



# Exome Sequencing Data Analysis and a Case-Control Study in Mexican Population Reveals Lipid Trait Associations of New and Known Genetic Variants in Dyslipidemia-Associated Loci

## OPEN ACCESS

### Edited by:

Tommaso Pippucci,  
Policlinico Sant'Orsola-Malpighi, Italy

### Reviewed by:

Alfonsina Tirozzi,  
University of Molise, Italy

Mark Z. Kos,  
The University of Texas Rio Grande  
Valley, United States

### \*Correspondence:

Emilio J. Córdova  
ecordova@inmegen.gob.mx  
Lorena Orozco  
lorozco@inmegen.gob.mx

†These authors share last authorship

### Specialty section:

This article was submitted to  
Genetics of Common and Rare  
Diseases,  
a section of the journal  
Frontiers in Genetics

Received: 02 November 2021

Accepted: 25 April 2022

Published: 20 May 2022

### Citation:

Jurado-Camacho PA, Cid-Soto MA,  
Barajas-Olmos F, García-Ortíz H,  
Baca-Peynado P,  
Martínez-Hernández A,  
Centeno-Cruz F, Contreras-Cubas C,  
González-Villalpando ME,  
Saldaña-Álvarez Y, Salas-Martínez G,  
Mendoza-Caamal EC,  
González-Villalpando C, Córdova EJ  
and Orozco L (2022) Exome  
Sequencing Data Analysis and a Case-  
Control Study in Mexican Population  
Reveals Lipid Trait Associations of New  
and Known Genetic Variants in  
Dyslipidemia-Associated Loci.  
Front. Genet. 13:807381.  
doi: 10.3389/fgene.2022.807381

Pedro A. Jurado-Camacho<sup>1,2</sup>, Miguel A. Cid-Soto<sup>1</sup>, Francisco Barajas-Olmos<sup>1</sup>,  
Humberto García-Ortíz<sup>1</sup>, Paulina Baca-Peynado<sup>1,2</sup>, Angélica Martínez-Hernández<sup>1</sup>,  
Federico Centeno-Cruz<sup>1</sup>, Cecilia Contreras-Cubas<sup>1</sup>, María Elena González-Villalpando<sup>3</sup>,  
Yolanda Saldaña-Álvarez<sup>1</sup>, Guadalupe Salas-Martínez<sup>1</sup>, Elvia C. Mendoza-Caamal<sup>4</sup>,  
Clicerio González-Villalpando<sup>3</sup>, Emilio J. Córdova<sup>5†\*</sup> and Lorena Orozco<sup>1†\*</sup>

<sup>1</sup>Immunogenomics and Metabolic Diseases Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico,

<sup>2</sup>Posgraduate in Biomedical Sciences, National Autonomous University of Mexico, Mexico City, Mexico, <sup>3</sup>Centro de Estudios en Diabetes, Unidad de Investigación en Diabetes y Riesgo Cardiovascular, Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública, Mexico City, Mexico, <sup>4</sup>Clinical Area, National Institute of Genomic Medicine, Mexico City, Mexico,

<sup>5</sup>Oncogenomics Consortium Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico

**Background:** Plasma lipid levels are a major risk factor for cardiovascular diseases. Although international efforts have identified a group of loci associated with the risk of dyslipidemia, Latin American populations have been underrepresented in these studies.

**Objective:** To know the genetic variation occurring in lipid-related loci in the Mexican population and its association with dyslipidemia.

**Methods:** We searched for single-nucleotide variants in 177 lipid candidate genes using previously published exome sequencing data from 2838 Mexican individuals belonging to three different cohorts. With the extracted variants, we performed a case-control study. Logistic regression and quantitative trait analyses were implemented in PLINK software. We used an LD pruning using a 50-kb sliding window size, a 5-kb window step size and a  $r^2$  threshold of 0.1.

**Results:** Among the 34251 biallelic variants identified in our sample population, 33% showed low frequency. For case-control study, we selected 2521 variants based on a minor allele frequency  $\geq 1\%$  in all datasets. We found 19 variants in 9 genes significantly associated with at least one lipid trait, with the most significant associations found in the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster on chromosome 11. Notably, all 11 variants associated with hypertriglyceridemia were within this cluster; whereas variants associated with hypercholesterolemia were located at chromosome 2 and 19, and for low high density lipoprotein cholesterol were in chromosomes 9, 11, and 19. No significant associated variants were found for low density lipoprotein. We found several novel variants associated with

different lipemic traits: rs3825041 in *BUD13* with hypertriglyceridemia, rs7252453 in *CILP2* with decreased risk to hypercholesterolemia and rs11076176 in *CETP* with increased risk to low high density lipoprotein cholesterol.

**Conclusions:** We identified novel variants in lipid-regulation candidate genes in the Mexican population, an underrepresented population in genomic studies, demonstrating the necessity of more genomic studies on multi-ethnic populations to gain a deeper understanding of the genetic structure of the lipemic traits.

**Keywords:** association study, genetic variants, dyslipidemia, exome analysis, mexican population

## INTRODUCTION

Currently, metabolic diseases have become one of the most challenging health concerns worldwide; they account for nearly 45% of all deaths worldwide (Forouzanfar et al., 2016). Prevention and management of these diseases are complicated, due to their long latency periods, numerous risk factors, and the presence of co-morbidities (Sevick et al., 2007). Among the known metabolic diseases, dyslipidemias comprise a well-established risk factor for cardiovascular diseases, which are the leading cause of death worldwide (Libby, 2002).

Although factors like diet and lifestyle (Wietlisbach et al., 1997) are recognized as important determinants in the clinical development of dyslipidemias, these diseases have also a strong genetic component (Gao et al., 2018). For instance, lipid disorders are highly prevalent in populations with Amerindian ancestry (up to 85.9%), compared with the prevalence among individuals from other ancestries (e.g., Caucasian 31.2% and African 41.1%) (Aguilar-Salinas et al., 2014; Noubiap et al., 2018). In addition, a large number of genome-wide association studies (GWAS) performed in populations of different ancestries have described more than 175 loci associated with dysregulated levels of plasma lipids. However, the largest body of information generated to date has relied mostly on evidence from Caucasian and Asian cohorts, with very few studies analyzing Latin American populations (Teslovich et al., 2010; Asselbergs et al., 2012; Willer et al., 2013; Wu et al., 2013; Surakka et al., 2015).

