Discovery and Refinement of Loci Associated with Lipid Levels

A full list of authors and affiliations appears at the end of the article.

[#] These authors contributed equally to this work.

Abstract

Low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides, and total cholesterol are heritable, modifiable, risk factors for coronary artery disease. To identify new loci and refine known loci influencing these lipids, we examined 188,578 individuals using genome-wide and custom genotyping arrays. We identify and annotate 157 loci associated with lipid levels at $P < 5 \times 10^{-8}$, including 62 loci not previously associated with lipid levels in humans. Using dense genotyping in individuals of European, East Asian, South Asian, and African ancestry, we narrow association signals in 12 loci. We find that loci associated with blood lipids are often associated with cardiovascular and metabolic traits including coronary artery disease, type 2 diabetes, blood pressure, waist-hip ratio, and body mass index. Our results illustrate the value of genetic data from individuals of diverse ancestries and provide insights into biological mechanisms regulating blood lipids to guide future genetic, biological, and therapeutic research.

Introduction

Blood lipids are heritable, modifiable, risk factors for coronary artery disease (CAD)[1,2], a leading cause of death[3]. Human genetic studies of lipid levels can identify targets for new therapies for cholesterol management and prevention of heart disease, and can complement animal studies[4,5]. Studies of naturally occurring genetic variation can proceed through large-scale association analyses focused on unrelated individuals or through investigation of Mendelian forms of dyslipidemia in families[6]. We previously identified 95 loci associated with blood lipids, accounting for ~10-12% of the total trait variance[4] and showed that variants with small effects can point to pathways and therapeutic targets that enable clinically-important changes in blood lipids[4,7].

Here, we report on studies of naturally occurring variation in 188,578 European-ancestry individuals and 7,898 non-European ancestry individuals. Our analyses identify 157 loci associated with lipid levels at $P < 5 \times 10^{-8}$, including 62 new loci. Thirty of the 62 loci do not include genes implicated in lipid biology by previous literature. We tested lipid-associated SNPs for association with mRNA expression levels, carried out pathway analyses to uncover relationships between loci, and compared the locations of lipid-associated SNPs with those of genes and other functional elements in the genome. These results provide direction for biological and therapeutic research into risk factors for CAD.

Results

Novel loci associated with blood lipid levels

We examined subjects of European ancestry, including 94,595 individuals from 23 studies

AUTHOR CONTRIBUTIONS

Writing and Analysis Group

G.R.A., M.B., L.A.C., P.D., P.W.F., S.K., K.L.M., E.I., G.M.P., S.S.R., S.R., M.S.S., E.M.S., S.S., C.J.W. (Lead). E.M.S. and S.S. performed meta-analysis and E.M.S., S.S., G.M.P., M.B., J.C., S.G., A.G., and S. K. performed bioinformatics analyses. E.M.S. and S.S. prepared the tables, figures and supplementary material. C.J.W. led the analysis and bioinformatics efforts. E.I. and K.M. led the biological interpretation of results. C.J.W. and G.R.A. wrote the manuscript. All analysis and writing group authors extensively discussed the analysis, results, interpretation and presentation of results.

All authors contributed to the research and reviewed the manuscript.

Design, management and coordination of contributing cohorts

(ADVANCE) T.L.A.; (AGES Reykjavik study) T.B.H., V.G.; (AIDHS/SDS) D.K.S.; (AMC-PAS) P.D., G.K.H.; (Amish GLGC) A.R.S.; (ARIC) E.B.; (B58C-WTCCC & B58C-T1DGC) D.P.S.; (B58C-Metabochip) C.M.L., C.P., M.I.M.; (BLSA) L.F.; (BRIGHT) P.B.M.; (CHS) B.M.P., J.I.R.; (CLHNS) A.B.F., K.L.M., L.S.A.; (CoLaus) P.V.; (deCODE) K.S., U.T.; (DIAGEN) P.E.S., S.R.B.; (DILGOM) S.R.; (DPS) M.U.; (DR's EXTRA) R.R.; (EAS) J.F.P.; (EGCUT (Estonian Genome Center of University of Tartu)) A.M.; (ELY) N.W.; (EPIC) N.W., K.K.; (EPIC_N_OBSET GWAS) E.H.Young; (ERF) C.M.V.; (ESS (Erasmus Stroke Study)) P.J.K.; (Family Heart Study FHS) I.B.B.; (FBPP) A.C., R.S.C., S.C.H.; (FENLAND) R.L., N.W.; (FIN-D2D 2007) A.K., L.M.; (FINCAVAS) M.K.; (Framingham) L.A.C., S.K., J.M.O.; (FRISCII) A.S., L.W.; (FUSION GWAS) K.L.M., M.B.; (FUSION stage 2) F.S.C., J.T., J.S.; (GenomEUTwin) J.B.W., N.G.M., K.O.K., V.S., J.K., A.J., D.I.B., N.P., T.D.S.; (GLACIER) P.W.F.; (Go-DARTS) A.D.M., C.N.P.; (GxE/Spanish Town) B.O.T., C.A.M., F.B., J.N.H., R.S.C.; (HUNT2) K.H.; (IMPROVE) U.D., A.H., E.T., S.E.H.; (InCHIANTI) S.B.; (KORAF4) C.G.;(LifeLines) B.H.W.; (LOLIPOP) J.S.K., J.C.C.; (LURIC) B.O.B.; W.M.; (MDC) L.C.G., S.K.; (METSIM) J.K., M.L.; (MICROS) P.P.P.; (MORGAM) D.A., J.F.; (MRC/UVRI GPC GWAS) P.K., G.A., J.S., E.H.Y.; (MRC National Survey of Health & Development) D.K.; (NFBC1986) M-R.J.; (NSPHS) U.G.; (ORCADES) H.C.; (PARC) Y.I.C., R.M.K., J.I.R.; (PIVUS) E.I., L.L.; (PROMIS) J.D., P.D., D.S.; (Rotterdam Study) A.H., A.G.U.; (SardiNIA) G.R.A.; (SCARFSHEEP) A.H., U.D.; (SEYCHELLES) M.B., M.B.; P.B.; (SUVIMAX) P.M.; (Swedish Twin Reg.) E.I., N.L.P.; (TAICHI) T.L.A., Y.I.C., C.A.H., T.Q., J.I.R., W.H.S.; (THISEAS) G.D., P.D.; (Tromsø) I.N.; (TWINGENE) U.D., E.I.; (ULSAM) E.I.; (Whitehall II) A.H., M.K.

Genotyping of contributing cohorts

(ADVANCE) D.A.; (AIDHS/SDS) L.F.B., M.L.G.; (AMC-PAS) P.D., G.K.H.; (B58C-WTCCC & B58C-T1DGC) W.L.M.; (B58C-Metabochip) M.I.M.; (BLSA) D.H.; (BRIGHT) P.B.M.; (CHS) J.I.R.; (DIAGEN) N.N., G.M.; (DILGOM) A.P.; (DR's EXTRA) T.A.L.; (EAS) J.F.W.; (EGCUT (Estonian Genome Center of University of Tartu)) T.E.; (EPIC) P.D.; (EPIC_N_SUBCOH GWAS) I.B.; (ERF) C.M.V.; (ESS (Erasmus Stroke Study)) C.M.V.; (FBPP) A.C., G.B.E.; (FENLAND) M.S.S.; (FIN-D2D 2007) A.J.S.; (FINCAVAS) T.L.; (Framingham) J.M.O.; (FUSION stage 2) L.L.B.; (GLACIER) I.B.; (Go-DARTS) C.G., C.N.P., M.I.M.; (IMPROVE) A.H.; (KORAF3) H.G., T.I.; (KORAF4) N.K.; (LifeLines) C.W.; (LOLIPOP) J.S.K., J.C.C.; (LURIC) M.E.K.; (MDC) B.F.V., R.D.; (MICROS) A.A.H.; (MORGAM) L.T., P.B.; (MRC/UVRI GPC GWAS) M.S.S.; (MRC National Survey of Health & Development) A.W., D.K., K.K.O.; (NFBC1986) A-L.H., M.J, M.M., P.E., S.V.; (NSPHS and FRISCII) Å.J.; (ORCADES) H.C.; (PARC) M.O.G., M.R.J., J.I.R.; (PIVUS) E.I., L.L.; (PROMIS) P.D., K.S.; (Rotterdam Study) A.G.U., F.R.; (SardiNIA) R.N.; (SCARFSHEEP) B.G., R.J.S.; (SEYCHELLES) F.M., G.B.E.; (Swedish Twin Reg.) E.I., N.L.P.; (TAICHI) D.A., T.L.A., E.K., T.Q., L.L.W.; (THISEAS) P.D.; (TWINGENE) A.H., E.I.; (ULSAM) E.I.; (WGHS) D.I.C., P.M.R.; (Whitehall II) A.H., C.L., M.K., M.K.

Phenotype definition of contributing cohorts

(ADVANCE) C.I.; (AGES Reykjavik study) T.B.H., V.G.; (AIDHS/SDS) L.F.B.; (AMC-PAS) J.J.K.; (Amish GLGC) A.R.S., B.D.M.; (B58C-WTCCC & B58C-T1DGC) D.P.S.; (B58C-Metabochip) C.P., E.H.; (BRIGHT) P.B.M.; (CHS) B.M.P.; (CoLaus) P.V.; (deCODE) G.I.E., H.H., I.O.; (DIAGEN) G.M.; (DILGOM) K.S.; (DPS) J.L.; (DR's EXTRA) P.K.; (EAS) J.L.B.; (EGCUT (Estonian Genome Center of University of Tartu)) A.M.; (EGCUT (Estonian Genome Center of University of Tartu)) K.F.; (ERF and Rotterdam Study) A.H.; (ERF) C.M.V; (ESS (Erasmus Stroke Study)) E.G.V., H.M.D., P.J.K.; (FBPP) A.C., R.S.C., S.C.H.; (FINCAVAS) T.V.N.; (Framingham) S.K., J.M.O.; (GenomEUTwin: MZGWA) J.B.W.; (GenomEUTwin-FINRISK) V.S.; (GenomEUTwin-FINTWIN) J.K., K.H.; (GenomEUTwin-GENMETS) A.J.; (GenomEUTwin-NLDTWIN) G.W.; (Go-DARTS) A.S.D., A.D.M., C.N.P., L.A.D.; (GxE/Spanish Town) C.A.M., F.B.; (IMPROVE) U.D.; A.H., E.T.; (KORAF3) C.M.; (KORAF4) A.D.; (LifeLines) L.J.; (LOLIPOP) J.S.K., J.C.C.; (LURIC) H.S.; (MDC) L.C.G.; (METSIM) A.S.; (MORGAM) G.C.; (MRC/UVRI GPC GWAS) R.N.N.; (MRC National Survey of Health & Development) D.K.; (NFBC1986) A.R., A-L.H., A.P., M-R.J.; (NSPHS and FRISCII) Å.J.; (NSPHS) U.G.; (ORCADES) S.H.W.; (PARC) Y.I.C., R.M.K.; (PIVUS) E.I., L.L.; (PROMIS) D.F.F.; (Rotterdam Study) A.H.; (SCARFSHEEP) U.D., B.G.; (SEYCHELLES) M.B., M.B., P.B.; (Swedish Twin Reg.) E.I., N.L.P.; (TAICHI) H.C., C.A.H., Y.H., E.K., S.L., W.H.S.; (THISEAS) G.D., M.D.; (Tromsø) T.W.; (TWINGENE) U.D., E.I.; (ULSAM) E.I.; (WGHS) P.M.R.; (Whitehall II) M.K.

