Scientific Reports

https://doi.org/10.1038/s41598-026-40367-z

Article in Press

Whole exome sequencing and 12-SNP LDL polygenic score in South Indian patients with familial hypercholesterolemia

Received: 4 August 2024

Accepted: 12 February 2026

Published online: 28 February 2026

Cite this article as: Abraham N., P P.V., Menon U. *et al.* Whole exome sequencing and 12-SNP LDL polygenic score in South Indian patients with familial hypercholesterolemia. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-40367-z

Nithya Abraham, Praveen V P, Usha Menon, Nisha Bhavani, Vasantha Nair, Marta Futema, Renjitha Bhaskaran, Ramesh Menon, Sarita Sekhar, Sajitha Krishnan, Harish Kumar & Devaki R. Nair

WHOLE EXOME SEQUENCING AND 12-SNP LDL POLYGENIC SCORE IN SOUTH INDIAN PATIENTS WITH FAMILIAL HYPERCHOLESTEROLEMIA

Nithya Abraham*[1], Praveen V P [1], Usha Menon [1], Nisha Bhavani [1],Vasantha Nair[1], Marta Futema[2], Renjitha Bhaskaran[3], Ramesh Menon [4], Sarita Sekhar[5], Sajitha Krishnan[6], Harish Kumar[1], Devaki R Nair[7]

1. Dept of Endocrinology and Diabetes, Amrita School of Medicine, Kochi, Amrita Vishwa Vidyapeetham, Kerala, India

2. Cardiovascular and Genomics Research Institute, School of Health & Medical Sciences, City St George's, University of London, London, UK

3. Dept of Statistics, Amrita School of Medicine, Kochi, Amrita Vishwa Vidyapeetham, Kerala, India

4. Medgenome Labs, Bangalore, India

5. Dept of Cardiology, Amrita School of Medicine, Kochi, Amrita Vishwa Vidyapeetham, Kerala, India

6. Dept of Biochemistry, Amrita School of Medicine, Kochi, Amrita Vishwa Vidyapeetham, Kerala, India

7. University college of London and Royal Free London NHS Foundation Trust, London

***Corresponding author**

Dr. Nithya Abraham
Assistant Professor
Dept of Endocrinology
Amrita School of Medicine, Kochi
Amrita Vishwa Vidyapeetham
Kerala, India
drnithya22@gmail.com, nithya21039@aims.amrita.edu, +91- 8136888325

ABSTRACT

Heterozygous familial hypercholesterolemia (FH), a monogenic cause for premature coronary artery disease (CAD) is often underdiagnosed. In individuals who meet the FH diagnostic criteria and lack pathogenic variants, polygenic factors are recognized as potential contributors. This study aimed to characterize the spectrum of genetic variants and determine the low-density lipoprotein polygenic risk score (LDL-PRS) among clinically diagnosed FH participants from South India. We recruited 116 unrelated participants with a pretreatment LDL- C concentration $\geq$ 190 mg/dl and a DLCN (Dutch Lipid Clinic Network) score $\geq$ 3. Targeted next-generation sequencing (NGS) of 23 lipid related genes and 12-SNP (Single nucleotide polymorphism) genotyping were performed. NGS identified 39 variants including 13 pathogenic and 26 variants of unknown significance (VUS) some of which were in non-classical genes: *ABCG5, ABCG8, APOE, PPP1R17, SREBF2*. Pathogenic variants were detected in 66.7% of those with definite FH,19.7% in probable FH and 2.7% in possible FH. Overall,66% were variant negative. Among variant negative (FH/V-) participants, 64% demonstrated high LDL-PRS, whereas 70% of variant positive participants also exhibited elevated scores; suggesting a contributory role of polygenic factors across both groups. Additionally, the observation that variant positive

individuals with high LDL-PRS have an increased risk of coronary artery disease (CAD) adds important nuance to risk stratification within genetically confirmed FH patients.

Confirmation of diagnosis by genetic testing is essential for the diagnosis of FH. Although LDL-PRS may offer little benefit in variant negative cases and improve CAD risk prediction in variant positive individuals, large scale studies are essential to validate its clinical utility and assess whether inclusion of additional LDL- raising SNPs could enhance the detection of polygenic FH in the Indian population.

Keywords

Familial Hypercholesterolemia

Monogenic

Variant

Polygenic

LDL- PRS

India

1. INTRODUCTION

Cardiovascular disease (CVD) is one of the leading causes of mortality in India, accounting for 28% of deaths [1]. Indians are more prone to develop coronary artery disease (CAD) at an earlier age with lower LDL levels compared to individuals in Western countries. The number of deaths attributed to CVD has risen from 2,266,000 in 1990 to 2,669,100 in 2004, representing a significant 17.8% increase in less than two decades [2]. Hypercholesterolemia is a significant risk factor contributing to CVD; however, its genetic underpinnings remain unexplored from the Indian subcontinent.

Heterozygous familial hypercholesterolemia (HeFH), the most prevalent genetic cause of premature CAD, is often unrecognised worldwide. The prevalence of FH among atherosclerotic CVD(ASCVD) is 1:17 (5.95%) than in general population [3]. Majority of HeFH cases (90-95%) result from pathogenic variations in the *LDLR* gene [4], 5% in *APOB* gene and 1% in *PCSK9* gene. Autosomal recessive variants in *LDLRAP1* and *ABCG5/8* genes are rare genetic causes. Those with elevated LDL-C levels and an FH-

associated variant carry a 2 to 3.5-fold increased risk of CAD compared to those without the variant [5,6]. Clinical diagnostic criteria for familial hypercholesterolemia (FH) have been established including the Simon Broome criteria in the UK [7], MEDPED in the USA [8] and the Dutch Lipid Clinic Network (DLCN) criteria in the Netherlands [9].

Limited literature exists regarding the genetic spectrum of FH in India, with a focus mainly on the three most mutated genes [10-14]. Despite the availability of high-throughput next-generation sequencing (NGS) techniques, only 20-30% of possible FH and 60-80% of those with definite FH are identified with an FH-causing variant [15-17]. This indicates that in nearly 60% of clinically diagnosed participants where variants have not been detected, either an unknown variant or a polygenic cause is suspected [18]. In recent years, large-scale genome-wide association studies have contributed to a better understanding of genetic variants influencing LDL-C levels. A high burden of common genetic variants, collectively increasing LDL-C by small fractions, can be summed together in a polygenic risk score (PRS) that mimics the FH phenotype [19]. It is crucial to differentiate

between monogenic and polygenic hypercholesterolemia for patient stratification and management, as monogenic FH confers a higher risk of CAD [20]. Although there are few Indian genetic studies on FH, the utility of LDL- PRS has not been explored among Indians.

This study from South India, hence aimed to study the genetic spectrum using targeted NGS of 23 candidate genes and determine the utility of 12 SNP-LDL-PRS among clinically diagnosed FH participants.

2. METHODS

SUBJECT RECRUITMENT

This observational cross-sectional study, was conducted in the Department of Endocrinology, at Amrita Institute of Medical Sciences, Kerala, India, in accordance with ICMR National Ethical Guidelines for Biomedical and Health Research. Informed consent was obtained from either the participant or their legal guardian. All the aspects of the study were approved by Amrita Institutional Ethics Committee (IRB-AIMS-2020-331). Thorough history-taking and physical examination was performed and the participants were then categorized based on the DLCN score. Those with a DLCN score $\geq$ 3 indicating

(possible/probable/definite) FH, with pretreatment LDL-C concentration ≥190 mg/dl were selected for the study. For patients already on lipid lowering treatment, we calculated an approximate untreated LDL-C by adjusting for the medication potency. Exclusion criteria encompassed individuals with uncontrolled hypothyroidism, chronic liver disease, nephrotic syndrome, hypertriglyceridemia > 500 mg/dL, and those on medications such as steroids or oral contraceptive pills. The enrolled participants underwent a comprehensive assessment that included a detailed family history of premature coronary artery disease (CAD) defined as <55 years in males and <60 years in females, clinical examination, lipid panel measurements, genetic testing (NGS 23 gene panel) and 12- SNP genotyping. The latter two were sent for analysis to Medgenome Labs, Bangalore, India.

NGS METHODOLOGY

Genomic DNA was extracted from whole blood samples using the QIAamp DNA Blood Mini Kit (Qiagen, Germany) and targeted next-generation sequencing (NGS) was performed. Targeted NGS (Next Generation Sequencing)

gene panel comprised of 23 genes*: LDLR, APOB, PCSK9, LDLRAP1, ABCA1, ABCG5, ABCG8, APOA1, APOE, GHR, LPL, SLCO1B1, APOA2, APOC2, CETP, GP1HBP1, SREBF2, APOA5, APOC3, EPHX2, ITIH4, LIPC, PPP1R17.* Selective amplification and sequencing of these genes were performed on Illumina HiseqX (Illumina, CA) platform. The average sequencing depth achieved was≥80-100x.The obtained sequencing reads were aligned to the human reference genome GRCh38.p13 using the BWA program [21,22]. The variant calling was performed using the GATK best practices pipeline with Sentieon (v201808.07). Subsequently, gene annotation of the variants was conducted using the VEP program [23] against the Ensembl release 99 human gene model [24]. Variants were further annotated for allele frequency using population databases such as GnomAD (v3.0), 1000 Genomes, Med Genome population-specific database, EVS, dbSNP (v151), and 1000 Japanese Genome [25,26]. Variants with a minor allele frequency (MAF) exceeding 0.05% in population databases were filtered out. Disease databases including OMIM (Feb 2020), ClinVar (Feb 2020), and HGMD (v2019.4) were also utilized for variant annotation. In-silico predictions were performed using Variant Effect Predictor,

Ensembl release 99, incorporating SIFT (version - 5.2.2), PolyPhen (2.2.2), dbNSFPv4.0 (LRT version - December 5, 2019), and MutationTaster2 (MT2). For analysis and interpretation, non-synonymous substitutions, small deletion-insertions, splice site variants, copy number variants (CNVs) in coding regions of transcripts, and functionally characterized variations in the regulatory regions (promoter and UTR) of genes were considered. CNV were detected using the Exome Depth (v1.1.10) method from targeted sequence data [27]. Bayesian factor > 15 was employed for interpretation, and silent variations that did not result in amino acid changes in the coding region were excluded. The pathogenicity of the prioritized variants was determined following the guidelines and recommendations of the American College of Medical Genetics and Genomics (ACMG) [28]. The variant annotations were sourced from ClinVar repository. The pathogenicity was also assessed using the consensus LDLR variant modifications to the existing ACMG guidelines, as recommended by ClinGen Familial Hypercholesterolemia Variant Curation Expert Panel [29]. Wherever possible, segregation studies using Sanger

sequencing were performed to determine the pathogenicity of the variant.