Recent findings have demonstrated that genetic factors associated with metabolic traits, including dyslipidemias, exhibited significant heterogeneity in allele frequency and in variant effects across groups with different ancestries (Klarin et al., 2018; Martagón et al., 2018). A study analyzing populations from multiple ethnicities has found important differences in the levels of association, allele frequencies, and haplotype distributions of several lipid loci (Wu et al., 2013). Several population-specific signals at these loci have been reported in non-European populations. For example, the association between the regulatory variant, rs12740374, in the *CELSR2/PSRC1/SORT1* locus, and low-density lipoprotein cholesterol (LDL-C) is higher in African-American individuals than in individuals of European descent (Buyske et al., 2012). Furthermore, the R230C variant of the *ABCA1* gene, which is associated with low levels of high-density lipoprotein cholesterol (HDL-C), is private to individuals with Amerindian ancestry (Acuña-Alonzo et al., 2010). Thus, the

distribution of variants located in lipid-related genes might vary between populations of different ethnicities.

The modern Mexican population is mainly composed of Mestizo individuals, who are the result of the recent admixture of original Amerindians, Europeans (mainly Spaniards) and, to a lesser extent, sub-Saharan Africans (Moreno-Estrada et al., 2014). The recent and complex admixture in the Mexican Mestizo population might have produced a high level of genetic heterogeneity in variants in lipid-related loci. Therefore, this study aimed to determine the frequency distribution of variants located at genes related to lipid traits in Mexican individuals, followed by an association testing and the identification of potential variants for lipemic traits, using previously published exome sequencing data (Estrada et al., 2014; García-Ortiz et al., 2021).

## MATERIAL AND METHODS

### Study Populations

This study included 2838 Mexican individuals belonging to the Mestizo cohorts Diabetes in Mexico Study (DMS,  $n = 968$ ) and Mexico City Diabetes Study (MCDS,  $n = 796$ ), published previously as part of the Slim Initiative in Genomic Medicine for the Americas (SIGMA) Type 2 Diabetes Consortium (Estrada et al., 2014), as well as to the indigenous cohort Metabolic Analysis in an Indigenous Sample (MAIS,  $n = 1074$ ) (García-Ortiz et al., 2021). The sample design was previously described in Estrada et al. and García-Ortiz et al. (Estrada et al., 2014; García-Ortiz et al., 2021). The MAIS sample belongs to 71 indigenous communities representing 60 ethnic groups from 10 linguistic families. All participants were unrelated volunteers and provided signed informed consent. This investigation was approved by the local ethics and research committees from the National Institute of Genomic Medicine and was conducted according to the principles of the Declaration of Helsinki. The genetic structure of DMS and MCDS population was previously described, with a mean (SD) proportion of Amerindian ancestry of  $66 \pm 17\%$ , whereas the proportion of Amerindian ancestry in the MAIS cohort was of  $93.2 \pm 8.7\%$  (Estrada et al., 2014; García-Ortiz et al., 2021).

### Lipid Measurements in Plasma

Levels of triglycerides (TG), total cholesterol (TC), and HDL-C were measured from blood samples collected after overnight

fasting, using a Synchron CX5 Analyzer System (Beckman Coulter Fullerton, CA, United States). LDL-C values were calculated with the Friedewald formula, excluding those samples with TG > 400 mg/dl (Warnick et al., 1990).

Each lipid disorder was diagnosed according to the American Heart Association and National Heart, Lung, and Blood Institute guidelines (AHA/NHLBI; <http://www.nhlbi.nih.gov>). An individual was diagnosed with a lipid disorder when serum levels showed any of the following: TG  $\geq$  150 mg/dl (hypertriglyceridemia; HTG), TC  $\geq$  200 mg/dl (hypercholesterolemia; HTC), LDL-C  $\geq$  130 mg/dl (elevated LDL), or HDL-C  $\leq$  50 mg/dl in females or  $\leq$  40 mg/dl in males (low HDL-C). Individuals with desirable lipid values were assigned as controls. Data on lipid-lowering medications were available for over 80% of the participants and adjustment was done for TG and TC binary phenotypes.

## Dataset Building and Single-Nucleotide Variant Annotations

First, we analyzed the Variant Call Format (VCF) files previously obtained from all participants. Quality controls for whole-exome sequencing (e.g., read depth, mean coverage, and missing rate) were described previously (Estrada et al., 2014). From these files, we extracted all the biallelic variants (mapped in the Genome Reference Consortium Human genome build 37) of 177 candidate genes for any lipid trait (**Supplementary Table S1**). These genes were selected from: 1) the Global Lipid Genetics Consortium ( $n = 165$ ) (Willer et al., 2013), 2) meta-analysis studies ( $n = 10$ ) (Teslovich et al., 2010; Weissglas-Volkov and Pajukanta, 2010), and 3) re-sequencing and clinical exome studies ( $n = 2$ ) (Estrada et al., 2014; Williams-Amy et al., 2014). We annotated the variants with the ENSEMBL Variant Effect Predictor (<https://www.ensembl.org/info/docs/tools/vep/index.html>; version 87) (McLaren et al., 2016). Finally, we constructed PLINK files that comprised clinical-demographic and genotyping data for the association analyses, which employed VCF tools (v0.1.12b) (Danecek et al., 2011).

## Statistical Analysis

A case-control study was conducted in each cohort to identify low-frequency SNVs [minor allele frequency (MAF) = 0.01–0.05] and common SNVs (MAF > 0.05) associated with lipid traits. To analyze associations between lipid components and alleles we estimated the odds ratio (OR) using logistic regression. We performed quantitative trait analyses with linear regression (Beta value). Both methods were performed using an additive model, adjusting by age, sex and the first 10 eigenvectors from the principal component analysis as covariates. All analyses were performed using PLINK v.1.9 software (Purcell et al., 2007).

Next, we performed a fixed effects model meta-analysis for each trait using a weighted inverse variance model in the software package METAL (Willer et al., 2010). Also, genomic control was applied to each study within METAL by adjusting for the genomic inflation factor prior to meta-analysis, to correct for possible residual population stratification.

Significant threshold was determined following the approach by Kanai et al. (Kanai et al., 2016). According to this approach an LD pruning was done with PLINK v.1.9 using a 50-kb sliding window size, a 5-kb window step size and a  $r^2$  threshold of 0.1. According with this, a total of 780 independent SNVs were identified giving a significant threshold of  $6.4 \times 10^{-5}$ .

