

RESEARCH ARTICLE

Multiple rare and common variants in *APOB* gene locus associated with oxidatively modified low-density lipoprotein levels

Eleonora Khlebus^{1,2*}, Vladimir Kutsenko^{1,3}, Alexey Meshkov¹, Alexandra Ershova¹, Anna Kiseleva¹, Anton Shevtsov³, Natalia Shcherbakova¹, Anastasiia Zharikova¹, Vadim Lankin⁴, Alla Tikhaze⁴, Irina Chazova⁴, Elena Yarovaya³, Oksana Drapkina¹, Sergey Boytsov⁴

1 Federal State Institution National Medical Research Center for Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Moscow, Russia, **2** Moscow Institute of Physics and Technology (State University), Moscow, Russia, **3** Lomonosov Moscow State University, Moscow, Russia, **4** Federal State Budget Organization National Medical Research Center of Cardiology of the Ministry of Healthcare of the Russian Federation, Moscow, Russia

* eleonora.khlebus@phystech.edu



OPEN ACCESS

Citation: Khlebus E, Kutsenko V, Meshkov A, Ershova A, Kiseleva A, Shevtsov A, et al. (2019) Multiple rare and common variants in *APOB* gene locus associated with oxidatively modified low-density lipoprotein levels. PLoS ONE 14(5): e0217620. <https://doi.org/10.1371/journal.pone.0217620>

Editor: Farook Thameem, University of Texas Health Science Center at San Antonio, UNITED STATES

Received: October 10, 2018

Accepted: May 15, 2019

Published: May 31, 2019

Copyright: © 2019 Khlebus et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data set containing fragment coordinates for TargetSeq Custom Enrichment Kit which was designed to target sequencing the region containing the complete genomic sequence of the *APOB* gene in locus 2p24-p23 (chr2: 20996301-21494945; GRCh37/hg19 reference human genome) are within the Supporting Information files. Public sharing of other data presented in the article is limited by law in Russia. According to the Article 12 «Cross-

Abstract

Oxidatively modified low-density lipoproteins (oxLDL) play an important role in the occurrence and progression of atherosclerosis. To identify the genetic factors influencing the oxLDL levels, we have genotyped 776 DNA samples of Russian individuals for 196,725 single-nucleotide polymorphisms (SNPs) using the Cardio-MetaboChip (Illumina, USA) and conducted genome-wide association study (GWAS). Fourteen common variants in the locus including *APOB* gene were significantly associated with the oxLDL levels ($P < 2.18 \times 10^{-7}$). These variants explained only 6% of the variation in the oxLDL levels. Then, we assessed the contribution of rare coding variants of *APOB* gene to the oxLDL levels. Individuals with the extreme oxLDL levels (48 with the lowest and 48 with the highest values) were selected for targeted sequencing of the region including *APOB* gene. To evaluate the contribution of the SNPs to the oxLDL levels we used various statistical methods for the association analysis of rare variants: WST, SKAT, and SKAT-O. We revealed that both synonymous and nonsynonymous SNPs affected the oxLDL levels. For the joint analysis of the rare and common variants, we conducted the SKAT-C testing and found a group of 15 SNPs significantly associated with the oxLDL levels ($P = 2.14 \times 10^{-9}$). Our results indicate that the oxLDL levels depend on both common and rare variants of the *APOB* gene.

Introduction

Atherosclerosis is a complex multifactorial disease that is a major cause of cardiovascular disorders and a leading cause of mortality in developed countries. Over the last few decades, it was also suggested that the low density lipoprotein (LDL) oxidation plays an important role in the development and progression of atherosclerosis and its complications [1–4].

border data transmission» of Federal Law of July 27, 2006 No.152 "Concerning Personal Data" (updated to September 1, 2015), we are not able to transfer the data. Data requests may be sent to Biobank at the National Medical Research Center for Preventative Medicine (biobank@gnicpm.ru), the corresponding author Eleonora Khlebus (eleonora.khlebus@phystech.edu), or principal investigator and head of the data steering committee Alexey Meshkov (meshkov@lipidclinic.ru).

Funding: This research was supported by grant of Russian Scientific Foundation 14-15-00245.

Competing interests: The authors have declared that no competing interests exist.

Under oxidative stress, accompanying the atherosclerosis development, the oxidized (lipoperoxides-containing) and/or oxidatively modified (containing the apoprotein modified by the secondary oxidation products—dicarbonyls) LDL particles accumulate in plasma [3, 4]. The oxidatively modified LDL (unlike oxidized LDL) is actively captured by macrophages, leading to the formation of the lipid enriched foam cells [5].

Previously, by performing the genome-wide association study (GWAS) in the Finnish population, it was shown that the oxLDL levels can be genetically determined [6]. However, human GWASs have important limitations. GWASs are typically focused on the common genetic variants with the minor allele frequency (MAF) greater than 0.05 and even the very large GWASs explain only a small fraction of the estimated heritability [7]. In the previously mentioned GWAS, the top significant SNP (rs676210) explained only 11% of the variation in the oxLDL [6]. In the current research, we have replicated the significant associations in an independent study of a Russian cohort and explained only 6% of the variation in the oxLDL levels.

One of the possible explanation of the 'missing heritability' is given by an underestimation of rare variants [8–10]. An association study of the low-frequency and rare coding variants with the blood lipids and coronary heart disease was performed previously [11]. The contribution of the rare variants to the low levels of high density lipoprotein (HDL) cholesterol was studied [12]. Targeted sequencing studies in subjects with the low cholesterol levels detected the rare mutations in *LDLR* [13], *PCSK9* [14], and *NPC1L1* [15] genes. Also, an effect of common and rare gene variants on plasma LDL cholesterol was assessed [16]. However, the distinct role of rare variants in the LDL oxidation has yet to be understood as well as their potential role in atherosclerosis. Therefore, to fill the heritability gap, we have verified the hypothesis that the rare (MAF<0.01) or low-frequency variants (MAF 0.01–0.05), which are not well covered by GWASs and not easily imputed, are also associated with the oxLDL levels.

To assess the contribution of the SNPs with the low MAF, the variants aggregation tests were developed. Instead of testing each SNP individually, these tests evaluate the cumulative effects of variants [17]. In our work, to find an association between the genetic variants and the oxLDL levels, we used the weighted-sum test (WST) [18], approaches based on the sequence kernel association test (SKAT) [19–22] and the adaptive combination of *P*-values method (ADA) [23, 24].

Here, we present the first results on a GWAS with oxLDL levels in Russian cohort. Furthermore, we report the association between the oxLDL levels and coding variants identified in *APOB* gene in the groups of patients with the extreme low and high oxLDL levels. Thus, for the first time in literature, we show the contribution of the common SNPs, as well as of the low-frequency and rare variants, to the oxLDL levels.

Materials and methods

For more information, see [S1 Appendix](#).

Study subjects

Study subjects were recruited from the Russian study "Approbation and implementation of new approaches to prevention, diagnosis, and treatment of atherosclerosis in outpatient settings by the example of the Western Administrative District of Moscow" from August to December 2009. OxLDL levels were measured for 776 patients with various cardiovascular risk according to the SCORE [25], these patients were selected for genotyping and GWAS analysis. DNA samples from 48 individuals with the highest oxLDL levels and 48 individuals with the lowest oxLDL levels out of total cohort were selected for targeted sequencing (Fig 1). The

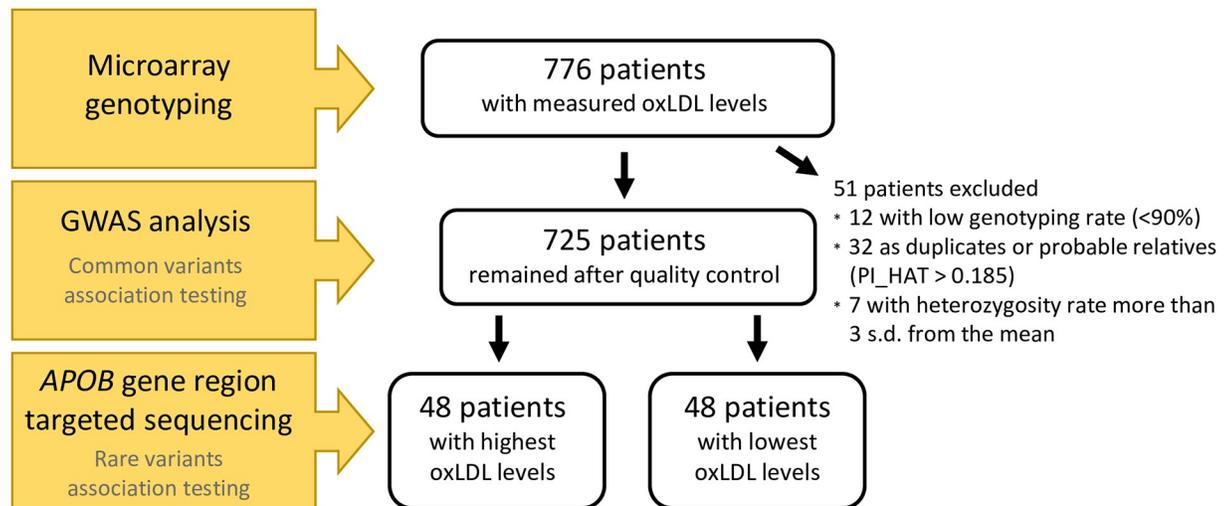


Fig 1. Cohort diagram indicating a number of patients participating into different steps of research.

<https://doi.org/10.1371/journal.pone.0217620.g001>

study was approved by the Russian Cardiology Research and Production Complex, A.L. Myasnikov Institute of Clinical Cardiology (Committee on the Ethics issues in clinical cardiology, protocol No.144, 27 April 2009). Written informed consent was obtained from all participants after approval by the ethical committee.

Laboratory tests

Blood samples were taken from each of the subjects in the morning after he or she had fasted overnight. Circulating serum oxLDL levels were assayed by Oxidized LDL ELISA kit (Merco-dia, Sweden) according to the protocol [26].

