Variant rs9644568 in the intergenic region downstream of the LPL gene is associated with high LDL cholesterol levels among Filipinos.

Eva Maria C. Cutiongco–de la Paz1,2,*, Jose B. Nevado, Jr.1, Lourdes Ella G. Santos3,

Aimee Yvonne Criselle L. Aman1,2, Jose Donato A. Magno3, Paul Ferdinand M. Reganit3, Richard Henry P. Tiongco II3, Deborah Ignacia D. Ona3, Felix Eduardo R. Punzalan3, Elmer Jasper B. Llanes3, Jaime Alfonso M. Aherrera3, Carla Mae O. Fernandez3, Lauro L. Abrahan IV3, Charlene F. Agustin3, Adrian John P. Bejarin1,2, and Rody G. Sy3

1Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila, Philippines
2Philippine Genome Center, University of the Philippines Diliman, Quezon City, Philippines
3Department of Medicine, Philippine General Hospital, University of the Philippines Manila, Philippines

*corresponding author

E-mail: eccutiongodelapaz@up.edu.ph

†These authors contributed equally to this work

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

High blood level of low-density lipoprotein cholesterol (LDL-C) is a major risk factor for cardiovascular disease. Although genetic variants linked to high LDL-C have been studied in other populations, there have been no previous studies among Filipinos. This study aims to determine the association of candidate genetic variants to high LDL-C. We performed an age- and sex-matched case-control study that compared Filipino participants with high LDL-C levels (n=60) with controls (n=60). DNA was extracted from blood samples and genotyped for candidate SNPs using a customized microarray chip. Logistic regression analyses were used to determine the composite association of genetic and clinical variables to the condition. Of the initial eleven SNPs associated with high LDL-C in univariate analyses, only the variant rs9644568 in the intergenic region downstream of the \textit{LPL} gene remained significantly associated with high LDL-C levels on multiple regression analysis and variable selection after adjustment for hypertension. The G allele was observed as the risk allele in a recessive model. The variant rs9644568-G in the \textit{LPL} gene was associated with high blood LDL-C levels among Filipinos. In combination with hypertension, this genetic profile may identify individuals who are susceptible to develop high LDL-C in this population.

* Corresponding author

E-mail: eccutiongodelapaz@up.edu.ph
Introduction

Low-density lipoprotein cholesterol (LDL-C) level is a risk factor and a traditional target for preventive cardiovascular disease (CVD) interventions. This association is enhanced in the presence of atherosclerosis [1], and the common view is that it is an active component of atherogenesis [2]. Several studies have since reported reduced adverse cardiovascular events after treatment with LDL-C-lowering drugs [2-4].

LDL-C levels may be affected by several factors, including genetics. A common familial form of LDL-C elevation, familial hypercholesterolemia (FH), has been well studied and three causative genes have already been identified, namely low-density lipoprotein receptor (LDLR), apolipoprotein B (APOB), and proprotein convertase subtilisin/kexin type 9 (PCSK9). In addition, several studies on hereditary dyslipidemias have shown a variety of genes involved, such as low density lipoprotein receptor adaptor protein 1 (LDLRAP1), ATP binding cassette subfamily G members 5 and 8 (ABCG5, ABCG8), angiopoietin like 3 (ANGPTL3), microsomal triglyceride transfer protein (MTTP), myosin regulatory light chain interacting protein (MYLIP), apolipoproteins A5, B, C2 and E (APOA5, APOB, APOC2, APOE), and lipoprotein lipase (LPL), which as expected, are involved in lipid metabolism, transport and processing [5,6].

Genotyping efforts in recent years attempted to find possible genetic susceptibility loci. For instance, genome-wide association studies (GWAS) have implicated LDLR, APOE, and 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR) as possible key regulators of lipid metabolism to high LDL-C [7-9]. Other genes that were considered include fatty acid desaturase 1 (FADS1), polypeptide N-acetylgalactosaminyltransferase 2 (GALNT2), and acetyl-coenzyme A carboxylase beta (ACACB) [10-12].

