Polygenic Versus Monogenic Causes of Hypercholesterolemia Ascertained Clinically

Jian Wang, Jacqueline S. Dron, Matthew R. Ban, John F. Robinson, Adam D. McIntyre, Maher Alazzam, Pei Jun Zhao, Allison A. Dilliott, Henian Cao, Murray W. Huff, David Rhainds, Cécile Low-Kam, Marie-Pierre Dubé, Guillaume Lettre, Jean-Claude Tardif, Robert A. Hegele

Objective—Next-generation sequencing technology is transforming our understanding of heterozygous familial hypercholesterolemia, including revision of prevalence estimates and attribution of polygenic effects. Here, we examined the contributions of monogenic and polygenic factors in patients with severe hypercholesterolemia referred to a specialty clinic.

Approach and Results—We applied targeted next-generation sequencing with custom annotation, coupled with evaluation of large-scale copy number variation and polygenic scores for raised low-density lipoprotein cholesterol in a cohort of 313 individuals with severe hypercholesterolemia, defined as low-density lipoprotein cholesterol >5.0 mmol/L (>194 mg/dL). We found that (1) monogenic familial hypercholesterolemia–causing mutations detected by targeted next-generation sequencing were present in 47.3% of individuals; (2) the percentage of individuals with monogenic mutations increased to 53.7% when copy number variations were included; (3) the percentage further increased to 67.1% when individuals with extreme polygenic scores were included; and (4) the percentage of individuals with an identified genetic component increased from 57.0% to 92.0% as low-density lipoprotein cholesterol level increased from 5.0 to >8.0 mmol/L (194 to >310 mg/dL).

Conclusions—In a clinically ascertained sample with severe hypercholesterolemia, we found that most patients had a discrete genetic basis detected using a comprehensive screening approach that includes targeted next-generation sequencing, an assay for copy number variations, and polygenic trait scores. (***Arterioscler Thromb Vasc Biol*. 2016;36:2439-2445. DOI: 10.1161/ATVBAHA.116.308027.)**

Key Words: cholesterol, LDL ■ hypercholesterolemia type IIA ■ hyperlipoproteinemias ■ mutation ■ sequence analysis, DNA

Familial hypercholesterolemia (FH) is a relatively common disorder, previously thought to have a monogenic basis.[1] The paradigmatic heterozygous form of FH (HeFH) is characterized by lifelong elevations in plasma low-density lipoprotein (LDL) cholesterol, typically >5.0 mmol/L (194 mg/dL), sometimes occurring with characteristic physical signs and frequently with a personal or family history of early cardiovascular disease (CVD).[1] Recent population-based surveys, including screening with DNA sequencing, suggest that HeFH has a prevalence of ≈1 in 217 individuals in Northern Europe.[2] Large-scale whole-exome sequencing efforts indicate that ≈4% of individuals with early coronary heart disease have HeFH resulting from one of several loss-of-function mutations in the *LDLR* gene encoding the LDL receptor.[3] Other large-scale sequencing efforts indicate that within subgroups of individuals with severe hypercholesterolemia, defined as untreated LDL cholesterol >5.0 mmol/L (>194 mg/dL), only ≈2% had a pathogenic mutation in an autosomal dominant FH gene.[4]

Furthermore, high-throughput DNA sequencing has shown that 20% to 40% of individuals with phenotypic HeFH have no mutation in canonical FH genes, such as *LDLR*, *APOB*, or *PCSK9*.[5] A few such individuals have rare mutations in minor genes, such as *APOE*, *ABCG5*, *ABCG8*, *LIPA*, or *STAP1*, underlying a phenotype that resembles FH.[6] Others carry a disproportionately high burden of multiple small-effect common variants (single nucleotide polymorphisms), each of which incrementally raises plasma LDL cholesterol by a fraction of a millimole per liter, but which cumulatively raise LDL cholesterol into the FH range.[5] Other individuals with apparent HeFH have none of the above causes identified, suggesting that mutations inaccessible by exome sequencing, such as intronic variants or copy number variations (CNVs),[7] mutations in as yet undefined genes, gene-by-gene

Received on: June 15, 2016; final version accepted on: October 10, 2016.

From the Robarts Research Institute (J.W., J.S.D., M.R.B., J.F.R., A.D.M., A.A.D., H.C., M.W.H., R.A.H.), Department of Biochemistry (J.S.D., M.A., A.A.D., M.W.H., R.A.H.), and Department of Medicine (P.J.Z., M.W.H., R.A.H.), Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada; Faculté de Médicine, Université de Montréal, Québec, Canada (M.-P.D., G.L., J.-C.T.); and Montréal Heart institute, Québec, Canada (D.R., C.L.-K., M.-P.D., G.L., J.-C.T.).

This manuscript was sent to Anne Tybjaerg-Hansen, Consulting Editor, for review by expert referees, editorial decision, and final disposition.

The online-only Data Supplement is available with this article at http://atvb.ahajournals.org/lookup/suppl/doi:10.1161/ATVBAHA.116.308027/-/DC1.

Correspondence to Robert A. Hegele, MD, Department of Medicine, Robarts Research Institute, The University of Western Ontario, 4288A-1151 Richmond St North, London, Ontario, Canada N6A 5K8. E-mail hegele@robarts.ca

Nonstandard Abbreviations and Acronyms	
CNV	copy number variation
CVD	cardiovascular disease
FH	familial hypercholesterolemia
HeFH	heterozygous familial hypercholesterolemia
LDL	low-density lipoprotein
MLPA	multiplex ligation primer amplification
NGS	next-generation sequencing
wGRS	weighted genetic risk score

interactions, gene-by-environment interactions, non-Mendelian mechanisms (eg, epigenetic imprinting), or purely environmental factors, could explain their phenotype. To the extent that a molecular diagnosis of HeFH is desirable,[1,8] for instance as a condition for third-party reimbursement of novel LDL-lowering therapies,[9,10] molecular screening may need to concurrently assess monogenic and polygenic determinants, as well as CNVs.

Next-generation sequencing (NGS) is a powerful technology that has the potential to address clinically relevant hypotheses related to genetic conditions, such as FH. We have developed an NGS panel for dyslipidemias, including FH, which simultaneously assesses the monogenic and polygenic determinants of severely elevated LDL cholesterol from batches of 24 clinical samples.[5] We assess major genes (*LDLR*, *APOB*, *PCSK9*, and *LDLRAP1*) and minor genes (*APOE*, *ABCG5*, *ABCG8*, *LIPA*, and *STAP1*) underlying monogenic FH, as well as single nucleotide polymorphism, genotypes for LDL cholesterol. We include an adjunctive method, namely multiplex ligation primer amplification (MLPA), to detect large-scale CNVs in the *LDLR* gene.[7] This process allows us to address a specific hypothesis, such as what proportion of patients referred to a specialty clinic with possible or probable FH have a monogenic versus polygenic basis. We applied this procedure to samples from 313 individuals referred with severe hypercholesterolemia, defined as LDL cholesterol >5.0 mmol/L (>194 mg/dL), and showed that 53.7% of individuals had monogenic FH mutations and another 13.4% had an extreme polygenic score for high LDL cholesterol. These proportions markedly exceed those observed in population- or cohort-based studies for patients with elevated LDL cholesterol levels, suggesting enrichment for genetic causes through the medical referral process.