LDL-PRS estimation

LDL- PRS was calculated by genotyping 12 SNP namely rs4299376, rs1367117, rs6511720, rs629301, rs2479409, rs1564348, rs1800562, rs3757354, rs11220462, rs8017377, rs7412 (*APOE*), and rs429358 (*APOE*), as demonstrated by Talmud et al [18]. The genotyping was performed using the Illumina Infinium Global Screening array-24 v3.0, (Illumina, California, USA). Samples with a minimum genotyping rate of 95% were included in the analysis. Genotype imputation was performed using Beagle version 5.0, utilizing the Genome Asia project reference panel [30,31]. The polygenic LDL score was generated by summing the number of LDL-raising alleles, and precision was enhanced by weighing each SNP according to its effect size. To determine the **weighted LDL-PRS,** the effect size used in our study was based on the beta coefficients reported by the Global Lipid Genetics Consortium (GLGC) for individuals of European ancestry [32]. While the polygenic risk score (PRS) does not provide

diagnostic results with specific cutoff values, we classified participants into groups based on their PRS deciles compared to a self-reported South Asian cohort from the UK Biobank [33]. The validity of this scoring system was evaluated among self-reported South Asian participants in the UK Biobank. For our analysis, an LDL-C PRS falling between decile 6 and 10 (0.88-1.4) was considered high, decile 4 to 6 (0.79-0.88) as intermediate, and decile 1 to 3 (-0.31-0.79) as low. A score falling between decile 6 and 10 (0.88-1.4) was categorized as a high polygenic score based on validation among South Asians by Gratton et al [33].

BIOCHEMICAL ANALYSIS

Lipid estimation was performed using the Roche Cobas C 8000 autoanalyzer (702 module, Germany) utilizing an enzymatic colorimetric method.

STATISTICAL TOOLS

Statistical analysis was performed using IBM SPSS (version 20, IBM Corporation, Armonk, NY, USA). Categorical

variables were presented as frequencies and percentages, while continuous variables were reported as mean, standard deviation, and median (interquartile range) in cases of non-normality. The participants were stratified based on sex, DLCN score, and genetic positivity. To assess the statistical significance of gender in relation to continuous variables, the Mann-Whitney U test was employed. The Kruskal-Wallis test was used to compare DLCN score/genetic positivity with continuous variables, while the Chi-square test was utilized for categorical variables. A p value <0.05 was considered statistically significant.

3 RESULTS

3.1 BASELINE CHARACTERISTICS

Of the 116 individuals identified as having FH, the majority were female. The median age at diagnosis was 53years(7-76years). Family history of premature CAD was reported in 73 participants (62.9%), while 29 (25%) self-reported a history of premature CAD (Table1). The median LDL-C for the entire cohort was 230mg/dl. Among those classified as definite, probable, and possible FH, the median LDL-C were

402.5 mg/dL, 252.5 mg/dL, 225 mg/dL respectively; demonstrating a significant difference (p<0.001) (Table 2). Tendon xanthomas were observed in 5, all of whom were classified as definite FH. Corneal arcus was noted in 20 and xanthelasma in 6 participants. Among those with xanthomas, one harboured VUS (variant of unknown significance). Clinical characteristics of the cohort is presented in Table 2.

3.2 MOLECULAR ANALYSIS

In the current cohort, NGS identified 39 variants, of which one-third were classified as pathogenic(P) or likely pathogenic (LP) (P=12, LP=1), while the remaining were VUS. Additionally, segregation analysis was conducted in 6 families (9 relatives). All identified variants were heterozygous, except for one individual with a homozygous intronic variant in the *LDLR* gene. Notably, two participants harboured double heterozygous VUS variant: *LDLR/LPL* and *LDLR/ABCG8*. All the pathogenic variants were detected in the *LDLR* and *ABCG8* gene, whereas the VUS variants were found in *APOB, PCSK9, APOE, ABCG5, ABCG8, PPP1R17, and SREBF2*. Excluding VUS, participants were classified into variant positive

(FH/V+) and negative (FH/V-) group. Variant positivity was observed in 66.7% of participants with definite FH(4P), 19.4% probable FH (7P), 2.7% possible FH(2P). The frequency of pathogenic variants correlated with increasing LDL-C levels: 0% (155-189 mg/dL), 6% (190-249 mg/dL), 19.2% (250-329 mg/dL) and 75% in LDL-C >330 mg/dL. As presented in Table 3, the variant positive group (FH/V+) presented at younger age, more frequently had xanthomas, and exhibited higher median LDL-C compared to the variant-negative cohort (FH/V-) (257 mg/dL vs. 214.25 mg/dL). Although not statistically significant, a higher prevalence of family history of premature CAD was noted among variant positives participants.

3.2 a PATHOGENIC VARIANTS

Majority of the variants were reported in the *LDLR* gene: one splice site, two nonsense, five missense, three frameshift. Additionally, two start-loss variants were detected in *ABCG8* gene (Table 4). A homozygous intronic variant was identified in the 3’ splice site region of the *LDLR* gene near intron 15(c.2312-2A>C) which is predicted to impact AG acceptor splice site of exon 16. This variant has previously been reported in Croatian population [34].

A novel frameshift variant in *LDLR* was identified in 2 unrelated probands. This heterozygous indel variation in exon 7 of *LDLR*, at position c.996_998delinsG results in premature truncation of the protein downstream of codon 333 by 24 amino acid. Another frameshift variant identified in exon 18 of LDLR at c.2557delG has previously been reported.

Other pathogenic *LDLR* variants identified in this cohort were **p.W4Ter, p.Q154Ter, p.C243R, p.G343S, p.V853SfsTer76, p.S177L and p.V797M** (Table 5) all of which have previously been described in literature. Tendon xanthomas were noted in *LDLR* gene variants, specifically p.S177L, p.V797M, p. Q154Ter, p.C698S and the 3'splice variant in intron 15. Interestingly, a pathogenic variant in exon 1 of *ABCG8* gene was detected in two unrelated probands. No pathogenic variants were identified in either *APOB or PCSK9*gene.

3.2b FH NON-CLASSICAL GENES and VUS

A total of 26 variants across *LDLR, APOB, PCSK9, APOE, ABCG5/8, SREBF2, and PPP1R17* genes were classified as VUS (Table 4). Although segregation analysis was

performed only in four families, most individuals with VUS had a personal or family history of cardiac events, consistent with a phenotype of FH.

Although 7 *LDLR* variants were categorized as VUS, in silico tools predicted it to be deleterious and damaging (Supplementary Table 1). These included p.T108K, p.C313S, p.E353K and p.R566Q. A 5’ splice site variant(c.1586+5G>A) in intron 10 of *LDLR* was identified in three unrelated participants. While its pathogenicity is currently conflicting in ClinVar database, segregation analysis in one family revealed the same variant in two children with FH features, supporting its potential pathogenicity. Similarly, p.C698S in exon 14 of *LDLR*, classified as VUS was reported in three unrelated participants. A different missense variant affecting the same codon (p.C698Y) has been previously reported [35]. Two participants with p.C698S variant exhibited FH phenotypes, with DLCN scores of 11 and 14; one had premature CAD. Cascade screening in the siblings of this proband revealed the same variant, supporting its pathogenicity.

In the *APOB* gene, 5 different missense variants were detected: p.M901R, p.R1224Q, p.L1358V, p.T3567M, p.R3886C. In silico prediction tools classified all five as potentially deleterious (Supplementary Table1). Segregation analysis in the participant with p.R3886C variant revealed the same novel missense variant in exon 26 of *APOB* in the proband's son who had an LDL-C of 180mg/dl.

VUS variants were also identified in *PCSK9* (p.A514T), *APOE* (p.Q291H), *SREBF2*(Sterol Regulatory Element Binding Protein 2) (p.A616S, p.D1118AfsTer9) (Supplementary Table1). Heterozygous variants in the *ABCG8* gene were relatively common. A heterozygous start-loss variant in exon 1 of the *ABCG8* gene (c.2T>C), predicted to be damaging by in silico tools, was detected in three unrelated probands, one of whom had premature CAD. Additionally, two distinct heterozygous variants were detected: one in *ABCG5* (intron 11, p.T94R) and another in p.N409T in *ABCG8*. It is worth noting that standard lipid assays do not differentiate between cholesterol and plant sterols. Therefore, measuring blood sitosterol levels is essential to identify sitosterolemia. However, this could not

be performed due to logistical constraints. Interestingly, a p.E57K variant in *PPP1R17 (protein phosphatase 1 regulatory subunit 17)* gene was detected in this cohort. This proband had an LDL-C of 250 mg/dL with strong family history of premature CAD and DLCN score of 6. Segregation analysis revealed the same variant in the proband's daughter who had an LDL-C of 160 mg/dl further supporting its association with FH. Notably,

3.3 POLYGENIC LDL SCORE

It is well established that LDL-C levels exhibit a polygenic trait, hence LDL-PRS was performed to assess if it could explain the elevated LDL levels observed in FH/V- participants. Due to insufficient DNA, LDL-PRS analysis was performed in 104 of the 116 participants. Among the 104 partcipants,15(15%) had low, 20(19%) had an intermediate and 69(66%) had a high LDL- PRS. Among the 72 FH/V- participants with available PRS data 46(64%) had a high PRS and 16(22%) had an intermediate PRS suggesting a polygenic basis (Supplementary figure1). Notably,70% (7/10) of FH/V+ with PRS data also had a high PRS indicating an additive effect of polygenic factors. Of the 104 participants with PRS data, 46 had high PRS

without a pathogenic variant and 7 had high PRS with a pathogenic variant (Figure 1). Although, the FH/V- had slightly higher mean LDL-PRS compared to FH/V-(0.94 ± 0.13 and 0.92± 0.14), this was not statistically significant unlike in other FH cohorts [18,32]. Similarly, no significant difference in LDL-C was observed between high and low PRS groups (238.12 ± 34.10 vs. 221.2 ± 11.9 mg/dL, p=0.35). However, a significant trend was noted where the proportion of individuals with high LDL-PRS increased with rising LDL-indicating a polygenic component at play (p=0.01) (Supplementary Figure 2). Interestingly, higher prevalence of CAD was observed among the FH/V+ with high LDL- PRS compared to FH/V+ with low LDL-PRS (57% vs 33%, p=0.05).