Significant threshold was determined following the approach by Kanai et al. (Kanai et al., 2016). According to this approach, high LD SNVs in 177 candidate lipids genes were filtered based on LD pruning performed in PLINK v.1.9 using a 50-kb sliding window size, a 5-kb window step size and a  $r^2$  threshold of 0.1. According with this, a total of 780 independent SNVs were identified that were used to establishing the genome-wide significant threshold *via* a straight-forward Bonferroni correction  $p < (6.4 \times 10^{-5})$ .

## RESULTS

### Study Participants

Our study population comprised 2838 individuals belonging to three different cohorts, the Diabetes in Mexico Study (DMS), the Mexico City Diabetes Study (MCDS) and the indigenous cohort Metabolic Analysis in an Indigenous Sample (MAIS) (**Table 1**). The DMS sample was composed of 968 individuals: 681 females (70.3%) and 287 males (29.7%). The mean age of the participants was  $54.1 \pm 9.8$  years and the mean body mass index BMI was  $28.4 \pm 4.9$  kg/m<sup>2</sup>. According to the AHA/NHLBI cutoff points, the lipid trait with the highest prevalence in this sample was low HDL-C (70.8%), followed by HTC (65.1%), HTG (54.9%), and high LDL-C (37%). The MCDS sample consisted of 796 individuals: 482 females (60.6%) and 314 males (39.4%) with a mean age of  $62.9 \pm 7.6$  years and a mean BMI of  $29.5 \pm 4.9$  kg/m<sup>2</sup>. The most frequent lipid alteration in this population was low HDL-C (82.8%), followed by HTG (58.8%), HTC (44.1%), and high LDL-C (40.0%). The MAIS sample was composed of 1074 individuals: 679 females (63.2%) and 395 males (36.8%) with a mean age of  $58.6 \pm 12.1$  years and a mean BMI of  $27.5 \pm 5.0$  kg/m<sup>2</sup>. Similar to the two previous cohorts, low HDL-C (75.2%) was the most common lipid trait, followed by HTG (66.8%), HTC (29.7%) and high LDL-C (14.9%) (**Table 1**).

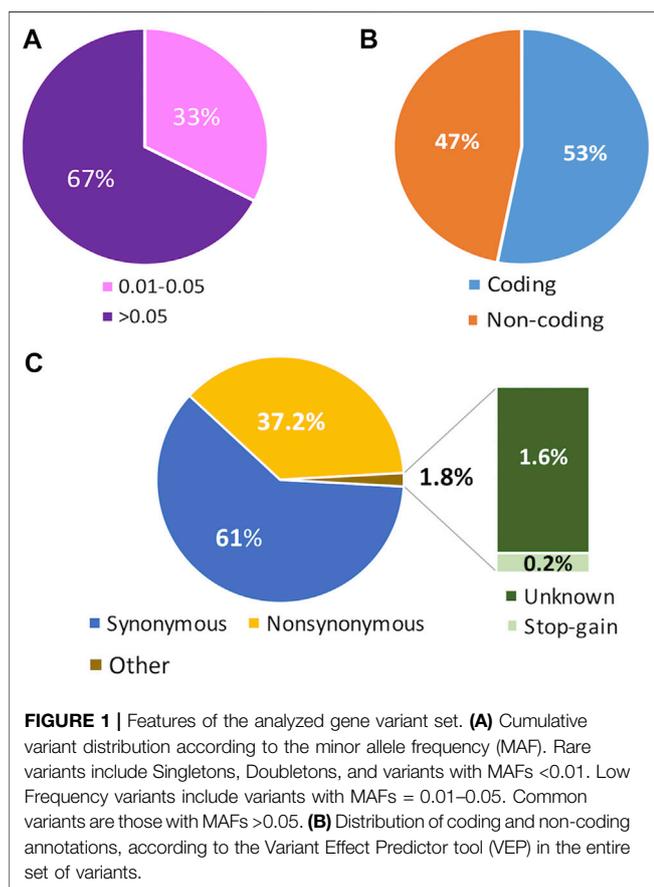
### Analysis of the Variation in Candidate Dyslipidemia-Related Genes

After analyzing the exome data from the three analyzed cohorts, we found a total of 34251 biallelic variants with a MAF above 1% within the 177 genes previously related to dyslipidemia. Among these variants, 33% were detected at low frequency (MAF = 0.01–0.05), and 67% were common (MAF > 0.05) (**Figure 1A**). According to their position in the gene, 53% of the variants were in coding regions, and the remaining 47% were found in non-coding regions (**Figure 1B**). According to the predicted annotation obtained with the Variant Effect Predictor web tool (<https://grch37.ensembl.org/Tools/VEP>) 61% of the variants found in the coding regions were synonymous, 37.2% were non-synonymous, 0.2% were stop-gain and 1.6% were unknown (**Figure 1C**).

**TABLE 1** | Clinical and demographic characteristics of the studied cohort.

Characteristic	DMS	MCDS	MAIS
	<i>n</i> = 968	<i>n</i> = 796	<i>n</i> = 1074
Women/Men (%)	70.3/29.7	60.6/39.4	63.2/36.8
Age (Years)	54.1 ± 9.8	62.9 ± 7.6	58.6 ± 12.1
BMI (kg/m <sup>2</sup> )	28.4 ± 4.9	29.5 ± 4.9	27.5 ± 5.0
FG (mg/dl)	125.2 ± 61.4	114.3 ± 54.2	114.2 ± 63.9
HTG (%)	54.9	58.8	66.8
Mean TG (mg/dl)	203.7 ± 129.3	176.3 ± 96.3	207.5 ± 119.5
HTC (%)	65.1	44.1	29.7
Mean TC (mg/dl)	201.1 ± 43.8	195.9 ± 36.5	182.8 ± 38.3
Low HDL-C (%)	70.8	82.8	75.2
Mean HDL-C (mg/dl)			
	Women	31.5 ± 8.5	40.6 ± 12.6
	Men	37.4 ± 13.0	38.7 ± 12.6
High LDL-C (%)	37	40.0	14.9
Mean <sup>1</sup> LDL-C (mg/dl)	120.7 ± 35.3	131.5 ± 33.6	103.7 ± 29.5

Data are presented as the mean ± SD, or the percentage, as indicated. BMI, body mass index; FG, fasting glucose; HTG, hypertriglyceridemia; TG, triglycerides; HTC, hypercholesterolemia; TC, total cholesterol; HDL-C, high density lipoprotein; LDL-C, Low density lipoprotein. <sup>1</sup>LDL-C values were calculated with the Friedewald equation.



