Investigation of variants identified in Caucasian genome-wide association studies for plasma HDL cholesterol and triglycerides levels in Mexican dyslipidemic study samples

Daphna Weissglas-Volkov, PhD, **Carlos A. Aguilar-Salinas, MD**, **Janet S. Sinsheimer, PhD**, **Laura Riba, PhD**, **Adriana Huertas-Vazquez, PhD**, **Maria L. Ordoñez-Sánchez, B.Sc**, **Rosario Rodriguez-Guillen, M.S**, **Rita M. Cantor, PhD**, **Teresa Tusie-Luna, MD, PhD**, and **Päivi Pajukanta, MD, PhD**

Department of Human Genetics, David Geffen School of Medicine at UCLA (D.W.-V., J.S.S, A.H.-V., R.M.C., P.P), Los Angeles, CA; Department of Endocrinology and Metabolism, (C.A.A.-S.), Department of Molecular Biology and Genomic Medicine Unit (L.R., M.L.O.-S., R. R.-G. T.T.-L.), Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubiran, Mexico City, Mexico; Instituto de Investigaciones Biomédicas de la UNAM (L.R., M.L.O.-S., R. R.-G. T.T.-L.), Salvador Zubiran, Mexico City, Mexico; Heart Institute, Cedars-Sinai Medical Center (A.H.-V.), Los Angeles, California

Abstract

Background—Although epidemiological studies have demonstrated an increased predisposition to low HDL cholesterol (HDL-C) and high triglyceride (TG) levels in the Mexican population, Mexicans have not been included in any of the previously reported genome-wide association studies (GWAS) for lipids.

Results—We investigated six SNPs associated with TGs, seven with HDL-C and one with both TGs and HDL-C in recent Caucasian GWAS in Mexican familial combined hyperlipidemia families and hypertriglyceridemia case-control study samples. These variants were within or near the genes ABCA1, ANGPTL3, APOA5, APOB, CETP, GALNT2, GCKR, LCAT, LIPC, LPL (2), MMAB-MVK, TRIB1 and XKR6-AMAC1L2. We performed a combined analysis of the family-based and case-control studies (n=2,298) using the Z-method to combine statistics. Ten of the SNPs were nominally significant and five were significant after Bonferroni correction (P = 2.20×10^{-3} – 2.6×10^{-11}) for the number of tests performed (APOA5, CETP, GCKR and GALNT2). Interestingly, our strongest signal was obtained for TGs with the minor allele of rs964184 (P=2.6×10^{-11}) in the APOA1/C3/A4/A5 gene cluster region that is significantly more common in Mexicans (27%) than in Caucasians (12%).

Conclusions—It is important to confirm whether known loci have a consistent effect across ethnic groups. We show replication of five Caucasian GWAS lipid associations in Mexicans. The remaining loci will require a comprehensive investigation to exclude or verify their significance in Mexicans. We also demonstrate that rs964184 has a large effect (OR=1.74) and is more frequent in the Mexican population, and thus it may contribute to the high predisposition to dyslipidemias in Mexicans.

Correspondence to Päivi Pajukanta, MD, PhD, Associate Professor, Department of Human Genetics, David Geffen School of Medicine at UCLA, Gonda Center, Room 6335B, 695 Charles E. Young Drive South, Los Angeles, California 90095-7088, USA. Phone: 1-310-267-2011, Fax: 1-310-794-5446, ppajukanta@mednet.ucla.edu.

Disclosures: No conflicts to disclose.

Keywords

Lipids; High-density lipoprotein cholesterol; triglycerides; genome-wide association studies; replication; diverse populations; Mexicans; Single-nucleotide polymorphism

Introduction

Unfavorable serum lipid levels are well-established risk factors for coronary artery disease (CAD [MIM 607339])[1] and highly common among Mexicans.[2] As several epidemiological studies have demonstrated that the Mexican population has a high predisposition to low HDL cholesterol (HDL-C) and high serum triglyceride (TG) levels,[3,4] investigation of the genetic factors conferring to these common forms of dyslipidemias in Mexicans is important. Unfortunately, this population has been underinvestigated in genetic studies and to date no genome-wide association studies (GWAS) for lipids have been performed in Mexicans.

Recent GWAS examining the concentrations of HDL-C and TGs in populations of European descent[5-7] identified SNPs at 15 loci as associated with HDL-C levels (ABCA1, ANGPTL4, APOA1/C3/A4/A5 gene cluster, CETP, FADS1/2/3, GALNT2, HNF4A, LCAT, LIPC, LIPG, LPL, MADD-FOLH1-NR1H3, MMAB-MVK, PLTP and TTC39B) and 12 loci with TGs (ANGPTL3, APOA1/C3/A4/A5 gene cluster, APOB, FADS1/2/3, GCKR, LPL, MLXIPL, NCAN-CILP2-PBX4, PLTP, APOE, TRIB1, XKR6-AMAC1L2). However, whether these variants also confer risk in populations with a different demographic history, such as Mexicans, remains unexplored. Thus we examine whether these known risk variants are involved in the increased susceptibility to dyslipidaemia in Mexicans.

