The Genetics of Ocular Endophenotypes of Primary Open Angle Glaucoma in Latinos

BY

DREW ROBERT NANNINI
B.S., University of Illinois at Urbana-Champaign, 2011
M.A., Kansas City University of Medicine and Biosciences, 2012
M.S., University of Illinois at Chicago, 2015

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Public Health Sciences
in the Graduate College of the
University of Illinois at Chicago, 2018

Chicago, Illinois

Defense Committee:

Supriya Mehta, Chair and Academic Advisor, Epidemiology and Biostatistics
Xiaoyi Raymond Gao, Thesis Advisor, Ophthalmology
Maria Argos, Epidemiology and Biostatistics
Ronald Hershow, Epidemiology and Biostatistics
Leslie Stayner, Epidemiology and Biostatistics

ACKNOWLEDGMENTS

I would first like to thank my committee members for their guidance throughout the development of this dissertation. I would like to especially thank Dr. Mehta for being my advisor the past five years and for her continuous words of encouragement and dedication to allow me to succeed as a graduate student. I would also like to thank Dr. Gao who has provided me with countless opportunities to grow as a researcher in his lab and for being a source of inspiration for crafting and executing quality research. I am also thankful for the research funding I received through the Mexican American Glaucoma Genetic Study that enabled me to start and finish this dissertation.

I would like to thank the study subjects who participated in this research for contributing their time to advance the fields of public health, genetics, and ophthalmology. I would also like to thank all of the staff members that aided in the collection and processing of the data used for these studies.

And lastly, I would like to thank my father, Kenneth Nannini, and my wife, Christina Nannini, for their unconditional love and support during my graduate education at the University of Illinois at Chicago.

CONTRIBUTION OF AUTHORS

Chapter 1 provides a brief introduction into the pathogenesis of glaucoma, genetics, and the Latino population, as well as highlights the significance of this dissertation. Chapter 2 is a literature review of the current landscape of the field of ocular genetics, with regard to the research questions asked, and provides context to the gaps in the literature. Chapter 3 describes the study population, genotype and phenotype data, and statistical methods to be used to answer the research questions. Chapter 4 presents findings for each of the research questions. The first analysis under this section has previously been published, for which I was first author. My research mentor, Dr. Xiaoyi Gao, generated the ancestral estimates for each study participant and aided in the interpretation of the results. I performed the statistical analyses and wrote up the manuscript. The second analysis under this section has previously been published, for which I was first author. My research mentor, Dr. Xiaoyi Gao, guided me through the analysis and aided in my conceptual understanding of the statistical methods. I performed the data analysis and wrote up the manuscript. The third analysis is unpublished work I conducted on my own. It is my intention this analysis will be published in the future as a co-authored manuscript. The fourth analysis has been accepted, for which I am first author. My research mentor, Dr. Xiaoyi Gao, helped me develop the idea for this analysis and provided constructive feedback. I performed the analysis and wrote up the manuscript. In Chapter 5, I present a summary of my findings, the public health implications of this dissertation, and future directions to extend this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AUC	Area under the curve
BMI	Body mass index
CCT	Central corneal thickness
CI	Confidence interval
CSF	Cerebrospinal fluid
EMMAX	Efficient mixed-model association expedited
ERF	Erasmus Rucphen Family
G×E	Gene-environment
GRS	Genetic risk score
GWAS	Genome-wide association study
IOP	Intraocular pressure
IPA	Ingenuity Pathway Analysis
LALES	Los Angeles Latino Eye Study
MAF	Minor allele frequency
MAGGS	Mexican American Glaucoma Genetic Study
mmHg	Millimeter of mercury
NA	Native American
NTG	Normal tension glaucoma
OR	Odds ratio
PACG	Primary angle closure glaucoma
POAG	Primary open angle glaucoma
PRIMUS	Pedigree Reconstruction and Identification of a Maximum Unrelated Set
Q-Q	Quantile-quantile
ROC	Receiver operating characteristic
RS	Rotterdam Study

LIST OF ABBREVIATIONS (continued)

SBP	Systolic blood pressure
SD	Standard deviation
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
VCDR	Vertical cup-disc ratio

SUMMARY

Population-based cross-sectional studies were conducted to evaluate the associations between genetic factors and quantitative traits of primary open angle glaucoma in a Latino population. Study participants received detailed ophthalmic and clinical examinations and a blood draw at the baseline visit. Analyses were performed to examine the association between genetic ancestry and intraocular pressure, to identify genetic variants associated with vertical cup-disc ratio, to identify gene-environment interactions between body mass index and genetic variants on vertical cup-disc ratio, and to construct and evaluate a genetic risk score for vertical cup-disc ratio.

African ancestry was significantly associated with intraocular pressure, with increasing proportions of African ancestry associated with increasing intraocular pressure in Latinos. A significant interaction between African ancestry and elevated systolic blood pressure was observed, with individuals with elevated systolic blood pressure experiencing a greater increase in intraocular pressure with increasing African ancestry.

Additionally, a novel genetic variant associated with vertical cup-disc ratio was identified in this Latino study sample. Previously identified genetic variants associated with vertical cup-disc ratio were also replicated.

Through gene-environment interaction analyses, several suggestive interactive associations between body mass index and genetic variants on vertical cup-disc ratio were identified. These associations represent biologically plausible candidate genomic regions for further investigation.

And lastly, significant associations between genetic risk scores and vertical cup-disc ratio were observed with higher genetic risk scores associated with larger vertical cup-disc ratio. Moreover, these genetic risk scores improved the discriminatory ability for primary open angle glaucoma.

Together, these findings address the current gaps in the literature regarding the genetic factors and biological mechanisms underlying the pathogenesis of primary open angle glaucoma in Latinos by conducting the first studies to examine the association between genetic factors and ocular quantitative traits of primary open angle glaucoma in this ethnic population.

I. INTRODUCTION

A. Glaucoma Background

Glaucoma is a leading cause of blindness worldwide. Characterized as a group of progressive optic neuropathies, glaucoma is a slow progressive disorder that results in the deterioration of the retinal ganglion cells and their axons, and eventually death of these cells. If left untreated, glaucoma can cause vision loss and irreversible blindness. Moreover, there is currently no known cure for glaucoma. Therefore, identifying risk factors associated with this disease will aid in furthering our understanding of the pathogenesis of glaucoma, and potentially, prevent future cases of this disease.

The manifestation of glaucoma is often asymptomatic early in the onset of the disease. Although individuals with glaucoma may experience symptoms, the disease may have progressed to advanced stages, resulting in irreversible damage to the optic nerve and potentially, vision loss. Given glaucoma is a slow progressive disorder that often does not exhibit symptoms, identifying risk factors associated with the development and progression of this disease may aid in reducing the incidence and severity of glaucoma.

Globally, glaucoma affects more than 70 million individuals, with Africa and Latin America and the Caribbean accounting for the highest and second highest prevalence of glaucoma cases, respectively.(1, 2) Due to the rapidly aging global population, recent estimates suggest that the number of individuals affected by glaucoma will increase to 76 million by 2020 and 111.8 million by 2040.(2) The largest projected increase is expected to occur in Africa, where the number of persons with glaucoma is estimated to double. Due to the substantial increase in the projected number of glaucoma cases within the next several decades, glaucoma is, and will continue to be, a significant public health issue.

As a heterogeneous group of conditions, several types of glaucoma exist. Primary open angle glaucoma (POAG) is the most prevalent form of glaucoma, affecting 44.7 million people

worldwide in 2010, with POAG most prevalent in Africa.(3) Within the United States, 2.7 million individuals have POAG, with projections estimating 7.3 million persons will have POAG by 2050.(4) The second most common form of glaucoma is primary angle closure glaucoma (PACG). Affecting 15.7 million people in the world in 2010, PACG is most prevalent in Asia, accounting for 86.5% of all PACG cases.(3) Most notably, China exhibited the greatest number of PACG cases, followed by India and Southeast Asia.(3) In addition to POAG and PACG, other forms of glaucoma include congenital, exfoliative, and pigmentary glaucoma, but are less prevalent.

While there are multiple types of glaucoma, the defining characteristic for this group of diseases is optic disc excavation. Also known as cupping, the deformation and remodeling of the optic nerve head is primarily due to the stress and strain produced by intraocular pressure (IOP).(1) The principal difference between the two main types of glaucoma relates to the anatomical outflow of the aqueous humor, the transparent fluid that fills the anterior chamber of the eye. In healthy eyes, the aqueous humor, produced by the ciliary body, flows unreservedly from the anterior chamber to the trabecular meshwork, a sieve-like tissue located in the iridocorneal angle, and is eventually drained into the blood stream. In POAG, the flow of the aqueous humor is internally blocked within the trabecular meshwork, preventing the drainage of the aqueous humor. In contrast, the trabecular meshwork is inaccessible during PACG due to the closure of the iridocorneal angle. For both POAG and PACG, the outflow of the aqueous humor is prevented, resulting in elevated IOP. The biomechanical stress from elevated IOP causes structural changes to ocular tissues, including the thinning of the retinal nerve fiber layer and narrowing of the neuroretinal rim (increasing the cup to disc ratio). In particular, these morphological changes to the optic nerve head result in the widening and deepening of the optic cup, or cupping. This cupping process damages the optic nerve fibers and eventually results in death of these fibers, severing the transmission of electrical impulses from the optic nerve to the brain and subsequently, vision loss.

Population-based epidemiological studies have illustrated that the prevalence of POAG varies by racial/ethnic group. Non-Hispanic Whites typically exhibit the lowest prevalence of POAG, followed by Latinos and Non-Hispanic Blacks who experience higher prevalences, respectively.(5-11) Moreover, individuals of African ancestry have 2.80 [95% confidence interval (CI): 1.83-4.06] greater odds, and individuals of Hispanic ethnicity have 2.00 [95% CI: 0.57-5.15] greater odds of POAG compared to persons of European ancestry.(2) When examining the effect of age on POAG stratified by racial/ethnic group, Hispanics and individuals of European ancestry experienced a steeper increase in POAG prevalence with increasing age, despite people of African ancestry having the highest prevalence of POAG across all age groups.(2) These observed racial differences indicate POAG development may be influenced by numerous factors.

Early studies investigating POAG indicated that genetic factors may play a significant role in the pathogenesis of this disease. These initial studies investigated the heritability of POAG in related individuals. Findings from the Baltimore Eye Study, a population-based prevalence study conducted in Baltimore, Maryland, identified a significant association between self-reported family history and POAG. Study participants with a first-degree family member with POAG experienced 2.85 [95% CI: 1.82-4.46] greater odds of POAG compared to individuals with a first-degree family member who did not have POAG.(5) Among first-degree family members, full siblings with a history of POAG had 3.69 [95% CI: 2.10-6.48] greater odds of POAG compared to those with no history of POAG. Moreover, children of the study participants with a history of POAG had 1.12 [95% CI: 0.26-4.86] greater odds of POAG compared to those with no history.(5) Detection bias, however, may have been present in this study due to higher odds of POAG among individuals who had prior knowledge of their glaucoma status compared to those who had no prior knowledge. Similarly, in the Barbados Eye Study, siblings of study participants exhibited 4.5 [95% CI: 2.2-9.1] greater odds of POAG for those with a sibling who had glaucoma compared to those who did not.(12) These familial findings, as well as the

observed racial/ethnic differences in POAG, suggest genetic factors may have a role in the development of POAG.

Researchers began to investigate the genetic architecture of POAG through the use of linkage analyses. While familial aggregation and segregation analyses previously aided in identifying an individual's risk based on familial disease status, and can determine the Mendelian inheritance pattern of a disease, these types of analyses do not identify the location of genetic loci associated with a given disease. Rather, linkage analyses have been successful in early genetics research in identifying disease-causing regions. For example, through linkage analyses, researchers have identified a genetic locus on chromosome 18 for bipolar disorder,(13) and a genetic locus on chromosome 13 for breast cancer, now known as BRCA2.(14) The use of linkage analyses to study the genetic architecture of POAG has led to the identification of numerous genetic loci.

The *MYOC* gene on chromosome 1 was the first gene to be associated with POAG.(15) Previously known as the trabecular meshwork inducible-glucocorticoid response protein, mutations in the myocilin protein are associated with juvenile or early adult-form of POAG, typically associated with elevated IOP, and frequently requires surgical intervention.(16) Transmitted as an autosomal dominant Mendelian disease, myocilin associated POAG accounts for 3-5% of POAG cases, making it the most common form of inherited glaucoma.(16) Moreover, 90% of carriers with mutations in *MYOC* develop POAG.(17) Compared to patients without *MYOC* associated glaucoma, in which the MYOC protein was present, this protein was absent in the aqueous humor of patients with *MYOC* associated glaucoma.(18) Additionally, findings indicated mutations in this gene may interfere with protein trafficking, resulting in the accumulation of misfolded proteins.(16) Despite these findings, how the effect of these misfolded proteins results in elevated IOP, and subsequent development of POAG, remains unknown.

The second gene associated with POAG was *OPTN*, located on chromosome 10 in the GLC1E locus.(19) Mutations in this gene, particularly the E50K variant, have been found to be associated with normal tension glaucoma (NTG), a form of POAG in which there is glaucomatous damage to the optic nerve in the absence of elevated IOP.(16) Patients with NTG who have the E50K mutation experience a severer form of glaucoma compared to those with NTG without the E50K mutation.(16) Moreover, among individuals with NTG, carriers of this variant typically are younger at the age of onset, exhibit more advanced optic nerve cupping, and require surgical interventions more frequently.(20) Similar to *MYOC*, the mechanistic role in which mutations in *OPTN* result in glaucoma remains unclear. One hypothesis suggests *OPTN* may play a role in the neuroprotection of retinal ganglion cell susceptibility to apoptosis. Overexpression of *OPTN* has been suggested to block the release of cytochrome c from mitochondria and aid in protecting cells from hydrogen peroxide induced cell death and mutations in *OPTN*, such as E50K, compromises the mitochondrial membrane, increasing the susceptibility to damage and cell death.(21)

The third gene found to be associated with POAG was *WDR36*, located on chromosome 5 in the GLC1G locus.(22) Findings from studies in families with autosomal dominant POAG identified this region, but were unsuccessful in determining the specific genetic variants in *WDR36* that resulted in POAG.(16) In one study, more severe disease was experienced by POAG individuals with *WDR36* variants compared to those without these variants, suggesting *WDR36* may play a role in the severity of disease, rather than the pathogenesis of POAG.(23)

Although these three genes, *MYOC*, *OPTN*, and *WDR36*, were found to be Mendelian forms of POAG, they only account for approximately 5% of POAG cases, suggesting a majority of POAG cases do not follow a simple Mendelian pattern.(24) As a complex disease, POAG is multifactorial in origin. Influenced by environmental, genetic, and lifestyle factors, the combination of these factors, potentially interacting with one another, may contribute to the pathogenesis of this disease. Deviating from the classical Mendelian paradigm, where a single

rare genetic variant causes disease, researchers began to hypothesize that a majority of POAG cases are a result of multiple common variants, each having a small to modest effect size on the development of the disease.

Advancements in technology over the past several decades have enabled researchers to more efficiently interrogate genetic variants across the entire human genome, and to better understand the genetic architecture of disease. Specifically, through genome-wide association studies (GWASs), researchers are able to analyze hundreds of thousands of common variants across the genome. These single nucleotide polymorphisms, or SNPs, are abundant throughout the genome and typically confer a small to modest effect size for a trait, in contrast to rare variants, which confer larger effect sizes. Moreover, compared to the methods used to identify the aforementioned POAG genes, GWASs assume no a priori knowledge of the genetic loci involved in the pathogenesis or biological pathways associated with the trait under study. Due to the unbiased nature of GWASs, this method can be used to identify novel genetic loci associated with a given phenotype throughout the genome.

Genome-wide association studies have identified numerous novel genetic loci associated with POAG. Most notably, these loci include *TMCO1*,(25-27) *AFAP1*,(25, 28) *CAV1-CAV2*,(29, 30) *FOXC1*,(25) *CDKN2B-AS1*,(25-27, 31, 32) *ABCA1*,(25, 28) *ATOH7*,(31) *ATXN2*,(25) *SIX1-SIX6*,(25, 27, 31, 32) *GAS7*,(25, 27) and *TXNRD2*.(25) Furthermore, the heritability of POAG from common genetic variants was estimated to be 0.42, and decreased to 0.36 after removing known genetic loci associated with the disease.(33) These findings indicate that common variants with small effect estimates contribute a large proportion to the heritability of POAG, and suggest additional genetic variants remain to be identified. Furthermore, previous GWASs of POAG were primarily conducted in European and Asian populations. As such, conducting GWASs in other racial populations will determine whether the effects of these variants are consistent across racial populations, as well as to identify novel genetic variants that may be specific to a given racial group.

POAG is a phenotypically heterogeneous disease. Although elevated IOP is often coupled with glaucomatous damage to the optic nerve in POAG, such damage can occur in the absence of heightened IOP and conversely, no damage to the optic nerve may occur in the presence of elevated IOP. Due to the wide range in phenotypic variation of POAG, examining quantitative traits may aid in reducing a heterogeneous disease into more homogenous traits. That is, identifying genetic factors that are associated with quantitative traits of a disease may further elucidate the biological mechanisms underlying the development of disease.

Arguably the two most clinically relevant quantitative traits of POAG, IOP and vertical cup-disc ratio (VCDR) serve as important physiological characteristics of POAG that are often measured during routine eye examinations. For healthy eyes, IOP typically ranges from 10 mmHg to 21 mmHg. Individuals with IOP greater than 21 mmHg exhibit a condition known as ocular hypertension, and are at a higher risk of developing glaucoma. Lowering IOP is currently the only available treatment for POAG. Moreover, as one of several measurements that can be derived from the examination of the optic nerve head, VCDR is the ratio of the vertical diameter of the optic cup to the vertical diameter of the optic disc. VCDR is often used to evaluate the extent of cupping of the optic nerve. Ranging from 0 to 1, a higher VCDR suggests possible glaucomatous damage to the optic nerve and potential vision loss. As important quantitative traits of POAG, identifying factors that influence both of these ocular parameters may aid in further our understanding of the development and progression of POAG.

B. Genetics and Public Health

With advancements in technology, our understanding of the role of genetics on population health has increased. From a public health standpoint, understanding the effects of genetics on human health may aid in preventing, identifying, and mitigating the impact of disease. Findings from genetic research can aid in developing and implementing primary prevention strategies to identify individuals at risk of developing a disease via screening for

known genetic variants associated with a given trait. These at-risk individuals can then take preventative measures to avert the development of disease, such as through behavioral modifications (e.g., diet, exercise, smoking, and alcohol consumption). Due to limited treatment options for POAG, understanding the attributable risk from genetic variants, in addition to environmental and lifestyle factors, may aid in identifying prevention strategies that are most effective in reducing the risk for this disease. Secondary prevention strategies may include identifying a biomarker that can be used to identify a disease early in its natural history, in which treatment can be administered to either forestall further development or to cure the disease. For example, a well-established biomarker can be measured during routine doctor examinations to identify a disease in the preclinical stage, in which treatment and prevention strategies can be initiated to either slow the progression or cure the disease. Through tertiary prevention strategies, treatments and therapies can be developed to target specific genes to mitigate the effect of disease and improve the likelihood of survival. One such example of tertiary prevention regarding the use of genetic research to reduce the burden of disease is gene therapy. For this type of therapy, a normally functioning gene is inserted into cells via a vector to replace a mutated gene. In recent years, the precision medicine initiative, in which personalized medicine is catered to an individual's genome, may further the use of tertiary prevention strategies in treating disease.

Combined with traditional risk factors, genetic factors provide additional information regarding an individual's risk of disease. The interplay between socioeconomic and lifestyle factors (e.g. income, education, smoking habits, diet, physical activity, etc.) with genetic factors may further aid in understanding the interactive effect of biological determinants of health with social determinants of health. Moreover, by understanding the interactive effect of genetic factors with environmental factors, individuals may be able to offset the risk from genetic factors through lifestyle modifications. At a local and federal level, public policy changes to address social factors, such as poverty and access to healthcare, may also aid in reducing risk of

disease. Through genetic research, we will better understand the genetic architecture of disease and the underlying biological mechanisms of disease. The application of such findings may be used to reduce the incidence and prevalence of disease via prevention strategies and public policy changes.

C. Latinos

Of the 281.4 million residents in the United States in 2000, 35.3 million or 13% of the total population was Hispanic.(34) In 2010, this ethnic group represented 16%, or 50.5 million, of the 308.7 million residents living in the United States.(35) From 2000 to 2010, the increase in the Hispanic population contributed to more than half of the growth in the total population of the United States, with the largest increase in the Mexican subpopulation. During this decade, this subpopulation experienced a 54% increase in the number of individuals living in the United States, increasing from 20.6 million to 31.8 million. Representing 63% of the Hispanic population, individuals of Mexican origin are the largest subpopulation.

Defined as a single ethnic group by the United States Census, Hispanics are a racially diverse group, composed of Mexicans, Cubans, Puerto Ricans, South or Central Americans, and other Spanish cultures. Due to the varying biogeographical ancestry within each subpopulation, Hispanics are genetically heterogeneous, with differing genomic variations unique to each subgroup. For example, population genetic structure estimates obtained from Hispanic subgroups showed Dominicans and Puerto Ricans exhibited the greatest proportions of African ancestry (41.8% and 23.6%, respectively), whereas Mexicans and Ecuadorians exhibited the lowest levels (5.6% and 7.3%, respectively).(36) The latter two groups also displayed the highest levels of Native American ancestry, 50.1% and 38.8%, respectively. Moreover, within each subpopulation, substantial variation in the amount of European, Native American, and African ancestries has been observed at the individual level. These findings provide insight into the genetic history of this diverse ethnic population, and may be used to

understand the effect of genetic variations on differences in health outcomes between such subpopulations and other ethnic groups.

Hispanics, specifically Mexicans, are an understudied population in ocular genetics research. While previous studies have identified numerous genetic factors associated with ocular diseases and related quantitative traits, these studies were primarily conducted in European and Asian populations. Conducting similar studies in a Mexican population allows for the opportunity to replicate previous findings (i.e., significance and direction of effect) in a different population, and might suggest such genetic factors are trans-ethnic, or the effect of the genetic factor is consistent across racial/ethnic groups. In addition, conducting genetic studies in this ethnic population will permit the identification of novel genetic factors associated with ocular disease and quantitative traits. Given the projected increase in the number of individuals affected by POAG, the significant increase in the Mexican population residing in the United States, and the lack of ocular genetic studies in this ethnic population, conducting genetic studies to identify genetic factors associated with quantitative traits of POAG may aid in discovering novel factors and related biological pathways, as well as elucidate racial differences for this disease.

D. Purpose of the Study

The objective of this dissertation is to address the current gaps in research regarding the genetic factors and biological mechanisms underlying the pathogenesis of POAG and related quantitative traits in Latinos. We will achieve this through the evaluation of four aims.

Aim 1: To determine whether there is an association between genetic ancestry and IOP in Latinos using data collected from the Los Angeles Latino Eye Study (LALES) and the Mexican American Glaucoma Genetic Study (MAGGS). Individual genetic ancestry estimates will be inferred from SNP data using the program STRUCTURE. Statistical analyses will be performed to evaluate the association between genetic ancestry and IOP. Effect modification

between genetic ancestry and blood pressure on IOP will also be examined. To the best of our knowledge, this study will be the first to investigate the association between genetic ancestry and IOP in a Latino population.

Aim 2: To conduct a GWAS on VCDR in a Latino population to investigate whether previously reported genetic variants identified in European and Asian populations are transferable to a Latino population and to identify novel variants. Pathway analysis will also be conducted to identify biological pathways associated with VCDR. Data from LALES and MAGGS will be used for this aim. This study will be the first GWAS conducted in Latinos on VCDR.

Aim 3: To conduct a genome-wide gene-environment interaction analysis of body mass index on VCDR in a Latino population to identify novel genetic variants associated with VCDR in subgroups of this population. Stratified analyses will be performed for identified genomic variants. For this aim, data from LALES and MAGGS will be used. This study will be the first genome-wide gene-environment interaction analysis of VCDR.

Aim 4: To construct genetic risk scores (GRSs) based on previously reported genetic variants for VCDR, and to assess the association between the GRSs on VCDR in a Latino population. Additionally, statistical analyses will evaluate the relationship between the VCDR GRSs with POAG. Receiver operating characteristic analyses will be performed to evaluate the ability of the VCDR GRSs to discriminate POAG status. Similarly, data to be used for these analyses were collected through LALES and MAGGS. To the best of our knowledge, this will be the first study to perform a VCDR GRS analysis in a Latino population.

E. Significance of the Study

Achievement of these aims will further the field of public health by conducting the first studies to examine the genetic basis of POAG quantitative traits in one of the largest minority groups in the United States. Additionally, the LALES and MAGGS combined are the largest

population-based epidemiological studies investigating ocular disease and visual-impairment in Latinos with available genetic data. Findings from these studies will further elucidate the biological mechanisms that regulate ocular phenotypes through the identification of novel associations between genetic factors and POAG quantitative traits. These results may contribute to the development of primary and secondary, and potentially tertiary, prevention strategies to reduce and prevent the development and progression of POAG in Latinos.

II. LITERATURE REVIEW

A. Intraocular Pressure and Race

Intraocular pressure is the amount of pressure created in the eye by the aqueous humor. The balance between the production of the aqueous humor by the ciliary body and the drainage of the aqueous humor into Schlemm's canal determines IOP. Abnormalities in the outflow of the aqueous humor typically result in elevated IOP. Specifically, a degenerative process in the trabecular meshwork, such as the depositing of extracellular material into the meshwork and underneath the endothelial lining of Schlemm's canal, results in resistance to the outflow of the aqueous humor.(37) As a consequence of the reduction in the outflow of the aqueous humor, IOP becomes elevated, causing excess pressure applied to the optic nerve head.

Current treatments to prevent damage to the optic nerve head and permanent vision loss among individuals at risk of POAG, or who have POAG, are directed towards lowering IOP. Numerous medical treatments have been developed to aid in the drainage of the aqueous humor and subsequently, lower IOP. These treatment options for obtaining and maintaining a healthy IOP include laser surgery, incisional surgery, and topical medication. Laser trabeculoplasty is the most widely used laser treatment and involves the use of light directed towards the trabecular meshwork cells, which activates these cells to remodel the local extracellular matrix, resulting in increased outflow of the aqueous humor.(1) With regard to incisional surgery, trabeculectomy is the most widely performed and involves the incision of a small hole in the sclera, or sclerectomy, that is covered by a scleral flap, allowing for the aqueous humor to flow from the anterior chamber to the subconjunctival space.(1, 38) Topical medications are applied to the ocular surface and either improve the outflow of the aqueous humor or reduce its production from the ciliary body. Despite the former two forms of treatment yielding equal or improved IOP management compared to topical medications, the use of surgery as the primary form of treatment is not widely accepted due to the potential risk of complications from these procedures.(1)

Intraocular pressure is currently the only known modifiable risk factor for POAG. Individuals with IOP > 21 mmHg have a condition called ocular hypertension. Randomized clinical trials have shown medical treatments not only lower IOP among individuals with ocular hypertension, but also prevent the development of POAG. One seminal study that evaluated the efficacy and safety of topical ocular hypotensive medication on delaying and preventing the development of POAG was The Ocular Hypertension Treatment Study.(39) In this randomized clinical trial, study participants with ocular hypertension, determined as IOP ranging from 24 to 32 mmHg in one eye, and ranging from 21 to 32 mmHg in the other, and had no other ocular diseases, were randomized to either the medication group or the observation group. Study subjects in the medication group sought to achieve an IOP of 24 mmHg or less as well as at least a 20% reduction in IOP from the baseline visit. Topical medications for the study subjects were modified to obtain these goals throughout the study. Commercially available topical hypotensive medications at the time of the study were distributed to study participants, with any newly developed medications during the study dispensed upon availability. Follow-up visits occurred every 6 months, in which the study participants received an ocular examination, with additional fundus examination and stereoscopic optic disc photographs every 12 months. The primary outcome for this study was POAG development in either or both eyes, characterized as visual field abnormality or optic disc damage consistent with POAG.

A total of 1,636 study participants were randomized 1:1 to either one of the two treatment groups. The median follow-up time for African Americans was 72 months and 78 months for other study participants. At baseline, the mean ± standard deviation (SD) IOP for participants in the medication group and observation group was 24.9 ± 2.6 mmHg and 24.9 ± 2.7 mmHg, respectively. A reduction in IOP among the medication group but not the observation group was observed, with IOP averages during follow-up of 19.3 ± 2.2 mmHg and 23.9 ± 2.9 mmHg, respectively. Moreover, the medication group exhibited a larger percentage reduction in IOP from baseline compared to the observation group, -22.5 ± 9.9% vs. -4.0 ± 11.6%,

respectively. With regard to POAG, 36 out of 817 participants in the medication group developed POAG, compared to 89 out of 819 participants in the observation group. After 60 months, the observation group exhibited a 9.5% cumulative probability of developing POAG compared to 4.4% for the medication group (hazard ratio = 0.40 [95% CI: 0.27-0.59]; $P < 0.0001$). This randomized clinical trial demonstrated use of topical ocular hypotensive medications aided in not only reducing IOP but also prevented the development of POAG among individuals with ocular hypertension.

Although findings from The Ocular Hypertension Treatment Study show topical hypotensive medications can lower IOP once elevated IOP is experienced, identifying factors that determine IOP may aid in understanding the biological processes regulating this trait, as well as potentially predict IOP. Numerous epidemiological studies have identified several factors associated with IOP among individuals of European, African, and Asian ancestries, including age, gender, body mass index, diabetes, blood pressure, myopia, and nuclear sclerosis, a type of age-related cataract.(40-42) In addition to these studies, one population-based cross-sectional study evaluated associations between biological factors and IOP in a Latino population.(43) Study participants from this study consisted of self-identified Latinos (primarily Mexican-American) from the Los Angeles Latino Eye Study. Study participants completed an in-home questionnaire and interview to capture age, gender, and history of medical conditions. Participants also underwent clinical and ocular examinations to obtain medical characteristics, including weight, height, pulse, systemic blood pressure, glucose levels, glycosylated hemoglobin, IOP, central corneal thickness (CCT), iris color, axial length, and nuclear cataract grade.

