Meta-GWAS of PCSK9 levels detects two novel loci at *APOB* and *TM6SF2*

Janne Pott[1,2], Jesper Gadin[3], Elizabeth Theusch[4], Marcus E. Kleber[5,6], Graciela E. Delgado[5], Holger Kirsten[1,2], Stefanie M. Hauck[7], Ralph Burkhardt[2,8,9], Hubert Scharnagl[10], Ronald M. Krauss[4,11], Markus Loeffler[1,2], Winfried März[5,10,12], Joachim Thiery[2,8,13], Angela Silveira[3], Ferdinand M. van 't Hooft[3], Markus Scholz[1,2,*]

[1]Institute for Medical Informatics, Statistics and Epidemiology, Medical Faculty, University of Leipzig, Leipzig, Germany

[2]LIFE Research Center for Civilization Diseases, Medical Faculty, University of Leipzig, Leipzig, Germany

[3]Division of Cardiovascular Medicine, Department of Medicine, Karolinska Institutet, Stockholm, Karolinska University Hospital Solna, Sweden

[4]Department of Pediatrics, University of California San Francisco, Oakland, CA, USA

[5]Vth Department of Medicine (Nephrology, Hypertensiology, Rheumatology, Endocrinology, Diabetology), Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[6]SYNLAB MVZ Humangenetik Mannheim, Mannheim, Germany

[7]Metabolomics and Proteomics Core and Research Unit Protein Science, Helmholtz Zentrum München, Neuherberg, Germany

[8]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital Leipzig

[9]Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Regensburg, Germany

[10]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria

[11]Department of Medicine, University of California San Francisco, Oakland, CA, USA

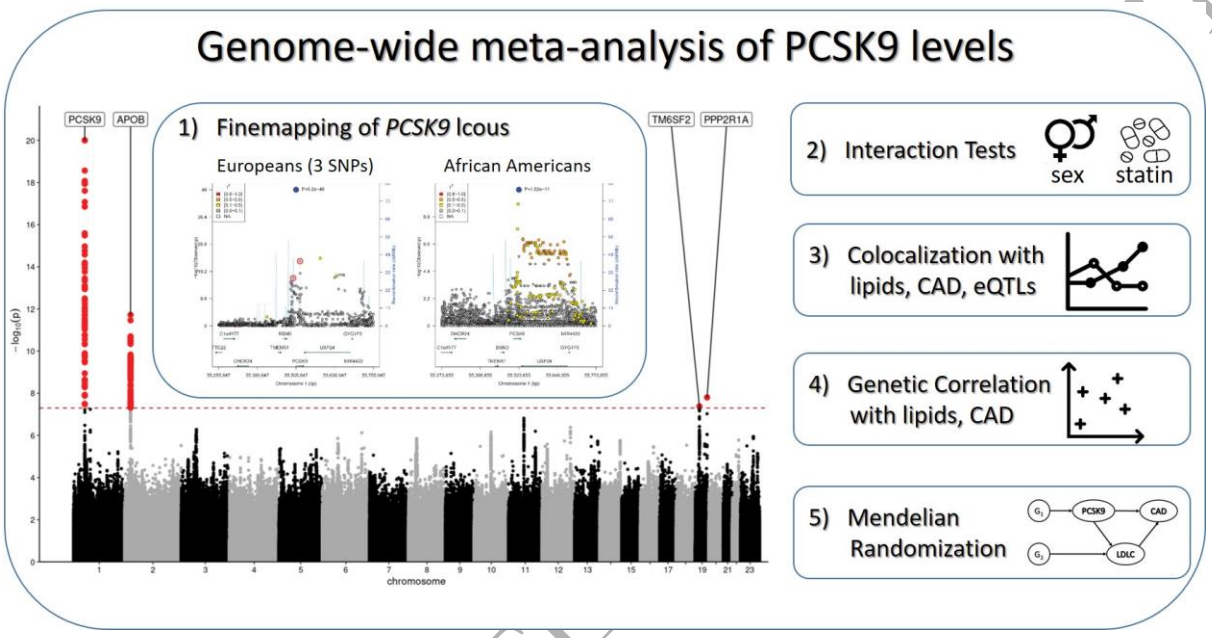
[12]SYNLAB Academy, SYNALB Holding Deutschland GmbH, Mannheim, Germany

[13]Faculty of Medicine, Kiel University, Kiel, Germany

* Corresponding author:

Markus Scholz (Institute for Medical Informatics, Statistics and Epidemiology, Haertelstr. 16-18, 04107 Leipzig, Germany; fax: +49 341 9716109; telephone number: +40 341 9716190; Email address: markus.scholz@imise.uni-leipzig.de)

Graphical Abstract



Abstract

Background: Proprotein convertase subtilisin/kexin type 9 (PCSK9) is a key player in lipid metabolism, as it degrades LDL receptors from hepatic cell membranes. So far, only variants of the *PCSK9* gene locus were found to be associated with PCSK9 levels. Here we aimed to identify novel genetic loci that regulate PCSK9 levels and how they relate to other lipid traits. Additionally, we investigated to what extend the causal effect of PCSK9 on coronary artery disease (CAD) is mediated by LDL-C.

Methods & Results: We performed a genome-wide association study meta-analysis of PCSK9 levels in up to 12,721 samples of European ancestry. The estimated heritability was 10.3%, which increased to 12.6% using only samples from patients without statin treatment. We

successfully replicated the known *PCSK9* hit consisting of three independent signals. Interestingly, in a study of 300 African Americans, we confirmed the locus with a different *PCSK9* variant. Beyond *PCSK9*, our meta-analysis detected three novel loci with genome-wide significance. Co-localization analysis with cis-eQTLs and lipid traits revealed biologically plausible candidate genes at two of them: *APOB* and *TM6SF2*. In a bivariate Mendelian Randomization analysis, we detected a strong effect of PCSK9 on LDL-C, but not vice versa. LDL-C mediated 63% of the total causal effect of PCSK9 on CAD.

Conclusion: Our study identified novel genetic loci with plausible candidate genes affecting PCSK9 levels. Ethnic heterogeneity was observed at the *PCSK9* locus itself. While the causal effect of PCSK9 on CAD is mainly mediated by LDL-C, an independent direct effect also occurs.

Introduction

Proprotein convertase subtilisin/kexin type 9 (PCSK9) is a key player in lipid metabolism by controlling cellular low-density lipoprotein-cholesterol (LDL-C) uptake. This regulation is based on PCSK9 binding to LDL receptors (LDLR) on hepatic cells followed by endocytosis and degradation of the PCSK9-LDLR complex (1). Understanding genetic and non-genetic regulators of PCSK9 levels is crucial for effective personalized treatment of lipid disorders. Statins are the most commonly used class of lipid-lowering drugs, but it is well known that their effect is attenuated to some extent by upregulation of PCSK9. Therefore, combined application of statins and PCSK9 inhibitors (e.g. alirocumab (2) or evolocumab (3)) are applied to counter this effect in cases of insufficient response to statin treatment.

The genetic regulation of PCSK9 is only partly unraveled. First, mutations in the *PCSK9* gene were discovered in 2003 and are related to autosomal dominant hypercholesterolemia (4). Since then, several gain-of-function and loss-of-function mutations of the *PCSK9* gene have been described (5,6). It is a genetically diverse locus, with differences of allele frequencies between Europeans and African-Americans, leading to a significant population differentiation (7). In a trans-ethnic fine-mapping study of lipid traits it was shown that *PCSK9* variants are associated with LDL-C in both European and African-American samples, but only one signal was associated in both, namely rs11591147 (8). PCSK9 levels are not routinely measured in large studies, and the two published genome-wide association studies (GWAS) of PCSK9 levels are of individuals of European ancestry (9,10). The variant rs11591147 was the lead SNP in both GWASs (9,10), and only the *PCSK9* locus was detected with sufficient significance. So far, no GWAS of PCSK9 levels in African-Americans has been published. While a recently conducted study estimated the PCSK9 heritability at 47% in Europeans (11), SNPs described so far only explained 4% of the trait's variance (9). This and the limited sample size of previous studies (N=1,215 and N=3,290) suggest that further genetic associations could be discovered. A

GWAS focusing on statin-induced change in PCSK9 levels (N=562) (12) revealed a significant hit in *CFAP44* (aka *WDR52*), which points towards an interaction of statin treatment and genetic regulation on PCSK9 levels.

In a Mendelian Randomization (MR) study we showed causal effects of PCSK9 levels on risk for several atherosclerotic phenotypes comprising coronary artery disease (CAD), and carotid plaques (9). However, mediating effects of LDL-C were not analyzed in this context.

In this study, we significantly increased the sample size of our previous work by performing the first genome-wide association meta-analysis (GWAMA) of PCSK9 levels in 12,721 participants of five independent studies of European descent. For African-Americans, only one study with a sample size of n=300 was available, which we used to validate the detected significantly associated loci in another ethnicity. The primary objective was to improve our understanding of the genetic regulation of PCSK9 levels by adding new genes and increasing the set of causal variants. We also analyzed statin-stratified subgroups to detect possible interactions of genetics and statin treatment. We used these results to unravel the causal relationships between lipid traits, PCSK9 and CAD by genetic correlation analyses, co-localization analyses, and Mendelian Randomization network analyses.

Results

Meta-GWAS & Finemapping results

We performed a GWAMA for PCSK9 levels in five independent studies: *LIFE-Heart* (13), *LIFE-Adult* (14), *LURIC* (15), *CAP* (16), and *TwinGene* (17) (n=12,721 individuals of European ancestry, n=10,186 without statin treatment). The average age ranged between 54 and 62 years and sex distributions were between 52% and 69% males. Detailed study characteristics are provided in the Supplemental Data and Table S1. Each cohort imputed genetic data to 1000 Genomes Phase 3 (18), and performed association analyses with PCSK9 levels using a standardized protocol (see Methods).

After applying SNP filters (MAF≥1%, info≥0.5, I^2≤0.9, number of studies≥2), about 8.7 million SNPs remained for a fixed-effect meta-analysis. No general signs of inflation were observed (λ=1.01 of fixed effects meta-analysis using genomic control corrected single study results, see Figure S1 for QQ-Plots). Genome-wide results of both the combined set and statin free subset are shown in Figure 1. We discovered 182 SNPs at four loci achieving genome-wide significance in at least one of the settings ($p<5x10^{-8}$). Further 125 SNPs at eight additional loci reached at least suggestive significance ($p<5x10^{-6}$). Summary statistics and annotations can be found in Table S2 – S5.

