Discovery and refinement of loci associated with lipid levels

Global Lipids Genetics Consortium*

Levels of low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides and total cholesterol are heritable, modifiable risk factors for coronary artery disease. To identify new loci and refine known loci influencing these lipids, we examined 188,577 individuals using genome-wide and custom genotyping arrays. We identify and annotate 157 loci associated with lipid levels at $P < 5 \times 10^{-8}$, including 62 loci not previously associated with lipid levels in humans. Using dense genotyping in individuals of European, East Asian, South Asian and African ancestry, we narrow association signals in 12 loci. We find that loci associated with blood lipid levels are often associated with cardiovascular and metabolic traits, including coronary artery disease, type 2 diabetes, blood pressure, waist-hip ratio and body mass index. Our results demonstrate the value of using genetic data from individuals of diverse ancestry and provide insights into the biological mechanisms regulating blood lipids to guide future genetic, biological and therapeutic research.

© 2013 Nature America, Inc. All rights reserved.

npg

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

Here we report on studies of naturally occurring variation in 188,577 European-ancestry individuals and 7,898 non-European-ancestry individuals. Our analyses identify 157 loci associated with lipid levels at $P < 5 \times 10^{-8}$, including 62 new loci. Thirty of the 62 loci do not include genes implicated in lipid biology by previous literature. We tested lipid-associated SNPs for association with mRNA expression levels, carried out pathway analyses to uncover relationships between loci and compared the locations of lipid-associated SNPs with those of genes and other functional elements in the genome. These results provide direction for biological and therapeutic research into risk factors for CAD.

RESULTS

New loci associated with blood lipid levels

We examined subjects of European ancestry, including 94,595 individuals from 23 studies genotyped with genome-wide association study (GWAS) arrays[4] and 93,982 individuals from 37 studies genotyped with the Metabochip array[8] (**Supplementary Fig. 1** and **Supplementary Table 1**). The Metabochip includes variants representing promising loci from our previous GWAS (14,886 SNPs) and from GWAS of other CAD risk factors and related traits (50,459 SNPs), variants from the 1000 Genomes Project[9] and focused resequencing[10] efforts in 64 previously associated loci (28,923 SNPs) and fine-mapping variants in 181 loci associated with other traits (93,308 SNPs). In cases where Metabochip and GWAS array data were available for the same individuals, we used Metabochip data to ensure that key variants were directly genotyped rather than imputed.

We excluded individuals known to be on lipid-lowering medications and evaluated the additive effect of each SNP on blood lipid levels after adjusting for age and sex. Genomic control values[11] for the initial meta-analyses were 1.10–1.15, low for a sample of this size, indicating that population stratification should have had only a minor impact on our results (**Supplementary Fig. 2**). After genomic control correction, 157 loci associated with blood lipid levels were identified ($P < 5 \times 10^{-8}$), including 62 newly associated loci (**Fig. 1**, **Tables 1–4** and **Supplementary Tables 2** and **3**). Loci were >1 Mb apart and nearly independent ($r^2 < 0.10$). Of the 62 newly associated loci, 24 demonstrated the strongest evidence of association with HDL cholesterol levels, 15 demonstrated the strongest evidence of association with LDL cholesterol levels, 8 demonstrated the strongest evidence of association with triglyceride levels, and 15 demonstrated the strongest evidence of association with total cholesterol (**Supplementary Fig. 3**). Several of these loci were validated by a similar extension based on published Global Lipids Genetics Consortium GWAS results[12].

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

Overlap of genetic discoveries and previous knowledge

To investigate connections between our new loci and known lipid biology, we first catalogued genes within 100 kb of the peak associated SNPs and searched PubMed and Online Mendelian Inheritance in Man (OMIM)

*Full lists of authors and affiliations appear at the end of the paper.

Received 12 October 2012; accepted 13 September 2013; published online 6 October 2013; doi:10.1038/ng.2797

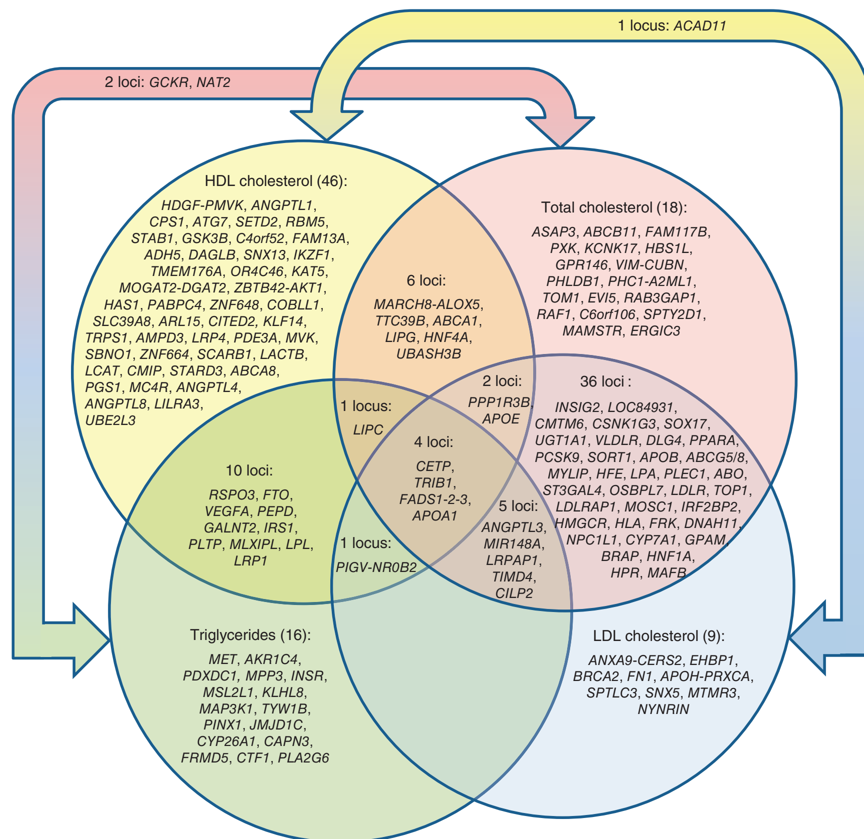

Figure 1 Overlap of loci associated with different lipid traits. The Venn diagram illustrates the number of loci that show association with multiple lipid traits. The number of loci primarily associated with only one trait is listed in parentheses after the trait name, and locus names are listed below. Loci that show association with two or more traits are shown in the appropriate segment.

for occurrences of these gene names and their aliases in the context of relevant keywords. After manual curation, we identified at least 1 strong candidate in 32 of the 62 loci (52%) (**Supplementary Table 4**). For the remaining 30 loci, we found no literature support for the role of a nearby gene in regulating blood lipid levels. This search highlighted genes whose connections to lipid metabolism have been extensively documented in mouse models (such as *VLDLR*[13] and *LRPAP1* (ref. 13)) and human cell lines (such as *VIM*[14]), as well as candidates whose connection to lipid levels is more recent, such as *VEGFA*. With respect to the latter, recent studies of *VEGFB* have suggested that vascular endothelial growth factors have an unexpected role in the targeting of lipids to peripheral tissues[15], which we corroborate by associating variants near *VEGFA* with blood triglyceride and HDL cholesterol levels.

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

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

Pathway analyses

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

Protein-protein interactions

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

Regulation of gene expression by associated variants

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

© 2013 Nature America, Inc. All rights reserved.
npg

Table 1 New loci primarily associated with HDL cholesterol levels obtained from joint GWAS and Metabochip meta-analysis

Locus	Marker name	Chr.	hg19 position (Mb)	Associated trait(s)	MAF	Minor/major allele	Effect of A1	Joint *n* (×1,000)	Joint *P* value
PIGV-NR0B2	rs12748152	1	27.14	HDL, LDL, TG	0.09	T/C	–0.051, 0.050, 0.037	187, 173, 178	1×10^{-15}, 3×10^{-12}, 1×10^{-9}
HDGF-PMVK	rs12145743	1	156.70	HDL	0.34	G/T	0.020	181	2×10^{-8}
ANGPTL1	rs4650994	1	178.52	HDL	0.49	G/A	0.021	187	7×10^{-9}
CPS1	rs1047891	2	211.54	HDL	0.33	A/C	–0.027	182	9×10^{-10}
ATG7	rs2606736	3	11.40	HDL	0.39	C/T	0.025	129	5×10^{-8}
SETD2	rs2290547	3	47.06	HDL	0.20	A/G	–0.030	187	4×10^{-9}
RBM5	rs2013208	3	50.13	HDL	0.50	T/C	0.025	170	9×10^{-12}
STAB1	rs13326165	3	52.53	HDL	0.21	A/G	0.029	187	9×10^{-11}
GSK3B	rs6805251	3	119.56	HDL	0.39	T/C	0.020	186	1×10^{-8}
C4orf52	rs10019888	4	26.06	HDL	0.18	G/A	–0.027	187	5×10^{-8}
FAM13A	rs3822072	4	89.74	HDL	0.46	A/G	–0.025	187	4×10^{-12}
ADH5	rs2602836	4	100.01	HDL	0.44	A/G	0.019	187	5×10^{-8}
RSPO3	rs1936800	6	127.44	HDL, TG[a]	0.49	C/T	0.020, –0.020	187, 168	3×10^{-10}, 3×10^{-8}
DAGLB	rs702485	7	6.45	HDL	0.45	G/A	0.024	187	6×10^{-12}
SNX13	rs4142995	7	17.92	HDL	0.38	T/G	–0.026	165	9×10^{-12}
IKZF1	rs4917014	7	50.31	HDL	0.32	G/T	0.022	187	1×10^{-8}
TMEM176A	rs17173637	7	150.53	HDL	0.12	C/T	–0.036	184	2×10^{-8}
MARCH8-ALOX5	rs970548	10	46.01	HDL, TC	0.26	C/A	0.026, 0.025	187, 187	2×10^{-10}, 8×10^{-9}
OR4C46	rs11246602	11	51.51	HDL	0.15	C/T	0.034	176	2×10^{-10}
KAT5	rs12801636	11	65.39	HDL	0.23	A/G	0.024	187	3×10^{-8}
MOGAT2-DGAT2	rs499974	11	75.46	HDL	0.19	A/C	–0.026	187	1×10^{-8}
ZBTB42-AKT1	rs4983559	14	105.28	HDL	0.40	G/A	0.020	184	1×10^{-8}
FTO	rs1121980	16	53.81	HDL, TG[b]	0.43	A/G	–0.020, 0.021	186, 155	7×10^{-9}, 3×10^{-8}
HAS1	rs17695224	19	52.32	HDL	0.26	A/G	–0.029	185	2×10^{-13}

