.052	0.915	1.210	0.511	A	0.652	European American
			0.922	0.617	1.378	0.693	A	0.838	African American
			1.028	0.659	1.601	0.902	A	0.537	Mexican American
rs6987702	*TRIB1*	8	1.007	0.869	1.166	0.932	T	0.728	European American
			1.610	1.027	2.524	0.038	T	0.288	African American
			1.196	0.904	1.581	0.210	T	0.434	Asian
rs1883025	*ABCA1*	9	0.818	0.694	0.964	0.027	A	0.26	European American
			0.871	0.569	1.334	0.526	A	0.35	African American
			0.934	0.568	1.535	0.787	A	<0.01	Mexican American
			0.403	0.085	1.909	0.252	A	0.231	Asian
rs1323432	*GRIN3A*	9	0.831	0.581	1.187	0.348	T	0.893	European American
			0.888	0.326	2.420	0.817	T	0.97	African American

			0.886	0.400	1.959	0.764	T	<0.01	Mexican American
rs3890182	*ABCA1*	9	0.942	0.762	1.165	0.612	A	0.117	European American
			1.041	0.669	1.622	0.857	A	0.124	African American
			0.901	0.373	2.181	0.817	A	>0.99	Mexican American
rs3905000	*ABCA1*	9	0.980	0.830	1.157	0.825	A	0.132	European American
			1.036	0.612	1.753	0.895	A	0.169	African American
			0.782	0.323	1.890	0.586	A	>0.99	Mexican American
			0.816	0.489	1.363	0.437	A	0.066	Asian
rs4149268	*ABCA1*	9	1.047	0.902	1.216	0.575	A	0.365	European American
			0.880	0.568	1.363	0.567	A	0.684	African American
			1.495	0.881	2.535	0.136	A	0.341	Mexican American
			0.669	0.394	1.134	0.902	A	0.298	Asian
rs4149274	*ABCA1*	9	1.072	0.810	1.419	0.625	A	0.293	Asian
rs471364	*TTC39B*	9	0.964	0.770	1.206	0.767	A	0.878	European American
			0.935	0.622	1.405	0.746	A	0.815	African American
			0.989	0.425	2.302	0.979	T	<0.01	Mexican American
rs10490924	*ARMS2*	10	1.526	1.289	1.807	$6.36X10^{-6}$	T	0.217	European American
			0.825	0.512	1.330	0.430	T	0.24	African American
			1.634	1.025	2.606	0.039	T	0.262	Mexican American
rs964184	*ZNF259/APOA1*	11	1.254	0.901	1.746	0.217	C	0.837	European American
			1.223	0.650	2.302	0.533	C	0.766	African American
			1.441	0.911	2.279	0.118	G	0.267	Mexican American
			0.890	0.666	1.189	0.429	C	0.786	Asian
rs174547	*FADS1*	11	1.038	0.808	1.335	0.786	T	0.656	European American
			0.878	0.529	1.458	0.615	T	0.911	African American
			1.138	0.687	1.884	0.617	T	0.427	Mexican American
			1.272	0.393	4.118	0.688	T	0.377	Asian
rs28927680	*BUD13*	11	0.871	0.705	1.075	0.237	C	0.072	European American
			0.845	0.550	1.297	0.441	C	0.167	African American
			0.742	0.401	1.373	0.342	G	0.867	Mexican American
rs3135506	*APOA5*	11	0.869	0.696	1.084	0.252	C	0.064	European American
			0.719	0.340	1.519	0.387	C	0.059	African American

			0.883	0.465	1.674	0.702	G	0.873	Mexican American
rs7120118	*NR1H3*	11	1.086	0.940	1.253	0.302	T	0.713	European American
			1.075	0.696	1.660	0.746	T	0.526	African American
			1.285	0.968	1.705	0.082	T	0.237	Asian
rs7395662	*MADD*	11	1.036	0.906	1.184	0.637	A	0.353	European American
			1.116	0.614	2.027	0.720	A	0.545	African American
			0.783	0.598	1.026	0.076	A	0.551	Asian
rs2338104	*KCTD10*	12	1.050	0.920	1.200	0.505	G	0.532	European American
			0.860	0.610	1.220	0.403	G	0.740	African American
			1.695	1.085	2.647	0.020	G	0.452	Mexican American
rs10468017	*LIPC*	15	1.050	0.896	1.231	0.579	T	0.282	European American
			0.752	0.349	1.624	0.469	T	0.182	African American
			1.44	0.848	2.445	0.176	T	0.173	Mexican American
			0.918	0.643	1.309	0.635	T	0.19	Asian
rs1800588	*LIPC*	15	0.830	0.691	0.998	0.068	T	0.223	European American
			0.971	0.650	1.450	0.884	T	0.505	African American
			1.082	0.696	1.683	0.727	T	0.504	Mexican American
rs261332	*LIPC*	15	0.767	0.599	0.982	0.052	A	0.2	European American
			1.112	0.566	2.182	0.758	A	0.25	African American
			0.395	0.045	3.452	0.401	A	0.119	Asian
rs757200	*CACNG3*	16	0.977	0.804	1.187	0.831	T	0.298	European American
			2.270	0.875	5.887	0.092	T	0.068	African American
			1.077	0.808	1.436	0.612	T	0.377	Asian
rs3764261	*CETP*	16	1.138	1.014	1.278	0.043	T	0.327	European American
			0.993	0.706	1.397	0.968	T	0.33	African American
			0.787	0.467	1.329	0.371	T	0.31	Mexican American
			1.066	0.741	1.534	0.730	T	0.169	Asian
rs9989419	*CETP*	16	0.970	0.867	1.086	0.629	A	0.394	European American
			0.990	0.714	1.372	0.951	A	0.565	African American
			0.754	0.464	1.223	0.253	A	0.282	Mexican American
			1.327	0.817	2.153	0.610	A	0.251	Asian
rs1864163	*CETP*	16	1.017	0.827	1.251	0.881	A	0.25	European American
			0.796	0.414	1.530	0.493	A	0.3	African American

			1.041	0.719	1.509	0.831	A	0.137	Asian
rs711752	*CETP*	16	1.101	0.914	1.326	0.353	A	0.43	European American
			0.665	0.315	1.405	0.285	A	0.26	African American
rs7205804	*CETP*	16	0.258	0.053	1.252	0.093	A	0.316	Asian
rs1566439	*CETP*	16	1.021	0.911	1.144	0.742	A	0.61	European American
			0.958	0.592	1.550	0.861	A	0.76	African American
			1.144	0.748	1.748	0.534	A	0.545	Mexican American
			1.038	0.793	1.358	0.788	A	0.4	Asian
rs1800775	*CETP*	16	1.00	0.870	1.160	0.986	C	0.519	European American
			1.570	1.030	2.380	0.037	C	0.415	African American
			1.130	0.700	1.830	0.597	C	0.464	Mexican American
			1.030	0.790	1.340	0.841	C	0.508	Asian
rs1800777	*CETP*	16	0.996	0.600	1.654	0.990	A	0.03	European American
rs255049	*DPEP3*	16	0.852	0.678	1.072	0.208	T	<0.010	European American
			1.047	0.655	1.672	0.849	T	0.891	Asian
rs12596776	*SLC12A3*	16	1.083	0.840	1.396	0.572	C	0.909	European American
			1.187	0.399	3.529	0.758	C	>0.99	African American
			1.014	0.442	2.330	0.974	G	<0.01	Mexican American
rs2217332	*HERPUD1*	16	1.124	0.927	1.362	0.274	T	0.148	European American
			0.936	0.372	2.354	0.888	T	0.114	African American
			1.108	0.725	1.694	0.634	T	0.071	Asian
rs2271293	*NUTF2*	16	1.144	0.921	1.421	0.263	A	0.115	European American
			0.946	0.487	1.835	0.869	A	0.073	African American
			1.242	0.683	2.256	0.477	A	0.147	Mexican American
			0.959	0.506	1.816	0.898	A	0.016	Asian
rs9891572	*METTL16*	17	1.120	0.938	1.338	0.248	T	0.15	European American
			1.488	0.746	2.965	0.259	T	0.182	African American
			0.978	0.749	1.277	0.872	T	0.356	Asian
rs2156552	*LIPG*	18	0.957	0.798	1.148	0.662	A	0.841	European American
			1.396	0.547	3.562	0.485	A	₲	African American
			0.972	0.381	2.477	0.952	T	<0.01	Mexican American
			1.445	0.462	4.517	0.527	A	0.176	Asian

rs4939883	*LIPG*	18	0.997	0.857	1.159	0.967	T	0.165	European American
			1.100	0.718	1.685	0.662	T	0.435	African American
			0.880	0.433	1.789	0.725	T	<0.01	Mexican American
			1.163	0.819	1.651	0.400	T	0.18	Asian
rs2304130	*ZNF101*	19	0.967	0.677	1.381	0.866	A	0.924	European American
			1.050	0.481	2.295	0.902	A	0.811	African American
			0.751	0.532	1.061	0.104	A	0.861	Asian
rs2967605	*RAB11B*	19	0.936	0.784	1.118	0.500	A	0.172	European American
			1.192	0.829	1.713	0.344	A	0.137	African American
			0.768	0.447	1.321	0.34	A	<0.01	Mexican American
			0.992	0.751	1.311	0.955	A	0.587	Asian
rs1800961	*HNF4A*	20	0.642	0.394	1.046	0.102	T	0.028	European American
			4.973	0.795	31.118	0.086	T	0.008	African American
			1.138	0.370	3.502	0.822	T	<0.01	Mexican American
			1.184	0.346	4.049	0.788	T	0.015	Asian
rs7679	*PCIF1*	20	1.014	0.757	1.359	0.931	T	0.819	European American
			0.868	0.474	1.590	0.647	T	0.952	African American
			1.276	0.663	2.453	0.466	T	0.857	Mexican American
			0.869	0.526	1.435	0.584	T	0.972	Asian

₲ Allele frequencies genotyped across more than one study were found to be greatly different

Supplementary Table 2: Genotyped SNPs by study site

SNP ID	Alleles	Chromosome	Nearest gene (location)	Previous association (PubMed ID)	Data available by study site				
					ARIC	CHS	EAGLE	SiMES	SP2
rs1061170	C/T	1	*CFH*	AMD (15761120; 15761122; 16849663; 22705344; 21665990; 20385826)	X	-	X	-	X
rs203674	G/T	1	*CFH*	AMD (18541031; 18043728)	-	-	-	-	X
rs3753394	T/C	1	*CFH*	AMD (22035603; 21111031; 18421087; 17167412)	-	X	-	X	X
rs3753396	A/G	1	*CFH*	AMD (18421087; 20157618; 18162041)	-	-	-	-	X
rs3766404	C/T	1	*CFH*	AMD (18043728; 15870199)	-	-	-	X	X
rs529825	A/G	1	*CFH*	AMD (21882633; 18043728)	-	-	-	-	X
rs6677604	A/G	1	*CFH*	AMD (23103884; 16998489; 18162041)	-	X	-	X	X
rs800292	A/G	1	*CFH*	AMD (23289807; 23260260; 23233260; 22618592; 22065928; 21909106)	-	X	-	X	X
rs6754295	T/G	2	*APOB*	Lipids (19060911; 19951432)	X	X	-	X	X
rs11726949	T/C	4	*CFI*	AMD	-	-	-	X	X

				(19603066; 18685559)					
rs10033900	T/C	4	*CFI*	AMD (22705344; 21665990; 20385826; 20087399; 23326517)	-	X	-	X	X
rs547154	T/G	6	*C2*	AMD (16936732; 16518403; 20157352; 19259132; 18493315)	-	-	-	X	X
rs12678919	A/G	8	*LPL*	AMD (22815349; 20385819) Lipids (20686565; 19060906)	-	-	X	X	X
rs10503669	A/C	8	*LPL*	Lipids (18193043; 21909109; 20370913)	X	-	X	-	-
rs2083637	T/C	8	*LPL*	Lipids (19060911; 18454146; 20160193)	X	X	-	X	X
rs2197089	T/C	8	*LPL*	Lipids (23344322; 21738485)	-	X	X	-	X
rs328	C/G	8	*LPL*	Lipids (22171074; 18193044; 17463246; 21738485; 21840003; 21316679; 20150529)	X	X	X	-	-
rs6586891	A/C	8	*LPL*	Lipids (21738485)	-	X	X	-	-
rs6987702	T/C	8	*TRIB1*	Lipids (19060911)	X	X	-	X	X
rs1883025	A/G	9	*ABCA1*	AMD (20385826; 20385819) Lipids	X	-	X	-	X

				(20686565; 19060906)					
rs1323432	T/C	9	*GRIN3A*	Lipids (18193043)	-	-	X	-	-
rs3890182	A/G	9	*ABCA1*	Lipids (20864672; 18193044; 21738485)	-	X	X	-	-
rs3905000	A/G	9	*ABCA1*	Lipids (19060911; 21347282)	X	X	X	X	X
rs4149268	A/G	9	*ABCA1*	Lipids (18193043; 19802338)	X	-	X	X	X
rs4149274	A/G	9	*ABCA1*	Lipids (18193043)	-	-	-	X	X
rs471364	A/G	9	*TTC39B*	Lipids (19060906)	-	X	X	-	X
rs10490924	T/G	10	*ARMS2*	AMD (22705344; 22694956; 21665990; 20861866; 20385826; 16080115)	X	-	X	-	-
rs964184	C/G	11	*ZNF259/APOA1*	AMD (20385819) Lipids (19060906; 20657596; 20686565; 20864672; 22359512)	-	-	X	X	X
rs174547	T/C	11	*FADS1*	AMD (20385826) Lipids (19060906; 21829377; 20972250; 20364269; 21738485)	-	-	X	-	X
rs28927680	C/G	11	*BUD13*	Lipids (18193044; 23634756; 21738485; 20018036)	X	X	X	-	-
rs3135506	C/G	11	*APOA5*	Lipids	X	X	X	-	-

				(18596051; 19018513; 20395964; 20429872; 21738485)					
rs7120118	T/C	11	*NR1H3*	Lipids (19060910)	X	X	-	X	X
rs7395662	A/G	11	*MADD*	Lipids (19060911)	X	X	-	X	X
rs2338104	C/G	12	*KCTD10/MVK*	Lipids (19060906; 18193043; 19060910; 21738485)	-	X	X	-	-
rs10468017	T/C	15	*LIPC*	AMD (20385826; 21665990; 23348725) Lipids (19060906; 21943158; 22359512)	X	-	X	X	X
rs1800588	T/C	15	*LIPC*	Lipids (18193044; 21347282; 21149302; 19802338)	X	-	X	-	-
rs261332	A/G	15	*LIPC*	Lipids (21738485; 21283740; 18193043)	X	-	-	-	X
rs757200	T/C	16	*CACNG3*	AMD (21169531)	-	X	-	X	X
rs3764261	T/G	16	*CETP*	AMD (20385819; 21665990; 20385826) Lipids (21738485; 18193043; 19060910; 19359809; 20686565; 23118302)	X	X	X	X	X
rs9989419	A/G	16	*CETP*	Lipids (18193043; 20031538;	X	X	X	X	X

				20864672; 21738485)					
rs1864163	A/G	16	*CETP*	Lipids (18193043; 20031564)	X	-	-	X	X
rs711752	A/G	16	*CETP*	Lipids (23675527; 20570915; 18193046)	X	-	-	-	-
rs7205804	A/G	16	*CETP*	Lipids (23675527; 18660489)	-	-	-	-	X
rs1566439	A/G	16	*CETP*	Lipids (18078817; 21738485; 18193043)	X	X	X	X	X
rs1800775	A/C	16	*CETP*	Lipids (17463246; 18193044; 20031564; 23675527; 20370913; 19197348)	X	-	X	X	X
rs1800777	A/G	16	*CETP*	Lipids (18660489; 20031564)	X	-	-	-	-
rs255049	T/C	16	*DPEP3*	Lipids (19060910)	X	-	-	X	X
rs12596776	C/G	16	*SLC12A3*	Lipids (21738485; 18193043)	X	-	X	-	-
rs2217332	T/C	16	*HERPUD1*	Lipids (20694148)	X	X	-	X	X
rs2271293	A/G	16	*NUTF2*	Lipids (19060911; 19060906; 21738485)	-	X	X	X	X
rs9891572	T/C	17	*METTL16*	Lipids (19060910)	X	X	-	X	X
rs2156552	A/T	18	*LIPG*	Lipids (18193043; 18193044; 20864672; 19148283; 19060910; 21738485)	-	X	X	-	X
rs4939883	T/C	18	*LIPG*	Lipids	X	X	X	X	X

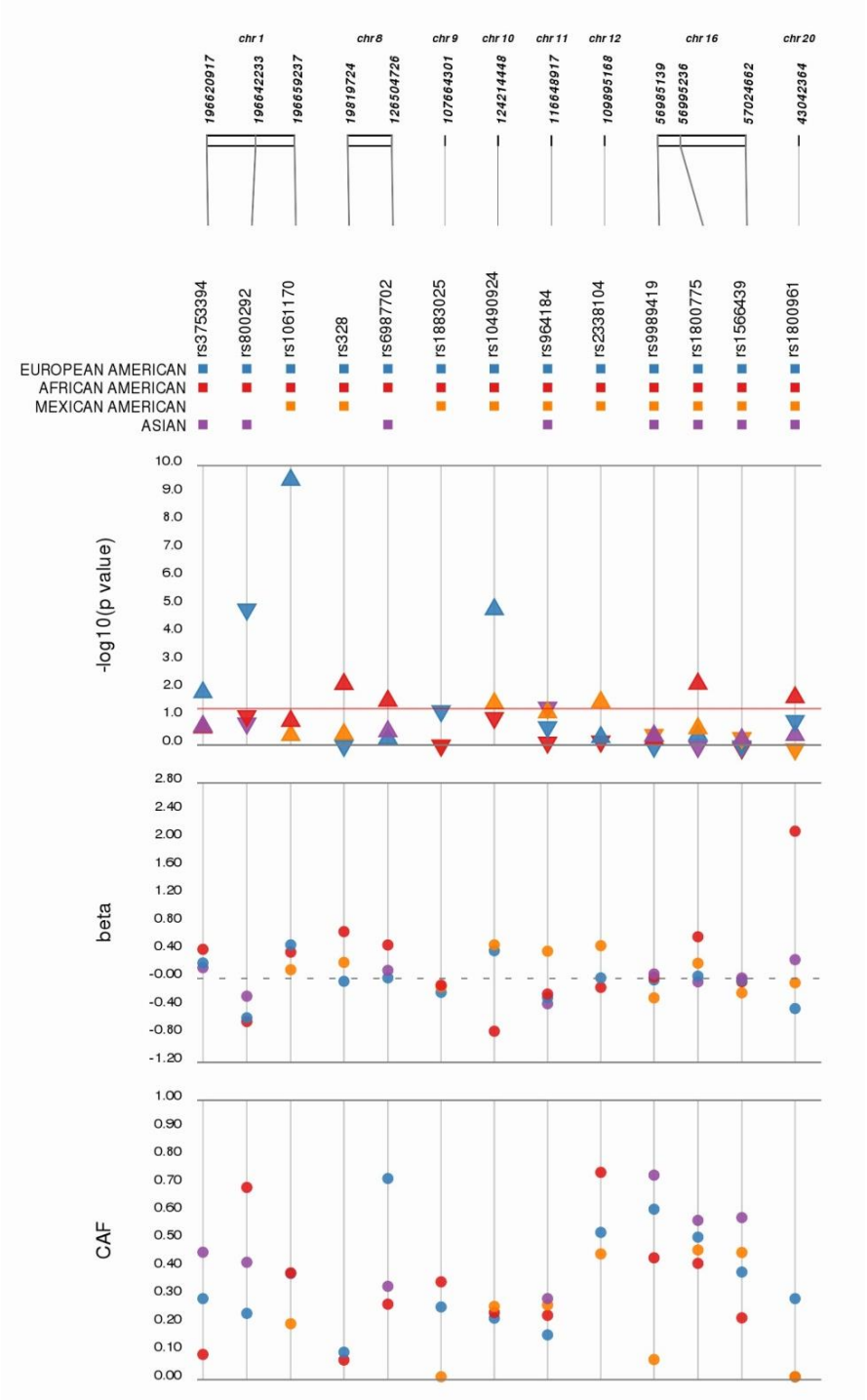
				(19060906; 19060911; 19802338; 20031538)					
rs2304130	A/G	19	*ZNF101*	Lipids (19060911; 20864672; 22359512)	-	X	-	X	X
rs2967605	A/G	19	*RAB11B*	Lipids (19060906; 21738485)	-	X	X	X	X
rs1800961	T/C	20	*HNF4A*	Lipids (19060906; 20686565; 21738485)	-	X	X	X	X
rs7679	T/C	20	*PCIF1*	Lipids (19060906; 21738485)	-	-	X	X	X

Supplementary Table 3. Model 3 significant AMD meta-analysis association results**. Each study site performed tests of association using logistic regression assuming an additive genetic model. Data were meta-analyzed using a fixed-effects inverse-variance weighted approach. Results are shown for nominally significant tests (p<0.05) adjusted for age, sex, body mass index, smoking status (ever/never), and HDL cholesterol (mg/dL) regardless of fasting status (Model 3).**

SNP ID	Gene	Chr	OR (95% CI)	Direction of Effect[a]	P-value	Coded Allele	CAF	Race/Ethnicity
rs1061170	*CFH*	1	1.62 (1.39-1.88)	+,.,+,.	4.9×10^{-10}	C	0.37	European American
			1.46 (0.80-2.69)	~~A~~	0.21		0.38	African American
			1.12 (0.48-1.60)	~~A~~	0.67		0.20	Mexican American
rs800292	*CFH*	1	0.57 (0.45-0.72)	~~A~~	9.2×10^{-6}	A	0.24	European American
			0.54 (0.29-1.02)	~~A~~	0.06		0.69	African American
			0.78 (0.57-1.06)	-,-,-,+	0.11		0.42	Asian
rs3753394	*CFH*	1	1.25 (1.02-1.51)	~~A~~	0.02	T	0.29	European American
			1.52 (0.61-3.74)	~~A~~	0.36		0.08	African American
			1.17 (0.85-1.59)	+,+,+,-	0.32		0.54	Asian
rs328	*LPL*	8	0.96 (0.79-1.16)	-,-,+,.	0.70	G	0.10	European American
			1.96 (1.17-3.25)	+,+,+,.	0.01		0.07	African American
			1.26 (0.50-3.16)	~~A~~	0.62		<0.01	Mexican American
rs6987702	*TRIB1*	8	1.01 (0.84-1.15)	-,+,.,.	0.89	T	0.73	European American
			1.62 (1.02-2.53)	+,+,.,.	0.04		0.29	African American
			1.12 (0.89-1.42)	-,+,+,+	0.49		0.43	Asian
rs1883025	*ABCA1*	9	0.82 (0.68-0.98)	-,.,-,.	0.04	A	0.26	European Americans
			0.91 (0.57-1.44)	-,.,-,.	0.68		0.35	African Americans
			0.89 (0.53-1.50)	~~A~~	0.68		<0.01	Mexican American
rs10490924	*ARMS2*	10	1.49 (1.25-1.77)	+,.,+,.	2.1×10^{-5}	T	0.22	European American
			0.47 (0.20-1.07)	~~A~~	0.07		0.24	African American
			1.62 (1.00-2.62)	~~A~~	0.05		0.26	Mexican American
rs964184	*ZNF259/APOA1*	11	0.76 (0.53-1.09)	~~A~~	0.15	G	0.84	European American
			0.80 (0.40-1.59)	~~A~~	0.53		0.77	African American
			1.47 (0.92-2.37)	~~A~~	0.10		0.27	Mexican American
			1.43 (1.03-1.98)	-,-,.,+	0.03		0.79	Asian

rs2338104	*KCTD10*	12	1.01 (0.88-1.16)	·,-,+,·	0.80	G	0.53	European American
			0.88 (0.61-1.25)	-,-,+,·	0.49		0.74	African American
			1.60 (1.01-2.54)	~~A~~	0.04		0.45	Mexican American
rs1800775	*CETP*	16	1.03 (0.88-1.21)	-,·,+,·	0.66	C	0.52	European American
			1.82 (1.16-2.85)	+,·,+,·	$9.8x10^{-3}$		0.42	African American
			1.24 (0.75-2.05)	~~A~~	0.38		0.46	Mexican American
			0.95 (0.70-1.30)	-,+,-,+	0.76		0.51	Asian
rs1800961	*HNF4A*	20	0.65 (0.39-1.06)	·,-,-,·	0.09	T	0.03	European American
			8.24 (1.23-55.2)	·,-,+,·	0.03		0.01	African American
			0.94 (.026-3.38)	~~A~~	0.93		<0.01	Mexican American
			1.31 (0.36-4.77)	~~A~~	0.68		0.02	Asian

~~ƨ~~ Direction of effect is given for ARIC, CHS, and EAGLE for EA, AA, and MA if data are available. Direction of effect is given for Asians for SiMES and SP2 1M, 550, and 610 platforms. Otherwise, the study site is set to missing (".").
~~A~~ Only a single study site is represented.



