ANGPTL8 protein-truncating variant and the risk of coronary disease, type 2 diabetes and adverse effects

Authors:

Pyry Helkkula, MEng¹, Tuomo Kiiskinen, MD¹,², Aki S. Havulinna, DSc (Tech)¹,², Juha Karjalainen, PhD¹,³,⁴, Seppo Koskinen, MD, PhD², Veikko Salomaa, MD, PhD², FinnGen, Mark J. Daly, PhD¹,³,⁴, Aarno Palotie, MD, PhD¹,³,⁴, Ida Surakka, PhD¹,⁶, Samuli Ripatti, PhD¹,³,⁴,⁷†

Affiliations:

1. Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland
2. Finnish Institute for Health and Welfare, Helsinki, Finland
3. Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
4. Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA
5. Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Analytic and Translational Genetics Unit, Department of Medicine, and the Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
6. Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA
7. Department of Public Health, University of Helsinki, Helsinki, Finland

¹Corresponding author:
Samuli Ripatti, PhD
Institute for Molecular Medicine Finland, HiLIFE
PO Box 20
FI-00014 University of Helsinki, Finland
and
Broad Institute of MIT and Harvard, Cambridge, MA

+358 40 567 0826
samuli.ripatti@helsinki.fi
Abstract

Protein-truncating variants (PTVs) affecting dyslipidemia risk may point to therapeutic targets for cardiometabolic disease. Our objective was to identify PTVs that associated with both lipid levels and cardiometabolic disease risk and assess their possible associations with risks of other diseases. To achieve this aim, we leveraged the enrichment of PTVs in the Finnish population and tested the association of low-frequency PTVs in 1,209 genes with serum lipid levels in the Finrisk Study (n = 23,435). We then tested which of the lipid-associated PTVs also associated with risks of cardiometabolic diseases or 2,264 disease endpoints curated in the FinnGen Study (n = 176,899). Three PTVs were associated with both lipid levels and the risk of cardiometabolic disease: triglyceride-lowering variants in ANGPTL8 (-24.0[-30.4 to -16.9] mg/dL per rs760351239-T allele, \( P = 3.4 \times 10^{-9} \)) and ANGPTL4 (-14.4[-18.6 to -9.8] mg/dL per rs746226153-G allele, \( P = 4.3 \times 10^{-9} \)) and the HDL cholesterol-elevating variant in LIPG (10.2[7.5 to 13.0] mg/dL per rs200435657-A allele, \( P = 5.0 \times 10^{-13} \)). The risk of type 2 diabetes was lower in carriers of ANGPTL8 (odds ratio [OR] = 0.67[0.47-0.92], \( P = 0.01 \)), ANGPTL4 (OR = 0.70[0.60-0.82], \( P = 1.4 \times 10^{-5} \ )) and LIPG (OR = 0.67[0.48-0.91], \( P = 0.01 \)) PTVs than in noncarriers. Moreover, the odds of coronary artery disease were 44% lower in carriers of a PTV in ANGPTL8 (OR = 0.56[0.38-0.83], \( P = 0.004 \)). Finally, the phenotype-wide scan of the ANGPTL8 PTV showed a markedly higher associated risk of esophagitis (585 cases, OR = 174.3[17.7-1715.1], \( P = 9.7 \times 10^{-6} \)) and sensorineural hearing loss (12,250 cases, OR = 2.45[1.63-3.68], \( P = 1.8 \times 10^{-5} \)). The ANGPTL8 PTV carriers were less likely to use statin therapy (53,518 cases, OR = 0.53[0.41-0.71], \( P = 1.2 \times 10^{-5} \)). Our findings provide genetic evidence of potential long-term efficacy and safety of therapeutic targeting of dyslipidemias.

Main

Dyslipidemia is a major risk factor for cardiovascular disease and is present in nearly half of type 2 diabetes patients[1]. For treating dyslipidemia, there are few alternatives to low-density lipoprotein (LDL) cholesterol-lowering therapy. Although common, this therapy often fails to treat the condition effectively, leaving patients with high risk of cardiovascular disease[2]. Therefore, a search for possible new drugs is necessary. This could be achieved, firstly, by developing drugs that modulate other blood lipid levels besides LDL cholesterol. Secondly developing new drugs that lower LDL cholesterol levels through different molecular mechanisms than existing ones. Thirdly, maximizing the safety of dyslipidemia drugs to enable their widespread and preventative use.

There is a limited number of drugs for treating dyslipidemia currently on the market or in development. Drugs targeting triglycerides, lipoprotein(a), high-density lipoprotein (HDL) cholesterol that are currently being developed have emerged from studies of PTVs[3-9], and it is uncertain whether they are safe enough to reach the market. Meanwhile, PCSK9 inhibitors and ezetimibe are the only common alternatives to statins for lowering LDL cholesterol levels. More options would be welcome given that statins have common side effects[10], and a combination therapy with ezetimibe does not reduce coronary atherosclerosis compared to patients on statin monotherapy[11]. On the other hand, PCSK9 inhibitors are unaffordable to many patients and do not benefit people with familial hypercholesterolemia, who have a defective copy of LDLR that PCSK9 inhibits[12, 13].

