A Genome-Wide Association Study of Central Corneal Thickness in Latinos

Xiaoyi Gao,^{1,2,*} W. James Gauderman,² Yutao Liu,^{3,4} Paul Marjoram,² Mina Torres,^{1,2} Talin Haritunians,⁵ Jane Z. Kuo,⁵ Yii-Der I. Chen,⁵ R. Rand Allingham,⁴ Michael A. Hauser,^{3,4} Kent D. Taylor,⁵ Jerome I. Rotter,⁵ and Rohit Varma^{1,2,*}

¹Department of Ophthalmology, Keck School of Medicine, University of Southern California, Los Angeles, California

²Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California

³Department of Medicine, Duke University, Durham, North Carolina

⁴Department of Ophthalmology, Duke University, Durham, North Carolina

⁵Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California

Correspondence: Xiaoyi Gao, Department of Ophthalmology and Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033; rgao@uic.edu.

Rohit Varma, Department of Ophthalmology and Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033; rvarma@uic.edu.

Current affiliation: *Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, Illinois.

Submitted: January 19, 2013

Accepted: March 6, 2013

Citation: Gao X, Gauderman WJ, Liu Y, et al. A genome-wide association study of central corneal thickness in Latinos.

Invest Ophthalmol Vis Sci.

2013;54:2435–2443. DOI:10.1167/iovs.13-11692

PURPOSE. Central corneal thickness (CCT) is a clinically important risk factor for primary open-angle glaucoma and keratoconus. Genetic factors controlling CCT in Latinos, the most populous minority population in the United States, are unclear. Here we describe the first genome-wide association study (GWAS) report of CCT in Latinos.

METHODS. We performed a GWAS for CCT on 1768 Latinos recruited in the Los Angeles Latino Eye Study (LALES) using Illumina's HumanOmniExpress BeadChip (~730K markers). To discover additional associated single-nucleotide polymorphisms (SNPs), we imputed SNPs based on the 1000 Genomes Project reference panels. All subjects were 40 years of age and older. We used linear regression with adjustment for age, sex, and principal components of genetic ancestry.

RESULTS. We replicated the involvement of several previously reported loci, such as *RXRA-COL5A1*, *FOXO1*, and *ZNF469*, for CCT in Latinos ($P < 0.002$). Moreover, we discovered novel SNPs, rs3118515, rs943423, rs3118594, and rs3132307, that reached GWAS significance ($P < 5 \times 10^{-8}$) in the uncharacterized *LOC100506532* (gene type: miscRNA) for CCT in Latinos. By conditional analysis, we demonstrate that rs3118515 in this gene is responsible for the GWAS signal in the chromosome 9 *RXRA-COL5A1* region in Latinos. Moreover, multiple sources of ENCODE evidence suggest that rs3118515 is in a regulatory region. Reverse-transcription PCR products indicated that transcripts of *LOC100506532* surrounding rs3118515 were expressed in human corneas.

CONCLUSIONS. We discovered novel SNPs for CCT in Latinos and provided the first reported evidence of the corneal expression of *LOC100506532*. These results help to further increase our understanding of the genetic architecture of CCT.

Keywords: central corneal thickness, Latino, GWAS, rs3118515, LOC100506532

The cornea is the transparent portion of the ocular wall through which light reaches the interior structures of the eye. The refractive power of the cornea is greater than that of the lens itself and, therefore, plays a critical role in focusing light on the retina.¹ It has become increasingly apparent that factors associated with the central corneal thickness (CCT) may be related to ocular health. Numerous epidemiologic studies have shown that thinner CCT is a clinically important risk factor for primary open-angle glaucoma (POAG), one of the leading causes of irreversible blindness globally.^{2–10} CCT is sometimes used as a quantitative endophenotype for POAG.¹¹ Moreover, a thinner CCT has been observed in many other ocular diseases, such as keratoconus, brittle cornea syndrome, Ehlers-Danlos syndrome, and osteogenesis imperfecta.^{12,13}

Although CCT is a normally distributed quantitative trait in the general population, variation in CCT exists between different ethnic groups.¹⁴ Epidemiologic studies investigating CCT, either as the primary or the secondary outcome, have been conducted in Caucasians,^{15–17} people of African de-

scent,¹⁸ Latinos,¹⁹ American Indians/Alaskan Natives,²⁰ and several Asian populations.^{21,22} These studies have shown clear ethnic-related differences in CCT. In general, populations of African descent have lower CCTs than that of Caucasians, whereas Latinos have slightly higher CCT values, and Asians have a broad variation in CCT.¹⁴

Evidence from familial and twin studies indicates that CCT is a highly heritable trait.^{23–26} It was estimated that the heritability of CCT was at least 0.6²³ and could be as high as 0.95,²⁴ which suggests that environmental factors play a relatively small role in CCT. Recently, a number of genetic loci from genome-wide association studies (GWAS) have been reported to be associated with CCT in Caucasians and Asians, including *AKAP13*,^{27,28} *COL5A1*,^{28,29} *RXRA-COL5A1*,^{28–30} *COL8A2*,²⁹ *FAM53B*,¹³ *FOXO1*,¹³ *IBTK*,²⁷ *LRRK1*,²⁷ and *ZNF469*.^{13,28–30} In a recently published meta-analysis of CCT on a large number of European and Asian individuals, 16 new loci were identified.³¹ Previous genetic studies for CCT have been conducted in individuals of either European or Asian ancestry. There may well be some

ethnic difference for CCT (significant single-nucleotide polymorphisms [SNPs] in one population may not be significant in other populations). To date there has not been a genetic study of CCT in Latinos; therefore, it is not known whether CCT is regulated by recently reported genetic variants or by as yet unidentified genetic factors.

Here we describe the first GWAS report of CCT in Latinos by using directly genotyped SNPs in conjunction with genotype imputation based on the 1000 Genome Project (1KGP) reference panels.