Supplementary Figure 1: Synthesis view plot of nominally significant (p < 0.05) meta-analysis association results for Model 3 which is adjusted for site of ascertainment, age, sex, BMI, smoking status, and HDL cholesterol regardless of fasting status. Included are the results of SNPs that were nominally significant in Model 2, but not in Model 3, for ease of comparison. The two models differed in HDL status (regardless of fasting vs. fasting) and use of Asian cohort (SiMES only vs. SiMES/SP2). P- values are represented by the colored arrows and are transformed by the –log10, with the threshold of p = 0.05 marked by the red line. Colored arrows also show the direction of effect (beta). P-values, beta's, and coded allele frequencies (CAF) are plotted by race/ethnicity.

Supplementary Table 4: Mitochondrial genetic association results for model adjusted by age, sex, BMI, and smoking status.

SNPID	Gene	OR	lower CI	upper CI	p-value	CA	CAF (%)	Race
mt10115	*ND3*	8.33	1.11	62.57	0.0393	C	0.004	Non-Hispanic White
		1.01	0.40	2.55	0.9917	C	0.298	Non-Hispanic Black
		—	—	—	0.9821	C	0.034	Mexican American
mt1018	*MTRNR1*	8.37	1.11	62.82	0.039	A	0.005	Non-Hispanic White
		0.72	0.31	1.68	0.451	A	0.515	Non-Hispanic Black
		—	—	—	0.9811	A	0.038	Mexican American
mt10400	*ND3*	0.44	0.03	5.66	0.5289	C	0.994	Non-Hispanic White
		0.34	0.03	3.45	0.3621	C	0.989	Non-Hispanic Black
		1.50	0.61	3.71	0.3748	C	0.731	Mexican American
mt10550	*NDL4*	1.14	0.60	2.15	0.6934	A	0.906	Non-Hispanic White
		—	—	—	0.9926	A	—	Non-Hispanic Black
		—	—	—	0.983	A	0.970	Mexican American
mt11177	*ND4*	1.01	0.41	2.50	0.9837	C	0.818	Mexican American
mt11251	*ND4*	0.68	0.44	1.04	0.0719	A	0.804	Non-Hispanic White
		—	—	—	0.9898	A	—	Non-Hispanic Black
		—	—	—	0.9916	A	0.985	Mexican American
mt11719	*ND4*	1.14	0.79	1.64	0.4887	A	0.534	Non-Hispanic White
		0.33	0.09	1.31	0.1158	A	0.949	Non-Hispanic Black
		1.03	0.22	4.86	0.9671	A	0.936	Mexican American
mt11947	*ND4*	0.51	0.18	1.46	0.2112	A	0.981	Non-Hispanic White
		—	—	—	0.9936	A	—	Non-Hispanic Black
		—	—	—	0.9909	A	—	Mexican American
mt12007	*ND4*	1.18	0.31	4.50	0.8104	A	0.017	Non-Hispanic White
		0.91	0.11	7.50	0.9312	A	0.062	Non-Hispanic Black
		2.47	1.18	5.18	0.0169	A	0.347	Mexican American
mt12414	*ND5*	1.05	0.35	3.16	0.931	C	0.022	Non-Hispanic White
		—	—	—	0.9902	C	—	Non-Hispanic Black
		—	—	—	0.9875	C	0.032	Mexican American
mt12705	*ND5*	1.09	0.56	2.11	0.8042	C	0.915	Non-Hispanic White

		1.71	0.45	6.41	0.429	C	0.084	Non-Hispanic Black
		0.87	0.40	1.91	0.73	C	0.321	Mexican American
mt13263	*ND5*	0.54	0.05	6.37	0.6231	A	0.995	Non-Hispanic White
		—	—	—	0.9937	A	—	Non-Hispanic Black
		2.26	0.75	6.84	0.1496	A	0.784	Mexican American
mt13506	*ND5*	1.57	0.51	4.82	0.4291	C	0.789	Non-Hispanic Black
		—	—	—	0.9911	C	—	Mexican American
mt13789	*ND5*	0.54	0.15	1.95	0.3502	C	0.159	Non-Hispanic Black
mt14178	*CYTB*	—	—	—	0.9784	C	—	Non-Hispanic White
		0.57	0.16	2.03	0.3849	C	0.158	Non-Hispanic Black
mt14318	*CYTB*	2.98	0.18	50.27	0.449	C	0.004	Non-Hispanic White
		0.46	0.15	1.39	0.1671	C	0.210	Mexican American
mt1438	*MTRNR1*	0.61	0.18	2.11	0.434	A	0.025	Non-Hispanic White
		1.19	0.32	4.48	0.7932	A	0.084	Non-Hispanic Black
		—	—	—	0.9911	A	—	Mexican American
mt14470	*ND6*	0.63	0.14	2.81	0.5447	C	0.023	Non-Hispanic White
		—	—	—	0.9929	C	—	Non-Hispanic Black
mt14560	*CYTB*	—	—	—	0.9856	A	—	Non-Hispanic White
		0.56	0.16	1.98	0.3645	A	0.157	Non-Hispanic Black
mt14668	*CYTB*	0.12	0.01	2.18	0.1522	C	—	Non-Hispanic White
		—	—	—	0.9899	C	—	Non-Hispanic Black
		0.38	0.09	1.62	0.1905	C	0.955	Mexican American
mt14766	*CYTB*	0.86	0.60	1.24	0.4241	C	0.472	Non-Hispanic White
		3.27	0.83	12.98	0.0914	C	0.051	Non-Hispanic Black
		0.82	0.18	3.81	0.8039	C	0.068	Mexican American
mt14905	*CYTB*	2.49	1.17	5.32	0.0185	A	0.104	Non-Hispanic White
		2.16	0.22	21.39	0.5114	A	0.121	Non-Hispanic Black
mt15043	*CYTB*	0.91	0.33	2.51	0.8549	A	0.036	Non-Hispanic White
		1.50	0.17	13.54	0.7158	A	0.018	Non-Hispanic Black
		0.68	0.27	1.68	0.3975	A	0.271	Mexican American

mt15326	*CYTB*	0.35	0.04	2.85	0.3277	A	0.010	Non-Hispanic White
		—	—	—	0.9911	A	—	Mexican American
mt15452	*CYTB*	1.47	0.96	2.24	0.0778	A	0.197	Non-Hispanic White
		—	—	—	0.9898	A	—	Non-Hispanic Black
		—	—	—	0.9874	A	0.015	Mexican American
mt15535	*CYTB*	1.01	0.41	2.50	0.9837	C	0.818	Mexican American
mt16111	*HVCR*	—	—	—	0.9822	A	0.009	Non-Hispanic White
		—	—	—	0.9846	A	0.018	Non-Hispanic Black
		2.90	1.38	6.11	0.0052	A	0.309	Mexican American
mt16189	*HVCR*	0.90	0.52	1.56	0.7108	C	0.148	Non-Hispanic White
		0.58	0.22	1.53	0.2723	C	0.356	Non-Hispanic Black
		0.86	0.38	1.95	0.7116	C	0.260	Mexican American
mt16271	*HVCR*	—	—	—	0.9852	C	—	Non-Hispanic White
		—	—	—	0.993	C	—	Non-Hispanic Black
mt16319	*HVCR*	0.42	0.05	3.52	0.4266	A	0.014	Non-Hispanic White
		3.23	0.33	31.60	0.3135	A	0.016	Non-Hispanic Black
		2.59	1.24	5.42	0.0117	A	0.348	Mexican American
mt16362	*HVCR*	1.70	0.93	3.10	0.084	C	0.085	Non-Hispanic White
		2.29	0.84	6.27	0.106	C	0.130	Non-Hispanic Black
		2.80	1.32	5.95	0.0074	C	0.424	Mexican American
mt16390	*Non-coding*	1.15	0.22	5.94	0.8646	A	0.016	Non-Hispanic White
		1.79	0.28	11.31	0.5385	A	—	Non-Hispanic Black
		—	—	—	0.9807	A	0.054	Mexican American
mt1736	*HVCR*	—	—	—	0.9857	A	0.997	Non-Hispanic White
		—	—	—	0.9906	A	—	Non-Hispanic Black
		0.40	0.19	0.83	0.0147	A	0.657	Mexican American
mt2092	*MTRNR2*	—	—	—	0.987	C	—	Non-Hispanic White
		—	—	—	0.9931	C	—	Non-Hispanic Black
		0.49	0.12	1.94	0.3079	C	0.949	Mexican American
mt3505	*ND1*	0.88	0.29	2.68	0.8243	A	0.978	Non-Hispanic White
		0.47	0.05	4.40	0.5058	A	0.985	Non-Hispanic Black

		—	—	—	0.9911	A	—	Mexican American
mt3552	*ND1*	—	—	—	0.9862	A	—	Non-Hispanic White
		—	—	—	0.991	A	—	Mexican American
mt3594	*ND1*	0.12	0.02	0.89	0.0376	C	0.996	Non-Hispanic White
		1.38	0.60	3.19	0.4466	C	0.490	Non-Hispanic Black
		—	—	—	0.9808	C	0.962	Mexican American
mt4769	*ND2*	1.06	0.22	5.15	0.9423	A	0.010	Non-Hispanic White
		—	—	—	0.9887	A	—	Non-Hispanic Black
		—	—	—	0.9911	A	—	Mexican American
mt4883	*ND2*	—	—	—	0.987	C	—	Non-Hispanic White
		—	—	—	0.993	C	—	Non-Hispanic Black
		0.54	0.14	2.11	0.3769	C	0.947	Mexican American
mt4977	*ND2*	—	—	—	0.9931	C	—	Non-Hispanic Black
		0.83	0.32	2.16	0.7001	C	0.180	Mexican American
mt5178	*ND2*	—	—	—	0.987	A	—	Non-Hispanic White
		—	—	—	0.993	A	—	Non-Hispanic Black
		1.85	0.47	7.20	0.3769	A	0.053	Mexican American
mt5442	*ND2*	0.92	0.11	7.53	0.936	C	0.060	Non-Hispanic Black
		—	—	—	0.991	C	—	Mexican American
mt6371	*CO1*	1.42	0.31	6.44	0.6476	C	0.980	Non-Hispanic White
mt825	*MTRNR1*	0.62	0.20	1.90	0.4004	A	0.214	Non-Hispanic Black
		—	—	—	0.991	A	—	Mexican American
mt8414	*ATP8*	—	—	—	0.987	C	—	Non-Hispanic White
		—	—	—	0.9931	C	—	Non-Hispanic Black
		0.54	0.14	2.11	0.3769	C	0.947	Mexican American
mt8468	*ATP8*	—	—	—	0.9869	C	—	Non-Hispanic White
		—	—	—	0.9599	C	0.793	Non-Hispanic Black
		—	—	—	0.9853	C	—	Mexican American
mt8703	*ATP6*	—	—	—	0.9857	C	—	Non-Hispanic White
		—	—	—	0.9895	C	—	Non-Hispanic Black

mt9042	*ATP6*	1.08	0.13	8.97	0.9423	C	0.940	Non-Hispanic Black
		—	—	—	0.991	C	—	Mexican American
mt9347	*COX3*	1.02	0.12	8.42	0.9889	A	0.943	Non-Hispanic Black
		—	—	—	0.9909	A	—	Mexican American
mt9950	*COX3*	—	—	—	0.9843	C	—	Non-Hispanic White
		0.92	0.11	7.53	0.9388	C	0.046	Non-Hispanic Black
		0.81	0.31	2.11	0.665	C	0.181	Mexican American

"—" denotes genetic association tests with uninterruptable results due to very few case counts

Supplementary Table 5: Complete results for association testing between AMD variants and quantitative traits.

SNP	CA	CAF	Trait	race_ethnicity	Beta	SE	CI lower	CI upper	p-value
rs1061170	C	0.37	BMI	Non-Hispanic Black	-1.20	0.51	-2.21	-0.19	0.02
rs547154	T	0.08	BMI	Mexican American	-1.18	0.61	-2.37	0.02	0.05
rs10490924	T	0.22	Lutein	Non-Hispanic White	-1.12	0.63	-2.36	0.13	0.08
rs547154	T	0.09	Lutein	Mexican American	3.11	1.83	-0.48	6.71	0.09
rs547154	T	0.11	LN Vitamin A	Non-Hispanic White	0.03	0.02	-0.01	0.07	0.09
rs547154	T	0.11	Lutein	Non-Hispanic White	1.19	0.86	-0.50	2.88	0.17
rs1061170	C	0.37	LN Vitamin A	Non-Hispanic Black	0.04	0.03	-0.02	0.10	0.17
rs547154	T	0.21	β-carotene	Non-Hispanic Black	-4.42	3.33	-10.97	2.13	0.19
rs1061170	C	0.36	LN α-carotene	Non-Hispanic White	0.04	0.03	-0.03	0.10	0.27
rs547154	T	0.08	LN α-carotene	Mexican American	0.09	0.08	-0.08	0.26	0.28
rs1061170	C	0.36	LN Vitamin A	Non-Hispanic White	0.01	0.01	-0.01	0.04	0.30
rs10490924	T	0.25	BMI	Non-Hispanic Black	-0.59	0.57	-1.72	0.54	0.30
rs547154	T	0.21	LN α-carotene	Non-Hispanic Black	-0.08	0.09	-0.25	0.09	0.35
rs547154	T	0.21	Lutein	Non-Hispanic Black	-1.41	1.54	-4.44	1.63	0.36
rs10490924	T	0.25	LN α-carotene	Non-Hispanic Black	0.08	0.08	-0.09	0.24	0.37
rs10490924	T	0.25	LN α-carotene	Non-Hispanic Black	0.07	0.08	-0.09	0.23	0.38
rs10490924	T	0.26	LN Vitamin A	Mexican American	0.02	0.02	-0.03	0.07	0.45
rs1061170	C	0.36	β-carotene	Non-Hispanic White	-0.87	1.18	-3.17	1.44	0.46
rs1061170	C	0.21	BMI	Mexican American	0.34	0.46	-0.58	1.25	0.47
rs547154	T	0.11	β-carotene	Non-Hispanic White	1.35	1.89	-2.36	5.07	0.48
rs10490924	T	0.25	LN Vitamin A	Non-Hispanic Black	0.02	0.03	-0.04	0.09	0.48
rs10490924	T	0.26	BMI	Mexican American	0.28	0.40	-0.50	1.06	0.48
rs1061170	C	0.36	Lutein	Non-Hispanic White	0.37	0.57	-0.76	1.49	0.52
rs1061170	C	0.37	LN α-carotene	Non-Hispanic Black	0.04	0.07	-0.10	0.19	0.55
rs1061170	C	0.36	BMI	Non-Hispanic White	-0.12	0.23	-0.56	0.32	0.60
rs1061170	C	0.21	β-carotene	Mexican American	-0.80	1.73	-4.21	2.60	0.64
rs1061170	C	0.37	β-carotene	Non-Hispanic Black	-1.32	2.87	-6.98	4.33	0.65
rs10490924	T	0.22	β-carotene	Non-Hispanic White	0.62	1.36	-2.05	3.28	0.65

rs10490924	T	0.26	Lutein	Mexican American	-0.51	1.19	-2.85	1.83	0.67
rs547154	T	0.11	LN α-carotene	Non-Hispanic White	0.02	0.05	-0.09	0.13	0.70
rs10490924	T	0.25	β-carotene	Non-Hispanic Black	-1.07	3.21	-7.38	5.24	0.74
rs10490924	T	0.22	LN Vitamin A	Non-Hispanic White	0.00	0.01	-0.02	0.03	0.75
rs1061170	C	0.21	LN α-carotene	Mexican American	0.02	0.06	-0.10	0.14	0.75
rs547154	T	0.08	LN Vitamin A	Mexican American	-0.01	0.04	-0.09	0.06	0.75
rs547154	T	0.21	BMI	Non-Hispanic Black	0.18	0.59	-0.99	1.34	0.77
rs10490924	T	0.22	BMI	Non-Hispanic White	0.08	0.26	-0.43	0.58	0.77
rs10490924	T	0.22	LN α-carotene	Non-Hispanic White	-0.01	0.04	-0.09	0.07	0.78
rs10490924	T	0.26	LN α-carotene	Mexican American	0.02	0.06	-0.09	0.12	0.78
rs10490924	T	0.26	LN α-carotene	Mexican American	0.01	0.06	-0.10	0.12	0.80
rs10490924	T	0.26	β-carotene	Mexican American	0.32	1.55	-2.72	3.36	0.84
rs10490924	T	0.25	Lutein	Non-Hispanic Black	-0.29	1.48	-3.21	2.63	0.84
rs1061170	C	0.21	LN Vitamin A	Mexican American	0.01	0.03	-0.05	0.06	0.86
rs1061170	C	0.37	Lutein	Non-Hispanic Black	-0.20	1.28	-2.72	2.33	0.88
rs547154	T	0.21	LN Vitamin A	Non-Hispanic Black	0.00	0.03	-0.06	0.07	0.90
rs10490924	T	0.22	LN α-carotene	Non-Hispanic White	0.00	0.04	-0.08	0.07	0.92
rs547154	T	0.11	BMI	Non-Hispanic White	0.02	0.36	-0.68	0.73	0.95
rs1061170	C	0.21	Lutein	Mexican American	-0.05	1.43	-2.86	2.76	0.97
rs547154	T	0.08	β-carotene	Mexican American	-0.04	2.40	-4.76	4.68	0.99

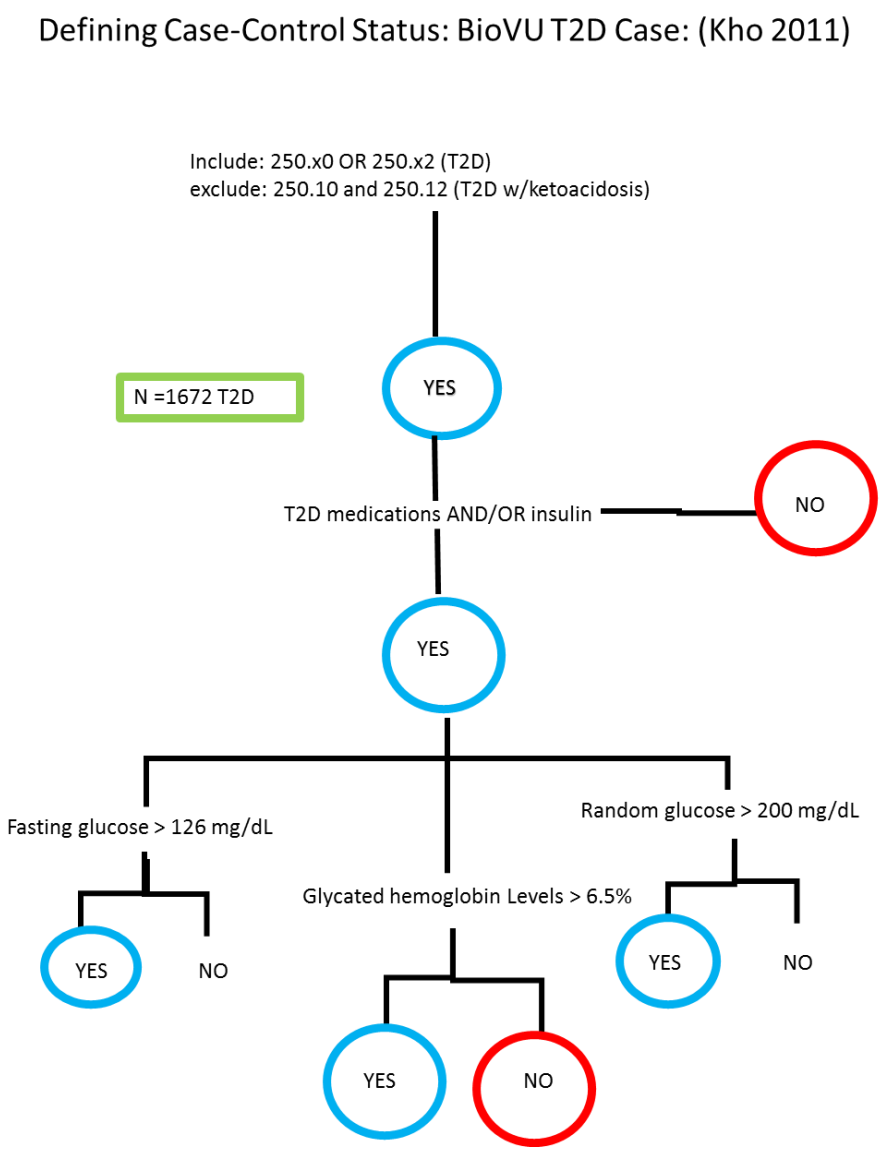
CA = coded allele
CAF = coded allele frequency
CI = confidence interval
SE = standard error
LN = natural log

Chapter 3 Appendix

Supplementary Table 6: List of diabetic retinopathy/diabetes medications

Drug name(s)	Generic drug name/active ingredient	Indications*
Avastin	Bevacizumab	Metastatic colorectal cancer, non-squamous non-small cell lung cancer, glioblastoma, metastatic renal cell carcinoma, and cervical cancer
Ozurdex (implant)	Dexamethasone	Macular edema following branch retinal vein occlusion or central retinal vein occlusion. Non-infectious uveitis and diabetic macular edema
Retisert (implant)	Fluocinolone acetonide	Non-infectious uveitis affecting posterior segment of eye
Iluvien (implant)	Fluocinolone acetonide	Diabetic macular edema
Zestril	Lisinopril	Hypertension, heart failure, acute myocardial infarction
Prinivil	Lisinopril	Hypertension, heart failure, acute myocardial infarction
Lucentis	Ranibizumab	Neovascular age-related macular degeneration, macular edema following retinal vein occlusion, diabetic macular edema

Indications is a limited list of FDA approved uses as stated on the Drugs@FDA website as of January 25th, 2014: http://www.accessdata.fda.gov/scripts/cder/drugsatfda/



Supplementary Figure 2: Flow diagram of the decision tree involved in determining Type-2 diabetes in EAGLE BioVU African Americans (n=1,672).