Besides the small number of alternatives, another issue is the long-term safety of drugs targeting dyslipidemia. Maximizing the long-term safety of these drugs is important because they are often used preventatively and for prolonged periods by wide sectors of the population. Given this use, the long-term, population-level side effects of these drugs are especially important to
minimize. Such an understanding could be achieved by investigating health impacts of genetic proxies for protein deficiencies that are associated with lipid levels. A study assessing this question exists for the chief lipoprotein(a)-modulating gene LPA[14]. However, it is less well known what are the health effects of the other major lipid-modifying genes targeted by drugs currently under development: ANGPTL3, ANGPTL4, APOC3 and CETP[3-8]. Although some of these drug targets have undergone clinical trials, they do not consider long-term health impacts and suffer from confounding factors, such as compound-specific off-target effects.

Considering these clinical needs, the present study aimed to further the treatment of dyslipidemia in two ways. The primary goal was to identify genetic mechanisms to find new and safe drug targets for treating dyslipidemia and cardiometabolic diseases. The secondary aim was to evaluate the long-term health consequences of existing dyslipidemia drug targets. Studying Finns provides a promising avenue for reaching both of these goals. The Finnish population isolate shows an enrichment of protein-truncating variants (PTVs)[15, 16] (stop-gained, frame-shift and essential splice-site mutations), thus enabling the detection of both new and previously-known therapeutic effects through a smaller sample size than in the non-Finnish European population. Thanks to national health records on Finns, it is also possible to screen for a wide range of long-term health impacts associated with these PTVs. When combined, the data on PTVs and health records provide us with exactly the type of long-term, population-wide, on-target side effect data that is currently lacking, as discussed above.

Results

An overview of our study’s analyses and their results is presented in Fig. 1. Our results are divided into three parts: the primary, secondary and tertiary analyses. Results of the primary analysis concern associations between PTVs and serum lipid levels. Results of the secondary analysis concern associations between lipid-associated PTVs and cardiometabolic disease risk. Finally, results of the tertiary analyses concern lipid and cardiometabolic disease-associated PTVs and their potential side effects.

Primary analyses

Protein-truncating variants and serum lipid levels

We found four PTVs to be associated with serum lipid levels. PTVs in CETP (rs751916721-T), LIPG (rs200435657-A), and ANGPTL8 (rs760351239-T) showed genome-wide significant associations with HDL cholesterol levels. PTVs in ANGPTL4 (rs746226153-G) and ANGPTL8 (rs760351239-T) had genome-wide significant associations with triglyceride levels (Tables 1–2). The four PTVs showed 27 to 210-fold enrichment in Finns compared to non-Finnish Europeans in the gnomAD database[17], version 2.1.1 (gnomad.broadinstitute.org) (Table 1).

To evaluate whether the associations between the serum lipid levels and the four PTVs were independent, we conducted conditional tests of independence and determined which credible sets the PTVs belonged to. In our conditional analysis, the associations between lipid levels and the PTVs in the Finrisk Study[18] data were not explained by previously-reported genome-wide significant variants[19-22] (Table 1 and Tables S3-S6 in the Supplement). The LIPG and ANGPTL8 PTVs alone formed credible sets with posterior probabilities above 99.9%. The CETP and ANGPTL4 PTVs were in high linkage disequilibrium with the noncoding variants rs566571297-T ($r^2 = 0.99$) and rs919624228-G ($r^2 = 0.97$) respectively. With their correlated variant pair, these two PTVs formed credible sets with posterior probabilities higher than 99%. Hence, all the four
PTVs were independently associated with changes in lipid levels with a very high probability (Tables S7-S10 in the Supplement).

To inspect if the lipid level associations with PTVs in Finns were concordant with other PTVs in the same genes, we surveyed the UK Biobank data. Using another rare non-Finnish European-enriched ANGPTL8 PTV (rs145464906-T) in Britons, we observed comparable HDL cholesterol and triglyceride level associations as with the PTV found in Finns (Table 2). PTVs in ANGPTL4, CETP or LIPG were not present in the UK Biobank data and thus we were not able to investigate the lipid level associations of other protein-truncating variation in these genes.

**Protein-truncating variants and serum lipid levels as a function of fasting time**

Next, we tested two hypotheses relating to the association of ANGPTL4 and ANGPTL8 with lipid levels as a function of fasting time. The first hypothesis was that ANGPTL4 and ANGPTL8 are associated with lower fasting and postprandial triglyceride levels respectively. We found that ANGPTL4 PTV heterozygotes had 16.6% (20.2 mg/dL, \( P = 8.7 \times 10^{-7} \)) lower triglyceride levels than non-carriers when the time since the last meal was from 4 to 8 hours. On the other hand, the heterozygotes’ triglyceride levels were not significantly lower than non-carriers’ when the fasting time was 3 hours or less (14.8% higher [20.6 mg/dL], \( P = 0.09 \)). The ANGPTL8 PTV heterozygotes in turn had 34.3% (47.8 mg/dL, \( P = 0.01 \)) lower triglyceride levels than non-carriers when the fasting time was up to 3 hours. From 4 to 8 hours after the last meal, the carriers had only 22.9% (27.9 mg/dL, \( P = 3.2 \times 10^{-4} \)) lower triglyceride levels than noncarriers.

