Meta-analysis of 49,549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in ANGPTL4 determining fasting TG levels

Mary F. Feitosa  
*Washington University School of Medicine in St. Louis*

Ingrid B. Borecki  
*Washington University School of Medicine in St. Louis*

L. Adrienne Cupples  
*Washington University School of Medicine in St. Louis*

et al

Follow this and additional works at: [https://digitalcommons.wustl.edu/open_access_pubs](https://digitalcommons.wustl.edu/open_access_pubs)

**Recommended Citation**

Feitosa, Mary F.; Borecki, Ingrid B.; Cupples, L. Adrienne; and et al., "Meta-analysis of 49,549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in ANGPTL4 determining fasting TG levels." *Journal of Medical Genetics*. 53, 441-449. (2016).  
[https://digitalcommons.wustl.edu/open_access_pubs/6320](https://digitalcommons.wustl.edu/open_access_pubs/6320)
SHORT REPORT

Meta-analysis of 49 549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in ANGPTL4 determining fasting TG levels

Elisabeth M van Leeuwen,1 Aniko Sabo,2 Joshua C Bis,3 Jennifer E Huffman,4,5 Ani Manichaikul,6 Albert V Smith,7,8 Mary F Feitosa,9 Serkalem Demissie,10 Peter K Joshi,11 Qing Duan,12 Jonathan Marten,12 Jan B van Klinken,13 Ida Surakka,14 Stella Trompet,20,21 Evangelos Evangelou,16,23 Leo-Pekka Lyytikäinen,24,25 Bamidele O Tayo,26 Joris Deelen,27 Peter J van der Most,15 Sander W van der Laan,28 Dan E Arking,29 Alanna Morrison,30 Abbas Dehghan,1 Oscar H Franco,26 Mika Kähönen,55,56 Terho Lehtimäki,24,25 Paul Elliott,57 Pim van der Harst,22,58 J Wouter Jukema,20 Dennis O Mook-Kanamori,19,59,60 Dorret I Boomsma,18 John C Chambers,16,17,42 Morris Swertz,58,61 Samuli Ripatti,39,40,41,63 Veronique Vitart,4 Ozren Polasek,39 Caroline Hayward,4 James G Wilson,65 James F Wilson,4,11 Vilmundur Gudnason,7,8 Stephen S Rich,6 Bruce M Psaty,3,66,67,68 Ingrid B Borecki,9 Eric Boerwinkle,2,30 Jerome I Rotter,69,70,71 L Adrienne Cupples,5,9 Cornelia M van Duijn1

ABSTRACT

Background So far, more than 170 loci have been associated with circulating lipid levels through genome-wide association studies (GWAS). These associations are largely driven by common variants, their function is often not known, and many are likely to be markers for the causal variants. In this study we aimed to identify more new rare and low-frequency functional variants associated with circulating lipid levels.

Methods We used the 1000 Genomes Project as a reference panel for the imputations of GWAS data from ∼60 000 individuals in the discovery stage and ∼90 000 samples in the replication stage.

Results Our study resulted in the identification of five new associations with circulating lipid levels at four loci.

All four loci are within genes that can be linked biologically to lipid metabolism. One of the variants, rs116843064, is a damaging missense variant within the ANGPTL4 gene.

Conclusions This study illustrates that GWAS with high-scale imputation may still help us unravel the biological mechanism behind circulating lipid levels.

INTRODUCTION

Genome-wide association studies (GWAS) for circulating lipid levels (high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC) and triglycerides (TG)) have identified over 170 loci. These...
studies have been based on imputations to the HapMap reference panel or primary versions of the 1000 Genomes Project (1kG) or genotyping on the Illumina Exome Chip. None has used imputations with the Phase 1 integrated release v3 of the 1kG which allows the imputation of rare and low-frequency functional variants and structural variations with more precision. Evidence of rare and low-frequency functional variants associated with circulating lipid levels comes from recent studies in which exome sequencing of the NPC1L1 gene identified rare variants associated with reduced LDL-C levels and reduced risk of coronary heart disease. Moreover, exome sequencing of LDLR and APOA5 identified rare variants associated with an increased LDL-C and increased TG levels and exome sequencing of APOC3 identified rare variants associated with reduced TG levels and reduced risk of coronary heart disease.

