

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

## Authors:

Pyry Helkkula, MEng<sup>1</sup>, Tuomo Kiiskinen, MD<sup>1,2</sup>, Aki S. Havulinna, DSc (Tech)<sup>1,2</sup>, Juha Karjalainen, PhD<sup>1,3,4</sup>, Seppo Koskinen, MD, PhD<sup>2</sup>, Veikko Salomaa, MD, PhD<sup>2</sup>, *FinnGen*, Mark J. Daly, PhD<sup>1,3,4</sup>, Aarno Palotie, MD, PhD<sup>1,4,5</sup>, Ida Surakka, PhD<sup>1,6</sup>, Samuli Ripatti, PhD<sup>1,4,7†</sup>

## 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

†Corresponding 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

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## 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 phenome-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 larger 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.

The Biobank Access Decisions for FinnGen Study samples and data utilized in FinnGen Data Freeze 4 include: THL Biobank BB2017\_55, BB2017\_111, BB2018\_19, BB\_2018\_34, BB\_2018\_67, BB2018\_71, BB2019\_7 Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, Auria Biobank AB17-5154, Biobank Borealis of Northern Finland\_2017\_1013, Biobank of Eastern Finland 1186/2018, Finnish Clinical Biobank Tampere MH0004, Central Finland Biobank 1-2017, and Terveystalo Biobank STB 2018001.

## 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.finnngen.fi/en/researchers/genotyping](http://www.finnngen.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 Finrisk 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/](http://www.cog-genomics.org/plink/1.9/)), file format, Python, version 3.6 ([www.python.org](http://www.python.org)) and the statsmodels Python package, version 0.8.0 ([www.statsmodels.org](http://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](http://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 blood stream 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](https://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

1. Kannel, W.B. and D.L. McGee, *Diabetes and cardiovascular disease. The Framingham study.* JAMA, 1979. **241**(19): p. 2035-8.
2. Fruchart, J.C., et al., *The Residual Risk Reduction Initiative: a call to action to reduce residual vascular risk in patients with dyslipidemia.* Am J Cardiol, 2008. **102**(10 Suppl): p. 1K-34K.
3. Dewey, F.E., et al., *Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease.* N Engl J Med, 2017. **377**(3): p. 211-221.
4. Dewey, F.E., et al., *Inactivating Variants in ANGPTL4 and Risk of Coronary Artery Disease.* N Engl J Med, 2016. **374**(12): p. 1123-33.
5. Gaudet, D., et al., *Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia.* N Engl J Med, 2015. **373**(5): p. 438-47.
6. Gaudet, D., et al., *Targeting APOC3 in the familial chylomicronemia syndrome.* N Engl J Med, 2014. **371**(23): p. 2200-6.
7. Graham, M.J., et al., *Cardiovascular and Metabolic Effects of ANGPTL3 Antisense Oligonucleotides.* N Engl J Med, 2017. **377**(3): p. 222-232.
8. Group, H.T.R.C., et al., *Effects of Anacetrapib in Patients with Atherosclerotic Vascular Disease.* N Engl J Med, 2017. **377**(13): p. 1217-1227.
9. Tsimikas, S., et al., *Lipoprotein(a) Reduction in Persons with Cardiovascular Disease.* N Engl J Med, 2020. **382**(3): p. 244-255.
10. Collins, R., et al., *Interpretation of the evidence for the efficacy and safety of statin therapy.* Lancet, 2016. **388**(10059): p. 2532-2561.
11. Kastelein, J.J., et al., *Simvastatin with or without ezetimibe in familial hypercholesterolemia.* N Engl J Med, 2008. **358**(14): p. 1431-43.
12. Hobbs, H.H., M.S. Brown, and J.L. Goldstein, *Molecular genetics of the LDL receptor gene in familial hypercholesterolemia.* Hum Mutat, 1992. **1**(6): p. 445-66.
13. Rader, D.J., J. Cohen, and H.H. Hobbs, *Monogenic hypercholesterolemia: new insights in pathogenesis and treatment.* J Clin Invest, 2003. **111**(12): p. 1795-803.
14. Emdin, C.A., et al., *Phenotypic Characterization of Genetically Lowered Human Lipoprotein(a) Levels.* J Am Coll Cardiol, 2016. **68**(25): p. 2761-2772.
15. Chheda, H., et al., *Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom.* Eur J Hum Genet, 2017. **25**(4): p. 477-484.
16. Lim, E.T., et al., *Distribution and medical impact of loss-of-function variants in the Finnish founder population.* PLoS Genet, 2014. **10**(7): p. e1004494.
17. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.
18. Vartiainen, E., et al., *Thirty-five-year trends in cardiovascular risk factors in Finland.* Int J Epidemiol, 2010. **39**(2): p. 504-18.
19. Klarin, D., et al., *Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program.* Nat Genet, 2018. **50**(11): p. 1514-1523.
20. Liu, D.J., et al., *Exome-wide association study of plasma lipids in >300,000 individuals.* Nat Genet, 2017. **49**(12): p. 1758-1766.
21. Lu, X., et al., *Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease.* Nat Genet, 2017. **49**(12): p. 1722-1730.