Supplementary Table 7: Study population characteristics of total POAG cases and controls among African Americans in EAGLE BioVU

	Total Cases > 20 yrs (SD)	Definite Cases	Potential Cases	Controls > 40 yrs (SD)
N	268	138	67	4813
Age (years)	71 (12.9)	--		57 (11.7)
Age at Diagnosis (years)	62 (12.5)	62.0 (12.0)	62.9 (12.7)	--
Age at Last Clinic (years)	--	--	--	54 (11.7)
Sex (% female)	61.6	63.7	65.6	60
Hypertensive (%)	48.3	55.1	59.7	46.6
BMI (kg/m^2)	30.1 (7.0)	30.1 (6.7)	29.0 (6.1)	30.1 (8.0)
Diastolic (mm/Hg)	76 (8.9)	74.5 (8.1)	77.7 (10.3)	80 (33.6)
Systolic (mm/Hg)	135 (14.2)	134.5 (14.1)	134 (14.7)	124 (26.2)
Cholesterol (mg/dL)	182 (48.2)	183 (40.6)	187 (58.0)	161 (65.2)
HDL (mg/dL)	52 (25.5)	52.5 (25.0)	58 (29.7)	53 (38.6)
LDL (mg/dL)	99 (49.3)	103 (42.9)	94.5 (45.7)	99 (50.7)
Triglycerides (mg/dL)	112 (70.8)	125 (76.3)	107 (65.8)	98 (67.8)

Median values were calculated for the following: Age was calculated from a given birth year. Age at POAG diagnosis was determined by the date of when POAG ICD-9 (365.11) was first mentioned in the records. Age at last clinic visit (LCV) was taken as the date of the last CPT mentioned in the records for controls. An individual was classified as hypertensive if he/she met one of three criteria: systolic blood pressure > 140 mm/Hg, diastolic blood pressure > 90 mm/Hg, or on hypertension medications all within a two year window of when they were diagnosed with POAG in cases and a two year window of their LCV date for controls. Blood pressure (systolic and diastolic), lipids (total cholesterol, high-density cholesterol, low-density cholesterol, and triglycerides), and body mass index (height and weight) were calculated from labs or measurements within two years of POAG diagnosis or LCV. Abbreviations: standard deviation (SD)

Chapter 4 Appendix

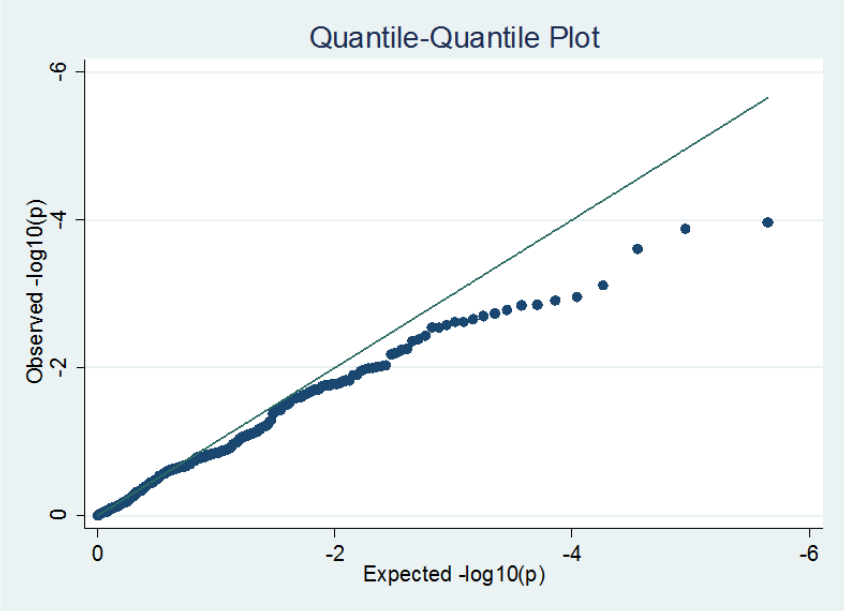
Supplementary Table 8: Ocular-related SNPs directly genotyped on the Metabochip

SNP	CHR	Closest Gene	Trait	PMID
rs994767	1	*ZC3H11B*	Ocular axial length	TBD
rs11755724	6	*RREB1*	AMD	20385826
rs2070600	6	*AGER*	Diabetic retinopathy	21067572
rs730497	7	*GCK*	HbA1c	19096518
rs10237118	7	*TRIM24*	Optic disc size (cup)	20395239
rs13266634	8	*EIF3H*	HbA1c	19734900
rs564398	9	*CDKN2B*	POAG	22428042
rs523096	9	*CDKN2B*	NT-Glaucoma	22792221
rs3217992	9	*CDKN2B*	NPG	22570617
rs2157719	9	*CDKN2B*	POAG	22570617
rs1412829	9	*CDKN2B*	NPG	22570617
rs1063192	9	*CDKN2B*	POAG	22419738
rs1063192	9	*CDKN2B*	POAG and CTD ratio	22419738 22570617
rs7894966	10	GAS7	IOP	22570627
rs7072268	10	*HK1*	HbA1c	19096518
rs3858145	10	*ATOH7*	Optic disc size (cup)	20395239
rs3793917	10	*ARMS2*	AMD	20385819
rs10490924	10	*ARMS2*	AMD	20385826
rs11812882	10	*CISD1*	Diabetic retinopathy	20871662
rs5742629	12	*IGF1*	Myopia (extreme)	22509095
rs10858945	12	*LOC338758*	Optic disc size (cup)	20395239
rs10483727	14	*AKR1B1P5*	POAG and Optic disc size (rim)	20548946 22419738 22570617
rs493258	15	*RPL28P4*	AMD	20385819
rs10468017	15	*LIPC*	AMD	20385826
rs3764261	16	*CETP*	AMD	20385819
rs10521145	16	*CCDC101*	Diabetic retinopathy	22427569
rs151227	16	*NUPR1*	Diabetic retinopathy	21441570
rs151230	16	*CCDC101*	Diabetic retinopathy	21441570
rs151229	16	*CCDC101*	Diabetic retinopathy	21441570
rs10521145	16	*CCDC101*	Diabetic retinopathy	21441570
rs11641853	16	*CCDC101*	Diabetic retinopathy	21441570
rs11074904	16	*SULT1A1*	Diabetic retinopathy	21441570
rs1109739	16	*16q12*	POAG	22661486
rs134173	22	*CHEK2*	CTD ratio	22570617

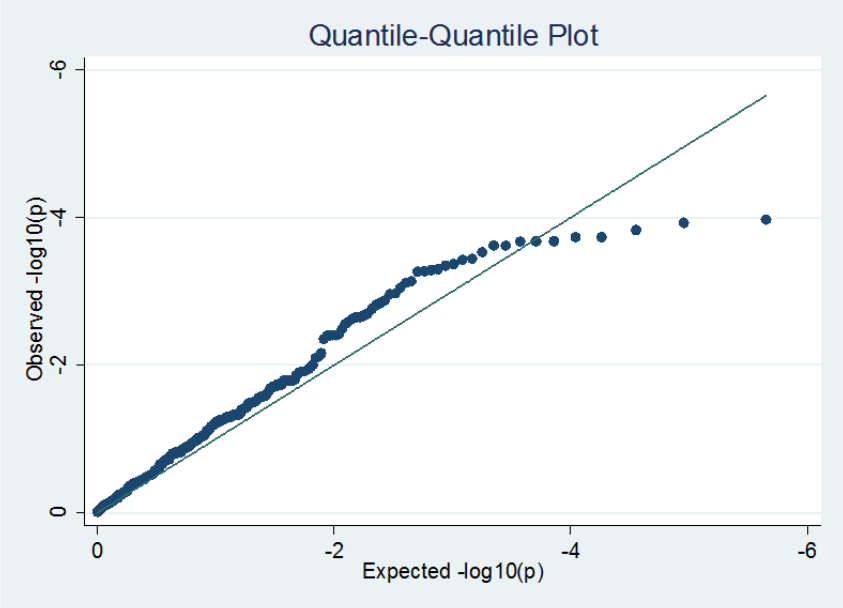
Supplementary Table 9: Twenty Most significant results for genetic association analysis of CDKN2B-AS1 region in African Americans cases (n=138) and controls (n=1,376). Logistic regression assuming an additive genetic model was performed for adjusted for age, sex, and PC.

SNP	Allele	MAF	Function Class	OR	95% CI	p-value
rs2065504	C	0.37	upstream intergenic	1.42	1.06-1.89	0.02
rs6475610	G	0.37	upstream intergenic	1.41	1.05-1.88	0.02
rs111690485	A	0.05	upstream intergenic	1.93	1.08-3.44	0.03
rs10217426	C	0.47	Intron	1.36	1.01-1.82	0.04
rs2383206	G	0.41	Intron	0.74	0.55-1.00	0.05
rs10965234	A	0.46	Intron	1.34	0.99-1.80	0.05
rs1537376	G	0.41	Intron	0.74	0.55-1.01	0.06
rs10965235	A	0.47	Intron	1.33	0.99-1.79	0.06
rs77728904	C	0.09	Intron	0.56	0.30-1.03	0.06
rs80166549	G	0.10	Intron	0.57	0.32-1.04	0.06
rs2383208	G	0.18	upstream intergenic	1.41	0.98-2.03	0.07
rs80202680	A	0.06	upstream intergenic	1.66	0.96-2.89	0.07
rs79985856	A	0.10	Intron	0.58	0.32-1.05	0.07
rs79182326	A	0.10	Intron	0.58	0.32-1.05	0.07
rs944797	G	0.41	Intron	0.76	0.56-1.03	0.08
rs77284052	A	0.10	Intron	0.59	0.33-1.06	0.08
rs77920300	A	0.10	Intron	0.59	0.33-1.06	0.08
rs17694493	G	0.11	Intron	0.62	0.36-1.07	0.09
rs17694572	A	0.10	Intron	0.61	0.35-1.08	0.09
rs2065505	G	0.25	upstream intergenic	1.32	0.95-1.82	0.09

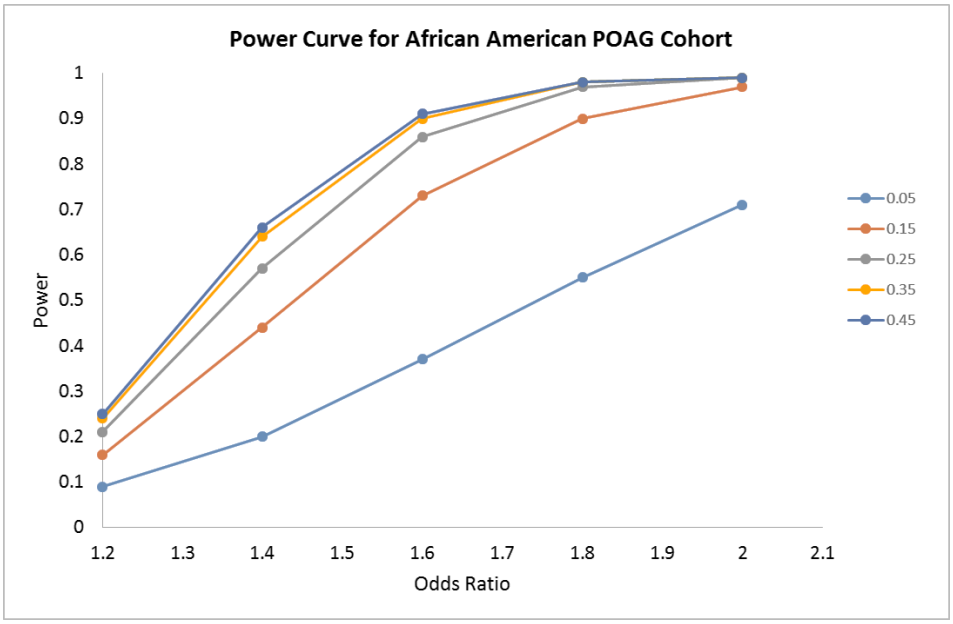
MAF = minor allele frequency
OR = odds ratio
CI = confidence interval



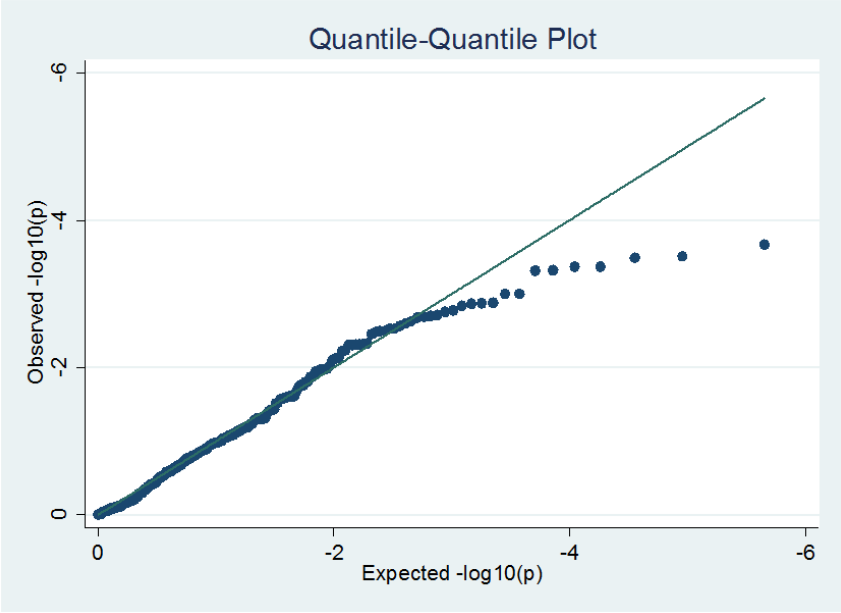
Supplementary Figure 4: Quantile-Quantile plots graphed in STATA 12.0 of p-values in the *CDKN2B-AS1* association analyses of logistic regression model adjusted by age, sex, and first three



Supplementary Figure 3: Quantile-Quantile plots graphed in STATA 12.0 of p-values in the *CDKN2B-AS1* association analyses of logistic regression model adjusted by age, sex, and median diastolic blood pressure.



Supplementary Figure 5: Power Calculations determined in Quanto for the EAGLE BioVU African American POAG cohort assuming 135 cases with a case: control ratio of 1:3, log-additive model, and a 2-sided t-test. The color coded lines represented the varying allele frequencies tested for in these power calculations.



Supplementary Figure 6: Quantile-Quantile plots graphed in STATA 12.0 of p-values in the *CDKN2B-AS1* association analyses of logistic regression model adjusted by age, sex, first 3 PC, and median diastolic blood pressure.

Supplementary Table 10: The one hundred most significant results for the genetic association analysis of the Metabochip and African Americans POAG cases (n=138) and controls (n=1,376). Logistic regression assuming an additive genetic model was performed for adjusted for age, sex, and PC.

CHR	SNP	Coded Allele	OR	95% CI-L	95% CI-U	P-value
3	rs4678836	A	1.96	1.42	2.70	4.31E-05
19	rs1671152	A	1.88	1.39	2.54	4.53E-05
21	rs9982695	A	2.01	1.44	2.82	4.59E-05
6	rs9479660	G	1.88	1.39	2.55	4.72E-05
6	rs11155927	G	1.93	1.41	2.66	5.28E-05
6	rs10456759	C	1.91	1.39	2.62	7.01E-05
6	rs9479657	G	1.81	1.35	2.44	9.21E-05
6	rs9479726	A	0.44	0.29	0.67	1.01E-04
4	rs3775202	G	1.75	1.31	2.35	1.82E-04
1	rs1572151	G	2.51	1.55	4.07	2.00E-04
6	chr6:25793471	C	1.75	1.30	2.37	2.71E-04
19	rs2910368	G	2.13	1.42	3.21	2.81E-04
16	chr16:52509162	A	1.99	1.37	2.89	3.04E-04
16	rs7190904	A	2.11	1.41	3.15	3.10E-04
15	rs11639241	C	1.76	1.29	2.39	3.21E-04
1	chr1:11778484	C	2.44	1.50	3.99	3.64E-04
1	rs17421462	A	2.44	1.50	3.99	3.64E-04
5	rs7714384	A	1.71	1.27	2.31	3.89E-04
13	rs449674	A	0.43	0.27	0.69	4.01E-04
6	chr6:25814153	A	1.84	1.31	2.57	4.07E-04
9	rs7863513	A	0.55	0.40	0.77	4.19E-04
19	chr19:19633722	A	0.51	0.35	0.74	4.39E-04
13	chr13:109680753	A	2.15	1.40	3.30	4.53E-04
5	rs4336354	G	2.21	1.42	3.45	4.64E-04
6	rs7454156	G	1.84	1.31	2.59	4.81E-04
17	rs9675320	A	0.58	0.43	0.79	4.82E-04
4	rs17028407	A	2.45	1.48	4.06	4.88E-04
16	chr16:52507417	A	1.85	1.31	2.62	4.89E-04
4	chr4:88283455	G	1.73	1.27	2.36	5.12E-04
19	rs2967732	C	1.90	1.32	2.74	5.15E-04
8	rs2158588	A	2.55	1.50	4.34	5.31E-04
16	chr16:52496582	A	1.79	1.29	2.49	5.44E-04
21	rs2832227	G	1.66	1.24	2.21	6.00E-04
13	rs17591848	G	1.83	1.29	2.60	6.51E-04
2	rs1370919	C	1.67	1.24	2.24	6.56E-04
6	chr6:25818066	G	0.59	0.44	0.80	6.75E-04

12	rs10860451	G	0.58	0.42	0.79	7.02E-04
11	chr11:72150910	G	1.67	1.24	2.25	7.04E-04
13	rs1322379	G	1.79	1.28	2.50	7.23E-04
6	chr6:25522825	A	0.58	0.42	0.80	7.46E-04
4	chr4:88242323	G	1.76	1.27	2.45	7.59E-04
6	chr6:25833213	A	0.60	0.44	0.81	7.68E-04
5	rs17066506	A	1.90	1.31	2.76	7.74E-04
16	rs4786689	A	0.56	0.40	0.79	7.74E-04
1	rs12119433	C	1.68	1.24	2.28	7.88E-04
22	rs9608416	A	1.80	1.28	2.55	8.19E-04
16	chr16:28662185	A	1.70	1.25	2.33	8.24E-04
19	chr19:19449565	A	2.22	1.39	3.56	8.87E-04
18	rs1786162	A	0.59	0.44	0.81	9.05E-04
16	chr16:28769235	C	1.66	1.23	2.23	9.22E-04
18	rs1786153	A	1.65	1.23	2.21	9.61E-04
16	chr16:28290744	A	1.70	1.24	2.33	9.77E-04
6	chr6:160868393	A	1.67	1.23	2.26	9.88E-04
6	rs6455482	G	1.77	1.26	2.48	9.90E-04
6	rs7774579	A	1.68	1.23	2.28	9.92E-04
2	chr2:21284826	G	2.29	1.40	3.76	1.03E-03
2	chr2:21292690	G	2.29	1.40	3.76	1.03E-03
19	chr19:19345008	G	1.96	1.31	2.94	1.05E-03
16	rs2008514	A	1.64	1.22	2.20	1.05E-03
16	chr16:28733454	G	1.64	1.22	2.20	1.08E-03
1	chr1:109506916	G	1.84	1.28	2.65	1.10E-03
5	rs1529707	A	0.60	0.44	0.81	1.10E-03
4	chr4:88259933	A	1.68	1.23	2.29	1.11E-03
2	rs3097385	G	1.85	1.28	2.67	1.12E-03
5	rs4700135	G	0.58	0.42	0.80	1.13E-03
19	chr19:19468633	G	2.18	1.36	3.50	1.15E-03
3	rs11916892	G	0.30	0.14	0.62	1.15E-03
2	chr2:21259122	A	2.29	1.39	3.76	1.16E-03
5	rs159584	A	2.46	1.43	4.23	1.19E-03
16	chr16:52499785	G	1.74	1.25	2.44	1.19E-03
17	rs11653150	A	1.63	1.21	2.19	1.20E-03
6	rs4713691	A	0.55	0.39	0.79	1.23E-03
16	chr16:28780899	A	1.64	1.21	2.21	1.25E-03
16	chr16:28772543	G	1.64	1.21	2.21	1.26E-03
6	rs11964613	A	0.55	0.38	0.79	1.27E-03
16	chr16:28779361	G	1.64	1.21	2.21	1.28E-03
2	rs11127229	A	1.90	1.29	2.82	1.29E-03
16	chr16:28782623	C	1.64	1.21	2.21	1.29E-03
2	chr2:21250393	G	2.26	1.37	3.73	1.33E-03

16	chr16:28775305	A	1.63	1.21	2.20	1.33E-03
2	rs6759676	G	0.61	0.45	0.82	1.33E-03
16	chr16:28770952	A	1.63	1.21	2.20	1.33E-03
5	rs9885411	G	2.53	1.43	4.46	1.34E-03
16	chr16:52499369	A	1.75	1.24	2.47	1.37E-03
16	chr16:28798966	A	1.63	1.21	2.20	1.37E-03
7	rs2041562	C	2.43	1.41	4.18	1.38E-03
1	chr1:62672716	A	2.05	1.32	3.17	1.38E-03
16	chr16:28790742	G	1.63	1.21	2.19	1.39E-03
7	rs17837626	C	2.09	1.33	3.28	1.39E-03
11	rs17197116	G	0.26	0.11	0.59	1.39E-03
16	rs3888190	A	1.63	1.21	2.19	1.40E-03
3	chr3:15794966	C	2.48	1.42	4.32	1.40E-03
5	rs10063054	C	1.60	1.20	2.14	1.41E-03
16	chr16:28788703	C	1.63	1.21	2.19	1.43E-03
16	chr16:28797632	A	1.63	1.21	2.19	1.43E-03
5	rs974359	A	1.65	1.21	2.25	1.44E-03
16	chr16:28776196	A	1.63	1.21	2.19	1.44E-03
16	chr16:28797209	C	1.63	1.21	2.19	1.46E-03
6	rs2524118	A	2.10	1.33	3.32	1.46E-03
6	chr6:25818742	C	1.63	1.21	2.19	1.47E-03

OR = odds ratio
CI-L = lower confidence interval
CI-U = upper confidence interval

Supplementary Table 11: The one hundred most significant results for the genetic association analysis of the Metabochip and African Americans POAG cases (n=138) and controls (n=1,376). Logistic regression assuming an additive genetic model was performed for adjusted for age, sex, PC, and median diastolic blood pressure.

CHR	SNP	Coded Allele	OR	95% CI-L	95% CI-U	P-value
1	chr1:228347779	A	2.373	1.563	3.603	5.00E-05
1	chr1:228354829	C	2.094	1.447	3.028	8.73E-05
21	rs9982695	A	2.09	1.446	3.02	8.74E-05
6	chr6:25793471	C	1.943	1.392	2.714	9.64E-05
4	rs3775202	G	1.921	1.383	2.666	9.70E-05
2	rs13423742	C	3.048	1.73	5.369	0.000115
6	rs7454156	G	2.078	1.427	3.026	0.000138
6	rs9479726	A	0.405	0.2536	0.6469	0.000155
19	rs1671152	A	1.916	1.367	2.686	0.000161
10	rs286489	A	1.901	1.358	2.66	0.00018
5	rs4336354	G	2.511	1.549	4.07	0.000187
16	rs7190904	A	2.293	1.483	3.546	0.000192
16	chr16:52507417	A	2.022	1.395	2.931	0.000202
5	rs7714384	A	1.858	1.336	2.584	0.000232
7	chr7:14708236	G	0.3397	0.1908	0.6047	0.000243
16	rs1424077	G	2.259	1.461	3.492	0.000246
6	rs6929849	A	0.5075	0.3532	0.7293	0.000246
4	rs17028407	A	2.735	1.596	4.688	0.000252
4	chr4:88283455	G	1.876	1.338	2.63	0.000266
19	rs2910368	G	2.299	1.468	3.602	0.000276
4	chr4:88259933	A	1.875	1.336	2.631	0.000279
13	rs1547918	G	2.302	1.468	3.609	0.00028
1	rs649214	G	1.848	1.326	2.577	0.000291
13	chr13:109680753	A	2.368	1.482	3.784	0.000314
16	chr16:52509162	A	2.085	1.398	3.111	0.000317
5	rs17066506	A	2.118	1.408	3.187	0.000319
4	rs6832117	A	1.809	1.306	2.505	0.00036
8	rs10956525	G	3.112	1.667	5.813	0.000368
1	chr1:109506916	G	2.062	1.384	3.072	0.000377
4	rs1408	G	1.854	1.317	2.611	0.000409
8	rs6415517	A	2.693	1.554	4.664	0.000411
4	chr4:88242323	G	1.915	1.335	2.747	0.000415
5	chr5:157782818	A	2.403	1.477	3.91	0.000418
3	rs4647226	A	3.01	1.629	5.559	0.000433
2	rs11127229	A	2.118	1.394	3.218	0.000439
4	chr4:88259209	A	1.85	1.313	2.607	0.00044

2	rs3768641	G	2.022	1.362	3	0.000472
13	rs449674	A	0.3867	0.227	0.6587	0.000473
22	rs9608416	A	2.001	1.355	2.953	0.000483
4	chr4:88272109	G	1.834	1.304	2.579	0.000491
6	rs241407	A	2.215	1.416	3.465	0.000494
8	rs2158588	A	2.755	1.552	4.891	0.000536
19	chr19:19449565	A	2.478	1.478	4.154	0.000579
15	rs11639241	C	1.815	1.29	2.553	0.000614
7	chr7:14720840	A	0.3619	0.2021	0.6479	0.000625
5	rs1529707	A	0.5403	0.3796	0.769	0.00063
13	rs1322379	G	1.902	1.315	2.75	0.000634
11	chr11:10305291	G	2.244	1.411	3.568	0.000635
2	rs1357011	C	2.641	1.51	4.617	0.000658
19	rs2967732	C	1.99	1.339	2.959	0.000669
5	rs10063054	C	1.758	1.27	2.434	0.00068
1	rs2422286	A	1.93	1.32	2.82	0.000688
6	rs12215670	G	2.639	1.505	4.628	0.000706
5	rs9885411	G	2.933	1.573	5.472	0.000716
7	rs4720833	A	0.4503	0.2836	0.715	0.000721
3	rs4678836	A	1.872	1.301	2.693	0.00073
5	chr5:157674339	A	1.777	1.273	2.481	0.000737
6	chr6:25818742	C	1.773	1.271	2.474	0.000748
6	rs3869129	A	0.5399	0.377	0.7733	0.000773
19	chr19:19468633	G	2.423	1.446	4.061	0.00078
2	rs3097385	G	2.002	1.335	3.003	0.000792
16	rs237174	A	1.925	1.313	2.822	0.000793
6	rs10456759	C	1.83	1.284	2.607	0.000823
10	rs11196187	A	2.694	1.506	4.819	0.000838
5	chr5:157668311	G	1.757	1.261	2.448	0.000874
4	rs223482	A	1.748	1.258	2.429	0.000886
13	rs3924002	G	2.137	1.365	3.346	0.000897
11	chr11:2876684	A	1.994	1.327	2.998	0.000902
17	rs2659015	A	2.046	1.341	3.124	0.000908
14	rs10135856	A	0.5781	0.4181	0.7993	0.000917
5	chr5:157767459	A	0.5537	0.3903	0.7857	0.000928
13	rs17591848	G	1.908	1.302	2.797	0.000929
5	chr5:157760095	G	0.5543	0.3907	0.7864	0.000945
5	chr5:157766571	G	0.5543	0.3907	0.7864	0.000945
5	chr5:157758859	A	0.5545	0.3909	0.7866	0.000948
15	chr15:73056964	A	2.521	1.457	4.362	0.00095
5	chr5:157758831	A	0.5554	0.3917	0.7877	0.000972
12	chr12:48529939	G	1.748	1.254	2.436	0.000975
9	rs7863513	A	0.54	0.3744	0.7789	0.000977