However, findings from studies of genetic polymorphisms cannot be generalized across populations due to inherent inter-ethnic differences in genetic profiles. To illustrate, among the Taiwanese population, a PCSK9 variant has been determined to have a causally significant association with high LDL-C levels; however, this association is not found among the Italian population [13,14]. In another study, a novel polymorphism of the ICAM1 was reported to be
associated with LDL-C in African Americans at conventional genome-wide significance, but not among Hispanics and East Asians [15]. Such discrepancies warrant verifying associations between genotypes and LDL-C levels in newly investigated populations. This could be due to independently evolved genetic differences with similar phenotypes due to metabolic proximity. Data from different ethnic populations will be able to contribute to a better understanding of diseases and can enhance customized clinical applications for individual populations.

Due to scarce data on genetic associations with LDL-C among Filipinos, this study aimed to identify candidate genetic markers associated with high LDL-C levels in this population. In addition to direct association, the variants were interpreted in the context of traditional risk factors to account for their significant, but potentially interdependent influences.

**Methodology**

**Study design and ethics clearance**

This is an age- and sex-matched case-control study performed from July 9, 2013 to March 31, 2017. This study complied to the Principles of the Declaration of Helsinki (2013) and all experimental protocols in this study were approved by the University of the Philippines Manila-Research Ethics Board (Study protocol code UPMREB-2012-0186-NIH). The investigators confirm that the methods complied with the relevant guidelines and regulations of UPMREB.

**Enrollment of participants and clinical data collection**

Participants were enrolled from the Philippine General Hospital, different local communities, and clinics in Metro Manila. They were unrelated, aged 18 years old and above, of Filipino descent up to the 3rd degree of consanguinity, and could provide independent consent. Cases were defined as (a) having at least one reading of serum LDL-C ≥160 mg/dl as defined by the NCEP ATP III guidelines; and (b) had not ever been on antilipidemics. Controls were participants with (a) at least one reading of serum LDL-C <160 mg/dl; and (b) without a previous diagnosis of dyslipidemia. Participants were excluded if they had end-stage renal disease, decompensated chronic liver disease, active malignancy, chronic lung disease, secondary hypertension, secondary dyslipidemia, or if pregnant. The demographic data and clinical characteristics were obtained from patient records and
verbal interviews. Clinical data, including lipid profile and serum creatinine, were also recorded. Written informed consent forms duly approved by the UPMREB was obtained from potential participants prior to recruitment to the study.

**DNA extraction and quantification**

DNA was extracted using the QiaAmp DNA Mini Kit following the spin protocol for blood buffy coat as specified in the manufacturer’s instruction manual. DNA was quantified using a spectrometer at 260nm and stored at -20°C until use. All DNA samples had A-260nm/A-280nm values above 1.80.

**Customized genotyping**

A customized GoldenGate Genotyping (GGGT) beadchip (Illumina, Inc., San Diego, CA, USA) was designed in 2012 using candidate SNPs from both coding and noncoding regions, including intergenic and intronic SNPs, which have shown evidence of association with high LDL-C levels and lipid metabolism. These were selected after an extensive search in the following databases: Pharmacogenomics Knowledgebase database (PharmGKB), National Human Genome Research Institute Genome-Wide Association Study Catalog (NHGRI GWAS Catalog), PubMed, and selected international patent databases (e.g., Patentscope and Espacenet). The selected SNPs were submitted to Illumina, Inc. for scoring to determine the suitability of the SNPs to discriminate for microarray assay and to estimate their specificity.

Customized genotyping of candidate SNPs was performed using DNA microarray technology following the GoldenGate Genotyping (GGGT) protocol specified in the manufacturer’s manual. While the microarray platform is optimally designed to detect bi-allelic SNPs, the study included tri- or quad-allelic SNPs using the reference allele and the most common alternative in cases where the variants correlate with significant clinical outcomes. After microarray processing, the beadchip was imaged on the HiScan System, and data from these images were analyzed using GenomeStudio software.
Data analysis

Quality control

GenomeStudio 2.0 and PLINK version 2.0.10 were used for data cleanup using the following criteria: call rates $\geq 95\%$, minor allele frequency (MAF) of $\geq 0.01$, genotype missingness of $\leq 0.05$, individual missingness $\leq 0.05$, and not having a Hardy-Weinberg equilibrium (HWE) of $p < 0.001$ in the control group.