Primary analysis from contributing cohorts

(ADVANCE) L.L.W.; (AIDHS/SDS) R.S.; (AMC-PAS) S.K.; (Amish GLGC) J.R.O., M.E.M.; (ARIC) K.A.V.; (B58C-Metabochip) C.M.L., E.H., T.F.; (B58C-WTCCC & B58C-T1DGC) D.P.S.; (BLSA) T.T.; (BRIGHT) T.J.; (CLHNS) Y.W.; (CoLaus) J.S.B.; (deCODE) G.T.; (DIAGEN) A.U.J.; (DILGOM) M.P.; (EAS) R.M.F.; (DPS) A.U.J.; (DR 'S EXTRA) A.U.J.; (EGCUT (Estonian Genome Center of University of Tartu)) E.M., K.F., T.E.; (ELY) D.G.; (EPIC) K.S., D.G.; (EPIC_N_OBSET GWAS) E.Y., C.L.; (EPIC_N_SUBCOH GWAS) N.W.; (ERF) A.I.; (ESS (Erasmus Stroke Study)) C.M.V., E.G.V.; (EUROSPAN) A.D.; (Family Heart Study FHS) I.B.B., M.F.F.; (FBPP) A.C., G.B.E.; (FENLAND) T.P., C.P.; (FENLAND GWAS) J.H.Z., J.L.; (FIN-D2D 2007) A.U.J.; (FINCAVAS) L.L.; (Framingham) L.A.C., G.M.P.; (FRISCII and NSPHS) Å.J.; (FUSION stage 2) T.M.T.; (GenomEUTwin-FINRISK) J.K.; (GenomEUTwin-FINTWIN) K.H.; (GenomEUTwin-GENMETS) I.S.; (GenomEUTwin-SWETWIN) P.K.M.; (GenomEUTwin-UK-TWINS) M.M.; (GLACIER) D.S.; (GLACIER) P.W.F.; (Go-DARTS) C.N.P., L.A.D.; (GxE/Spanish Town) C.D.P.; (HUNT) A.U.J.; (IMPROVE) R.J.S.; (InCHIANTI) T.T.; (KORAF3) M.M.; (KORAF4) A.P.; (LifeLines) I.M.N.; (LOLIPOP) W.Z.; (LURIC) M.E.K.; (MDC) B.F.V.; (MDC) P.F., R.D.; (METSIM) A.U.J.; (MRC/UVRI GPC GWAS) R.N.N.; (MRC National Survey of Health & Development) A.W., J.L.; (NFBC1986) M.K., I.S., S.K.S.; (NSPHS and FRISCII) Å.J.; (PARC) X.L.; (PIVUS) C.S., E.I.; (PROMIS) J.D., D.F.F., K.S.; (Rotterdam Study) A.I.; (SardiNIA) C.S., J.L.B., S.S.; (SCARFSHEEP) R.J.S.; (SEYCHELLES) G.B.E., M.B.; (SUVIMAX) T.J.; (Swedish Twin Reg.) C.S., E.I.; (TAICHI) D.A., T.L.A., H.C., M.G., C.A.H., T.Q., L.L.W; (THISEAS) S.K.; (Tromsø) A.U.J.; (TWINGENE) A.G., E.I.; (ULSAM) C.S., E.I., S.G.; (WGHS) D.I.C.; (Whitehall II) S.S.

genotyped with GWAS arrays[4] and 93,982 individuals from 37 studies genotyped with the Metabochip array[8] (Supplementary Table 1 and Supplementary Fig. 1). The Metabochip includes variants representing promising loci from our previous GWAS (14,886 SNPs) and from GWAS of other CAD risk factors and related traits (50,459 SNPs), variants from the 1000 Genomes Project[9] and focused resequencing[10] efforts in 64 previously associated loci (28,923 SNPs), and fine-mapping variants in 181 loci associated with other traits (93,308 SNPs). In cases where Metabochip and GWAS array data were available for the same individuals, we used Metabochip data to ensure key variants were directly genotyped, rather than imputed.

We excluded individuals known to be on lipid lowering medications and evaluated the additive effects of each SNP on blood lipid levels after adjusting for age and sex. Genomic control values[11] for the initial meta-analyses were $1.10 - 1.15$, low for a sample of this size, indicating that population stratification should have only a minor impact on our results (Supplementary Fig. 2). After genomic control correction, 157 loci associated with blood lipid levels were identified ($P < 5 \times 10^{-8}$), including 62 new loci (Tables 1A-D, Figure 1, Supplementary Tables 2 and 3). Loci were >1 Mb apart and nearly independent ($r^2 < 0.10$). Of the 62 novel loci, 24 demonstrated the strongest evidence of association with HDL cholesterol, 15 with LDL cholesterol, 8 with triglyceride levels, and 15 with total cholesterol (Supplementary Fig. 3). Several of these loci were validated by a similar extension based on GLGC GWAS results [12].

The effects of newly identified loci were generally smaller than in earlier GWAS (Supplementary Fig. 4). For the 62 newly identified variants, trait variance explained in the Framingham offspring were 1.6% for HDL cholesterol, 2.1% for triglycerides, 2.4% for LDL cholesterol, and 2.6% for total cholesterol.

Overlap of genetic discoveries and prior knowledge

To investigate connections between our new loci and known lipid biology, we first catalogued genes within 100 kb of the peak associated SNPs and searched PubMed and OMIM for occurrences of these gene names and their aliases in the context of relevant keywords. After manual curation, we identified at least one strong candidate in 32 of the 62 loci (52%) (Supplementary Table 4). For the remaining 30 loci, we found no literature support for the role of a nearby gene on blood lipid levels. This search highlighted genes whose connections to lipid metabolism have been extensively documented in mouse models (such as *VLDLR*[13] and *LRPAP1*[13]) and human cell lines (such as *VIM*[14]), as well as candidates whose connection to lipid levels is more recent, such as *VEGFA*. For the latter, recent studies of *VEGFB* have suggested that vascular endothelial growth factors have an unexpected role in the targeting of lipids to peripheral tissues[15], which we corroborate by associating variants near *VEGFA* with blood triglyceride and HDL levels.

Multiple types of evidence supported several literature candidates (Supplementary Table 2). For example, *VLDLR* is categorized by Gene Ontology[16] in the retinoid × nuclear receptor (RXR) activation pathway, which also includes genes (*APOB, APOE, CYP7A1, APOA1, HNF1A, HNF4A*) in previously implicated loci[4]. However, since these additional sources of evidence build on overlapping knowledge they are not truly independent.

To estimate the probability of finding 32 literature supported candidates after automated search and manual review of results, we repeated our text-mining literature search using 100 permutations of SNPs matched for allele frequency, distance to the nearest gene, and number of linkage disequilibrium proxies. To approximate hand-curation of the text-mining results, we focused on genes implicated by 3 or more publications (25 in observed data, 8.7 on average in control SNP sets, $P = 8 \times 10^{-8}$).

Pathway analyses

We performed a gene-set enrichment analysis, using MAGENTA[17], to evaluate over-representation of biological pathways among associated loci. Across the 157 loci, MAGENTA identified 71 enriched pathways. These pathways included at least one gene in 20 of our newly identified loci (Supplementary Table 5). Examples include *DAGLB* (connected to previously associated loci by genes in the triglyceride lipase activity pathway), *INSIG2* (connected by the cholesterol and steroid metabolic process pathways), *AKR1C4* (connected by the steroid metabolic process and bile acid biosynthesis pathways), *VLDLR* (connected by the retinoic × receptor activation and lipid transport pathways, among others), *PPARA, ABCB11*, and *UGT1A1* (three genes assigned to pathways implicated in activation of nuclear hormone receptors, which play an important role in lipid metabolism through the transcriptional regulation of genes in sterol metabolic pathways[18]). Among the 16 loci where literature review and pathway analysis both suggested a candidate, the predictions overlapped 14 times (Supplementary Table 2; by chance, we expect 6.6 overlapping predictions, $P = 1 \times 10^{-5}$).

Protein-protein interactions

We assessed evidence for physical interactions between proteins encoded near our associated SNPs using DAPPLE[19]. We found an excess of direct protein-protein interactions for genes in loci associated with LDL (10 interactions, $P = 0.0002$), HDL (8 interactions, $P = 0.002$), and total cholesterol (6 interactions, $P = 0.017$), but not for triglycerides (2 interactions, $P = 0.27$) (Supplementary Fig. 5). Most of the interactions involved genes at known loci (such as the interaction network connecting *PLTP, APOE, APOB*, and *LIPC*) or highlighted the same genes as literature and pathway analyses (such as those connecting *VLDLR, APOE, APOB, CETP*, and *LPL*). Among novel loci, we identified a link between *AKT1* and *GSK3B. GSK3B* has been shown to play a role in energy metabolism[20] and its activity is regulated by *AKT1* through phosphorylation[21]. Literature review also supported a role in blood lipid levels for these two genes.

Regulation of gene expression by associated variants

Many complex trait associated variants act through the regulation of gene expression. We examined whether our 62 novel variants were associated with expression levels of nearby genes in liver, omental fat, or subcutaneous fat. Fifteen were associated with expression of a nearby transcript with $P < 5 \times 10^{-8}$ (Supplementary Table 6) and, in seven, the lipid-associated variant was in strong disequilibrium with the strongest expression-quantitative trait locus (eQTL) for the region ($r^2 > 0.8$). In three of these loci, literature search also prioritized candidate genes. In all three, eQTL analysis and literature review identified the

same candidate (*DAGLB, SPTLC3,* and *PXK, P* = 0.05). For the remaining four loci (near *RBM5, ADH5, TMEM176A*, and *GPR146*), analysis of expression levels identified candidates that were not supported by literature or pathway analyses.

Coding variation

In some loci where previous coding variant association studies were inconclusive, we now find convincing evidence of association, demonstrating the benefits of the large sample sizes achievable by collaboration. For example, in the *APOH* locus[22], our most strongly associated variant is rs1801689 (*APOH* C325G, $P = 1 \times 10^{-11}$ for LDL cholesterol). Overall, at 15 of the 62 new loci, there is at least one nonsynonymous variant within 100kb and in strong ($r^2 > 0.8$) linkage disequilibrium with the index SNP (Supplementary Table 7)(18 loci with no restrictions on distance). This ~30% overlap between associated loci and coding variation is similar to that in other complex traits[9]. Unexpectedly, in the 11 loci where a candidate was suggested by literature review and by coding variation, the two coincided seven times (P = 0.03 compared to expected chance overlap of 3.8 times); thus, agreement between literature and coding variation was less significant than for eQTL and pathway analysis or protein-protein interactions.

Overlap between association signals and regulators of transcription in liver

Despite our efforts, 18 of the 62 new loci remain without prioritized candidate genes. The liver is an important hub of lipid biosynthesis and there is evidence that lipid loci might be associated with changes in gene regulation in liver cells[23]. Using ENCODE data[23], we evaluated whether associated SNPs overlapped experimentally annotated functional elements identified in HepG2 cells, a commonly used model of human hepatocytes. To determine significance, we generated 100,000 lists of permuted SNPs, matched for minor allele frequency, distance to the nearest gene, and number of SNPs in $r^2 > 0.8$ (described in Methods). In HepG2 cells, lipid-associated SNPs were enriched in eight of the 15 functional chromatin states defined by Ernst *et al.*[24] ($P < 1 \times 10^{-5}$; Supplementary Table 8). The strongest enrichment was in regions with "strong enhancer activity" (3.7-fold enrichment, $P = 2 \times 10^{-25}$; Supplementary Table 9). In the other eight cell types examined by Ernst *et al.*, no more than three functional chromatin states showed evidence for enrichment (and, when present, enrichment was weaker).