The second hypothesis was that the ANGPTL8 PTV is associated with lower fasting LDL cholesterol and triglyceride levels. We observed that ANGPTL8 PTV heterozygotes had 13.4% (18.0 mg/dL, \( P = 4.5 \times 10^{-5} \)) lower LDL cholesterol levels than noncarriers with a fasting time between 4 and 8 hours. Moreover, the heterozygotes had 48.6% (55.1 mg/dL, \( P = 9.7 \times 10^{-4} \)) and 26.2% (36.9 mg/dL, \( P = 0.007 \)) lower triglyceride and LDL cholesterol levels respectively than noncarriers when fasting for 9 hours or longer. The triglyceride and LDL cholesterol levels as a function of ANGPTL4 and ANGPTL8 PTV genotype and fasting time are shown in Fig. 2 and Fig. S1 in the Supplement respectively.

**Secondary analyses**

**Lipid-associated protein-truncating variants and cardiometabolic disease risk**

Given that the primary analysis suggested that ANGPTL4 and ANGPTL8 PTVs may reduce triglyceride levels, we tested these PTVs’ impact on type 2 diabetes (T2D), coronary artery disease (CAD) risk in 176,899 individuals from the FinnGen Study. The ANGPTL8 PTV was associated (two-tailed \( P < 0.05 \)) with lower odds of CAD (OR = 0.56[0.38-0.83], \( P = 0.004 \)) and T2D (OR = 0.65[0.47-0.92], \( P = 0.01 \)). The ANGPTL4 PTV heterozygotes had 30% lower odds of T2D (OR = 0.70[0.60-0.82], \( P = 1.4 \times 10^{-5} \)). In tests between cardiometabolic diseases and the HDL cholesterol-increasing PTVs, we observed that the LIPG PTV was associated with 33% lower odds of T2D (OR = 0.67[0.48-0.91], \( P = 0.01 \)). Finally, we tested if the four PTVs identified in our primary analysis were associated with available traditional non-lipid risk factors for CAD in the FinnGen Study (Table S11 in the Supplement).

Finally, as in the primary analysis, we checked if our findings for the Finnish-enriched ANGPTL8 PTV (rs760351239-T) were congruent with the other ANGPTL8 PTV enriched in UK Biobank. In specific, we tested if the non-Finnish European-enriched ANGPTL8 PTV (rs145464906-T) was associated with T2D and CAD risk. Although rs145464906-T in the UK Biobank data alone was not significantly associated with the risk of T2D or CAD, a meta-analysis of the rs760351239-T and rs145464906-T PTVs strengthened both of these associations (Table 3).
Tertiary analyses

Phenome-wide associations of protein-truncating variants in ANGPTL4, ANGPTL8 and LIPG

To evaluate a wide range of possible long-term health effects of the PTVs associated with both lipid levels and cardiometabolic disease risk, we tested the association between these PTVs and 2,264 curated disease endpoints. In addition to T2D risk, the ANGPTL4 PTV was phenome-wide significantly associated (two-tailed $P < 2.2 \times 10^{-5}$) with lower risk of insulin treatment for diabetes (OR = 0.70[0.60-0.81], $P = 5.9 \times 10^{-6}$), and with endocrine and metabolic diseases (OR = 0.78[0.70-0.87], $P = 1.5 \times 10^{-5}$). The ANGPTL8 PTV was phenome-wide significantly associated with 47% lower odds of statin therapy (OR = 0.53[0.40-0.71], $P = 1.2 \times 10^{-5}$). However, the ANGPTL8 PTV carriers had substantially higher risks of sensorineural hearing loss (OR = 2.45[1.63-3.68], $P = 1.8 \times 10^{-5}$) and esophagitis (OR = 174.3[17.7-1715.1], $P = 9.7 \times 10^{-6}$) (Table 4). See Online Tables 3-6 for the complete phenome-wide association statistics of the PTVs.

Discussion

In this study, we first presented associations between serum lipid levels and PTVs in four genes; ANGPTL4, ANGPTL8, CETP and LIPG. We then showed that carriers of PTVs in ANGPTL8, ANGPTL4 and LIPG had lower risk for T2D and that the risk of CAD was lower in carriers of a PTV in ANGPTL8. Finally, we showed that the PTV carriers in ANGPTL8 had elevated risk of esophagitis and sensorineural hearing loss.

First, our study points to ANGPTL8 as a potential new therapeutic target for lowering triglyceride levels. This is interesting because while CETP, LIPG and ANGPTL4 are well-known lipid genes, the association between a PTV in ANGPTL8 and triglyceride levels has not previously been reported in humans. ANGPTL8 is known to inhibit triglyceride hydrolyzing lipoprotein lipase[23, 24]. In our data, the association between ANGPTL8 PTV and lower triglyceride levels had the strongest effects soon after a meal in a post-prandial state, and after prolonged fasting. These findings are in line with animal studies which show that Angptl8 is secreted after feeding and has a very short half-life[25, 26], and that hepatic very low-density lipoprotein (VLDL) secretion is decreased in Angptl8 knockout mice[27].