Statistical analysis

Chi-squared test and Fisher exact test, as applicable, were used for categorical data, and the student $t$-test was used for quantitative data were performed to assess for significant differences between alleles and genotypes (i.e., allelic and genotypic association tests, which include additive, dominant and recessive models). These models were identified based on the distribution of the case and control genotypes. Tests for allelic and genotypic associations were done by setting the cut-off at Bonferroni-corrected $p < 0.05$.

Upon determination of genotypic effects, the genotypes were recorded according to the most significant model, with imputations made for missing data. Univariate analysis was done to determine the association of SNPs with high LDL-C phenotype. A multiple logistic regression using stepwise backward elimination was performed with nominal $p \leq 0.05$ and clinical variable with $p \leq 0.20$.

Results

After data quality control, the final set of samples for analysis included 120 participants consisting of 60 cases and 60 controls (Fig 1A). Using age- and sex-matched case-control design, the 120 participants were genotyped using 132 genomic variants.

Of the 132 variants selected for the study, 50 variants were removed after quality control with PLINK: 33 SNPs failed the Hardy-Weinberg test ($p > 0.001$) among controls, ten SNPs had missing genotype data for more than 5% of the individuals, and one SNP was removed because the minor allele frequency was lower than 1% (Fig 1B). Eleven variants were significantly associated with high LDL-C after adjusting for multiple testing (i.e., Bonferroni-adjusted $\alpha = 0.05$).
FIG 1. Schematic diagram of data processing and analysis for the high LDL-C group. A total of 183 participants (A) and 132 SNPs (B) were analyzed to determine the association of genetic variants with high LDL-C and lipid metabolism. Abbreviations: MIND – individual missingness; SNP – single nucleotide polymorphism; HWE – Hardy-Weinberg equilibrium; GENO – genotypic missingness.

Table 1 summarizes the clinical characteristics of the included participants. Their crude odds ratios are in Table S1. Participants in the high LDL-C group tend to have more comorbidities that showed significance such as hypertension ($p < 0.0001$), BMI $\geq 25$ kg/m$^2$ ($p = 0.056$), and ischemic heart disease ($p = 0.002$). The cases also seemed to have slightly elevated creatinine levels.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants with high LDL-C (n = 60)</th>
<th>Participants with normal LDL-C (n = 60)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ≥ 60 years, %</td>
<td>40.00</td>
<td>40.00</td>
<td>1.000</td>
</tr>
<tr>
<td>Sex, % male</td>
<td>38.33</td>
<td>38.33</td>
<td>1.000</td>
</tr>
<tr>
<td>Co-morbidities, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>86.67</td>
<td>28.33</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>BMI $\geq 25$ kg/m$^2$</td>
<td>43.33</td>
<td>26.67</td>
<td>0.056</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>20.00</td>
<td>0.00</td>
<td>&gt; 0.999</td>
</tr>
<tr>
<td>Previous stroke</td>
<td>1.67</td>
<td>1.67</td>
<td>1.000</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>45.00</td>
<td>13.33</td>
<td>0.002</td>
</tr>
<tr>
<td>eGFR &lt; 60 ml/min</td>
<td>10.00</td>
<td>5.00</td>
<td>0.215</td>
</tr>
<tr>
<td>Lifestyle factors, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker (both current and past)</td>
<td>28.33</td>
<td>25.00</td>
<td>0.638</td>
</tr>
<tr>
<td>Alcohol drinker (&gt;2 shots per day)</td>
<td>50.00</td>
<td>46.67</td>
<td>0.618</td>
</tr>
<tr>
<td>Blood chemistry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine, mg/dl</td>
<td>0.90 (0.42)</td>
<td>0.80 (0.19)</td>
<td>0.094</td>
</tr>
<tr>
<td>Total cholesterol, mg/dl</td>
<td>273.79 (44.54)</td>
<td>193.01 (25.86)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Triglycerides, mg/dl</td>
<td>159.47 (68.76)</td>
<td>91.70 (36.67)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>HDL, mg/dl</td>
<td>52.46 (14.19)</td>
<td>56.94 (13.50)</td>
<td>0.079</td>
</tr>
<tr>
<td>LDL, mg/dl</td>
<td>191.88 (41.34)</td>
<td>120.62 (21.87)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>eGFR, ml/min</td>
<td>87.72 (24.42)</td>
<td>93.33 (18.77)</td>
<td>0.161</td>
</tr>
</tbody>
</table>
Abbreviations: SD – standard deviation; BMI – body mass index; eGFR – estimated glomerular filtration rate; HDL – high-density lipoprotein; LDL – low-density lipoprotein

