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


Available at http://clok.uclan.ac.uk/16075/


It is advisable to refer to the publisher’s version if you intend to cite from the work.
http://dx.doi.org/10.1038/ncomms6068

For more information about UCLan’s research in this area go to
http://www.uclan.ac.uk/researchgroups/ and search for <name of research Group>.

For information about Research generally at UCLan please go to
http://www.uclan.ac.uk/research/

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the http://clok.uclan.ac.uk/policies/
Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins


Received 27 Feb 2014 | Accepted 22 Aug 2014 | Published 28 Oct 2014

These authors jointly supervised the work. Correspondence and requests for materials should be addressed to J.W.L. (email: j.w.jukema@lumc.nl) or M.J.C. to (email: m.j.caulfield@qmul.ac.uk). Members of Welcome Trust Case Control Consortium are listed at the end of the paper. These authors jointly supervised the work. Correspondence and requests for materials should be addressed to J.W.L. (email: j.w.jukema@lumc.nl) or M.J.C. to (email: m.j.caulfield@qmul.ac.uk).
the 3-hydroxyethyl-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% and a 20–30% reduction of cardiovascular events. Despite the clinical efficacy of statins in a wide range of patients, interindividual variability exists with regard to LDL-C lowering response as well as efficacy in reducing major cardiovascular events. 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.

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 reduction following rosuvastatin therapy. In the CARDs and ASCOT studies, single nucleotide polymorphisms (SNPs) at LPA (rs10455872) and APOE (rs445925 and rs4420638) were associated with LDL-C response to atorvastatin treatment. A combined GWAS in three statin trials identified a SNP within CLMN (rs8014194) that is associated with the magnitude of statin-induced reduction in plasma cholesterol. However, two other GWAS identified no genetic determinants of LDL-C response to statin therapy at a genome-wide significant level.

On the basis of these studies, as well as previous candidate gene studies, the only genetic variants that have been consistently identified to be associated with variation in LDL-C response to statin therapy, irrespective of statin formulation, are located at or nearby APOE and LPA. To determine whether additional loci may influence LDL-C response to statins, we formed the Genomic Investigation of Statin Therapy (GIST) consortium and conducted a pharmacogenetic meta-analysis using GWAS data sets from randomized controlled trials (RCTs) and observational studies. We identify two loci not previously identified in GWAS, SORT1/CELSR2/PSRC1 and SLCO1B1. In addition, we confirm the associations within the APOE and LPA genes. These findings will extend the knowledge of the pharmacogenetic architecture of statin response.

Results

First-stage meta-analysis. The GIST consortium includes 6 RCTs (n = 8,421 statin recipients) and 10 observational studies (n = 10,175 statin recipients) that participated in the first stage (see Methods; Supplementary Tables 1 and 2; Supplementary Notes 1 and 2). To search for genetic variants associated with differential LDL-C response to statin therapy, each study independently performed a GWAS among statin users, using the difference between the natural log-transformed LDL-C levels on- and off-treatment as the response variable (see Methods). The first-stage meta-analysis identified 13 SNPs, that attained genome-wide significance (P < 5 × 10⁻⁸) for association with LDL-C response to statin treatment (Table 1). The most significant association was for a SNP on chromosome 5 (rs13166647, MAF = 0.230, β = −0.253, s.e. = 0.046, P = 4.50 × 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 × 10⁻⁴ from 158 loci for further investigation in three additional studies comprising up to 22,318 statin-treated subjects (see Methods; Supplementary Tables 1 and 5; 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 × 10⁻⁷ and 2.26 × 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: SORT1/CELSR2/PSRC1 (rs646776, β = −0.013, s.e. = 0.002, P = 1.05 × 10⁻⁹ and rs12740374, β = −0.013, s.e. = 0.002, P = 1.05 × 10⁻⁹; Fig 2c) and SLCO1B1 (rs2900478, β = 0.016, s.e. = 0.003, P = 1.22 × 10⁻⁹; Fig 2d), indicating an additional 1.5% increase per allele in LDL-C lowering effect for carriers of the SORT1/CELSR2/PSRC1 SNP and a 1.6% smaller LDL-C lowering per minor allele for carriers of the SLCO1B1 SNP.

The six next-ranked SNPs with P values just below 5 × 10⁻⁸ in the combined meta-analysis, including the two SNPs at RICTOR (rs13166647 and 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, SORT1/CELSR2/PSRC1 and SLCO1B1.

