Published in final edited form as:

J Hum Genet. 2017 October ; 62(10): 895–901. doi:10.1038/jhg.2017.55.

Replication and fine-mapping of genetic predictors of lipid traits in African-Americans

QiPing Feng, PhD[1], **Wei-Qi Wei, MMed PhD**[2], **Rebecca T Levinson, BS**[3], **Jonathan D Mosley, MD**[1], and **C Michael Stein, MBChB**[1,4]

[1]Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

[2]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN

[3]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN

[4]Department of Pharmacology, Vanderbilt University, Nashville, TN

Abstract

Background—Circulating lipid concentrations are among the strongest modifiable risk factors for coronary artery disease (CAD). Most genetic studies have focused on Caucasian populations with little information available for populations of African ancestry.

Methods—Using a cohort of ~2,800 African-Americans (AAs) from a biobank at Vanderbilt University (BioVU), we sought to trans-ethnically replicate genetic variants reported by the Global Lipids Genetics Consortium to be associated with lipid traits in Caucasians, followed by fine-mapping those loci using all available variants on the MetaboChip.

Results—In AAs, we replicated one of 56 SNPs for total cholesterol (TC) (rs6511720 in *LDLR*, $p=2.15\times10^{-8}$), one of 63 SNPs for high-density lipoprotein cholesterol (HDL-C) (rs3764261 in *CETP*, $p=1.13\times10^{-5}$), two of 46 SNPs for low-density lipoprotein cholesterol (LDL-C) (rs629301 in *CELSR2/SORT1*, $p=1.11\times10^{-5}$ and rs6511720 in *LDLR*, $p=2.47\times10^{-5}$) and one of 34 SNPs for TG (rs645040 in *MSL2L1*, $p=4.29\times10^{-4}$). Using all available variants on MetaboChip for fine mapping, we identified additional variants associated with TC (*APOE*), HDL-C (*LPL* and *CETP*) and LDL-C (*APOE*). Furthermore, we identified two loci significantly associated with non-HDL-C: *APOE/APOC1/TOMM40* and *PCSK9*.

Conclusions—The genetic architecture of lipid traits in AAs differs substantially from that in Caucasians and it remains poorly characterized.

Keywords

lipid; non-HDL cholesterol; African-Americans

Disclosures

There is no potential conflict of interest.

Introduction

Circulating lipid concentrations are among the strongest modifiable risk factors for coronary artery disease (CAD). The wide inter-individual variation observed in lipid concentrations is influenced by both environmental and genetic factors. Large genome-wide association studies (GWAS) have identified >150 loci associated with concentrations of one or more lipid traits (1,2). Nevertheless, only 10–15% of overall variation in lipid concentrations can be attributed to genetic variants identified, despite an estimated heritability of 40–50% (3,4).

Understanding the genetic architecture for lipid traits is critical for developing risk prediction approaches to CAD. Most genetic studies have focused on Caucasian populations and limited efforts have been made to generalize Caucasian findings to populations of African ancestry such as African-Americans (AAs). Earlier studies suggested that less than half of lipid associated genetic findings replicated in African populations (5). The Global Lipids Genetics Consortium (GLGC) attempted to replicate their Caucasian meta-analyses findings in AAs (2); only one locus for high-density lipoprotein cholesterol (HDL-C), *CETP*, three loci for low-density lipoprotein cholesterol (LDL-C), *SORT1*, *LDLR* and *APOE*, and none for triglycerides (TG) replicated. There has been little additional work to further replicate and fine-map LDL-C, HDL-C, TG and total cholesterol (TC) loci in AA populations.

Rare (<1%) and low-frequency (1–5%) variants are inadequately represented on common GWAS genotyping platforms. Given that rare variants play a dominant role in the etiology of Mendelian disorders, they may also convey larger risk than common variants for common phenotypes. The increased genetic diversity in the AA population provides an opportunity to study low-frequency variants related to lipid metabolism and the potential to identify novel drug targets to lower CAD risk. One of such recent example is variation in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene. By fine-mapping the *PCSK9* gene locus in an AA cohort, Cohen *et al.* identified loss-of-function variants associated with markedly decreased LDL-C levels and with decreased CAD risk (6–8). This finding triggered the development of a new class of lipid lowering drugs (6–8), the PCSK9 inhibitors, which have recently been approved by the FDA (9–11). In Caucasians, efforts have been made to identify and understand low-frequency variants for LDL-C (12); however, less is known in AAs particularly for other lipid traits, such as HDL-C,TC, and TG although they are associated with CAD risk (13).

Non-HDL cholesterol (non-HDL-C) is a measure of cholesterol containing all atherogenic particles. Observations from several large longitudinal cohorts show that non-HDL-C is a better predictor of CAD than LDL-C (14–17). In the third Adult Treatment Panel (ATPIII) guidelines of the US National Cholesterol Education Program, non-HDL-C was introduced as a secondary target of therapy in patients with TG greater than 200 mg/dL. By sequencing Caucasians, *LDLR* and *ASGR1* variants were associated with both non-HDL-C and CAD risk (18,19). However, no genome-wide scale analysis has been done to understand genetic predictors for non-HDL-C in AAs. Another lipid trait, remnant cholesterol, is a measure of cholesterol content of triglyceride-rich lipoproteins and includes very low-density lipoprotein, intermediate-density lipoprotein and chylomicron remnants (20,21). Elevated

remnant cholesterol levels have been associated with increased risk of CAD independent of HDL-C (22). However, little is known about the genetic predictors of remnant cholesterol level in AAs.

In this study of AAs, we set out to trans-ethnically replicate genetic variants reported by the GLGC to be associated with lipid traits in Caucasians (1,2), followed by fine-mapping of those loci using all available variants on the MetaboChip, a platform designed for metabolic traits (23,24) and enriched for rare and low-frequency variants.

