Genetic Variants Influencing Circulating Lipid Levels and Risk of Coronary Artery Disease

Dawn M. Waterworth, Sally L. Ricketts, Kijoung Song, Li Chen, Jing Hua Zhao, Samuli Ripatti, Yurii S. Aulchenko, Weihua Zhang, Xin Yuan, Noha Lim, Jian'an Luan, Sofie Ashford, Eleanor Wheeler, Elizabeth H. Young, David Hadley, John R. Thompson, Peter S. Braund, Toby Johnson, Maksim Struchalin, Ida Surakka, Robert Luben, Kay-Tee Khaw, Sheila A. Rodwell, Ruth J.F. Loos, S. Matthijs Boekholdt, Michael Inouye, Panagiotis Deloukas, Paul Elliott, David Schlessinger, Serena Sanna, Angelo Scuteri, Anne Jackson, Karen L. Mohlke, Jaako Tuomilehto, Robert Roberts, Alexandre Stewart, Y. Antero Kesäniemi, Robert W. Mahley, Scott M. Grundy, Wellcome Trust Case Control Consortium, Wendy McArdle, Lon Cardon, Gérard Waeber, Peter Vollenweider, John C. Chambers, Michael Boehnke, Gonçalo R. Abecasis, Veikko Salomaa, Marjo-Riitta Järvelin, Aimo Ruokonen, Inês Barroso, Stephen E. Epstein, Hakon H. Hakonarson, Daniel J. Rader, Muredach P. Reilly, Jacqueline C.M. Witteman, Alistair S. Hall, Nilesh J. Samani, David P. Strachan, Philip Barter, Cornelia M. van Duijn, Jaspal S. Kooner, Leena Peltonen, Nicholas J. Wareham, Ruth McPherson, Vincent Mooser, Manjinder S. Sandhu

Objective—Genetic studies might provide new insights into the biological mechanisms underlying lipid metabolism and risk of CAD. We therefore conducted a genome-wide association study to identify novel genetic determinants of low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides.

Methods and Results—We combined genome-wide association data from 8 studies, comprising up to 17 723 participants with information on circulating lipid concentrations. We did independent replication studies in up to 37 774 participants from 8 populations and also in a population of Indian Asian descent. We also assessed the association between single-nucleotide polymorphisms (SNPs) at lipid loci and risk of CAD in up to 9 633 cases and 38 684 controls. We identified 4 novel genetic loci that showed reproducible associations with lipids (probability values, 1.6×10^{-8} to 3.1×10^{-10}). These include a potentially functional SNP in the *SLC39A8* gene for HDL-C, an SNP near the *MYLIP/GMPR* and *PPP1R3B* genes for LDL-C, and at the *AFF1* gene for triglycerides. SNPs showing strong statistical association with 1 or more lipid traits at the *CELSR2*, *APOB*, *APOE-C1-C4-C2* cluster, *LPL*, *ZNF259-APOA5-A4-C3-A1* cluster and *TRIB1* loci were also associated with CAD risk (probability values, 1.1×10^{-3} to 1.2×10^{-9}).

Conclusion—We have identified 4 novel loci associated with circulating lipids. We also show that in addition to those that are largely associated with LDL-C, genetic loci mainly associated with circulating triglycerides and HDL-C are also associated with risk of CAD. These findings potentially provide new insights into the biological mechanisms underlying lipid metabolism and CAD risk. **(*Arterioscler Thromb Vasc Biol*. 2010;30:2264-2276.)**

Key Words: coronary artery disease ■ epidemiology ■ lipids ■ genetics

Circulating levels of blood lipids have been consistently associated with risk of coronary artery disease (CAD).[1] However, whereas low-density lipoprotein cholesterol (LDL-C) is known to cause atherosclerosis and CAD, the role of circulating high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG) in the development of atherosclerosis and CAD remains uncertain.[2,3] In this context, the integration of population genetics and epidemiological approaches could help assess the etiologic role of HDL-C and TG levels in atherosclerosis and CAD.[4] By identifying novel genetic determinants of blood lipids, these integrated approaches can also help provide new insights into the biological mechanisms regulating lipid metabolism and identify potentially novel therapeutic targets for CAD.[5–7]

See accompanying article on page 2084

Recent genome-wide association (GWA) studies have identified several new loci that influence circulating levels of blood

Received on: December 14, 2009; final version accepted on: May 19, 2010.

From the Genetics Division, GlaxoSmithKline R&D, King of Prussia, Pa (D.M.W., K.S., X.Y., N.L., L. Cardon, V.M.); Department of Public Health and Primary Care, Strangeways Research Laboratory (S.L.R., S.A., E.H.Y., R.L., K.-T.K., M.S.S.) and Department of Public Health and Primary Care (S.A.R.), University of Cambridge, Cambridge, United Kingdom; Division of Cardiology, University of Ottawa Heart Institute, Ottawa, Ontario, Canada (L. Chen, R.R., A. Stewart, R.M.); Medical Research Council Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom (J.H.Z., J.L., R.J.F.L., N.J.W., M.S.S.); Institute for Molecular Medicine (FIMM), University of Helsinki, Finland (S.R., I.S., L.P.); Departments of Epidemiology (Y.S.A., A.H., C.M.v.D.) and Forensic Molecular Biology (M.S.), Erasmus University Medical Center, Rotterdam, the

lipids, with around 30 genetic loci showing reproducible statistical associations with circulating HDL-C, LDL-C and TG.[5–12] However, given that a substantial proportion of the genetic variance for these traits remains unexplained, these loci are likely to represent only a small proportion of all genetic determinants involved in lipid metabolism. We therefore conducted an extended GWA study of LDL-C, HDL-C, and TG levels to identify novel genetic determinants of these traits and validated our associations in independent populations, including a population of Indian Asian descent. We also examined the association between genetic variants showing reproducible statistical association with lipid levels with risk of CAD.

Methods

Study Populations

GWA Metaanalysis of Circulating Lipid Traits

We used data from 8 study populations comprising up to 17 723 participants of white European descent. These are the EPIC-Norfolk subcohort (up to 2269 participants), the EPIC-Norfolk obese set (up to 1009 participants), the British 1958 birth cohort (Wellcome Trust Case Control Consortium controls—up to 1458 participants), the CoLaus study (up to 5226 participants), the Genetic Epidemiology of Metabolic Syndrome (GEMS) study (up to 1665 participants), a sample from the London Life Sciences Population (LOLIPOP) study (up to 813 participants), controls from the FUSION type 2 diabetes study (up to 1099 participants), and the SardiNIA Study of Aging (up to 4184 participants). Individual studies have been described in detail in recent reports and are summarized briefly in Supplemental Table I, available online at http://atvb.ahajournals.org.[6,9,12] Selected descriptive characteristics of all study populations are also provided in Supplemental Table I. We used only data from the GEMS study[13,14] for our GWA analysis of LDL-C, as the study comprises cases and controls of dyslipidaemia, defined by high and low percentiles of HDL-C and TG, respectively (Supplemental Table I).

Lipid Replication Analyses

Our replication set encompassed individuals of white European descent from 8 studies comprising up to 37 774 participants. Individual studies are summarized in Supplemental Table II. Briefly, the replication set comprised the EPIC-Norfolk cohort[15] (up to 19 793 individuals who had DNA and lipid measurements available and did not overlap with the EPIC-Norfolk subcohort or obese set), controls from the Ottawa Heart Study[16] (up to 1445 participants), Fenland study[17] (up to 1402 participants), an additional subset of the LOLIPOP study[9] (up to 710 participants), British 1958 Birth cohort T1DGC controls[18] (2527 participants—there is no overlap with the control set from the original (Wellcome Trust Case Control Consortium substudy), Northern Finland Birth Cohort 1966[10] (up to 5138 participants), National FINRISK Study[7] (up to 910 participants), and Rotterdam study[7] (up to 5849 participants).

Case-Control Studies for CAD

For our CAD metaanalysis, we combined data from 9 studies comprising up to 9633 cases and 38 684 controls. These studies included 2 nonoverlapping case-control studies of CAD derived from the EPIC-Norfolk cohort,[15] (Wellcome Trust Case Control Consortium CAD study,[19,20] Ottawa Heart Study,[16] MEDSTAR and PENN CATH studies,[21] and nested CAD case-control studies derived from the CoLaus study,[6] GEMS study[13,14] and Rotterdam study.[7] Details of these studies are provided in Supplemental Table III.

Studies of Indian Asian Ethnicity

To examine the consistency of our novel association signals for lipids in an ethnically distinct population, we used 4 nonoverlapping subsets of Indian Asian participants from the LOLIPOP study,[9] collectively comprising up to 9665 participants (see Supplemental Table II for details).

Local ethics committees approved all studies, and all participants gave written informed consent.

Genotyping

GWA Genotyping

The 8 studies used in the GWA metaanalysis of lipid traits have been genotyped with different genome-wide single-nucleotide polymorphism (SNP) chips (for details see Supplemental Table I). To enable us to combine data from all studies for our GWA metaanalysis, we used information on SNP genotypes in our samples and HapMap II data to statistically predict (impute) all SNP genotypes for all individuals. These genome-wide imputation analyses were con-

(Continued). Netherlands; Department of Epidemiology and Public Health (W.Z., J.C.C.), National Heart and Lung Institute (P.E., J.S.K.) and Department of Biostatistics and Epidemiology, School of Public Health (M.-R.J.), Imperial College London, London, United Kingdom; Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom (E.W., E.H.Y., M.I., P.D., I.B., L.P., M.S.S.); Division of Community Health Sciences, St George's, University of London, London, United Kingdom (D.H.); Departments of Health Sciences and Genetics, University of Leicester, Leicester, United Kingdom (J.R.T., P.S.B.); William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom (T.J.); National Public Health Institute, Biomedicum, Helsinki, Finland (I.S.); Medical Research Council Centre for Nutritional Epidemiology in Cancer Prevention and Survival, Departments of Vascular Medicine (S.M.B.) and Cardiology (S.M.B.), Academic Medical Center, Amsterdam, the Netherlands; Laboratory of Genetics (D.S.) and Gerontology Research Center (A. Scuteri), National Institute on Aging, National Institutes of Health, Baltimore, MD; Istituto di Neurogenetica e Neurofarmacologia, CNR, Monserrato, Cagliari, Italy (S.S., A. Scuteri); Department of Biostatistics (A.J., M.B.) and Center for Statistical Genetics (G.R.A.), Department of Biostatistics, University of Michigan, Ann Arbor, Mich; Department of Genetics, University of North Carolina, Chapel Hill, NC (K.L.M.); National Public Health Institute, Department of Epidemiology and Health Promotion, Helsinki, Finland (J.T.); Department of Internal Medicine (Y.A.K.), Institute of Health Sciences (M.-R.J.), Biocenter Oulu (Y.A.K., M.-R.J.), and Institute of Diagnostics, Clinical Chemistry (A.R.), University of Oulu, Oulu, Finland; Gladstone Institute of Neurological Disease and Gladstone Institute of Cardiovascular Disease, San Francisco, Calif (R.W.M.); Center for Human Nutrition, Department of Clinical Nutrition, University of Texas Southwestern Medical Center, Dallas, Texas (S.M.G.); Avon Longitudinal Study of Parents and Children, University of Bristol, Bristol, United Kingdom (W.M.); Department of Internal Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland (G.W., P.V.); National Institute for Health and Welfare, Helsinki, Finland (V.S., M.-R.J.); Cardiovascular Research Institute, MedStar Research Institute, Washington Hospital Center, Washington, DC (S.E.E.); Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA (H.H.H.); Institute for Translational Medicine and Therapeutics, School of Medicine (D.J.R.) and the Cardiovascular Institute (D.J.R.), University of Pennsylvania, Philadelphia, Pa; LIGHT, University of Leeds, Leeds, United Kingdom (A.S.H.); Heart Research Institute, Camperdown, Sydney, New South Wales, Australia (P.B.); Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Mass (L.P.).

Dr Rodwell and Dr Peltonen are deceased.

Dr Waterworth and Dr Ricketts contributed equally to this work.

Correspondence to Manjinder S. Sandhu, Department of Public Health and Primary Care, Strangeways Research Laboratory, University of Cambridge, Cambridge CB1 8RN, United Kingdom. E-mail manj.sandhu@srl.cam.ac.uk

Arterioscler Thromb Vasc Biol **is available at http://atvb.ahajournals.org** **DOI: 10.1161/ATVBAHA.109.201020**

ducted in each study independently using either IMPUTE[22] or MACH[12] (Supplemental Table I).

Replication Genotyping for Lipid SNPs

SNPs taken forward for replication for the 3 lipid traits were genotyped on the EPIC-Norfolk cohort using either the iPLEX Sequenom MassARRAY platform or allelic discrimination on an ABI 7900 instrument (TaqMan, Applied Biosystems, Warrington, UK). Criteria for genotyping quality are outlined in Supplemental Table II. For the remaining 7 replication studies (Supplemental Table II), genotypes were available *in silico* using data from genome-wide SNP chips or imputation analyses (Supplemental Table II).

Genotyping of Case-Control Studies for CAD

Genotypes were available for *in silico* testing of lipid SNPs for association with CAD risk for the 9 case-control studies described above (see Supplemental Table III for details).

Studies of Indian Asian Ethnicity

Genotypes were available for *in silico* testing of SNPs with circulating lipid levels for the 4 nonoverlapping subsets of the LOLIPOP study (Supplemental Table II).

Statistical Analyses

GWA Metaanalysis of Circulating Lipid Traits

Sample and SNP quality control criteria and statistical analysis for each lipid trait was done within each study independently (Supplemental Table I). For the initial GWA screen, analyses were done within study using a uniform analytic strategy. All lipid traits were natural log transformed before GWA analysis across studies. The choice of natural log transformation was guided by the shape of the phenotype distributions across studies, to minimize skew while also retaining a link to the original data—particularly for studies comprising selected populations. This transformation also provided an interpretable regression coefficient. Analyses were conducted using an additive model adjusted for age, sex, and geographical/population covariables where appropriate. Association analysis for both imputed and genotyped SNPs was done using SNPTEST[22] (with the full posterior probability genotype distribution) or MERLIN.[12] Only SNPs with a minor allele frequency of 1% or more and with a posterior-probability score more than 0.90 were considered for these imputed association analyses. Criteria for imputation quality and genomic control parameters are outlined in Supplemental Table I.

We conducted a GWA metaanalysis by combining summary data from each of the 8 studies using a fixed-effects model and inverse-variance weighted averages of β-coefficients with Stata, version 8.2. This provided us with a combined estimate of the overall β-coefficient and its standard error. Between-study heterogeneity was assessed with the χ^2 test. To optimize data quality, we analyzed only SNPs that passed sample and SNP quality control criteria in each of the 8 studies and that had a measure of association (β-coefficient and standard error) in all 8 studies (see above for details). Data for 2 155 369 autosomal SNPs were available for analysis of circulating HDL-C levels, 2 154 923 for LDL-C, and 2 155 784 SNPs for TG. We also calculated an inflation factor (λ) for each study, which was estimated from the mean of the χ^2 tests generated on all SNPs that were tested (Supplemental Table I). The overall genomic control parameters[23] were 1.08, 1.07, and 1.06 in our metaanalysis for HDL-C, LDL-C and TG, respectively. These results suggest that unmodeled relatedness or population stratification is unlikely to materially influence our results.

For the 3 lipid traits (HDL-C, LDL-C and TG), we examined SNPs only at known, previously reported, and novel loci that had a combined $P<1\times10^{-5}$ (an arbitrary statistical threshold) in the metaanalysis and that did not show any heterogeneity among studies ($P<0.1$).

Replication Analyses for Lipid SNPs

For each novel locus, the SNP showing the strongest statistical association was taken forward for replication in stage 2. These comprised 40 SNPs in total: 11 for HDL-C, 13 for LDL-C, 15 for TG, and 1 for both HDL-C and TG. We conducted replication analyses in the EPIC-Norfolk cohort using linear regression using natural log–transformed lipid levels and an additive model with adjustment for age and sex. We combined these data with *in silico* replication sets from the other 7 studies using metaanalysis, as above, to obtain an overall estimate of association in the combined datasets. These analyses comprised adjustment for age, sex and population variables, as relevant (Supplemental Table II).

Association Analyses for CAD Risk

Association analyses were done using either SNPTEST,[22] PLINK[24] or ProbABEL[25] for 8 studies with genome-wide SNP data available. Analyses for these studies comprised adjustment for at least age, sex and population variables (where relevant—see Supplemental Table III). For the EPIC-Norfolk-2 case-control study, we used logistic regression and a log additive model adjusted for age and sex to test for association of novel lipid SNPs with CAD. We combined summary estimates (log odds ratios and standard errors) for each of the 9 studies using metaanalysis, as above, to obtain a combined estimate of the association between SNPs and risk of CAD for a log additive model.

Studies of Indian Asian Ethnicity

For association analyses of potentially novel lipid loci in Indian Asian individuals, we combined data from the 4 LOLIPOP subsets (Supplemental Table II) using the metaanalytic strategy outlined above. We then conducted a formal assessment of the heterogeneity (Q statistic) between the 2 ethnic groups for our potentially novel lipid loci by comparing overall summary estimates from our white European and Indian Asian studies using metaanalysis, as above.

Results

Known Genes Influencing Lipids

For LDL-C, HDL-C, and TG levels, the strongest statistical association signals in our GWA metaanalysis were at loci previously implicated in lipid metabolism or those recently identified as potential lipid genes.[5] Table 1 lists the 28 SNPs with strongest statistical associations for the 3 lipid traits ($P<1\times10^{-5}$) with no detectable heterogeneity among studies ($P>0.1$) at these genetic loci. Eighteen SNPs at these known loci reached genome-wide statistical association ($P<5\times10^{-8}$) in our data. As expected, genes showing strong statistical associations with LDL-C included *APOB*, the *APOE-C1-C4-C2* cluster, *CELSR2*, *HMGCR*, *LDLR*, *PCSK9,* and *CILP2*, whereas *CETP*, *LIPC*, *LIPG*, *LPL, ABCA1, LCAT, GALNT2,* and *MMAB/MVK* showed clear statistical associations with HDL-C. Likewise, the *ZNF259-APOA5-A4-C3-A1* cluster, *LPL*, *ANGPTL3*, *GCKR, TRIB1,* and *MLXIPL* genetic regions were strongly associated with TG levels. Several genetic loci were associated with more than 1 lipid trait, including an SNP at the *APOE-C1-C4-C2* cluster, which showed a strong association with all 3 traits (Table 1).