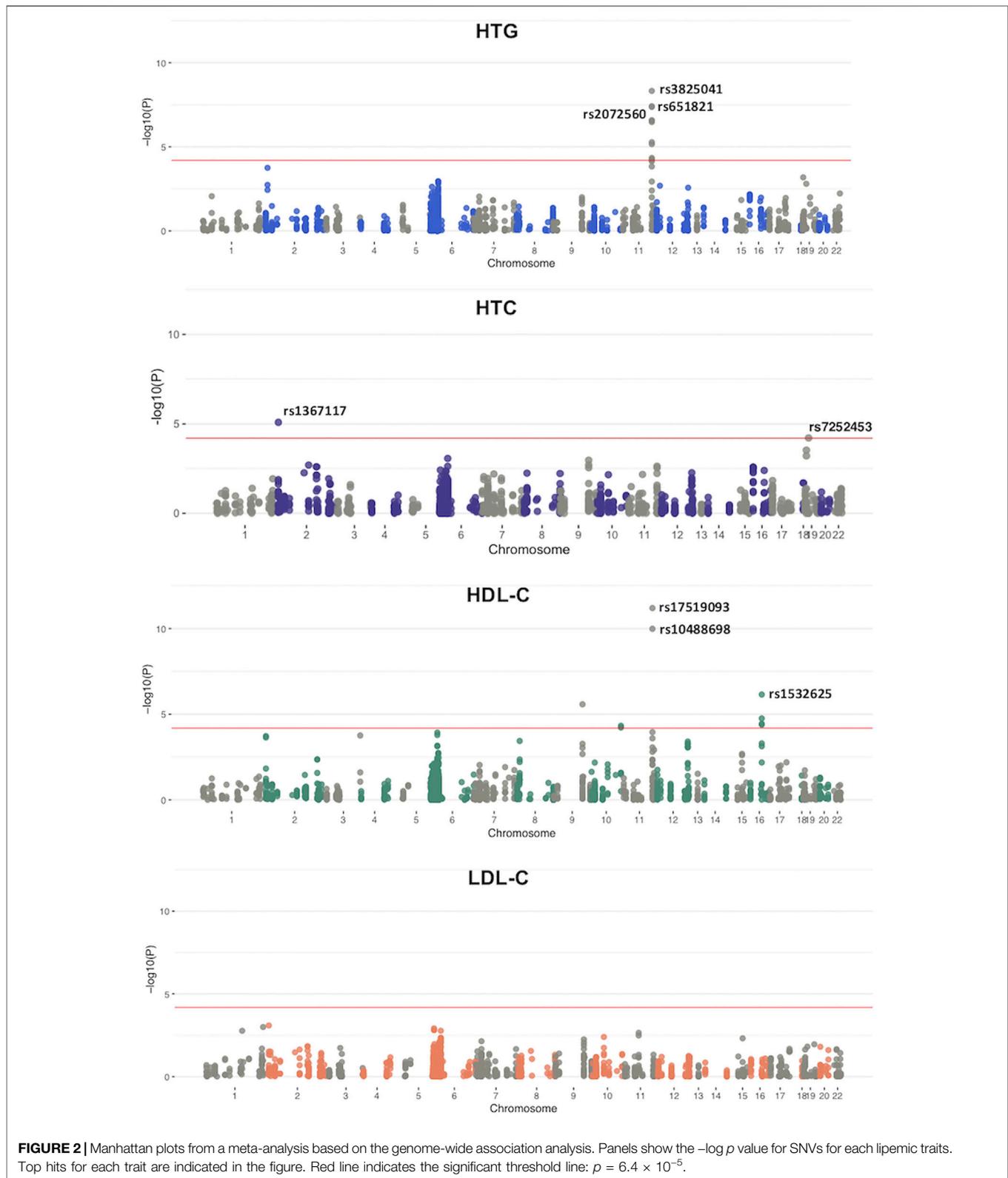
## Low-Frequency and Common Variant Association Study

For the case-control study we used a meta-analysis approach to evaluate low-frequency and common variants pruned for LD (MAF >0.01; *n* = 780). We explored associations between individual variants and each different lipid phenotype with

binary logistic regression. Quantile-quantile (QQ) plots for each analyzed phenotype showed a genomic inflation factor for HTG, HTC, low HDL-C and high LDL-C of  $\lambda = 1.0, 1.02, 1.03,$  and  $0.91$  respectively (**Supplementary Figure S1**). A total of 19 variants in 9 genes achieved a  $p < 6.4 \times 10^{-5}$  or higher significant association with at least one of the analyzed lipid components (**Figure 2**). Remarkably, the 11 variants that were associated with HTG were found on chromosome 11, located in *BUD13*, *APOA5*, *APOC3*, *APOA1*, and *APOA4* genes (**Supplementary Table S2**). In the case of HTC, associated variants were found on chromosome 2 and 19, in *APOB* and *CILP2* genes; whereas in the case of low HDL-C, two of the associated variants were found on chromosome 11, in *BUD13* gene, four on chromosome 16 in *CETP* and one on chromosome 9 in *ABCA1* (**Supplementary Table S2**). We did not find any variant associated with high LDL-C.