The strongest association was observed at 1p32.3, within the *PCSK9* gene. This locus has been previously described for associations with PCSK9 levels (10,9), but also for lipids (19) and coronary artery disease (20). The other three genome-wide significant hits were novel: at 2q24.1 (lead SNP rs673548), 19p13.11 (rs58542926), and 19q13.41 (rs71180459). To better define potential causal genes, we performed per locus conditional-joint (COJO) analyses to identify independent variants, and validated the results in 300 African-American samples. Next, credible sets (CS) containing the causal variant with 99% certainty were determined for

each independent SNP, and checked for interactions regarding sex and statin treatment. Finally, we tested for colocalization with cis-eQTLs, lipids and CAD association signals (see Methods). We present the results per locus.

PCSK9 locus (1p32.3)

We detected 71 SNPs with genome-wide significance at the 1p32.3 locus. In the conditional-joint analysis, three independent signals were detected: the lead SNPs were rs11591147, rs2495477, and rs11206510 (see Table 1 for summary statistics of COJO and Figure 2 for regional association of the conditioned estimates, and Figure S2 for forest plots). Since both rs11591147 and rs2495477 had a posterior probability (PP) > 0.99, their CS contained only the respective variant. The lead SNP rs11591147 codes for the well-known missense mutation R46L, increasing PCSK9 degradation. Rs2495477 is an intronic variant, which, according to CADD evaluation, influences the splicing process. Therefore, both are functionally plausible causal variants for PCSK9 levels. For the third independent signal, rs11206510, 21 SNPs form the respective 99% CS. Besides rs11206510, which is known for its association with CAD, two SNPs are plausible causal variants: rs11583680, a missense mutation (A53V), and rs45448095, a 5' UTR modifier. Both have regulatory consequences according to CADD evaluation (see Table S6 and S7 for CS and CADD annotation).

In the attempt to validate the locus in African-American samples from *CAP*, rs11591147 turned out to be monomorphic in this ethnicity. The other two SNPs did not reach the nominal threshold, but rs11206510 had concordant effect direction. Statistics are provided in Table S8. Analyzing the PCSK9 locus in more detail, we found one variant significantly associated in African Americans, namely rs28362263 causing the loss-of-function mutation A443T (21) in PCSK9 (β=0.36, p=1.0x10^{-11}). This SNP was monomorphic in our European GWAMA and it had been reported for association with LDL-C in multi-ethnic studies (6,22), but not with

PCSK9. We provided summary statistics in African-Americans for this locus in Table S9 and a regional association plot in Figure S3.

The lead SNP rs11591147 showed a significant interaction with statin treatment. Still being genome-wide significant in the statin free subset ($\beta_{no\ statin} = -0.320, p_{no\ statin} = 9.06 \times 10^{-47}$), there was no effect in participants under statin treatment ($\beta_{statin} = -0.076, p_{statin} = 0.199$). Effect sizes differed significantly controlling for a 5% FDR ($p_{diff} = 1.24 \times 10^{-4}, q_{diff} = 0.035$, see Table S10 and Figure S4). However, this interaction was not confirmed when using association with statin-induced changes in plasma PCSK9 in European ancestry *CAP* participants (p=0.502), who all had plasma PCSK9 measured before and during statin treatment. Five SNPs of the rs11206510 credible set were suggestively significantly associated in the statin treatment group, and the effects were more than twice of those for the statin-free subset. However, formal interaction analysis did not pass the FDR controlled threshold. In the sex-stratified analysis, rs2495477 showed a trend towards a stronger effect in men, but again this difference was not significant after FDR control.

The unconditioned association signals at 1p32.3 were co-localized with total cholesterol (TC), LDL-C and CAD (PP4>0.75), but not with eQTLs of PCSK9 expression levels (max. PP4=0.110 in transformed fibroblasts). Using results of the conditional-joint analysis, there was a strong co-localization of *PCSK9* eQTLs with rs2495477 in testis and whole blood, and with rs11206510 in brain (cerebellum and cerebellar hemisphere), lung, and nerve (tibial) tissues. Similar results were obtained when using statistics of the statin free sub-group. Colocalization results are summarized in Tables 2, S11 and Figure 3.

APOB locus (2p24.1)

There were 168 genome-wide significant SNPs at 2p24.1, but the conditional analysis revealed only one independent signal, rs673548 (PP=0.350, see Figures S5 and S6 for forest plot and

RA plot). The 99%-CS of this locus contained 68 variants. The second highest posterior probability was observed for rs676210 (PP=0.209), which is in high LD with the lead SNP (LD r^2=0.99). It is a missense mutation of *APOB* (P2739L) with high CADD score, and therefore a likely causal variant for the observed association. Both variants showed the same effect direction in African-Americans, and rs673548 reached nominal significance (p=0.040).

We did not observe any significant interactions for the CS-SNPs of this locus (lead SNP: $\beta_{no\ statin} = 0.038, \beta_{statin} = 0.030,\ p_{diff} = 0.682$). Co-localization analyses revealed an overlap with high-density lipoprotein-cholesterol (HDL-C) and triglyceride (TG) (PP4=0.985 and 0.975, respectively), but not with TC and LDL-C (both PP3=1), indicating different SNPs driving the associations of these traits compared to PCSK9. In the analyses of eQTL data, we only detected co-localizations when using statistics of the statin free subgroup. Here, we observed co-localizations between PCSK9 associations and *APOB* eQTLs in colon sigmoid (PP4=0.928), brain (substantia nigra, PP4=0.869), artery (tibial, PP4=0.843)) and esophagus (muscularis and gastroesophageal junction, PP4=0.970 and PP4=0.847, respectively).

TM6SF2 locus (19p13.11)

There was one genome-wide and 21 suggestive significant SNPs at 19p13.11 in the statin-free setting, while in the combined analysis this locus reached only suggestive significance. Only the lead-SNP rs58542926 represented an independent signal, and its 99% credible set contains 171 SNPs. The lead variant codes for a missense mutation in *TM6SF2* (E167K, loss of function mutation) and was associated with a negative effect on PCSK9 levels. We therefore considered it the most likely causal variant at this locus and *TM6SF2* as the respective candidate gene.

Although the locus was more strongly associated in statin free participants than in the combined analysis, there was no significant SNP x statin interaction (rs58542926 : $\beta_{no\ statin} = -0.052, \beta_{statin} = -0.020,\ p_{diff} = 0.200$). However, there were four SNPs with nominal

significant interaction, and all of them were also associated with statin-induced changes in plasma PCSK9 levels in *CAP*. TC, LDL-C, and TG associations co-localized with the PCSK9 association here (PP4>0.98), while there was no co-localization with HDL-C (PP1=0.956). We did not observe any co-localization with *TM6SF2* eQTLs, but with *ATP13A1* in skin tissue (sun exposed, PP4=0.873), and with *MAU2* in whole blood (PP4=0.850).

PPP2R1A locus (19q13.14)

At the fourth locus, only one SNP reached genome-wide and three SNPs suggestive significance (lead SNP rs71180459). The 99%-CS of rs71180459 was large, containing 4,109 of the 4,370 variants of the region, of which 107 had a CADD score >10 (see Figure S7 and Table S6). Three of the SNPs are in weak LD (r^2=0.3) with a variant reported for HDL-C levels (23) (reported gene *FPR3*).

In an enrichment analysis using all eQTLs in LD (r^2>0.3) and nearby genes of our four genome-wide loci, we detected one pathway that connected the gene *PPP2R1A* from this locus with *APOB*: "Platelet sensitization by LDL" (pathway ID R-HSA-432142, Enrichment OR=21.1, p=0.0042).

The locus did not reach suggestive significance in the statin free subset, and only 153 variants of the CS reached nominal significance here. In the statin-treated subset, 390 SNPs were nominally associated, of which two displayed significant SNP x statin interaction. However, these two SNPs were not associated in the combined analysis or the statin free analysis. While we could not confirm statin interaction of these two SNPs in *CAP*, 17 other SNPs reached nominal significance here.

In the sex-stratified analyses of the credible set, 589 SNPs were associated in men, while only 101 did so in women. This resulted in 40 SNPs with significant sex interaction, which had all sex-specific effect differences and effect directions

We detected no co-localization between our PCSK9 signal and any lipid trait, CAD, or eQTLs of any gene. The highest PP4 was observed for a lncRNA (CTC-471J1.10, PP4=0.304 in thyroid tissue), while the posterior probability for co-localization with PPP2R1A was 0.237 in colon tissue.

Heritability and genetic correlation

The five independent and genome-wide significant SNPs of the combined setting explained about 3.5% of PCSK9 variance (4.0% in the statin-free). By expanding the significance threshold to 15 SNPs with FDR<1%, the explained variance increased to 6.7% (7.4% with 13 SNPs in the statin-free setting). Using the online tool LDHub (24) we estimated the heritability of PCSK9 levels and genetic correlations with other lipid traits (25). The estimated heritability was h^2=10.3% (standard error 3.7%; 12.6% with SE 4.8% in the statin-free setting), which is in a similar magnitude as that of other lipid traits according to LDHub (LDL-C: 10.7%; TC: 13.7%). There was a significant genetic correlation between PCSK9 and TC (r_g=0.39, p=0.005) and LDL-C (r_g=0.34, p=0.020), but not with HDL-C or TG (see Table 2, data from Teslovich et al. (26)). We also analyzed lipid metabolites of Kettunen et al. (27) for genetic correlation and found 29 of the 107 traits significantly correlated (see Table S12). The strongest correlation was observed with free cholesterol in large LDL (r_g=0.82, p=0.002). There was also a significant correlation with Apolipoprotein B (r_g=0.54, p=0.025).

In our GWAS catalog look-up, we found 2,554 variants associated with TC, LDL-C, HDL-C or TG. For 2,099 of these SNPs distributed over 368 distinct loci, PCSK9 association statistics were available in our study. Using our results of combined and statin-free analyses, we detected 30 genome-wide, 16 suggestive, and 527 nominal significant SNPs (at 3, 1, and 102 independent loci, respectively). Therefore, 28% (106 out of 368) of known lipid loci are also associated with PCSK9 on at least a nominal level (enrichment p=8.52×10^{-50}). There were 31% of all LDL-C loci co-associated, whereas for HDL-C only 21% were co-associated (enrichment

p=1.64x10^{-31} and p=1.46x10^{-16}, respectively). Statistics for all 2,099 SNPs are given in Table S13.