Our goal in this study was to identify rare and low-frequency functional variants associated with circulating lipid levels in a larger sample size compared with the exome sequencing of candidate gene approach. To this end, we imputed genotypes for study samples participating in the cohorts of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium using the Phase 1 integrated release v.3 of the 1kG and conducted a meta-analysis of about approximately 60,000 individuals, followed by a replication in an independent set of 90,000 individuals.

METHODS
Please see online supplementary methods for complete descriptions of the methods. In summary, for the discovery stage of this project, we used the data from 20 cohorts of the CHARGE consortium (see online supplementary methods). All cohorts were imputed with reference to the 1kG reference panel (version Phase 1 integrated release V3). The total number of individuals in the discovery stage was 59,409 for HDL-C, 48,780 for LDL-C, 60,024 for TC and 49,549 for TG. Online supplementary tables S1 and S2 contain the baseline characteristics per cohort and more details about SNP genotyping and genotype imputations. Within each cohort, each variant was tested for association with each of the lipid traits, assuming an additive genetic model. The association results of all cohorts for all variants were combined using inverse variance weighting. We used the following filters for the variants: 0.3 < R² (measurement for the imputation quality) ≤ 1.0 and expected minor allele count (expMA ≥ 2 × MAF (minor allele frequency) × R² × sample size) > 10 prior to meta-analysis. After meta-analysis of all available variants, we excluded the variants that were not present in at least four cohorts, to prevent false positive findings. In order to select only variants that were independently associated with each of the lipid traits, we used the genome-wide complex trait analysis (GCTA) tool, V1.13. To identify novel loci we selected from the list of variants identified by GCTA, those variants which are within 0.5 Mb of a loci previously published by Teslovich et al. which resulted in a sample size of about 80% (see online supplementary figure S5 and supplementary table S3). We also significantly replicated 43.4% of the loci described by the Global Lipids Genetics Consortium (GLGC) despite a sample size of about 30% (see online supplementary figure S6 and supplementary table S4).

A conditional and joint analysis using GCTA identified 185 independent variants for HDL-C, 174 for LDL-C, 214 for TC and 119 for TG. Next, we excluded all variants that were not genome-wide significant (p value ≤ 5 × 10⁻⁸) in the initial discovery stage, which resulted in 56 variants for HDL-C, 50 for LDL-C, 66 for TC and 37 for TG. And we excluded all variants which are within 0.5 Mb of a loci previously published by Teslovich et al. or GLGC, which resulted in three variants for HDL-C, three for LDL-C, seven for TC and six for TG. These variants are located at 17 different loci and include one deletion (figure 1 and table 1). These 19 variants were selected for replication. The total number of individuals in the replication stage was 84,598, 72,486, 83,739 and 73,519 for HDL-C, LDL-C, TC and TG, respectively (see online supplementary tables S1 and S2 for baseline characteristics and information about SNP genotyping and imputation details). The sample size in the replication stage was larger than the initial discovery sample for 17 out of the 19 variants. The frequencies of the variants were similar between the discovery and replication cohorts. The directions of effect were the same in the discovery and replication cohorts for 16 out of the 19 variants (see online supplementary figure S7). We used a Bonferroni corrected threshold for significance (p value ≤ 2.63 × 10⁻⁸). Five out of the 19 variants were significantly replicated (table 1): rs6457374 (TC), rs186696265 (LDL-C and TC), rs77697917 (HDL-C) and rs116843064 (TG). The frequency of these variants ranged between 0.012 and 0.249 within the discovery sample. Online supplementary table S5 shows the heterogeneity for the 19 variants after the meta-analysis of all discovery cohorts and of all replication cohorts. We also meta-analysed all variants in the individuals of the discovery cohorts and replication cohorts combined (table 1 and see online supplementary tables S5 and S6) and per ethnicity (see online supplementary table S6) using a fixed-effect meta-analysis approach. We found that the five significantly replicated variants we identified in this study are only significant within the European samples, thereby noticing that there are much more European samples in this study, compared with the African and Asian samples. When using a random-effect meta-analysis to account for the multiple ethnicities in our sample (see online supplementary table S7), we found that of the five replicated variants, one attained genome-wide significance (p value ≤ 5 × 10⁻⁸) and the other four nominal significance (p value ≤ 0.05).