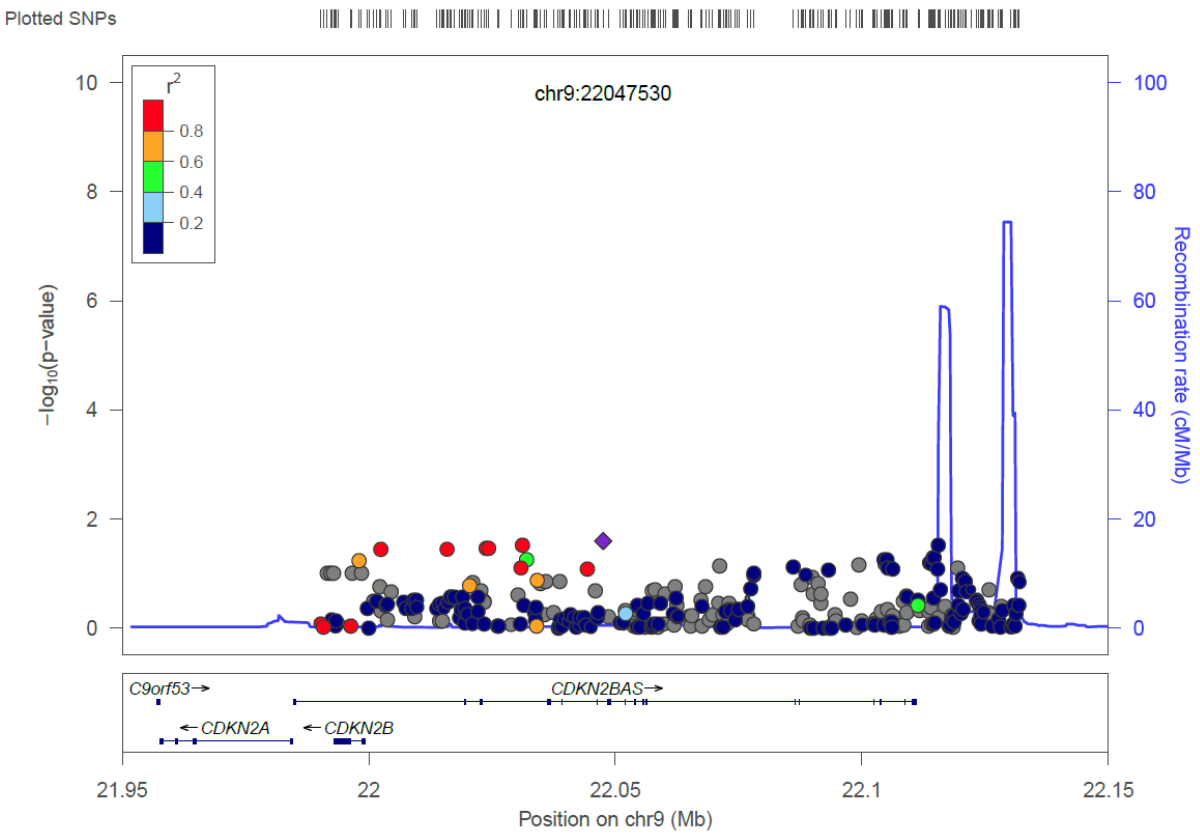
10	rs4347309	G	1.792	1.266	2.536	0.000989
6	chr6:25818066	G	0.5682	0.4058	0.7957	0.001001
8	rs11863	A	1.767	1.258	2.48	0.001009
5	rs4700135	G	0.5332	0.3661	0.7765	0.001043
10	rs853928	A	0.4842	0.3138	0.7472	0.001049
5	chr5:157736208	A	0.5515	0.386	0.7879	0.001077
5	chr5:157759964	A	0.5573	0.3925	0.7913	0.001079
1	rs11810369	A	1.726	1.244	2.395	0.001081
6	rs9479660	G	1.752	1.252	2.453	0.001087
18	rs1786153	A	1.723	1.243	2.389	0.001103
17	rs16956560	G	1.754	1.251	2.457	0.001104
5	chr5:157764878	G	0.5579	0.3928	0.7924	0.001115
6	rs3201892	C	0.5271	0.3585	0.7749	0.001126
2	rs13394146	A	1.738	1.246	2.425	0.001141
2	rs4849816	A	1.734	1.245	2.417	0.001145
6	rs6455482	G	1.851	1.277	2.682	0.001149
2	rs12613548	A	0.4604	0.2884	0.7351	0.001158
1	rs2039988	A	2.045	1.327	3.15	0.001176
12	chr12:48488687	A	1.76	1.251	2.476	0.001182
6	chr6:25833213	A	0.5745	0.4109	0.8031	0.001185
8	rs11784268	C	1.871	1.281	2.734	0.001203

OR = odds ratio
CI-L = lower confidence interval
CI-U = upper confidence interval

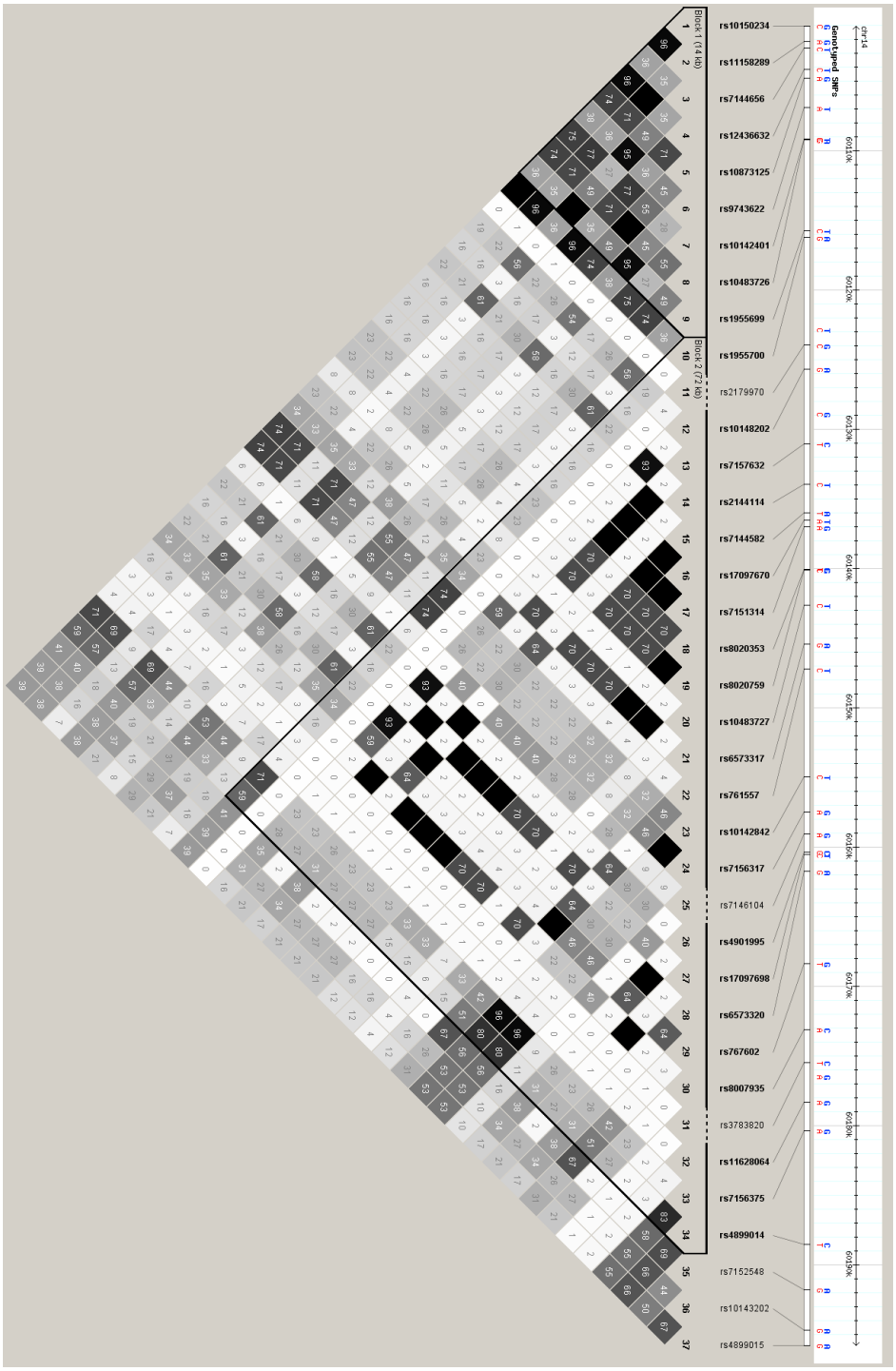
Supplementary Table 12: The ten most significant results for the genetic association analysis of the Metabochip and African Americans POAG cases (n=138) and controls (n=1,376). Logistic regression assuming an additive genetic model was performed for adjusted for age, sex, and PC.

CHR	SNP	Gene	Coded Allele	MAF	OR	95% CI	p-value
3	rs4678836	*ARPP21-STAC*	A	0.25	1.95	1.41-2.70	4.31×10^{-5}
19	rs1671152	*GP6*	A	0.32	1.87	1.38-2.53	4.53×10^{-5}
21	rs9982695	*C21orf33*	A	0.24	2.01	1.43-2.81	4.59×10^{-5}
6	rs9479660	*RGS17-OPRM1*	G	0.26	1.88	1.38-2.54	4.72×10^{-5}
6	rs11155927	*RGS17-OPRM1*	G	0.21	1.93	1.40-2.66	5.28×10^{-5}
6	rs10456759	*LINC01108-intergenic*	C	0.26	1.90	1.38-2.62	7.01×10^{-5}
6	rs9479657	*RGS17-OPRM1*	G	0.31	1.81	1.34-2.44	9.21×10^{-5}
6	rs9479726	*RGS17-OPRM1*	A	0.24	0.44	0.29-0.66	1.01×10^{-4}
4	rs3775202	*VEGFC*	G	0.43	1.75	1.30-2.34	1.80×10^{-4}
1	rs1572151	*MTHFR*	G	0.05	2.50	1.54-4.07	2.00×10^{-4}

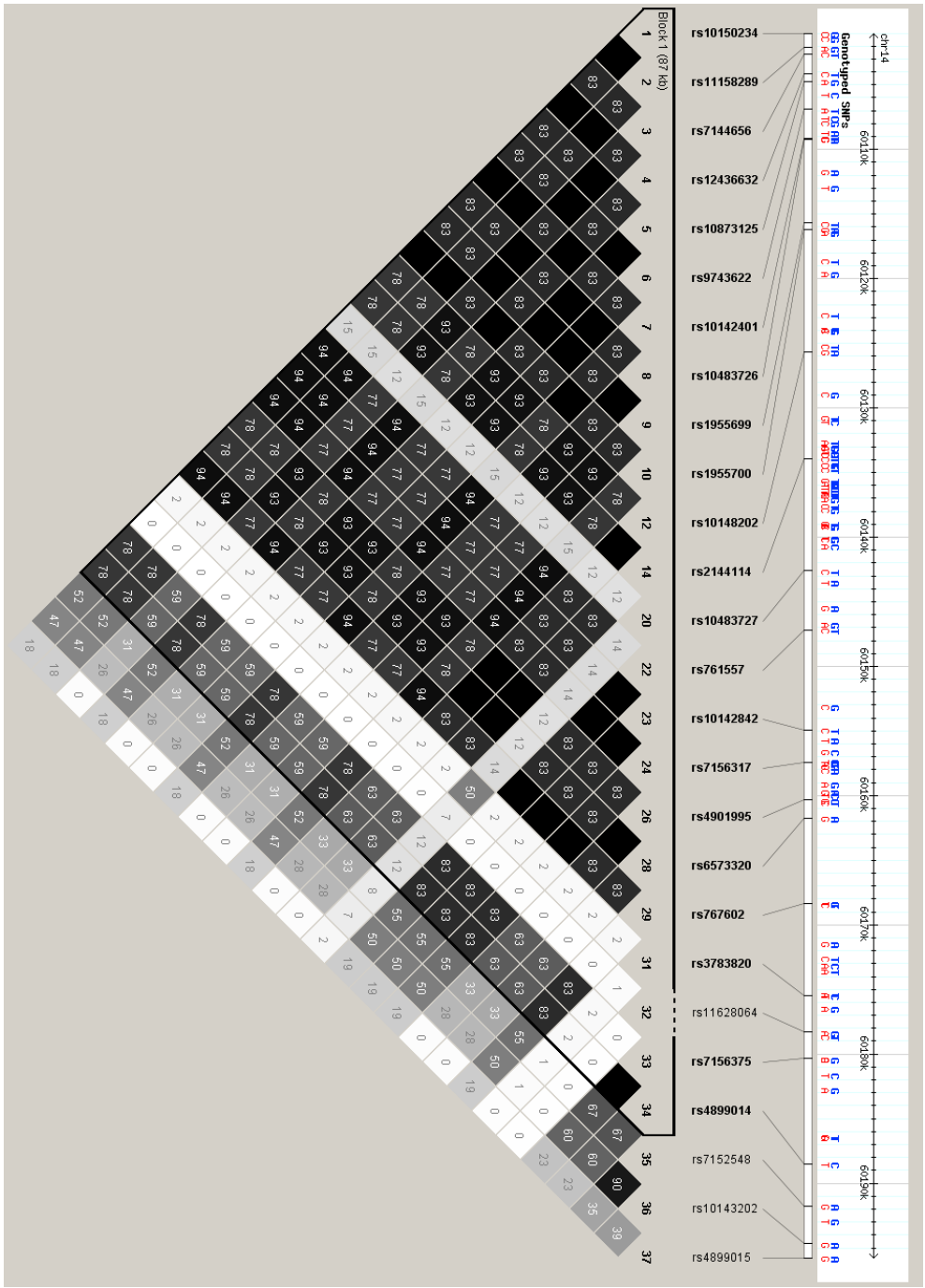
MAF = minor allele frequency
OR = odds ratio
CI = confidence interval



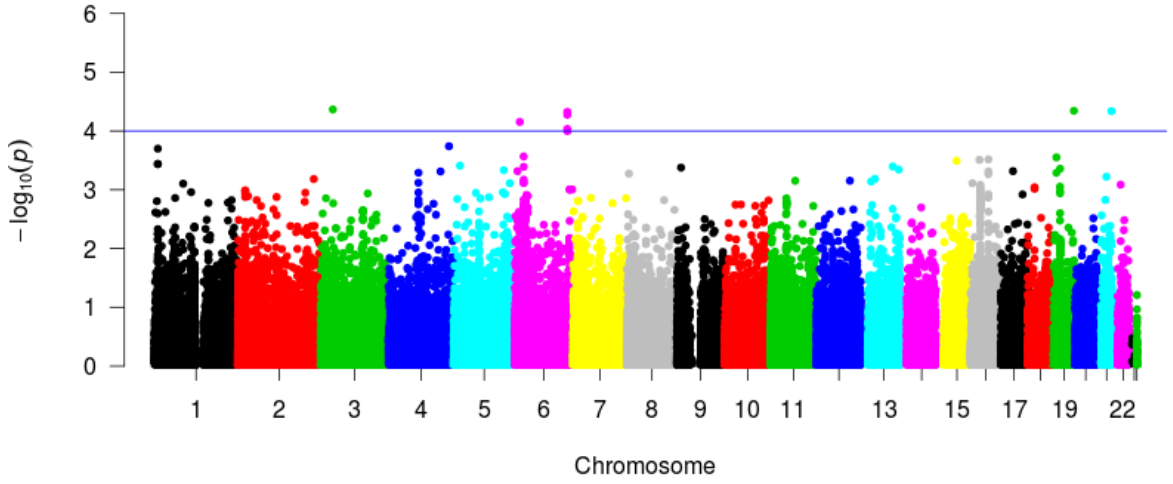
Supplementary Figure 7: Locus Zoom regional association plot for POAG in African Americans for *CDKN2B-AS1*. Vertical axis is $-\log_{10}$ of the p-value, the horizontal axis is the chromosomal position. Each dot represents a SNP tested for association with POAG in 138 cases and 1,376 controls. Approximate linkage disequilibrium between the most significant SNP and the other SNPs in the plot is shown by the r^2 legend with LD calculations from 1000 Genomes CEU.



Supplementary Figure 8: Haploview (v. 4.2) LD plot of the Chromosome 14 *SIX6* (rs10483727) region in HapMap III ASW population. Plot generated for an approximately 100kb window around rs10483727. Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$.



Supplementary Figure 9: Haploview (v. 4.2) LD plot of the Chromosome 14 *SIX6* (rs10483727) region in HapMap III CEU population. Plot generated for an approximately 100kb window around rs10483727. Displayed are SNP rs numbers listed 5` to 3`. Tiled blocks represent pairwise tests of LD. The linkage score measured by the correlation coefficient (r^2) of each pairwise comparison is displayed in the block with white blocks representing $r^2 = 0$, shades of gray representing $0 < r^2 < 1$, and black representing $r^2 = 1$.



Supplementary Figure 10: Manhattan plot of EAGLE BioVU African American POAG genetic association results. Logistic regression assuming an additive genetic model was performed for 138 cases and 1,376 controls adjusted by age, sex, and principal components. P-values ($-\log_{10}$) on the y-axis) for each test of association are plotted by chromosome (x-axis). The blue line depicts a suggestive significance threshold of $p = 10^{-4}$.

Chapter 5 Appendix

Supplementary Table 13: Twenty most significant results for diabetic retinopathy African American Metabochip genetic association. Logistic regression assuming an additive genetic model was performed for 119 cases and 434 controls adjusted by age and sex.

CHR	SNP	Gene	Allele	OR	CI_L	CI_U	p-value
8	rs1487170	SNTB1-HAS2	G	2.71	1.66	4.42	6.37E-05
4	rs11732574	CYTL1-STK32B	G	2.23	1.50	3.32	7.35E-05
12	rs12309053	SPATS2	A	0.53	0.38	0.73	1.27E-04
4	rs12506426	LDB2	A	0.47	0.31	0.69	1.47E-04
11	rs1991320	PARVA	G	1.85	1.34	2.55	1.66E-04
1	rs1329817	POU3F1-RRAGC	A	1.83	1.33	2.53	2.24E-04
5	rs7723568	IRX1-LINCO1020	C	0.55	0.40	0.76	2.45E-04
2	chr2:28459238	-	G	1.82	1.32	2.51	2.52E-04
7	rs7790518	C1GALT1	C	2.23	1.45	3.44	2.69E-04
4	rs228618	MANBA	A	1.77	1.30	2.41	2.74E-04
10	rs2483769	CXCL12-TMEM72-AS1	G	1.94	1.36	2.79	2.92E-04
3	chr3:38767092	-	A	1.83	1.32	2.54	2.94E-04
8	rs7831823	SLC26A7	A	2.80	1.60	4.90	3.09E-04
3	rs6792467	ITGA9	A	0.41	0.26	0.67	3.22E-04
3	rs3792426	SLITRK3	G	0.52	0.36	0.74	3.49E-04
3	rs9846696	SGOL1-ZNF385D	C	2.54	1.52	4.24	3.83E-04
2	rs7591633	LINC01122	A	1.78	1.29	2.44	3.89E-04
9	rs7873748	DENND4C	A	1.99	1.36	2.91	4.12E-04
9	rs2815175	CNTLN	G	1.92	1.34	2.77	4.13E-04
6	rs6935299	CASC6-EPHA7	A	1.89	1.33	2.69	4.21E-04

REFERENCES

Ach, T., Huisingh, C., McGwin, G., Messinger, J.D., Zhang, T., Bentley, M.J., Gutierrez, D.B., Ablonczy, Z., Smith, R.T., Sloan, K.R., et al. (2014). Quantitative autofluorescence and cell density maps of the human retinal pigment epithelium. Invest. Ophthalmol. Vis. Sci. *55*, 4832–4841.

Adam, M.F., Belmouden, A., Binisti, P., Brézin, A.P., Valtot, F., Béchetoille, A., Dascotte, J.C., Copin, B., Gomez, L., Chaventré, A., et al. (1997). Recurrent mutations in a single exon encoding the evolutionarily conserved olfactomedin-homology domain of TIGR in familial open-angle glaucoma. Hum. Mol. Genet. *6*, 2091–2097.

Age-Related Eye Disease Study 2 (AREDS2) Research Group, Chew, E.Y., SanGiovanni, J.P., Ferris, F.L., Wong, W.T., Agron, E., Clemons, T.E., Sperduto, R., Danis, R., Chandra, S.R., et al. (2013). Lutein/zeaxanthin for the treatment of age-related cataract: AREDS2 randomized trial report no. 4. JAMA Ophthalmol. *131*, 843–850.

Age-Related Eye Disease Study Research Group (1999). The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. Control. Clin. Trials *20*, 573–600.

Age-Related Eye Disease Study Research Group (2001). A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report no. 9. Arch. Ophthalmol. *119*, 1439–1452.

Age-Related Eye Disease Study Research Group, SanGiovanni, J.P., Chew, E.Y., Clemons, T.E., Ferris, F.L., Gensler, G., Lindblad, A.S., Milton, R.C., Seddon, J.M., and Sperduto, R.D. (2007). The relationship of dietary carotenoid and vitamin A, E, and C intake with age-related macular degeneration in a case-control study: AREDS Report No. 22. Arch. Ophthalmol. *125*, 1225–1232.

Ajit Bolar, N., Vanlander, A.V., Wilbrecht, C., Van der Aa, N., Smet, J., De Paepe, B., Vandeweyer, G., Kooy, F., Eyskens, F., De Latter, E., et al. (2013). Mutation of the iron-sulfur cluster assembly gene IBA57 causes severe myopathy and encephalopathy. Hum. Mol. Genet. *22*, 2590–2602.

Alward, W.L.M., Kwon, Y.H., Khanna, C.L., Johnson, A.T., Hayreh, S.S., Zimmerman, M.B., Narkiewicz, J., Andorf, J.L., Moore, P.A., Fingert, J.H., et al. (2002). Variations in the myocilin gene in patients with open-angle glaucoma. Arch. Ophthalmol. *120*, 1189–1197.

Anderson, D., and Patella, V. (1998). Automated Static Perimetry (St Louis, Mosby).

Anderson, N.B., Bulatao, R.A., Cohen, B., and National Research Council (US) Panel on Race, E. (2004). Race/Ethnicity, Socioeconomic Status, and Health.

Angi, M., Kalirai, H., Coupland, S.E., Damato, B.E., Semeraro, F., and Romano, M.R. (2012). Proteomic Analyses of the Vitreous Humour. Mediators Inflamm. *2012*.

Antonetti, D.A., Barber, A.J., Bronson, S.K., Freeman, W.M., Gardner, T.W., Jefferson, L.S., Kester, M., Kimball, S.R., Krady, J.K., LaNoue, K.F., et al. (2006). Diabetic Retinopathy Seeing Beyond Glucose-Induced Microvascular Disease. Diabetes *55*, 2401–2411.

Antonetti, D.A., Klein, R., and Gardner, T.W. (2012). Diabetic retinopathy. N. Engl. J. Med. *366*, 1227–1239.

Arakawa, S., Takahashi, A., Ashikawa, K., Hosono, N., Aoi, T., Yasuda, M., Oshima, Y., Yoshida, S., Enaida, H., Tsuchihashi, T., et al. (2011). Genome-wide association study identifies two susceptibility loci for exudative age-related macular degeneration in the Japanese population. Nat. Genet. *43*, 1001–1004.

Arar, N.H., Freedman, B.I., Adler, S.G., Iyengar, S.K., Chew, E.Y., Davis, M.D., Satko, S.G., Bowden, D.W., Duggirala, R., Elston, R.C., et al. (2008). Heritability of the severity of diabetic retinopathy: the FIND-Eye study. Invest. Ophthalmol. Vis. Sci. *49*, 3839–3845.

Arevalo, M.A., Santos-Galindo, M., Lagunas, N., Azcoitia, I., and Garcia-Segura, L.M. (2011). Selective estrogen receptor modulators as brain therapeutic agents. J. Mol. Endocrinol. *46*, R1–R9.

Arthur, S.N., Aldridge, A.J., De León-Ortega, J., McGwin, G., Xie, A., and Girkin, C.A. (2006). Agreement in assessing cup-to-disc ratio measurement among stereoscopic optic nerve head photographs, HRT II, and Stratus OCT. J. Glaucoma *15*, 183–189.

Asbell, P.A., Dualan, I., Mindel, J., Brocks, D., Ahmad, M., and Epstein, S. (2005). Age-related cataract. Lancet *365*, 599–609.

Asman, P., and Heijl, A. (1992). Glaucoma Hemifield Test. Automated visual field evaluation. Arch. Ophthalmol. *110*, 812–819.