Materials and Methods

Materials and Methods are available in the online-only Data Supplement.

Results

Demographics

Baseline demographic features of the Ontario hypercholesterolemia cohort, overall and subdivided by sex, are shown in Table 1. In total, 91.1% of individuals were self-reported as white, with the remainder being of South Asian, Chinese, African, or unspecified ethnic background. The mean±standard deviation (SD) age was 51.0±15.1 years (range, 18.1–88.8 years), and the mean±SD untreated LDL cholesterol level was 6.78±1.75 mmol/L (262±72 mg/dL), with range 5.01 to 13.3 mmol/L (194–514 mg/dL). According to Dutch Lipid Clinic Network criteria, 65.5% of individuals had definite or probable HeFH. Personal and family (parental) history of early CVD (defined as onset <55 years and <60 years in men and women, respectively) were seen in 17.9% and 46.9% of subjects overall.

Monogenic Variants

Likely or definite causal variants in *LDLR*, *APOB*, and *PCSK9* genes detected in this study are listed in Table I in the online-only Data Supplement. The key findings summarized by mutation type are shown in Table 2. Overall, 148 of 313 (47.3%) individuals had at least 1 likely or definite causal mutation detected by NGS. All these variants were confirmed with Sanger sequencing. A further 20 of 313 (6.4%) had a heterozygous pathogenic CNV detected by MLPA, increasing the proportion of individuals with a mutation to 168 of 313 (53.7%).

Most FH mutation-positive Ontario subjects had a heterozygous *LDLR* gene mutation (141/168 or 83.9%), whereas 15 (8.9%) and 2 (1.2%) of the subjects had mutations in *APOB* and *PCSK9* genes, respectively. No potential causative variants were found in *STAP1*, *APOE*, *ABCG5*, *ABCG8*, *LDLRAP1*, or *LIPA*. Twenty of all FH mutation-positive subjects (11.9%) had a large-scale CNV in the *LDLR* gene. Although no patient had homozygous FH based on our selection criteria, 10 FH mutation-positive individuals (6.0%) were found to have 2 variant alleles; of the 20 variant alleles in this pool, 14, 4, and 2 were in *LDLR*, *APOB*, and *PCSK9* genes, respectively (Table I in the online-only Data Supplement). The mean LDL cholesterol in carriers of 2 mutant alleles was nonsignificantly higher than that in the rest of the study sample (Table 2; Figure 1A). Simple heterozygotes for *APOB* or *PCSK9* mutations had significantly lower mean LDL cholesterol than simple heterozygotes for *LDLR* mutations (Figure 1A).

Among the 168 FH mutation-positive individuals were 105 unique mutations, 90 of which (85.7%) were within the *LDLR* gene. Among all mutations found in this study, 16 were novel, of which 12 were within the *LDLR* gene: 2 splicing, 5 frameshift, 0 nonsense, and 5 missense mutations (Table I in the online-only Data Supplement). When considering only *LDLR* gene mutations, there were no differences in mean LDL cholesterol levels between subgroups of individuals with different mutation types (Figure 1A).

There was a stepwise significant increase in the proportion of individuals with monogenic mutations according to plasma LDL cholesterol stratum (Figure 1B). For individuals with LDL cholesterol 5.00 to 5.99 mmol/L (194–231 mg/dL), 6.00 to 6.99 mmol/L (232–270 mg/dL), 7.00 to 7.99 mmol/L (271–309 mg/dL), and ≥8.00 mmol/L (≥310 mg/dL), respectively, 42.1%, 40.4%, 69.8%, and 88.0% were positive for a rare mutation.

Polygenic Trait Scores

Of 145 FH mutation-negative Ontario individuals, 42 (29.0%) had an extreme LDL weighted genetic risk score (wGRS) >1.96, defined in Table II in the online-only Data

Table 1. Demographics of the Ontario Hypercholesterolemia Sample

	Overall (N=313)	Men (N=130)	Women (N=183)
Age, y	51.0±15.1	49.0±13.4	52.4±16.0
Female, %	58.5	…	…
Body mass index, kg/m^2	27.8±5.8	28.6±5.6	27.2±6.0
Total cholesterol, mmol/L	8.91±1.87	8.57±1.70	9.14±1.95
LDL cholesterol, mmol/L	6.78±1.75	6.56±1.64	6.93±1.81
HDL cholesterol, mmol/L	1.36±0.50	1.26±0.61	1.43±0.40
Triglyceride, mmol/L	1.79±0.89	1.82±0.73	1.76±0.99
Personal history of CVD,* %	17.9	25.0	12.9
Family history of CVD,* %	46.9	50.0	44.7
Definite or probable HeFH (DLCN criteria),* %	65.5	63.3	67.1
LDL polygenic risk score, unweighted (uwGRS)	13.9±1.77	14.0±1.82	13.8±1.74
LDL polygenic risk score, weighted (wGRS)	1.75±0.24	1.75±0.25	1.75±0.23

All values represent the mean±standard deviation. CVD indicates cardiovascular disease (onset <55 y in men; <60 y in women); DLCN, Dutch Lipid Clinics Network; HDL, high-density lipoprotein; HeFH, heterozygous familial hypercholesterolemia; LDL, low-density lipoprotein; uwGRS, unweighted genetic risk score; and wGRS, weighted genetic risk score.

*Based on complete data from 145 individuals.

Supplement. This was significantly higher than the proportion of FH mutation-positive individuals (11.9%) who had such an extreme wGRS (odds ratio, 3.02; 95% confidence interval, 1.61–5.68; P<0.0001). Similarly, 11.8% of 1092 individuals in the 1000 Genomes Project had an extreme LDL wGRS >1.96 (Figure 2A). We examined the distribution of wGRS in various cohorts (Figure 2B) and found no difference between 1000 Genomes Project and Ontario FH mutation-positive individuals (mean±SD scores, 1.66±0.27 and 1.68±0.23, respectively, NS). In contrast, the distribution of wGRS in Ontario FH mutation-negative individuals was markedly shifted to the right, with a mean±SD score of 1.84±0.21. Although the absolute differences in mean scores were modest, as is typical for polygenic effects, the differences in overall distribution of scores were highly significant between FH mutation-negative and both FH mutation-positive (P=4.3×10^{-10}) and 1000 Genomes Project control individuals (P=2.9×10^{-18}).