4. DISCUSSION

With the emergence of NGS, there has been an increase in variant detection rate over the recent years facilitating improved cascade screening and implementation of preventive treatment strategies. In India, over the last few years, efforts to improve FH detection have gained momentum. However, due to cost constraints and limited

resource settings, FH remains undiagnosed in India. This study represents the first in the country to comprehensively study the spectrum of genetic variants and utility of LDL-PRS score among clinically diagnosed FH participants.

In line with FH cohorts worldwide [34-36], the most frequent pathogenic variants were found in the *LDLR* gene (85%). However, the spectrum was heterogeneous compared to previous Indian studies [10-14,35-38]. A novel frameshift variant, K333DfsTer 24, located in exon 7 of *LDLR*, was detected in two unrelated families with a strong family history of premature CAD and a mean LDL of 250 mg/dl. Additionally, a homozygous 3' splice site variant in intron 15, previously reported in a Croatian cohort [34], but identified for the first time in India, was found in a 57-year-old individual presenting with significant xanthomas. Despite a strong family history of premature CAD among the siblings, the index case had not experienced any CVD events.

Variants identified in exon1, 4, 5, 7, 16, 18, and intron 10 of the *LDLR* gene in the current study have been previously reported in the literature [34-40]. Segregation analysis

was performed in six families. A 7-year-old child with tendon xanthomas was found to harbour the p.S177L variant in *LDLR*, which was also identified in his heterozygous father. Both of his siblings showed a similar phenotype with LDL-C levels of 250 mg/dl. This variant has been previously documented in Indian cohorts [37]. Additionally, the p.G343S variant in exon 7 of *LDLR* was detected in a 38-year-old mother with premature CAD. Her two daughters, aged 8 and 10, also carried the same variant. The p.V797M variant in exon 16 of *LDLR* which has been previously reported in India, was frequent in this cohort [37]. One individual with this variant developed premature CAD at 37 years of age, and had a DLCN score of 16. Furthermore, a 24-year-old woman with classic features and elevated LDL-C (405 mg/dL) carried the p.Q154Ter nonsense variant in exon 4 of LDLR. Although this variant has not been reported in India, it has been described in a Polish FH cohort [39].

The two variants, p.C698S in exon 14 of the *LDLR* gene and the 5' splice variant in intron 10 of *LDLR*, have been classified as VUS (Variants of Uncertain Significance) in the ClinVar database, the latter variant been previously

reported from India [37]. As mentioned earlier, the probands carrying these variants exhibited typical FH features, with segregation studies confirming the presence of the same variant among family members. This suggests that the variants in exon 14 and intron 10 of the *LDLR* gene might be reconsidered as likely pathogenic (data uploaded in Clinvar repository). However, functional studies are necessary to validate this.

The second most frequently identified variant was in the *APOB* gene. Consistent with studies from Indian [37,40] and global [41] FH cohorts, the most common *APOB* variant was found in exon 26. Two unrelated individuals each carrying a distinct missense variant in exon 26 of *APOB* (p.T3567M and p.R3886C), presented with premature CAD. A segregation analysis of the proband with p.R3886C variant revealed its presence in his son, who had elevated LDL-C levels (170 mg/dL). However, additional functional studies are needed to prove its pathogenicity. Additionally, *APOB* variants were identified in exons 18, 23 and 25, with predictive algorithms indicating pathogenic effects (Supplementary Table 1). In the *PCSK9* gene, a single variant in exon 10 (p.A514T) was

detected. While this variant has been reported globally [42], this is the first documentation in an Indian cohort. The affected individual had an LDL-C level of 239 mg/dL and a strong family history of CAD.

It is well established that 5–30% of individuals with FH harbor variants beyond the three classical FH-associated genes [43]. Recent studies from India have reported variants in non-classical genes [37,40]. In the current cohort, variants in non-classical genes identified were: *ABCG5, ABCG8, APOE, PPP1R17,* and *SREBF2*.

A heterozygous start-loss variant in exon 1 (p.M1) of *ABCG5*, was identified in three unrelated probands, one of whom had a DLCN score of 6 and premature CAD. Additionally, two other variants were detected in *ABCG5*: (exon 3, intron 11) and *ABCG8* (exon 9). Recent studies suggest that heterozygous carriers of mutant *ABCG5/8* may exacerbate the FH phenotype and act as modifying factors [44-46], increasing cardiovascular risk.

Two variants in *SREBF2 gene* were identified in exons 10 and 19 with one affected individual exhibiting premature CAD and a DLCN score of 6. A recent study from South

India also reported a *SREBF2* variant in FH individuals [40], while Muller et al [47] identified missense variants in exons 6, 7, and 10. These variants impair the *SREBP-2* cleavage, thereby reducing the transcriptional activation of sterol regulatory element-containing genes, including *LDLR*. A novel missense variant in exon 4 of *APOE* (p.Q291H) was identified in a 39-year-old participant with an LDL-C level of 230 mg/dL. Marduel et al. [48] previously demonstrated the segregation of the *Leu167del* variant in exon 4 of *APOE* within a family affected by hypercholesterolemia. This variant is known to impair LDL binding to the LDL receptor (*LDLR*), thereby impairing LDL uptake [49]. Other rare *APOE* variants (p.Glu21Lys, p.Pro102Arg, and p.Leu270Glu) have also been described in the literature [50,51]. Furthermore, two novel variants were identified in the *PPP1R17* gene: a missense variant in exon 3 and a copy number variation (duplication) spanning exons 1–5. This gene encodes a protein phosphatase inhibitor that serves as a substrate for cGMP-dependent protein kinase. Notably, an allele of this gene at promoter SNP (-1323T>C) has been associated with increased risk of hypercholesterolemia [52]. To our knowledge, this is the

first report of *PPP1R17* variants in an FH cohort, warranting study into its role in lipoprotein metabolism.

Only limited studies from India have confirmed molecular diagnosis of FH [10-14,37,40]. Among the earlier reports, Soutar et al (1989) and Rubinsztein et al (1992) described variants in CpG regions(P664L) of *LDLR* gene, marking one of the first FH variants among Indians residing in South Africa. Subsequent studies in the same community revealed additional CpG hotspot variants in exon 4,9and 16 of *LDLR* [12]. A decade later, Ashavaid et al in 2000, reported variants in Exon 3 and 4 in *LDLR* among 25 FH participants [10,11]. Further studies by Arul Jothi et al [13] and Kulkarni et al [14] reported variants in exon4; intron7 and exon3; exon10 respectively. Notably, exon 3 and 4 had emerged as the most frequently reported sites of *LDLR* among Indians, likely due to selective screening of these hotspots.

More recent Indian studies [37,40] have expanded the mutational spectrum, reporting variants involving classical and non-classical genes. In the current cohort, only few *LDLR* variants: p.S177L(exon 4); p.V797M (exon 16); p.C243R(exon 5); 5'splice site variant in intron 10 had

previously been reported among Indians. These observations underscore the genetic heterogeneity of FH in India and highlight the importance of broader genetic screening among clinically diagnosed FH cases to uncover novel variants.

The detection rate and genetic spectrum in FH cohorts vary worldwide, ranging from 2% to 53.7%, largely due to differences in patient selection criteria, ethnic backgrounds and the methodologies used across studies [5,34-36,38]. Additionally, unexplained variations may be influenced by polygenic and epigenetic factors. The modest genetic detection rate (11.2%) in our cohort may be contributed by several factors including inclusion of higher proportion of participants with possible FH, potential exclusion of large copy number variations (CNVs) in the NGS analysis, ethnic variability, or alternative etiologies such as polygenic cause. In comparison, Khera et al reported variant detection rate of 1.7% among individuals with LDL>190 mg/dl [5].

It is widely recognized that [18,19] variant-negative FH often result from polygenic etiology. In our cohort, 64% variant-negative participants had a high polygenic score,

potentially accounting for the elevated LDL-C levels. This supports a polygenic etiology, which still warrants treatment even in the absence of an identifiable variant. Among those with identified variants, 70% had high PRS. In comparison, Trinder et al demonstrated that 28% and 15.3% of FH/V+ had high polygenic scores above 80th and 90th percentile of population reference [53]. Unlike Western cohorts where variant-negative individuals exhibit higher polygenic scores than variant-positives [18,19], our study did not find a statistically significant difference between the two groups. This discrepancy may stem from several factors: 1.the relatively small sample size may have limited the statistical power to detect the differences, 2. ethnic differences could affect LDL-PRS distribution, 3. given that all participants in our cohort had elevated LDL-C, it is possible that both the groups may have accumulated multiple LDL-C raising alleles.

Recent studies have demonstrated that individuals carrying FH variants who possess high PRS are at higher risk of CAD compared to those with FH variant but low PRS [53,54]. Consistent with these studies, our study also showed a significant similar pattern. This suggests that

even among FH variant carriers, a high polygenic burden could help identify those at elevated risk, underscoring the need for comprehensive risk assessment. While a monogenic FH variant confers elevated LDL-C levels and CVD risk, polygenic hypercholesterolemia also contributes to increased risk, albeit to a lesser degree. This underscores the relevance of polygenic score in clinical care. Hence, the coexistence of both these factors appears concerning and merits attention. The cardiovascular risk in FH is heterogenous and several FH risk prediction algorithms such as the FH-Risk-Score [55], SAFEHEART-RE (Spanish Familial Hypercholesterolemia Cohort) [56] have been developed to predict cardiovascular mortality in individuals with FH.

Currently PRS is used in research settings and has not been recommended by guidelines for clinical care. In this context, integrating LDL-PRS into risk stratification models could serve as a valuable tool for risk stratification and early intervention [57]. Nonetheless, larger studies are needed to confirm the predictive value of LDL-PRS and to evaluate its cost-effectiveness in routine clinical practice in the Indian population. Future genome-wide polygenic

scores incorporating hundreds of variants might enhance the predictive power, but their effectiveness depends on the availability of large population- specific datasets particularly for South Asians including Indians.