Subfraction analyses. To extend our results for the novel GWAS finding SORT1/CELSR2/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 SORT1 rs646776 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...
Table 1 | Genome-wide significant associations in stage 1, stage 2 and combined meta-analysis.

<table>
<thead>
<tr>
<th>Chr</th>
<th>Position</th>
<th>Lead SNP</th>
<th>Gene</th>
<th>Coding allele</th>
<th>Noncoding allele</th>
<th>Phase</th>
<th>N</th>
<th>Frequency-coding allele</th>
<th>Beta*</th>
<th>s.e.</th>
<th>% Extra reduction</th>
<th>P value</th>
</tr>
</thead>
</table>
| 1   | 109620053 | rs646776  | SORT1/CELSR2/PSRC1 | C | T | Stage 1 | 16,697 | 0.230 | -0.015 | 0.003 | 1.5 | 6.70 × 10⁻⁷ 
|     |          |          |       |             |                 | Stage 2 | 21,902 | 0.216 | -0.010 | 0.003 | 1.0 | 2.43 × 10⁻⁴ 
|     |          |          |       |             |                 | Combined | 38,599 | 0.013 | 0.002 | 1.3 | 1.05 × 10⁻⁵ 
| 6   | 160930108 | rs10455872 | LPA  | G | A | Stage 1 | 12,981 | 0.069 | 0.041 | 0.006 | 4.1 | 1.95 × 10⁻⁹ 
|     |          |          |       |             |                 | Stage 2 | 18,075 | 0.087 | 0.059 | 0.005 | 5.9 | 7.14 × 10⁻⁶ 
|     |          |          |       |             |                 | Combined | 31,056 | 0.052 | 0.044 | 5.2 | 7.41 × 10⁻⁶ 
| 12  | 21260064  | rs2900478 | SLCO1B1 | A | T | Stage 1 | 16,749 | 0.165 | 0.016 | 0.003 | -1.6 | 2.26 × 10⁻⁶ 
|     |          |          |       |             |                 | Stage 2 | 7,504 | 0.164 | 0.017 | 0.006 | -1.7 | 3.54 × 10⁻⁴ 
|     |          |          |       |             |                 | Combined | 24,253 | 0.016 | 0.003 | -1.6 | 1.22 × 10⁻⁵ 
| 19  | 50107480  | rs445925  | APOE  | A | G | Stage 1 | 13,909 | 0.098 | -0.043 | 0.005 | 4.3 | 1.58 × 10⁻⁸ 
|     |          |          |       |             |                 | Stage 2 | 3,613 | 0.157 | -0.088 | 0.011 | 8.8 | 1.41 × 10⁻⁸ 
|     |          |          |       |             |                 | Combined | 17,522 | -0.051 | 0.005 | 5.1 | 8.52 × 10⁻⁹ 

Chr, chromosome; SNP, single nucleotide polymorphism. *Beta for difference between the natural log-transformed on- and off-treatment low-density lipoprotein cholesterol (LDL-C) levels adjusted for natural log-transformed off-treatment LDL-C, age, sex- and study-specific covariates. The beta reflects the fraction of differential LDL-C lowering in carriers versus non-carriers of the SNP; a negative beta indicates a better statin response (stronger LDL-C reduction); a positive beta a worse statin response. Betas and P values were generated using linear regression analysis. 

Effects of off-treatment LDL-C. To demonstrate that our findings for LDL-C response to statins 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-measured LDL-C response in the GIST meta-analysis (P < 1 × 10⁻⁸), which were available in the CARDs data set. Correcting our effect size estimate further and modelling measurement noise at baseline reduced the apparent effect only slightly for all the markers, suggesting that there is little effect of measurement noise. Next, within the JUPITER trial, additional analyses were performed to determine whether there was an interaction between LDL-C change and statin or placebo allocation. Supplementary Table 8 shows significant P values for interaction (all <5 × 10⁻²) for SNPs at the four genome-wide significant loci in the GIST meta-analysis, also suggesting that genetic effects on baseline LDL-C as manifested in the placebo group contribute at most only in part to genetic effects on LDL-C response in the statin group.

Genome-Wide Conditional Analysis. To investigate whether there were multiple SNPs within any gene and multiple loci associated with differential LDL-C lowering to statin therapy, we performed a conditional analysis across the genome using the summary statistics of the combined meta-analysis. The results of the Genome-Wide Conditional Analysis (GWCA; see Methods; Supplementary Table 9) showed 14 SNPs independently associated with statin response and these explained ~5% of the variation in LDL-C response to statin treatment. Of the 14 independent SNPs, 6 were genome-wide significant in the combined GWAS meta-analysis (Supplementary Table 5).