Table 2. Genetic Diversity in the Ontario Hypercholesterolemia Sample

Mutation Type	Overall (N=313)	LDL Cholesterol, mmol/L	Number of Unique Mutations	Novel to This Study
FH mutation-positive	168	7.36±2.17	105	16
LDLR (single mutations)	141	7.40±1.97	90	12
Splicing	20	7.65±1.78	12	2
Frameshift	14	7.84±2.20	10	5
Copy number variation	20	7.11±1.77	8	0
Nonsense	13	7.81±1.65	9	0
Missense (and small in-frame deletion)	74	7.07±1.61	51	5
APOB (missense)	15	6.51±1.30	2	0
PCSK9 (missense)	2	6.38±0.45	2	0
Two mutations	10	7.96±4.50	11	5*
FH mutation-negative	145	6.23±1.27	…	…
LDL wGRS <1.96	103	6.20±1.22	…	…
LDL wGRS ≥1.96	42	6.28±1.39	…	…

All values represent the mean±standard deviation. *APOB* indicates gene encoding apolipoprotein B; FH, familial hypercholesterolemia; LDL, low-density lipoprotein; *LDLR*, gene encoding LDL receptor; *PSCK9*, gene encoding proprotein convertase subtilisin/kexin type 9; wGRS, weighted genetic risk score; and 2 mutations, individuals with 2 likely or definite FH-causing alleles.

**LDLR*, I623T was counted once in the missense mutation category.

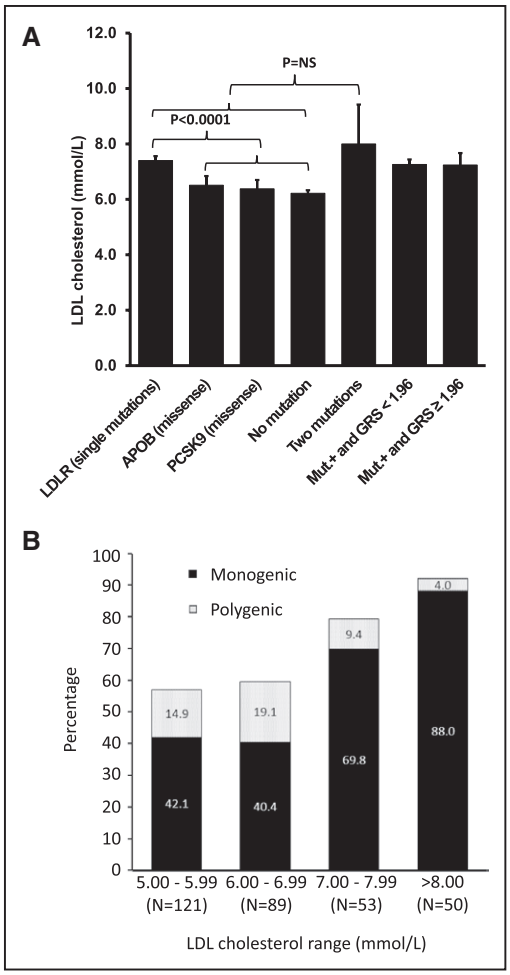


Figure 1. Monogenic and polygenic components of hypercholesterolemia in the Ontario cohort. **A**, Low-density lipoprotein (LDL) cholesterol levels (mean±standard deviations) according to monogenic variant genotype in the Ontario severe hypercholesterolemia sample. Individuals are classified according to presence of 0, 1, or 2 variants (mutations) detected by next-generation sequencing (NGS) and multiplex ligation primer amplification (MLPA) in *LDLR*, *APOB*, or *PSCK9* genes. Familial hypercholesterolemia mutation-positive individuals (Mut.+) are further subgrouped according to extreme weighted genetic risk score (GRS) <1.96 and ≥1.96, respectively. Comparisons of mean LDL cholesterol levels between selected genotype classes are shown, with nominal significant $P<0.05$. **B**, Percentages of individuals with severe hypercholesterolemia within different LDL cholesterol ranges (numbers of individuals shown), classified as having a monogenic variant detected by NGS and MLPA or a polygenic basis defined as an extreme weighted GRS ≥1.96 (≥90th percentile for elevated LDL cholesterol).

We then evaluated the mean LDL cholesterol levels in mutation-negative individuals according to extreme wGRS and found no difference (Table 2). We also found no difference in mean LDL cholesterol levels in FH mutation-positive individuals according to extreme wGRS (Figure 1A).

Combining Monogenic and Polygenic Determinants

Overall, the percentage of Ontario individuals with severe hypercholesterolemia who had an identifiable probable genetic cause detected with NGS was 148 of 313 (47.3%), which increased to 168 of 313 (53.7%) with MLPA results. This increased further to 210 of 313 (67.1%) when individuals with an extreme wGRS were included. Because LDL cholesterol levels increased, so did the percentage of individuals with an identifiable genetic cause. Specifically, for individuals with LDL cholesterol 5.00 to 5.99 mmol/L (194–231 mg/dL), 6.00 to 6.99 mmol/L (232–270 mg/dL), 7.00 to 7.99 mmol/L (271–309 mg/dL), and ≥8.00 mmol/L (≥310 mg/dL), respectively, 57.0%, 59.5%, 79.2%, and 92.0% had a genetic basis for their elevated LDL cholesterol (Figure 1B). Furthermore, polygenic determinants were less common among individuals with the highest LDL cholesterol levels (Figure 1B).

Discussion

High-throughput NGS has transformed our understanding of HeFH. Here, we applied targeted NGS with custom annotation, coupled with MLPA evaluation of large-scale CNV and