MATERIALS AND METHODS

Ethics Statement

This research was approved by the University of Southern California Health Sciences Campus and Cedars-Sinai Medical Center Institutional Review Boards and all clinical investigation was conducted according to the principles expressed in the Declaration of Helsinki.

Study Sample

We conducted this research using 1768 GWAS genotyped samples from the Los Angeles Latino Eye Study (LALES), a population-based study of 6357 Latinos, living in six census tracts in the city of La Puente, Los Angeles County, California. All subjects were 40 years of age and older. Written, informed consent was obtained from all participants. The data obtained from LALES have been used to document the prevalence, incidence, and impact of visual impairment in Latinos.³²

Measurements of CCT

CCT measurements were obtained using a commercial ultrasonic computation module (A-scan/pachymeter DGH 4000B SBH IOL Computation Module; DGH Tech, Inc., Exton, PA). Three measurements of each eye were taken and averaged to yield a single final value. We used the average CCT between the right eye and left eye CCT measurements for all the downstream analyses. If only one eye's CCT measurement was available, it was used as a substitute for the average. We used the inverse normal transformation to convert the data to the standard normal distribution.

Genotyping and Quality Control

We genotyped 1815 Latinos recruited in LALES using Illumina OmniExpress BeadChip Kit (730,525 markers; Illumina, Inc., San Diego, CA). We also included 76 duplicate samples to verify genotyping reproducibility. The genotyping was performed at the Genotyping Laboratory of the Medical Genetics Institute and Clinical and Translational Science Institute (CTSI) of the Cedars-Sinai Medical Center. SNPs were called using commercial software (Illumina GenomeStudio, v2011.1; Illumina, Inc.). Individuals were excluded if genotyping call rates were <97%. The average call rate was >99.32%. The reproducibility was >99.99%. Genotypes for 647,630 well-clustered SNPs were released by the genotyping laboratory. We used PLINK (v1.07)³³ to perform quality control. Individuals with inconsistency between reported sex and genetic sex ($n = 12$) and unexpected duplicates ($n = 24$) were dropped from the analysis. We also removed CCT outliers ($n = 3$) and individuals with a missing CCT phenotype ($n = 8$) in our genotyped samples. In the end, 1768 individuals remained in the final analysis, among which 1644 unrelated subjects were used as a stage 1 discovery set, and 124 first-degree relatives from 59 families were used as a stage 2 replication set. Markers

were excluded if minor allele frequency (MAF) < 0.01, call rates < 95%, or if Hardy-Weinberg equilibrium P values < 10^{-6} . This resulted in 587,456 SNPs in the final analysis. SNPs were coded on the forward strand to facilitate the imputation process.

Genotype Imputation

To interrogate additional SNPs not directly genotyped, we conducted genotype imputation using MACH^{34,35} and the 1KGP reference panels. MACH (v1.0.16.c) and 1KGP Phase I (α) phased haplotypes were downloaded from the MACH software's website (see Web Resources). MACH uses a Markov-chain algorithm and has been shown to be one of the leading algorithms for genotype imputation.³⁵ The 1KGP reference panels, with the inclusion of whole-genome sequencing data, contain a large number of variants: 38.9 million. We used the AMR+CEU+YRI reference panel (a combination of Mexican, Colombian, Puerto Rican, CEPH, and Yoruba haplotypes) since we have shown that this panel gave the highest genotype imputation accuracy for Latinos.³⁶

We used the standard genotype imputation approach (in MACH) and specified 50 iterations of the Markov sampler and 400 haplotypes when updating the phase for each individual. Imputed genotypes were coded as allelic dosages (fractional counts ranging from 0–2). Imputed SNPs with a MACH Rsq (an estimate of the squared correlation between true genotypes and estimated allelic dosage³⁵) < 0.80 and MAF < 0.01 were removed. In all, 6,290,547 imputed SNPs remained in the analysis.

Statistical Analysis

Principal components (PCs) of genetic ancestry were inferred using EIGENSOFT.³⁷ To make comparisons to reference populations of known ancestry, we included all the unrelated North Europeans (CEU, $n = 60$), West Africans (YRI, $n = 60$), and East Asians (CHB, $n = 45$; JPT, $n = 44$) in the HapMap Phase 3 project³⁸ and Native Americans ($n = 105$).³⁹ The first four PCs were retained and used as covariates in the downstream association analysis. Moreover, the genomic control (GC) inflation factor⁴⁰ was calculated and a quantile-quantile (Q-Q) probability plot was generated to visualize the distribution of the test statistics. In the stage 1 discovery, association analysis was conducted using linear regression with adjustment for age, sex, and principal components of genetic ancestry, and assuming an additive genetic effects model. Genotyped SNPs were analyzed using PLINK software.³³ In stage 2 replication, analyses of the association between SNPs and CCT were conducted using a linear mixed-effects model (Proc Mixed procedure of SAS v9.2; SAS Institute, Cary, NC), with adjustment for age, sex, and principal components of genetic ancestry. The empirical “sandwich” estimator and compound symmetry covariance structure were used. Fixed-effects meta-analyses of stage 1 and stage 2 data using inverse-variance weighting were performed using METAL.⁴¹ Imputed SNPs were analyzed using the mach2qtl software (see Web Resources) and genotype imputation uncertainty was accounted for by using allelic dosage. SNPs with values of $P < 5 \times 10^{-8}$ were declared genome-wide significant. For replicating previously published loci, we applied the *simpleM* method^{42,43} for multiple testing correction. Conditional association analysis was performed using the --condition option (in PLINK) for directly genotyped SNPs and R for imputed SNPs, including the lead SNP being conditioned upon as a covariate in the regression model. R software (provided in the public domain by R Foundation for Statistical Computing, Vienna, Austria, available at <http://www.r-project.org/>) was used for graphing.⁴⁴