Methods

Study Population

The study was approved by the Institutional Review Board of Vanderbilt University. The study cohort consisted of third party-identified AAs older than 18 years of age who had lipid measurements and genotyping information available in the Vanderbilt DNA biobank (BioVU). BioVU accrues DNA samples from blood drawn for routine clinical testing after these samples are scheduled to be discarded. BioVU and sample handling have been previously described (25). Samples and existing genotypes in BioVU are linked to a de-identified version of each individual's electronic health record (EHR).

Phenotyping – Extraction of lipid measurements from electronic health records

Lipid measurements were extracted from each individual's EHR, and median LDL-C, HDL-C, TC and TG calculated. Lipid measurements after statin exposure were excluded from the analyses. We manually reviewed charts for individuals having extreme lipid levels to exclude data entry errors. We further calculated non-HDL-C for each individual by subtracting HDL-C from TC and remnant cholesterol by subtracting both HDL-C and LDL-C from TC.

Genotyping

Genotyping was performed using the Illumina MetaboChip (23,24) a custom BeadChip targeting 196,725 genetic variants. Genotyping data were curated for quality control using PLINK (26). We removed SNPs with a call rate less than 95% and SNPs that deviated significantly from Hardy–Weinberg equilibrium (HWE) ($p \leq 1.0 \times 10^{-6}$). We removed samples: (1) with per-individual call rate <95%; (2) with mismatch between genetic and EHR sex; (3) with a cryptic relationship closer than a third-degree relative. Twenty-nine individuals were removed from the analysis.

Candidate SNPs

The GLGC previously identified genetic variants associated with four lipid traits (1,2). We extracted genotype information for these variants, including 63 variants for HDL-C, 46 variants for LDL-C, 56 variants for TC and 34 variants for TG. We also extracted all available genetic variants (from transcription start to transcription end) for GLGC-identified genes, including 3939 variants for HDL-C, 1995 variants for LDL-C, 2430 variants for TC and 3888 variants for TG.

Statistical analyses

Data were adjusted for population stratification using principal component analyses implemented in EIGENSOFT4.2 (27,28). BioVU contains observer-reported ancestry information obtained at the time of clinical visit. This information has been confirmed to achieve high accuracy (29).

Genetic association analysis was performed using PLINK v1.07 (26). Median lipid levels were natural log transformed. We tested the association between genetic variants and lipid traits (HDL-C, LDL-C, TC, TG, non-HDL-C, and remnant cholesterol). An additive inheritance model was assumed and tested using a linear regression model. Analyses were adjusted for age (at the time of median lipid measurement), sex, and 6 principal components (PCs) for ancestry. We conducted 3-tiers of analyses: (1) testing lipid-associated candidate SNPs reported by the GLGC; (2) testing SNPs within the lipid-associated gene identified by the GLGC – a gene locus was defined as the range of the gene transcript; and (3) testing all available SNPs on the MetaboChip. The associations were further conditioned on the lead SNPs within the regions. Specifically, we conditioned on rs28362286 for *PCSK9* region, rs7412 for *APOE* region, rs34065661 for *CETP* region and rs4389957 for *LPL* region. All analyses were adjusted for multiple testing accordingly. The regional association plots were generated using LocusZoom (30) (http://locuszoom.sph.umich.edu/locuszoom/). With adjustment for multiple testing, we defined levels of significance as follows: (1) for candidate SNPs: 8.9×10^{-4} for TC, 7.9×10^{-4} for HDL-C, 1.1×10^{-3} for LDL-C and 1.5×10^{-3} for TG; (2) for SNPs in candidate gene regions: 2.1×10^{-5} for TC, 1.3×10^{-5} for HDL-C, 2.5×10^{-5} for LDL-C and 1.3×10^{-5} for TG ; (3) for the entire MetaboChip analyses: 2.8×10^{-7} for all lipid levels.

Heritability analyses

We used Genome-wide Complex Trait Analysis (GCTA, version 1.24.7) (31) to estimate the polygenic variance attribute to all genotyped MetaboChip SNPs. We excluded SNPs: (1) with MAF less than 0.01, (2) with HWE less than 0.000001, (3) with SNP call rate less than 0.02. We also removed individuals with call rate less than 0.95. We utilized the identified set of SNPs (n=114,451) to calculate a genetic relatedness matrix. Heritability estimates were adjusted for age, sex, and 20 PCs.

Results

Lipid levels in AAs genotyped on the Metabochip were available for TC (n= 2778), LDL-C (n=2438), HDL-C (n=2550), TG (n=2690), non-HDL-C (n=2468), and remnant cholesterol (n=2262). Demographic characteristics are summarized in Table 1.

Replication of GLGC reported genetic variants

First, we set out to trans-ethnically replicate the GLGC variants (1,2) in the BioVU AA cohort. After correcting for multiple testing, we replicated one of 56 SNPs for TC (rs6511720 in *LDLR*, $p=2.15\times10^{-8}$), one of 63 SNPs for HDL-C (rs3764261 in *CETP*, $p=1.13\times10^{-5}$), two of 46 SNPs for LDL-C (rs629301 in *CELSR2/SORT1*, $p=1.11\times10^{-5}$ and rs6511720 in *LDLR*, $p=2.47\times10^{-5}$) and one of 34 SNPs for TG (rs645040 in *MSL2L1*,

p=4.29×10^{-4}) (Table 2). These variants were all associated with predicted lipid changes in the same direction as observed in Caucasians (1). Variants in *SORT1*, *CETP* and *LDLR* had similar effect size in AAs as in Caucasians; however, rs645040 in *MSL2L1* was associated with larger changes in TG in AAs than in Caucasians – the minor allele was associated with −8.56 mg/dL TG change in AAs, but only −2.2 mg/dL change in Caucasians (1) (Table 2).