Recently Identified and Novel Lipid Genes

We also found statistical evidence for potentially novel loci that may influence circulating levels of blood lipids. Supplemental Table IV lists SNPs showing the strongest statistical association at 40 potentially novel loci with $P<1\times10^{-5}$ and no detectable heterogeneity among studies ($P>0.1$). None of these SNPs reached genome-wide statistical association in our metaanalysis. Therefore, to help validate statistical associations found at these potentially novel loci, we examined whether these SNPs showed statistical associations in additional population-based cohorts as part of a replication study (complete results for all 40 SNPs at this stage 2 replication validation step are given in Supplemental Table V).

From our stage 2 analysis, we identified SNPs at 8 loci that showed evidence for independent replication

Table 1. Statistical Associations Between SNPs Showing the Strongest Association Signal ($P<1\times10^{-5}$ With No Heterogeneity Among Studies [$P>0.1$]) in Stage 1 With 1 or More Lipid Traits at Known Lipid Loci

						HDL-C	
SNP	Chr	Pos (Mb)*	Nearest Locus or Loci	Effect Allele†	Effect Allele Frequency†	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity
LDL-C								
rs11206510	1	55.2	*PCSK9*	T	0.77	0.002 (0.004)	0.52	0.35
rs660240	1	109.5	*CELSR2*	A	0.21	0.005 (0.004)	0.22	0.31
rs515135	2	21.2	*APOB*	A	0.19	0.003 (0.004)	0.47	0.52
rs12916	5	74.7	*HMGCR*	T	0.62	0.001 (0.003)	0.80	0.57
rs2954021	8	126.6	*TRIB1*	G	0.50	0.011 (0.003)	1.3×10^{-4}	0.24
rs1558861	11	116.1	*BUD13, ZNF259, APOA5-A4-C3-A1*	T	0.94	0.031 (0.006)	1.7×10^{-7}	0.09
rs2738459	19	11.1	*LDLR*	C	0.48	−0.003 (0.004)	0.34	0.66
rs10401969	19	19.3	*SF4–CILP2*	T	0.91	−0.007 (0.006)	0.26	0.07
rs4420638	19	50.1	*APOE-C1-C4-C2*	G	0.18	**−0.021 (0.004)**	**2.0×10^{-7}**	**0.72**
HDL-C								
rs10489615	1	226.6	*GALNT2*	G	0.60	**0.018 (0.003)**	**3.8×10^{-9}**	**0.19**
rs11902417	2	21.1	*APOB*	G	0.78	**−0.017 (0.003)**	**3.7×10^{-7}**	**0.35**
rs325	8	19.9	*LPL*	T	0.89	**−0.047 (0.005)**	**7.8×10^{-25}**	**0.23**
rs3890182	9	104.7	*ABCA1*	G	0.88	**0.022 (0.004)**	**4.7×10^{-7}**	**0.74**
rs964184	11	116.2	*ZNF259, APOA5-A4-C3-A1*	G	0.12	**−0.029 (0.004)**	**1.6×10^{-11}**	**0.17**
rs9943753	12	108.3	*MYO1H, KCTD10, UBE3B, MMAB, MVK*	G	0.63	**0.016 (0.003)**	**3.2×10^{-6}**	**0.94**
rs261334	15	56.5	*LIPC*	G	0.20	**0.034 (0.004)**	**4.9×10^{-22}**	**0.72**
rs9989419	16	55.5	*CETP*	G	0.60	**0.035 (0.003)**	**1.3×10^{-32}**	**0.28**
rs12449157	16	66.3	*GFOD2–LCAT*	G	0.17	**0.019 (0.004)**	**2.3×10^{-7}**	**0.56**
rs2156552	18	45.4	*LIPG*	T	0.81	**0.028 (0.004)**	**1.7×10^{-12}**	**0.73**
TG								
rs1168013	1	62.7	*DOCK7, ANGPTL3*	G	0.65	0.0001 (0.003)	0.97	0.88
rs6544366	2	21.1	*APOB*	T	0.22	0.016 (0.003)	5.3×10^{-7}	0.34
rs1260333	2	27.7	*GCKR*	C	0.55	0.005 (0.003)	0.08	0.87
rs1178979	7	72.3	*BAZ1B, BCL7B, TBL2, MLXIPL*	A	0.80	−0.010 (0.004)	8.0×10^{-3}	0.83
rs10105606	8	19.9	*LPL*	C	0.68	−0.023 (0.003)	1.7×10^{-14}	0.68
rs2954029	8	126.6	*TRIB1*	T	0.46	0.012 (0.003)	4.5×10^{-5}	0.33
rs4938303	11	116.1	*BUD13, ZNF259, APOA5-A4-C3-A1*	T	0.75	0.018 (0.003)	9.6×10^{-8}	0.37
rs16965220	16	55.6	*CETP, LOC100130044, NLRC5*	C	0.68	−0.006 (0.003)	0.04	0.07
rs2304130	19	19.7	*CILP2–ZNF101*	G	0.09	0.004 (0.006)	0.55	0.78

The genome-wide association metaanalyses (stage 1) for HDL-C and TG are based on data from 7 study populations comprising up to 16 056 and 16 058 participants, respectively. For LDL-C, the metaanalysis is based on data from 8 study populations comprising up to 17 543 participants. Chr indicates chromosome; Pos, position. The strongest SNP association for each lipid trait with no heterogeneity among studies ($P>0.1$) is denoted by bold typeface.

*Based on NCBI Build 35.

†Based on Study 1 (EPIC-Norfolk subcohort). Effect allele corresponds to a forward strand of NCBI Build 36.3.

‡β-Coefficients represent the change in circulating lipid level (natural log) per additional effect allele, adjusted for age and sex.

($P<0.05$) with 1 or more lipid traits and that showed directional consistency with the discovery studies and no material heterogeneity among studies ($P>0.1$). Table 2 summarizes the results for these SNPs. In a combined analysis of all studies (including the discovery GWA studies), 6 of these loci reached genome-wide statistical association ($P<5\times10^{-8}$). These were SNPs at the *MYLIP/GMPR* and *PPP1R3B* loci for LDL-C; at the *SLC39A8*, *TTC39B,* and *FADS1* loci for HDL-C; and at *FADS1* for TG. We note that recently published reports have also found SNPs at the *TTC39B* and *FADS1-FAD2-FADS3* loci to be associated with circulating HDL-C and HDL-C/TG levels, respectively.[5,7] As they were included in our replication strategy, we have retained these SNPs in Table 2 to

Table 1. Continued

TG			LDL-C		
β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity
0.016 (0.008)	0.04	0.30	**0.026 (0.004)**	**1.2×10^{-10}**	**0.69**
−0.004 (0.008)	0.56	0.67	**−0.044 (0.004)**	**1.2×10^{-26}**	**0.67**
−0.009 (0.008)	0.25	0.67	**−0.038 (0.004)**	**2.4×10^{-20}**	**0.15**
0.003 (0.006)	0.64	0.79	**−0.023 (0.003)**	**1.4×10^{-11}**	**0.67**
−0.039 (0.006)	6.3×10^{-11}	0.13	**−0.017 (0.003)**	**1.4×10^{-7}**	**0.40**
−0.142 (0.012)	2.0×10^{-30}	5.4×10^{-3}	**−0.031 (0.006)**	**2.0×10^{-6}**	**0.28**
0.007 (0.007)	0.31	0.58	**−0.018 (0.004)**	**6.6×10^{-6}**	**0.40**
0.095 (0.013)	8.4×10^{-14}	0.04	**0.046 (0.007)**	**9.5×10^{-12}**	**0.74**
0.042 (0.008)	**5.5×10^{-7}**	**0.14**	**0.059 (0.004)**	**1.7×10^{-40}**	**0.16**
−0.023 (0.006)	2.4×10^{-4}	0.24	0.004 (0.003)	0.25	0.99
0.036 (0.007)	2.7×10^{-7}	0.20	0.011 (0.004)	4.0×10^{-3}	0.16
0.097 (0.010)	4.9×10^{-24}	0.08	−0.005 (0.005)	0.34	0.19
0.013 (0.009)	0.16	0.69	0.004 (0.005)	0.43	0.84
0.142 (0.009)	9.0×10^{-53}	1.3×10^{-3}	0.022 (0.005)	6.4×10^{-6}	5.6×10^{-3}
−0.005 (0.007)	0.51	0.89	−0.005 (0.004)	0.20	0.92
0.019 (0.007)	0.01	0.79	−0.002 (0.004)	0.65	0.59
0.003 (0.006)	0.67	0.38	−0.002 (0.003)	0.58	0.82
−0.018 (0.008)	0.02	0.10	0.004 (0.004)	0.31	0.24
−0.020 (0.008)	0.02	0.87	0.011 (0.004)	0.01	0.20
0.035 (0.007)	**6.4×10^{-8}**	**0.88**	**0.009 (0.003)**	6.7×10^{-3}	0.78
−0.036 (0.007)	**1.9×10^{-7}**	**0.20**	**−0.011 (0.004)**	3.8×10^{-3}	0.14
−0.054 (0.006)	**1.7×10^{-19}**	**0.22**	**−0.003 (0.003)**	0.36	0.99
0.054 (0.008)	**2.3×10^{-12}**	**0.71**	**−0.012 (0.004)**	2.5×10^{-3}	0.82
0.067 (0.006)	**3.6×10^{-25}**	**0.14**	**0.0003 (0.003)**	0.94	0.37
−0.040 (0.006)	**1.8×10^{-11}**	**0.13**	**−0.015 (0.003)**	9.2×10^{-7}	0.14
−0.067 (0.007)	**4.1×10^{-21}**	**0.52**	**−0.008 (0.004)**	0.02	0.28
0.028 (0.006)	**9.6×10^{-6}**	**0.16**	**0.009 (0.003)**	0.01	0.92
−0.070 (0.013)	**3.9×10^{-8}**	**0.36**	**−0.036 (0.007)**	1.1×10^{-7}	0.51

present the relevant data and confirm statistical associations at the genome-wide level in our combined analysis.

MYLIP-GMPR

SNP rs2142672 showed strong statistical association with LDL-C levels. The C allele (frequency, 74%) is associated with relatively higher levels of circulating LDL-C. The SNP lies in a distinct block of high linkage disequilibrium (LD) between 2 genes— myosin regulatory light chain interacting protein (*MYLIP*) and guanosine monophosphate reductase (*GMPR*) on chromosome 6p23 (Supplemental Figure If). The illustration suggests that this SNP is correlated with other SNPs that also show similar patterns of statistical association

Table 2. SNPs at Novel or Recently Identified Lipid Loci That Show Replication of Statistical Associations With Circulating Lipid Levels

						Stage 2				Combined			
SNP	Chr	Pos (Mb)*	Nearest Locus or Loci	Effect Allele†	Effect Allele Frequency†	No. of Participants	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity	No. of Participants	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity
LDL-C													
rs2142672	6	16.3	*MYLIP, GMPR*	C	0.74	28 112	0.010 (0.003)	2.0×10^{-4}	0.30	45 655	0.013 (0.002)	2.7×10^{-9}	0.30
rs456598	6	160.5	*IGF2R, SLC22A1*	G	0.87	19 882	−0.010 (0.004)	0.01	0.79	37 425	−0.015 (0.003)	8.4×10^{-7}	0.58
rs2126259	8	9.2	*PPP1R3B*	A	0.10	28 145	−0.014 (0.004)	9.5×10^{-5}	0.13	45 688	−0.018 (0.003)	1.4×10^{-9}	0.10
HDL-C													
rs13107325	4	103.5	*SLC39A8*	T	0.08	22 128	−0.017 (0.006)	2.1×10^{-3}	0.24	38 184	−0.023 (0.004)	1.6×10^{-8}	0.37
rs643531	9	15.3	*TTC39B*	C	0.14	34 152	−0.009 (0.003)	2.6×10^{-3}	0.60	50 208	−0.013 (0.002)	4.1×10^{-8}	0.24
rs174548	11	61.3	*FADS1*	G	0.30	33 930	−0.008 (0.002)	7.6×10^{-5}	0.78	49 986	−0.011 (0.002)	9.9×10^{-10}	0.59
TG													
rs442177	4	88.4	*AFF1*	A	0.60	28 676	0.014 (0.004)	1.2×10^{-3}	0.99	44 734	0.019 (0.004)	1.5×10^{-7}	0.71
rs6867983	5	55.9	*C5orf35*	T	0.14	23 957	0.014 (0.007)	0.04	0.15	40 015	0.024 (0.005)	6.1×10^{-6}	0.10
rs174548	11	61.3	*FADS1*	G	0.30	31 066	0.019 (0.004)	2.0×10^{-5}	0.98	47 124	0.024 (0.004)	8.9×10^{-11}	0.78

The replication analysis (stage 2) is based on data from up to 8 study populations comprising up to 37 774 participants. The combined analysis is based on data from up to 15 studies from stages 1 and 2 and comprising up to 55 497 participants. Chr indicates chromosome; Pos, position.

*Based on NCBI Build 35.

†Based on Study 1 (EPIC-Norfolk subcohort). Effect allele corresponds to a forward strand of NCBI Build 36.3.

‡β-Coefficients represent the change in circulating lipid level (natural log) per additional effect allele, adjusted for age and sex.

and that cluster around the *MYLIP* gene (Supplemental Figure If). Our data confirm results from a recent study that also identifies this locus as one that influences LDL-C.[26] A recent report has also implicated MYLIP (IDOL) in the regulation of circulating LDL-C levels, by its induction of low-density lipoprotein receptor degradation via ubiquitination.[27]

PPP1R3B

SNP rs2126259 lies upstream of the protein phosphatase 1, regulatory (inhibitor) subunit 3B gene (*PPP1R3B*) on chromosome 8p23 and is statistically associated with circulating LDL-C levels. The A allele (10% frequency) is associated with relatively lower levels of circulating LDL-C. The LD structure of this region is modest with surrounding SNPs also showing statistical association (Supplemental Figure Ig). The PPP1R3B protein is involved in the regulation of glycogen metabolism in both muscle and liver.[28] It is possible that its association with circulating LDL-C levels is a reflection of downstream effects on the bioavailability of TG. In addition, in support of our findings for LDL-C, this locus has also been shown to be associated with very-low-density lipoprotein-cholesterol levels.[26]

TTC39B

The C allele (14% frequency) of rs643531 at the tetratricopeptide repeat domain 39B (*TTC39B*) locus on chromosome 9p22 is associated with lower HDL-C levels (Table 2). SNP rs643531 lies within intron 1 of the *TTC39B* gene in a modest LD block that does not contain any other known or putative genes (Supplemental Figure Ic). Again, several highly correlated SNPs in this region show statistical associations with HDL-C levels (Supplemental Figure Ic) in our genome-wide scan. Our data support results from a recent report showing statistical association between an SNP—rs471364— at this locus and HDL-C levels.[5] The 2 SNPs (rs471364 and rs643531) are correlated at an r^2 of 0.74 and show directional consistent associations. The function of the *TTC39B* gene in humans is presently unknown.

SLC39A8

SNP rs13107325 at the solute carrier family 39 (zinc transporter) member 8 (*SLC39A8*) locus on chromosome 4q22 shows strong statistical association with circulating levels of HDL-C (Table 2). It is a nonsynonymous SNP located in exon 8 of the *SLC39A8* gene, which produces a change in amino acid from alanine to threonine. The T allele (frequency 8%) is associated with relatively lower levels of circulating HDL-C and is not materially correlated with any other SNP across 100 kb of genomic sequence spanning the *SLC39A8* gene in HapMap (Supplemental Figure Id). This gene encodes a zinc transporter that has been shown to function in the cellular importation of zinc at the onset of inflammation, and its expression is induced by TNF-α.[29] It is possible that the SLC39A8 molecule might be associated with HDL-C in an inflammatory context.

FADS1

SNP rs174548 at the fatty acid desaturase 1 (*FADS1*) locus on chromosome 11q12 shows strong statistical association with both HDL-C and TG levels. The G allele (30% frequency) is associated with relatively higher TG and lower HDL-C levels (Table 2). The genomic context of this locus is illustrated in Supplemental Figure Ia and Ib. SNP rs174548 lies in a block of clear LD that also contains the *C11orf9/10*, *FEN1* and *FADS2/3* genes. Several highly correlated SNPs within this LD block show statistical association with these traits in our

Table 3. Statistical Associations Between Stage 1 SNPs at Putative Novel or Recently Identified Lipid Loci With Circulating Lipid Levels in Individuals of Indian Asian Ethnicity and a Combined Analysis of All Studies

							Stage 3		Combined			
SNP	Chr	Pos (Mb)*	Nearest Locus or Loci	Effect Allele†	Effect Allele Frequency†	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity	β-Coefficient (SE)‡	*P* Value	*P* Value for Heterogeneity Between Ethnic Groups	*P* Value for Overall Heterogeneity Between Studies
LDL-C												
rs2142672	6	16.3	*MYLIP, GMPR*	C	0.68	0.002 (0.005)	0.72	0.93	0.011 (0.002)	2.2×10^{-8}	0.04	0.28
rs456598	6	160.5	*IGF2R, SLC22A1*	G								
rs2126259	8	9.2	*PPP1R3B*	A	0.13	−0.023 (0.007)	9.6×10^{-4}	0.02	−0.019 (0.003)	6.5×10^{-12}	0.54	0.03
HDL-C												
rs13107325	4	103.5	*SLC39A8*	T								
rs643531	9	15.3	*TTC39B*	C	0.07	−0.013 (0.007)	0.07	0.77	−0.013 (0.002)	7.3×10^{-9}	0.99	0.39
rs174548	11	61.3	*FADS1*	G	0.17	−0.017 (0.004)	1.1×10^{-4}	0.24	−0.011 (0.002)	1.2×10^{-12}	0.19	0.44
TG												
rs442177	4	88.4	*AFF1*	A	0.50	0.027 (0.007)	2.9×10^{-4}	0.27	0.020 (0.003)	3.1×10^{-10}	0.31	0.62
rs6867983	5	55.9	*C5orf35*	T	0.12	0.016 (0.012)	0.16	0.97	0.023 (0.005)	2.6×10^{-6}	0.56	0.21
rs174548	11	61.3	*FADS1*	G	0.17	0.041 (0.010)	2.6×10^{-5}	0.13	0.026 (0.003)	4.5×10^{-14}	0.09	0.44

The Stage 3 analysis for circulating HDL-C, LDL-C, and TG levels is based on data from up to 9665 participants from 4 subsets of the LOLIPOP study (see Supplemental Table II for details). Chr indicates chromosome; Pos, position.

