Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1038/ncomms6068</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:13454864">http://nrs.harvard.edu/urn-3:HUL.InstRepos:13454864</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins


Statins effectively lower cholesterol levels in large studies and the observed interindividual response variability may be partially explained by genetic variation. Here we perform a pharmacogenetic meta-analysis of genome-wide association studies (GWAS) in studies addressing the LDL cholesterol response to statins, including up to 18,596 statin-treated subjects. We validate the most promising signals in a further 22,318 statin recipients and identify two loci, partially explained by genetic variation. Here we perform a pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins.
The 3-hydroxymethyl-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, also known as statins, are widely prescribed and are highly effective in the management and prevention of cardiovascular disease. Statin therapy results in a lowering of low-density lipoprotein cholesterol (LDL-C) levels by up to 55%\(^1\) and a 20–30% reduction of cardiovascular events\(^2\). Despite the clinical efficacy of statins in a wide range of patients\(^2\), interindividual variability exists with regard to LDL-C-lowering response as well as efficacy in reducing major cardiovascular events\(^3\). The suggestion that some of this variability may be due, in part, to common pharmacogenetic variation is supported by previous studies that have identified genetic variants associated with differential LDL-C response to statin therapy\(^4,6\).

A small number of genome-wide association studies (GWAS) have previously identified loci associated with statin response on a genome-wide level. A GWAS in the JUPITER trial identified three genetic loci, ABCG2 (rs2199936), LPA (rs10455872) and APOE (rs7412), that were associated with percentage LDL-C lowering effect compared with non-carriers. The minor allele of ABCG2 (rs445925 and rs4420638) were associated with LDL-C response to atorvastatin treatment\(^8\). A combined GWAS in three statin trials identified a SNP within chromosome 19, at \((rs445925, \text{minor allele frequency} = 0.098, \quad \beta = -0.043, \quad \text{s.e.} = 0.005, \quad P = 1.58 \times 10^{-18}; \quad \text{Fig. 2a})\), indicating that carriers of the rs445925 SNP respond to statins with an additional 4.3% increase per allele in LDL-C lowering effect compared with non-carriers. The second strongest association was with a SNP at LPA on chromosome 6 \((rs10455872, \quad \text{MAF} = 0.069, \quad \beta = 0.041, \quad \text{s.e.} = 0.006, \quad P = 1.95 \times 10^{-11}; \quad \text{Fig. 2b})\), indicating a 5.9% smaller LDL-C lowering per minor allele for carriers of the SNP compared with non-carriers. Associations at both loci have previously been described\(^7,8\). A third genome-wide significant association was found with a SNP at RICTOR on chromosome 5 \((rs13166647, \quad \text{MAF} = 0.230, \quad \beta = -0.253, \quad \text{s.e.} = 0.046, \quad P = 4.50 \times 10^{-8})\), although genotypes for this SNP were only available in two studies within the first stage \((n = 2,144)\).

**Second-stage meta-analysis.** We selected 246 SNPs with \(P < 5 \times 10^{-8}\) from 158 loci for further investigation in three additional studies comprising up to 22,318 statin-treated subjects (see Methods; Supplementary Tables 1 and 2; Supplementary Note 3). This second stage confirmed the genome-wide significant associations between variations within the APOE and LPA loci and LDL-C response, as observed in the first stage (Table 1; Supplementary Fig. 2; Supplementary Table 5). In addition, SNPs at two new loci with \(P\) values between 6.70 \(\times 10^{-7}\) and 2.26 \(\times 10^{-6}\) in the first phase were shown to be significantly associated with statin-induced LDL-C lowering after statin treatment in the total combined meta-analysis at a genome-wide level: \(\text{SORT1/CERS2/PSRC1} (rs646776, \quad \beta = -0.013, \quad \text{s.e.} = 0.002, \quad P = 1.05 \times 10^{-9});\) \(\text{rs12740374,} \quad \beta = -0.013, \quad \text{s.e.} = 0.002, \quad P = 1.05 \times 10^{-9})\).  

The six next-ranked SNPs with \(P\) values just below \(5 \times 10^{-8}\) in the combined meta-analysis, including the two SNPs at RICTOR \((rs13166647 \quad \text{and} \quad rs13172966)\), were selected for additional genotyping in the Scandinavian ASCOT participants (see Methods). None of these six SNPs reached genome-wide significance after this additional genotyping (Supplementary Table 6). Therefore, our overall genome-wide significant findings were the SNPs at APOE, LPA, \text{SORT1/CERS2/PSRC1} and \text{SLCO1B1}.