Atherosclerosis progression is accompanied by oxidative stress, while malondialdehyde (MDA) and other low molecular weight dicarbonyls can accumulate in blood plasma, including MDA homologue glyoxal and MDA isomer methylglyoxal [4]. Glyoxal and methylglyoxal, like MDA, can cause the atherogenic modification of LDL [4, 26, 27]. Previously we have determined that by using Oxidized LDL ELISA kit (Merco-dia, Sweden) we can measure not any oxidized LDL, but mainly MDA-modified LDL (S1 Fig) [28]. Further under oxidatively modified LDL (oxLDL) we mean only these MDA-modified LDL.

Total cholesterol (TC), triglycerides (TG), HDL, apolipoprotein A1 (ApoA1), apolipoprotein B (ApoB), high-sensitivity C-reactive protein (CRP), and lipoprotein (a) levels were measured using an automatic biochemistry analyzer ARCHITECT c8000 (Abbott Laboratories, USA). LDL was calculated according to the Friedewald formula [29] and in the case of $TG > 4.5$ mmol/l LDL was estimated by the direct method using the same analyzer ARCHITECT c8000.

Measurement of intima-medial thickness and plaque parameters

High-resolution B-mode ultrasonography was performed with a 11-3 MHz linear-type probe (PHILIPS iE33 ultrasound system, Philips Inc., Eindhoven, Netherlands). All measurements were taken in the common carotid artery, carotid bulb and proximal segment of internal carotid artery. Three different longitudinal views (anterior oblique, lateral, and posterior oblique) of both carotid systems and transverse views of all plaques were obtained. A more detailed procedure was described in an earlier publication [30]. The individual value of mean intima-

medial thickness (mean-IMT) was the mean of mean-IMTs of the right and left carotid arteries.

The presence of atherosclerotic plaque was estimated at 6 sites of carotid pool: the whole length of both common carotid arteries, both bifurcations, and both internal carotid arteries. Plaque was defined as a focal structure that encroached into the arterial lumen by at least 0.5 mm or 50% of the surrounding IMT value or demonstrated a thickness of ≥ 1.5 mm as measured from the media-adventitia interface to the intima-lumen interface [31]. Plaque number was defined as the sum total of the plaques. An individual value of total stenosis was defined as a sum of maximum reductions in the percent diameter stenosis of all carotid plaques.

Microarray-based genotyping and quality control

Genotyping was performed by using Cardio-MetaboChip (Illumina, USA), designed to test 196,725 SNPs. Quality control for patients and SNPs was conducted by PLINK (v 1.07) [32]. SNPs which passed quality control had parameters: call rate >0.95 , MAF >0.05 , Hardy-Weinberg equilibrium $P > 1.0 \times 10^{-5}$. Also, we excluded duplicates or probable relatives based on pairwise identity by state (PI_HAT > 0.185) and samples with heterozygosity rates more than 3 s.d. from the mean. To determine whether our sequence variations were caused by the population stratification, we assayed genotyping data using the principal component analysis (PCA). The PCA revealed no evidence of differences in genetic ancestry between samples (S2 Fig).

Target enrichment and DNA sequencing

Targeted sequencing was performed using the TargetSeq Custom Enrichment Kit (Thermo Fisher Scientific, USA) on the SOLiD 5500W system (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. TargetSeq Custom Enrichment Kit was designed to target the region containing the complete genomic sequence of the *APOB* gene in locus 2p24-p23 (chr2: 20996301-21494945; GRCh37/hg19 reference human genome). The kit consists of 536 fragments accounting for a total of 391 833 bp. Unique probes were designed using the Sequence Search and Alignment by Hashing Algorithm (SSAHA) [33]. Capture design coordinates are provided in S1 Dataset.

Variant calling, postprocessing, and multiple alignment

Mapping and variant calling were performed with LifeScope Genomic Analysis Software for SOLiD Next-Generation Sequencing (GRCh37/hg19 reference human genome). We used Samtools [34] for duplicate reads removal. To get coverage data we used BedTools [35]. We filtered variants using 10x coverage threshold. Each base with low-quality (Phred score < 30) sequence was also removed. Value 30 means that read assigned a Phred mapping quality with this score has a 1 in 1000 chance of being misaligned [36]. We used ANNOVAR software [37] for SNPs annotation. Functional effects of the detected variants were assessed with the SIFT [38] and PolyPhen-2 [39] algorithms. Multiple protein alignment was obtained using MUSCLE [40] and visualized using Jalview [41].

Statistical analysis

Univariable statistical analysis was performed with Statistica v.8.0. *P*-value of less than 0.05 was considered to be statistically significant. Data were presented as median (25th–75th percentile). *P*-value for quantitative parameters was calculated using a non-parametric Mann-

Whitney test. *P*-value for quality parameters was calculated using Yates' corrected χ^2 test. If the sample size included five subjects or fewer, a two-tailed Fisher's exact test was used. Correlation analysis was performed by Spearman's rank correlation test. For estimation of variation in the oxLDL levels non-adjusted R^2 was used.

Association of common SNPs, obtained by microarray genotyping, with oxLDL levels in GWAS was tested using the PLINK software (v 1.07) [32]. Also, we examined the SNPs association with the other parameters: oxLDL/LDL and oxLDL/ApoB ratios, TC, TG, LDL, HDL, CRP, ApoA1 and ApoB.

For the analysis of rare and low-frequency variants obtained by targeted sequencing, we used weighted-sum test (WST) [18] and sequence kernel association test (SKAT) [19]. Also, we used optimal SKAT (SKAT-O) method which is based on combination of burden test and SKAT [20]. As covariates we used age, sex, smoking status, body mass index, waist, HDL, CRP, lipoprotein (a), hypertension, myocardial infarction, diabetes mellitus, stroke, and statins use. For the joint analysis of rare, low-frequency and common variants, we applied combined SKAT (SKAT-C) [21]. For the analysis of rare and low-frequency variants by WST we used the custom script in programming language R, for methods SKAT, SKAT-O and SKAT-C R-package [42]. Variants selection was performed by using results of adaptive combination of *P*-values (ADA) method [24] and SKAT backward elimination (BE-SKAT) test [22]. To select a subgroup of associated variants, as suggested at [43], we used the elastic net [44] from R-package [45] with AUC-ROC-based cross-validation.

Results and discussion

GWAS for finding variants affecting the oxLDL levels

We studied 776 DNA samples of Russian individuals with measured oxLDL levels for 196,725 SNPs genotyped by the Cardio-MetaboChip (Illumina, USA). Based on the quality control, we selected 725 patients and 101,704 SNPs. Clinical and laboratory characteristics of the patients are shown in Table 1. Under oxLDL we mean the MDA-modified LDL [28].

The oxLDL levels varied from 21.03 to 163.72 U/dl (with the median of 68.5) and correlated with the levels of the TC, TG, LDL, CRP and ApoB, and with the ultrasound markers of atherosclerosis (S1 Table). GWAS analysis was performed for oxLDL levels and for biochemical parameters correlated with it.

The Manhattan plot (Fig 2), quantile-quantile plot (S3 Fig) and regional plot (S4 Fig) [46] illustrate the results of the association analyses performed for the oxLDL levels. We identified 14 significant SNPs ($P < 2.18 \times 10^{-7}$) on the chromosome 2 (S2 Table). Twelve out of fourteen significant variants were localized in the intronic regions of *APOB* gene and in the intergenic regions near this gene. Other two SNPs were *APOB* nonsynonymous (NS) exonic variants: rs1042034 (p.Ser4338Asn) and rs676210 (p.Pro2739Leu). The association of the variants from chromosome 2 with the oxLDL/LDL and oxLDL/ApoB ratios was also found. GWAS results for other biochemical parameters were not significant in this study.

All 14 significant SNPs, found in our study and associated with the oxLDL levels, are in concordance with the variants obtained earlier in the Finnish population and replicated in the German cohorts [6]. In the latter study, only one independent SNP (rs676210) was declared as a true functional variant, which explained only 11% of the oxLDL variation. Present study showed that our significant variants explained only 6% of oxLDL variation ($R^2 = 6\%$). To this end, we suggested that the rare and low-frequency variants could be crucial for the LDL oxidation. We used the targeted sequencing of *APOB* gene locus to find other functional variants explaining an additional variation in the oxLDL levels.

Table 1. Characteristics of 725 patients cohort and individuals with low and high oxLDL levels.

Parameter	Total cohort, n = 725	Low oxLDL levels group, n = 48	High oxLDL levels group, n = 48	P-value*
Age, years	57 (51-63)	59 (50-64)	56 (51-62)	0.703
Men, n (%)	206 (28.4)	21 (43.8)	17 (35.4)	0.531
Smoking, n (%)	108 (14.9)	7 (14.6)	12 (25.0)	0.306
BMI, kg/m ²	28.7 (25.8-32.2)	28.7 (24.9-31.2)	28.3 (26.1-32.0)	0.613
Waist, cm	93 (84-102)	94 (84-101)	90 (85-102)	0.946
Total cholesterol, mmol/l	5.90 (5.17-6.88)	4.25 (3.90-4.88)	7.66 (7.01-8.25)	1.17 × 10 ⁻¹⁴
Triglycerides, mmol/l	1.60 (1.14-2.19)	1.22 (0.84-1.71)	2.20 (1.67-2.91)	1.55 × 10 ⁻⁷
LDL, mmol/l	3.60 (2.92-4.39)	2.31 (1.90-2.60)	4.86 (3.93-5.61)	2.04 × 10 ⁻¹⁴
HDL, mmol/l	1.34 (1.14-1.58)	1.29 (1.10-1.50)	1.20 (1.06-1.48)	0.322
Apolipoprotein B, mg/dl	102 (87-123)	72 (64-80)	142 (125-154)	7.88 × 10 ⁻¹⁴
Apolipoprotein A1, mg/dl	162 (144-184)	151 (139-176)	161 (141-181)	0.347
Lipoprotein (a), mg/dl	11.4 (5.2-33.5)	8.1 (3.4-22.8)	13.4 (6.2-33.4)	0.078
oxLDL, U/dl	68.50 (55.72-85.64)	35.29 (31.02-39.64)	118.35 (113.47-124.12)	3.15 × 10 ⁻¹⁷
oxLDL/LDL, U/dl per mmol per l	19.32 (16.21-23.00)	15.75 (12.64-18.53)	23.79 (20.87-28.00)	8.49 × 10 ⁻¹³
oxLDL/ApoB, U/dl per mg per dl	0.67 (0.56-0.78)	0.48 (0.42-0.55)	0.82 (0.75-1.01)	4.05 × 10 ⁻¹⁵
CRP, mg/dl	0.24 (0.13-0.42)	0.19 (0.09-0.37)	0.28 (0.16-0.45)	0.031
Hypertension, n (%)	586 (80.8)	39 (81.3)	37 (77.1)	0.802
Myocardial infarction, n (%)	71 (9.8)	7 (14.6)	7 (14.6)	0.772
Diabetes mellitus, n (%)	113 (15.6)	9 (18.8)	6 (12.5)	0.544
Stroke, n (%)	16 (2.2)	3 (6.3)	4 (8.3)	1.000
Statins, n (%)	163 (22.5)	19 (39.6)	13 (27.1)	0.279
Total stenosis, %	80 (25-130)	63 (23-115)	75 (25-150)	0.362
Plaque number	3 (1-4)	2 (1-4)	3 (1-5)	0.315
Mean-IMT, mm	0.71 (0.62-0.85)	0.71 (0.63-0.85)	0.71 (0.61-0.86)	0.916