*Based on NCBI Build 35.

†Effect allele corresponds to forward strand of NCBI Build 36.3, and effect allele frequency is based on the control subset of LOLIPOP participants genotyped by the Wellcome chip (Supplemental Table II).

‡β-Coefficients represent the change in circulating lipid level (natural log) per additional effect allele, adjusted for age and sex.

genome-wide scan (Supplemental Figure Ia and Ib), including 2 SNPs in the 3′ untranslated region of the *FADS1* gene. Our study supports findings from a recent report showing that an SNP—rs174547— at this locus is also associated with both HDL-C and TG levels.[5] SNP rs174548 in our study is highly correlated (r^2 0.8) to SNP rs174547 and shows directionally consistent associations. Fatty acid desaturases are involved in the metabolism of polyunsaturated fatty acids in humans, and SNPs at the *FADS1/2* gene cluster have been linked to changes in the fatty acid composition of serum phospholipids in humans.[30]

Examination of Lipid Associations in an Indian Asian Population

In an exploratory analysis, and to provide a wider context for our studies, we examined whether our replicated loci (from Table 2) were also associated with the relevant lipid traits in a population of Indian Asian descent—stage 3. Table 3 shows the results of these analyses. Only 6 of the 8 SNPs were available for analysis in this population (SNPs at 2 loci—*SLC39A8* and *IGF2R/SLC22A1*—were not present or poorly imputed—Supplemental Information). As expected, given the low statistical resolution for this study, we found evidence for independent replication ($P<0.05$) at only 3 loci in Indian Asian individuals— *PPP1R3B* for circulating LDL-C levels, *FADS1* for circulating HDL-C and TG levels, and *AFF1* for circulating TG levels. However, association signals for the other 3 loci (*TTC39B*, *C5orf35* and *MYLIP/GMPR*) were directionally consistent between ethnic groups (Table 3).

In a combined analysis of all studies (stages 1, 2, and 3), we identified an additional locus that reached genome-wide statistical association—*AFF1*—a novel locus for circulating TG ($P=3.1\times10^{-10}$ (Table 3)). SNP rs442177 lies in intron 10 of the *AFF1* gene on chromosome 4q21 in a modest LD block with correlated SNPs showing similar levels of statistical association (Supplemental Figure Ie). The A allele (60% frequency in white European populations) is associated with relatively higher levels of circulating TG. The *AFF1* gene encodes a protein involved in the regulation of cyclin-dependent kinase inhibitor *CDKN1B* and may therefore be involved in cell cycle regulation.[31] Its function with respect to TG metabolism is unknown. For 3 of our 8 novel lipid loci, we did not observe a statistical association with lipids in the Indian Asian population, and for 2 SNPs data were not available (Table 3). These observations could denote limited statistical resolution, differences in LD patterns in Indian Asians compared with Europeans, or that there are no association signals at these loci in Indian Asian populations. However, in further exploratory analyses, examining association signals across a 10-kb region spanning these 5 SNPs in our studies, we found evidence indicating that additional association signals may be present at some of these loci (see Supplemental Information).

Association With Risk of CAD

Given the causal link between circulating LDL-C levels and risk of CAD, and the consistent associations between circulating levels of TG and HDL-C with subsequent risk of CAD, we assessed the association between these known, recently identified, and potentially novel genetic lipid loci and risk of CAD. Table 4 shows the association between these 36 SNPs

Table 4. Statistical Associations Between Stage 1 SNPs at Known/Recently Identified and Putative Novel Loci With Risk of CAD in Individuals of White European Ethnicity

SNP	Chr	Pos (Mb)*	Nearest Locus or Loci	Effect Allele†	Effect Allele Frequency†	No. of Cases/Controls	Odds Ratio (95% Confidence Interval)‡	*P* Value	*P* Value for Heterogeneity
rs11206510	1	55.2	*PCSK9*	T	0.77	6988/19 945	1.07 (1.01 to 1.13)	0.02	0.31
rs1168013	1	62.7	*DOCK7, ANGPTL3*	G	0.65	7002/20 029	0.96 (0.91 to 1.00)	0.06	0.07
rs660240	**1**	**109.5**	*CELSR2*	**A**	**0.21**	**6207/19 638**	**0.85 (0.80 to 0.90)**	**5.5×10^{-8}**	**0.06**
rs10489615	1	226.6	*GALNT2*	G	0.60	6330/19 838	0.97 (0.93 to 1.02)	0.28	0.21
rs11902417	2	21.1	*APOB*	G	0.78	7002/20 040	1.01 (0.96 to 1.07)	0.63	0.06
rs6544366	2	21.1	*APOB*	T	0.22	6976/20 018	0.98 (0.93 to 1.04)	0.56	0.13
rs515135	**2**	**21.2**	*APOB*	**A**	**0.19**	**6449/20 031**	**0.90 (0.85 to 0.96)**	**1.1×10^{-3}**	**0.01**
rs1260333	2	27.7	*GCKR*	C	0.55	6955/19 981	0.94 (0.90 to 0.98)	6.9×10^{-3}	0.20
rs442177	4	88.4	*AFF1*	A	0.60	8187/32 167	1.02 (0.98 to 1.06)	0.40	0.41
rs13107325	4	103.5	*SLC39A8*	T	0.08	4328/20 585	0.89 (0.79 to 0.99)	0.04	0.69
rs6867983	5	55.9	*C5orf35*	T	0.14	8744/32 520	1.02 (0.97 to 1.08)	0.40	0.49
rs12916	5	74.7	*HMGCR*	T	0.62	6928/19 936	0.94 (0.90 to 0.99)	0.01	0.76
rs2142672	6	16.3	*MYLIP, GMPR*	C	0.74	8125/32 346	1.03 (0.98 to 1.07)	0.26	0.90
rs456598	6	160.5	*IGF2R, SLC22A1*	G	0.87	5593/22 096	0.90 (0.84 to 0.97)	7.3×10^{-3}	0.39
rs1178979	7	72.3	*BAZ1B, BCL7B, TBL2, MLXIPL*	A	0.80	6990/20 026	1.03 (0.97 to 1.09)	0.31	0.36
rs2126259	8	9.2	*PPP1R3B*	A	0.10	8258/32 517	1.01 (0.95 to 1.08)	0.73	0.09
rs325	**8**	**19.9**	*LPL*	**T**	**0.89**	**6881/19 882**	**1.20 (1.11 to 1.30)**	**3.2×10^{-6}**	**0.20**
rs10105606	8	19.9	*LPL*	C	0.68	6825/19 797	1.07 (1.02 to 1.12)	5.6×10^{-3}	0.67
rs2954029	8	126.6	*TRIB1*	T	0.46	6997/20 734	0.93 (0.89 to 0.97)	7.4×10^{-4}	0.30
rs2954021	**8**	**126.6**	*TRIB1*	**G**	**0.50**	**7018/20 765**	**0.92 (0.88 to 0.96)**	**2.1×10^{-4}**	**0.56**
rs643531	9	15.3	*TTC39B*	C	0.14	9075/34 589	0.98 (0.93 to 1.04)	0.50	0.37
rs3890182	9	104.7	*ABCA1*	G	0.88	7003/20 036	0.96 (0.90 to 1.03)	0.28	0.23
rs174548	11	61.3	*FADS1*	G	0.30	9068/34 364	1.01 (0.97 to 1.06)	0.52	0.45
rs4938303	11	116.1	*BUD13, ZNF259, APOA5-A4-C3-A1*	T	0.75	6601/19 638	0.93 (0.89 to 0.98)	9.5×10^{-3}	0.11
rs1558861	11	116.1	*BUD13, ZNF259, APOA5-A4-C3-A1*	T	0.94	4654/13 359	0.88 (0.77 to 1.00)	0.04	0.75
rs964184	**11**	**116.2**	*ZNF259, APOA5-A4-C3-A1*	**G**	**0.12**	**6958/20 001**	**1.22 (1.14 to 1.30)**	**1.2×10^{-9}**	**0.42**
rs9943753	12	108.3	*MYO1H, KCTD10, UBE3B, MMAB, MVK*	G	0.63	3540/15 657	1.01 (0.94 to 1.09)	0.75	0.51
rs261334	15	56.5	*LIPC*	G	0.20	6414/19 980	1.05 (0.99 to 1.11)	0.10	0.71
rs9989419	16	55.5	*CETP*	G	0.60	6991/20 009	1.01 (0.96 to 1.06)	0.71	0.16
rs16965220	16	55.6	*CETP, LOC100130044, NLRC5*	C	0.68	7001/20 029	1.00 (0.95 to 1.05)	0.94	0.49
rs12449157	16	66.3	*GFOD2–LCAT*	G	0.17	6935/19 961	0.96 (0.90 to 1.02)	0.16	0.73
rs2156552	18	45.4	*LIPG*	T	0.81	6991/20 034	0.97 (0.91 to 1.03)	0.32	0.25
rs2738459	19	11.1	*LDLR*	C	0.48	3540/15 657	0.96 (0.89 to 1.03)	0.23	0.05
rs10401969	19	19.3	*SF4–CILP2*	T	0.91	6723/19 654	1.08 (0.97 to 1.20)	0.17	0.58
rs2304130	19	19.7	*CILP2–ZNF101*	G	0.09	3540/15 657	1.02 (0.90 to 1.16)	0.74	0.22
rs4420638	**19**	**50.1**	*APOE-C1-C4-C2*	**G**	**0.18**	**7004/20 033**	**1.17 (1.10 to 1.24)**	**1.5×10^{-6}**	**0.02**

The metaanalysis is based on data from nine studies comprising up to 9633 cases and 38 684 controls. SNPs showing both genome-wide statistical association with lipids (Table 1) and statistical association with CAD risk after adjustment for multiple testing ($P<0.0013$ after testing 36 SNPs) are denoted by bold typeface.

*Based on NCBI Build 35.

†Based on Study 1 (EPIC-Norfolk subcohort). Effect allele corresponds to forward strand of NCBI Build 36.3.

‡Odds ratios are based on the additive model.

linked to lipid metabolism in our data and risk of CAD in up to 9633 cases and 38 684 controls.

As expected, and given the prior associations between these loci and blood lipids, a much greater proportion of these SNPs showed statistical associations with CAD risk at $P<0.05$ than expected by chance alone, taking into account any correlated SNPs (Table 4). We identified 6 genetic loci that showed both genome-wide statistical association with blood lipids and statistical association with CAD after adjustment for multiple testing ($P<0.0013$ after testing 36 SNPs) (Table 4). Specifically, we confirm the association between variation at the *CELSR2*[20] and *APOB* genes and variation at the *APOE-C1-C4-C2* cluster, which influence mainly LDL-C levels, and risk of CAD (Figure 1). None of the genetic variants largely or specifically associated with HDL-C showed statistical association with CAD risk after correction for multiple testing. Notably, SNPs at the *ZNF259-APOA5-A4-C3-A1* cluster—which reached genome-wide statistical association—and at the *TRIB1* and *LPL* loci, which show strongest association with TG levels (Figure 2), were also statistically associated with risk of CAD after adjustment for multiple testing (Table 4 and Figure 1). The direction of association with CAD risk for all of these SNPs was consistent with their association with lipid levels (Table 1). However, several of the SNPs at these loci were associated with more than 1 lipid trait (Figure 2 and Table 1). Of note, only SNPs at *CELSR2* and *APOB* showed specific associations with LDL-C. By contrast, only SNPs at the *LPL* locus showed clear associations with HDL-C and TG, but they were not associated with LDL-C in our studies (Figure 2 and Table 1).

Discussion

Our studies have identified 3 novel loci (*PPP1R3B* for LDL-C, *SLC39A8* for HDL-C, and AFF1 for TG) associated with variation in circulating LDL-C, HDL-C, and TG. We also provide strong statistical evidence for 6 loci that influence levels of blood lipids and risk of CAD. In addition to those that are largely associated with LDL-C concentrations, we show that genetic loci mainly associated with circulating TG are also associated with risk of CAD. Collectively, these studies potentially provide new insights into biological regulation of lipid metabolism and the etiology of CAD.

We provide robust statistical evidence for the association of 3 novel genetic loci with circulating LDL-C, HDL-C, and TG levels, in addition to confirming the recently reported novel associations for circulating LDL-C with SNPs at *MYLIP/GMPR*, HDL-C levels with SNPs at the *TTC39B* locus, and for both circulating HDL-C and TG levels at the *FADS1* locus.[5,26] The function of the 3 novel loci, *PPP1R3B*, *SLC39A8*, and *AFF1*, in lipid metabolism is not known. However, consistent with our results, a recent study has shown that *PPP1R3B* is also associated with very-low-density lipoprotein-cholesterol levels[26]. Fine-mapping and functional studies, including large-scale resequencing to help identify common and rare functional variants,[32] might help clarify the role of proteins encoded by these genes in lipid metabolism and relevant disorders.

Recent reports have identified several potentially novel loci for circulating lipids.[5,7] One of these reports presents an updated metaanalysis and has used the same threshold for stage 2 SNP selection as our study ($P<1\times10^{-5}$).[5] By using this arbitrary threshold for selection, we will undoubtedly have missed some additional novel loci. For example, 1 report identified an SNP—rs1501908—that lies between the *TIMD4* and *HAVCR1* loci and is reproducibly associated with circulating levels of LDL-C. We selected this SNP for stage 2 replication testing, but it did not reach statistical association in our stage 2 samples ($P=0.1$) (Supplemental Tables IV and V). However, the association signal in our data is directionally consistent with that found in the original report. These findings suggest that further novel loci involved in the regulation of blood lipids exist—providing opportunities for additional insights into lipid biology and potential therapeutic targets—and therefore highlight the need for a more comprehensive analysis of all available studies to gain appropriate statistical resolution to identify these loci.

We identified 6 lipid genes that show strong statistical association with CAD risk. Associations at these loci were directionally consistent with their associations with blood lipids. Three of these are predominantly associated with circulating LDL-C levels—*APOB*, *APOE* cluster, and *CELSR2*—reiterating the causal link between LDL-C and CAD,[33] and as previously reported.[12,20,34] Interestingly, our variant at the APOE locus, *rs4420638*, was correlated (data not shown) with 1 of the canonical APOE variants ($r^2=0.71$ with E4) but showed no correlation with E2 ($r^2=0.018$), suggesting that other independent variants may contribute to the variation in LDL-C and risk of CAD.

Whereas previous studies and reviews have provided only suggestive or inconsistent evidence for the *APOB* locus and CAD risk,[12,34] we confirm that common variation at the *APOB* locus is associated with risk of CAD—in line with the effect of rare, highly deleterious mutations at this gene.[35] However, some genes known to be implicated in Mendelian forms of hypercholesterolemia and more recently myocardial infarction, including *LDLR* and *PCSK9*[21,36,37], showed only suggestive evidence for association with CAD risk in our data. Because of limited statistical power to detect associations, larger scale studies of these and other genetic variants that influence LDL-C levels may help reliably determine their association with CAD risk.