Chr., chromosome; A1, minor allele; A2, major allele; TG, triglycerides; TC, total cholesterol. Effect sizes are given with respect to the minor allele (A1) in s.d. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest *P* value is listed first.
[a]The secondary trait was most strongly associated with a different SNP: rs719726 (within 1 Mb of rs1936800, $r^2 = 0.74$). [b]The secondary trait was most strongly associated with a different SNP, rs9930333 (within 1 Mb of rs1121980, $r^2 = 0.99$).

searches also prioritized candidate genes. In all three, eQTL analysis and literature review identified the same candidate (*DAGLB*, *SPTLC3* and *PXK*; $P = 0.05$). For the remaining four loci (near *RBM5*, *ADH5*, *TMEM176A* and *GPR146*), analysis of expression levels identified candidates that were not supported by literature or pathway analyses.

Coding variation

In some loci where previous association studies of coding variants were inconclusive, we now found convincing evidence of association, demonstrating the benefits of the large sample sizes achievable through collaboration. For example, in the *APOH* locus[22], our most strongly associated variant was rs1801689 (APOH p.Cys325Gly; $P = 1 \times 10^{-11}$ for LDL cholesterol levels). Overall, at 15 of the 62 new loci, there was at least 1 nonsynonymous variant within 100 kb of and in strong LD ($r^2 > 0.8$) with the index SNP (**Supplementary Table 7**) (18 loci when there was no restriction on distance). This ~30% overlap between associated loci and coding variation is similar to that for other complex traits[9]. Unexpectedly, in the 11 loci where a candidate

Table 2 New loci primarily associated with LDL cholesterol levels obtained from joint GWAS and Metabochip meta-analysis

Locus	Marker name	Chr.	hg19 position (Mb)	Associated trait(s)	MAF	Minor/major allele	Effect of A1	Joint *n* (×1,000)	Joint *P* value
ANXA9-CERS2	rs267733	1	150.96	LDL	0.16	G/A	–0.033	165	5×10^{-9}
EHBP1	rs2710642	2	63.15	LDL	0.35	G/A	–0.024	173	6×10^{-9}
INSIG2	rs10490626	2	118.84	LDL, TC[a]	0.08	A/G	–0.051, –0.042	173, 184	2×10^{-12}, 6×10^{-9}
LOC84931	rs2030746	2	121.31	LDL, TC	0.40	T/C	0.021, 0.020	173, 187	9×10^{-9}, 4×10^{-8}
FN1	rs1250229	2	216.30	LDL	0.27	T/C	–0.024	173	3×10^{-8}
CMTM6	rs7640978	3	32.53	LDL, TC	0.09	T/C	–0.039, –0.038	172, 186	1×10^{-8}, 2×10^{-8}
ACAD11	rs17404153	3	132.16	LDL, HDL[b]	0.14	T/G	–0.034, –0.028	172, 187	2×10^{-9}, 5×10^{-9}
CSNK1G3	rs4530754	5	122.86	LDL, TC	0.46	G/A	–0.028, –0.023	173, 187	4×10^{-12}, 2×10^{-9}
MIR148A	rs4722551	7	25.99	LDL, TG[c], TC	0.20	C/T	0.039, 0.023, 0.029	173, 178, 187	4×10^{-14}, 9×10^{-11}, 7.0×10^{-9}
SOX17	rs10102164	8	55.42	LDL, TC	0.21	A/G	0.032, 0.030	173, 187	4×10^{-11}, 5×10^{-11}
BRCA2	rs4942486	13	32.95	LDL	0.48	T/C	0.024	172	2×10^{-11}
APOH-PRXCA	rs1801689	17	64.21	LDL	0.04	C/A	0.103	111	1×10^{-11}
SPTLC3	rs364585	20	12.96	LDL	0.38	A/G	–0.025	172	4×10^{-10}
SNX5	rs2328223	20	17.85	LDL	0.21	C/A	0.03	171	6×10^{-9}
MTMR3	rs5763662	22	30.38	LDL	0.04	T/C	0.077	163	1×10^{-8}

Chr., chromosome; A1, minor allele; A2, major allele; TG, triglycerides; TC, total cholesterol. Effect sizes are given with respect to the minor allele (A1) in s.d. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest *P* value is listed first.
[a]The secondary trait was most strongly associated with a different SNP, rs17526895 (within 1 Mb of rs10490626, $r^2 = 0.98$). [b]The secondary trait was most strongly associated with a different SNP, rs13076253 (within 1 Mb of rs17404153, $r^2 = 0.00$). [c]The secondary trait was most strongly associated with the different SNP rs4719841 (within 1 Mb of rs4722551, $r^2 = 0.10$).

© 2013 Nature America, Inc. All rights reserved.

npg

Table 3 New loci primarily associated with total cholesterol levels obtained from joint GWAS and Metabochip meta-analysis

Locus	Marker name	Chr.	hg19 position (Mb)	Associated trait(s)	MAF	Minor/major allele	Effect of A1	Joint n (×1,000)	Joint P value
ASAP3	rs1077514	1	23.77	TC	0.15	C/T	–0.03	184	6×10^{-9}
ABCB11	rs2287623	2	169.83	TC	0.41	G/A	0.027	184	4×10^{-12}
FAM117B	rs11694172	2	203.53	TC	0.25	G/A	0.028	187	2×10^{-9}
UGT1A1	rs11563251	2	234.68	TC, LDL	0.12	T/C	0.037, 0.034	187, 173	1×10^{-9}, 5×10^{-8}
PXK	rs13315871	3	58.38	TC	0.10	A/G	–0.036	187	4×10^{-8}
KCNK17	rs2758886	6	39.25	TC	0.30	A/G	0.023	187	3×10^{-8}
HBS1L	rs9376090	6	135.41	TC	0.28	C/T	–0.025	187	3×10^{-9}
GPR146	rs1997243	7	1.08	TC	0.16	G/A	0.033	183	3×10^{-10}
VLDLR	rs3780181	9	2.64	TC, LDL	0.08	G/A	–0.044, –0.044	186, 172	7×10^{-10}, 2×10^{-9}
VIM-CUBN	rs10904908	10	17.26	TC	0.43	G/A	0.025	187	3×10^{-11}
PHLDB1	rs11603023	11	118.49	TC	0.42	T/C	0.022	187	1×10^{-8}
PHC1-A2ML1	rs4883201	12	9.08	TC	0.12	G/A	–0.035	187	2×10^{-9}
DLG4	rs314253	17	7.09	TC, LDL	0.37	C/T	–0.023, –0.024	184, 170	3×10^{-10}, 3×10^{-10}
TOM1	rs138777	22	35.71	TC	0.36	A/G	0.021	185	5×10^{-8}
PPARA	rs4253772	22	46.63	TC, LDL[a]	0.11	T/C	0.032, 0.031	185, 171	1×10^{-8}, 3×10^{-8}

Chr., chromosome; A1, minor allele; A2, major allele; TC, total cholesterol. Effect sizes are given with respect to the minor allele (A1) in s.d. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest P value is listed first.
[a]At one locus, the secondary trait was most strongly associated with a different SNP, rs4253776 (within 1 Mb of rs4253772, r^2 = 0.95).

© 2013 Nature America, Inc. All rights reserved. npg

was suggested by literature review and by examination of coding variation, the candidates from these methods coincided 7 times (P = 0.03 compared to the expected overlap by chance of 3.8 times); thus, agreement between literature review and examination of coding variation was less significant than for eQTL studies and analyses of pathways or protein-protein interactions.

Overlap between association signals and regulators of transcription in liver

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

We proceeded to investigate the overlap between lipid-associated loci and functional marks in HepG2 cells in more detail (**Supplementary Table 9**). Notable regulatory elements showing significant overlap with lipid-associated loci included histone marks associated with active regulatory regions (acetylation of histone H3 at lysine 27 (H3K27ac), $P = 3 \times 10^{-20}$; acetylation of histone H3 at lysine 9 (H3K9ac), $P = 3 \times 10^{-22}$), promoters (trimethylation of histone H3 at lysine 4 (H3K4me3), $P = 2 \times 10^{-15}$; dimethylation of histone H3 at lysine 4 (H3K4me2), $P = 8 \times 10^{-12}$), transcribed regions (trimethylation of histone H3 at lysine 36 (H3K36me3), $P = 4 \times 10^{-14}$), indicators of open chromatin (FAIRE (formaldehyde-assisted isolation of regulatory elements), $P = 5 \times 10^{-9}$; DNase I sensitivity, $P = 2 \times 10^{-4}$) and regions that interact with the transcription factors HNF4A ($P = 6 \times 10^{-10}$) and CEBP/B ($P = 1 \times 10^{-5}$). Overall, 56 of our 62 new loci contained at least 1 SNP that overlapped a functional mark[24] and/or chromatin state[23] highlighted in **Supplementary Table 9**, including all but 3 of the loci where no candidates were suggested by literature review or analyses of pathways, coding variation or gene expression (**Supplementary Table 10**).