5. **CONCLUSION**

This study offers valuable insights into the genetic and polygenic architecture of familial hypercholesterolemia in the South Indian population, highlighting the diagnostic utility of NGS and LDL-PRS scoring. It reveals the heterogenous spectrum of variants. Although the overall variant detection rate in our cohort is low, NGS proved effective in identifying variants in non-classical genes. The finding that a significant proportion of clinically diagnosed FH participants lacked pathogenic variants, yet exhibited a high LDL-PRS, underscores the growing relevance of polygenic risk assessment, especially in variant negative cases. Additionally, the observation that FH/V+ individuals with high LDL-PRS are at increased risk of CAD supports the role of LDL-PRS in refining risk stratification among genetically confirmed FH cases. However, larger studies

are essential to validate its utility and to explore whether additional LDL raising SNPs could enhance the detection of polygenic FH in the Indian population.

LIMITATION

The limitations of the study, included the relatively small sample, incomplete screening of relatives particularly those with VUS. In a few cases, the lack of pretreatment LDL-C data could have impacted the accuracy of DLCN scoring. Additionally, the adoption of a polygenic score derived from a South Asian UK group may not fully represent the genetic framework of Indian population, addressing the need to develop a country-specific polygenic score.

DATA AVAILABILITY STATEMENT: The datasets generated during the current study are available in the Clinvar repository. The Clinvar accession numbers are **SCV005205817 to SCV005205823 and SCV005205792 to SCV005205813**

REFERENCES

1. World Health Organization. Cardiovascular diseases. World Health Organization. 2021. Available at: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

2. Yusuf S. et al. “Global burden of cardiovascular diseases. Part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization,” *Circulation*, vol. **104**, no. 22, pp. 2746–2753, 2001.

3. Hu P.et al. Prevalence of Familial Hypercholesterolemia Among the General Population and Patients with Atherosclerotic Cardiovascular Disease: A Systematic Review and Meta-Analysis. *Circulation.* 2020 Jun **2**;141(22):1742-1759.

4. Goldstein, J. K., Hobbs, H. H. & Brown, M. S. Familial hypercholesterolemia. In *The Metabolic and Molecular Bases of Inherited Disease* 8th edn (eds Scriver, C. R. *et al.*) 2863–2913 (McGraw-Hill, 2001).

5. Khera AV.et al. Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. *J Am Coll Cardiol*. 2016 Jun **7**;67(22):2578-89

6. Tada H.et al. Impact of clinical signs and genetic diagnosis of familial hypercholesterolemia on the prevalence of coronary artery disease in patients with severe hypercholesterolemia. *Eur Heart J*. 2017;**38**(20):1573-1579.

7. Risk of fatal coronary heart disease in familial hypercholesterolemia. Scientific steering committee on behalf of the Simon Broome register group*. BMJ*. 1991; **303**:893–6.doi:10.1136/bmj.303.6807.893

8. Williams RR. et al. Diagnosing heterozygous familial hypercholesterolemia using new practical criteria validated by molecular genetics. *The American journal of cardiology*. 1993; **72**:171–6.

9. WHO Human Genetics Programme. (1999). Familial hypercholesterolemia (FH): report of a second WHO consultation,

Geneva, 4 September 1998. World Health Organization. https://iris.who.int/handle/10665/66346

10. Ashavaid,T., Altaf,A., Nair K. Molecular basis of familial hypercholesterolemia: an Indian experience. *Indian J Clin Biochem*. 2000; **15**:11–19

11. Ashavaid,TF., Kondkar AA., Nair KG. Identification of two LDL Receptor mutations causing familial hypercholesterolemia in Indian subjects by a simplified rapid PCR-heteroduplex method. *Clin Chem*. 2000; **46**:1183–1185.

12. Kotze MJ. et al. CpG hotspot mutations at the LDL receptor locus are a frequent cause of familial hypercholesterolemia among South African Indians. *Clin Genet*. 1997; **51**:394–398.

13. ArulJothi KN. et al. Molecular analysis of the LDLR gene in coronary artery disease patients from the Indian population. *Clin Biochem*. 2016; **49**:669–674.

14. Kulkarni SS. et al. Mutation analysis of the LDL receptor gene in Indian families with familial hypercholesterolemia. *Asian J Med Sci*. 2011; **2**:82–86.

15. Taylor A. et al. Mutation detection rate and spectrum in familial hypercholesterolemia patients in the UK pilot cascade project. *Clin Genet* 2010; **77**: 572–80.

16. Humphries SE. et al. Genetic causes of familial hypercholesterolemia in patients in the UK: relation to plasma lipid levels and coronary heart disease risk. *J Med Genet* 2006; **43**: 943–49.

17. Graham CA. et al. Genetic screening protocol for familial hypercholesterolemia which includes splicing defects gives an improved mutation detection rate. *Atherosclerosis* 2005; **182**: 331–40

18. Talmud PJ. et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolemia: a case-control study. *Lancet* 2013; **381**: 1293–301
19. Futema M . et al. Clinical utility of the polygenic LDL-C SNP score in familial hypercholesterolemia*. Atherosclerosis* 2018 Oct; **277**:457-463.
20. Khera AV. et al. Whole genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. *Circulation* 2019; **139**:1593-1602
21. Freed D. et al. The Sentieon Genomics Tools-A fast and accurate solution to variant calling from next-generation sequence data. BioRxiv:115717, 2017.
22. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*2010; **26**(5):589-95
23. McLaren W.et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 2010; **26**(16):2069-70,
24. Zerbino D. R. et al. Ensembl 2018. *Nucleic Acids Res.*, **46**(D1): D754-D761
25. Auton A. et al. 1000 Genomes Project Consortium; A global reference for human genetic variation. *Nature*. 2015 Oct **1**;526(7571):68-74.
26. Karczewski. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020 ;**581**, 434- 443
27. Plagnol V. et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 2012; **28**:2747-2754.
28. Richards S. et al. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical

Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015 May;**17**(5):405-24.

29. Chora J R et al. ClinGen Familial Hypercholesterolemia Expert Panel. The Clinical Genome Resource (ClinGen) Familial Hypercholesterolemia Variant Curation Expert Panel consensus guidelines for LDLR variant classification. *Genet Med.* 2022 Feb;**24**(2):293-306. doi: 10.1016/j.gim.2021.09.012.
30. Browning ,BL., Zhou, Y., and Browning SR. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet* 2018; **103**(3):338-348. Doi: 10.1016/j.ajhg.2018.07.015.
31. GenomeAsia100K Consortium. The Genome Asia 100K Project enables genetic discoveries across Asia. *Nature* 2019; **576**, 106–111.
32. Teslovich TM.et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010. **466**, 707–713
33. Gratton J.et al. LDL-C Concentrations and the 12-SNP LDL-C Score for Polygenic Hypercholesterolaemia in Self-Reported South Asian, Black and Caribbean Participants of the UK Biobank. *Front Genet*.2022;**13**:845498.
34. Pecin I.et al. Mutation detection in Croatian patients with familial hypercholesterolemia. *Ann Hum Genet*. 2013 Jan;**77**(1):22-30
35. Pek SLT. et al. Spectrum of mutations in index patients with familial hypercholesterolemia in Singapore: Single center study. *Atherosclerosis.* 2018 Feb; **269**:106-116
36. Pereira E. et al. Recurrent and novel LDL receptor gene mutations causing heterozygous familial hypercholesterolemia in La Habana. *Hum Genet*. 1995 Sep;**96** (3):319-22.
37. Setia N. et al. Genetic analysis of familial hypercholesterolemia in Asian Indians: A single-center study . *J Clin Lipidol* 2020 Jan-Feb;**14**(1):35-45

38. Wang J.et al. Polygenic Versus Monogenic Causes of Hypercholesterolemia Ascertained Clinically. Arterioscler Thromb Vasc Biol.2016; **36**:2439–2445
39. Chmara M. et al. Molecular characterization of Polish patients with familial hypercholesterolemia: novel and recurrent LDLR mutations.*J Appl Genet.*2010 ;**51**(1):95-106. DOI: 10.1007/bf03195716.
40. Pillai,KKB., Shah,SAV.,Reddy, LL., Ashavaid ,TF., Vishwanathan S. Targeted exome sequencing in South Indian patients with Familial hypercholesterolemia. *Clin Chim Acta*. 2022 Feb **15**; 527:47-55. doi: 10.1016/j.cca.2021.12.022.
41. Alves AC. et al. Further evidence of novel APOB mutations as a cause of familial hypercholesterolaemia. *Atherosclerosis*. 2018 Oct; **277**:448-456. Doi: 10.1016/j.atherosclerosis.2018.06.819.
42. Le QT. et al. Plasma Membrane Tetraspanin CD81 Complexes with Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) and Low-Density Lipoprotein Receptor (LDLR), and Its Levels Are Reduced by PCSK9. *J Biol Chem*. 2015 Sep **18**;290(38):23385-400
43. Wiegman A.et al; European Atherosclerosis Society Consensus Panel. Familial hypercholesterolaemia in children and adolescents: gaining decades of life by optimizing detection and treatment. *Eur Heart J*. 2015 Sep **21**;36(36):2425-37. doi: 10.1093/eurheartj/ehv157
44. Tada H. et al: Rare and Deleterious Mutations in ABCG5/ABCG8 Genes Contribute to Mimicking and Worsening of Familial Hypercholesterolemia Phenotype. *Circ J* 2019, **83**:1917–1924
45. Reynolds TM: Sitosterolaemia: A rare cause of accelerated atherosclerosis. *J Clin Pathol* 2018, **71**:863
46. Helgadottir A. et al: Genetic variability in the absorption of dietary sterols affects the risk of coronary artery disease. *Eur Heart J 2*020, **41**:2618–2628.