Previous findings. In Supplementary Table 10, we performed a look-up in our GWAS meta-analysis for SNPs previously described in the literature (NHGRI Catalogue11 of Published GWAS and Candidate gene studies) to be associated with statin response, besides the loci associated at a genome-wide level in the current study. None of these SNPs was associated with statin response in our GWAS after correcting 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, SORT1, 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 resource12 (see Methods). Five genes (APOE, BRCA1, GRPELI, ADRB2 and ETV1) showed convincing evidence of statin responsiveness on the basis of greater than twofold differential expression in response to statin treatment. Eight genes showed suggestive evidence (1.5- to 2-fold change; TOMM40, SREBP1, PSRC1, BCL3, BCAM, ANK3, SIVA1 and RANBP9; Supplementary Data 3).

Finally, involvement in statin response was investigated at a pathway level using GeneGo Metacore (Thomson Reuters13). 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.

We have performed a meta-analysis of GWAS including more...
the GWAS study performed in CARDS and ASCOT\(^8\) and of the SNP rs7412 in JUPITER\(^2\). Since the apoE e2 protein results in increased hepatic cholesterol synthesis, it may also predispose to stronger inhibition of cholesterol synthesis by statin treatment\(^8,10\).

Three independent SNPs at \(LPA\) were significantly associated with LDL-C response to statins. The minor G allele of the lead SNP rs10455872 was associated with smaller LDL-C reduction than the major allele. This result was similar to the previous GWAS findings for this SNP in the JUPITER trial and the combined ASCOT and CARDS study\(^2,8\). The rs10455872 SNP was strongly associated with the KIV-2 copy number variant in \(Lp(a)\), which encodes variability in apo(a) size and is responsible for \(\sim 30\%\) of variance in \(Lp(a)\) levels\(^8,13\). Furthermore, rs10455872 was shown to be strongly associated with plasma \(Lp(a)\) levels\(^8,16\). Standard assays of LDL-C, as well as the Friedewald formula, include cholesterol that resides in \(Lp(a)\)\(^8\). Carriers of this \(LPA\) variant are characterized by higher \(Lp(a)\) levels and a larger proportion of their measured LDL-C resides in \(Lp(a)\) particles\(^8,10\). Since statin therapy does not reduce the number of \(Lp(a)\) particles\(^13\), their presence attenuates the measured LDL-C response to statins.

Two SNPs at \(SORT1/CELSR2/PSRC1\) (rs646776 and rs12740374) on chromosome 1p were associated with an enhanced statin LDL-C response. A similar association was previously observed in a large candidate gene study in HPS\(^5\); however, we demonstrate this finding now first at a genome-wide significance level. The minor allele of rs12740374 has been shown to generate a binding site for the transcription factor C/EBP\(^a,18\). Transcription results in upregulation of hepatic expression of three genes at this locus, \(SORT1\), \(CELSR2\) and \(PSRC1\) (ref. 18), which we also showed in our eQTL analysis (Supplementary Data 1). Of these, \(SORT1\) is most notable, in that it encodes the multifunctional intracellular trafficking protein sortilin, which has been shown to bind tightly to apoB\(^19\). Sortilin-induced lowering of plasma LDL-C results from two mechanisms: reduced secretion of apoB-containing precursors and, perhaps of greater importance, increased hepatic LDL uptake via binding to sortilin at the cell surface, with subsequent internalization and lysosomal degradation\(^19\). Notably, the minor allele of rs646776 is preferentially associated with lower levels of small and very small LDL (Table 2), suggesting that sortilin is of particular importance for regulating levels of these particles\(^18\). Smaller LDL subfractions have been shown to be relatively enriched in particles with reduced LDL receptor binding affinity and cellular uptake\(^20\), a property that may contribute to their associations with increased risk for cardiovascular disease\(^21\). This property may also underlie the diminished efficacy of statins for reduction of these particles (Supplementary Fig. 3)\(^22\), since statins act to reduce LDL-C levels to a large extent by increasing LDL receptor expression as a result of upregulation of the transcription factor SREBP2, whereas \(SORT1\) is not regulated by this mechanism. Hence, the greater statin-mediated reduction of LDL-C among carriers of the rs646776 minor allele could be attributed to relative depletion of LDL particles dependent on sortilin for clearance and hence a residually greater proportion of those LDL particles whose uptake is more dependent on the LDL receptor than on sortilin.