BMI—body mass index, LDL—low-density lipoprotein, HDL—high-density lipoprotein, CRP—high-sensitivity C-reactive protein, oxLDL—oxidatively modified low-density lipoprotein, ApoB—Apolipoprotein B, Mean-IMT—mean intima-medial thickness. Data are presented as numbers (percentages) in cases of categorical data and median (25th–75th percentile) in cases of continuous data.

* P-value (difference between low and high oxLDL levels groups) for quantitative parameters was calculated for non-parametric Mann-Whitney test, for quality parameters P-value was calculated using two-tailed Fisher’s exact test if it is available, otherwise—for Yates’ corrected χ^2 test.

<https://doi.org/10.1371/journal.pone.0217620.t001>

Targeted sequencing for finding variants of *APOB*

For this part of our study, we selected those 96 patients out of 725 individuals who had an extremely high (H group) and low (L group) oxLDL levels: 48 individuals with the lowest oxLDL levels (with the median of 35.29 U/dl) and 48 individuals with the highest ones (with the median of 118.35 U/dl). These groups differed significantly by oxLDL levels ($P = 3.15 \times 10^{-17}$ by χ^2 test). H group was matched to the L group with respect to age, sex, smoking status, body mass index, waist, levels of HDL, ApoA1, and lipoprotein (a). The patients were also matched for the prevalence of hypertension, myocardial infarction, diabetes mellitus, stroke, and statin use. The severity of carotid atherosclerosis according to the results of carotid ultrasound were comparable between the study groups. Individuals from H group demonstrated significantly higher levels of TC, TG, LDL, ApoB, and CRP. The characteristics of two groups are provided in Table 1.

For targeted sequencing we designed the locus 2p24-p23 which included *APOB* gene (chr2: 20996301-21494945; GRCh37/hg19 reference human genome) and 14 significant variants according to our GWAS analysis. We conducted a targeted sequencing of designed locus and

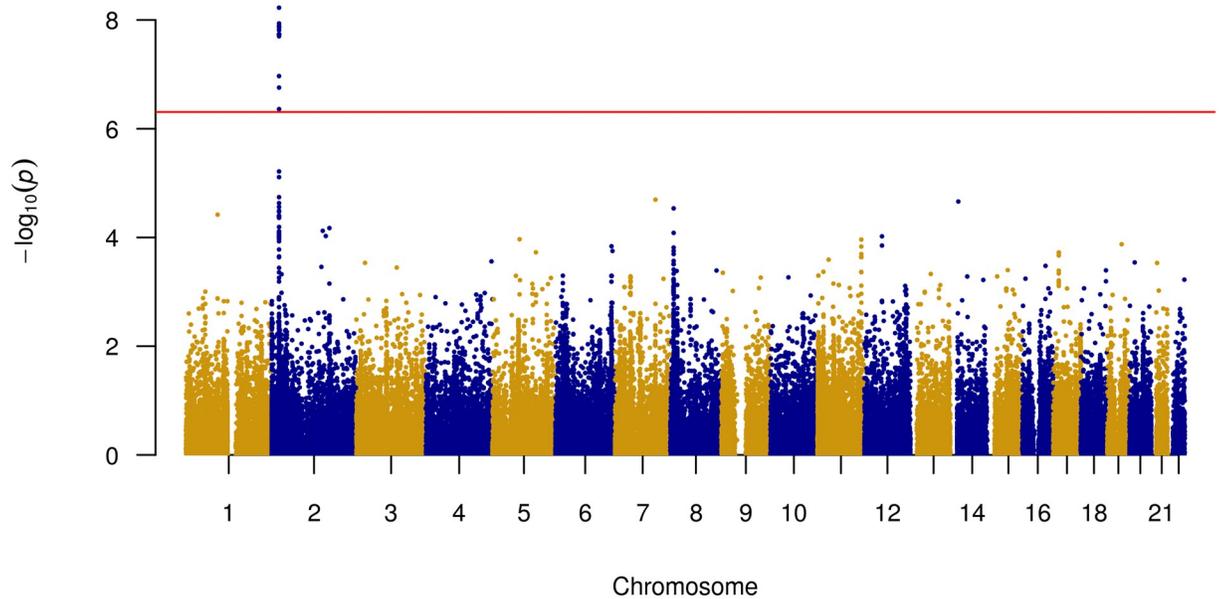


Fig 2. Manhattan plot of GWAS results showing highest SNPs associated with oxLDL levels. The dots represent the SNPs organized by the chromosomal order and position. The y-axis describes the statistical significance (expressed as $-\log_{10}$ of the P -values).

<https://doi.org/10.1371/journal.pone.0217620.g002>

identified a total of 1,992 SNPs; 30 SNPs were exonic (Table 2) with 23 SNPs leading to the NS amino acid change including one nonsense mutation.

Eleven variants were present only in the L group. Six of them were rare NS variants: rs1801703 (p.Val4128Met), rs1042023 (p.Gln3432Glu), rs148170480 (p.Val2539Ile), rs368970025 (p.Val1955Met), rs374427541 (p.Pro1343Ser), rs12713843 (p.Arg1128His). Two NS variants unique for the L group were not singletons, they were found more than in one patient. Rare NS variant rs1801703 (p.Val4128Met) was found in two individuals, and the NS low-frequency SNP rs533617 (p.His1923Arg) was presented in three individuals. Moreover, two patients from the L group had synonymous change rs1799812 (p.Leu3377Leu) not presented in the H group.

Eight SNPs were unique to the H group. All NS variants out of them were rare singletons: five previously reported (p.Ser4392Asn, rs61743502 (p.Val4265Ala), rs12713540 (p.Ser3801Thr), rs12713844 (p.Asp1113His), rs12714097 (p.Pro877Leu)) and two novel variants (p.Lys4463Glu and p.Lys2467Ter). Also, two patients from the H group had synonymous change rs72653093 (p.Leu2538Leu) presented only in this group of patients.

For each variant out of 30 exonic variants obtained by the targeted sequencing, the MAF was received from the Exome Aggregation Consortium (ExAC) database (<http://exac.broadinstitute.org/>). There were two novel variants with no MAF data in our SNPs set. In order to account for these variants in the variant association testing, their MAF was defined as the minimal MAF observed in our SNPs set of 30 exonic variants.

Additionally, to exclude linked variants we performed linkage disequilibrium (LD) analysis based on 1000 Genomes Project data (S5 Fig). Two common variants rs676210 and rs1042034 in the considered SNPs set were in a perfect LD according to the public data; moreover, they were in a perfect LD according to our study. We excluded rs1042034 variant. The other four variants (rs6413458 and rs1801702, rs1799812 and rs1801703), which are in LD according to 1000 Genomes Project data, had the low MAF and were not linked in our study. We did not

Table 2. Exonic SNPs of APOB found by targeted sequencing in patients with high and low oxLDL levels.