**Subfraction analyses.** To extend our results for the novel GWAS finding \text{SORT1/CERS2/PSRC1}, we performed additional association analyses, using measurements of cholesterol levels in four LDL subfractions (large, medium, small and very small) from two of the trials in GIST, CAP and PRINCE (Table 2; see Methods). The minor allele of \text{SORT1} rs646777 was associated with greater statin-induced reductions in levels of all LDL subfractions, and there was a nonsignificant trend for larger effect sizes and greater statistical significance for lowering of small and very small LDL (Table 2). In contrast, the APOE SNP associated with greater

---

**Figure 1 | Results of the GWAS meta-analysis.** Manhattan plot presenting the \(-\log_{10} P\) values from the combined meta-analysis \((n = 40,914)\) on LDL-C response after statin treatment. \(P\) values were generated using linear regression analysis.
LDL-C response to statins (rs445925) showed a small and non-significant association with change in very small LDL (Table 2). For the minor allele of rs2900478 (SLCO1B1), the borderline significant association with smaller magnitude of LDL-C reduction showed a trend for preferential association with larger versus smaller LDL subfractions. The lack of association of rs10455872 (LPA) with changes in LDL subfractions is consistent with evidence discussed below that this locus affects levels of lipoprotein(a) (Lp(a)) and not LDL particles. Using generalized estimating equations, we tested the association of log change in each of the LDL subfractions with interactions of the four SNPs. For very small LDL, the association with the rs646776 minor allele was significantly different from that of the other minor alleles ($P = 0.03$ after adjustment for multiple testing).

**Effects of off-treatment LDL-C.** To demonstrate that our findings for LDL-C response to statin treatment are unlikely to be explained through associations with baseline LDL-C levels, we performed a number of additional analyses (see Methods). First, Supplementary Table 7 shows regression coefficients for baseline- and modelling measurement noise at baseline reduced the evidence discussed below that this locus affects levels of lipoprotein(a) (Lp(a)) and not LDL particles. Using generalized estimating equations, we tested the association of log change in each of the LDL subfractions with interactions of the four SNPs. For very small LDL, the association with the rs646776 minor allele was significantly different from that of the other minor alleles ($P = 0.03$ after adjustment for multiple testing).

**Functional analyses.** Functional characterization of the 246 SNPs selected for the second stage was performed using a range of bioinformatics tools (see Methods). A total of 420 expression quantitative trait loci (eQTL) associations were identified across a wide range of tissues (Supplementary Data 1), which comprised 67 independent gene eQTL associations. Eleven genes, including APOE, SORTI, CELSR2 and PSRC1, showed eQTLs in liver, which considering its primary role in mediating statin-induced LDL reduction may be particularly relevant to statin response. Putative gene eQTLs were combined with genes annotated to variants in linkage disequilibrium (LD) with LDL-C response-associated variants, resulting in a list of 185 candidate gene loci, defined by 2,681 SNPs (Supplementary Data 2 and 3). To identify statin responsive genes among the candidate loci, gene expression data measured in response to statin treatment in a range of cell lines was retrieved from the Connectivity Map resource (12) (see Methods). Five genes (APOE, BRCA1, GPR1L, ADHR2 and ETV1) showed convincing evidence of statin responsiveness on the basis of greater than two-fold differential expression in response to statin treatment. Eight genes showed suggestive evidence (1.5- to 2-fold change; TOMM40, SREBP1, PSRC1, BCL3, BCAM, ANK3, SIV1 and RANBP9; Supplementary Data 3). Finally, involvement in statin response was investigated at a pathway level using GeneGo Metacore (Thomson Reuters 13). Briefly, 87 literature-reported genes linked to statin response were combined with the 185 candidate gene loci reported here.
with the lead SNP (shown in purple). The RefSeq genes in the region are shown in the lower panel.

Figure 2 | Regional association plots of the genome-wide significant associations with LDL-C response after statin treatment. The plots show the genome-wide significant associated loci in the combined meta-analysis (n = 40,914), the APOE locus (a), the LPA locus (b), the SORT1/CELSR2/PSRC1 locus (c) and the SLC01B1 locus (d) (generated using LocusZoom (http://genome.sph.umich.edu/wiki/LocusZoom)). The colour of the SNPs is based on the LD with the lead SNP (shown in purple). The RefSeq genes in the region are shown in the lower panel. P values were generated using linear regression analysis.