We also provide compelling statistical evidence that genetic variants at loci predominantly associated with both circulating blood TG and HDL-C are also associated with risk of CAD—specifically at the *ZNF259-APOA5-A4-C3-A1* cluster, *TRIB1*, and *LPL* loci. The *TRIB1* locus is a recently identified lipid gene that predominantly influences TG but is also associated with LDL-C and HDL-C. One previous report has shown suggestive evidence for an association between a SNP at this locus and CAD risk.[12] We provide convincing evidence for association with CAD risk in our studies. The *LPL* variant in our study (rs325) is in perfect LD with the known S447X variant—a gain-of-function mutation, which causes a 2–amino acid truncation in the enzyme and increases its activity.[38] Our data are consistent with these observations and suggest that LPL activity may be causally linked to CAD risk. By contrast, the CAD-risk variant we identified in the *ZNF259-APOA5-A4-C3-A1* cluster was largely uncorrelated

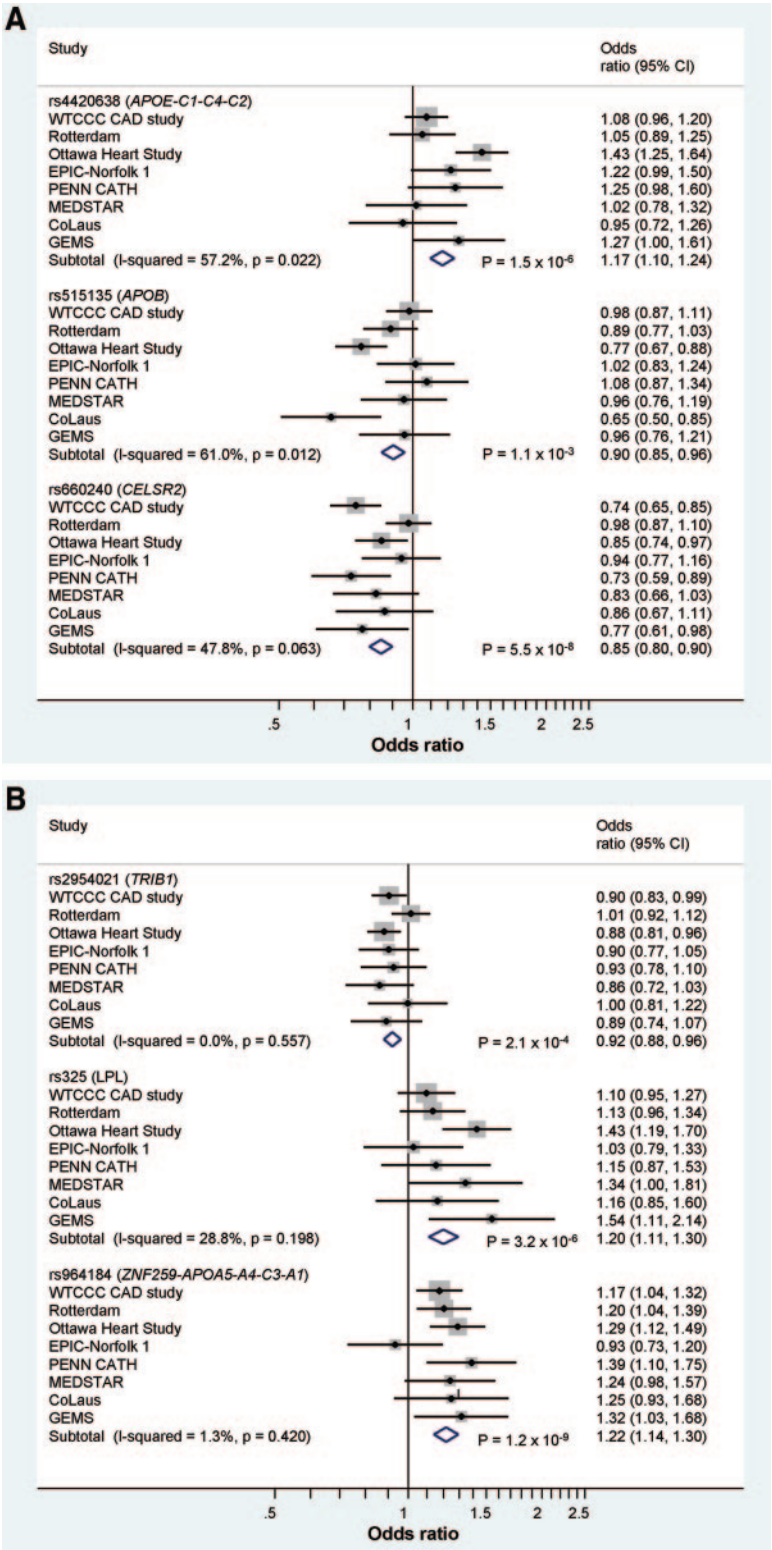


Figure 1. Associations between SNPs at loci predominantly associated with circulating levels of LDL-C (A) and TG/HDL-C (B) with risk of CAD in 8 studies comprising up to 7,018 cases and 20 765 controls (see Methods for details). There was no material heterogeneity among studies for these associations (Table 4).

with variants at the *APOA5* and related genes that have been previously linked to TG levels and CAD risk.[39] A recent systematic review of known genetic variants at the *LPL* locus and CAD risk provided only suggestive evidence for association with CAD risk.[40] Similarly, previous reports have provided only weak and inconsistent evidence to suggest that variation at the *APOA5* cluster is implicated in CAD risk.[39,41] Collectively, our data, based on an unbiased analytic framework, confirm that the *LPL, TRIB1,* and *ZNF259-APOA5-A4-C3-A1* cluster are CAD susceptibility loci.

Importantly, consistent with the biological role of the *LPL* locus,[42] SNPs at this locus were not associated with LDL-C levels in our large scale analysis, suggesting that the association with CAD risk is independent of LDL-C. However,

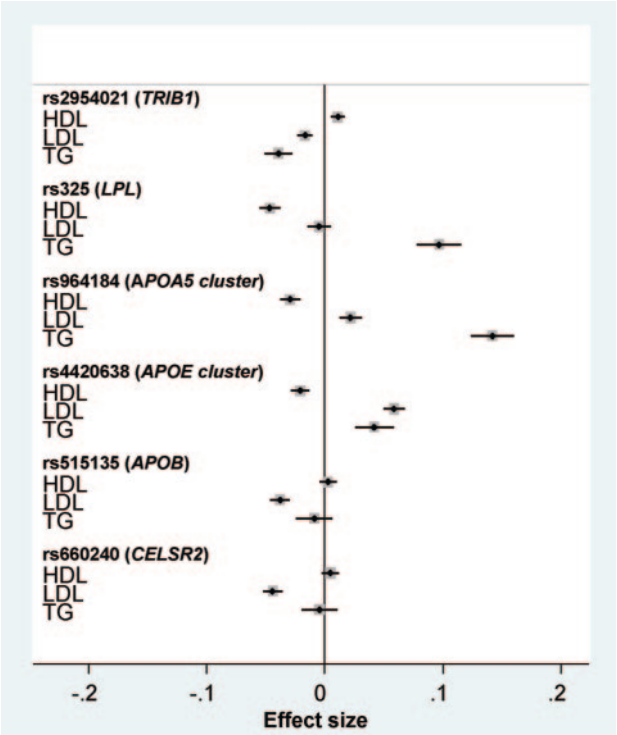


Figure 2. Associations between SNPs and circulating lipids at known or recently identified loci that show statistical association with risk of CAD in Figure 1 and Table 4. Associations and effect sizes are based on stage 1 metaanalyses and natural log–transformed data (see Table 1 for details).

other loci that were associated with CAD risk showed robust associations with potentially multiple lipid traits (including LDL-C). By contrast, some loci showed comparable magnitudes of association with 1 or more lipid traits but showed inconsistent magnitudes of association with CAD risk. These differences might be due to limited statistical power or the differential impact of comparable differences in these lipids on risk of CAD. Statistical analyses adjusting for these intermediate phenotypes (lipid levels) when examining SNP-CAD risk associations may help disentangle the impact of these genetic variants on lipid levels and CAD risk. However, these analyses would require large-scale prospective studies with information on genetic variants, biomarkers and subsequent disease risk, which are not available across most of the studies used in the current analysis.

These lipid and CAD risk loci may also have pleiotropic actions.[5] As a result, interpretation of interpretation of these findings is complex. From a qualitative perspective, these findings may suggest that some, but not all, biological mechanisms involved in TG and HDL-C regulation and metabolism or their correlates (including atherogenic very-low-density lipoprotein remnant lipoproteins[43]) may be implicated in the etiology of CAD.[43–45] In this context, these data suggest that therapeutic approaches that target specific lipid pathways might have a potentially greater impact on reducing risk of CAD—particularly in the context of our findings for the *LPL* locus.

None of the genetic loci showing reproducible and specific association with HDL-C levels (including *CETP*), showed strong evidence for association with CAD risk. Because of limited statistical power, in terms of the expected magnitudes of the associations among HDL-C levels, HDL-C SNPs, and CAD risk,[46] we may have missed HDL-C genetic loci that also show association with CAD risk. Furthermore, the functional relationship of HDL-C to CAD risk is inherently complex, and plasma concentrations of HDL-C are not always a reliable marker of reverse cholesterol transport or other biological functions of HDL, including antiinflammatory effects.[47,48]

We assessed the generalizability of our novel SNP-lipid associations in a population of Indian Asian ethnicity and found that there was directional consistency between the 2 populations for statistically associated SNPs, with no strong heterogeneity between the 2 ethnic groups. However, we only had limited statistical resolution to detect any differences in the magnitudes of these associations between ethnic groups. It will be important to fully characterize the associations among all known genetic regulators of blood lipids and their link to CAD risk in this and other ethnically distinct populations. Importantly, genetic epidemiological approaches may help determine whether the marked differences in the prevalence of some metabolic diseases among populations have a genetic basis.[49]

In conclusion, our studies have identified 4 novel loci associated with variation in circulating lipid concentrations and indicate that, with the caveats outlined above, genetic variants that influence lipid concentrations (primarily those that are associated with circulating LDL-C or specific metabolic and regulatory pathways for both TG and HDL-C) are also associated with risk of CAD. These findings potentially provide new insights into the biological mechanisms underlying lipid metabolism and CAD risk.

Sources of Funding

This work was supported by the United Kingdom Medical Research Council, Wellcome Trust (WT), British Heart Foundation, European Commission, and GlaxoSmithKline. Specifically, we acknowledge use of genotype data from the 1958 British birth cohort DNA collection, funded by Medical Research Council Grant G0000934 and Wellcome Trust Grant 068545/Z/02. Dr Barroso and Dr Wheeler acknowledge support from European Union FP6 funding (contract no LSHM-CT-2003-503041). Dr Sandhu and Dr Ricketts are funded by the British Heart Foundation (PG/08/094) and Medical Research Council (G0801566). Dr Barroso (077016/Z/05/Z), Dr Peltonen, Dr Inouye, and Dr Deloukas are funded by the Wellcome Trust. Some computation was done on the Vital-IT system at the Swiss Institute of Bioinformatics. The Ottawa Heart Study is funded by HSFO NA6001 and CIHR. The GEMS study was sponsored in part by GlaxoSmithKline. The CoLaus study was supported by grants from GlaxoSmithKline, the Swiss National Science Foundation (Grant 33CSCO-122661) and the Faculty of Biology and Medicine of Lausanne. This research used resources provided by the Type 1 Diabetes Genetics Consortium (T1DGC), a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development, and Juvenile Diabetes Research Foundation International (JDRF) and was supported by U01 DK062418. The T1DGC GWAS project was funded by the National Institute of Diabetes and Digestive and Kidney Diseases and JDRF and was coordinated by the JDRF/WT Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (Cambridge, United Kingdom), which is funded by the JDRF International, the Wellcome Trust, and the National Institute for Health Research Cambridge Biomedical Research Center. The Cambridge Institute for Medical Research is in receipt of a Wellcome Trust Strategic Award (079895). Recruitment of the PennCATH cohort was supported by the Cardiovascular Institute of the University of Pennsylvania. Genotyping for the PennCATH cohort was

performed at the Center for Applied Genomics at the Children's Hospital of Philadelphia and supported by GlaxoSmithKline through an Alternate Drug Discovery Initiative research alliance award (to M.P.R. and D.J.R.) with the University of Pennsylvania School of Medicine. The Rotterdam Study was supported by the Netherlands Organisation for Scientific Research (NWO Groot, 175.010.2005.011, 911.03.012), the Research Institute for Diseases in the Elderly (RIDE2, 01493015), and the Netherlands Genomic Initiative (NGI/NWO) (050-060-810). NFBC 1966 received financial report from the Academy of Finland (project Grants 104781, 120315, 1114194, Center of Excellence in Complex Disease Genetics), University Hospital Oulu, Biocenter, University of Oulu, Oulu, Finland, National Heart, Lung, and Blood Institute Grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council (G0500539, PrevMetSyn/Medical Research Council), and the Wellcome Trust (GR069224), United Kingdom. The DNA extractions, sample quality controls, biobank upkeep, and aliquotting were performed in the National Public Health Institute, Biomedicum, Helsinki, Finland, supported financially by the Academy of Finland and Biocentrum Helsinki.

Disclosures

Drs Waterworth, Song, Yuan, Lim, Chen, and Mooser are full-time employees of GlaxoSmithKline. Dr Sandhu, McPherson, Reilly, and Rader have received research funding from GlaxoSmithKline. Dr Grundy has consulted for GlaxoSmithKline. YAK has received research funding from, has provided Continuing Medical Education on behalf of, and has acted as a consultant to AstraZeneca, Laboratories Fournier, Merck/Schering Plough, Novartis, Pfizer, and Sanofi-Aventis, and also owns some Orion-Pharma stocks. Dr Barter has received consulting fees from Abbott, AstraZeneca, BMS, CSL, Genfit, Merck, Pfizer, Resverlogix and Roche; lecture fees from Abbott, AstraZeneca, Merck, Pfizer, and Roche; and grant support from Pfizer. Drs Vollenweider and Waeber received grant money from GlaxoSmithKline to fund the CoLaus study.

References

1. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J III. Factors of risk in the development of coronary heart disease–six year follow-up experience. The Framingham Study. *Ann Intern Med*. 1961;55:33–50.
2. Barkowski RS, Frishman WH. HDL metabolism and CETP inhibition. *Cardiol Rev*. 2008;16:154–162.
3. Singh IM, Shishehbor MH, Ansell BJ. High-density lipoprotein as a therapeutic target: a systematic review. *J Am Med Assoc*. 2007;298:786–798.
4. Smith GD, Timpson N, Ebrahim S. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. *Ann Med*. 2008;40:524–541.
5. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burtt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS, Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL, Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet*. 2009;41:56–65.
6. Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, Song K, Yuan X, Johnson T, Ashford S, Inouye M, Luben R, Sims M, Hadley D, McArdle W, Barter P, Kesaniemi YA, Mahley RW, McPherson R, Grundy SM, Bingham SA, Khaw KT, Loos RJ, Waeber G, Barroso I, Strachan DP, Deloukas P, Vollenweider P, Wareham NJ, Mooser V. LDL-cholesterol concentrations: a genome-wide association study. *Lancet*. 2008;371:483–491.
7. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Doring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet*. 2009;41:47–55.
8. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008;40:189–197.
9. Kooner JS, Chambers JC, Aguilar-Salinas CA, Hinds DA, Hyde CL, Warnes GR, Gomez Perez FJ, Frazer KA, Elliott P, Scott J, Milos PM, Cox DR, Thompson JF. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet*. 2008;40:149–151.
10. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, Jones CG, Zaitlen NA, Varilo T, Kaakinen M, Sovio U, Ruokonen A, Laitinen J, Jakkula E, Coin L, Hoggart C, Collins A, Turunen H, Gabriel S, Elliot P, McCarthy MI, Daly MJ, Jarvelin MR, Freimer NB, Peltonen L. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet*. 2009;41:35–46.
11. Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, Ahmadi K, Dobson RJ, Marcano AC, Hajat C, Burton P, Deloukas P, Brown M, Connell JM, Dominiczak A, Lathrop GM, Webster J, Farrall M, Spector T, Samani NJ, Caulfield MJ, Munroe PB. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet*. 2008;82:139–149.
12. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008;40:161–169.
13. Stirnadel H, Lin X, Ling H, Song K, Barter P, Kesaniemi YA, Mahley R, McPherson R, Waeber G, Bersot T, Cohen J, Grundy S, Mitchell B, Mooser V, Waterworth D. Genetic and phenotypic architecture of metabolic syndrome-associated components in dyslipidemic and normolipidemic subjects: the GEMS Study. *Atherosclerosis*. 2008;197:868–876.
14. Wyszynski DF, Waterworth DM, Barter PJ, Cohen J, Kesaniemi YA, Mahley RW, McPherson R, Waeber G, Bersot TP, Sharma SS, Nolan V, Middleton LT, Sundseth SS, Farrer LA, Mooser V, Grundy SM. Relation between atherogenic dyslipidemia and the Adult Treatment Program-III definition of metabolic syndrome (Genetic Epidemiology of Metabolic Syndrome Project). *Am J Cardiol*. 2005;95:194–198.
15. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, Wareham N. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer*. 1999;80(suppl 1):95–103.
16. Stewart AF, Dandona S, Chen L, Assogba O, Belanger M, Ewart G, LaRose R, Doelle H, Williams K, Wells GA, McPherson R, Roberts R. Kinesin family member 6 variant Trp719Arg does not associate with angiographically defined coronary artery disease in the Ottawa Heart Genomics Study. *J Am Coll Cardiol*. 2009;53:1471–1472.
17. Lindgren CM, Heid IM, Randall JC, Lamina C, Steinthorsdottir V, Qi L, Speliotes EK, Thorleifsson G, Willer CJ, Herrera BM, Jackson AU, Lim N, Scheet P, Soranzo N, Amin N, Aulchenko YS, Chambers JC, Drong A, Luan J, Lyon HN, Rivadeneira F, Sanna S, Timpson NJ, Zillikens MC, Zhao JH, Almgren P, Bandinelli S, Bennett AJ, Bergman RN, Bonnycastle LL, Bumpstead SJ, Chanock SJ, Cherkas L, Chines P, Coin L, Cooper C, Crawford G, Doering A, Dominiczak A, Doney AS, Ebrahim S, Elliott P, Erdos MR, Estrada K, Ferrucci L, Fischer G, Forouhi NG, Gieger C, Grallert H, Groves CJ, Grundy S, Guiducci C, Hadley D, Hamsten A, Havulinna AS, Hofman A, Holle R, Holloway JW, Illig T, Isomaa B, Jacobs LC, Jameson K, Jousilahti P, Karpe F, Kuusisto J, Laitinen J, Lathrop GM, Lawlor DA, Mangino M, McArdle WL, Meitinger T, Morken MA, Morris AP, Munroe P, Narisu N, Nordstrom A, Nordstrom P, Oostra BA, Palmer CN, Payne F, Peden JF, Prokopenko I, Renstrom F, Ruokonen A, Salomaa V, Sandhu MS, Scott LJ, Scuteri A, Silander K, Song K, Yuan X, Stringham HM, Swift AJ, Tuomi T, Uda M,

Vollenweider P, Waeber G, Wallace C, Walters GB, Weedon MN, Witteman JC, Zhang C, Zhang W, Caulfield MJ, Collins FS, Davey SG, Day IN, Franks PW, Hattersley AT, Hu FB, Jarvelin MR, Kong A, Kooner JS, Laakso M, Lakatta E, Mooser V, Morris AD, Peltonen L, Samani NJ, Spector TD, Strachan DP, Tanaka T, Tuomilehto J, Uitterlinden AG, van Duijn CM, Wareham NJ, Hugh W, Waterworth DM, Boehnke M, Deloukas P, Groop L, Hunter DJ, Thorsteinsdottir U, Schlessinger D, Wichmann HE, Frayling TM, Abecasis GR, Hirschhorn JN, Loos RJ, Stefansson K, Mohlke KL, Barroso I, McCarthy MI. Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet.* 2009;5:e1000508.

18. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*. 2009;41:703–707.
19. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–678.
20. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, Konig IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443–453.
21. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Ardissino D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Mannucci PM, Schwartz SM, Siscovick DS, Yee J, Friedlander Y, Elosua R, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Kathiresan S, Meigs JB, Williams G, Nathan DM, MacRae CA, O'Donnell CJ, Salomaa V, Havulinna AS, Peltonen L, Melander O, Berglund G, Voight BF, Kathiresan S, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Musunuru K, Daly MJ, Purcell S, Voight BF, Purcell S, Nemesh J, Korn JM, McCarroll SA, Schwartz SM, Yee J, Kathiresan S, Lucas G, Subirana I, Elosua R, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Samani NJ, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Schunkert H, Erdmann J, Linsel-Nitschke P, Lieb W, Ziegler A, Konig I, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Schunkert H, Samani NJ, Erdmann J, Ouwehand W, Hengstenberg C, Deloukas P, Scholz M, Cambien F, Reilly MP, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney J, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Epstein SE, Rader DJ, Scheffold T, Berger K, Stoll M, Huge A, Girelli D, Martinelli N, Olivieri O, Corrocher R, Morgan T, Spertus JA, McKeown P, Patterson CC, Schunkert H, Erdmann E, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Engert JC, Do R, Xie C, Anand S, Kathiresan S, Ardissino D, Mannucci PM, Siscovick D, O'Donnell CJ, Samani NJ, Melander O, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Altshuler D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet*. 2009;41:334–341.
22. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39:906–913.
23. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55:997–1004.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
25. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics*. 2007;23:1294–1296.
26. Chasman DI, Pare G, Mora S, Hopewell JC, Peloso G, Clarke R, Cupples LA, Hamsten A, Kathiresan S, Malarstig A, Ordovas JM, Ripatti S, Parker AN, Miletich JP, Ridker PM. Forty-three loci associated with plasma lipoprotein size, concentration and cholesterol content in genome-wide analysis. *PLoS Genet.* 2009;5:e1000730.
27. Zelcer N, Hong C, Boyadjian R, Tontonoz P. LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science*. 2009;325:100–104.
28. Munro S, Cuthbertson DJ, Cunningham J, Sales M, Cohen PT. Human skeletal muscle expresses a glycogen-targeting subunit of PP1 that is identical to the insulin-sensitive glycogen-targeting subunit G(L) of liver. *Diabetes*. 2002;51:591–598.
29. Besecker B, Bao S, Bohacova B, Papp A, Sadee W, Knoell DL. The human zinc transporter SLC39A8 (Zip8) is critical in zinc-mediated cytoprotection in lung epithelia. *Am J Physiol Lung Cell Mol Physiol*. 2008;294:L1127–L1136.
30. Schaeffer L, Gohlke H, Muller M, Heid IM, Palmer LJ, Kompauer I, Demmelmair H, Illig T, Koletzko B, Heinrich J. Common genetic variants of the FADS1 FADS2 gene cluster and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. *Hum Mol Genet*. 2006;15:1745–1756.
31. Xia ZB, Popovic R, Chen J, Theisler C, Stuart T, Santillan DA, Erfurth F, Diaz MO, Zeleznik-Le NJ. The MLL fusion gene, MLL-AF4, regulates cyclin-dependent kinase inhibitor CDKN1B (p27kip1) expression. *Proc Natl Acad Sci U S A*. 2005;102:14028–14033.
32. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9:356–369.
33. Grundy SM. Promise of low-density lipoprotein-lowering therapy for primary and secondary prevention. *Circulation*. 2008;117:569–573.
34. Benn M. Apolipoprotein B levels, APOB alleles, and risk of ischemic cardiovascular disease in the general population, a review. *Atherosclerosis*. 2009;206:17–30.
35. Cambien F, Tiret L. Genetics of cardiovascular diseases: from single mutations to the whole genome. *Circulation*. 2007;116:1714–1724.
36. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med*. 2006;354:1264–1272.
37. Soutar AK, Naoumova RP. Mechanisms of disease: genetic causes of familial hypercholesterolemia. *Nat Clin Pract Cardiovasc Med.* 2007;4:214–225.
38. Nierman MC, Rip J, Kuivenhoven JA, Sakai N, Kastelein JJ, de Sain-van der Velden MG, Stroes ES, Prinsen BH. Enhanced apoB48 metabolism in lipoprotein lipase X447 homozygotes. *Atherosclerosis*. 2007;194:446–451.
39. Vaessen SF, Schaap FG, Kuivenhoven JA, Groen AK, Hutten BA, Boekholdt SM, Hattori H, Sandhu MS, Bingham SA, Luben R, Palmen JA, Wareham NJ, Humphries SE, Kastelein JJ, Talmud PJ, Khaw KT. Apolipoprotein A-V, triglycerides and risk of coronary artery disease: the prospective Epic-Norfolk Population Study. *J Lipid Res*. 2006;47:2064–2070.
40. Sagoo GS, Tatt I, Salanti G, Butterworth AS, Sarwar N, van MM, Jukema JW, Wiman B, Kastelein JJ, Bennet AM, de FU, Danesh J, Higgins JP. Seven lipoprotein lipase gene polymorphisms, lipid fractions, and coronary disease: a HuGE association review and meta-analysis. *Am J Epidemiol*. 2008;168:1233–1246.
41. Lai CQ, Parnell LD, Ordovas JM. The APOA1/C3/A4/A5 gene cluster, lipid metabolism and cardiovascular disease risk. *Curr Opin Lipidol*. 2005;16:153–166.
42. Rahalkar AR, Giffen F, Har B, Ho J, Morrison KM, Hill J, Wang J, Hegele RA, Joy T. Novel LPL mutations associated with lipoprotein lipase deficiency: two case reports and a literature review. *Can J Physiol Pharmacol*. 2009;87:151–160.
43. Miller M, Ginsberg HN, Schaefer EJ. Relative atherogenicity and predictive value of non-high-density lipoprotein cholesterol for coronary heart disease. *Am J Cardiol*. 2008;101:1003–1008.
44. 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. *J Am Med Assoc*. 2007;298:299–308.
45. Criqui MH. Triglycerides and coronary heart disease revisited (again). *Ann Intern Med*. 2007;147:425–427.
46. Thompson A, Di AE, Sarwar N, Erqou S, Saleheen D, Dullaart RP, Keavney B, Ye Z, Danesh J. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. *J Am Med Assoc*. 2008;299:2777–2788.
47. Movva R, Rader DJ. Laboratory assessment of HDL heterogeneity and function. *Clin Chem*. 2008;54:788–800.
48. Barter PJ, Puranik R, Rye KA. New insights into the role of HDL as an anti-inflammatory agent in the prevention of cardiovascular disease. *Curr Cardiol Rep*. 2007;9:493–498.
49. Godsland IF, Johnston DG, Chaturvedi N. Mechanisms of disease: lessons from ethnicity in the role of triglyceride metabolism in ischemic heart disease. *Nat Clin Pract Endocrinol Metab*. 2007;3:530–538.

Supplement Material

Exploratory analyses

To see whether there were other statistical signals present at the *SLC39A8*, *TTC39B*, *C5orf35, MYLIP/GMPR* and *IGF2R/SLC22A1* loci in individuals of Indian Asian descent, we examined association summary statistics across a 10 kb region spanning SNPs from **Table 3** in our studies. There were no association signals present at the *SLC39A8* or *TTC39B* loci for circulating HDL-c levels (data not shown). The *SLC39A8* SNP rs13107325 is relatively rare (8% frequency) in white European individuals so it may not be present in individuals of Indian Asian descent. Indeed it is not present in individuals of Chinese, Japanese or Yoruban ethnicity in HapMap. Likewise, the frequency of SNP rs643531 at the *TTC39B* locus is 7% in Indian Asian individuals compared to 14% in white European participants (**Tables 2 and 3**). The signal for this SNP was directionally consistent with data from white Europeans, so it is likely that we did not have had the statistical resolution to reliably detect an association in Indian Asians. For the *C5orf35* locus we observed a nominal statistical signal at SNP rs4301182 for association with TG at P = 0.03 (β coefficient 0.031 (standard error 0.015)) that is directionally consistent with that seen for SNP rs6867983 in our white European samples. SNP rs4301182 is correlated to SNP rs6867983 at an r^2 of 0.81 in individuals of white European ethnicity in HapMap. For circulating LDL-c levels we did not observe any signals at the *MYLIP/GMPR* locus in Indian Asian individuals, however for the *IGF2R/SLC22A1* locus there were putative statistical signals at two uncorrelated SNPs—rs2065396 and rs3777409 (P = 0.03 and P = 0.04 respectively). This suggests that there may be associations present at this locus in Indian Asians.

Supplementary References

1. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18(6):499-502.

2. Day N, Oakes S, Luben R et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. Br J Cancer 1999;80 Suppl 1:95-103.

3. Sandhu MS, Waterworth DM, Debenham SL et al. LDL-cholesterol concentrations: a genome-wide association study. Lancet 2008;371(9611):483-491.

4. Anon. Genetic information from the British 1958 birth cohort. DNA collection. http://www.b58cgene.sgul.ac.uk/collection.php. 2009.

5. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.

6. Marques-Vidal P, Pecoud A, Hayoz D et al. Prevalence and characteristics of vitamin or dietary supplement users in Lausanne, Switzerland: the CoLaus study. Eur J Clin Nutr 2009;63(2):273-281.

7. Stirnadel H, Lin X, Ling H et al. Genetic and phenotypic architecture of metabolic syndrome-associated components in dyslipidemic and normolipidemic subjects: the GEMS Study. Atherosclerosis 2008;197(2):868-876.

8. Wyszynski DF, Waterworth DM, Barter PJ et al. Relation between atherogenic dyslipidemia and the Adult Treatment Program-III definition of metabolic syndrome (Genetic Epidemiology of Metabolic Syndrome Project). Am J Cardiol 2005;95(2):194-198.

9. Kooner JS, Chambers JC, Aguilar-Salinas CA et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. Nat Genet 2008;40(2):149-151.

10. Pilia G, Chen WM, Scuteri A et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. PLoS Genet 2006;2(8):e132.

11. Scuteri A, Sanna S, Chen WM et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet 2007;3(7):e115.

12. Willer CJ, Sanna S, Jackson AU et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 2008;40(2):161-169.

13. Scott LJ, Mohlke KL, Bonnycastle LL et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316(5829):1341-1345.

14. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Stat Med 2005;24(19):2911-2935.

15. Stewart AF, Dandona S, Chen L et al. Kinesin family member 6 variant Trp719Arg does not associate with angiographically defined coronary artery disease in the Ottawa Heart Genomics Study. J Am Coll Cardiol 2009;53(16):1471-1472.

16. Lindgren. Genome wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. PLoS Genet 2009 Jun;5(6):e1000508.

17. Barrett JC, Clayton DG, Concannon P et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009 *in press*.

18. Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. Acta Paediatr Scand 1969;193:Suppl.

19. Jarvelin MR, Sovio U, King V et al. Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. Hypertension 2004;44(6):838-846.

20. Sabatti C, Service SK, Hartikainen AL et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nat Genet 2009;41(1):35-46.

21. Vartiainen E, Jousilahti P, Alfthan G, Sundvall J, Pietinen P, Puska P. Cardiovascular risk factor changes in Finland, 1972-1997. Int J Epidemiol 2000;29(1):49-56.

22. Aulchenko YS, Ripatti S, Lindqvist I et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat Genet 2009;41(1):47-55.

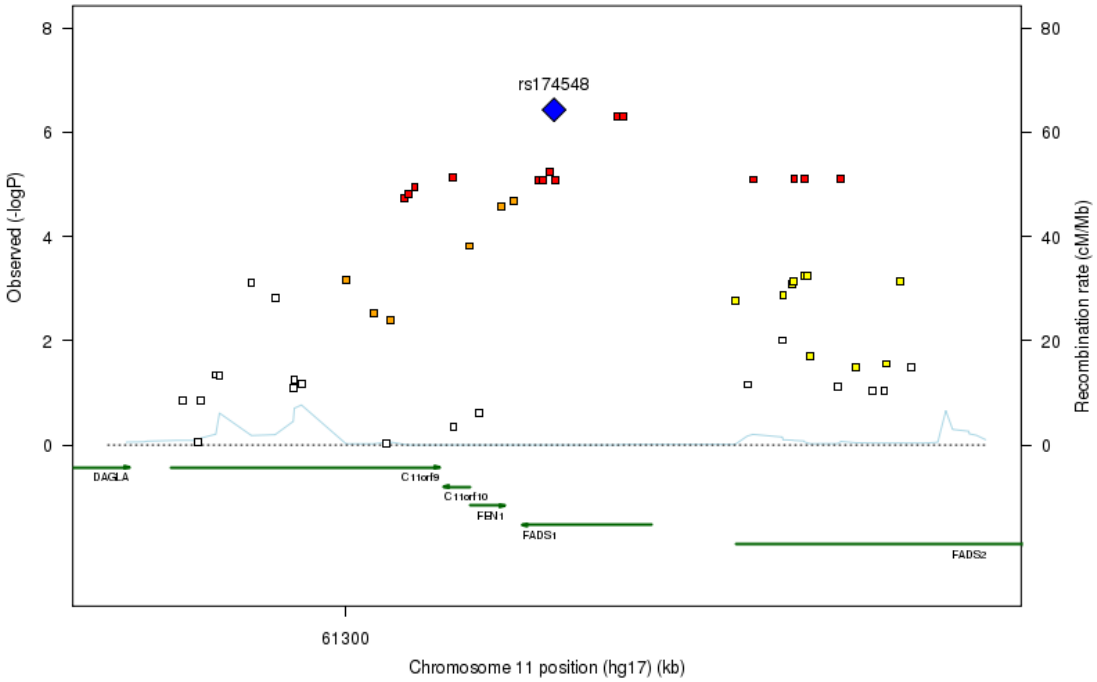
23. Hofman A, Breteler MM, van Duijn CM et al. The Rotterdam Study: objectives and design update. Eur J Epidemiol 2007;22(11):819-829.

24. Samani NJ, Erdmann J, Hall AS et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357(5):443-453.

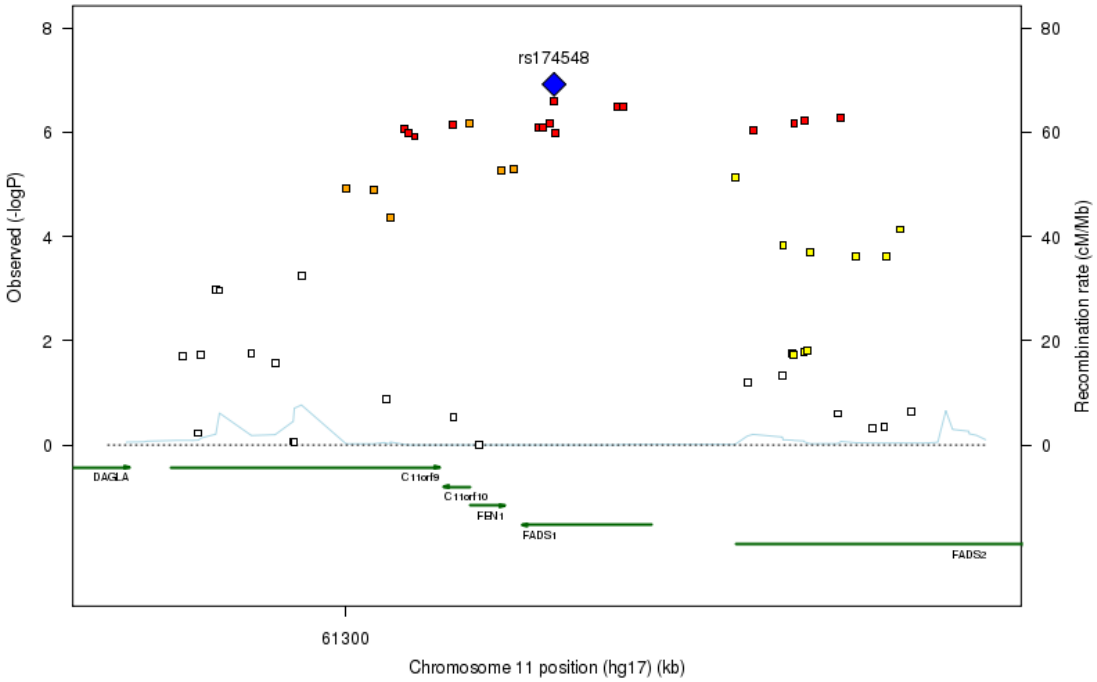
25. Helgadottir A, Thorleifsson G, Manolescu A et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316(5830):1491-1493.