Notably, the strong association of rs646776 with statin-induced reductions in small and very small LDL particles contrasts to the weaker associations of changes in these particles with rs445925, likely the result of differing mechanisms underlying the effects of these SNPs on statin response. As noted above, rs445925 is a proxy for the SNP defining the apoE e2 protein variant that is thought to predispose to heightened statin response as a result of greater statin inhibition of cholesterol synthesis and hence upregulation of SREBP and LDL receptor activity.

The \(SLCOIB1\) rs2900478 minor allele was associated with a smaller LDL-C reduction in response to statin treatment. \(SLCOIB1\) encodes the organic anion-transporting polypeptide OATP1B1 and facilitates the hepatic uptake of statins\(^23\). SNP rs2900478 is in strong LD \((r^2 = 0.89)\) with rs4149056, which represents the Val174Ala substitution resulting in complete loss of function. In the HPS trial, which used simvastatin, this candidate gene SNP was associated with a 1% lower LDL-C reduction per allele\(^24\). Single-dose studies have shown that the observed area under the curve of plasma level of active simvastatin after a dose of 40 mg was 221% higher in rs4149056 CC homozygotes compared with rs4149056 TT homozygotes, as compared with atorvastatin 20 mg (144% higher for CC versus TT)\(^24\). This finding results from the slower hepatic uptake of statins caused by the genetic variant, which would also be expected to result in a reduction in the cholesterol-lowering effect\(^25\). In a GWAS of the genetic risk factors for simvastatin-induced myopathy, \(SLCOIB1\) showed the strongest association\(^25\). Homozygous carriers of the \(SLCOIB1\) variant had a 1.69 times higher risk for myopathy compared with non-carriers. This might have led to a decrease in study medication adherence, and consequently a decreased effect on LDL-C in carriers of this SNP. In addition, previous analysis in the GoDARTS study showed that the effect of the \(SLCOIB1\) gene on statin efficacy was abolished after removal of individuals who showed signs of intolerance\(^26\).

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 \(P < 5 \times 10^{-8}\) that were not GWAS significant on single-SNP analysis (HGDD, RNF175, ISA1C-HTR1A, GLIS3-SLC1A1, LOC100128657, NKX2-3-SLC25A28 and PELD\(2\) ). These findings require replication in independent, larger data sets. The significant SNPs in the GWCA analysis explained \(\sim 5\%\) 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 \(SLCOIB1\) variant on statin pharmacogenetics is known to be highly dependent on statin type and dose\(^24,27\). 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 statin dosages and types comparable within the studies. To correct for between-study variance, we used a fixed effect meta-analysis with inverse variance weighting. Since we observed that the \(SLCOIB1\) gene was genome-wide significantly associated with LDL lowering, this highlights the thoroughness of our analytical approach, in which the analyses were correctly adjusted for the type and dose of statins used (Supplementary Table 3). Moreover, a comparison of the estimates of the SNPs between the RCTs (where there are no intra-individual differences in dosages) with the estimates of the SNPs in the observational studies showed large homogeneity between the estimates in the various study designs (Supplementary Fig. 2), indicating that our adjustment for dosage seems to be sufficient within this study.
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/CELSR2/PSRC1 and SLCO1B1, as well as a moderate but genome-wide significant effect on LDL-C response. The minor alleles of the APOE rs449525 and SORT1/CELSR2/PSRC1 rs464776 SNPs were associated with a larger statin response, whereas the minor alleles of the LPA rs10455872 and SLCO1B1 rs2900478 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, CARD, PRINC, PROSPER and TINT) and 10,175 statin-treated subjects from 10 observational studies (AGES, ARIC, BioYU, CHS, FHS, GoDARTS I, GoDARTS II, Health ABC, HVH and MESA). Further details of the first stage were 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 the change in the natural log of the ratio of the LDL-C concentrations in the treated vs untreated subject. 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 model was assumed for the SNPs. For imputed SNPs, regression analysis was performed on expected allele dosage.