DISCUSSION
We conducted a GWAS that included GWAS data imputed to the 1kG to identify rare and low-frequency, potentially functional, variants associated with circulating lipid levels. To this end, we imputed genotypes in approximately 60,000 individuals from 20 cohorts in the CHARGE consortium with the 1kG
reference panel. The meta-analysis, followed by GCTA analysis revealed 19 associations with MAF ranging from 0.01 to 0.48. Of the 19 associations, we were able to replicate five in an independent sample of approximately 90,000 individuals.

One of the five associations we identified is between TG and rs116843064, an exonic variant in the ANGPTL4 gene on chromosome 19 (figure 2C). This missense variant changes the amino acid glutamic acid into lysine (Glu40Lys) and is predicted to be damaging for the structure and function of the protein by Polyphen28 MutationTaster9 and likelihood ratio test (LRT).10 ANGPTL4 is significantly associated with the Kyoto Encyclopedia of Genes and Genomes (KEGG) term fatty acid metabolism, the GO process lipid storage and the gene ontology (GO) cellular component lipid particle (p value=1.10×10−5 and 2.87×10−18, respectively, genenetwork.nl). ANGPTL4 has been associated with HDL-C before using the GWAS approach4 and with TG before using an exome sequencing approach11 and more recently using the GWAS approach.1 We therefore do not claim this finding as novel, though this is the smallest study in which this variant was genome-wide significantly associated with TG and replicated in an independent sample.

The second new finding we identified is the association between TC and rs6457374, an intergenic variant located on chromosome 6 between the genes HLA-C and HLA-B (figure 2A). Both genes are associated with the KEGG term ATP binding cassette (ABC) transporters (p value of 4.29×10−5 and 3.84×10−5 for HLA-C and HLA-B, respectively, genenetwork.nl) which is in line with, among others, a previously published association between TC and an exonic variant in the ABCA6 gene which is also an ABC transporter.12 ABC transporters transport a wide variety of substrates across extracellular and intracellular membranes, including lipids.13

The third finding of this study is the association between HDL-C and rs77697917, an intergenic variant on chromosome 17 between the genes SOST and DUSP3 (figure 2B). DUSP3 is associated with the regulation and function of carbohydrate-responsive element-binding protein (ChREBP) in the liver (p value=3.03×10−5, genenetwork.nl). ChREBP mediates the activation of several regulatory enzymes involved in lipogenesis.14−18 This variant is in high linkage disequilibrium (D′=0.936) in the 1 kG with rs72836561, an exonic variant in the gene CD300LG (MAF=0.027, β=−2.437, se=0.381, p value=1.51×10−10 in the discovery stage). This missense variant changes the amino acid arginine into cysteine (Arg82Cys) and is predicted to be damaging for the structure and function of the protein by Polyphen28 MutationTaster9 and LRT.10 This amino acid polymorphism has been associated with HDL-C in exome-wide association studies19 and TG in GWAS1 before.

The fourth variant we identified is rs186696265, which is located on chromosome 6 and associated with LDL-C and TC (figure 2D, E). This intergenic variant is between the LPA (Lipoprotein, Lp(A)) gene and the PLG (Plasminogen) gene. The LPA gene has been associated before with LDL-C and TC before.2 The reported lead SNP was rs1564348, which in the newer human genome versions is annotated to the LPA Member 1) gene instead of the LPA (Lipoprotein, Lp(A)) gene. This explains why we again identified a locus near the LPA gene, which has been identified by others as well.7

Fourteen out of the 19 variants were not replicated despite similar sample sizes and similar frequencies within the replication stage as compared with the discovery stage. Of those 14 variants, 11 exhibited effect sizes in the same direction in both stages. A possible explanation might be that the replication sample size is much larger compared with that of the discovery sample size. Two variants might have lacked significant replication due to small sample size, rs60839105 and rs151198427.