Expression Analysis

To examine the expression status of transcripts surrounding rs3118515 in ocular tissues, we designed RT-PCR experiments for a known transcript of uncertain coding potential (i.e., TCONS_00015892) (forward 5'-CCT CAT TCC TTC TCC CTT CC-3' and reverse 5'-GTG GAA CTG GGA CAA TGG AG-3'), using Primer3 software v.0.4.0.^{45,46} The expected PCR product of TCONS_00015892 is 224 bp in length. Nonglaucomatous ocular tissues were obtained from human donor eyes through North Carolina Eye Bank (Winston-Salem, NC). These tissues included cornea ($n = 4$), trabecular meshwork ($n = 5$), retina ($n = 2$), and optic nerve head ($n = 1$). Total RNA was extracted using *mir*VanamiRNA isolation kit (Life Technologies, Grand Island, CA). Reverse transcription reaction was performed (SuperScript III reverse transcriptase; Life Technologies). The PCR reaction was done next (Apex Hot Start *Taq* DNA polymerase; Genesee Scientific, San Diego, CA). All PCR reactions were done according to the recommended standard protocol. The PCR reaction for an unrelated gene, *GALC* (galactosylceramidase, forward 5'-GGA AGA AGC TTT GGT CTT CTGA-3' and reverse 5'-CCC AGA CAG GAG ATT CTA CCAC-3'), was also performed to ensure the quality of cDNA samples. The expected PCR product of *GALC* was 228 bp in length. The PCR products were examined using gel electrophoresis with 2% agarose gel.

RESULTS

The characteristics of the study sample are presented in Table 1. The study subjects were Latinos recruited from the Los Angeles Latino Eye Study (LALES). Overall, 1768 individuals (1644 in stage 1 discovery and 124 first-degree relatives from 59 families in stage 2 replication) contributed information to the analysis of CCT. The overall mean (SD) age was 57.1 (10.6) and 56.6 (12.0) years for the discovery and replication data sets, respectively. The proportion of females was 54.9% in the discovery sample and 70.9% in the replication sample. The average CCT of both eyes was used as the phenotype and had a mean (SD) of 551.6 (33.5) μm (range: 451–671 μm) and 548.6 (33.6) μm (range: 454–640 μm) for the discovery and replication data sets, respectively.

Results From Measured Genotypes

The GC inflation factor⁴⁰ was moderate, $\lambda = 1.03$. A Q-Q plot of SNP association P values (GC corrected) is shown in Supplementary Figure S1. The Q-Q plot suggests that the observed P values do not deviate from their expected distribution, except at the extreme tail. Overall, the Q-Q plot indicated good control of population stratification for our Latino subjects.

A Manhattan plot of genome-wide P values from stage 1 discovery association analyses is shown in Figure 1. The results for top hits are summarized in Table 2. One SNP, rs3118515 ($P = 7.79 \times 10^{-9}$, GRCh37/hg19 position 137436314), on chromosome 9q34.3 passed the commonly used GWAS significance threshold ($P < 5 \times 10^{-8}$).⁴⁷ The minor allele A of rs3118515

(MAF = 0.262) was associated with reduced CCT with β (SE) = -0.23 (0.04). Another SNP, rs943423 ($P = 1.03 \times 10^{-7}$, GRCh37/hg19 position 137437183), 869 bp downstream of rs3118515 reached borderline GWAS significance. The minor allele C of rs943423 (MAF = 0.259) was associated with reduced CCT with β (SE) = -0.21 (0.04). Both rs3118515 and rs943423 reside in the uncharacterized gene, *LOC100506532* (gene type: miRNA; source: NCBI), which encompassed 18-kb DNA in the region situated 85.6 kb 3' of the *RXRA* gene (retinoic X receptor alpha) and 96.5 kb 5' of the *COL5A1* gene (collagen, type V, alpha 1). We then performed an in silico replication of rs3118515 using pedigree data from stage 2 replication and linear mixed-effects model to account for relatedness. The estimated effect of the minor allele showed the same direction of association for the discovery and replication sets. A meta-analysis of the discovery and replication data sets strengthened the association with rs3118515, $P = 8.25 \times 10^{-10}$. Therefore, rs3118515 represents a novel association with CCT in Latinos. Moreover, after meta-analysis of stage 1 discovery and stage 2 replication results, rs943423 also gained GWAS significance ($P = 1.41 \times 10^{-8}$). There was modest linkage disequilibrium (LD) between rs3118515 and rs943423 ($r^2 = 0.60$) in our Latino data set. rs3118515 explained 2.8% of the variance in CCT. Neither rs3118515 nor rs943423 has previously been reported for association with CCT.

Results From 1KGP Imputed Data

For this stage of the genetic association analysis, we used all available unrelated individuals, including unrelated individuals within pedigrees (final $n = 1707$). Instead of meta-analyzing stage 1 and stage 2 results, this approach balanced the needs to improve study power and to reduce computing complexity. Analysis of imputed SNPs did not identify any associations that were stronger than SNPs that were directly genotyped. A regional SNP association plot for the *LOC100506532* region is shown in Figure 2A. Except for SNPs in strong or moderate LD with rs3118515, most SNPs in the region remained insignificant. Other than the two significant measured SNPs (rs3118515 and rs943423), two additional imputed SNPs in this region were identified: rs3118594 with $P = 9.03 \times 10^{-9}$ and rs3132307 with $P = 1.57 \times 10^{-8}$. The genomic control inflation factor for the association of the imputed SNPs was $\lambda = 1.02$.

Conditional Association

Conditional analysis is often used to determine whether additional associated SNPs exist and which SNP(s) contribute most to the association.^{48–50} We conducted a conditional analysis for every SNP in the *RXRA-COL5A1* region (from the 5' start of *RXRA* to the 3' end of *COL5A1*), including the most significantly associated SNP, rs3118515, as a covariate and have summarized the results in Figure 2B. In Figure 2B, it is clear that previously significant SNPs become nonsignificant after conditioning on rs3118515, which suggests that rs3118515 is the dominant contributor to association with CCT including SNPs that have been previously reported.