*Significance set at $p < 0.05$ using Chi-squared test for age and sex, conditional logistic regression for comorbidities and lifestyle factors, and Student’s $t$ test for continuous variables.

There were 11 variants found to have a statistically significant association with LDL-C levels on the genotypic association test after adjusting for multiple testing (Bonferroni-adjusted $\alpha = 0.00061$) (Table S2). These variants underwent logistic regression analysis for the computation of their OR and to determine which will retain significance upon adjusting for other variables (Table 2). One SNP had clinically significant odds ratios (OR $\geq 2.5$) and was statistically significant ($p < .001$) on simple logistic regression analysis: rs12279250 in NELL1. This SNP seemed to display a dominant genotypic effect.

Clinical variables with a significant difference between the two groups, $p \leq 0.20$: hypertension, BMI $\geq 25$ kg/m$^2$, and ischemic heart disease were included together with the 11 variants with nominal $p < 0.05$ in the full model for the multiple regression analysis (Table 3). On variable selection using the stepwise algorithm ($p < 0.05$), only the presence of hypertension and rs9644568 retained their significant association with high LDL-C levels under an additive model with allele G as the risk allele (AG vs. GG: adjusted OR = 2.44, 95% CI = 0.41-20.60; $p = 0.355$; AA vs GG: adjusted OR = 5.33, 95% CI = 1.88, 18.47; $p = 0.003$).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Implicated Gene</th>
<th>Variant Role</th>
<th>Risk Allele</th>
<th>Frequency</th>
<th>Genotype</th>
<th>Crude Odds Ratio (95% CI)</th>
<th>$p$-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9644568</td>
<td>8</td>
<td>LPL</td>
<td>Intergenic (regulates LPL as by GTex Portal)</td>
<td>G</td>
<td>0.60</td>
<td>AG vs AA</td>
<td>2.36 (0.50, 17.27)</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GG vs AA</td>
<td>13.00 (3.27, 87.36)</td>
<td>0.001</td>
</tr>
<tr>
<td>rs1045642</td>
<td>7</td>
<td>ABCB1</td>
<td>synonymous variant</td>
<td>A</td>
<td>0.325</td>
<td>AG vs AA</td>
<td>2.43 (1.03, 5.94)</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GG vs AA</td>
<td>7.33 (2.56, 23.07)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>rs12279250</td>
<td>11</td>
<td>NELL1</td>
<td>intron variant</td>
<td>G</td>
<td>0.54</td>
<td>GG vs AG/AA</td>
<td>2.20 (1.01, 4.96)</td>
<td>0.051</td>
</tr>
<tr>
<td>rs1532085</td>
<td>15</td>
<td>LIPC</td>
<td>unknown</td>
<td>A</td>
<td>0.20</td>
<td>AG vs GG</td>
<td>2.44 (1.05, 5.80)</td>
<td>0.039</td>
</tr>
</tbody>
</table>