Figure 1 Manhattan plots for HDL-C (A), LDL-C (B), TC (C) and TG (D) after the meta-analysis of all discovery cohorts. Variants that were present in at least four cohorts and that are not within 0.5 Mb of a previously published locus23 were included. The black line indicates the genome-wide significant line (5×10−8), the black and red dots the variants identified by GCTA which are not genome-wide significant and which are genome-wide significant, respectively. HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.
Table 1 The results for the 19 variants after the meta-analysis of all discovery cohorts, all replication cohorts and all cohorts combined

<table>
<thead>
<tr>
<th>Trait</th>
<th>Chr:Position</th>
<th>rs identifier</th>
<th>nearest gene</th>
<th>A1/A2</th>
<th>Freq</th>
<th>N</th>
<th>β</th>
<th>SE</th>
<th>p Value</th>
<th>Freq</th>
<th>N</th>
<th>β</th>
<th>SE</th>
<th>p Value</th>
<th>Freq</th>
<th>N</th>
<th>β</th>
<th>SE</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDL-C</td>
<td>3:72 067 255</td>
<td>rs75909755</td>
<td>PROL2-EIF4E3</td>
<td>T/C</td>
<td>0.03</td>
<td>62 607</td>
<td>1.593</td>
<td>0.275</td>
<td>7.27E-09</td>
<td>0.03</td>
<td>86 252</td>
<td>−0.019</td>
<td>0.031</td>
<td>5.45E-01</td>
<td>0.03</td>
<td>0.002</td>
<td>0.031</td>
<td>9.57E-01</td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>6:31 272 261</td>
<td>rs6457374</td>
<td>HLA-B</td>
<td>T/C</td>
<td>0.75</td>
<td>46 839</td>
<td>2.339</td>
<td>0.339</td>
<td>5.32E-12</td>
<td>0.81</td>
<td>74 417</td>
<td>−0.057</td>
<td>0.016</td>
<td>4.23E-04</td>
<td>0.81</td>
<td>0.062</td>
<td>0.016</td>
<td>1.18E-04</td>
<td></td>
</tr>
<tr>
<td>LDL-C</td>
<td>6:31 325 323</td>
<td>rs9266229</td>
<td>HLA-B</td>
<td>C/G</td>
<td>0.53</td>
<td>37 981</td>
<td>−2.201</td>
<td>0.344</td>
<td>1.62E-10</td>
<td>0.41</td>
<td>61 582</td>
<td>−0.025</td>
<td>0.014</td>
<td>7.37E-02</td>
<td>0.41</td>
<td>−0.029</td>
<td>0.014</td>
<td>4.04E-02</td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>6:36 648 275</td>
<td>–</td>
<td>CDKN1A</td>
<td>CAG/C</td>
<td>0.45</td>
<td>53 425</td>
<td>−0.019</td>
<td>0.003</td>
<td>7.63E-09</td>
<td>0.49</td>
<td>59 018</td>
<td>−0.003</td>
<td>0.004</td>
<td>5.20E-01</td>
<td>0.46</td>
<td>−0.013</td>
<td>0.003</td>
<td>5.93E-07</td>
<td></td>
</tr>
<tr>
<td>HDL-C</td>
<td>3:72 067 255</td>
<td>rs75909755</td>
<td>PROL2-EIF4E3</td>
<td>T/C</td>
<td>0.03</td>
<td>62 607</td>
<td>1.593</td>
<td>0.275</td>
<td>7.27E-09</td>
<td>0.03</td>
<td>86 252</td>
<td>−0.019</td>
<td>0.031</td>
<td>5.45E-01</td>
<td>0.03</td>
<td>0.002</td>
<td>0.031</td>
<td>9.57E-01</td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>6:31 272 261</td>
<td>rs6457374</td>
<td>HLA-B</td>
<td>T/C</td>
<td>0.75</td>
<td>46 839</td>
<td>2.339</td>
<td>0.339</td>
<td>5.32E-12</td>
<td>0.81</td>
<td>74 417</td>
<td>−0.057</td>
<td>0.016</td>
<td>4.23E-04</td>
<td>0.81</td>
<td>0.062</td>
<td>0.016</td>
<td>1.18E-04</td>
<td></td>
</tr>
<tr>
<td>LDL-C</td>
<td>6:31 325 323</td>
<td>rs9266229</td>
<td>HLA-B</td>
<td>C/G</td>
<td>0.53</td>
<td>37 981</td>
<td>−2.201</td>
<td>0.344</td>
<td>1.62E-10</td>