TABLE 1. Characteristics of the Study Sample

Study	Sample Size	Females, %	Age (y), Mean (SD)	CCT (μm) Mean (SD)	CCT (μm) Range
Stage 1 discovery	1644	54.9	57.1 (10.5)	551.6 (33.4)	451–671
Stage 2 replication	124 (59 families)	70.9	56.6 (12.0)	548.6 (33.6)	454–640
Total	1768	56.0	57.1 (10.6)	551.4 (33.5)	451–671

Sample size $n = 1768$ (individuals who have complete phenotype, covariates, and genotype data).

FIGURE 1. Manhattan plot of the genome-wide P values. The results for the 587,456 genotyped SNPs are plotted as $-\log_{10}(P \text{ value})$ by genomic position.

Interrogation of Published Loci for CCT

To determine whether the associations observed in populations of European and Asian ancestry are relevant in the Latino population, we evaluated 35 previously reported CCT SNPs with $P < 5 \times 10^{-8}$. The results are summarized in Table 3. We observed that most of these SNPs have values of $P < 0.05$ in LALES (24/35 = 69%). Moreover, the directions of association are consistent with earlier reports. Among the 18 reports that provide the directions of effect alleles for CCT, all were replicated in our analysis. It is unlikely that this degree of directional consistency is due to chance ($P = 3.81 \times 10^{-6}$). To address multiple testing correction, we applied the *simpleM* method^{42,43} and estimated that the effective number of independent tests for the 35 SNPs was 25. Replication of these data for nine SNPs, those residing in or nearby *RXRA-COL5A1*, *COL5A1*, *FOXO1*, and *ZNF469*, survived Bonferroni correction for multiple tests (marked in boldface in Table 3). We were unable to replicate the association of CCT with *COL8A2*, *FAM53B*, and *AKAP13*, even when we provided the most significant hits (directly genotyped) ± 100 kb of the previously reported SNPs in Table 3. Interestingly, the *COL8A2* locus could not be replicated in Europeans either.³⁰ Compared with a similar attempt to replicate known CCT loci in samples of European ancestry,¹¹ we replicated more SNPs in our Latino samples. Lu et al.³¹ recently reported a new set of loci for CCT. Again, we see highly consistent directions of association in our data (Supplementary Data, Supplementary Table S2).

Biological Evidence

A UCSC Genome Browser screen shot of the 9q34.3 region is shown in Supplementary Figure S2. rs3118515 and rs943424 are highlighted in red and orange, respectively. rs3118515 overlaps with peaks of ENCODE⁵¹ functional assays, such as

histone modification peaks, DNase I hypersensitive sites, and transcription factor ChIP-seq peaks (*c-fos*, *c-Jun*, *JunD*, and *FOSL2*; gray boxes; darkness shows signal strength), which suggests that rs3118515 is a functional SNP for CCT.

Reverse transcription polymerase chain reaction (RT-PCR) products for transcripts of *LOC100506532* surrounding rs3118515 in a variety of normal human ocular tissues are shown in Figure 3. The expected 224-bp RT-PCR product is seen in human cornea (lane C3 of Fig. 3A). The quality of cDNA was assessed using cDNA primers for the *GALC* (galactosylceramidase) gene in Figure 3B. It is clear that *GALC* is consistently expressed in all cDNA samples. This result is the first reported evidence that transcripts of *LOC100506532* are expressed in human corneas.

DISCUSSION

This is the first reported GWAS to investigate the genetic determinants of CCT in a Latino population. We discovered rs3118515 in the uncharacterized *LOC100506532* gene as a novel SNP for CCT. A second SNP, rs943423, in the same gene, was also significantly associated after combination of the discovery and replication data sets. We have replicated the previous associations of *RXRA-COL5A1*, *FOXO1*, and *ZNF469* with CCT. Genotype imputation based on the 1KGP reference panels significantly increased the number of SNPs that we could interrogate but did not identify more highly associated SNPs. Two imputed SNPs, rs3118594 and rs3132307, upstream of rs3118515, also reached GWAS significance.

The SNP rs3118515 resides in the uncharacterized gene, *LOC100506532*, located between *RXRA* and *COL5A1*. Multiple SNPs within the 200-kb region of *RXRA-COL5A1* (Table 3) were reported by Vitart et al.,²⁸ Vithana et al.,²⁹ and Hoehn et

TABLE 2. Genome-Wide Significant Association With CCT in Latinos

SNP ID	Chr	Position	Gene	A1/A2	MAF	Discovery		Replication		Meta-Analysis
						β	P	β	P	P
rs3118515	9	137,436,314	LOC100506532 , <i>RXRA</i> , <i>COL5A1</i>	A/G	0.262	−0.23	7.79E-09	−0.40	0.01	8.25E-10
rs943423	9	137,437,183	LOC100506532 , <i>RXRA</i> , <i>COL5A1</i>	C/T	0.259	−0.21	1.03E-07	−0.35	0.02	1.41E-08

The closest gene to the SNP is listed first and is in boldface if the SNP locates within the gene. Chr, chromosome; A1/A2, allele 1/allele 2.