Initial fine mapping of 65 lipid-associated loci

Previous fine mapping of five LDL cholesterol–associated loci found that variants with the strongest association were often substantially different in frequency and effect size from those identified by GWAS[10]. Metabochip genotypes enabled us to carry out an initial fine-mapping analysis for 65 loci: 60 selected for fine mapping on

Table 4 New loci primarily associated with triglyceride levels obtained from joint GWAS and Metabochip meta-analysis

Locus	Marker name	Chr.	hg19 position (Mb)	Associated trait(s)	MAF	Minor/major allele	Effect of A1	Joint N (×1,000)	Joint P value
LRPAP1	rs6831256	4	3.47	TG, TC[a], LDL[a]	0.42	G/A	0.026, 0.025, 0.022	177, 187, 173	2×10^{-12}, 1×10^{-10}, 2×10^{-8}
VEGFA	rs998584	6	43.76	TG, HDL	0.49	A/C	0.029, –0.026	175, 184	3×10^{-15}, 2×10^{-11}
MET	rs38855	7	116.36	TG	0.47	G/A	–0.019	178	2×10^{-8}
AKR1C4	rs1832007	10	5.25	TG	0.18	G/A	–0.033	178	2×10^{-12}
PDXDC1	rs3198697	16	15.13	TG	0.43	T/C	–0.020	176	2×10^{-8}
MPP3	rs8077889	17	41.88	TG	0.22	C/A	0.025	176	1×10^{-8}
INSR	rs7248104	19	7.22	TG	0.42	A/G	–0.022	176	5×10^{-10}
PEPD	rs731839	19	33.90	TG, HDL	0.35	G/A	0.022, –0.022	176, 185	3×10^{-9}, 3×10^{-9}

Chr., chromosome; A1, minor allele; A2, major allele; TG, triglycerides; TC, total cholesterol. Effect sizes are given with respect to the minor allele (A1) in s.d. For loci associated with two or more traits at genome-wide significance, the trait corresponding to the strongest P value is listed first.
[a]At one locus, secondary traits were most strongly associated with a different SNP, rs6818397 (within 1 Mb of rs6831256, r^2 = 0.18).

the basis of our previous study[4] and 5 nominated for fine mapping because of association with other traits.

For each of these loci, we identified the most strongly associated Metabochip variant and evaluated whether it (i) reached genome-wide significant evidence for association (to avoid chance fluctuations in regions where the signal was relatively weak) and (ii) was different from the GWAS index SNP in terms of frequency and effect size (operationalized to $r^2 < 0.8$ with the GWAS index SNP). In the European samples, fine mapping identified eight loci where the fine-mapping signal was clearly different from the GWAS signal (**Supplementary Table 11**). The two largest differences were at the loci near *PCSK9* (top GWAS variant with MAF (f) = 0.24, $P = 9 \times 10^{-24}$; fine-mapping variant with $f = 0.03$, $P = 2 \times 10^{-136}$) and *APOE* (GWAS variant $f = 0.20$, $P = 3 \times 10^{-44}$; fine-mapping variant $f = 0.07$, $P = 3 \times 10^{-651}$), consistent with results from Sanna *et al.*[10]. Large differences were also observed near *LRP4* (GWAS $f = 0.17$, $P = 8 \times 10^{-14}$; fine-mapping $f = 0.35$, $P = 1 \times 10^{-26}$), *IGF2R* (GWAS $f = 0.16$, $P = 7 \times 10^{-9}$; fine-mapping $f = 0.37$, $P = 2 \times 10^{-13}$), *NPC1L1* (GWAS $f = 0.27$, $P = 2 \times 10^{-5}$; fine-mapping $f = 0.24$, $P = 1 \times 10^{-12}$), *ST3GAL4* (GWAS $f = 0.26$, $P = 2 \times 10^{-6}$; fine-mapping $f = 0.07$, $P = 6 \times 10^{-11}$), *MED1* (GWAS $f = 0.37$, $P = 3 \times 10^{-5}$; fine-mapping $f = 0.24$, $P = 2 \times 10^{-10}$) and *COBLL1* (GWAS $f = 0.12$, $P = 2 \times 10^{-6}$; fine-mapping $f = 0.11$, $P = 6 \times 10^{-9}$). Thus, although the large changes observed by Sanna *et al.*[10] after fine mapping are by no means unique, they are not typical. Except for the p.Arg46Leu variant encoded in *PCSK9*, the variants showing the strongest association in fine-mapped loci all had MAF > 0.05.

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

Association of lipid-related loci with metabolic and cardiovascular traits

To evaluate the role of the 157 loci identified here in related traits, we evaluated the most strongly associated SNPs for each locus in genetic studies of CAD (n = 114,590, including 37,653 cases)[27,28], type 2 diabetes (T2D; n = 47,117, including 8,130 cases)[29], body mass index (BMI; n = 123,865 individuals)[30] and waist-hip ratio (WHR; n = 77,167 individuals)[31], systolic and diastolic blood pressure (SBP and DBP; n = 69,395 individuals)[32] and fasting glucose levels (n = 46,186 non-diabetic individuals)[33]. We observed an excess of SNPs nominally associated ($P < 0.05$) with all these traits, including a 5.1-fold excess for CAD (40 nominally significant loci; $P = 2 \times 10^{-19}$), a 4.1-fold excess for BMI (32 loci; $P = 1 \times 10^{-11}$), a 3.7-fold excesses for DBP (29 loci; $P = 1 \times 10^{-9}$), a 3.4-fold excess for WHR (27 loci; $P = 1 \times 10^{-9}$), a 2.5-fold excess for SBP (20 loci; $P = 1 \times 10^{-4}$), a 2.3-fold excess for T2D (18 loci; $P = 0.001$) and a 2.2-fold excess for fasting glucose levels (17 loci; $P = 3 \times 10^{-3}$) (**Supplementary Table 12**). Interestingly, for the new loci, we observed greater overlap with BMI, SBP and DBP (nine overlapping loci each) than with CAD (eight overlapping loci). Of the new loci, the two SNPs showing the strongest association with CAD mapped near *RBM5* (rs2013208: $P_{HDL} = 9 \times 10^{-12}$, $P_{CAD} = 7 \times 10^{-5}$) and *CMTM6* (rs7640978: $P_{LDL} = 1 \times 10^{-8}$, $P_{CAD} = 4 \times 10^{-4}$).

We tested whether the LDL cholesterol–, total cholesterol– or triglyceride-increasing allele or the HDL cholesterol–decreasing allele was associated with increased risk of cardiovascular disease or related metabolic outcomes; the direction of effect of each locus was categorized according to the primary association signal at the locus, as in **Tables 1**–**4**. We observed association with increased CAD risk (104/149; $P = 1 \times 10^{-6}$), SBP (96/155; $P = 2.7 \times 10^{-3}$) and WHR adjusted for BMI (92/154; $P = 0.019$). There were many instances where a single locus was associated with many traits. These included variants near *FTO*, consistent with previous reports[34]; near *VEGFA* (associated with triglyceride levels, CAD, T2D, SBP and DBP); near *SLC39A8* (associated with HDL cholesterol levels, BMI, SBP and DBP); and near *MIR581* (associated with HDL cholesterol levels, BMI, T2D and DBP). In some cases, such as *FTO*, a strong association with BMI or another phenotype generated weaker association signals for other metabolic traits[34]. In other cases, such as *SORT1*, a primary effect on lipid levels might mediate secondary association with other traits, such as CAD[7].

Association of lipid traits with CAD

Epidemiological studies consistently show that high total cholesterol and LDL cholesterol levels are associated with increased risk of CAD, whereas high HDL cholesterol levels are associated with reduced risk of CAD[35]. In genetic studies, the connection between LDL cholesterol levels and CAD is clear, whereas the results for HDL cholesterol levels are more equivocal[36–38]. In our data, trait-increasing alleles at the loci showing the strongest association with LDL cholesterol levels (31 loci), triglyceride levels (30 loci) or total cholesterol levels (38 loci) were associated with increased risk of CAD ($P = 2 \times 10^{-12}$, 2×10^{-16} and 0.006, respectively). Conversely, trait-decreasing alleles at loci showing the strongest association with HDL cholesterol levels (64 loci) were associated with increased CAD risk at $P = 0.02$. When we focused on loci uniquely associated with LDL cholesterol levels (12 loci where $P > 0.05$ for other lipids), triglyceride levels (6 loci) or HDL cholesterol levels (14 loci), only the association with LDL cholesterol remained significant ($P = 0.03$).

To better explore how associations with individual lipid levels were related to CAD risk, we used linear regression to test whether association with lipid levels could predict impact on CAD risk. In this analysis, the effect on CAD of 149 lipid-associated loci (CAD results were not available for 8 SNPs) was correlated with LDL cholesterol (Pearson's $r = 0.74$; $P = 7 \times 10^{-6}$) and triglyceride (Pearson's $r = 0.46$; $P = 0.02$) effect sizes but not with HDL cholesterol effect sizes (Pearson's $r = -9 \times 10^{-4}$; $P = 0.99$; **Supplementary Fig. 6**). Because most variants affect multiple lipid fractions (**Fig. 1**), dissecting the relationship between lipid level and CAD effects requires multivariate analysis. In a companion manuscript in this issue, we use multivariate analysis and detailed examination of triglyceride-associated loci to show that increased LDL cholesterol and triglyceride levels but not HDL cholesterol levels appear to be causally related to CAD risk[39].