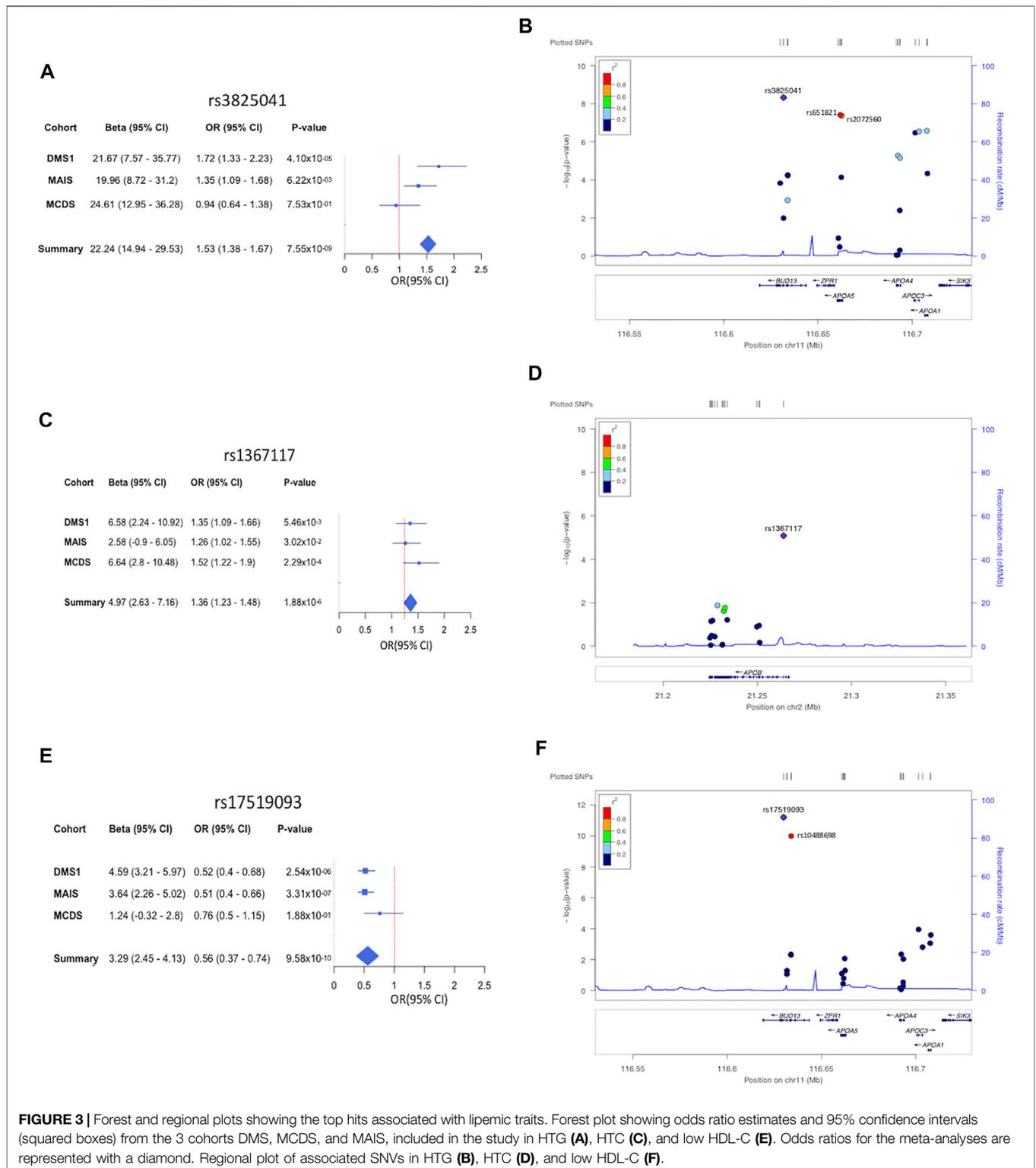
The 11 variants associated with HTG were all located in the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster (**Figure 2**; **Supplementary Table S2**). The strongest signal for the variants associated with increased risk to HTG was detected for rs3825041 in *BUD13* (OR = 1.53,  $p = 7.55 \times 10^{-9}$ ;  $\beta = 22.24$  mg/dl) (**Figure 3**; **Supplementary Table S2**), followed by rs651821 (OR = 1.48,  $p = 5.70 \times 10^{-8}$ ;  $\beta = 21.40$  mg/dl) and rs2072560 in *APOA5* (OR = 1.48,  $p = 6.18 \times 10^{-8}$ ;  $\beta = 21.65$  mg/dl), rs2070665 in *APOA1* (OR = 1.41,  $p = 3.79 \times 10^{-7}$ ;  $\beta = 17.71$  mg/dl), rs5128 in *APOC3* (OR = 1.40,  $p = 4.22 \times 10^{-7}$ ;  $\beta = 18.31$  mg/dl), rs5104 (OR = 1.34,  $p = 6.1 \times 10^{-6}$ ;  $\beta = 19.61$  mg/dl) and rs5092 in *APOA4* (OR = 1.33,  $p = 8.07 \times 10^{-6}$ ;  $\beta = 18.97$  mg/dl) and rs11820589 in *BUD13* (OR = 1.35,  $p = 5.71 \times 10^{-5}$ ;  $\beta = 21.92$  mg/dl). In contrast, the variants rs4520 in *APOC3* (OR = 0.74,  $p = 2.96 \times 10^{-7}$ ;  $\beta = -19.03$  mg/dl) rs5070 in *APOA1* (OR = 0.78,  $p = 4.27 \times 10^{-5}$ ;  $\beta = -15.75$  mg/dl) and rs10488698 in *BUD13* (OR = 0.73,  $p = 5.80 \times 10^{-5}$ ;  $\beta = -20.98$  mg/dl) were associated with protection against HTG. All 11 variants were also significantly associated with TG levels in the quantitative trait analysis (**Supplementary Figure S2**; **Supplementary Table S2**).

Moreover, the variant rs1367117 in *APOB* (OR = 1.36,  $p = 1.88 \times 10^{-6}$ ;  $\beta = 4.89$ ) was significantly associated with HTC and with a



significant increase in serum TC levels in the quantitative trait analysis (Figure 3; Supplementary Table S2). On the other hand, the variant rs7252453 in *CILP2* ( $OR = 0.73$ ,  $p = 4.40 \times 10^{-5}$ ;  $\beta =$

$-5.17$ ) was significantly associated with protection against HTC (Supplementary Figure S3; Supplementary Table S2). Regarding low HDL-C, the variants rs17519093 and



rs10488698 in *BUD13* (OR = 0.56,  $p = 9.58 \times 10^{-10}$ ,  $\beta = 3.29$ ; OR = 0.58,  $p = 7.34 \times 10^{-9}$ ;  $\beta = 3.12$ , respectively) and rs1532625 in *CETP* (OR = 0.72,  $p = 6.56 \times 10^{-6}$ ;  $\beta = 2.21$ ) were significantly associated with decreased risk in both the binary and the

quantitative trait analysis (Figure 3; Supplementary Figure S4; Supplementary Table S2). Finally, the variants rs708272, rs11076176, and rs289714 in *CETP* (OR = 1.32,  $p = 1.75 \times 10^{-5}$ ,  $\beta = -2.12$ ; OR = 1.32,  $p = 3.64 \times 10^{-5}$ ,  $\beta = -1.83$ ; OR = 1.32,  $p = 4.03$

$\times 10^{-5}$ ,  $\beta = -1.75$ , respectively) as well as rs9282541 in *ABCA1* (OR = 1.67,  $p = 3.34 \times 10^{-5}$ ;  $\beta = -2.50$ ) showed a significant association with increased risk to low HDL-C in both the binary and the quantitative trait analysis (**Supplementary Figure S4; Supplementary Table S2**).