Of the original 6,357 study participants from the baseline study, 5,958 Latinos remained for subsequent analysis. Subjects were primarily female (58%), with a mean (± SD) age of 54.9 ± 10.9 years, and with a mean IOP of 14.5 ± 3.2 mmHg. Overall, the mean IOP significantly increased with increasing age ($P < 0.0001$). Specifically, IOP increased from 14.0 ± 2.8 mmHg

at ages 40 to 49 years to 14.6 ± 3.2 mmHg, 14.8 ± 3.4 mmHg, and 15.1 ± 3.7 mmHg for ages 50 to 59, 60 to 69, and 70 or older, respectively. Additionally, women had significantly higher IOP compared to men, 14.6 ± 3.1 mmHg vs. 14.3 ± 3.3 mmHg ($P < 0.0001$). Variables significantly associated with higher IOP ($P < 0.001$) from univariate analysis of categorical variables were: female gender, hypertension, diabetes, higher nuclear cataract grade, and dark brown iris color. Family history of glaucoma, history of tobacco use, history of alcohol use, history of steroid use, cardiovascular disease, and history of eye trauma were not significantly associated with IOP. In the univariate analysis of continuous variables, IOP was significantly ($P < 0.001$) associated with: older age, higher body mass index (BMI), higher systolic blood pressure (SBP), higher diastolic blood pressure, faster pulse, thicker CCT, and higher glycosylated hemoglobin. Axial length was not significantly associated with IOP. After performing stepwise multivariable regression, age, gender, diabetes, nuclear cataract grading, iris color, BMI, SBP, diastolic blood pressure, and CCT remained as significant predictors of IOP. Together, these predictors explained approximately 10% of the variation in IOP, with CCT and SBP explaining a majority of the variation (4% for each variable). Findings from this study are consistent with results from previous studies.(40, 41)

While demographic, clinical, and ocular factors have been found to be associated with IOP, these features explain only a small proportion of the variation in IOP. As such, the identification of additional factors may aid in further explaining this trait. Comparison of large epidemiological studies conducted in different ethnic populations have revealed racial differences in IOP, suggesting genetic ancestry may influence IOP. The Baltimore Eye Study was one of the first seminal studies to report on the differences in ocular parameters by race.(44, 45) Between January 1985 and November 1988, a cluster sampling strategy, allowing for an equal balance of study subjects by race, was used to select 16 locations in Baltimore, Maryland. Individuals in these locations were screened and those who were eligible received an ophthalmologic examination, including three IOP measurements in each eye, in which the

median measurement was defined as the measurement for each eye, and the higher of the two measurements was used as the final value for each person. Of the 1,770 subjects referred for ophthalmic examination, 766 black participants and 659 white participants completed testing. In comparing the IOP measurements between the two races, white subjects had a higher mean IOP compared to the black subjects (17.17 ± 3.35 mmHg vs. 16.00 ± 4.18 mmHg; $P < 0.001$).

A second seminal study, the Barbados Eye Study, was conducted between April 1988 and May 1992.(46) In Barbados, West Indies, a random sample of natural born citizens aged 40 to 84 years was selected from the national registry. Responding study participants received an ophthalmic examination, during which various ocular parameters were measured, including three IOP measurements per eye, with the average of these measurements yielding the final IOP value for each study participant for each eye. In contrast to the Baltimore Eye Study, black participants in the Barbados Eye Study experienced a higher IOP for both eyes (right: 18.1 ± 4.8 mmHg; left: 18.1 ± 4.8 mmHg) compared to whites (right: 16.2 ± 3.1 mmHg; left: 16.0 ± 2.8 mmHg). Additionally, the percentage of black participants with IOP > 21 mmHg was higher compared to white participants, 16.5% vs. 4.6%. Differences in the study populations may explain the discrepancy in results between the Baltimore Eye Study and the Barbados Eye Study. For example, as a Caribbean island, the population of Barbados primarily originated from West Africa. Moreover, compared to African Americans from the United States, who are an admixture of European and African ancestries, Barbadians have a larger proportion of genetic ancestry from Africa.(47) As such, individuals with a greater amount of African ancestry may have a higher frequency of genetic variants associated with higher IOP than those with a lower proportion of African ancestry.

The Blue Mountains Eye Study, another seminal ocular investigation, was a population-based study consisting primarily of Caucasian individuals in an urban community of Sydney, Australia between January 1992 and January 1994.(9) Study participants were identified through census data and were invited to receive an eye examination, during which a single IOP

measurement was taken in either one or both eyes. Of the 3,641 study participants with reliable IOP measurements, the average IOP for the right eye was 16.1 ± 2.9 mmHg and 16.0 ± 2.9 mmHg for the left eye. Furthermore, ocular hypertension was present in 3.7% of the study participants.

In addition to the LALES, one other population-based study has been conducted on Hispanics. The Proyecto VER study consisted of a census-based random sample selection of Hispanics in Nogales and Tucson, Arizona from April 1997 to September 1999.(8) Eligible participants received an in-home interview and eye examination at a nearby clinic, during which IOP measurements were taken. The average IOP for the 4,774 subjects included in the final analysis was 15.6 ± 3.2 mmHg. In comparison, the average IOP among LALES subjects was 14.5 ± 3.2 mmHg.(43)

Lastly, a cross-sectional study in Japan randomly selected study participants from the general population stratified by sex and age as part of the National Institute for Longevity Sciences – The Longitudinal Study of Aging program.(48) Three IOP measurements were taken in each eye for each study participant, and the mean measurements for the right eye yielded the final value. Among the 1,317 study subjects included in the final analysis, the average (SD) IOP was 13.4 (2.6) mmHg. The average IOP in this Japanese population was the lowest of the ethnic populations reported in these previous studies. These population-based studies demonstrate racial differences in IOP and suggest genetic factors within each racial group may influence IOP determination.

Despite these observed racial differences, the study populations and corresponding methodologies for each study varied. Compared to the other seminal studies, the Japanese cross-sectional study had the smallest proportion of female study participants, while the Baltimore Eye Study and Proyecto VER had the highest proportions. Interestingly, an inverse association between age and IOP was observed in the Japanese study, although the exact reason for this remains unclear. Moreover, the age distributions for these studies differed from

one another. For example, study subjects in LALES had a mean age of 54.9 years compared to a mean age of 65 years for study participants in the Blue Mountains Eye Study. Additionally, the eligibility criteria for each study differed. That is, study participants had to be 40 years of age or older for LALES, whereas study participants for the Blue Mountains Eye Study had to be 49 years of age or older. In addition, genetic variants, such as differences in allele frequencies for IOP related variants, may explain the racial differences in IOP. Apropos to IOP measurements, several studies took three IOP measurements, either taking the mean or median of the three values as the final measurement, while one study took a single IOP measurement. Although taking a single IOP measurement may reduce time during a clinical visit, as well as simplify data management and statistical analysis compared to three measurements, fluctuations in IOP and intra-observer variation may affect the reliability of IOP measurements. A previous study investigating the reliability of IOP measurements in epidemiological studies determined using the median of three IOP measurements yielded more reliable estimates compared to a single IOP measurement.(49) Specifically, the inter-observer and intra-observer variation decreased by 11% and 9%, respectively, when the median of three IOP measurements was taken compared to a single measurement. Moreover, the median of the three measurements was preferred over the mean as to reduce the influence of IOP outlier measurements. Differences in both the study sample characteristics and study methodology may aid in explaining the observed racial variation in IOP across these study samples.

Racial differences in other quantitative traits associated with IOP have also been observed. In addition to being a risk factor for numerous diseases, including obesity and cardiovascular disease, hypertension is an important contributor to IOP. With a prevalence of 29.1% for adults in the United States, hypertension is a significant public health issue.(50) Additionally, hypertension disproportionally affects Non-Hispanic Blacks compared to other racial groups, with this racial group experiencing the highest prevalence (42.1%), followed by Non-Hispanic Whites (28.0%), Hispanics (26.0%), and Asians (24.7%). In addition to being

associated with IOP, chronic elevated blood pressure may result in atherosclerosis and disruption of the autoregulatory mechanisms controlling the blood flow to the optic nerve head. These biological changes can result in a reduction in perfusion and subsequent damage to the retinal ganglion cells, potentially leading to POAG.(51) Taken together, individuals of African ancestry exhibit elevated IOP and a high prevalence of hypertension, both of which contribute to POAG development. Latinos, a racial group comprised of African ancestry, exhibit intermediate measures of IOP and hypertension, and offer a unique opportunity to evaluate the interplay between these two traits with regard to genetic ancestry.

As social constructs, race / ethnicity attempt to classify individuals based on biological traits and cultural factors but fail to capture the genetic heterogeneity in populations, most notably in admixed populations. Advancements in technology and genetics have led to the development of ancestral informative markers and genome-wide data, enabling researchers the ability to estimate the genetic ancestry of individuals. For example, Kumar et al. estimated genetic ancestry for study participants from the Coronary Artery Risk Development in Young Adults study using genome-wide genotype data and the program ADMIXTURE, as well as ancestral informative markers for several replication datasets.(52) The investigators observed an 8.14 decrease in forced expiratory volume in 1 second per 1% increase in African ancestry. The findings from this study suggest genetic ancestry may serve as a proxy for an aggregate measure of genetic factors that are differentially distributed based on biogeographic separation, and may better refine risk estimates as compared to racial / ethnic classifications.(52) Moreover, genetic ancestry may further aid in elucidating observed racial differences for diseases and quantitative traits.

Latinos are a three-way admixed population of African, European, and Native American ancestry, with considerable variation in the proportion of ancestries among Latino individuals.(36, 53, 54) Recently, a study conducted in Latinos found that those with a greater proportion of Native American ancestry had higher odds of severe diabetic retinopathy

compared to those with a lower proportion of Native American ancestry.(55) That is, for each percent increase in Native American ancestry, the odds of severe diabetic retinopathy was 1.02 [95% CI: 1.01-1.03]. To further aid in the interpretability of this finding, the investigators dichotomized genetic ancestry at 50% Native American ancestry and observed Latinos with higher Native American ancestry (≥ 50%) had 1.87 [95% CI: 1.26-2.78] greater odds of severe diabetic retinopathy, compared to Latinos with lower Native American ancestry (< 50%). Given the observed racial differences in IOP, as well as the genetic heterogeneity in Latinos, determining whether and to what extent there is an association between IOP and genetic ancestry may aid in understanding the racial differences in IOP. Therefore, we aim to evaluate the association between IOP and genetic ancestry in Latinos. We hypothesize increasing African ancestry will be associated with higher IOP in Latinos. Moreover, based on previous literature, we hypothesize the association between African ancestry and intraocular pressure will be modified by a significant interaction between African ancestry and blood pressure.

B. The Genetics of Vertical Cup-Disc Ratio

Situated at the posterior side of the eye, the optic disc is the location at which the retinal ganglion cell axons converge and exit the retina. Also called the optic nerve head, the optic disc is the starting point of the optic nerve, a pair of nerves that transmit visual information from the eye to the brain. This anatomical structure is routinely assessed during eye examinations to identify and monitor possible glaucomatous damage. Damage to the optic nerve head or the optic nerve itself may result in visual impairment and potentially vision loss. As such, regularly examining the optic disc for changes in the morphology is important in early detection of ocular disease, including POAG.

One optic disc parameter often estimated is the ratio of the vertical diameter of the optic cup to the vertical diameter of the optic disc, also called the vertical cup-disc ratio (VCDR). Among individuals with glaucoma, a progressive thinning of the neuroretinal rim occurs due to

the loss of the retinal ganglion cell axons, in addition to the vasculature and glial cells.(56) This results in the enlarging of the optic cup, a process also known as optic nerve cupping or cupping, and can result in visual field loss. Although damage to the optic nerve, and subsequent vision loss, is irreversible, identifying factors associated with this process, specifically factors associated with VCDR, may aid in elucidating the underlying biological mechanisms of POAG, and potentially lead to public health strategies and therapies to prevent or slow the progression of this disease.

Previous epidemiological studies have illustrated differences in VCDR by ethnic population. Varma et al. conducted a population-based epidemiological study to investigate the effect of race, age, and gender on various ocular parameters using data from the Baltimore Eye Study.(57) Simultaneous stereoscopic optic disc photographs were taken and an image analyzer system was used to calculate various optic parameters, including disc area, cup area, and VCDR. A total of 4,877 healthy individuals, consisting of 2,097 black participants and 2,780 white participants, were included in the study. Due to media opacities, procedural complications, and ocular difficulties, optic disc photographs were available for 3,387 study participants: 1,534 black participants and 1,853 white participants. In comparing the VCDR measurements between the two racial groups, the black study participants had significantly larger VCDR measurements compared to the white study participants, with a mean VCDR of 0.56 [95% CI: 0.55-0.57] for the black participants and 0.49 [95% CI: 0.48-0.50] for the white participants (P = 0.0001). Albeit this finding, no additional variables were controlled for in the analysis that may confound the observed association, such as age and gender. For example, in the total study population, white participants tended to be older compared to black participants.(45) Due to the unequal distribution of age between the two racial groups, age may confound the observed association. Findings from this study suggest differences in VCDR measurements may vary by ethnic population.

Expanding the number of ethnic populations, Knight et al. conducted a cross-sectional, multi-ethnic observational study to evaluate the association between race and numerous ocular parameters.(58) Using spectral-domain optical coherence tomography, a next generation technology that provides detailed tissue structure information, several optic nerve head characteristics were calculated from these results, including disc area, rim area, and VCDR. After adjusting for age, study participants of African descent (n = 51) exhibited the largest VCDR measurements, with an average (standard error) VCDR of 0.51 (0.02). Study participants identified as Hispanic (n = 35) and Chinese (n = 63) experienced the second largest VCDR measurements, with an average VCDR of 0.44 (0.03) and 0.44 (0.02), respectively, followed by participants of European ancestry (n = 122) with an average VCDR of 0.41 (0.01). Although the difference in VCDR between Hispanic, Chinese, and European individuals may not be clinically meaningful as compared to African individuals, these results highlight the racial differences in VCDR. However, the number of study participants in each racial group was small, potentially limiting the generalizability and affecting the stability of these findings. Additionally, while a significant difference between race and VCDR was observed, this analysis adjusted only for age. As such, differences in demographic and environmental factors between racial groups may contribute to the observed association, including intraocular pressure and body mass index. Despite these limitations, this study demonstrates variation in VCDR by race and suggests genetic factors may contribute to VCDR.

To investigate the role of genetics on VCDR, several studies estimated the heritability, or the proportion of variation in a trait due to genetic factors, for VCDR. These studies estimated the heritability of VCDR to be 48% - 66%.(59, 60) Combined with the observed racial differences in VCDR, these findings suggest genetic factors may influence VCDR determination. Given the apparent genetic nature of VCDR, several genome-wide association studies were conducted to identify specific genetic variants associated with VCDR.

The first GWAS to investigate the association between genetic variants and VCDR was conducted in European participants primarily from the Rotterdam Study.(61) This study conducted a two stage GWAS, consisting of a discovery and replication dataset. The discovery dataset included 5,312 study participants from the Rotterdam Study cohort I (RS-I) and 2,048 study participants from the Rotterdam Study cohort II (RS-II). The replication datasets consisted of 1,966 study participants from the Rotterdam Study cohort III (RS-III), 1,646 individuals from the Erasmus Rucphen Family (ERF) study and 843 participants from the TwinsUK cohort. Performing linear regression, adjusting for age, gender, and optic disc area, two SNPs reached genome-wide significance ($P < 5\times10^{-8}$) in the discovery dataset. The most significant association was rs1063192 in *CDKN2B* on 9q21 (minor allele frequency (MAF) = 0.45) and was associated with a 0.014 decrease in VCDR per copy of the G allele. The second most significant SNP was rs10483727 in *SIX1* on 14q22-23 (MAF = 0.40) and was associated with a 0.014 increase in VCDR per copy of the T allele. During the meta-analysis of the discovery and replication datasets, these two SNPs became more significant and conferred similar magnitude and direction of effects. In addition to these two SNPs, four other SNPs reached genome-wide significance, including rs17146964 in *SCYL1* on 11q13 (MAF (G) = 0.21; β = -0.014), rs1547014 in *CHEK2* on 22q12.1 (MAF (T) = 0.29; β = -0.011), rs1900004 in *ATOH7/PBLD* on 10q21.3-q22.1 (MAF (T) = 0.22; β = -0.013), and rs1926320 in *DCLK1* on 13q13 (MAF (C) = 0.24; β = 0.012). Three additional SNPs demonstrated suggestive associations ($P < 1 \times 10^{-6}$) with VCDR, including rs8068952 in *BCAS3* on 17q23 (MAF (G) = 0.23; β = -0.012), rs12025126 in *RERE* on 1p36.2-p36.1 (MAF (C) = 0.29; β = -0.011), and rs2159128 in *ARID3A* on 19p13.3 (MAF (G) = 0.13; β = -0.019). Moreover, of these loci, 4 were found to be associated with POAG (9q21, 14q22-23, 11q13, and 10q21.3-q22.1).

This first GWAS of VCDR identified 6 loci, though there are several limitations for this study. First, the method of VCDR measurement varied across cohorts. For RS-I, RS-II, and TwinsUK, optic nerve imaging was performed using digitized stereoscopic images, whereas RS-

III and ERF used confocal scanning laser ophthalmoscopy. Although the authors state previous studies have shown a high correlation (Pearson's correlation coefficient: 0.75) between these two measurements, the differences in the measurement methodology may lead to measurement error, potentially affecting the observed associations.(62) Second, this study consisted solely of participants of European descent. As such, these findings may not be generalizable to other ethnic populations. Although this study identified genetic variants associated with VCDR, the six genome-wide significant SNPs from the meta-analysis only account for 2.2% of the variation in VCDR. This suggests additional genetic factors remain to be identified.

A second GWAS meta-analysis was conducted using data from the International Glaucoma Genetics Consortium.(63) Using data from 14 studies from the United States, Europe, Australia, and Asia, a four stage meta-analysis was performed to identify additional variants associated with VCDR, and to examine the association between these genetic variants and glaucoma. In the first stage of the study, the investigators performed a meta-analysis of data from 10 study populations, consisting of 21,094 individuals of European ancestry. For stage two, the researchers assessed the transferability of significant findings identified from stage one in 6,784 study participants from four Asian cohorts. In stage three, analyses were conducted to determine whether the identified variants are independent of other ocular parameters, including disc area and spherical equivalent. Lastly, the fourth phase involved gene-based tests and pathway analysis to identify biological pathways associated with VCDR.

During the first stage of the investigation, 440 genome-wide significant SNPs over 15 genomic loci were identified. These regions include *CDC7/TGFBR3*, *COL8A1*, *DUSP1*, *EXOC2*, *CDKN2BAS*, *ATOH7*, *PLCE1*, *SSSCA1*, *ADAMTS8*, *RPAP3*, *TMTC2*, *SIX1/6*, *SALL1*, *BMP2*, and *CHEK2*. Of these regions, the top SNP for nine of the loci are intergenic, whereas the remaining six are either intronic or in a 5' upstream region of a gene. In stage 2, the investigators assessed the association of the top SNP in each identified locus in Asian subjects and found eight were nominally associated ($P < 0.05$) with VCDR. Moreover, these variants

exhibited consistent direction of effect, as well as similar effect sizes. Of the remaining seven non-significant loci, five SNPs exhibited consistent direction of effect. When these ethnic populations were combined and analyzed together, three additional loci reached GWAS significance. These genetic loci were *RERE*, *HSF2*, and *CARD10*, of which the first was intronic and the last two were intergenic. Of the 18 genome-wide significant loci, ten were novel for VCDR.

Among the 18 novel loci, 4 loci have also been associated with optic disc area. Due to the correlation between optic disc area and VCDR, the investigators adjusted for optic disc area in the association between these variants and VCDR during stage 3. After controlling for optic disc area, the association in the *CDC7/TGFBR3* locus became non-significant with a reduction in the effect size, suggesting the observed association between this locus and VCDR was primarily due to the association with optic disc area. A reduction in the effect estimate was also observed for the *ATOH7* locus, but the association remained statistically significant. Furthermore, there was no change in the significance levels for the 10 novel loci, suggesting these associations are primarily due to VCDR. When these associations were adjusted for spherical equivalent, no major changes occurred. Additionally, the 18 identified loci accounted for 5.1-5.9% of the variation in VCDR in the European cohorts, after adjusting for age and sex, of which 1.6-1.8% of the variation was explained by the novel loci. When examining the association between these loci with POAG, the investigators replicated two previous associations in the *CDKN2BAS* and *SIX1/6* regions, and identified 6 novel associations with POAG among the remaining GWAS significant VCDR loci.

To further identify genetic loci associated with VCDR that were not found directly through single variant association testing, the authors performed gene-based testing and pathway analysis using the software VEGAS.(64) Mapping all SNPs within a gene, as well as SNPs ± 50kb of a gene, the authors identified two additional genes, *PITPNB* and *REEP5*, associated with VCDR. While *REEP5* conferred no association with POAG, *PITPNB* was marginally

associated with POAG. Through pathway analysis using Pathway-VEGAS, the investigators found ‘negative regulation of cyclin-dependent protein kinase activity’, a pathway related to cell growth, as the only significant pathway associated with VCDR.

While identifying novel genetic variants associated with VCDR, a limitation exists in the pathway analysis. In the current study, the program VEGAS generates gene based *P* values by summing the SNP-based chi-square statistics for SNPs within a given gene boundary, accounting for linkage disequilibrium. VEGAS, however, lacks power to detect associations for SNPs with low MAF, potentially due to the lack of correlation between these low frequency SNPs and other SNPs.(65) Due to the potential for SNPs with low MAF to be causal variants, methods to better analyze such variants are needed. As such, use of alternative pathway analysis tools may aid in identifying additional biological pathways associated with VCDR.

A third GWAS meta-analysis, consisting of 19 studies, was recently conducted to investigate genetic variants associated with intraocular pressure and optic disc characteristics, including VCDR.(27) Similar to the previous GWAS, this study used a multi-stage approach using European and Asian populations. In the meta-analysis of European individuals (n = 23,899), 21 genomic loci reached genome-wide significance, of which 5 were novel associations, including *RPE65*, *F5*, *PDZD2*, *RREB1*, and *DGKB*. The remaining 16 regions have previously been associated with VCDR or optic disc area. In the meta-analysis of individuals of Asian descent (n = 8,373), one of the five novel loci, *RREB1*, was nominally associated with VCDR. In the combined meta-analysis (n = 32,272), four additional novel loci were genome-wide significant, including *VCAN*, *PSCA*, *ENO4*, and *RBM23*. In total, 9 novel loci were found to be associated with VCDR. Among these loci, *F5* has been reported to be associated with optic disc area. To determine whether this association was solely due to VCDR or was primarily associated with disc area, the investigators further adjusted for disc area and observed a decrease in both the effect estimate and significance, suggesting the latter of the

two explanations. With regard to determining the influence of these variants on POAG, only *SIX6* and *CDKN2B-AS1* were associated with POAG.

The investigators further performed gene-based testing using VEGAS2(66) and biological functional enrichment with DEPICT.(67) Through VEGAS2, two additional loci, *RARB* and *HORMAD2-AS1*, were significantly associated with VCDR. These loci, however, were previously found to be associated with optic disc area, and thus these associations may be driven by their relationship with disc area.(68) Using the program DEPICT, several pathways related to metabolic processes were suggestive of associations when significant SNPs associated with VCDR, cup area, and disc area were analyzed together, including increased circulating leptin level, increased insulin sensitivity, and abnormal fat cell morphology.

Analogous to the previous meta-analysis, this GWAS meta-analysis is not without limitations with regard to the pathway analysis. First, the program DEPICT was run using only the top SNPs associated with traits. True causal variants with lower levels of significance may be excluded from such analyses and thus, would reduce the power to detect biological meaningful associations. As such, methods to include additional variants and their corresponding effect estimates are needed. Second, the authors analyzed SNPs associated with VCDR, cup area, and disc area together. As such, the identified pathways may not be independently associated with VCDR, but may rather be associated with the other two traits.

These GWASs have been successful in identifying genetic loci associated with VCDR, elucidating the biological mechanisms influencing VCDR determination. Although GWASs are powerful tools to improve our understanding of the genetic architecture of complex diseases, compared to Mendelian diseases, most genetic variants identified through GWASs confer only a small to modest effect size. Moreover, both individually and combined, the heritability of identified variants for complex traits is small. Attempting to uncover the ‘missing heritability’ of complex traits is important because individual differences in genetic factors are known to be related to disease susceptibility, and by understanding such genetic variation may lead to

improvements in preventing, diagnosing, and treating disease.(69) Several theories have been suggested to explain this missing heritability, including the lack of genetic variants yet to be identified, the exclusion and poor detection of rare variants on current genotyping chips, the absence of structural variants on these chips, and the low power to detect gene-gene and gene-environment interactions. Addressing these sources of missing heritability may aid in further elucidating the genetic architecture of disease and subsequently, improve health outcomes.

One conclusion from early GWASs indicated that for most complex diseases, multiple loci contribute to the trait understudy, each explaining a small proportion of variation.(70) Although the proportion of variation explained is small, this observation implied additional genetic variants remain to be identified and the discovery of such variants through larger sample sizes will further increase this proportion.(71) By increasing the sample size, studies will have more power to identify additional genetic variants with small effect sizes, as well as those with low frequencies, and thus, may aid in uncovering the missing heritability of traits. For example, 40 SNPs were associated with human height in 2008, explaining 5% of the heritability but by 2014, approximately 700 SNPs were identified, explaining 20% of the heritability, with thousands of variants expected to be identified over the next several years due to increases in sample size, further explaining the remaining heritability.(70) In addition to increasing sample size, the overwhelming majority of GWASs have been conducted in European populations, indicating genetic studies in Non-European populations may further identify additional variants associated with traits.(69) As two options to identify additional variants, and subsequently uncover the missing heritability of traits, future GWASs should attempt to have large sample sizes and be conducted in Non-European populations.

The contribution of variants with low minor allele frequencies (< 5%) and rare variants (< 1%) has been speculated to address the missing heritability. Moreover, such genetic variants have been suggested to have a greater impact on explaining familial risk and predicting individual disease risk than common variants identified through GWASs.(72) These variants,

however, are not commonly available on genome-wide genotype arrays and are not identified in other genetic studies due to modest effect sizes.(69) The primary approach to identify rare variants is through sequencing, either the entire genome or targeted segments. Current methods to identify rare variants involve sequencing targeted regions of the genome that are strongly and consistently associated with a given trait through GWASs. Through advancements in technology, the cost of whole genome sequencing has decreased, and with the use of custom genotyping arrays, including arrays that target low frequency and rare variants, additional heritability may be explained by these genetic variants.

Structural variations in the human genome, such as insertions, deletions, inversions, and translocations, have also been hypothesized to further explain the remaining heritability. Similar to SNPs, rare structural variations confer large effect sizes, whereas more common variations exhibit more modest effect sizes.(69) Moreover, like low frequency and rare genetic variants, structural variations are not frequently ascertained on GWAS arrays. The inclusion of such genomic alterations into current genetic studies may further elucidate the missing heritability of numerous traits.

Gene-environment (G×E) interactions represent an additional strategy to explain the missing heritability. Complex traits and diseases are likely to be influenced by multiple genes and environmental factors, as well as the interplay between the two. GWASs identify those genetic variants that exhibit significant main effects, whereas variants that require an interacting factor are missed, potentially limiting our understanding of disease. From an epidemiological prospective, investigating G×E interactions will aid in our understanding of how genetic susceptibility in a subgroup of a population may predispose such individuals to heightened environmental effects, or how exposure to environmental factors may lead to increased or decreased gene expression that may ultimately result in disease.(73) Genetic variants that confer a heightened effect in only one subgroup of the population, or have genetic effects in opposite directions, may produce weak marginal genetic effects in standard genome-wide

association studies.(74) Testing for G×E interactions has been considered a promising approach to further improve our understanding of the genetic architecture of complex diseases by identifying novel genetic variants missed during standard GWAS analyses, and may further explain the missing heritability.(74)

The investigations of G×E interactions have yielded novel genetic variants associated with complex diseases that remained elusive when only the main genetic effect was considered. For example, a genome-wide gene-environment study investigating the interactive effects of genetic variants and coffee consumption on Parkinson's disease identified novel associations with SNPs in *GRIN2A*.(75) Specifically, heavy coffee drinkers exhibited a greater reduction in the odds of Parkinson's disease per copy of the T allele for rs4998386 (OR (standard error) = 0.43 (0.07), $P = 6 \times 10^{-7}$) compared to light coffee drinkers (OR (standard error) = 0.84 (0.11), $P = 0.19$). A second genome-wide gene-environment interaction analysis investigated the interactive effects of tobacco smoke and genetic variants on lung cancer.(76) The T allele of rs4589502 in the *SMAD6-SMAD3* region yielded significant interactive effects with smoking, with a lower risk among non-smokers (OR = 0.74 [95% CI: 0.64, 0.85], $P = 1.62 \times 10^{-5}$) and a harmful effect for smokers (OR = 1.14 [95% CI: 1.00, 1.29], $P = 4.61 \times 10^{-2}$). And lastly, a third genome-wide G×E interaction study was conducted to identify genetic variants associated with asthma-related BMI increase.(77) Among those study subjects with asthma, the odds of being obese increased by 1.89 (OR = 1.89 [95% CI: 1.39, 2.57], $P = 4.34 \times 10^{-5}$) fold for each copy of the A allele for rs2107212 in the *KRT23-KRT39* region, and among those non-asthmatic study subjects, the odds of being obese decreased by 0.89 (OR = 0.89 [95% CI: 0.77, 1.03], $P = 0.12$) fold for each A allele. These studies demonstrate the utility of incorporating environmental factors in GWASs to identify novel variants, subsequently increasing our understanding of the genetic architecture and decreasing the amount of missing heritability for complex traits. Moreover, these findings may lead to public health prevention strategies, such as identifying subgroups in a population who genetically are at a greater risk of disease and providing

targeted interventions to reduce risk among such individuals. Additionally, genetic variants identified in G×E interaction analyses that either diminish or enhance an environmental factor may be used as a biomarker for pharmacogenetic prevention and treatment strategies, such as determining which medications are most effective for patients.(75)

Genome-wide association studies have provided valuable insights into the genetic architecture of common diseases. In order to continue to identify novel genetic variants and uncover the remaining missing heritability of traits, GWASs can be enhanced by increasing the sample size of studies, conducting studies in non-European populations, incorporating low frequency and rare variants, analyzing structural variants, and investigating G×E interactions. As one of the goals of genetic research, findings from GWASs should lead to translational advances directly influencing prevention and treatment strategies for disease. For example, findings from GWASs initiated the medication repositioning for components of the IL-23 pathway to treatments for psoriasis, ankylosing spondylitis, and inflammatory bowel disease.(70) The identification of genetic variants through GWASs have elucidated the biological mechanisms influencing complex traits, aided in phenotype prediction, and yielded additional treatments for diseases. Apropos to VCDR, additional GWASs, including studies in different ethnic populations and the assessment of G×E interactions, may lead to effective prevention strategies for POAG.