FIGURE 2. Regional SNP association plot for the *LOC100506532* region. (A) The lead SNP, rs3118515, is plotted in blue with its *P* value indicated ($n = 1707$ unrelated individuals). Genotyped and imputed SNPs are plotted as diamonds and circles, respectively. The degree of pairwise LD between the lead and neighboring SNPs is plotted in different colors: red, strong LD with $r^2 \geq 0.8$; orange, moderate LD with $0.5 \leq r^2 < 0.8$; yellow, weak LD with $0.2 \leq r^2 < 0.5$; white, no LD with $r^2 < 0.2$. Genes in the region are indicated with green arrows showing the strand orientation. The associated SNPs lie within *LOC100506532* surrounded by *RXRA* and *COL5A1*. SNP positions are according to GRCh37/hg19. All *P* values are nonconditioned. (B) Regional association plot conditional on the lead SNP, rs3118515.

al.,³⁰ and, thus, this study represents an independent confirmation of this region in this Latino population. Interestingly, rs3118515 was not investigated in previous GWAS studies of CCT. This may have occurred since rs3118515 is not included in earlier GWAS genotyping arrays (the Illumina Human610-Quad²⁹) or, on the HapMap Phase II reference panels, if genotype imputation was performed.^{28,30} At present, it is not clear whether the association of rs3118515 with CCT is specific to the Latino population. It would be interesting to know the association strength of rs3118515, either directly genotyped or imputed, in earlier published studies. Results of our conditional association analysis suggest that this SNP provides a major portion of the association with CCT in the *RXRA-COL5A1* region in Latinos. In Supplementary Figure S2, it is also interesting to notice that the previous GWAS hits for CCT in the *RXRA-COL5A1* region are either upstream or downstream of rs3118515, approximately 4 kb away. Previously reported GWAS SNPs (highlighted in green) associated with CCT do not overlap with ENCODE functional assays, such as DNase I-hypersensitive clusters and transcription factor

ChIP-seq peaks. rs3118515 (highlighted in red) lies within peak binding regions for transcription factors, *c-fos*, *c-Jun*, *JunD*, and *FOSL2*, as well as DNase I hypersensitive sites. Multiple sources of ENCODE evidence suggest that rs3118515 is in a regulatory region and is an excellent functional candidate for CCT. Moreover, RT-PCR products indicated that transcripts of *LOC100506532* surrounding rs3118515 are expressed in human corneas. This is the first time that a variant within *RXRA-COL5A1* has been linked with biological evidence for CCT.

rs3118515 was not captured by strong LD with previously reported CCT SNPs in *RXRA-COL5A1*. A summary of the LD r^2 (derived from 1KGP data) among all the genome-wide significant SNPs within *RXRA-COL5A1* is shown in Supplementary Table S1. Among the eight SNPs reported in Caucasians and Asians/Chinese, that is, rs1409832, rs4842044, rs1536478, rs3118516, rs3118518, rs3132306, rs1536482, and rs3132304 (rs3118518 and rs3132304 were reported in the supplementary appendix by Cornes et al.²⁷), all have weak to no LD with rs3118515, except that rs1409832 at 8 kb upstream shows moderate LD with r^2 values of 0.64 and 0.73 in CEU (CEPH in Utah residents) and CHB (Han Chinese in Beijing, China), respectively. It is interesting to note that previously reported rs4842044 and rs1536478 have complete LD with each other in CEU and CHB. rs3118516, rs3118518, rs3132306, and rs1536482 also have complete LD or nearly complete LD in CEU and CHB. Furthermore, none of the eight previously reported SNPs provides biological evidence or expression support with CCT. Therefore, the association of rs3118515 with CCT cannot be explained by previously reported SNPs within the *RXRA-COL5A1* region.

The strengths of our study were that we used both population- and family-based study designs, which complement each other. Additionally, we used genotype imputation based on the 1KGP reference panel, which offers extensive genomic coverage for nearly all common variants. We carefully designed our imputation procedure in Latinos³⁶ and retained only SNPs of high imputation quality. Our sample size was sufficient to achieve 95% power to detect variants with MAF 0.26 explaining 2.8% (like rs3118515) of trait variation (two-sided test, $\alpha_p = 5 \times 10^{-8}$, calculated by Quanto⁵²⁻⁵⁴). With advances in understanding the functional elements in the human genome,⁵¹ the newly discovered rs3118515 for CCT resides in a regulator region based on the ENCODE evidence. We also demonstrated by expression analysis that transcripts harboring rs3118515 were expressed in human corneas. This is the first time that the association of the 200-kb *RXRA-COL5A1* region with CCT was narrowed down to a variant with strong biological relevance. Like many other GWAS, this study is not without limitations. rs3118515 explained only a small portion, less than 3%, of the variance in CCT in our data, suggesting that important loci remain to be determined. We are extending our genotyping effort to the entire LALES cohort. As our genotyping progresses, we will have more power to detect additional loci. The expression was observed in one of the four human corneas. Unfortunately, we do not have any CCT measurement for these corneas. The ages for these four cornea samples are 77, 49, 61, and 53 years. The first two are males and the last two are females. The postmortem delay time between death and sample collection was within 6.5 hours. The primary causes of death were intracranial hemorrhage; renal cancer with metastases to liver, bone, and peritoneum; stroke with intracranial hemorrhage; and breast cancer with metastases to liver and brain. It is not clear whether any of these clinical factors was related with the specific corneal expression.

In summary, in this first GWAS of CCT in Latinos, we discovered the involvement of novel SNPs, both directly