## DISCUSSION

The contribution of genetic variation to human diseases is widely recognized (Xue et al., 2012). In particular, loss-of-function or modifier variants are usually related to changes in the biological activity of the corresponding gene. Several genomic regions are currently recognized as drivers of dyslipidemias in populations from several ethnicities, although in some populations, such as those of Amerindian or African origin, genomic studies have been scarce. Furthermore, many of these loci have displayed highly different associations with these entities across different ancestries (Lek et al., 2016; Martin et al., 2017). Therefore, lipid-associated genes should be analyzed in all populations.

Here, by analyzing SNVs within 177 candidate genes, we found several associations with different types of dyslipidemias in Mexican Mestizos, an admixed population with strong Amerindian (51%) and European (46%) components (Norris et al., 2018).

In our population we were able to replicate several associations with different lipemic traits previously reported in other ethnicities. For example, variants in the gene cluster *APOA1/C3/A4/A5-ZPR1-BUD13*, such as the regulatory SNVs rs5128 in *APOC3*, and rs651821 in *APOA5*, as well as the missense SNV rs2072560 also within *APOA5*, and the intronic variant rs2070665 in *APOA1*, were all associated with HTG in our population, as they are in European, Asian, and African populations (Feng et al., 2016; Fu et al., 2015; Jasim et al., 2018; Ken-Dror et al., 2010; Song et al., 2015; Zhou et al., 2013). Likewise, the association of the missense SNV rs1367117 in *APOB* with HTC has also been reported in populations of European ancestry (Lu et al., 2010), whereas the association of the missense SNVs rs10488698 in *BUD13* and rs9282541 in *ABCA1* with low HDL-C has also been observed in Asian and Latino American populations, respectively (Zhang et al., 2017; Acuña-Alonso et al., 2010). We also observed novel lipemic trait-associated variants in *BUD13*, such as rs3825041 and rs17519093, both localized within introns, associated with HTG and high levels of HDL-C, respectively. Using public data contained in the Common Metabolic Diseases Knowledge Portal (<https://t2d.hugeamp.org/>), we were able to confirm the associations of rs3825041 and rs17519093 with different lipemic traits. Thus, one of the genomic regions most consistently associated with lipid traits in human populations of diverse ethnic origins is the cluster *APOA1/C3/A4/A5-ZPR1-BUD13* (Teslovich et al., 2010; Willer et al., 2013; Parra et al., 2017; Bai et al., 2019). Importantly, all 11 variants associated with HTG in our study, were in this cluster. Among them, the novel variant rs3825041 in *BUD13* showed the strongest association. Notably, this variant showed high LD with rs651821 ( $r^2 = 0.84$ ) and rs2072560 ( $r^2 = 0.89$ ) within *APOA5*, which have both been previously reported as associated with HTG in several

populations of different ancestries, including in Mexicans (Ken-Dror et al., 2010; Kim et al., 2019). Taken together, these data provide more insights about the variants at the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster as a relevant risk factor for dyslipidemias such as HTG and low HDL-C, and highlight the notion that these could be biomarkers for susceptibility to these traits.

Others novel association signals were observed with the synonymous SNV rs7252453 in *CILP2* and decreased risk to HTC and the intronic SNV rs11076176 in *CETP* and increased risk to low HDL-C serum levels. On *CETP* we also observed an association of the intronic variants rs708272 and rs289714 with high risk to low HDL-C, as well as the association of the intronic variant rs1532625 with protection to this dyslipidemia. These findings are in line with those reported previously in Mexican individuals (Vargas-Alarcon et al., 2018), in a meta-analysis involving six independent Hispanic cohorts (Gao et al., 2018) and in Chinese population (Guo et al., 2015).

In summary, despite our selecting candidate genes were previously associated with dyslipidemia in other populations, we were able to find additional variants showing the strongest associations with lipid traits in Mexican individuals. These differences support the notion that high allelic heterogeneity exists in lipid loci across populations. Remarkably, the new associations of variants in genes previously related to dyslipidemia, points out the importance of studying different ethnicities, since different associated variants within the same genes could be particular to one or another population ancestry. It is also worth to note that several of the risk variants previously associated with lipemic traits in different ethnic groups, including European, Asian and African populations were replicated in our study. Taken together, our results suggest that genetic architecture of dyslipidemias is partially share among different ethnic groups.

## DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the European Variation Archive repository (EVA <https://www.ebi.ac.uk/eva/>), accession number PRJEB52611.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the local ethics and research committees from the National Institute of Genomic Medicine. The patients/participants provided their written informed consent to participate in this study.

## AUTHORS CONTRIBUTIONS

LO and EC contributed to the conception and design of the work. PJ-C, FB-O, HG-O, and MC-S performed the bioinformatic

analyses for variant calling. PB-P, AM-H, MG-V, and CG-V contributed in acquisition of data. FC-C, CC-C, YS-A, GS-M, and EM-C participated in the association studies. PJ-C, LO, and EC wrote the manuscript.

## FUNDING

This work was supported by the National Institute of Genomic Medicine (grant number INM 2014); and the CONACyT (grant numbers SEP-2014-243587 and PDCPN\_2016\_3251).