We proceeded to investigate the overlap between lipid loci and functional marks in HepG2 cells in more detail (Supplementary Table 9). Notable regulatory elements showing significant overlap with lipid loci included histone marks associated with active regulatory regions (H3K27ac, $P = 3 \times 10^{-20}$; H3K9ac, $P = 3 \times 10^{-22}$), promoters (H3K4me3, $P = 2 \times 10^{-15}$, H3K4me2, $P = 8 \times 10^{-12}$), transcribed regions (H3K36me3, $P = 4 \times 10^{-14}$), indicators of open chromatin (FAIRE, $P = 5 \times 10^{-9}$; DNase, $P = 2 \times 10^{-4}$), and regions that interact with transcription factors HNF4A ($P = 6 \times 10^{-10}$) and CEBP/B ($P = 1 \times 10^{-5}$). Overall, 56 of our 62 new loci contained at least one SNP that overlaps a functional mark[24] and/or chromatin state[23] highlighted in Supplementary Table 9, including all but 3 of the loci where no candidates were suggested by literature review or analyses of pathways, coding variation, or gene expression (Supplementary Table 10).

Initial fine-mapping of 65 lipid-associated loci

Previous fine-mapping of five LDL-associated lipid loci found that variants showing the strongest association were often substantially different in frequency and effect size from those identified in GWAS[10]. Metabochip genotypes enabled us to carry out an initial fine-mapping analysis for 65 loci: 60 selected for fine-mapping based on our previous study[4] and 5 nominated for fine-mapping because of association to other traits.

For each of these loci, we identified the most strongly associated Metabochip variant and evaluated whether it (a) reached genome-wide significant evidence for association (to avoid chance fluctuations in regions where the signal was relatively weak) and (b) was different from the GWAS index SNP in terms of frequency and effect size (operationalized to $r^2 < 0.8$ with the GWAS index SNP). In the European samples, fine-mapping identified eight loci where the fine-mapping signal was clearly different from the GWAS signal (Supplementary Table 11). The two largest differences were at the loci near *PCSK9* (top GWAS variant with minor allele frequency f = 0.24 and $P = 9 \times 10^{-24}$; fine-mapping variant with f = 0.03, $P = 2 \times 10^{-136}$) and *APOE* (GWAS variant f = 0.20, $P = 3 \times 10^{-44}$, fine-mapping variant f = 0.07, $P = 3 \times 10^{-651}$), consistent with Sanna *et al*[10]. Large differences were also observed near *LRP4* (GWAS f = 0.17, $P = 8 \times 10^{-14}$; fine-mapping f = 0.35, $P = 1 \times 10^{-26}$), *IGF2R* (GWAS f = 0.16, $P = 7 \times 10^{-9}$; fine-mapping f = 0.37, $P = 2 \times 10^{-13}$), *NPC1L1* (GWAS f = 0.27, $P = 2 \times 10^{-5}$; fine-mapping f = 0.24, $P = 1 \times 10^{-12}$), *ST3GAL4* (GWAS f = 0.26, $P = 2 \times 10^{-6}$; fine-mapping f = 0.07, $P = 6 \times 10^{-11}$), *MED1* (GWAS f = 0.37, $P = 3 \times 10^{-5}$; fine-mapping f = 0.24, $P = 2 \times 10^{-10}$), and *COBLL1* (GWAS f = 0.12, $P = 2 \times 10^{-6}$; fine-mapping f = 0.11, $P = 6 \times 10^{-9}$). Thus, although the large changes observed by Sanna *et al*[10] after fine-mapping are by no means unique, they are not typical. Except for the R46L variant in *PCSK9*, the variants showing strongest association in fine-mapped loci all had minor allele frequency > .05.

We also attempted fine-mapping in African (N=3,263), East Asian (N=1,771), and South Asian (N=4,901) ancestry samples. Despite comparatively small samples, ancestry-specific analyses identified SNPs clearly distinct from the original GWAS variant in five loci (Supplementary Table 11). These were: *APOE*, consistent with European ancestry analyses above; three loci where differences in linkage disequilibrium between populations enabled fine-mapping in African (*SORT1, LDLR*) or East Asian (*APOA5*) ancestry samples; and *CETP*, where an African-specific variant was present. For *CETP, SORT1*, and *APOA5*, results are consistent with other fine-mapping and functional studies[7,7,25,26].

Association of lipid loci with metabolic and cardiovascular traits

To evaluate the role of the 157 loci identified here on related traits, we evaluated the most strongly associated SNPs for each locus in genetic studies of coronary artery disease (CAD, N=114,590 including 37,653 cases)[27,28], type 2 diabetes (T2D, N=47,117 including 8,130 cases)[29], body mass index (BMI, N=123,865 individuals)[30] and waist-hip ratio (WHR, N=77,167 individuals)[31], systolic and diastolic blood pressure (SBP and DBP, N=69,395 individuals)[32], and fasting glucose (N=46,186 non-diabetics)[33]. We observed an excess of SNPs nominally associated ($P < 0.05$) with all these traits: a 5.1 fold excess for CAD (40 nominally significant loci, $P = 2 \times 10^{-19}$), a 4.1 fold excess for BMI (32 loci, $P = 1 \times 10^{-11}$),

3.7 fold excesses for DBP (29 loci, $P = 1\times10^{-9}$), a 3.4 fold excess for WHR (27 loci, $P = 1\times10^{-9}$), a 2.5 fold excess for SBP (20 loci, $P = 1\times10^{-4}$), a 2.3 fold excess for T2D (18 loci, $P = 0.001$), and a 2.2 fold excess for fasting glucose (17 loci, $P = 3\times10^{-3}$) (Supplementary Table 12). Interestingly, among the novel loci, we observed greater overlap with BMI, SBP, and DBP (9 overlapping loci each) than with CAD (8 overlapping loci). Among new loci, the two SNPs showing strongest association to CAD map near *RBM5* (rs2013208, $P_{HDL} = 9\times10^{-12}$, $P_{CAD} = 7\times10^{-5}$) and *CMTM6* (rs7640978, $P_{LDL} = 1\times10^{-8}$, $P_{CAD} = 4\times10^{-4}$).

We tested whether the LDL-, total cholesterol- or triglyceride-increasing allele, or HDL-decreasing allele was associated with increased risk of cardiovascular disease or related metabolic outcomes; the direction of effect of each locus was categorized according to the primary association signal at the locus, as in Tables 1A-D. We observed association with increased CAD risk (104/149, $P = 1\times10^{-6}$), SBP (96/155, $P = 2.7\times10^{-3}$) and WHR adjusted for BMI (92/154, $P = 0.019$). There were many instances where a single locus was associated with many traits. These included variants near *FTO*, consistent with previous reports[34]; near *VEGFA* (associated with triglyceride levels, CAD, T2D, SBP, and DBP), near *SLC39A8* (associated with HDL cholesterol, BMI, SBP, and DBP), and near *MIR581* (associated with HDL cholesterol, BMI, T2D, and DBP). In some cases, like *FTO*, a strong association with BMI or another phenotype generates weaker association signals for other metabolic traits[34]. In other cases, like *SORT1*, a primary effect on lipid levels may mediate secondary association with other traits, like CAD[7].

Association of individual lipids with coronary artery disease

Epidemiological studies consistently show high total cholesterol and LDL cholesterol levels are associated with increased risk of CAD, whereas high HDL cholesterol levels are associated with reduced risk of CAD[35]. In genetic studies, the connection between LDL cholesterol and CAD is clear, whereas the results for HDL cholesterol levels are more equivocal[36-38]. In our data, trait increasing alleles at the loci showing strongest association with LDL cholesterol (31 loci), triglycerides (30 loci), or total cholesterol (38 loci) were associated with increased risk of CAD ($P = 2\times10^{-12}$, $P = 2\times10^{-16}$, and $P = 0.006$). Conversely, trait decreasing alleles at loci showing the strongest association with HDL cholesterol (64 loci), were associated with increased CAD risk with $P = 0.02$. When we focused on loci uniquely associated with LDL cholesterol (12 loci where $P > .05$ for other lipids), triglycerides (6 loci), or HDL cholesterol (14 loci), only the LDL association remained significant ($P = 0.03$).

To better explore how associations with individual lipid levels related to CAD risk, we used linear regression to test whether association with lipid levels could predict impact on CAD risk. In this analysis, the effect on CAD of 149 lipid loci (CAD results were not available for 8 SNPs) was correlated with LDL (Pearson r=0.74, $P = 7\times10^{-6}$) and triglyceride (Pearson r=0.46, $P = 0.02$) effect sizes, but not HDL effect sizes (Pearson r=-9×10^{-4}, $P = 0.99$; Supplementary Fig. 6). Since most variants affect multiple lipid fractions (Figure 1), dissecting the relationship between lipid level and CAD effects requires multivariate analysis. In a companion manuscript, we use multivariate analysis and detailed examination

of triglyceride associated loci to show that increased LDL and triglyceride levels, but not HDL, appear causally related to CAD risk.

Evidence for additional loci, not yet reaching genome-wide significance

To evaluate evidence for loci not yet reaching genome-wide significance, we compared direction of effect in GWAS and Metabochip analyses of non-overlapping samples, outside the 157 genome-wide significant loci. Among independent variants ($r^2 < 0.1$) with $P < 0.1$ in the GWAS-only analysis, a significant excess were concordant in direction of effect for HDL (62.9% in 1,847 SNPs, $P < 10^{-16}$), LDL (58.6% of 1,730 SNPs, $P < 10^{-16}$), triglyceride levels (59.1% of 1,783 SNPs, $P < 10^{-16}$), and total cholesterol (61.0% of 1,904 SNPs, $P < 10^{-16}$), suggesting many additional loci to be discovered in future studies.

Discussion

Molecular understanding of the genes and pathways that modify blood lipid levels in humans will facilitate the design of new therapies for cardiovascular and metabolic disease. This understanding can be gained from studies of model organisms, *in vitro* experiments, bioinformatic analyses, and human genetic studies. Here, we demonstrate association between blood lipid levels and 62 new loci, bringing the total number of lipid-associated loci to 157 (See Tables 1A-D and Figure 1). All but one of the loci identified here include protein-coding genes within 100 kb of the SNP showing strongest association. While 38 of the 62 new loci include genes whose role in blood lipid levels is supported by literature review or analysis of curated pathway databases, the remainder includes only genes whose role on blood lipid levels has not been documented.

In total, there are 240 genes within 100 kb of one of our 62 new lipid-associated loci – providing a daunting challenge for future functional studies. Prioritizing on the basis of literature review, pathway analysis, regulation of mRNA expression levels, and protein altering variants suggests that 70 genes in 44 of the 62 new loci might be the focus of the first round of functional studies (summarized in Supplementary Table 2). While we found significant overlap, different sources of prioritization sometimes disagreed. This result suggests that truly understanding causality will be very challenging. The Supplementary Note includes an interpreted digest of genes highlighted by our study. Clearly, a range of approaches will be needed to follow-up these findings. To illustrate possibilities, consider U. S. Patent Application #20,090,036,394 disclosing that, in the mouse, knockout of *Gpr146* modifies blood lipid levels. Here, we show that variants near the human homologue of this gene, *GPR146*, are associated with levels of total cholesterol – providing an added incentive for studies of GPR146 inhibitors in humans. *GPR146* encodes a G-protein coupled receptor – an attractive pharmaceutical target – so it is tempting to speculate that, one day, pharmaceutical inhibition of GPR146 may modify cholesterol levels and reduce risk of heart disease.