<td>0.41</td>
<td>61 582</td>
<td>−0.025</td>
<td>0.014</td>
<td>7.37E-02</td>
<td>0.41</td>
<td>−0.029</td>
<td>0.014</td>
<td>4.04E-02</td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>6:36 648 275</td>
<td>–</td>
<td>CDKN1A</td>
<td>CAG/C</td>
<td>0.45</td>
<td>53 425</td>
<td>−0.019</td>
<td>0.003</td>
<td>7.63E-09</td>
<td>0.49</td>
<td>59 018</td>
<td>−0.003</td>
<td>0.004</td>
<td>5.20E-01</td>
<td>0.46</td>
<td>−0.013</td>
<td>0.003</td>
<td>5.93E-07</td>
<td></td>
</tr>
<tr>
<td>HDL-C</td>
<td>3:72 067 255</td>
<td>rs75909755</td>
<td>PROL2-EIF4E3</td>
<td>T/C</td>
<td>0.03</td>
<td>62 607</td>
<td>1.593</td>
<td>0.275</td>
<td>7.27E-09</td>
<td>0.03</td>
<td>86 252</td>
<td>−0.019</td>
<td>0.031</td>
<td>5.45E-01</td>
<td>0.03</td>
<td>0.002</td>
<td>0.031</td>
<td>9.57E-01</td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>6:31 272 261</td>
<td>rs6457374</td>
<td>HLA-B</td>
<td>T/C</td>
<td>0.75</td>
<td>46 839</td>
<td>2.339</td>
<td>0.339</td>
<td>5.32E-12</td>
<td>0.81</td>
<td>74 417</td>
<td>−0.057</td>
<td>0.016</td>
<td>4.23E-04</td>
<td>0.81</td>
<td>0.062</td>
<td>0.016</td>
<td>1.18E-04</td>
<td></td>
</tr>
<tr>
<td>LDL-C</td>
<td>6:31 325 323</td>
<td>rs9266229</td>
<td>HLA-B</td>
<td>C/G</td>
<td>0.53</td>
<td>37 981</td>
<td>−2.201</td>
<td>0.344</td>
<td>1.62E-10</td>
<td>0.41</td>
<td>61 582</td>
<td>−0.025</td>
<td>0.014</td>
<td>7.37E-02</td>
<td>0.41</td>
<td>−0.029</td>
<td>0.014</td>
<td>4.04E-02</td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>6:36 648 275</td>
<td>–</td>
<td>CDKN1A</td>
<td>CAG/C</td>
<td>0.45</td>
<td>53 425</td>
<td>−0.019</td>
<td>0.003</td>
<td>7.63E-09</td>
<td>0.49</td>
<td>59 018</td>
<td>−0.003</td>
<td>0.004</td>
<td>5.20E-01</td>
<td>0.46</td>
<td>−0.013</td>
<td>0.003</td>
<td>5.93E-07</td>
<td></td>
</tr>
<tr>
<td>HDL-C</td>
<td>3:72 067 255</td>
<td>rs75909755</td>
<td>PROL2-EIF4E3</td>
<td>T/C</td>
<td>0.03</td>
<td>62 607</td>
<td>1.593</td>
<td>0.275</td>
<td>7.27E-09</td>
<td>0.03</td>
<td>86 252</td>
<td>−0.019</td>
<td>0.031</td>
<td>5.45E-01</td>
<td>0.03</td>
<td>0.002</td>
<td>0.031</td>
<td>9.57E-01</td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>6:31 272 261</td>
<td>rs6457374</td>
<td>HLA-B</td>
<td>T/C</td>
<td>0.75</td>
<td>46 839</td>
<td>2.339</td>
<td>0.339</td>
<td>5.32E-12</td>
<td>0.81</td>
<td>74 417</td>
<td>−0.057</td>
<td>0.016</td>
<td>4.23E-04</td>
<td>0.81</td>
<td>0.062</td>
<td>0.016</td>
<td>1.18E-04</td>
<td></td>
</tr>
<tr>
<td>LDL-C</td>
<td>6:31 325 323</td>
<td>rs9266229</td>
<td>HLA-B</td>
<td>C/G</td>
<td>0.53</td>
<td>37 981</td>
<td>−2.201</td>
<td>0.344</td>
<td>1.62E-10</td>
<td>0.41</td>
<td>61 582</td>
<td>−0.025</td>
<td>0.014</td>
<td>7.37E-02</td>
<td>0.41</td>
<td>−0.029</td>
<td>0.014</td>
<td>4.04E-02</td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>6:36 648 275</td>
<td>–</td>
<td>CDKN1A</td>
<td>CAG/C</td>
<td>0.45</td>
<td>53 425</td>
<td>−0.019</td>
<td>0.003</td>
<td>7.63E-09</td>
<td>0.49</td>
<td>59 018</td>
<td>−0.003</td>
<td>0.004</td>
<td>5.20E-01</td>
<td>0.46</td>
<td>−0.013</td>
<td>0.003</td>
<td>5.93E-07</td>
<td></td>
</tr>
</tbody>
</table>