TABLE 3. Comparison With Previously Reported SNPs for CCT

Previously Reported										LALES				Consistency				Most Significant Hits ± 100 kb	
SNP ID	Chr	Position	Genes	Effect	Allele	Freq	β	Reference(s)	A1/A2	Freq	β	P	Imputed	of Direction	SNP ID	Position	P		
rs7367703	1	36555758	COL8A2				-4.43	29	T/C	0.07	-0.04	5.08E-01		NA	rs11577123	36600773	1.76E-02		
rs7550047	1	36567343	COL8A2				-4.42	29	A/G	0.91	-0.07	2.24E-01	Y	NA	rs11577123	36600773	1.76E-02		
rs96067	1	36571920	COL8A2				-4.80	29	C/T	0.39	-0.06	6.98E-02		NA	rs11577123	36600773	1.76E-02		
rs1538138	6	82794594	IBTK	T	0.18-0.30		-4.23	27	T/C	0.14	-0.13	9.98E-03		Y	rs16893934	82856409	3.15E-05		
rs4718428	7	66421446	C7orf42	G	0.46-0.74		-3.18	27	G/T	0.44	0.08	1.16E-02	Y	Y	rs2901311	66415069	3.75E-03		
rs1324183	9	13557491	9p23	A	0.21-0.27		-3.37	27	C/A	0.87	-0.05	3.83E-01	Y	Y	rs9969775	13571933	2.17E-02		
rs1409832	9	137428425	RXR4-COL5A1				-3.95	29	G/T	0.23	-0.17	1.54E-05		NA	rs3118515	137436314	3.50E-09		
rs4842044	9	137431904	RXR4-COL5A1				-4.67	29	T/C	0.53	-0.17	1.62E-06	Y	NA	rs3118515	137436314	3.50E-09		
rs1536478	9	137432248	RXR4-COL5A1				-4.63	29	A/G	0.53	-0.17	1.58E-06	Y	NA	rs3118515	137436314	3.50E-09		
rs3118516	9	137439792	RXR4-COL5A1	A	0.34		-0.15	30	G/A	0.75	-0.21	7.69E-07	Y	Y	rs3118515	137436314	3.50E-09		
rs3132306	9	137440212	RXR4-COL5A1	T	0.66		0.15	30	T/C	0.74	-0.20	4.86E-07	Y	Y	rs3118515	137436314	3.50E-09		
rs1536482	9	137440528	RXR4-COL5A1	G	0.34		0.22	28,30 (A, Freq = 0.33, β = -0.15)	G/A	0.73	-0.22	5.71E-08	Y	Y	rs3118515	137436314	3.50E-09		
rs7044529	9	137568051	COL5A1					29	T/C	0.21	-0.12	5.29E-03		NA	rs10858254	137485476	2.13E-04		
rs1006368	10	126346603	FAM53B	A/G				13	C/T	0.80	-0.01	8.64E-01	Y	NA	rs7911370	126320992	1.80E-02		
rs11245330	10	126380338	FAM53B	A/G				13	A/G	0.20	0.00	9.74E-01		NA	rs7911370	126320992	1.80E-02		
rs1034200	13	23228691	FGF9-FTHL7				0.14	28	T/G	0.24	0.07	1.02E-01		NA	rs4081797	23234909	4.21E-03		
rs2755237	13	41109429	FOXO1	A/C				13	C/A	0.14	-0.14	5.32E-03		NA	rs2721051	41110884	5.67E-04		
rs6496932	15	85825567	PDE8A-AKAP13	A/G				13	A/G	0.08	-0.22	5.67E-04		NA	rs2721051	41110884	5.67E-04		
rs1828481	15	85840912	AKAP13	C	0.45-0.56		0.13	28	A/C	0.51	0.05	1.77E-01	Y	NA	rs7183764	85903051	3.78E-02		
rs7172789	15	85843517	AKAP13	C	0.45-0.56		3.14	27	C/T	0.49	0.05	1.42E-01		Y	rs7183764	85903051	3.78E-02		
rs930847	15	101558562	LRRK1	G	0.17-0.39		3.72	27	C/A	0.18	0.08	7.82E-02		Y	rs6598411	101542877	1.14E-02		
rs4965359	15	101585336	LRRK1	A	0.40-0.67		-3.50	27	G/A	0.45	0.08	2.67E-02		Y	rs6598411	101542877	1.14E-02		
rs12447690	16	88298124	ZNF469	T/C			0.17	13,29,30 (G, β = -0.18), 30 (T, Freq = 0.64, β = 0.16)	C/T	0.36	-0.08	3.34E-02		Y	rs9938149	88331640	9.05E-04		
rs7500824	16	88299491	ZNF469	A	0.36		-0.16	30	A/G	0.22	-0.09	2.69E-02		Y	rs9938149	88331640	9.05E-04		
rs7405095	16	88307825	ZNF469	A	0.36		-0.16	30	G/A	0.80	-0.13	5.56E-03	Y	Y	rs9938149	88331640	9.05E-04		
rs7501109	16	88320862	ZNF469	C	0.64		0.16	30	C/G	0.80	-0.14	4.06E-03	Y	Y	rs9938149	88331640	9.05E-04		
rs7501402	16	88320911	ZNF469	A	0.36		-0.16	30	T/C	0.64	-0.10	1.42E-02	Y	Y	rs9938149	88331640	9.05E-04		
rs6540223	16	88321436	ZNF469	T	0.64		0.16	30	T/C	0.80	-0.13	5.32E-03	Y	Y	rs9938149	88331640	9.05E-04		
rs12448211	16	88330513	ZNF469	A	0.62		0.16	30	A/G	0.65	-0.10	7.76E-03	Y	Y	rs9938149	88331640	9.05E-04		
rs9938149	16	88331640	ZNF469	A/C				13,29,30 (A, Freq = 0.62, β = 0.16)	C/A	0.21	-0.15	9.05E-04		Y	rs9938149	88331640	9.05E-04		
rs9922572	16	88334112	ZNF469	A	0.34		-0.14	30	C/A	0.81	-0.15	1.70E-03	Y	Y	rs9938149	88331640	9.05E-04		
rs9925231	16	88338107	ZNF469				-4.79	29	C/T	0.68	-0.10	1.12E-02	Y	NA	rs9938149	88331640	9.05E-04		
rs7204132	16	88344517	ZNF469				-4.95	29	G/T	0.83	-0.12	1.60E-02	Y	NA	rs9938149	88331640	9.05E-04		
rs9927272	16	88346709	ZNF469				-3.95	29	A/G	0.66	-0.04	2.98E-01	Y	NA	rs9938149	88331640	9.05E-04		

For previously reported SNPs, if multiple references are available, extra information on the allele, frequency, and effective size is given in parentheses. For LALES, the frequency of allele 1 is given. If the SNP is directly genotyped, allele 1 is the minor allele and β models the effect on CCT per increase of one minor allele. If the SNP is imputed, a negative β means allele 2 is associated with reduced trait values. All imputed SNPs had MACH Rsq ≥ 0.80. *simpleM* estimated the effective number of independent tests to be 25 and hence the Bonferroni correction gave 0.05/25 = 0.002. *P* values that survive the multiple testing correction are shown in bold. The most significant hits ±100 kb of the reported SNPs tested are also listed. SNP positions are according to GRCh37/hg19. Freq, frequency.