Mendelian Randomization Analyses

We aimed at distinguishing between direct and indirect effect mediated by LDL-C of PCSK9 on CAD using Mendelian Randomization (28). To rule out a possible reverse causality, we first performed a bidirectional MR of PCSK9 and LDL-C levels (see Methods and Figure S8).

As expected, we found a significant causal effect of PCSK9 on LDL-C (β=1.770, p=4.44x10^{-144}, see Table 3 and Figure S9) when using the three independent SNPs of the *PCSK9* locus as instruments for PCSK9 plasma levels. The opposite direction did not reach significance when 24 instruments of LDL-C were used ($\beta_{m=24} = 0.015, p = 0.179$). For sensitivity, we repeated the analysis with the subset of SNPs explaining more than 0.5% of LDL-C variance. Here, a significant reverse causality was observed ($\beta_{m=3} = 0.029, p = 0.016$) suggesting feedback loops between PCSK9 and LDL-C. We obtained the same results when restricting to statin free subjects ($\beta_{m=24} = 0.022, p = 0.071, \beta_{m=3} = 0.043, p = 0.001$).

In our mediation analysis of PCSK9 on CAD, we estimated an indirect effect as product of the PCSK9 effect on LDL-C and the LDL-C effect on CAD: β_{indir}=0.632 (p=5.44x10^{-7}). The direct causal effect is the difference of total and indirect effect of PCSK9 on CAD, which was significant too (β_{dir}==0.364, p=0.045). Similar results were obtained when restricting to statin-free subjects or using other instruments for LDL-C (see Table S14). We conclude that 63% of the causal PCSK9 effect on CAD is mediated by LDL-C, but there is also a significant direct effect.

Discussion

We performed the first genome-wide meta-analysis of PCSK9 blood levels in a sample size of 12,721, which is six-times that of the largest single study published so far. We detected three novel loci with genome-wide significance and added further independent variants to the already known locus at 1p32.3. Based on these variants, we estimated the direct and LDL-C mediated causal effects of PCSK9 on CAD showing that both are significant and in the same order of magnitude.

We confirmed the known association at 1p32.3 at the *PCSK9* gene locus. By our fine-mapping approach, we identified three independent signals representing different modes of action. The lead SNP is the known missense mutation R46L increasing the rate of PCSK9s degradation (5,29,9). In a phenome-wide association study, this SNP was linked to disorders of lipid metabolism (hyperlipidemia, hypercholesterolemia) (30). The second independent variant was located in the fifth intron of PCSK9 and was predicted to modify RNA splicing. One possible functional mechanism could be that the modified splicing results in reduced mRNA levels. This is supported by our co-localization analyses, in which the conditioned estimates of this hit co-localized with eQTLs of PCSK9 gene expression in whole blood and testis. The third credible set included three plausible candidates. While the lead SNP rs11206510 is known for its association to CAD (20), the highest CADD score was observed for rs11583680, coding for a missense mutation in PCSK9 (A53V). Another causal candidate was rs45448095, which was reported to be in high LD with an in-frame leucine insertion associated with lower LDL-C levels (31,32). The mentioned leucine insertion was not included in our study data due to low MAF. The SNP itself was located in the 5'UTR region regulating transcription according to CADD annotation. Analyzing ethnic heterogeneity of this locus in 300 African-Americans, we observed a different lead SNP for this locus, rs28362263. It is a loss-of-function mutation that increases the chance for furin cleavage by which large parts of the catalytic domain are lost

(21). We observed a positive SNP effect here, which could be explained by the antibody used in the ELISA that detects both the furin-cleaved and intact versions of PCSK9 in the circulation and hence only partially relates to PCSK9 activity. This variant is not present in our GWAMA, since it is monomorphic in Europeans. Conversely, our independent variant rs11591147 was monomorphic in African-Americans demonstrating the necessity of further trans-ethnic analyses of this locus with larger sample sizes.

A novel association was found at 2p24.1 around the *APOB* gene, which has been reported for associations with TC, LDL-C, HDL-C and TG in GWAMAs (19) and with disorders of lipid metabolism in a PheWAS analysis (30). Of note, we only observed co-localization with HDL-C and TG. Due to access to ApoB100 data in *LIFE-Heart*, we performed co-localization analysis with this trait but could not find any evidence for co-localization (33), indicating that our hit does not act via ApoB100 plasma levels (data not shown). The credible set of the lead variant contained 69 SNPs, of which the missense mutation P2739L affecting the secondary structure of ApoB100 is the most plausible causal variant (34). ApoB100 is targeted by biotherapeutics and small molecule drugs (35) to treat familiar hypercholesterolemia, a genetic disorder caused by mutations in the *PCSK9* gene, among others. PCSK9 binds to ApoB100 in LDL-C (36,37). This mutation could change the binding affinity resulting in higher levels of free PCSK9, i.e. more available PCSK9 antibody binding sites in the immunoassays. More experimental data are required to corroborate this mechanism in more detail.

We detected another genome-wide significant signal at 19p13.11 with 171 SNPs in the 99% credible set of the lead variant. This locus was described for its associations with plasma lipoprotein concentrations (38), and both *SUGP1* and *TM6SF2* were suggested as causal genes (39). However, the gene expression associations of both genes was not co-localized with our PCSK9 signal. Subsequent studies identified the E167K missense mutation in *TM6SF2* as the most likely causal variant for observed associations (40,41). A PheWAS of this variant had

detected associations to various liver diseases, decreased neutrophil count, haemoglobin traits, and platelet traits (42). In addition, this variant was shown to increase the SREBP-1c expression levels (43), which is an important transcription factor of HMGCR, LDLR, and PCSK9. This might also explain the observed trend to a statin interaction. Functional studies in human hepatoma cells (44) and *Tm6sf2* KO-mice (45) confirmed that *TM6SF2* inhibition influences the secretion and/or lipidation of triglyceride-rich lipoproteins by the liver. In a model of 3D spheroids from primary human hepatocytes it was shown that *TM6SF2* E167K increases hepatocyte fat content by reducing ApoB particle secretion, and that genes of the cholesterol metabolism were differentially expressed compared to wild type *TM6SF2* (46). Here, we observed that three lipid traits (triglyceride, cholesterol and LDL-cholesterol) were co-localized with our PCSK9 association signal. Of note, Smagris and co-workers (45) observed a decrease in plasma PCSK9 concentration in *Tm6sf2*$^{-/-}$ mice compared to wild-type mice, but this reduction did not reach statistical significance. Overall, these observations suggest that TM6SF2 influences the secretion of PCSK9 by hepatocytes and that TM6SF2 could be a potential target for future drug development. The mechanism responsible for the involvement of TM6SF2 in PCSK9 secretion remains to be further elucidated.

Finally, we found an association at 19q13.41. The credible set of the lead variant was large so that it was impossible to pin down the causal variant and candidate gene with sufficient certainty. Possible candidate genes were from the *FPR* family and *PPP2R1A*. FPRs have been linked to HDL-C levels (23), as their activation of neutrophils is canceled by ApoA1 (47). We compared our in-house annotation with online-tool FUMA (Functional Mapping and Annotation) (48), which listed eight *ZNFs* as possible candidate genes, which all were included in the co-localization analyses as nearby genes. However, our association signal did not co-localize, neither with *FPRs*, *ZNFs*, nor with HDL-C levels. *PPP2R1A* codes for the structural subunit of the protein phosphatase 2 (PP2A), necessary as scaffold for the regulatory and

catalytic subunits of PP2A (49). It lies in a shared pathway with APOB enhancing platelet aggregation ("Platelet sensitization by LDL", pathway ID R-HSA-432142). In addition, PP2A interacts with SREBP-2 and changes its phosphorylation status. It is an important modification to enable SREBP-2 to act as transcription factor for LDLR and PCSK9 (50,51). The observed sex- and statin-interactions further supports *PPP2R1A* as candidate gene, as it is also a putative biomarker for endometrial cancer (52), and has been discussed to interact with statin (53). However, this mechanism requires further experimental validation.

Based on the newly identified variants, we analyzed the causal relationships between PCSK9, LDLC and CAD in more detail by Mendelian Randomization Analyses. For bivariate MR approaches, it is recommended to use instruments that have a direct biological relationship to the respective exposure, which can be assumed for the three independent variants of *PCSK9*. Using these instruments, we found a significant causal effect of PCSK9 on LDL-C. However, for the reverse direction of LDL-C on PCSK9 the choice of instruments was more challenging. In a first attempt, we used 24 SNPs reported to be associated with LDL-C, with sufficient quality in our data and not associated with PCSK9 levels. No significant causal effect was detected with this choice. In a second attempt, we used the three strongest LDL-C signals in terms of explained LDL-C variance, namely variants at sortilin 1 (*SORT1*), *APOE* and *LDLR*. This resulted in a weak causal effect. However, a pleiotropic effect of those instruments cannot completely be excluded. *SORT1* is not only involved in LDL-metabolism, but also in PCSK9 secretion (54), and both APOE and LDLR mediate the effect of PCSK9 on hepatic lipid production (55).

We also observed a strong mediating effect of LDL-C on the causal relationship of PCSK9 and CAD explaining about 63% of the total effect of PCSK9 on CAD. However, the direct effect reached significance too, which is in line with previous findings of PCSK9 as an independent risk factor of CAD risk (56). Underlying mechanisms of this direct effect need to be

investigated in further studies, also taking into account the particle sizes of LDL and LDL subclasses.

In conclusion, we detected four independent loci associated with PCSK9 levels. The strongest hit at the *PCSK9* locus showed both, a considerable locus heterogeneity as demonstrated by the identification of three independent variants and an ethnic heterogeneity requiring further investigations. While no clear candidate gene could be assigned to the 19q13.41 locus, the remaining two hits are in plausible genes involved in lipid metabolism. This suggests that PCSK9 levels is also a polygenic trait not only regulated by variants in its gene, but also by genes effecting secretion of or binding affinity to PCSK9. Our Mendelian Randomization analysis suggests that the causal effect of PCSK9 on LDL-C levels is much larger than a possible reverse direction. Finally, the causal effect of PCSK9 on CAD is mainly mediated by LDL-C, but an independent direct effect also occurs.