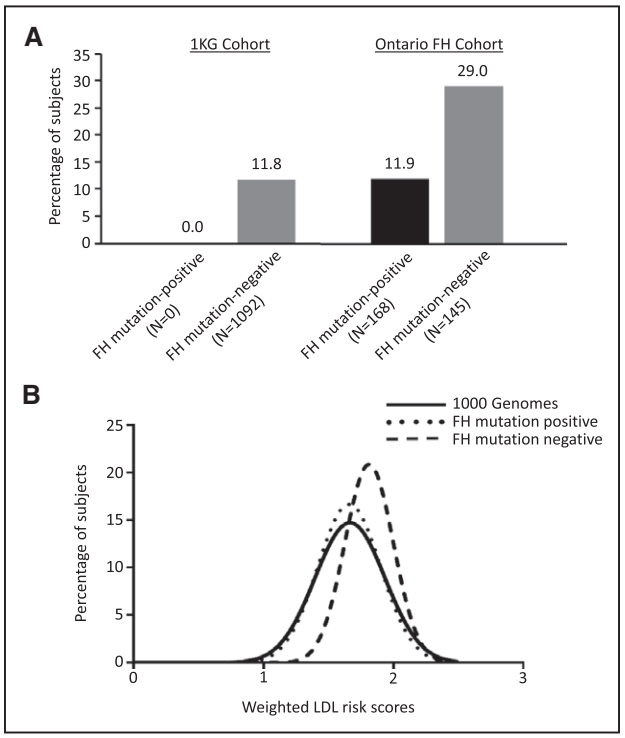


Figure 2. Polygenic component of hypercholesterolemia in the Ontario cohort. **A**, Percentages of individuals from various samples and subgroups with an extreme weighted genetic risk score (wGRS) ≥1.96 (≥90th percentile for elevated low-density lipoprotein [LDL] cholesterol), including the 1000 Genomes (1KG, Release 1; http://www.1000genomes.org/) control cohort and individuals with and without a monogenic cause for familial hypercholesterolemia (FH) detected by next-generation sequencing or multiplex ligation primer amplification in the Ontario cohort. Total numbers of individuals and percentages with extreme wGRSs are shown. FH mutation-negative with severe hypercholesterolemia had significantly higher odds of an extreme wGRS than mutation-positive individuals (odds ratio, 3.02; 95% confidence interval, 1.61–5.68; $P<0.0001$) and normal controls from the 1KG cohort. **B**, Distribution of wGRSs for elevated LDL cholesterol in individuals from the 1KG cohort, and Ontario individuals with severe hypercholesterolemia who are FH mutation-positive or negative. There is no significant difference in the distribution of wGRSs between 1KG and FH mutation-positive individuals, whereas the distribution of wGRSs in FH mutation-negative individuals differs significantly from both these other groups ($P=2.9\times10^{-18}$ and $P=4.3\times10^{-10}$, respectively).

polygenic GRS assessment in a cohort of 313 individuals with severe hypercholesterolemia, in whom FH was the likely clinical diagnosis. We found that (1) monogenic FH-causing mutations detected by targeted NGS were present in 47.3% of individuals; (2) the percentage of individuals with monogenic mutations increased to 53.7% when heterozygous CNVs were included; (3) 83.9% of monogenic mutations were within the *LDLR* gene; (4) the percentage of individuals with a genetic component further increased to 67.1% when individuals with extreme wGRS were included; (5) the percentage of individuals with an identified genetic component increased from 57.0% to 92.0% as LDL cholesterol level increased from 5.0 to >8.0 mmol/L (194 to >310 mg/dL); and (6) individuals with *LDLR* gene mutations had higher mean LDL cholesterol levels than individuals either with *APOB* or *PSCK9* mutations or those with an extreme wGRS.

Within this clinical cohort, NGS alone underestimated the number of individuals with a genetic basis for severe hypercholesterolemia. About half of individuals were FH mutation-positive solely based on NGS results, but this increased to more than two-thirds of individuals when CNVs and extreme polygenic wGRS were considered. We suggest that these additional genetic determinants should be considered in routine molecular assessment of patients with severe hypercholesterolemia. The actual proportion of patients with each type of genetic determinant will vary between cohorts and populations. At present, at least 2 methods are required: our NGS method (see Materials and Methods) and pipeline reports both small monogenic variants and polygenic risk scores, whereas detection of CNVs requires that the independent MLPA method is run in parallel. However, we note that there has been marked progress in improving bioinformatic annotation tools to predict CNVs from NGS data.[11] It is likely that in the near future, a single platform, namely NGS plus bioinformatics, will permit reporting of small-scale sequence variants, large-scale CNVs, and genetic risk scores. MLPA could then be reserved for confirming predicted CNVs from NGS results.

A recent large-scale sequencing study performed in individuals with LDL cholesterol >4.91 mmol/L (>190 mg/dL) ascertained through CVD case–control and population-based samples showed, perhaps surprisingly, prevalence of FH mutation-positive individuals of only ≈2%.[4] This is much lower than the 47.3% of individuals with LDL cholesterol >5 mmol/L observed in our clinically ascertained cohort. Why is there such a wide disparity in the prevalence of FH mutation-positive individuals? The most likely explanation is the difference in ascertainment strategy.

Individuals in our study were all medically identified with severe hypercholesterolemia, resulting in referral for assessment and treatment advice. From the outset, they represented a selected group of individuals, in whom FH was already considered as possible on clinical grounds. In contrast, agnostic screening in broader, nonselected cohorts, when genetic dyslipidemia is not at top-of-mind, could be expected to yield a lower prevalence of FH mutation-positive individuals for the LDL cholesterol level, especially if family history was not considered.[4] Similar disparities of frequencies of FH mutation-positive individuals have been observed previously.[12] For instance, 35% to 65% of patients in other clinically ascertained hypercholesterolemia cohorts of various ethnic backgrounds were FH mutation-positive by DNA sequencing,[13–15] in contrast to only 2.4% of individuals with total cholesterol >7 mmol/L (>271 mg/dL) ascertained agnostically in a population-based study.[16] This disparity was proposed to be related to enrichment for FH mutation-positive individuals through the medical referral process combined with unavailability of other prioritizing clinical data, such as family history in population-based samples.[6] Also, unmeasured types of variation, such as CNVs and polygenic effects, may be present in a high proportion of hypercholesterolemic individuals from population-based samples.

Our findings also indicate that as LDL cholesterol increases, so does the likelihood that the patient is FH mutation-positive. A critical point of inflection seems to be LDL cholesterol >8.0 mmol/L (>310 mg/dL), where almost all patients have a monogenic basis, whereas almost none have a polygenic basis. In contrast, at the lower end of the severe hypercholesterolemia range, that is, LDL cholesterol 5.0 to 6.9 mmol/L, ≈15% to 20% of individuals have an extreme wGRS. Among individuals with a monogenic basis for their severe hypercholesterolemia, we found no differences between genetic subgroups. Interestingly, several patients carried 2 mutant alleles, of which 1 always affected *LDLR*: in 4 cases, the second mutation was also in the *LDLR* gene (ie, compound heterozygotes), whereas in 6 cases the second mutation was in either *APOB* or *PSCK9* genes (ie, double heterozygotes). Such genotypes have been occasionally reported in some cases of phenotypic homozygous FH.[17] Although individuals in this study with this genotype tended to have higher LDL cholesterol levels, the difference was not significant, although this may have reflected the small subgroup size. Alternatively, the mutations found in this subgroup may only have had marginal additive effects on LDL cholesterol levels.

Overall, the odds of FH mutation-negative individuals in this cohort having an extreme wGRS is ≈3-fold greater than both the general population and the FH mutation-positive individuals. This is similar to the ≈2-fold greater odds of an extreme wGRS found FH mutation-negative individuals in a comparable FH cohort studied using an analogous but different risk score.[5] An extreme wGRS thus identifies a subset of patients, although by no means all, who are at increased genetic risk for severe hypercholesterolemia. Interestingly, an extreme wGRS did not seem to modulate the severity of the LDL cholesterol phenotype in either FH mutation-positive or mutation-negative subjects, although again this could simply reflect small subgroups and insufficient statistical power. In our 145 FH-mutation negative patients, there was no correlation between LDL cholesterol level and wGRS (r=0.05; P=NS). Nonetheless, the strong association of wGRS with high LDL cholesterol levels shown by discrete statistical analysis (ie, odds ratio, 3.02; P<0.0001) confirms the validity of this score as a genetic marker of risk of elevated LDL cholesterol. However, demonstrating a continuous positive proportional linear relationship between wGRS and LDL cholesterol levels may require a much larger sample size. Alternatively, it is possible that in FH mutation-negative individuals, those with a low wGRS have another unidentified cause for elevated LDL cholesterol, independent of wGRS. Furthermore, in FH mutation-positive individuals, the influence on phenotype of the monogenic mutation could trump that of the polygenic wGRS.