The variants in bold are the significantly replicated variants. A1 is allele 1 and A2 is allele 2, Freq is the frequency of A1, β is the effect of A1. HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.
Both variants only pass quality control in the cohorts in the discovery stage that contain individuals of African ancestry (see online supplementary figure S7). Although there are several cohorts with individuals of African ancestry in the replication stage, both variants did not pass quality control in most cohorts which leads to the conclusion that these variants might be population-specific. This is also suggested by the 1 kG data (Phase 3) as the frequency of the C-allele is 92% in African samples and 100% in the European samples for rs60839105 and the frequency of the G-allele is 86% in the African samples and 100% in the European samples for rs151198427. Imputations of cohorts with individuals of African ancestry with the African Genome Variation Project20 might confirm the association of rs60839105 with HDL-C and rs151198427 with TC.

To our knowledge, this is the first GWAS of circulating lipid levels using the Phase 1 integrated release V3 of the 1 kG, therefore we cannot compare the positive replication rate with other studies. However, we did replicate 88.1% of the findings of Teslovich et al2 and 43.4% of the findings of GLGC3 despite our smaller sample. A high replication rate is expected based on the high overlap of our samples with the samples of Teslovich et al2 and with the samples of GLGC3 though it indicates that when using the 1000 Genomes instead of the HapMap reference panel, we can achieve a high replication rate using a smaller sample size. We also tried to replicate findings from