Since only a small fraction of the common variants (>10 million) are directly or indirectly (imputed) genotyped in current GWAS, these lipid-associated SNPs might not represent the actual functional variants but rather be in linkage-disequilibrium (LD) with the causal variants. As patterns of LD vary between populations,[8] studies in Mexicans can also assist in fine-mapping the actual susceptibility variant(s). Furthermore, it is necessary to establish whether confirmed loci have a consistent effect across ethnic groups if they are to be used in cardiovascular risk assessment.[9]

In the current study, we investigated common variants in Mexicans (MAF > 10%) within or near the genes ABCA1, ANGPTL3, APOA1/C3/A4/A5 gene cluster, APOB, CETP, GALNT2, GCKR, LCAT, LIPC, LPL, MMAB-MVK, TRIB1, and XKR6-AMAC1L2. Seven of these genes were previously known to be involved in lipid metabolism: ATP-binding cassette A1 (ABCA1), necessary for the efflux of cholesterol from cells to HDL particles; apolipoproteins (APOA1/C3/A4/A5 and B), the structural components and ligands of lipoprotein particles; lecithin-cholesterol acyltransferase (LCAT) that catalyzes the formation of cholesteryl esters in HDL particles; cholesterylester transfer protein (CETP) that transfers cholesteryl esters from HDL particles to apo-B containing lipoproteins in exchange for TGs; and the lipases, hepatic lipase (LIPC) and lipoprotein lipase (LPL), that hydrolyze TGs on lipoproteins.[10]

Of the remaining six genes, the function of TRIB1 and XKR6-AMAC1L2 is poorly understood. The angiopoietin-like protein 3 (ANGPTL3) is a secreted protein that inhibits LPL activity posttranscriptionaly.[11] GALNT2 encodes the polypeptide N-acetylgalactosaminyltransferase 2, involved in the first step of O-linked glycosylation of proteins. GALNT2 may thus affect HDL-C through the glycosylation of proteins involved in lipid metabolism. For instance, LCAT, apoC-III, VLDL and LDL receptors are all O-glycosylated.[12] The glucokinase regulatory protein (GCKR) regulates glucokinase (GCK), a key regulator of glucose storage and disposal in the liver.[13] MVK and MMAB are both regulated by the sterol-responsive

element-binding protein 2 (SREBP2) through a shared common promoter.[12] SREBP2 is a transcription factor that controls cholesterol homeostasis. MVK encodes the mevalonate kinase, which catalyzes an early step in the biosynthesis of cholesterol.[12] Methylmalonic aciduria (MMAB) encodes an enzyme involved in the formation of adenosylcobalamin, necessary for degradation of cholesterol.[12] However, the exact mechanisms by which these novel genes influence HDL and/or TG metabolism remain to be clarified in future studies.

We investigated these GWAS loci in two independent Mexican dyslipidemic study samples comprised of a hypertriglyceridemia case-control sample and families with familial combined hyperlipidemia (FCHL [MIM 144250]). FCHL is a common genetic dyslipidemia affecting 1%-6% of the general population, and 10%-20% of subjects with premature CAD.[14,15] FCHL is primarily characterized by elevated levels of serum TGs, total cholesterol (TC) or both, [16-18] with additional component traits including low HDL-C levels[19]. As all GWAS for lipids to date examined study samples that were not ascertained for dyslipidemia, this is also the first study to investigate these variants in dyslipidemic study samples at increased risk to develop CAD.

Methods

Study Participants

The study design was approved by the ethics committees of the participating centers. All subjects provided a written informed consent. Clinical characteristics of the study samples are shown in Supplementary table 1.

A total of 73 extended Mexican FCHL families ($n = 877$) were included in this study. The families were recruited at the Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ) in Mexico City, as described previously.[20] None of the family members were using lipid lowering medication when the blood sample was taken. The 1,421 hypertriglyceridemic cases ($n = 727$) and controls ($n = 694$) were also recruited at the INCMNSZ in Mexico City. The inclusion criteria were fasting serum TGs > 2.3 mmol/L (200 mg/dL) for the cases and < 1.7 mmol/L (150 mg/dL) for the controls.[21] Exclusion criteria were T2DM or morbid obesity (BMI > 40 kg/m^2), and the use of lipid lowering drugs for the controls. The case-control samples were also classified as low HDL-C if serum HDL-C levels were < 1.04 mmol/L (40 mg/dL) ($n = 480$) and normal HDL-C if serum HDL-C levels were ≥ 1.3 mmol/L (50 mg/dL) ($n = 492$).[21] Measurements of fasting TG and HDL-C levels were performed with commercially available standardized methods in both the families and case-control study samples.[20]

SNP selection and Genotyping

Sixteen SNPs surpassing genome-wide significant threshold[22] ($P < 5 \times 10^{-8}$) for either HDL-C or TGs in Caucasian GWAS[5-7] and with minor allele frequency (MAF) ≥ 0.1 based on the HapMap[23] Mexican-American data were genotyped using the SNPlex genotyping platform (Applied Biosystems). These variants were within or near the genes ABCA1, ANGPTL3, ANGPTL4, APOA5, APOB, CETP, GALNT2, GCKR, LCAT, LIPC, LPL (2), MMAB-MVK, PLTP, TRIB1 and XKR6-AMAC1L2 (Supplementary table 2). Two of the SNPs, rs2967605 near ANGPTL4 and rs7679 near PLTP, had less than 90% genotype call rate and were excluded from subsequent analyses. All other SNPs had at least 90% genotype call rate and were in Hardy-Weinberg equilibrium (P > 0.05) in the normotriglyceridemic controls, as well as in 171 unrelated FCHL family members. Only 1.35 Mendelian errors per SNP were found in 451 genotyped parent-offspring pairs (overall error rate = 0.003) with the PEDSTATS program[24] and these were excluded from subsequent analyses.