47. Muller PY., Miserez AR. Identification of mutations in the gene encoding sterol regulatory element binding protein (SREBP)-2 in hypercholesterolaemia subjects. *Journal of Medical Genetics* .2002;**39:**271-275
48. Marduel M. et al. Description of a Large Family with Autosomal Dominant Hypercholesterolemia Associated with the APOE p. Leu167del Mutation. *Hum Mutat*. 2013 Jan;**34**(1):83-7. doi: 10.1002/humu.22215.
49. Awan Z. et al. APOE p. Leu167del mutation in familial hypercholesterolemia. *Atherosclerosis.* 2013 Dec;**231**(2):218-22
50. Van den Maagdenberg AM. et al. Characterization of five new mutants in the carboxyl-terminal domain of human apolipoprotein E: no cosegregation with severe hyperlipidemia. *Am J Hum Genet*. 1993;**52**(5):937–46
51. Wardell MR. et al. Two apolipoprotein E5 variants illustrate the importance of the position of additional positive charge on receptor-binding activity*. J Lipid Res.* 1991;**32**(3):521–8.
52. Ono S.et al. A promoter SNP (-1323T>C) in G-substrate gene (GSBS) correlates with hypercholesterolemia. *J Hum Genet*. 2003;**48**(9):447-450.
53. Trinder M. et al. Risk of Premature Atherosclerotic Disease in Patients with Monogenic Versus Polygenic Familial Hypercholesterolemia. *J Am Coll Cardiol.* 2019 Jul, **74** (4) 512–522.https://doi.org/10.1016/j.jacc.2019.05.043
54. Paquette M.et al. Polygenic risk score predicts prevalence of cardiovascular disease in patients with familial hypercholesterolemia. *J Clin Lipidol*. 2017 May-Jun**;11**(3):725-732.e5. doi: 10.1016/j.jacl.2017.03.019.
55. Paquette M. et al. Familial Hypercholesterolemia-Risk-Score: A New Score Predicting Cardiovascular Events and Cardiovascular Mortality

in Familial Hypercholesterolemia. *Arterioscler Thromb Vasc Biol*. 2021 Oct;**41**(10):2632-2640. doi: 10.1161/ATVBAHA.121.316106.

56. Gallo A et al. SAFEHEART risk equation and cholesterol-year-score are powerful predictors of cardiovascular events in French patients with familial hypercholesterolemia. *Atherosclerosis,* Volume **306**,41-49.

57. Bosco G., Mszar R., Piro S., Sabouret P., Gallo P. Cardiovascular Risk Estimation and Stratification Among Individuals with Hypercholesterolemia. *Curr Atheroscler Rep* **26**, 537–548 (2024). https://doi.org/10.1007/s11883-024-01225-3

ACKNOWLEDGEMENTS: We thank Amrita Vishwa Vidyapeetham University for the research grant and the staff of Department of Genetics at Amrita School of medicine, Kochi.

AUTHOR CONTRIBUTIONS

Nithya Abraham : Conceptualization, Data curation, Writing- Original draft; **Praveen V P**- Conceptualization, Methodology, Writing- review and editing; **Usha Menon**- Conceptualization; **Nisha Bhavani**- Conceptualization, Review; **Vasantha Nair**- Conceptualization ;**Marta Futema**- Validation, Writing- review and editing; **Renjitha Bhaskaran**- Statistical analysis; **Ramesh Menon**-Genetic testing; **Sarita Sekhar**- Conceptualization; **Sajitha Krishnan** - Methodology,

Supervision ;**Harish Kumar**- Conceptualization; **Devaki R Nair**- Conceptualization, Writing- review and editing.

COMPETING INTERESTS STATEMENT- The authors declare no conflict of interest.

ETHICAL APPROVAL-The study involving human participants were approved by approved by Amrita Institutional Ethics Committee (IRB-AIMS-2020-331) in accordance with ICMR National Ethical Guidelines for Biomedical and Health Research. Informed consent was obtained from either the participant or their legal guardian.

FIGURE LEGENDS

Table 1: **CLINICAL CHARACTERISTICS OF THE COHORT**

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, DLCN: Dutch Lipid Clinic Network Criteria, FH: Familial hypercholesterolemia

To assess the statistical significance of gender in relation to continuous variables, the Mann-Whitney U test was employed, while the Chi-square test was used to compare with categorical variables.

Table 2: CLINICAL CHARACTERISTICS (based on DLCN score)

DLCN scores: Definite FH >8, Probable FH 6-8, Possible FH 3-5

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, V+: variant positive, V -: variant negative, VUS: Variant of unknown significance, DLCN- Dutch Lipid Clinic Network Criteria.

The Kruskal-Wallis test and Chi-square test was used to compare DLCN score with continuous variables and categorical variables respectively.

Table 3: COMPARISON BETWEEN GENETIC GROUPS

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, DLCN: Dutch Lipid Clinic Network, V+: Variant positive, V -: Variant negative, VUS: Variant of unknown significance.

[a] 1 had CNV in the PPR17

The Kruskal-Wallis test and Chi -square test was used to compare genetic positivity groups with continuous variables and categorical variables respectively.

Table 4: CLINICAL AND GENETIC INFORMATION OF INDEX PARTICIPANTS WITH VARIANT POSITIVITY

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDL-PRS: LDL polygenic score

LDLR: Low density lipoprotein receptor, APO B: Apolipoprotein B, PCSK9: Proprotein convertase subtilisin/kexin type 9, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2, CNV copy number variation

Table 5: DETAILS OF PATHOGENIC VARIANTS

LDLR: Low density lipoprotein receptor, APO B: Apolipoprotein B, PCSK9: Proprotein convertase subtilisin/kexin type 9, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2

Figure 1: UPSET PLOT: OVERLAP BETWEEN GENETIC VARIANTS AND LDL-PRS IN COHORT(n=104) ^

VUS: Variant of unknown significance, PRS: polygenic score

ABBREVIATIONS:

HeFH- Heterozygous familial hypercholesterolemia, CAD- Coronary artery disease, *LDLR:* Low density lipoprotein receptor, APOB: Apolipoprotein B, PCSK*9:* Proprotein convertase subtilisin/ kexin type 9, ABCG*5/8*: ATP-binding cassette sub-family G member 5/8, *PPP1R17*: Protein phosphatase 1 regulatory subunit 17, *SREBF2:* Sterol regulatory element-binding protein 2

Table 1: CLINICAL CHARACTERISTICS OF THE COHORT(n=116)

	Total	Male	Female	*p* value
Mean AGE (years)	52.94± 10.10	51.97± 11.34	53.39± 9.32	0.448

SEX (n)	116	46	70	
Median TC (mg/ dL) +	297 (280 ,317)	294.5(275, 311.75)	301.50 (281.50, 321.50)	0.16
Median LDL (mg/ dL) +	230 (217, 252)	229.50(211.50, 250)	230 (219.75, 252. 25)	0.38
Median TG (mg/ dL) +	253 (105,207)	176(125, 215.25)	141(101.5, 202)	0.08
Median HDL (mg/ dL) +	43 (38, 50)	42(37, 46.50)	46(38,52)	0.07
Median VLDL (mg/dl) +	30 (21,41)	35(25.75, 44.25)	28(20,40)	0.04
Premature CAD, n (%)	29(25%)	14(30.4%)	15(21.4%)	0.14
Family history of CAD, n (%)	73(62.9%)	31(67.3%)	42(60%)	0.57
Tendon xanthomas(n)	5	2	3	0.99
Xanthelasma(n)	6	3	3	0.68
DLCN score (%)				
Definite FH, n (%)	6(5.2%)	2(4.4%)	4(5.7%)	0.81
Probable FH, n (%)	36(31%)	14(30.4%)	22 (31.4%)	
Possible FH, n (%)	74(63.8%)	30 (65.2%)	44 (62.9%)	

+TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, DLCN: Dutch Lipid Clinical Network Criteria., FH: Familial hypercholesterolemia

To assess the statistical significance of gender in relation to continuous variables, the Mann-Whitney U test was employed, while the Chi-square test was used to compare with categorical variables.

Table 2: CLINICAL CHARACTERISTICS (based on DLCN score #)

	DEFINITE (>8) (n=6)	PROBABLE (6-8) (n=36)	POSSIBLE (3-5) (n=74)	*p* value
Mean age(years)	38.7± 14.00	52.83± 11.58	53.62± 8.61	0.03
Sex (M: F)	2:4	13:21	31:45	0.60
Median TC $^{+}$ (mg/dL)	516.5(418,571)	316(293.50,336.75)	289(274,310)	<0.001
Median LDL^{+} (mg/dL)	402.5(348.75,438.25)	252.5(229,265.5)	225(212,233)	<0.001
Median TG^{+} (mg/dL)	136.5(101.152.5)	161(111.75,221.75)	154(103,207)	0.38
Median HDL^{+} (mg/dL)	45(41.25,58.5)	42.5(36.5,51.5)	43(38,50)	0.73
Median VLDL^{+} (mg/dL)	27(21.75, 30)	32(24.25,44)	30(20,41)	0.42
Personal history of CAD, n (%)	2(33.3%)	19(55.9%)	8(10.5%)	<0.001
Family history of Hypercholesterolemia	4(66.66%)	23(67.6%)	31(40.8%)	0.009
Family history of CAD, n (%)	6(100%)	28(82.4%)	39(51.3%)	<0.001
Tendon xanthoma, n (%)	5(83.3%)	0	0	0.000
Xanthelasma, n (%)	2(33.3%)	2(5.6%)	2(2.6%)	<0.001
Genetic positivity				

V+, n (%)	4(66.7%)	7(19.4%)	2(2.7%)	<0.001
VUS, n (%)	2(33.3%)	8(22.2%)	16(21.6%)	
V-, n (%)	0	21(58.4%)	56(75.7%)	

+ TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, V+: variant positive, V -: variant negative, VUS: Variant of unknown significance, DLCN- Dutch Lipid Clinic Network Criteria.

#DLCN scores: Definite FH >8, Probable FH 6-8, Possible FH 3-5

The Kruskal-Wallis test and Chi-square test was used to compare DLCN score with continuous variables and categorical variables respectively.

TABLE 3: COMPARISON BETWEEN GENETIC GROUPS

	V+(n =13)	VUS [a] (n=26)	V- (n=77)	*p* value
Mean Age(years)	43.14± 14.81	55.44± 9.65	54.05± 8.18	<0.001
Sex (M: F)	6:7	7:19	33:44	0.18
Premature CAD	46.2% (6)	30.8% (8)	19.5% (15)	0.09
Family history of premature CAD	69.2% (9)	57.7% (15)	63.6% (49)	0.76
Family history of hypercholesterolemia	76.9% (10)	61.5% (16)	41.6% (32)	0.10
Median TC (mg/dL) [+]	321(297,397)	307(284.75,327.5)	292.5(275.25,313.75)	0.016
Median LDL (mg/dL) [+]	257(228.5,325.5)	215.5(230,254)	214.25(226,240.75)	0.005
Median TG (mg/dL) [+]	124(81,183)	141.5(107.25,246)	159.5(106,206.75)	0.169
Median HDL (mg/dL) [+]	41(37.5,46)	45.5(39,51.25)	43(37,50)	0.322

Median VLDL (mg/dL) +	25(16,36)	28(21.75,49)	31.5(21.25,41)	0.214
DLCN				
Definite	30.8% (4)	7.7% (2)	0	<0.001
Probable	53.8% (7)	30.8% (8)	27.3% (21)	
Possible	15.4% (2)	61.5% (16)	72.7% (56)	
Tendon xanthoma	30.8% (4)	3.8% (1)	0	0.000
Xanthelasma	15.4% (2)	15.4% (4)	0	0.000
Mean LDL-PRS (n)	0.92± 0.14 (n= 10)	0.95± 0.03 (n=22)	0.94± 0.13 (n=72)	0.781
LDL- PRS	(n=10)	(n=22)	(n=72)	
LOW	20% (2)	13.6% (3)	13.9% (10)	
INTERMEDIATE	10% (1)	13.6% (3)	22.2% (16)	0.80
HIGH	70% (7)	72.7% (16)	63.9% (46)	

V+- variant positive- variant negative

+TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, DLCN: Dutch Lipid Clinic Network Criteria., V+: variant positive, V -: variant negative, VUS: Variant of unknown significance.