Our results also show that carriers of a Finnish-enriched ANGPTL8 PTV had 44% lower odds of coronary artery disease and 35% lower odds of type 2 diabetes. This is in line with genetic evidence of triglycerides being a causal factor for CAD risk[28]. Examining another non-Finnish European-enriched ANGPTL8 PTV (rs145464906-T) in the UK Biobank data, we observed that also this PTV was associated with lower triglyceride levels. In addition, a meta-analysis of the ANGPTL8 PTVs in FinnGen and UK Biobank data increased the statistical significance of T2D and CAD risk associations observed in FinnGen data alone.

Third, the ANGPTL8 PTV carriers had an elevated risk of esophagitis and sensorineural hearing loss compared to noncarriers. Esophagitis is inflammation that damages the esophageal epithelium, which is usually a consequence of epithelial injury. When gastrointestinal reflux injures esophageal tissue, a pro-inflammatory response is caused by increased nuclear factor kappa-B kinase (NF-κB) mediated gene expression[29-31]. ANGPTL8 has been shown to affect NF-κB by suppressing the activation of its subunit gamma IKKγ[32] in humans. This mechanistic understanding is congruent with the observed high risk of esophagitis in ANGPTL8 PTV carriers. We have no mechanistic explanation for the increased risk of sensorineural hearing loss which was associated with the ANGPTL8 PTV, however this association was large enough to be of interest.
Thus, the increased risks of sensorineural hearing loss and esophagitis are findings with important implications for the safety of ANGPTL8 inhibition.

Lastly, in addition to this main contribution on the potential benefits and side effects of the ANGPTL8 PTV, our study also found associations between the ANGPTL4 PTV, the LIPG PTV and a lower risk of T2D. An earlier report also found that ANGPTL4 PTV carriers had a lower risk of T2D than noncarriers[33]. The association between the LIPG PTV and a lower T2D risk supports previous findings showing that endothelial lipase, encoded by LIPG, is correlated with T2D risk[34, 35] and has pro-inflammatory properties[36-38]. Given that inflammation in adipose tissue increases insulin resistance[39-41], it may be expected that endothelial lipase affects T2D risk.

Our study has also some limitations. First, our study was observational in nature and consequently unable to directly reveal causal effects between genes and outcomes. Although we found robust lipid level associations for PTVs that were congruent with previous reports, we cannot confirm the causality between the PTVs and the observed disease risk associations. To confirm the causality between the associated disease risks and the Finnish-enriched PTVs, our findings need to be studied using human cell lines. However, utilizing samples from the bottlenecked Finnish population offered us considerable boosts in statistical power and testing these variants in non-bottlenecked populations would have only been possible with one or two orders of magnitude larges sample sizes.

In summary, we identified PTVs in ANGPTL4 and ANGPTL8 that were associated with lower triglyceride levels and PTVs in CETP and LIPG that were associated with higher HDL cholesterol levels. The carriers of PTVs in ANGPTL8, ANGPTL4 and LIPG had lower risk of T2D and that the risk of CAD was lower in carriers of a PTV in ANGPTL8. Finally, we showed that the PTV carriers in ANGPTL8 had elevated risk of esophagitis and sensorineural hearing loss than noncarriers. These findings point to potential target genes for development of novel preventive medication against T2D and CAD and highlight the utility of bottleneck populations in search of associations between protein-truncating variation and biomarkers.

Methods

We identified PTVs associated with both lipid levels and cardiometabolic disease risk, and then examined their associations with other disease endpoints. In our primary analysis, we studied 23,435 Finns to find PTVs associated with serum lipid levels. Using data from the FinnGen Study, which is based on 176,899 individuals, we then studied the association between these mutations and risk of T2D and CAD. We refer to this as our secondary analysis. In our tertiary and final analysis, using FinnGen data, we assessed the long-term health impacts of the PTVs with cardiometabolic disease risk associations by screening them for modified risk of 2,264 diseases. An overview of our study is depicted in Fig. 1.

Study populations

We used three different data sets for our study: the Finrisk Study cohorts, the FinnGen Study and UK Biobank. In total, the Finrisk Study dataset contained 23,435 chip-genotyped and imputed samples, selected randomly from the Finnish population in 1992, 1997, 2002, 2007 and 2012[18]. The baseline characteristics of the Finrisk Study participants are shown in Table S1 in the Supplement.

The FinnGen Study contains biobank data and national health registry data for 176,899 individuals. The health registry information of participants from the Finrisk and FinnGen Study was followed up until 31.12.2017. The details of the individual FinnGen cohorts are shown in Table S2
in the Supplement. All Finrisk and FinnGen Study participants were of Finnish descent. The genotyping and imputation data release of UK Biobank data was from 5th March 2018.

All participants gave written informed study-specific consent. Patients and control subjects in the FinnGen Study provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, older Finnish research cohorts, collected prior the start of FinnGen Study (August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea, the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) approved the FinnGen Study protocol No. HUS/990/2017.

The FinnGen project is approved by Finnish Institute for Health and Welfare (THL), approval number THL/2031/6.02.00/2017, amendments THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019), Digital and population data service agency VRK43431/2017-3, VRK/6909/2018-3, the Social Insurance Institution (KELA) KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019 and Statistics Finland TK-53-1041-17.