## REFERENCES

- Acuña-Alonzo, V., Flores-Dorantes, T., Kruit, J. K., Villarreal-Molina, T., Arellano-Campos, O., Hünemeier, T., et al. (2010). A Functional ABCA1 Gene Variant Is Associated with Low HDL-Cholesterol Levels and Shows Evidence of Positive Selection in Native Americans. *Hum. Mol. Genet.* 19 (14), 2877–2885. doi:10.1093/hmg/ddq173
- Aguilar-Salinas, C. A., Tusie-Luna, T., and Pajukanta, P. (2014). Genetic and Environmental Determinants of the Susceptibility of Amerindian Derived Populations for Having Hypertriglyceridemia. *Metabolism Clin. Exp.* 63, 887–894. doi:10.1016/j.metabol.2014.03.012
- Asselbergs, F. W., Guo, Y., Van Iperen, E. P. A., Sivapalaratnam, S., Tragante, V., Lanktree, M. B., et al. (2012). Large-scale Gene-Centric Meta-Analysis across 32 Studies Identifies Multiple Lipid Loci. *Am. J. Hum. Genet.* 91 (5), 823–838. doi:10.1016/j.ajhg.2012.08.032
- Bai, W., Kou, C., Zhang, L., You, Y., Yu, W., Hua, W., et al. (2019). Functional Polymorphisms of the APOA1/C3/A4/A5-ZPR1-BUD13 Gene Cluster Are Associated with Dyslipidemia in a Sex-specific Pattern. *PeerJ* 6, e6175. doi:10.7717/peerj.6175
- Buyse, S., Wu, Y., Carty, C. L., Cheng, L., Assimes, T. L., Dumitrescu, L., et al. (2012). Evaluation of the MetaboChip Genotyping Array in African Americans and Implications for Fine Mapping of Gwas-Identified Loci: The PAGE Study. *PLoS ONE* 7 (4), e35651. doi:10.1371/journal.pone.0035651
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The Variant Call Format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Estrada, K., Aukrust, I., Bjørkhaug, L., Burt, N. P., Mercader, J. M., García-Ortiz, H., et al. (2014). Association of a Low-Frequency Variant in HNF1A with Type 2 Diabetes in a Latino Population the SIGMA Type 2 Diabetes Consortium. *JAMA - J. Am. Med. Assoc.* 311, 2305–2314. doi:10.1001/jama.2014.6511
- Feng, D. W., Ma, R. L., Guo, H., He, J., Yan, Y. Z., Muratbek, et al. (2016). Association of APOA1 Gene Polymorphisms (Rs670, Rs5069, and Rs2070665) with Dyslipidemia in the Kazakhs of Xinjiang. *Genet. Mol. Res.* 15. doi:10.4238/gmr.15028094
- Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., et al. (2016). Global, Regional, and National Comparative Risk Assessment of 79 Behavioural, Environmental and Occupational, and Metabolic Risks or Clusters of Risks, 1990–2015: a Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* 388, 1659–1724. doi:10.1016/S0140-6736(16)31679-8
- Fu, Q., Tang, X., Chen, J., Su, L., Zhang, M., Wang, L., et al. (2015). Effects of Polymorphisms in APOA4-APOA5-ZNF259-BUD13 Gene Cluster on Plasma Levels of Triglycerides and Risk of Coronary Heart Disease in a Chinese Han Population. *PLoS ONE* 10. doi:10.1371/journal.pone.0138652
- Gao, C., Tabb, K. L., Dimitrov, L. M., Taylor, K. D., Wang, N., Guo, X., et al. (2018). Exome Sequencing Identifies Genetic Variants Associated with Circulating Lipid Levels in Mexican Americans: The Insulin Resistance Atherosclerosis Family Study (IRASFS). *Sci. Rep.* 8. doi:10.1038/s41598-018-23727-2
- García-Ortiz, H., Barajas-Olmos, F., Contreras-Cubas, C., Cid-Soto, M. Á., Córdova, E. J., Centeno-Cruz, F., et al. (2021). The Genomic Landscape of Mexican Indigenous Populations Brings Insights into the Peopling of the Americas. *Nat. Commun.* 12, 5942–5954. doi:10.1038/s41467-021-26188-w
- Guo, S., Hu, Y., Ding, Y., Liu, J., and Zhang, M. (2015). Association between Eight Functional Polymorphisms and Haplotypes in the Cholesterol Ester Transfer Protein (CETP) Gene and Dyslipidemia in National Minority Adults in the Far West Region of China. *Int. J. Environ. Res. Public Health* 12, 15979–15992. doi:10.3390/ijerph121215036
- Jasim, A. A., Al-Bustan, S. A., Al-Kandari, W., Al-Serri, A., and Al-Askar, H. (2018). Sequence Analysis of APOA5 Among the Kuwaiti Population Identifies Association of Rs2072560, Rs2266788, and Rs662799 with TG and VLDL Levels. *Front. Genet.* 9. doi:10.3389/fgene.2018.00112
- Kanai, M., Tanaka, T., and Okada, Y. (2016). Empirical Estimation of Genome-wide Significance Thresholds Based on the 1000 Genomes Project Data Set. *J. Hum. Genet.* 61 (10), 861–866. PMID: 27305981; PMCID: PMC5090169. doi:10.1038/jhg.2016.72
- Ken-Dror, G., Goldbourt, U., and Dankner, R. (2010). Different Effects of Apolipoprotein A5 SNPs and Haplotypes on Triglyceride Concentration in Three Ethnic Origins. *J. Hum. Genet.* 55. doi:10.1038/jhg.2010.27
- Kim, H. K., Anwar, M. A., and Choi, S. (2019). Association of BUD13-Znf259-APOA5-APOA1-SIK3 Cluster Polymorphism in 11q23.3 and Structure of APOA5 with Increased Plasma Triglyceride Levels in a Korean Population. *Sci. Rep.* 9. doi:10.1038/s41598-019-44699-x
- Klarin, D., Damrauer, S. M., Cho, K., Sun, Y. V., Teslovich, T. M., Honerlaw, J., et al. (2018). Genetics of Blood Lipids Among ~300,000 Multi-Ethnic Participants of the Million Veteran Program. *Nat. Genet.* 50, 1514–1523. doi:10.1038/s41588-018-0222-9
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of Protein-Coding Genetic Variation in 60,706 Humans. *Nature* 536, 285–291. doi:10.1038/nature19057
- Libby, P. (2002). Inflammation in Atherosclerosis. *Nature* 420, 868–874. doi:10.1038/nature01323
- Lu, Y., Feskens, E. J., Boer, J. M., Imholz, S., Verschuren, W. M., Wijmenga, C., et al. (2010). Exploring Genetic Determinants of Plasma Total Cholesterol Levels and Their Predictive Value in a Longitudinal Study. *Atherosclerosis* 213, 200–205. doi:10.1016/j.atherosclerosis.2010.08.053
- Martagón, A. J., Bello-Chavolla, O. Y., Arellano-Campos, O., Almeda-Valdés, P., Walford, G. A., Cruz-Bautista, I., et al. (2018). Mexican Carriers of the HNF1A P. E508K Variant Do Not Experience an Enhanced Response to Sulfonylureas. *Diabetes Care* 41, 1726–1731. doi:10.2337/dc18-0384
- Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., et al. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* 107, 788–789. doi:10.1016/j.ajhg.2017.03.004
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. *Genome Biol.* 17 (1), 122. doi:10.1186/s13059-016-0974-4
- Moreno-Estrada, A., Gignoux, C. R., Fernández-López, J. C., Zakharia, F., Sikora, M., Contreras, A. v., et al. (2014). The Genetics of Mexico Recapitulates Native American Substructure and Affects Biomedical Traits. *Science* 344, 1280–1285. doi:10.1126/science.1251688
- Norris, E. T., Wang, L., Conley, A. B., Rishishwar, L., Mariño-Ramírez, L., Valderrama-Aguirre, A., et al. (2018). Genetic Ancestry, Admixture and