[a]1 had CNV in the PPR17 gene

The Kruskal-Wallis test and Chi – square test was used to compare genetic positivity groups with continuous variables and categorical variables respectively.

TABLE 4

Clinical and genetic information of index participants with variant positivity (P+VUS)

No:	Age (years)/Sex	TC (mg/dl)	LDL (mg/dl)	Xanthomas/CAD	Dyslipidemia in family	DLCN score	DLCN category	Gene affected	Amino acid change	Clinvar: P/LP/ VUS	Segregation Study (genetics for the same variant)	LDL-PRS
1.	38/F	280	240.0	N/Y	Y	6	Probable	***LDLR: EXON7***	p.G343S	P	2 daughters+	0.719
2.	69/M	309	257.0	N/Y	N	7	Probable	***LDLR EXON 5***	p.C243R	P	-	1.012
3	41/M	245	209.0	N/Y	Y	6	Probable	***LDLR EXON16***	p.V797M	P	-	0.664
4	43/M	296	255.0	N/N	Y	6	Probable	***LDLR EXON1***	p.W4Ter	P	Daughter had LDL 250mg/dl, genetics not done	0.904
5	66/F	360	257.0	N/Y	N	7	Probable	***LDLR EXON 18***	p.V853 SfsTer76	P	-	0.902
6	41/F	394	331.0	Y/Y	Y	16	Definite	***LDLR EXON 16***	p.V797M	P	-	0.976
7	24/F	543	450.0	Y/N	Y	16	Definite	***LDLR EXON4***	p.Q154Ter	P	-	0.812

8	57/F	581	402.0	Y/N	Y	16	Definite	***LDLR INTRON 15 Homozygous***	c.2312-2A>C(3'splice sie)	P	-	1.052
9	7/M	300	280.0	Y/N	Y	11	Definite	***LDLR EXON4***	p.S177L	LP	Father,2 brothers +	-
10	47/F	269	230.0	N/N	Y	4	Possible	***LDLR EXON 7 (NOVEL)***	p.K333DfsTer24	P	-	-
11	39/M	353	267.0	N/N	Y	6	Probable	***LDLR EXON 7 (NOVEL)***	p.K333DfsTer24	P	-	1.126
12	54/F	400	320.0	Y/Y	Y	15	Definite	***LDLR EXON14***	p.C698S	VUS	-	0.946
13	33/F	490	403.0	N/N	Y	9	Definite	***LDLR EXON14***	p.C698S	VUS	Sister+	1.004
14	39/M	326	276.0	N/N	Y	7	Probable	***LDLR EXON14***	p.C698S	VUS		0.849
15	59/F	321	242.0	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	Son+	-
16	46/M	298	220.0	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	-	1.22
17	52/F	307	227.0	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	-	1.005
18	61/F	293	240.0	N/Y	Y	7	Probable	***LDLR EXON4***	p.T108K	VUS	-	0.962

19	61/M	269	202.0	N/N	Y	4	Possible	***LDLR EXON7***	p.E353K	VUS	-	-
20	61/F	317	238.0	N/N	Y	4	Possible	***LDLR EXON6***	p.C313S	VUS	-	-
21	58/M	314	226.0	N/N	N	3	Possible	***LDLR EXON13***	p.R566Q	VUS	-	0.949
22	54/M	308	239.0	N/N	Y	4	Possible	***PCSK9***	p.A514T	VUS	-	1.017
23	70/F	269	220.0	N/N	N	3	Possible	***APOB EXON23***	p.R1224Q	VUS	-	0.643
24	64/F	290	206.0	N/Y	Y	6	Probable	***APOB EXON18***	p.M901R	VUS	-	1.079
25	56/M	310	224.0	N/N	Y	4	Possible	***APOB EXON25***	p.L1358V	VUS	-	0.988
26	50/F	361	252.0	N/Y	Y	8	Probable	***APOB EXON26***	p.T3567M	VUS	-	0.583
27	55/F	241	220.0	N/N	Y	5	Possible	***APOB EXON26***	p.R3886C	VUS	Son+	0.886
28	50/F	305	253.0	N/N	Y	6	Probable	***ABCG8 EXON1 (start loss variation)***	p.M1	P	-	-
29	51/F	357	260.0	N/Y	N	7	Probable	***ABCG8 EXON1 (start loss variation)***	p.M1	P	-	0.927
30	41/F	344	278.0	N/N	N	5	Possible	***ABCG8 EXON9***	p.N409T	VUS	-	0.897

31	65/F	311	230.0	N/N	Y	4	Possible	***ABCG5 EXON3***	p.T94R	VUS	-	0.977
32	56/F	284	217.0	N/N	Y	4	Possible	***ABCG5 INTRON 11***	c.1650-2A>T 3' splice site	VUS	Son negative	0.87
33	39/F	298	231.0	N/N	N	3	Possible	***APOE EXON4***	p.Q291H	VUS	-	0.996
34	59/M	322	238.0	N/Y	Y	6	Probable	***SREBF2 EXON19***	c.3353_3359del(p.D1118AfsTer9)	VUS	-	-
35	48/F	285	210.0	N/N	Y	4	Possible	***SREBF2 EXON10***	p.A616S	VUS	-	1.31
36	60/F	295	211.0	N/N	N	3	Possible	***PPP1R17 (CNV)***	chr7: g(31687307_31687214)_(31708456_31707203) dup	VUS	-	1.1
37	41/F	369	298.0	N/N	Y	6	Probable	***PPP1R17 exon 3***	p.E57K	VUS	Daughter+	0.897
38	67/F	280	220.0	N/Y	Y	6	Probable	***LDLR(EXON7)/ABCG8(EXON1)***	LDLR: p.E353K/ ABCG8: p.M1	VUS/P	-	0.748
39	58/F	370	288.0	N/Y	N	7	Probable	***LDLR(EXON10)/LPL(EXON6)***	LDLR: p.V523L/ LPL: p.R333C	VUS	-	-

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDL-PRS: LDL polygenic score

LDLR: Low density lipoprotein receptor, APO B: Apolipoprotein B, PCSK9: Proprotein convertase subtilisin/kexin type 9, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2, CNV copy number variation

TABLE 5

DETAILS OF PATHOGENIC VARIANTS

	Gene	Location/ Variant type	Nucleotide change (Amino acid change)	ClinVar ID	MAF %(GnomAD)	SIFT/LRT	Polyphen 2	Mutation Taster
P1	***LDLR (homozygous)***	Intron 15	c. 2312-2A>C 3' splice site g. 11128006A>C	1789452	NA	-	-	+
P2	***LDLR (P)***	Exon1/ Nonsense	c.11G>A (p.Trp4Ter) g.11089559G>A	250973	0.0006%	-	-	+
P3	***LDLR (LP)***	Exon4/ missense	c.530C>T (p.Ser177Leu) g.11105436C>T	3686	0.003%	+/+	+	+
P4	***LDLR (P)***	Exon 4 nonsense	c.460C>T (p.Gln154Ter) g.11105366C>T	251237	NA	-	-	+
P5	***LDLR (P)***	Exon5/ missense	c.727T>C (p.Cys243Arg) g.11106597T>C	251425	NA	+/+	+	+
P6	***LDLR (P)***	Exon7/ missense	c.1027G>A (p.Gly343Ser) g.11110738G>A)	183106	0.001%	+/+	+	+
P7	***LDLR (P)***	Exon 7/ frameshift	c.996_998delinsG (p. Lys333AspfsTer24) g.11110706: TTAA>TG	NOVEL	NA	-	-	+