Table 2 | Associations of the minor alleles of rs646776, rs445925, rs2900478 and rs10455872 with changes in LDL-C and LDL subfractions in response to statin in the combined CAP and PRINCE studies.

<table>
<thead>
<tr>
<th>Change*</th>
<th>SORT1/CELSR2/PSRC1</th>
<th>APOE</th>
<th>LPA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs646776 (MAF 0.2)</td>
<td>rs445925 (MAF 0.086)</td>
<td>rs2900478 (MAF 0.16)</td>
</tr>
<tr>
<td>LDL-C total</td>
<td>−0.023 0.008 0.003</td>
<td>−0.046 0.018 0.008</td>
<td>0.010 0.005 0.004</td>
</tr>
<tr>
<td>Large LDL-C</td>
<td>−0.028 0.014 0.042</td>
<td>−0.075 0.029 0.009</td>
<td>0.02 0.008 0.01</td>
</tr>
<tr>
<td>Medium LDL-C</td>
<td>−0.027 0.015 0.075</td>
<td>−0.079 0.032 0.012</td>
<td>0.016 0.009 0.07</td>
</tr>
<tr>
<td>Small LDL-C</td>
<td>−0.047 0.018 0.009</td>
<td>−0.071 0.037 0.050</td>
<td>0.002 0.010 0.83</td>
</tr>
<tr>
<td>Very small LDL-C</td>
<td>−0.034 0.009 0.00006</td>
<td>−0.022 0.017 0.202</td>
<td>0.001 0.005 0.90</td>
</tr>
</tbody>
</table>

LDL-C, low-density lipoprotein cholesterol; MAF, minor allele frequency.

*Change in (ln treatment) – ln (baseline) models adjusted for log (baseline variable), age, sex, body mass index, smoking(y/n) and study (CAP versus PRINCE). Betas and P values were assessed using a generalized estimating equation method.

( Supplementary Data 3). A conservative network of direct interactions was constructed between query genes (Supplementary Data 4). The network included 24 genes located in the LDL-C-associated loci (Supplementary Fig. 4). Collectively, our functional and pathway analysis confirms a strong biological and functional role in statin response for several strongly associated gene loci, including APOE/TOMM40/PVRL2 and SORT1/CELSR2/PSRC2.

Discussion

We have performed a meta-analysis of GWAS including more than 40,000 subjects, investigating genetic variants associated with variation in LDL-C lowering on statin treatment independent from associations with baseline LDL-C. We identified four loci at genome-wide significance, including the previously identified APOE and LPA, and the novel GWAS loci SORT1/CELSR2/PSRC1 and SLC01B1.

Nine SNPs in the APOE gene region reached genome-wide significance for LDL-C response. The minor allele of the lead SNP rs445925, which is a proxy for the apoE ε2 protein variant defining SNP rs7412 (ref. 14), was associated with a larger LDL-C-lowering response to statins compared with carriers of the major allele. The magnitude and direction of the effect size was similar to previously reported findings for the rs445925 variant in
degradation. Notably, the minor allele of rs646776 is
importance, increased hepatic LDL uptake via binding to sortilin
lowering of plasma LDL-C results from two mechanisms: reduced
Friedewald formula, include cholesterol that resides in Lp(a)6,8.

This property may also underlie the diminished efficacy of statins
SLCO1B1 gene was genome-wide significantly associated with
statin efficacy was abolished after removal of individuals who showed signs of intolerance26.

GWCA identified three independent loci in the APOE gene
region and two loci in the LPA gene region (Supplementary Table 9). GWCA also showed several other loci with
P < 5 × 10−8 that were not GWAS significant on single-SNP
analysis (HGD, RNFL75, ISCA1L-HTR1A, GLIS3-SLC1A1,
LOC100128657, NKX2-3-SLC2A2 and PEL2). These findings
will require replication in independent, larger data sets. The
significant SNPs in the GWCA analysis explained ~5% of the
variation in LDL-C response to statin treatment. Whether this 5%
is clinically relevant should be investigated by other studies. For
example, it would be of interest to investigate whether this
differential LDL-C lowering is also associated with differential
event reduction by statin treatment.