22. Willer, C.J., et al., *Discovery and refinement of loci associated with lipid levels*. Nat Genet, 2013. **45**(11): p. 1274-1283.
23. Chi, X., et al., *ANGPTL8 promotes the ability of ANGPTL3 to bind and inhibit lipoprotein lipase*. Mol Metab, 2017. **6**(10): p. 1137-1149.
24. Haller, J.F., et al., *ANGPTL8 requires ANGPTL3 to inhibit lipoprotein lipase and plasma triglyceride clearance*. J Lipid Res, 2017. **58**(6): p. 1166-1173.
25. Dang, F., et al., *Fasting and Feeding Signals Control the Oscillatory Expression of Angptl8 to Modulate Lipid Metabolism*. Sci Rep, 2016. **6**: p. 36926.
26. Fu, Z., A.B. Abou-Samra, and R. Zhang, *A lipasin/Angptl8 monoclonal antibody lowers mouse serum triglycerides involving increased postprandial activity of the cardiac lipoprotein lipase*. Sci Rep, 2015. **5**: p. 18502.
27. Wang, Y., et al., *Mice lacking ANGPTL8 (Betatrophin) manifest disrupted triglyceride metabolism without impaired glucose homeostasis*. Proc Natl Acad Sci U S A, 2013. **110**(40): p. 16109-14.
28. Do, R., et al., *Common variants associated with plasma triglycerides and risk for coronary artery disease*. Nat Genet, 2013. **45**(11): p. 1345-52.
29. Fang, Y., et al., *Gastroesophageal reflux activates the NF-kappaB pathway and impairs esophageal barrier function in mice*. Am J Physiol Gastrointest Liver Physiol, 2013. **305**(1): p. G58-65.
30. Goldman, A., et al., *Characterization of squamous esophageal cells resistant to bile acids at acidic pH: implication for Barrett's esophagus pathogenesis*. Am J Physiol Gastrointest Liver Physiol, 2011. **300**(2): p. G292-302.
31. Rafiee, P., et al., *Effect of curcumin on acidic pH-induced expression of IL-6 and IL-8 in human esophageal epithelial cells (HET-1A): role of PKC, MAPKs, and NF-kappaB*. Am J Physiol Gastrointest Liver Physiol, 2009. **296**(2): p. G388-98.
32. Zhang, Y., et al., *ANGPTL8 negatively regulates NF-kappaB activation by facilitating selective autophagic degradation of IKKgamma*. Nat Commun, 2017. **8**(1): p. 2164.
33. Gusarova, V., et al., *Genetic inactivation of ANGPTL4 improves glucose homeostasis and is associated with reduced risk of diabetes*. Nat Commun, 2018. **9**(1): p. 2252.
34. Badellino, K.O., et al., *Endothelial lipase concentrations are increased in metabolic syndrome and associated with coronary atherosclerosis*. PLoS Med, 2006. **3**(2): p. e22.
35. Shiu, S.W., et al., *Type 2 diabetes mellitus and endothelial lipase*. Atherosclerosis, 2008. **198**(2): p. 441-7.
36. Hara, T., et al., *Targeted deletion of endothelial lipase increases HDL particles with anti-inflammatory properties both in vitro and in vivo*. J Lipid Res, 2011. **52**(1): p. 57-67.
37. Kempe, S., et al., *NF-kappaB controls the global pro-inflammatory response in endothelial cells: evidence for the regulation of a pro-atherogenic program*. Nucleic Acids Res, 2005. **33**(16): p. 5308-19.
38. Kojima, Y., et al., *Endothelial lipase modulates monocyte adhesion to the vessel wall. A potential role in inflammation*. J Biol Chem, 2004. **279**(52): p. 54032-8.
39. Festa, A., et al., *Chronic subclinical inflammation as part of the insulin resistance syndrome: the Insulin Resistance Atherosclerosis Study (IRAS)*. Circulation, 2000. **102**(1): p. 42-7.
40. Uysal, K.T., et al., *Protection from obesity-induced insulin resistance in mice lacking TNF-alpha function*. Nature, 1997. **389**(6651): p. 610-4.
41. Xu, H., et al., *Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance*. J Clin Invest, 2003. **112**(12): p. 1821-30.