The previous VCDR GWASs, however, were conducted in populations of European and Asian descent, potentially limiting the generalizability of these results. Moreover, because variants occur with varying frequency by ethnic population, conducting studies in different ethnic populations will aid in: (1) determining whether previously reported genetic variants exhibit consistent direction and size of effect across ethnic groups, and (2) identifying novel variants. To date, no GWAS on VCDR has been reported in Latinos. We aim to perform the first GWAS on VCDR in Latinos. We hypothesize we will replicate numerous previously reported VCDR SNPs identified in studies conducted in individuals of European and Asian descent in a Latino population, demonstrate the transferability of genetic variants across ethnicities, and potentially

identify novel genetic variants associated with VCDR. Using pathway analysis tools, we will identify novel pathways associated with VCDR. Findings from this study may aid in further elucidating the racial differences in VCDR.

Moreover, the incorporation of gene-environment interactions into genetic studies are now beginning to be conducted. There is compelling evidence from epidemiological studies that lower body mass index is associated with larger VCDR. A population-based, cross-sectional study of Chinese adults observed a significant positive association between neuroretinal rim area and BMI, with higher BMI associated with larger neuroretinal rim area.(78) Population-based studies of adult Malay individuals identified lower BMI was associated with larger VCDR, after adjusting for age, sex, IOP, and other covariates.(79, 80) An additional population-based study conducted in a Korean population identified a similar association with VCDR being negatively associated with BMI, after adjusting for covariates.(81) Moreover, a recent population-based study among Japanese individuals observed a significant association between BMI and VCDR, with male subjects exhibiting larger VCDR with lower BMI.(82) These findings suggest higher BMI may be protective against the thinning of the neuroretinal rim and cupping, and subsequent damage, to the optic nerve.

Despite this association, the underlying biological mechanism remains unclear. One proposed theory suggests BMI may influence cerebrospinal fluid (CSF), a clear body fluid that surrounds the brain and spinal cord acting as a cushion. Cerebrospinal fluid pressure was previously found to be positively associated with BMI.(83) Moreover, the optic nerve is exposed to IOP anteriorly and CSF posteriorly and changes in the translaminar pressure, the difference in pressure between IOP and the CSF pressure, may damage the optic nerve and consequently contribute to glaucomatous damage.(83) For example, if the relative pressure from CSF is higher, swelling of the optic nerve may occur.(84) Conversely, if the relative pressure from IOP is higher, then cupping of the optic nerve may occur. In both situations, an imbalance in the homeostasis of IOP and CSF pressure can lead to damage to the optic nerve. Moreover,

findings from a case-control study comparing POAG cases with nonglaucomatous controls observed CSF pressure was significantly negatively associated with VCDR ($P < 0.0001$).(84) That is, lower CSF pressure was associated with larger VCDR. The investigators also observed higher translaminar pressure was significantly associated with larger VCDR ($P < 0.0001$). And lastly, POAG cases had a significantly lower mean CSF pressure compared to nonglaucomatous controls, 9.2 ± 2.9 mmHg vs 13.0 ± 4.2 mmHg, respectively ($P < 0.00005$). These findings suggest higher cerebrospinal fluid pressure, which aids in counterbalancing IOP and reducing the translaminar pressure difference, may reduce possible glaucomatous damage to the optic nerve and subsequently glaucoma.(83) Taken together, higher BMI may protect against glaucomatous damage to the optic nerve, e.g. larger VCDR, by equalizing IOP through CSF.

Given the epidemiological evidence as well as the biological plausibility for the role BMI has on VCDR, BMI is a strong candidate for conducting a gene-environment interaction analysis of VCDR. Therefore, we aim to perform the first genome-wide gene-environment interaction analysis of body mass index on VCDR in Latinos. We hypothesize we will identify novel genetic variants associated with VCDR that will either increase or decrease VCDR in subsets of Latino study participants based on BMI. Findings from the examination of G×E interactions for VCDR may aid in further uncovering the missing heritability for this quantitative trait of glaucoma and potentially may lead to public health screening and intervention strategies to reduce the prevalence of this disease.

C. Genetic Risk Scores and Vertical Cup-Disc Ratio

The morphology of the optic disc is commonly assessed during routine ophthalmic examination to monitor and diagnose multiple ocular diseases, including glaucoma. In particular, VCDR is a useful clinical measurement to identify glaucomatous damage to the optic nerve.

Accordingly, identifying factors that affect VCDR will not only aid in uncovering the biological mechanisms regulating this ocular trait, but may also assist in predicting ocular disease.

Large-scale epidemiological studies have identified numerous factors associated with a larger VCDR, including higher IOP and lower BMI.(8, 80, 81) Additional factors have been identified but findings from studies have yielded inconsistent results. Despite several studies identifying male gender to be positively associated with VCDR,(80, 81) other studies found no significant differences between genders.(57, 85) Similarly, multiple studies identified increasing age was associated with VCDR,(8, 80, 81, 86) whereas other studies found no association.(57, 85) Diastolic blood pressure has been reported to be both positively(81) and negatively(80) associated with VCDR. Although several conventional risk factors have been identified, these ocular and systematic traits explain less than 4% of the variation in VCDR, indicating that other factors may contribute to this ocular trait.(80)

Vertical cup-disc ratio has also demonstrated to be partially determined by genetic factors, with heritability estimates of 48% to 66%.(59, 60) Genome-wide association studies have identified numerous loci associated with VCDR, including *ATOH7*, *SCYL1, SIX1*, and *CHEK2*.(27, 61, 63) Despite the identification of multiple VCDR associated genetic variants, each variant confers only a small to modest effect and individually have limited predictive power. Genetic risk scores (GRSs) evaluate the joint genetic effect of individual genetic variants by aggregating these effects into a single measure. A previous study examined whether such a polygenetic model exists for VCDR and investigated whether the polygenetic model can predict POAG.(87) In a three-phased study design, the investigators first used the Rotterdam Study I (RS-I), Rotterdam Study II (RS-II), and Rotterdam Study III (RS-III) to develop a GRS for VCDR. During the second phase, the investigators assessed whether the constructed GRS can predict VCDR in an independent cohort, the Erasmus Rucphen Family (ERF) study. And in the third phase, the investigators determined whether this GRS for VCDR could predict POAG using POAG cases and controls from the RS-I cohort. The Rotterdam Studies, RS-I, RS-II, and RS-III,

were originally conducted in Ommoord, Netherlands, and consisted of 7,983 (55 years of age and older), 3,011 (55 years of age and older), and 3,392 (45 years of age and older) study participants, respectively. Additionally, the ERF study is a family based study consisting of a genetically isolated population of more than 3,000 participants from the Netherlands. Ophthalmic assessment for RS-I and RS-II was conducted at baseline and at follow-up where multiple ocular measurements were obtained, including simultaneous stereoscopic photography analyzed by ImageNet. RS-III and ERF followed similar procedures, except Heidelberg Retina Tomograph 2 was used for optic nerve head imaging. In both instances, the built-in software calculated VCDR.

Study participants in RS-I, RS-II, and RS-III were genotyped on the Illumina Infinium II HumanHap array. In comparison, numerous genotyping platforms were used for the ERF study participants, including Illumina 6k, Illumina 318k, Illumina 370k, and Affymetrix 250k. These genotypes were merged and imputed using reference panels from the Northern Europeans from Utah population in HapMap. If measurements were available for both eyes, the measurement from one eye was randomly selected for further analysis. Risk scores were generated based off of SNP significance levels from a meta-analysis GWAS of the discovery cohorts (e.g., $P < 10^{-10}$, $P < 10^{-9}$, $P < 10^{-8}$, etc.). For each threshold, the effect size from each SNP was multiplied by the number of risk alleles and the average of all weighted scores yielded the final risk score per study subject. Linear regression was performed with VCDR as the outcome and the GRS as the main effect, adjusting for age, gender, and optic disc area, and the proportion of variation in VCDR explained by the GRSs were calculated. In a similar manner, logistic regression was performed to assess the relationship between the GRSs and POAG.

In the final analysis, 9,326 subjects were included in the discovery dataset and 1,646 subjects in the replication dataset. With age and gender in the model, only 0.3% of the variation in VCDR was explained. When the GRS containing SNPs with $P < 10^{-10}$ was entered into the model with age and gender, an additional 0.1% of VCDR variation was explained. However, as

the number of SNPs in the GRS increased, the amount of variation also increased up until $P < 10^{-2}$. At a threshold of $P < 0.2$, the inclusion of the GRS in the model with age and gender explained an additional 1.0% of the total variation in VCDR. To determine if VCDR and POAG share a common genetic component, the authors tested the association between the GRSs for VCDR and POAG. Together, age and gender explained 4% of the variation in POAG. When the GRS from the first threshold ($P < 10^{-10}$) was introduced into the model, only an additional 0.3% of the variation was explained. Analogous to VCDR, as the number of SNPs in the GRS increased, the amount of variation explained increased as well. The GRSs continued to explain additional variation in POAG up until $P < 0.3$, explaining 4.7% of the variation, which was more explained variation than both age and gender combined. Results from this study suggest a polygenetic model exists for VCDR and this polygenetic model may share common genetic origins with POAG. These findings further exemplify the utility of genetic variants with smaller effect sizes and higher significance levels, i.e. $P < 0.2$ and $P < 0.3$, in elucidating the genetic architecture of traits beyond those genetic variants reaching traditional genetic significance levels ($P < 5 \times 10^{-8}$). Moreover, this suggests additional genetic variants contributing to these traits remain to be identified.

A second study in a multiethnic Asian population performed a similar analysis using previously reported SNPs for VCDR and IOP.(88) The Singapore Epidemiology of Eye Diseases study was a population-based cross-sectional study of Malays, Indians, and Chinese living in Singapore. Study participants received a standardized interview and ocular examination. Optic nerve head imaging was performed using the Heidelberg Retina Tomograph II, and VCDR was calculated using the built-in software. The mean VCDR from both eyes were used for downstream analysis. Study subjects were genotyped using the Illumina Human610-Quad BeadChip, and genotype imputation was performed using MaCH with CEU + JPT (Japanese in Tokyo, Japan) + CHB (Han Chinese in Beijing, China) + YRI (Yoruba in Ibadan, Nigeria) HapMap reference panels. Using 18 previously reported loci associated with VCDR at the time

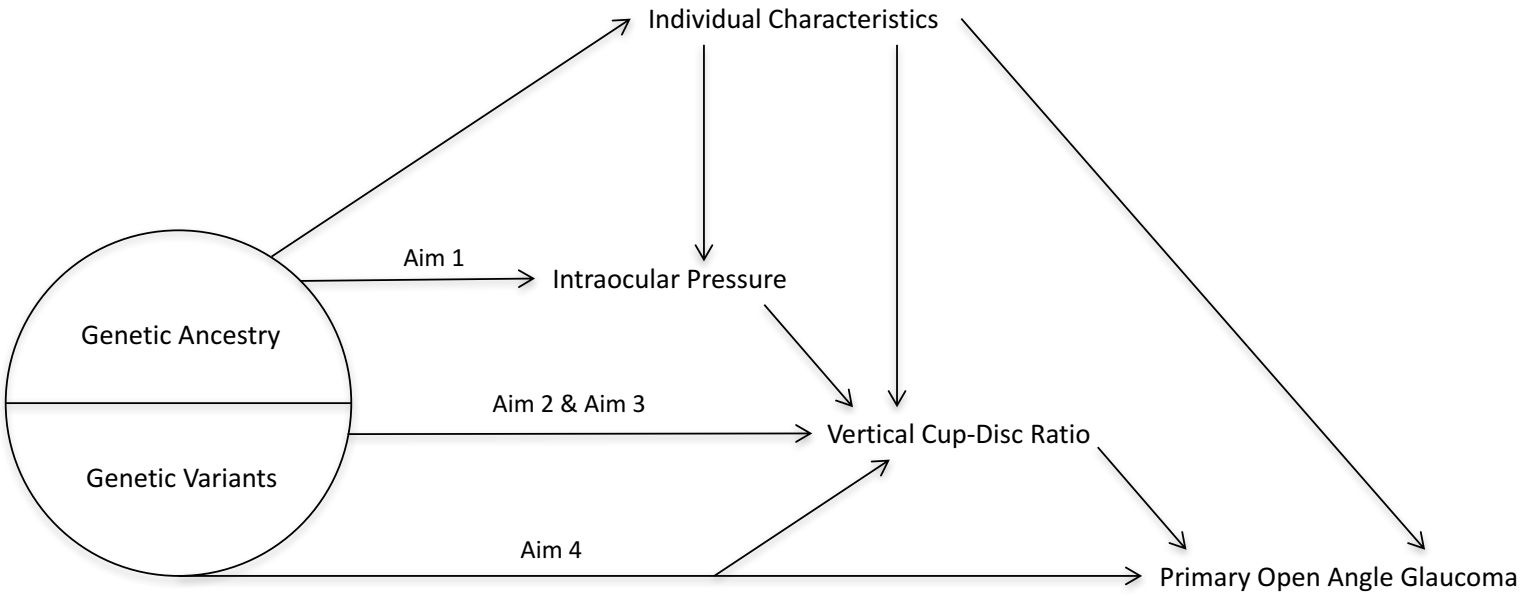
of the study, the investigators identified the most significant SNP within ± 100 kb of the index SNP in their dataset and generated a VCDR GRS weighted by their effect estimates for POAG.(63) The investigators also constructed a GRS for IOP using 7 previously associated SNPs using a similar approach.(89) Alleles for the SNPs were coded in a manner that risk alleles conferred an increased risk for POAG and the summation of these weighted SNPs was taken for each study participant. Logistic regression was conducted to examine the association between tertiles of VCDR GRS on POAG, adjusting for numerous covariates. To evaluate the discriminatory ability for POAG between traditional clinical risk factors and the VCDR and IOP GRSs, receiver operating characteristic analyses were performed and area under the curve was calculated. Overall, a higher VCDR GRS was significantly associated with greater odds of POAG ($P_{trend} = 7.9 \times 10^{-5}$). Compared to individuals in the lowest tertile of the VCDR GRS, those in the highest tertile had 2.31 [95% CI: 1.50-3.55] greater odds of POAG. The addition of the VCDR and IOP GRS to a model with traditional risk factors conferred a borderline significant increase in the AUC ($P = 0.06$), increasing from 0.72 [95% CI: 0.67-0.76] to 0.74 [95% CI: 0.70-0.78], representing a moderately discriminating capacity for POAG. Findings from this study demonstrate the combined effect of VCDR SNPs is associated with POAG for which individuals with more VCDR risk alleles exhibit greater odds of POAG.

While these studies provide evidence of the cumulative effect of VCDR genetic variants on VCDR and POAG, these studies were conducted in individuals of European and Asian descent and thus, may not be generalizable to other ethnic groups. Latinos, an underrepresented ethnic group in ocular genetics research, also exhibit a high prevalence of POAG and identifying factors associated with this disease may aid in reducing the occurrence of POAG.(10) As such, examining the association between an aggregate measure of genetic risk and VCDR among Latinos will further our understanding of the determinants of this trait. Additionally, the generation of GRSs for an endophenotype of POAG will enable an opportunity to evaluate whether the addition of this genetic information improves the discriminatory ability

for POAG as compared to traditional risk factors. We aim to construct GRSs based on VCDR associated SNPs and evaluate the associations between the GRSs on VCDR and POAG in a Latino population. We hypothesize the GRSs will be associated with both VCDR and POAG and will further increase the discriminatory ability for POAG.

In summary, Figure 1 depicts the conceptual model for this dissertation and corresponding aims. As the two most clinically relevant quantitative traits of POAG, this dissertation will identify genetic factors associated with IOP and VCDR by determining whether there is an association between IOP and genetic ancestry, identifying novel genetic variants and replicating previous associations with VCDR via a GWAS, further identifying novel genetic variants associated with VCDR through a genome-wide gene-environment interaction analysis of BMI, and constructing and evaluating genetic risk scores for VCDR, and determining whether these GRSs improve the discriminatory ability for POAG. Findings from this dissertation will further our understanding of the genetic architecture of these quantitative traits and the biological mechanisms influencing POAG pathogenesis and progression.

Figure 1. Conceptual model and dissertation aims for the associations between genetic factors and quantitative ocular traits of primary open angle glaucoma.



Aim 1: Genetic Ancestry and Intraocular Pressure
Aim 2: Genome-Wide Association Study of Vertical Cup-Disc Ratio
Aim 3: Genome-Wide Gene-Environment Interaction Analysis of Body Mass Index and Vertical Cup-Disc Ratio
Aim 4: Genetic Risk Score of Vertical Cup-Disc Ratio and Association with Primary Open Angle Glaucoma

III. MATERIALS AND METHODS

A. Ethics Statement

The institutional review boards at the University of Illinois at Chicago, the University of Southern California Health Sciences Campus, and the Los Angeles Biomedical Research Institute at Harbor-University of California, Los Angeles approved the following research. All clinical investigations were conducted according to the principles outlined in the Declaration of Helsinki. All study participants provided written, informed consent.

B. Study Sample

This dissertation research was conducted using data collected from the Los Angeles Latino Eye Study (LALES), the largest population-based epidemiological study exploring the prevalence of visual functioning, ocular disorders, and visual impairment in Latinos.(90) Latinos living in and around 6 census tracts of La Puente, Los Angeles County, California were recruited for this study. This study area was selected due to the high proportion of Latino residents and the age distribution of Latinos in this geographical area was similar to that of Latinos in the United States. Eligibility for enrollment into the study was determined by: (1) self-description of Latino or Latino heritage, (2) at least 40 years of age during the primarily assessment, and (3) a resident living in the identified census tracts. Initially, 10,663 individuals were screened for enrollment, of which 7,789 met the eligibility criteria for the study. Eligible study participants were administered an in-home questionnaire, collecting information regarding general health, demographic information, medication use, and history of ocular disease. Study participants then received an in-clinic medical and ocular examination, as well as an in-clinic interview. During the clinical examinations, numerous clinical and ocular parameters were obtained, including height, weight, blood pressure, visual acuity, intraocular pressure, central corneal thickness, and fundus photographs. Blood was drawn at the baseline clinical visit and stored for additional analyses. Those study participants unable to complete the in-clinic examinations were asked to undergo

an in-home examination. A total of 6,357 Latinos completed an ocular examination and were included for downstream analysis, yielding a participation rate of 82% (6,357/7,789). Of those not included in the final study sample, 908 refused to participate and 524 completed the in-home interview but not a clinical examination. Study participants who completed an in-home interview were younger and more likely to be female compared to those who did not complete an in-home interview. All study participants were 40 years of age or older at the time of the baseline examination.

C. Ocular Phenotype Measurements

1. Intraocular Pressure

Measurements of IOP were obtained using Goldmann applanation tonometry (Haag-Streit, Bern, Switzerland) for each study participant in a randomly assigned order. Three measurements were collected from each eye, first measured in the right eye and then the left eye. The average of the three measurements for each eye was calculated, yielding a single IOP measurement for each eye. The attainment of three IOP measurements has previously been shown to yield more reliable results, both intra-observer and inter-observer, compared to a single measurement.(49) The left and right eye measurements were averaged to obtain the final IOP measurement for each study subject. If IOP measurements were collected for only one eye, the average of these measurements was taken and used as the final measurement.

2. Vertical Cup-Disc Ratio

Stereoscopic optic disc photography was performed using the Topcon TRC 50EX Retinal Camera (Topcon Corporation of America, Paramus, NJ) with Ektachrome 100 film (Kodak, Rochester, NY). Photographs were examined using a stereoscopic viewer (Asahi viewer, Pentax, Englewood, CO). A board-certified ophthalmologist determined VCDR measurements for the right and left eyes. The average between the right and left eyes was

calculated and used as the final measurement for study participants. If a VCDR measurement was available for only one eye, this measurement was substituted for the final VCDR measurement.

3. Primary Open Angle Glaucoma

The identification of POAG cases occurred in a 3-phase process. In the first phase, two glaucoma specialists evaluated all clinical data, including family history of glaucoma, any history of glaucoma, treatment of glaucoma, and history of other ocular disease. During the second phase, the two glaucoma specialists assessed visual field and optic disc photographs to determine either the presence or absence of POAG in each eye using pre-specified definitions of POAG.(10) The specialists evaluated these data for study participants independently of each other. For the instances in which the specialists were in agreement, the diagnosis was assigned to each specific eye for each study participant. In the third phase, if there was disagreement between the two specialists, a third glaucoma specialist reviewed the data to make a diagnosis. Agreement of two out of three specialists was required in order to diagnose POAG for each eye.

Primary open angle glaucoma cases were defined in several ways. The main definition for POAG cases included visual field abnormality and evidence of optic disc damage consistent with glaucoma with an open angle in at least one eye. POAG cases were also defined as having one of the following in the presence of an open angle: 1) visual acuity ≤ 20/200 with a cup-disc ratio of 1.0 at the end stage of POAG; 2) irregular visual field in at least 1 test that is consistent with glaucomatous visual field defects with no indication of optic disc damage; 3) optic disc damage consistent with glaucoma with no indication of visual field irregularity; and 4) visual field irregularities and optic disc damage groupings that are consistent with glaucoma.(10)

D. Genotype Data

1. Genotyping and Quality Control

For the first aim, a total of 3,929 Latinos recruited from the LALES were genotyped through the Mexican American Glaucoma Genetic Study (MAGGS). Study participants were genotyped using the Illumina OmniExpress BeadChip Kit (730,525 markers; Illumina, Inc., San Diego, CA). The Genotyping Laboratory of the Institute for Translational Genomics and Population Sciences at the Los Angeles Biomedical Research Institute at Harbor-UCLA conducted the genotyping for this study. Single nucleotide polymorphisms (SNPs) were called using the software Illumina GenomeStudio (v2011.1; Illumina, Inc., San Diego, CA). A genotyping call rate of less than 97% was used to exclude study participants from further analysis due to low quality genotype data. The program PLINK (v1.07)(91) was used to further perform quality control on the genotype data. Single nucleotide polymorphisms were removed from downstream analysis if the genotyping call rate was less than 95%, the minor allele frequency (MAF) was < 1%, or the Hardy-Weinberg equilibrium *P* values < 10^{-6}. Study participants were additionally removed if there were inconsistencies between reported sex and genetically inferred sex, unexpected duplicates, and participants receiving any glaucoma medical treatment or IOP lowering medication. After applying the above quality control parameters, 619,712 SNPs and 3,541 individuals remained for further analysis.

For the second, third, and fourth aims, additional Latino genotype samples became available and were included for downstream analysis. The genotype data and quality control parameters have been described elsewhere.(92) A total of 4,996 Latinos recruited from the LALES were genotyped through the MAGGS. Study participants were genotyped using either the Illumina OmniExpress BeadChip Kit (730,525 markers; Illumina, Inc., San Diego, CA; n = 4,278) or the Illumina Hispanic/SOL BeadChip (~2.5 million markers; Illumina, Inc. San Diego, CA; n = 718). A genotyping call rate of less than 97% was used to exclude study participants from further analysis due to low quality genotype data. The program PLINK (v1.90)(93) was

used to further perform quality control on the genotype data. Overlapping SNPs between the two chips were retained for analysis. Single nucleotide polymorphisms were excluded if the genotyping call rate was < 95%, MAF < 1%, or the Hardy-Weinberg equilibrium *P* values < 10^{-6}. Study participants were removed if there were inconsistencies between reported sex and genetically inferred sex and unexpected duplicates. These quality control parameters yielded 576,798 SNPs and 4,549 study participants for downstream analysis. To aid in the genotype imputation process, SNPs were coded on the forward strand.

2. Genetic Ancestry Estimation

For aim 1, estimation of genetic ancestry was performed using the program STRUCTURE, a Bayesian clustering algorithm using a Markov Chain Monte Carlo method.(94, 95) In order to make inferences of the genetic ancestry estimates from the Latino sample, three reference populations of known ancestry were included during the ancestry estimation process. Specifically, unrelated individuals of Northern European ancestry from Utah, US (n = 87) and West Africa (Yoruba in Ibadan, Nigeria; n = 88) from the 1000 Genomes Project (96) and Native Americans (n = 105).(97) After merging these datasets together, 5,000 random autosomal SNPs were selected to estimate genetic ancestry. The random selection of SNPs across the genome has been shown to reliably estimate genetic ancestry in various ancestral populations.(98) We used 10,000 burn-ins and 10,000 iterations when running STRUCTURE. Prior to running the program, we specified 3 reference populations. After running the program, each study participant received estimated proportions of European, African, and Native American ancestry with the summation of these ancestries equaling 1. This procedure was performed five times and the average of these runs for each inferred ancestry was used for downstream analysis.

3. Genotype Imputation

For aims 2, 3, and 4, genotype imputation was performed to increase the genomic coverage by imputing SNPs not directly genotyped on the genotyping chips. To reduce the computational time for the imputation process, the program SHAPEIT2 (99) was used to phase the genotype data. Using the 1000 Genomes Project reference panels (phase 1, version 3), genotype imputation was performed using the program Minimac3.(100) The reference panels used for imputation contain 39.7 million variants, substantially increasing the number of variants to be analyzed in downstream analyses. For this Latino sample, reference panels consisting of CEU + YRI + AMR (Caucasians of European ancestry, Yoruba, and a combination of Mexican, Puerto Rican, and Colombian haplotypes) were used for imputation. A previous study demonstrated this reference panel yielded the highest imputation genotype quality for this ethnic population and as such, was used for these aims.(101)

Low quality imputed SNPs (i.e., Rsq < 0.80) and SNPs with low frequency (i.e., MAF < 1%) were excluded. Rsq is the squared correlation between the true, unobserved genotype and the imputed genotype and represents the quality of imputed SNPs. Imputed genotypes were coded as allelic dosages, estimated allele counts ranging from 0 to 2. After removal of low quality and low frequency SNPs, 6,844,888 SNPs remained for further analysis.

4. Construction of Genetic Risk Scores

For aim 4, unweighted and weighted genetic risk scores were constructed based on previously reported VCDR SNPs.(27, 61, 63) Risk alleles were defined as those alleles that result in an increase in VCDR. Given that several SNPs have been reported in multiple VCDR GWASs and to avoid duplication of SNPs in the GRSs, weights from the study with the largest sample size was used for these SNPs. This resulted in 68 previously reported SNPs to be used for the construction of the unweighted and weighted GRSs. Using a previous candidate gene approach,(88) we also constructed unweighted and weighted GRSs based on the lead SNP

(most significant SNP) from our GWAS results within ± 100 kb of the 68 previously reported SNPs. Moreover, unweighted and weighted GRSs were generated from our genome-wide association data using all independent SNPs (PLINK pruned at $r^2 = 0.2$) with $P < 1 \times 10^{-3}$.(102) The unweighted GRS was constructed as the summation of the risk alleles for these 68 SNPs, assuming each risk allele confers the same effect on VCDR. The weighted GRS was constructed by multiplying the VCDR risk allele by the effect estimate as reported in the respective study. The individual weighted genetic variants were then summed together to obtain the final weighted GRS.

E. Statistical Analysis

1. Genetic Ancestry and Intraocular Pressure

Summary statistics (i.e., frequency distributions and means) for the following clinical, socioeconomic, and genetic variables are reported: age, sex, body mass index (BMI), systolic blood pressure (SBP), type 2 diabetes (T2D), IOP, central corneal thickness (CCT), income, education, smoking status, proportion of African ancestry, and proportion of Native American ancestry. To avoid multicollinearity, the proportion of European ancestry was not included in the analysis due to this proportion equaling one minus the proportions of African ancestry and Native American ancestry. Simple linear regression was conducted to investigate the association between IOP and each variable. Multiple linear regression was performed to examine the relationship between the proportions of genetic ancestry and IOP, adjusting for potential confounders. For the simple linear regression and multiple linear regression models, IOP was the main response variable with Native American ancestry and African ancestry as the main effects. In order to find the most parsimonious model for the association under study, backwards selection was performed to retain significant covariates at a significance level of $P \leq 0.05$. Based on the observed differences in the prevalence of hypertension by race, potential effect modification was explored by introducing an interaction term between elevated SBP and

genetic ancestry in the final model. As a clinically meaningful categorization, systolic blood pressure was dichotomized to classify individuals as normotensive (SBP < 140 mmHg) or hypertensive (SBP ≥ 140 mmHg). Stratum specific estimates were reported if the interaction term was significant at a threshold of $P \leq 0.05$. We further conducted quantile regression to examine the effect of genetic ancestry on IOP across the range of observed genetic ancestry estimates at conditional quantiles, as compared to ordinary least-squares regression, which models the conditional mean. All statistical analyses were conducted using SAS v9.2 (SAS Inc., Cary, NC).

2. Genome-Wide Association Study of Vertical Cup-Disc Ratio

The program EIGENSOFT (103) was used to infer principal components of genetic ancestry for the Latino study participants. To make comparisons between our study sample with known ancestral populations, we incorporated reference panels containing unrelated Western Africans (Yoruba in Ibadan, Nigeria; n = 88) and Northern Europeans from Utah, US (n = 87) from the 1000 Genomes Project (96) and Native Americans (n = 105).(97) Retained as covariates for downstream analysis were the first four principle components. We assessed the effect of population stratification by calculating the genomic control inflation factor (104) and the distribution of the test statistics was visually inspected via a quantile-quantile (Q-Q) plot.