GG vs AA: adjusted OR = 10.66, 95% CI = 2.15-81.05; $p = 0.008$ (Table 4).
<table>
<thead>
<tr>
<th>Factors</th>
<th>Genotype</th>
<th>Adjusted OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td></td>
<td>2.86 (1.25, 6.73)</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.72 (1.71, 14.22)</td>
<td></td>
</tr>
<tr>
<td>BMI ≥ 25 kg/m²</td>
<td></td>
<td>0.75 (0.17, 3.56)</td>
<td>0.706</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.41 (1.02, 13.43)</td>
<td></td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td></td>
<td>0.37 (0.06, 1.95)</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.79 (2.38, 21.20)</td>
<td></td>
</tr>
<tr>
<td>rs9644568 AG vs AA</td>
<td>G</td>
<td>1.92 (0.17, 33.73)</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38.82 (2.08, 1758.82)</td>
<td>0.030</td>
</tr>
<tr>
<td>rs1045642 AG vs AA</td>
<td>G</td>
<td>4.96 (0.75, 40.90)</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.82 (1.53, 358.85)</td>
<td>0.032</td>
</tr>
<tr>
<td>rs12279250 AA vs AG/GG</td>
<td></td>
<td>2.37 (0.35, 15.77)</td>
<td>0.359</td>
</tr>
<tr>
<td>rs1532085 AG vs GG</td>
<td></td>
<td>0.66 (0.07, 5.18)</td>
<td>0.697</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.90 (0.46, 40.06)</td>
<td>0.221</td>
</tr>
<tr>
<td>rs4149268 AG vs AA</td>
<td>G</td>
<td>4.89 (1.00, 27.94)</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.36 (2.73, 270.92)</td>
<td>0.007</td>
</tr>
<tr>
<td>rs4299376 AG vs GG</td>
<td></td>
<td>0.25 (0.01, 3.89)</td>
<td>0.333</td>
</tr>
</tbody>
</table>

**TABLE 3. Genetic variants and clinical factors in the multiple regression full model.**

- *Significance set at p < 0.05 on conditional logistic regression analysis*
<table>
<thead>
<tr>
<th>Variable</th>
<th>Genotype</th>
<th>Adjusted OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9644568</td>
<td>AG vs AA</td>
<td>2.44 (0.41, 20.60)</td>
<td>0.355</td>
</tr>
<tr>
<td></td>
<td>GG vs AA</td>
<td>10.66 (2.15, 81.05)</td>
<td>0.008</td>
</tr>
<tr>
<td>Hypertension</td>
<td>-</td>
<td>14.23 (5.53, 40.81)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

BMI = body mass index; CI = confidence interval; OR = odds ratio

**TABLE 4. Genetic variants and clinical factors remaining after variable selection.**

**Discussion**

Genetic determinants of high LDL-C may be helpful in the understanding and optimum management of cardiovascular risk. As associated variations often differ across populations, especially for complex conditions, studies on Filipinos must be undertaken considering that cardiovascular disease remains as the top cause of mortality in the country. This study analyzes candidate genetic variants of high LDL-C among Filipino patients. In addition, a regression model of candidate genetic variants incorporating traditional clinical risk factors was generated to assess their relative effects on LDL-C profiles.

Univariate analyses found 11 variants associated with LDL-C levels. The reported genes include lipoprotein lipase (LPL), ATP binding cassette subfamily B member (ABCB1), lipase C (LIPC), ATP binding cassette subfamily A member (ABCA1), ATP binding cassette subfamily G member 8 (ABCG8), apolipoprotein B (APOB), and cadherin EGF LAG seven-pass G-type receptor 2 (CELSR2). Expectedly, these are known genes whose roles in LDL-C metabolism are well established — from lipid transport, cellular metabolism and homeostasis, cellular translocation, and storage [16-21].
A major finding in this study is the variant rs9644568 near the \textit{LPL} gene. Logistic regression with adjustment for hypertension shows a significant association of rs9644568 to high LDL-C levels after adjusting for hypertension among Filipinos. The G allele of rs9644568 confers a 2.44-fold increased risk for high LDL-C in heterozygous individuals and a 10.66-fold increased risk in homozygous individuals. This suggests a possible additive or a multiplicative model of transmission. This is similar to the findings of a genome-wide association study among Europeans, which consistently demonstrated the significant association of G allele of rs4977574 with high LDL-C levels and an increased risk of myocardial infarction (MI) [22]. In contrast, the minor G allele protects LDL-C in another study that included Hispanics, East Asians and also Europeans [23]. A similar study found the minor G allele to be associated with decreased risk for coronary artery disease and peripheral artery disease in a Chinese population with diabetes [24].