42. Howie, B.N., P. Donnelly, and J. Marchini, *A flexible and accurate genotype imputation method for the next generation of genome-wide association studies*. PLoS Genet, 2009. **5**(6): p. e1000529.
43. Friedewald, W.T., R.I. Levy, and D.S. Fredrickson, *Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge*. Clin Chem, 1972. **18**(6): p. 499-502.
44. Cholesterol Treatment Trialists, C., et al., *Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials*. Lancet, 2010. **376**(9753): p. 1670-81.
45. Denny, J.C., et al., *Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data*. Nat Biotechnol, 2013. **31**(12): p. 1102-10.
46. Chang, C.C., et al., *Second-generation PLINK: rising to the challenge of larger and richer datasets*. Gigascience, 2015. **4**: p. 7.
47. Benner, C., et al., *FINEMAP: efficient variable selection using summary data from genome-wide association studies*. Bioinformatics, 2016. **32**(10): p. 1493-501.
48. Kersten, S., et al., *Characterization of the fasting-induced adipose factor FIAF, a novel peroxisome proliferator-activated receptor target gene*. J Biol Chem, 2000. **275**(37): p. 28488-93.
49. Ren, J., et al., *Long-term coronary heart disease risk associated with very-low-density lipoprotein cholesterol in Chinese: the results of a 15-Year Chinese Multi-Provincial Cohort Study (CMCS)*. Atherosclerosis, 2010. **211**(1): p. 327-32.
50. Nordestgaard, B.G., *A Test in Context: Lipid Profile, Fasting Versus Nonfasting*. J Am Coll Cardiol, 2017. **70**(13): p. 1637-1646.
51. Zhou, W., et al., *Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies*. Nat Genet, 2018. **50**(9): p. 1335-1341.

## 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.

Following biobanks are acknowledged for collecting the FinnGen project samples: Auria Biobank (<https://www.auria.fi/biopankki>), THL Biobank (<https://thl.fi/fi/web/thl-biopankki>), Helsinki Biobank (<https://www.terveyskyla.fi/helsinginbiopankki>), Biobank Borealis of Northern Finland (<https://www oulu.fi/university/node/38474>), Finnish Clinical Biobank Tampere ([https://www.tays.fi/enUS/Research\\_and\\_development/Finnish\\_Clinical\\_Biobank\\_Tampere](https://www.tays.fi/enUS/Research_and_development/Finnish_Clinical_Biobank_Tampere)), Biobank of Eastern Finland (<https://itasuomenbiopankki.fi>), Central Finland Biobank (<https://www.ksshp.fi/fi-FI/Potilaalle/Biopankki>), Finnish Red Cross Blood Service Biobank (<https://www.veripalvelu.fi/verenluovutus/biopankkitoiminta>) and Terveystalo Biobank (<https://www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/>). All Finnish Biobanks are members of BBMRI.fi infrastructure ([www.bbmri.fi](http://www.bbmri.fi)). We also thank study participants for their generous participation at THL Biobank, the National FINRISK study and UK Biobank. This research has been conducted using the UK Biobank Resource with the application number 22627.