Materials and Methods

Data Availability Statement

Data related to this project, including summary level data from the meta-analyses, can be found online at https://www.health-atlas.de/studies/44.

Studies

Five independent studies contributed to this GWAMA: *LIFE-Heart* (13), *LIFE-Adult* (14), *LURIC* (15), *CAP* (16), and *TwinGene* (17), reaching a total sample size of 12,721 participants of European ancestry (n=10,186 without statin treatment), and 300 participants of African-American ancestry (*CAP*, all without statin treatment). Given the small sample size for African-Americans, we refrained from carrying out a trans-ethnic GWAMA and instead performed a GWAMA in Europeans, and then validated the detected GWAMA lead SNPs in African-Americans. Detailed study characteristics are provided in the Supplemental Data and in Table S1.

All studies meet the ethical standards of the Declaration of Helsinki and were approved by relevant institutional review boards. Written informed consent including agreement with genetic analyses was obtained from all participants in all studies.

PCSK9 measurement

In *LIFE-Heart* and *LURIC*, total PCSK9 levels were analyzed in plasma samples using a commercial assay (Quantikine Human PCSK9 immunoassay, R&D Systems). Data were log-transformed for further analyses (9,57). In *LIFE-Adult*, relative quantification of EDTA plasma PCSK9 was assessed by normalized protein expression units from the Olink target 96 multiplex platform (Olink Proteomics AB; CVD panel III). In *CAP*, a colorimetric ELISA assay using an AX213 antibody was used to measure PCSK9 in plasma (measured by BG Medicine in

collaboration with Merck) (12), while in *TwinGene* serum PCSK9 concentration was determined as described in (58).

Genotyping, imputation, and study level quality control

Genotyping arrays and pre-imputation QC per study are shown in Table S1. Sample and SNP quality control (QC) of genotype raw data was performed at the discretion of the single study analysts (see Supplemental Data). SNP-QC measures included call rate, violation of Hardy-Weinberg equilibrium and minor allele frequency or monomorphism. For imputation, all studies used 1000 Genomes Phase 3 (18).

Analysis plan

Single study analysts were requested to follow a standardized analysis plan sent to all contributing studies. Genome-wide associations were estimated using linear regression analyses assuming additive genetic models (SNP dosage), adjusting for sex, age, statin treatment and current smoking. Principal components were included in the regression model if considered necessary by the respective study analyst (*CAP*). X-chromosomal SNPs were analyzed assuming total X inactivation (i.e. male genotypes were coded as A=0, B=2 and female genotypes were coded as AA=0, AB=1 and BB=2). Association analyses were performed with PLINK2 (*LIFE-Adult*, *LIFE-Heart*, and *LURIC*) or mach2qtl (*CAP*). Due to dizygotic twin pairs in *TwinGene*, associations were calculated with a mixed linear model adjusting for the estimated genetic relationship matrix. As recommended, the genetic relationship matrix was estimated leaving out the chromosome under analysis (GCTA MLMA LOCO (59,60)).

We also requested sub-group analyses of subjects stratified for statin medication on genome-wide scale, if available (n=10,186 without statin treatment). The same regression model except for considering statin treatment as covariate and the same software were used for that purpose.

Meta-Analysis

FileQC

Single study GWAS results were harmonized centrally by a pre-meta QC of summary files per study. First, we excluded SNPs with missing values in allelic information (effect allele, effect allele frequency) or statistics (beta estimate, standard error, imputation quality score). Further SNP filtering criteria were minor allele frequency (MAF) <1%, imputation info score <0.5, and minor allele count (MAC) ≤6.

We used the R package "EasyQC" (61) to filter SNPs with mismatching alleles or chromosomal position with respect to the reference (1000 Genomes Phase 3, Version 5 (2015) for European samples (18)), and with high deviation of study to reference allele frequency (difference >0.2). Finally, the alleles were harmonized so that the same effect allele was used in all studies.

Meta GWAS

For the meta-analysis, single study results were combined using a fixed-effect model assuming homogenous genetic effects across studies. Heterogeneity of study results was assessed by I^2 statistics. We filtered SNPs with minimum of imputation info-score across studies<0.5, I^2>90%, or number of studies with association statistics <2. The genome-wide and suggestive significance level was set to $\alpha_{gw} = 5 \times 10^{-8}$ and $\alpha_{sug} = 1 \times 10^{-6}$, respectively. SNPs were considered for down-stream analyses if reaching at least suggestive significance. A locus was defined as the set of associated SNPs in physical proximity (+/- 500 kb) to the respective regional lead SNP. Pairwise LDs were calculated using data from 1000 Genomes Phase 3, Version 5 (2015) for European samples (18).

We annotated all variants reaching at least suggestive significance with the following bioinformatics resources: nearby Ensembl genes (+/-250 kb) (62), variants reported in the

GWAS Catalog in linkage disequilibrium (LD, r^2>0.3) (63), and expression quantitative traits in LD (r^2>0.3) (64–68). We then used the nearby genes and eQTL genes to test for pathway enrichments (retrieved from DOSE (69) and Reactome (70)).

Secondary Analyses

Conditional & Joint Analyses (COJO)

We used the tool GCTA (version 1.92.0beta3) (60) to test for secondary signals at each locus. First, we applied GCTA's stepwise model selection algorithm (cojo-slct) to identify independent variants. In case of more than one independent variant, we performed conditional association analysis adjusting for the respective other independent variants (cojo-cond) (71), resulting in conditional statistics for all SNPs at the locus (fixed genomic range). As reference panel we used the genetic data of the combination of *LIFE-Adult* and *LIFE-Heart* (n=13,369).

We looked-up all independent SNPs in the *CAP* African-American samples (before statin-treatment) and checked for concordant effect direction. In addition, we also analyzed the *PCSK9* locus in African-Americans in more detail and tested for ethnicity specific causal variants.

Credible Set Analyses

For each independent variant, we performed a credible set analyses to determine the likely causal SNPs within a region of +/-500 kb of the respective SNP (72,73). In case of more than one independent signal per locus, we used the respective conditional statistics. R-package "gtx" was used to derive Approximate Bayes Factors (ABF) from the (conditional) effect estimates and standard errors. Standard deviation priors were chosen per locus in dependence on the respective distribution of effect sizes (Difference of 97.5 and 2.5 percentile divided by 2*1.96). Results varied between 0.007 (locus 19q13.41) and 0.014 (2p24.1). The ABFs were used to calculate the posterior probability that a variant drives the association signal. We ordered

variants by their posterior probability and calculated for each SNP the cumulative posterior in descending order until 99% was achieved, resulting in a set of SNPs containing the causal variant with 99% certainty.

We annotated the variants of the 99% credible sets (CS) with the above mentioned bioinformatics resources. We also added CADD scores as measure of deleteriousness (74) and considered variants with CADD>15 and high posterior probability (PP) as possible causal SNPs. A gene was considered a plausible candidate if containing such a likely causal variant.

Interaction Analyses

Finally, we performed statin- and sex-stratified analyses for all SNPs within the 99% CSs. We compared effect sizes of SNPs by testing their differences against zero (t-test, pre-filtering for SNPs associated at least on nominal level in one of the strata, followed by Benjamini & Hochberg FDR 5%) (75). For the sex-stratified analysis, we used data from *LIFE* and *TwinGene* (n=5,345 and n=4,880 for men and women, respectively). For the statin-stratified analysis, we used data from participants without statin treatment (subsets of *LIFE-Heart*, *LIFE-Adult*, and *LURIC*; all *TwinGene*, n= 9,623) and with statin medication (subsets of *LIFE-Heart*, and *LIFE-Adult*; n=1,589). Since *CAP* assessed PCSK9 levels prior and during statin treatment, we calculated associations with the respective differences in order to validate observed SNP x statin interactions (n=563).

Co-localization Analyses

For each independent locus, we performed a pairwise co-localization test (76) between our GWAS results and literature GWAS results for total cholesterol (TC), LDL-C, high-density lipoprotein-cholesterol (HDL-C) and triglycerides (TG) (19), cis-eQTLs (64) (GTEx v7), and CAD (20). In more detail, the co-localization method evaluates whether two trait associations share the same causal variant (76). Five hypotheses are tested (H_0: no association with either

trait; H_1: association with trait 1, not with trait 2; H_2: association with trait 2, not with trait 1; H_3: association with trait 1 and 2, two independent SNPs; H_4: association with trait 1 and 2, shared SNP). As threshold for co-localization a posterior probability of ≥0.75 for H_4 was applied. As before, we used all SNPs within the 500kb window around the independent lead SNPs. For the loci 1p32.3 and 2p24.1, we used eQTL-statistics of the candidate genes only (*PCSK9*, and *APOB*) as determined by credible set analysis. Since multiple independent variants were detected at the 1p32.3 locus, we performed co-localization analyses for each of them separately using conditional estimates. This allows identification of signals likely acting via eQTLs. For the loci 19p11.13 and 19q13.41, we analyzed all genes of our annotation for co-localizing eQTLs (SNPs in the 99% credible sets with CADD>10, genes nearby (+/-250 kb) or known cis effects). Therefore, we tested 40 genes at 19p13.11, and 78 genes at 19q13.41.

Heritability, Genetic Correlation & Look-up of lipid loci

We estimated the heritability of PCSK9 in three modes: using only the genome-wide significant and independent SNPs (n=6 SNPs), using SNPs with FDR<1% (n=15 SNPs, pairwise LD r^2<0.1), and using the online tool LDHub (24) (n= 1,085,662 SNPs). We repeated this in the statin-free set (n=5, n=13, and n=1,087,910 SNPs in the three modes, respectively). In addition, we estimated the genetic correlation (25) of PCSK9 with lipid traits (26) and metabolites (27) and we checked known lipid loci for association with PCSK9 levels. For this analysis, we searched the GWAS Catalog (63) for loci associated with TC (trait ID in the experimental factor ontology: EFO_0004574), LDL-C (EFO_0004611), HDL-C (EFO_0004612) and TG (EFO_0004530). We downloaded all variants reported for these four traits (download date 23.01.2020) and excluded those not achieving genome-wide significance and duplicates (m=2554 SNPs after filtering). Due to filter criteria in the meta-analysis, no PCSK9 statistics were available for 458 of these SNPs.