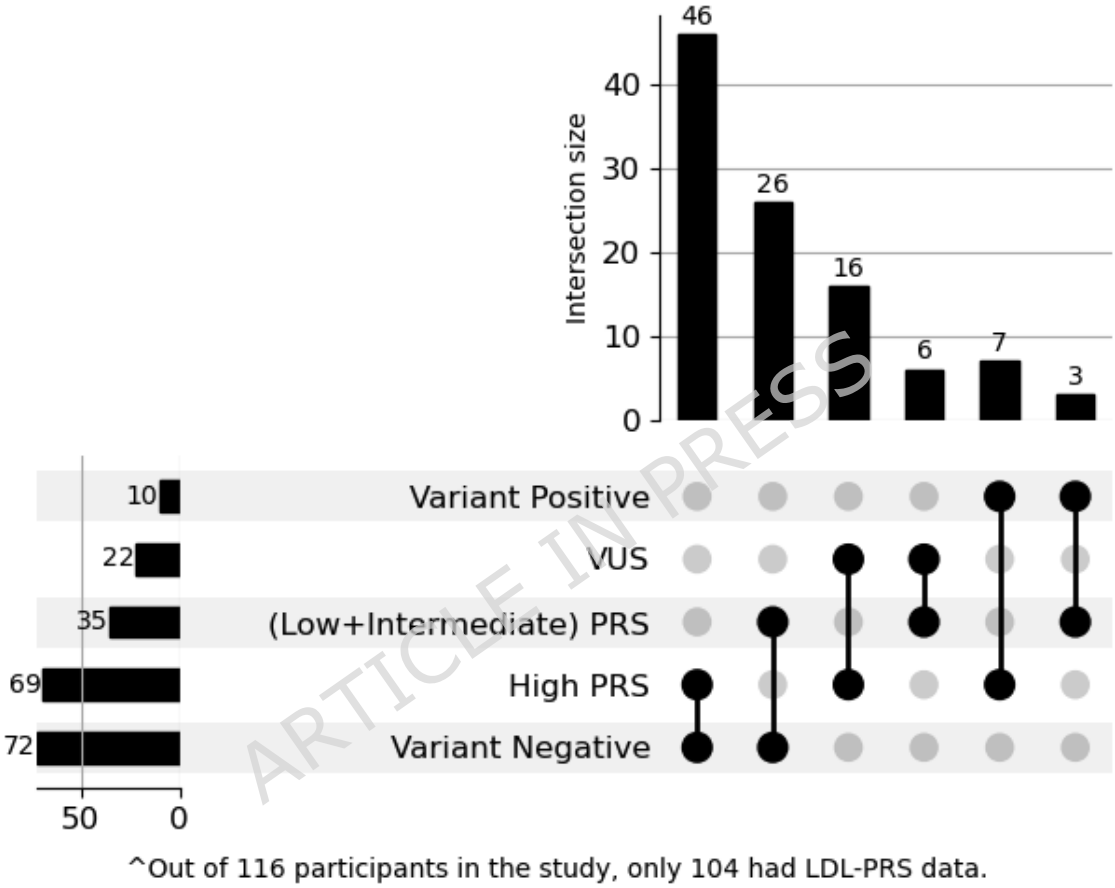
P8	***LDLR (P)***	Exon 7 Frameshift	c.996_998delinsG (p.Lys333AspfsTer24) g.11110707_11110709delinsG	NOVEL	NA	-	-	+
P9	***LDLR (P)***	Exon16/ missense	c.2389G>A (p.Val797Met) g.11128085G>A	226393	0.001%	-	-	+
P10	***LDLR (P)***	Exon16/ missense	c.2389G>A (p.Val797Met) g.11128085G>A)	226393	0.001%	-	-	+
P11	***LDLR (P)***	Exon18/ frameshift	c.2557delG (p.Val853SerfsTer76)	2627788	NA	-	-	+
P12	***ABCG8 (P)***	Exon1/ Start loss variation	c.2T>C (p.Met1) g.43839055T>C	1677567	0.001%	+	-	+
P 13	***ABCG8 (P)***	Exon1/ Start loss variation	c.2T>C (p.Met1) g.43839055T>C	1677567	0.001%	+	-	+

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDLR: Low density lipoprotein receptor ABCG8: ATP-binding cassette sub-family G member 8, NA: not available

MAF: Minor allele frequency

Figure 1: Upset plot: Overlap between genetic variants and LDL-PRS in cohort(n=104) ^



VUS: Variant of Known significance, PRS: polygenic score

Table 1: CLINICAL CHARACTERISTICS OF THE COHORT(n=116)

	Total	Male	Female	*p* value
Mean AGE (years)	52.94± 10.10	51.97± 11.34	53.39± 9.32	0.448
SEX (n)	116	46	70	
Median TC (mg/ dL) +	297 (280 ,317)	294.5(275, 311.75)	301.50 (281.50, 321.50)	0.16
Median LDL (mg/ dL) +	230 (217, 252)	229.50(211.50, 250)	230 (219.75, 252. 25)	0.38
Median TG (mg/ dL) +	253 (105,207)	176(125, 215.25)	141(101.5, 202)	0.08
Median HDL (mg/ dL) +	43 (38, 50)	42(37, 46.50)	46(38,52)	0.07
Median VLDL (mg/dl) +	30 (21,41)	35(25.75, 44.25)	28(20,40)	0.04
Premature CAD, n (%)	29(25%)	14(30.4%)	15(21.4%)	0.14
Family history of CAD, n (%)	73(62.9%)	31(67.3%)	42(60%)	0.57
Tendon xanthomas(n)	5	2	3	0.99
Xanthelasma(n)	6	3	3	0.68
DLCN score (%)				
Definite FH, n (%)	6(5.2%)	2(4.4%)	4(5.7%)	

Probable FH, n (%)	36(31%)	14(30.4%)	22 (31.4%)	0.81
Possible FH, n (%)	74(63.8%)	30 (65.2%)	44 (62.9%)	

+TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, DLCN: Dutch Lipid Clinic Network Criteria., FH: Familial hypercholesterolemia

To assess the statistical significance of gender in relation to continuous variables, the Mann-Whitney U test was employed, while the Chi-square test was used to compare with categorical variables.

Table 2: CLINICAL CHARACTERISTICS (based on DLCN score #)

	DEFINITE (>8) (n=6)	PROBABLE (6-8) (n=36)	POSSIBLE (3-5) (n=74)	*p* value
Mean age(years)	38.7± 14.00	52.83± 11.58	53.62± 8.61	0.03
Sex (M: F)	2:4	13:21	31:45	0.60
Median TC + (mg/dL)	516.5(418,571)	316(293.50,336.75)	289(274,310)	<0.001
Median LDL+ (mg/dL)	402.5(348.75,438.25)	252.5(229,265.5)	225(212,233)	<0.001
Median TG+ (mg/dL)	136.5(101.152.5)	161(111.75,221.75)	154(103,207)	0.38
Median HDL+ (mg/dL)	45(41.25,58.5)	42.5(36.5,51.5)	43(38,50)	0.73
Median VLDL+ (mg/dL)	27(21.75, 30)	32(24.25,44)	30(20,41)	0.42
Personal history of CAD, n (%)	2(33.3%)	19(55.9%)	8(10.5%)	<0.001
Family history of Hypercholesterolemia	4(66.66%)	23(67.6%)	31(40.8%)	0.009

Family history of CAD, n (%)	6(100%)	28(82.4%)	39(51.3%)	<0.001
Tendon xanthoma, n (%)	5(83.3%)	0	0	0.000
Xanthelasma, n (%)	2(33.3%)	2(5.6%)	2(2.6%)	<0.001
Genetic positivity				
V+, n (%)	4(66.7%)	7(19.4%)	2(2.7%)	<0.001
VUS, n (%)	2(33.3%)	8(22.2%)	16(21.6%)	
V-, n (%)	0	21(58.4%)	56(75.7%)	

+ TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, V+: variant positive, V -: variant negative, VUS: Variant of unknown significance, DLCN Dutch Lipid Clinic Network Criteria.

#DLCN scores: Definite FH >8, Probable FH 6-8, Possible FH 3-5

The Kruskal-Wallis test and Chi-square test was used to compare DLCN score with continuous variables and categorical variables respectively.

TABLE 3: COMPARISON BETWEEN GENETIC GROUPS

	V+(n =13)	VUS [a](n=26)	V- (n=77)	*p* value
Mean Age(years)	43.14± 14.81	55.44± 9.65	54.05± 8.18	<0.001
Sex (M: F)	6:7	7:19	33:44	0.18
Premature CAD	46.2% (6)	30.8% (8)	19.5% (15)	0.09
Family history of premature CAD	69.2% (9)	57.7% (15)	63.6% (49)	0.76
Family history of hypercholesterolemia	76.9% (10)	61.5% (16)	41.6% (32)	0.10
Median TC (mg/dL) [+]	321(297,397)	307(284.75,327.5)	292.5(275.25,313.75)	0.016

Median LDL (mg/dL) +	257(228.5,325.5)	215.5(230,254)	214.25(226,240.75)	0.005
Median TG (mg/dL) +	124(81,183)	141.5(107.25,246)	159.5(106,206.75)	0.169
Median HDL (mg/dL) +	41(37.5,46)	45.5(39,51.25)	43(37,50)	0.322
Median VLDL (mg/dL) +	25(16,36)	28(21.75,49)	31.5(21.25,41)	0.214
DLCN				
Definite	30.8% (4)	7.7% (2)	0	<0.001
Probable	53.8% (7)	30.8% (8)	27.3% (21)	
Possible	15.4% (2)	61.5% (16)	72.7% (56)	
Tendon xanthoma	30.8% (4)	3.8% (1)	0	0.000
Xanthelasma	15.4% (2)	15.4% (4)	0	0.000
Mean LDL-PRS (n)	0.92± 0.14 (n= 10)	0.95± 0.03 (n=22)	0.94± 0.13 (n=72)	0.781
LDL- PRS	(n=10)	(n=22)	(n=72)	
LOW	20% (2)	13.6% (3)	13.9% (10)	
INTERMEDIATE	10% (1)	13.6% (3)	22.2% (16)	0.80
HIGH	70% (7)	72.7% (16)	63.9% (46)	

V+: variant positive, V-: variant negative

+TC, HDL- C, LDL – C: to convert mg/ dL to mmol/ L multiply by 0.02586 and TG: to convert mg/ dl to mmol / L multiply by 0.01129

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, DLCN: Dutch Lipid Clinic Network Criteria., V+: variant positive, V -: variant negative, VUS: Variant of unknown significance.

[a]1 had CNV in the PPR17 gene

The Kruskal-Wallis test and Chi – square test was used to compare genetic positivity groups with continuous variables and categorical variables respectively.