Genomic position (GRCh37/ hg19)	Exon	Amino acid change	SNP ID	SIFT/ Poly-Phen-2	MAF (ExAC)	**Low group	**High group	Previously published with respect to
21224853	29	p.Ala4481Thr	rs1801695	D / B	0.02407	47,1,0	48,0,0	HDL [47]; oxLDL [6, 28]; Dementia [48]
21224854*	29	p.Gln4480Gln			0.000008267	47,1,0	48,0,0	No
21224907	29	p.Lys4463Glu		D / B		48,0,0	47,1,0	No
21225119	29	p.Ser4392Asn		T / B	0.00004974	48,0,0	47,1,0	Familial hypercholesterolemia [49]
21225281	29	p.Ser4338Asn	rs1042034	T / B	0.7057	7,24,17	1,6,41	HDL, TG [50]; LDL [51–53]; Ischemic Stroke [54]; Hyperlipidemia [55]; oxLDL [6]; TC [52, 53, 56, 57]
21225485	29	p.Arg4270Thr	rs1801702	T / B	0.0456	46,2,0	45,3,0	TC, LDL [58]
21225500	29	p.Val4265Ala	rs61743502	T / B	0.005139	48,0,0	47,1,0	Familial hypercholesterolemia [59]
21225753	29	p.Glu4181Lys	rs1042031	T / B	0.153	41,7,0	37,9,2	ADH [60]; Carotid plaques [61]; Hipertension, TG [62]; Serum lipid levels [63]; Familial hypercholesterolemia [64, 65]; HDL [58]; HDL, LDL [66]; Breast Cancer [67]; Ischemic Stroke [68]
21225912	29	p.Val4128Met	rs1801703	T / B	0.006171	46,2,0	48,0,0	Exceptional longevity [69]
21228339	26	p.Ser3801Thr	rs12713540	T / P	0.001046	48,0,0	47,1,0	PDR [70]
21228827	26	p.Arg3638Gln	rs1801701	T / B	0.06887	43,5,0	42,6,0	LDL [51]; TG [63]; CAD [71]; Familial hypercholesterolemia [64]; Ischemic Stroke [68]
21229446	26	p.Gln3432Glu	rs1042023	T / B	0.007588	47,1,0	48,0,0	LDL receptor binding [72, 73]; Familial hypercholesterolemia [74]
21229609	26	p.Leu3377Leu	rs1799812		0.006083	46,2,0	48,0,0	ADH [60]
21231524	26	p.Pro2739Leu	rs676210	D / D	0.2928	17,24,7	41,6,1	MI [75]; oxLDL [6, 28]; TC [56]; Lipid metabolism phenotypes [76]; Familial hypercholesterolemia [64]
21231592	26	p.Ile2716Ile	rs6413458		0.01952	47,1,0	45,3,0	Lp-PLA2 [77]; ADH [60]
21232125*	26	p.Val2539Ile	rs148170480	T / B	0.002399	47,1,0	48,0,0	No
21232128*	26	p.Leu2538Leu	rs72653093		0.001509	48,0,0	46,2,0	No
21232195	26	p.Thr2515Thr	rs693		0.3899	17,24,7	7,22,19	Ischemic Stroke [54]; LDL [78–81]; Carotid plaques [61]; TC [79]; TG [80]; Breast Cancer [67]
21232341	26	p.Lys2467Ter		T / -		48,0,0	47,1,0	No
21233877*	26	p.Val1955Met	rs368970025	T / B	0.00008238	47,1,0	48,0,0	No
21233972*	26	p.His1923Arg	rs533617	D / D	0.03116	45,3,0	48,0,0	LDL, TC [82, 83]; oxLDL [6, 28]
21234915	26	p.Leu1609Leu	rs72653083		0.001726	47,1,0	48,0,0	No
21236221	25	p.Pro1343Ser	rs374427541	D / D	0.00003295	47,1,0	48,0,0	No
21238367*	22	p.Arg1128His	rs12713843	D / B	0.003708	47,0,1	48,0,0	LDL, TC [84]
21238413*	22	p.Asp1113His	rs12713844	D / P	0.006858	48,0,0	47,1,0	Spastic Paraplegia [85]; Familial hypercholesterolemia [86]
21245813*	18	p.Asn902Asn	rs1801700		0.03445	46,2,0	39,9,0	ADH [60]
21245889*	18	p.Pro877Leu	rs12714097	D / D	0.000479	48,0,0	47,1,0	Familial hypercholesterolemia [87]
21249716	15	p.Val730Ile	rs12691202	T / B	0.02504	47,1,0	44,4,0	ADH [60]
21250914	14	p.Ala618Val	rs679899	T / D	0.4857	19,16,13	30,15,3	FH [65]; MI [75]; LDL [51]; Hyperlipidemia [55]; Chronic kidney disease [88]; Coronary heart disease [89]; oxLDL [6, 28]
21263900	4	p.Thr98Ile	rs1367117	T / B	0.2619	29,14,5	18,24,6	LDL, TC [50, 90]; Lipid metabolism phenotypes [76]

SIFT score: T—Tolerated, D—Deleterious; PolyPhen-2 score: B—Benign, P—Possibly damaging, D—Probably damaging; ADH—autosomal dominant hypercholesterolemia, CAD—coronary artery disease, HDL—high density lipoprotein, LDL—low density lipoprotein, Lp-PLA2—lipoprotein-associated phospholipase A2, MI—myocardial infarction, oxLDL—oxidatively modified low density lipoprotein, PDR—proliferative diabetic retinopathy, TC—total cholesterol, TG—triglycerides.

*SNPs significantly associated with oxLDL levels by BE-SKAT;

**Number of patients with regarded variants genotypes presented as homozygous reference allele, heterozygous, homozygous alternative allele.

<https://doi.org/10.1371/journal.pone.0217620.t002>

exclude any of these variants. Thus, we excluded only one variant rs1042034 and for the further variant association testing, we used 29 out of 30 exonic SNPs.

Variant association testing

Initially, to find the association between the genetic variants and the oxLDL levels we used the WST [18], SKAT (with/without adjustment for covariates) [19] and optimal SKAT (SKAT-O) method which is based on combination of burden test and SKAT (with/without adjustment for covariates) [20]. As covariates we used age, sex, smoking status, body mass index, waist, HDL, CRP, lipoprotein (a), hypertension, myocardial infarction, diabetes mellitus, stroke and statins use.

First of all, we analyzed the rare SNPs out of 29 variants. We considered (i) the subgroup of variants defined as deleterious/damaging (DD) by the SIFT and PolyPhen-2 algorithms as the most plausible functional variants, (ii) the subgroup of NS variants, (iii) the entire group of rare variants (with synonymous and NS amino acid change). The same testing steps were conducted for the rare and low-frequency variants together. The results of the analysis are presented in Table 3.

As follows from Table 3, the WST did not overcome the required level of significance in none of the groups, which could be due to the large number of multidirectional and neutral variants. However, the SKAT-O adjusted for covariates did not overcome the level of significance in the group of rare variants as well. The SKAT-O analysis showed that the best combination of the burden test and SKAT method was the SKAT test itself. The behavior of SKAT-O depending on the combination coefficient is shown on S6 Fig. Thus, in the following analysis, we used only the SKAT, because the corresponding *P*-value was calculated more accurately. As it can be seen in Table 3, the addition of covariates decreased the *P*-value.

Testing with the SKAT adjusted for the covariates showed no association between the rare DD variants and the oxLDL levels (*P* = 0.358). The *P*-value of the former with addition of the low-frequency DD variant rs533617 (p.His1923Arg) decreased to *P* = 0.004. The functional prediction algorithms could fail the correct identification of the any NS sequence variations, as it was shown for the SNPs associated with the LDL and the dominant hypercholesterolemia [91]. Therefore, the consideration of other variants seemed reasonable.

The NS variants that alter an amino acid can change the protein function [38], thus such variants are more likely to affect the phenotype compared to the synonymous variants. However, even if the SNP does not change amino acid, it can still affect the gene function by altering the mRNA stability or splicing [92, 93]. The analysis of the NS variants subgroup and of the subgroup with the synonymous and the NS variants showed a significant association with the oxLDL levels according to SKAT adjusted for covariates. The *P*-value was the least in the subgroup with the synonymous and the NS variants. This indicates that both NS SNPs and

Table 3. Association analysis between exonic *APOB* variants and oxLDL levels.

Test	Rare variants (MAF<0.01)					Rare and low-frequency variants (MAF<0.05)				
	WST	SKAT	SKAT with covariates	SKAT-O	SKAT-O with covariates	WST	SKAT	SKAT with covariates	SKAT-O	SKAT-O with covariates
Deleterious/damaging*	0.244	0.464	0.358	0.645	0.443	0.623	0.037	0.004	0.043	0.006
Nonsynonymous	0.829	0.164	0.048	0.253	0.089	0.813	0.164	0.034	0.278	0.065
Synonymous & nonsynonymous	0.789	0.087	0.034	0.126	0.064	0.173	0.017	0.003	0.034	0.006

*By SIFT and PolyPhen-2 algorithms.

<https://doi.org/10.1371/journal.pone.0217620.t003>

synonymous contribute to the oxLDL levels. Additionally, the *P*-value decreased in all subgroups of the rare variants after adding the variants with the higher MAF—low-frequency variants.

Then, we conducted the joint analysis of all 29 variants which included rare, low-frequency and common variants. For this purpose, we used the SKAT-C method which showed the significant association of all 29 SNPs with oxLDL levels ($P = 2.7 \times 10^{-8}$). However, some of these variants could be neutral, so they should be excluded from the analysis. Methods for finding the causal variants differ depending on variants frequency. Therefore, we divided the SNPs set of 29 variants into two subgroups: rare with low-frequency variants and common variants. The scheme of the joint testing of these SNPs is shown on the Fig 3.

For the subgroup of rare and low-frequency variants we applied the ADA [24] and BE-SKAT [21] tests. ADA test chose the variant rs533617 (p.His1923Arg) with the $P = 0.3$. We used ADA implementation based on the Fisher’s test. We presume that the small size of patients group caused the high *P*-values in the univariate analysis by the Fisher’s test, so this could cause the failure of the ADA test. As noted in [23], the BE-SKAT is less conservative than the ADA test, i.e. this method can select more variants. However, the BE-SKAT can leave more neutral variants than ADA. The BE-SKAT selected 9 variants: p.Gln4480Gln, rs148170480