Mendelian Randomization

By Mendelian Randomization analyses, we aimed at determining the causal effect of PCSK9 on CAD. Moreover, we distinguish between direct effects and an indirect effect mediated by LDL-C as suggested by (28) (see Figure S8 for a graphical visualization). We first performed a bidirectional Mendelian randomization of PCSK9 and LDL-C levels in order to rule out a possible reverse causality.

As instruments for PCSK9, we used the three independent variants of *PCSK9* (G1). For LDL-C, we used 24 LDL-C associated SNPs not associated with PCSK9 levels (G2). For sensitivity, we repeated the analysis with three SNPs explaining more than 0.5% of LDL-C variance each and using the summary statistics of the statin-free subset. Summary statistics for LDL-C and CAD were obtained from Surakka et al. (19) and Nikpay et al. (20), respectively.

Thus, we estimated four causal effects $\alpha = \hat{\beta}_{IVW}$(PCSK9→LDL-C), $\beta = \hat{\beta}_{IVW}$(LDL-C→CAD), $\gamma = \hat{\beta}_{IVW}$(LDL-C→PCSK9), and $\tau = \hat{\beta}_{IVW}$(PCSK9→CAD), using G1 and G2 as instruments by inverse-variance weighting (IVW) as implemented in the R package "MendelianRandomization".

Then we estimated the indirect effect as product of α and β, and compared it to τ by formal t-statistics to test for differences:

$$\hat{\beta}_{indirect}(\text{PCSK9} \rightarrow \text{CAD}) = \alpha \times \beta,$$

$$SE\left(\hat{\beta}_{indirect}\right) = \sqrt{\alpha^2 \times SE(\beta) + \beta^2 \times SE(\alpha)}$$

$$\hat{\beta}_{direct}(\text{PCSK9} \rightarrow \text{CAD}) = \tau - \hat{\beta}_{indirect}(\text{PCSK9} \rightarrow \text{CAD}),$$

Acknowledgements

TwinGene was supported by grants from the Leducq Foundation (13CVD03), Swedish Research Council (12660) and the Swedish Heart-Lung Foundation (201202729). We thank Robert John Konrad and his team at Ely Lilly and Company for the PCSK9 measurements in serum samples from the *TwinGene* cohort.

LIFE-Heart and *LIFE-Adult* are funded by the Leipzig Research Center for Civilization Diseases (LIFE). LIFE is funded by means of the European Union, by the European Regional Development Fund (ERDF) and by means of the Free State of Saxony within the framework of the excellence initiative. PCSK9 measurements in *LIFE-Adult* were funded by the HI-MAG project "Serum proteome biomarkers as mediators of cardiometabolic disease development" of the Medical Faculty of the University Leipzig and the Helmholtz Zentrum München. This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (grant # 01ZX1906B). We thank Sylvia Henger for data quality control of *LIFE-Adult* and *LIFE-Heart*, Kay Olischer and Annegret Unger for technical assistance regarding *LIFE-Heart*, and Kerstin Wirkner for running the *LIFE-Adult* study center. We thank all study participants of the *LIFE-Adult* study whose personal dedication and commitment have made this project possible. *LIFE-Adult* genotyping (round 3) was done at the Cologne Center for Genomics (CCG, University of Cologne, Peter Nürnberg and Mohammad R. Toliat). For *LIFE-Adult* genotype imputation, compute infrastructure provided by ScaDS (Dresden/Leipzig Competence Center for Scalable Data Services and Solutions) at the Leipzig University Computing Centre was used.

LURIC was supported by the 7th Framework Program (integrated project AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739) of the European Union. We thank the *LURIC* study team who were either temporarily or permanently involved

in patient recruitment as well as sample and data handling, in addition to the laboratory staff at the Ludwigshafen General Hospital and the Universities of Freiburg and Ulm, Germany.

CAP was supported by NIH grant NIH U01 HL069757 and a Merck Investigator Initiated Research Grant.

Conflicts of interest

Winfried März is employed with SYNLAB Holding Deutschland GmbH. Grants and personal fees from AMGEN, BASF, Sanofi, Siemens Diagnostics, Aegerion Pharmaceuticals, Astrazeneca, Danone Research, Numares, Pfizer, Hoffmann LaRoche: personal fees from MSD, Alexion; grants from Abbott Diagnostics, all outside the submitted work.

Marcus Kleber received lecture fees from Bayer and SYNLAB outside the submitted work and is employed with SYNLAB Holding Deutschland GmbH.

Hubert Scharnagl received grants and personal fees from AMGEN, Sanofi, Abbott, numares, and Unilever, all outside the submitted work.

REFERENCES

1. Zhang, D.-W., Garuti, R., Tang, W.-J., Cohen, J.C. and Hobbs, H.H. (2008) Structural requirements for PCSK9-mediated degradation of the low-density lipoprotein receptor. *Proc. Natl. Acad. Sci. U.S.A.,* **105,** 13045–13050, 10.1073/pnas.0806312105.

2. Cannon, C.P., Cariou, B., Blom, D., McKenney, J.M., Lorenzato, C., Pordy, R., Chaudhari, U. and Colhoun, H.M. (2015) Efficacy and safety of alirocumab in high cardiovascular risk patients with inadequately controlled hypercholesterolaemia on maximally tolerated doses of statins: the ODYSSEY COMBO II randomized controlled trial. *Eur. Heart. J.,* **36,** 1186–1194, 10.1093/eurheartj/ehv028.

3. Sabatine, M.S., Giugliano, R.P., Wiviott, S.D., Raal, F.J., Blom, D.J., Robinson, J., Ballantyne, C.M., Somaratne, R., Legg, J. and Wasserman, S.M. *et al.* (2015) Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. *N. Engl. J. Med.,* **372,** 1500–1509, 10.1056/NEJMoa1500858.

4. Abifadel, M., Varret, M., Rabès, J.-P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L. and Erlich, D. *et al.* (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat. Genet.,* **34,** 154–156, 10.1038/ng1161.

5. Dron, J.S. and Hegele, R.A. (2017) Complexity of mechanisms among human proprotein convertase subtilisin-kexin type 9 variants. *Curr. Opin. Lipidol.,* **28,** 161–169, 10.1097/MOL.0000000000000386.

6. Kotowski, I.K., Pertsemlidis, A., Luke, A., Cooper, R.S., Vega, G.L., Cohen, J.C. and Hobbs, H.H. (2006) A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. *Am. J. Hum. Genet.,* **78,** 410–422, 10.1086/500615.

7. Ding, K. and Kullo, I.J. (2008) Molecular population genetics of PCSK9: a signature of recent positive selection. *Pharmacogenet. Genomics,* **18,** 169–179, 10.1097/FPC.0b013e3282f44d99.

8. Wu, Y., Waite, L.L., Jackson, A.U., Sheu, W.H.-H., Buyske, S., Absher, D., Arnett, D.K., Boerwinkle, E., Bonnycastle, L.L. and Carty, C.L. *et al.* (2013) Trans-ethnic fine-mapping of lipid loci identifies population-specific signals and allelic heterogeneity that increases the trait variance explained. *PLoS Genet.* First published on March 21, 2013, 10.1371/journal.pgen.1003379.

9. Pott, J., Schlegel, V., Teren, A., Horn, K., Kirsten, H., Bluecher, C., Kratzsch, J., Loeffler, M., Thiery, J. and Burkhardt, R. *et al.* (2018) Genetic Regulation of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) Plasma Levels and Its Impact on Atherosclerotic Vascular Disease Phenotypes. *Circ. Genom. Precis. Med.,* **11,** e001992, 10.1161/CIRCGEN.117.001992.

10. Chernogubova, E., Strawbridge, R., Mahdessian, H., Mälarstig, A., Krapivner, S., Gigante, B., Hellénius, M.-L., Faire, U. de, Franco-Cereceda, A. and Syvänen, A.-C. *et al.* (2012) Common and low-frequency genetic variants in the PCSK9 locus influence circulating PCSK9 levels. *Arterioscler. Thromb. Vasc. Biol.,* **32,** 1526–1534, 10.1161/ATVBAHA.111.240549.

11. Enkhmaa, B., Kim, K., Zhang, W., Prakash, N., Truax, K., Anuurad, E. and Berglund, L. (2020) PCSK9 in African Americans and Caucasians in Relation to Lp(a) Level, Apo(a) Size and Heritability. *J. Endocr. Soc.,* **4,** bvaa073, 10.1210/jendso/bvaa073.

12. Theusch, E., Medina, M.W., Rotter, J.I. and Krauss, R.M. (2014) Ancestry and other genetic associations with plasma PCSK9 response to simvastatin. *Pharmacogenet. Genomics,* **24,** 492–500, 10.1097/FPC.0000000000000081.

13. Scholz, M., Henger, S., Beutner, F., Teren, A., Baber, R., Willenberg, A., Ceglarek, U., Pott, J., Burkhardt, R. and Thiery, J. (2020) Cohort profile: The Leipzig Research Center for Civilization Diseases-Heart study (LIFE-Heart). *Int. J. Epidemiol.,* 10.1093/ije/dyaa075.

14. Loeffler, M., Engel, C., Ahnert, P., Alfermann, D., Arelin, K., Baber, R., Beutner, F., Binder, H., Brähler, E. and Burkhardt, R. *et al.* (2015) The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health,* **15,** 691, 10.1186/s12889-015-1983-z.

15. Winkelmann, B.R., März, W., Boehm, B.O., Zotz, R., Hager, J., Hellstern, P. and Senges, J. (2001) Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics,* **2,** S1-73, 10.1517/14622416.2.1.S1.

16. Simon, J.A., Lin, F., Hulley, S.B., Blanche, P.J., Waters, D., Shiboski, S., Rotter, J.I., Nickerson, D.A., Yang, H. and Saad, M. *et al.* (2006) Phenotypic predictors of response to simvastatin therapy among African-Americans and Caucasians: the Cholesterol and Pharmacogenetics (CAP) Study. *Am. J. Cardiol.,* **97,** 843–850, 10.1016/j.amjcard.2005.09.134.

17. Magnusson, P.K.E., Almqvist, C., Rahman, I., Ganna, A., Viktorin, A., Walum, H., Halldner, L., Lundström, S., Ullén, F. and Långström, N. *et al.* (2013) The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res. Hum. Genet.,* **16,** 317–329, 10.1017/thg.2012.104.

18. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A. and Abecasis, G.R. (2015) A global reference for human genetic variation. *Nature,* **526,** 68–74, 10.1038/nature15393.

19. Surakka, I., Horikoshi, M., Mägi, R., Sarin, A.-P., Mahajan, A., Lagou, V., Marullo, L., Ferreira, T., Miraglio, B. and Timonen, S. *et al.* (2015) The impact of low-frequency and rare variants on lipid levels. *Nat. Genet.,* **47,** 589–597, 10.1038/ng.3300.

20. Nikpay, M., Goel, A., Won, H.-H., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P. and Hopewell, J.C. *et al.* (2015) A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.,* **47,** 1121–1130, 10.1038/ng.3396.

21. Benjannet, S., Rhainds, D., Hamelin, J., Nassoury, N. and Seidah, N.G. (2006) The proprotein convertase (PC) PCSK9 is inactivated by furin and/or PC5/6A: functional consequences of natural mutations and post-translational modifications. *J. Biol. Chem.* First published on August 15, 2006, 10.1074/jbc.M606495200.

22. Hoffmann, T.J., Theusch, E., Haldar, T., Ranatunga, D.K., Jorgenson, E., Medina, M.W., Kvale, M.N., Kwok, P.-Y., Schaefer, C. and Krauss, R.M. *et al.* (2018) A large electronic health record-based genome-wide study of serum lipids. *Nat. Genet.,* **50,** 401–413, 10.1038/s41588-018-0064-5.

23. Klarin, D., Damrauer, S.M., Cho, K., Sun, Y.V., Teslovich, T.M., Honerlaw, J., Gagnon, D.R., DuVall, S.L., Li, J. and Peloso, G.M. *et al.* (2018) Genetics of blood lipids among ~300,000 multi-ethnic

participants of the Million Veteran Program. *Nat. Genet.,* **50,** 1514–1523, 10.1038/s41588-018-0222-9.

24. Zheng, J., Erzurumluoglu, A.M., Elsworth, B.L., Kemp, J.P., Howe, L., Haycock, P.C., Hemani, G., Tansey, K., Laurin, C. and Pourcain, B.S. *et al.* (2017) LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics,* **33,** 272–279, 10.1093/bioinformatics/btw613.

25. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., Duncan, L., Perry, J.R.B., Patterson, N. and Robinson, E.B. *et al.* (2015) An atlas of genetic correlations across human diseases and traits. *Nat. Genet.,* **47,** 1236–1241, 10.1038/ng.3406.

26. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I. and Willer, C.J. *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature,* **466,** 707–713, 10.1038/nature09270.

27. Kettunen, J., Demirkan, A., Würtz, P., Draisma, H.H.M., Haller, T., Rawal, R., Vaarhorst, A., Kangas, A.J., Lyytikäinen, L.-P. and Pirinen, M. *et al.* (2016) Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat. Commun.,* **7,** 11122, 10.1038/ncomms11122.

28. Burgess, S., Daniel, R.M., Butterworth, A.S. and Thompson, S.G. (2015) Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int. J. Epidemiol.,* **44,** 484–495, 10.1093/ije/dyu176.

29. Dewpura, T., Raymond, A., Hamelin, J., Seidah, N.G., Mbikay, M., Chrétien, M. and Mayne, J. (2008) PCSK9 is phosphorylated by a Golgi casein kinase-like kinase ex vivo and circulates as a phosphoprotein in humans. *FEBS J.,* **275,** 3480–3493, 10.1111/j.1742-4658.2008.06495.x.

30. Safarova, M.S., Satterfield, B.A., Fan, X., Austin, E.E., Ye, Z., Bastarache, L., Zheng, N., Ritchie, M.D., Borthwick, K.M. and Williams, M.S. *et al.* (2019) A phenome-wide association study to discover pleiotropic effects of PCSK9, APOB, and LDLR. *NPJ. Genom. Med.* First published on February 11, 2019, 10.1038/s41525-019-0078-7.

31. Yue, P., Averna, M., Lin, X. and Schonfeld, G. (2006) The c.43_44insCTG variation in PCSK9 is associated with low plasma LDL-cholesterol in a Caucasian population. *Hum. Mutat.,* **27,** 460–466, 10.1002/humu.20316.

32. Abifadel, M., Rabès, J.-P., Devillers, M., Munnich, A., Erlich, D., Junien, C., Varret, M. and Boileau, C. (2009) Mutations and polymorphisms in the proprotein convertase subtilisin kexin 9 (PCSK9) gene in cholesterol metabolism and disease. *Hum. Mutat.,* **30,** 520–529, 10.1002/humu.20882.

33. Dittrich, J., Beutner, F., Teren, A., Thiery, J., Burkhardt, R., Scholz, M. and Ceglarek, U. (2019) Plasma levels of apolipoproteins C-III, A-IV, and E are independently associated with stable atherosclerotic cardiovascular disease. *Atherosclerosis,* **281,** 17–24, 10.1016/j.atherosclerosis.2018.11.006.

34. Wojczynski, M.K., Gao, G., Borecki, I., Hopkins, P.N., Parnell, L., Lai, C.-Q., Ordovas, J.M., Chung, B.H. and Arnett, D.K. (2010) Apolipoprotein B genetic variants modify the response to fenofibrate: a GOLDN study. *J. Lipid. Res.,* **51,** 3316–3323, 10.1194/jlr.P001834.

35. Finan, C., Gaulton, A., Kruger, F.A., Lumbers, R.T., Shah, T., Engmann, J., Galver, L., Kelley, R., Karlsson, A. and Santos, R. *et al.* (2017) The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.,* **9,** 10.1126/scitranslmed.aag1166.

36. Kosenko, T., Golder, M., Leblond, G., Weng, W. and Lagace, T.A. (2013) Low density lipoprotein binds to proprotein convertase subtilisin/kexin type-9 (PCSK9) in human plasma and inhibits PCSK9-mediated low density lipoprotein receptor degradation. *J. Biol. Chem.,* **288,** 8279–8288, 10.1074/jbc.M112.421370.

37. Lagace, T.A. (2014) PCSK9 and LDLR degradation: regulatory mechanisms in circulation and in cells. *Curr. Opin. Lipidol.,* **25,** 387–393, 10.1097/MOL.0000000000000114.

38. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L. and Mora, S. *et al.* (2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.,* **45,** 1274–1283, 10.1038/ng.2797.

39. Kim, M.J., Yu, C.-Y., Theusch, E., Naidoo, D., Stevens, K., Kuang, Y.-L., Schuetz, E., Chaudhry, A.S. and Medina, M.W. (2016) SUGP1 is a novel regulator of cholesterol metabolism. *Hum. Mol. Genet.,* **25,** 3106–3116, 10.1093/hmg/ddw151.

40. Holmen, O.L., Zhang, H., Fan, Y., Hovelson, D.H., Schmidt, E.M., Zhou, W., Guo, Y., Zhang, J., Langhammer, A. and Løchen, M.-L. *et al.* (2014) Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. *Nat. Genet.,* **46,** 345–351, 10.1038/ng.2926.

41. Kozlitina, J., Smagris, E., Stender, S., Nordestgaard, B.G., Zhou, H.H., Tybjærg-Hansen, A., Vogt, T.F., Hobbs, H.H. and Cohen, J.C. (2014) Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat. Genet.,* **46,** 352–356, 10.1038/ng.2901.

42. Chen, V.L., Chen, Y., Du, X., Handelman, S.K. and Speliotes, E.K. (2020) Genetic variants that associate with cirrhosis have pleiotropic effects on human traits. *Liver Int.* First published on January 01, 2020, 10.1111/liv.14321.

43. Chen, L., Du, S., Lu, L., Lin, Z., Jin, W., Hu, D., Jiang, X., Xin, Y. and Xuan, S. (2017) The additive effects of the TM6SF2 E167K and PNPLA3 I148M polymorphisms on lipid metabolism. *Oncotarget,* **8,** 74209–74216, 10.18632/oncotarget.18474.

44. Mahdessian, H., Taxiarchis, A., Popov, S., Silveira, A., Franco-Cereceda, A., Hamsten, A., Eriksson, P. and van't Hooft, F. (2014) TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. *Proc. Natl. Acad. Sci. U.S.A.,* **111,** 8913–8918, 10.1073/pnas.1323785111.

45. Smagris, E., Gilyard, S., BasuRay, S., Cohen, J.C. and Hobbs, H.H. (2016) Inactivation of Tm6sf2, a Gene Defective in Fatty Liver Disease, Impairs Lipidation but Not Secretion of Very Low Density Lipoproteins. *J. Biol. Chem.,* **291,** 10659–10676, 10.1074/jbc.M116.719955.

46. Prill, S., Caddeo, A., Baselli, G., Jamialahmadi, O., Dongiovanni, P., Rametta, R., Kanebratt, K.P., Pujia, A., Pingitore, P. and Mancina, R.M. *et al.* (2019) The TM6SF2 E167K genetic variant induces lipid biosynthesis and reduces apolipoprotein B secretion in human hepatic 3D spheroids. *Sci. Rep.,* **9,** 11585, 10.1038/s41598-019-47737-w.

47. Liao, X.-L., Lou, B., Ma, J. and Wu, M.-P. (2005) Neutrophils activation can be diminished by apolipoprotein A-I. *Life Sci.,* **77,** 325–335, 10.1016/j.lfs.2004.10.066.

48. Watanabe, K., Taskesen, E., van Bochoven, A. and Posthuma, D. (2017) Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* First published on November 28, 2017, 10.1038/s41467-017-01261-5.

49. Sents, W., Ivanova, E., Lambrecht, C., Haesen, D. and Janssens, V. (2013) The biogenesis of active protein phosphatase 2A holoenzymes: a tightly regulated process creating phosphatase specificity. *FEBS J.,* **280,** 644–661, 10.1111/j.1742-4658.2012.08579.x.

50. Rice, L.M., Donigan, M., Yang, M., Liu, W., Pandya, D., Joseph, B.K., Sodi, V., Gearhart, T.L., Yip, J. and Bouchard, M. *et al.* (2014) Protein phosphatase 2A (PP2A) regulates low density lipoprotein uptake through regulating sterol response element-binding protein-2 (SREBP-2) DNA binding. *J. Biol. Chem.,* **289,** 17268–17279, 10.1074/jbc.M114.570390.