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](http://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.<sup>a</sup>**

Locus and variant	Chromosome position <sup>b</sup>	Type of mutation	MAF in Finrisk Study %	Enrichment <sup>c</sup>	Primary association	<i>P</i>	Conditional <i>P</i> <sup>d</sup>
<i>ANGPTL4</i> <i>rs746226153-G</i>	19:8364556	Frameshift	0.48	27.2	Triglycerides	$4.3 \times 10^{-9}$	$3.7 \times 10^{-8}$
<i>ANGPTL8</i> <i>rs760351239-T</i>	19:11240228	Stop-gained	0.15	85.0	Triglycerides	$3.4 \times 10^{-9}$	-
<i>CETP</i> <i>rs751916721-T</i>	16:56962097	Splice-site	0.12	46.3	HDL cholesterol	$1.6 \times 10^{-21}$	$5.9 \times 10^{-19}$
<i>LIPG</i> <i>rs200435657-A</i>	18:49565316	Splice-site	0.20	210.1	HDL cholesterol	$5.0 \times 10^{-13}$	$3.6 \times 10^{-12}$

Abbreviations: MAF, minor-allele frequency.

<sup>a</sup> 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.

<sup>b</sup> Chromosome numbers and positions refer to genome build GRCh38.

<sup>c</sup> The allele frequency enrichment in Finns with respect to non-Finnish Europeans according to the gnomAD database, version 2.1.1 (gnomad.broadinstitute.org).

<sup>d</sup> The conditional P value is the largest P value from association tests conditioning on other previously reported genome-widely significant markers in the same gene. The conditional P value for rs760351239-T is not available because no genome-widely significant 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.<sup>a</sup>**

Locus and variant	Study	Triglycerides		LDL cholesterol		HDL cholesterol		Total cholesterol	
		Effect (95% CI) <sup>b</sup> mg/dL	P	Effect (95% CI) <sup>b</sup> mg/dL	P	Effect (95% CI) <sup>b</sup> mg/dL	P	Effect (95% CI) <sup>b</sup> mg/dL	P
<i>ANGPTL4</i> <i>rs746226153-G</i>	Finrisk	-14.4 (-18.6 to -9.8)	$4.3 \times 10^{-9}$	2.4 (-2.1 to 7.0)	0.30	1.5 (-0.2 to 3.3)	0.09	-0.2 (-5.3 to 4.9)	0.94
<i>ANGPTL8</i> <i>rs760351239-T</i>	Finrisk	-24.0 (-30.4 to -16.9)	$3.4 \times 10^{-9}$	-16.8 (-24.8 to -8.7)	$4.5 \times 10^{-5}$	9.1 (6.1 to 12.3)	$4.6 \times 10^{-9}$	-14.2 (-23.2 to -5.1)	$2.2 \times 10^{-3}$
<i>ANGPTL8</i> <i>rs145464906-T</i>	UK Biobank	-18.9 (-21.2 to -15.1)	$3.3 \times 10^{-25}$	-4.5 (-7.6 to -1.4)	$4.2 \times 10^{-3}$	6.1 (4.8 to 7.4)	$7.4 \times 10^{-20}$	-1.0 (-5.2 to 3.1)	0.62
<i>CETP</i> <i>rs751916721-T</i>	Finrisk	5.1 (-5.2 to 16.8)	0.34	-7.0 (-16.1 to 2.1)	0.13	16.7 (13.2 to 20.1)	$1.6 \times 10^{-21}$	12.0 (1.9 to 22.2)	0.02
<i>LIPG</i> <i>rs200435657-A</i>	Finrisk	3.7 (-4.6 to 12.8)	0.40	4.1 (-3.4 to 11.7)	0.28	10.2 (7.5 to 13.0)	$5.0 \times 10^{-13}$	17.3 (9.0 to 25.5)	$4.2 \times 10^{-5}$