The study sample was divided into a discovery dataset (stage 1), consisting of unrelated individuals and a replication dataset (stage 2), consisting of related individuals. For both stage 1 and stage 2 analyses, an additive genetic model was assumed, where the effect size increases linearly with each additional risk allele. Due to a non-normal distribution of VCDR measurements upon visual examination of the histogram, VCDR measurements were inverse normally transformed. Linear regression was first conducted to investigate the association between SNPs and VCDR among study participants in the discovery dataset, adjusting for age, sex, and the first four principal components using PLINK (v1.90).(93) For the replication dataset,

we used linear mixed-effects models to examine the association between SNPs and VCDR, adjusting for age, sex, and the first four principal components of genetic ancestry and to account for relatedness between individuals using SAS (SAS Inc., Cary, NC). Compound symmetry was used for the covariance matrix and the empirical "sandwich" estimator was used during the replication analyses. We also used the software EMMAX (Efficient Mixed-Model Association Expedited)(105) to perform linear regression on both directly genotyped and imputed SNPs on the full dataset (the discovery and replication datasets combined), adjusting for age, sex, the first four principal components of genetic ancestry, and kinship. Compared to PLINK, which does not take into account any relatedness between study participants, EMMAX uses linear-mixed effects models to adjust for relatedness and population stratification. To account for genotype imputation uncertainty for imputed SNPs, allelic dosage was used in EMMAX. During the discovery dataset, SNPs with a $P < 1 \times 10^{-6}$ were retained and analyzed in the replication dataset. SNPs were declared genome-wide significant if $P < 5 \times 10^{-8}$ and suggestive if $P < 1 \times 10^{-6}$ during the full study sample analysis. We used the program simpleM to identify the number of independent tests to correct for multiple testing when replicating previously published loci.(106-108) For identified genomic regions during linear regression, we performed conditional analyses by including the lead SNP into the regression model as a covariate. This approach aids in determining whether additional genetic variants are associated with VCDR separate of the lead SNP. All graphing was performed using R (109) and LocusZoom (hg19 / 1000 Genomes Project 2014, AMR).(110)

We also performed pathway analysis to identify biological pathways associated with VCDR using the maximum number of unrelated study participants from the full dataset. The program Pedigree Reconstruction and Identification of a Maximum Unrelated Set (PRIMUS) was used to identify the maximum number of unrelated study participants in this Latino study sample.(111) This program uses pairwise identity by descent estimates from PLINK to generate undirected graphs representing family networks or pedigrees where nodes and edges denote

individuals and the pairwise relationship, respectively. Within each family network, the program selects the maximum unrelated individuals and combines these individuals to obtain the final list of study participants. Using directly genotyped SNPs, we mapped SNPs based on GRCh37/hg19 genomic positions to autosomal genes and also included a ± 50 kb gene boundary to capture regulatory elements and other functional elements associated with gene regulation. The program SKAT-O (112) was used to perform gene-set associations, adjusting for age, sex, and the first four principal components of genetic ancestry. Enrichment of biological pathways with genes associated with VCDR was conducted using the commercial software Ingenuity Pathway Analysis (IPA) by QIAGEN to evaluate canonical pathways.(113) Pathways were declared significant if the *P* value was ≤ 0.05 after correcting for multiple testing.

3. Genome-Wide Gene-Environment Interaction Analysis of Body Mass Index and Vertical Cup-Disc Ratio

The program EIGENSOFT (103) was used to infer principal components of genetic ancestry for the Latino study participants. To make comparisons between our study sample with known ancestral populations, we incorporated reference panels containing unrelated Western Africans (Yoruba in Ibadan, Nigeria; n = 88) and Northern Europeans from Utah, US (n = 87) from the 1000 Genomes Project (96) and Native Americans (n = 105).(97) Retained as covariates for downstream analysis were the first four principle components. We assessed the effect of population stratification by calculating the genomic control inflation factor (104) and generated a quantile-quantile (Q-Q) plot to visually inspect the distribution of the test statistics.

Traditional G×E interaction analysis consists of an exhaustive examination of an interaction for each SNP. For quantitative traits, this traditional approach has shown to have poor power and as such, alternative 2-step approaches were developed. Kooperberg and Leblanc developed a 2-step method that first screens all SNPs based on the genetic marginal effect at a predetermined significance threshold.(114) SNPs that pass this significance threshold

are formally tested for G×E interaction with a Bonferroni correction significance level. This approach reasons that most variants involved in interactions will exhibit some genetic marginal effect. Pare´ et al developed an alternative 2-step approach that first screens for variance heterogeneity across genotypes by using Levene's test for homogeneity of variance for all SNPs.(115) A subset of SNPs that pass a significance threshold are formally tested for G×E interaction with a Bonferroni corrected *P* value. This approach rationalizes that if the effect of a quantitative trait loci depends on an environmental factor, the variability of the quantitative trait in those with a risk allele will differ from the variability of the quantitative trait in those without the risk allele. A third 2-step approach is a modified approach of Pare´'s method in which the residuals from a linear regression model of the quantitative trait regressed on the environmental factor are first obtained and then examined using Levene's test of variance heterogeneity.(116) Screened genetic variants that reach a significance threshold are then examined for G×E interaction in the second step. This approach aids to remove the correlation between the variance estimator and the G×E interaction analysis in the presence of a marginal environmental effect. The final 2-step approach recently proposed by Zhang et al first estimates the residuals from the quantitative trait regressed on the environmental factor and then combines the *P* values from the marginal genetic scan with the *P* values from the test of variance heterogeneity from the residuals using Fisher's method.(116) Gene-environment interaction analyses are then conducted on the screened SNPs. This approach combines all the information from the previously proposed methods.

To evaluate each of these G×E methods for a quantitative trait, a simulation study was performed to compare the Type 1 error rate and power of these methods.(116) Results from the simulation demonstrated all the methods maintained similar Type 1 error rates when there was no marginal effect for the environmental factor but when the marginal effect of the environmental variable increased, the Type 1 error rate increased for the method proposed by Pare´et al. During the power comparison in the presence of one interaction term, Pare´'s

method and the traditional G×E interaction analysis were the least and second least powerful methods, respectively. Furthermore, the power of these methods was independent of the marginal genetic effect. The power of the approaches proposed by Kooperberg and Leblanc and Zhang et al were the most and second most powerful methods, respectively. Furthermore, the power of these approaches increased as the marginal genetic effect increased, with the Kooperberg and Leblanc method remaining more powerful across various genetic effect thresholds. Results from this study demonstrate the Kooperberg and Leblanc is the most powerful method to detect G×E interactions for a quantitative trait outcome in the presence of one interaction term and the power from this method increases with increasing marginal genetic effect. Based on the findings from this study, the Kooperberg and Leblanc method will be used to identify G×E interactions for the genome-wide gene-environment interaction analysis of BMI and VCDR.

The study sample for this analysis consisted of the maximum number of unrelated study participants from the full dataset, as previously described. Due to a non-normal distribution of VCDR measurements upon visual examination of the histogram, VCDR measurements were inverse normally transformed. Height and weight measurements obtained during clinical visits were used to calculate BMI. BMI was used to classify study participants as under / normal weight (BMI < 25 kg/m^2) and overweight / obese (BMI ≥ 25 kg/m^2). The program G×EScan was used to perform statistical analyses.(116) Linear regression was first conducted in Step 1 to investigate the association between SNPs and VCDR adjusting for age, sex, and the first four principal components of genetic ancestry. Single nucleotide polymorphisms with a *P* value ≤ 0.05 were carried forward to be formally analyzed for G×E interactions in Step 2. During this step, linear regression was performed to examine interactions between SNPs and BMI by including a SNP×BMI interaction term for all SNPs identified in Step 1, adjusting for age, sex, and the first four principal components. Interaction terms with a *P* value < 0.05 / M, where M represents the number of SNPs identified in Step 1, were declared genome-wide significant in

Step 2. Stratified analyses were conducted for SNPs with an interaction *P* value < 1×10^{-4}. For the analysis of imputed SNPs, allelic dosage was used in G×EScan to account for genotype imputation uncertainty. Conditional analyses were performed for the top identified genomic regions during linear regression by including the lead SNP into the regression model as a covariate. This approach aids in determining whether additional genetic variants are associated with VCDR separate of the lead SNP. All graphing was performed using R (109) and LocusZoom (hg19 / 1000 Genomes Project 2014, AMR).(110)

4. Genetic Risk Scores of Vertical Cup-Disc Ratio

Univariate analyses (i.e., frequency distributions and means) were conducted to describe the study sample characteristics for the maximum number of unrelated study participants as previously described. In addition to the GRSs, clinical variables, such as age, sex, BMI, SBP, CCT, IOP, and T2D, were included in this analysis based on previous associations with VCDR and potential socioeconomic and environmental confounders, including income, education, and smoking status, were also included. Measurements for VCDR were inverse normally transformed due to a non-normal distribution of this trait upon visual inspection of the histogram. Simple linear regression analyses were performed to assess the association between each of the aforementioned variables with VCDR. Multiple linear regression analyses were performed to examine the association between GRSs and VCDR, adjusting for significant covariates. For selection of the final regression model, stepwise selection was performed, retaining significant variables with a $P \leq 0.05$. The additional amount of variance explained by the GRSs were reported.

To assess the relationship between the GRSs and POAG, logistic regression analyses were conducted. Quintiles of GRSs were generated to compare study participants with the lowest number of risk alleles to participants with greater number of risk alleles on the odds of POAG. Receiver operating characteristic (ROC) analyses were performed and the area under

the curve (AUC) was calculated and used to compare the improvement in the discriminatory ability of GRSs for POAG. We first calculated the AUC for a model containing only age and sex. We then included the GRSs into the model with age and sex, calculated the AUC, and compared the two AUCs. All statistical analyses were conducted using SAS v9.4 (SAS Inc, Cary, NC) and graphing was performed using R.(109)

IV. RESULTS

Results and discussion for aim 1 (genetic ancestry and intraocular pressure) have previously been published and is cited below (see Appendix A for copyright statement).

Nannini D, Torres M, Chen YD, et al. African Ancestry Is Associated with Higher Intraocular Pressure in Latinos. Ophthalmology 2016;123(1):102-8.

Results and discussion for aim 2 (genome-wide association study of vertical cup-disc ratio) have previously been published and is cited below (see Appendix B for written permission).

Nannini DR, Torres M, Chen Y-DI, et al. A genome-wide association study of vertical cup-disc ratio in a Latino population. *Invest Ophthalmol Vis Sci*. 2017;58:87-95.

A. Genetic Ancestry and Intraocular Pressure

1. Study Sample

Table I summarizes the study sample characteristics, as well as the simple linear regression results, for the variables included in this investigation. The mean ± standard deviation age of the study sample is 54.9 (10.5) years, and 40.6% of the study participants are males. The mean ± standard deviation of IOP, BMI, SBP, and CCT is 14.6 ± 2.8 mmHg, 31.0 ± 5.6 kg/m^2, 124.0 ± 19.0 mmHg, and 550.3 ± 33.7 μm, respectively. The average proportion of African ancestry and Native American ancestry is 3.1% ± 4.1% and 44.1% ± 14.7%, respectively. Additionally, for simple linear regression modeling, a majority of the covariates are significantly associated with IOP, including age ($P < 0.0001$), gender ($P < 0.0001$), BMI ($P < 0.0001$), SBP ($P < 0.0001$), CCT ($P < 0.0001$), T2D ($P < 0.0001$), smoking status ($P = 0.006$), income ($P = 0.045$), and African ancestry ($P = 0.002$). Education level and Native American ancestry are not associated with IOP.

TABLE I. STUDY SAMPLE CHARACTERISTICS AND SIMPLE LINEAR REGRESSION RESULTS

	Participants (n = 3,541)	*P*
IOP, mmHg	14.6 (2.8)	-
Age, yrs	54.9 (10.5)	<0.0001
Gender, male	40.6%	<0.0001
BMI, kg/m^2	31.0 (5.6)	<0.0001
SBP, mmHg	124.0 (19.0)	<0.0001
CCT, μm	550.3 (33.7)	<0.0001
T2D, yes	28.3%	<0.0001
Smoking Status		0.006
Never	62.2%	
Former	24.6%	
Current	13.2%	
Education level, yrs		0.72
≤ 6	44.6%	
7-11	21.9%	
≥ 12	33.5%	
Income level[a]		0.045
<$20,000	50.1%	
$20,000-$40,000	35.8%	
>$40,000	14.1%	
NA ancestry, %	44.1 (14.7)	0.96
African ancestry, %	3.1 (4.1)	0.002

Abbreviations: IOP, intraocular pressure; yrs, years; BMI, body mass index; SBP, systolic blood pressure; CCT, central corneal thickness; T2D, type 2 diabetes; NA, Native American.
[a] Missing income for 440 study participants.

2. Multiple Linear Regression Results

Table II presents the results from the multiple linear regression modeling. Model 1 shows the results from the full linear regression model containing all covariates. With all of the covariates entered into the model, only age (P = 0.0018), gender (P = 0.0001), BMI (P = 0.0002), SBP (P < 0.0001), CCT (P < 0.0001), T2D (P < 0.0001), and African ancestry (P = 0.0017) were associated with IOP. Smoking status, income, education, and Native American ancestry were not significant in the full model at a significance cutoff of $P \leq 0.05$. Age, gender, BMI, SBP, CCT, T2D, and African ancestry were significant after performing backwards selection. After adjusting for these significant predictors, African ancestry remained significantly

associated with IOP (*P* = 0.0005), shown in Model 2. As such, for every 1% increase in African ancestry in Latinos, there is a 0.038 mmHg increase in IOP. Additionally, African ancestry remained significant (*P* = 0.0037) when IOP was inverse normally transformed.

TABLE II. MULTIPLE LINEAR REGRESSION RESULTS FOR GENETIC ANCESTRY

	Model 1		Model 2	
	Beta	*P*	Beta	*P*
Age	0.015	0.0018	0.017	0.0002
Gender	-0.38	0.0001	-0.40	<0.0001
BMI	0.033	0.0002	0.035	<0.0001
SBP	0.023	<0.0001	0.022	<0.0001
CCT	0.019	<0.0001	0.018	<0.0001
T2D	0.68	<0.0001	0.66	<0.0001
Smoking Status	-0.12	NS	-	-
Income[a]	-0.021	NS	-	-
Education	0.061	NS	-	-
NA ancestry	0.003	NS	-	-
African ancestry	0.035	0.0017	0.038	0.0005

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; CCT, central corneal thickness; T2D, type 2 diabetes; NA, Native American.
[a] Missing income for 440 study participants.

3. Interaction Results

Blood pressure is often dichotomized into hypertension and normal tension to aid in interpretation. Hence, SBP was dichotomized into elevated SBP (SBP ≥ 140 mmHg) and normal pressure (SBP < 140 mmHg). Analysis of effect modification between African ancestry and elevated SBP resulted in a significant interaction term. Table III presents the final model for the association between African ancestry and IOP including the interaction term (*P* = 0.037). When stratified by elevated SBP, among individuals with normal blood pressure, for every 1% increase in African ancestry, there is a 0.033 mmHg increase in IOP (*P* = 0.003). In comparison, among

elevated SBP individuals, every 1% increase in African ancestry results in a 0.105 mmHg increase, a 3-fold increase in IOP (*P* = 0.008). Furthermore, African ancestry and its interaction with elevated SBP remained significant when IOP was inverse normally transformed with *P* = 0.021 and 0.044, respectively.

TABLE III. MULTIPLE LINEAR REGRESSION WITH AN INTERACTION TERM[a]

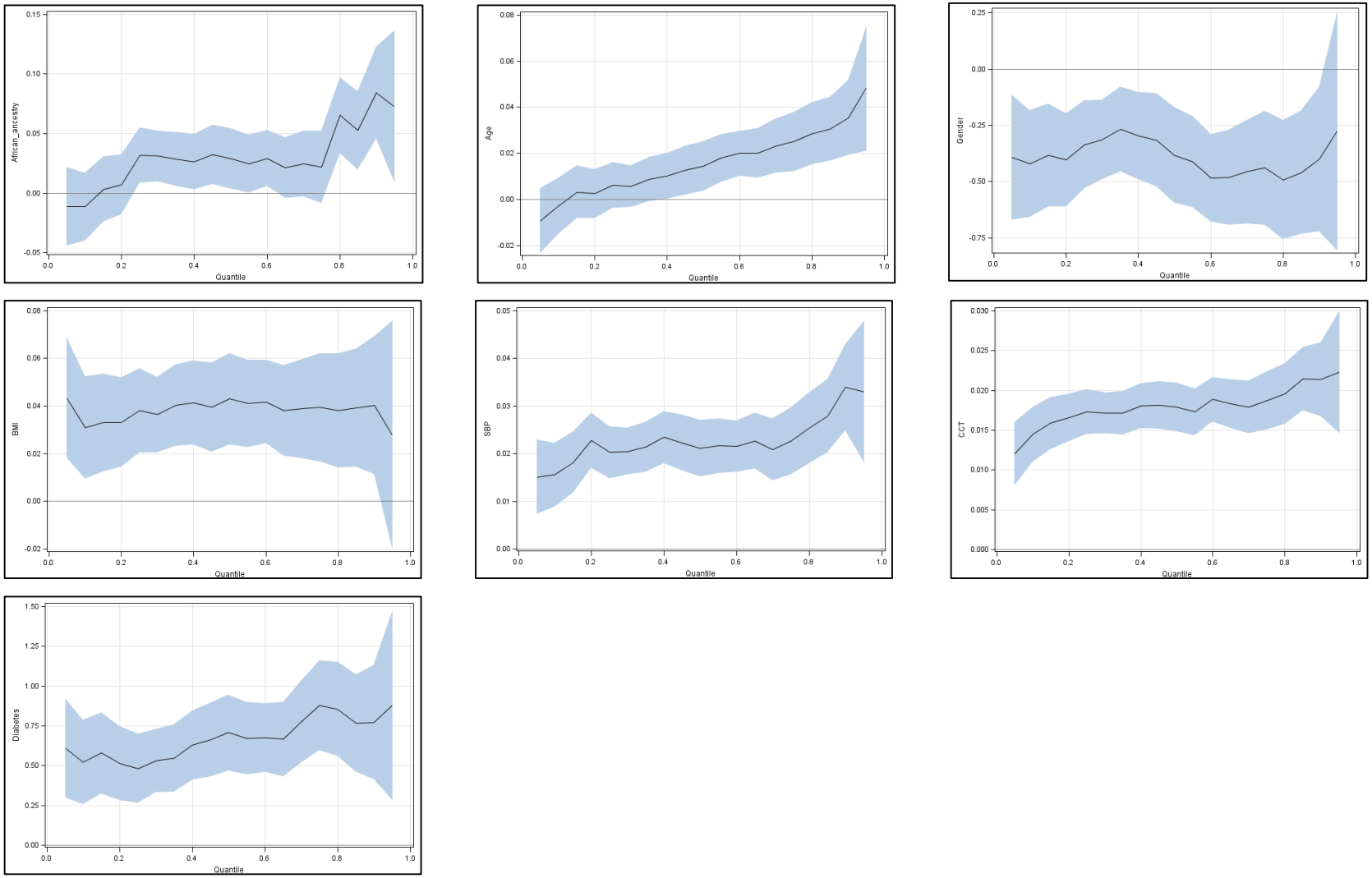
	Beta	*P*
Age	0.024	<0.0001
Gender	-0.36	<0.0001
BMI	0.042	<0.0001
Elevated SBP	0.52	0.001
CCT	0.018	<0.0001
T2D	0.71	<0.0001
African ancestry	0.033	0.004
African ancestry + Elevated SBP	0.077	0.037

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; CCT, central corneal thickness; T2D, type 2 diabetes.
[a] Systolic blood pressure dichotomized into normal pressure (systolic blood pressure < 140 mmHg) and elevated SBP (systolic blood pressure ≥ 140 mmHg).

4. Quantile Regression Results

Figure 2 presents the plots generated from quantile regression analysis for 7 covariates, namely, African ancestry, age, gender, BMI, SBP, CCT, and T2D. For each of the 7 coefficients, we plot 19 quantile regression estimates for the quantiles ranging from 0.05 to 0.95. As shown in the African ancestry plot, the overall pattern depicts a positive effect of African ancestry on IOP, especially in the upper tail of the ancestry-IOP distribution. The effect of African ancestry on IOP in the upper tail can be 3 times greater than that in the middle range of the distribution (0.075 vs. 0.025 mmHg). Furthermore, the upper tail of African ancestry has a larger effect on

Figure 2. Estimated parameters by quantile with 95% confidence interval.



Quantile regression plots for covariates in the final model, including African ancestry, age, gender, body mass index (BMI), systolic blood pressure (SBP), central corneal thickness (CCT), and diabetes. The x-axis and y-axis denote the quantile scale and the effect of a covariate on intraocular pressure (IOP) for a given quantile, respectively.

IOP than any other quantitative covariates, namely, age, BMI, SBP, and CCT. Interesting patterns were also observed for the other covariates included in the analysis. In general, advancing age has a steady positive effect on IOP. Males have lower IOP than females for any chosen quantile. The effect of BMI seems to be relatively flat, with a 0.04 mmHg increase in IOP in nearly all the quantiles. Both SBP and CCT show patterns of increasing effect on IOP, especially for the upper tail of the distribution. Across all quantiles in the distribution, T2D has a large and positive effect on IOP.

5. Discussion

To our knowledge, we are the first to report the relationship between African ancestry and IOP in a Latino population. Using data collected from LALES, we included 3,541 Latinos in this investigation. We identified a significant association between African ancestry and IOP in this sample of Latinos. After adjusting for covariates, increasing proportion of African ancestry was associated with increasing IOP in Latinos. Furthermore, the main association between IOP and African ancestry was modified by a significant interaction between African ancestry and elevated SBP. When stratified by elevated SBP, individuals with elevated SBP experienced a greater increase in IOP with increasing African ancestry. Both African ancestry and its interaction with elevated SBP serve as novel risk factors for IOP in Latinos and help to explain the variation in IOP at the individual level.

In addition to generalized linear models, we investigated the relationship with IOP using quantile regression. Compared with ordinary least-squares regression, which models the conditional mean of a response variable, quantile regression models the conditional quantiles of the response variable and is more robust. As such, this regression method gives more detailed patterns by providing a more complete picture of the relationship between variables and hence is more suitable when the change in response varies by quantiles. For the association between

IOP and African ancestry, there was a positive effect of African ancestry on IOP, especially at the upper tail of the ancestry-IOP distribution.

We also ran ADMIXTURE,(117) an independent genetic ancestry estimation program that can use a large amount of SNPs, to compare the association results of African ancestry and IOP. In general, the genetic ancestry estimates between STRUCTURE and ADMIXTURE were in agreement with a correlation coefficient greater than 0.97. Additionally, when genetic estimates from ADMIXTURE were substituted in for the STRUCTURE estimates, the association between IOP and African ancestry, as well as the interaction term, remained significant (data not shown). The large correlation coefficient, as well as similar significance levels, suggests our results are consistent and robust.

In the final model, African ancestry was significantly associated with a modest increase in IOP per percent increase in African ancestry. Gender and T2D, in comparison, had greater effects on IOP in this study. Previous studies have reported large effect sizes for both diabetes and gender on IOP, consistent with the effect sizes in our study.(118, 119) In relation to the rest of the variables included in this study, African ancestry exhibited a greater effect size than age, BMI, SBP, and CCT. These results illustrate that no single risk factor can fully explain an increase in IOP. Considering these findings, African ancestry represents a novel risk factor for increased IOP.

Although self-identified race is often used as a surrogate for genetic ancestry, for admixed populations this may not accurately capture the totality of an individual's genetic ancestry. Genetic ancestry has led to a better explanation of the phenotype variation at the individual level in Latinos. For example, Gao et al. identified a significant association between Native American ancestry and severe diabetic retinopathy in Latinos.(55) The current study presents another example that genetic ancestry should be considered in risk estimation in this admixed population. Halder et al. found that biogeographical ancestry was not a better predictor for cardiovascular disease risk than self-reported race and there were no differences in the

predictive power between both categorizations for 12 of the 15 outcomes when comparing African Americans and European Americans.(120) Similarly, Girkin et al. concluded that the added value of biogeographical ancestry over self-identified race is questionable for ocular phenotypes among African and European descent groups.(121) According to the classifications for race in the Census, which include White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander, Latinos do not have a clear racial category that best describes this population.(122) Although Halder et al. and Girkin et al. conclude that biogeographical ancestry is not a better predictor than self-identified race for various phenotypes, these studies performed analyses in European and African descendants, subjects from clear-cut race categories. Furthermore, Girkin et al. noted that the limited degree of admixture in the ADAGES cohort may have confined the ability to detect association with genetic ancestry within their African descendant groups. Both our investigation (55) and other studies (54, 97) confirmed that Latinos are typically a three-way admixture of Native American, European, and African ancestry. As such, instead of classifying our subjects as a single Latino group or misclassifying them into distinct racial groups, genetic ancestry is the only way to faithfully capture the ancestral makeup of this highly admixed population.

Our discovery is also consistent with traditional epidemiology results in IOP from different ethnic groups. Racial differences have been observed with regard to the variation in IOP, with individuals of African descent having higher IOP and Non-Hispanic Whites experiencing lower IOP, with Latinos and Native Americans tending to have IOP levels in between these two groups.(7, 8, 46, 123) Differences in the prevalence of hypertension by race have been observed, with Non-Hispanic Blacks having the highest prevalence of hypertension, followed by Non-Hispanic Whites, Hispanics, and Asians, respectively.(50) Additionally, a recent meta-analysis identified an increased risk of primary open angle glaucoma for patients with hypertension compared to those without.(124) It is unlikely that our association signals are owing to differences in glaucoma prevalence among different ancestral populations. The

prevalence of open angle glaucoma in the LALES cohort is 4.74%, of which 82% of cases have normal-tension with IOP measurements less than or equal to 21 mmHg.(10) Furthermore, our sample is from a single cohort of Latinos of Mexican origin rather than a combination of several distinct racial groups. Taken together, Latinos with a high percent of African ancestry, and who also have elevated SBP, maybe at risk for increased IOP and potentially, higher risk of glaucoma.

However, this study is not without limitations. Genetic ancestry estimates were determined using statistical methods and may contain errors. These errors, however, are likely to be random and irrelevant to the phenotype analyzed, potentially biasing our estimates toward the null. Additionally, while IOP increased with increasing African ancestry in general, other factors can cause elevated IOP other than ancestry. For example, in all other ethnic populations, there are individuals that experience high IOP, which clearly is not due to African ancestry. This study included participants who self-identified as Latino, primarily individuals of Mexican origin. It would be interesting to know how African ancestry plays a role in Latinos of different origins, such as in Puerto Ricans, who tend to have a higher proportion of African ancestry compared to Latinos of Mexican origin.(53) More research is needed to examine the relationship between genetic ancestry and IOP in this context. Both IOP and SBP are challenging phenotypes and fluctuate over time. As such, depending on the time of day these measurements were taken, temporal changes in SBP may misclassify an individual as either normal pressure or hypertensive based on a single reading and fluctuations in IOP may affect the study estimates. We do not have medical records for systemic medications for hypertension and steroid preparations. However, LALES cohort represents a low-income population of Latinos. Socioeconomic factors, such as income, may act as a barrier for accessing affordable medication for systematic diseases for this study population. Nevertheless, we excluded subjects with glaucoma treatment or IOP lowering medication. This study only used a subset (all genotyped subjects available so far) of LALES. We are extending our genotyping effort to the

entire LALES cohort through the MAGGS project. As genotyping progresses, we will have more power to detect additional association signals. To further evaluate any possible selection bias, we compared the equality of the IOP distribution with that for the entire LALES cohort using the Kolmogorov-Smirnov test. The *P* value was non-significant (*P* = 0.35), suggesting that our genotyped sample is an unbiased representation of all LALES subjects.

In conclusion, we identified a positive association between African ancestry and IOP in Latinos, after adjusting for known risk factors. Additionally, we identified a significant interaction between African ancestry and elevated SBP in regards to IOP with individuals with elevated SBP experiencing a larger increase in IOP with increasing African ancestry. To our knowledge, this is the first time that African ancestry and its interaction with elevated SBP have been associated with greater IOP in Latinos and they represent novel risk factors for elevated IOP in this admixed population.

B. Genome-Wide Association Study of Vertical Cup-Disc Ratio

1. Study Sample

Table IV presents the descriptive statistics for the overall study population, as well as the discovery (stage 1) and replication (stage 2) sets separately. For the entire study sample, the mean (SD) age was 54.8 (10.6) years, with the mean age of the discovery and replication sets as 54.2 (9.9) years and 56.9 (12.5) years, respectively. The proportion of females in the entire study was 58.9%: 56.5% in the discovery set and 68.0% in the replication set. Together, the average VCDR (SD) was 0.34 (0.18; range, 0.10–0.90), with the average of the discovery and replication sets as 0.34 (0.18; range, 0.10–0.90) and 0.35 (0.19; range, 0.10–0.90), respectively.

2. Genome-Wide Association Results

The genomic control inflation factor (104) was moderate, $\lambda = 1.03$. Figure 3 displays the Q-Q plot of the observed *P* values versus the expected *P* values. As seen in the plot, the

observed P values do not deviate from the null, except at the extreme tail. Both the genomic control inflation factor and the Q-Q plot indicate proper control of population stratification in this sample of Latino individuals.

TABLE IV. DESCRIPTIVE STATISTICS OF THE STUDY SAMPLE

Study	Sample Size	Females	Age, y, Mean (SD)	VCDR, Mean (SD)	VCDR Range
Discovery Set, stage 1	3,596	56.5 %	54.2 (9.9)	0.34 (0.18)	0.10 – 0.90
Replication Set, stage 2	941	68.0 %	56.9 (12.5)	0.35 (0.19)	0.10 – 0.90
Total	4,537	58.9 %	54.8 (10.6)	0.34 (0.18)	0.10 – 0.90

Figure 3. A quantile-quantile (Q-Q) plot of the $-\log_{10}$ (*P* values) for the 576,798 genotyped SNPs analyzed in the discovery set.