Statistical analyses

The case-control subjects were analyzed by logistic regression analysis for the additive model, as implemented in the PLINK v1.06 software.[25] In the Mexican FCHL sample, quantitative lipid levels were used in order to utilize data from all available family members. Association analysis was performed utilizing the quantitative transmission disequilibrium test (QTDT) implemented in the genetic analysis package SOLAR[26,27] using an additive model with age and sex as covariates. TG values were log transformed to approach a normal distribution. Furthermore, the t-distribution rather than the normal distribution option of SOLAR was used to allow for robust estimation of the mean and variance even if the trait distribution deviates from normality.[28]

We did not employ a joint analysis with pooled genotype data of the families and case-control samples because different ascertainment procedures were used to collect these study samples. Accordingly, the case-control sample was analyzed using affection status whereas in the families the entire lipid distribution was utilized. We then performed a combined analysis of the family-based and case-control studies ($n = 2,298$) using the Z-method to combine statistics. [29,30] The Z-statistics were summed and weighted by the square-root of the proportion of individuals examined in each study.[30]

Genotype data for the GWAS associated SNPs were extracted from the International HapMap Project[23] (http://www.hapmap.org) Phase 3 dataset for CEU (Northern and Western European ancestry from the CEPH collection) and MEX (Mexican ancestry from Los Angeles, CA). Only genotypes of founders were used for the estimation of allele frequencies, and accordingly 112 founders of European ancestry and 50 founders of Mexican ancestry were included in the analyses. Differences in allele frequencies between the Mexican-American and CEU founders of HapMap were determined by a chi-square test. We also compared the allele frequencies of the 112 CEU founders of HapMap to 150 of the Mexican controls from Mexico city (TGs < 150 mg/dL and HDL-C > 50 mg/dL) using a chi-square test.

Genotype data in the surrounding regions of the GWAS associated SNPs (±100 kb) were extracted from HapMap in similar manner. Only SNPs having frequencies in both populations were included (130 SNPs per region on average), and only genotypes of founders were used in the LD analysis. LD was measured in r^2 using the "matrix" option of the PLINK v1.05 software.

We estimated the proportion of cases and controls exceeding the clinical thresholds[21] for low HDL-C (< 40 mg/dl) and high TGs (> 200 mg/dl) as a function of the allelic dosage score for HDL-C and TGs, respectively.[5] Allelic dosage scores were calculated for SNPs associated with P-value ≤ 0.01 (Table 1 and Supplementary table 2) by weighting the counts of the risk allele (0, 1, or 2) by the beta-coefficient of the logistic-regression model (log odds-ratio) and summing across all SNPs.[5] We performed a test of trend across the genotype scores using the "prop.trend.test" function in R (R.2.8.0).

Results

We tested whether 15 SNPs surpassing genome-wide significant threshold[22] for either HDL-C or TGs in Caucasian GWAS[5-7] are associated in Mexicans. Each variant was tested for association with HDL-C or TGs based on the Caucasian GWAS genome-wide significant results. Accordingly, 8 SNPs were analyzed for HDL-C and 7 for TGs (rs964184 near the APOA5 was analyzed with both HDL-C and TGs). We considered the attempts to replicate variants identified for HDL-C and TGs as separate experiments, and therefore we adjusted the HDL-C association results for 8 independent tests and TGs for 7 independent tests based on the number of variants analyzed for each trait.

Three SNPs were associated with HDL-C after Bonferroni correction for 8 independent tests (Bonferroni significant level P = 6.25×10^{-3}): rs3764261 residing 2.5kb upstream of CETP (P = 2.35×10^{-8}), rs964184 11kb upstream of APOA5 (P = 1.06×10^{-3}), and rs4846914 in the first intron of GALNT2 (P = 2.20×10^{-3}). Two SNPs were significant with TGs after Bonferroni correction for 7 independent tests (Bonferroni significant level P = 7.14×10^{-3}): rs964184 residing near APOA5 (P = 2.60×10^{-11}), and rs1260326 a missense variant in GCKR (P = 1.26×10^{-3}). These signals were all for the same risk allele as in Caucasians (Table 1).

Evaluation of global ancestry with the GWAS variants and lipid traits

Mexicans are an admixed population, descended from a recent mix of Amerindian and European ancestry with a small proportion of African ancestry.[31,32] Population admixture may confound allelic association if both the trait distribution and the allele frequency differ between ancestries. Family-based association has the advantage that it is robust to spurious associations due to population admixture[33]. However, to minimize the possibility of spurious associations due to admixture in the case-control study sample, we evaluated whether the trait distributions and/or the effect of the GWAS variants differs with ancestry estimates in 588 Mexican family members for which we had individual ancestry (IA) estimates available (Supplementary methods). As both the FCHL families and the hypertriglyceridemic cases and controls were recruited by the same dyslipidemia clinic, INCMNSZ, in Mexico City we anticipated that the proportion of ancestry admixture in the families would also be representative of the cases and controls

To investigate the relationship between the GWAS variants and ancestry, we examined the effect of genotype × ancestry interaction on the traits by inclusion of the genotypes, coded as 0-2 copies of the minor allele, IAs and their interaction term in the model (Supplementary methods).[34] We did not observe a significant interaction between the genotype and IA on HDL-C levels with any of the eight HDL-C SNPs (P > 0.06) or on TGs with any of the seven TG SNPs (P > 0.09) (Supplementary table 2 and Supplementary figure 1). The non-significant interaction results suggest that the strength of association is not different between the ancestry backgrounds, and hence ethnic admixture should not influence the strength of the association with these SNPs even when using a nonfamily-based robust approach. Similarly, neither HDL-C nor TG levels were significantly associated with ancestry in the Mexican families (P > 0.6) (Supplementary figure 1). Furthermore, to ensure that FCHL is not significantly associated with ancestry, we also examined the relation between the IA proportions and TC levels, which was also used to define the FCHL families. As for TGs, TC levels were also not significantly associated with IA (P = 0.4) (Supplementary figure 1). Taken together, these data show that the associations and the distribution of lipid traits are not influenced by ancestry in Mexican families, suggesting that admixed ancestry should not influence our association analyses in the unrelated study sample.