Each locus typically includes many strongly associated (and potentially causal) variants. Our fine-mapping results illustrate how genetic analysis of large samples and individuals of diverse ancestry can help focus the search for causal variants. In our fine-mapping analysis of 65 lipid-associated loci, we were able to separate the strongest signal in a region from the

prior GWAS signal in 12 instances. In three of these 12 instances, fine-mapping was enabled by analysis of a few thousand African or East Asian ancestry individuals, whereas in the remaining instances, fine-mapping was possible through examination of nearly 100,000 European ancestry samples. A more detailed fine-mapping exercise, including imputation of variants from emerging very large reference panels, may help refine the location of additional signals.

Lipid-associated loci were strongly associated with CAD, T2D, BMI, SBP, and DBP. In univariate analyses, we found that impact on LDL and triglycerides all predicted association with CAD, but HDL did not. In a more detailed multivariate investigation, a companion manuscript shows that our data is consistent with the hypothesis that both LDL and triglycerides, but not HDL, are causally related to CAD risk. HDL, LDL, and triglycerides levels summarize aggregate levels of different lipid particles, each with potentially distinct consequences for CAD risk. We evaluated association of our loci with lipid subfractions in 2,900 individuals from the Framingham Heart Study (Supplementary Table 13, Supplementary Fig. 7) and with sphingolipids, which are components of lipid membranes in cells, in 4,034 individuals from five samples of European ancestry[39] (Supplementary Table 14). The results suggest HDL-associated variants can have a markedly different impact on these sub-phenotypes. For example, among HDL loci, variants near *LIPC* were strongly associated with plasmalogen levels ($P < 10^{-40}$), variants near *ABCA1* were associated with sphingomyelin levels ($P < 10^{-5}$), and variants near *CETP* – which show the strongest association with HDL cholesterol overall – were associated with neither of these. Detailed genetic dissection of these sub-phenotypes in larger samples, could lead to functional groupings of HDL-associated variants that reconcile the results of genetic studies (which show no clear connection between HDL cholesterol-associated variants and CAD risk) and epidemiologic studies (which show clear association between plasma HDL levels and CAD risk).

In summary, we report the largest genetic association study of blood lipid levels yet conducted. The large number of loci identified, the many candidate genes they contain, and the diverse proteins they encode generate new leads and insights into lipid biology. It is our hope that the next round of genetic studies will build on these results, using new sequencing, genotyping, and imputation technologies to examine rare loss-of-function alleles and other variants of clear functional impact to accelerate the translation of these leads into mechanistic insights and improved treatments for CAD.

Online Methods

Samples studied

We collected summary statistics for Metabochip SNPs from 45 studies. Among these, 37 studies consisted primarily of individuals of European ancestry (see Supplementary Table 1 and Supplementary Note for details), including both population-based studies and case-control studies of CAD and T2D. Another 8 studies consisted primarily of individuals with non-European ancestry: two studies of South Asian descent, AIDHS/SDS (N=1,516) and PROMIS (N=3,385); two studies of East Asian descent, CLHNS (N=1,771) and TAI-CHI (N=7044); and five studies of recent African ancestry, MRC/UVRI GPC (N=1,687) from

Uganda, SEY (N=426) from the Caribbean, and FBPP (N=1,614, TG results unavailable), GXE (N=397), and SPT (N=838) from the United States (more details in Supplementary Table 1 and Supplementary Note).

Genotyping

We genotyped 196,710 genetic variants prioritized on the basis of prior GWAS for cardiovascular and metabolic phenotypes using the Illumina iSelect Metabochip[8] genotyping array. To design the Metabochip, we used our previous GWAS of ~100,000 individuals[4] to prioritize 5,023 SNPs for HDL cholesterol, 5,055 for LDL cholesterol, 5,056 for triglycerides, and 938 for total cholesterol. These independent SNPs represent most loci with $P < .005$ in our original GWAS for HDL cholesterol, LDL cholesterol and triglycerides and with $P < .0005$ for total cholesterol. An additional 28,923 SNPs were selected for fine-mapping of 65 previously identified lipid loci. The Metabochip also included 50,459 SNPs prioritized based on GWAS of non-lipid traits and 93,308 SNPs selected for fine-mapping of loci associated with non-lipid traits (5 of these loci were associated with blood lipids by the analyses described here).

Phenotypes

Blood lipid levels were typically measured after > 8 hours of fasting. Individuals known to be on lipid-lowering medication were excluded when possible. LDL cholesterol levels were directly measured in 10 studies (24% of total study individuals) and estimated using the Friedewald formula[40] in the remaining studies. Trait residuals within each study cohort were adjusted for age, age^2, and sex, and then quantile normalized. Explicit adjustments for population structure using principal component[41] or mixed model approaches[42] were carried out in 24 studies (35% of individuals); all studies were adjusted using genomic control prior to meta-analysis[11]. In studies ascertained on diabetes or CVD status, cases and controls were analyzed separately (Supplementary Table 1). All meta-analyses were limited to a single ancestral group (e.g. European only).

Primary statistical analysis

Individual SNP association tests were performed using linear regression with the inverse normal transformed trait values as the dependent variable and the expected allele count for each individual as the independent variable. These analyses were performed using PLINK (26 samples, 53% of the total number of individuals), SNPTEST (4 samples, 20% of individuals), EMMAX (9 samples, 14% of individuals), Merlin (4 samples, 9% of individuals), GENABEL (1 sample, 3% of individuals), and MMAP (1 sample, 1% of individuals) (Supplementary Table 1).

Meta-analysis

Meta-analysis was performed using the Stouffer method[43,44], with weights proportional to the square root of the sample size for each sample. To correct for inflated test statistics due to potential population stratification, we first applied genomic control to each sample and then repeated the procedure with initial meta-analysis results. For GWAS samples, we used all available SNPs when estimating the median test statistic and inflation factor λ. For

Metabochip samples, we used a subset of SNPs (N = 7,168) that had *P*-values > 0.50 for all lipid traits in the original GWAS, expecting that the majority of these would not be associated with lipids and would behave as null variants in the Metabochip samples. Signals were considered to be novel if they reached a *P*-value $< 5 \times 10^{-8}$ in the combined GWAS and Metabochip meta-analysis and were >1 Mb away from the nearest previously described lipid locus and other novel loci. We used only European samples for the discovery of novel genome-wide significant loci. The non-European samples were meta-analyzed and examined only for fine-mapping analyses.

Quality control

To flag potentially erroneous analyses, we carried out a series of quality control steps. Average standard errors for association statistics from each study were plotted against study sample size to identify outlier studies. We inspected allele frequencies to ensure all analyses used the same strand assignment of alleles. We evaluated whether reported statistics and allelic effects were consistent with published findings for known loci. Genomic control values for study specific analyses were inspected, and all were <1.20. Finally, within each study, we excluded variants for which the minor allele was observed <7 times.

Proportion of trait variance explained

We estimated the increase in trait variance explained by novel loci in the Framingham cohort (N=7,132) using three models for each trait-residual: 1) lead and secondary SNPs from the previously published loci[4] and 2) previously published lipid loci plus newly reported loci; and 3) newly reported loci. We regressed lipid residuals on these sets of SNPs using the lme kinship package in R.

Initial automated review of the published literature

An initial list of candidates within each locus was generated with Snipper (http://csg.sph.umich.edu/boehnke/snipper/) and then subjected to manual review. For each locus, Snipper first generates a list of nearby genes and then checks for the co-occurrence of the corresponding gene names and selected search terms ("cholesterol", "lipids", "HDL", "LDL", or "triglycerides") in published literature and OMIM. We supplemented this approach with traditional literature searches using PubMed and Google.

Generating permuted sets of non-associated SNPs

To estimate the expected chance overlap between literature searches and our loci, we generated lists of permuted SNPs. To generate these lists, we first identified all non-associated lipid SNPs (P > 0.10 for any of the 4 lipid traits) and created bins based on 3 statistics: minor allele frequency, distance to the nearest gene, and number of SNPs with $r^2 >$ 0.8. For each index SNP, we identified 500 non lipid-associated SNPs that fell within the same 3 bins and randomly selected one SNP for each permuted list.

Pathway analyses

To investigate if lipid-associated variants overlapped previously annotated pathways, we used gene set enrichment analysis (GSEA), as implemented in MAGENTA[17] using the

meta-analysis of all studies, including GWAS and Metabochip SNPs. Briefly, MAGENTA first assigns SNPs to a given gene when within 110 kb upstream or 40 kb downstream of transcript boundaries. The most significant SNP *P*-value within this interval is then adjusted for confounders (gene size, marker density, LD) to create a gene association score. When the same SNP is assigned to multiple genes, only the gene with the lowest score is kept for downstream analyses. Subsequently, MAGENTA attaches pathway terms to each gene using several annotation resources, including GO, PANTHER, Ingenuity, and KEGG. Finally, the genes are ranked on their gene association score, and a modified GSEA test is used to test the null hypothesis that all gene score ranks above a given rank cutoff are randomly distributed with regard to a given pathway term (and compared to multiple randomly sampled gene sets of identical size). We evaluated enrichment by using a rank cutoff of 5% of the total number of genes. A minimum of 10,000 gene set permutations were performed, and up to 1,000,000 permutations for GSEA *P*-values below 1×10^{-4}.

We used the Disease Association Protein–Protein Link Evaluator package (DAPPLE; http://www.broadinstitute.org/mpg/dapple/dapple.php) to examine evidence for protein-protein interaction networks connecting genes across different lipid loci. This analysis included the 62 novel loci as well as the 95 previously known loci; we focus our discussion on pathways that included one or more genes from novel loci.

Cis-expression quantitative trait locus analysis

To determine whether lipid-associated SNPs might act as *cis*-regulators of nearby genes, we examined association with expression levels of 39,280 transcripts in 960 human liver samples, 741 human omental fat samples, and 609 human subcutaneous fat samples. Tissue samples were collected postmortem or during surgical resection from donors; tissue collection, DNA and RNA isolation, expression profiling, and genotyping were performed as described[45]. MACH was used to obtain imputed genotypes for ~2.6 million SNPs in the HapMap release 22 for each of the samples. We examined the correlation between each of the 62 new index SNPs and all transcripts within 500 kb of the SNP position, performing association analyses as previously described[46].

Functional annotation of associated variants

We attempted to identify lipid-associated SNPs that fall in important regulatory domains. We initially created a list of all potentially causal variants by selecting index SNPs at loci identified in this study or in Teslovich *et al*[4]. We then selected any variant in strong linkage disequilibrium ($r^2 > 0.8$ from 1000 Genomes or HapMap) with each index SNP. We compared the position of the index SNPs and their proxies to previously described functional marks[23,24]. To assess the expected overlap with functional marks, we created 100,000 permuted sets of non-associated SNPs (see above) and evaluated permuted SNP lists for overlap with functional domains. We estimated a *P*-value for each functional domain as the proportion of permuted sets with an equal or greater number of loci overlapping functional domains (for large *P*-values). For small *P*-values we used a normal approximation to the empirical overlap distribution to estimate *P*-values.