In the current study, we combined the results of 6 randomized
clinical trials and 10 observational studies in the first stage. This
approach resulted also in combining several types of statins, since
different statins were studied in the trials and within the
observational studies (Supplementary Table 2). This, and the
variation in statin dosage during follow-up for an individual, is a
limitation of the current study, since, for example, the impact of
the SLCO1B1 variant on statin pharmacogenetics is known to be
highly dependent on statin type and dose24,27. To overcome this
limitation, the individual study analyses were adjusted for statin
dose. Dividing the actual statin dose given by the statin-specific
dose equivalent (Supplementary Table 3) gives the statin-adjusted
equivalent based on the daily dosages required to achieve a mean
30% LDL-C reduction. Using this table, we made the different
statins equivalent based on the daily dosages required to achieve a mean
30% LDL-C reduction. Using this table, we made the different
statins equivalent based on the daily dosages required to achieve a mean
30% LDL-C reduction. Using this table, we made the different
statins equivalent based on the daily dosages required to achieve a mean
30% LDL-C reduction. Using this table, we made the different

Another possible limitation of the current study is the influence of the identified genetic variants on baseline LDL-C levels. In pharmacogenetic studies investigating the LDL-C-lowering response to statins, it is important to eliminate the effect of association between the genetic variant and baseline LDL-C levels, since those findings may confound the response to treatment associations. Previous large GWAS studies have shown strong associations between baseline LDL-C levels and genetic variants in SORT1/CERS2/PSRC1, APOE and LPA. To eliminate those possible confounding effects, our response to treatment analyses were adjusted for baseline LDL-C levels. In addition, additional analysis in CARDS and JUPITER suggests no or little influence of genetic associations with baseline LDL-C on the genetic effects on LDL-C-lowering response.

In conclusion, this study is the largest meta-analysis of GWAS for LDL-C response to statin therapy conducted to date. Our results demonstrate that apart from the previously identified APOE and LPA loci, two new loci, SORT1/CERS2/PSRC1 and SLC20A11, also have a modest but genome-wide significant effect on LDL-C response. The minor alleles of the APOE rs445925 and SORT1/CERS2/PSRC1 rs646776 SNPs were associated with a larger statin response, whereas the minor alleles of the LPA rs10455872 and SLC20A11 rs2990478 SNPs were associated with a smaller statin response. Our findings advance the understanding of the pharmacogenetic architecture of statin response.

Methods

Study populations. The meta-analysis was conducted in the GIST consortium, which includes data from 8 randomized controlled statin trials (RCTs) and 11 prospective, population-based studies. The initial analysis (first stage) was performed in 8,421 statin-treated subjects from 6 RCTs (ASCOT, CARDS, CAP, PRINC, PROSPER and TNT) and 10,175 statin-treated subjects from 10 observational studies (AGES, ARIC, BioVU, CHS, HHS, GoDARTS I, GoDARTS II, Health ABC, HVH and MESA). Further information (second stage) was performed in 21,975 statin-treated subjects from two randomized trials (HPS and JUPITER) and one observational study (Rotterdam Study). Six SNPs were additionally genotyped in the Scandinavian participants of the ASCOT Study. The details of the first- and second-stage studies can be found in the Supplementary Tables 1 and 2 and Supplementary Notes 1 and 2.

Subjects. Response to statin treatment was studied in statin-treated subjects only and not in those treated with placebo. Subjects included in the observational studies’ analysis should be treated with statins and have LDL-C measurements before and after start of statin treatment. Subjects of reported or suspected non-European ancestry were excluded. All participants gave written informed consent and the study was approved by all institutional ethics committees.