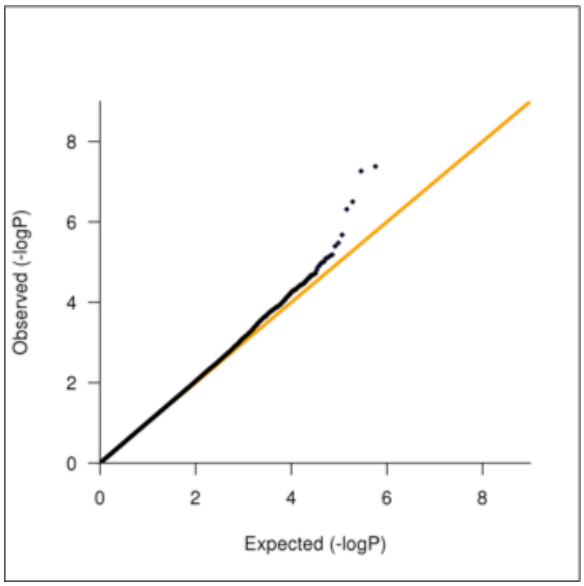
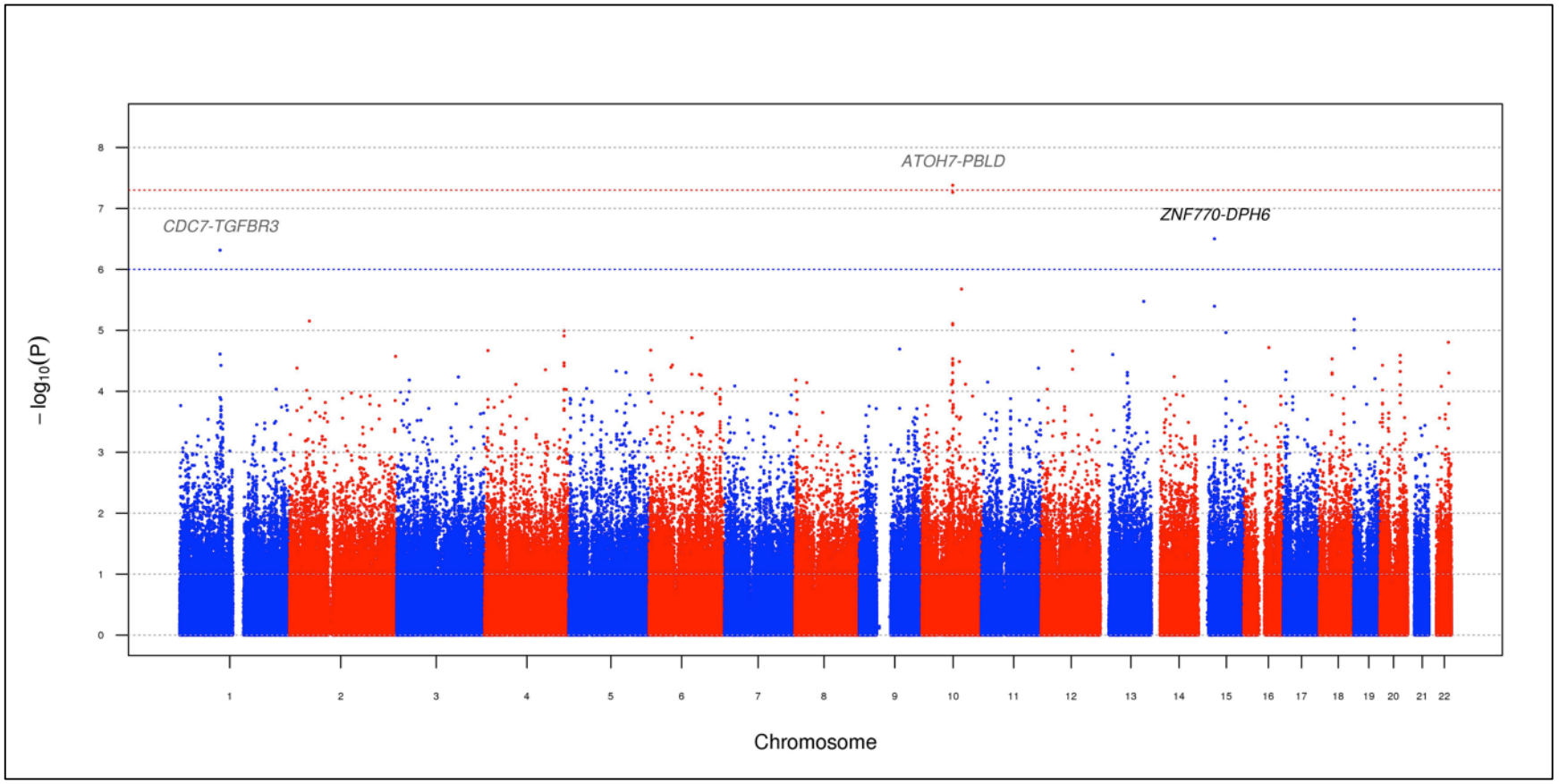


Figure 4 displays a Manhattan plot of the genome-wide *P* values from the discovery set association analysis. The results for the top SNPs ($P < 1 \times 10^{-6}$) are summarized in Table V.

One SNP, rs1900005 ($P = 4.17 \times 10^{-8}$, GRCh37/hg19 position 69,998,055), on chromosome 10q21.3 reached the GWAS significance level $P < 5 \times 10^{-8}$. The minor allele A (MAF = 0.36) was associated with a reduction in VCDR with β (SE) = -0.13 (0.02). This SNP is located 6.2 kb upstream of the *ATOH7* (atonal bHLH transcription factor 7) gene and 44 kb downstream of the *PBLD* (phenazine biosynthesis-like protein domain containing) gene. The second most significant SNP, rs7916697 ($P = 5.44 \times 10^{-8}$, GRCh37/hg19 position 69,991,853), is located 6.2 kb upstream of rs1900005 and was borderline GWAS significant. This SNP is situated in the 5' untranslated region (5' UTR) of *ATOH7*. The minor allele A of rs7916697 (MAF = 0.38) is associated with a decrease in VCDR with β (SE) = -0.13 (0.02). The third most significant SNP is rs16960773 ($P = 3.15 \times 10^{-7}$, GRCh37/hg19 position 35,604,502), located on chromosome 15q14 and situated 324 kb upstream of *ZNF770* (zinc finger protein 770) and 53.2 kb downstream of *DPH6* (diphthamine biosynthesis 6). The minor allele G (MAF = 0.08) is associated with a reduction in VCDR with β (SE) = -0.23 (0.04). The last most significant SNP is rs1192419 ($P = 4.85 \times 10^{-7}$, GRCh37/hg19 position 92,080,059), located on chromosome 1p22.1 and positioned 88.7 kb downstream of *CDC7* (cell division cycle 7) and 65.8 kb downstream of *TGFBR3* (transforming growth factor, beta receptor III). The minor allele A (MAF = 0.29) of rs1192419 is associated with an increase in VCDR with β (SE) = 0.13 (0.03).

We then analyzed these top SNPs in the replication set using linear mixed-effects models. As displayed in Table V, the direction of effect for the top four SNPs are consistent with the directions observed in the discovery set. The associations for three SNPs (rs1192419, rs7916697, and rs1900005) were strengthened when the full study sample was analyzed. In addition to rs1900005 remaining significant, rs7916697 also became genome-wide significant ($P = 1.97 \times 10^{-11}$) and rs1192419 became borderline genome-wide significant ($P = 9.56 \times 10^{-8}$) after analyzing the discovery and replication sets together. The SNPs rs7916697 and rs1900005 have previously been reported to be associated with VCDR in European and Asian individuals and the results reported in this study confirm these associations in a Latino

Figure 4. Manhattan plot displaying the $-\log_{10}$ (*P* values) for the association between VCDR and the 576,798 SNPs in the discovery set (stage 1).



The red and blue horizontal dotted lines indicate genome-wide significant associations ($P = 5 \times 10^{-8}$) and suggestive associations ($P = 1 \times 10^{-6}$), respectively. Previously reported and novel loci associated with VCDR are shown in dark grey (*CDC7-TGFBR3* and *ATOH7-PBLD*) and black (*ZNF770-DPH6*), respectively. SNPs are plotted by genomic position.

TABLE V. SUMMARY RESULTS FOR THE TOP RANKING GENOTYPED SNPS ASSOCIATED WITH VCDR IN LATINOS

						Discovery		Replication		Entire Sample
SNP	Chr	Position	Gene	A1/A2	MAF	β	*P*	β	*P*	*P*
rs1192419	1	92,080,059	*CDC7-TGFBR3*	A/G	0.29	0.13	4.85E-07	0.10	6.25E-02	9.56E-08
rs7916697	10	69,991,853	***ATOH7***	A/G	0.38	-0.13	5.44E-08	-0.20	4.07E-05	1.97E-11
rs1900005	10	69,998,055	*ATOH7-PBLD*	A/C	0.36	-0.13	4.17E-08	-0.21	1.97E-05	6.41E-12
rs16960773	15	35,604,502	*ZNF770-DPH6*	G/A	0.08	-0.23	3.15E-07	-0.11	1.95E-01	9.63E-07

SNPs with $P < 1 \times 10^{-6}$ in the discovery set are included in the table and were analyzed in the replication set.

Abbreviations: Chr, chromosome; A1/A2, allele 1/allele 2; MAF, minor allele frequency. Gene name is in boldface if the SNP is located inside the gene. SNP positions are according to GRCh37/hg19.

population.

3. Results From Imputed SNPs

To interrogate additional SNPs not directly genotyped, we performed genotype imputation on the full study sample. After retaining SNPs of high quality (Rsq ≥ 0.80), no additional genomic regions reached GWAS significance. Within the *ATOH7-PBLD* region, numerous imputed SNPs reached genome-wide significance, including several SNPs more significant than those directly genotyped. The most significant SNP in this region is rs56238729 ($P = 1.22 \times 10^{-13}$, Rsq = 0.98), located 9.8 kb downstream of rs7916697. This SNP represents a novel association with VCDR. The regional SNP association plot for the *ATOH7-PBLD* region is presented in Figure 5. Genotyped SNPs are plotted as squares, and imputed SNPs as circles.

4. Conditional Analysis

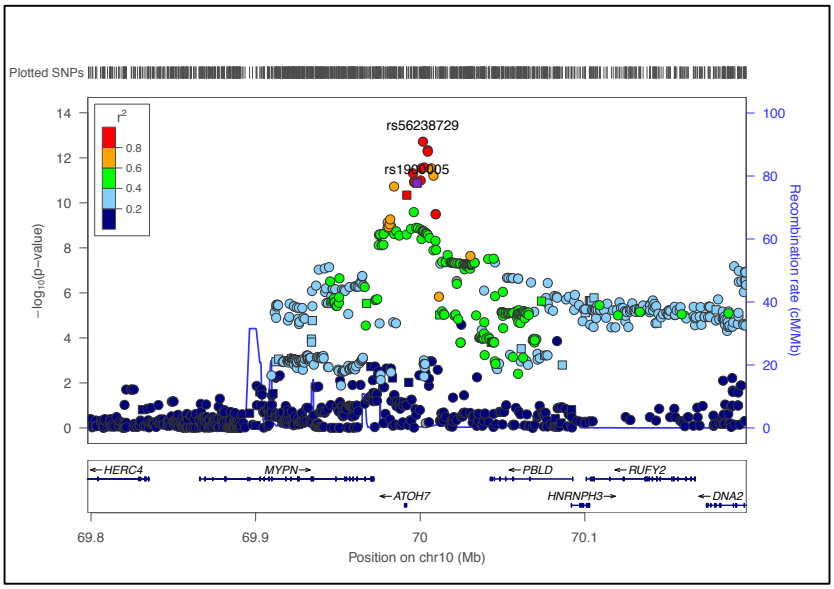
To determine whether additional SNPs contribute to the VCDR association, we conducted conditional analysis in the *ATOH7-PBLD* region. As shown in Figure 5, conditioning on the most significant SNP, rs56238729, by including this SNP as a covariate into linear regression models resulted in all immediate SNP associations to be reduced toward the null with no other SNP remaining significant. These data suggest rs56238729 is the leading SNP of the VCDR association and further nullifies the associations of surrounding SNPs in the *ATOH7-PBLD* region, including SNPs that have previously been reported.

5. Analysis of Previously Reported Loci for VCDR

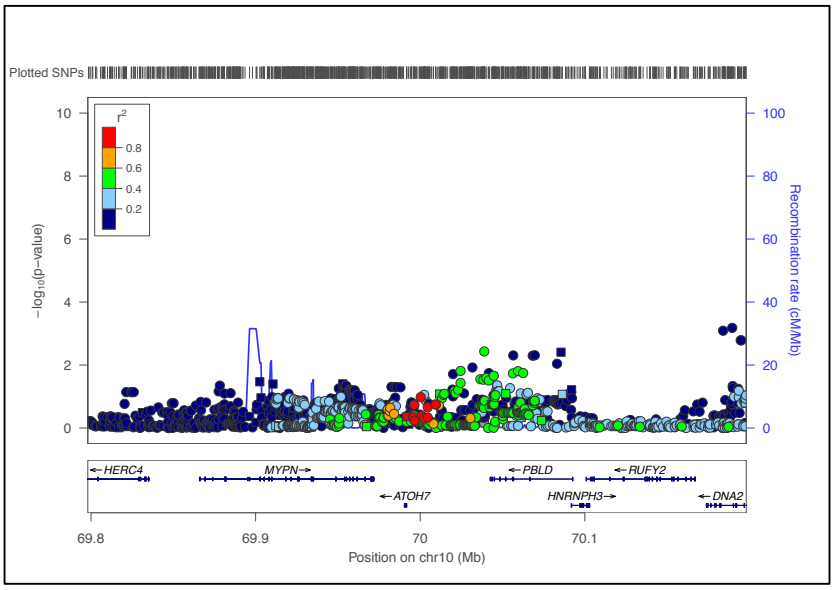
We investigated previously reported VCDR loci identified in populations of European and Asian descent to determine whether these associations are consistent in a Latino population. Table VI summarizes these results. Of the previously reported SNPs, 25 SNPs exhibited high imputation quality (Rsq ≥ 0.80) and when further analyzed, we observed 32% (8/25) having a *P*

Figure 5. Regional SNP association plots for the *ATOH7-PBLD* region.

(A)



(B)



(A) The most significant directly genotyped SNP, rs1900005, using the entire study sample is plotted in purple. Genotyped and imputed SNPs (Rsq ≥ 0.80) are plotted as squares and circles, respectively. Genes are shown below the SNPs and the arrows indicate the strand orientation for each gene. The color-coding in the plot represents the level of linkage disequilibrium with rs1900005. After imputation, rs56238729 is the most significant SNP in the *ATOH7-PBLD* region. (B) Regional SNP association plot conditioning on rs56238729, the most significant SNP in the *ATOH7-PBLD* region.

TABLE VI. COMPARISON WITH PREVIOUSLY REPORTED SNPS ASSOCIATED WITH VCDR IN LATINOS

				Previously Reported				Latinos						Most Significant Hits ± 100 kb		
SNP ID	Chr	Position	Genes Nearby	Effect Allele	Freq	β	Reference	A1/A2	AF1	β	*P*	Imp	Con of Dir	SNP ID	Position	*P*
rs301801	1	8495945	*RERE*	C/T	0.33	0.01	(63)	T/C	0.79	0.01	6.67E-01	Y	N	rs2784739	8497558	1.66E-01
rs12025126	1	8759554	*RERE*	C	0.28	-0.01	(61)	T/C	0.59	0.00	8.85E-01	Y	Y	rs4908777	8804237	1.00E-01
rs4658101	1	92077409	*CDC7-TGFBR3*	A/G	0.18	0.02	(63)	A/G	0.29	0.12	**4.58E-07**	Y	Y	rs1192419	92080059	*9.56E-08*
rs2623325	3	99131755	*COL8A1*	A/C	0.13	0.02	(63)	C/A	0.77	-0.04	8.67E-02	Y	Y	rs1157333	99071153	*3.86E-04*
rs17658229	5	172191052	*DUSP1*	C/T	0.05	-0.02	(63)	T/C	0.99	0.05	5.88E-01	Y	Y	rs2291045	172233065	3.05E-02
rs17756712	6	625071	*EXOC2*	G/A	0.18	0.01	(63)	A/G	0.82	-0.05	6.73E-02	Y	Y	rs4960092	597871	5.33E-02
rs868153	6	122389955	*HSF2*	G/T	0.36	-0.01	(63)	T/G	0.74	0.05	5.94E-02	Y	Y	rs1521224	122294945	*6.23E-04*
rs1063192	9	22003367	*CDKN2B*	G	0.45	-0.01	(61)	G/A	0.19	-0.09	**9.86E-04**	N	Y	rs1063192	22003367	*9.86E-04*
rs7865618	9	22031005	*CDKN2BAS*	G/A	0.43	-0.01	(63)	G/A	0.19	-0.09	**7.25E-04**	Y	Y	rs1063192	22003367	*9.86E-04*
rs1900005	10	69998055	*ATOH7*	A/C	0.23	-0.02	(63)	A/C	0.36	-0.15	**6.41E-12**	N	Y	rs1900005	69998055	*6.41E-12*
rs1900004	10	70000881	*ATOH7-PBLD*	T	0.22	-0.01	(61)	C/T	0.64	0.15	**1.50E-12**	Y	Y	rs1900005	69998055	*6.41E-12*
rs7072574	10	96036306	*PLCE1*	A/G	0.33	0.01	(63)	G/A	0.72	-0.02	4.87E-01	Y	Y	rs11187842	96052511	5.63E-03
rs17146964	11	65249145	*SCYL1*	G	0.21	-0.01	(61)	A/G	0.86	0.03	2.89E-01	N	Y	rs619586	65266169	3.92E-02
rs1346	11	65337251	*SSSCA1*	T/A	0.19	-0.01	(63)	A/T	0.86	0.04	2.30E-01	Y	Y	rs619586	65266169	3.92E-02
rs4936099	11	130280725	*ADAMTS8*	C/A	0.42	-0.01	(63)	C/A	0.29	-0.01	8.02E-01	Y	Y	rs10736582	130241003	1.51E-02
rs11168187	12	48044011	*RPAP3*	G/A	0.16	-0.01	(63)	A/G	0.87	0.01	8.12E-01	N	Y	rs757282	48130578	*4.83E-06*
rs10862688	12	83922912	*TMTC2*	G/A	0.45	0.01	(63)	A/G	0.75	0.00	9.19E-01	N	Y	rs904091	83950668	*5.02E-04*
rs1926320	13	36652617	*DCLK1*	C	0.24	0.01	(61)	T/C	0.69	-0.02	3.72E-01	Y	Y	rs1887829	36721724	1.24E-01
rs4901977	14	60789176	*SIX1-SIX6*	T/C	0.31	0.01	(63)	C/T	0.75	-0.05	5.27E-02	N	Y	rs4901977	60789176	5.27E-02
rs10483727	14	61072875	*SIX1*	T	0.4	0.01	(61)	T/C	0.35	0.03	1.33E-01	N	Y	rs1010053	61005625	6.08E-02
rs1345467	16	51482321	*SALL1*	G/A	0.27	0.01	(63)	G/A	0.17	0.06	2.63E-02	N	Y	rs8053277	51469726	1.38E-02
rs8068952	17	59286644	*BCAS3*	G	0.24	-0.01	(61)	G/C	0.19	-0.03	3.51E-01	Y	Y	rs7212615	59323318	1.26E-02
rs2159128	19	950380	*ARID3A*	G	0.13	-0.02	(61)	G/T	-	-	-	NA	NA	rs1056144	974967	2.11E-01
rs6054374	20	6578556	*BMP2*	T/C	0.42	-0.01	(63) (61, 63) (T/C, Freq = 0.3, β = -0.013)	C/T	0.46	0.06	1.02E-02	Y	Y	rs6140015	6487524	*4.37E-04*
rs1547014	22	29100711	*CHEK2*	T	0.29	-0.01		T/C	0.33	-0.05	3.43E-02	N	Y	rs4035540	29087041	3.44E-03
rs5756813	22	38175477	*CARD10*	G/T	0.39	0.01	(63)	G/T	0.41	0.01	5.74E-01	Y	Y	rs5750472	38081747	8.86E-02

Abbreviations: Chr, chromosome; Freq, frequency; A1/A2, allele 1 / allele 2; Imp, Imputed; Con of Dir, Consistency of Direction. Additional allele, frequency, and effect size information is given in parentheses for SNPs with multiple references. The frequency of allele 1 is given for our Latino sample and is modeled as the effect allele. To correct for multiple testing for correlated SNPs, the program simpleM was used, and identified 23 independent tests, resulting in a Bonferroni correction *P* value of 0.05/23 = 2.17E-03. Shown in bold are *P* values meeting this threshold. Using directly genotyped SNPs, the most significant hit within ± 100 kb of previously reported SNPs are listed (*P* values < 2.17E-03 are italicized). All SNPs, except rs2159128, had an Rsq ≥ 0.80 from Minimac3. Most SNPs exhibited consistent direction of effect, except for rs301801. SNP positions are according to GRCh37/hg19.

< 0.05 in our Latino population. Furthermore, we observed consistent directions of associations with all previous SNPs, except one (rs301801). To account for multiple testing and to avoid penalties associated with a traditional Bonferroni correction, we used simpleM (106-108) to calculate the effective number of independent tests. This method identified 23 independent tests, resulting in a Bonferroni correction significance level of $P = 2.17 \times 10^{-3}$. We were able to replicate five index SNPs in three regions (*CDC7-TGFBR3*, *CDKN2B-CDKN2BAS*, and *ATOH7-PBLD*) after multiple testing correction (shown in boldface in Table VI). To replicate genomic regions associated with VCDR, we extracted the most significant directly genotyped SNP within ± 100 kb of the previously reported SNPs. This approach identified an additional five SNPs in five regions surviving the multiple testing correction, replicating the associations between VCDR and the *COL8A1*, *HSF2*, *RPAP3*, *TMTC2*, and *BMP2* regions.

6. Pathway Analysis

To determine whether canonical pathways were enriched with genes associated with VCDR, we performed pathway analysis using IPA. After adjusting for multiple testing, the only pathway significantly associated with VCDR was the ‘‘pathogenesis of multiple sclerosis (MS)’’ pathway ($P = 7.41 \times 10^{-3}$). Of the nine genes comprising this pathway, five genes from our dataset overlapped with these genes, including *CCL3*, *CCL4*, *CXCL9*, *CXCL10*, and *CXCL11*.

7. Discussion

This study represents the first GWAS conducted on VCDR in Latinos. We identified two genome-wide significant SNPs, rs1900005 and rs7916697, associated with VCDR, confirming the involvement of the *ATOH7-PBLD* region. We also identified suggestive associations in the *CDC7-TGFBR3* and *ZNF770-DPH6* regions. We discovered a novel SNP, rs56238729, in the *ATOH7-PBLD* region to be significantly associated with VCDR in Latinos after genotype imputation from the 1000 Genomes Project reference panels. Moreover, we were able to

replicate genomic regions previously associated with VCDR, including *COL8A1*, *HSF2*, *RPAP3*, *TMTC2*, and *BMP2*. Results from our pathway analysis identified one canonical pathway associated with VCDR.

The most significant SNPs in our study reside in the *ATOH7-PBLD* region. Previous GWAS studies, including Ramdas et al (61) and Springelkamp et al,(63) identified many SNPs in the *ATOH7-PBLD* region to be associated with VCDR. In addition to VCDR, earlier GWAS have also associated this region with optic disc area,(61, 68, 125, 126) cup area,(68, 125) and POAG.(61) Both rs1900005 and rs7916697, the most and second most significant SNPs in our study respectively, have previously been associated with VCDR.(63) In particular, rs7916697 resides in the 5' UTR region of *ATOH7*, a single exon gene that plays a role in retinal ganglion cell development.(127) Moreover, rs7916697 has been associated with a reduction in optic disc area (126) and was identified to have a significant interactive effect with rs1063192 in an Afro-Caribbean population, resulting in a reduction in POAG risk.(128) Taken together, our results are consistent with prior studies of these SNPs being strongly associated with glaucoma related quantitative traits and may have a biological role in the pathogenesis of POAG.

The third and fourth most significant SNPs indicate suggestive associations in the *CDC7-TGFBR3* and *ZNF770-DPH6* regions, respectively. Similar to the previous region, the *CDC7-TGFBR3* region has been reported to be associated with VCDR,(63) optic disc area,(61, 126) and POAG.(129) Moreover, expression of both *CDC7* and *TGFBR3* have been observed in numerous human ocular tissues, most notably the optic disc and optic nerve.(129) The SNP rs1192419 has specifically been associated with VCDR (63) and an increase in disc area.(68) Additionally, common variants within the genomic region on chromosome 15q14, in which *ZNF770-DPH6* resides in, has previously been associated with refractive error and myopia.(130) This study independently confirms the associations of rs1900005 and rs7916697 in the *ATOH7-PBLD* region with VCDR in a sample of Latinos and suggests additional loci in the *CDC7-TGFBR3* and *ZNF770-DPH6* regions.

While ophthalmologists routinely assess the VCDR to diagnose and monitor the progression of POAG, the cupping of the optic nerve may not solely be a result of glaucoma and may result from other conditions, such as optic neuritis.(131) Our pathway results implicate an association between the pathogenesis of MS and VCDR. Multiple sclerosis is a demyelinating disease of the central nervous system that commonly affects vision. Patients with MS were found to have a higher VCDR compared to healthy controls, suggesting enlarging of the optic cup due to the thinning of the retinal nerve fiber layer may be explained by the predilection of the disease to afflict the optic nerves.(132) Moreover, several of the genes included in this pathway code for chemokines that were shown to be at higher concentrations in the aqueous humor of glaucomatous eyes compared to cataract controls.(133) Collectively, our pathway results suggest the biological mechanisms influencing VCDR and MS may share common genetic constituents.

This is the first GWAS of VCDR in Latinos and several limitations exist. First, Latinos are historically an understudied population. As far as we know, our dataset is currently the only Latino genetic dataset with ophthalmic phenotypes. Furthermore, the three-way admixture of Latinos makes it even more challenging in genetics research.(55, 101, 134) We performed a fixed-effects meta-analysis for the discovery and replication sets on the top genotyped SNPs using METAL (135) and obtained similar results as EMMAX (Appendix C). These results suggest population stratification and genetic relatedness were properly controlled for in our analysis, despite Latinos being a three-way admixed population. However, we emphasize the need for replication in an independent Latino cohort. And secondly, we did not conduct secondary analyses adjusting for disc area, an ocular parameter known to be correlated with VCDR.(136) Unfortunately, at the time of the VCDR data collection, disc area was not collected. Given a previously reported reduction in significance for VCDR associated variants after adjusting for disc area,(63) a similar trend may be observed for our results.

In conclusion, in the first GWAS of VCDR in Latinos, we discovered a novel SNP that is

significantly associated with VCDR in Latinos. In addition, two SNPs reached genome-wide significance, replicating associations in the *ATOH7-PBLD* region. We were also able to replicate associations with several previously reported genomic regions for VCDR in this population. Our pathway results identified a novel association between the pathogenesis of multiple sclerosis and VCDR, suggesting potential shared genetic factors influencing both VCDR and MS. The findings from this study suggest that many genetic factors influencing VCDR are shared among ethnic populations.

C. Genome-Wide Gene-Environment Interaction Analysis of Body Mass Index and Vertical Cup-Disc Ratio

1. Study Sample

Table VII presents the characteristics of the study sample included for this analysis. In this Latino sample, 420 (10.7%) study participants are under / normal weight with a mean BMI (standard deviation, SD) of 23.3 kg/m^2 (1.5; range: 14.6-25.0) and 3,507 (89.3%) are overweight / obese with a mean BMI (SD) of 31.8 kg/m^2 (5.1; range: 25.0-60.8), with an overall mean BMI (SD) of 30.9 kg/m^2 (5.5; range: 14.6-60.8). The proportion of females among under / normal weight subjects was 56.7% and 57.5% for overweight / obese subjects, with an overall proportion of 57.5%. The mean age (SD) for the study sample was 54.7 (10.4) years, with a mean age of 56.3 (12.2) years for under / normal weight study participants and 54.6 (10.2) years for overweight / obese participants. Lastly, the overall mean VCDR (SD) for the study sample was 0.34 (0.18) with an average VCDR of 0.36 (0.20; range: 0.10-0.90) for under / normal weight individuals and 0.34 (0.18; range: 0.10-0.90) for overweight / obese participants.

2. Genome-Wide Association Results

The genomic control inflation factor (104) for the overall study sample during step 1 was moderate, $\lambda = 1.03$. The Q-Q plot of the observed *P* values verses the expected *P* values for the

TABLE VII. CHARACTERISTICS OF THE STUDY SAMPLE

Study	Sample Size	Females, %	Age, y, Mean (SD)	BMI, Mean (SD)	BMI Range	VCDR, Mean (SD)	VCDR Range
Under / Normal Weight	420	56.7	56.3 (12.2)	23.3 (1.5)	14.6-25.0	0.36 (0.20)	0.10-0.90
Overweight / Obese	3,507	57.5	54.6 (10.2)	31.8 (5.1)	25.0-60.8	0.34 (0.18)	0.10-0.90
Total	3,927	57.5	54.7 (10.4)	30.9 (5.5)	14.6-60.8	0.34 (0.18)	0.10-0.90

marginal genetic effects for all genotyped SNPs in step 1 is presented in Figure 6. As displayed in the plot, except at the extreme tail, the observed *P* values do not deviate from the null. Taken together, the genomic control inflation factor and the Q-Q plot indicate proper control of population stratification.

Figure 6. A quantile-quantile plot of the $-\log_{10}(P$ values) for the 576,798 genotyped SNPs analyzed in step 1.

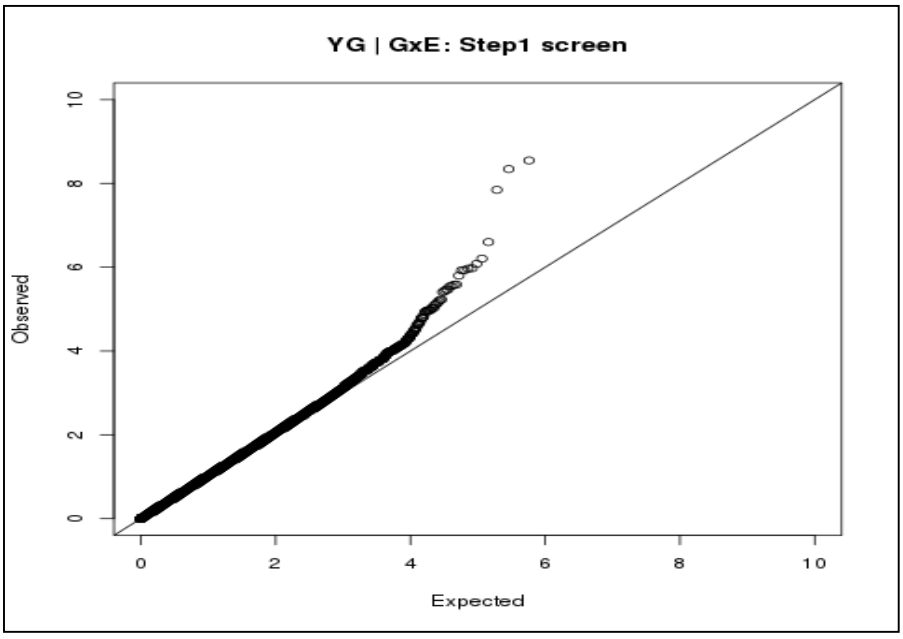


Figure 7 presents the Manhattan plot for the marginal genetic *P* values for SNPs analyzed in step 1. Of the 576,798 SNPs analyzed in step 1, 30,712 SNPs exhibited a marginal genetic $P \leq 0.05$ and were analyzed for SNP×BMI interactions in step 2. Figure 7 also presents the Manhattan plot for the SNP×BMI interaction *P* values from step 2. The results for the top SNPs ($P < 1\times10^{-4}$) are summarized in Table VIII. No SNP reached genome-wide significance

after correcting for multiple testing ($P = 1.63\times10^{-6}$) during step 2. The most significant SNP×BMI interaction signal is located at 13q33.3 ($P = 1.90\times10^{-6}$) situated 140 kb downstream of *TNFSF13B* and 147 kb upstream of *MYO16*. The second most significant SNP is located at 2q12.1 ($P = 3.70\times10^{-6}$) located in *IL1RL1*. The third most significant SNP is located at 4q23 ($P = 3.90\times10^{-6}$), located 1 kb upstream of *ADH1B* and 14 kb downstream of *ADH1C*.