FIGURE 3. Reverse-transcription PCR analysis of *LOC100506532* in human ocular tissues. Lanes L to ONH denote molecular-weight marker, cornea 1, cornea 2, cornea 3, cornea 4, trabecular meshwork (TM) 1, TM2, TM3, TM4, TM 5, retina 1, retina 2, and optic nerve head, respectively. The 224-bp reverse-transcript PCR product indicated that transcripts of *LOC100506532* surrounding rs3118515 were observed in human cornea; that is, lane C3 of (A). The quality of cDNA was assessed using cDNA primers for the *GALC* (galactosylceramidase) gene in (B). The consistent expression of *GALC* in all cDNA samples further supported the corneal expression of *LOC100506532*.

genotyped and imputed, that reached genome-wide significance. Conditional analysis indicates that, at least in this Latino population, the responsible SNP in the chromosome 9 *RXRA-COL5A1* region is rs3118515 and that this may be a functional variant. In addition, we replicated the involvement of several previously reported genes for CCT in this population. These results help to further increase our understanding of the genetic architecture of CCT, demonstrating the utility of using Latinos for GWAS analyses.

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes Project Phase I haplotypes, <http://www.sph.umich.edu/csg/abecasis/MACH/download/1000G-PhaseI-Interim.html>
EIGENSOFT, <http://genepath.med.harvard.edu/sreich/Software.htm>
UCSC Genome Browser, <http://genome.ucsc.edu/>
MACH, <http://www.sph.umich.edu/csg/abecasis/MACH/download/>
mach2qtl, <http://www.sph.umich.edu/csg/abecasis/MACH/download/>
METAL, <http://www.sph.umich.edu/csg/abecasis/metal/>
NCBI, <http://www.ncbi.nlm.nih.gov/>
PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>
QUANTO, <http://hydra.usc.edu/gxe/>
simpleM, <http://simplem.sourceforge.net/>

Acknowledgments

The authors thank the study participants in LALES and study staff who helped with the data collection.

Supported in part by National Institutes of Health Grants

U10EY011753 (RV), R01EY022651 (XG), R21HL115606 (WJG), the National Institute of Diabetes and Digestive and Kidney Disease Grant DK063491 to the Southern California Diabetes Endocrinology Research Center, and of the Cedars-Sinai Board of Governor's Chair in Medical Genetics (JIR). DNA handling and genotyping at Cedars-Sinai Medical Center was supported in part by the National Center for Research Resources Grant UL1RR033176, now at the National Center for Advancing Translational Sciences, CTSI Grant UL1TR000124.

Disclosure: **X. Gao**, None; **W.J. Gauderman**, None; **Y. Liu**, None; **P. Marjoram**, None; **M. Torres**, None; **T. Haritunians**, None; **J.Z. Kuo**, None; **Y.-D.I. Chen**, None; **R.R. Allingham**, None; **M.A. Hauser**, None; **K.D. Taylor**, None; **J.I. Rotter**, None; **R. Varma**, None

References

- Lang GK. *Ophthalmology: A Pocket Textbook Atlas*. 2nd ed. New York: Thieme Medical; 2007.
- Gordon MO, Beiser JA, Brandt JD, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Arch Ophthalmol*. 2002;120:714–720; discussion 829–830.
- Leske MC, Heijl A, Hyman L, Bengtsson B, Dong L, Yang Z. Predictors of long-term progression in the early manifest glaucoma trial. *Ophthalmology*. 2007;114:1965–1972.
- Miglior S, Pfeiffer N, Torri V, Zeyen T, Cunha-Vaz J, Adamsons I. Predictive factors for open-angle glaucoma among patients with ocular hypertension in the European Glaucoma Prevention Study. *Ophthalmology*. 2007;114:3–9.
- Fan BJ, Wiggs JL. Glaucoma: genes, phenotypes, and new directions for therapy. *J Clin Invest*. 2010;120:3064–3072.
- Francis BA, Varma R, Vigen C, et al. Population and high-risk group screening for glaucoma: the Los Angeles Latino Eye Study. *Invest Ophthalmol Vis Sci*. 2011;52:6257–6264.