Evaluation of the GWAS allele frequencies in Europeans and Mexicans

The contribution of a variant to disease susceptibility in a population depends on its frequency. Hence, we evaluated whether the allele frequencies of the variants differ between Europeans and Mexicans by comparing the genotype data of 112 CEU founders with European ancestry from the HapMap project[23] to150 of the Mexican controls (TGs < 150 mg/dL and HDL-C > 50 mg/dL) (Table 2). Interestingly, our strongest association signal (P = 2.60×10^{-11}) (Table 1) was obtained with TGs for the minor allele of rs964184 near the APOA1/C3/A4/A5 gene cluster that is significantly more common in Mexicans (27%) than in Europeans (12%) (P = 5.89×10^{-5}) (Table 2). Similarly, the common allele of rs10468017 near LIPC that was nominally significant with HDL-C (P = 9.19×10^{-3}) (Supplementary table 2), is also more prevalent in Mexicans (81%) than in Europeans (69%) (P = 1.74×10^{-3}) (Table 2). Conversely, the frequency of the risk allele of rs7819412 near XKR6-AMAC1L2 that did not replicate in

this study was significantly less frequent in Mexicans than Europeans (27% versus 51%, P = 1.04×10^{-8}) (Table 2).

Since the allele frequencies of the Mexican controls might not represent the frequencies in the general population, as is the case for the HapMap samples, we also performed this analysis by comparing the genotype data of the European CEU founders to 50 Mexican-American founders from the HapMap project that identified themselves as having grandparents who were born in Mexico. The results of the two analyses were in good agreement (Table 2) except for rs1883025 near ABCA1 that is significantly more frequent in the HapMap Mexican-American sample when compared to the European CEU founders but not with the Mexican controls from Mexico City (Table 2).

Comparison of LD in the surrounding regions of the GWAS variants between the European and the Mexican populations

The function of most variants identified in GWAS is unknown, and most of these variants reside outside coding regions.[35] Hence these SNPs might not represent the actual susceptibility variants but rather be in LD (i.e tag) with the functional variants. Studies in Mexicans also offer the opportunity to examine the surrounding LD of the associated alleles and assist in fine-mapping of the actual susceptibility variant, as LD patterns in the area of the associated variants is likely to be different in diverse populations. We therefore evaluated whether the LD structures surrounding the Caucasians GWAS variants differ in the Mexican population. We examined the pairwise LD of each variant with its neighboring SNPs (±100 kb) in the Mexican and European populations using the HapMap data. For each variant we compared the concordance of SNPs they tagged with $r^2 \geq 0.7$ in Mexicans relative to Europeans. We observed differences in tagging for five of the GWAS variants in Mexicans (rs10096633, rs2083637, rs7819412, rs4846914 and rs1883025) as these five SNPs tagged less SNPs in Mexicans than in Caucasians (Supplementary figure 2), suggesting that Mexicans may be informative in fine-mapping studies of these GWAS loci (LPL(2), XKR6-AMAC1L2, GALNT2 and ABCA1). However, overall the percentage of agreement was high (78 % on average), with the least agreement obtained for rs10096633 downstream of LPL (20%) (Supplementary figure 2). Interestingly, also the association signals (10 tests with P < 0.05) were all for the same risk allele as in Caucasians[5-7] except for rs10096633 that was nominally significant with TGs (P = 8.08×10^{-3}) for the opposite allele as in Caucasians (Supplementary table 2). Hence this observed difference in the pairwise LD may explain the allelic discrepancy of rs10096633 between Mexicans and Europeans.

The proportion of individuals with low HDL-C or high TGs as a function of an allelic dosage score

Since plasma HDL-C and TGs levels are polygenic traits, we also examined the cumulative effect of the associated risk alleles ('allelic dosage') on the proportion of individuals exceeding clinical thresholds for low HDL-C (< 40 mg/dl) and high TG (> 200 mg/dl).[21] We calculated allelic dosage scores in the hypertriglyceridemic cases and controls for SNPs associated with P-value ≤ 0.01 (Table 1 and Supplementary table 2) Accordingly, SNPs near CETP, APOA5, GALNT2 and LIPC were included in the low HDL-C analysis and SNPs near APOA5, GCKR and LPL in the high TG analysis. The proportion of individuals exceeding clinical thresholds for low HDL-C and high TG levels increased with the genotype score (figure 1). To evaluate the incline in proportions we also performed a test of trend across the genotype scores. The trends were statistically significant for both the low HDL-C (P = 1.89×10^{-8}) and the high TG proportions (P = 5.07×10^{-13}).

Discussion

Hispanic populations, such as Mexicans, have been mainly ignored in large scale genetic studies for cardiovascular risk factors as well as other common diseases, although the Mexican population has an increased predisposition to several forms of dyslipidemias (low HDL-C and high TGs) and premature CAD.[2-4] The current study evaluated whether 15 associations identified in recent Caucasian GWAS[5-7] with HDL-C or TGs also confer risk in the Mexican population. Association at nominal level ($P \leq 0.05$) was observed for 10 of the loci, of which half were also significant after adjusting for multiple comparisons. Importantly, we also demonstrate that two associated SNPs are more prevalent in the Mexican population, and thus may confer to the increased predisposition to these forms of dyslipidemia in Mexicans.

Mexicans descend from a recent mix of Amerindian and European ancestry which led to marked differences in allelic frequencies and patterns of LD across their genome.[31,32] Our strongest association signal ($P = 2.60 \times 10^{-11}$) was obtained for the minor allele of rs964184 near the APOA1/C3/A4/A5 cluster that is significantly more prevalent in Mexicans (27%) than in Caucasians (12%). The strength of this association even surpassed the stringent criteria for declaring genome-wide significance[22] ($P < 5 \times 10^{-8}$) with only 2,298 individuals studied. This finding clearly demonstrates that even if a variant has an effect in all ancestry groups, it might be more prevalent in one population. Furthermore, this finding suggests that depending on its frequency, the contribution of each locus to disease susceptibility may vary between populations, and hence the implications of SNPs to public health can differ between populations.