The role of polygenic risk in severe hypercholesterolemia requires further evaluation to address several questions. For instance, several risk scores for elevated LDL cholesterol have been proposed[18]; which one performs best clinically? Are different scores required for different ethnic groups? For the same degree of LDL cholesterol elevation, is there a difference in prognosis or treatment response based on whether the patient has monogenic or polygenic hypercholesterolemia? For instance, when LDL cholesterol is >5.0 mmol/L (>194 mg/dL) and a monogenic FH mutation is present, there is a 22-fold increased risk of early CVD.[4] However, even when a monogenic cause is absent, CVD risk in such individuals, whose LDL cholesterol is elevated presumably because of polygenic factors, is increased 6-fold,[4] which is not inconsiderable. Given this difference in risk, should more intensive efforts be made to identify FH mutation-positive individuals or is the difference between a 22-fold and 6-fold increase clinically irrelevant? Should intervention be guided predominantly by the degree of LDL cholesterol elevation irrespective of the genetic basis? A mitigating factor is that baseline untreated LDL cholesterol levels may not always be available: in such cases, documenting a patient's genetic architecture with respect to elevated LDL cholesterol may still be important.

Finally, about one-third (32.9%) of individuals in this study with severe hypercholesterolemia had no identifiable genetic cause, although this was reduced to 8.0% in individuals with LDL cholesterol >8.0 mmol/L (>310 mg/dL). It is possible, although unlikely, that as yet undiscovered monogenic causes of hypercholesterolemia could explain the phenotype in some patients.[19,20] Other possible determinants include (1) additional as yet uncharacterized polygenic determinants, some of which may be ethnicity specific; (2) higher order or nonlinear interactions between 2 or more genetic variants that act synergistically to raise LDL cholesterol; (3) the impact of non-Mendelian factors, such as mitochondrial or epigenetic effects; or (4) strong predominant environmental factors or environmental factors that require interactions with either known or unknown genetic determinants to influence the phenotype.

Thus, in a real-world sample, we show that comprehensive screening of patients with severe hypercholesterolemia should include targeted resequencing, preferably using an NGS platform, plus an assay to detect CNVs, plus a set of single nucleotide polymorphisms for a genetic risk score based on the LDL cholesterol-raising alleles. Inclusion of CNVs and a wGRS substantially increases the proportion of individuals with a genetic basis for their clinically ascertained severe hypercholesterolemia, although the actual proportions will vary between cohorts and subpopulations. Stratification by genetic cause can be used in prospective observational or intervention studies, whereas individuals with severe hypercholesterolemia in the absence of a defined genetic cause can be subjects for further investigation to characterize the basis of their phenotype. Finally, because single nucleotide polymorphisms for elevated LDL cholesterol also cluster in families, we think that cascade screening of relatives should proceed whether a proband has elevated LDL cholesterol on either a monogenic or a polygenic basis.

Sources of Funding

Dr Hegele is supported by the Jacob J. Wolfe Distinguished Medical Research Chair, the Edith Schulich Vinet Research Chair in Human Genetics, and the Martha G. Blackburn Chair in Cardiovascular Research and has received operating grants from the Canadian Institutes of Health Research (Foundation Grant), the Heart and Stroke Foundation of Ontario (G-15-0009214), and Genome Canada through Genome Quebec (award 4530).

Disclosures

Dr Hegele is a consultant for Boston Heart Diagnostics.

References

1. Nordestgaard BG, Chapman MJ, Humphries SE, et al; European Atherosclerosis Society Consensus Panel. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *Eur Heart J*. 2013;34:3478–3490a. doi: 10.1093/eurheartj/eht273.
2. Benn M, Watts GF, Tybjærg-Hansen A, Nordestgaard BG. Mutations causative of familial hypercholesterolaemia: screening of 98 098 individuals from the Copenhagen General Population Study estimated a prevalence of 1 in 217. *Eur Heart J*. 2016;37:1384–1394. doi: 10.1093/eurheartj/ehw028.
3. Do R, Stitziel NO, Won HH, et al; NHLBI Exome Sequencing Project. Exome sequencing identifies rare *LDLR* and *APOA5* alleles conferring risk for myocardial infarction. *Nature*. 2015;518:102–106. doi: 10.1038/nature13917.
4. Khera AV, Won HH, Peloso GM, et al. Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. *J Am Coll Cardiol*. 2016;67:2578–2589. doi: 10.1016/j.jacc.2016.03.520.
5. Talmud PJ, Shah S, Whittall R, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet*. 2013;381:1293–1301. doi: 10.1016/S0140-6736(12)62127-8.
6. Hegele RA, Ban MR, Cao H, McIntyre AD, Robinson JF, Wang J. Targeted next-generation sequencing in monogenic dyslipidemias. *Curr Opin Lipidol*. 2015;26:103–113. doi: 10.1097/MOL.0000000000000163.
7. Wang J, Ban MR, Hegele RA. Multiplex ligation-dependent probe amplification of *LDLR* enhances molecular diagnosis of familial hypercholesterolemia. *J Lipid Res*. 2005;46:366–372. doi: 10.1194/jlr.D400030-JLR200.
8. Sniderman AD, Tsimikas S, Fazio S. The severe hypercholesterolemia phenotype: clinical diagnosis, management, and emerging therapies. *J Am Coll Cardiol*. 2014;63:1935–1947. doi: 10.1016/j.jacc.2014.01.060.
9. Hegele RA, Gidding SS, Ginsberg HN, McPherson R, Raal FJ, Rader DJ, Robinson JG, Welty FK. Nonstatin low-density lipoprotein-lowering therapy and cardiovascular risk reduction-statement from ATVB Council. *Arterioscler Thromb Vasc Biol*. 2015;35:2269–2280. doi: 10.1161/ATVBAHA.115.306442.
10. Hegele RA. Improving the monitoring and care of patients with familial hypercholesterolemia. *J Am Coll Cardiol*. 2016;67:1286–1288. doi: 10.1016/j.jacc.2016.01.041.
11. Samarakoon PS, Sorte HS, Stray-Pedersen A, Rødningen OK, Rognes T, Lyle R. cnvScan: a CNV screening and annotation tool to improve the clinical utility of computational CNV prediction from exome sequencing data. *BMC Genomics*. 2016;17:51. doi: 10.1186/s12864-016-2374-2.
12. Tybjaerg-Hansen A, Jensen HK, Benn M, Steffensen R, Jensen G, Nordestgaard BG. Phenotype of heterozygotes for low-density lipoprotein receptor mutations identified in different background populations. *Arterioscler Thromb Vasc Biol*. 2005;25:211–215. doi: 10.1161/01.ATV.0000149380.94984.f0.
13. Futema M, Plagnol V, Whittall RA, Neil HA, Humphries SE; Simon Broome Register Group; UK10K. Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia. *J Med Genet*. 2012;49:644–649. doi: 10.1136/jmedgenet-2012-101189.
14. Ahmad Z, Adams-Huet B, Chen C, Garg A. Low prevalence of mutations in known loci for autosomal dominant hypercholesterolemia in a multiethnic patient cohort. *Circ Cardiovasc Genet*. 2012;5:666–675. doi: 10.1161/CIRCGENETICS.112.963587.
15. Maglio C, Mancina RM, Motta BM, Stef M, Pirazzi C, Palacios L, Askaryar N, Borén J, Wiklund O, Romeo S. Genetic diagnosis of familial hypercholesterolaemia by targeted next-generation sequencing. *J Intern Med*. 2014;276:396–403. doi: 10.1111/joim.12263.