Genotyping and quality control
The Finrisk Study samples were genotyped using the HumanCoreExome BeadChip, Human610-Quad BeadChip, Affymetrix6.0 and Infinium HumanOmniExpress (Illumina Inc., San Diego, CA, USA) chips and a Finnish-ancestry-specific imputation panel consisting of 2,690 deep-coverage (25-30x) whole-genome and 5,092 whole-exome sequences. In the primary analysis, the 1,377 PTVs were located in 1,209 genes and had a MAF between 0.1 and 5%. These PTVs in the Finrisk Study cohorts were imputed and had a IMPUTE2[42] genotype information score with a mean of 0.95 (standard deviation of 0.05) and a minimum of 0.75. The FinnGen Study samples were genotyped with various Illumina and a custom AxiomGT1 Affymetrix array (www.finngen.fi/en/researchers/genotyping). All the lipid-associated PTVs were directly genotyped in at least 62.9% of the FinnGen Study individuals with the AxiomGT1 Affymetrix array. The genotypes of the PTVs in our secondary and tertiary analyses that were not genotyped on chip were imputed using a genotype panel that consisted of 3,775 deep-coverage (25-30x) whole-genome sequenced individuals of Finnish ancestry. The PTVs in our secondary and tertiary analyses had IMPUTE2[42] genotype information scores above 0.92 in the FinnGen data. Detailed description of the genotyping methods, genotype imputation and quality-control procedures are described in the Supplement.

Study outcomes
A blood sample and the self-reported fasting time since the previous meal at the time of blood sample collection of each Finrisk Study participant were collected during a clinical visit. The total cholesterol, HDL cholesterol and triglyceride levels were measured directly from serum or plasma and LDL cholesterol was either directly measured or estimated using the Friedewald formula[43]. In UK Biobank the blood lipid levels were measured from serum directly. LDL and total cholesterol levels of individuals with lipid-lowering therapy were divided by 0.7 and 0.8[44] respectively.
Information on diagnoses in the FinnGen data were collected and confirmed by examining national healthcare registries and recorded using the International Classification of Diseases [ICD] revisions 8-10. Purchase information on prescription drugs since 1995 were obtained from the Finnish social insurance institution (KELA) reimbursement records and coded using the Anatomical Therapeutic Chemical [ATC] classification. All FinnGen Study participants’ healthcare registry information were followed up until 31.12 2017. Cancer diagnoses and causes of death were obtained from their respective national registries. The clinical expert groups of the FinnGen Study have defined disease events using ICD and ATC codes. For a complete list of the considered clinical endpoints and corresponding ICD and ATC codes, see Online Tables 1-3.

For the T2D and CAD statuses in UK Biobank participants we used PheCodes[45] 250.2 and 411.4 respectively.

Study design and statistical analyses

Primary analysis
In this analysis, we sought to identify associations between lipid levels and PTVs in the Finnrisk Study. The lipid levels tested were plasma or serum levels of HDL cholesterol, LDL cholesterol, total cholesterol and logarithmically transformed triglycerides on a natural log scale. Our model was additive, and included age, age squared, sex, collection year, and ten principal components of ancestry as fixed-effects covariates. To correct for the effect of adiposity on triglyceride levels, we adjusted the triglyceride association tests for waist-to-hip ratio as well. We considered the genome-wide association significance threshold of a two-sided P value of less than $5.0 \times 10^{-8}$ to be significant. Genetic association analyses were carried out using the PLINK[46], version v1.90b3.45 (www.cog-genomics.org/plink/1.9/), file format, Python, version 3.6 (www.python.org) and the statsmodels Python package, version 0.8.0 (www.statsmodels.org). In our scan, we only considered variants with a MAF 0.1-5% to account for adequate statistical power and the expected low frequency of high-impact alleles. To assess the functional validity of the associations, we performed conditional analyses with previously associated variants in the gene (Tables S3-S6 in the Supplement) and determined the 95% credible sets of variants in each gene locus (Tables S7-S10 in the Supplement). The credible sets were determined using FINEMAP[47], version 1.4.3 (www.finemap.me).

Next, we carried out a post-hoc analysis to test the hypotheses that ANGPTL4 and ANGPTL8 PTVs are associated with triglycerides and LDL cholesterol levels as a function of fasting time. These hypotheses were based on animal studies that we reviewed after finding four PTVs associated with lipids in our genome-wide analysis. Firstly, we found studies that showed that in mice, Angptl4 and Angptl8 inhibit lipoprotein lipase (LPL) as a function of fasting time[25, 26, 48]. LPL inhibition is important because it is the mechanism by which several triglyceride-lowering drugs currently under development work[3, 4, 7]. If the association between triglyceride levels and ANGPTL8 and ANGPTL4 PTVs depends on fasting time, then fasting time dependent effects on triglyceride levels of these PTVs are very relevant for our study, which aims to assess the effect of these PTVs on hypertriglyceridemia risk.

Literature similarly suggested that, to determine whether ANGPTL8 is a viable drug target, it might also be important to test its association with LDL cholesterol levels as a function of fasting time. A mouse study found that ANGPTL8 modulates VLDL secretion[27]. Given our goal of testing the viability of ANGPTL8 as a drug target, the effect of ANGPTL8 on VLDL levels would be important to assess in humans because high VLDL levels are associated with a higher risk of CAD[49]. We could not directly test whether human ANGPTL8 PTV carriers had lower VLDL levels than noncarriers, so we instead tested the association between the ANGPTL8 PTV and levels of
LDL cholesterol and triglycerides after fasting. Our rationale for this test was that fasting triglyceride and LDL cholesterol levels in particular can be considered a proxy for VLDL. This is due to the deficiency of chylomicron and chylomicron remnant particles in the bloodstream in a fasted state[50].