The association signals ($P \leq 0.05$) were all for the same risk allele as in Caucasians[5-7] except for rs10096633, downstream of the LPL gene, that was significant for the opposite allele as in Caucasians.[6] Although the allele frequency of rs10096633 was not significantly different between Mexicans and Europeans ($P = 0.8$), we show that the LD structure surrounding rs10096633 varies between the two populations. This finding suggests that rs10096633 is not the causal variant and that given the altered LD, future studies in Mexicans could assist in fine-mapping the actual functional variant in this locus. Conversely, SNP rs964184 near the APOA1/C3/A4/A5 gene cluster may be the actual susceptibility variant, as in both populations this SNP is a singleton (all pairwise $r^2 < 0.7$) and strongly associated with TGs despite the large difference in allele frequency. Nevertheless, deep resequencing in the rs964184 region is needed in order to conclude that this SNP is the sole signal.

Although the genes of the APOA1/C3/A4/A5 cluster (≥ 11 kb away) are the most plausible candidates in the region of rs964184, this variant resides in the 5′ UTR of a zinc finger protein (ZPR1) that may be involved in cell proliferation and signal transduction.[36] Hence, functional studies are warranted to establish the actual causal gene and the consequences of this variant. Regarding the other significant (i.e. Bonferroni correction) variants, SNP rs4846914 resides in the first intron of GALNT2. There are several SNPs redundant to rs4846914 ($r^2 \geq 0.7$) all located within the first intron. First introns are known to harbor many regulatory elements.[37] Accordingly, all of these variants in LD should be tested for potential regulatory effect. Similarly, rs3764261 resides 2.5kb upstream of CETP. This SNP is in high LD with variants located 1-8 kb from the transcriptional start site of CETP. Thus, these variants and their haplotypes should also be investigated in functional studies. On the other hand, SNP rs1260326 is a missense variant (P446L) in GCKR, and recent in-vitro experiments have shown that this amino-acid substitution affects glucose and TG levels through increased GCK activity in the liver.[13]

To screen for association signals, we also used family-based association method that is robust to population admixture,[33] such as in Mexicans.[31,32] Furthermore, we examined whether the

associations were influenced by global ancestry using European/Amerindian informative markers. To reduce sources of heterogeneity, both the dyslipidemic families and case/control study sample were recruited by the same dyslipidemia clinic. We recognize that differences in ascertainment criteria may have potentially caused heterogeneity. For instance the common allele of rs2083637 near the LPL gene was associated with HDL-C ($P = 3.85 \times 10^{-3}$) in the case-control samples, as in Caucasians,[7] while the opposite allele was implicated in the FCHL families ($P = 2.65 \times 10^{-2}$), suggesting that a different variant(s) in LPL may be involved in FCHL.[38] Nevertheless, we feel that a more reliable conformation was obtained by utilizing a variety of resources that includes dyslipidemic families with multiple affected individuals and unrelated hypertriglyceridemic case/control individuals.

As our sample size was not sufficiently powered for rare variants and small genetic effects (OR <1.3) we have not attempted to investigate confirmed associations with MAF less than 10% in the Mexican-American data of HapMap. Hence, given our relatively small sample size the remaining loci and those that did not replicate in this study will require comprehensive investigation using larger samples to exclude or verify their significance in Mexicans. Future GWAS in Mexicans should further assist in fine-mapping the causal variant(s) in these known loci as well as reveal novel susceptibility alleles.

In summary, this is the first study to present association signals of susceptibility variants from Caucasian GWAS for lipids in Mexicans. Furthermore, thus far all GWAS for lipids have examined study samples that were not ascertained for dyslipidemia. Hence, to the best of our knowledge this is also the first study to report the effect of these variants in dyslipidemic individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson T, Flegal K, Ford E, Furie K, Go A, Greenlund K, Haase N, Hailpern S, Ho M, Howard V, Kissela B, Kittner S, Lackland D, Lisabeth L, Marelli A, McDermott M, Meigs J, Mozaffarian D, Nichol G, O'Donnell C, Roger V, Rosamond W, Sacco R, Sorlie P, Stafford R, Steinberger J, Thom T, Wasserthiel-Smoller S, Wong N, Wylie-Rosett J, Hong Y. Heart disease and stroke statistics -- 2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 2009;119:480–6. [PubMed: 19171871]

2. Stevens G, Dias R, Thomas K, Rivera J, Carvalho N, Barquera S, Hill K, Ezzati M. Characterizing the epidemiological transition in Mexico: national and subnational burden of diseases, injuries, and risk factors. PLoS Med 2008;5:e125. [PubMed: 18563960]

3. Aguilar-Salinas C, Olaiz G, Valles V, Torres J, Gomez Perez F, Rull J, Rojas R, Franco A, Sepulveda J. High prevalence of low HDL cholesterol concentrations and mixed hyperlipidemia in a Mexican nationwide survey. J Lipid Res 2001;42:1298–307. [PubMed: 11483632]

4. Aguilar-Salinas CA, Rojas R, Gómez-Pérez FJ, Valles V, Ríos-Torres JM, Franco A, Olaiz G, Rull JA, Sepúlveda J. High prevalence of metabolic syndrome in Mexico. Archives of Medical Research 2004;35:76–81. [PubMed: 15036804]