26. Kathiresan S, Voight BF, Purcell S et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009;41(3):334-341.

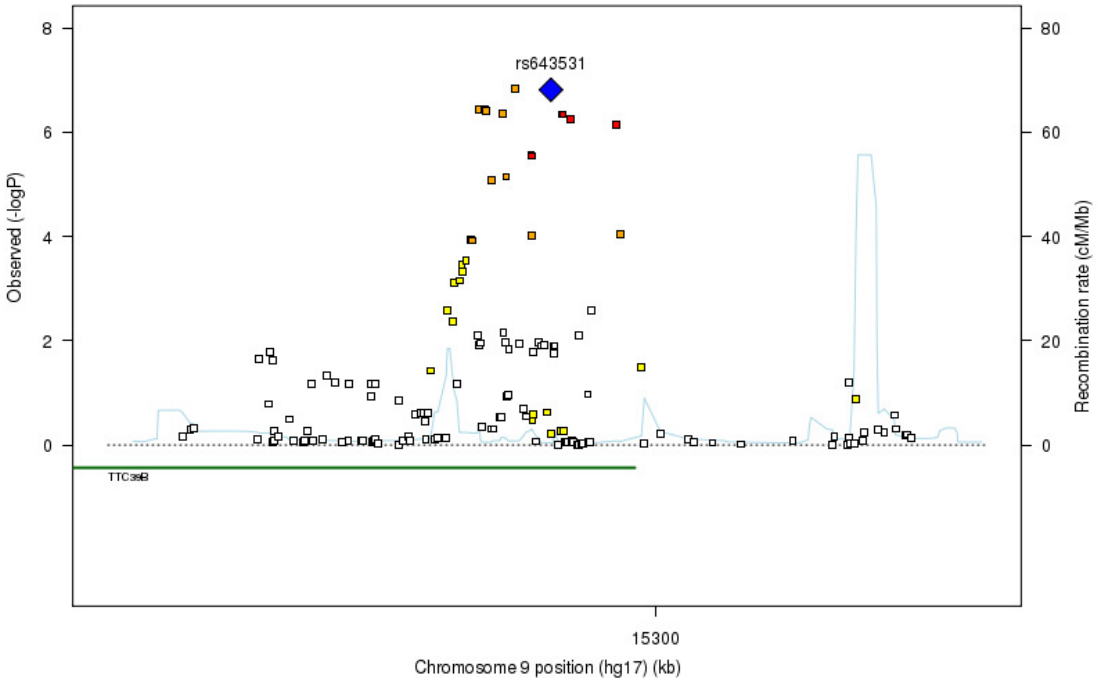
27. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24(24):2938-2939.



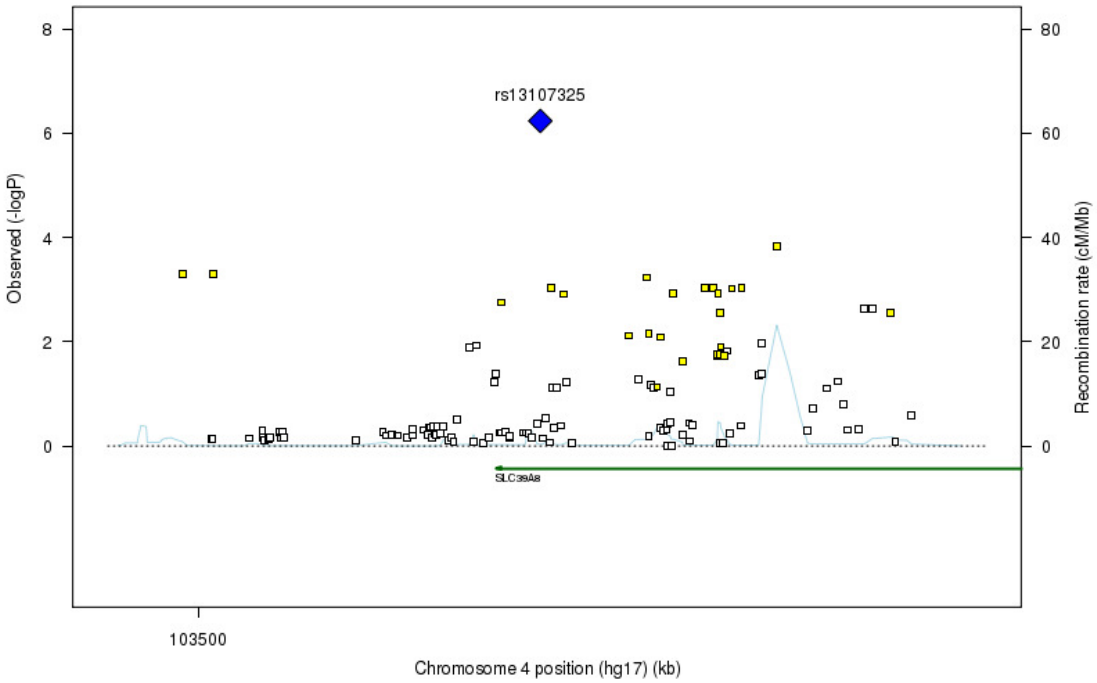
Supplementary Figure I (a) HDL-c and SNP rs174548 at *FADS1* (chr11 co-ordinates 61277924-61377923)



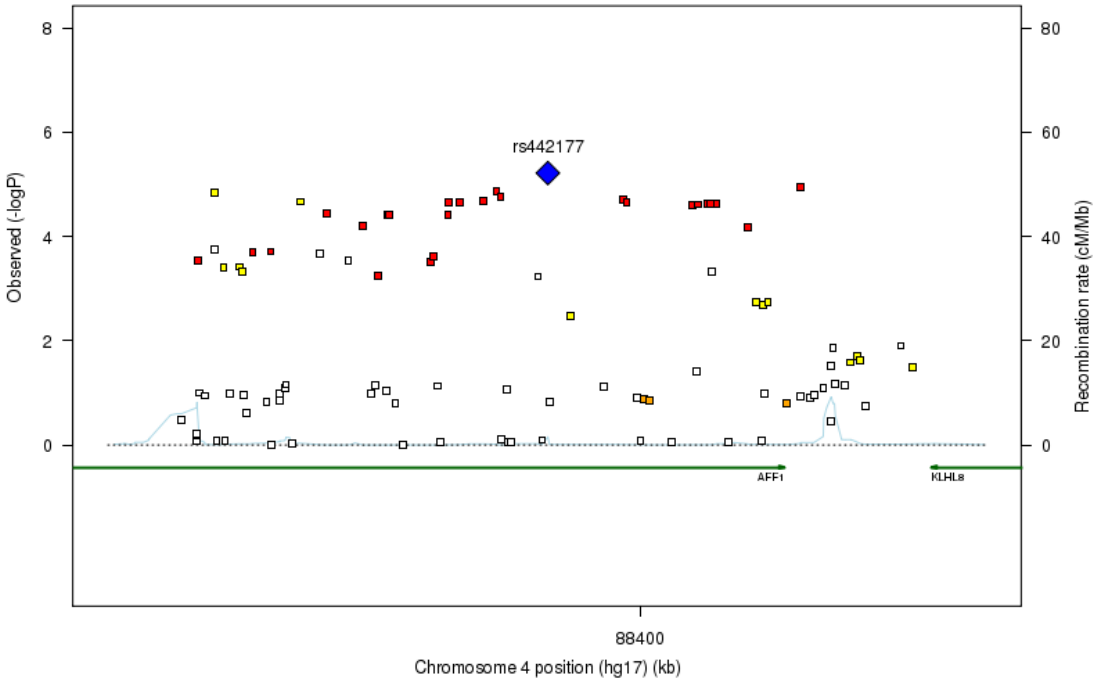
Supplementary Figure I (b) TG and rs174548 at *FADS1* (chr11 co-ordinates 61277924-61377923)



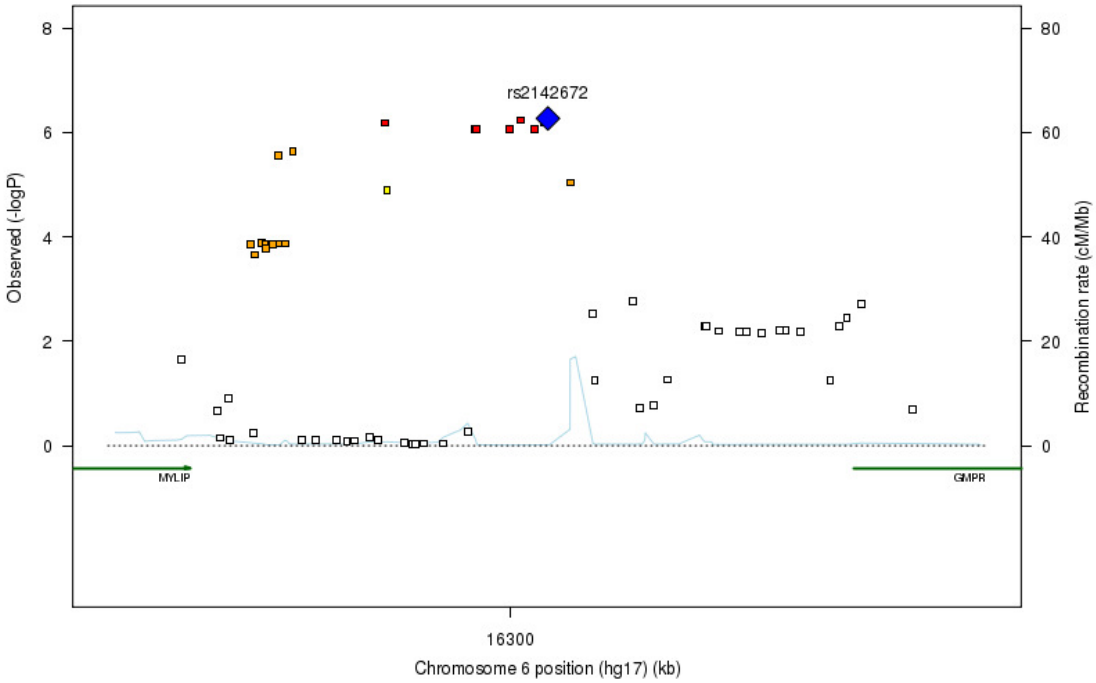
Supplementary Figure I (c) HDL-c and SNP rs643531 at *TTC39B* (chromosome 9 co-ordinates 15236034-15336033)



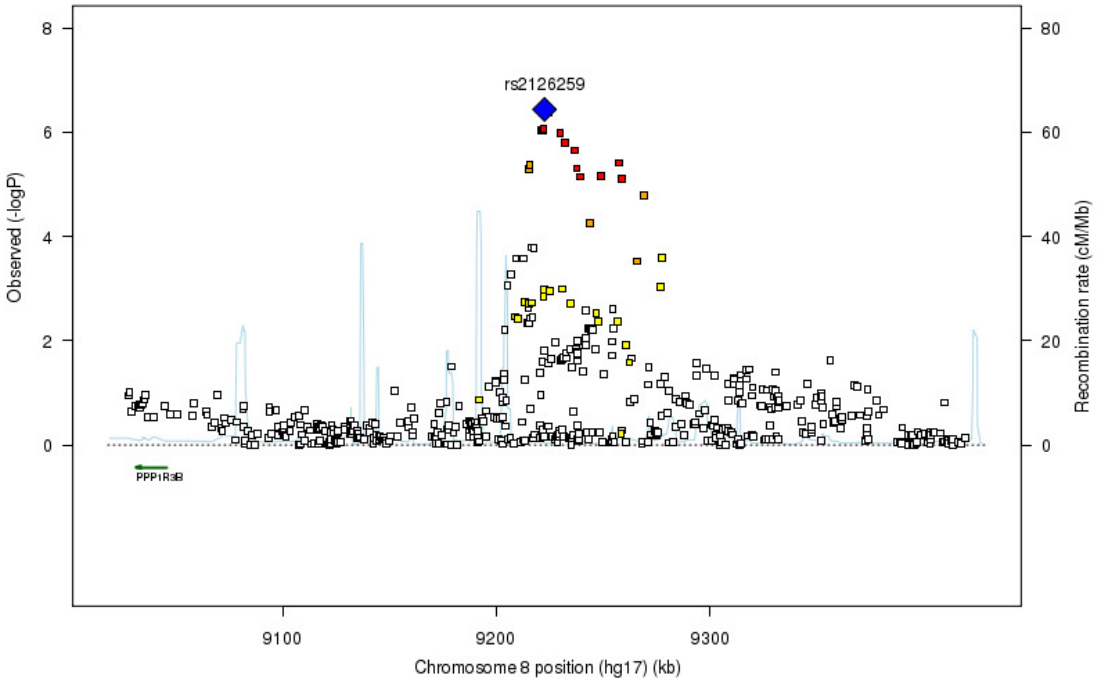
Supplementary Figure I (d) HDL-c and SNP rs13107325 at *SLC39A8* (chromosome 4 co-ordinates 103495887-103595886)



Supplementary Figure I (e) TG and SNP rs442177 at *AFF1* (chromosome 4 co-ordinates 88337440-88437439



Supplementary Figure I (f) LDL-c and SNP rs2142672 at *MYLIP/GMPR* (chromosome 6 co-ordinates 16255173-16355172)



Supplementary Figure I (g) LDL-c and SNP rs2126259 at *PPP1R3B* (chromosome 8 co-ordinates 9022556-9422555)

Supplementary Figure I Regional plots of meta-analytical Stage 1 association results at six novel loci that show genome-wide statistical association with one or more circulating lipid traits after Stage 2 and 3 replication. Plots of 100 kb genomic sequence (a-f) and 400 kb (g) spanning each SNP are aligned with association signals for SNPs. The SNP with the strongest association signal at Stage 1 is represented by a blue diamond. Each square represents a SNP and their colour represents the extent of linkage disequilibrium with the target SNP ($r^2 \geq 0.8$ = red, $r^2 \geq 0.5$ & < 0.8 = orange, $r^2 \geq 0.2$ & < 0.5 = yellow, $r^2 < 0.2$ = white). The recombination rates and hotspots are defined by blue lines and are compiled from genotyping data. Positions of genes and genomic co-ordinates are shown below the plots. Plots were generated using SNAP (http://www.broad.mit.edu/mpg/snap)[27] and based on HapMap release 21, NCBI B35 assembly, dbSNP build 126, [CEPH Utah trios].

Supplementary Table I. Study characteristics and details of analysis metrics and methods for individual studies in Stage 1

STUDY	EPIC-Norfolk subcohort	EPIC-Norfolk obese set	British 1958 birth cohort (WTCCC controls)	CoLaus	GEMS	LOLIPOP (white European sample 1)	SardiNIA	FUSION controls
LIPID MEASUREMENTS								
Sample	Non-fasting fresh blood	Non-fasting fresh blood	Non-fasting fresh blood	Fasting (12 hr) fresh blood	Fasting (12 hr) fresh blood	Fasting (8 hr) fresh blood	Fasting (overnight) fresh blood	Fasting fresh blood
Assay/Analsyer	Serum total cholesterol, HDL-c and triglycerides measured using RA-1000 analyser (Bayer Diagnostics, Basingstoke, UK). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using RA-1000 analyser (Bayer Diagnostics, Basingstoke, UK). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Olympus model AU640 autoanalyser (Olympus Inc, Center Valley, PA, USA). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Modular P apparatus (Roche Diagnostics, Basel, Switzerland). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Hitachi 704 analyser (Hitachi, Tokyo, Japan). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Olympus model AU640 autoanalyser (Olympus Inc, Center Valley, PA, USA). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides determined using standard enzymatic methods. LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides determined using standard enzymatic methods. LDL-c levels calculated using Friedewald formula[1]
SAMPLES								
Description	Subcohort randomly selected from EPIC-Norfolk cohort , UK using a random selection algorithm	Case series of obese individuals (BMI ≥ 30kg/m^2) derived from EPIC-Norfolk cohort, UK	UK National population sample	Cohort study of randomly selected individuals from Lausanne, Switzerland	Case-control study of dyslipidaemia	Population-based cohort of individuals of white European and Indian Asian descent recruited from West London, UK	Longitudinal study of aging-related quantitative traits in the Ogliastra region of Sardinia, Italy	Non-diabetic control participants from diabetes case-control study, Finland
Samples with phenotype: N all (%males/%females)	2,269 (46/54)	1,009 (42/58)	up to 1,458 (49/51)	up to 5,226 (47/53)	1,665 (59/41)	up to 813 (77/23)	4,184 (43/57)	up to 1,099 (52/48)
Age (mean (sd), years)	59.2 (9.0)	59.2 (8.8)	44.9 (0.4)	53.4 (10.8)	52.4 (9.5)	53.3 (10.4)	43.1 (17.5)	60.1 (11.5)
BMI (mean (sd), kg/m2)	26.3 (3.8)	32.9 (2.8)	27.2 (4.9)	25.9 (4.6)	28.5 (3.6)	27.8 (4.8)	25.3 (4.7)	27.0 (3.9)
HDL-c (median (iqr), mmol/L)	1.4 (1.1-1.6)	1.2 (1.0-1.5)	1.5 (1.3-1.8)	1.6 (1.3-1.9)	NA *	1.3 (1.1-1.5)	1.6 (1.4-1.9)	1.4 (1.2-1.7)
LDL-c (median (iqr), mmol/L)	3.9 (3.3-4.6)	4.0 (3.4-4.7)	3.3 (2.8-3.9)	3.4 (2.8-4.1)	3.5 (2.9-4.2)	3.6 (3.0-4.2)	3.2 (2.6-3.8)	3.7 (3.1-4.3)
Triglycerides (median (iqr), mmol/L)	1.5 (1.1-2.1)	2.0 (1.4-2.7)	1.7 (1.1-2.6)	1.1 (0.8-1.6)	NA *	1.3 (0.9-1.9)	0.8 (0.5-1.1)	1.2 (0.9-1.6)

* For the GEMS study, case and control participants were determined by triglyceride and HDL-c levels (see references[7,8] for details). These traits therefore have truncated distributions and are not reported.