Abbreviations: CI, confidence interval.

<sup>a</sup> 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.

<sup>b</sup> 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.<sup>a</sup>**

Locus and variant	Cohort	Type 2 diabetes; 23,338 cases in FinnGen				Coronary artery disease; 18,295 cases in FinnGen			
		Allele frequency <sup>b</sup>		OR <sup>c</sup> (95% CI)	<i>P</i>	Allele frequency <sup>b</sup>		OR <sup>c</sup> (95% CI)	<i>P</i>
		Cases %	Controls %			Cases %	Controls %		
<i>ANGPTL4</i> <i>rs746226153-G</i>	FinnGen	0.42	0.54	0.70 (0.60-0.82)	$1.4 \times 10^{-5}$	0.52	0.54	1.00 (0.83-1.21)	1.00
<i>ANGPTL8</i> <i>rs760351239-T</i>	FinnGen	0.09	0.12	0.65 (0.47-0.92)	0.01	0.09	0.12	0.56 (0.38-0.83)	0.004
<i>ANGPTL8</i> <i>rs145464906-T</i>	UK Biobank	0.04	0.05	0.79 (0.48-1.30)	0.35	0.03	0.05	0.73 (0.44-1.19)	0.21
<i>ANGPTL8</i> PTV meta-analysis <sup>d</sup>	FinnGen and UK Biobank	0.06	0.07	0.69 (0.52-0.92)	0.01	0.05	0.07	0.62 (0.45-0.84)	0.002
<i>CETP</i> <i>rs751916721-T</i>	FinnGen	0.10	0.011	0.82 (0.58-1.16)	0.26	0.12	0.11	0.93 (0.62-1.37)	0.70
<i>LIPG</i> <i>rs200435657-A</i>	FinnGen	0.10	0.14	0.67 (0.48-0.91)	0.01	0.13	0.13	0.71 (0.50-1.02)	0.06

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

<sup>a</sup> Shown in the table are the association statistics between the serum lipid-associated PTVs and type 2 diabetes and coronary artery disease risk.

<sup>b</sup> 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[45] 250.2 and 411.4 for type 2 diabetes and coronary artery disease respectively.

<sup>c</sup> In the FinnGen data the 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. The odds ratios for the associations between the disease risks and PTVs in UK Biobank were obtained from Zhou et al.[51]

<sup>d</sup> 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.<sup>a</sup>**

Phenotype	<i>ANGPTL8</i> PTV rs760351239-T				<i>P</i>
	No. of Cases	Allele Frequency <sup>b</sup> Cases      Controls %		OR <sup>c</sup> (95% CI)	
Esophagitis	585	0.77	0.12	174.3 (17.7-1715.1)	$9.7 \times 10^{-6}$
Statin medication	53,518	0.10	0.13	0.53 (0.40-0.71)	$1.2 \times 10^{-5}$
Sensorineural hearing loss	12,250	0.22	0.11	2.45 (1.63-3.68)	$1.8 \times 10^{-5}$