Ayala-Haedo, J.A., Gallins, P.J., Whitehead, P.L., Schwartz, S.G., Kovach, J.L., Postel, E.A., Agarwal, A., Wang, G., Haines, J.L., Pericak-Vance, M.A., et al. (2010). Analysis of single nucleotide polymorphisms in the NOS2A gene and interaction with smoking in age-related macular degeneration. Ann. Hum. Genet. *74*, 195–201.

Balasubbu, S., Sundaresan, P., Rajendran, A., Ramasamy, K., Govindarajan, G., Perumalsamy, N., and Hejtmancik, J.F. (2010). Association analysis of nine candidate gene polymorphisms in Indian patients with type 2 diabetic retinopathy. BMC Med. Genet. *11*, 158.

Ballon, B.J., Echelman, D.A., Shields, M.B., and Ollie, A.R. (1992). Peripheral visual field testing in glaucoma by automated kinetic perimetry with the Humphrey Field Analyzer. Arch. Ophthalmol. *110*, 1730–1732.

Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinforma. Oxf. Engl. *21*, 263–265.

Bateson, W., Saunders, E.R., and Punnett, R.C. Report II. Experimental studies in the physiology of heredity.

Biesterbos, J.W.H., Dudzina, T., Delmaar, C.J.E., Bakker, M.I., Russel, F.G.M., von Goetz, N., Scheepers, P.T.J., and Roeleveld, N. (2013). Usage patterns of personal care products: important factors for exposure assessment. Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc. *55*, 8–17.

Biswas, S., Xia, S., and Lin, S. (2014). Detecting rare haplotype-environment interaction with logistic Bayesian LASSO. Genet. Epidemiol. *38*, 31–41.

Bonomi, L., Marchini, G., Marraffa, M., Bernardi, P., Morbio, R., and Varotto, A. (2000). Vascular risk factors for primary open angle glaucoma: the Egna-Neumarkt Study. Ophthalmology *107*, 1287–1293.

Braghini, C.A., Neshich, I.A.P., Neshich, G., Soardi, F.C., de Mello, M.P., Costa, V.P., de Vasconcellos, J.P.C., and de Melo, M.B. (2013). New mutation in the myocilin gene segregates with juvenile-onset open-angle glaucoma in a Brazilian family. Gene *523*, 50–57.

Brian, G., and Taylor, H. (2001). Cataract blindness--challenges for the 21st century. Bull. World Health Organ. *79*, 249–256.

Brown, D.R., and Topcu, M. (2003). Willingness to participate in clinical treatment research among older African Americans and Whites. The Gerontologist *43*, 62–72.

Buckingham, B.P., Inman, D.M., Lambert, W., Oglesby, E., Calkins, D.J., Steele, M.R., Vetter, M.L., Marsh-Armstrong, N., and Horner, P.J. (2008). Progressive ganglion cell degeneration precedes neuronal loss in a mouse model of glaucoma. J. Neurosci. Off. J. Soc. Neurosci. *28*, 2735–2744.

Budde, W.M. (2000). Heredity in primary open-angle glaucoma. Curr. Opin. Ophthalmol. *11*, 101–106.

Buentello-Volante, B., Rodriguez-Ruiz, G., Miranda-Duarte, A., Pompa-Mera, E.N., Graue-Wiechers, F., Bekker-Méndez, C., Ayala-Ramirez, R., Quezada, C., Rodríguez-Loaiza, J.L., and Zenteno, J.C. (2012). Susceptibility to advanced age-related macular degeneration and alleles of complement factor H, complement factor B, complement component 2, complement component 3, and age-related maculopathy susceptibility 2 genes in a Mexican population. Mol. Vis. *18*, 2518–2525.

Buggio, L., Pagliardini, L., Gentilini, D., De Braud, L., Viganò, P., and Vercellini, P. (2014). A rare familial case of endometriosis with very severe gynecological and obstetric complications: novel genetic variants at a glance. Gynecol. Obstet. Invest. *77*, 201–204.

Buitendijk, G.H.S., Rochtchina, E., Myers, C., van Duijn, C.M., Lee, K.E., Klein, B.E.K., Meuer, S.M., de Jong, P.T.V.M., Holliday, E.G., Tan, A.G., et al. (2013). Prediction of age-related macular degeneration in the general population: the Three Continent AMD Consortium. Ophthalmology *120*, 2644–2655.

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., et al. (2011a). Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. Nat. Genet. *43*, 574–578.

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., et al. (2011b). Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. Nat. Genet. *43*, 574–578.

Buyske, S., Wu, Y., Carty, C.L., Cheng, I., Assimes, T.L., Dumitrescu, L., Hindorff, L.A., Mitchell, S., Ambite, J.L., Boerwinkle, E., et al. (2012a). Evaluation of the metabochip genotyping array in African Americans and implications for fine mapping of GWAS-identified loci: the PAGE study. PloS One *7*, e35651.

Buyske, S., Wu, Y., Carty, C.L., Cheng, I., Assimes, T.L., Dumitrescu, L., Hindorff, L.A., Mitchell, S., Ambite, J.L., Boerwinkle, E., et al. (2012b). Evaluation of the Metabochip Genotyping Array in African Americans and Implications for Fine Mapping of GWAS-Identified Loci: The PAGE Study. PloS One *7*, e35651.

Calafat, A.M., Ye, X., Wong, L.-Y., Reidy, J.A., and Needham, L.L. (2008). Exposure of the U.S. population to bisphenol A and 4-tertiary-octylphenol: 2003-2004. Environ. Health Perspect. *116*, 39–44.

Calkins, D.J. (2012). Critical pathogenic events underlying progression of neurodegeneration in glaucoma. Prog. Retin. Eye Res. *31*, 702–719.

Calkins, D.J. (2013). Age-Related Changes in the Visual Pathways: Blame It on the Axon. Invest. Ophthalmol. Vis. Sci. *54*, ORSF37–ORSF41.

Cano, M., Wang, L., Wan, J., Barnett, B.P., Ebrahimi, K., Qian, J., and Handa, J.T. (2014). Oxidative stress induces mitochondrial dysfunction and a protective unfolded protein response in RPE cells. Free Radic. Biol. Med. *69*, 1–14.

Canter, J.A., Olson, L.M., Spencer, K., Schnetz-Boutaud, N., Anderson, B., Hauser, M.A., Schmidt, S., Postel, E.A., Agarwal, A., Pericak-Vance, M.A., et al. (2008). Mitochondrial DNA Polymorphism A4917G Is Independently Associated with Age-Related Macular Degeneration. PLoS ONE *3*, e2091.

Carlton, V.E.H., Harris, B.Z., Puffenberger, E.G., Batta, A.K., Knisely, A.S., Robinson, D.L., Strauss, K.A., Shneider, B.L., Lim, W.A., Salen, G., et al. (2003). Complex inheritance of familial hypercholanemia with associated mutations in TJP2 and BAAT. Nat. Genet. *34*, 91–96.

Cassin, B., and Solomon, S.D. Dictionary of Eye Terminology (Gainesville, Florida: Triad Publishing Company).

Chakravarthy, U., Augood, C., Bentham, G.C., de Jong, P.T.V.M., Rahu, M., Seland, J., Soubrane, G., Tomazzoli, L., Topouzis, F., Vingerling, J.R., et al. (2007). Cigarette smoking and age-related macular degeneration in the EUREYE Study. Ophthalmology *114*, 1157–1163.

Chakravarthy, U., Wong, T.Y., Fletcher, A., Piault, E., Evans, C., Zlateva, G., Buggage, R., Pleil, A., and Mitchell, P. (2010). Clinical risk factors for age-related macular degeneration: a systematic review and meta-analysis. BMC Ophthalmol. *10*, 31.

Chandrasekaran, S., Cumming, R.G., Rochtchina, E., and Mitchell, P. (2006). Associations between elevated intraocular pressure and glaucoma, use of glaucoma medications, and 5-year incident cataract: the Blue Mountains Eye Study. Ophthalmology *113*, 417–424.

Chang, M., Ned, R.M., Hong, Y., Yesupriya, A., Yang, Q., Liu, T., Janssens, A.C.J.W., and Dowling, N.F. (2011). Racial/ethnic variation in the association of lipid-related genetic variants with blood lipids in the US adult population. Circ. Cardiovasc. Genet. *4*, 523–533.

Chang, M.-H., Lindegren, M.L., Butler, M.A., Chanock, S.J., Dowling, N.F., Gallagher, M., Moonesinghe, R., Moore, C.A., Ned, R.M., Reichler, M.R., et al. (2009). Prevalence in the United States of selected candidate gene variants: Third National Health and Nutrition Examination Survey, 1991-1994. Am. J. Epidemiol. *169*, 54–66.

Chang, T.C., Congdon, N.G., Wojciechowski, R., Muñoz, B., Gilbert, D., Chen, P., Friedman, D.S., and West, S.K. (2005). Determinants and heritability of intraocular pressure and cup-to-disc ratio in a defined older population. Ophthalmology *112*, 1186–1191.

Charlesworth, J., Kramer, P.L., Dyer, T., Diego, V., Samples, J.R., Craig, J.E., Mackey, D.A., Hewitt, A.W., Blangero, J., and Wirtz, M.K. (2010). The path to open-angle glaucoma gene discovery: endophenotypic status of intraocular pressure, cup-to-disc ratio, and central corneal thickness. Invest. Ophthalmol. Vis. Sci. *51*, 3509–3514.

Chen, C., Bartenhagen, C., Gombert, M., Okpanyi, V., Binder, V., Röttgers, S., Bradtke, J., Teigler-Schlegel, A., Harbott, J., Ginzel, S., et al. (2013). Next-generation-sequencing-based risk stratification and identification of new genes involved in structural and sequence variations in near haploid lymphoblastic leukemia. Genes. Chromosomes Cancer *52*, 564–579.

Chen, D., Zhang, Z., Mao, C., Zhou, Y., Yu, L., Yin, Y., Wu, S., Mou, X., and Zhu, Y. (2014). ANRIL inhibits p15(INK4b) through the TGFβ1 signaling pathway in human esophageal squamous cell carcinoma. Cell. Immunol. *289*, 91–96.

Chen, J.L., Luviano, D.M., Chen, J.C., Yu, F., and Sarraf, D. (2009). Comparison of diabetic retinopathy phenotype between Latinos and Blacks. J. Diabetes Complications *23*, 371–375.

Chen, J.Z., Gokden, N., Greene, G.F., Mukunyadzi, P., and Kadlubar, F.F. (2002). Extensive somatic mitochondrial mutations in primary prostate cancer using laser capture microdissection. Cancer Res. *62*, 6470–6474.

Chen, W., Stambolian, D., Edwards, A.O., Branham, K.E., Othman, M., Jakobsdottir, J., Tosakulwong, N., Pericak-Vance, M.A., Campochiaro, P.A., Klein, M.L., et al. (2010). Genetic variants near TIMP3 and high-density lipoprotein-associated loci influence susceptibility to age-related macular degeneration. Proc. Natl. Acad. Sci. U. S. A. *107*, 7401–7406.

Cheng, I., Kocarnik, J.M., Dumitrescu, L., Lindor, N.M., Chang-Claude, J., Avery, C.L., Caberto, C.P., Love, S.-A., Slattery, M.L., Chan, A.T., et al. (2013). Pleiotropic effects of genetic risk variants for other cancers on colorectal cancer risk: PAGE, GECCO and CCFR consortia. Gut.

Cheng, M., Guo, Z., Li, H., Li, Z., Li, C., and Geng, C. (2014). Identification of sequence polymorphisms in the mitochondrial displacement loop as risk factors for sporadic and familial breast cancer. Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med. *35*, 4773–4777.

Cheng, Y.J., Gregg, E.W., Geiss, L.S., Imperatore, G., Williams, D.E., Zhang, X., Albright, A.L., Cowie, C.C., Klein, R., and Saaddine, J.B. (2009). Association of A1C and fasting plasma glucose levels with diabetic retinopathy prevalence in the U.S. population: Implications for diabetes diagnostic thresholds. Diabetes Care *32*, 2027–2032.

Cherecheanu, A.P., Garhofer, G., Schmidl, D., Werkmeister, R., and Schmetterer, L. (2013). Ocular perfusion pressure and ocular blood flow in glaucoma. Curr. Opin. Pharmacol. *13*, 36–42.

Cheung, C.M.G., Tai, E.S., Kawasaki, R., Tay, W.T., Lee, J.L., Hamzah, H., and Wong, T.Y. (2012). Prevalence of and risk factors for age-related macular degeneration in a multiethnic Asian cohort. Arch. Ophthalmol. *130*, 480–486.

Chew, E.Y., Klein, M.L., Clemons, T.E., Agrón, E., Ratnapriya, R., Edwards, A.O., Fritsche, L.G., Swaroop, A., Abecasis, G.R., and Age-Related Eye Disease Study Research Group (2014). No clinically significant

association between CFH and ARMS2 genotypes and response to nutritional supplements: AREDS report number 38. Ophthalmology *121*, 2173–2180.

Chong, E.W.-T., Wong, T.Y., Kreis, A.J., Simpson, J.A., and Guymer, R.H. (2007). Dietary antioxidants and primary prevention of age related macular degeneration: systematic review and meta-analysis. BMJ *335*, 755.

Chuang, J.-Z., Zhao, Y., and Sung, C.-H. (2007). SARA-regulated vesicular targeting underlies formation of the light-sensing organelle in mammalian rods. Cell *130*, 535–547.

Cipriani, V., Leung, H.-T., Plagnol, V., Bunce, C., Khan, J.C., Shahid, H., Moore, A.T., Harding, S.P., Bishop, P.N., Hayward, C., et al. (2012). Genome-wide association study of age-related macular degeneration identifies associated variants in the TNXB-FKBPL-NOTCH4 region of chromosome 6p21.3. Hum. Mol. Genet. *21*, 4138–4150.

Colak, E., Kosanović-Jaković, N., Zorić, L., Radosavljević, A., Stanković, S., and Majkić-Singh, N. (2011). The association of lipoprotein parameters and C-reactive protein in patients with age-related macular degeneration. Ophthalmic Res. *46*, 125–132.

Congdon, N., O'Colmain, B., Klaver, C.C.W., Klein, R., Muñoz, B., Friedman, D.S., Kempen, J., Taylor, H.R., and Mitchell, P. (2004). Causes and prevalence of visual impairment among adults in the United States. Arch. Ophthalmol. *122*, 477–485.

Conley, Y.P., Jakobsdottir, J., Mah, T., Weeks, D.E., Klein, R., Kuller, L., Ferrell, R.E., and Gorin, M.B. (2006). CFH, ELOVL4, PLEKHA1 and LOC387715 genes and susceptibility to age-related maculopathy: AREDS and CHS cohorts and meta-analyses. Hum. Mol. Genet. *15*, 3206–3218.

Contreras, A.V., Zenteno, J.C., Fernández-López, J.C., Rodríguez-Corona, U., Falfán-Valencia, R., Sebastian, L., Morales, F., Ochoa-Contreras, D., Carnevale, A., and Silva-Zolezzi, I. (2014). CFH haplotypes and ARMS2, C2, C3, and CFB alleles show association with susceptibility to age-related macular degeneration in Mexicans. Mol. Vis. *20*, 105–116.

Corbie-Smith, G., Thomas, S.B., Williams, M.V., and Moody-Ayers, S. (1999). Attitudes and Beliefs of African Americans Toward Participation in Medical Research. J. Gen. Intern. Med. *14*, 537–546.

Corbie-Smith, G., Thomas, S.B., and St George, D.M.M. (2002). Distrust, race, and research. Arch. Intern. Med. *162*, 2458–2463.

Courtenay, M.D., Cade, W., Schwartz, S.G., Kovach, J.L., Agarwal, A., Wang, G., Haines, J.L., Pericak-Vance, M.A., and Scott, W.K. (2014). Set-based joint test of interaction between SNPs in the VEGF pathway and exogenous estrogen finds association with age-related macular degeneration. Invest. Ophthalmol. Vis. Sci.

Crawford, D., Goodloe, R., Farber-Eger, E., Boston, J., Pendergrass, S., Haines, J., Ritchie, M., and Bush, W. Leveraging epidemiologic and clinical collections for genomic studies of complex traits. Submitted.

Crawford, D.C., Goodloe, R., Brown-Gentry, K., Wilson, S., Roberson, J., Gillani, N.B., Ritchie, M.D., Dilks, H.H., and Bush, W.S. (2013). Characterization of the Metabochip in diverse populations from the

International HapMap Project in the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) project. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 188–199.

Cristina Kenney, M., Chwa, M., Atilano, S.R., Falatoonzadeh, P., Ramirez, C., Malik, D., Tarek, M., Cáceres-Del-Carpio, J., Nesburn, A.B., Boyer, D.S., et al. (2014). Inherited mitochondrial DNA variants can affect complement, inflammation and apoptosis pathways: insights into mitochondrial-nuclear interactions. Hum. Mol. Genet. *23*, 3537–3551.

Crosslin, D.R., McDavid, A., Weston, N., Nelson, S.C., Zheng, X., Hart, E., de Andrade, M., Kullo, I.J., McCarty, C.A., Doheny, K.F., et al. (2012). Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. Hum. Genet. *131*, 639–652.

Crosslin, D.R., McDavid, A., Weston, N., Zheng, X., Hart, E., de Andrade, M., Kullo, I.J., McCarty, C.A., Doheny, K.F., Pugh, E., et al. (2013). Genetic variation associated with circulating monocyte count in the eMERGE Network. Hum. Mol. Genet. *22*, 2119–2127.

Cubey, R. bevi. (1976). Glaucoma following the application of corticosteroid to the skin of the eyelids. Br. J. Dermatol. *95*, 207–208.

Curcio, C.A., and Allen, K.A. (1990). Topography of ganglion cells in human retina. J. Comp. Neurol. *300*, 5–25.

Cutter, J., Tan, B.Y., and Chew, S.K. (2001). Levels of cardiovascular disease risk factors in Singapore following a national intervention programme. Bull. World Health Organ. *79*, 908–915.

Danias, J., Lee, K.C., Zamora, M.-F., Chen, B., Shen, F., Filippopoulos, T., Su, Y., Goldblum, D., Podos, S.M., and Mittag, T. (2003). Quantitative analysis of retinal ganglion cell (RGC) loss in aging DBA/2NNia glaucomatous mice: comparison with RGC loss in aging C57/BL6 mice. Invest. Ophthalmol. Vis. Sci. *44*, 5151–5162.

Darwin, C. On the Origin of Species by Means of Natural Selection,.

Das, S.K., and Elbein, S.C. (2006). The Genetic Basis of Type 2 Diabetes. Cellscience *2*, 100–131.

David S. Friedman (last), O'Colmain, B., and Ilona Mestril (2012). 2012 Fifth Edition of Vision Problems in the U.S.

DeAngelis, M.M., Ji, F., Kim, I.K., Adams, S., Capone, A., Ott, J., Miller, J.W., and Dryja, T.P. (2007). Cigarette smoking, CFH, APOE, ELOVL4, and risk of neovascular age-related macular degeneration. Arch. Ophthalmol. *125*, 49–54.

Delaney, J.T., Ramirez, A.H., Bowton, E., Pulley, J.M., Basford, M.A., Schildcrout, J.S., Shi, Y., Zink, R., Oetjens, M., Xu, H., et al. (2012). Predicting clopidogrel response using DNA samples linked to an electronic health record. Clin. Pharmacol. Ther. *91*, 257–263.

Denny, J.C., Ritchie, M.D., Crawford, D.C., Schildcrout, J.S., Ramirez, A.H., Pulley, J.M., Basford, M.A., Masys, D.R., Haines, J.L., and Roden, D.M. (2010). Identification of genomic predictors of atrioventricular conduction: using electronic medical records as a tool for genome science. Circulation *122*, 2016–2021.

Denny, J.C., Crawford, D.C., Ritchie, M.D., Bielinski, S.J., Basford, M.A., Bradford, Y., Chai, H.S., Bastarache, L., Zuvich, R., Peissig, P., et al. (2011). Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical records for genome- and phenome-wide studies. Am. J. Hum. Genet. *89*, 529–542.

Despriet, D.D.G., Klaver, C.C.W., Witteman, J.C.M., Bergen, A.A.B., Kardys, I., de Maat, M.P.M., Boekhoorn, S.S., Vingerling, J.R., Hofman, A., Oostra, B.A., et al. (2006). Complement factor H polymorphism, complement activators, and risk of age-related macular degeneration. JAMA J. Am. Med. Assoc. *296*, 301–309.

Dewan, A., Liu, M., Hartman, S., Zhang, S.S.-M., Liu, D.T.L., Zhao, C., Tam, P.O.S., Chan, W.M., Lam, D.S.C., Snyder, M., et al. (2006). HTRA1 promoter polymorphism in wet age-related macular degeneration. Science *314*, 989–992.

DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, Mexican American Type 2 Diabetes (MAT2D) Consortium, Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples (T2D-GENES) Consortium, Mahajan, A., Go, M.J., Zhang, W., Below, J.E., Gaulton, K.J., et al. (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat. Genet. *46*, 234–244.

Dielemans, I., Vingerling, J.R., Wolfs, R.C., Hofman, A., Grobbee, D.E., and de Jong, P.T. (1994). The prevalence of primary open-angle glaucoma in a population-based study in The Netherlands. The Rotterdam Study. Ophthalmology *101*, 1851–1855.

Ding, K., de Andrade, M., Manolio, T.A., Crawford, D.C., Rasmussen-Torvik, L.J., Ritchie, M.D., Denny, J.C., Masys, D.R., Jouni, H., Pachecho, J.A., et al. (2013). Genetic variants that confer resistance to malaria are associated with red blood cell traits in African-Americans: an electronic medical record-based genome-wide association study. G3 Bethesda Md *3*, 1061–1068.

Dumitrescu, L., Carty, C.L., Taylor, K., Schumacher, F.R., Hindorff, L.A., Ambite, J.L., Anderson, G., Best, L.G., Brown-Gentry, K., Bůžková, P., et al. (2011). Genetic Determinants of Lipid Traits in Diverse Populations from the Population Architecture using Genomics and Epidemiology (PAGE) Study. PLoS Genet. *7*.

Ebrahimi, K.B., and Handa, J.T. (2011). Lipids, Lipoproteins, and Age-Related Macular Degeneration. J. Lipids *2011*, 1–14.

Edwards, A.O., Ritter, R., 3rd, Abel, K.J., Manning, A., Panhuysen, C., and Farrer, L.A. (2005). Complement factor H polymorphism and age-related macular degeneration. Science *308*, 421–424.

Eichler, E.E., Flint, J., Gibson, G., Kong, A., Leal, S.M., Moore, J.H., and Nadeau, J.H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. Nat. Rev. Genet. *11*, 446–450.

Emanuele, N., Sacks, J., Klein, R., Reda, D., Anderson, R., Duckworth, W., Abraira, C., and Veterans Affairs Diabetes Trial Group (2005). Ethnicity, race, and baseline retinopathy correlates in the veterans affairs diabetes trial. Diabetes Care *28*, 1954–1958.

Emanuele, N., Moritz, T., Klein, R., Davis, M.D., Glander, K., Khanna, A., Thottapurathu, L., Bahn, G., Duckworth, W., Abraira, C., et al. (2009). Ethnicity, race, and clinically significant macular edema in the Veterans Affairs Diabetes Trial (VADT). Diabetes Res. Clin. Pract. *86*, 104–110.

Fagundes, N.J.R., Kanitz, R., Eckert, R., Valls, A.C.S., Bogo, M.R., Salzano, F.M., Smith, D.G., Silva, W.A., Zago, M.A., Ribeiro-dos-Santos, A.K., et al. (2008). Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. Am. J. Hum. Genet. *82*, 583–592.

Findl, O., Rainer, G., Dallinger, S., Dorner, G., Polak, K., Kiss, B., Georgopoulos, M., Vass, C., and Schmetterer, L. (2000). Assessment of optic disk blood flow in patients with open-angle glaucoma. Am. J. Ophthalmol. *130*, 589–596.

Fingert, J.H., Stone, E.M., Sheffield, V.C., and Alward, W.L.M. (2002). Myocilin glaucoma. Surv. Ophthalmol. *47*, 547–561.

Flammer, J. (1994). The vascular concept of glaucoma. Surv. Ophthalmol. *38 Suppl*, S3–S6.

Flammer, J., Orgül, S., Costa, V.P., Orzalesi, N., Krieglstein, G.K., Serra, L.M., Renard, J.-P., and Stefánsson, E. (2002). The impact of ocular blood flow in glaucoma. Prog. Retin. Eye Res. *21*, 359–393.

Fong, D.S., Aiello, L., Gardner, T.W., King, G.L., Blankenship, G., Cavallerano, J.D., Ferris, F.L., and Klein, R. (2004). Retinopathy in Diabetes. Diabetes Care *27*, 84S – 87.