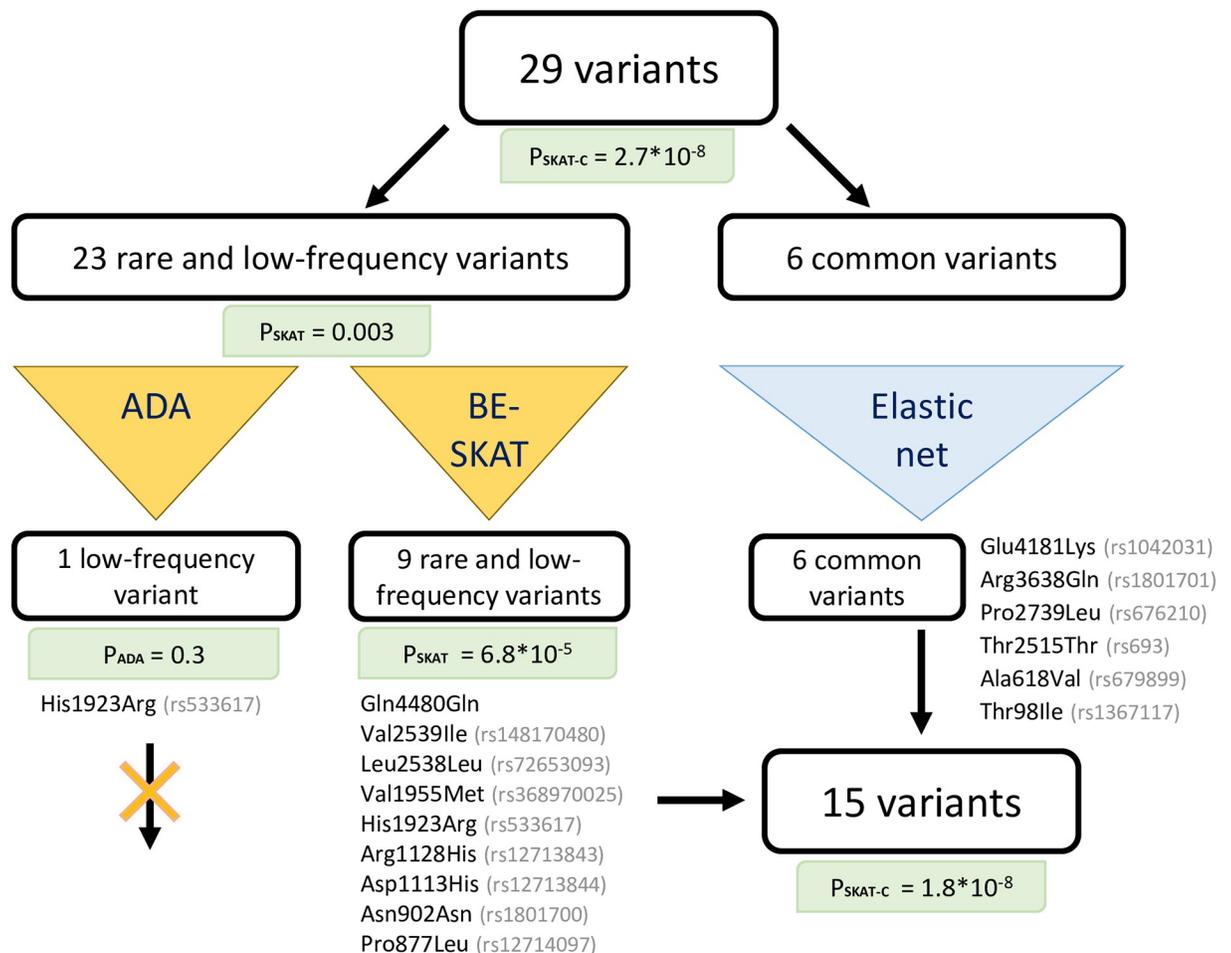


Fig 3. Joint analysis of rare, low-frequency and common exonic variants obtained by targeted sequencing.

<https://doi.org/10.1371/journal.pone.0217620.g003>

(p.Val2539Ile), rs72653093 (p.Leu2538Leu), rs368970025 (p.Val1955Met), rs533617 (p.His1923Arg), rs12713843 (p.Arg1128His), rs12713844 (p.Asp1113His), rs1801700 (p.Asn902Asn), rs12714097 (p.Pro877Leu). The initial *P*-value of the SKAT was reduced from 0.003 in 23 variants group to 6.8×10^{-5} in 9 variants group.

For the subgroup of common variants, we applied an elastic net with cross-validation based on the AUC-ROC metric [44]. According to the present analysis, all considered common variants were used in the model. Thereby, we included all 6 common variants in further analysis.

Finally, we combined the analysis results of both subgroups (rare and low-frequency variants and common variants) and revealed a set of 15 variants for which the SKAT-C method rejected the null hypothesis of the absence of association at a required significance level ($P = 1.8 \times 10^{-8}$). Thus, as the result of our statistical analysis, predominantly neutral variants were excluded from the analysis, reducing the number of variants and the *P*-value.

Six common variants included in the set of 15 SNPs, associated with oxLDL levels by the SKAT-C, previously were associated with the LDL, TG and other lipid parameters and different cardiovascular events (see Table 2). They may affect the oxLDL levels but significant associations were confirmed in the GWAS separately only for the rs676210 variant.

Rare and low-frequency variants were more interesting in the context of finding the 'missing heritability' and explaining the variation in oxLDL levels. The SNP rs533617 (p.His1923Arg) was a potentially causal variant by the statistical analysis. It was presented only in three individuals from the L group and was predicted to be the DD. However, this variant was observed only in patients with the SNP rs676210 significantly associated with the oxLDL by GWAS. For this reason, we could neither confirm nor entirely exclude the contribution of variant rs533617 to the decreased oxLDL levels.

The synonymous variant rs72653093 (p.Leu2538Leu) was presented in two individuals from the H group. It was included in the SNPs set obtained by BE-SKAT and was not published previously. The other non singleton variants rs1801703 (p.Val4128Met) and rs1799812 (p.Leu3377Leu), found only in the L group, were excluded by the BE-SKAT as neutral variants. This could be due to the covariates adjustment or due to the small sample size.

The single patient in the L group was homozygous for variant rs12713843 (p.Arg1128His). This SNP was predicted to be deleterious by the SIFT but benign by PolyPhen-2. This difference may be due to the difference in mathematical algorithms underlying methods. Since SIFT and PolyPhen-2 algorithms do not provide the evidence that this SNP might be the deleterious one for sure, their predictions should be interpreted with caution. Previously the variant rs12713843 was associated with lower LDL and lower TC levels [84]. The carrier of this SNP had lower levels of these parameters in comparison to the median value in the L group (TC = 3.76 mmol/l, LDL = 1.29 mmol/l, both lower than the 25th percentile) and the minimal oxLDL level among the entire cohort (oxLDL = 21.03 U/dl). Thus, the rs12713843 variant may be causative variant for decreasing oxLDL level. However, this decrease may not be caused by the oxidation, but it is due to the reduced LDL and TC levels, described in [84]. Additionally, the carrier of this SNP had the other variant rs676210, associated with the decreased oxLDL level.

NS singletons rs12713844 (p.Asp1113His) and rs12714097 (p.Pro877Leu) presented in the H group were selected by the BE-SKAT. Both were predicted to be deleterious and both previously reported with respect to the familial hypercholesterolemia (Table 2). Novel variants p.Lys4463Glu and p.Lys2467Ter, presented only in the H group, were excluded by the BE-SKAT as being neutral.

	<i>H. sapiens</i>	<i>P. troglodytes</i>	<i>M. mulatta</i>	<i>O. cuniculus</i>	<i>B. taurus</i>	<i>C. familiaris</i>	<i>E. telfairi</i>	<i>M. musculus</i>	<i>R. norvegicus</i>	<i>G. gallus</i>	<i>X. tropicalis</i>	<i>D. rerio</i>	SIFT	PolyPhen-2	
Low oxLDL	Ala4481Thr	A	A	A	A	A	A	V	S	S	S	-	D	B	
	Val4128Met	V	V	L	V	M	L	I	M	V	V	Y	T	T	B
	Gln3432Glu	Q	Q	Q	Q	Q	Q	Q	H	H	N	S	K	T	B
	Val2539Ile*	V	V	V	V	V	V	I	V	V	I	L	A	T	B
	Val1955Met*	V	V	M	L	M	P	M	P	L	K	N	P	T	B
	His1923Arg*	H	H	H	H	H	H	H	H	H	H	H	H	D	D
	Pro1343Ser	P	P	P	P	P	P	P	P	P	P	P	P	D	D
	Arg1128His*	R	R	R	R	R	R	R	R	R	R	R	A	D	B
Low and High oxLDL	Ser4338Asn*	S	N	N	Q	D	N	N	K	K	A	K	N	T	B
	Arg4270Thr	R	R	R	I	M	G	G	R	R	K	K	R	T	B
	Glu4181Lys*	E	E	E	E	G	E	T	K	K	E	A	T	T	B
	Arg3638Gln*	R	R	R	Q	Q	Q	Q	Q	Q	Q	T	Q	T	B
	Pro2739Leu*	P	P	P	P	P	P	P	P	P	P	P	P	D	D
	Val730Ile	V	V	V	V	V	V	V	V	V	V	V	V	T	B
	Ala618Val*	A	A	A	A	A	A	A	A	A	A	A	A	T	D
	Thr98Ile*	T	T	T	I	T	T	T	T	T	S	T	V	T	B
High oxLDL	Lys4463Glu	K	K	K	K	K	K	T	M	R	R	K	-	D	B
	Ser4392Asn	S	S	S	S	T	N	S	S	S	T	N	-	T	B
	Val4265Ala	V	V	V	I	M	V	E	L	V	V	I	I	T	B
	Ser3801Thr	S	S	S	S	S	S	S	S	S	I	I	-	T	P
	Lys2467Ter	K	K	K	K	R	R	K	K	K	K	K	R	T	-
	Asp1113His*	D	D	D	D	D	D	D	D	D	A	D	E	D	P
	Pro877Leu*	P	P	P	P	-	P	P	P	P	P	P	P	D	D

Fig 4. Evolutionary sequence conservation and predicted functional effects of nonsynonymous sequence variations in *APOB*. Conservation for each variation found in the low oxLDL subjects only (top), the both low and high oxLDL subjects (middle), or high oxLDL levels group (bottom). An asterisk (*) indicates the SNPs significantly associated with oxLDL levels by the BE-SKAT. The alignment includes *Homo sapiens* (human), *Pan troglodytes* (chimpanzee), *Macaca mulatta* (Rhesus monkey), *Oryctolagus cuniculus* (rabbit), *Bos taurus* (cow), *Canis lupus familiaris* (dog), *Echinops telfairi* (hedgehog), *Mus musculus* (house mouse), *Rattus norvegicus* (rat), *Gallus gallus* (chicken), *Xenopus tropicalis* (frog), *Danio rerio* (zebrafish). The predicted effect of each variant on protein function is indicated (right columns). SIFT: T—tolerated; D—deleterious. PolyPhen-2: B—benign; P—possibly damaging; D—probably damaging.

<https://doi.org/10.1371/journal.pone.0217620.g004>

APOB gene evolutionary conservation analysis

To examine the relationship between the evolutionary conservation and functional effects of the NS sequence variations, identified by targeted sequencing, we aligned the human ApoB amino acid sequence with the sequences from 11 other species (Fig 4).

Two SNPs rs533617 (p.His1923Arg) and rs374427541 (p.Pro1343Ser) in the L group were highly conserved from human to zebrafish and one (rs12713843 (p.Arg1128His)) to frog. There were three substituted amino acids conserved in primates and some other mammals: rs1801695 (p.Ala4481Thr), rs1042023 (p.Gln3432Glu), rs148170480 (p.Val2539Ile).

In the H group two changed amino acid residues rs12713540 (p.Ser3801Thr) and rs12713844 (p.Asp1113His) were completely conserved from primates to rats, and rs12714097 (p.Pro877Leu) was conserved from human to zebrafish, except cow.