Evidence for additional loci not yet reaching genome-wide significance

To evaluate evidence for loci not yet reaching genome-wide significance, we compared the directions of effect in GWAS and Metabochip analyses of non-overlapping samples outside the 157 genome-wide significant loci. For independent variants ($r^2 < 0.1$) with association $P < 0.1$ in the GWAS-only analysis, a significant excess was concordant in the direction of effect for HDL cholesterol levels (62.9% in 1,847 SNPs; $P < 1 \times 10^{-16}$), LDL cholesterol levels (58.6% of 1,730 SNPs; $P < 1 \times 10^{-16}$), triglyceride levels (59.1% of 1,783 SNPs;

© 2013 Nature America, Inc. All rights reserved. npg

$P < 1 \times 10^{-16}$) and total cholesterol levels (61.0% of 1,904 SNPs; $P < 1 \times 10^{-16}$), suggesting that there are many additional loci to be discovered in future studies.

DISCUSSION

Molecular understanding of the genes and pathways that modify blood lipid levels in humans will facilitate the design of new therapies for cardiovascular and metabolic disease. This understanding can be gained from studies of model organisms, *in vitro* experiments, bioinformatic analyses and human genetic studies. Here we demonstrate association between blood lipid levels and 62 new loci, bringing the total number of lipid-associated loci to 157 (**Tables 1–4** and **Fig. 1**). All but one of the loci identified here include protein-coding genes within 100 kb of the SNP showing the strongest association. Whereas 38 of the 62 new loci include genes whose role in the regulation of blood lipid levels is supported by literature review or analysis of curated pathway databases, the remainder include only genes whose role in such regulation has not been documented.

In total, there are 240 genes within 100 kb of 1 of our 62 new lipid-associated loci—providing a daunting challenge for future functional studies. Prioritizing on the basis of literature review, pathway analysis, regulation of mRNA expression levels and protein-altering variants suggests that 70 genes in 44 of the 62 new loci might be the focus of the first round of functional studies (summarized in **Supplementary Table 2**). Although we found significant overlap, different sources of prioritization sometimes disagreed. This result suggests that truly understanding causality will be very challenging. We include an interpreted digest of genes highlighted by our study in the **Supplementary Note**. Clearly, a range of approaches will be needed to follow up these findings. To illustrate possibilities, consider US Patent Application 20090036394 disclosing that, in the mouse, knockout of *Gpr146* modifies blood lipid levels. Here we show that variants near the human homolog of this gene, *GPR146*, are associated with the levels of total cholesterol—providing an added incentive for studies of GPR146 inhibitors in humans. *GPR146* encodes a G protein–coupled receptor, an attractive pharmaceutical target, so it is tempting to speculate that, one day, pharmaceutical inhibition of GPR146 may modify cholesterol levels and reduce risk of heart disease.

Each associated locus typically includes many strongly associated (and potentially causal) variants. Our fine-mapping results illustrate how genetic analysis of large samples and individuals of diverse ancestry can help focus the search for causal variants. In our fine-mapping analysis of 65 lipid-associated loci, we were able to separate the strongest signal in a region from the previous GWAS-identified signal in 12 instances. In 3 of these 12 instances, fine mapping was enabled by the analysis of a few thousand individuals of African or East Asian ancestry, whereas, in the remaining instances, fine mapping was possible through the examination of nearly 100,000 individuals of European ancestry. A more detailed fine-mapping exercise, including imputation of variants from emerging, very large reference panels, may help refine the locations of additional signals.

Lipid-associated loci were strongly associated with CAD, T2D, BMI, SBP and DBP. In univariate analyses, we found that effects on LDL cholesterol and triglyceride levels all predicted association with CAD, but HDL cholesterol levels did not. In a companion paper, more detailed multivariate investigation shows that our data are consistent with the hypothesis that both LDL cholesterol and triglyceride levels but not HDL cholesterol levels are causally related to CAD risk. HDL cholesterol, LDL cholesterol and triglyceride levels summarize aggregate levels of different lipid particles, each with potentially distinct consequences for CAD risk. We evaluated the association of our loci with lipid subfractions in 2,900 individuals from the Framingham Heart Study (**Supplementary Fig. 7** and **Supplementary Table 13**) and with sphingolipids, which are components of lipid membranes in cells, in 4,034 individuals from 5 samples of European ancestry[40] (**Supplementary Table 14**). The results suggest that HDL cholesterol–associated variants can have a markedly different impact on these subphenotypes. For example, among HDL cholesterol–associated loci, variants near *LIPC* were strongly associated with plasmalogen levels ($P < 1 \times 10^{-40}$), variants near *ABCA1* were associated with sphingomyelin levels ($P < 1 \times 10^{-5}$), and variants near *CETP*, which show the strongest association with HDL cholesterol levels overall, were associated with neither of these. Detailed genetic dissection of these subphenotypes in larger samples could lead to functional groupings of HDL cholesterol–associated variants that reconcile the results of genetic studies (which show no clear connection between HDL cholesterol–associated variants and CAD risk) and epidemiological studies (which show clear association between plasma HDL cholesterol levels and CAD risk).

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

URLs. Summary results for our studies are available. We hope that they will facilitate continued research into the genetics of blood lipid levels and, eventually, help identify improved treatments for CAD. To browse the full result set, go to http://www.sph.umich.edu/csg/abecasis/public/lipids2013/. Snipper, http://csg.sph.umich.edu/boehnke/snipper/; DAPPLE, http://www.broadinstitute.org/mpg/dapple/dapple.php.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We especially thank the more than 196,000 volunteers who participated in our study. Detailed acknowledgment of funding sources is provided in the **Supplementary Note**.

AUTHOR CONTRIBUTIONS

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

All authors contributed to the research and reviewed the manuscript.

Design, management and coordination of contributing cohorts: (ADVANCE) T.L.A.; (AGES Reykjavik study) T.B.H. and V.G.; (AIDHS/SDS) D.K.S.; (AMC-PAS) P.D. and G.K.H.; (Amish GLGC) A.R.S.; (ARIC) E.B.; (B58C-WTCCC and B58C-T1DGC) D.P.S.; (B58C-Metabochip) C.M.L., C. Power and M.I.M.; (BLSA) L.F.; (BRIGHT) P.B.M.; (CARDIOGRAM) N.S.; (CHS) B.M.P. and J.I.R.; (CLHNS) A.B.F., K.L.M. and L.S.A.; (CoLaus) P.V.; (CROATIA-Vis) C.H. and I.R.; (deCODE) K. Stefansson and U.T.; (DIAGEN) P.E.H.S. and S.R.B.; (DILGOM) S.R.;

© 2013 Nature America, Inc. All rights reserved.
npg

(DPS) M.U.; (DR's EXTRA) R.R.; (EAS) J.F.P.; (EGCUT (Estonian Genome Center of the University of Tartu)) A.M.; (ELY) N.J.W.; (ENGAGE) N.B.F.; (EPIC) N.J.W. and K.-T.K.; (EPIC_N_OBSET GWAS) E.H.Y.; (ERF) C.M.v.D.; (ESS (Erasmus Stroke Study)) P.J.K.; (Family Heart Study (FHS)) I.B.B.; (FBPP) A.C., R.S.C. and S.C.H.; (FENLAND) R.J.F.L. and N.J.W.; (FIN-D2D 27) A.K. and L.M.; (FINCAVAS) M. Kähönen; (Framingham) L.A.C., S. Kathiresan and J.M.O.; (FRISCII) A. Siegbahn and L.W.; (FUSION GWAS) K.L.M. and M. Boehnke; (FUSION stage 2) F.S.C., J.T. and J. Saramies; (GenomEUTwin) J.B.W., N.G.M., K.O.K., V.S., J. Kaprio, A.J., D.I.B., N.L.P. and T.D.S.; (GLACIER) P.W.F., G.H.; (Go-DARTS) A.D.M. and C.N.A.P.; (GxE/Spanish Town) B.O.T., C.A.M., F.B., J.N.H. and R.S.C.; (HUNT2) K. Hveem; (IMPROVE) U.d.F., A. Hamsten, E.T. and S.E.H.; (InCHIANTI) S.B.; (KORAF4) C.G.; (LifeLines) B.H.R.W.; (LOLIPOP) J.S.K. and J.C.C.; (LURIC) B.O.B. and W.M.; (MDC) L.C.G. and S. Kathiresan; (MEDSTAR) M.S.B., S.E.E.; (METSIM) J. Kuusisto and M.L.; (MICROS) P.P.P.; (MORGAM) D. Arveiler and J.F.; (MRC/UVRI GPC GWAS) P. Kaleebu, G.A., J. Seeley and E.H.Y.; (MRC National Survey of Health and Development) D.K.; (NFBC1986) M.-R.J.; (NSPHS) U.G.; (ORCADES) H.C.; (PARC) Y.-D.I.C., R.M.K. and J.I.R.; (PennCath) D.J.R. and M.P.R.; (PIVUS) E.I. and L.L.; (PROMIS) J.D., P.D. and D. Saleheen; (Rotterdam Study) A. Hofman and A.G.U.; (SardiNIA) G.R.A.; (SCARFSHEEP) A. Hamsten and U.d.F.; (SEYCHELLES) M. Burnier, M. Bochud and P. Bovet; (SUVIMAX) P.M.; (Swedish Twin Registry) E.I. and N.L.P.; (TAICHI) T.L.A., Y.-D.I.C., C.A.H., T.Q., J.I.R. and W.H.-H.S.; (THISEAS) G.D. and P.D.; (Tromsø) I.N.; (TWINGENE) U.d.F. and E.I.; (ULSAM) E.I.; and (Whitehall II) A. Hingorani and M. Kivimaki.