Fontana, L., Poinoosawmy, D., Bunce, C.V., O'Brien, C., and Hitchings, R.A. (1998). Pulsatile ocular blood flow investigation in asymmetric normal tension glaucoma and normal subjects. Br. J. Ophthalmol. *82*, 731–736.

Foong, A.W.P., Saw, S.-M., Loo, J.-L., Shen, S., Loon, S.-C., Rosman, M., Aung, T., Tan, D.T.H., Tai, E.S., and Wong, T.Y. (2007). Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES). Ophthalmic Epidemiol. *14*, 25–35.

Freeman, E.E., Muñoz, B., Rubin, G., and West, S.K. (2007). Visual field loss increases the risk of falls in older adults: the Salisbury eye evaluation. Invest. Ophthalmol. Vis. Sci. *48*, 4445–4450.

Freeman, E.E., Roy-Gagnon, M.-H., Descovich, D., Massé, H., and Lesk, M.R. (2013). The heritability of glaucoma-related traits corneal hysteresis, central corneal thickness, intraocular pressure, and choroidal blood flow pulsatility. PloS One *8*, e55573.

Freimuth, V.S., Quinn, S.C., Thomas, S.B., Cole, G., Zook, E., and Duncan, T. (2001). African Americans' views on research and the Tuskegee Syphilis study. Soc. Sci. Med. *52*, 797–808.

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

Friedman, N.E., and Kaiser, P. (2007). Essentials of Ophthalmology (Philadelphia, PA: Elsevier Inc).

Friedman, D.S., Katz, J., Bressler, N.M., Rahmani, B., and Tielsch, J.M. (1999). Racial differences in the prevalence of age-related macular degeneration: the Baltimore Eye Survey. Ophthalmology *106*, 1049–1055.

Friedman, D.S., Wolfs, R.C.W., O'Colmain, B.J., Klein, B.E., Taylor, H.R., West, S., Leske, M.C., Mitchell, P., Congdon, N., and Kempen, J. (2004a). Prevalence of open-angle glaucoma among adults in the United States. Arch. Ophthalmol. *122*, 532–538.

Friedman, D.S., O'Colmain, B.J., Muñoz, B., Tomany, S.C., McCarty, C., de Jong, P.T.V.M., Nemesure, B., Mitchell, P., and Kempen, J. (2004b). Prevalence of age-related macular degeneration in the United States. Arch. Ophthalmol. *122*, 564–572.

Fritsche, L.G., Loenhardt, T., Janssen, A., Fisher, S.A., Rivera, A., Keilhauer, C.N., and Weber, B.H.F. (2008). Age-related macular degeneration is associated with an unstable ARMS2 (LOC387715) mRNA. Nat. Genet. *40*, 892–896.

Fritsche, L.G., Chen, W., Schu, M., Yaspan, B.L., Yu, Y., Thorleifsson, G., Zack, D.J., Arakawa, S., Cipriani, V., Ripke, S., et al. (2013). Seven new loci associated with age-related macular degeneration. Nat. Genet.

Fritsche, L.G., Fariss, R.N., Stambolian, D., Abecasis, G.R., Curcio, C.A., and Swaroop, A. (2014). Age-Related Macular Degeneration: Genetics and Biology Coming Together. Annu. Rev. Genomics Hum. Genet.

Fu, Y.-P., Hallman, D.M., Gonzalez, V.H., Klein, B.E.K., Klein, R., Hayes, M.G., Cox, N.J., Bell, G.I., and Hanis, C.L. (2010). Identification of Diabetic Retinopathy Genes through a Genome-Wide Association Study among Mexican-Americans from Starr County, Texas. J. Ophthalmol. *2010*.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., et al. (2002). The Structure of Haplotype Blocks in the Human Genome. Science *296*, 2225–2229.

Galea, S., and Tracy, M. (2007). Participation Rates in Epidemiologic Studies. Ann. Epidemiol. *17*, 643–653.

Gehlbach, P., Li, T., and Hatef, E. (2009). Statins for age-related macular degeneration. Cochrane Database Syst. Rev. Online CD006927.

Gemmy Cheung, C.M., Li, X., Cheng, C.-Y., Zheng, Y., Mitchell, P., Wang, J.J., Jonas, J.B., Nangia, V., and Wong, T.Y. (2012). Prevalence and Risk Factors for Age-Related Macular Degeneration in Indians: A Comparative Study in Singapore and India. Am. J. Ophthalmol.

Geraldes, P., Hiraoka-Yamamoto, J., Matsumoto, M., Clermont, A., Leitges, M., Marette, A., Aiello, L.P., Kern, T.S., and King, G.L. (2009). Activation of PKC-δ and SHP-1 by hyperglycemia causes vascular cell apoptosis and diabetic retinopathy. Nat. Med. *15*, 1298–1306.

Geyer, O., Wolf, A., Levinger, E., Harari-Shacham, A., Walton, D.S., Shochat, C., Korem, S., and Bercovich, D. (2011). Genotype/phenotype correlation in primary congenital glaucoma patients from different ethnic groups of the Israeli population. Am. J. Ophthalmol. *151*, 263–271.e1.

Ghanbari, H., Dalloul, G., Hasan, R., Daccarett, M., Saba, S., David, S., and Machado, C. (2009). Effectiveness of implantable cardioverter-defibrillators for the primary prevention of sudden cardiac death in women with advanced heart failure: a meta-analysis of randomized controlled trials. Arch. Intern. Med. *169*, 1500–1506.

Gold, B., Merriam, J.E., Zernant, J., Hancox, L.S., Taiber, A.J., Gehrs, K., Cramer, K., Neel, J., Bergeron, J., Barile, G.R., et al. (2006). Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. Nat. Genet. *38*, 458–462.

Goldberg, J., Flowerdew, G., Smith, E., Brody, J.A., and Tso, M.O. (1988). Factors associated with age-related macular degeneration. An analysis of data from the first National Health and Nutrition Examination Survey. Am. J. Epidemiol. *128*, 700–710.

Gong, J., Schumacher, F., Lim, U., Hindorff, L.A., Haessler, J., Buyske, S., Carlson, C.S., Rosse, S., Bůžková, P., Fornage, M., et al. (2013). Fine Mapping and Identification of BMI Loci in African Americans. Am. J. Hum. Genet. *93*, 661–671.

Gordon MO, Beiser JA, Brandt JD, and et al (2002). The ocular hypertension treatment study: Baseline factors that predict the onset of primary open-angle glaucoma. Arch. Ophthalmol. *120*, 714–720.

Goto, A., Akahori, M., Okamoto, H., Minami, M., Terauchi, N., Haruhata, Y., Obazawa, M., Noda, T., Honda, M., Mizota, A., et al. (2009). Genetic analysis of typical wet-type age-related macular degeneration and polypoidal choroidal vasculopathy in Japanese population. J. Ocul. Biol. Dis. Infor. *2*, 164–175.

Goulart, A.C., Germer, S., Rexrode, K.M., Martin, M., and Zee, R.Y.L. (2008). Polymorphisms in advanced glycosylation end product-specific receptor (AGER) gene, insulin resistance, and type 2 diabetes mellitus. Clin. Chim. Acta Int. J. Clin. Chem. *398*, 95–98.

Grassi, M.A., Fingert, J.H., Scheetz, T.E., Roos, B.R., Ritch, R., West, S.K., Kawase, K., Shire, A.M., Mullins, R.F., and Stone, E.M. (2006). Ethnic variation in AMD-associated complement factor H polymorphism p.Tyr402His. Hum. Mutat. *27*, 921–925.

Grassi, M.A., Tikhomirov, A., Ramalingam, S., Below, J.E., Cox, N.J., and Nicolae, D.L. (2011). Genome-wide meta-analysis for severe diabetic retinopathy. Hum. Mol. Genet. *20*, 2472–2481.

Grassmann, F., Friedrich, U., Fauser, S., Schick, T., Milenkovic, A., Schulz, H.L., von Strachwitz, C.N., Bettecken, T., Lichtner, P., Meitinger, T., et al. (2015). A Candidate Gene Association Study Identifies DAPL1 as a Female-Specific Susceptibility Locus for Age-Related Macular Degeneration (AMD). Neuromolecular Med.

Groot, K.R., Sevilla, L.M., Nishi, K., DiColandrea, T., and Watt, F.M. (2004). Kazrin, a novel periplakin-interacting protein associated with desmosomes and the keratinocyte plasma membrane. J. Cell Biol. *166*, 653–659.

Gunter, E., Lewis, B.G., and Koncikowski, S.M. (1996). Laboratory Procedures Used for the Third National Health and Nutrition Examination Survey (NHANES III), 1988-1994.

Gupta, P.D., Johar, K., Nagpal, K., and Vasavada, A.R. (2005). Sex hormone receptors in the human eye. Surv. Ophthalmol. *50*, 274–284.

Haas, P., Kubista, K.E., Krugluger, W., Huber, J., and Binder, S. (2015). Impact of visceral fat and pro-inflammatory factors on the pathogenesis of age-related macular degeneration. Acta Ophthalmol. (Copenh.).

Haiman, C.A., Fesinmeyer, M.D., Spencer, K.L., Buzková, P., Voruganti, V.S., Wan, P., Haessler, J., Franceschini, N., Monroe, K.R., Howard, B.V., et al. (2012). Consistent directions of effect for established type 2 diabetes risk variants across populations: the population architecture using Genomics and Epidemiology (PAGE) Consortium. Diabetes *61*, 1642–1647.

Haines, J.L., Hauser, M.A., Schmidt, S., Scott, W.K., Olson, L.M., Gallins, P., Spencer, K.L., Kwan, S.Y., Noureddine, M., Gilbert, J.R., et al. (2005). Complement factor H variant increases the risk of age-related macular degeneration. Science *308*, 419–421.

Hammes, H.-P. (2005). Pericytes and the pathogenesis of diabetic retinopathy. Horm. Metab. Res. Horm. Stoffwechselforschung Horm. Métabolisme *37 Suppl 1*, 39–43.

Harris, C.A., Henttu, P., Parker, M.G., and Sumpter, J.P. (1997). The estrogenic activity of phthalate esters in vitro. Environ. Health Perspect. *105*, 802–811.

Heijl, A., Leske, M.C., Bengtsson, B., Hyman, L., Bengtsson, B., Hussein, M., and Early Manifest Glaucoma Trial Group (2002). Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. Arch. Ophthalmol. *120*, 1268–1279.

Hietala, K., Forsblom, C., Summanen, P., and Groop, P.-H. (2008). Heritability of Proliferative Diabetic Retinopathy. Diabetes *57*, 2176–2180.

Hildebrand, G., Darius, and Fielder, A., R. (2003). Anatomy and Physiology of the Retina. In Ocular Circulation (chapter 33), (Mosby, St Louis: Springer), pp. 747–784.

Hill, S.E., Rodeheffer, C.D., Griskevicius, V., Durante, K., and White, A.E. (2012). Boosting beauty in an economic decline: mating, spending, and the lipstick effect. J. Pers. Soc. Psychol. *103*, 275–291.

Hirschhorn, J.N., and Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. *6*, 95–108.

Hoffmann, T.J., Tang, H., Thornton, T.A., Caan, B., Haan, M., Millen, A.E., Thomas, F., and Risch, N. (2014). Genome-wide association and admixture analysis of glaucoma in the Women's Health Initiative. Hum. Mol. Genet. *23*, 6634–6643.

Hollenbeck, P.J., and Saxton, W.M. (2005). The axonal transport of mitochondria. J. Cell Sci. *118*, 5411–5419.

Holliday, E.G., Smith, A.V., Cornes, B.K., Buitendijk, G.H.S., Jensen, R.A., Sim, X., Aspelund, T., Aung, T., Baird, P.N., Boerwinkle, E., et al. (2013). Insights into the Genetic Architecture of Early Stage Age-Related Macular Degeneration: A Genome-Wide Association Study Meta-Analysis. PLoS ONE *8*.

Howard, K.P., Klein, B.E.K., Lee, K.E., and Klein, R. (2014). Measures of body shape and adiposity as related to incidence of age-related eye diseases: observations from the Beaver Dam Eye Study. Invest. Ophthalmol. Vis. Sci. *55*, 2592–2598.

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

Hughes, A.E., Orr, N., Patterson, C., Esfandiary, H., Hogg, R., McConnell, V., Silvestri, G., and Chakravarthy, U. (2007). Neovascular age-related macular degeneration risk based on CFH, LOC387715/HTRA1, and smoking. PLoS Med. *4*, e355.

Hughes, K., Yeo, P.P., Lun, K.C., Thai, A.C., Sothy, S.P., Wang, K.W., Cheah, J.S., Phoon, W.O., and Lim, P. (1990). Cardiovascular diseases in Chinese, Malays, and Indians in Singapore. II. Differences in risk factor levels. J. Epidemiol. Community Health *44*, 29–35.

Hughes, K., Aw, T.C., Kuperan, P., and Choo, M. (1997). Central obesity, insulin resistance, syndrome X, lipoprotein(a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore. J. Epidemiol. Community Health *51*, 394–399.

Hyttinen, V., Kaprio, J., Kinnunen, L., Koskenvuo, M., and Tuomilehto, J. (2003). Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. Diabetes *52*, 1052–1055.

Imperatore, G., Hanson, R.L., Pettitt, D.J., Kobes, S., Bennett, P.H., and Knowler, W.C. (1998). Sib-pair linkage analysis for susceptibility genes for microvascular complications among Pima Indians with type 2 diabetes. Pima Diabetes Genes Group. Diabetes *47*, 821–830.

Institute of Medicine (US) Committee on Women's Health Research (2010). Women's Health Research: Progress, Pitfalls, and Promise (Washington (DC): National Academies Press (US)).

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

Ivers, R.Q., Cumming, R.G., Mitchell, P., and Attebo, K. (1998). Visual impairment and falls in older adults: the Blue Mountains Eye Study. J. Am. Geriatr. Soc. *46*, 58–64.

Jacobson, N., Andrews, M., Shepard, A.R., Nishimura, D., Searby, C., Fingert, J.H., Hageman, G., Mullins, R., Davidson, B.L., Kwon, Y.H., et al. (2001). Non-secretion of mutant proteins of the glaucoma gene myocilin in cultured trabecular meshwork cells and in aqueous humor. Hum. Mol. Genet. *10*, 117–125.

Jakobsdottir, J., Conley, Y.P., Weeks, D.E., Mah, T.S., Ferrell, R.E., and Gorin, M.B. (2005). Susceptibility genes for age-related maculopathy on chromosome 10q26. Am. J. Hum. Genet. *77*, 389–407.

Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.-C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., et al. (2008). Genotype, haplotype and copy-number variation in worldwide human populations. Nature *451*, 998–1003.

Jeff, J.M., Ritchie, M.D., Denny, J.C., Kho, A.N., Ramirez, A.H., Crosslin, D., Armstrong, L., Basford, M.A., Wolf, W.A., Pacheco, J.A., et al. (2013). Generalization of Variants Identified by Genome-Wide Association Studies for Electrocardiographic Traits in African Americans. Ann. Hum. Genet.

Jeganathan, V.S.E., Sabanayagam, C., Tai, E.S., Lee, J., Lamoureux, E., Sun, C., Kawasaki, R., and Wong, T.Y. (2009). Retinal Vascular Caliber and Diabetes in a Multiethnic Asian Population. Microcirculation *16*, 534–543.

Jiang, X., Varma, R., Wu, S., Torres, M., Azen, S.P., Francis, B.A., Chopra, V., and Nguyen, B.B.-T. (2012). Baseline risk factors that predict the development of open-angle glaucoma in a population: the Los Angeles Latino Eye Study. Ophthalmology *119*, 2245–2253.

Jones, M.M., Manwaring, N., Wang, J.J., Rochtchina, E., Mitchell, P., and Sue, C.M. (2007). Mitochondrial DNA haplogroups and age-related maculopathy. Arch. Ophthalmol. *125*, 1235–1240.

De Jong, P.T.V.M. (2006). Age-related macular degeneration. N. Engl. J. Med. *355*, 1474–1485.

Kaiser, J., and Servick, K. (2015). Biomedical initiatives get top billing in budget rollout. Science *347*, 601.

Kang, P., Tian, C., and Jia, C. (2012). Association of RAGE gene polymorphisms with type 2 diabetes mellitus, diabetic retinopathy and diabetic nephropathy. Gene *500*, 1–9.

Kass, M.A., Heuer, D.K., Higginbotham, E.J., Johnson, C.A., Keltner, J.L., Miller, J.P., Parrish, R.K., Wilson, M.R., and Gordon, M.O. (2002). The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Arch. Ophthalmol. *120*, 701–713; discussion 829–830.

Kathiresan, S., Willer, C.J., Peloso, G.M., Demissie, S., Musunuru, K., Schadt, E.E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T., et al. (2009). Common variants at 30 loci contribute to polygenic dyslipidemia. Nat. Genet. *41*, 56–65.

Katz, R.V., Green, B.L., Kressin, N.R., Claudio, C., Wang, M.Q., and Russell, S.L. (2007). Willingness of minorities to participate in biomedical studies: confirmatory findings from a follow-up study using the Tuskegee Legacy Project Questionnaire. J. Natl. Med. Assoc. *99*, 1052–1060.

Kaur, I., Katta, S., Reddy, R.K., Narayanan, R., Mathai, A., Majji, A.B., and Chakrabarti, S. (2010). The involvement of complement factor B and complement component C2 in an Indian cohort with age-related macular degeneration. Invest. Ophthalmol. Vis. Sci. *51*, 59–63.

Kawasaki, R., Wang, J.J., Aung, T., Tan, D.T.H., Mitchell, P., Sandar, M., Saw, S.-M., Wong, T.Y., and Singapore Malay Eye Study Group (2008). Prevalence of age-related macular degeneration in a Malay population: the Singapore Malay Eye Study. Ophthalmology *115*, 1735–1741.

Kawasaki, R., Yasuda, M., Song, S.J., Chen, S.-J., Jonas, J.B., Wang, J.J., Mitchell, P., and Wong, T.Y. (2010). The Prevalence of Age-Related Macular Degeneration in Asians: A Systematic Review and Meta-Analysis. Ophthalmology *117*, 921–927.

Keebler, M.E., Sanders, C.L., Surti, A., Guiducci, C., Burtt, N.P., and Kathiresan, S. (2009). Association of blood lipids with common DNA sequence variants at 19 genetic loci in the multiethnic United States National Health and Nutrition Examination Survey III. Circ. Cardiovasc. Genet. *2*, 238–243.

Kenney, M.C., Chwa, M., Atilano, S.R., Pavlis, J.M., Falatoonzadeh, P., Ramirez, C., Malik, D., Hsu, T., Woo, G., Soe, K., et al. (2013). Mitochondrial DNA variants mediate energy production and expression levels for CFH, C3 and EFEMP1 genes: implications for age-related macular degeneration. PloS One *8*, e54339.

Kenney, M.C., Chwa, M., Atilano, S.R., Falatoonzadeh, P., Ramirez, C., Malik, D., Tarek, M., Del Carpio, J.C., Nesburn, A.B., Boyer, D.S., et al. (2014). Molecular and bioenergetic differences between cells with African versus European inherited mitochondrial DNA haplogroups: implications for population susceptibility to diseases. Biochim. Biophys. Acta *1842*, 208–219.

Kho, A.N., Pacheco, J.A., Peissig, P.L., Rasmussen, L., Newton, K.M., Weston, N., Crane, P.K., Pathak, J., Chute, C.G., Bielinski, S.J., et al. (2011). Electronic medical records for genetic research: results of the eMERGE consortium. Sci. Transl. Med. *3*, 79re1.

Kho, A.N., Hayes, M.G., Rasmussen-Torvik, L., Pacheco, J.A., Thompson, W.K., Armstrong, L.L., Denny, J.C., Peissig, P.L., Miller, A.W., Wei, W.-Q., et al. (2012). Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. J. Am. Med. Inform. Assoc. JAMIA *19*, 212–218.

Klein, R., and Klein, B. (1995). Vision disorders in diabetes.

Klein, R., Klein, B.E., Moss, S.E., Davis, M.D., and DeMets, D.L. (1984). The Wisconsin epidemiologic study of diabetic retinopathy. II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. Arch. Ophthalmol. *102*, 520–526.

Klein, R., Davis, M.D., Magli, Y.L., Segal, P., Klein, B.E., and Hubbard, L. (1991). The Wisconsin age-related maculopathy grading system. Ophthalmology *98*, 1128–1134.

Klein, R., Klein, B.E.K., Jensen, S.C., and Meuer, S.M. (1997). The Five-year Incidence and Progression of Age-related Maculopathy: The Beaver Dam Eye Study. Ophthalmology *104*, 7–21.

Klein, R., Klein, B.E.K., Tomany, S.C., and Cruickshanks, K.J. (2003a). The association of cardiovascular disease with the long-term incidence of age-related maculopathy: the Beaver Dam Eye Study. Ophthalmology *110*, 1273–1280.

Klein, R., Klein, B.E.., Marino, E.K., Kuller, L.H., Furberg, C., Burke, G.L., and Hubbard, L.D. (2003b). Early age-related maculopathy in the cardiovascular health study. Ophthalmology *110*, 25–33.

Klein, R., Klein, B.E.K., Knudtson, M.D., Cotch, M.F., Wong, T.Y., Liu, K., Burke, G.L., Saad, M.F., Jacobs, D.R., Jr, and Sharrett, A.R. (2007a). Subclinical atherosclerotic cardiovascular disease and early age-related macular degeneration in a multiracial cohort: the Multiethnic Study of Atherosclerosis. Arch. Ophthalmol. *125*, 534–543.

Klein, R., Klein, B.E.K., Knudtson, M.D., Cotch, M.F., Wong, T.Y., Liu, K., Burke, G.L., Saad, M.F., Jacobs, D.R., Jr, and Sharrett, A.R. (2007b). Subclinical atherosclerotic cardiovascular disease and early age-

related macular degeneration in a multiracial cohort: the Multiethnic Study of Atherosclerosis. Arch. Ophthalmol. *125*, 534–543.

Klein, R., Knudtson, M.D., Klein, B.E.K., Wong, T.Y., Cotch, M.F., Liu, K., Cheng, C.Y., Burke, G.L., Saad, M.F., Jacobs, D.R., Jr, et al. (2008). Inflammation, complement factor h, and age-related macular degeneration: the Multi-ethnic Study of Atherosclerosis. Ophthalmology *115*, 1742–1749.

Klein, R., Chou, C.-F., Klein, B.E.K., Zhang, X., Meuer, S.M., and Saaddine, J.B. (2011). Prevalence of age-related macular degeneration in the US population. Arch. Ophthalmol. *129*, 75–80.

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., et al. (2005). Complement factor H polymorphism in age-related macular degeneration. Science *308*, 385–389.

Klein R, C.C. (2011). PRevalence of age-related macular degeneration in the us population. Arch. Ophthalmol. *129*, 75–80.

Klein R, Clegg L, Cooper LS, and et al (1999). PRevalence of age-related maculopathy in the atherosclerosis risk in communities study. Arch. Ophthalmol. *117*, 1203–1210.

Kleinstein, R.N., Jones, L.A., Hullett, S., Kwon, S., Lee, R.J., Friedman, N.E., Manny, R.E., Mutti, D.O., Yu, J.A., and Zadnik, K. (2003). Refractive error and ethnicity in children. Arch. Ophthalmol. *121*, 1141–1147.

Kocarnik, J.M., Park, S.L., Han, J., Dumitrescu, L., Cheng, I., Wilkens, L.R., Schumacher, F.R., Kolonel, L., Carlson, C.S., Crawford, D.C., et al. (2014). Replication of Associations Between GWAS SNPs and Melanoma Risk in the Population Architecture Using Genomics and Epidemiology (PAGE) Study. J. Invest. Dermatol.

Kohane, I.S. (2011). Using electronic health records to drive discovery in disease genomics. Nat. Rev. Genet. *12*, 417–428.

Van Koolwijk, L.M.E., Despriet, D.D.G., van Duijn, C.M., Pardo Cortes, L.M., Vingerling, J.R., Aulchenko, Y.S., Oostra, B.A., Klaver, C.C.W., and Lemij, H.G. (2007). Genetic contributions to glaucoma: heritability of intraocular pressure, retinal nerve fiber layer thickness, and optic disc morphology. Invest. Ophthalmol. Vis. Sci. *48*, 3669–3676.

Kullberg, C.E., and Arnqvist, H.J. (1995). Good blood glucose control characterizes patients without retinopathy after long diabetes duration. Diabet. Med. J. Br. Diabet. Assoc. *12*, 314–320.

Kuzin, A.A., Varma, R., Reddy, H.S., Torres, M., Azen, S.P., and Los Angeles Latino Eye Study Group (2010). Ocular biometry and open-angle glaucoma: the Los Angeles Latino Eye Study. Ophthalmology *117*, 1713–1719.

Kwon, Y.H., Fingert, J.H., Kuehn, M.H., and Alward, W.L.M. (2009). Primary open-angle glaucoma. N. Engl. J. Med. *360*, 1113–1124.