TABLE 4

Clinical and genetic information of index participants with variant positivity (P+VUS)

No:	Age (years)/Sex	TC (mg/dl)	LDL (mg/dl)	Xanthomas/CAD	Dyslipidemia in family	DLCN score	DLCN category	Gene affected	Amino acid change	Clinvar: P/LP/ VUS	Segregation Study (genetics for the same variant)	LDL-PRS
1.	38/F	280	240	N/Y	Y	6	Probable	***LDLR: EXON7***	p.G343S	P	2 daughters+	0.719
2.	69/M	309	257	N/Y	N	7	Probable	***LDLR EXON 5***	p.C243R	P	-	1.012
3	41/M	245	209	N/Y	Y	6	Probable	***LDLR EXON16***	p.V797M	P	-	0.664
4	43/M	296	255	N/N	Y	6	Probable	***LDLR EXON1***	p.W4Ter	P	Daughter had LDL 250mg/dl, genetics not done	0.904
5	66/F	360	257	N/Y	N	7	Probable	***LDLR EXON 18***	p.V853 SfsTer76	P	-	0.902
6	41/F	394	331	Y/Y	Y	16	Definite	***LDLR EXON 16***	p.V797M	P	-	0.976

7	24/F	543	450	Y/N	Y	16	Definite	***LDLR EXON4***	p.Q154Ter	P	-	0.812
8	57/F	581	402	Y/N	Y	16	Definite	***LDLR INTRON 15 Homozygous***	c.2312-2A>C(3'splice sie)	P	-	1.052
9	7/M	300	280	Y/N	Y	11	Definite	***LDLR EXON4***	p.S177L	LP	Father,2 brothers +	-
10	47/F	269	230	N/N	Y	4	Possible	***LDLR EXON 7 (NOVEL)***	p.K333DfsTer24	P	-	-
11	39/M	353	267	N/N	Y	6	Probable	***LDLR EXON 7 (NOVEL)***	p.K333DfsTer24	P	-	1.126
12	54/F	400	320	Y/Y	Y	15	Definite	***LDLR EXON14***	p.C698S	VUS	-	0.946
13	33/F	490	403	N/N	Y	9	Definite	***LDLR EXON14***	p.C698S	VUS	Sister+	1.004
14	39/M	326	276	N/N	Y	7	Probable	***LDLR EXON14***	p.C698S	VUS		0.849
15	59/F	321	242	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	Son+	-
16	46/M	298	220	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	-	1.22
17	52/F	307	227	N/N	Y	4	Possible	***LDLR INTRON 10***	c.1586+5G>A	VUS	-	1.005

18	61/F	293	240	N/Y	Y	7	Probable	***LDLR EXON4***	p.T108K	VUS	-	0.962
19	61/M	269	202	N/N	Y	4	Possible	***LDLR EXON7***	p.E353K	VUS	-	-
20	61/F	317	238	N/N	Y	4	Possible	***LDLR EXON6***	p.C313S	VUS	-	-
21	58/M	314	226	N/N	N	3	Possible	***LDLR EXON13***	p.R566Q	VUS	-	0.949
22	54/M	308	239	N/N	Y	4	Possible	***PCSK9***	p.A514T	VUS	-	1.017
23	70/F	269	220	N/N	N	3	Possible	***APOB EXON23***	p.R1224Q	VUS	-	0.643
24	64/F	290	206	N/Y	Y	6	Probable	***APOB EXON18***	p.M901R	VUS	-	1.079
25	56/M	310	224	N/N	Y	4	Possible	***APOB EXON25***	p.L1358V	VUS	-	0.988
26	50/F	361	252	N/Y	Y	8	Probable	***APOB EXON26***	p.T3567M	VUS	-	0.583
27	55/F	241	220	N/N	Y	5	Possible	***APOB EXON26***	p.R3886C	VUS	Son+	0.886
28	50/F	305	253	N/N	Y	6	Probable	***ABCG8 EXON1 (start loss variation)***	p.M1	P	-	-
29	51/F	357	260	N/Y	N	7	Probable	***ABCG8 EXON1 (start loss variation)***	p.M1	P	-	0.927

30	41/F	344	278	N/N	N	5	Possible	***ABCG8 EXON9***	p.N409T	VUS	-	0.897
31	65/F	311	230	N/N	Y	4	Possible	***ABCG5 EXON3***	p.T94R	VUS	-	0.977
32	56/F	284	217	N/N	Y	4	Possible	***ABCG5 INTRON 11***	c.1650-2A>T 3' splice site	VUS	Son negative	0.87
33	39/F	298	231	N/N	N	3	Possible	***APOE EXON4***	p.Q291H	VUS	-	0.996
34	59/M	322	238	N/Y	Y	6	Probable	***SREBF2 EXON19***	c.3353_3359del(p.D1118AfsTer9)	VUS	-	-
35	48/F	285	210	N/N	Y	4	Possible	***SREBF2 EXON10***	p.A616S	VUS	-	1.31
36	60/F	295	211	N/N	N	3	Possible	***PPP1R17 (CNV)***	chr7: g (31687307_31687214)_(31708456_31707203) dup	VUS	-	1.1
37	41/F	369	298	N/N	Y	6	Probable	***PPP1R17 exon 3***	p.E57K	VUS	Daughter+	0.897
38	67/F	280	220	N/Y	Y	6	Probable	***LDLR(EXON7)/ABCG8(EXON 1)***	*LDLR*: p.E353K/ *ABCG8:* *p*.M1	VUS/P	-	0.748
39	58/F	370	288	N/Y	N	7	Probable	***LDLR(EXON10)/LPL(EXON6)***	*LDLR*: p.V523L/ *LPL*: p.R333C	VUS	-	-

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDL-PRS: LDL polygenic score

LDLR: Low density lipoprotein receptor, APO B: Apolipoprotein B, PCSK9: Proprotein convertase subtilisin/kexin type 9, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2, CNV copy number variation

TABLE 5

DETAILS OF PATHOGENIC VARIANTS

	Gene	Location/ Variant type	Nucleotide change (Amino acid change)	ClinVar ID	MAF %(GnomAD)	SIFT/LRT	Polyphen 2	Mutation Taster
P1	***LDLR (homozygous)***	Intron 15	c. 2312-2A>C (3’ splice site) g. 11128006A>C	1789452	NA	-	-	+
P2	***LDLR (P)***	Exon1/ Nonsense	c.11G>A (p.Trp4Ter) g.11089559G>A	250973	0.0006%	-	-	+
P3	***LDLR (LP)***	Exon4/ missense	c.530C>T (p.Ser177Leu) g.11105436C>T	3686	0.003%	+/+	+	+
P4	***LDLR (P)***	Exon 4 nonsense	C 460C>T (p.Gln154Ter) g.11105366C>T	251237	NA	-	-	+
P5	***LDLR (P)***	Exon5/ missense	c.727T>C (p.Cys243Arg) g.11106597T>C	251425	NA	+/+	+	+
P6	***LDLR (P)***	Exon7/ missense	c.1027G>A (p.Gly343Ser) g.11110738G>A)	183106	0.001%	+/+	+	+
P7	***LDLR (P)***	Exon 7/ frameshift	c.996_998delinsG (p. Lys333AspfsTer24)	NOVEL	NA	-	-	+

			g.11110706: TTAA>TG					
P8	***LDLR (P)***	Exon 7 Frameshift	c.996_998delinsG (p. Lys333AspfsTer24) g.11110707_11110709delinsG	NOVEL	NA	-	-	+
P9	***LDLR (P)***	Exon16/ missense	c.2389G>A (p.Val797Met) g.11128085G>A	226393	0.001%	-	-	+
P10	***LDLR (P)***	Exon16/ missense	c.2389G>A (p.Val797Met) g.11128085G>A)	226393	0.001%	-	-	+
P11	***LDLR (P)***	Exon18/ frameshift	c.2557delG (p. Val853SerfsTer76)	2627788	NA	-	-	+
P12	***ABCG8 (P)***	Exon1/ Start loss variation	c.2T>C (p.Met1) g.43839055T>C	1677567	0.001%	+	-	+
P 13	***ABCG8 (P)***	Exon1/ Start loss variation	c.2T>C (p.Met1) g.43839055T>C	1677567	0.001%	+	-	+

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDLR: Low density lipoprotein receptor ABCG8: ATP-binding cassette sub-family G member 8, NA: not available

MAF: Minor allele frequency

FIGURE LEGENDS

Table 1: CLINICAL CHARACTERISTICS OF THE COHORT

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, DLCN: Dutch Lipid Clinic Network, FH: Familial hypercholesterolemia

To assess the statistical significance of gender in relation to continuous variables, the Mann-Whitney U test was employed, while the Chi-square test was used to compare with categorical variables.

Table 2: CLINICAL CHARACTERISTICS (based on DLCN score)

DLCN scores: Definite FH >8, Probable FH 6-8, Possible FH 3-5

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, V+: variant positive, V -: variant negative, VUS: Variant of unknown significance, DLCN- Dutch Lipid Clinic Network Criteria.

The Kruskal-Wallis test and Chi-square test was used to compare DLCN score with continuous variables and categorical variables respectively.

TABLE 3: COMPARISON BETWEEN GENETIC GROUPS

TC: total cholesterol, LDL: Low density lipoprotein, TG: Triglycerides, HDL: High density lipoprotein, VLDL: Very low-density lipoprotein, CAD: Coronary artery disease, DLCN: Dutch Lipid Clinic Network, V+: Variant positive, V -: Variant negative, VUS: Variant of unknown significance.

[a] 1 had CNV in the PPR17

The Kruskal-Wallis test and Chi -square test was used to compare genetic positivity groups with continuous variables and categorical variables respectively.

TABLE 4: CLINICAL AND GENETIC INFORMATION OF INDEX PARTICIPANTS WITH VARIANT POSITIVITY

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDL-PRS: LDL polygenic score

LDLR: Low density lipoprotein receptor, APO B: Apolipoprotein B, PCSK9: Proprotein convertase subtilisin/kexin type 9, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2, CNV copy number variation

TABLE 5: DETAILS OF PATHOGENIC VARIANTS

P: Pathogenic, LP: Likely pathogenic, VUS: Variant of unknown significance

LDLR: Low density lipoprotein receptor ABCG8: ATP-binding cassette sub-family G member 8, NA: not available

MAF: Minor allele frequency

Figure 1: UPSET PLOT: OVERLAP BETWEEN GENETIC VARIANTS AND LDL-PRS IN COHORT(n=104) ^

VUS: Variant of Known significance, PRS: polygenic score

Supplementary table 1: DETAILS OF VUS

LDLR: Low-density lipoprotein receptor, Apo B: Apolipoprotein B, Apo E: Apolipoprotein E, ABCG5: ATP-binding cassette sub-family G member 5, ABCG8: ATP-binding cassette sub-family G member 8, PPP1R17: Protein phosphatase 1 regulatory subunit 17, SREBF2: Sterol regulatory element-binding protein 2, LPL: Lipoprotein lipase

VUS: Variant of unknown significance, ACMG: American college of Medical Genetics and Genomics, MAF: Minor allele frequency

Supplementary figure 1: DISTRIBUTION OF LDL-PRS AMONG VARIANTS

V+: Variant positive, V -: Variant negative, VUS: Variant of unknown significance, LDL PRS: LDL polygenic risk score.

Supplementary figure 2: CORRELATION BETWEEN LDL-C AND LDL-PRS

PRS: polygenic score

ABBREVIATIONS:

HeFH- Heterozygous familial hypercholesterolemia, CAD- Coronary artery disease, *LDLR:* Low density lipoprotein receptor, APOB: Apolipoprotein B, PCSK*9:* Proprotein convertase subtilisin/ kexin type 9, ABCG*5/8*: ATP-binding cassette sub-family G member 5/8, *PPP1R17*: Protein phosphatase 1 regulatory subunit 17, *SREBF2:* Sterol regulatory element-binding protein 2