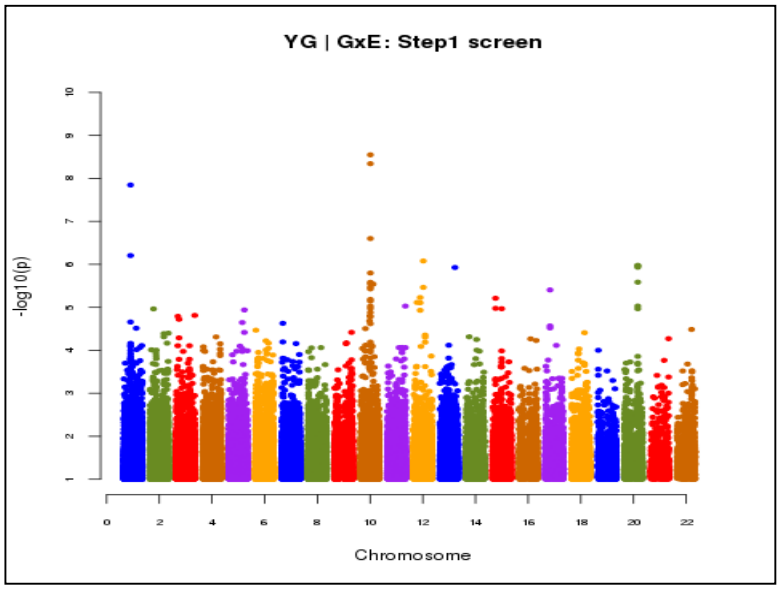
To further explore and understand the biological mechanisms of the observed associations, stratified analyses were conducted by BMI classification for the top ranking SNPs. Table IX presents the individual SNP effect sizes by BMI stratum (under / normal weight and overweight / obese) for the top SNPs. For most SNPs, except rs12683130 and rs11695757, the marginal genetic associations were more significant among under / normal weight participants than overweight / obese participants. Interestingly, only 3 of the 11 top SNPs (4q23, rs1037439, and rs1389595) are associated with higher VCDR in under / normal weight subjects compared to overweight / obese subjects, while the remaining 8 SNPs are associated with lower VCDR. For the top three SNPs, among under / normal weight subjects, the minor allele A for 13q33.3 and the minor allele T for 2q12.1 are associated with a reduction in VCDR ($\beta = -0.37$ and $\beta = -0.36$, respectively), whereas the minor allele T for 4q23 is associated with an increase in VCDR ($\beta = 0.40$). Among overweight / obese subjects 13q33.3 (A) and 2q12.1 (T) were also associated with a reduction in VCDR ($\beta = -0.01$ and $\beta = -0.02$, respectively) and 4q23 (T) was associated with an increase in VCDR ($\beta = 0.01$).

3. Results from Imputed SNPs

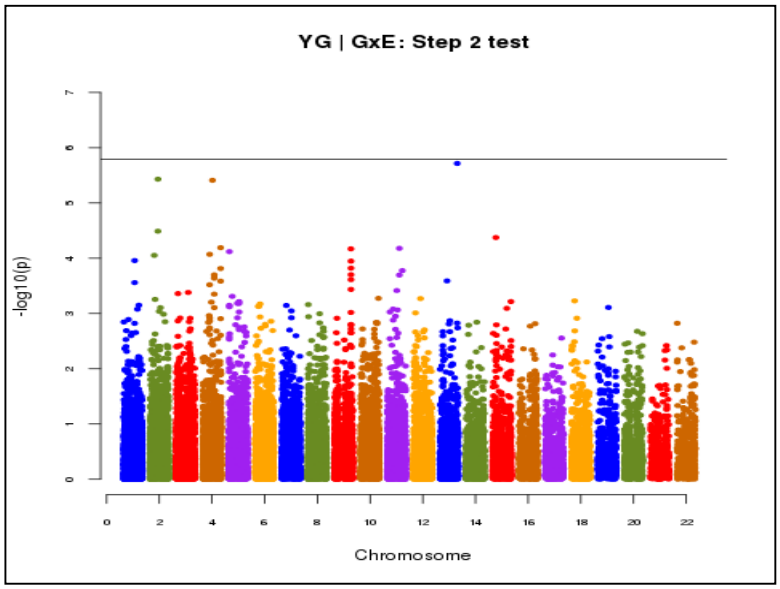
We performed genotype imputation to assess the association of SNPs not directly genotyped. Of the 6.8 million SNPs analyzed in step 1, 368,832 SNPs exhibited a marginal genetic $P \leq 0.05$ and were formally tested for SNP×BMI interactions in step 2. No additional genomic regions reached GWAS significance after correcting for multiple testing.

Figure 7. Manhattan plots displaying the $-\log_{10}(P$ values) from step 1 and step 2 of the genome-wide gene-environment interaction analysis for genotyped SNPs.

(A)



(B)



(A) Step 1 results for the association between the 576,798 SNPs and VCDR, adjusting for age, sex, and the first four principal components of genetic ancestry. (B) Step 2 results for the association between SNP×BMI and VCDR for the 30,712 SNPs with a $P \leq 0.05$ in step 1. The horizontal black line represents the Bonferroni significance threshold for the SNP×BMI interaction term ($P = 1.63\times10^{-6}$). Genome-wide association results for SNPs are plotted by genomic position.

TABLE VIII. SUMMARY STATISTICS FOR THE TOP RANKING INTERACTIVE SNPS WITH BMI ASSOCIATED WITH VCDR IN LATINOS

SNP	Chr	Position	Gene	A1/A2	MAF	$\beta_{interaction}$	$P_{interaction}$
13q33.3	13	-	*TNFSF13B-MYO16*	A/G	0.36	0.35	1.90×10^{-6}
2q12.1	2	-	***IL1RL1***	T/C	0.26	0.35	3.70×10^{-6}
4q23	4	-	*ADH1B-ADH1C*	T/G	0.23	-0.38	3.90×10^{-6}
rs12712142	2	102960584	***IL1RL1***	A/C	0.27	0.32	3.20×10^{-5}
rs1037439	15	37026565	***C15orf41***	G/A	0.27	-0.34	4.20×10^{-5}
rs2016910	4	183959226	*DCTD-WWC2*	G/T	0.42	0.29	6.50×10^{-5}
rs7129973	11	88915570	***TYR***	G/A	0.33	0.30	6.60×10^{-5}
rs12683130	9	125684827	***ZBTB26***	C/T	0.27	0.31	6.80×10^{-5}
rs1389595	5	2902855	*C5orf38-IRX1*	T/C	0.33	-0.31	7.60×10^{-5}
rs1844635	4	71265172	***PROL1***	C/A	0.45	0.27	8.50×10^{-5}
rs11695757	2	55158486	***EML6***	C/T	0.01	0.94	8.90×10^{-5}

Abbreviations: Chr, chromosome; A1/A2, allele 1/allele 2; MAF, minor allele frequency. Genes are bolded for SNPs located inside a gene. SNP positions are according to GRCh37/hg19.

4. Conditional Analysis

We conducted conditional analyses among under / normal weight participants for the top three identified regions to determine whether additional genetic variants contribute to the VCDR associations. Figure 8 presents the regional plots, as well as conditional analyses, for the top three identified regions. During stratified analysis, one imputed SNP was more significant than 13q33.3 in the *TNFSF13B-MYO16* region. After conditioning on the most significant SNP, the significance of the associations of the SNPs in the surrounding region reduced, suggesting the imputed SNP is the leading SNP of the VCDR association among under / normal weight individuals in this region.

For the *IL1RL1* region on chromosome 2, 2q12.1 remained the most significant SNP after imputation and after conditioning on 2q12.1, all of the associations for the neighboring SNPs moved towards the null, indicating 2q12.1 is the leading VCDR SNP in under / normal weight subjects in the *IL1RL1* region. Similarly, 4q23 was the most significant SNP after

TABLE IX. STRATIFIED ANALYSIS FOR THE TOP RANKING INTERACTIVE SNPS WITH BMI ASSOCIATED WITH VCDR IN LATINOS

					Under / Normal Weight			Overweight / Obese		
SNP	Chr	Position	Gene	A1/A2	MAF	β	*P*	MAF	β	*P*
13q33.3	13	-	*TNFSF13B-MYO16*	A/G	0.34	-0.37	1.70×10^{-6}	0.36	-0.01	7.50×10^{-1}
2q12.1	2	-	***IL1RL1***	T/C	0.29	-0.36	4.22×10^{-6}	0.26	-0.02	4.91×10^{-1}
4q23	4	-	*ADH1B-ADH1C*	T/G	0.26	0.40	3.17×10^{-6}	0.23	0.01	6.88×10^{-1}
rs12712142	2	102960584	***IL1RL1***	A/C	0.30	-0.33	3.45×10^{-5}	0.26	-0.02	5.64×10^{-1}
rs1037439	15	37026565	***C15orf41***	G/A	0.27	0.38	7.64×10^{-6}	0.27	0.03	1.98×10^{-1}
rs2016910	4	183959226	*DCTD-WWC2*	G/T	0.39	-0.32	1.81×10^{-5}	0.42	-0.02	3.41×10^{-1}
rs7129973	11	88915570	***TYR***	G/A	0.32	-0.32	2.47×10^{-5}	0.33	-0.03	2.65×10^{-1}
rs12683130	9	125684827	***ZBTB26***	C/T	0.26	-0.28	8.11×10^{-4}	0.27	0.10	1.02×10^{-4}
rs1389595	5	2902855	*C5orf38-IRX1*	T/C	0.29	0.32	6.14×10^{-5}	0.34	0.02	3.84×10^{-1}
rs1844635	4	71265172	***PROL1***	C/A	0.43	-0.30	1.88×10^{-5}	0.45	-0.03	1.71×10^{-1}
rs11695757	2	55158486	***EML6***	C/T	0.03	-0.58	1.52×10^{-2}	0.01	0.37	2.42×10^{-4}

Abbreviations: Chr, chromosome; Freq, frequency; A1/A2, allele 1 / allele 2; MAF, minor allele frequency.

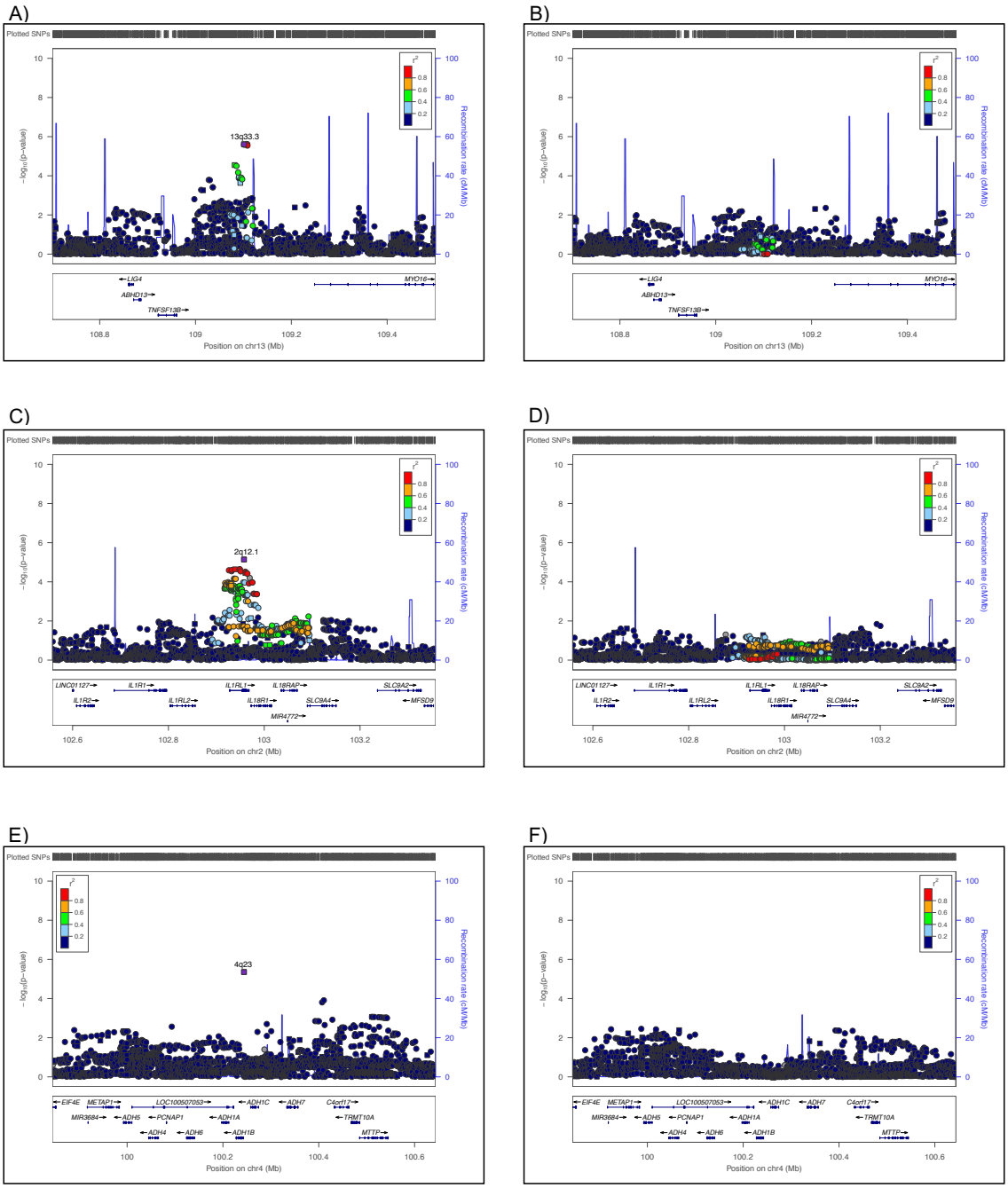
imputation in the *ADH1B-ADH1C* region and after conditioning on this SNP, no additional genetic variants were associated with VCDR, suggesting 4q23 is the lead VCDR SNP among under / normal weight participants in the *ADH1B-ADH1C* region.

5. Discussion

We performed the first genome-wide gene-environment interaction analysis of body mass index on vertical cup-disc ratio. In this study, we identified several suggestive interactive associations between SNPs and BMI on VCDR using a Latino population. Despite the lack of significant findings, we identified several biologically plausible candidate genomic regions for further examination. Moreover, in an attempt to uncover the remaining missing heritability of complex traits, such as VCDR, this study exemplifies the potential utility of G×E studies to further identify genetic variants associated with such traits.

The most significant SNP, 13q33.3, resides in the *TNFSF13B-MYO16* region. Tumor Necrosis Factor Superfamily Member 13b, also known as BAFF, plays an important role in B cell proliferation and differentiation, functions as a T-cell co-stimulatory molecule, and elevated levels of BAFF have been associated with several autoimmune diseases, including systemic lupus erythematosus and rheumatoid arthritis.(137) Additionally, expression of TNFSF13B was elevated in adipocytes during differentiation and expression was augmented by TNF-α treatment, indicating TNFSF13B is an adipokine, a group of cytokines that modulate inflammation. BAFF was also found to be produced by astrocytes, up-regulated in multiple sclerosis lesions, and elevated in the CSF of patients with systemic lupus erythematosus.(138, 139) Moreover, *MYO16* has previously been reported to be associated with childhood obesity related traits in a Hispanic population.(140) En masse, these results indicate this genetic loci is associated with inflammatory processes via fat cells and the inflammatory responses may result in damage to the optic nerve.

Figure 8. Regional SNP association plots for the *TNFSF13B-MYO16*, *IL1RL1*, and *ADH1B-ADH1C* regions among under / normal weight participants.



Genotyped and imputed SNPs are plotted as squares and circles, respectively. Genes are shown below the SNPs and the arrows indicate the strand orientation for each gene. The color-coding in the plots represent the level of linkage disequilibrium with the top genotype SNP, plotted in purple, in each region. (A) The most significant SNP in *TNFSF13B-MYO16* is an imputed SNP and (B) no additional SNPs remained associated with VCDR after conditioning on this SNP. (C) 2q12.1 was the most significant SNP in the *IL1RL1* region and (D) after conditioning on this SNP, the surrounding associations reduced towards the null. (E) The most significant SNP in *ADH1B-ADH1C* is 4q23 and (F) no additional SNPs remained associated with VCDR after conditioning on this SNP.

The second most significant SNP is located in *IL1RL1*. The lead SNP, 2q12.1, has previously been associated with eosinophil count, asthma, atopic asthma, and IgE in European and Asian study subjects,(141) and was further replicated in a Hispanic study population.(142) The *IL1RL1* gene is a member of the interleukin 1 receptor family with several SNPs having previously been reported as protein quantitative trait loci in human cerebrospinal fluid, as well as numerous inflammatory diseases and immune factors.(143) Additionally, *IL1RL1* is associated with the Th2 immune response and is a receptor for IL-33.(144) Interleukin 33 is expressed in glia cells in the central nervous system and retina and is elevated in advanced age-related macular degeneration.(145) Interleukin 33 is also expressed in human preadipocytes and adipocytes, is negatively associated with BMI, and has been hypothesized to exhibit protective effects in preventing obesity-induced inflammation.(146-148) Similar to the previous loci, these findings demonstrate *IL1RL1* plays a role in inflammatory responses and the protective effect from the associated immune factors may serve to prevent damage to the optic nerve head and subsequent development of glaucoma.

The third most significant SNP, 4q23, is in an intergenic region between *ADH1B*-*ADH1C*, members of the alcohol dehydrogenase family. Expression of *ADH1B* in adipose tissue has previously been reported to be inversely associated with BMI, waist circumference, and fasting glucose in a Mexican American population.(149) Findings from this study suggest elevated levels of ADH1B in human adipocytes may enhance energy mobilization by promoting efficient metabolism of alcohol into energy, potentially regulating the storage of fat. Additionally, a previous GWAS identified *ADH1C* to be associated with clusterin in CSF.(150) As a glycoprotein that aides in apoptosis and cell homeostasis, accumulation of clusterin has been reported in the optic nerve head of eyes with glaucoma.(151) Moreover, clusterin and activation of the complement system have been found to play a role in the pathologic process of exfoliation glaucoma and the *CLU* gene interacts with other genes implicated with primary open

angle glaucoma, including *GAS7* and *CAV1*.(152, 153) Overall, this suggests this genomic region may modulate BMI and development of glaucoma.

There are several limitations to the study needed to be addressed. First, the sample size used in the current study is small, especially for the investigation of G×E interactions and as such, this analysis is underpowered to identify significant GWAS interactions. And second, we are unable to replicate our findings in a similar population. Currently, our dataset is the only Latino dataset that contains both genetic and ophthalmic data and consequently, replicating our G×E findings in a population with similar characteristics is presently challenging. Despite these limitations, this study identified suggestive associations that are biologically relevant and warrant further investigation.

In summary, we conducted the first genome-wide gene-environment interaction analysis of body mass index on vertical cup-disc ratio. We identified several suggestive associations that exhibit biological relevance and may serve as candidate genomic regions for further investigation. Moreover, these findings demonstrate the importance of exploring G×E interactions to further uncover the missing heritability of complex traits, as well as the potential to better classify individuals at higher risk for disease.

D. Genetic Risk Scores of Vertical Cup-Disc Ratio

1. Study Sample

Table X presents a summary of the study sample characteristics and simple linear regression results between VCDR and the variables included in this study. The mean (standard deviation, SD) for the untransformed VCDR of the study sample is 0.3 (0.2). Among the study participants, 42.5% are males, 25.0% are diabetics, and 5.7% have POAG. The mean (SD) of age, CCT, IOP, and weighted GRS is 56.7 (10.4) years, 550.1 (33.7) μm, 14.7 (3.0) mmHg, and 0.8 (0.1), respectively. During univariate linear regression, numerous variables are significantly associated with VCDR, including age ($P < 0.0001$), gender ($P = 0.0022$), SBP ($P < 0.0001$), IOP

($P < 0.0001$), T2D ($P = 0.0052$), income ($P = 0.0479$), POAG status ($P < 0.0001$), and weighted GRS ($P < 0.0001$). Additionally, BMI, CCT, smoking status, and education are not associated with VCDR.

TABLE X. SUMMARY STATISTICS AND SIMPLE LINEAR REGRESSION RESULTS

Characteristic	Participants (n = 4,018)	*P*
VCDR	0.3 (0.2)	-
Age, year	56.7 (10.4)	<0.0001
Gender, male	42.5%	0.0022
BMI, kg/m^2	30.9 (5.5)	0.0798
SBP, mmHg	123.8 (19.1)	<0.0001
CCT, μm	550.1 (33.7)	0.8259
IOP, mmHg	14.7 (3.0)	<0.0001
T2D, yes	25.0%	0.0052
Smoking status		0.9565
Never	60.7%	
Former	25.4%	
Current	13.9%	
Education, yr		0.0785
≤ 6	44.6%	
7-11	21.9%	
≥ 12	33.4%	
Income[a]		0.0479
< $20,000	50.0%	
$20,000-$40,000	35.9%	
> $40,000	14.1%	
POAG		<0.0001
Cases	5.7%	
Controls	94.3%	
Weighted GRS	0.8 (0.1)	<0.0001

Abbreviations: VCDR, vertical cup-disc ratio; BMI, body mass index; SBP, systolic blood pressure; CCT, central corneal thickness; IOP, intraocular pressure; T2D, type 2 diabetes; POAG, primary open angle glaucoma; GRS, genetic risk score.
[a] Missing income for 513 study participants.

2. Genetic Risk Score and VCDR

Table XI displays the SNPs and weights used to construct the GRS from previously reported SNPs. Table XII presents the multiple linear regression results from model building. With all of the possible risk factors entered into a full linear regression model, only age ($P < 0.0001$), gender ($P = 0.0016$), CCT ($P = 0.0154$), IOP ($P < 0.0001$), education ($P = 0.0106$), and the weighted GRS ($P < 0.0001$) remained in the model at a significance cutoff of $P \leq 0.05$ during stepwise selection. The base multiple linear regression model including age, gender, CCT, IOP, and education accounts for 4.30% of the total variance for VCDR. An additional 2.74% of the variance of VCDR is explained by the weighted GRS, yielding a total of 7.04% explained by the model. The unweighted GRS yielded similar results ($\beta = 0.02$, $P < 0.0001$, 2.60% additional variance explained).

3. Genetic Risk Score and POAG

Multiple logistic regression analyses using quintiles of weighted GRS evaluated the association of the weighted GRS on POAG. After performing stepwise regression, age ($P < 0.0001$), gender ($P = 0.0318$), CCT ($P = 0.0015$), IOP ($P < 0.0001$), SBP ($P = 0.0408$), and the weighted GRS ($P = 0.0011$) remained significantly associated with POAG. Figure 9 shows the distribution of the weighted GRS in the study sample and odds ratios of POAG comparing each of the upper GRS quintiles with the lowest, adjusting for age, gender, CCT, IOP, and SBP. Compared to the lowest quintile, both the highest and second highest quintiles had significantly higher odds of POAG, OR = 1.75 (95% CI: [1.09, 2.81]; $P = 0.0212$) and OR = 2.15 (95% CI: [1.34, 3.45]; $P = 0.0015$), respectively. Analysis of the unweighted GRS yielded similar estimates and significance levels. The highest and second highest quintiles of the unweighted GRS had significantly higher odds of POAG compared to the lowest quintile, OR = 2.00 (95% CI: [1.24, 3.22]; $P = 0.0042$) and OR = 1.91 (95% CI: [1.18, 3.10]; $P = 0.0087$), respectively.

TABLE XI. PREVIOUSLY REPORTED SINGLE NUCLEOTIDE POLYMORPHISMS INCLUDD IN GENETIC RISK SCORES FOR VERTICAL CUP-DISC RATIO

Chr	Nearest Gene	SNP	Alleles	VCDR-increasing allele	Frequency of VCDR-increasing allele	Weight	Reference
1	*RERE*	rs301801	T/C	C	0.21	0.008	(63)
1	*RERE*	rs12025126	T/C	T	0.59	0.011	(61)
1	*RPE65*	rs1925953	A/T	T	0.68	0.006	(27)
1	*CDC7/TGFBR3*	rs1192414	A/G	A	0.28	0.014	(27)
1	*CDC7/TGFBR3*	rs4658101	A/G	A	0.29	0.013	(27)
1	*F5*	rs10753787	T/C	C	0.75	0.007	(27)
3	*FLNB*	rs6764184	G/T	T	0.29	0.007	(27)
3	*COL8A1*	rs2623325	C/A	A	0.23	0.016	(63)
3	*COL8A1*	rs6804624	T/C	C	0.37	0.008	(27)
3	*COL8A1*	rs1997404	T/G	G	0.26	0.008	(27)
5	*PDZD2*	rs72759609	T/C	T	0.91	0.012	(27)
5	*VCAN*	rs7717697	T/C	T	0.64	0.007	(27)
5	*DUSP1*	rs17658229	T/C	T	0.99	0.02	(63)
5	*DUSP1*	rs114503346	C/T	C	0.99	0.021	(27)
5	*DUSP1*	rs35084382	T/C	T	0.99	0.018	(27)
6	*EXOC2*	rs17756712	A/G	G	0.18	0.01	(63)
6	*RREB1*	rs4960295	G/A	A	0.51	0.007	(27)
6	*HSF2*	rs868153	T/G	T	0.74	0.007	(63)
7	*DGKB*	rs10274998	C/T	T	0.41	0.008	(27)
8	*CRISPLD1*	rs117598310	G/T	T	0.05	0.009	(27)
8	*PSCA*	rs2920293	C/G	C	0.47	0.006	(27)
9	*CDKN2B*	rs1063192	G/A	A	0.81	0.014	(61)
9	*CDKN2BAS*	rs7865618	G/A	A	0.81	0.013	(63)
9	*CDKN2B/AS1*	rs2157719	C/T	T	0.81	0.013	(27)
9	*CDKN2B/AS1*	rs1360589	C/T	T	0.82	0.013	(27)
10	*ATOH7*	rs7916697	A/G	G	0.62	0.018	(27)
10	*ATOH7*	rs7916410	T/C	C	0.64	0.018	(27)
10	*ATOH7*	rs1900005	A/C	C	0.64	0.018	(63)
10	*ATOH7/PBLD*	rs1900004	C/T	C	0.64	0.013	(61)
10	*PLCE1*	rs3891783	C/G	G	0.37	0.007	(27)
10	*PLCE1*	rs1830890	A/G	G	0.28	0.006	(27)
10	*PLCE1*	rs7072574	G/A	A	0.28	0.009	(63)
10	*ENO4*	rs1681739	C/T	T	0.31	0.006	(27)
11	*SCYL1*	rs17146964	A/G	A	0.86	0.014	(61)
11	*SSSCA1*	rs1346	A/T	A	0.86	0.013	(27)
11	*ADAMTS8*	rs4936099	C/A	A	0.71	0.007	(27)
12	*RPAP3*	rs11168187	A/G	A	0.87	0.009	(63)
12	*TMTC2*	rs10862688	A/G	G	0.25	0.008	(63)
12	*TMTC2*	rs442376	T/C	C	0.60	0.011	(27)
12	*TMTC2*	rs482507	C/T	T	0.59	0.011	(27)
12	*TMTC2*	rs324780	G/A	A	0.61	0.011	(27)
12	*FAM101A*	rs7311936	G/C	G	0.67	0.006	(27)
13	*DCLK1*	rs7323428	G/T	T	0.31	0.007	(27)
13	*DCLK1*	rs1926320	T/C	C	0.32	0.012	(61)
14	*SIX1/6*	rs4901977	C/T	T	0.25	0.011	(63)

TABLE XI. PREVIOUSLY REPORTED SINGLE NUCLEOTIDE POLYMORPHISMS INCLUDED IN GENETIC RISK SCORES FOR VERTICAL CUP-DISC RATIO (continued)

Chr	Nearest Gene	SNP	Alleles	VCDR-increasing allele	Frequency of VCDR-increasing allele	Weight	Reference
14	*SIX6*	rs4436712	G/T	T	0.34	0.009	(27)
14	*SIX6*	rs8015152	C/T	T	0.27	0.01	(27)
14	*SIX1*	rs10483727	T/C	T	0.35	0.012	(61)
14	*SIX6*	rs34935520	G/A	G	0.35	0.009	(27)
15	*FAM169B*	rs6598351	C/T	T	0.12	0.006	(27)
15	*ASB7*	rs60779155	G/A	A	0.34	0.01	(27)
15	*ASB7*	rs34222435	C/T	T	0.33	0.01	(27)
15	*ASB7*	rs4299136	G/C	C	0.33	0.01	(27)
16	*SALL1*	rs11646917	G/T	G	0.71	0.009	(27)
16	*SALL1*	rs4784295	C/G	C	0.18	0.009	(27)
16	*SALL1*	rs1345467	G/A	G	0.17	0.009	(27)
17	*BCAS3*	rs8068952	G/C	C	0.81	0.012	(61)
19	*ARID3A*	rs2159128	G/T	T	0.14	0.019	(61)
20	*BMP2*	rs6054374	C/T	C	0.46	0.007	(63)
20	*BMP2*	rs6054375	G/T	G	0.47	0.01	(27)
20	*BMP2*	rs6107845	G/A	G	0.48	0.009	(27)
22	*CHEK2*	rs1547014	T/C	C	0.67	0.013	(63)
22	*CHEK2*	rs5762752	C/G	G	0.65	0.011	(27)
22	*CHEK2*	rs5752773	G/C	C	0.67	0.012	(27)
22	*CHEK2*	rs738722	T/C	C	0.69	0.012	(27)
22	*CARD10*	rs2092172	G/A	A	0.18	0.009	(27)
22	*CARD10*	rs56385951	G/A	A	0.09	0.011	(27)
22	*CARD10*	rs5756813	G/T	G	0.41	0.008	(63)

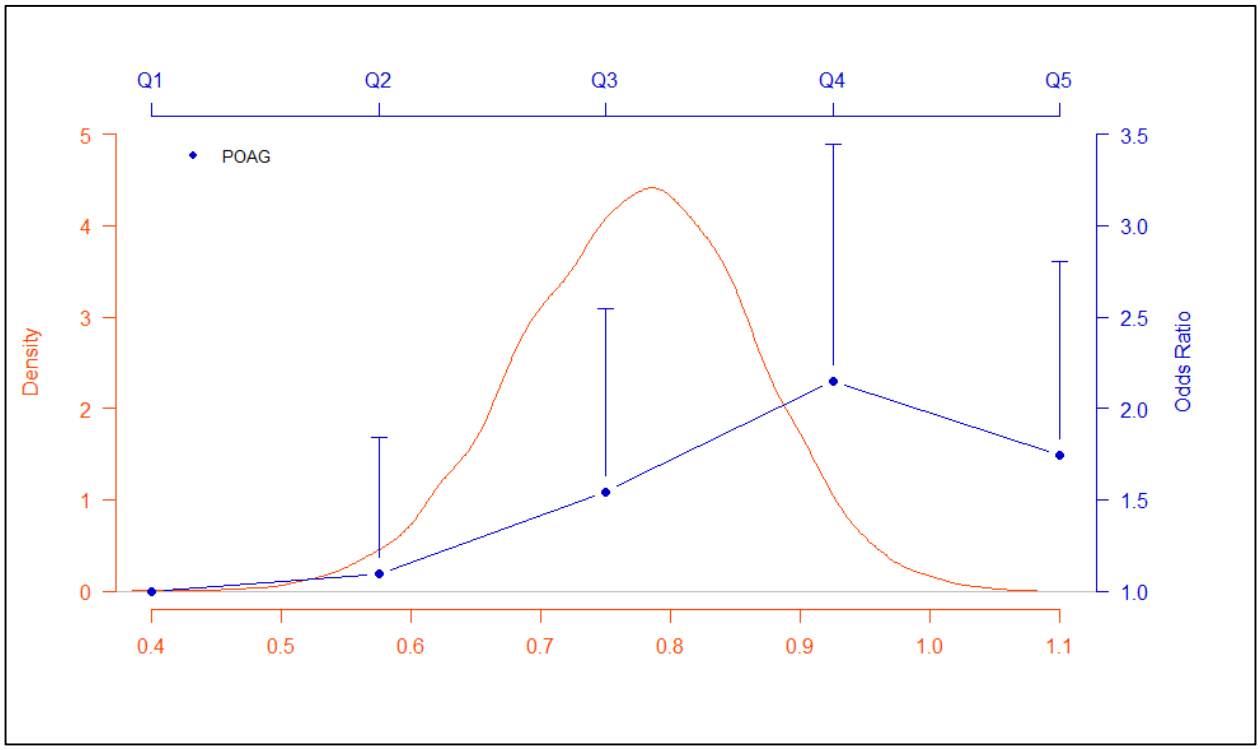
We conducted ROC analyses to examine the discriminatory power of the unweighted and weighted GRSs on POAG status. Figure 10 presents the ROC curves for models without and with the weighted GRS constructed from previously reported VCDR SNPs, the weighted GRS generated from lead SNPs, and the weighted GRS derived from our genome-wide association data. The AUC is 0.728 (95% CI: [0.694, 0.761]) for the model with only age and gender. When the weighted GRS from previously reported SNPs was added into the model, there was a non-significant increase in the AUC to 0.735 (95% CI: [0.701, 0.768]; *P* = 0.150). When the weighted GRS using the lead SNPs was added to the model, there was a significant increase in the AUC to 0.755 (95% CI: [0.722, 0.787]; *P* = 0.002). In contrast, the addition of the GRS derived from our own genome-wide association data resulted in a significant increase in the AUC to 0.809 (95% CI: [0.781, 0.837]; *P* < 0.0001). Similar associations and significance levels were obtained for the unweighted GRS for these analyses (Figure 11).