STUDY	EPIC-Norfolk subcohort	EPIC-Norfolk obese set	British 1958 birth cohort (WTCCC controls)	CoLaus	GEMS	LOLIPOP (white European sample 1)	SardiNIA	FUSION controls
GENOTYPING								
Genotyping platform & SNP panel	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Illumina Human Hap300
Genotyping centre	Affymetrix	Affymetrix	Affymetrix	Affymetrix Expression Analysis	Affymetrix	Affymetrix	SardiNIA Research Laboratory, Lanusei, Italy	Center for Inherited Disease Research
Genotyping calling algorithm	BRLMM	BRLMM	CHIAMO	BRLMM	BRLMM	BRLMM	BRLMM	Beadstudio
SAMPLE QC								
Call rate	≥ 94%	≥ 94%	≥ 97%	≥ 90%	≥ 90%	≥ 95%	> 95%	> 97.5%
Heterozygosity	<23% or >30%	<23% or >30%	<23% or >30%	NA	NA	NA	NA	NA
Other exclusions	1) Ethnic outliers 2) >5.0% discordance in SNP pairs with r^2= 1 in HapMap 3) Related individuals (>70% concordance with another DNA) selected based on sample call rate) 4) Duplicate (>99% concordance with another DNA)	1) Ethnic outliers 2) >5.0% discordance in SNP pairs with r^2= 1 in HapMap 3) Related individuals (>70% concordance with another DNA) selected based on sample call rate) 4) Duplicate (>99% concordance with another DNA)	1) Ethnic outliers 2) External discordance with genotype or phenotype data 3) Related individuals (>86% concordance with another DNA) 4) Duplicate (>99% concordance with another DNA	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Verified sex using X-linked markers 2) Verified relationships and checked for duplicates using RELPAIR	1) Likely first or second degree relatives based on genotypes, using RELPAIR
SNP QC filters (prior to imputation)					.			
MAF	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 5%	≥ 1%
HWE	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-4}$	$P > 10^{-7}$	$P > 10^{-7}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$
Call rate	≥ 90%	≥ 90%	≥ 95%	≥ 90%	≥ 90%	≥ 90%	> 90%	> 90%
Other							Mendel errors > 2	Mendel + duplicate inconsistencies > 3
SNP number in dataset post-QC	382,036	382,036	419,829	384,079	407,071	400,602	356,359	304,463
IMPUTATION STATS								
Imputation software	IMPUTE	IMPUTE	IMPUTE	IMPUTE	IMPUTE	IMPUTE	MACH	MACH
Imputation quality metrics	proper_info ≥ 0.40	proper_info ≥ 0.40	proper_info ≥ 0.40	proper_info ≥ 0.40	proper_info ≥ 0.40	proper_info ≥ 0.40	r2hat ≥ 0.3	r2hat ≥ 0.3
Other SNP QC filters applied?	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	NMI inconsistencies ≥ 5 MAF ≥ 1%	MAF ≥ 1%

STUDY	EPIC-Norfolk subcohort	EPIC-Norfolk obese set	British 1958 birth cohort (WTCCC controls)	CoLaus	GEMS	LOLIPOP (white European sample 1)	SardiNIA	FUSION controls
DATA ANALYSIS								
Number of SNPs in analysis	up to 2,378,402	up to 2,377,144	up to 2,411,229	up to 2,341, 936	2,401,159	up to 2,393,773	2,248,915	2,414,220
Trait transformation	natural log	natural log	natural log	natural log	natural log	natural log	natural log	natural log
Adjustments	age, sex	age, sex	sex	age, sex, 10 geographical principle components, adjustment for participants taking lipid-lowering medication *	age, sex	age, sex, top five principle components	age, sex	age, sex
Analysis method	additive test	additive test	additive test	additive test	additive test	additive test	additive test	additive test
Software for analysis	SNPTEST	SNPTEST	SNPTEST	SNPTEST	SNPTEST	SNPTEST	MERLIN	MERLIN
Genomic Control Lambda (HDL, Trigs, LDL)	1.02, 1.01, 1.02	1.02, 1.00, 1.01	1.01, 1.00, 1.01	1.03, 1.04, 1.01	NA, NA, 1.03	1.02, 1.03, 1.03	1.14, 1.09, 1.07	1.00, 1.01, 1.03
					.			
REFERENCES								
Reference cohort	2	2	4	6	7, 8	9	10	13
Reference genotyping	3	3	5	3	3	9	11, 12	13

* For CoLaus LDL-c levels were corrected for participants taking lipid-lowering medication by multiplying by a constant of 1.375. This was derived from an estimated average reduction of 37.5% obtained for circulating LDL-c levels, based on the most commonly prescribed medication[14].

Supplementary Table II. Study characteristics and details of analysis metrics and methods for individual studies in replication analysis Stages 2 and 3

	Stage 2								Stage 3			
STUDY	**EPIC-Norfolk cohort**	**Ottawa Heart Study controls**	**Fenland**	**British 1958 birth cohort (T1DGC controls)**	**Northern Finland birth cohort 1966**	**National FINRISK Study**	**Rotterdam Study**	**LOLIPOP (white European sample 2)**	**LOLIPOP (Indian Asian samples) ***			
LIPID MEASUREMENTS												
Sample	Non-fasting fresh blood	Fasting (11h) fresh blood	Fasting (overnight) fresh blood	Non-fasting fresh blood	Fasting (overnight) fresh blood	Non-fasting fresh blood	Non-fasting (HDL-c & TC) and fasting (overnight) (TG) fresh blood	Fasting (8 hr) fresh blood	Fasting (8 hr) fresh blood			
Assay/Analsyer	Serum total cholesterol, HDL-c and triglycerides measured using RA-1000 analyser (Bayer Diagnostics, Basingstoke, UK). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using LX-20 Clinical Chemistry Analyzer (Beckman Coulter). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Dimension® RxL System (Siemens, Surrey, UK). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Olympus model AU640 autoanalyser (Olympus Inc, Center Valley, PA, USA). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Hitachi 911 Clinical Chemistry Analyzer (Boehringer Mannheim). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Hitachi 911 Clinical Chemistry Analyzer (Boehringer Mannheim). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using standard protocol. LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Olympus model AU640 autoanalyser (Olympus Inc, Center Valley, PA, USA). LDL-c levels calculated using Friedewald formula[1]	Serum total cholesterol, HDL-c and triglycerides measured using Olympus model AU640 autoanalyser (Olympus Inc, Center Valley, PA, USA). LDL-c levels calculated using Friedewald formula[1]			
SAMPLES												
Description	Population-based cohort of white European men and women recruited from Norfolk, UK	Control participants	Population based cohort of men and women recruited from Cambridgeshire, UK	UK National population sample	Population-based birth cohort from Northern Finland	Population-based cohort from Finland	Population-based cohort of men and women recruited from Rotterdam, The Netherlands	Population-based cohort of individuals of white European and South Asian descent recruited from West London, UK	Population-based cohort of individuals of white European and Indian Asian descent recruited from West London, UK			
Samples with phenotype: N all (%males/%females)	up to 19,793 (45/55)	up to 1,445 (52/48)	up to 1,402 (44/56)	up to 2,527 (48/52)	up to 5,138 (48/52)	up to 910 (65/35)	up to 5,849 (41/59)	up to 710 (100/0)	[a] up to 2,094 (100/0)	[b] up to 569 (100/0)	[c] up to 3,025 (82/18)	[d] up to 3,977 (86/14)
Age (mean (sd), years)	58.6 (9.4)	75.0 (5.0)	45.0 (7.3)	45.3 (0.3)	31 (0.0)	59.6 (10.7)	69.4 (9.1)	56.0 (8.9)	48.1 (10.4)	50.8 (8.2)	59.4 (9.6)	52.5 (10.2)
BMI (mean (sd), kg/m2)	25.9 (3.6)	26.0 (4.0)	27.1(4.9)	27.5 (5.0)	24.7 (4.2)	NA	26.3 (3.7)	28.6 (5.3)	26.8 (4.3)	27.1 (4.0)	27.6 (4.5)	26.8 (4.2)
HDL-c (median (iqr), mmol/L)	1.4 (1.1-1.7)	1.5 (0.4)†	1.4 (1.2-1.7)	1.5 (1.3-1.8)	1.5 (1.3-1.8)	1.4 (1.1-1.6)	1.3 (0.4)†	1.2 (1.0-1.4)	1.2 (1.0-1.4)	1.2 (1.0-1.4)	1.1 (1.0-1.3)	1.2 (1.0-1.4)
LDL-c (median (iqr), mmol/L)	3.9 (3.2-4.6)	3.6 (0.9)†	3.3 (2.8-3.9)	3.3 (2.8-3.9)	2.9 (2.4-3.5)	3.5 (2.9-4.0)	3.7 (0.9)†	3.0 (2.4-3.7)	3.3 (2.8-3.9)	3.4 (2.9-4.0)	2.6 (2.0-3.3)	3.3 (2.7-3.9)
Triglycerides (median (iqr), mmol/L)	1.5 (1.1-2.1)	1.4 (0.9)†	1.0 (0.7-1.5)	1.7 (1.1-2.6)	1.0 (0.7-1.4)	1.3 (1.0-1.9)	1.5 (0.8)†	1.4 (1.0-2.1)	1.5 (1.1-2.1)	1.6 (1.1-2.4)	1.5 (1.1-2.1)	1.5 (1.1-2.2)

	Stage 2								Stage 3			
STUDY	**EPIC-Norfolk cohort**	**Ottawa Heart Study controls**	**Fenland**	**British 1958 birth cohort (T1DGC controls)**	**Northern Finland birth cohort 1966**	**National FINRISK Study**	**Rotterdam Study**	**LOLIPOP (white European sample 2)**	**LOLIPOP (Indian Asian samples)**			
GENOTYPING												
Genotyping platform & SNP panel	Sequenom iPLEX, Taqman	Affymetrix 500K & 6.0 Array set	Affymetrix 500K Array Set	Illumina Human Hap 550	Illumina Infinium 370cnvDuo array	Illumina Human Hap 610K	Illumina Human Hap 550	Custom Perlegen chip	Illumina Human Hap 300	Custom Perlegen chip	Wellcome data	Wellcome data
Genotyping centre	MRC Epidemiology Unit	Canadian Cardiovascular Genetic Centre	Affymetrix	Illumina	Broad	Sanger	Rotterdam	Perlegen	Decode	Perlegen	Decode	Decode
Genotyping calling algorithm	NA	BRLMM & BIRDSEED	BRLMM	Illuminus	Beadstudio	Illuminus	Beadstudio	Perlegen algorithm	Beadstudio	Perlegen algorithm	Beadstudio	Beadstudio
SAMPLE QC												
Call rate	NA	> 98%	≥ 95%	> 98%	≥ 95%	≥ 95%	> 97.5%	> 98%	>98%	>98%	>98%	>98%
Heterozygosity	NA	NA	Heterozygosity upper-bound 0.28822 lowerbound 0.27348	NA	NA	NA	FDR ≥ 1%	NA	NA	NA	NA	NA
Other exclusions	NA	Ethnic outliers	Failed relatedness and duplicate check	NA	Failed relatedness, duplicate check, ethnic outliers	Failed relatedness, duplicate check, ethnic outliers	1) Ethnic outliers 2) Related individuals 3) Duplicate (>95% concordance with another DNA)		Duplicates, Missing rate > 5%		Duplicates	Duplicates
SNP QC filters (prior to imputation)												
MAF	NA	NA	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%
HWE	NA	NA	$P > 10^{-6}$	$P > 5.7 \times 10^{-7}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$	$P > 10^{-6}$
Call rate	NA	NA	≥ 90%	≥ 95%	≥ 98%	≥ 99%	≥ 98%	≥ 95%	≥95%	≥95%	≥95%	≥95%
Other												
IMPUTATION STATS												
Imputation software	NA	IMPUTE	IMPUTE	MACH	MACH	MACH	MACH	MACH	MACH			
Imputation quality metrics	NA	proper_info ≥ 0.40	proper_info ≥ 0.40	r2hat ≥ 0.3	r2hat ≥ 0.3	r2hat ≥ 0.3	r2hat ≥ 0.3	r2hat ≥ 0.3	r2hat ≥ 0.3			
Other SNP QC filters applied?	HWE P > 0.05 MAF ≥ 1%	MAF ≥ 1% call rate ≥ 80%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%			

	Stage 2								Stage 3
STUDY	**EPIC-Norfolk cohort**	**Ottawa Heart Study controls**	**Fenland**	**British 1958 birth cohort (T1DGC controls)**	**Northern Finland birth cohort 1966**	**National FINRISK Study**	**Rotterdam Study**	**LOLIPOP (white European sample 2)**	**LOLIPOP (Indian Asian samples)**
DATA ANALYSIS									
Trait transformation	natural log	natural log	natural log	natural log	natural log	natural log	natural log	natural log	natural log
Adjustments	age, sex	age, sex	age, sex	age, sex	age, sex	age, sex	age, sex	age, sex, top five population components	age, sex, top five population components
Analysis method	linear regression	additive test	additive test	additive test	linear regression	linear regression	additive test	additive test	additive test
Software for analysis	STATA 10.1	SNPTEST	SNPTEST	ProbABEL	PLINK	ProbABEL	ProbABEL	mach2qtl	mach2qtl
REFERENCES									
Reference study	2	15	16	4	18, 19	21	23	9	9
Reference genotyping	NA	15	16	17	20	22	22	9	9

* The LOLIPOP samples of Indian Asian ethnicity comprised four subsets genotyped with different genome-wide chips (a, b, c and d). The subsets c and d are case and control participants from a case-control study of CHD nested within the LOLIPOP cohort.

† Values are mean (standard deviation).

Supplementary Table III. Study characteristics and details of analysis metrics and methods for individual studies in CAD risk meta-analysis

STUDY	EPIC-Norfolk-1	EPIC-Norfolk-2	Ottawa Heart Study	WTCCC CAD study	PENN CATH	MEDSTAR	CoLaus	GEMS	Rotterdam
SAMPLES									
Description	IHD case-control set of participants with genome-wide data available nested within EPIC-Norfolk cohort, UK	IHD case-cohort set nested within EPIC-Norfolk cohort, UK	CAD case-control study of participants recruited from the Champlain region including Ottawa	CAD case-control study of participants recruited nationally from the UK	CAD case-control study nested within cohort of participants attending University of Pennsylvania Medical Center, PA, USA, for angiography	CAD case-control study nested within cross sectional cohort of participants attending Washington Hospital Center, Washington DC, USA, for angiography	CAD case-cohort set nested within CoLaus study	CAD case-cohort set nested within GEMS study	CAD case-cohort set nested within Rotterdam study
N cases / N controls	421 / 2698	2,625 / 18,634	1,542 / 1,455	1926 / 2938	933 / 468	943 / 483	202 / 5233	262 / 1580	779 / 5,195
Case/control definition	Cases: individuals who have prevalent or incident IHD as either self-reported at baseline healthcheck or death certificate/hospital discharge notes. For the latter, IHD is defined as ICD9 410-414 or ICD10 I20-I25. Controls: individuals without prevalent or incident IHD. All participants with prevalent or incident stroke were excluded.	Cases: individuals who have prevalent or incident IHD as either self-reported at baseline healthcheck or death certificate/hospital discharge notes. For the latter, IHD is defined as ICD9 410-414 or ICD10 I20-I25. Controls: individuals without prevalent or incident IHD. All participants with prevalent or incident stroke were excluded.	Cases: individuals exhibited symptomatic CAD before the age of 55 years in males and 65 in females. CAD defined as greater than 50% stenosis of a coronary artery was confirmed by coronary angiography. Cases with a history of diabetes were excluded. Controls: asymptomatic elderly individuals were selected as controls(≥65 for males, ≥70 for females).	Cases: individuals with premature CAD (a validated history of either MI or coronary revascularisation before age 66 years) and one or more first degree relatives with CAD. Controls: unrelated individuals from 1958 British birth cohort and UK Blood Services.	Cases: individuals who had one or more coronary vessels with ≥ 50% stenosis equally selected among stable CAD cases without history of MI and CAD cases with a history of MI. Controls: men aged > 40 and women aged > 45 showing no evidence of CAD.	Cases: individuals who had one or more coronary vessels with ≥ 50% stenosis equally selected among stable CAD cases without history of MI and CAD cases with a history of MI. Controls: individuals aged > 45 with no evidence of CAD.	Cases: individuals who reported previous acute MI, cardiac catheterisation, coronary artery bypass or angina. Controls: individuals without any of above case definitions.	Cases: individuals who reported previous CAD. Controls: individuals who reported no current or previous CAD.	Cases: individuals with incident CAD. Controls: individuals without previous or current CAD at follow-up.
Sex (%males / %females)									
Cases:	67 / 33	65 / 35	76/24	79 / 21	76 / 24	72 / 28	72 / 28	74 / 26	51 / 49
Controls:	41 / 59	42 / 58	52/48	49 / 51	48 / 52	51 / 49	46 / 54	57 / 43	36 / 63
Age (mean (sd), years)									
Cases:	64.1 (7.4)	64.4 (8.0)	48.7 (7.3)	45.7 (9.7)	56.8 (9.2)	54.6 (7.3)	62.6 (8.4)	55.0 (7.5)	72.5 (9.5)
Controls:	58.0 (8.8)	57.4 (9.1)	75.0 (5.0)	49.8 (7.7)	61.7 (9.6)	60.0 (8.9)	62.5 (8.4)	52.0 (9.7)	68.7 (9.0)
BMI (mean (sd), kg/m^2)									
Cases:	29.3 (4.3)	27.0 (3.8)	28.6 (4.9)	27.6 (4.2)	29.8 (5.6)	31.6 (6.8)	28.3 (4.4)	28.5 (3.4)	26.3 (3.6)
Controls:	28.2 (4.7)	25.9 (3.7)	26.0 (4.0)	-	29.0 (6.4)	31.4 (7.7)	26.8 (4.5)	28.5 (3.7)	26.3 (3.7)