Association with lipid subfractions

Lipoprotein fractions for Women's Genome Health Study (WGHS) samples (N = 23170) were measured using the LipoProtein-II assay (Liposcience Inc. Raleigh, NC) and Framingham Heart Study Offspring samples (N = 2900) were measured with the LipoProtein-I assay (Liposcience Inc. Raleigh, NC)[47]. Additional information on sub-fraction measurements can be found in Supplementary Fig. 7. Log transformations were used for non-normalized traits. All models were adjusted for age, sex, and PCs. The genetic association analysis of WGHS used SNP genotypes imputed from the HapMap r22 CEU reference panel using MACH. 16,730 out of 23,170 WGHS participants were fasting for 8 hours prior to blood draw (72.2%).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Cristen J. Willer[#1,2,3,4,†], Ellen M. Schmidt[#2], Sebanti Sengupta[#4], Gina M. Peloso[5,6,7], Stefan Gustafsson[8,9], Stavroula Kanoni[10], Andrea Ganna[8,9,11], Jin Chen[4], Martin L. Buchkovich[12], Samia Mora[13,14], Jacques S. Beckmann[15,16], Jennifer L. Bragg-Gresham[4], Hsing-Yi Chang[17], Ayşe Demirkan[18], Heleen M. Den Hertog[19], Ron Do[6], Louise A. Donnelly[20], Georg B. Ehret[21,22], Tõnu Esko[7,23,24], Mary F. Feitosa[25], Teresa Ferreira[26], Krista Fischer[23], Pierre Fontanillas[7], Ross M. Fraser[27], Daniel F. Freitag[28], Deepti Gurdasani[10,28], Kauko Heikkilä[29], Elina Hyppönen[30], Aaron Isaacs[18,31], Anne U. Jackson[4], Åsa Johansson[32,33], Toby Johnson[34,35], Marika Kaakinen[36,37], Johannes Kettunen[38,39], Marcus E. Kleber[40,41], Xiaohui Li[42], Jian'an Luan[43], Leo-Pekka Lyytikäinen[44,45], Patrik K.E. Magnusson[11], Massimo Mangino[46], Evelin Mihailov[23,24], May E. Montasser[47], Martina Müller-Nurasyid[48,49,50], Ilja M. Nolte[51], Jeffrey R. O'Connell[47], Cameron D. Palmer[7,52,53], Markus Perola[23,38,39], Ann-Kristin Petersen[48], Serena Sanna[54], Richa Saxena[55], Susan K. Service[56], Sonia Shah[57], Dmitry Shungin[58,59,60], Carlo Sidore[4,54,61], Ci Song[8,9,11], Rona J. Strawbridge[62,63], Ida Surakka[38,39], Toshiko Tanaka[64], Tanya M. Teslovich[4], Gudmar Thorleifsson[65], Evita G. Van den Herik[19], Benjamin F. Voight[66,67], Kelly A. Volcik[68], Lindsay L. Waite[69], Andrew Wong[70], Ying Wu[12], Weihua Zhang[71,72], Devin Absher[69], Gershim Asiki[73], Inês Barroso[10,74], Latonya F. Been[75], Jennifer L. Bolton[27], Lori L Bonnycastle[76], Paolo Brambilla[77], Mary S. Burnett[78], Giancarlo Cesana[79], Maria Dimitriou[80], Alex S.F. Doney[20], Angela Döring[81,82], Paul Elliott[37,83], Stephen E. Epstein[78], Gudmundur Ingi Eyjolfsson[84], Bruna Gigante[85], Mark O. Goodarzi[86], Harald Grallert[87], Martha L. Gravito[75], Christopher J. Groves[88], Göran Hallmans[89], Anna-Liisa Hartikainen[90], Caroline Hayward[91], Dena Hernandez[92], Andrew A. Hicks[93], Hilma Holm[65], Yi-Jen Hung[94], Thomas Illig[87,95], Michelle R. Jones[86], Pontiano Kaleebu[73], John J.P. Kastelein[96], Kay-Tee Khaw[97], Eric Kim[42], Norman Klopp[87,95], Pirjo Komulainen[98], Meena Kumari[57], Claudia Langenberg[43], Terho Lehtimäki[44,45], Shih-Yi Lin[99], Jaana Lindström[100], Ruth J.F. Loos[43,101,102,103], François Mach[21], Wendy L McArdle[104],

Christa Meisinger[81], Braxton D. Mitchell[47], Gabrielle Müller[105], Ramaiah Nagaraja[106], Narisu Narisu[76], Tuomo V.M. Nieminen[107,108,109], Rebecca N. Nsubuga[73], Isleifur Olafsson[110], Ken K. Ong[43,70], Aarno Palotie[38,111,112], Theodore Papamarkou[10,28,113], Cristina Pomilla[10,28], Anneli Pouta[90,114], Daniel J. Rader[115,116], Muredach P. Reilly[115,116], Paul M. Ridker[13,14], Fernando Rivadeneira[117,118,119], Igor Rudan[27], Aimo Ruokonen[120], Nilesh Samani[121,122], Hubert Scharnagl[123], Janet Seeley[73,124], Kaisa Silander[38,39], Alena Stančáková[125], Kathleen Stirrups[10], Amy J. Swift[76], Laurence Tiret[126], Andre G. Uitterlinden[117,118,119], L. Joost van Pelt[127,128], Sailaja Vedantam[7,52,53], Nicholas Wainwright[10,28], Cisca Wijmenga[128,129], Sarah H. Wild[27], Gonneke Willemsen[130], Tom Wilsgaard[131], James F. Wilson[27], Elizabeth H. Young[10,28], Jing Hua Zhao[43], Linda S. Adair[132], Dominique Arveiler[133], Themistocles L. Assimes[134], Stefania Bandinelli[135], Franklyn Bennett[136], Murielle Bochud[137], Bernhard O. Boehm[138,139], Dorret I. Boomsma[130], Ingrid B. Borecki[25], Stefan R. Bornstein[140], Pascal Bovet[137,141], Michel Burnier[142], Harry Campbell[27], Aravinda Chakravarti[22], John C. Chambers[71,72,143], Yii-Der Ida Chen[144,145], Francis S. Collins[76], Richard S. Cooper[146], John Danesh[28], George Dedoussis[80], Ulf de Faire[85], Alan B. Feranil[147], Jean Ferrières[148], Luigi Ferrucci[64], Nelson B. Freimer[56,149], Christian Gieger[48], Leif C. Groop[150,151], Vilmundur Gudnason[152], Ulf Gyllensten[32], Anders Hamsten[62,63,153], Tamara B. Harris[154], Aroon Hingorani[57], Joel N. Hirschhorn[7,52,53], Albert Hofman[117,119], G. Kees Hovingh[96], Chao Agnes Hsiung[155], Steve E. Humphries[156], Steven C. Hunt[157], Kristian Hveem[158], Carlos Iribarren[159], Marjo-Riitta Järvelin[36,37,83,114,160], Antti Jula[161], Mika Kähönen[162], Jaakko Kaprio[29,38,163], Antero Kesäniemi[164], Mika Kivimaki[57], Jaspal S. Kooner[72,143,165], Peter J. Koudstaal[19], Ronald M. Krauss[166], Diana Kuh[70], Johanna Kuusisto[167], Kirsten O. Kyvik[168,169], Markku Laakso[167], Timo A. Lakka[98,170], Lars Lind[171], Cecilia M. Lindgren[26], Nicholas G. Martin[172], Winfried März[41,123,173], Mark I. McCarthy[26,88], Colin A. McKenzie[174], Pierre Meneton[175], Andres Metspalu[23,24], Leena Moilanen[176], Andrew D. Morris[20], Patricia B. Munroe[34,35], Inger Njølstad[131], Nancy L. Pedersen[11], Chris Power[30], Peter P. Pramstaller[93,177,178], Jackie F. Price[27], Bruce M. Psaty[179,180], Thomas Quertermous[134], Rainer Rauramaa[98,181], Danish Saleheen[28,182,183], Veikko Salomaa[184], Dharambir K. Sanghera[75], Jouko Saramies[185], Peter E.H. Schwarz[140,186], Wayne H-H Sheu[187], Alan R. Shuldiner[47,188], Agneta Siegbahn[8,33,171], Tim D. Spector[46], Kari Stefansson[65,189], David P. Strachan[190], Bamidele O. Tayo[146], Elena Tremoli[191], Jaakko Tuomilehto[100,192,193,194], Matti Uusitupa[195,196], Cornelia M. van Duijn[18,31], Peter Vollenweider[197], Lars Wallentin[33,171], Nicholas J. Wareham[43], John B. Whitfield[172], Bruce H.R. Wolffenbuttel[128,198], Jose M. Ordovas[199,200,201], Eric Boerwinkle[68], Colin N.A. Palmer[20], Unnur Thorsteinsdottir[65,189], Daniel I. Chasman[13,14], Jerome I. Rotter[42], Paul W. Franks[58,60,202], Samuli Ripatti[10,38,39], L. Adrienne Cupples[5,203], Manjinder S. Sandhu[10,28], Stephen S. Rich[204], Michael Boehnke#[4], Panos Deloukas#[10], Sekar Kathiresan#[6,7,205,206], Karen L. Mohlke#[12,†], Erik Ingelsson#[8,9,26,†], Gonçalo R. Abecasis#[4,†], and The Global Lipids Genetics Consortium

Affiliations

[1]Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, Michigan 48109, USA [2]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan 48109, USA [3]Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA [4]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, USA [5]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts 02118, USA [6]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA [7]Broad Institute, Program in Medical and Population Genetics, Cambridge, Massachusetts 02142, USA [8]Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala, Sweden [9]Science for Life Laboratory, Uppsala University, Uppsala, Sweden [10]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, CB10 1SA, Hinxton, United Kingdom [11]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden [12]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 USA [13]Division of Preventive Medicine, Brigham and Women's Hospital, 900 Commonwealth Ave., Boston MA 02215, USA [14]Harvard Medical School, Boston MA 02115, USA [15]Service of Medical Genetics, Lausanne University Hospital, Lausanne, Switzerland [16]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland [17]Division of Preventive Medicine and Health Services Research, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan [18]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands [19]Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands [20]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School. Dundee, DD1 9SY, United Kingdom [21]Cardiology, Department of Specialities of Medicine, Geneva University Hospital, Rue Gabrielle-Perret-Gentil 4, 1211 Geneva 14, Switzerland [22]Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA [23]Estonian Genome Center of the University of Tartu, Tartu, Estonia [24]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia [25]Department of Genetics, Washington University School of Medicine, USA [26]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, United Kingdom [27]Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland, United Kingdom [28]Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom [29]Hjelt Institute, Department of Public Health, University of Helsinki, Finland [30]Centre For Paediatric Epidemiology and Biostatistics/MRC Centre of Epidemiology for Child Health, University College of London Institute of Child Health, London, United Kingdom [31]Centre for Medical Systems Biology, Leiden, the Netherlands [32]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden [33]Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden [34]Genome Centre, Barts and The