Genotyping of contributing cohorts: (ADVANCE) D. Absher; (AIDHS/SDS) L.F.B. and M.L.G.; (AMC-PAS) P.D. and G.K.H.; (B58C-WTCCC and B58C-T1DGC) W.L.M.; (B58C-Metabochip) M.I.M.; (BLSA) D.H.; (BRIGHT) P.B.M.; (CHS) J.I.R.; (DIAGEN) N.N. and G.M.; (DILGOM) A. Palotie; (DR's EXTRA) T.A.L.; (EAS) J.F.W.; (EGCUT (Estonian Genome Center of the University of Tartu)) T.E.; (EPIC) P.D.; (EPIC_N_SUBCOH GWAS) I.B.; (ERF) C.M.v.D.; (ESS (Erasmus Stroke Study)) C.M.v.D.; (FBPP) A.C. and G.B.E.; (FENLAND) M.S.S.; (FIN-D2D 27) A.J.S.; (FINCAVAS) T.L.; (Framingham) J.M.O.; (FUSION stage 2) L.L.B.; (GLACIER) I.B.; (Go-DARTS) C.J.G., C.N.A.P. and M.I.M.; (IMPROVE) A. Hamsten; (KORAF3) H.G. and T.I.; (KORAF4) N.K.; (LifeLines) C.W.; (LOLIPOP) J.S.K. and J.C.C.; (LURIC) M.E.K.; (MDC) B.F.V. and R.D.; (MICROS) A.A.H.; (MORGAM) L.T. and P. Brambilla; (MRC/UVRI GPC GWAS) M.S.S.; (MRC National Survey of Health and Development) A.W., D.K. and K.K.O.; (NFBC1986) A.-L.H., M.-R.J., M.M., P.E. and S.V.; (NSPHS and FRISCII) Å.J.; (ORCADES) H.C.; (PARC) M.O.G., M.R.J. and J.I.R.; (PIVUS) E.I. and L.L.; (PROMIS) P.D. and K. Stirrups; (Rotterdam Study) A.G.U. and F.R.; (SardiNIA) R.N.; (SCARFSHEEP) B.G. and R.J.S.; (SEYCHELLES) F.M. and G.B.E.; (Swedish Twin Registry) E.I. and N.L.P.; (TAICHI) D. Absher, T.L.A., E.K., T.Q. and L.L.W.; (THISEAS) P.D.; (TWINGENE) A. Hamsten and E.I.; (ULSAM) E.I.; (WGHS) D.I.C., S.M. and P.M.R.; and (Whitehall II) A. Hingorani, C.L., M. Kumari and M. Kivimaki.

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

Primary analysis from contributing cohorts: (ADVANCE) L.L.W.; (AIDHS/SDS) R.S.; (AMC-PAS) S. Kanoni; (Amish GLGC) J.R.O. and M.E.M.; (ARIC) K.A.V.; (B58C-Metabochip) C.M.L., E.H. and T.F.; (B58C-WTCCC and B58C-T1DGC) D.P.S.; (BLSA) T.T.; (BRIGHT) T.J.; (CLHNS) Y.W.; (CoLaus) J.S.B.; (deCODE) G.T.; (DIAGEN) A.U.J.; (DILGOM) M.P.; (EAS) R.M.F.; (DPS) A.U.J.; (DR's EXTRA) A.U.J.; (EGCUT (Estonian Genome Center of the University of Tartu)) E.M., K.F. and T.E.; (ELY) D.G.; (EPIC) K. Stirrups and D.G.; (EPIC_N_OBSET GWAS) E.H.Y. and C.L.; (EPIC_N_SUBCOH GWAS) N.W.; (ERF) A.I.; (ESS (Erasmus Stroke Study)) C.M.v.D. and E.G.V.d.H.; (EUROSPAN) A. Demirkan; (Family Heart Study (FHS)) I.B.B. and M.F.F.; (FBPP) A.C. and G.B.E.; (FENLAND) T.P. and C. Pomilla; (FENLAND GWAS) J.H.Z. and J. Luan; (FIN-D2D 27) A.U.J.; (FINCAVAS) L.-P.L.; (Framingham) L.A.C. and G.M.P.; (FRISCII and NSPHS) Å.J.; (FUSION stage 2) T.M.T.; (GenomEUTwin-FINRISK) J. Kettunen; (GenomEUTwin-FINTWIN) K. Heikkilä; (GenomEUTwin-GENMETS) I.S.; (GenomEUTwin-SWETWIN) P.K.E.M.; (GenomEUTwin-UK-TWINS) M.M.; (GLACIER) D. Shungin; (GLACIER) P.W.F.; (Go-DARTS) C.N.A.P. and L.A.D.; (GxE/Spanish Town) C.D.P.; (HUNT) A.U.J.; (IMPROVE) R.J.S.; (InCHIANTI) T.T.; (KORAF3) M.M.-N.; (KORAF4) A.-K.P.; (LifeLines) I.M.N.; (LOLIPOP) W.Z.; (LURIC) M.E.K.; (MDC) B.F.V.; (MDC) P.F. and R.D.; (METSIM) A.U.J.; (MRC/UVRI GPC GWAS) R.N.N.; (MRC National Survey of Health and Development) A.W. and J. Luan; (NFBC1986) M. Kaakinen, I.S. and S.K.S.; (NSPHS and FRISCII) Å.J.; (PARC) X.L.; (PIVUS) C. Song and E.I.; (PROMIS) J.D., D.F.F. and K. Stirrups; (Rotterdam Study) A.I.; (SardiNIA) C. Sidore, J.L.B.-G. and S. Sanna; (SCARFSHEEP) R.J.S.; (SEYCHELLES) G.B.E. and M. Bochud; (SUVIMAX) T.J.; (Swedish Twin Registry) C. Song and E.I.; (TAICHI) D. Absher, T.L.A., H.-Y.C., M.O.G., C.A.H., T.Q. and L.L.W.; (THISEAS) S. Kanoni; (Tromsø) A.U.J.; (TWINGENE) A.G. and E.I.; (ULSAM) C. Song, E.I. and S.G.; (WGHS) D.I.C.; and (Whitehall II) S. Shah.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

© 2013 Nature America, Inc. All rights reserved.

1. Kannel, W.B., Dawber, T.R., 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.* **55**, 33–50 (1961).
2. Castelli, W.P. Cholesterol and lipids in the risk of coronary artery disease—the Framingham Heart Study. *Can. J. Cardiol.* **4** (suppl. A), 5A–10A (1988).
3. Lloyd-Jones, D. *et al.* Heart disease and stroke statistics—2010 update: a report from the American Heart Association. *Circulation* **121**, e46–e215 (2010).
4. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
5. Barter, P.J. & Rye, K.A. Cholesteryl ester transfer protein (CETP) inhibition as a strategy to reduce cardiovascular risk. *J. Lipid Res.* **53**, 1755–1766 (2012).
6. Rahalkar, A.R. & Hegele, R.A. Monogenic pediatric dyslipidemias: classification, genetics and clinical spectrum. *Mol. Genet. Metab.* **93**, 282–294 (2008).
7. Musunuru, K. *et al.* From noncoding variant to phenotype via *SORT1* at the 1p13 cholesterol locus. *Nature* **466**, 714–719 (2010).
8. Voight, B.F. *et al.* The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet.* **8**, e1002793 (2012).
9. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
10. Sanna, S. *et al.* Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. *PLoS Genet.* **7**, e1002198 (2011).
11. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
12. Asselbergs, F.W. *et al.* Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am. J. Hum. Genet.* **91**, 823–838 (2012).
13. Welch, C.L. *et al.* Genetic regulation of cholesterol homeostasis: chromosomal organization of candidate genes. *J. Lipid Res.* **37**, 1406–1421 (1996).
14. Sarria, A.J., Panini, S.R. & Evans, R.M. A functional role for vimentin intermediate filaments in the metabolism of lipoprotein-derived cholesterol in human SW-13 cells. *J. Biol. Chem.* **267**, 19455–19463 (1992).
15. Hagberg, C.E. *et al.* Vascular endothelial growth factor B controls endothelial fatty acid uptake. *Nature* **464**, 917–921 (2010).
16. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
17. Segrè, A.V., Groop, L., Mootha, V.K., Daly, M.J. & Altshuler, D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* **6**, e1001058 (2010).
18. Fitzgerald, M.L., Moore, K.J. & Freeman, M.W. Nuclear hormone receptors and cholesterol trafficking: the orphans find a new home. *J. Mol. Med. (Berl.)* **80**, 271–281 (2002).
19. Rossin, E.J. *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet.* **7**, e1001273 (2011).
20. Plyte, S.E., Hughes, K., Nikolakaki, E., Pulverer, B.J. & Woodgett, J.R. Glycogen synthase kinase-3: functions in oncogenesis and development. *Biochim. Biophys. Acta* **1114**, 147–162 (1992).