We examined the serum lipid level associations of other PTVs in the same genes as in our findings, using UK Biobank data. For this we used another rare non-Finnish European-enriched ANGPTL8 PTV (rs145464906-T). PTVs in ANGPTL4, CETP or LIPG were not present in the UK Biobank data and thus we were not able to analyze the lipid level associations of protein-truncating variation in these genes.

Secondary and tertiary analyses

In the secondary analyses, we examined associations between cardiometabolic disease risk and the lipid-associated PTVs identified in the primary analysis. The analyses included two cardiometabolic diseases: T2D and CAD. The associations (two-tailed P < 0.05) were tested on data from the FinnGen Study.

As in the primary analysis, we checked if our findings for the Finnish-enriched ANGPTL8 PTV (rs760351239-T) were consistent with another rare ANGPTL8 PTV in UK Biobank. We tested if the non-Finnish European-enriched ANGPTL8 PTV (rs145464906-T) was associated with T2D and CAD risk. In addition, we conducted an inverse-variance-weighted meta-analysis of the rs760351239-T and rs145464906-T PTVs in the FinnGen and UK Biobank data.

In the tertiary analyses we examined the other health impacts of the PTVs associated with both serum lipid levels and the risk of T2D or CAD. We screened these PTVs broadly for modified risk of 2,264 diseases in the FinnGen data. We regarded a two-sided P value below $2.2 \times 10^{-5}$ (Bonferroni-corrected threshold for 2,264 traits) to be statistically significant.

In both the secondary and tertiary analyses in the FinnGen data, the odds ratios for disease outcomes were estimated using SAIGE[51], version 0.35.8.8 (www.github.com/weizhouUMICH/SAIGE/releases/tag/0.35.8.8). Age, sex, genotyping batch and ten principal components of ancestry and the kinship matrix were included as fixed-effects covariates. See Online Tables 1-3 for the disease definitions and Online Tables 4-6 for the associations between disease risks and the PTVs. The details of computing the associations in the FinnGen data are described in the Supplement. The T2D and CAD risk associations were also computed using SAIGE and were obtained from Zhou et al.[51]
References

Acknowledgements

We would like to thank Heidi Silvennoinen and Katri Silvennoinen for proofreading the manuscript, and Sari Kivikko, Huei-Yi Shen and Ulla Tuomainen for management assistance. Many thanks go to Kalle Pärn, Marita A. Isokallio, Javier Nunez Fontarnau and Priit Palta for the genotype imputation of Finrisk Study participants.


The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and by eleven industry partners (AbbVie Inc, AstraZeneca UK Ltd, Biogen MA Inc, Celgene Corporation, Celgene International II Sàrl, Genentech Inc, Merck Sharp & Dohme Corp, Pfizer Inc., GlaxoSmithKline, Sanofi, Maze Therapeutics Inc., Janssen Biotech Inc).

This work was supported by the Sigrid Jusélius Foundation (to S.R., A.P.); University of Helsinki HiLIFE Fellow grants 2017-2020 (to S.R.); Academy of Finland Center of Excellence in Complex Disease Genetics (grant number 312062 to S.R., 312074 to A.P., 312075 to M.J.D.); Academy of Finland (grant number 285380 to S.R., 128650 to A.P., 321356 to A.S.H.); Foundation and the Horizon 2020 Research and Innovation Programme (grant number 667301 [COSYN] to A.P.); The Finnish Foundation for Cardiovascular Research (to S.R., A.P., V.S.); Precision Health Scholars Award from the University of Michigan Medical School (to I.S.); Doctoral Programme in Population Health, University of Helsinki (to P.H., T.K.); Emil Aaltonen Foundation (to P.H.). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

V.S. has consulted for Novo Nordisk and Sanofi and received honoraria from these companies (unrelated to the present study). He also has ongoing research collaboration with Bayer Ltd. (unrelated to the present study). A.P. is a member of the Pfizer Genetics Scientific Advisory Panel. The remaining authors declared no relevant competing interests.