7. Herndon LW, Weizer JS, Stinnett SS. Central corneal thickness as a risk factor for advanced glaucoma damage. *Arch Ophthalmol*. 2004;122:17–21.
8. Jiang X, Varma R, Wu S, et al. Baseline risk factors that predict the development of open-angle glaucoma in a population: the Los Angeles Latino Eye Study. *Ophthalmology*. 2012;119:2245–2253.
9. Leske MC, Wu SY, Hennis A, Honkanen R, Nemesure B. Risk factors for incident open-angle glaucoma: the Barbados Eye Studies. *Ophthalmology*. 2008;115:85–93.
10. Mehdizadeh A, Hoseinzadeh A, Fazelzadeh A. Central corneal thickness as a risk factor for glaucoma. *Med Hypotheses*. 2007;69:1205–1207.
11. Ulmer M, Li J, Yaspan BL, et al. Genome-wide analysis of central corneal thickness in primary open-angle glaucoma cases in the NEIGHBOR and GLAUGEN consortia. *Invest Ophthalmol Vis Sci*. 2012;53:4468–4474.
12. Dimasi DP, Chen JY, Hewitt AW, et al. Novel quantitative trait loci for central corneal thickness identified by candidate gene analysis of osteogenesis imperfecta genes. *Hum Genet*. 2010;127:33–44.
13. Lu Y, Dimasi DP, Hysi PG, et al. Common genetic variants near the Brittle Cornea Syndrome locus ZNF469 influence the blinding disease risk factor central corneal thickness. *PLoS Genet*. 2010;6:e1000947.
14. Dimasi DP, Burdon KP, Craig JE. The genetics of central corneal thickness. *Br J Ophthalmol*. 2010;94:971–976.
15. Gros-Otero J, Arruabarrena-Sanchez C, Teus M. Central corneal thickness in a healthy Spanish population. *Arch Soc Esp Ophthalmol*. 2011;86:73–76.
16. Eysteinsson T, Jonasson F, Sasaki H, et al. Central corneal thickness, radius of the corneal curvature and intraocular pressure in normal subjects using non-contact techniques: Reykjavik Eye Study. *Acta Ophthalmol Scand*. 2002;80:11–15.
17. Wolfs RC, Klaver CC, Vingerling JR, Grobbee DE, Hofman A, de Jong PT. Distribution of central corneal thickness and its association with intraocular pressure: the Rotterdam Study. *Am J Ophthalmol*. 1997;123:767–772.
18. Nemesure B, Wu SY, Hennis A, Leske MC. Corneal thickness and intraocular pressure in the Barbados Eye Studies. *Arch Ophthalmol*. 2003;121:240–244.
19. Hahn S, Azen S, Ying-Lai M, Varma R. Central corneal thickness in Latinos. *Invest Ophthalmol Vis Sci*. 2003;44:1508–1512.
20. Torres RJ, Jones E, Edmunds B, Becker T, Gioffi GA, Mansberger SL. Central corneal thickness in Northwestern American Indians/Alaskan Natives and comparison with White and African-American persons. *Am J Ophthalmol*. 2008;146:747–751.
21. Zhang H, Xu L, Chen C, Jonas JB. Central corneal thickness in adult Chinese. Association with ocular and general parameters. The Beijing Eye Study. *Graefes Arch Clin Exp Ophthalmol*. 2008;46:587–592.
22. Suzuki S, Suzuki Y, Iwase A, Araie M. Corneal thickness in an ophthalmologically normal Japanese population. *Ophthalmology*. 2005;112:1327–1336.
23. Alsibirk PH. Corneal thickness. II. Environmental and genetic factors. *Acta Ophthalmol (Copenh)*. 1978;56:105–113.
24. Toh T, Liew SH, MacKinnon JR, et al. Central corneal thickness is highly heritable: the twin eye studies. *Invest Ophthalmol Vis Sci*. 2005;46:3718–3722.
25. Zheng Y, Ge J, Huang G, et al. Heritability of central corneal thickness in Chinese: the Guangzhou Twin Eye Study. *Invest Ophthalmol Vis Sci*. 2008;49:4303–4307.
26. Landers JA, Hewitt AW, Dimasi DP, et al. Heritability of central corneal thickness in nuclear families. *Invest Ophthalmol Vis Sci*. 2009;50:4087–4090.
27. Cornes BK, Khor CC, Nongpiur ME, et al. Identification of four novel variants that influence central corneal thickness in multi-ethnic Asian populations. *Hum Mol Genet*. 2012;21:437–445.
28. Vitart V, Bencic G, Hayward C, et al. New loci associated with central cornea thickness include COL5A1, AKAP13 and AVGR8. *Hum Mol Genet*. 2010;19:4304–4311.
29. Vithana EN, Aung T, Khor CC, et al. Collagen-related genes influence the glaucoma risk factor, central corneal thickness. *Hum Mol Genet*. 2011;20:649–658.
30. Hoehn R, Zeller T, Verhoeven VJ, et al. Population-based meta-analysis in Caucasians confirms association with COL5A1 and ZNF469 but not COL8A2 with central corneal thickness. *Hum Genet*. 2012;131:1783–1793.
31. Lu Y, Vitart V, Burdon KP, et al. Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nat Genet*. 2013;45:155–163.
32. Varma R, Paz SH, Azen SP, et al. The Los Angeles Latino Eye Study: design, methods, and baseline data. *Ophthalmology*. 2004;111:1121–1131.
33. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
34. Li Y, Ding J, Abecasis GR. Program No. 2290 for the 2006 ASHG Annual Meeting. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am J Hum Genet*. 2006;S79.
35. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol*. 2010;34:816–834.
36. Gao X, Haritunians T, Marjoram P, et al. Genotype imputation for Latinos using the HapMap and 1000 Genomes Project reference panels. *Front Genet*. 2012;3:117.
37. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet*. 2006;2:e190.
38. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010;467:52–58.
39. Mao X, Bigham AW, Mei R, et al. A genomewide admixture mapping panel for Hispanic/Latino populations. *Am J Hum Genet*. 2007;80:1171–1178.
40. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55:997–1004.
41. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–2191.
42. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol*. 2008;32:361–369.
43. Gao X. Multiple testing corrections for imputed SNPs. *Genet Epidemiol*. 2011;35:154–158.
44. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2009.
45. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Totowa, NJ: Humana Press; 2000:365–386.
46. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40:e115.
47. Hindorf L, Junkins H, Hall P, Mehta J, Manolio T. *A Catalog of Published Genome-Wide Association Studies*. Available at: <http://www.genome.gov/gwastudies/>. Accessed June 1, 2011.
48. Haritunians T, Jones MR, McGovern DP, et al. Variants in ZNF365 isoform D are associated with Crohn's disease. *Gut*. 2011;60:1060–1067.

49. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707-713.
50. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466:714-719.
51. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57-74.
52. Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. *Stat Med*. 2002;21:35-50.
53. Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. *Am J Epidemiol*. 2002;155:478-484.
54. Gauderman WJ. Candidate gene association analysis for a quantitative trait, using parent-offspring trios. *Genet Epidemiol*. 2003;25:327-338.