5. Kathiresan S, Willer C, Peloso G, Demissie S, Musunuru K, Schadt E, Kaplan L, Bennett D, Li Y, Tanaka T, Voight B, Bonnycastle L, Jackson A, Crawford G, Surti A, Guiducci C, Burtt N, Parish S, Clarke R, Zelenika D, Kubalanza K, Morken M, Scott L, Stringham H, Galan P, Swift A, Kuusisto J, Bergman R, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta K, Dupuis J, de Bakker P, O'Donnell C, Chambers J, Kooner J, Hercberg S, Meneton P, Lakatta E, Scuteri A, Schlessinger D, Tuomilehto J, Collins F, Groop L, Altshuler D, Collins R, Lathrop G, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas J, Boehnke M, Abecasis G, Mohlke K, Cupples L. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet 2009;41:56–65. [PubMed: 19060906]

6. Sabatti C, Service SK, Hartikainen A, Pouta A, Ripatti S, Brodsky J, Jones CG, Zaitlen NA, Varilo T, Kaakinen M, Sovio U, Ruokonen A, Laitinen J, Jakkula E, Coin L, Hoggart C, Collins A, Turunen H, Gabriel S, Elliot P, McCarthy MI, Daly MJ, Järvelin M, Freimer NB, Peltonen L. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nature Genetics 2009;41:35–46. [PubMed: 19060910]

7. Aulchenko Y, Ripatti S, Lindqvist I, Boomsma D, Heid I, Pramstaller P, Penninx B, Janssens A, Wilson J, Spector T, Martin N, Pedersen N, Kyvik K, Kaprio J, Hofman A, Freimer N, Jarvelin M, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks A, Falchi M, Willemsen G, Hottenga J, de Geus E, Montgomery G, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands E, Uitterlinden A, Witteman J, Oostra B, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Doring A, Wichmann H, Smit J, McCarthy M, van Duijn C, Peltonen L. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat Genet 2009;41:47–55. [PubMed: 19060911]

8. A haplotype map of the human genome. Nature 2005;437:1299–320. [PubMed: 16255080]

9. Ioannidis JP. Prediction of Cardiovascular Disease Outcomes and Established Cardiovascular Risk Factors by Genome-Wide Association Markers. Circ Cardiovasc Genet 2009;2:7–15. [PubMed: 20031562]

10. Rader DJ. Molecular regulation of HDL metabolism and function: implications for novel therapies. J Clin Invest 2006;116:3090–3100. [PubMed: 17143322]

11. Romeo S, Yin W, Kozlitina J, Pennacchio LA, Boerwinkle E, Hobbs HH, Cohen JC. Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. J Clin Invest 2009;119:70–9. [PubMed: 19075393]

12. Holleboom AG, Vergeer M, Hovingh GK, Kastelein JJ, Kuivenhoven JA. The value of HDL genetics. Current Opinion in Lipidology 2008;19:385–394. [PubMed: 18607186]

13. Beer NL, Tribble ND, McCulloch LJ, Roos C, Johnson PR, Orho-Melander M, Gloyn AL. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. Hum Mol Genet 2009;18:4081–4088. [PubMed: 19643913]

14. Grundy S, Greenland P, Herd A, Huebsch J, Jones R, Mitchell J, Schlant R. Cardiovascular and risk factor evaluation of healthy American adults. A statement for physicians by an Ad Hoc Committee appointed by the Steering Committee, American Heart Association. Circulation 1987;75:1340A–1362A.

15. Hopkins P, Heiss G, Ellison R, Province M, Pankow J, Eckfeldt J, Hunt S. Coronary artery disease risk in familial combined hyperlipidemia and familial hypertriglyceridemia: a case-control comparison from the National Heart, Lung, and Blood Institute Family Heart Study. Circulation 2003;108:519–23. [PubMed: 12847072]

16. Goldstein J, Hazzard W, Schrott H, Bierman E, Motulsky A. Genetics of hyperlipidemia in coronary heart disease. Trans Assoc Am Physicians 1972;85:120–38. [PubMed: 4349620]

17. Genest J, Martin-Munley S, McNamara J, Ordovas J, Jenner J, Myers R, Silberman S, Wilson P, Salem D, Schaefer E. Familial lipoprotein disorders in patients with premature coronary artery disease. Circulation 1992;85:2025–33. [PubMed: 1534286]

18. Nikkila E, Aro A. Family study of serum lipids and lipoproteins in coronary heart-disease. Lancet 1973;1:954–9. [PubMed: 4121585]

19. Hunt S, Wu L, Hopkins P, Stults B, Kuida H, Ramirez M, Lalouel J, Williams R. Apolipoprotein, low density lipoprotein subfraction, and insulin associations with familial combined hyperlipidemia.

Study of Utah patients with familial dyslipidemic hypertension. Arteriosclerosis 1989;9:335–44. [PubMed: 2497719]

20. Weissglas-Volkov D, Huertas-Vazquez A, Suviolahti E, Lee J, Plaisier C, Canizales-Quinteros S, Tusie-Luna T, Aguilar-Salinas C, Taskinen M, Pajukanta P. Common hepatic nuclear factor-4alpha variants are associated with high serum lipid levels and the metabolic syndrome. Diabetes 2006;55:1970–7. [PubMed: 16804065]

21. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). Jama 2001;285:2486–97. [PubMed: 11368702]

22. Pe'er I, Yelensky R, Altshuler D, Daly M. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 2008;32:381–5. [PubMed: 18348202]

23. Frazer K, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61. [PubMed: 17943122]

24. Wigginton J. Abecasis GPEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. Bioinformatics 2005;21:3445–7. [PubMed: 15947021]

25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, Maller J, Sklar P, de Bakker P, Daly M. Sham PPLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75. [PubMed: 17701901]

26. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 1998;62:1198–211. [PubMed: 9545414]

27. Havill L, Dyer T, Richardson D, Mahaney M, Blangero J. The quantitative trait linkage disequilibrium test: a more powerful alternative to the quantitative transmission disequilibrium test for use in the absence of population stratification. BMC Genet 2005;6:S91. [PubMed: 16451707]