16. Norsworthy PJ, Vandrovcova J, Thomas ER, Campbell A, Kerr SM, Biggs J, Game L, Soutar AK, Smith BH, Dominiczak AF, Porteous DJ, Morris AD, Scotland G, Aitman TJ. Targeted genetic testing for familial hypercholesterolaemia using next generation sequencing: a population-based study. *BMC Med Genet*. 2014;15:70. doi: 10.1186/1471-2350-15-70.
17. Cuchel M, Bruckert E, Ginsberg HN, et al; European Atherosclerosis Society Consensus Panel on Familial Hypercholesterolaemia. Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Atherosclerosis Society. *Eur Heart J*. 2014;35:2146–2157. doi: 10.1093/eurheartj/ehu274.
18. Futema M, Shah S, Cooper JA, et al. Refinement of variant selection for the LDL cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries. *Clin Chem*. 2015;61:231–238. doi: 10.1373/clinchem.2014.231365.
19. Stitziel NO, Peloso GM, Abifadel M, et al. Exome sequencing in suspected monogenic dyslipidemias. *Circ Cardiovasc Genet*. 2015;8:343–350. doi: 10.1161/CIRCGENETICS.114.000776.
20. Futema M, Plagnol V, Li K, et al; Simon Broome Consortium; UK10K Consortium. Whole exome sequencing of familial hypercholesterolaemia patients negative for *LDLR/APOB/PCSK9* mutations. *J Med Genet*. 2014;51:537–544. doi: 10.1136/jmedgenet-2014-102405.

Highlights

- In a clinically ascertained cohort of 313 Ontario patients with hypercholesterolemia, a targeted next-generation sequencing panel identified monogenic mutations in 53.7% and an extreme polygenic risk score in an additional 13.4% of individuals.
- The percentage of individuals with an identified genetic component increased from 57.0% to 92.0% as LDL cholesterol level increased from 5.0 to >8.0 mmol/L (194 to >310 mg/dL).
- Next-generation sequencing based on evaluation of hypercholesterolemia needs to consider both monogenic and polygenic causes.

Arteriosclerosis, Thrombosis, and Vascular Biology

JOURNAL OF THE AMERICAN HEART ASSOCIATION

Polygenic Versus Monogenic Causes of Hypercholesterolemia Ascertained Clinically

Jian Wang, Jacqueline S. Dron, Matthew R. Ban, John F. Robinson, Adam D. McIntyre, Maher Alazzam, Pei Jun Zhao, Allison A. Dilliott, Henian Cao, Murray W. Huff, David Rhainds, Cécile Low-Kam, Marie-Pierre Dubé, Guillaume Lettre, Jean-Claude Tardif and Robert A. Hegele

Arterioscler Thromb Vasc Biol. 2016;36:2439-2445; originally published online October 20, 2016;
doi: 10.1161/ATVBAHA.116.308027
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1079-5642. Online ISSN: 1524-4636

Materials and Methods

Subjects

Three hundred and thirteen consecutive unrelated adults aged ≥18 years from Ontario, Canada referred with possible heterozygous familial hypercholesterolemia (HeFH) were included; patients known to have homozygous FH were excluded. Untreated fasting lipid profiles were recorded, according to our established clinic procedure.[1, 2] All subjects had untreated Friedewald-determined plasma low-density lipoprotein (LDL) cholesterol >5.0 mmol/L (194 mg/dL), plus either a personal or family history of early cardiovascular disease (CVD), plus family history of hyperlipidemia. Physical findings were not uniformly recorded. Diagnosis of HeFH was made using the Dutch Lipid Clinic Networks (DLCN) criteria.[3] All subjects provided signed informed consent and the protocol was approved by the Western University Ethics Review Board (number 07290E). Of note, 55 of the patients in this study were also part of the analysis conducted by Futema et al.[4]

As a reference group for the genetic risk score (GRS) analysis, Phase 1 of the 1000 Genomes Project (1KG) was used. This cohort is comprised of 1,092 individuals of varying ethnicities over the age of 18 who are self-reported as healthy. Based on these reports and lack of additional phenotype information, we assume some fluctuations in LDL cholesterol levels, but by and large, we consider the majority of the population as normolipidemic.

Genetic analyses

Genomic DNA was isolated from whole blood as described.[1] Genomic libraries of indexed and pooled patient samples were generated for target candidate genes in lipid metabolism, including coding regions, ≥150 base pairs (bp) at intron-exon boundaries and >1000 bp of the 5' untranslated region of the known causative genes for FH, namely *LDLR*, *APOB*, *PCSK9*, *STAP1*, *APOE*, and *LDLRAP1* on the LipidSeq Panel, as described.[5, 6] *ABCG5*, *ABCG8*, and *LIPA*, some mutations in which can cause similar phenotypes in some patients, were also included. The reagents also allow capture of the 10 genotypes for common single nucleotide polymorphisms (SNPs) statistically associated with LDL cholesterol levels in the general population[7]; the trait-raising alleles are tallied to create a polygenic trait score.[6] Prepared sample libraries were assayed in the MiSeq personal sequencer (Illumina, San Diego CA) as described.[5] The method has average >300-fold coverage for each base. Samples were also run using multiplex ligation primer amplification (MLPA) for coding regions of the *LDLR* gene, as described.[8] Sanger sequencing was used to confirm variants detected by next-generation sequencing (NGS).

Annotation and evaluation of observed variants

FASTQ files derived from the MiSeq output were processed individually using a custom automated workflow in CLC Genomics Workbench version 8.5.1 (CLCbio, Aarhus, Denmark) for sequence alignment, variant calling, producing a variant call format (vcf) file, and target region coverage statistics. Variant annotation was performed using ANNOVAR (http://www.biobase-international.com/product/annovar) with customized scripts.