Fine-mapping of GLGC loci significantly associated with lipid traits

Second, we tested all genetic variants present on the Metabochip within loci identified by the GLGC as associated with lipids. Locus fine mapping yielded additional associations (Table 3), which were not previously reported as leading SNPs from GLGC analyses in Caucasians.

1. Associations with TC.One additional *PCSK9* variant, rs28362286, associated with TC (Table 3, p=1.50×10^{-10}). A well-characterized *APOE* variant, rs7412, was also associated with TC with genome-wide significance (p=3.09×10^{-22}).
2. Associations with HDL-C.Six additional variants were associated with HDL-C in AAs – one in *LPL* and five in *CETP*. Two SNPs in *CETP* achieved genome-wide significance: rs7499892 (p=1.51×10^{-10}) was previously reported in several Caucasian GWAS studies (32,33), and rs34065661 (p=1.53×10^{-13}), a missense variant, was the variant most significantly associated HDL-C in our analyses (Table 3). We queried the lead *CETP* SNP (rs34065661) and the linked SNPs in Genotype-Tissue Expression (GTEx) database. Rs711752, which associated with rs34065661 (D′=1.0, R^2=0.192, Supplementary Table 2), was found to be an eQTL variant and carriers of the variant had significantly lower CETP expression (p=1.6×10^{-5}) compared to non-carriers (Figure 2A).
3. Associations with LDL-C. Three variants were associated with LDL-C. Rs28362261 (p=2.08×10^{-5}) and rs28362286 (p=1.99×10^{-13}) in the *PCSK9* locus, and rs7412 (p=2.48×10^{-44}) in the *APOE* locus, which accounted for −22.08 mg/dL LDL-C change, comparable to −22.52 mg/dL from NHANES III(34). Using the GTEx database, rs7412 is predicted to significantly alter APOE expression (p=7.3×10^{-6}) (Figure 2B).
4. No additional associations were observed for TG

Association between variants in the entire Metabochip and GLGC lipid traits

Third, by analyzing the entire MetaboChip, we identified additional variants significantly associated with lipids (Supplementary Table 3). For each identified gene, we defined a gene region as the gene transcript ± 50kb, and generated regional association plots (Figure 1). We further identified 8 SNPs in the CETP region associated with HDL-C and 7 SNPs in the APOE region associated with LDL-C (Supplementary Table 3). The eight *CETP* variants are in strong linkage disequilibrium (Supplementary Figure 1). Most associations were attributed to their linkage with the lead SNPs in the regions (Supplementary Table 3). After conditioning on the lead SNPs, two variants remained significantly associated with HDL-C (rs4783961 in *CETP* region and rs4389957 in *LPL/SLC18A2* region, Table 4), and one variant significantly associated with LDL-C (rs611917 in CELSR2/SORT1 region) (Table 4).

No additional associated variant was identified for TC and TG.

Genetic predictors for non-HDL-C

We tested the association between all MetaboChip variants and non-HDL-C levels, and identified two loci that contained 5 SNPs significantly associated with non-HDL-C levels (Table 5). All five SNPs were also associated with either TC or LDL in previous analyses (Table 5).

Genetic predictors for remnant cholesterol

We further tested the association with remnant cholesterol, and observed no association in AAs (data not shown).

Heritability of lipid traits explained by MetaboChip

By estimating the percentage of heritability explained by available genotypes in MetaboChip, we found that additive genetic components explained 22.52 ± 8.86 % of TC (p=0.0048), 19.22 ± 9.59% of HDL-C (p=0.023) and 28.51 ± 9.74 % of LDL-C (p=0.0013), 34.35 ± 10.9 % of non-HDL-C (p=0.00063), but only 8.25 ± 8.51 % of TG (p=0.159).

Discussion

In this study, we sought to replicate variants associated with lipid traits identified by the GLGC and to fine-map significantly associated loci in ~2,800 AAs. There are two major findings of the study: (1) relatively few lipid-associated variants identified by the GLGC replicated in AAs; and (2) we identified additional variants associated with TC (*APOE*), HDL-C (*LPL* and *CETP*) and LDL-C (*APOE*), and two loci significantly associated with non-HDL-C (*APOE/APOC1/TOMM40* and *PCSK9)*.

The GLGC previously identified 157 loci associated with one or more lipid traits from a cohort of predominantly Caucasians; only a few variants replicated in AAs. Fine-mapping and analyses using the entire MetaboChip identified a few additional genetic associations with GLGC lipid traits in AAs and also 5 SNPs associated with non-HDL-C.

In addition to the complexity of the genetic architecture in individuals of African ancestry, several reasons could explain why so few GLGC variants replicated in AAs. It is likely that because the 157 loci account for a small portion of overall inter-individual variability of lipids in Caucasians, the contribution of those variants could be even harder to detect if their frequency is lower in AAs than in Caucasians. Nevertheless, we were able to confirm the role of *APOE*, *CETP*, *PCSK9* and *LPL* in regulating lipid levels in AA population:

Rs7412 in *APOE* is one of the most promising predictors of LDL-C and TC levels in AAs. APOE is a ligand for LDL receptor and therefore is involved in the removal of LDL from the circulation. The rs7412 SNP changes the amino acid at position 158 in APOE from Arg to Cys and in the GTEx database, rs7412 carriers have lower APOE expression compared to non-carriers. Previous reports from NHANES III suggested that rs7412 was associated with 22.52 mg/dL lower LDL-C concentrations and 20.68 mg/dL lower TC concentrations per

minor allele (34). In AAs, rs7412 was associated with 22.08 mg/dL lower LDL-C, similar to findings in other populations.

Human cholesteryl ester transfer (CETP) protein plays a crucial role in lipid metabolism by mediating the transfer of cholesteryl esters from HDL to apolipoprotein (apo) B rich lipoproteins in exchange for TG. High CETP activity contributes to an unfavorable plasma lipoprotein profile by lowering HDL-C and increasing LDL-C. In the current analyses, variants in the *CETP* region were identified with genome-wide significance. Further efforts to sequence *CETP* in AAs may provide additional insight into the gene function and its genetic structure.