21. Toker, A. & Cantley, L.C. Signalling through the lipid products of phosphoinositide-3-OH kinase. *Nature* **387**, 673–676 (1997).
22. Kaprio, J., Ferrell, R.E., Kottke, B.A., Kamboh, M.I. & Sing, C.F. Effects of polymorphisms in apolipoproteins E, A-IV, and H on quantitative traits related to risk for cardiovascular disease. *Arterioscler. Thromb.* **11**, 1330–1348 (1991).
23. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49 (2011).
24. The ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **9**, e1001046 (2011).
25. Buyske, S. *et al.* Evaluation of the metabochip genotyping array in African Americans and implications for fine mapping of GWAS-identified loci: the PAGE study. *PLoS ONE* **7**, e35651 (2012).
26. Palmen, J. *et al.* The functional interaction on *in vitro* gene expression of *APOA5* SNPs, defining haplotype *APOA52*, and their paradoxical association with plasma triglyceride but not plasma apoAV levels. *Biochim. Biophys. Acta* **1782**, 447–452 (2008).
27. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43**, 333–338 (2011).
28. Coronary Artery Disease (C4D) Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat. Genet.* **43**, 339–344 (2011).
29. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).
30. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
31. Heid, I.M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42**, 949–960 (2010).
32. Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).
33. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
34. Freathy, R.M. *et al.* Common variation in the *FTO* gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. *Diabetes* **57**, 1419–1426 (2008).
35. Clarke, R. *et al.* Cholesterol fractions and apolipoproteins as risk factors for heart disease mortality in older men. *Arch. Intern. Med.* **167**, 1373–1378 (2007).
36. Willer, C.J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* **40**, 161–169 (2008).
37. Voight, B.F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* **380**, 572–580 (2012).
38. Frikke-Schmidt, R. *et al.* Association of loss-of-function mutations in the *ABCA1* gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *J. Am. Med. Assoc.* **299**, 2524–2532 (2008).
39. Do, R. *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat. Genet.* doi:10.1038/ng.2795 (6 October 2013).
40. Demirkan, A. *et al.* Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet.* **8**, e1002490 (2012).

© 2013 Nature America, Inc. All rights reserved.

Cristen J Willer[1–4,207,208], Ellen M Schmidt[2,207], Sebanti Sengupta[4,207], Gina M Peloso[5–7], Stefan Gustafsson[8,9], Stavroula Kanoni[10], Andrea Ganna[8,9,11], Jin Chen[4], Martin L Buchkovich[12], Samia Mora[13,14], Jacques S Beckmann[15,16], Jennifer L Bragg-Gresham[4], Hsing-Yi Chang[17], Ayşe Demirkan[18], Heleen M Den Hertog[19], Ron Do[6], Louise A Donnelly[20], Georg B Ehret[21,22], Tõnu Esko[7,23,24], Mary F Feitosa[25], Teresa Ferreira[26], Krista Fischer[23], Pierre Fontanillas[7], Ross M Fraser[27], Daniel F Freitag[28], Deepti Gurdasani[10,28], Kauko Heikkilä[29], Elina Hyppönen[30], Aaron Isaacs[18,31], Anne U Jackson[4], Åsa Johansson[32,33], Toby Johnson[34,35], Marika Kaakinen[36,37], Johannes Kettunen[38,39], Marcus E Kleber[40,41], Xiaohui Li[42], Jian'an Luan[43], Leo-Pekka Lyytikäinen[44,45], Patrik K E Magnusson[11], Massimo Mangino[46], Evelin Mihailov[23,24], May E Montasser[47], Martina Müller-Nurasyid[48–50], Ilja M Nolte[51], Jeffrey R O'Connell[47], Cameron D Palmer[7,52,53], Markus Perola[23,38,39], Ann-Kristin Petersen[48], Serena Sanna[54], Richa Saxena[6], Susan K Service[55], Sonia Shah[56], Dmitry Shungin[57–59], Carlo Sidore[4,54,60], Ci Song[8,9,11], Rona J Strawbridge[61,62], Ida Surakka[38,39], Toshiko Tanaka[63], Tanya M Teslovich[4], Gudmar Thorleifsson[64], Evita G Van den Herik[19], Benjamin F Voight[65,66], Kelly A Volcik[67], Lindsay L Waite[68], Andrew Wong[69], Ying Wu[12], Weihua Zhang[70,71], Devin Absher[68], Gershim Asiki[72], Inês Barroso[10,73,74], Latonya F Been[75], Jennifer L Bolton[27], Lori L Bonnycastle[76], Paolo Brambilla[77], Mary S Burnett[78], Giancarlo Cesana[79], Maria Dimitriou[80], Alex S F Doney[20], Angela Döring[81,82], Paul Elliott[37,83], Stephen E Epstein[78], Gudmundur Ingi Eyjolfsson[84], Bruna Gigante[85], Mark O Goodarzi[86], Harald Grallert[87], Martha L Gravito[75], Christopher J Groves[88], Göran Hallmans[89], Anna-Liisa Hartikainen[90], Caroline Hayward[91], Dena Hernandez[92], Andrew A Hicks[93], Hilma Holm[64], Yi-Jen Hung[94], Thomas Illig[87,95], Michelle R Jones[86], Pontiano Kaleebu[72], John J P Kastelein[96], Kay-Tee Khaw[97], Eric Kim[42], Norman Klopp[87,95], Pirjo Komulainen[98], Meena Kumari[56], Claudia Langenberg[43], Terho Lehtimäki[44,45], Shih-Yi Lin[99], Jaana Lindström[100], Ruth J F Loos[43,101–103], François Mach[21], Wendy L McArdle[104], Christa Meisinger[81], Braxton D Mitchell[47], Gabrielle Müller[105], Ramaiah Nagaraja[106], Narisu Narisu[76], Tuomo V M Nieminen[107–109], Rebecca N Nsubuga[72], Isleifur Olafsson[110], Ken K Ong[43,69], Aarno Palotie[38,111,112], Theodore Papamarkou[10,28,113], Cristina Pomilla[10,28], Anneli Pouta[90,114], Daniel J Rader[115,116], Muredach P Reilly[115,116], Paul M Ridker[13,14], Fernando Rivadeneira[117–119], Igor Rudan[27], Aimo Ruokonen[120], Nilesh Samani[121,122], Hubert Scharnagl[123], Janet Seeley[72,124], Kaisa Silander[38,39], Alena Stancáková[125], Kathleen Stirrups[10], Amy J Swift[76], Laurence Tiret[126], Andre G Uitterlinden[117–119], L Joost van Pelt[127,128], Sailaja Vedantam[7,52,53], Nicholas Wainwright[10,28], Cisca Wijmenga[128,129], Sarah H Wild[27], Gonneke Willemsen[130], Tom Wilsgaard[131], James F Wilson[27], Elizabeth H Young[10,28], Jing Hua Zhao[43], Linda S Adair[132], Dominique Arveiler[133], Themistocles L Assimes[134], Stefania Bandinelli[135], Franklyn Bennett[136], Murielle Bochud[137], Bernhard O Boehm[138,139], Dorret I Boomsma[130], Ingrid B Borecki[25], Stefan R Bornstein[140], Pascal Bovet[137,141], Michel Burnier[142], Harry Campbell[27], Aravinda Chakravarti[22], John C Chambers[70,71,143], Yii-Der Ida Chen[144,145], Francis S Collins[76], Richard S Cooper[146], John Danesh[28], George Dedoussis[80], Ulf de Faire[85], Alan B Feranil[147], Jean Ferrières[148], Luigi Ferrucci[63], Nelson B Freimer[55,149], Christian Gieger[48], Leif C Groop[150,151], Vilmundur Gudnason[152], Ulf Gyllensten[32], Anders Hamsten[61,62,153], Tamara B Harris[154], Aroon Hingorani[56], Joel N Hirschhorn[7,52,53], Albert Hofman[117,119], G Kees Hovingh[96],

Chao Agnes Hsiung[155], Steve E Humphries[156], Steven C Hunt[157], Kristian Hveem[158], Carlos Iribarren[159], Marjo-Riitta Järvelin[36,37,83,114,160], Antti Jula[161], Mika Kähönen[162], Jaakko Kaprio[29,38,163], Antero Kesäniemi[164], Mika Kivimaki[56], Jaspal S Kooner[71,143,165], Peter J Koudstaal[19], Ronald M Krauss[166], Diana Kuh[69], Johanna Kuusisto[167], Kirsten O Kyvik[168,169], Markku Laakso[167], Timo A Lakka[98,170], Lars Lind[171], Cecilia M Lindgren[26], Nicholas G Martin[172], Winfried März[41,123,173], Mark I McCarthy[26,88], Colin A McKenzie[174], Pierre Meneton[175], Andres Metspalu[23,24], Leena Moilanen[176], Andrew D Morris[20], Patricia B Munroe[34,35], Inger Njølstad[131], Nancy L Pedersen[11], Chris Power[30], Peter P Pramstaller[93,177,178], Jackie F Price[27], Bruce M Psaty[179,180], Thomas Quertermous[134], Rainer Rauramaa[98,181], Danish Saleheen[28,182,183], Veikko Salomaa[184], Dharambir K Sanghera[75], Jouko Saramies[185], Peter E H Schwarz[140,186], Wayne H-H Sheu[187], Alan R Shuldiner[47,188], Agneta Siegbahn[8,33,171], Tim D Spector[46], Kari Stefansson[64,189], David P Strachan[190], Bamidele O Tayo[146], Elena Tremoli[191], Jaakko Tuomilehto[100,192–194], Matti Uusitupa[195,196], Cornelia M van Duijn[18,31], Peter Vollenweider[197], Lars Wallentin[33,171], Nicholas J Wareham[43], John B Whitfield[172], Bruce H R Wolffenbuttel[128,198], Jose M Ordovas[199–201], Eric Boerwinkle[67], Colin N A Palmer[20], Unnur Thorsteinsdottir[64,189], Daniel I Chasman[13,14], Jerome I Rotter[42], Paul W Franks[57,59,202], Samuli Ripatti[10,38,39], L Adrienne Cupples[5,203], Manjinder S Sandhu[10,28], Stephen S Rich[204], Michael Boehnke[4,208], Panos Deloukas[10,208], Sekar Kathiresan[6,7,205,206,208], Karen L Mohlke[12,208], Erik Ingelsson[8,9,26,208] & Gonçalo R Abecasis[4,208]