Figure 2  The regional association results of the initial meta-analysis of all discovery cohorts for (A) TC on chromosome 6, (B) HDL-C on chromosome 17, (C) TG on chromosome 19, (D) LDL-C on chromosome 6 and (E) TC on chromosome 6. HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.
exome sequencing of candidate genes. The p.Arg406X mutation in the NPC1L1 gene (rs145297799), which was reported to be associated with reduced LDL-C levels and reduced risk of coronary heart disease, is not available in the 1kG reference panel and, therefore, we were not able to replicate this finding. Do et al. described the exome sequencing of the genes LDLR and APOA5 and identified rare variants associated with an increased risk of myocardial infarction, increased LDL-C and TG levels. Of those rare variants, only two in the LDLR gene and seven in the APOA5 gene exist in our discovery meta-analysis. Both LDLR variants are associated with TG in our discovery meta-analysis (rs34282181, β = −0.093, SEβ = 0.023, p value = 4.827 × 10^-5) and rs2075291, β = 0.219, SEβ = 0.046, p value = 2.092 × 10^-6), but not significantly associated with LDL-C (rs34282181, β = −3.939, SEβ = 1.861, p value = 0.034 and rs2075291, β = −2.316, SEβ = 3.001, p value = 0.440). None of the seven APOA5 variants were significantly associated with TG or LDL-C in our discovery meta-analysis (lowest p value is for LDL-C with rs72658860, β = −18.430, SEβ = 7.140, p value = 9.848 × 10^-4). The third published finding we tried to replicate, was the association between APOC3 and TG levels. Of the seven variants reported, only one existed in our discovery meta-analysis (chromosome 11, position 116 701 354), which is associated with TG (β = −0.343, SEβ = 0.113, p value = 2.311 × 10^-6). Those authors also reported an association between a APOA5 variant (rs3135506) and TG as the most significant finding. This variant was also significantly associated with TG in our discovery meta-analysis (β = 0.129, SEβ = 0.007, p value = 1.999 × 10^-8). These replication efforts demonstrate that many of the published results of exome sequencing can be replicated through the use of 1kG imputations.

In conclusion, we identified and replicated five variants associated with circulating lipid levels. These variants are in genes that can be linked biologically to lipid metabolism. Although there were a large number of variants that did not replicate at the accepted genome-wide significance threshold, the low-cost, hypothesis-free approach that we applied uncovered five variants. This study, therefore, illustrates that GWAS may still help us unravel the biological mechanisms behind circulating lipid levels.

**Author affiliations**

1Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
3Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
4Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
5Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
6Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland
7Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland
8Public Health Sciences, Loyola University Chicago Stritch School of Medicine, Maywood, USA
9Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
10Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands
11Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
12McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, USA
13Human Genomics Center, The University of Texas School of Public Health, Houston, USA
14Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
15Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
16Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Aberdeen, UK
17British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK
18Department of Biostatistics, University of Washington, Seattle, USA
19Brigham and Women’s Hospital, Boston, USA
20Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Bethesda, USA
21Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK
22Faculty of Medicine, University of Split, Split, Croatia
23Department of Health, National Institute for Health and Welfare, Helsinki, Finland
24Cardiovascular Science, National Heart and Lung Institute, Imperial College London, London, UK
25Imperial College Healthcare NHS Trust, Imperial College London, London, UK
26Department of Psychiatry, VU University Medical Center Amsterdam/GGGZinGeest and EMGO+ Institute for Health and Care Research and Neuroscience Campus Amsterdam, Amsterdam, The Netherlands
27Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK
28Department of Nephrology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
29Department of Cardiology, Turku University Hospital, Turku, Finland
30Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
31School of Medicine, University of Tampere, Tampere, Finland
32Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, Jamaica
33Laboratory of Clinical Chemistry and Hematology, University Medical Center Utrecht, Utrecht, The Netherlands
34Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
35Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland
36Department of Clinical Physiology, University of Tampere School of Medicine, Tampere, Finland
37Laboratory of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, University of Cambridge, Cambridge, UK
38Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
39Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands
40Epidemiology Section, Department of BESC, King Faisal Medical Hospital and Research Centre, Riyadh, Saudi Arabia
41Genomics Coordination Center, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
42Public Health, University of Helsinki, Helsinki, Finland
43Welcome Trust Sanger Institute, UK

Acknowledgements The authors especially thank all volunteers who participated in the study. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions to the ERFS study and to P Verart for her help in genealogy, J Veerger for the supervision of the laboratory work and P Snijders for his help in data collection.