London School of Medicine and Dentistry, Queen Mary University of London, London, UK [35]Clinical Pharmacology, NIHR Cardiovascular Biomedical Research Unit, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry Queen Mary University of London, London, UK [36]Biocenter Oulu, University of Oulu, Oulu, Finland [37]Institute of Health Sciences, University of Oulu, Finland [38]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Finland [39]Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland [40]Department of Internal Medicine II – Cardiology, University of Ulm Medical Centre, Ulm, Germany [41]Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Ludolf-Krehl-Strasse 7-11, 68167 Mannheim, Germany [42]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA [43]MRC Epidemiology Unit, Institute of Metabolic Science, Box 285, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, United Kingdom [44]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland [45]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere 33014, Finland [46]Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom [47]Division of Endocrinology, Diabetes, and Nutrition, Department of Medicine, University of Maryland, School of Medicine, Baltimore, Maryland [48]Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg 85764, Germany [49]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians University, Munich, Germany [50]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University of Munich, Munich, Germany [51]Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands [52]Division of Endocrinology, Children's Hospital Boston, Massachusetts 02115, USA [53]Division of Genetics, Program in Genomics, Children's Hospital, Boston, Massachusetts 02115, USA [54]Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato, 09042, Italy [55]Massachusetts General Hospital/Broad Institute, Harvard University, Cambridge, MA, USA [56]Center for Neurobehavioral Genetics, The Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, USA [57]Genetic Epidemiology Group, Deparment of Epidemiology and Public Health, UCL, London WC1E 6BT, United Kingdom [58]Department of Clinical Sciences, Genetic & Molecular Epidemiology Unit, Lund University Diabetes Center, Scania University Hosptial, Malmö, Sweden [59]Department of Odontology, Umeå University, Umeå, Sweden [60]Department of Public Health and Primary Care, Unit of Medicine, Umeå University, Umeå, Sweden [61]Dipartimento di Scienze Biomediche, Universita di Sassari, 07100 SS, Italy [62]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden [63]Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden [64]Clinical Research Branch, National Institute Health, Baltimore, MD, USA [65]deCODE Genetics/Amgen, 101 Reykjavik, Iceland [66]Department of Genetics, University of Pennsylvania - School of Medicine, Philadelphia PA, 19104, USA [67]Department of Systems Pharmacology and Translational Therapeutics, University

of Pennsylvania - School of Medicine, Philadelphia PA, 19104, USA [68]Human Genetics Center, University of Texas Health Science Center - School of Public Health, Houston, TX 77030, USA [69]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA [70]MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London, WC1B 5JU, United Kingdom [71]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom [72]Ealing Hospital, Southall, Middlesex UB1 3HW, United Kingdom [73]MRC/UVRI Uganda Research Unit on AIDS, Entebbe, Uganda [74]University of Cambridge Metabolic Research Laboratories and NIHR Cambridge Biomedical Research Centre, Level 4, Institute of Metabolic Science Box 289 Addenbrooke's Hospital Cambridge CB2 OQQ, UK [75]Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA [76]Genome Technology Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA [77]Department of Experimental Medicine, University of Milano Bicocca, Italy [78]MedStar Health Research Institute, 6525 Belcrest Road, Suite 700, Hyattsville, MD 20782, USA [79]Research Centre on Public Health, University of Milano Bicocca, Italy [80]Department of Dietetics-Nutrition, Harokopio University, 70 El. Venizelou Str, Athens, Greece [81]Institute of Epidemiology I, Helmholtz Zentrum München, Neuherberg 85764, Germany [82]Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg 85764, Germany [83]Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment and Health, School of Public Health, Imperial College London, UK [84]The Laboratory in Mjodd, 108 Reykjavik, Iceland [85]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden [86]Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA [87]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg 85764, Germany [88]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, OX3 7LJ, United Kingdom [89]Department of Public Health and Clinical Medicine, Nutritional research, Umeå University, Umeå, Sweden [90]Department of Clinical Sciences/Obstetrics and Gynecology, Oulu University Hospital, Oulu, Finland [91]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland, United Kingdom [92]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA [93]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy - Affiliated Institute of the University of Lübeck, Lübeck, Germany [94]Division of Endocrinology & Metabolism, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan [95]Hannover Unified Biobank, Hannover Medical School, Hannover 30625, Germany [96]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands [97]Clinical Gerontology Unit, University of Cambridge, Cambridge, United Kingdom [98]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland [99]Division of Endocrine and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, School of Medicine, National Yang-Ming University, Taipei, Taiwan [100]Diabetes Prevention Unit, National Institute for

Health and Welfare, 00271 Helsinki, Finland [101]The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, USA [102]The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at ount Sinai, New York, USA [103]The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York [104]School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, United Kingdom [105]Institute for Medical Informatics and Biometrics, University of Dresden, Medical Faculty Carl Gustav Carus, Fetscherstrasse 74, 01307 Dresden, Germany [106]Laboratory of Genetics, National Institute on Aging, Baltimore, MD21224, USA [107]Department of Clinical Pharmacology, University of Tampere School of Medicine, Tamperew 33014, Finland [108]Department of Internal Medicine, Päijät-Häme Central Hospital, Lahti, Finland [109]Division of Cardiology, Helsinki University Central Hospital, Helsinki, Finland [110]Department of Clinical Biochemistry, Landspitali University Hospital, 101 Reykjavik, Iceland [111]Department of Medical Genetics, Haartman Institute, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland [112]Genetic Epidemiology Group, Wellcome Trust Sanger Institute, Hinxton, Cambridge, United ingdom [113]Department of Statistical Sciences, University College of London, London, United Kingdom [114]National Institute for Health and Welfare, Oulu, Finland [115]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, 3400 Civic Center Blvd, Building 421, Translational Research Center, Philadelphia, PA 19104-5158, USA [116]Division of Translational Medicine and Human Genetics, Perelman School of Medicine at the University of Pennsylvania, 3400 Civic Center Blvd, Building 421, Translational Research Center, Philadelphia, PA 19104-5158, USA [117]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands [118]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands [119]Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging NCHA), Leiden, The Netherlands [120]Department of Clinical Sciences/ Clinical Chemistry, University of Oulu, Oulu, Finland [121]National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester LE3 9QP, UK [122]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, LE3 9QP, UK [123]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Austria [124]School of International Development, University of East Anglia, Norwich NR4 7TJ, United Kingdom [125]University of Eastern Finland and Kuopio University Hospital, 70210 Kuopio, Finland [126]INSERM UMRS 937, Pierre and Marie Curie University, Paris, France [127]Department of Laboratory Medicine, University of Groningen, University Medical Center Groningen, The Netherlands [128]LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, The Netherlands [129]Department of Genetics, University of Groningen, University Medical Center Groningen, The Netherlands [130]Department of Biological Psychology, VU Univ, Amsterdam, The Netherlands [131]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway [132]Department of Nutrition,

University of North Carolina, Chapel Hill, NC, USA [133]Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Faculty of Medicine, Strasbourg, France [134]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA [135]Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy [136]Chemical Pathology, Department of Pathology, University of the West Indies, Mona, Kingston 7, Jamaica [137]Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Route de la Corniche 10, 1010 Lausanne, Switzerland [138]Division of Endocrinology and Diabetes, Department of Internal Medicine, Ulm University Medical Centre, Ulm, Germany [139]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore [140]Department of Medicine III, University of Dresden, Medical Faculty Carl Gustav Carus, Fetscherstrasse 74, 01307 Dresden, Germany [141]Ministry of Health, Victoria, Republic of Seychelles [142]Service of Nephrology, Lausanne University Hospital, Lausanne, Switzerland [143]Imperial College Healthcare NHS Trust, London, United Kingdom [144]Division of Reproductive Endocrinology, Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Los Angeles, California, USA [145]Department of Medicine, University of California Los Angeles, Los Angeles, California, USA [146]Department of Preventive Medicine and Epidemiology, Loyola University Medical School, Maywood, Illinois 60153, USA [147]Office of Population Studies Foundation, University of San Carlos, Talamban, Cebu City, Philippines [148]Department of Cardiology, Toulouse University School of Medicine, Rangueil Hospital, Toulouse, France [149]Department of Psychiatry, University of California, Los Angeles, USA [150]Department of Clinical Sciences, Lund University, SE-20502, Malmö, Sweden [151]Department of Medicine, Helsinki University Hospital, FI-00029 Helsinki, Finland [152]Icelandic Heart Association, Kopavogur, Iceland [153]Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden [154]Laboratory of Epidemiology, Demography, and Biometry, National Institute on Ageing, Bethesda, MD, USA [155]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan [156]Cardiovascular Genetics, BHF Laboratories, Institute Cardiovascular Science, University College London, London, United Kingdom [157]Cardiovascular Genetics, University of Utah School of Medicine, Salt Lake City, UT, USA [158]HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway [159]Kaiser Permanente, Division of Research, Oakland, CA, USA [160]Unit of Primary Care, Oulu University Hospital, Oulu, Finland [161]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Turku, Finland [162]Department of Clinical Physiology, University of Tampere School of Medicine, Tampere 33014, Finland [163]Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland [164]Institute of Clinical Medicine, Department of Medicine, University of Oulu and Clinical Research Center, Oulu University Hospital, Oulu, Finland [165]National Heart & Lung Institute, Imperial College London, Hammersmith Hospital, London, United Kingdom [166]Children's Hospital Oakland Research Institute, 5700 Martin Luther King Junior Way, Oakland, CA 94609, USA [167]Department of Medicine, University of Eastern Finland and

Kuopio University Hospital, 70210 Kuopio, Finland [168]Institute of Regional Health Services Research, University of Southern Denmark, Odense, Denmark [169]Odense Patient data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark [170]Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Finland [171]Department of Medical Sciences, Uppsala University, Uppsala, Sweden [172]Queensland Institute of Medical Research, Locked Bag 2000, Royal Brisbane Hospital, Queensland 4029, Australia [173]Synlab Academy, Synlab Services GmbH,Gottlieb-Daimler-Straße 25, 68165 Mannheim, Germany [174]Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, Kingston 7, Jamaica [175]U872 Institut National de la Santé et de la Recherche Médicale, Centre de Recherche des Cordeliers, 75006 Paris, France [176]Department of Medicine, Kuopio University Hospital, Kuopio, Finland [177]Department of Neurology, General Central Hospital, Bolzano, Italy [178]Department of Neurology, University of Lübeck, Lübeck, Germany [179]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA, USA [180]Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA [181]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland [182]Center for Non-Communicable Diseases, Karachi, Pakistan [183]Department of Medicine, University of Pennsylvania, USA [184]Unit of Chronic Disease Epidemiology and Prevention, National Institute for Health and Welfare, Helsinki, Finland [185]South Karelia Central Hospital, Lappeenranta, Finland [186]Paul Langerhans Institute Dresden, German Center for Diabetes Research (DZD), Dresden, Germany [187]Division of Endocrine and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan [188]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland [189]Faculty of Medicine, University of Iceland, 101 Reykjavík, Iceland [190]Division of Population Health Sciences and Education, St George's, University of London, Cranmer Terrace, London SW17 0RE, United Kingdom [191]Department of Pharmacological Sciences, University of Milan, Monzino Cardiology Center, IRCCS, Milan, Italy [192]Centre for Vascular Prevention, Danube-University Krems, 3500 Krems, Austria [193]King Abdulaziz University, Faculty of Medicine, Jeddah 21589, Saudi Arabia [194]Red RECAVA Grupo RD06/0014/0015, Hospital Universitario La Paz, 28046 [195]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Finland [196]Research Unit, Kuopio University Hospital, Kuopio, Finland [197]Department of Medicine, Lausanne University Hospital, Switzerland [198]Department of Endocrinology, University of Groningen, University Medical Center Groningen, The Netherlands [199]Department of Cardiovascular Epidemiology and Population Genetics, National Center for rdiovascular Investigation, Madrid, Spain [200]IMDEA-Alimentacion, Madrid, Spain [201]Nutrition and Genomics Laboratory, Jean Mayer-USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA [202]Department of Nutrition, Harvard School of Public Health, Boston, MA, USA [203]Framingham Heart Study, Framingham, MA, USA [204]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA

[205]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts 02114, USA [206]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA

ACKNOWLEDGEMENTS

REFERENCES

1. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. Annals of Internal Medicine. 1961; 55:33–50. [PubMed: 13751193]

2. Castelli WP. Cholesterol and lipids in the risk of coronary artery disease--the Framingham Heart Study. Canadian Journal of Cardiology. 1988; 4(Suppl A):5A–10A. [PubMed: 3282627]

3. Lloyd-Jones D, et al. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation. 2010; 121:e46–e215. [PubMed: 20019324]

4. Teslovich TM, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010; 466:707–13. [PubMed: 20686565]

5. Barter PJ, Rye KA. Cholesteryl ester transfer protein (CETP) inhibition as a strategy to reduce cardiovascular isk. Journal of Lipid Research. 2012

6. Rahalkar AR, Hegele RA. Monogenic pediatric dyslipidemias: classification, genetics and clinical spectrum. Molecular Genetics and Metabolism. 2008; 93:282–94. [PubMed: 18023224]

7. Musunuru K, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature. 2010; 466:714–9. [PubMed: 20686566]

8. Voight BF, et al. The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, nd Anthropometric Traits. PLoS Genetics. 2012 (in press).