Outcome measurements. The response to statin treatment was defined as the difference between the natural log-transformed on- and off-treatment LDL-C levels. The beta of the corresponding regression reflects thus the fraction of different LDL lowering in carriers versus non-carriers of the SNP. For observational studies, the on-treatment LDL-C levels were taken into account for all kinds of prescribed statins, at any dosage, for any indication and for at least 4 weeks before measurement. Characteristics of on- and off-treatment LDL-C levels and statins used in each study are shown in Supplementary Table 2. For each individual, at least one off-treatment LDL-C measurement and at least one on-treatment LDL-C measurement were required. When multiple on- or off-treatment measurements were available, the mean of the cholesterol measurements was used. Subjects with missing on- or off-treatment measurements were excluded, with the exception of the GoDARTS cohorts for which missing off-treatment LDL-C levels were estimated using imputation methods (Supplementary Note 2). In the HPS, proportional LDL-C response was defined by the changes in natural log lipid levels from the screening visit before starting statin therapy to the randomization visit. To control for possible associations with off-treatment LDL-C levels, analyses were adjusted for the natural log-transformed off-treatment LDL-C level. An additive genetic model was assumed when tested using a linear regression model. For imputed SNPs, regression analysis was performed on expected allele dosage. Analyses were additionally adjusted for age, sex and study-specific covariates (for example, ancestry principal components or country). Analyses in the observational studies were, if available, additionally adjusted for the statin dose by the natural logarithm of the dose equivalent as defined in Supplementary Table 3. This table shows the dose equivalent per statin type; dividing the statin dosage of an individual by the dose equivalent shown in Supplementary Table 3 will give the adjusted statin dosage.

Quality control and meta-analysis. Centrally, within each study, SNPs with MAF < 1% or imputation quality < 0.3 were excluded from the analysis. QQ-plots were assessed for each study to identify locus-specific outliers (Supplementary Fig. 1). The software package METAL was used for performing the meta-analysis (http://www.sph.umich.edu/csg/abecasis/Metal/index.html). A fixed effects, inverse variance weighted approach was used. Using an inverse variance weighted meta-analysis will give smaller weights to studies with large s.e. To correct for possible population stratification, genomic control was performed by adjusting the within-study findings and the meta-analysis results for the genomic inflation factor.

Second stage. SNPs with P values < 5 \times 10^{-8} in the first-stage meta-analysis were selected for further investigation in a second stage. A maximum of two SNPs per locus were selected, based on statistical significance, except for the APOE locus, for which all genome-wide significant associated SNPs were selected for validation. A total of 10 SNPs, within 10 independent loci, were included in the second stage, which was performed in the JUPITER trial, HPS study and the Rotterdam Study, which all had GWAS data and response to statin treatment available. For 2 of the 246 SNPs, a proxy was used in the JUPITER trial, and 31 SNPs were not available, nor was a proxy SNP. HPS, HPS provided data on 151 directly genotyped SNPs from GWAS and IPLEX experiments, including 48 of the second stage and 103 proxy SNPs (r^2 > 0.8). Analysis in HPS was not adjusted for baseline LDL-C levels. In addition, the number of subjects with data varied from SNP-to-SNP and ranged from ~4,000 for variants with GWAS data to ~18,000 for some candidate genes. Results of the first and second stage were combined using fixed effects, inverse variance weighted meta-analysis and analysed by METAL. As a third stage, six SNPs with P values 5 \times 10^{-8} < P < 5 \times 10^{-7} in the combined meta-analysis were selected for additional genotyping in the Scandinavian participants of the ASCOT study. Kaspar assays were designed for four of the SNPs using the KBioscience Primerpick software, and oligos were provided by Integrated DNA Technologies (http://eu.idtdna.com/site). Full Kaspar methodology is available from LGC SNP genotyping (http://www.lgcgenomics.com/genotyping/kasp-genotyping-reagents/). Two SNPs (rs981844 and rs13166647) were genotyped using Taqman assays supplied by Life Technologies (http://www.lifetechologies.com/uk/en/home.html) using the standard Taqman protocol. Results of the additional genotyping were combined with results from the first and second stages using a fixed effects, inverse variance weighted meta-analysis and analysed by METAL.

Determination of changes in LDL subfractions. LDL subclasses were analysed as described previously using non-denaturing gradient gel electrophoresis of fasting plasma samples taken at baseline and after 6 weeks of simvastatin 40 mg per day (CAP study, n = 579) or 12 weeks of pravastatin 40 mg per day (PRINCE study, n = 1,284). Aliquots of 0.1 ml of whole plasma were mixed 1:1 with a sampling buffer of 20% sucrose and 0.25% bromophenol blue. Electrophoresis of samples and size calibration standards was performed using 2–14% polyacrylamide gradients at 150 V for 3 h following a 15 min pre-run at 75 V. Gels were stained with 0.07% Sudan black for 1 h and stored in a 0.81% acetic acid, 4% methanol solution until they were scanned by computer-assisted densitometry for determination of areas of LDL IVb (22.0–23.2 nm), LDL IVA (23.3–24.1 nm), LDL IIB (24.2–24.6 nm), LDL IIC (24.7–25.5 nm), LDL IIB (25.6–26.4 nm), LDL IId (26.5–27.1 nm) and LDL I (27.2–28.5 nm). The cholesterol concentrations of the subfractions (mg dl^{-1} plasma) were determined by multiplying percent of the total stained LDL area for each subfraction by the LDL-C for that sample. For genetic association analysis, subfractions were grouped into large LDL (LDL I + IIA), small LDL (LDL IId + IIB) and small LDL (LDL IIda). LDL IVb + IVA + IVb) as described previously. A generalized estimating equation method was used to test the association of log change with the interaction of the four SNPs by LDL subfraction.