28. Almasy L, Gur R, Haack K, Cole S, Calkins M, Peralta J, Hare E, Prasad K, Pogue-Geile M, Nimgaonkar V, Gur R. A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. Am J Psychiatry 2008;165:1185–92. [PubMed: 18628350]

29. Whitlock MC. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. Journal of Evolutionary Biology 2005;18:1368–1373. [PubMed: 16135132]

30. Skol A, Scott L, Abecasis G, Boehnke M. Optimal designs for two-stage genome-wide association studies. Genet Epidemiol 2007;31:776–88. [PubMed: 17549752]

31. Price A, Patterson N, Yu F, Cox D, Waliszewska A, McDonald G, Tandon A, Schirmer C, Neubauer J, Bedoya G, Duque C, Villegas A, Bortolini M, Salzano F, Gallo C, Mazzotti G, Tello-Ruiz M, Riba L, Aguilar-Salinas C, Canizales-Quinteros S, Menjivar M, Klitz W, Henderson B, Haiman C, Winkler C, Tusie-Luna T, Ruiz-Linares A, Reich D. A genomewide admixture map for Latino populations. Am J Hum Genet 2007;80:1024–36. [PubMed: 17503322]

32. Salari K, Burchard E. Latino populations: a unique opportunity for epidemiological research of asthma. Paediatr Perinat Epidemiol 2007;21:15–22. [PubMed: 17935571]

33. Abecasis G, Cardon L, Cookson W. A general test of association for quantitative traits in nuclear families. Am J Hum Genet 2000;66:279–92. [PubMed: 10631157]

34. Deo RC, Reich D, Tandon A, Akylbekova E, Patterson N, Waliszewska A, Kathiresan S, Sarpong D, Taylor HA, Wilson JG. Genetic differences between the determinants of lipid profile phenotypes in African and European Americans: the Jackson Heart Study. PLoS Genetics 2009;5:e1000342. [PubMed: 19148283]

35. Manolio TA, Brooks LD, Collins FSA. HapMap harvest of insights into the genetics of common disease. J Clin Invest 2008;118:1590–1605. [PubMed: 18451988]

36. Galcheva-Gargova Z, Gangwani L, Konstantinov KN, Mikrut M, Theroux SJ, Enoch T, Davis RJ. The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells. Molecular Biology of the Cell 1998;9:2963–2971. [PubMed: 9763455]

37. Kalari KR, Casavant M, Bair TB, Keen HL, Comeron JM, Casavant TL, Scheetz TE. First exons and introns--a survey of GC content and gene structure in the human genome. In Silico Biology 2006;6:237–242. [PubMed: 16922687]

38. Nevin D, Brunzell J, Deeb S. The LPL gene in individuals with familial combined hyperlipidemia and decreased LPL activity. Arterioscler Thromb 1994;14:869–73. [PubMed: 8199176]
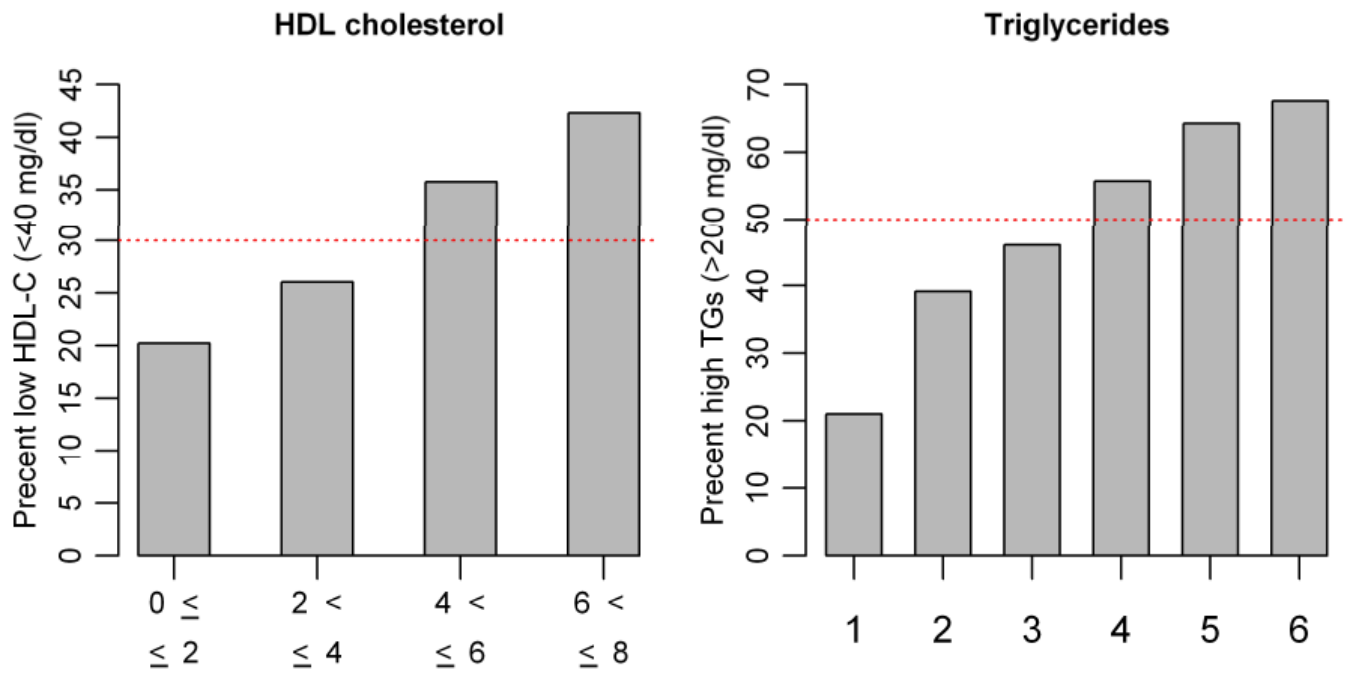
Figure 1.
The proportion of cases and controls exceeding the clinical thresholds based on the NCEP: ATPIII criteria[21] for low HDL-C (< 40 mg/dl) and high TGs (> 200 mg/dl) as a function of the allelic dosage score for HDL-C and TGs, respectively.[5] Allelic dosage scores are plotted on the x axis, and the proportion of individuals with low HDL-C (left) and high TGs (right) on the y axis. We modeled four alleles for low HDL-C and three for high TGs, associated at < 0.01 level (Supplementary table 2). The dashed line indicates the expected proportions given the number of cases and controls. Both trends were highly significant ($P < 1.89 \times 10^{-8}$).