9. The 1000 Genomes Project. A map of human genome variation from population scale sequencing. Nature. 2010; 467

10. Sanna S, et al. Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. PLoS Genet. 2011; 7:e1002198. [PubMed: 21829380]

11. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]

12. Asselbergs FW, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. Am J Hum Genet. 2012; 91:823–38. [PubMed: 23063622]

13. Welch CL, et al. Genetic regulation of cholesterol homeostasis: chromosomal organization of candidate genes. Journal of Lipid Research. 1996; 37:1406–21. [PubMed: 8827514]

14. Sarria AJ, Panini SR, Evans RM. A functional role for vimentin intermediate filaments in the metabolism of lipoprotein-derived cholesterol in human SW-13 cells. Journal of Biological Chemistry. 1992; 267:19455–63. [PubMed: 1527066]

15. Hagberg CE, et al. Vascular endothelial growth factor B controls endothelial fatty acid uptake. Nature. 2010; 464:917–21. [PubMed: 20228789]

16. Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics. 2000; 25:25–9. [PubMed: 10802651]

17. Segre AV, Groop L, Mootha VK, Daly MJ, Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 2010; 6

18. Fitzgerald ML, Moore KJ, Freeman MW. Nuclear hormone receptors and cholesterol trafficking: the orphans find a new home. J Mol Med (Berl). 2002; 80:271–81. [PubMed: 12021839]

19. Rossin EJ, et al. Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS Genet. 2011; 7:e1001273. [PubMed: 21249183]

20. Plyte SE, Hughes K, Nikolakaki E, Pulverer BJ, Woodgett JR. Glycogen synthase kinase-3: functions in oncogenesis and development. Biochimica et Biophysica Acta. 1992; 1114:147–62. [PubMed: 1333807]

21. Toker A, Cantley LC. Signalling through the lipid products of phosphoinositide-3-OH kinase. Nature. 1997; 387:673–6. [PubMed: 9192891]

22. Kaprio J, Ferrell RE, Kottke BA, Kamboh MI, Sing CF. Effects of polymorphisms in apolipoproteins E, A-IV, and H on quantitative traits related to risk for cardiovascular disease. Arteriosclerosis and Thrombosis. 1991; 11:1330–48. [PubMed: 1911720]

23. Ernst J, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011; 473:43–9. [PubMed: 21441907]

24. The ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 2011; 9:e1001046. [PubMed: 21526222]

25. Buyske S, et al. Evaluation of the metabochip genotyping array in African Americans and implications for fine mapping of GWAS-identified loci: the PAGE study. PLoS ONE. 2012; 7:e35651. [PubMed: 22539988]

26. Palmen J, et al. The functional interaction on in vitro gene expression of APOA5 SNPs, defining haplotype APOA52, and their paradoxical association with plasma triglyceride but not plasma apoAV levels. Biochimica et Biophysica Acta. 2008; 1782:447–52. [PubMed: 18395529]

27. Schunkert H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nature Genetics. 2011; 43:333–8. [PubMed: 21378990]

28. The Coronary Artery Disease (C4D) Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nature Genetics. 2011; 43:339–44. [PubMed: 21378988]

29. Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nature Genetics. 2010; 42:579–89. [PubMed: 20581827]

30. Speliotes EK, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass ndex. Nature Genetics. 2010; 42:937–48. [PubMed: 20935630]

31. Heid IM, et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nature Genetics. 2010; 42:949–60. [PubMed: 20935629]

32. Ehret GB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011; 478:103–9. [PubMed: 21909115]

33. Dupuis J, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nature Genetics. 2010; 42:105–16. [PubMed: 20081858]

34. Freathy RM, et al. Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. Diabetes. 2008; 57:1419–26. [PubMed: 18346983]

35. Clarke R, et al. Cholesterol fractions and apolipoproteins as risk factors for heart disease mortality in older men. Archives of Internal Medicine. 2007; 167:1373–8. [PubMed: 17620530]

36. Willer CJ, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nature Genetics. 2008; 40:161–9. [PubMed: 18193043]

37. Voight BF, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012

38. Frikke-Schmidt R, et al. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. JAMA. 2008; 299:2524–32. [PubMed: 18523221]

39. Demirkan A, et al. Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. PLoS Genet. 2012; 8:e1002490. [PubMed: 22359512]

40. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein holesterol in plasma, without use of the preparative ultracentrifuge. Clinical Chemistry. 1972; 18:499–502. [PubMed: 4337382]

41. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics. 2006; 38:904–9. [PubMed: 16862161]

42. Kang HM, et al. Variance component model to account for sample structure in genome-wide association studies. Nature Genetics. 2010; 42:348–54. [PubMed: 20208533]

43. Stouffer, SA.; Suchman, EA.; DeVinney, LC.; Star, SA.; Williams, RMJ. Adjustment During Army Life. Princeton University Press.; Princeton, NJ.: 1949.

44. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–1. [PubMed: 20616382]

45. Keating BJ, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS ONE. 2008; 3:e3583. [PubMed: 18974833]

46. Schadt EE, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol. 2008; 6:e107. [PubMed: 18462017]

47. Chasman DI, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet. 2009; 5:e1000730. [PubMed: 19936222]

A SNAPSHOT OF LIPID GENETICS

1 locus: *ACAD11*

2 loci: *GCKR, NAT2*

HDL cholesterol (46):

HDGF-PMVK, ANGPTL1, CPS1, ATG7, SETD2, RBM5, STAB1, GSK3B, C4orf52, FAM13A, ADH5, DAGLB, SNX13, IKZF1, TMEM176A, OR4C46, KAT5, MOGAT2-DGAT2, ZBTB42-AKT1, HAS1, PABPC4, ZNF648, COBLL1, SLC39A8, ARL15, CITED2, KLF14, TRPS1, AMPD3, LRP4, PDE3A, MVK, SBNO1, ZNF664, SCARB1, LACTB, LCAT, CMIP, STARD3, ABCA8, PGS1, MC4R, ANGPTL4, ANGPTL8, LILRA3, UBE2L3

Total cholesterol (18):

ASAP3, ABCB11, FAM117B, PXK, KCNK17, HBS1L, GPR146, VIM-CUBN, PHLDB1, PHC1-A2ML1, TOM1, EVI5, RAB3GAP1, RAF1, C6orf106, SPTY2D1, MAMSTR, ERGIC3

6 loci:

MARCH8-ALOX5, TTC39B, ABCA1, LIPG, HNF4A, UBASH3B

2 loci:

PPP1R3B, APOE

36 loci:

INSIG2, LOC84931, CMTM6, CSNK1G3, SOX17, UGT1A1, VLDLR, DLG4, PPARA, PCSK9, SORT1, APOB, ABCG5/8, MYLIP, HFE, LPA, PLEC1, ABO, ST3GAL4, OSBPL7, LDLR, TOP1, LDLRAP1, MOSC1, IRF2BP2, HMGCR, HLA, FRK, DNAH11, NPC1L1, CYP7A1, GPAM, BRAP, HNF1A, HPR, MAFB

1 locus:

LIPC

4 loci:

CETP, TRIB1, FADS1-2-3, APOA1

10 loci:

RSPO3, FTO, VEGFA, PEPD, GALNT2, IRS1, PLTP, MLXIPL, LPL, LRP1

5 loci:

ANGPTL3, MIR148A, LRPAP1, TIMD4, CILP2

1 locus:

PIGV-NR0B2

Triglycerides (16):

MET, AKR1C4, PDXDC1, MPP3, INSR, MSL2L1, KLHL8, MAP3K1, TYW1B, PINX1, JMJD1C, CYP26A1, CAPN3, FRMD5, CTF1, PLA2G6

LDL cholesterol (9):

ANXA9-CERS2, EHBP1, BRCA2, FN1, APOH-PRXCA, SPTLC3, SNX5, MTMR3, NYNRIN

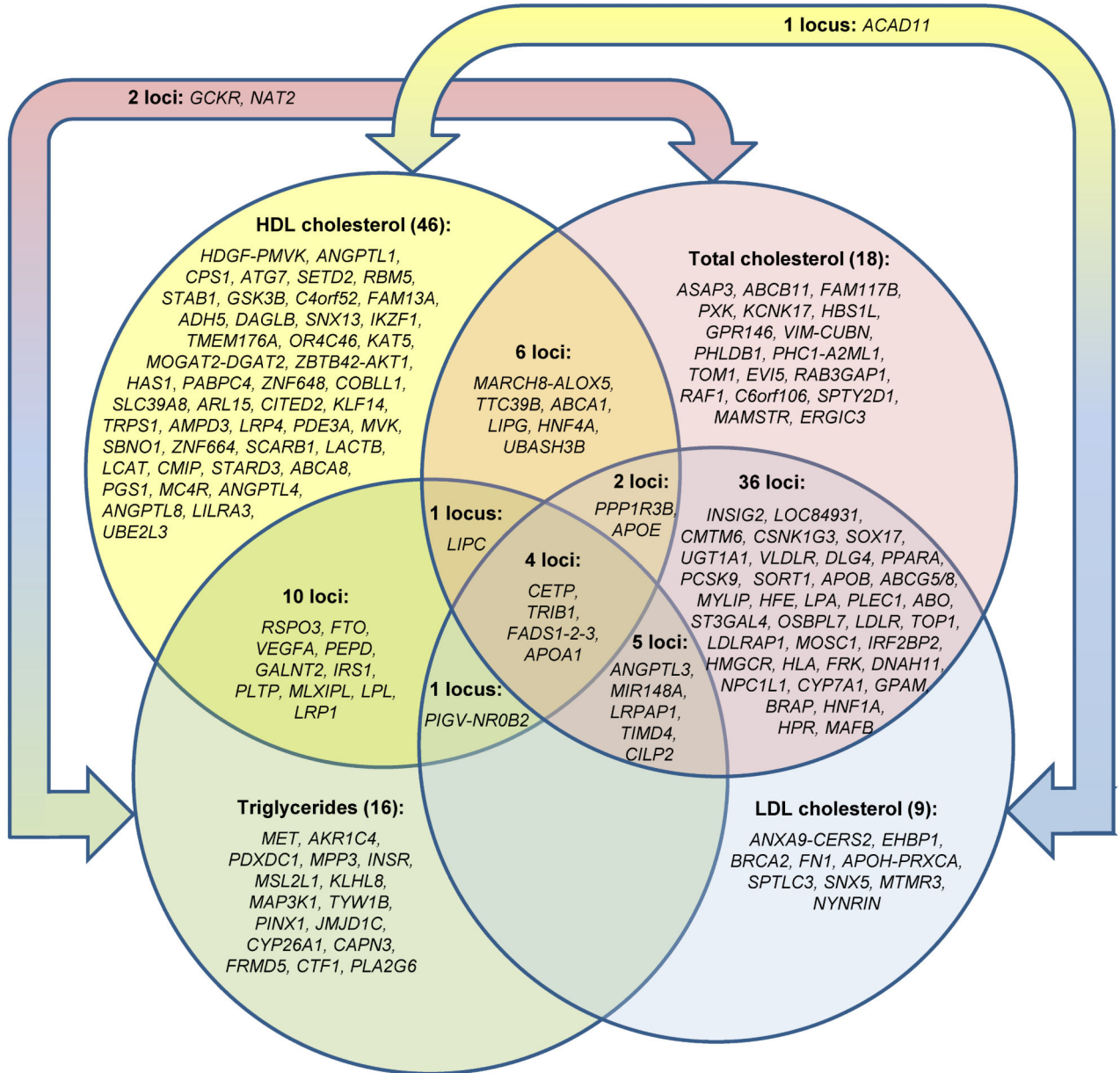**FIGURE 1. Overlap between loci associated with different lipid traits**

This Venn Diagram illustrates the number of loci that show association with multiple lipid traits. The number of loci primarily associated with only one trait is listed in parentheses after the trait name and the locus name is listed below in italics. Loci that show association with two or more traits are shown in the appropriate section.