[1]Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, Michigan, USA. [2]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA. [3]Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA. [4]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA. [5]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [6]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [7]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. [8]Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala, Sweden. [9]Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [10]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [11]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [12]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA. [13]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [14]Harvard Medical School, Boston, Massachusetts, USA. [15]Service of Medical Genetics, Lausanne University Hospital, Lausanne, Switzerland. [16]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland. [17]Division of Preventive Medicine and Health Services Research, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan. [18]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands. [19]Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands. [20]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School. Dundee, UK. [21]Cardiology, Department of Specialities of Medicine, Geneva University Hospital, Geneva, Switzerland. [22]Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [23]Estonian Genome Center of the University of Tartu, Tartu, Estonia. [24]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [25]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA. [26]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [27]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK. [28]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [29]Hjelt Institute, Department of Public Health, University of Helsinki, Helsinki, Finland. [30]Centre for Paediatric Epidemiology and Biostatistics/Medical Research Council (MRC) Centre of Epidemiology for Child Health, University College London Institute of Child Health, London, UK. [31]Centre for Medical Systems Biology, Leiden, The Netherlands. [32]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [33]Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden. [34]Genome Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [35]Clinical Pharmacology, National Institute for Health Research (NIHR) Cardiovascular Biomedical Research Unit, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [36]Biocenter Oulu, University of Oulu, Oulu, Finland. [37]Institute of Health Sciences, University of Oulu, Oulu, Finland. [38]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [39]Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland. [40]Department of Internal Medicine II–Cardiology, University of Ulm Medical Centre, Ulm, Germany. [41]Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany. [42]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [43]MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Hills Road, Cambridge, UK. [44]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland. [45]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland. [46]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [47]Department of Medicine, Division of Endocrinology, Diabetes and Nutrition, University of Maryland, School of Medicine, Baltimore, Maryland, USA. [48]Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany. [49]Department of Medicine I, University Hospital Grosshadern, Ludwig Maximilians University of Munich, Munich, Germany. [50]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig Maximilians University of Munich, Munich, Germany. [51]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [52]Division of Endocrinology, Children's Hospital Boston, Boston, Massachusetts, USA. [53]Division of Genetics, Program in Genomics, Children's Hospital, Boston, Massachusetts, USA. [54]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Monserrato, Italy. [55]Center for Neurobehavioral Genetics, The Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, California, USA. [56]Genetic Epidemiology Group, Department of Epidemiology and Public Health, University College London, London, UK. [57]Department of Clinical Sciences, Genetic & Molecular Epidemiology Unit, Lund University Diabetes Center, Scania University Hospital, Malmö, Sweden. [58]Department of Odontology, Umeå University, Umeå, Sweden. [59]Department of Public Health and Primary Care, Unit of Medicine, Umeå University, Umeå, Sweden. [60]Dipartimento di Scienze Biomediche, Universita di Sassari, Sassari, Italy. [61]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden. [62]Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. [63]Clinical Research Branch, US National Institutes of Health, Baltimore, Maryland, USA. [64]deCODE Genetics/Amgen, Reykjavik, Iceland. [65]Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. [66]Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. [67]Human Genetics Center, University of Texas Health Science Center, School of Public Health, Houston, Texas, USA. [68]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA. [69]MRC Unit for Lifelong Health and Ageing, London, UK. [70]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [71]Ealing Hospital, Southall, UK. [72]MRC/Uganda Virus Research Institute Uganda Research Unit on AIDS, Entebbe, Uganda. [73]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [74]NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [75]Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA. [76]Genome Technology Branch, National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. [77]Department of Experimental Medicine, University of Milano-Bicocca, Milan, Italy. [78]MedStar Health Research Institute, Hyattsville, Maryland, USA. [79]Research Centre on Public Health, University of Milano-Bicocca, Milan, Italy. [80]Department of Dietetics-Nutrition, Harokopio University, Athens, Greece. [81]Institute of Epidemiology I, Helmholtz Zentrum München, Neuherberg, Germany. [82]Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany. [83]Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment

© 2013 Nature America, Inc. All rights reserved.

npg

and Health, School of Public Health, Imperial College London, London, UK. [84]The Laboratory in Mjodd, Reykjavik, Iceland. [85]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [86]Department of Medicine, Division of Endocrinology, Diabetes and Metabolism, Cedars-Sinai Medical Center, Los Angeles, California, USA. [87]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany. [88]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [89]Department of Public Health and Clinical Medicine, Nutritional Research, Umeå University, Umeå, Sweden. [90]Department of Clinical Sciences/Obstetrics and Gynecology, Oulu University Hospital, Oulu, Finland. [91]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK. [92]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, Maryland, USA. [93]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy (Affiliated Institute of the University of Lübeck). [94]Division of Endocrinology & Metabolism, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan. [95]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany. [96]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands. [97]Clinical Gerontology Unit, University of Cambridge, Cambridge, UK. [98]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland. [99]Department of Internal Medicine, Division of Endocrine and Metabolism, Taichung Veterans General Hospital, School of Medicine, National Yang-Ming University, Taipei, Taiwan. [100]Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland. [101]The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, New York, USA. [102]The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, New York, USA. [103]The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York, New York, USA. [104]School of Social and Community Medicine, University of Bristol, Bristol, UK. [105]Institute for Medical Informatics and Biometrics, University of Dresden, Medical Faculty Carl Gustav Carus, Dresden, Germany. [106]Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA. [107]Department of Clinical Pharmacology, University of Tampere School of Medicine, Tampere, Finland. [108]Department of Internal Medicine, Päijät-Häme Central Hospital, Lahti, Finland. [109]Division of Cardiology, Helsinki University Central Hospital, Helsinki, Finland. [110]Department of Clinical Biochemistry, Landspitali University Hospital, Reykjavik, Iceland. [111]Department of Medical Genetics, Haartman Institute, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland. [112]Genetic Epidemiology Group, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. [113]Department of Statistical Sciences, University College London, London, UK. [114]National Institute for Health and Welfare, Oulu, Finland. [115]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, Translational Research Center, Philadelphia, Pennsylvania, USA. [116]Division of Translational Medicine and Human Genetics, Perelman School of Medicine at the University of Pennsylvania, Translational Research Center, Philadelphia, Pennsylvania, USA. [117]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands. [118]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands. [119]Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), Leiden, The Netherlands. [120]Department of Clinical Sciences/Clinical Chemistry, University of Oulu, Oulu, Finland. [121]National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK. [122]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK. [123]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria. [124]School of International Development, University of East Anglia, Norwich, UK. [125]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. [126]INSERM UMRS 937, Pierre and Marie Curie University, Paris, France. [127]Department of Laboratory Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [128]LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [129]Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [130]Department of Biological Psychology, VU University, Amsterdam, The Netherlands. [131]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. [132]Department of Nutrition, University of North Carolina, Chapel Hill, North Carolina, USA. [133]Department of Epidemiology and Public Health, University of Strasbourg, Faculty of Medicine, Strasbourg, France. [134]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. [135]Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy. [136]Chemical Pathology, Department of Pathology, University of the West Indies, Mona, Jamaica. [137]Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Lausanne, Switzerland. [138]Department of Internal Medicine, Division of Endocrinology and Diabetes, Ulm University Medical Centre, Ulm, Germany. [139]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. [140]Department of Medicine III, University of Dresden, Medical Faculty Carl Gustav Carus, Dresden, Germany. [141]Ministry of Health, Victoria, Republic of Seychelles. [142]Service of Nephrology, Lausanne University Hospital, Lausanne, Switzerland. [143]Imperial College Healthcare National Health Service (NHS) Trust, London, UK. [144]Division of Reproductive Endocrinology, Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Los Angeles, California, USA. [145]Department of Medicine, University of California, Los Angeles, Los Angeles, California, USA. [146]Department of Preventive Medicine and Epidemiology, Loyola University Medical School, Maywood, Illinois, USA. [147]Office of Population Studies Foundation, University of San Carlos, Talamban, Cebu City, Philippines. [148]Department of Cardiology, Toulouse University School of Medicine, Rangueil Hospital, Toulouse, France. [149]Department of Psychiatry, University of California, Los Angeles, Los Angeles, California, USA. [150]Department of Clinical Sciences, Lund University, Malmö, Sweden. [151]Department of Medicine, Helsinki University Hospital, Helsinki, Finland. [152]Icelandic Heart Association, Kopavogur, Iceland. [153]Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden. [154]Laboratory of Epidemiology, Demography and Biometry, National Institute on Ageing, Bethesda, Maryland, USA. [155]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan. [156]Cardiovascular Genetics, British Heart Foundation Laboratories, Institute Cardiovascular Science, University College London, London, UK. [157]Cardiovascular Genetics, University of Utah School of Medicine, Salt Lake City, Utah, USA. [158]Nord-Trøndelag Health Study (HUNT) Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway. [159]Kaiser Permanente, Division of Research, Oakland, California, USA. [160]Unit of Primary Care, Oulu University Hospital, Oulu, Finland. [161]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Turku, Finland. [162]Department of Clinical Physiology, University of Tampere School of Medicine, Tampere, Finland. [163]Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland. [164]Institute of Clinical Medicine, Department of Medicine, University of Oulu and Clinical Research Center, Oulu University Hospital, Oulu, Finland. [165]National Heart & Lung Institute, Imperial College London, Hammersmith Hospital, London, UK. [166]Children's Hospital Oakland Research Institute, Oakland, California, USA. [167]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. [168]Institute of Regional Health Services Research, University of Southern Denmark, Odense, Denmark. [169]Odense Patient Data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark. [170]Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Kuopio, Finland. [171]Department of Medical Sciences, Uppsala University, Uppsala, Sweden. [172]Queensland Institute of Medical Research, Royal Brisbane Hospital, Brisbane, Queensland, Australia. [173]Synlab Academy, Synlab Services, Mannheim, Germany. [174]Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, Jamaica. [175]INSERM U872, Centre de Recherche des Cordeliers, Paris, France. [176]Department of Medicine, Kuopio University Hospital, Kuopio, Finland. [177]Department of Neurology, General Central Hospital, Bolzano, Italy. [178]Department of Neurology, University of Lübeck, Lübeck, Germany. [179]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, Washington, USA. [180]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA. [181]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. [182]Center for Non-Communicable Diseases, Karachi, Pakistan. [183]Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [184]Unit of Chronic Disease Epidemiology and Prevention, National Institute for Health and Welfare, Helsinki, Finland. [185]South Karelia Central Hospital, Lappeenranta, Finland. [186]Paul Langerhans Institute Dresden, German Center for Diabetes Research (DZD), Dresden, Germany. [187]Department of Internal Medicine, Division of Endocrine and Metabolism, Taichung Veterans General Hospital, Taichung, Taiwan. [188]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland, USA. [189]Faculty of Medicine, University of Iceland, Reykjavík, Iceland. [190]Division of Population Health Sciences and Education, St. George's, University of London, London, UK. [191]Department of Pharmacological Sciences, University of Milan, Monzino Cardiology Center, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Milan, Italy. [192]Centre for Vascular Prevention, Danube University Krems, Krems, Austria. [193]King Abdulaziz University, Faculty of Medicine, Jeddah, Saudi Arabia. [194]La Red Temática de Investigación Cooperativa en Enfermedades Cardiovasculares (RECAVA) Grupo RD06/0014/0015, Hospital Universitario La Paz, Madrid, Spain. [195]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland. [196]Research Unit, Kuopio University Hospital, Kuopio, Finland. [197]Department of Medicine, Lausanne University Hospital, Lausanne, Switzerland. [198]Department of Endocrinology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [199]Department of Cardiovascular Epidemiology and Population Genetics, National Center for Cardiovascular Investigation, Madrid, Spain. [200]Madrid Institute for Advanced Studies (Instituto Madrileño de Estudios Avanzados)–Alimentacion, Madrid, Spain. [201]Nutrition and Genomics Laboratory, Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts, USA. [202]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [203]Framingham Heart Study, Framingham, Massachusetts, USA. [204]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA. [205]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA. [206]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [207]These authors contributed equally to this work. [208]These authors jointly directed this work. Correspondence should be addressed to C.J.W. (cristen@umich.edu), K.L.M. (mohlke@med.unc.edu), E.I. (erik.ingelsson@medsci.uu.se) or G.R.A. (goncalo@umich.edu).