Data availability

The FinnGen data may be accessed through Finnish Biobanks’ FinnBB portal (www.finbb.fi) and THL Biobank data through THL Biobank (https://thl.fi/en/web/thl-biobank).
Table 1. Associations between protein-truncating variants and serum lipid levels in the Finrisk Study.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Locus and variant</th>
<th>Chromosome position\textsuperscript{b}</th>
<th>Type of mutation</th>
<th>MAF in Finrisk Study $%$</th>
<th>Enrichment\textsuperscript{c}</th>
<th>Primary association</th>
<th>$P$</th>
<th>Conditional $P$\textsuperscript{d}</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGPTL4 rs7462262153-G</td>
<td>19:8364556</td>
<td>Frameshift</td>
<td>0.48</td>
<td>27.2</td>
<td>Triglycerides</td>
<td>$4.3 \times 10^{-9}$</td>
<td>$3.7 \times 10^{-8}$</td>
</tr>
<tr>
<td>ANGPTL8 rs760351239-T</td>
<td>19:11240228</td>
<td>Stop-gained</td>
<td>0.15</td>
<td>85.0</td>
<td>Triglycerides</td>
<td>$3.4 \times 10^{-9}$</td>
<td>-</td>
</tr>
<tr>
<td>CETP rs751916721-T</td>
<td>16:56962097</td>
<td>Splice-site</td>
<td>0.12</td>
<td>46.3</td>
<td>HDL cholesterol</td>
<td>$1.6 \times 10^{-21}$</td>
<td>$5.9 \times 10^{-19}$</td>
</tr>
<tr>
<td>LIPG rs200435657-A</td>
<td>18:49565316</td>
<td>Splice-site</td>
<td>0.20</td>
<td>210.1</td>
<td>HDL cholesterol</td>
<td>$5.0 \times 10^{-13}$</td>
<td>$3.6 \times 10^{-12}$</td>
</tr>
</tbody>
</table>

Abbreviations: MAF, minor-allele frequency.

\textsuperscript{a} Shown in the table are the four PTVs associated genome-wide significantly (two-sided $P < 5 \times 10^{-8}$) with at least one serum lipid level measure in the Finrisk Study cohorts. The tested serum lipid levels were LDL, HDL and total cholesterol, as well as natural logarithm transformed triglycerides.

\textsuperscript{b} Chromosome numbers and positions refer to genome build GRCh38.

\textsuperscript{c} The allele frequency enrichment in Finns with respect to non-Finnish Europeans according to the gnomAD database, version 2.1.1 (gnomad.broadinstitute.org).

\textsuperscript{d} The conditional P value is the largest P value from association tests conditioning on other previously reported genome-wide significantly markers in the same gene. The conditional P value for rs760351239-T is not available because no genome-wide significantly genetic variant in ANGPTL8 has been reported before.
Table 2. Associations between protein-truncating variants and serum lipid levels in the Finrisk Study and UK Biobank.

<table>
<thead>
<tr>
<th>Locus and variant</th>
<th>Study</th>
<th>Triglycerides</th>
<th>LDL cholesterol</th>
<th>HDL cholesterol</th>
<th>Total cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Effect (95% CI)(^b)</td>
<td>Effect (95% CI)(^b)</td>
<td>Effect (95% CI)(^b)</td>
<td>Effect (95% CI)(^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mg/dL</td>
<td>mg/dL</td>
<td>mg/dL</td>
<td>mg/dL</td>
</tr>
<tr>
<td>ANGPTL4 rs746226153-G</td>
<td>Finrisk</td>
<td>-14.4</td>
<td>2.4</td>
<td>1.5</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-18.6 to -9.8)</td>
<td>(-2.1 to 7.0)</td>
<td>(-0.2 to 3.3)</td>
<td>(-5.3 to 4.9)</td>
</tr>
<tr>
<td>ANGPTL8 rs760351239-T</td>
<td>Finrisk</td>
<td>-24.0</td>
<td>-16.8</td>
<td>9.1</td>
<td>-14.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-30.4 to -16.9)</td>
<td>(-24.8 to -8.7)</td>
<td>(6.1 to 12.3)</td>
<td>(-23.2 to -5.1)</td>
</tr>
<tr>
<td>ANGPTL8 rs145464906-T</td>
<td>UK Biobank</td>
<td>-18.9</td>
<td>-4.5</td>
<td>6.1</td>
<td>-1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-21.2 to -15.1)</td>
<td>(-7.6 to -1.4)</td>
<td>(4.8 to 7.4)</td>
<td>(-5.2 to 3.1)</td>
</tr>
<tr>
<td>CETP rs751916721-T</td>
<td>Finrisk</td>
<td>5.1</td>
<td>-7.0</td>
<td>16.7</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-5.2 to 16.8)</td>
<td>(-16.1 to 2.1)</td>
<td>(13.2 to 20.1)</td>
<td>(1.9 to 22.2)</td>
</tr>
<tr>
<td>LIPG rs200435657-A</td>
<td>Finrisk</td>
<td>3.7</td>
<td>4.1</td>
<td>10.2</td>
<td>17.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-4.6 to 12.8)</td>
<td>(-3.4 to 11.7)</td>
<td>(7.5 to 13.0)</td>
<td>(9.0 to 25.5)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval.  
\(^a\) All association tests in were adjusted for age, age squared, sex. In the Finrisk Study and UK Biobank association tests were additionally adjusted for the study cohort and ten principal components, and assessment center and 40 principal components respectively. In analyses of serum triglycerides, the triglyceride levels were natural logarithm transformed and waist-to-hip ratio was included as an additional covariate. Values for LDL and total cholesterol were adjusted for the use of lipid-lowering medication by dividing by 0.7 and 0.8 respectively.  
\(^b\) The effect is the difference per each copy of the minor-allele in units in mg/dL of triglycerides, LDL, HDL and total cholesterol. To convert the values for cholesterol to mmol/L, multiply by 0.0259. To convert the values for triglycerides to mmol/L, multiply by 0.0113.
Table 3. Association between protein-truncating variants and cardiometabolic disease risk in the FinnGen Study and UK Biobank.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Locus and variant</th>
<th>Cohort</th>
<th>Type 2 diabetes; 23,338 cases in FinnGen</th>
<th>Coronary artery disease; 18,295 cases in FinnGen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Allele frequency\textsuperscript{b}</td>
<td>OR\textsuperscript{c} (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>ANGPTL4</td>
<td>FinnGen</td>
<td>0.42</td>
<td>0.54</td>
</tr>
<tr>
<td>rs746226153-G</td>
<td></td>
<td>0.09</td>
<td>0.12</td>
</tr>
<tr>
<td>ANGPTL8</td>
<td>UK Biobank</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>rs760351239-T</td>
<td></td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>ANGPTL8 PTV meta-analysis\textsuperscript{d}</td>
<td>FinnGen and UK Biobank</td>
<td>0.10</td>
<td>0.011</td>
</tr>
<tr>
<td>CETP rs751916721-T</td>
<td>FinnGen</td>
<td>0.10</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Abbreviations: OR, odds ratio; CI, confidence interval.