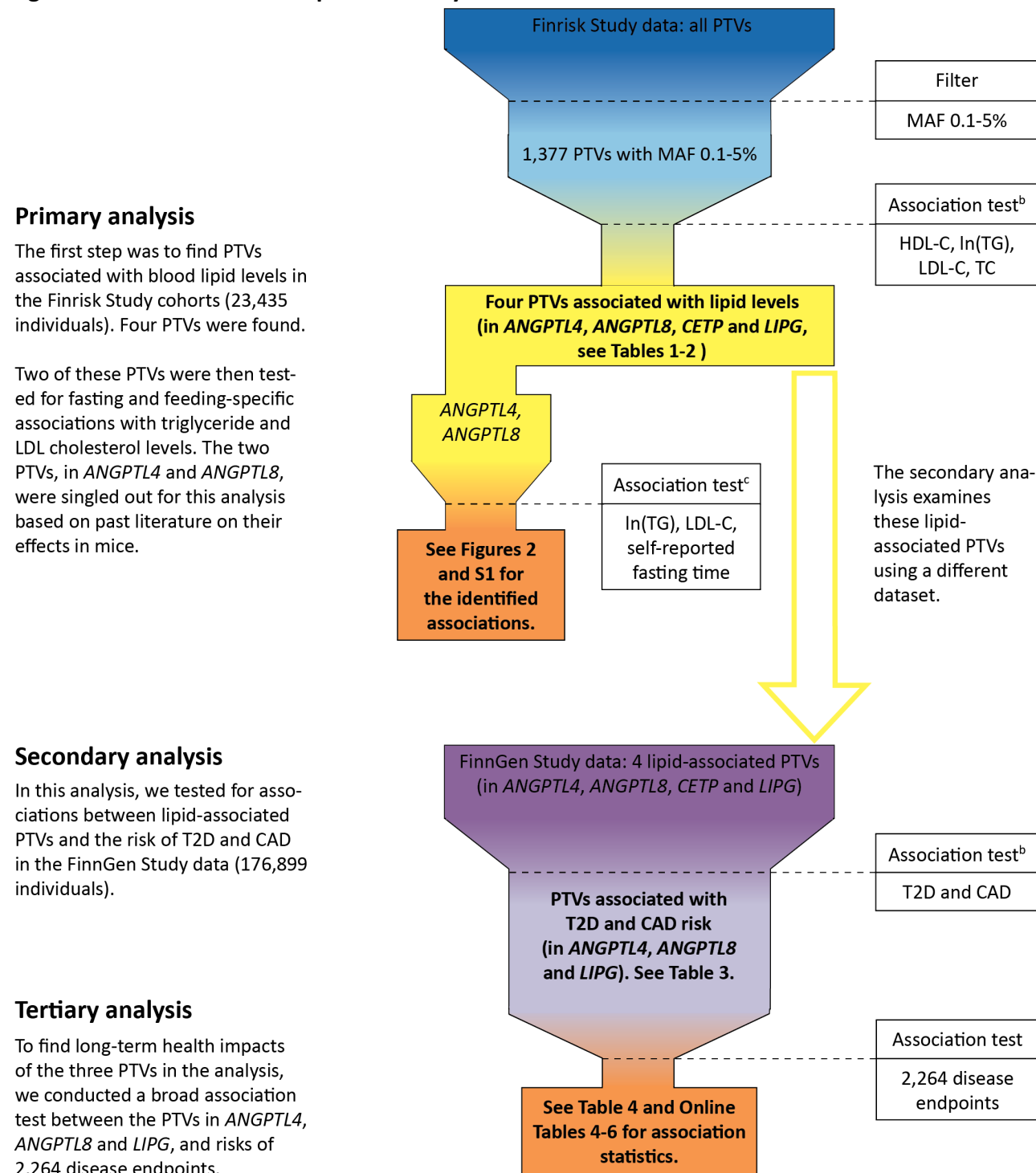
Abbreviations: PTV, protein-truncating variant; OR, odds ratio; CI, confidence interval.

<sup>a</sup> Shown in the table are the phenome-wide significant associations ( $P < 2.2 \times 10^{-5}$ ) for the rs760351239-T allele in the FinnGen Study.

<sup>b</sup> The allele frequencies are reported in percent.

<sup>c</sup> 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.<sup>a</sup>**



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.