## ACKNOWLEDGMENTS

PJ-C is a doctoral student of the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received a CONACyT fellowship (619190).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.807381/full#supplementary-material>

- Health Determinants in Latin America. *BMC Genomics* 19. doi:10.1186/s12864-018-5195-7
- Noubiap, J. J., Bigna, J. J., Nansseu, J. R., Nyaga, U. F., Balti, E. V., Echouffo-Tcheugui, J. B., et al. (2018). Prevalence of Dyslipidaemia Among Adults in Africa: a Systematic Review and Meta-Analysis. *Lancet Glob. Health* 6, e998–e1007. doi:10.1016/S2214-109X(18)30275-4
- Parra, E. J., Mazurek, A., Gignoux, C. R., Sockell, A., Agostino, M., Morris, A. P., et al. (2017). Admixture Mapping in Two Mexican Samples Identifies Significant Associations of Locus Ancestry with Triglyceride Levels in the BUD13/ZNF259/APOA5 Region and Fine Mapping Points to Rs964184 as the Main Driver of the Association Signal. *PLoS ONE* 12. doi:10.1371/journal.pone.0172880
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795
- Sevick, M. A., Trauth, J. M., Ling, B. S., Anderson, R. T., Piatt, G. A., Kilbourne, A. M., et al. (2007). Patients with Complex Chronic Diseases: Perspectives on Supporting Self-Management. *J. General Intern. Med.* 22, 438–444. doi:10.1007/s11606-007-0316-z
- Song, Y., Zhu, L., Richa, M., Li, P., Yang, Y., and Li, S. (2015). Associations of the APOC3 Rs5128 Polymorphism with Plasma APOC3 and Lipid Levels: A Meta-Analysis. *Lipids Health Dis.* 14. doi:10.1186/s12944-015-0027-0
- Surakka, I., Horikoshi, M., Mägi, R., Sarin, A. P., Mahajan, A., Lagou, V., et al. (2015). The Impact of Low-Frequency and Rare Variants on Lipid Levels. *Nat. Genet.* 47, 589–597. doi:10.1038/ng.3300
- Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., et al. (2010). Biological, Clinical and Population Relevance of 95 Loci for Blood Lipids. *Nature* 466, 707–713. doi:10.1038/nature09270
- Vargas-Alarcon, G., Perez-Mendez, O., Herrera-Maya, G., Garcia-Sanchez, C., Martinez-Rios, M. A., and Peña-Duque, M. A. (2018). CETP and LCAT Gene Polymorphisms Are Associated with High-Density Lipoprotein Subclasses and Acute Coronary Syndrome. *Lipids* 53, 157–166. doi:10.1002/lipd.12017
- Warnick, G. R., Knopp, R. H., Fitzpatrick, V., and Branson, L. (1990). Estimating Low-Density Lipoprotein Cholesterol by the Friedewald Equation Is Adequate for Classifying Patients on the Basis of Nationally Recommended Cutpoints. *Clin. Chem.* 36, 15–19. doi:10.1093/clinchem/36.1.15
- Weissglas-Volkov, D., and Pajukanta, P. (2010). Genetic Causes of High and Low Serum HDL-Cholesterol. *J. Lipid Res.* 51, 2032–2057. doi:10.1194/jlr.R004739
- Wietlisbach, V., Paccaud, F., Rickenbach, M., and Gutzwiller, F. (1997). Trends in Cardiovascular Risk Factors (1984-1993) in a Swiss Region: Results of Three Population Surveys. *Prev. Med.* 26, 523–533. doi:10.1006/pmed.1997.0167
- Willer, C. J., Li, Y., and Abecasis, G. R. (2010). METAL: Fast and Efficient Meta-Analysis of Genomewide Association Scans. *Bioinformatics* 26, 2190–2191. doi:10.1093/bioinformatics/btq340
- Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., et al. (2013). Discovery and Refinement of Loci Associated with Lipid Levels. *Nat. Genet.* 45, 1274–1285. doi:10.1038/ng.2797
- Williams Amy, A. L., Jacobs Suzanne, S. B. R., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., et al. (2014). Sequence Variants in SLC16A11 Are a Common Risk Factor for Type 2 Diabetes in Mexico. *Nature* 506, 97–101. doi:10.1038/nature12828
- Wu, Y., Waite, L. L., Jackson, A. U., Sheu, W. H. H., Buyske, S., Absher, D., et al. (2013). Trans-Ethnic Fine-Mapping of Lipid Loci Identifies Population-specific Signals and Allelic Heterogeneity that Increases the Trait Variance Explained. *PLoS Genet.* 9, e1003379. doi:10.1371/journal.pgen.1003379
- Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E. V., Mort, M., et al. (2012). Deleterious- and Disease-Allele Prevalence in Healthy Individuals: Insights from Current Predictions, Mutation Databases, and Population-Scale Resequencing. *Am. J. Hum. Genet.* 91, 1022–1032. doi:10.1016/j.ajhg.2012.10.015
- Zhang, L., You, Y., Wu, Y., Zhang, Y., Wang, M., Song, Y., et al. (2017). Association of BUD13 Polymorphisms with Metabolic Syndrome in Chinese Population: a Case-Control Study. *Lipids Health Dis.* 16, 127. doi:10.1186/s12944-017-0520-8
- Zhou, L., He, M., Mo, Z., Wu, C., Yang, H., Yu, D., et al. (2013). A Genome Wide Association Study Identifies Common Variants Associated with Lipid Levels in the Chinese Population. *PLoS ONE* 8. doi:10.1371/journal.pone.0082420

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jurado-Camacho, Cid-Soto, Barajas-Olmos, García-Ortiz, Baca-Peynado, Martínez-Hernández, Centeno-Cruz, Contreras-Cubas, González-Villalpando, Saldaña-Álvarez, Salas-Martínez, Mendoza-Caamal, González-Villalpando, Córdova and Orozco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.