© 2013 Nature America, Inc. All rights reserved.
npg

ONLINE METHODS

Samples studied. We collected summary statistics for Metabochip SNPs from 45 studies. Of these, 37 studies consisted primarily of individuals of European ancestry (see **Supplementary Table 1** and the **Supplementary Note** for details), including both population-based studies and case-control studies of CAD and T2D. Another 8 studies consisted primarily of individuals with non-European ancestry, including 2 studies of individuals of South Asian descent, AIDHS/SDS (n = 1,516) and PROMIS (n = 3,385); 2 studies of individuals of East Asian descent, CLHNS (n = 1,771) and TAI-CHI (n = 7,044); and 5 studies of individuals of recent African ancestry, MRC/UVRI GPC (n = 1,687) from Uganda, SEY (n = 426) from the Caribbean, and FBPP (n = 1,614; triglyceride results unavailable), GXE (n = 397) and SPT (n = 838) from the United States (more details in **Supplementary Table 1** and the **Supplementary Note**). Each contributing study individually obtained ethics approval for their data generation and analyses.

Genotyping. We genotyped 196,710 genetic variants prioritized on the basis of previous GWAS for cardiovascular and metabolic phenotypes using the Illumina iSelect Metabochip[8] genotyping array. To design the Metabochip, we used our previous GWAS of ~100,000 individuals[4] to prioritize 5,023 SNPs for HDL cholesterol, 5,055 SNPs for LDL cholesterol, 5,056 SNPs for triglycerides and 938 SNPs for total cholesterol. These independent SNPs represent most loci with P < 0.005 in our original GWAS for HDL cholesterol, LDL cholesterol and triglycerides and with P < 0.0005 for total cholesterol. An additional 28,923 SNPs were selected for fine mapping of 65 previously identified lipid loci. The Metabochip also included 50,459 SNPs prioritized on the basis of GWAS of non-lipid traits and 93,308 SNPs selected for fine mapping of loci associated with non-lipid traits (5 of these loci were associated with blood lipids by the analyses described here).

Phenotypes. Blood lipid levels were typically measured after >8 h of fasting. Individuals known to be on lipid-lowering medication were excluded when possible. LDL cholesterol levels were directly measured in ten studies (24% of total study individuals) and were estimated using the Friedewald formula[41] in the remaining studies. Trait residuals within each study cohort were adjusted for age, age^2 and sex and were then quantile normalized. Explicit adjustments for population structure using principal-component[42] or mixed-model approaches[43] were carried out in 24 studies (35% of study individuals); all studies were adjusted using genomic control before meta-analysis[11]. In studies ascertained on diabetes or cardiovascular disease status, cases and controls were analyzed separately (**Supplementary Table 1**). All meta-analyses were limited to a single ancestry group (for example, European only).

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

Meta-analysis. Meta-analysis was performed using the Stouffer method[44,45] with weights proportional to the square root of the sample size for each sample. To correct for inflated test statistics due to potential population stratification, we first applied genomic control to each sample and then repeated the procedure with initial meta-analysis results. For GWAS samples, we used all available SNPs when estimating the median test statistic and inflation factor λ. For Metabochip samples, we used a subset of SNPs (n = 7,168) that had P values of >0.50 for all lipid traits in the original GWAS, expecting that the majority of these would not be associated with lipids and would behave as null variants in the Metabochip samples. Signals were considered to be novel if they reached a P value of $<5 \times 10^{-8}$ in the combined GWAS and Metabochip meta-analysis and were >1 Mb away from the nearest previously described lipid-associated locus and other new loci. We used only European samples for the discovery of new genome-wide significant loci. Non-European samples were used only for meta-analysis and examination of fine-mapping analyses.

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

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

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

Generating permuted sets of non-associated SNPs. To estimate the expected chance overlap between literature searches and our loci, we generated lists of permuted SNPs. To generate these lists, we first identified all non-associated lipid-related SNPs (P > 0.10 for any of the four lipid traits) and created bins on the basis of three statistics: MAF, distance to the nearest gene and number of SNPs with r^2 > 0.8. For each index SNP, we identified 500 non-lipid-associated SNPs that fell within the same 3 bins and randomly selected 1 SNP for each permuted list.

Pathway analyses. To investigate whether lipid-associated variants overlapped previously annotated pathways, we used gene set enrichment analysis (GSEA), as implemented in MAGENTA[17] using the meta-analysis of all studies, including GWAS and Metabochip SNPs. Briefly, MAGENTA first assigns SNPs to a given gene when within 110 kb upstream or 40 kb downstream of transcript boundaries. The most significant SNP P value within this interval is then adjusted for confounders (gene size, marker density and LD) to create a gene association score. When the same SNP is assigned to multiple genes, only the gene with the lowest score is kept for downstream analyses. Subsequently, MAGENTA attaches pathway terms to each gene using several annotation resources, including GO, PANTHER, Ingenuity and KEGG. Finally, the genes are ranked on the basis of their gene association scores, and a modified GSEA test is used to test the null hypothesis that all gene score ranks above a given rank cutoff are randomly distributed with regard to a given pathway term (and compared to multiple randomly sampled gene sets of identical size). We evaluated enrichment using a rank cutoff of 5% of the total number of genes. A minimum of 10,000 gene set permutations were performed, and up to 1,000,000 permutations were performed for GSEA P values below 1×10^{-4}.

We used the Disease Association Protein-Protein Link Evaluator package (DAPPLE) to examine evidence for protein-protein interaction networks connecting genes across different lipid-related loci. This analysis included the 62 new loci as well as the 95 previously known loci; we focus our discussion on pathways that included 1 or more genes from new loci.

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

© 2013 Nature America, Inc. All rights reserved.

 doi:10.1038/ng.2797

index SNPs and all transcripts within 500 kb of the SNP position, performing association analyses as previously described[47].

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

Association with lipid subfractions. Lipoprotein fractions for Women's Genome Health Study (WGHS) samples (n = 23,170) were measured using the LipoProtein-II assay (Liposcience), and Framingham Heart Study Offspring samples (n = 2,900) were measured with the LipoProtein-I assay (Liposcience)[48]. Additional information on subfraction measurements can be found in **Supplementary Figure 7**. Log transformations were used for non-normalized traits. All models were adjusted for age, sex and principal components. The genetic association analysis of WGHS used SNP genotypes imputed from the HapMap release 22 CEU (Utah residents of Northern and Western European ancestry) reference panel using MACH. Of the 23,170 WGHS participants, 16,730 were fasting for 8 h before blood draw (72.2%).

41. Friedewald, W.T., Levy, R.I. & Fredrickson, D.S. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* **18**, 499–502 (1972).
42. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
43. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348–354 (2010).
44. Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A. & Williams, R.M.J. *Adjustment During Army Life* (Princeton University Press, Princeton, NJ., 1949).
45. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
46. Keating, B.J. *et al.* Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS ONE* **3**, e3583 (2008).
47. Schadt, E.E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
48. Chasman, D.I. *et al.* Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet.* **5**, e1000730 (2009).

© 2013 Nature America, Inc. All rights reserved.

npg