GWAS analysis. Each study independently performed the GWAS on the difference between natural log-transformed on- and off-treatment LDL-C levels. 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 model was assumed for the SNPs. 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 by each study to identify between-study differences (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 $P < 5 \times 10^{-8}$ 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 52 SNPs, within 11 independent loci, were selected for 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 24 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 data provided on 151 directly genotyped SNPs from GWAS and IPLEX expression, including 48 of the second stage and 103 proxy SNPs ($r^2 > 0.8$). Analysis in HPS was not adjusted for in baseline LDL-C levels. In addition, the number of subjects with data varied from SNP-to-SNP and ranges 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 $P < 5 \times 10^{-8} < P < 5 \times 10^{-7}$ 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 Primerpicker 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.lifetechnologies.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 (PRINC study, n = 1,208). 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 at 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 I, LDL IIa, LDL IIb, LDL IIIa, LDL IIIb and LDL V (25.6–26.4 nm). LDL IIa and LDL IIb (24.2–24.6 nm), LDL IIIa (24.7–25.5 nm), LDL I, LDL IVb (22.0–23.2 nm), LDL IIb and LDL IIIa were grouped into large LDL (LDL I + IIa), medium LDL (LDL IIb), small LDL (LDL IIIa) and very small LDL (LDL 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 levels 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 performed in CAPDS, 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 $\beta = \alpha + (1 - \rho) \gamma$, where $\rho$ is the correlation between the on- and off-treatment measurements.
β is the coefficient of regression for on-treatment trait value on genotype adjusted for measured off-treatment value, ρ is the coefficient of regression of baseline LDL on genotype, and the total variation that could be explained at a locus may be estimated from CARDS GWAS data. SNPs on different chromosomes or more than 10 Mb distance 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^-8. 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 collinearity, 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 r^2 > 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 based on direct interactions only. By including direct interactions between these genes were combined with all genes in associated loci (including genes in LD) and a network was constructed based on direct interactions only, we created a conservative 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 (r^2 = 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 between eSNPs (eSNP correlations, including the following tissues: fresh lymphocytes41,42, fresh leukocytes31, leukocyte samples in individuals with Celiac disease43, whole-blood samples33–36, lymphoblastoid cell lines (LCL) derived from asthmatic children34,38, HapMap LCL from three populations38, a separate study on HapMap CEU CEPH, additional LCL of population samples41–43 (Mangravite et al., unpublished), CD19 + B cells44, primary phytohaemagglutinin-stimulated T cells41, CD4 + T cells45, peripheral blood monocytes44,45, CD11 + dendritic cells before and after Mycobacterium tuberculosis infection46,47, intestinal and subcutaneous adipose43,44, stomach49, cerebellum53,55, three additional large studies of brain regions including prefrontal cortex, cerebral cortex and cerebellum, respectively55, liver45,59,59, osteoblast50, ileum56, lung56, skin56,63 and primary fibroblasts51. Micro-RNA QTLs were also queried for LCL47 and gluteal and abdominal adipose45. The collected eSNP results met the criteria for genetic expression levels as defined in the original papers. In each case where a LDL-C-associated SNP or proxy was associated with a transcript, we further examined the strongest eSNP for that transcript within the data set (best eSNP), and the LD between the best eSNP and GIST-selected eSNPs to estimate the concordance of the LDL-C and expression signals.

Statin response connectivity map analysis. The Connectivity Map (Cmap) data set is available at the Broad Institute (www.broadinstitute.org/cmap) and contains more than 7,000 expression profiles representing 1,309 compounds used on five different cultured human cancer cell lines (MCF7, MDA-MB-231, HEP-2, PC3 and SKMEL5). We selected (prostate tumour-derived) PC3 cells as they showed the most responsiveness to statins at a genome-wide level. Four statins were included in our analysis, including pravastatin, atorvastatin, simvastatin and rosuvastatin. PC3 Instance reference files for each statin treatment were extracted (as defined by Lamb et al.52), that is, a treatment associated to its control pair. Transcripts were considered to show evidence of differential expression with a fold change > 2. A fold change > 1.5 was considered to be suggestive of differential expression only.

Exploration of functional impact among directly and indirectly associated variants. Genes and variants across all LDL-C-associated loci were investigated for evidence of functional perturbation using a range of bioinformatics tools and databases. Variants showing LD (CEU r^2 > 0.8) with associated variants were explored for impact on coding gene function using Annovar48 and regulatory function using a combination of HaploReg53 and RegulomeDB54, which both draw on comprehensive data from the Encyclopedia of DNA Elements (ENCODE)55 and the NIH Roadmap Epigenomics consortium56. 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).


Acknowledgements

Data from the Heart Protection Study have been contributed by Heart Protection Study Collaborative Group9.