PCSK9 is an enzyme that regulates LDL-C levels by promoting the degradation of LDL receptors that are responsible for the removal of LDL particles from the circulation into the liver. Rs28362286 is a well-characterized African-specific variant in *PCSK9* and has been associated with altered PCSK9 function, lower LDL-C concentrations, and reduced CAD(7). The variant results in a truncated PCSK9 protein with reduced activity that is therefore less efficient in degrading LDL receptors; consequently, more LDL is removed from the circulation. In the current cohort, the variant was significantly associated with TC, LDL-C, and non-HDL-C concentrations. It is possible that additional ancestry-specific *PCSK9* variants contribute further to inter-individual variation in lipid concentrations but high-density genotyping or sequencing will be needed to fully elucidate the function of ancestry specific variants.

In addition to LDL-C, HDL-C, TC, and TG, we further sought to identify the genetic determinants of non-HDL-C. A previous study identified a splice region variant in *LDLR* associated with lower non-HDL-C in a Caucasian population (18). We failed to identify any *LDLR* or other novel variant that predicted non-HDL-C levels in AAs. We did observe that variants in *PCSK9* and *APOE* associated with other lipid traits were also significantly associated with non-HDL-C. This is not unexpected since non-HDL-C contains LDL-C, TC, and remnant cholesterol. The extent to which genetic associations with non-HDL-C are driven by association with LDL-C or TC, or by an unidentified relationship with remnant cholesterol, is unclear.

Furthermore, no genetic predictor was identified for remnant cholesterol. A large-scale Mendelian randomization study suggested that remnant cholesterol is a causal risk factor for ischemic heart disease independent of HDL levels (22). However, no genetic factor has been reported to affect remnant cholesterol levels exclusively and in clinical practice, remnant cholesterol varies more than other lipid traits over years (22); therefore, identification of its genetic predictors will be challenging.

Although we replicated some genetic associations for TC, HDL-C and LDL-C, no significant association was observed for TG, even after scanning the entire MetaboChip. Compared to other lipid traits, TG is affected more by environmental factors, such as diet and medications (35). The GLGC previously estimated that known genetic predictors explained ~15% of overall variance for TC, LDL-C and HDL-C, and ~11% for TG (1,2). Compared to the other lipid traits, less TG heritability was explained by genotypes on the

MetaboChip. However, given that the MetaboChip was designed to capture most loci associated with metabolic traits in Caucasians, it may not be the best tool to estimate overall heritability of lipid traits in AAs.

We acknowledge several limitations: (1) Using the MetaboChip, we only fine-mapped regions known to be GWAS-significant regions for metabolic traits. Because most of the early GWAS that informed the construction of the MetaboChip were conducted in Caucasians, African-exclusive loci will be underrepresented. Pilot studies from the African Genome Variation Project have identified a substantial proportion of unshared (11–23%) and novel (16–24%) variants in populations of African-ancestry(36). (2) Power to detect associations with rare variants was limited. With a cohort of ~2,800 AA individuals, there was adequate power to detect associations for variants with >1% frequency and an effect size larger than 5 mg/dL. Nevertheless, power was limited for rare or private variants. Sequencing the whole exome or whole genome in an adequate number of individuals of African ancestry would assist in understanding the genetic architecture of lipid metabolism in African populations.

In conclusion, we identified both known and novel genetic variants which significantly associated lipids traits in AAs. The observations will contribute to understand genetic architecture of lipid, especially in minorities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by NIH grant (GM120523, GM109145, HL133786 and 5T32GM080178-09), American Heart Association (16FTF30130005 and 16SDG27490014) and Vanderbilt Faculty Research Scholar Fund. The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's resources, BioVU and the Synthetic Derivative, which are supported by institutional funding and by the Vanderbilt National Center for Advancing Translational Science grant 2UL1 TR000445-06 from NCATS/NIH. Existing genotypes in BioVU were funded by NIH grants RC2GM092618 from NIGMS/OD and U01HG004603 from NHGRI/NIGMS.

References

1. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010; 466(7307):707–13. [PubMed: 20686565]
2. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013 Nov; 45(11):1274–83. [PubMed: 24097068]
3. de Miranda Chagas SV, Kanaan S, Chung Kang H, Cagy M, de Abreu RE, da Silva LA, et al. Environmental factors, familial aggregation and heritability of total cholesterol, low density lipoprotein-cholesterol and high density lipoprotein-cholesterol in a Brazilian population assisted by the Family Doctor Program. Public Health. 2011 Jun; 125(6):329–37. [PubMed: 21571348]
4. Pilia G, Chen W-M, Scuteri A, Orrú M, Albai G, Dei M, et al. Heritability of cardiovascular and personality traits in 6, 148 Sardinians. PLoS Genet. 2006 Aug 25.2(8):e132. [PubMed: 16934002]
5. Dumitrescu L, Carty CL, Taylor K, Schumacher FR, Hindorff LA, Ambite JL, et al. Genetic Determinants of Lipid Traits in Diverse Populations from the Population Architecture using Genomics and Epidemiology (PAGE) Study. PLoS Genet. 2011 Jun 30.7(6):e1002138. [PubMed: 21738485]

6. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat Genet. 2005 Feb; 37(2):161–5. [PubMed: 15654334]
7. Cohen JC, Boerwinkle E, Mosley TH, Hobbs HH. Sequence Variations in PCSK9, Low LDL, and Protection against Coronary Heart Disease. N Engl J Med. 2006 Mar 23; 354(12):1264–72. [PubMed: 16554528]
8. Cohen JC. Emerging LDL therapies: Using human genetics to discover new therapeutic targets for plasma lipids. J Clin Lipidol. 2013 Jun; 7(3 Suppl):S1–5. [PubMed: 23642322]
9. Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, et al. Efficacy and Safety of Alirocumab in Reducing Lipids and Cardiovascular Events. N Engl J Med. 2015 Mar 15.0(0) null.
10. Sabatine MS, Giugliano RP, Wiviott SD, Raal FJ, Blom DJ, Robinson J, et al. Efficacy and Safety of Evolocumab in Reducing Lipids and Cardiovascular Events. N Engl J Med. 2015 Mar 15.0(0) null.
11. Cholesterol-lowering drugs get FDA advisory approval - CNN.com [Internet]. CNN. [cited 2015 Jul 21]. Available from: http://www.cnn.com/2015/06/09/health/cholesterol-lowering-drug-fda/index.html
12. Futema M, Plagnol V, Li K, Whittall RA, Neil HAW, Seed M, et al. Whole exome sequencing of familial hypercholesterolaemia patients negative for LDLR/APOB/PCSK9 mutations. J Med Genet. 2014 Aug; 51(8):537–44. [PubMed: 24987033]
13. Dewey FE, Gusarova V, O'Dushlaine C, Gottesman O, Trejos J, Hunt C, et al. Inactivating Variants in ANGPTL4 and Risk of Coronary Artery Disease. N Engl J Med. 2016 Mar 24; 374(12):1123–33. [PubMed: 26933753]
14. Cui Y, Blumenthal RS, Flaws JA, Whiteman MK, Langenberg P, Bachorik PS, et al. Non-high-density lipoprotein cholesterol level as a predictor of cardiovascular disease mortality. Arch Intern Med. 2001 Jun 11; 161(11):1413–9. [PubMed: 11386890]
15. Pischon T, Girman CJ, Sacks FM, Rifai N, Stampfer MJ, Rimm EB. Non-high-density lipoprotein cholesterol and apolipoprotein B in the prediction of coronary heart disease in men. Circulation. 2005 Nov 29; 112(22):3375–83. [PubMed: 16316964]
16. Ingelsson E, Schaefer EJ, Contois JH, McNamara JR, Sullivan L, Keyes MJ, et al. Clinical utility of different lipid measures for prediction of coronary heart disease in men and women. JAMA. 2007 Aug 15; 298(7):776–85. [PubMed: 17699011]
17. Rana JS, Boekholdt SM, Kastelein JJP, Shah PK. The Role of Non-HDL Cholesterol in Risk Stratification for Coronary Artery Disease. Curr Atheroscler Rep. 2011 Dec 28; 14(2):130–4.
18. Gretarsdottir S, Helgason H, Helgadottir A, Sigurdsson A, Thorleifsson G, Magnusdottir A, et al. A Splice Region Variant in LDLR Lowers Non-high Density Lipoprotein Cholesterol and Protects against Coronary Artery Disease. PLOS Genet. 2015 Sep 1.11(9):e1005379. [PubMed: 26327206]
19. Nioi P, Sigurdsson A, Thorleifsson G, Helgason H, Agustsdottir AB, Norddahl GL, et al. Variant ASGR1 Associated with a Reduced Risk of Coronary Artery Disease. N Engl J Med. 2016 Jun 2; 374(22):2131–41. [PubMed: 27192541]
20. Nordestgaard BG, Benn M, Schnohr P, Tybjaerg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. JAMA. 2007 Jul 18; 298(3):299–308. [PubMed: 17635890]
21. Chapman MJ, Ginsberg HN, Amarenco P, Andreotti F, Borén J, Catapano AL, et al. Triglyceride-rich lipoproteins and high-density lipoprotein cholesterol in patients at high risk of cardiovascular disease: evidence and guidance for management. Eur Heart J. 2011 Jun; 32(11):1345–61. [PubMed: 21531743]
22. Varbo A, Benn M, Tybjærg-Hansen A, Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG. Remnant Cholesterol as a Causal Risk Factor for Ischemic Heart Disease. J Am Coll Cardiol. 2013 Jan 29; 61(4):427–36. [PubMed: 23265341]
23. Crawford DC, Goodloe R, Brown-Gentry K, Wilson S, Roberson J, Gillani NB, et al. Characterization of the Metabochip in diverse populations from the International HapMap Project in the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) project. Pac Symp Biocomput Pac Symp Biocomput. 2013:188–99. [PubMed: 23424124]

24. Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, et al. The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, and Anthropometric Traits. PLoS Genet. 2012 Aug 2.8(8):e1002793. [PubMed: 22876189]
25. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther. 2008; 84(3):362–9. [PubMed: 18500243]
26. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3): 559–75. [PubMed: 17701901]
27. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006 Dec. 2(12):e190. [PubMed: 17194218]
28. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug; 38(8):904–9. [PubMed: 16862161]
29. Dumitrescu L, Ritchie MD, Brown-Gentry K, Pulley JM, Basford M, Denny JC, et al. Assessing the accuracy of observer-reported ancestry in a biorepository linked to electronic medical records. Genet Med. 2010; 12(10):648–50. [PubMed: 20733501]
30. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010 Sep 15; 26(18):2336–7. [PubMed: 20634204]
31. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011 Jan 7; 88(1):76–82. [PubMed: 21167468]
32. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet. 2009 Nov.5(11):e1000730. [PubMed: 19936222]
33. Zemunik T, Boban M, Lauc G, Janković S, Rotim K, Vatavuk Z, et al. Genome-wide association study of biochemical traits in Korcula Island, Croatia. Croat Med J. 2009 Feb; 50(1):23–33. [PubMed: 19260141]
34. Chang M, Yesupriya A, Ned RM, Mueller PW, Dowling NF. Genetic variants associated with fasting blood lipids in the U.S. population: Third National Health and Nutrition Examination Survey. BMC Med Genet. 2010; 11:62. [PubMed: 20406466]
35. Tannock, L., Bhat, A. Risk Assessment and Guidelines for the Management of High Triglycerides. In: De Groot, LJ.Chrousos, G.Dungan, K.Feingold, KR.Grossman, A.Hershman, JM., et al., editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc; 2000. [cited 2017 Apr 13]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK326745/
36. Gurdasani D, Carstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, et al. The African Genome Variation Project shapes medical genetics in Africa. Nature. 2015 Jan 15; 517(7534):327–32. [PubMed: 25470054]
37. Wu Y, Marvelle AF, Li J, Croteau-Chonka DC, Feranil AB, Kuzawa CW, et al. Genetic association with lipids in Filipinos: waist circumference modifies an APOA5 effect on triglyceride levels. J Lipid Res. 2013 Nov; 54(11):3198–205. [PubMed: 24023260]
38. Rasmussen-Torvik LJ, Pacheco JA, Wilke RA, Thompson WK, Ritchie MD, Kho AN, et al. High Density GWAS for LDL Cholesterol in African Americans Using Electronic Medical Records Reveals a Strong Protective Variant in APOE. Clin Transl Sci. 2012; 5(5):394–399. [PubMed: 23067351]
39. Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. Circ Cardiovasc Genet. 2012 Apr 1; 5(2):257–64. [PubMed: 22331829]
40. Kettunen J, Tukiainen T, Sarin A-P, Ortega-Alonso A, Tikkanen E, Lyytikäinen L-P, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet. 2012 Mar; 44(3):269–76. [PubMed: 22286219]

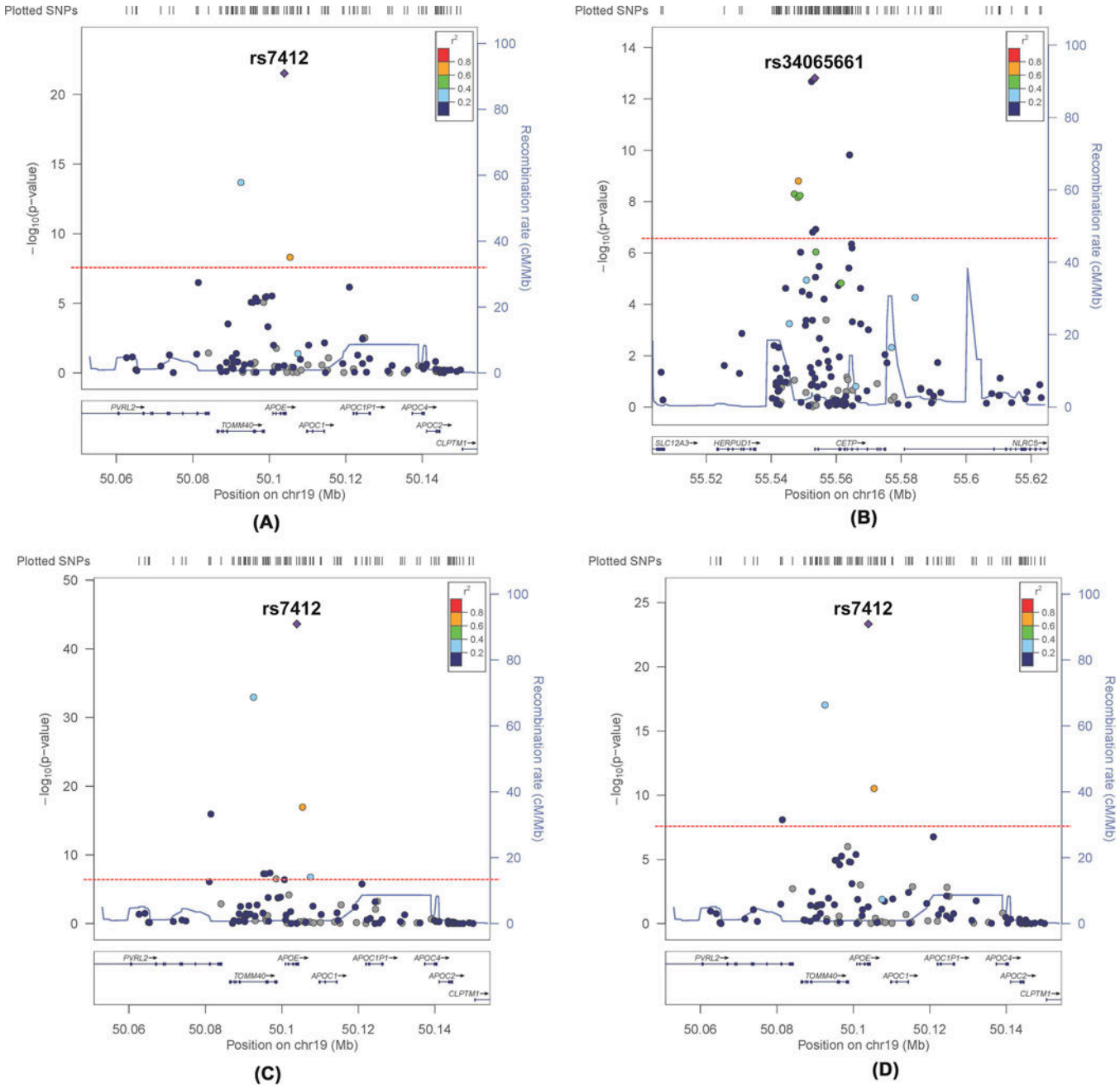

Figure 1.

Regional association plots of the genome-wide significant associations with lipids traits in AA cohort.

The plots show the genome-wide significant associated loci in BioVU AA cohort (generated using LocusZoom (http://locuszoom.sph.umich.edu/locuszoom/), including the *APOE* locus in association with TC (**A**), LDL-C (**C**) and nonHDL-C (**D**), the *CETP* locus in association with HDL-C (**B**). The RefSeq genes in the region are shown in lower panel. The red line represents genome-wide significant cutoff for MetaboChip (2.7×10^{-7}). P-value were generated using linear regression analysis.