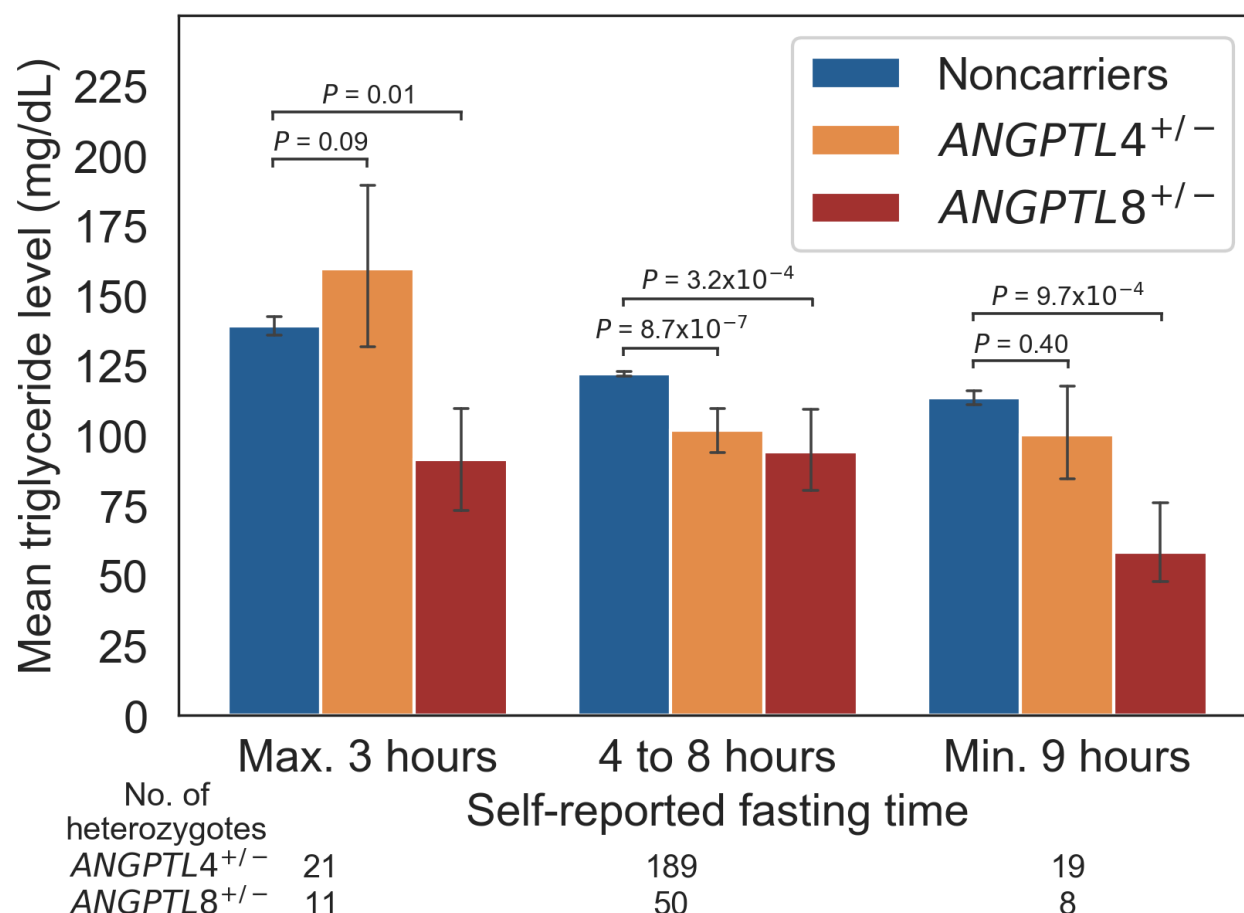
<sup>a</sup> Shown in the figure is an overview of the present study. The final results are shown in boxes with a bolded font.

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

<sup>c</sup> 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.<sup>a</sup>**



<sup>a</sup> 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.