PROSPER/PHASE The Prospective Study of Pravastatin in the Elderly at Risk (PROSPER) trial was supported by an investigator-initiated grant from Bristol-Myers Squibb, USA. The study was conducted, analysed and reported independently of the company. The GWAS project PHASE has received funding from the European Union’s Seventh Framework Programme (FP7/2007–2013) under grant agreement HEALTH-F2-2009-223004. A part of the genotyping was funded by The Netherlands Consortium for Healthy Ageing (NGI: 05060810). Professor Dr JWJ is an established clinical investigator of The Netherlands Heart Foundation (2001 D 032).

ASCOT The Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) trial was funded by an investigator-initiated grant from Pfizer, USA. The study was investigator led and was conducted, analysed and reported independently of the company. The GWAS project PHASE has received funding from the European Union’s Seventh Framework Programme (FP7/2007–2013) under grant agreement HEALTH-F2-2009-223004. A part of the genotyping was funded by The Netherlands Consortium for Healthy Ageing (NGI: 05060810). Professor Dr JWJ is an established clinical investigator of The Netherlands Heart Foundation (2001 D 032).

CARDS We thank the other investigators, the staff and the participants of the CARDS study. A full list of CARDS investigators can be found in our previous papers. CARDS was funded by grants to the Universities of London and Manchester by Pfizer, UK and the Department of Health. The study was investigator led and was conducted, analysed and reported independently of the company. The GWAS project PHASE has received funding from Pfizer, the UK Department of Health and the National Institutes of Health: grant U19 HL069977 from the National Heart, Lung, and Blood Institute, and grant UL1TR000124 from the National Center for Advancing Translational Sciences.

TN7: The TN7 study was funded by Pfizer, who also provided support for genotyping.
AGES: This study has been funded by NIH contract N01-AG-1-2100, the NIA Intramural Research Program, Hjartavænd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN2682010001EC, HHSN268201000006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C and HHSN268201100012C), R01HL87641, R01HL59367 and R01HL086949; National Human Genome Research Institute contract U01HG004402, and National Institutes of Health contract U01HG002055. We thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number ULIR002505, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

BioVU: BioVU receives support through the National Center for Research Resources UL1 RR024757, which is now the National Center for Advancing Translational Sciences, 2 UL1 TR000445. Genotyping was supported via grant U01-HG04603 from the National Human Genome Research Institute and RC2-MG092188 from the National Institute of General Medical Sciences.

CHS: This CHS research was supported by NHLBI contracts HHSN268201000036C, HHSN268200800007C, N01HC5522, N01HC58079, N01HC58508, N01HC58081, N01HC58502, N01HC58308 and N01HC58068; and NHLBI grants HL080295, HL07652, HL015756 and HL103612 with additional contribution from the National Institute on Aging, the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG02629 from the National Institute on Aging (NIA). A full list of CHS investigators and institutions can be found at http://www.chs-nihbi.org/pi.htm. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Framingham HS: The Framingham Heart Study work was supported by the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine (Contract No. 01-HC-21595), its contract with Affymetrix, Inc. for genotyping services (Contract No. N02-HL-6-478) and based on analyses from Framingham Heart Study investigators participating in the SNP Health Association Resource (SHArE) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGa-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center, also supported by R01HL103612 (PI psaty, subcontract PI, Vasan).

GoDARTS: We are grateful to all the participants who took part in this study, to the general practitioners, to the Scottish School of Primary Care for their help in recruiting the participants and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The Welcome Trust provides support for Wellcome Trust United Kingdom Type 2 Diabetes Case Control Collection (GoDARTS) and informatics support is provided by the Chief Scientist Office. The Welcome Trust funds the Scottish Health Informatics Programme, provides core support for the Wellcome Trust Centre for Health Informatics in Oxford and funds the Wellcome Trust Case Control Consortium. This research was specifically funded by Diabetes UK (084727/Z/08/Z, 085475/Z/08/Z and 085475/B/08/Z). Genotyping of the GoDARTS population was funded by NIA grants R01AG62103 and R01AG62106. The genome-wide association studies of LDL cholesterol response to statins.


Additional information

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 S.R.H. 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/
Membership of Wellcome Trust Case Control Consortium


Data and Analysis Group


DNA, Genotyping, Data QC and Informatics Group


Publications Committee

Christopher G. Mathew, Jenefer M. Blackwell, Matthew A. Brown, Aiden Corvin, Mark I. McCarthy, Chris C.A. Spencer.