Effect of off-treatment LDL-C. Effects of genetic variation on treatment response as measured by on-treatment LDL-C could be mediated through effects on the off-treatment LDL-C. To evaluate whether genetic on-treatment LDL-C likely reflects residual effect on off-treatment LDL-C, it is necessary to adjust for the off-treatment LDL-C levels and to correct the maximum likelihood estimate of the adjusted effect of genotype on on-treatment value for the noise in off-treatment values (the noise is both random measurement error and intra-individual variation in usual LDL-C). This analysis was only carried out in CARDS, where baseline measurements were available. From the rules of path analysis, we calculated the direct effect of genotype on an on-treatment trait value as \beta = ax \times (1 - p), where...
β is the coefficient of regression for on-treatment trait value on genotype adjusted for measured off-treatment value, s is the coefficient of regression of baseline LDL on genotype, and δ is the regression coefficient between baseline difference in LDL-C and its difference for measured off-treatment value. For these calculations, we used ρ = 0.8 as a plausible value for the intraclass correlation based on the within-person correlation in LDL-C values taken over two off-treatment visits in CARDs. The interaction of candidate SNPs with statin versus placebo allocation was assessed in the JUPITER trial, since this study was not involved in the first-stage meta-analysis. Regression models were applied to the combined population of statin- and placebo-treated subjects by including extra terms encoding placebo allocation and the product of placebo allocation with SNP minor allele dose.

GWCA using Genome-Complex Trait Analysis. There may be multiple causal variants in a gene and the total variation that could be explained at a locus may be underestimated if only the most significant SNP in the region is selected. To identify independent SNPs, we ideally can perform a conditional analysis, starting with the top associated SNP, across the whole genome followed by a stepwise procedure of selecting additional SNPs, one by one, according to their conditional P values. Such a strategy would allow the discovery of more than two associated SNPs at a locus. To identify independent SNPs across the genome-wide data, we used an approximate conditional and joint analysis approach implemented in Genome-Complex Trait Analysis (GCTA) software (http://www.complextraitgenomics.com/software/gcta/). We used summary-level statistics from the first- and second-stage genome-wide meta-analyses between corrections between SNPs estimated from CARDs GWAS data. SNPs on different chromosomes or more than 10 Mb distant are assumed to be in linkage equilibrium. The model selection process in GCTA starts with the most significant SNP in the single-SNP meta-analysis across the whole genome with P value < 5 × 10⁻⁶. In the next step, it calculates the P values of all the remaining SNPs conditional on the top SNP that have already been selected in the model. To avoid problems due to colinearity, if the squared multiple correlations between a SNP to be tested and the selected SNPs(s) is larger than a cut-off value, such as 0.9, the conditional P value for that SNP will be set to 1. Select the SNPs with minimum conditional P value that is lower than the cut-off P value. Fit all the selected SNPs jointly in a model and drop the SNPs with the P value that is greater than the cut-off P value. This process is repeated until no SNPs can be added or removed from the model.