Both groups included three common variants highly conserved from human to zebrafish: rs676210 (p.Pro2739Leu), rs12691202 (p.Val730Ile) and rs679899 (p.Ala618Val).

It was initially assumed that highly conserved NS variation at amino acid residues would be found only in either L or H group, as demonstrated previously for the genetic variation in *NPC1L1* [15]. However, we found several variants common to both extremes. Also, it was shown previously for the variants of *PCSK9* [91] and *APOB* [51].

It is also interesting that two SNPs from the L group were highly conserved in all regarded species and had the DD status. While we suggested that the low oxLDL level is a healthier trait in comparison with the increased oxLDL, high evolutionary sequence conservation suggests that the variants, responsible for the lowering of the oxLDL levels, also cause the loss of any other important functions. Thus, the positive effect of variants which decrease oxLDL level and deleterious effect revealed by *in silico* prediction methods remain controversial.

Conclusion

The major finding of this study is association of the oxLDL levels with both common and rare variants of the *APOB* gene. For the first time, we showed the association of the rare variants with the circulating oxLDL levels. Additionally, we performed the joint analysis of rare, low-frequency and common *APOB* variants. Whereas the group of variants associated with oxLDL levels was revealed, the rare or novel variants still should be interpreted cautiously. It is necessary to evaluate the variant and the gene in the context of the patient's and family's history, physical examinations, and previous laboratory tests to distinguish between variants that cause the patient's disorder and those that are benign. Functional evaluation can also provide a more definitive assessment of the variant pathogenicity.

APOB gene variants can change secondary structure of the human ApoB and the LDL particle size in comparison with the wild type LDL particle [94]. These changes may also impact the LDL oxidation. To determine the molecular biological effects of the genetic variant as well as the interactions with other variants and to understand the mechanisms of oxidation, mathematical modeling and high-throughput technologies such as mass spectrometry or infrared spectroscopy analysis could be applied. Such studies of *APOB* variants may provide useful information to understand their role in the LDL oxidation and atherosclerosis progression.

Supporting information

S1 Fig. Binding specificity of native low density lipoprotein (nLDL), malondialdehyde-modified LDL (MDA-LDL), glyoxal-modified LDL (G-LDL) and methylglyoxal-modified LDL (MG-LDL) to the mAb-4E6 antibody.

(TIF)

S2 Fig. Principal component analysis (PCA) for GWAS samples. A. The colors ranged from red (high oxLDL levels) to blue (low oxLDL levels). The matching suggested minimal evidence of population stratification. B. The colors indicate 48 individuals with the highest oxLDL levels (red) and 48 individuals with the lowest oxLDL levels (blue).
(TIFF)

S3 Fig. Quantile-quantile plot of the observed and expected P -values obtained from GWAS with oxLDL levels showing a clear deviation from normal distribution.
(TIF)

S4 Fig. Regional plot of the *APOB* region with SNPs significantly associated with oxLDL levels. Dots show P -values from GWAS on the $-\log_{10}$ scale (vertical axis), and the chromosomal position (horizontal axis). Pairwise linkage disequilibrium (r^2) from the most significant SNP (rs11902417) is color-coded. The light blue curve shows the local recombination rate based on 1000 Genomes EUR data (release Nov 24, 2014).
(TIF)

S5 Fig. Heatmap of pairwise linkage disequilibrium (LD) statistics for *APOB* gene exonic variants. The colors of the cells refer to r^2 values and show the amount of LD between two markers. LD statistics between a pair of SNPs was calculated using haplotype CEU data from the 1000 Genomes Project.
(TIF)

S6 Fig. P -value of SKAT-O with covariates on synonymous and nonsynonymous variants. The plot illustrates the behavior of SKAT-O and its dependence on the combination coefficient of SKAT and Burden test. If the coefficient is equal to null, SKAT-O behaves as SKAT. If the coefficient is equal to one, SKAT-O behaves as Burden test.
(TIF)

S1 Dataset. Capture design coordinates for targeted sequencing. TargetSeq Custom Enrichment Kit was designed to target the region containing the complete genomic sequence of the *APOB* gene in locus 2p24-p23 (chr2: 20996301-21494945; GRCh37/hg19 reference human genome).
(DOC)

S1 Appendix. Materials and methods for more details.
(PDF)

S1 Table. Correlation of oxLDL levels with biochemical parameters and ultrasound markers of atherosclerosis.
(PDF)

S2 Table. Fourteen the most significant SNPs obtained by GWAS for oxLDL levels in 725 samples. All SNPs are localized on Chromosome 2 in *APOB* gene region.
(PDF)

Acknowledgments

We thank the participants for taking part in the study. The authors acknowledge the contributions of I. Zhanin for the help with targeted sequencing, E. Khrameeva for the help with genotyping and for her advice concerning the use of the PLINK, and S. Smetnev for the technical expertise.

Author Contributions

Conceptualization: Eleonora Khlebus, Vladimir Kutsenko, Alexey Meshkov, Alexandra Ershova, Elena Yarovaya, Oksana Drapkina, Sergey Boytsov.

Data curation: Eleonora Khlebus.

Formal analysis: Eleonora Khlebus, Vladimir Kutsenko, Alexandra Ershova, Anton Shevtsov, Natalia Shcherbakova, Anastasiia Zharikova.

Funding acquisition: Alexey Meshkov, Vadim Lankin, Irina Chazova, Oksana Drapkina, Sergey Boytsov.

Investigation: Alexandra Ershova, Anna Kiseleva, Vadim Lankin, Alla Tikhaze.

Methodology: Eleonora Khlebus, Vladimir Kutsenko, Alexey Meshkov, Alexandra Ershova, Anna Kiseleva, Vadim Lankin, Alla Tikhaze, Elena Yarovaya.

Project administration: Eleonora Khlebus, Alexey Meshkov.

Resources: Alexey Meshkov, Irina Chazova, Oksana Drapkina, Sergey Boytsov.

Software: Eleonora Khlebus, Vladimir Kutsenko, Anton Shevtsov.

Supervision: Alexey Meshkov, Elena Yarovaya.

Validation: Eleonora Khlebus, Anna Kiseleva.

Visualization: Eleonora Khlebus, Vladimir Kutsenko.

Writing – original draft: Eleonora Khlebus, Vladimir Kutsenko, Alexandra Ershova.

Writing – review & editing: Eleonora Khlebus, Vladimir Kutsenko, Alexey Meshkov, Alexandra Ershova, Anna Kiseleva, Vadim Lankin, Alla Tikhaze, Elena Yarovaya.

References

1. Di Pietro N, Formoso G, Pandolfi A. Physiology and pathophysiology of oxLDL uptake by vascular wall cells in atherosclerosis. *Vascul Pharmacol*. 2016; 84:1–7. <https://doi.org/10.1016/j.vph.2016.05.013> PMID: 27256928
2. Linton MF, Yancey PG, Davies SS, Jerome WGJ, Linton EF, Vickers KC. The role of lipids and lipoproteins in atherosclerosis. 2015;.
3. Lankin VZ, Tikhaze AK, Kukharchuk VV, Konovalova GG, Pisarenko OI, Kamynnyi AI, et al. Antioxidants decreases the intensification of low density lipoprotein in vivo peroxidation during therapy with statins. *Mol Cell Biochem*. 2003; 249(1–2):129–140. <https://doi.org/10.1023/A:1024742907379> PMID: 12956408
4. Lankin VZ, Tikhaze AK. Role of oxidative stress in the genesis of atherosclerosis and diabetes mellitus: a personal look back on 50 years of research. *Curr Aging Sci*. 2017; 10(1):18–25. <https://doi.org/10.2174/1874609809666160926142640> PMID: 27677837
5. Lankin VZ, Tikhaze AK, Kumskova EM. Macrophages actively accumulate malonyldialdehyde-modified but not enzymatically oxidized low density lipoprotein. *Mol Cell Biochem*. 2012; 365(1-2):93–98. <https://doi.org/10.1007/s11010-012-1247-5> PMID: 22307746
6. Mäkelä KM, Seppälä I, Hernesniemi JA, Lyytikäinen LP, Oksala N, Kleber ME, et al. Genome-wide association study pinpoints a new functional apolipoprotein B variant influencing oxidized low-density lipoprotein levels but not cardiovascular events: AtheroRemo Consortium. *Circ Cardiovasc Genet*. 2013; 6(1):73–81. <https://doi.org/10.1161/CIRCGENETICS.112.964965> PMID: 23247145
7. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA*. 2012; 109(4):1193–1198. <https://doi.org/10.1073/pnas.1119675109> PMID: 22223662
8. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev*. 2009; 19(3):212–219. <https://doi.org/10.1016/j.gde.2009.04.010> PMID: 19481926