51. Jeong, H.J., Lee, H.-S., Kim, K.-S., Kim, Y.-K., Yoon, D. and Park, S.W. (2008) Sterol-dependent regulation of proprotein convertase subtilisin/kexin type 9 expression by sterol-regulatory element binding protein-2. *J. Lipid. Res.,* **49,** 399–409, 10.1194/jlr.M700443-JLR200.

52. Remmerie, M. and Janssens, V. (2019) PP2A: A Promising Biomarker and Therapeutic Target in Endometrial Cancer. *Front. Oncol.* First published on June 04, 2019, 10.3389/fonc.2019.00462.

53. Hafizz, A.M.H.A., Zin, R.R.M., Aziz, N.H.A., Kampan, N.C. and Shafiee, M.N. (2020) Beyond lipid-lowering: role of statins in endometrial cancer. *Mol. Biol. Rep.* First published on September 08, 2020, 10.1007/s11033-020-05760-5.

54. Gustafsen, C., Kjolby, M., Nyegaard, M., Mattheisen, M., Lundhede, J., Buttenschøn, H., Mors, O., Bentzon, J.F., Madsen, P. and Nykjaer, A. *et al.* (2014) The hypercholesterolemia-risk gene SORT1 facilitates PCSK9 secretion. *Cell. Metab.,* **19,** 310–318, 10.1016/j.cmet.2013.12.006.

55. Tavori, H., Giunzioni, I., Predazzi, I.M., Plubell, D., Shivinsky, A., Miles, J., Devay, R.M., Liang, H., Rashid, S. and Linton, M.F. *et al.* (2016) Human PCSK9 promotes hepatic lipogenesis and atherosclerosis development via apoE- and LDLR-mediated mechanisms. *Cardiovasc. Res.,* **110,** 268–278, 10.1093/cvr/cvw053.

56. Leander, K., Mälarstig, A., Van't Hooft, F.M., Hyde, C., Hellénius, M.-L., Troutt, J.S., Konrad, R.J., Öhrvik, J., Hamsten, A. and Faire, U. de (2016) Circulating Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Predicts Future Risk of Cardiovascular Events Independently of Established Risk Factors. *Circulation.* First published on February 19, 2016, 10.1161/CIRCULATIONAHA.115.018531.

57. Rogacev, K.S., Heine, G.H., Silbernagel, G., Kleber, M.E., Seiler, S., Emrich, I., Lennartz, S., Werner, C., Zawada, A.M. and Fliser, D. *et al.* (2016) PCSK9 Plasma Concentrations Are Independent of GFR and Do Not Predict Cardiovascular Events in Patients with Decreased GFR. *PLoS ONE,* **11,** e0146920, 10.1371/journal.pone.0146920.

58. Alborn, W.E., Cao, G., Careskey, H.E., Qian, Y.-W., Subramaniam, D.R., Davies, J., Conner, E.M. and Konrad, R.J. (2007) Serum proprotein convertase subtilisin kexin type 9 is correlated directly with serum LDL cholesterol. *Clin. Chem.,* **53,** 1814–1819, 10.1373/clinchem.2007.091280.

59. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M. and Price, A.L. (2014) Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.,* **46,** 100–106, 10.1038/ng.2876.

60. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.,* **88,** 76–82, 10.1016/j.ajhg.2010.11.011.

61. Winkler, T.W., Day, F.R., Croteau-Chonka, D.C., Wood, A.R., Locke, A.E., Mägi, R., Ferreira, T., Fall, T., Graff, M. and Justice, A.E. *et al.* (2014) Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.,* **9,** 1192–1212, 10.1038/nprot.2014.071.

62. Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A. and Girón, C.G. *et al.* (2018) Ensembl 2018. *Nucleic Acids Res.,* **46,** D754-D761, 10.1093/nar/gkx1098.

63. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E. and Sollis, E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.,* **47,** D1005-D1012, 10.1093/nar/gky1120.

64. Battle, A., Brown, C.D., Engelhardt, B.E. and Montgomery, S.B. (2017) Genetic effects on gene expression across human tissues. *Nature,* **550,** 204–213, 10.1038/nature24277.

65. Joehanes, R., Zhang, X., Huan, T., Yao, C., Ying, S.-X., Nguyen, Q.T., Demirkale, C.Y., Feolo, M.L., Sharopova, N.R. and Sturcke, A. *et al.* (2017) Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. *Genome. Biol.,* **18,** 16, 10.1186/s13059-016-1142-6.

66. Kirsten, H., Al-Hasani, H., Holdt, L., Gross, A., Beutner, F., Krohn, K., Horn, K., Ahnert, P., Burkhardt, R. and Reiche, K. *et al.* (2015) Dissecting the genetics of the human transcriptome identifies novel trait-related trans-eQTLs and corroborates the regulatory relevance of non-protein coding loci†. *Hum. Mol. Genet.,* **24,** 4746–4763, 10.1093/hmg/ddv194.

67. Westra, H.-J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K. and Powell, J.E. *et al.* (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.,* **45,** 1238–1243, 10.1038/ng.2756.

68. Xia, K., Shabalin, A.A., Huang, S., Madar, V., Zhou, Y.-H., Wang, W., Zou, F., Sun, W., Sullivan, P.F. and Wright, F.A. (2012) seeQTL: a searchable database for human eQTLs. *Bioinformatics,* **28,** 451–452, 10.1093/bioinformatics/btr678.

69. Yu, G., Wang, L.-G., Yan, G.-R. and He, Q.-Y. (2015) DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics,* **31,** 608–609, 10.1093/bioinformatics/btu684.

70. Yu, G. and He, Q.-Y. (2016) ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.,* **12,** 477–479, 10.1039/c5mb00663e.

71. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Madden, P.A.F., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N. and Loos, R.J. *et al.* (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.,* **44,** 369-375, S1-3, 10.1038/ng.2213.

72. Wakefield, J. (2007) A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am. J. Hum. Genet.,* **81,** 208–227, 10.1086/519024.

73. Wakefield, J. (2009) Bayes factors for genome-wide association studies: comparison with P-values. *Genet. Epidemiol.,* **33,** 79–86, 10.1002/gepi.20359.

74. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. and Kircher, M. (2019) CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.,* **47,** D886-D894, 10.1093/nar/gky1016.

75. Winkler, T.W., Justice, A.E., Cupples, L.A., Kronenberg, F., Kutalik, Z. and Heid, I.M. (2017) Approaches to detect genetic effects that differ between two strata in genome-wide meta-analyses: Recommendations based on a systematic evaluation. *PLoS ONE,* **12,** e0181038, 10.1371/journal.pone.0181038.

76. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C. and Plagnol, V. (2014) Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.,* **10,** e1004383, 10.1371/journal.pgen.1004383.

Figure Legends

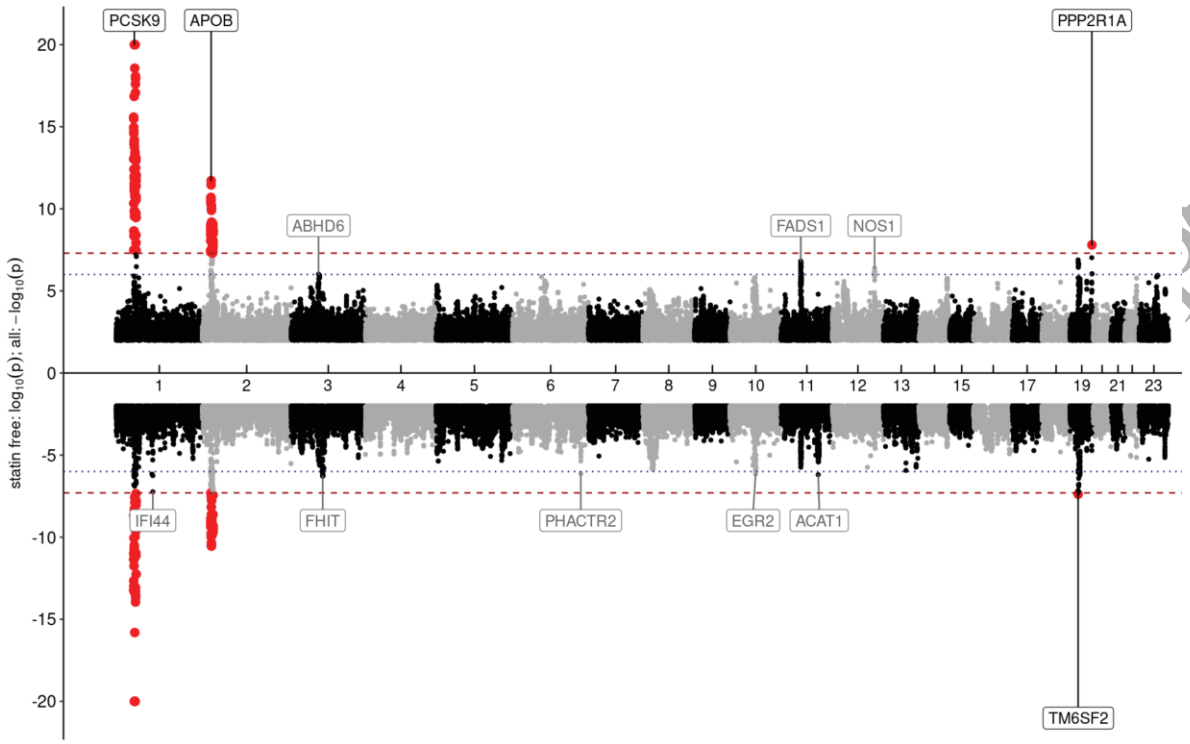


Figure 1: Miami Plot. Distribution of log10-transformed p-values of our GWAMA for all (top, - log10) and the statin free subset (bottom, + log10). The red dashed and blue dotted lines mark genome-wide significance (α=5x10-8) and suggestive significance (α=1x10-6), respectively. The Y-axis is limited to [-20, 20] (max. and min. original y-value: 48.0 for all, -45.8 for sub). Four distinct loci with genome-wide significance were found and their candidate genes are given in black for the best associating group. Candidate genes of suggestive loci are shown in grey. QQ-Plots are given in the Figure S1.