TABLE 1A

Novel Loci Primarily Associated with HDL Cholesterol Obtained from Joint GWAS and Metabochip Meta-analysis

Locus	MarkerName	Chr	hg19 Position (Mb)	Associated trait(s)	MAF	Minor/major Allele	Effect of A1	Joint N (in 1000s)	Joint *P*-value
PIGV-NR0B2	rs12748152	1	27.14	HDL, LDL, TG	.09	T/C	-.051/.050/.037	187/173/178	$1\times10^{-15}/3\times10^{-12}/1\times10^{-9}$
HDGF-PMVK	rs12145743	1	156.70	HDL	.34	G/T	.020	181	2×10^{-8}
ANGPTL1	rs4650994	1	178.52	HDL	.49	G/A	.021	187	7×10^{-9}
CPS1	rs1047891	2	211.54	HDL	.33	A/C	-.027	182	9×10^{-10}
ATG7	rs2606736	3	11.40	HDL	.39	C/T	.025	129	5×10^{-8}
SETD2	rs2290547	3	47.06	HDL	.20	A/G	-.030	187	4×10^{-9}
RBM5	rs2013208	3	50.13	HDL	.50	T/C	.025	170	9×10^{-12}
STAB1	rs13326165	3	52.53	HDL	.21	A/G	.029	187	9×10^{-11}
GSK3B	rs6805251	3	119.56	HDL	.39	T/C	.020	186	1×10^{-8}
C4orf52	rs10019888	4	26.06	HDL	.18	G/A	-.027	187	5×10^{-8}
FAM13A	rs3822072	4	89.74	HDL	.46	A/G	-.025	187	4×10^{-12}
ADH5	rs2602836	4	100.01	HDL	.44	A/G	.019	187	5×10^{-8}
RSPO3	rs1936800	6	127.44	HDL, TG[a]	.49	C/T	.020/-.020	187/168	$3\times10^{-10}/3\times10^{-8}$
DAGLB	rs702485	7	6.45	HDL	.45	G/A	.024	187	7×10^{-12}
SNX13	rs4142995	7	17.92	HDL	.38	T/G	-.026	165	9×10^{-12}
IKZF1	rs4917014	7	50.31	HDL	.32	G/T	.022	187	1×10^{-8}
TMEM176A	rs17173637	7	150.53	HDL	.12	C/T	-.036	184	2×10^{-8}
MARCH8-ALOX5	rs970548	10	46.01	HDL, TC	.26	C/A	.026/-.026	187/187	$2\times10^{-10}/8\times10^{-9}$
OR4C46	rs11246602	11	51.51	HDL	.15	C/T	.034	176	2×10^{-10}
KAT5	rs12801636	11	65.39	HDL	.23	A/G	.024	187	3×10^{-8}
MOGAT2-DGAT2	rs499974	11	75.46	HDL	.19	A/C	-.026	187	1×10^{-8}
ZBTB42-AKT1	rs4983559	14	105.28	HDL	.40	G/A	.020	184	1×10^{-8}
FTO	rs1121980	16	53.81	HDL, TG	.43	A/G	-.020/-.021	186/155	$7\times10^{-9}/3\times10^{-8}$
HAS1	rs17695224	19	52.32	HDL	.26	A/G	-.029	185	2×10^{-13}

Chr, chromosome;MAF, minor allele frequency; A1, minor allele; A2, major allele.Effect sizes are given with respect to the minor allele (A1) in SD units. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest *P*-value is listed first. At one locus, the secondary trait was most strongly associated with a different SNP:

[a] rs719726 (within 1Mb of rs1936800, $r^2 = 0.74$).

TABLE 1B

Novel Loci Primarily Associated with LDL Cholesterol Obtained from Joint GWAS and Metabochip Meta-analysis

Locus	MarkerName	Chr	hg19 Position (Mb)	Associated trait(s)	MAF	Minor/major Allele	Effect of A1	Joint N (in 1000s)	Joint P-value
ANXA9-CERS2	rs267733	1	150.96	LDL	.16	G/A	-.033	165	5×10^{-9}
EHBP1	rs2710642	2	63.15	LDL	.35	G/A	-.024	173	6×10^{-9}
INSIG2	rs10490626	2	118.84	LDL, TC[b]	.08	A/G	-.051/.042	173/184	$2\times10^{-12}/6\times10^{-9}$
LOC84931	rs2030746	2	121.31	LDL, TC	.40	T/C	.021/.020	173/187	$9\times10^{-9}/4\times10^{-8}$
FN1	rs1250229	2	216.30	LDL	.27	T/C	-.024	173	3×10^{-8}
CMTM6	rs7640978	3	32.53	LDL, TC	.09	T/C	-.039/-.038	172/186	1×10^{-8}
ACAD11	rs17404153	3	132.16	LDL, HDL[c]	.14	T/G	-.034/.028	172/187	$2\times10^{-9}/5\times10^{-9}$
CSNK1G3	rs4530754	5	122.86	LDL, TC	.46	G/A	-.028/-.023	173/187	$4\times10^{-12}/2\times10^{-9}$
MIR148A	rs4722551	7	25.99	LDL, TG[d], TC	.20	C/T	.039/.029/.023	173/187/178	$4\times10^{-14}/9\times10^{-11}/7.0\times10^{-9}$
SOX17	rs10102164	8	55.42	LDL, TC	.21	A/G	.032/.030	173/187	$4\times10^{-11}/5\times10^{-11}$
BRCA2	rs4942486	13	32.95	LDL	.48	T/C	.024	172	2×10^{-11}
APOH-PRXCA	rs1801689	17	64.21	LDL	.04	C/A	.103	111	1×10^{-11}
SPTLC3	rs364585	20	12.96	LDL	.38	A/G	-.025	172	4×10^{-10}
SNX5	rs2328223	20	17.85	LDL	.21	C/A	.03	171	6×10^{-9}
MTMR3	rs5763662	22	30.38	LDL	.04	T/C	.077	163	1×10^{-8}

Chr, chromosome;MAF, minor allele frequency; A1, minor allele; A2, major allele.Effect sizes are given with respect to the minor allele (A1) in SD units. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest P-value is listed first. At three loci, secondary traits were most strongly associated with different SNPs.

[b] rs17526895 (within 1Mb of rs10490626, $r^2 = 0.98$);

[c] rs13076253 (within 1Mb of rs17404153, $r^2 = 0.00$);

[d] rs4719841 (within 1Mb of rs4722551, $r^2 = 0.10$).

TABLE 1C

Novel Loci Primarily Associated with Total Cholesterol Obtained from Joint GWAS and Metabochip Meta-analysis

Locus	MarkerName	Chr	hg19 Position (Mb)	Associated trait(s)	MAF	Minor/major Allele	Effect of A1	Joint N (in 1000s)	Joint P-value
ASAP3	rs1077514	1	23.77	TC	.15	C/T	−0.03	184	6×10^{-9}
ABCB11	rs2287623	2	169.83	TC	.41	G/A	0.027	184	4×10^{-12}
FAM117B	rs11694172	2	203.53	TC	.25	G/A	0.028	187	2×10^{-9}
UGT1A1	rs11563251	2	234.68	TC, LDL	.12	T/C	0.037/0.034	187/173	$1\times10^{-9}/5\times10^{-8}$
PXK	rs13315871	3	58.38	TC	.10	A/G	−0.036	187	4×10^{-8}
KCNK17	rs2758886	6	39.25	TC	.30	A/G	0.023	187	3×10^{-8}
HBS1L	rs9376090	6	135.41	TC	.28	T/C	−0.025	187	3×10^{-9}
GPR146	rs1997243	7	1.08	TC	.16	G/A	0.033	183	3×10^{-10}
VLDLR	rs3780181	9	2.64	TC, LDL	.08	G/A	−0.044/−0.044	186/172	$7\times10^{-10}/2\times10^{-9}$
VIM-CUBN	rs10904908	10	17.26	TC	.43	G/A	0.025	187	3×10^{-11}
PHLDB1	rs11603023	11	118.49	TC	.42	T/C	0.022	187	1×10^{-8}
PHC1-A2ML1	rs4883201	12	9.08	TC	.12	G/A	−0.035	187	2×10^{-9}
DLG4	rs314253	17	7.09	TC, LDL	.37	C/T	−0.023/−0.024	184/170	$3\times10^{-10}/3\times10^{-10}$
TOM1	rs138777	22	35.71	TC	.36	A/G	0.021	185	5×10^{-8}
PPARA	rs4253772	22	46.63	TC, LDL ^{e}	.11	T/C	0.032/−0.031	185/171	$1\times10^{-8}/3\times10^{-8}$

Chr, chromosome;MAF, minor allele frequency; A1, minor allele; A2, major allele.Effect sizes are given with respect to the minor allele (A1) in SD units. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest P-value is listed first. At one locus, the secondary trait was most strongly associated with a different SNP:

^{e} rs4253776 (within 1Mb of rs4253772, r^2 = 0.95).

TABLE 1D

Novel Loci Primarily Associated with Triglycerides Obtained from Joint GWAS and Metabochip Meta-analysis

Locus	MarkerName	Chr	hg19 Position (Mb)	Associated trait(s)	MAF	Minor/major Allele	Effect of A1	Joint N (in 1000s)	Joint P-value
LRPAP1	rs6831256	4	3.47	TG, TC^f,LDL^f	.42	G/A	0.026/−0.022/−0.025	177/173/187	2×10^{-12}/1×10^{-10}/2×10^{-8}
VEGFA	rs998584	6	43.76	TG, HDL	.49	A/C	0.029/−0.026	175/184	3×10^{-15}/2×10^{-11}
MET	rs38855	7	116.36	TG	.47	G/A	−0.019	178	2×10^{-8}
AKR1C4	rs1832007	10	5.25	TG	.18	G/A	−0.033	178	2×10^{-12}
PDXDC1	rs3198697	16	15.13	TG	.43	T/C	−0.020	176	2×10^{-8}
MPP3	rs8077889	17	41.88	TG	.22	C/A	0.025	176	1×10^{-8}
INSR	rs7248104	19	7.22	TG	.42	A/G	−0.022	176	5×10^{-10}
PEPD	rs731839	19	33.90	TG, HDL	.35	G/A	0.022/−0.022	176/185	3×10^{-9}/3×10^{-9}

Chr, chromosome;MAF, minor allele frequency; A1, minor allele; A2, major allele.Effect sizes are given with respect to the minor allele (A1) in SD units. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest *P*-value is listed first. At one locus, secondary traits were most strongly associated with a different SNP:

^{f}rs6818397 (within 1 Mb of rs6831256, r^2 = 0.18).