The SNP rs9644568 is located on chromosome 8p21.3 in an intergenic region downstream of the gene \textit{LPL} (UCSC Genome Browser, \url{http://genome.ucsc.edu}. Accessed October 10, 2023). Although the variant is >100 kb far from the gene LPL, expression from several tissues suggests that it regulates the gene (GTEx Portal. \url{https://gtexportal.org}. Accessed October 10, 2023).

This gene is predominantly expressed in the heart, skeletal muscles, and adipose tissues\textsuperscript{25}. \textit{LPL} encodes a rate-limiting enzyme called lipoprotein lipase, which plays an important role in lipid metabolism, as well as in lipid intake and clearance. Lipoprotein lipase aids in the breakdown of the circulating triglyceride-rich lipoproteins in the bloodstream and the uptake of free fatty acids into peripheral organs for energy storage and consumption [26-30]. Alterations in LPL activity could result in dyslipidemia, characterized by increased plasma LDL-C, TG, and reduced plasma HDL-C levels [31]. This association of the variant is interesting, as prior data, to our knowledge, do not indicate rs9644568 as a risk marker for high LDL levels or as having the highest odds ratio in any non-Filipino population. This can imply that the variant is a unique risk marker for Filipinos, independent of other components of metabolic syndrome, i.e., hypertension and diabetes mellitus.

To support the significance of \textit{LPL} as a candidate locus, the study also indicated that another downstream \textit{LPL} variant, rs7016880, was associated with LDL-C. Interestingly, rs7016880 shares
many characteristics with rs9644568 – being located as a downstream non-coding regulator, having a possible multiplicative model, with a high odds ratio. This may provide greater confidence for the veracity of the findings in this study.

As for the other variants, the adjustments due to hypertension and the rs9644568 seem to mask their statistical significance. This has several implications. First, rs9644568 is rarely reported as a marker for LDL-C in the literature, but the statistical significance of rs9644568 stood out prominently in the current study, thus implying that the SNP is unique among the Filipino population. Second, hypertension might have masked the statistical significance of some of the candidate genes in the metabolic syndrome context. This is unsurprising as some variants have been previously associated with high blood pressure. To illustrate, ABCBI was reported to be associated with hypertension through dietary salt transport and retention [32]. In a study among Han Chinese, variants of ABCAI, rs2472510 and rs2515614, were associated with hypertension, with decreased expression of ABCAI linked with hypertension risk [33]. On the other hand, meta-analysis of 2799 cases and 6794 controls showed an association of CYP3A5 rs776746 with hypertension [34]. A later study on a small sample of Han Chinese also reported a significant association of CYP3A5 rs776746 with the elevated risk of hypertension [35].

Several variants are mostly known genes whose roles in LDL-C metabolism are well established [16-21]. The G allele of rs1532085 in LIPC with a protective effect for LDL-C levels in Chinese populations [36], is similar to the findings in this study. In a large GWAS study involving >100,000 Caucasians, rs4299376 in ABCG8 was found to be one of the strongly associated genes with LDL-C [38]. The variant rs676210 in APOB has been identified as a susceptibility locus for very low-density lipoprotein levels in the Women's Genome Health Study (WGHS) involving 17,296 women of European ancestry [38]. This association was further reported in the Chinese Yugur population[39]. A GWAS study of 11,685 showed a significant association of CELSR2 rs660240 to LDL-C levels where the G allele is associated with risk and this was confirmed in a meta-analysis of Below et al. (2016) [40,41].

However, some genes can be considered as attractive at best at this point. One intriguing
finding is the direct association of ABCB1 with LDL-C. There is scarce data on this association, as most studies focused on its association with response to statins [42,43]. The variant ABCB1 rs1045642 is by itself associated with statin response [44]. The variant rs12279250 in NELL1, a growth factor of still uncertain significance, has been associated with increased triglyceride levels [45,46], which may lead to a subsequent increase in LDL-C levels. Despite the study identifying rs4149268 in ABCA1 as associated with LDL-C, most studies reported an association with HDL-C levels instead of LDL-C [47]. This may need to be verified, but it is noteworthy to investigate if such an association remains among Filipinos. As for rs776746 in CYP3A5, there is little study on the direct association with LDL-C levels. This could be due to the assumption that cytochrome P proteins are mainly associated with drugs. In fact, most studies focused on response to cholesterol-lowering drugs [48,49]. As the present study focused on participants without drug intake, a possible role for CYP3A5 on LDL-C level could be plausible. Similarly, there is minimal information on rs2686586. Even its association with the gene GOLIM4 is questionable, as no evidence of regulation nor proximity (>50kb) is noted in the literature. Nonetheless, the fact that it is significantly associated with LDL-C needs to be verified.