Pathway analysis and construction of a statin response network. Genes showing evidence of association (based on direct association or LD (HapMap CEU r² > 0.8)) were reviewed for evidence of involvement in statin response at a pathway level using GeneGo MetaCore (Thomson Reuters (portal.genego.com)). A statin response network was constructed in two stages. First, all genes with a literature-reported involvement in statin response (based on Medical Subject Headings (MeSH)) were identified using GeneGo MetaCore (Supplementary Data 3). Second, these genes were combined with all genes in associated loci (including genes in LD) and a network was constructed based on direct interactions only. By using direct interactions only, we created a comprehensive network of direct gene interactions that have been consistently linked to statin response in the literature.

eQTL analysis. LDL-C-associated index SNPs (246 SNPs) were used to identify 1,443 LD proxy SNPs displaying complete LD (r² = 1) across four HapMap builds in European ancestry samples (CEU) using the SNP tool (http://www.broadinstitute.org/mpg/snap/). The primary index SNPs and LD proxies were selected against a collected database of expression SNP (eSNP) results, including the following tissues: fresh lymphocytes, fresh leukocytes, leukocyte samples in individuals with Celiac disease, whole-blood samples, lymphoblastoid cell lines (LCL) derived from asthmatic children, LCL derived from three populations, a separate study on HapMap CEU LCL, additional LCL of population samples, primary phytohaemagglutinin-stimulated T cells, CD1+ T cells, peripheral blood monocytes, CD11+ dendritic cells before and after Mycobacterium tuberculosis infection, oral mucosal and subcutaneous adipose, stomach, endomysium, endocardium, ERα+ and ERβ+ breast cancer tumour cells, brain cortex, cerebellum, liver, colon, and the NIH Roadmap Epigenomics consortium. Building on the functional annotation, we also identified variants that were shown to mediate eQTLs. Genes in associated loci were also used to query the NIH connectivity map for evidence of differential expression in PC3 cell lines treated with pravastatin, simvastatin and rosuvastatin. By combining a wide range of functional data and pathway support, we were able to build up a view of genes with the highest level of support in statin response.

References


34. Fehrmann, R. S. et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet. 7, e1002197 (2011).


Author contributions


Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: B.M.P. serves on the Data and Safety Monitoring Board of a clinical trial funded by the device manufacturer (Zoll LifeCor). N.P. and A.S. received funding from Pfizer for the extended follow-up of the ASCOT UK participants. D.I.C. and P.M.R. received research support for independent genetic analysis in JUPITER from AstraZeneca. F.N. and B.J.B. have employment, stock and stock options in AstraZeneca, a for-profit company engaged in the discovery, development, manufacture and marketing of proprietary therapeutics such as rosuvastatin, but do not consider that this creates any conflict of interest with the subject-matter of this publication. R.M.K. serves on the Merck Global Atherosclerosis Advisory Board. The remaining authors declare no competing financial interests.

Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/


This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© 2014 Macmillan Publishers Limited. All rights reserved.
Membership of Wellcome Trust Case Control Consortium


Data and Analysis Group

Chris C.A. Spencer77, Gavin Band77, Céline Bellenguez77, Colin Freeman77, Garrett Hellenthal77, Eleni Giannoulatou77, Matti Pirinen77, Richard Pearson77, Amy Strange77, Zhan Su77, Damjan Vukcevic77, Peter Donnelly77,78.

DNA, Genotyping, Data QC and Informatics Group


Publications Committee

Christopher G. Mathew92, Jenefer M. Blackwell80,81, Matthew A. Brown83, Aiden Corvin86, Mark I. McCarthy98, Chris C.A. Spencer77.

77Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, UK; 78Department of Statistics, University of Oxford, Oxford, UK; 79Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 80Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia, Australia; 81Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge, UK; 82Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King’s College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London, UK; 83University of Queensland Diamantina Institute, Princess Alexandra Hospital, University of Queensland, Brisbane, Queensland, Australia; 84Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK; 85Department of Epidemiology and Public Health, University College London, London, UK; 86Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland; 87Molecular and Physiological Sciences, The Wellcome Trust, London, UK; 88Centre for Digestive Diseases, Queen Mary University of London, London, UK; 89Digestive Diseases Centre, Leicester Royal Infirmary, Leicester, UK; 90Department of Clinical Pharmacology, Old Road Campus, University of Oxford, Oxford, UK; 91Clinical Neurosciences, St George’s University of London, London, UK; 92King’s College London, Department of Medical and Molecular Genetics, School of Medicine, Guy’s Hospital, London, London, UK; 93Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK; 94King’s College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London, UK; 95University of Cambridge, Department of Clinical Neurosciences, Addenbrooke’s Hospital, Cambridge, UK; 96NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK; 97Department of Molecular Neuroscience, Institute of Neurology, Queen Square, London, UK; 98Oxford Centre for Diabetes, Endocrinology and Metabolism (ICDEM), Churchill Hospital, Oxford, UK.