TABLE XII. MULTIPLE LINEAR REGRESSION RESULTS FOR GENETIC RISK SCORE

	Model 1		Model 2	
Characteristic	Beta	*P*	Beta	*P*
Age	0.009	<0.0001	0.01	<0.0001
Gender	0.103	0.0033	0.096	0.0016
BMI	-0.0045	NS	-	-
SBP	-0.0002	NS	-	-
CCT	-0.0013	0.0111	-0.0011	0.0154
IOP	0.0533	<0.0001	0.0519	<0.0001
T2D	0.043	NS	-	-
Smoking Status	-0.0133	NS	-	-
Income[a]	-0.0309	NS	-	-
Education	0.0417	0.0305	0.044	0.0106
Weighted GRS	1.767	<0.0001	1.831	<0.0001

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; CCT, central corneal thickness; IOP, intraocular pressure; T2D, type 2 diabetes; GRS, genetic risk score.
[a] Missing income for 513 study participants.

Figure 9. Distribution of weighted genetic risk score from previously reported SNPs and association with primary open angle glaucoma.

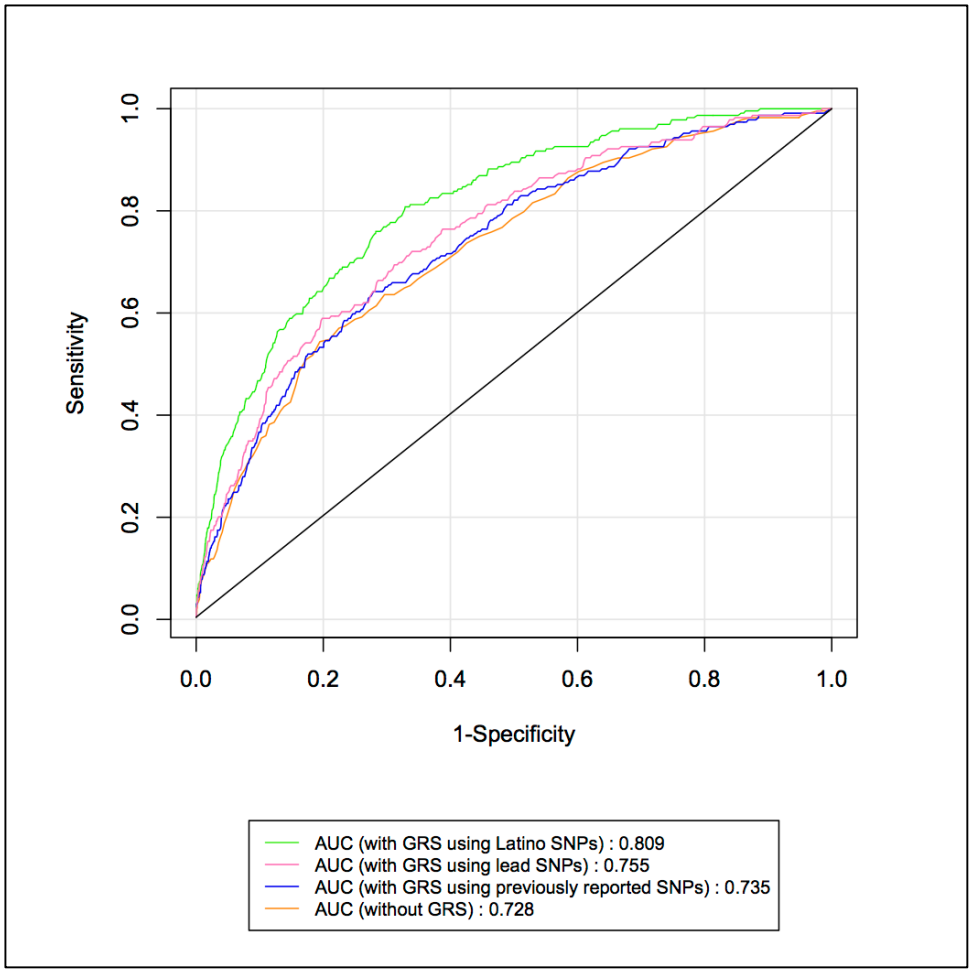


Distribution of the weighted GRS and odds ratios of POAG comparing each of the four upper GRS quintiles to the lowest quintile, adjusting for age, gender, CCT, IOP, and SBP. Vertical lines of each point (OR) represents the upper 95% confidence interval.

4. Discussion

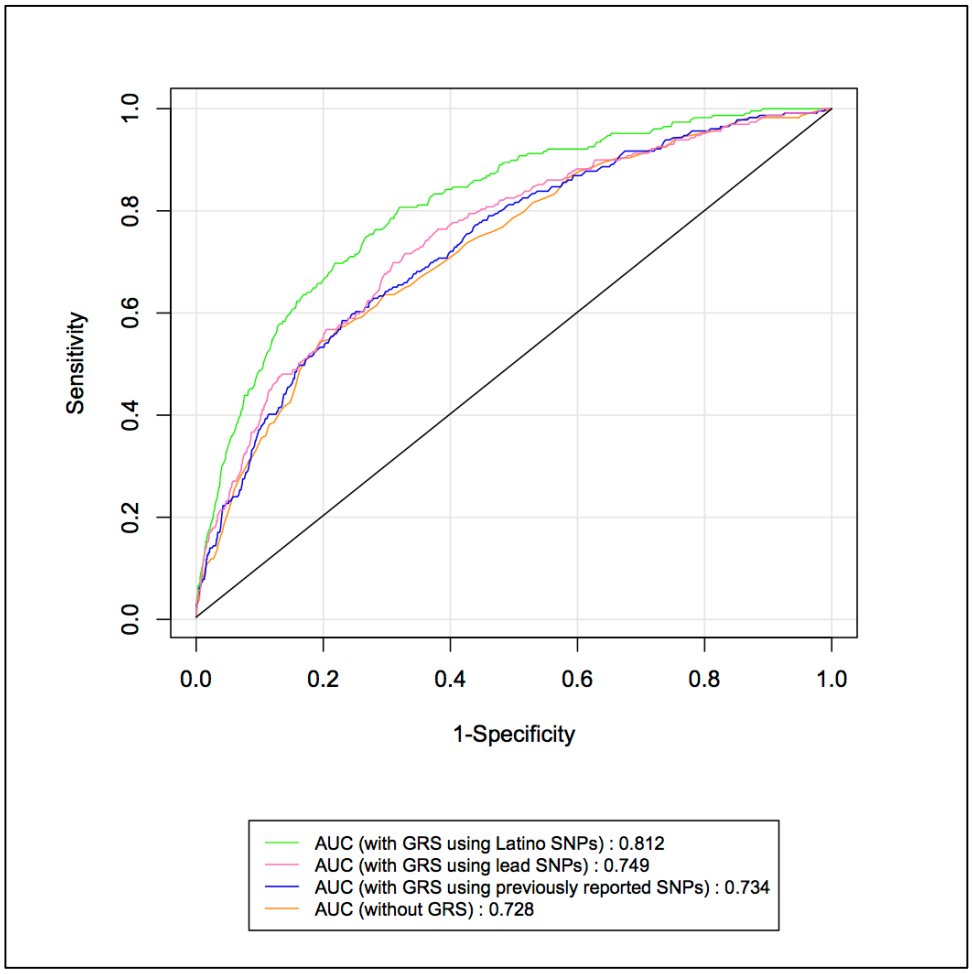
In this study, we constructed genetic risk scores based on SNPs previously associated with VCDR and evaluated whether these GRSs were associated with VCDR and POAG and increased the discriminatory ability for POAG. We observed significant associations between the GRSs and VCDR, indicating a higher GRS was associated with a larger vertical cup-disc ratio. These associations remained significant after the inclusion of traditional risk factors, explaining an additional 2.74% of the variation in VCDR. Moreover, compared to the lowest quintile of the GRSs, study participants in the highest two quintiles experienced significantly higher odds of POAG. We show the inclusion of ethnic specific GRSs significantly increased

Figure 10. Receiver operating characteristic curves predicting primary open angle glaucoma for weighted GRS.



The curves are based on logistic regression models adjusting for age and gender without and with the weighted GRS constructed from previously reported VCDR SNPs, the weighted GRS generated from lead SNPs, and the weighted GRS derived from our genome-wide association data. AUC represents the area under the curve, with a larger AUC representing better classification of POAG status. The addition of weighted GRS derived from our own genome-wide association data significantly improved the discriminatory ability for POAG ($P < 0.0001$).

Figure 11. Receiver operating characteristic curves predicting primary open angle glaucoma for unweighted GRS.



The curves are based on logistic regression models adjusting for age and gender without and with the unweighted GRS constructed from previously reported VCDR SNPs, the unweighted GRS generated from lead SNPs, and the unweighted GRS derived from our genome-wide association data. AUC represents the area under the curve, with a larger AUC representing better classification of POAG status. The addition of unweighted GRS derived from our own genome-wide association data significantly improved the discriminatory ability for POAG ($P < 0.0001$).

the discriminatory power for POAG. Additionally, we obtained similar results for the unweighted GRS. To our knowledge, we are the first to report these associations in a Latino population.

The success of genome-wide association studies in identifying genetic variants indicates that multiple genetic loci, rather than a single gene, contribute to the susceptibility of a given phenotype. Despite the modest effect of individual variants, creating an aggregated score allows for the evaluation of the combined genetic effect of these variants on a trait. The utility of GRSs in the fields of public health and medicine has the potential to significantly reduce the incidence of disease by being used as a screening tool to identify individuals at a greater risk of a disorder. Genetic risk scores can be used to identify subgroups in a population that are at a higher risk for a disorder, so targeted public health interventions can be directed towards them. In a similar manner, GRSs aid in the movement towards personalized medicine. By assessing an individual's GRS before the development of disease, early interventions (e.g., dietary, behavioral, etc.) can be implemented to counterbalance the genetic risk.(154) En masse, GRSs provide an opportunity to be a useful tool in summarizing an individual's genetic susceptibility to a trait and may be potentially used for reducing the occurrence of disease.

Primary open angle glaucoma is a heterogeneous disease, both genetically and phenotypically. As such, investigating quantitative traits and the corresponding genetic variants will aid in understanding the biological mechanisms underlying this disease. We observed significant associations between the GRSs and POAG, with higher GRSs associated with greater odds of POAG. To further examine the utility of GRSs, we performed several ROC analyses to evaluate whether the inclusion of the GRSs improved POAG discriminatory ability. We observed a moderate increase in the AUC after including the GRSs with traditional risk factors, although the increase was minor, potentially limiting the utility of such genetic risk scores in a clinical setting. A study conducted in a multiethnic Asian population observed a borderline significant improvement in the discriminatory ability for glaucoma when IOP and VCDR GRSs were included into a model with traditional risk factors.(88) Specifically, the AUC

estimate for POAG exhibited a modest improvement when the IOP and VCDR GRSs were included with traditional risk factors, increasing from 0.72 to 0.74 (AUC difference = 0.02; *P* = 0.06).(88) In our study of Latinos, we observed similar AUC estimates for both the unweighted and weighted GRSs, demonstrating the consistent modest improvement in AUC from previously reported SNPs. Furthermore, using GRSs constructed from our own genome-wide association data, we observed significant increases in the AUC with the addition of more SNPs.(102) This suggests additional genetic variants, besides those previously reported, with low effect estimates further aid in the discriminating ability for POAG by incorporating additional genetic information into the model. Moreover, ethnic specific weights for the construction of GRSs may further aid in improving disease prediction. Together, despite the marginal increases in AUC in both the current and previous studies using published genetic variants, these findings suggest GRSs constructed from quantitative traits of POAG can aid in increasing the discriminatory ability for this disease, including variants with lower effect estimates. Moreover, due to the polygenetic nature of POAG, further identification of genetic variants associated with the pathogenesis of POAG may aid in improving the predictive power, and clinical utility, of GRSs.

The strengths of this study include the generation of GRSs consisting of VCDR SNPs identified to date. Also, we observed significant associations with GRSs constructed from SNPs identified primarily in European populations in our study sample consisting of Latinos and as such, these results may be generalizable to other ethnic populations. There are several limitations however. First, we used previously reported genetic variants identified in GWAS thus far, which explain only a small amount of variation in VCDR, which resulted in a moderate improvement in POAG prediction. The GRS used were constructed from a limited number of genetic variants from European and Asian populations, and may not be transferable to other racial groups. Unweighted GRS, however, are preferred to weighted GRS when the existing studies are comprised of different ethnicities compared to the population under study.(155) We obtained similar results for both the weighted and unweighted GRSs, demonstrating the

robustness of our results and the potential transferability of a GRS for VCDR. Additionally, the GRSs were constructed based on SNPs identified using traditional GWAS significance thresholds, which may not have captured variants with weaker effect sizes. When we constructed GRSs using a larger number of variants weighted by Latino specific estimates, we observed a better classification of POAG, suggesting that increasing the number of SNPs and applying population specific weights can lead to better predictions for POAG.

In summary, we observed GRSs composed of 68 previously reported VCDR SNPs were significantly associated with VCDR in a Latino population. Moreover, the GRSs were significantly associated with POAG, with individuals with higher GRSs experiencing greater odds of POAG. Inclusion of ethnic specific GRSs constructed using a larger number of SNPs significantly improved the discriminatory ability for POAG. The application of GRSs as a population-based evaluation tool can potentially yield significant reductions in disease incidence. By quantifying an individual's genetic risk before disease development, early interventions can be adopted to counterbalance this genetic risk.

V. CONCLUSION

A. Summary of Main Findings

Despite being the largest minority group in the United States, Latinos are an understudied population in ocular genetics research. Previous studies attempting to identify genetic factors associated with ocular diseases and related quantitative traits were primarily conducted in European and Asian populations. Moreover, Latinos are genetically heterogeneous, enabling the investigation of the effect of genetic ancestry on ocular characteristics. As such, this dissertation sought to address the current gaps in the literature regarding the genetic factors and biological mechanisms underlying the pathogenesis of POAG in Latinos by conducting the first studies to examine the association between genetic factors and ocular quantitative traits of POAG in this population.

Racial differences in IOP have been observed, with individuals of African descent exhibiting higher IOP when compared to European and Hispanic individuals. As a three-way admixture of European, Native American, and African ancestries, we examined the association between genetic ancestry and IOP in Latinos. We observed a significant association between African ancestry and IOP, with increasing proportions of African ancestry associated with higher IOP. Due to the documented racial differences of hypertension, we examined potential effect modification between genetic ancestry and elevated systolic blood pressure on IOP. We observed a significant interaction between African ancestry and elevated systolic blood pressure. Those individuals with elevated systolic blood pressure exhibited higher IOP with increasing African ancestry compared to those who do not have elevated systolic blood pressure. These findings are consistent with traditional epidemiological studies, and suggest African ancestry, and its interaction with elevated systolic blood pressure, are novel risk factors for intraocular pressure in Latinos.

Genome-wide association studies have identified numerous genetic loci associated with vertical cup-disc ratio, but these GWAS were conducted in European and Asian populations. To

replicate previously reported loci, as well as identify novel genetic variants, we conducted the first GWAS of VCDR in Latinos. We identified two significant and two suggestive SNPs associated with VCDR. After performing genotype imputation, we identified a novel SNP in the *ATOH7-PBLD* region. Additionally, we replicated eight previously reported regions and identified a novel association between the pathogenesis of multiple sclerosis and VCDR during pathway analysis. Such results suggest the genetic factors influencing VCDR exhibit consistent associations across ethnic populations.

Although genome-wide association studies have identified genetic variants associated with numerous complex traits, such findings only explain a small proportion of the variation in these traits. One strategy to uncover the remaining missing heritability is investigating the effect of environmental factors on genetic variants through gene-environment interactions. Given the compelling evidence of an association between BMI and VCDR, as well as previous genetic variants explaining only a small proportion of VCDR, we conducted the first genome-wide gene-environment interaction analysis of BMI on VCDR in a Latino population. We identified several suggestive interactive associations between SNPs and BMI on VCDR that represent candidate genomic regions for further investigation. During stratified analyses, a majority of the SNPs were significant among under / normal weight study participants, and exhibited protective effects against larger VCDR. Specifically, the top three regions, i.e. *TNFSF13B-MYO16*, *IL1RL1*, and *ADH1B-ADH1C*, exhibit biological relevance related to inflammatory processes. These findings demonstrate the utility of exploring G×E interactions to further uncover the missing heritability of quantitative traits for ocular disease, and the potential to classify individuals who have a higher risk for disease.

The findings from genome-wide association studies, while informative in our understanding of the genetic architecture of numerous traits, confer only a modest effect on a given trait and have limited predictive power. As such, genetic risk scores enable the examination of the cumulative genetic effect on traits. We constructed genetic risk scores for

VCDR and evaluated the association between GRSs and VCDR and determined if the GRSs improved the discriminatory power for POAG. We observed significant associations between GRSs and VCDR, indicating higher GRSs were associated with larger VCDR. Moreover, the GRSs were associated with POAG, with individuals in the highest two quintiles experiencing greater odds of POAG, and the inclusion of the ethnic specific GRSs improved the discriminatory ability for POAG. Such findings suggest evaluation of an individual's cumulative genetic risk may aid in identifying individuals at a greater risk for disease, in which early interventions may be initiated to counterbalance the genetic risk.

B. Contributions of Knowledge

Through this dissertation, we report several of the first studies to investigate the association between genetic factors and ocular traits in a Latino population. To the best of our knowledge, we are the first to investigate the association between genetic ancestry and IOP in this population, identifying a significant association between African ancestry and IOP. We also are the first to perform a GWAS of VCDR in a Latino population, replicating previous findings from European and Asian study samples. Moreover, we are the first to conduct a genome-wide gene-environment interaction analysis of BMI on VCDR, and identified several suggestive and biologically relevant associations. And lastly, we are the first to construct and evaluate VCDR GRSs and demonstrate improvements in the discriminatory ability of POAG in a Latino sample.

C. Public Health Relevance

Examining the role of genetics in human disease not only aids in our understanding of the genetic architecture of these traits, but can also translate into public health prevention strategies that may assist in preventing, identifying, and mitigating disease. Findings from the current studies may aid in developing primary prevention strategies that attempt to identify individuals who are genetically at a higher risk for POAG. From the analysis of genetic ancestry

and IOP, genetic screening can be conducted to identify Latinos with high proportions of African ancestry who are at a higher genetic predisposition for elevated IOP. Strategies to monitor IOP, such as receiving routine eye examinations, can then be recommended to prevent glaucomatous damage. Moreover, Latinos with high proportions of African ancestry who also have elevated systolic blood pressure could also be recommended to receive routine examinations to monitor IOP, and potentially medication, such as diuretics or angiotensin–converting enzyme inhibitors, to lower systolic blood pressure. Results from the GWAS of VCDR aided in confirming the transferability of genetic variants across different ethnic groups. Such findings suggest the underlying biological mechanisms influencing VCDR determination may be similar across ethnic groups and strategies that target these mechanisms to prevent the enlargement of VCDR in one ethnic population may be applied to other ethnic populations. Findings from the genome-wide gene-environment interaction analysis of BMI on VCDR indicate anthropometric measurements, such as BMI, can modify the effect of genetic variants on VCDR. This suggests environmental factors, such as anthropometric measurements, may be used to classify individuals who are genetically at a higher risk for enlarged VCDR, and potentially can be used to counterbalance the effect from genetic variants on VCDR determination. Lastly, findings from the GRSs and VCDR and POAG analysis demonstrate that construction of a cumulative genetic measure can be used as a screening tool to characterize and identify individuals with an elevated aggregate genetic susceptibility for larger VCDR, in which early interventions can be adopted to counterbalance this genetic risk. Moreover, these results suggest the possibility of generating VCDR GRSs to predict future POAG, in which proactive measures can be taken to mitigate the genetic risk for this disease.

D. Future Directions

While this dissertation furthers our understanding of the genetic architecture of complex ocular traits by extending ocular genetic studies to include an ethnically diverse population,

many opportunities exist to extend the scope of this work. First, additional studies in independent and ethnically similar populations are needed to replicate the findings presented here. To the best of our knowledge, the dataset used for this dissertation is currently the only Latino dataset that contains both ophthalmological and genetic data. As such, replicating the findings presented here is advised to validate these results and to continue scientific discovery. However, given the growing number of large-scale studies, including the UK Biobank and the All of Us project, opportunities to replicate our findings will increasingly become feasible. Second and in tandem with the previous point, genetic studies ought to expand to include diverse populations not only to evaluate the transferability of findings from one population to the next, but to better identify the genetic determinants of complex traits for the human population as a whole.(156) Inclusion of diverse populations will also lead to reductions in the occurrence of disease and improvement in health outcomes for all ethnic groups while avoiding any health disparities owing to a bias towards one ethnic population in genetic studies. Third, compared to European studies, this current work uses a smaller sample size, limiting the ability to detect causal variants in this study population. Future studies should attempt to increase sample size or perform meta-analyses to aid in further identifying genetic variants associated with complex traits, especially for genome-wide gene-environment interaction analyses. And lastly, the current work focused exclusively on common genetic variants, while omitting rare and structural variants. With the advent of low-cost sequencing and computational advancements, analysis of such genetic factors will further uncover the missing heritability of complex traits and subsequently improve our understanding of the biological mechanisms underlying such traits, including primary open angle glaucoma.

CITED LITERATURE

1. Weinreb RN, Leung CK, Crowston JG, Medeiros FA, Friedman DS, Wiggs JL, et al. Primary open-angle glaucoma. Nat Rev Dis Primers. 2016;2:16067.

2. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology. 2014;121(11):2081-90.

3. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 2006;90(3):262-7.

4. Vajaranant TS, Wu S, Torres M, Varma R. The changing face of primary open-angle glaucoma in the United States: demographic and geographic changes from 2011 to 2050. Am J Ophthalmol. 2012;154(2):303-14 e3.

5. Tielsch JM, Katz J, Sommer A, Quigley HA, Javitt JC. Family history and risk of primary open angle glaucoma. The Baltimore Eye Survey. Arch Ophthalmol. 1994;112(1):69-73.

6. Coffey M, Reidy A, Wormald R, Xian WX, Wright L, Courtney P. Prevalence of glaucoma in the west of Ireland. Br J Ophthalmol. 1993;77(1):17-21.

7. Wensor MD, McCarty CA, Stanislavsky YL, Livingston PM, Taylor HR. The prevalence of glaucoma in the Melbourne Visual Impairment Project. Ophthalmology. 1998;105(4):733-9.

8. Quigley HA, West SK, Rodriguez J, Munoz B, Klein R, Snyder R. The prevalence of glaucoma in a population-based study of Hispanic subjects: Proyecto VER. Arch Ophthalmol. 2001;119(12):1819-26.

9. Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. Ophthalmology. 1996;103(10):1661-9.

10. Varma R, Ying-Lai M, Francis BA, Nguyen BB, Deneen J, Wilson MR, et al. Prevalence of open-angle glaucoma and ocular hypertension in Latinos: the Los Angeles Latino Eye Study. Ophthalmology. 2004;111(8):1439-48.

11. Leske MC, Connell AM, Schachat AP, Hyman L. The Barbados Eye Study. Prevalence of open angle glaucoma. Arch Ophthalmol. 1994;112(6):821-9.

12. Nemesure B, Leske MC, He Q, Mendell N. Analyses of reported family history of glaucoma: a preliminary investigation. The Barbados Eye Study Group. Ophthalmic Epidemiol. 1996;3(3):135-41.

13. Stine OC, Xu J, Koskela R, McMahon FJ, Gschwend M, Friddle C, et al. Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. Am J Hum Genet. 1995;57(6):1384-94.

14. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995;378(6559):789-92.

15. Stone EM, Fingert JH, Alward WL, Nguyen TD, Polansky JR, Sunden SL, et al.

Identification of a gene that causes primary open angle glaucoma. Science. 1997;275(5300):668-70.

16. Allingham RR, Liu Y, Rhee DJ. The genetics of primary open-angle glaucoma: a review. Exp Eye Res. 2009;88(4):837-44.

17. Fingert JH, Heon E, Liebmann JM, Yamamoto T, Craig JE, Rait J, et al. Analysis of myocilin mutations in 1703 glaucoma patients from five different populations. Hum Mol Genet. 1999;8(5):899-905.

18. Jacobson N, Andrews M, Shepard AR, Nishimura D, Searby C, Fingert JH, et al. Non-secretion of mutant proteins of the glaucoma gene myocilin in cultured trabecular meshwork cells and in aqueous humor. Hum Mol Genet. 2001;10(2):117-25.

19. Rezaie T, Child A, Hitchings R, Brice G, Miller L, Coca-Prados M, et al. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science. 2002;295(5557):1077-9.

20. Aung T, Rezaie T, Okada K, Viswanathan AC, Child AH, Brice G, et al. Clinical features and course of patients with glaucoma with the E50K mutation in the optineurin gene. Invest Ophthalmol Vis Sci. 2005;46(8):2816-22.

21. De Marco N, Buono M, Troise F, Diez-Roux G. Optineurin increases cell survival and translocates to the nucleus in a Rab8-dependent manner upon an apoptotic stimulus. J Biol Chem. 2006;281(23):16147-56.

22. Monemi S, Spaeth G, DaSilva A, Popinchalk S, Ilitchev E, Liebmann J, et al. Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. Hum Mol Genet. 2005;14(6):725-33.

23. Hauser MA, Allingham RR, Linkroum K, Wang J, LaRocque-Abramson K, Figueiredo D, et al. Distribution of WDR36 DNA sequence variants in patients with primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 2006;47(6):2542-6.

24. Fingert JH. Primary open-angle glaucoma genes. Eye (Lond). 2011;25(5):587-95.

25. Bailey JN, Loomis SJ, Kang JH, Allingham RR, Gharahkhani P, Khor CC, et al. Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. Nat Genet. 2016;48(2):189-94.

26. Burdon KP, Macgregor S, Hewitt AW, Sharma S, Chidlow G, Mills RA, et al. Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. Nat Genet. 2011;43(6):574-8.

27. Springelkamp H, Iglesias AI, Mishra A, Hohn R, Wojciechowski R, Khawaja AP, et al. New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. Hum Mol Genet. 2017;26(2):438-53.

28. Gharahkhani P, Burdon KP, Fogarty R, Sharma S, Hewitt AW, Martin S, et al. Common variants near ABCA1, AFAP1 and GMDS confer risk of primary open-angle glaucoma. Nat Genet. 2014;46(10):1120-5.

29. Thorleifsson G, Walters GB, Hewitt AW, Masson G, Helgason A, DeWan A, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. Nat Genet. 2010;42(10):906-9.

30. Wiggs JL, Kang JH, Yaspan BL, Mirel DB, Laurie C, Crenshaw A, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma in Caucasians from the USA. Hum Mol Genet. 2011;20(23):4707-13.

31. Ramdas WD, van Koolwijk LM, Lemij HG, Pasutto F, Cree AJ, Thorleifsson G, et al. Common genetic variants associated with open-angle glaucoma. Hum Mol Genet. 2011;20(12):2464-71.

32. Wiggs JL, Yaspan BL, Hauser MA, Kang JH, Allingham RR, Olson LM, et al. Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. PLoS Genet. 2012;8(4):e1002654.

33. Cuellar-Partida G, Craig JE, Burdon KP, Wang JJ, Vote BJ, Souzeau E, et al. Assessment of polygenic effects links primary open-angle glaucoma and age-related macular degeneration. Sci Rep. 2016;6:26885.

34. Guzmán B. The Hispanic Population - Census 2000 Brief. U.S. Department of Commerce: Economics and Statistics Administration; 2001.

35. Ennis SR, Ríos-Vargas M, Albert NG. The Hispanic Population - 2010 Censes Briefs. U.S. Department of Commerce: Economics and Statistics Administration; 2011.

36. Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, et al. Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Natl Acad Sci U S A. 2010;107 Suppl 2:8954-61.

37. Gupta A, Singh-Radcliff N. Pharmacology in anesthesia practice. New York: Oxford University Press; 2013. xviii, 462 p. p.