In this study, clinical variables that show a significant association with high LDL-C phenotype are hypertension, BMI $\geq 25$ kg/m$^2$, and ischemic heart disease (IHD). Several studies consistently show strong correlations between these comorbidities and LDL-C levels. The link between LDL-C and hypertension was elucidated by the group of Laaksonen et al. (2008). They reported that elevated serum levels of LDL-C increase the risk of hypertension in Finnish men, which explained the strong association of dyslipidemia with incident hypertension [50]. Meanwhile, a previous study performed in the US population suggested a linear association of LDL-C concentrations and BMI [51]. The association between high LDL-C levels and high IHD risk is further substantiated by the study of Breuer (2005), focusing on improving cardiovascular disease outcomes by lowering LDL-C levels [52]. Moreover, dyslipidemia and hypertension are known as major risk factors for cardiovascular diseases; with these two comorbidities reported to be frequently diagnosed in patients with ischemic heart disease [53].
This study presents an opportunity to identify individuals at risk for developing high LDL-C levels by looking at their genetic profiles to complement clinical correlates. Estimated associations between variant rs9644568 and hypertension were very high, but the confidence intervals around them were too wide to rule out the possibility of substantial associations. Further studies with larger sample sizes are required to validate the conclusions of this study.

Conclusion

Genetic determinants largely influence the development of high LDL-C levels. This study found an association of variant rs9644568 in the intergenic region downstream of the *LPL* gene, with high LDL-C levels among Filipinos. Adding a genetic component may improve screening and assessment of patients at risk for developing high LDL-C levels.
References


Acknowledgements

The authors would like to express their sincere gratitude to the administrative staff and researchers of the Microarray Laboratory of the Institute of Human Genetics, National Institutes of Health-University of the Philippines Manila, the Philippine Genome Center, and the clinical fellows of the Department of Medicine for actively being involved in the project. We are extending our gratitude to our sponsor, the Philippine Council for Health Research and Development of the Department of Science and Technology (PCHRD-DOST).

Author Contributions

EMCutongco-De La Paz, conceptualized the study and designed the methodology, acquired funding and led the project administration and supervision, conducted the research and investigation process, performed the data curation and analysis, interpreted the data, wrote the original draft, reviewed and edited the manuscript for intellectual content; JBNevado conceptualized the study and designed the methodology, acquired funding and led the project administration and supervision, conducted the research and investigation process, performed the data curation and analysis, interpreted the data, wrote the original draft, reviewed and edited the manuscript for intellectual content; LEGSantos, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; AYCLAman, acquired funding and led the project administration and supervision, conducted the research and investigation process, performed the data curation and analysis, interpreted the data, wrote the original draft; JDAMagno, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; PMMReganit, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; RHPTiongco, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; DIDOna, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; FERPunzalan, conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; EJBLlanes,
conceptualized the study and designed the methodology, conducted the research and investigation process, interpreted the data; JMAherrera, conducted the research and investigation process, interpreted the data; CMOFernandez, performed the data curation and analysis, interpreted the data, reviewed and edited the manuscript for intellectual content; LLAbraham, conducted the research and investigation process, interpreted the data; CFAgustin, conducted the research and investigation process, interpreted the data; AJPBejarin, conducted the research and investigation process, performed the data curation and analysis, interpreted the data, wrote the original draft; RGSy, conceptualized the study and designed the methodology, acquired funding and led the project administration and supervision, conducted the research and investigation process, interpreted the data, reviewed and edited the manuscript for intellectual content.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare no competing interests.

Funding Source

The present study was sponsored by a grant-in-aid from the PCHRD-DOST.