9. Bomba L, Walter K, Soranzo N. The impact of rare and low-frequency genetic variants in common disease. *Genome Biol.* 2017; 18(1):77. <https://doi.org/10.1186/s13059-017-1212-4> PMID: 28449691
10. Lusis AJ. Genetics of atherosclerosis. *Trends Genet.* 2012; 28(6):267–275. <https://doi.org/10.1016/j.tig.2012.03.001> PMID: 22480919
11. Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitzel NO, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet.* 2014; 94(2):223–232. <https://doi.org/10.1016/j.ajhg.2014.01.009> PMID: 24507774
12. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* 2004; 305(5685):869–872. <https://doi.org/10.1126/science.1099870> PMID: 15297675
13. Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science.* 1986; 232(4746):34–47. <https://doi.org/10.1126/science.3513311> PMID: 3513311
14. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet.* 2005; 37(2):161–165. <https://doi.org/10.1038/ng1509> PMID: 15654334
15. Cohen JC, Pertsemlidis A, Fahmi S, Esmail S, Vega GL, Grundy SM, et al. Multiple rare variants in *NPC1L1* associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci USA.* 2006; 103(6):1810–1815. <https://doi.org/10.1073/pnas.0508483103> PMID: 16449388
16. Burnett JR, Hooper AJ. Common and rare gene variants affecting plasma LDL cholesterol. *Clin Biochem Rev.* 2008; 29(1):11–26. PMID: 18566665
17. Rytova AI, Khlebus EY, Shevtsov AE, Kutsenko VA, Shcherbakova NV, Zharikova AA, et al. Modern probabilistic and statistical approaches to search for nucleotide sequence options associated with integrated diseases. *Russ J Genet.* 2017; 53(10):1091–1104. <https://doi.org/10.1134/S1022795417100088>
18. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet.* 2009; 5(2):e1000384. <https://doi.org/10.1371/journal.pgen.1000384> PMID: 19214210
19. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet.* 2011; 89(1):82–93. <https://doi.org/10.1016/j.ajhg.2011.05.029> PMID: 21737059
20. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet.* 2012; 91(2):224–237. <https://doi.org/10.1016/j.ajhg.2012.06.007> PMID: 22863193
21. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet.* 2013; 92(6):841–853. <https://doi.org/10.1016/j.ajhg.2013.04.015> PMID: 23684009
22. Ionita-Laza I, Capanu M, De Rubeis S, McCallum K, Buxbaum JD. Identification of rare causal variants in sequence-based studies: methods and applications to VPS13B, a gene involved in Cohen syndrome and autism. *PLoS Genet.* 2014; 10(12):e1004729. <https://doi.org/10.1371/journal.pgen.1004729> PMID: 25502226
23. Lin WY. Beyond rare-variant association testing: pinpointing rare causal variants in case-control sequencing study. *Sci Rep.* 2016; 6:21824. <https://doi.org/10.1038/srep21824> PMID: 26903168
24. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, et al. Pathway analysis by adaptive combination of P-values. *Genet Epidemiol.* 2009; 33(8):700–709. <https://doi.org/10.1002/gepi.20422> PMID: 19333968
25. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J.* 2003; 24(11):987–1003. [https://doi.org/10.1016/s0195-668x\(03\)00114-3](https://doi.org/10.1016/s0195-668x(03)00114-3) PMID: 12788299
26. Lankin V, Konovalova G, Tikhaze A, Shumaev K, Kumskova E, Viigimaa M. The initiation of free radical peroxidation of low-density lipoproteins by glucose and its metabolite methylglyoxal: a common molecular mechanism of vascular wall injury in atherosclerosis and diabetes. *Mol Cell Biochem.* 2014; 395(1–2):241–252. <https://doi.org/10.1007/s11010-014-2131-2> PMID: 24997046
27. Lankin V, Tikhaze A, Kapel'ko V, Shepel'kova G, Shumaev K, Panasenko O, et al. Mechanisms of oxidative modification of low density lipoproteins under conditions of oxidative and carbonyl stress. *Biochemistry (Moscow).* 2007; 72(10):1081–1090. <https://doi.org/10.1134/S0006297907100069>
28. Khlebus EY, Meshkov AN, Lankin VZ, Orlovsky AA, Kiseleva AV, Shcherbakova NV, et al. Lipid profile and genetic markers associated with the level of oxidized low density lipoproteins. *Russian Journal of Cardiology.* 2017; 10(150):49–54. <https://doi.org/10.15829/1560-4071-2017-10-49-54>

29. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem.* 1972; 18(6):499–502. PMID: [4337382](#)
30. Ershova AI, Balakhonova TV, Meshkov AN, Rozhkova TA, Boytsov SA. Ultrasound markers that describe plaques are more sensitive than mean intima-media thickness in patients with familial hypercholesterolemia. *Ultrasound Med Biol.* 2012; 38(3):417–422. <https://doi.org/10.1016/j.ultrasmedbio.2011.11.014> PMID: [22261515](#)
31. Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Desvarieux M, et al. Mannheim intima-media thickness consensus. *Cerebrovasc Dis.* 2004; 18(4):346–349. <https://doi.org/10.1159/000081812> PMID: [15523176](#)
32. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81(3):559–575. <https://doi.org/10.1086/519795> PMID: [17701901](#)
33. Ning Z, Cox AJ, Mullikin JC. SSAHA: a fast search method for large DNA databases. *Genome Res.* 2001; 11(10):1725–1729. <https://doi.org/10.1101/gr.194201> PMID: [11591649](#)
34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: [19505943](#)
35. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010; 26(6):841–842. <https://doi.org/10.1093/bioinformatics/btq033> PMID: [20110278](#)
36. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.* 1998; 8(3):175–185. <https://doi.org/10.1101/gr.8.3.175> PMID: [9521921](#)
37. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38(16):e164. <https://doi.org/10.1093/nar/gkq603> PMID: [20601685](#)
38. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009; 4(7):1073–1081. <https://doi.org/10.1038/nprot.2009.86> PMID: [19561590](#)
39. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010; 7(4):248–249. <https://doi.org/10.1038/nmeth0410-248> PMID: [20354512](#)
40. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5):1792–1797. <https://doi.org/10.1093/nar/gkh340> PMID: [15034147](#)
41. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 2009; 25(9):1189–1191. <https://doi.org/10.1093/bioinformatics/btp033> PMID: [19151095](#)
42. Lee S, Miropolsky L, Wu M. Skat: Snp-Set (Sequence) Kernel Association Test R package version 1.1.2; 2015.
43. Malovini A, Bellazzi R, Napolitano C, Guffanti G. Multivariate methods for genetic variants selection and risk prediction in cardiovascular diseases. *Front Cardiovasc Med.* 2016; 3:17. <https://doi.org/10.3389/fcvm.2016.00017> PMID: [27376073](#)
44. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Statist Soc B.* 2005; 67(2):301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
45. Friedman J, Hastie T, Tibshirani R. glmnet: Lasso and elastic-net regularized generalized linear models. R package version. 2009; 1(4).
46. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics.* 2010; 26(18):2336–2337. <https://doi.org/10.1093/bioinformatics/btq419> PMID: [20634204](#)
47. Edmondson AC, Braund PS, Stylianou IM, Khera AV, Nelson CP, Wolfe ML, et al. Dense genotyping of candidate gene loci identifies variants associated with high-density lipoprotein cholesterol. *Circ Cardiovasc Genet.* 2011; 4(2):145–155. <https://doi.org/10.1161/CIRCGENETICS.110.957563> PMID: [21303902](#)
48. Reynolds CA, Hong MG, Eriksson UK, Blennow K, Wiklund F, Johansson B, et al. Analysis of lipid pathway genes indicates association of sequence variation near SREBF1/TOM1L2/ATPAF2 with dementia risk. *Hum Mol Genet.* 2010; 19(10):2068–2078. <https://doi.org/10.1093/hmg/ddq079> PMID: [20167577](#)
49. Prus Y, Sergienko I, Malyshev P, Komar O, Popova A, NA S. A rare genetic mutation in patients with heterozygous familial hypercholesterolemia. *Atherosclerosis and Dyslipidaemias.* 2017; 2(27):84–90.

50. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010; 466(7307):707–713. <https://doi.org/10.1038/nature09270> PMID: 20686565
51. Benn M, Stene MC, Nordestgaard BG, Jensen GB, Steffensen R, Tybjaerg-Hansen A. Common and rare alleles in apolipoprotein B contribute to plasma levels of low-density lipoprotein cholesterol in the general population. *J Clin Endocrinol Metab*. 2008; 93(3):1038–1045. <https://doi.org/10.1210/jc.2007-1365> PMID: 18160469
52. Bryant EK, Dressen AS, Bunker CH, Hokanson JE, Hamman RF, Kamboh MI, et al. A multiethnic replication study of plasma lipoprotein levels-associated SNPs identified in recent GWAS. *PLoS One*. 2013; 8(5):e63469. <https://doi.org/10.1371/journal.pone.0063469> PMID: 23717430
53. Tukiainen T, Kettunen J, Kangas AJ, Lyytikäinen LP, Soininen P, Sarin AP, et al. Detailed metabolic and genetic characterization reveals new associations for 30 known lipid loci. *Hum Mol Genet*. 2012; 21(6):1444–1455. <https://doi.org/10.1093/hmg/ddr581> PMID: 22156771
54. Zhou F, Guo T, Zhou L, Zhou Y, Yu D. Variants in the *APOB* gene was associated with Ischemic Stroke susceptibility in Chinese Han male population. *Oncotarget*. 2018; 9(2):2249–2254. <https://doi.org/10.18632/oncotarget.23369> PMID: 29416768
55. Gu QL, Han Y, Lan YM, Li Y, Kou W, Zhou YS, et al. Association between polymorphisms in the *APOB* gene and hyperlipidemia in the Chinese Yugur population. *Braz J Med Biol Res*. 2017; 50(11):e6613. <https://doi.org/10.1590/1414-431X20176613> PMID: 28902930
56. Kim DS, Burt AA, Ranchalis JE, Jarvik ER, Rosenthal EA, Hatsukami TS, et al. Novel gene-by-environment interactions: *APOB* and *NPC1L1* variants affect the relationship between dietary and total plasma cholesterol. *J Lipid Res*. 2013; 54(5):1512–1520. <https://doi.org/10.1194/jlr.P035238> PMID: 23482652
57. Kulminski AM, Culminskaya I, Arbeevev KG, Ukraintseva SV, Stallard E, Arbeeveva L, et al. The role of lipid-related genes, aging-related processes, and environment in healthspan. *Aging Cell*. 2013; 12(2):237–246. <https://doi.org/10.1111/acer.12046> PMID: 23320904
58. Liao YC, Lin HF, Rundek T, Cheng R, Hsi E, Sacco RL, et al. Multiple genetic determinants of plasma lipid levels in Caribbean Hispanics. *Clin Biochem*. 2008; 41(4-5):306–312. <https://doi.org/10.1016/j.clinbiochem.2007.11.011> PMID: 18078817
59. Radovica-Spalvina I, Latkovskis G, Silamikelis I, Fridmanis D, Elbere I, Ventins K, et al. Next-generation-sequencing-based identification of familial hypercholesterolemia-related mutations in subjects with increased LDL-C levels in a latvian population. *BMC Med Genet*. 2015; 16(1):86. <https://doi.org/10.1186/s12881-015-0230-x> PMID: 26415676
60. Huijgen R, Sjouke B, Vis K, de Randamie JS, Defesche JC, Kastelein JJ, et al. Genetic variation in *APOB*, *PCSK9*, and *ANGPTL3* in carriers of pathogenic autosomal dominant hypercholesterolemic mutations with unexpected low LDL-C Levels. *Hum Mutat*. 2012; 33(2):448–455. <https://doi.org/10.1002/humu.21660> PMID: 22095935
61. Starcevic JN, Letonja MS, Praznikar ZJ, Makuc J, Vujkovic AC, Petrovic D. Polymorphisms XbaI (rs693) and EcoRI (rs1042031) of the ApoB gene are associated with carotid plaques but not with carotid intima-media thickness in patients with diabetes mellitus type 2. *Vasa*. 2014; 43(3):171–180. <https://doi.org/10.1024/0301-1526/a000346>
62. Ríos-González BE, Ibarra-Cortés B, Ramírez-López G, Sánchez-Corona J, Magaña-Torres MT. Association of polymorphisms of genes involved in lipid metabolism with blood pressure and lipid values in mexican hypertensive individuals. *Dis Markers*. 2014; 2014:150358. <https://doi.org/10.1155/2014/150358> PMID: 25587205
63. Al-Bustan SA, Alnaqeeb MA, Annice BG, Ebrahim GA, Refai TM. Genetic association of *APOB* polymorphisms with variation in serum lipid profile among the Kuwait population. *Lipids Health Dis*. 2014; 13(1):157. <https://doi.org/10.1186/1476-511X-13-157> PMID: 25292352
64. Chiou KR, Charng MJ. Common mutations of familial hypercholesterolemia patients in Taiwan: characteristics and implications of migrations from southeast China. *Gene*. 2012; 498(1):100–106. <https://doi.org/10.1016/j.gene.2012.01.092> PMID: 22353362
65. Muiya P, Wakil S, Al-Najai M, Meyer BF, Al-Mohanna F, Alshahid M, et al. Identification of loci conferring risk for premature CAD and heterozygous familial hyperlipidemia in the *LDLR*, *APOB* and *PCSK9* genes. *Int J Diabetes Mellit*. 2009; 1(1):16–21. <https://doi.org/10.1016/j.ijdm.2009.05.003>
66. Gu W, Zhang M, Wen S. Association between the *APOB* XbaI and EcoRI polymorphisms and lipids in Chinese: a meta-analysis. *Lipids Health Dis*. 2015; 14(1):123. <https://doi.org/10.1186/s12944-015-0125-z> PMID: 26446158
67. Liu X, Wang Y, Qu H, Hou M, Cao W, Ma Z, et al. Associations of polymorphisms of rs693 and rs1042031 in apolipoprotein B gene with risk of breast cancer in Chinese. *Jpn J Clin Oncol*. 2013; 43(4):362–368. <https://doi.org/10.1093/jjco/hyt018> PMID: 23444115