There is no consensus on the procedure to attribute causality or pathogenicity to variants detected by NGS.[9, 10] Fortunately, variants detected in FH genes have had a long history of archiving and annotation, as well as abundant publications of functional consequences.[11] For instance, >1700 individual variants previously reported as being causative in FH are reported in the Human Gene Mutation Database (HGMD; http://www.biobase-

international.com/product/hgmd) and the University College London (UCL) FH mutation database (http://www.ucl.ac.uk/ldlr/Current/); these are the reference databases for all variants detected by our procedure.

Annotated coding and noncoding (±10 base pair from adjacent exon) variants in vcfs were first filtered to select the rare variants according to minor allele frequencies (MAF) <1% in 1KG (http://www.1000genomes.org/), Exome Variant Server (EVS; http://evs.gs.washington.edu/EVS/) or Exome Aggregation Consortium (ExAC; http://exac.broadinstitute.org/) databases. Polymorphism Phenotype Version 2 (PolyPhen-2)[12] Sorting Intolerant from Tolerant (SIFT)[13] and Combined Annotation Dependent Depletion (CADD)[14] scores were used to evaluate the deleteriousness of the filtered coding variants. Splicing Based Analysis of Variants (SPANR)[15] and Automated Splice Site and Exon Definition Analyses (ASSEDA; www.http://splice.uwo.ca) were used to identify rare deleterious splicing variants.

Novel variants found in this study were determined to be likely causative when: 1) they had no listed allele frequencies in 1KG, EVS or ExAC databases, no rsID in the dbSNP database, and/or were not reported in HGMD or UCL FH databases; 2) for coding variants, a deleterious score from $\geq$2 in silico algorithms; and 3) for non-coding variants, a deleterious score for $\geq$1 in silico algorithm. Copy number variants (CNVs) detected by MLPA were similarly searched for in HGMD and UCL FH databases. Hereafter, we will use the term "mutation" interchangeably with "rare definite or very likely causative variant" for the sake of brevity. As controls for our annotation pipeline, we used sequence data from the 1KG database. Rare variants are summarized in Supplemental Table I.

Polygenic trait scores

A set of 10 genetic markers associated with raising plasma LDL cholesterol were selected from the original lipid genome-wide association study (GWAS) (Supplemental Table II)[7], as the association signals from these loci have persisted as the strongest signals in the subsequent Global Lipid Genetic Consortium GWAS reports.[16, 17] Both weighted and unweighted GRS (wGRS and uwGRS, respectively) were calculated; for the former, the weighting factors were the published beta-coefficients for per-allele change in LDL cholesterol.[7]. We chose the 90th percentile for wGRS - i.e. $\geq$1.96 - as the definition for an extreme score from the subjects in the 1KG database. The 90th percentile for the uwGRS from 1KG was 16/20.

Statistical analysis

All statistical comparisons were conducted using SAS version 9.3 (SAS Institute, Cary NC). Between-group differences in quantitative traits means were evaluated using unpaired Student's t-test assuming unequal variances. For discrete traits, χ^2 analysis was used and odds ratios (ORs) were calculated using the case-control method with the FREQ procedure. Statistical significance for all comparisons was defined as a two-tailed P-value <0.05.

Supplemental References

1. Johansen CT, Wang J, Lanktree MB, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet*. 2010;42:684-687
2. Johansen CT, Wang J, Lanktree MB, et al. An increased burden of common and rare lipid-associated risk alleles contributes to the phenotypic spectrum of hypertriglyceridemia. *Arterioscler Thromb Vasc Biol*. 2011;31:1916-1926
3. van Aalst-Cohen ES, Jansen AC, Tanck MW, Defesche JC, Trip MD, Lansberg PJ, Stalenhoef AF, Kastelein JJ. Diagnosing familial hypercholesterolaemia: The relevance of genetic testing. *Eur Heart J*. 2006;27:2240-2246
4. Futema M, Shah S, Cooper JA, et al. Refinement of variant selection for the ldl cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries. *Clin Chem*. 2015;61:231-238
5. Johansen CT, Dube JB, Loyzer MN, MacDonald A, Carter DE, McIntyre AD, Cao H, Wang J, Robinson JF, Hegele RA. Lipidseq: A next-generation clinical resequencing panel for monogenic dyslipidemias. *J Lipid Res*. 2014;55:765-772
6. Hegele RA, Ban MR, Cao H, McIntyre AD, Robinson JF, Wang J. Targeted next-generation sequencing in monogenic dyslipidemias. *Curr Opin Lipidol*. 2015;26:103-113
7. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet*. 2009;41:56-65
8. Wang J, Ban MR, Hegele RA. Multiplex ligation-dependent probe amplification of LDLR enhances molecular diagnosis of familial hypercholesterolemia. *J Lipid Res*. 2005;46:366-372
9. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. 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;17:405-424
10. Manrai AK, Ioannidis JP, Kohane IS. Clinical genomics: From pathogenicity claims to quantitative risk estimates. *JAMA*. 2016;315:1233-1234
11. Talmud PJ, Futema M, Humphries SE. The genetic architecture of the familial hyperlipidaemia syndromes: Rare mutations and common variants in multiple genes. *Curr Opin Lipidol*. 2014;25:274-281
12. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248-249
13. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4:1073-1081
14. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46:310-315
15. Xiong HY, Alipanahi B, Lee LJ, et al. Rna splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science*. 2015;347:1254806
16. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707-713
17. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013;45:1274-1283

Supplementary Table I. Summary of Causative Mutations in Patients Presenting with Low-Density Lipoprotein Cholesterol > 5 mmol/L