38. Weinreb RN, Khaw PT. Primary open-angle glaucoma. Lancet. 2004;363(9422):1711-20.

39. Kass MA, Heuer DK, Higginbotham EJ, Johnson CA, Keltner JL, Miller JP, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Arch Ophthalmol. 2002;120(6):701-13; discussion 829-30.

40. Klein BE, Klein R, Linton KL. Intraocular pressure in an American community. The Beaver Dam Eye Study. Invest Ophthalmol Vis Sci. 1992;33(7):2224-8.

41. Wu SY, Leske MC. Associations with intraocular pressure in the Barbados Eye Study. Arch Ophthalmol. 1997;115(12):1572-6.

42. Xu L, Li J, Zheng Y, Cui T, Zhu J, Ma K, et al. Intraocular pressure in Northern China in an urban and rural population: the Beijing eye study. Am J Ophthalmol. 2005;140(5):913-5.

43. Memarzadeh F, Ying-Lai M, Azen SP, Varma R, Los Angeles Latino Eye Study G.

Associations with intraocular pressure in Latinos: the Los Angeles Latino Eye Study. Am J Ophthalmol. 2008;146(1):69-76.

44. Sommer A, Tielsch JM, Katz J, Quigley HA, Gottsch JD, Javitt J, et al. Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey. Arch Ophthalmol. 1991;109(8):1090-5.

45. Tielsch JM, Sommer A, Katz J, Royall RM, Quigley HA, Javitt J. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. JAMA. 1991;266(3):369-74.

46. Leske MC, Connell AM, Wu SY, Hyman L, Schachat AP. Distribution of intraocular pressure. The Barbados Eye Study. Arch Ophthalmol. 1997;115(8):1051-7.

47. Murray T, Beaty TH, Mathias RA, Rafaels N, Grant AV, Faruque MU, et al. African and non-African admixture components in African Americans and an African Caribbean population. Genet Epidemiol. 2010;34(6):561-8.

48. Nomura H, Ando F, Niino N, Shimokata H, Miyake Y. The relationship between age and intraocular pressure in a Japanese population: the influence of central corneal thickness. Curr Eye Res. 2002;24(2):81-5.

49. Dielemans I, Vingerling JR, Hofman A, Grobbee DE, de Jong PT. Reliability of intraocular pressure measurement with the Goldmann applanation tonometer in epidemiological studies. Graefes Arch Clin Exp Ophthalmol. 1994;232(3):141-4.

50. Nwankwo T, Yoon SS, Burt V, Gu Q. Hypertension among adults in the United States: National Health and Nutrition Examination Survey, 2011-2012. NCHS Data Brief. 2013(133):1-8.

51. Memarzadeh F, Ying-Lai M, Chung J, Azen SP, Varma R, Los Angeles Latino Eye Study G. Blood pressure, perfusion pressure, and open-angle glaucoma: the Los Angeles Latino Eye Study. Invest Ophthalmol Vis Sci. 2010;51(6):2872-7.

52. Kumar R, Seibold MA, Aldrich MC, Williams LK, Reiner AP, Colangelo L, et al. Genetic ancestry in lung-function predictions. N Engl J Med. 2010;363(4):321-30.

53. Salari K, Choudhry S, Tang H, Naqvi M, Lind D, Avila PC, et al. Genetic admixture and asthma-related phenotypes in Mexican American and Puerto Rican asthmatics. Genet Epidemiol. 2005;29(1):76-86.

54. Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, et al. A genomewide admixture map for Latino populations. Am J Hum Genet. 2007;80(6):1024-36.

55. Gao X, Gauderman WJ, Marjoram P, Torres M, Chen YD, Taylor KD, et al. Native American ancestry is associated with severe diabetic retinopathy in Latinos. Invest Ophthalmol Vis Sci. 2014;55(9):6041-5.

56. Kwon YH, Fingert JH, Kuehn MH, Alward WL. Primary open-angle glaucoma. N Engl J Med. 2009;360(11):1113-24.

57. Varma R, Tielsch JM, Quigley HA, Hilton SC, Katz J, Spaeth GL, et al. Race-, age-,

gender-, and refractive error-related differences in the normal optic disc. Arch Ophthalmol. 1994;112(8):1068-76.

58. Knight OJ, Girkin CA, Budenz DL, Durbin MK, Feuer WJ, Cirrus OCTNDSG. Effect of race, age, and axial length on optic nerve head parameters and retinal nerve fiber layer thickness measured by Cirrus HD-OCT. Arch Ophthalmol. 2012;130(3):312-8.

59. Klein BE, Klein R, Lee KE. Heritability of risk factors for primary open-angle glaucoma: the Beaver Dam Eye Study. Invest Ophthalmol Vis Sci. 2004;45(1):59-62.

60. Charlesworth J, Kramer PL, Dyer T, Diego V, Samples JR, Craig JE, et al. 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. 2010;51(7):3509-14.

61. Ramdas WD, van Koolwijk LM, Ikram MK, Jansonius NM, de Jong PT, Bergen AA, et al. A genome-wide association study of optic disc parameters. PLoS Genet. 2010;6(6):e1000978.

62. Ikram MK, Borger PH, Assink JJ, Jonas JB, Hofman A, de Jong PT. Comparing ophthalmoscopy, slide viewing, and semiautomated systems in optic disc morphometry. Ophthalmology. 2002;109(3):486-93.

63. Springelkamp H, Hohn R, Mishra A, Hysi PG, Khor CC, Loomis SJ, et al. Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process. Nat Commun. 2014;5:4883.

64. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010;87(1):139-45.

65. Huang H, Chanda P, Alonso A, Bader JS, Arking DE. Gene-based tests of association. PLoS Genet. 2011;7(7):e1002177.

66. Mishra A, Macgregor S. VEGAS2: Software for More Flexible Gene-Based Testing. Twin Res Hum Genet. 2015;18(1):86-91.

67. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6:5890.

68. Springelkamp H, Mishra A, Hysi PG, Gharahkhani P, Hohn R, Khor CC, et al. Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genet Epidemiol. 2015;39(3):207-16.

69. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.

70. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet. 2017;101(1):5-22.

71. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012;90(1):7-24.

72. McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. Hum Mol Genet. 2008;17(R2):R156-65.

73. Gauderman WJ, Mukherjee B, Aschard H, Hsu L, Lewinger JP, Patel CJ, et al. Update on the State of the Science for Analytical Methods for Gene-Environment Interactions. American Journal of Epidemiology. 2017;186(7):762-70.

74. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP. Finding novel genes by testing G x E interactions in a genome-wide association study. Genet Epidemiol. 2013;37(6):603-13.

75. Hamza TH, Chen H, Hill-Burns EM, Rhodes SL, Montimurro J, Kay DM, et al. Genome-wide gene-environment study identifies glutamate receptor gene GRIN2A as a Parkinson's disease modifier gene via interaction with coffee. PLoS Genet. 2011;7(8):e1002237.

76. Zhang R, Chu M, Zhao Y, Wu C, Guo H, Shi Y, et al. A genome-wide gene-environment interaction analysis for tobacco smoke and lung cancer susceptibility. Carcinogenesis. 2014;35(7):1528-35.

77. Wang L, Murk W, DeWan AT. Genome-Wide Gene by Environment Interaction Analysis Identifies Common SNPs at 17q21.2 that Are Associated with Increased Body Mass Index Only among Asthmatics. PLoS One. 2015;10(12):e0144114.

78. Xu L, Wang YX, Wang S, Jonas JB. Neuroretinal rim area and body mass index. PLoS One. 2012;7(1):e30104.

79. Zheng Y, Cheung CY, Wong TY, Mitchell P, Aung T. Influence of height, weight, and body mass index on optic disc parameters. Invest Ophthalmol Vis Sci. 2010;51(6):2998-3002.

80. Amerasinghe N, Wong TY, Wong WL, Mitchell P, Shen SY, Loon SC, et al. Determinants of the optic cup to disc ratio in an Asian population: the Singapore Malay Eye Study (SiMES). Arch Ophthalmol. 2008;126(8):1101-8.

81. Kim YJ, Kim JM, Shim SH, Bae JH, Park KH, Epidemiologic Survey Committee of the Korean Ophthalmological S. Associations between Optic Cup-to-disc Ratio and Systemic Factors in the Healthy Korean Population. Korean J Ophthalmol. 2015;29(5):336-43.

82. Fukuoka H, Tange C, Otsuka R, Ando F, Shimokata H. The impact of anthropometric and ocular parameters on optic cup-to-disc ratio. BMJ Open Ophthalmology. 2017;1.

83. Berdahl JP, Fleischman D, Zaydlarova J, Stinnett S, Allingham RR, Fautsch MP. Body mass index has a linear relationship with cerebrospinal fluid pressure. Invest Ophthalmol Vis Sci. 2012;53(3):1422-7.

84. Berdahl JP, Allingham RR, Johnson DH. Cerebrospinal fluid pressure is decreased in primary open-angle glaucoma. Ophthalmology. 2008;115(5):763-8.

85. Ramrattan RS, Wolfs RC, Jonas JB, Hofman A, de Jong PT. Determinants of optic disc characteristics in a general population: The Rotterdam Study. Ophthalmology. 1999;106(8):1588-96.

86. Healey PR, Mitchell P, Smith W, Wang JJ. The influence of age and intraocular pressure

on the optic cup in a normal population. J Glaucoma. 1997;6(5):274-8.

87. Ramdas WD, Amin N, van Koolwijk LM, Janssens AC, Demirkan A, de Jong PT, et al. Genetic architecture of open angle glaucoma and related determinants. J Med Genet. 2011;48(3):190-6.

88. Tham YC, Liao J, Vithana EN, Khor CC, Teo YY, Tai ES, et al. Aggregate Effects of Intraocular Pressure and Cup-to-Disc Ratio Genetic Variants on Glaucoma in a Multiethnic Asian Population. Ophthalmology. 2015;122(6):1149-57.

89. Hysi PG, Cheng CY, Springelkamp H, Macgregor S, Bailey JN, Wojciechowski R, et al. Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma. Nat Genet. 2014;46(10):1126-30.

90. Varma R, Paz SH, Azen SP, Klein R, Globe D, Torres M, et al. The Los Angeles Latino Eye Study: design, methods, and baseline data. Ophthalmology. 2004;111(6):1121-31.

91. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75.

92. Gao X, Nannini DR, Corrao K, Torres M, Chen YI, Fan BJ, et al. Genome-wide association study identifies WNT7B as a novel locus for central corneal thickness in Latinos. Hum Mol Genet. 2016;25(22):5035-45.

93. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.

94. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003;164(4):1567-87.

95. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945-59.

96. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.

97. Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, et al. A genomewide admixture mapping panel for Hispanic/Latino populations. Am J Hum Genet. 2007;80(6):1171-8.

98. Gao X, Starmer J. Human population structure detection via multilocus genotype clustering. BMC Genet. 2007;8:34.

99. Delaneau O, Marchini J, Genomes Project C, Genomes Project C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun. 2014;5:3934.

100. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284-7.

101. Gao X, Haritunians T, Marjoram P, McKean-Cowdin R, Torres M, Taylor KD, et al. Genotype Imputation for Latinos Using the HapMap and 1000 Genomes Project Reference Panels. Front Genet. 2012;3:117.

102. Nannini DR, Torres M, Chen YI, Taylor KD, Rotter JI, Varma R, et al. A Genome-Wide Association Study of Vertical Cup-Disc Ratio in a Latino Population. Invest Ophthalmol Vis Sci. 2017;58(1):87-95.

103. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006;2(12):e190.

104. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55(4):997-1004.

105. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 2010;42(4):348-54.

106. Gao X. Multiple testing corrections for imputed SNPs. Genet Epidemiol. 2011;35(3):154-8.

107. Gao X, Becker LC, Becker DM, Starmer JD, Province MA. Avoiding the high Bonferroni penalty in genome-wide association studies. Genet Epidemiol. 2010;34(1):100-5.

108. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol. 2008;32(4):361-9.

109. R Development Core Team. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing, editor. Vienna, Austria2008.

110. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336-7.

111. Staples J, Qiao D, Cho MH, Silverman EK, University of Washington Center for Mendelian G, Nickerson DA, et al. PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. Am J Hum Genet. 2014;95(5):553-64.

112. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet. 2012;91(2):224-37.

113. QIAGEN. Ingenuity Pathway Analysis

114. Kooperberg C, Leblanc M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. Genet Epidemiol. 2008;32(3):255-63.

115. Paré G, Cook NR, Ridker PM, Chasman DI. On the Use of Variance per Genotype as a Tool to Identify Quantitative Trait Interaction Effects: A Report from the Women's Genome Health Study. PLOS Genetics. 2010;6(6):e1000981.

116. Zhang P, Lewinger JP, Conti D, Morrison JL, Gauderman WJ. Detecting Gene-Environment Interactions for a Quantitative Trait in a Genome-Wide Association Study. Genet Epidemiol. 2016;40(5):394-403.

117. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655-64.

118. Tomoyose E, Higa A, Sakai H, Sawaguchi S, Iwase A, Tomidokoro A, et al. Intraocular pressure and related systemic and ocular biometric factors in a population-based study in Japan: the Kumejima study. Am J Ophthalmol. 2010;150(2):279-86.

119. Han YS, Lee JW, Lee JS. Intraocular pressure and influencing systemic health parameters in a Korean population. Indian J Ophthalmol. 2014;62(3):305-10.

120. Halder I, Kip KE, Mulukutla SR, Aiyer AN, Marroquin OC, Huggins GS, et al. Biogeographic ancestry, self-identified race, and admixture-phenotype associations in the Heart SCORE Study. Am J Epidemiol. 2012;176(2):146-55.

121. Girkin CA, Nievergelt CM, Kuo JZ, Maihofer AX, Huisingh C, Liebmann JM, et al. Biogeographic Ancestry in the African Descent and Glaucoma Evaluation Study (ADAGES): Association With Corneal and Optic Nerve Structure. Invest Ophthalmol Vis Sci. 2015;56(3):2043-9.

122. Humes KR, Jones NA, Ramirez RR, United States. Bureau of the Census. Overview of race and Hispanic origin : 2010. Washington, D.C.: U.S. Dept. of Commerce, Economics and Statistics Administration, U.S. Census Bureau; 2011. 23 p. p.

123. Kass MA, Zimmerman TJ, Alton E, Lemon L, Becker B. Intraocular pressure and glaucoma in the Zuni indians. Arch Ophthalmol. 1978;96(12):2212-3.

124. Zhao D, Cho J, Kim MH, Guallar E. The association of blood pressure and primary open-angle glaucoma: a meta-analysis. Am J Ophthalmol. 2014;158(3):615-27 e9.

125. Macgregor S, Hewitt AW, Hysi PG, Ruddle JB, Medland SE, Henders AK, et al. Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. Hum Mol Genet. 2010;19(13):2716-24.

126. Khor CC, Ramdas WD, Vithana EN, Cornes BK, Sim X, Tay WT, et al. Genome-wide association studies in Asians confirm the involvement of ATOH7 and TGFBR3, and further identify CARD10 as a novel locus influencing optic disc area. Hum Mol Genet. 2011;20(9):1864-72.

127. Brown NL, Dagenais SL, Chen CM, Glaser T. Molecular characterization and mapping of ATOH7, a human atonal homolog with a predicted role in retinal ganglion cell development. Mamm Genome. 2002;13(2):95-101.

128. Cao D, Jiao X, Liu X, Hennis A, Leske MC, Nemesure B, et al. CDKN2B polymorphism is associated with primary open-angle glaucoma (POAG) in the Afro-Caribbean population of Barbados, West Indies. PLoS One. 2012;7(6):e39278.

129. Li Z, Allingham RR, Nakano M, Jia L, Chen Y, Ikeda Y, et al. A common variant near

TGFBR3 is associated with primary open angle glaucoma. Hum Mol Genet. 2015;24(13):3880-92.

130. Solouki AM, Verhoeven VJ, van Duijn CM, Verkerk AJ, Ikram MK, Hysi PG, et al. A genome-wide association study identifies a susceptibility locus for refractive errors and myopia at 15q14. Nat Genet. 2010;42(10):897-901.

131. Rebolleda G, Noval S, Contreras I, Arnalich-Montiel F, Garcia-Perez JL, Munoz-Negrete FJ. Optic disc cupping after optic neuritis evaluated with optic coherence tomography. Eye (Lond). 2009;23(4):890-4.

132. Syc SB, Warner CV, Saidha S, Farrell SK, Conger A, Bisker ER, et al. Cup to disc ratio by optical coherence tomography is abnormal in multiple sclerosis. J Neurol Sci. 2011;302(1-2):19-24.

133. Chua J, Vania M, Cheung CM, Ang M, Chee SP, Yang H, et al. Expression profile of inflammatory cytokines in aqueous from glaucomatous eyes. Mol Vis. 2012;18:431-8.

134. Nannini D, Torres M, Chen YD, Taylor KD, Rotter JI, Varma R, et al. African Ancestry Is Associated with Higher Intraocular Pressure in Latinos. Ophthalmology. 2016;123(1):102-8.

135. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190-1.

136. Ramdas WD, Wolfs RC, Hofman A, de Jong PT, Vingerling JR, Jansonius NM. Heidelberg Retina Tomograph (HRT3) in population-based epidemiology: normative values and criteria for glaucomatous optic neuropathy. Ophthalmic Epidemiol. 2011;18(5):198-210.

137. Kim YH, Choi BH, Cheon HG, Do MS. B cell activation factor (BAFF) is a novel adipokine that links obesity and inflammation. Exp Mol Med. 2009;41(3):208-16.

138. Krumbholz M, Theil D, Derfuss T, Rosenwald A, Schrader F, Monoranu CM, et al. BAFF is produced by astrocytes and up-regulated in multiple sclerosis lesions and primary central nervous system lymphoma. J Exp Med. 2005;201(2):195-200.

139. George-Chandy A, Trysberg E, Eriksson K. Raised intrathecal levels of APRIL and BAFF in patients with systemic lupus erythematosus: relationship to neuropsychiatric symptoms. Arthritis Res Ther. 2008;10(4):R97.

140. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One. 2012;7(12):e51954.

141. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet. 2009;41(3):342-7.

142. Jain D, Hodonsky CJ, Schick UM, Morrison JV, Minnerath S, Brown L, et al. Genome-wide association of white blood cell counts in Hispanic/Latino Americans: the Hispanic Community Health Study/Study of Latinos. Hum Mol Genet. 2017;26(6):1193-204.

143. Sasayama D, Hattori K, Ogawa S, Yokota Y, Matsumura R, Teraishi T, et al. Genome-wide quantitative trait loci mapping of the human cerebrospinal fluid proteome. Hum Mol Genet. 2017;26(1):44-51.

144. Akhabir L, Sandford A. Genetics of interleukin 1 receptor-like 1 in immune and inflammatory diseases. Curr Genomics. 2010;11(8):591-606.

145. Xi H, Katschke KJ, Jr., Li Y, Truong T, Lee WP, Diehl L, et al. IL-33 amplifies an innate immune response in the degenerating retina. J Exp Med. 2016;213(2):189-207.

146. Wood IS, Wang B, Trayhurn P. IL-33, a recently identified interleukin-1 gene family member, is expressed in human adipocytes. Biochem Biophys Res Commun. 2009;384(1):105-9.

147. Hasan A, Al-Ghimlas F, Warsame S, Al-Hubail A, Ahmad R, Bennakhi A, et al. IL-33 is negatively associated with the BMI and confers a protective lipid/metabolic profile in non-diabetic but not diabetic subjects. BMC Immunol. 2014;15:19.

148. Miller AM, Asquith DL, Hueber AJ, Anderson LA, Holmes WM, McKenzie AN, et al. Interleukin-33 induces protective effects in adipose tissue inflammation during obesity in mice. Circ Res. 2010;107(5):650-8.

149. Winnier DA, Fourcaudot M, Norton L, Abdul-Ghani MA, Hu SL, Farook VS, et al. Transcriptomic identification of ADH1B as a novel candidate gene for obesity and insulin resistance in human adipose tissue in Mexican Americans from the Veterans Administration Genetic Epidemiology Study (VAGES). PLoS One. 2015;10(4):e0119941.

150. Deming Y, Xia J, Cai Y, Lord J, Holmans P, Bertelsen S, et al. A potential endophenotype for Alzheimer's disease: cerebrospinal fluid clusterin. Neurobiol Aging. 2016;37:208 e1- e9.

151. Nikolskaya T, Nikolsky Y, Serebryiskaya T, Zvereva S, Sviridov E, Dezso Z, et al. Network analysis of human glaucomatous optic nerve head astrocytes. BMC Med Genomics. 2009;2:24.

152. Doudevski I, Rostagno A, Cowman M, Liebmann J, Ritch R, Ghiso J. Clusterin and complement activation in exfoliation glaucoma. Invest Ophthalmol Vis Sci. 2014;55(4):2491-9.

153. van Koolwijk LM, Ramdas WD, Ikram MK, Jansonius NM, Pasutto F, Hysi PG, et al. Common genetic determinants of intraocular pressure and primary open-angle glaucoma. PLoS Genet. 2012;8(5):e1002611.

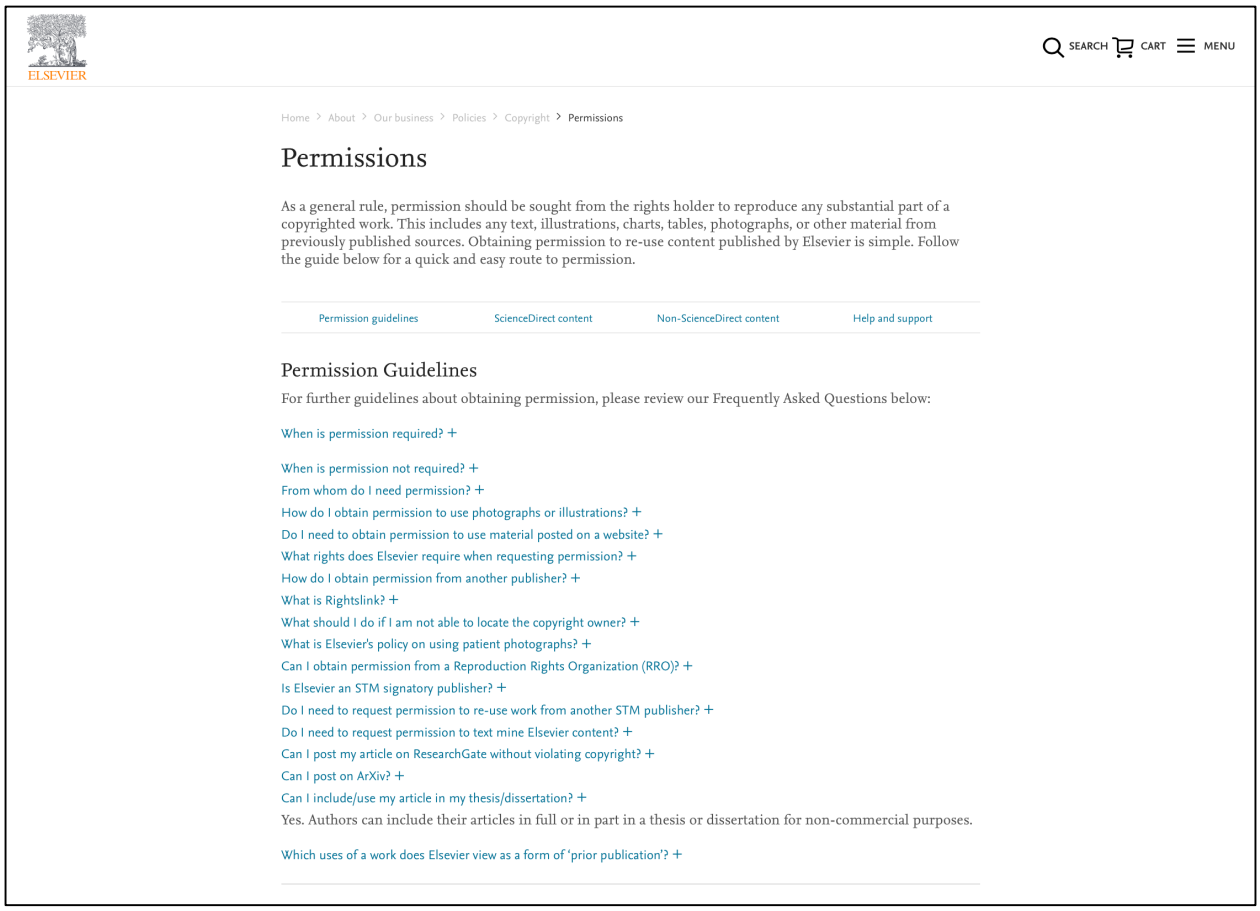
154. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet. 2017;49(3):403-15.

155. Smith JA, Ware EB, Middha P, Beacher L, Kardia SL. Current Applications of Genetic Risk Scores to Cardiovascular Outcomes and Subclinical Phenotypes. Curr Epidemiol Rep. 2015;2(3):180-90.

156. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nat Rev Genet. 2010;11(5):356-66.

Appendices

Appendix A. Copyright statement from Elsevier.



ELSEVIER

SEARCH CART MENU

Home > About > Our business > Policies > Copyright > Permissions

Permissions

As a general rule, permission should be sought from the rights holder to reproduce any substantial part of a copyrighted work. This includes any text, illustrations, charts, tables, photographs, or other material from previously published sources. Obtaining permission to re-use content published by Elsevier is simple. Follow the guide below for a quick and easy route to permission.

Permission guidelines | ScienceDirect content | Non-ScienceDirect content | Help and support

Permission Guidelines

For further guidelines about obtaining permission, please review our Frequently Asked Questions below:

When is permission required? +

When is permission not required? +
From whom do I need permission? +
How do I obtain permission to use photographs or illustrations? +
Do I need to obtain permission to use material posted on a website? +
What rights does Elsevier require when requesting permission? +
How do I obtain permission from another publisher? +
What is Rightslink? +
What should I do if I am not able to locate the copyright owner? +
What is Elsevier's policy on using patient photographs? +
Can I obtain permission from a Reproduction Rights Organization (RRO)? +
Is Elsevier an STM signatory publisher? +
Do I need to request permission to re-use work from another STM publisher? +
Do I need to request permission to text mine Elsevier content? +
Can I post my article on ResearchGate without violating copyright? +
Can I post on ArXiv? +
Can I include/use my article in my thesis/dissertation? +

Yes. Authors can include their articles in full or in part in a thesis or dissertation for non-commercial purposes.

Which uses of a work does Elsevier view as a form of 'prior publication'? +

Accessed at: http://www.elsevier.com/permissions

Appendix B: Written permission from IOVS.

Dear Drew,

Permission is hereby granted to include the following article in your PhD dissertation for the University of Illinois at Chicago:

Nannini DR, Torres M, Chen Y-DI, et al. A genome-wide association study of vertical cup-disc ratio in a Latino population. *Invest Ophthalmol Vis Sci*. 2017;58:87-95.

A reprint of this material must include a full article citation.

Best regards,

Debbie Chin
ARVO Journals

Association for Research in Vision and Ophthalmology
1801 Rockville Pike, Suite 400
Rockville, MD 20852 USA
Direct: +1.240.221.2926 | Main: +1.240.221.2900 | Fax: +1.240.221.0370
www.arvo.org

Appendix C. Comparison between METAL and EMMAX for the top genotyped VCDR SNPs.

SNP	METAL	EMMAX
rs1192419	9.02×10^{-08}	9.56×10^{-08}
rs7916697	1.80×10^{-11}	1.97×10^{-11}
rs1900005	9.76×10^{-12}	6.41×10^{-12}
rs16960773	2.89×10^{-07}	9.63×10^{-07}

VITA

NAME: Drew Robert Nannini

EDUCATION: B.S., Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2011

M.A., Bioethics, Kansas City University of Medicine and Biosciences, Kansas City, Missouri, 2012

M.S., Epidemiology, University of Illinois at Chicago, Chicago, Illinois, 2015

TEACHING: Department of Epidemiology and Biostatistics, University of Illinois at Chicago, Teaching Assistant, 2015-2017

HONORS: 19th Summer Institute in Statistical Genetics Scholarship, 2014

ABSTRACTS: Gao, X., Nannini, D., Torres, M., Goulee, K., Chen, Y., Taylor, K., Rotter, J., and Varma, R. (2015). Genetic variants associated with intraocular pressure in Latinos. The Association for Research in Vision and Ophthalmology. (Denver, CO).

Gao, X., Nannini, D., Torres, M., Chen, Y., Taylor, K., Rotter, J., and Varma, R. (2015). African ancestry is associated with intraocular pressure in Latinos. American Society of Human Genetics (Baltimore, MD).

Nannini, D., Goulee, K., Torres, M., Chen, Y.-D., Taylor, K., Gauderman, J., Rotter, J.I., Varma, R., and Gao, X.R. (2016). A genome-wide association study of central corneal thickness in Latinos. The Association for Research in Vision and Ophthalmology. (Seattle, WA).

Gao, X., Nannini, D., Torres, M., Chen, Y.I., Taylor, K.D., Rotter, J.I., and Varma, R. (2016). A genome-wide association study of vertical cup-disc ratio in Latinos. American Society of Human Genetics Meeting. (Vancouver, Canada).

Gao, X.R., Nannini, D., Torres, M., and Varma, R. (2017). European ancestry is associated with lower risk of primary open angle glaucoma in Latinos. The Association for Research in Vision and Ophthalmology. (Baltimore, MD).

Nannini, D.R., Kim, H., Fan, F., Gao, X.R., and Group, M.A.G.G.S. (2017). Genetic risk score is associated with vertical cup-disc ratio and improves prediction of primary open angle glaucoma in Latinos. The American Society of Human Genetics. (Orlando, FL).

PUBLICATIONS: Aghababaei N, Wasserman JA, Nannini D. The religious person revisited: cross-cultural evidence from the HEXACO model of personality structure. Mental Health, Religion & Culture 2014;17(1):24-9.

Jason Adam W, Naser A, Drew N. Culture, Personality, and Attitudes

Toward Euthanasia: A Comparative Study of University Students in Iran and the United States. OMEGA - Journal of Death and Dying 2015;72(3):247-70.

Nannini D, Torres M, Chen YD, et al. African Ancestry Is Associated with Higher Intraocular Pressure in Latinos. Ophthalmology 2016;123(1):102-8.

Gao X, Nannini DR, Corrao K, et al. Genome-wide association study identifies WNT7B as a novel locus for central corneal thickness in Latinos. Hum Mol Genet 2016;25(22):5035-45.

Nannini DR, Torres M, Chen YI, et al. A Genome-Wide Association Study of Vertical Cup-Disc Ratio in a Latino Population. Invest Ophthalmol Vis Sci 2017;58(1):87-95.