68. Au A, Griffiths LR, Irene L, Kooi CW, Wei LK. The impact of APOA5, APOB, APOC3 and ABCA1 gene polymorphisms on ischemic stroke: Evidence from a meta-analysis. *Atherosclerosis*. 2017; 265:60–70. <https://doi.org/10.1016/j.atherosclerosis.2017.08.003> PMID: 28865324
69. Cash TP, Pita G, Domínguez O, Alonso MR, Moreno LT, Borrás C, et al. Exome sequencing of three cases of familial exceptional longevity. *Aging Cell*. 2014; 13(6):1087–1090. <https://doi.org/10.1111/ace.12261> PMID: 25116423
70. Ung C, Sanchez AV, Shen L, Davoudi S, Ahmadi T, Navarro-Gomez D, et al. Whole exome sequencing identification of novel candidate genes in patients with proliferative diabetic retinopathy. *Vision Res*. 2017; 139:168–176. <https://doi.org/10.1016/j.visres.2017.03.007> PMID: 28431867
71. Xiao D, Huang K, Chen Q, Huang B, Liu W, Peng Y, et al. Four Apolipoprotein B gene polymorphisms and the risk for coronary artery disease: a meta-analysis of 47 studies. *Genes & Genomics*. 2015; 37(7):621–632. <https://doi.org/10.1007/s13258-015-0292-3>
72. Gaffney D, Hoffs MS, Cameron IM, Stewart G, O'Reilly DSJ, Packard CJ. Influence of polymorphism Q3405E and mutation A3371V in the apolipoprotein B gene on LDL receptor binding. *Atherosclerosis*. 1998; 137(1):167–174. [https://doi.org/10.1016/S0021-9150\(97\)00242-6](https://doi.org/10.1016/S0021-9150(97)00242-6) PMID: 9568749
73. Pullinger CR, Love JA, Liu W, Hennessy LK, Ghassemzadeh M, Newcomb KC, et al. The apolipoprotein B Q3405E polymorphism has no effect on its low-density-lipoprotein receptor binding affinity. *Hum Genet*. 1996; 98(6):678–680. <https://doi.org/10.1007/s004390050283> PMID: 8931699
74. Maurer F, Pradervand S, Guilleret I, Nanchen D, Maghraoui A, Chapatte L, et al. Identification and molecular characterisation of Lausanne Institutional Biobank participants with familial hypercholesterolaemia—a proof-of-concept study. *Swiss Med Wkly*. 2016; 146:w14326. <https://doi.org/10.4414/asmw.2016.14326> PMID: 27497240
75. Liu C, Yang J, Han W, Zhang Q, Shang X, Li X, et al. Polymorphisms in ApoB gene are associated with risk of myocardial infarction and serum ApoB levels in a Chinese population. *Int J Clin Exp Med*. 2015; 8(9):16571–16577. PMID: 26629186
76. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-Three Loci Associated with Plasma Lipoprotein Size, Concentration, and Cholesterol Content in Genome-Wide Analysis. *PLoS Genet*. 2009; 5(11):e1000730. <https://doi.org/10.1371/journal.pgen.1000730> PMID: 19936222
77. Chu AY, Guilianini F, Grallert H, Dupuis J, Ballantyne CM, Barratt BJ, et al. Genome-Wide Association Study Evaluating Lp-PLA2 Mass and Activity at Baseline and Following Rosuvastatin Therapy. *Circ Cardiovasc Genet*. 2012; 5(6):676–685. <https://doi.org/10.1161/CIRCGENETICS.112.963314> PMID: 23118302
78. Sabatti C, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, Jones CG, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet*. 2009; 41(1):35–46. <https://doi.org/10.1038/ng.271> PMID: 19060910
79. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet*. 2009; 41(1):47–55. <https://doi.org/10.1038/ng.269> PMID: 19060911
80. Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008; 40(2):189–197. <https://doi.org/10.1038/ng.75> PMID: 18193044
81. Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. 2007; 316(5829):1331–1336. <https://doi.org/10.1126/science.1142358> PMID: 17463246
82. Kanoni S, Masca NG, Stirrups KE, Varga TV, Warren HR, Scott RA, et al. Analysis with the exome array identifies multiple new independent variants in lipid loci. *Hum Mol Genet*. 2016; 25(18):4094–4106. <https://doi.org/10.1093/hmg/ddw227> PMID: 27466198
83. Tada H, Won HH, Melander O, Yang J, Peloso GM, Kathiresan S. Multiple associated variants increase the heritability explained for plasma lipids and coronary artery disease. *Circ Cardiovasc Genet*. 2014; 7(5):583–587. <https://doi.org/10.1161/CIRCGENETICS.113.000420> PMID: 25170055
84. Dewey FE, Murray MF, Overton JD, Habegger L, Leader JB, Fetterolf SN, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science*. 2016; 354(6319):aaf6814. <https://doi.org/10.1126/science.aaf6814> PMID: 28008009
85. Peddareddygar LR, Grewal RP. Identification of Novel Mutations in Spatacsin and Apolipoprotein B Genes in a Patient with Spastic Paraplegia and Hypobetalipoproteinemia. *Case Rep Genet*. 2015; 2015. <https://doi.org/10.1155/2015/219691> PMID: 26064709
86. Alves AC, Etxebarria A, Soutar AK, Martin C, Bourbon M. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet*. 2014; 23(7):1817–1828. <https://doi.org/10.1093/hmg/ddt573> PMID: 24234650

87. Vandrovцова J, Thomas ER, Atanur SS, Norsworthy PJ, Neuwirth C, Tan Y, et al. The use of next-generation sequencing in clinical diagnosis of familial hypercholesterolemia. *Genet Med*. 2013; 15(12):948–957. <https://doi.org/10.1038/gim.2013.55> PMID: 23680767
88. Yoshida T, Kato K, Yokoi K, Watanabe S, Metoki N, Satoh K, et al. Association of candidate gene polymorphisms with chronic kidney disease in Japanese individuals with hypertension. *Hypertens Res*. 2009; 32(5):411–418. <https://doi.org/10.1038/hr.2009.22> PMID: 19282863
89. Junyent M, Tucker KL, Shen J, Lee YC, Smith CE, Mattei J, et al. A composite scoring of genotypes discriminates coronary heart disease risk beyond conventional risk factors in the Boston Puerto Rican Health Study. *Nutr Metab Cardiovasc Dis*. 2010; 20(3):157–164. <https://doi.org/10.1016/j.numecd.2009.03.016> PMID: 19501493
90. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013; 45(11):1274–1283. <https://doi.org/10.1038/ng.2797> PMID: 24097068
91. Kotowski IK, Pertsemlidis A, Luke A, Cooper RS, Vega GL, Cohen JC, et al. A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. *Am J Hum Genet*. 2006; 78(3):410–422. <https://doi.org/10.1086/500615> PMID: 16465619
92. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Reviews Genetics*. 2002; 3(4):285–298. <https://doi.org/10.1038/nrg775> PMID: 11967553
93. Chamary JV, Hurst LD. Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome biology*. 2005; 6(9):R75. <https://doi.org/10.1186/gb-2005-6-9-r75> PMID: 16168082
94. Fernandez-Higuero JA, Etxebarria A, Benito-Vicente A, Alves AC, Arrondo JL, Ostolaza H, et al. Structural analysis of *APOB* variants, p.(Arg3527Gln), p.(Arg1164Thr) and p.(Gln4494del), causing Familial Hypercholesterolaemia provides novel insights into variant pathogenicity. *Sci Rep*. 2015; 5:18184. <https://doi.org/10.1038/srep18184> PMID: 26643808