Table 1

Association of Caucasian GWAS lipid signals in Mexican FCHL families and hypertriglyceridemia case-control study samples.

Trait	Chr.	SNP[16;24;8]	Nearest gene of interest	Mexican FCHL families			Mexican Cases and controls				Mexican Combined		Risk Allele[e]	
				MAF[a]	z^b	P	MAF[c]	z^b	P	OR[d] (95% CI)	z^b	P	Mex	Eur
HDL	1q42	rs4846914	GALNT2	0.45	2.19	2.84×10^{-2}	0.39/0.44	2.16	3.06×10^{-2}	1.24 (1.02-1.51)	3.06	2.20×10^{-3}	G	G
HDL	11q23	rs964184	APOA5	0.38	0.93	3.51×10^{-1}	0.30/0.39	3.53	4.16×10^{-4}	1.43 (1.17-1.74)	3.27	1.06×10^{-3}	G	G
HDL	16q13	rs3764261[f]	CETP	0.32	-2.58	9.93×10^{-3}	0.41/0.28	-5.18	2.21×10^{-7}	0.58 (0.47-0.71)	-5.58	2.35×10^{-8}	C	C
TG	2p23	rs1260326	GCKR	0.36	0.61	5.42×10^{-1}	0.29/0.36	3.57	3.55×10^{-4}	1.37 (1.15-1.64)	3.22	1.26×10^{-3}	T	T
TG	11q23	rs964184	APOA5	0.38	3.07	2.13×10^{-3}	0.27/0.40	6.03	1.62×10^{-9}	1.74 (1.45-2.08)	6.67	2.60×10^{-11}	G	G

[a]MAF indicates the minor allele frequency in 140 founders.

[b]Z score of the minor allele for low HDL-C or high TGs.

[c]MAF indicates the minor allele frequency in controls/cases of the corresponding trait.

[d]OR (95% CI) indicates the odds-ratio and 95% confidence interval.

[e]Risk allele indicates the allele which is associated with unfavorable lipid outcome (low HDL-C or high TGs) in Europeans (Eur)[5-7] and Mexicans (Mex).

[f]The SNP rs3764261 which is in perfect LD ($r^2 = 1$) with rs173539 in the HapMap CEU data, was genotyped instead of rs173539.

Table 2

Allele frequencies of the lipid associated GWAS variants in Mexican, Mexican-American and European samples.

SNP[5-7]	Nearest gene of interest	Reference/other allele[a]	Mexican controls[b]	Mexican-American[d]	European[c]	Mex vs. Eur[e]	Mex-Am vs. Eur[f]
rs10096633	LPL	C/T	T (0.15)	T (0.13)	T (0.14)	ns	ns
rs10468017	LIPC	C/T	T (0.19)	T (0.14)	T (0.31)	0.002	0.001
rs10889353	ANGPTL3	A/C	C (0.42)	C (0.33)	C (0.33)	ns	ns
rs1260326	GCKR	T/C	T (0.39)	T (0.31)	T (0.42)	ns	ns
rs1883025	ABCA1	C/T	A (0.24)	A (0.35)	A (0.20)	ns	0.003
rs2083637	LPL	A/G	C (0.26)	C (0.23)	C (0.27)	ns	ns
rs2271293	LCAT	G/A	A (0.14)	A (0.15)	A (0.09)	ns	ns
rs2338104	MMAB, MVK	C/G	G (0.45)	G (0.45)	C (0.49)	ns	ns
rs2954029	TRIB1	A/T	T (0.39)	T (0.39)	T (0.38)	ns	ns
rs3764261	CETP	C/A	A (0.43)	A (0.29)	A (0.34)	0.04	ns
rs4846914	GALNT2	G/A	G (0.37)	G (0.51)	G (0.43)	ns	ns
rs7557067	APOB	G/A	G (0.29)	G (0.28)	G (0.20)	ns	ns
rs7819412	XKR6-AMAC1L2	G/A	A (0.27)	A (0.33)	A (0.51)	1.04×10^{-8}	0.003
rs964184	APOA1/C3/A4/A5	G/C	G (0.27)	G (0.27)	G (0.12)	5.89×10^{-5}	9.75×10^{-4}

[a]Allele are in + strand relative to the human reference sequence.

[b]Minor allele (MAF) in150 of the Mexican controls (TGs < 150 mg/dL and HDL-C > 50 mg/dL).

[c]Minor allele (MAF) in 50 Mexican-American founders from the HapMap project.

[d]Minor allele (MAF) in 112 CEU founders with European ancestry from the HapMap project.

[e]P-value results for differences in allele frequency between the Mexican controls (Mex) and CEU founders (Eur) of HapMap determined by a chi-square test.

[f]P-value results for differences in allele frequency between Mexican-American founders of HapMap (Mex-Am) and CEU founders (Eur) of HapMap determined by a chi-square test. ns, indicates non-significant results (P>0.05)