The authors thank Behroz Alizadeh, Annemieke Boeije, Marcel Buinenberg, Monnie Fester, ilja Wolle, Lude Franke, Mila Vainohamaduri for their help in creating the GWAS database, and Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, Judith Vonk for their work related to data collection and validation. The authors are grateful to the study participants, the staff from the LifeLines Cohort Study and the contributing research centres delivering data to LifeLines and the participating general practitioners and pharmacists.

MESA and the MESA SHARE genotyping project are conducted and supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079 and UL1-TR-000040 from the National Heart, Lung, and Blood Institute (NHLBI). Funding for MESA SHARE genotyping was provided by NHLBI Contract N02-HL-6-4278. The provision of genotyping data was supported as part of the National Center for Advancing Translational Sciences, CTSA grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CROATIA-Korcula, CROATIA-Split and CROATIA-Vis (CR-Korcula, CR-Split, CR-Vis) were funded by the Medical Research Council UK, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROS Project (contract no. LSHG-CT-2006-018947) and the Croatian Science Foundation (grant 8875).

The ERFS study as a part of EUROS Project (European Special Populations Research Network) was supported by European Commission FP6 STR grant number 018947 (LSHG-CT-2006-018947) and also received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme “Quality of Life and Management of the Living Resources” of 5th Framework Programme (no. QLG2-CT-2002-01254). The ERFS study was further supported by ENGAGE consortium and CBMS. High-throughput analysis of the ERFS data was supported by joint grant from the Netherlands Organisation for Organising and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing in ERFS was supported by the ZonMW grant (project 91110125). The Family Heart Study was supported by the by grants R01-HL-087700, R01-HL-088215 and R01-HL-117078 from the National Heart, Lung, and Blood Institute.

Generation Scotland received core funding from the Chief Scientist Office of the Scottish Government Health Directorate C2Z/16/6 and the Scottish Funding Council HR3006. Genotyping of the SS:SHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the UK’s Medical Research Council.

The LifeLines Cohort Study, and generation and management of GWAS genotype data for the LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation.

The Leiden Longevity Study (LLS) has received funding from the European Union’s Seventh Framework Programme (FP7/2007–2011) under grant agreement no 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SentierGenom IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), UnileverColworth and by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). The LOLLIP pop study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0061966,G0070031), the Wellcome Trust (084732/2/08Z) the NIHR

The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 750.100.2005.011, 913-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RI2E), the Netherlands Geriatric (NGL) Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organisation for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

Participating centres of TRAILS include the University Medical Center and University of Groningen, the Erasmus University Medical Center Rotterdam, the University of Utrecht, the Radboud Medical Centre Nijmegen, and the Pammasia Bavo group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMW Risk Behavior and Dependence grants 60-60060-97-118; ZonMW Culture and Health grant 261-98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481-08-013 and 481-11-001), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), and the participating universities.

Competing interests
BMP serves on the DSMB of a clinical trial of a device funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. OHF works in ErasmusACE, a centre for ageing research across the life course funded by Nestlé Nutrition (Nestec); Metagenics; and AXA. Nestlé Nutrition (Nestec); Metagenics; and AXA had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review or approval of the manuscript.

Patient consent
Obtained.

Ethics approval
All studies were performed with the approval of the local Medical Ethics Committees.

Provenance and peer review
Not commissioned; externally peer reviewed.

Data sharing statement
All meta-analysis results and the results of the follow-up analysis of this project are available.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
Meta-analysis of 49 549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in *ANGPTL4* determining fasting TG levels


*J Med Genet* 2016 53: 441-449 originally published online April 1, 2016
doi: 10.1136/jmedgenet-2015-103439

Updated information and services can be found at:
[http://jmg.bmj.com/content/53/7/441](http://jmg.bmj.com/content/53/7/441)

**These include:**

**References**

This article cites 20 articles, 3 of which you can access for free at:
[http://jmg.bmj.com/content/53/7/441#BIBL](http://jmg.bmj.com/content/53/7/441#BIBL)

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)