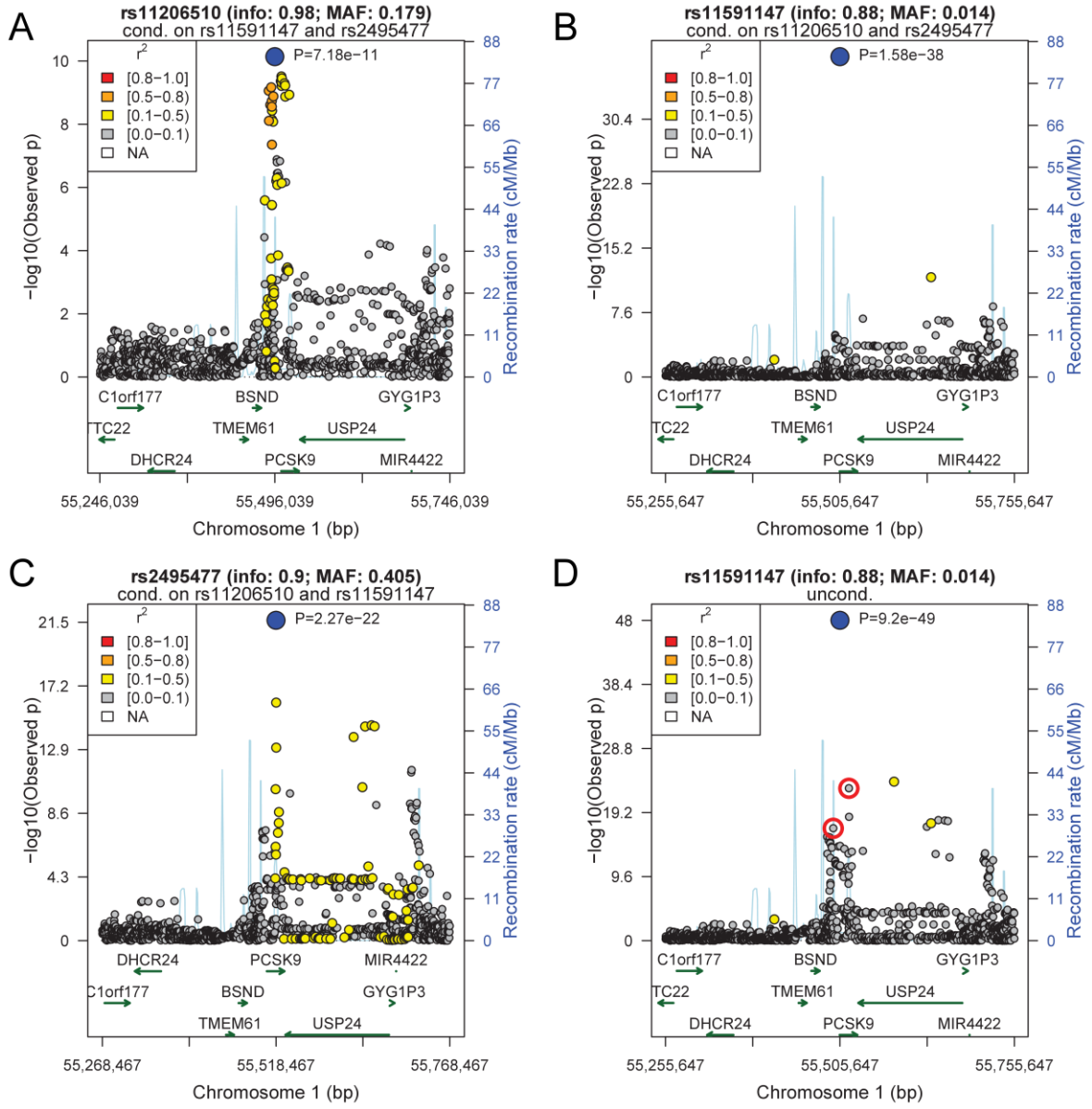


Figure 2: Regional association (RA) plots of the locus 1p32.3. The respective lead SNP is colored blue, the other SNPs are colored according to their LD with the lead SNP (using 1000 Genomes Phase 3, Europeans only). A) – C) RA plots of the independent variants with statistics conditioned on the respective other independent variants. D) RA Plot of the unconditional statistics with rs11591147 as lead SNP. Red circles indicate the two other independent signals, rs2495477 and rs11306510.

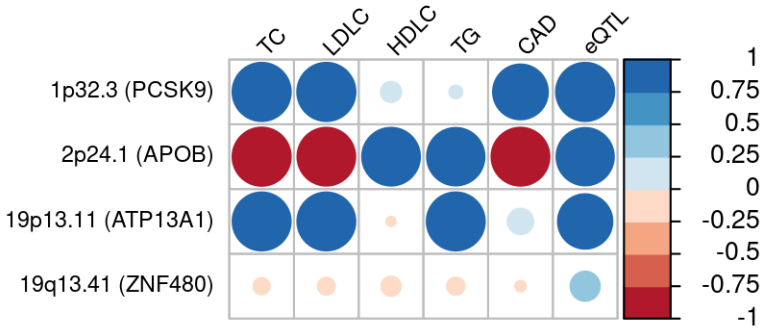


Figure 3: Results of pairwise co-localization for the genome-wide significant loci and other lipid traits, CAD and eQTLs. Maximum of PP4 and PP3 are shown. Blue indicates a high posterior probability of co-localization (PP4, both traits share the same causal variant), and red a high posterior probability for two independent associations at this locus (PP3). White cells indicate that there is no signal for the trait compared. For the eQTL comparison, only the maximum PP4 value over independent SNPs and eQTL tissues was displayed per locus

Tables

Cytoband / Candidate Gene	#SNPs in 99% CS (95%)	SNP	EA / OA	MAF	Info	I^2	Beta	p-value	CADD score
1p32.3 *PCSK9*	1 (1)	**rs11591147**	T/G	0.014	0.880	0.793	-0.293 -0.269	9.20×10^{-49} 2.18×10^{-39}	17.1
	1 (1)	**rs2495477**	G/A	0.405	0.899	0.554	-0.046 -0.045	1.40×10^{-23} 2.18×10^{-22}	14.7
	21 (17)	**rs11206510**	C/T	0.179	0.976	0.000	-0.048 -0.037	1.40×10^{-17} 4.36×10^{-11}	0.073
2p24.1 *APOB*	68 (34)	**rs673548**	A/G	0.216	0.996	0.000	0.036	1.92×10^{-12}	5.1
		rs676210*	A/G	0.215	1.000	0.000	0.036	3.47×10^{-12}	27.1
19p13.11	171 (61)	**rs58542926**	T/C	0.088	0.963	0.338	-0.048	4.17×10^{-8}	23.2
19q13.41	4109 (3118)	**rs71180459**	T/TG	0.103	0.904	0.000	0.040	1.58×10^{-8}	0

Table 1: Independent genome-wide significant SNPs of our GWAMA. For 1p32.3, beta estimates and p-values are given for the univariate (GWAS, first line) and joint analysis (COJO-cond, second line). Lead SNPs per credible set are marked in bold. For *APOB*, the most likely causal variant is provided in the second line. Results of all associated variants and SNPs in credible sets are shown in Tables S2 and S6. * LD between rs673548 – rs676210: $r^2 = 0.988$.

Trait	r_g	$p(r_g)$	1p32.3	2p24.1	19p13.11	19q13.41
TC	**0.392**	0.005	**1.000**	0.000	**0.993**	0.006
LDL-C	**0.325**	0.017	**1.000**	0.001	**0.992**	0.006
HDL-C	0.230	0.061	0.122	**0.985**	0.013	0.019
TG	0.123	0.414	0.049	**0.974**	**0.994**	0.006
CAD	0.021	0.825	**0.888**	0.000	0.193	0.004

Table 2: Genetic relationships between PCSK9, CAD and lipid traits at genome-wide and locus-specific level. Genetic correlations (r_g) with PCSK9 were estimated using our PCSK9 summary statistics and those of lipid traits and CAD available from Teslovich et al. (26) and Nikpay et al. (20). LDHub was used for analyses. Co-localization (columns 4-7, Bayesian posterior probabilities for PP4) was determined using summary statistics of Surakka et al. (19) (lipids) and Nikpay et al. (20). High PP4 indicates co-localization. Significant genetic correlations ($p(r_g)<0.05$) and clear evidence for co-localization (PP4>0.75) are marked in bold. Further results can be found in Tables S11 and S12.

Parameter (see Figure S8)	X	Y	causal estimate	se	p-value	# SNPs
α	PCSK9	LDL-C	1.770	0.069	$4.4x10^{-144}$	3*
β	LDL-C	CAD	0.357	0.070	$3.2x10^{-07}$	24+
τ	PCSK9	CAD	0.996	0.131	$3.2x10^{-14}$	3*
Indir. effect (α * β) mediated by LDL-C	PCSK9	CAD	0.632	0.126	$5.4x10^{-7}$	
Direct effect (τ- α * β)	PCSK9	CAD	0.364	0.182	$4.5x10^{-2}$	

Table 3: Results of Mendelian Randomization and Mediation analyses. Parameters correspond to Figure S8. Analyzed exposure and outcome are denoted with X and Y, respectively. Causal estimates, their standard errors (SE), and p-values were obtained using the IVW approach. We used the summary statistics of our meta-analysis, Surakka et al. (19), and Nikpay et al. (20) for instrument effects on PCSK9, LDL-C, and CAD, respectively.

* Independent SNPs at *PCSK9* (rs11591147, rs2495477 and rs11206510)

+ Independent SNPs associated with LDL-C but not associated with PCSK9

Abbreviations

APOB Apolipoprotein B

CAD Coronary artery disease

CADD Combined annotation dependent depletion

COJO conditional-joint

CS Credible set

ELISA Enzyme-linked Immunosorbent Assay

eQTLs Expression quantitative trait loci

FDR False discovery rate

FPR Formyl Peptide Receptor

GWAMA genome-wide association meta-analysis

GWAS Genome-wide association study

HDL-C high-density lipoprotein-cholesterol

HMGCR HMG-CoA reductase

LD Linkage disequilibrium

LDL-C Low-density lipoprotein-cholesterol

LDLR LDL receptor

MAF Minor allele frequency

MR Mendelian Randomization

PCSK9 Proprotein convertase subtilisin/kexin type 9

PP posterior probability

PP1/PP2A Protein phosphatase 1/2A

PPP2R1A Protein Phosphatase 2 Scaffold Subunit A alpha

SNPs Single nucleotide polymorphism

TC Total cholesterol

TG triglyceride

TM6SF2 Transmembrane 6 Superfamily Member 2