Lachin, J.M., Genuth, S., Nathan, D.M., Zinman, B., and Rutledge, B.N. (2008). Effect of glycemic exposure on the risk of microvascular complications in the diabetes control and complications trial--revisited. Diabetes *57*, 995–1001.

Ladapo, J.A., Kymes, S.M., Ladapo, J.A., Nwosu, V.C., and Pasquale, L.R. (2012). Projected clinical outcomes of glaucoma screening in African American individuals. Arch. Ophthalmol. *130*, 365–372.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Larsson, K., Ljung Björklund, K., Palm, B., Wennberg, M., Kaj, L., Lindh, C.H., Jönsson, B.A.G., and Berglund, M. (2014). Exposure determinants of phthalates, parabens, bisphenol A and triclosan in Swedish mothers and their children. Environ. Int. *73*, 323–333.

Van Leeuwen, R., Klaver, C.C.W., Vingerling, J.R., Hofman, A., and de Jong, P.T.V.M. (2003). The risk and natural course of age-related maculopathy: follow-up at 6 1/2 years in the Rotterdam study. Arch. Ophthalmol. *121*, 519–526.

Van Leeuwen, R., Klaver, C.C.W., Vingerling, J.R., Hofman, A., van Duijn, C.M., Stricker, B.H.C., and de Jong, P.T.V.M. (2004). Cholesterol and age-related macular degeneration: is there a link? Am. J. Ophthalmol. *137*, 750–752.

Leibowitz, H.M., Krueger, D.E., Maunder, L.R., Milton, R.C., Kini, M.M., Kahn, H.A., Nickerson, R.J., Pool, J., Colton, T.L., Ganley, J.P., et al. (1980). The Framingham Eye Study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975. Surv. Ophthalmol. *24*, 335–610.

Leske, M.C., Connell, A.M., Schachat, A.P., and Hyman, L. (1994). The Barbados Eye Study. Prevalence of open angle glaucoma. Arch. Ophthalmol. *112*, 821–829.

Leske, M.C., Connell, A.M., Wu, S.Y., Hyman, L.G., and Schachat, A.P. (1995). Risk factors for open-angle glaucoma. The Barbados Eye Study. Arch. Ophthalmol. *113*, 918–924.

Leske, M.C., Heijl, A., Hyman, L., Bengtsson, B., Dong, L., Yang, Z., and EMGT Group (2007). Predictors of long-term progression in the early manifest glaucoma trial. Ophthalmology *114*, 1965–1972.

Leske, M.C., Wu, S.-Y., Hennis, A., Honkanen, R., Nemesure, B., and BESs Study Group (2008). Risk factors for incident open-angle glaucoma: the Barbados Eye Studies. Ophthalmology *115*, 85–93.

Leslie, R.D., and Pyke, D.A. (1982). Diabetic retinopathy in identical twins. Diabetes *31*, 19–21.

Levine, R.A., Demirel, S., Fan, J., Keltner, J.L., Johnson, C.A., and Kass, M.A. (2006). Asymmetries and Visual Field Summaries as Predictors of Glaucoma in the Ocular Hypertension Treatment Study. Invest. Ophthalmol. Vis. Sci. *47*, 3896–3903.

Liang, F.-Q., and Godley, B.F. (2003). Oxidative stress-induced mitochondrial DNA damage in human retinal pigment epithelial cells: a possible mechanism for RPE aging and age-related macular degeneration. Exp. Eye Res. *76*, 397–403.

Lin, L.L.K., Shih, Y.F., Hsiao, C.K., and Chen, C.J. (2004). Prevalence of myopia in Taiwanese schoolchildren: 1983 to 2000. Ann. Acad. Med. Singapore *33*, 27–33.

Lions Clubs International Foundation, N.E.I. (2007). 2005 Survey of public knowledge, attitudes, and practices related to eye health and disease.

Liu, K., Chen, L.J., Tam, P.O.S., Shi, Y., Lai, T.Y.Y., Liu, D.T.L., Chiang, S.W.Y., Yang, M., Yang, Z., and Pang, C.P. (2013a). Associations of the C2-CFB-RDBP-SKIV2L Locus with Age-related Macular Degeneration and Polypoidal Choroidal Vasculopathy. Ophthalmology *120*, 837–843.

Liu, W., Liu, Y., Challa, P., Herndon, L.W., Wiggs, J.L., Girkin, C.A., Allingham, R.R., and Hauser, M.A. (2012). Low prevalence of myocilin mutations in an African American population with primary open-angle glaucoma. Mol. Vis. *18*, 2241–2246.

Liu, Y., Hauser, M.A., Akafo, S.K., Qin, X., Miura, S., Gibson, J.R., Wheeler, J., Gaasterland, D.E., Challa, P., Herndon, L.W., et al. (2013b). Investigation of known genetic risk factors for primary open angle glaucoma in two populations of African ancestry. Invest. Ophthalmol. Vis. Sci. *54*, 6248–6254.

Livingstone, S.J., Levin, D., Looker, H.C., Lindsay, R.S., Wild, S.H., Joss, N., Leese, G., Leslie, P., McCrimmon, R.J., Metcalfe, W., et al. (2015). Estimated life expectancy in a Scottish cohort with type 1 diabetes, 2008-2010. JAMA *313*, 37–44.

Long, J., Edwards, T., Signorello, L.B., Cai, Q., Zheng, W., Shu, X.-O., and Blot, W.J. (2012). Evaluation of Genome-wide Association Study-identified Type 2 Diabetes Loci in African Americans. Am. J. Epidemiol.

Macgregor, S., Hewitt, A.W., Hysi, P.G., Ruddle, J.B., Medland, S.E., Henders, A.K., Gordon, S.D., Andrew, T., McEvoy, B., Sanfilippo, P.G., et al. (2010). Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. Hum. Mol. Genet. *19*, 2716–2724.

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

Malinowski, J., Farber-Eger, E., and Crawford, D.C. (2014). Development of a data-mining algorithm to identify ages at reproductive milestones in electronic medical records. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 376–387.

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

Matise, T.C., Ambite, J.L., Buyske, S., Carlson, C.S., Cole, S.A., Crawford, D.C., Haiman, C.A., Heiss, G., Kooperberg, C., Marchand, L.L., et al. (2011a). The Next PAGE in Understanding Complex Traits: Design for the Analysis of Population Architecture Using Genetics and Epidemiology (PAGE) Study. Am. J. Epidemiol. *174*, 849–859.

Matise, T.C., Ambite, J.L., Buyske, S., Carlson, C.S., Cole, S.A., Crawford, D.C., Haiman, C.A., Heiss, G., Kooperberg, C., Marchand, L.L., et al. (2011b). The Next PAGE in understanding complex traits: design for the analysis of Population Architecture Using Genetics and Epidemiology (PAGE) Study. Am. J. Epidemiol. *174*, 849–859.

McCarty, C., Chisholm, R., Chute, C., Kullo, I., Jarvik, G., Larson, E., Li, R., Masys, D., Ritchie, M., Roden, D., et al. (2011). The eMERGE Network: A consortium of biorepositories linked to electronic medical records data for conducting genomic studies. BMC Med. Genomics *4*, 13.

McGlynn, M.M., Ehrlich, J.R., Marlow, E.D., Chee, R.-I., Silva, F.Q., Van Tassel, S.H., and Radcliffe, N.M. (2013). Association of blood and ocular perfusion pressure with structural glaucomatous progression by flicker chronoscopy. Br. J. Ophthalmol. *97*, 1569–1573.

Mimivati, Z., Nurliza, K., Marini, M., and Liza-Sharmini, A. (2014). Identification of MYOC gene mutation and polymorphism in a large Malay family with juvenile-onset open angle glaucoma. Mol. Vis. *20*, 714–723.

Mitchell, S.L., Goodloe, R., Brown-Gentry, K., Pendergrass, S.A., Murdock, D.G., and Crawford, D.C. (2014a). Characterization of mitochondrial haplogroups in a large population-based sample from the United States. Hum. Genet.

Mitchell, S.L., Hall, J.B., Goodloe, R.J., Boston, J., Farber-Eger, E., Pendergrass, S.A., Bush, W.S., and Crawford, D.C. (2014b). Investigating the relationship between mitochondrial genetic variation and cardiovascular-related traits to develop a framework for mitochondrial phenome-wide association studies. BioData Min. *7*, 6.

Monemi, S., Spaeth, G., DaSilva, A., Popinchalk, S., Ilitchev, E., Liebmann, J., Ritch, R., Héon, E., Crick, R.P., Child, A., et al. (2005). Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. Hum. Mol. Genet. *14*, 725–733.

Morgan-Davies, J., Taylor, N., Hill, A.R., Aspinall, P., O'Brien, C.J., and Azuara-Blanco, A. (2004). Three dimensional analysis of the lamina cribrosa in glaucoma. Br. J. Ophthalmol. *88*, 1299–1304.

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

Morrison, J.C., Johnson, E.C., Cepurna, W., and Jia, L. (2005). Understanding mechanisms of pressure-induced optic nerve damage. Prog. Retin. Eye Res. *24*, 217–240.

Moss, A.J., Zareba, W., Hall, W.J., Klein, H., Wilber, D.J., Cannom, D.S., Daubert, J.P., Higgins, S.L., Brown, M.W., and Andrews, M.L. (2002). Prophylactic Implantation of a Defibrillator in Patients with Myocardial Infarction and Reduced Ejection Fraction. N. Engl. J. Med. *346*, 877–883.

Mueller, E.E., Schaier, E., Brunner, S.M., Eder, W., Mayr, J.A., Egger, S.F., Nischler, C., Oberkofler, H., Reitsamer, H.A., Patsch, W., et al. (2012). Mitochondrial Haplogroups and Control Region Polymorphisms in Age-Related Macular Degeneration: A Case-Control Study. PLoS ONE *7*.

Munaut, C., Lambert, V., Noël, A., Frankenne, F., Deprez, M., Foidart, J.M., and Rakic, J.M. (2001). Presence of oestrogen receptor type beta in human retina. Br. J. Ophthalmol. *85*, 877–882.

Munch, I.C., Sander, B., Kessel, L., Hougaard, J.L., Taarnhøj, N.C.B.B., Sørensen, T.I.A., Kyvik, K.O., and Larsen, M. (2007). Heredity of small hard drusen in twins aged 20-46 years. Invest. Ophthalmol. Vis. Sci. *48*, 833–838.

Nachat, R., Cipolat, S., Sevilla, L.M., Chhatriwala, M., Groot, K.R., and Watt, F.M. (2009). KazrinE is a desmosome-associated liprin that colocalises with acetylated microtubules. J. Cell Sci. *122*, 4035–4041.

Naj, A.C., Scott, W.K., Courtenay, M.D., Cade, W.H., Schwartz, S.G., Kovach, J.L., Agarwal, A., Wang, G., Haines, J.L., and Pericak-Vance, M.A. (2013). Genetic factors in nonsmokers with age-related macular degeneration revealed through genome-wide gene-environment interaction analysis. Ann. Hum. Genet. *77*, 215–231.

Nakano, M., Ikeda, Y., Tokuda, Y., Fuwa, M., Omi, N., Ueno, M., Imai, K., Adachi, H., Kageyama, M., Mori, K., et al. (2012). Common variants in CDKN2B-AS1 associated with optic-nerve vulnerability of glaucoma identified by genome-wide association studies in Japanese. PloS One *7*, e33389.

Nakata, I., Yamashiro, K., Yamada, R., Gotoh, N., Nakanishi, H., Hayashi, H., Akagi-Kurashige, Y., Tsujikawa, A., Otani, A., Saito, M., et al. (2012). Significance of C2/CFB variants in age-related macular degeneration and polypoidal choroidal vasculopathy in a Japanese population. Invest. Ophthalmol. Vis. Sci. *53*, 794–798.

Nang, E.E.K., Khoo, C.M., Tai, E.S., Lim, S.C., Tavintharan, S., Wong, T.Y., Heng, D., and Lee, J. (2009). Is There a Clear Threshold for Fasting Plasma Glucose That Differentiates Between Those With and Without Neuropathy and Chronic Kidney Disease? The Singapore Prospective Study Program. Am. J. Epidemiol. *169*, 1454–1462.

Narayan K, B.J. (2003). LIfetime risk for diabetes mellitus in the united states. JAMA J. Am. Med. Assoc. *290*, 1884–1890.

National Alliance for Eye and Vision Research (2006). The Silver Book: Vision Loss: Chronic Disease and Medical Innovationn in an Aging nation.

Neale, B.M., Fagerness, J., Reynolds, R., Sobrin, L., Parker, M., Raychaudhuri, S., Tan, P.L., Oh, E.C., Merriam, J.E., Souied, E., et al. (2010). Genome-wide association study of advanced age-related macular degeneration identifies a role of the hepatic lipase gene (LIPC). Proc. Natl. Acad. Sci. U. S. A. *107*, 7395–7400.

Nejad, T.M., Foster, C., and Gongal, D. (2014). Finite element modelling of cornea mechanics: a review. Arq. Bras. Oftalmol. *77*, 60–65.

Newton, K.M., Peissig, P.L., Kho, A.N., Bielinski, S.J., Berg, R.L., Choudhary, V., Basford, M., Chute, C.G., Kullo, I.J., Li, R., et al. (2013). Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. J. Am. Med. Inform. Assoc. JAMIA *20*, e147–e154.

Ng, M.C.Y., Shriner, D., Chen, B.H., Li, J., Chen, W.-M., Guo, X., Liu, J., Bielinski, S.J., Yanek, L.R., Nalls, M.A., et al. (2014). Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. PLoS Genet. *10*, e1004517.

Ng, T.K., Chen, L.J., Liu, D.T.L., Tam, P.O.S., Chan, W.M., Liu, K., Hu, Y.J., Chong, K.K.L., Lau, C.S.L., Chiang, S.W.Y., et al. (2008). Multiple gene polymorphisms in the complement factor h gene are associated with exudative age-related macular degeneration in chinese. Invest. Ophthalmol. Vis. Sci. *49*, 3312–3317.

Nonyane, B.A.S., Nitsch, D., Whittaker, J.C., Sofat, R., Smeeth, L., Chakravarthy, U., and Fletcher, A.E. (2010). An ecological correlation study of late age-related macular degeneration and the complement factor H Y402H polymorphism. Invest. Ophthalmol. Vis. Sci. *51*, 2393–2402.

Norton, B., and Pearson, E.S. (1976). A note on the background to, and refereeing of, R. A. Fisher's 1918 paper "On the correlation between relatives on the supposition of Mendelian inheritance." Notes Rec. R. Soc. Lond. *31*, 151–162.

OCOMM.OCPT Disability Planner: Special Rules For People Who Are Blind Or Have Low Vision.

Oetjens, M., Bush, W.S., Birdwell, K.A., Dilks, H.H., Bowton, E.A., Denny, J.C., Wilke, R.A., Roden, D.M., and Crawford, D.C. (2014). Utilization of an EMR-biorepository to identify the genetic predictors of calcineurin-inhibitor toxicity in heart transplant recipients. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 253–264.

Ojaimi, E., Nguyen, T.T., Klein, R., Islam, F.M.A., Cotch, M.F., Klein, B.E.K., Wang, J.-J., and Wong, T.Y. (2011). Retinopathy signs in people without diabetes: the multi-ethnic study of atherosclerosis. Ophthalmology *118*, 656–662.

Okubo, T., Yokoyama, Y., Kano, K., and Kano, I. (2001). ER-dependent estrogenic activity of parabens assessed by proliferation of human breast cancer MCF-7 cells and expression of ERα and PR. Food Chem. Toxicol. *39*, 1225–1232.

De Oliveira, A.P.C., and Kasahara, N. (2014). Correlation between ocular perfusion pressure fluctuation and glaucoma severity. Int. Ophthalmol.

Ortego, J., Escribano, J., and Coca-Prados, M. (1997). Cloning and characterization of subtracted cDNAs from a human ciliary body library encoding TIGR, a protein involved in juvenile open angle glaucoma with homology to myosin and olfactomedin. FEBS Lett. *413*, 349–353.

Osman, W., Low, S.-K., Takahashi, A., Kubo, M., and Nakamura, Y. (2012). A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. Hum. Mol. Genet. *21*, 2836–2842.

Palmer, N.D., McDonough, C.W., Hicks, P.J., Roh, B.H., Wing, M.R., An, S.S., Hester, J.M., Cooke, J.N., Bostrom, M.A., Rudock, M.E., et al. (2012). A genome-wide association search for type 2 diabetes genes in African Americans. PloS One *7*, e29202.

Pan, C.-W., Cheung, C.Y., Aung, T., Cheung, C.-M., Zheng, Y.-F., Wu, R.-Y., Mitchell, P., Lavanya, R., Baskaran, M., Wang, J.J., et al. (2013). Differential associations of myopia with major age-related eye diseases: the Singapore Indian Eye Study. Ophthalmology *120*, 284–291.

Pan, C.-W., Klein, B.E.K., Cotch, M.F., Shrager, S., Klein, R., Folsom, A., Kronmal, R., Shea, S.J., Burke, G.L., Saw, S.-M., et al. Racial Variations in the Prevalence of Refractive Errors in the United States: The Multi-Ethnic Study of Atherosclerosis. Am. J. Ophthalmol.

Pascolini, D., and Mariotti, S.P. (2012). Global estimates of visual impairment: 2010. Br. J. Ophthalmol. *96*, 614–618.

Pasquale, L.R., Rosner, B.A., Hankinson, S.E., and Kang, J.H. (2007). Attributes of female reproductive aging and their relation to primary open-angle glaucoma: a prospective study. J. Glaucoma *16*, 598–605.

Pasquale, L.R., Loomis, S.J., Weinreb, R.N., Kang, J.H., Yaspan, B.L., Bailey, J.C., Gaasterland, D., Gaasterland, T., Lee, R.K., Scott, W.K., et al. (2013). Estrogen pathway polymorphisms in relation to primary open angle glaucoma: An analysis accounting for gender from the United States. Mol. Vis. *19*, 1471–1481.

Patterson, N., Price, A.L., and Reich, D. (2006). Population Structure and Eigenanalysis. PLoS Genet *2*, e190.

Peissig, P.L., Rasmussen, L.V., Berg, R.L., Linneman, J.G., McCarty, C.A., Waudby, C., Chen, L., Denny, J.C., Wilke, R.A., Pathak, J., et al. (2012). Importance of multi-modal approaches to effectively identify cataract cases from electronic health records. J. Am. Med. Inform. Assoc. JAMIA *19*, 225–234.

Pendergrass, S.A., Dudek, S.M., Crawford, D.C., and Ritchie, M.D. (2010). Synthesis-View: visualization and interpretation of SNP association results for multi-cohort, multi-phenotype data and meta-analysis. BioData Min. *3*, 10.

Polansky, J.R., Fauss, D.J., Chen, P., Chen, H., Lütjen-Drecoll, E., Johnson, D., Kurtz, R.M., Ma, Z.D., Bloom, E., and Nguyen, T.D. (1997). Cellular pharmacology and molecular biology of the trabecular meshwork inducible glucocorticoid response gene product. Ophthalmol. J. Int. Ophtalmol. Int. J. Ophthalmol. Z. Für Augenheilkd. *211*, 126–139.

Profirovic, J., Strekalova, E., Urao, N., Krbanjevic, A., Andreeva, A., Varadarajan, S., Fukai, T., Hen, R., Ushio-Fukai, M., and Voyno-Yasenetskaya, T.A. (2013). A novel regulator of angiogenesis in endothelial cells: 5-hydroxytriptamine 4 receptor. Angiogenesis *16*, 15–28.

Psaty, B.M., O'Donnell, C.J., Gudnason, V., Lunetta, K.L., Folsom, A.R., Rotter, J.I., Uitterlinden, A.G., Harris, T.B., Witteman, J.C.M., and Boerwinkle, E. (2009). Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. Circ. Cardiovasc. Genet. *2*, 73–80.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

Quigley, H.A., and Broman, A.T. (2006). The number of people with glaucoma worldwide in 2010 and 2020. Br. J. Ophthalmol. *90*, 262–267.

Quigley, H.A., and Vitale, S. (1997). Models of open-angle glaucoma prevalence and incidence in the United States. Invest. Ophthalmol. Vis. Sci. *38*, 83–91.

Quigley, H.A., West, S.K., Rodriguez, J., Munoz, B., Klein, R., and Snyder, R. (2001). The prevalence of glaucoma in a population-based study of Hispanic subjects: Proyecto VER. Arch. Ophthalmol. *119*, 1819–1826.

Ramdas, W.D., van Koolwijk, L.M.E., Ikram, M.K., Jansonius, N.M., de Jong, P.T.V.M., Bergen, A.A.B., Isaacs, A., Amin, N., Aulchenko, Y.S., Wolfs, R.C.W., et al. (2010). A genome-wide association study of optic disc parameters. PLoS Genet. *6*, e1000978.

Ramirez, A.H., Shi, Y., Schildcrout, J.S., Delaney, J.T., Xu, H., Oetjens, M.T., Zuvich, R.L., Basford, M.A., Bowton, E., Jiang, M., et al. (2012). Predicting warfarin dosage in European-Americans and African-Americans using DNA samples linked to an electronic health record. Pharmacogenomics *13*, 407–418.

Rasmussen-Torvik, L.J., Pacheco, J.A., Wilke, R.A., Thompson, W.K., Ritchie, M.D., Kho, A.N., Muthalagu, A., Hayes, M.G., Armstrong, L.L., Scheftner, D.A., et al. (2012). High density GWAS for LDL cholesterol in African Americans using electronic medical records reveals a strong protective variant in APOE. Clin. Transl. Sci. *5*, 394–399.

Reidy, A., Minassian, D.C., Vafidis, G., Joseph, J., Farrow, S., Wu, J., Desai, P., and Connolly, A. (1998). Prevalence of serious eye disease and visual impairment in a north London population: population based, cross sectional study. BMJ *316*, 1643–1646.

Rein, D.B., Zhang, P., Wirth, K.E., Lee, P.P., Hoerger, T.J., McCall, N., Klein, R., Tielsch, J.M., Vijan, S., and Saaddine, J. (2006). The economic burden of major adult visual disorders in the United States. Arch. Ophthalmol. *124*, 1754–1760.

Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G.P., and Mariotti, S.P. (2004). Global data on visual impairment in the year 2002. Bull. World Health Organ. *82*, 844–851.

Restrepo, N.A., Spencer, K.M., Goodloe, R., Garrett, T.A., Heiss, G., Bůžková, P., Jorgensen, N., Jensen, R.A., Matise, T.C., Hindorff, L.A., et al. (2014). Genetic determinants of age-related macular degeneration in diverse populations from the PAGE Study. Invest. Ophthalmol. Vis. Sci.

Restrepo, N.A., Mitchell, S.L., Goodloe, R.J., Murdock, D.G., Haines, J.L., and Crawford, D.C. (2015). Mitochondrial variation and the risk of age-related macular degeneration across diverse populations. Pac. Symp. Biocomput. Pac. Symp. Biocomput. *20*, 243–254.

Reynolds, R., Rosner, B., and Seddon, J.M. (2010). Serum lipid biomarkers and hepatic lipase gene associations with age-related macular degeneration. Ophthalmology *117*, 1989–1995.

Rezaie, T., Child, A., Hitchings, R., Brice, G., Miller, L., Coca-Prados, M., Héon, E., Krupin, T., Ritch, R., Kreutzer, D., et al. (2002). Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science *295*, 1077–1079.

Richardson, A.J., Islam, F.M.A., Guymer, R.H., and Baird, P.N. (2009). Analysis of Rare Variants in the Complement Component 2 (C2) and Factor B (BF) Genes Refine Association for Age-Related Macular Degeneration (AMD). Invest. Ophthalmol. Vis. Sci. *50*, 540–543.

Ridker, P.M., Paré, G., Parker, A.N., Zee, R.Y.L., Miletich, J.P., and Chasman, D.I. (2009). Polymorphism in the CETP gene region, HDL cholesterol, and risk of future myocardial infarction: Genomewide analysis among 18 245 initially healthy women from the Women's Genome Health Study. Circ. Cardiovasc. Genet. *2*, 26–33.

Ritchie, M.D., Denny, J.C., Crawford, D.C., Ramirez, A.H., Weiner, J.B., Pulley, J.M., Basford, M.A., Brown-Gentry, K., Balser, J.R., Masys, D.R., et al. (2010). Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. Am. J. Hum. Genet. *86*, 560–572.

Ritchie, M.D., Denny, J.C., Zuvich, R.L., Crawford, D.C., Schildcrout, J.S., Bastarache, L., Ramirez, A.H., Mosley, J.D., Pulley, J.M., Basford, M.A., et al. (2013). Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. Circulation *127*, 1377–1385.

Rivera, A., Fisher, S.A., Fritsche, L.G., Keilhauer, C.N., Lichtner, P., Meitinger, T., and Weber, B.H.F. (2005). Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum. Mol. Genet. *14*, 3227–3236.