\textsuperscript{a} Shown in the table are the association statistics between the serum lipid-associated PTVs and type 2 diabetes and coronary artery disease risk.

\textsuperscript{b} The allele frequencies are reported in percent. The rs145464906-T allele frequencies in cases and controls were calculated in 343,687 unrelated individuals in UK Biobank and are based on PheCodes\textsuperscript{[45]} 250.2 and 411.4 for type 2 diabetes and coronary artery disease respectively.

\textsuperscript{c} In the FinnGen data the odds ratios were calculated using SAIGE\textsuperscript{[51]} saddle-point approximation-based score test after adjustment for age, sex, genotyping batch and ten principal components of ancestry. The odds ratios for the associations between the disease risks and PTVs in UK Biobank were obtained from Zhou et al.\textsuperscript{[51]}

\textsuperscript{d} The rs760351239-T and rs145464906-T PTVs present in FinnGen and UK Biobank genotype data respectively have been meta-analyzed.
Table 4. Phenome-wide associations with the **ANGPTL8** protein-truncating variant in the FinnGen Study.\(^a\)

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>ANGPTL8 PTV rs760351239-T</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Cases</td>
<td>Allele Frequency(^b)</td>
<td>OR(^c) (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td></td>
</tr>
<tr>
<td>Esophagitis</td>
<td>585</td>
<td>0.77</td>
<td>0.12</td>
<td>174.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(17.7-1715.1)</td>
</tr>
<tr>
<td>Statin medication</td>
<td>53,518</td>
<td>0.10</td>
<td>0.13</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.40-0.71)</td>
</tr>
<tr>
<td>Sensorineural hearing loss</td>
<td>12,250</td>
<td>0.22</td>
<td>0.11</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.63-3.68)</td>
</tr>
</tbody>
</table>

Abbreviations: PTV, protein-truncating variant; OR, odds ratio; CI, confidence interval.

\(^a\) Shown in the table are the phenome-wide significant associations (P < 2.2 × 10^-5) for the rs760351239-T allele in the FinnGen Study.

\(^b\) The allele frequencies are reported in percent.

\(^c\) Odds ratios were calculated using SAIGE\(^{[51]}\) saddle-point approximation-based score test after adjustment for age, sex, genotyping batch and ten principal components of ancestry.
Figure 1. An overview of the present study.a

**Primary analysis**

The first step was to find PTVs associated with blood lipid levels in the Finrisk Study cohorts (23,435 individuals). Four PTVs were found.

Two of these PTVs were then tested for fasting and feeding-specific associations with triglyceride and LDL cholesterol levels. The two PTVs, in ANGPTL4 and ANGPTL8, were singled out for this analysis based on past literature on their effects in mice.

**Secondary analysis**

In this analysis, we tested for associations between lipid-associated PTVs and the risk of T2D and CAD in the FinnGen Study data (176,899 individuals).

**Tertiary analysis**

To find long-term health impacts of the three PTVs in the analysis, we conducted a broad association test between the PTVs in ANGPTL4, ANGPTL8 and LIPG, and risks of 2,264 disease endpoints.

Abbreviations: PTV, protein-truncating variant; MAF, minor-allele frequency; HDL-C, high-density lipoprotein cholesterol level; ln(TG), natural logarithm of triglyceride level; LDL-C, low-density lipoprotein cholesterol level; TC, total cholesterol level; T2D, type 2 diabetes; CAD, coronary artery disease.

a Shown in the figure is an overview of the present study. The final results are shown in boxes with a bolded font.

b Results for these association tests were also compared with other PTVs in UK Biobank.

c These association tests were post-hoc in nature.
Figure 2. Mean serum triglyceride levels of ANGPTL4 and ANGPTL8 PTV heterozygotes at their self-reported fasting time in the Finrisk Study.\textsuperscript{a}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\end{figure}

\textsuperscript{a}The figure shows the mean serum triglyceride level by ANGPTL4 and ANGPTL8 PTV carrier status with respect to self-reported fasting time. The number of ANGPTL4 and ANGPTL8 heterozygotes for each fasting time interval are reported below the fasting time legend. The height of the bars indicate means and the error bars 95\% confidence intervals. The P values are for the two-sided Welch’s t-test of the natural logarithm of triglyceride levels between noncarriers and heterozygotes.