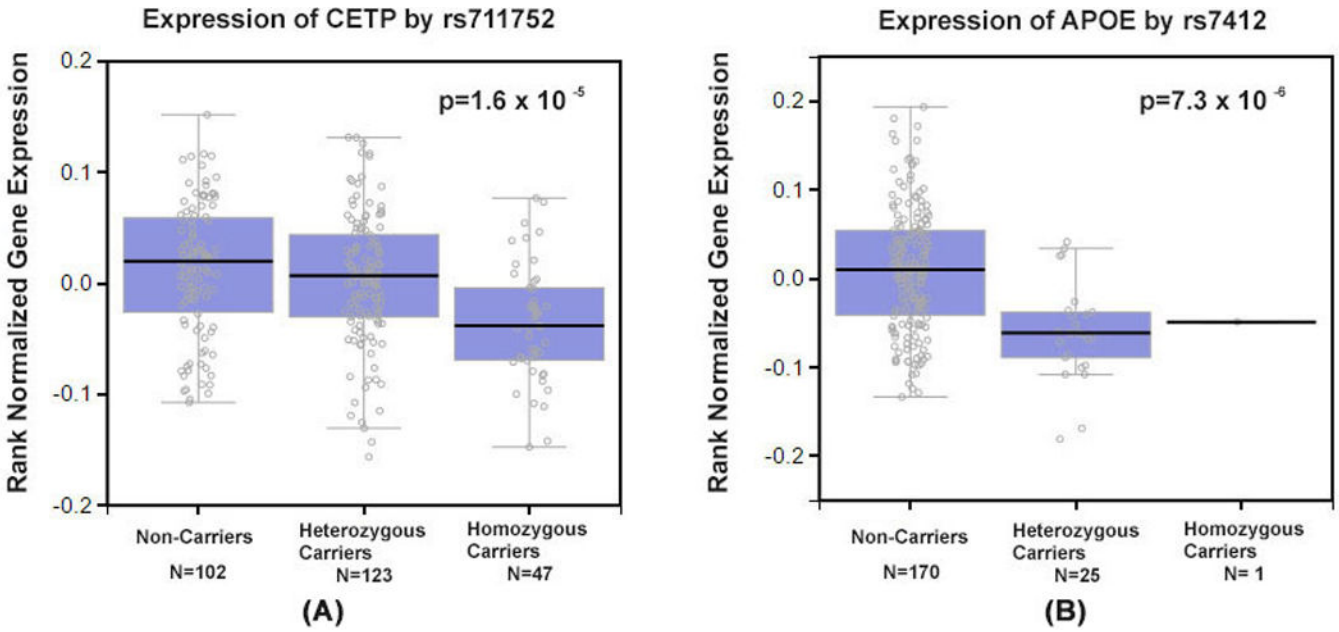

Figure 2.

Association plots for gene expression and genetic variants.

The plots show the associations between gene expression and genetic variants, including rs711752 in *CETP* gene (**A**) and rs7412 in *APOE* gene (**B**).

Table 1

Cohort characteristics

	N of individuals included	N of female (%)	Lipid levels (mg/dL)	Age at measurement	N of statin users (%)
Total Cholesterol (TC)	2778	1740 (62.6%)	185.3 ± 45.2	44.7 ± 15.4	1168 (42.0%)
High-density lipoprotein cholesterol (HDL-C)	2550	1624 (63.7%)	53.2 ± 17.5	44.9 ± 15.2	1091 (42.8%)
Low-density lipoprotein cholesterol (LDL-C)	2438	1548 (63.5%)	109.9 ± 39.9	45.2 ± 15.1	1041 (42.7%)
Triglycerides (Trigs)	2690	1681 (62.5%)	119.9 ± 87.6	44.8 ± 15.1	1139 (42.3%)
Non-HDL cholesterol (NonHDL-C)	2468	1576 (63.9%)	133.9 ± 44.8	45.5 ± 15.0	1066 (43.2%)

Data are shown as number (%) and mean ± standard deviation

Table 2

Assocaition between Global Lipid Genetics Consortium SNPs and lipid traits[1]

Lipid Traits	Chr.	rsNo.	Minor Allele	Minor Allele Freq.			BETA	STAT	Effect Size (mg/dL)			Effect of Minor Allele (mg/dL)[1 2]		P	Gene name
				Current Cohort	African American[3]	Caucasian[4]			Minor/Minor	Minor/Major	Major/Major	Caucasian	African American		
Total Cholesterol	19	rs6511720	T	0.1346	0.123	0.096	−0.05221	−5.617	172.4 ± 30.3	177.9 ± 44.2	188.1 ± 45.5		−9.71	2.15E-08	LDLR
High-density lipoprotein cholesterol	16	rs3764261	T	0.3133	0.311	0.328	0.04315	4.4	56.0 ± 19.4	54.2 ± 16.9	51.9 ± 17.1	3.39	2.16	1.13E-05	CETP
Low-density lipoprotein cholesterol	1	rs629301	G	0.3439	0.418	0.283	− 0.048	− 4.405	102.5± 38.5	108.4 ± 38.0	113.1 ± 41.8	−5.65	−5.12	1.11E-05	CELSR2/SORT1
	19	rs6511720	T	0.1346	0.123	0.096	−0.06387	−4.227	95.2 ± 25.4	104.8 ± 40.2	111.6 ± 39.8	−6.99	−7.2	2.47E-05	LDLR
Triglycerides	3	rs645040	G	0.2915	0.262	0.232	−0.05823	−3.527	104.7 ± 60.2	116.5 ± 95.5	123.7 ± 87.6	−2.2	−8.56	0.000429	MSL2L1

[1]The associations have been adjusted for age, gender and PC1-6

[2]Effects of minor allele in AAs are estimated as an additive effect of the minor allele. Effect of minor allele in Caucasians are cited from Teslovich et al. (2010)

[3]African American: African ancestry in southwest US (ASW)from 1000 genome project.