Roden, D.M., Pulley, J.M., Basford, M.A., Bernard, G.R., Clayton, E.W., Balser, J.R., and Masys, D.R. (2008). Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin. Pharmacol. Ther. *84*, 362–369.

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

Roy, M.S., Klein, R., O'Colmain, B.J., Klein, B.E.K., Moss, S.E., and Kempen, J.H. (2004). The prevalence of diabetic retinopathy among adult type 1 diabetic persons in the United States. Arch. Ophthalmol. *122*, 546–551.

Rudnicka, A.R., Mt-Isa, S., Owen, C.G., Cook, D.G., and Ashby, D. (2006). Variations in primary open-angle glaucoma prevalence by age, gender, and race: a Bayesian meta-analysis. Invest. Ophthalmol. Vis. Sci. *47*, 4254–4261.

Rudolf, M., Malek, G., Messinger, J.D., Clark, M.E., Wang, L., and Curcio, C.A. (2008). Sub-retinal drusenoid deposits in human retina: organization and composition. Exp. Eye Res. *87*, 402–408.

Ryu, E., Fridley, B.L., Tosakulwong, N., Bailey, K.R., and Edwards, A.O. (2010). Genome-wide association analyses of genetic, phenotypic, and environmental risks in the age-related eye disease study. Mol. Vis. *16*, 2811–2821.

Sabatti, C., Service, S.K., Hartikainen, A.-L., Pouta, A., Ripatti, S., Brodsky, J., Jones, C.G., Zaitlen, N.A., Varilo, T., Kaakinen, M., et al. (2009). Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nat. Genet. *41*, 35–46.

Sadigh, S., Luo, X., Cideciyan, A.V., Sumaroka, A., Boxley, S.L., Hall, L.M., Sheplock, R., Feuer, W.J., Stambolian, D.S., and Jacobson, S.G. (2014). Drusen and Photoreceptor Abnormalities in African-Americans with Intermediate Non-neovascular Age-related Macular Degeneration. Curr. Eye Res. 1–9.

SanGiovanni, J.P., Arking, D.E., Iyengar, S.K., Elashoff, M., Clemons, T.E., Reed, G.F., Henning, A.K., Sivakumaran, T.A., Xu, X., DeWan, A., et al. (2009). Mitochondrial DNA variants of respiratory complex I that uniquely characterize haplogroup T2 are associated with increased risk of age-related macular degeneration. PloS One *4*, e5508.

Schachar, R.A. (2006). The mechanism of accommodation and presbyopia. Int. Ophthalmol. Clin. *46*, 39–61.

Schache, M., and Baird, P.N. (2013). The Australian twin registry as a resource for genetic studies into ophthalmic traits. Twin Res. Hum. Genet. Off. J. Int. Soc. Twin Stud. *16*, 52–57.

Schmidl, D., Schmetterer, L., Garhöfer, G., and Popa-Cherecheanu, A. (2015). Gender differences in ocular blood flow. Curr. Eye Res. *40*, 201–212.

Schmidt, S., Hauser, M.A., Scott, W.K., Postel, E.A., Agarwal, A., Gallins, P., Wong, F., Chen, Y.S., Spencer, K., Schnetz-Boutaud, N., et al. (2006). Cigarette smoking strongly modifies the association of LOC387715 and age-related macular degeneration. Am. J. Hum. Genet. *78*, 852–864.

Scientific Committee on Consumer Safety Europe (2010). Opinion on triclosan (antimicrobial restistance).

Scott, L.J., Mohlke, K.L., Bonnycastle, L.L., Willer, C.J., Li, Y., Duren, W.L., Erdos, M.R., Stringham, H.M., Chines, P.S., Jackson, A.U., et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science *316*, 1341–1345.

Seddon, J.M., Cote, J., Page, W.F., Aggen, S.H., and Neale, M.C. (2005). The US Twin Study of Age-Related Macular Degeneration: Relative Roles of Genetic and Environmental Influences. Arch Ophthalmol *123*, 321–327.

Shah, S., Chatterjee, A., Mathai, M., Kelly, S.P., Kwartz, J., Henson, D., and McLeod, D. (1999). Relationship between corneal thickness and measured intraocular pressure in a general ophthalmology clinic. Ophthalmology *106*, 2154–2160.

Shalev, V., Sror, M., Goldshtein, I., Kokia, E., and Chodick, G. (2011). Statin use and the risk of age related macular degeneration in a large health organization in Israel. Ophthalmic Epidemiol. *18*, 83–90.

Sheu, S.-J., Liu, N.-C., Ou, C.-C., Bee, Y.-S., Chen, S.-C., Lin, H.-C., and Chan, J.Y.H. (2013a). Resveratrol stimulates mitochondrial bioenergetics to protect retinal pigment epithelial cells from oxidative damage. Invest. Ophthalmol. Vis. Sci. *54*, 6426–6438.

Sheu, W.H.-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.K., et al. (2013b). Genome-wide association study in a Chinese population with diabetic retinopathy. Hum. Mol. Genet. *22*, 3165–3173.

Siatkowski, R.M., Cotter, S.A., Crockett, R.S., Miller, J.M., Novack, G.D., and Zadnik, K. (2008). Two-year multicenter, randomized, double-masked, placebo-controlled, parallel safety and efficacy study of 2% pirenzepine ophthalmic gel in children with myopia. J. AAPOS Off. Publ. Am. Assoc. Pediatr. Ophthalmol. Strabismus Am. Assoc. Pediatr. Ophthalmol. Strabismus *12*, 332–339.

Sim, X., Ong, R.T.-H., Suo, C., Tay, W.-T., Liu, J., Ng, D.P.-K., Boehnke, M., Chia, K.-S., Wong, T.-Y., Seielstad, M., et al. (2011). Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. *7*, e1001363.

Smith, W., Mitchell, P., and Wang, J. (1997). Gender, oestrogen, hormone replacement and age-related macular degeneration: Results from the Blue Mountains Eye Study. Aust. N. Z. J. Ophthalmol. *25*, 13–15.

Smythies, J. (1996). A note on the concept of the visual field in neurology, psychology, and visual neuroscience. Perception *25*, 369–371.

Sobrin, L., Ripke, S., Yu, Y., Fagerness, J., Bhangale, T.R., Tan, P.L., Souied, E.H., Buitendijk, G.H.S., Merriam, J.E., Richardson, A.J., et al. (2012). Heritability and Genome-Wide Association Study to Assess Genetic Differences between Advanced Age-related Macular Degeneration Subtypes. Ophthalmology *119*, 1874–1885.

Spencer, K.L., Hauser, M.A., Olson, L.M., Schmidt, S., Scott, W.K., Gallins, P., Agarwal, A., Postel, E.A., Pericak-Vance, M.A., and Haines, J.L. (2007). Protective effect of complement factor B and complement component 2 variants in age-related macular degeneration. Hum. Mol. Genet. *16*, 1986–1992.

Spencer, K.L., Glenn, K., Brown-Gentry, K., Haines, J.L., and Crawford, D.C. (2012a). Population differences in genetic risk for age-related macular degeneration and implications for genetic testing. Arch. Ophthalmol. *130*, 116–117.

Spencer, K.L., Glenn, K., Brown-Gentry, K., Haines, J.L., and Crawford, D.C. (2012b). Population differences in genetic risk for age-related macular degeneration and implications for genetic testing. Arch. Ophthalmol. *130*, 116–117.

Steck, A.K., Vehik, K., Bonifacio, E., Lernmark, A., Ziegler, A.-G., Hagopian, W.A., She, J., Simell, O., Akolkar, B., Krischer, J., et al. (2015). Predictors of Progression From the Appearance of Islet Autoantibodies to Early Childhood Diabetes: The Environmental Determinants of Diabetes in the Young (TEDDY). Diabetes Care.

Stein, J.D., Vanderbeek, B.L., Talwar, N., Nan, B., Musch, D.C., and Zacks, D.N. (2011a). Rates of nonexudative and exudative age-related macular degeneration among Asian American ethnic groups. Invest. Ophthalmol. Vis. Sci. *52*, 6842–6848.

Stein, J.D., Kim, D.S., Niziol, L.M., Talwar, N., Nan, B., Musch, D.C., and Richards, J.E. (2011b). Differences in rates of glaucoma among Asian Americans and other racial groups, and among various Asian ethnic groups. Ophthalmology *118*, 1031–1037.

Steinberg, K.K., Sanderlin, K.C., Ou, C.Y., Hannon, W.H., McQuillan, G.M., and Sampson, E.J. (1997). DNA banking in epidemiologic studies. Epidemiol. Rev. *19*, 156–162.

Steinmetz, R., Brown, N.G., Allen, D.L., Bigsby, R.M., and Ben-Jonathan, N. (1997). The environmental estrogen bisphenol A stimulates prolactin release in vitro and in vivo. Endocrinology *138*, 1780–1786.

Stone, E.M., Fingert, J.H., Alward, W.L., Nguyen, T.D., Polansky, J.R., Sunden, S.L., Nishimura, D., Clark, A.F., Nystuen, A., Nichols, B.E., et al. (1997). Identification of a gene that causes primary open angle glaucoma. Science *275*, 668–670.

Sudchada, P., and Scarpace, K. (2014). Transcription factor 7-like 2 polymorphisms and diabetic retinopathy: a systematic review. Genet. Mol. Res. GMR *13*, 5865–5872.

Suzuki, Y., Shirato, S., Taniguchi, F., Ohara, K., Nishimaki, K., and Ohta, S. (1997). Mutations in the TIGR gene in familial primary open-angle glaucoma in Japan. Am. J. Hum. Genet. *61*, 1202–1204.

Tamm, E.R., Russell, P., Epstein, D.L., Johnson, D.H., and Piatigorsky, J. (1999). Modulation of myocilin/TIGR expression in human trabecular meshwork. Invest. Ophthalmol. Vis. Sci. *40*, 2577–2582.

Tan, C.E., Emmanuel, S.C., Tan, B.Y., and Jacob, E. (1999). Prevalence of diabetes and ethnic differences in cardiovascular risk factors. The 1992 Singapore National Health Survey. Diabetes Care *22*, 241–247.

Tan, J.S.L., Mitchell, P., Rochtchina, E., and Wang, J.J. (2007a). Statins and the long-term risk of incident age-related macular degeneration: the Blue Mountains Eye Study. Am. J. Ophthalmol. *143*, 685–687.

Tan, J.S.L., Mitchell, P., Smith, W., and Wang, J.J. (2007b). Cardiovascular risk factors and the long-term incidence of age-related macular degeneration: the Blue Mountains Eye Study. Ophthalmology *114*, 1143–1150.

Tan, J.S.L., Wang, J.J., Flood, V., Rochtchina, E., Smith, W., and Mitchell, P. (2008). Dietary antioxidants and the long-term incidence of age-related macular degeneration: the Blue Mountains Eye Study. Ophthalmology *115*, 334–341.

Tanaka, K., Nakayama, T., Yuzawa, M., Wang, Z., Kawamura, A., Mori, R., Nakashizuka, H., Sato, N., and Mizutani, Y. (2011). Analysis of candidate genes for age-related macular degeneration subtypes in the Japanese population. Mol. Vis. *17*, 2751–2758.

Tedeschi-Blok, N., Buckley, J., Varma, R., Triche, T.J., and Hinton, D.R. (2007). Population-based study of early age-related macular degeneration: role of the complement factor H Y402H polymorphism in bilateral but not unilateral disease. Ophthalmology *114*, 99–103.

Teo, Y.-Y., Small, K.S., and Kwiatkowski, D.P. (2010). Methodological challenges of genome-wide association analysis in Africa. Nat. Rev. Genet. *11*, 149–160.

Thakkinstian, A., McEvoy, M., Chakravarthy, U., Chakrabarti, S., McKay, G.J., Ryu, E., Silvestri, G., Kaur, I., Francis, P., Iwata, T., et al. (2012). The association between complement component 2/complement factor B polymorphisms and age-related macular degeneration: a HuGE review and meta-analysis. Am. J. Epidemiol. *176*, 361–372.

The Eye Diseases Prevalence Research Group (2004). The Prevalence of Diabetic Retinopathy Among Adults in the United States. Arch Ophthalmol *122*, 552–563.

Thorleifsson, G., Walters, G.B., Hewitt, A.W., Masson, G., Helgason, A., DeWan, A., Sigurdsson, A., Jonasdottir, A., Gudjonsson, S.A., Magnusson, K.P., et al. (2010). Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. Nat. Genet. *42*, 906–909.

Tian, J., Yu, W., Qin, X., Fang, K., Chen, Q., Hou, J., Li, J., Chen, D., Hu, Y., and Li, X. (2012). Association of genetic polymorphisms and age-related macular degeneration in chinese population. Invest. Ophthalmol. Vis. Sci. *53*, 4262–4269.

Tielsch, J.M., Katz, J., Sommer, A., Quigley, H.A., and Javitt, J.C. (1994). Family history and risk of primary open angle glaucoma. The Baltimore Eye Survey. Arch. Ophthalmol. *112*, 69–73.

Tikellis, G., Sun, C., Gorin, M.B., Klein, R., Klein, B.E.K., Larsen, E.K.M., Siscovick, D.S., Hubbard, L.D., and Wong, T.Y. (2007). Apolipoprotein e gene and age-related maculopathy in older individuals: the cardiovascular health study. Arch. Ophthalmol. *125*, 68–73.

Tilleul, J., Richard, F., Puche, N., Zerbib, J., Leveziel, N., Sahel, J.A., Cohen, S.Y., Korobelnik, J.-F., Feingold, J., Munnich, A., et al. (2013). Genetic association study of mitochondrial polymorphisms in neovascular age-related macular degeneration. Mol. Vis. *19*, 1132–1140.

Tomany, S.C., Wang, J.J., Van Leeuwen, R., Klein, R., Mitchell, P., Vingerling, J.R., Klein, B.E.K., Smith, W., and De Jong, P.T.V.M. (2004). Risk factors for incident age-related macular degeneration: pooled findings from 3 continents. Ophthalmology *111*, 1280–1287.

Tserentsoodol, N., Gordiyenko, N.V., Pascual, I., Lee, J.W., Fliesler, S.J., and Rodriguez, I.R. (2006). Intraretinal lipid transport is dependent on high density lipoprotein-like particles and class B scavenger receptors. Mol. Vis. *12*, 1319–1333.

Tsugawa, Y., Mukamal, K.J., Davis, R.B., Taylor, W.C., and Wee, C.C. (2012). Should the hemoglobin A(1c) diagnostic cutoff differ between blacks and whites?: a cross-sectional study. Ann. Intern. Med. *157*, 153–159.

Twigg, J., and Majima, S. (2014). Consumption and the constitution of age: expenditure patterns on clothing, hair and cosmetics among post-war “baby boomers.” J. Aging Stud. *30*, 23–32.

Udar, N., Atilano, S.R., Memarzadeh, M., Boyer, D.S., Chwa, M., Lu, S., Maguen, B., Langberg, J., Coskun, P., Wallace, D.C., et al. (2009). Mitochondrial DNA Haplogroups Associated with Age-Related Macular Degeneration. Invest. Ophthalmol. Vis. Sci. *50*, 2966–2974.

Ulrich, J.N., Spannagl, M., Kampik, A., and Gandorfer, A. (2008). Components of the fibrinolytic system in the vitreous body in patients with vitreoretinal disorders. Clin. Experiment. Ophthalmol. *36*, 431–436.

Vajaranant, T.S., Wu, S., Torres, M., and Varma, R. (2012a). The Changing Face of Primary Open-Angle Glaucoma in the United States: Demographic and Geographic Changes from 2011-2050. Am. J. Ophthalmol. *154*, 303–314.e3.

Vajaranant, T.S., Wu, S., Torres, M., and Varma, R. (2012b). A 40-Year Forecast of the Demographic Shift in Primary Open-Angle Glaucoma in the United States. Invest. Ophthalmol. Vis. Sci. *53*, 2464–2466.

Varma, R., Steinmann, W.C., and Scott, I.U. (1992). Expert agreement in evaluating the optic disc for glaucoma. Ophthalmology *99*, 215–221.

Venter, J.C., 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., et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Vidal-Sanz, M., Salinas-Navarro, M., Nadal-Nicolás, F.M., Alarcón-Martínez, L., Valiente-Soriano, F.J., de Imperial, J.M., Avilés-Trigueros, M., Agudo-Barriuso, M., and Villegas-Pérez, M.P. (2012). Understanding glaucomatous damage: anatomical and functional data from ocular hypertensive rodent retinas. Prog. Retin. Eye Res. *31*, 1–27.

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

Vitale S, E.L. (2008). PRevalence of refractive error in the united states, 1999-2004. Arch. Ophthalmol. *126*, 1111–1119.

Vitale S, S.R. (2009). INcreased prevalence of myopia in the united states between 1971-1972 and 1999-2004. Arch. Ophthalmol. *127*, 1632–1639.

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

Voight, B.F., Kang, H.M., Ding, J., Palmer, C.D., Sidore, C., Chines, P.S., Burtt, N.P., Fuchsberger, C., Li, Y., Erdmann, J., et al. (2012). The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. *8*, e1002793.

Volcik, K.A., Ballantyne, C.M., Braun, M.C., Coresh, J., Mosley, T.H., and Boerwinkle, E. (2008). Association of the complement factor H Y402H polymorphism with cardiovascular disease is dependent upon hypertension status: The ARIC study. Am. J. Hypertens. *21*, 533–538.

Wang, Z., and Li, H. (2004). [Sample size requirements for association studies on gene-gene interaction in case-control study]. Zhonghua Liu Xing Bing Xue Za Zhi Zhonghua Liuxingbingxue Zazhi *25*, 623–626.

Wang, G., Spencer, K.L., Scott, W.K., Whitehead, P., Court, B.L., Ayala-Haedo, J., Mayo, P., Schwartz, S.G., Kovach, J.L., Gallins, P., et al. (2010). Analysis of the indel at the ARMS2 3'UTR in age-related macular degeneration. Hum. Genet. *127*, 595–602.

Wang, J.J., Buitendijk, G.H.S., Rochtchina, E., Lee, K.E., Klein, B.E.K., van Duijn, C.M., Flood, V.M., Meuer, S.M., Attia, J., Myers, C., et al. (2013). Genetic Susceptibility, Dietary Antioxidants, and Long-term Incidence of Age-related Macular Degeneration in Two Populations. Ophthalmology.

Wang, Q., Liu, M., Li, X., Chen, L., and Tang, H. (2009). Kazrin F is involved in apoptosis and interacts with BAX and ARC. Acta Biochim. Biophys. Sin. *41*, 763–772.

Wang, S.Y., Singh, K., and Lin, S.C. (2012). Prevalence and predictors of depression among participants with glaucoma in a nationally representative population sample. Am. J. Ophthalmol. *154*, 436–444.e2.

Waryah, A.M., Narsani, A.K., Sheikh, S.A., Shaikh, H., and Shahani, M.Y. (2013). The novel heterozygous Thr377Arg MYOC mutation causes severe Juvenile Open Angle Glaucoma in a large Pakistani family. Gene *528*, 356–359.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., et al. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. *42*, D1001–D1006.

Wendler, D., Kington, R., Madans, J., Wye, G.V., Christ-Schmidt, H., Pratt, L.A., Brawley, O.W., Gross, C.P., and Emanuel, E. (2006). Are Racial and Ethnic Minorities Less Willing to Participate in Health Research? PLoS Med. *3*.

Wickham, L.A., Gao, J., Toda, I., Rocha, E.M., Ono, M., and Sullivan, D.A. (2000). Identification of androgen, estrogen and progesterone receptor mRNAs in the eye. Acta Ophthalmol. Scand. *78*, 146–153.

Wiggs, J.L., Yaspan, B.L., Hauser, M.A., Kang, J.H., Allingham, R.R., Olson, L.M., Abdrabou, W., Fan, B.J., Wang, D.Y., Brodeur, W., et al. (2012). Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. PLoS Genet. *8*, e1002654.

Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath, S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M., et al. (2008a). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat. Genet. *40*, 161–169.

Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath, S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M., et al. (2008b). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat. Genet. *40*, 161–169.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinforma. Oxf. Engl. *26*, 2190–2191.

Wolfs, R.C.W., Borger, P.H., Ramrattan, R.S., Klaver, C.C.W., Hulsman, C.A.A., Hofman, A., Vingerling, J.R., Hitchings, R.A., and Jong, P.T.V.M. de (2000). Changing Views on Open-Angle Glaucoma: Definitions and Prevalences—The Rotterdam Study. Invest. Ophthalmol. Vis. Sci. *41*, 3309–3321.

Wong, T.Y., Foster, P.J., Hee, J., Ng, T.P., Tielsch, J.M., Chew, S.J., Johnson, G.J., and Seah, S.K. (2000). Prevalence and risk factors for refractive errors in adult Chinese in Singapore. Invest. Ophthalmol. Vis. Sci. *41*, 2486–2494.

WONG, T.Y., KLEIN, R., ISLAM, F.M.A., COTCH, M.F., FOLSOM, A.R., KLEIN, B.E.K., SHARRETT, A.R., and SHEA, S. (2006). Diabetic Retinopathy in a Multi-ethnic Cohort in the United States. Am. J. Ophthalmol. *141*, 446–455.

Woo, W.W., Lim, K.A., Yang, H., Lim, X.Y., Liew, F., Lee, Y.S., and Saw, S.M. (2004). Refractive errors in medical students in Singapore. Singapore Med. J. *45*, 470–474.

Writing Group for the DCCT/EDIC Research Group, Orchard, T.J., Nathan, D.M., Zinman, B., Cleary, P., Brillon, D., Backlund, J.-Y.C., and Lachin, J.M. (2015). Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. JAMA *313*, 45–53.

Wu, L., Tao, Q., Chen, W., Wang, Z., Song, Y., Sheng, S., Li, P., and Zhou, J. (2013). Association between polymorphisms of complement pathway genes and age-related macular degeneration in a Chinese population. Invest. Ophthalmol. Vis. Sci. *54*, 170–174.

Yang, Z., Camp, N.J., Sun, H., Tong, Z., Gibbs, D., Cameron, D.J., Chen, H., Zhao, Y., Pearson, E., Li, X., et al. (2006). A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. Science *314*, 992–993.

Yu, D.Y., and Cringle, S.J. (2001). Oxygen distribution and consumption within the retina in vascularised and avascular retinas and in animal models of retinal disease. Prog. Retin. Eye Res. *20*, 175–208.

Yu, Y., Bhangale, T.R., Fagerness, J., Ripke, S., Thorleifsson, G., Tan, P.L., Souied, E.H., Richardson, A.J., Merriam, J.E., Buitendijk, G.H.S., et al. (2011). Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. Hum. Mol. Genet. *20*, 3699–3709.

Zarkada, G., Heinolainen, K., Makinen, T., Kubota, Y., and Alitalo, K. (2015). VEGFR3 does not sustain retinal angiogenesis without VEGFR2. Proc. Natl. Acad. Sci. U. S. A. *112*, 761–766.

Zeggini, E., Weedon, M.N., Lindgren, C.M., Frayling, T.M., Elliott, K.S., Lango, H., Timpson, N.J., Perry, J.R.B., Rayner, N.W., Freathy, R.M., et al. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science *316*, 1336–1341.

Zhang, J., Guo, Z., Bai, Y., Cui, L., Zhang, S., and Xu, J. (2013). Identification of sequence polymorphisms in the displacement loop region of mitochondrial DNA as a risk factor for renal cell carcinoma. Biomed. Rep. *1*, 563–566.

Ziskind, A., Bardien, S., van der Merwe, L., and Webster, A.R. (2008). The frequency of the H402 allele of CFH and its involvement with age-related maculopathy in an aged Black African Xhosa population. Ophthalmic Genet. *29*, 117–119.

(1987). Diabetes Control and Complications Trial (DCCT): Results of Feasibility Study. The DCCT Research Group. Diabetes Care *10*, 1–19.

(1989). The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am. J. Epidemiol. *129*, 687–702.

(2000a). Risk factors associated with age-related macular degeneration. A case-control study in the age-related eye disease study: Age-Related Eye Disease Study Report Number 3. Ophthalmology *107*, 2224–2232.

(2000b). The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration.The AGIS Investigators. Am. J. Ophthalmol. *130*, 429–440.

(2003). The International HapMap Project. Nature *426*, 789–796.

(2004). Centers for Disease Control and Prevention Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94 Bethesda, MD.

CDC - About Vision Health - Common Eye Disorders - Vision Health Initiative (VHI).

FastStats.

CDC - 2011 National Diabetes Fact Sheet - Publications - Diabetes DDT.

CDC - Just the Facts - National Agenda for Public Health Action - Publications - Diabetes DDT.

AKT3 v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma) [Homo sapiens] - Gene - NCBI.

US Department of Health and Human Services (DHHS) (1996). Centers for Disease Control and Prevention Third National Health and Nutrition Examination Survey, 1988-94, Plan and Operations Procedures Manuals[CD-ROM]National Center for Health Statistics, Centers for Disease Control and Prevention,Hyattsville, MD.