STUDY	EPIC-Norfolk-1	EPIC-Norfolk-2	Ottawa Heart Study	WTCCC CAD study	PENN CATH	MEDSTAR	CoLaus	GEMS	Rotterdam
GENOTYPING									
Genotyping platform & SNP panel	Affymetrix 500K Array Set	Sequenom iPLEX	Affymetrix 500K & 6.0 Array set	Affymetrix 500K Array Set	Affymetrix Array 6.0	Affymetrix Array 6.0	Affymetrix 500K Array Set	Affymetrix 500K Array Set	Illumina Infinium-II HumanHap550 SNP array (v3)
Genotyping centre	Affymetrix	MRC Epidemiology Unit	Canadian Cardiovascular Genetic Centre	Affymetrix	Broad Institute	Broad Institute	Affymetrix Expression Analysis	Affymetrix	Rotterdam
Genotyping calling algorithm	BRLMM	NA	BRLMM & BIRDSEED	CHIAMO	BIRDSEED	BIRDSEED	BRLMM	BRLMM	Beadstudio
SAMPLE QC									
Call rate	≥ 94%	NA	> 98%	≥ 97%	≥ 90%	≥ 90%	≥ 90%	≥ 90%	≥ 97.5
Heterozygosity	<23% or >30%	NA	NA	<23% or >30%	NA	NA	NA	NA	FDR ≥ 1%
Other exclusions	1) Ethnic outliers 2) >5.0% discordance in SNP pairs with r^2= 1 in HapMap 3) Related individuals (>70% concordance with another DNA) selected based on sample call rate) 4) Duplicate (>99% concordance with another DNA)	NA	Ethnic outliers	1) Ethnic outliers 2) External discordance with genotype or phenotype data 3) Related individuals (>86% concordance with another DNA) 4) Duplicate (>99% concordance with another DNA)	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Sex inconsistency with genetic data from X-linked SNPs 2) Inconsistent genotypes when compared with duplicate samples	1) Ethnic outliers 2) Related individuals 3) Duplicate (>95% concordance with another DNA)
SNP QC filters (prior to imputation)									
MAF	≥ 1%	NA	NA	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥1%
HWE	$P > 10^{-6}$	NA	NA	$P > 10^{-4}$	$P > 10^{-4}$	$P > 10^{-4}$	$P > 10^{-7}$	$P > 10^{-7}$	$P \geq 10^{-6}$
Call rate	≥ 90%	NA	NA	≥ 98%	≥ 90%	≥ 90%	≥ 90%	≥ 90%	≥ 98%
Other									
IMPUTATION STATS									
Imputation software	IMPUTE	NA	IMPUTE	IMPUTE	MACH	MACH	IMPUTE	IMPUTE	MACH
Imputation quality metrics	proper_info ≥ 0.40	NA	proper_info ≥ 0.40	proper_info ≥ 0.40	r2hat ≥ 0.3	r2hat ≥ 0.3	proper_info ≥ 0.40	proper_info ≥ 0.40	r2hat ≥ 0.3
Other SNP QC filters applied?	MAF ≥ 1%	HWE P > 0.05 MAF ≥ 1%	MAF ≥ 1% call rate ≥ 80%	MAF ≥ 1% Call rate > 80%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%	MAF ≥ 1%

STUDY	EPIC-Norfolk-1	EPIC-Norfolk-2	Ottawa Heart Study	WTCCC CAD study	PENN CATH	MEDSTAR	CoLaus	GEMS	Rotterdam
DATA ANALYSIS									
Adjustments	age, sex, BMI	age, sex	sex	age, sex	age, sex	age, sex	age, sex	age, sex	age, sex
Analysis method	additive test	logistic regression	additive test	additive test	logistic regression	logistic regression	logistic regression	logistic regression	logistic regression
Software for analysis	SNPTEST	STATA 10.1	SNPTEST	PLINK	PLINK	PLINK	PLINK	PLINK	ProbABEL
REFERENCES									
Reference study	2	2	15	24	25	26	6	7, 8	23
Reference genotyping	3	NA	15	5	26	26	3	3	22

Supplementary Table IV. Statistical associations between SNPs showing evidence of association with circulating lipid levels in Stage 1 ($P < 1 \times 10^{-5}$) with no heterogeneity among studies ($P > 0.1$) at novel loci

						Stage 1		
SNP	**Chr**	**Pos (Mb) ***	**Nearest locus (loci)**	**Effect allele †**	**Effect allele freq †**	**β-coefficient (se) ‡**	**P-value**	**P-value for heterogeneity**
LDL-c								
rs4407201	2	130.2	*RAB6C*	T	0.10	0.028 (0.006)	8.9×10^{-6}	0.69
rs264715	3	65.9	*MAGI1*	G	0.03	0.050 (0.011)	2.5×10^{-6}	0.75
rs16855082	4	42.7	*GRXCR1*	T	0.04	0.040 (0.008)	7.8×10^{-7}	0.95
rs1501908	5	156.3	*TIMD4*	C	0.35	-0.017 (0.003)	1.1×10^{-7}	0.17
rs2142672	6	16.3	*MYLIP, GMPR*	C	0.74	0.018 (0.004)	5.4×10^{-7}	0.62
rs9275427	6	32.8	*HLA-DQA2*	T	0.16	0.020 (0.004)	6.7×10^{-6}	0.37
rs456598	6	160.5	*IGF2R, SLC22A1*	G	0.87	-0.023 (0.005)	2.9×10^{-6}	0.61
rs2126259	8	9.2	*PPP1R3B*	A	0.10	-0.028 (0.006)	3.6×10^{-7}	0.48
rs4135183	9	110.1	*TXN*	A	0.13	-0.023 (0.005)	5.5×10^{-6}	0.28
rs17136824	10	12.5	*CAMK1D*	G	0.80	-0.019 (0.004)	4.4×10^{-6}	0.41
rs17697566	10	49.1	*FRMPD2*	T	0.82	0.023 (0.005)	3.8×10^{-6}	0.44
rs11227299	11	65.3	*DKFZp761E198*	G	0.35	-0.016 (0.003)	5.7×10^{-6}	0.15
rs4307732	11	125.8	*ST3GAL4*	G	0.89	-0.025 (0.006)	7.9×10^{-6}	0.98
HDL-c								
rs10749832	1	41.7	*HIVEP3*	G	0.40	0.013 (0.003)	7.1×10^{-6}	0.56
rs3923229	2	85.5	*TGOLN2, RETSAT*	A	0.74	0.016 (0.003)	3.6×10^{-6}	0.82
rs7560571	2	95.1	*MAL*	G	0.35	-0.014 (0.003)	4.9×10^{-6}	0.39
rs13088837	3	63.4	*SYNPR*	G	0.24	0.016 (0.003)	6.7×10^{-6}	0.82
rs13107325	4	103.5	*SLC39A8*	T	0.08	-0.030 (0.006)	5.8×10^{-7}	0.72
rs643531	9	15.3	*TTC39B*	C	0.14	-0.021 (0.004)	1.5×10^{-7}	0.49
rs174548	11	61.3	*FADS1*	G	0.30	-0.016 (0.003)	3.7×10^{-7}	0.67
rs4943343	13	35.3	*DCLK1*	T	0.08	-0.028 (0.006)	5.9×10^{-6}	0.99

rs11642898	16	19.6	*C16orf62*	T	0.14	0.022 (0.005)	1.2×10^{-6}	0.49
rs756942	17	65.1	*MAP2K6*	T	0.24	0.015 (0.003)	4.6×10^{-6}	0.21
rs12980205	19	48.7	*PHLDB3*	T	0.23	-0.020 (0.004)	2.5×10^{-6}	0.48
rs1981383	21	42.2	*C2CD2*	G	0.48	-0.015 (0.003)	7.4×10^{-6}	0.13
TG								
rs1251474	1	76.2	*LOC729766*	C	0.64	-0.029 (0.006)	4.3×10^{-6}	0.50
rs2791544	1	216.1	*LYPLAL1*	C	0.62	0.029 (0.006)	1.8×10^{-6}	0.64
rs10497518	2	179.3	*TTN*	G	0.03	0.066 (0.015)	6.9×10^{-6}	0.78
rs1378168	3	16.0	*ANKRD28*	G	0.28	-0.033 (0.007)	9.4×10^{-7}	0.70
rs4390955	3	18.7	*SATB1*	G	0.27	-0.029 (0.007)	6.7×10^{-6}	0.53
rs9309766	3	77.4	*ROBO2*	T	0.59	-0.027 (0.006)	8.0×10^{-6}	0.81
rs442177	4	88.4	*AFF1*	A	0.60	0.028 (0.006)	6.0×10^{-6}	0.40
rs17378274	4	162.4	*FSTL5*	C	0.43	-0.029 (0.006)	2.9×10^{-6}	0.14
rs6867983	5	55.9	*C5orf35*	T	0.14	0.039 (0.008)	3.9×10^{-6}	0.50
rs636202	6	139.9	*CITED2*	T	0.49	0.028 (0.006)	6.4×10^{-6}	0.74
rs10111012	8	110.2	*TRHR*	T	0.89	-0.049 (0.011)	8.1×10^{-6}	0.37
rs17663474	8	117.7	*EIF3H*	C	0.91	0.049 (0.011)	6.9×10^{-6}	0.73
rs2499902	10	11.5	*CUGBP2*	T	0.23	-0.036 (0.008)	9.5×10^{-6}	0.84
rs174548	11	61.3	*FADS1*	G	0.30	0.035 (0.007)	1.2×10^{-7}	0.69
rs4906879	15	24.2	*GABRB3*	G	0.56	-0.041 (0.008)	5.9×10^{-7}	0.34
rs12913924	15	55.6	*CGNL1*	G	0.82	-0.040 (0.009)	2.7×10^{-6}	0.93

The genome-wide association meta-analyses (Stage 1) for HDL-c and TG are based on data from seven study populations comprising up to 16,056 and 16,058 participants, respectively. For LDL-c, the meta-analysis is based on data from eight study populations comprising up to 17,543 participants.

* Based on NCBI Build 35.

† Based on Study 1 (EPIC-Norfolk sub-cohort). Effect allele corresponds to forward strand of NCBI Build 36.3.

‡ Beta-coefficients represent the change in circulating lipid level (natural log) per additional effect allele, adjusted for age and sex.

Supplementary Table V. Replication of statistical associations between SNPs at novel loci with circulating lipid levels

						Stage 2				Combined			
SNP	Chr	Pos (Mb) *	Nearest locus (loci)	Effect allele †	Effect allele freq †	No. participants	β - coefficient (se) ‡	P-value	P-value for heterogeneity	No. participants	β- coefficient (se) ‡	P-value	P-value for heterogeneity
LDL-c													
rs4907201	2	130.2	*RAB6C*	T	0.10	25,457	0.002 (0.004)	0.57	0.26	43,000	0.011 (0.004)	3.1×10^{-3}	0.05
rs264715	3	65.9	*MAGI1*	G	0.03	12,101	-0.001 (0.010)	0.93	0.47	29,644	0.024 (0.007)	1.2×10^{-3}	0.08
rs16855082	4	42.7	*GRXCR1*	T	0.04	7,661	0.008 (0.013)	0.53	0.76	25,204	0.030 (0.007)	6.9×10^{-6}	0.74
rs1501908	5	156.3	*TIMD4*	C	0.35	14,074	-0.006 (0.004)	0.10	0.87	31,617	-0.012 (0.002)	4.9×10^{-7}	0.19
rs2142672	6	16.3	*MYLIP, GMPR*	C	0.74	28,112	0.010 (0.003)	2.0×10^{-4}	0.30	45,655	0.013 (0.002)	2.7×10^{-9}	0.30
rs9275427	6	32.8	*HLA-DQA2*	T	0.16	12,777	0.007 (0.004)	0.07	0.68	30,320	0.013 (0.003)	1.4×10^{-5}	0.28
rs456598	6	160.5	*IGF2R, SLC22A1*	G	0.87	19,882	-0.010 (0.004)	0.01	0.79	37,425	-0.015 (0.003)	8.4×10^{-7}	0.58
rs2126259	8	9.2	*PPP1R3B*	A	0.10	28,145	-0.014 (0.004)	9.5×10^{-5}	0.13	45,688	-0.018 (0.003)	1.4×10^{-9}	0.10
rs4135183	9	110.1	*TXN*	A	0.13	12,777	-0.004 (0.005)	0.39	0.81	30,320	-0.014 (0.004)	1.3×10^{-4}	0.17
rs7136824	10	12.5	*CAMK1D*	G	0.80	22,855	0.001 (0.003)	0.76	0.70	40,398	-0.007 (0.003)	9.3×10^{-3}	0.03
rs7697566	10	49.1	*FRMPD2*	T	0.82	18,299	-0.003 (0.004)	0.52	0.70	35,842	0.008 (0.003)	0.01	0.03

rs11227299	11	65.3	*DKFZp761E198*	G	0.35	18,821	0.002 (0.003)	0.60	0.39	36,364	-0.006 (0.002)	0.01	5.0×10^{-3}
rs4307732	11	125.8	*ST3GAL4*	G	0.89	7,661	-0.004 (0.007)	0.63	0.42	25,204	-0.017 (0.005)	1.2×10^{-4}	0.55
HDL-c													
rs10749832	1	41.7	*HIVEP3*	G	0.40	12,634	0.005 (0.003)	0.14	0.65	28,690	0.009 (0.002)	1.5×10^{-5}	0.46
rs3923229	2	85.5	*TGOLN2, RETSAT*	A	0.74	30,386	0.001 (0.002)	0.60	0.81	46,442	0.006 (0.002)	2.5×10^{-3}	0.14
rs7560571	2	95.1	*MAL*	G	0.35	17,809	-0.004 (0.003)	0.12	0.25	33,865	-0.009 (0.002)	2.3×10^{-5}	0.11
rs10088837	3	63.4	*SYNPR*	G	0.24	25,554	0.001 (0.003)	0.63	0.15	41,610	0.007 (0.002)	1.9×10^{-3}	0.04
rs13107325	4	103.5	*SLC39A8*	T	0.08	22,128	-0.017 (0.006)	2.1×10^{-3}	0.24	38,184	-0.023 (0.004)	1.6×10^{-8}	0.37
rs643531	9	15.3	*TTC39B*	C	0.14	34,152	-0.009 (0.003)	2.6×10^{-3}	0.60	50,208	-0.013 (0.002)	4.1×10^{-8}	0.24
rs174548	11	61.3	*FADS1*	G	0.30	33,930	-0.008 (0.002)	7.6×10^{-5}	0.78	49,986	-0.011 (0.002)	9.9×10^{-10}	0.59
rs4943343	13	35.3	*DCLK1*	T	0.08	11,398	0.003 (0.007)	0.63	0.70	27,454	-0.014 (0.005)	2.6×10^{-3}	0.18
rs11642898	16	19.6	*C16orf62*	T	0.14	30,722	0.002 (0.003)	0.61	0.35	46,778	0.008 (0.003)	2.0×10^{-3}	0.01
rs756942	17	65.1	*MAP2K6*	T	0.24	16,536	-0.002 (0.003)	0.60	0.27	32,592	0.006 (0.002)	6.0×10^{-3}	4.6×10^{-3}
rs12980205	19	48.7	*PHLDB3*	T	0.23	16,536	0.002 (0.003)	0.66	0.21	32,592	-0.007 (0.003)	0.01	4.9×10^{-3}
rs1981383	21	42.2	*C2CD2*	G	0.48	25,218	0.002 (0.002)	0.50	0.02	41,274	-0.004 (0.002)	0.05	8.5×10^{-5}

TG

rs1251474	1	76.2	*LOC729766*	C	0.64	27,759	-0.001 (0.004)	0.86	0.28	43,817	-0.010 (0.004)	5.1×10^{-3}	0.02
rs2791544	1	216.1	*LYPLAL1*	C	0.62	13,039	0.005 (0.006)	0.45	0.03	29,097	0.018 (0.004)	6.8×10^{-5}	0.02
rs10497518	2	179.3	*TTN*	G	0.03	14,365	-0.021 (0.014)	0.14	0.98	30,423	0.021 (0.010)	0.04	0.05
rs1378168	3	16.0	*ANKRD28*	G	0.28	20,602	-0.0004 (0.006)	0.95	0.49	36,660	-0.014 (0.004)	1.5×10^{-3}	0.04
rs4490955	3	18.7	*SATB1*	G	0.27	20,701	-0.010 (0.006)	0.07	0.66	36,759	-0.018 (0.004)	1.8×10^{-5}	0.36
rs9309766	3	77.4	*ROBO2*	T	0.59	14,368	0.006 (0.006)	0.33	0.82	30,426	-0.010 (0.004)	0.02	0.08
rs442177	4	88.4	*AFF1*	A	0.60	28,676	0.014 (0.004)	1.2×10^{-3}	0.99	44,734	0.019 (0.004)	1.5×10^{-7}	0.71
rs17378274	4	162.4	*FSTL5*	C	0.43	13,039	-0.001 (0.006)	0.89	0.07	29,097	-0.015 (0.004)	6.5×10^{-4}	2.6×10^{-3}
rs6867983	5	55.9	*C5orf35*	T	0.14	23,957	0.014 (0.007)	0.04	0.15	40,015	0.024 (0.005)	6.1×10^{-6}	0.10
rs646202	6	139.9	*CITED2*	T	0.49	13,039	0.006 (0.006)	0.39	0.88	29,097	0.017 (0.004)	1.1×10^{-4}	0.49
rs10111012	8	110.2	*TRHR*	T	0.89	21,927	-0.014 (0.011)	0.22	0.28	37,985	-0.031 (0.008)	5.7×10^{-5}	0.11
rs17663474	8	117.7	*EIF3H*	C	0.91	13,057	-0.0002 (0.013)	0.98	0.59	29,115	0.028 (0.008)	7.5×10^{-4}	0.17
rs2499902	10	11.5	*CUGBP2*	T	0.23	19,342	0.007 (0.006)	0.28	0.72	35,400	-0.009 (0.005)	0.06	0.03
rs174548	11	61.3	*FADS1*	G	0.30	31,066	0.019 (0.004)	2.0×10^{-5}	0.98	47,124	0.024 (0.004)	8.9×10^{-11}	0.78
rs4906879	15	24.2	*GABRB3*	G	0.56	19,198	0.016 (0.005)	2.3×10^{-3}	0.20	35,256	-0.001 (0.004)	0.89	3.4×10^{-6}

rs12913924	15	55.6	*CGNL1*	G	0.82	22,496	-0.010 (0.006)	0.09	0.56	38,554	-0.021 (0.005)	3.7×10^{-5}	0.33	

The replication analysis (Stage 2) is based on data from up to eight study populations comprising up to 37,774 participants.
The combined analysis is based on data from up to 15 studies from Stages 1 and 2 and comprising up to 55,497 participants.
* Based on NCBI Build 35.
† Based on Study 1 (EPIC-Norfolk sub-cohort). Effect allele corresponds to forward strand of NCBI Build 36.3.
‡ Beta-coefficients represent the change in circulating lipid level (natural log) per additional effect allele, adjusted for age and sex.