[4]Caucassian : Utah resident with Northern and Western European ancestry (CEU) from 1000 genome project.

Table 3

Additional variants associated with lipid traits by fine-mapping candidate loci

(1) Additional Associations between Total Cholesterol and GLOBAL region (p<2.1E-05)*														
CHR	SNP	Minor Allele	MAF	Original association test$			Condition on lead SNP#			gene	Function	R^2,	D′,	citations
				BETA	STAT	P	BETA	STAT	P			to lead SNPs%	to lead SNPs%	
1		A	0.009733	−0.1968	−6.431	1.50E-10	–	–	–	PCSK9	stop-gain	–	–	
19		T	0.1061	−0.0986	−9.788	3.09E-22	–	–	–	APOE		–	–	37–40
(2) Additional Associations between HDL-C and GLOBAL region (p<1.3E-05)*														
CHR	SNP	Minor Allele	MAF	Original association test$			Condition on lead SNP#			gene	Function	R^2, to lead SNPs %	D′, to lead SNPs %	citations
				BETA	STAT	P	BETA	STAT	P					
8		C	0.138	−0.06145	−4.489	7.49E-06	–	–	–	LPL		–	–	
16		G	0.06532	0.1335	7.428	1.53E-13	–	–	–	CETP	missense	–	–	
16	rs17231534	A	0.1208	0.06186	4.458	8.67E-06	0.06919	5.036	5.12E-07	CETP		0.0100	1.0000	
16	rs7499892	T	0.3597	−0.05995	−6.433	1.51E-10	−0.04863	−5.166	2.59E-07	CETP		0.0300	0.8760	32,33
16	rs9930761	C	0.1085	0.07369	5.064	4.42E-07	0.08011	5.564	2.93E-08	CETP		0.0050	0.7500	
16	rs5883	T	0.1071	0.07295	4.996	6.29E-07	0.07916	5.48	4.70E-08	CETP	synonynous	0.0050	0.7490	
(3) Additioanal Association between LDL-C and GLOBAL region (p<2.5E-05)*														
CHR	SNP	Minor Allele	MAF	Original association test$			Condition on lead SNP#			gene	Function	R^2, to lead SNPs %	D′, to lead SNPs %	citations
				BETA	STAT	P	BETA	STAT	P					
1		A	0.009733	−0.3722	−7.394	1.99E-13	–	–	–	PCSK9	stop-gain	–	–	
1	rs28362261	G	0.01687	−0.1743	−4.265	2.08E-05	−0.1822	−4.509	6.83E-06	PCSK9	missense	0.0000	1.0000	
19		T	0.1061	−0.2285	−14.27	2.48E-44	–	–	–	APOE		–	–	37–40

*All lipid traits were natural log transformed.

$The associtions have been adjusted for age, gender and 6 PCs

#The associtions have been adjusted for age, gender, 6 PCs and lead SNPs in the regions (rs28362286 for PCSK9, rs7412 for APOE, rs34065661 for CETP, bold)

%The linkage to lead SNPs in the region (rs28362286 for PCSK9, rs7412 for APOE, rs34065661 for CETP, bold)

MAF - minor allele Frequency, BETA - regression coefficient, STAT - Coefficient t-statistic

Table 4

Additional variants associated with lipid traits by scanning entire MetaboChip

(1) Additional Associations between HDL-C *												
CHR	SNP	MAF	Minor Allele	Original association test$			Condition on lead SNP#			gene	R^2, to lead SNPs %	D', to lead SNPs %
				BETA	STAT	P	BETA	STAT	P			
8	[redacted]	0.1232	G	−0.07882	−5.533	3.50E-08	–		–	close to LPL and SLC18A2	–	–
16	rs4783961	0.4356	A	0.06607	7.386	2.09E-13	0.04976	5.293	1.32E-07	CETP	0.0910	1.0000
(2) Additioanal Association between LDL-C *												
CHR	SNP	MAF	Minor Allele	Original association test $			Condition on lead SNP #			gene	R^2, to lead SNPs %	D', to lead SNPs %
				BETA	STAT	P	BETA	STAT	P			
1	[redacted]	0.3359	G	−0.0555	−5.155	2.76E-07	–		–	CELSR2 (SORTI region)	–	–

*All lipid traits were natural log transformed.

$The associtions have been adjusted for age, gender and 6 PCs

#The associtions have been adjusted for age, gender, 6 PCs and lead SNPs in the regions (rs28362286 for PCSK9, rs7412 for APOE, rs34065661 for CETP, rs4389957 for LPL)

%The linkage to lead SNPs in the region (rs28362286 for PCSK9, rs7412 for APOE, rs34065661 for CETP and rs4389957 for LPL, bold)

MAF - minor allele Frequency, BETA - regression coefficient, STAT - Coefficient t-statistic

Table 5

Genetic variants associated with nonHDL cholesterol *

CHR	SNP	MAF	Minor Allele	BETA	STAT	P	Gene	Also Associated with
1	rs28362286	0.009733	A	−0.367	−8.029	1.5E-15	PCSK9	TC,
19	rs7254892	0.1274	A	−0.08772	−6.588	5.5E-11	APOE	LDL
19	rs61679753	0.11	A	−0.1424	−9.878	1.4E-22	APOE	TC,
19	rs7412	0.1061	T	−0.1719	−11.95	5.6E-32	APOE	TC,LDL
19	rs75627662	0.1643	T	−0.09318	−7.656	2.8E-14	APOE	TC,LDL

* All lipid traits were natural log transformed.

The associations have been adjusted for age, gender and 6 PCs