Gene	Location	Nucleotide change	Mutation name	Number of occurrences in patients	PMID / Novel
LDLR	Intron 2	c.190+4A>T	IVS2+4A>T	1	16250003
LDLR	Intron 2	c.191-2A>G	IVS2-2A>G	1	9259195
LDLR	Intron 3	c.313+1G>A	IVS3+1G>A	2	7718019
LDLR	Intron 3	c.313+2T>C	IVS3+2T>C	3	7616128
LDLR	Intron 3	c.314-2A>C	IVS3-2A>C	5	11668627
LDLR	Intron 4	c.695-1G>A	IVS4-1G>A	1	Novel
LDLR	Intron 7	c.1060+2T>G	IVS7+2T>G	1	Novel
LDLR	Intron 8	c.1187-10G>A	IVS8-10G>A	2	11668627
LDLR	Intron 9	c.1359-1G>A	IVS9-1G>A	1	9254862
LDLR	Intron 12	c.1845+1G>A	IVS12+1G>A	1	15556094
LDLR	Intron 12	c.1845+15C>A	IVS12+15C>A	1	15576851
LDLR	Intron 12	c.1846-1G>A	IVS12-1G>A	1	8828981
LDLR	Exon 1	c.16_17insTTCCT	p.W6Ffs*201	1	Novel
LDLR	Exon 2	c.233delG	p.R78Lfs*127	1	Novel
LDLR	Exon 4	c.647_648insT	p.C216Cfs*1	1	Novel
LDLR	Exon 4	c.652delG	p.G218Vfs*46	1	7649546
LDLR	Exon 4	c.680_681delAC	p.D227Gfs*11	1	8093663
LDLR	Exon 6	c.820delA	p.T274Hfs*95	4	19026292
LDLR	Exon 6	c.905delG	p.C302Sfs*67	1	Novel
LDLR	Exon 9	c.1205_1206delTC	p.F402Ffs*37	1	11040093
LDLR	Exon 10	c.1476_1477delCT	p.S493Cfs*41	2	7866407
LDLR	Exon 17	c.2411_2412insG	p.V806Gfs*10	1	Novel
LDLR	Promoter, Exon 1	>15kb del Promoter & Exon 1	del Promoter & Exon 1	12	15576851
LDLR	Exons 1, 2, 3, 4	del Exons 1-4	del Exons 1-4	1	ND
LDLR	Exons 3, 4, 5, 6	del Exons 3-6	del Exons 3-6	1	ND
LDLR	Exons 6	del Exon 6	del Exon 6	1	ND
LDLR	Exon 7	dup Exon 7	dup Exon 7	1	ND
LDLR	Exons 11, 12	del Exons 11, 12	del Exons 11, 12	1	ND
LDLR	Exons 16, 17, 18	del Exons 16-18	del Exons 16-18	1	ND
LDLR	Exons 17, 18	del Exons 17, 18	del Exons 17, 18	2	ND

LDLR	Exon 2	c.168_170delTGA	p.D57del	1	11668627
LDLR	Exon 2	c.131G>A	p.W44X	1	1301956
LDLR	Exon 3	c.337G>T	p.E113X	2	1301956
LDLR	Exon 4	c.460C>T	p.Q154X	1	10206683
LDLR	Exon 4	c.501C>A	p.C167X	1	7616128
LDLR	Exon 8	c.1176C>A	p.C392X	1	8831933
LDLR	Exon 10	c.1467C>G	p.Y489X	1	7833932
LDLR	Exon 12	c.2043C>A	p.C681X	4	3025214
LDLR	Exon 15	c.2167G>T	p.E723X	1	Amsellem (2000) LDLR LSDB, 757
LDLR	Exon 15	c.2215C>T	p.Q739X	1	12417285
LDLR	Exon 2	c.81C>G	p.C27W	1	1301956
LDLR	Exon 2	c.187T>C	p.C63R	1	11668627
LDLR	Exon 3	c.259T>G	p.W87G	2	2318961
LDLR	Exon 3	c.268G>A	p.D90N	1	9259195
LDLR	Exon 3	c.299A>T	p.D100V	1	Novel
LDLR	Exon 3	c.301G>A	p.E101K	4	1301940
LDLR	Exon 4	c.427T>A	p.C143S	1	Amsellem (2000) LDLR LSDB, 697
LDLR	Exon 4	c.504C>A	p.D168E	1	15823276
LDLR	Exon 4	c.517T>C	p.C173R	1	1301956
LDLR	Exon 4	c.551G>A	p.C184Y	2	9678702
LDLR	Exon 4	c.590G>A	p.C197Y	1	1301956
LDLR	Exon 4	c.662A>G	p.D221G	1	1301956
LDLR	Exon 4	c.681C>G	p.D227E	2	2569482
LDLR	Exon 4	c.682G>A	p.E228K	1	2318961
LDLR	Exon 6	c.858C>A	p.S286R	3	1301956
LDLR	Exon 6	c.862G>A	p.E288K	1	10090484
LDLR	Exon 7	c.979C>T	p.H327Y	1	9259195
LDLR	Exon 7	c.986G>A	p.C329Y	1	9452118
LDLR	Exon 7	c.1003G>A	p.G335S	2	1301956
LDLR	Exon 7	c.1027G>A	p.G343S	1	1301956
LDLR	Exon 8	c.1085A>C	p.D362A	1	11810272
LDLR	Exon 8	c.1091G>A	p.C364Y	1	Novel
LDLR	Exon 8	c.1102T>C	p.C368R	1	9452094

LDLR	Exon 8	c.1151A>C	p.Q384P	1	11810272
LDLR	Exon 8	c.1186G>A	p.G396S	1	7573037
LDLR	Exon 9	c.1241T>G	p.L414R	1	9452118
LDLR	Exon 9	c.1255T>G	p.Y419D	1	Novel
LDLR	Exon 9	c.1285G>A	p.V429M	2	2569482
LDLR	Exon 9	c.1325A>G	p.Y442C	1	15576851
LDLR	Exon 9	c.1329G>T	p.W443C	1	11810272
LDLR	Exon 10	c.1436T>C	p.L479P	1	11313767
LDLR	Exon 10	c.1444G>A	p.D482N	1	8535447
LDLR	Exon 10	c.1567G>A	p.V523M	3	2088165
LDLR	Exon 11	c.1592T>G	p.M531R	1	Novel
LDLR	Exon 11	c.1618G>A	p.A540T	2	9544745
LDLR	Exon 11	c.1646G>A	p.G549D	1	2088165
LDLR	Exon 12	c.1745T>C	p.L582P	5	11668627
LDLR	Exon 12	c.1775G>A	p.G592E	2	1301956
LDLR	Exon 12	c.1816G>T	p.A606S	1	9544745
LDLR	Exon 12	c.1833G>C	p.L611F	1	11810272
LDLR	Exon 13	c.1868T>C	p.I623T	1	Novel
LDLR	Exon 13	c.1897C>T	p.R633C	1	9259195
LDLR	Exon 14	c.2000G>A	p.C667Y	2	2318961
LDLR	Exon 14	c.2029T>C	p.C677R	1	1301956
LDLR	Exon 14	c.2054C>T	p.P685L	2	2726768
LDLR	Exon 14	c.2096C>T	p.P699L	2	7489239
LDLR	Exon 14	c.2120A>T	p.D707V	1	20809525
LDLR	Exon 15	c.2242G>A	p.D748N	1	25487149
LDLR	Exon 16	c.2343G>T	p.E781D	1	11668627
LDLR	Exon 17	c.2475C>A	p.N825K	3	11668640
APOB	Exon 26	c.10580G>A	p.R3527Q	14	2563166
APOB	Exon 26	c.10579C>T	p.R3527W	1	7627691
PCSK9	Exon 8	c.1251C>A	p.H417Q	1	16465619
PCSK9	Exon 9	c.1405C>T	p.R469W	1	16211558
LDLR, *LDLR*	*LDLR* Exons 2 & 4	*LDLR* c.165C>G; *LDLR* c.520G>T	*LDLR* p.[G55G(;) E174X]	1	Novel; Novel
LDLR, *LDLR*	*LDLR* Exons 3 & 6	*LDLR* del Exon 3 & Exon 6	*LDLR* del Exon 3(;) del Exon 6	1	ND; ND

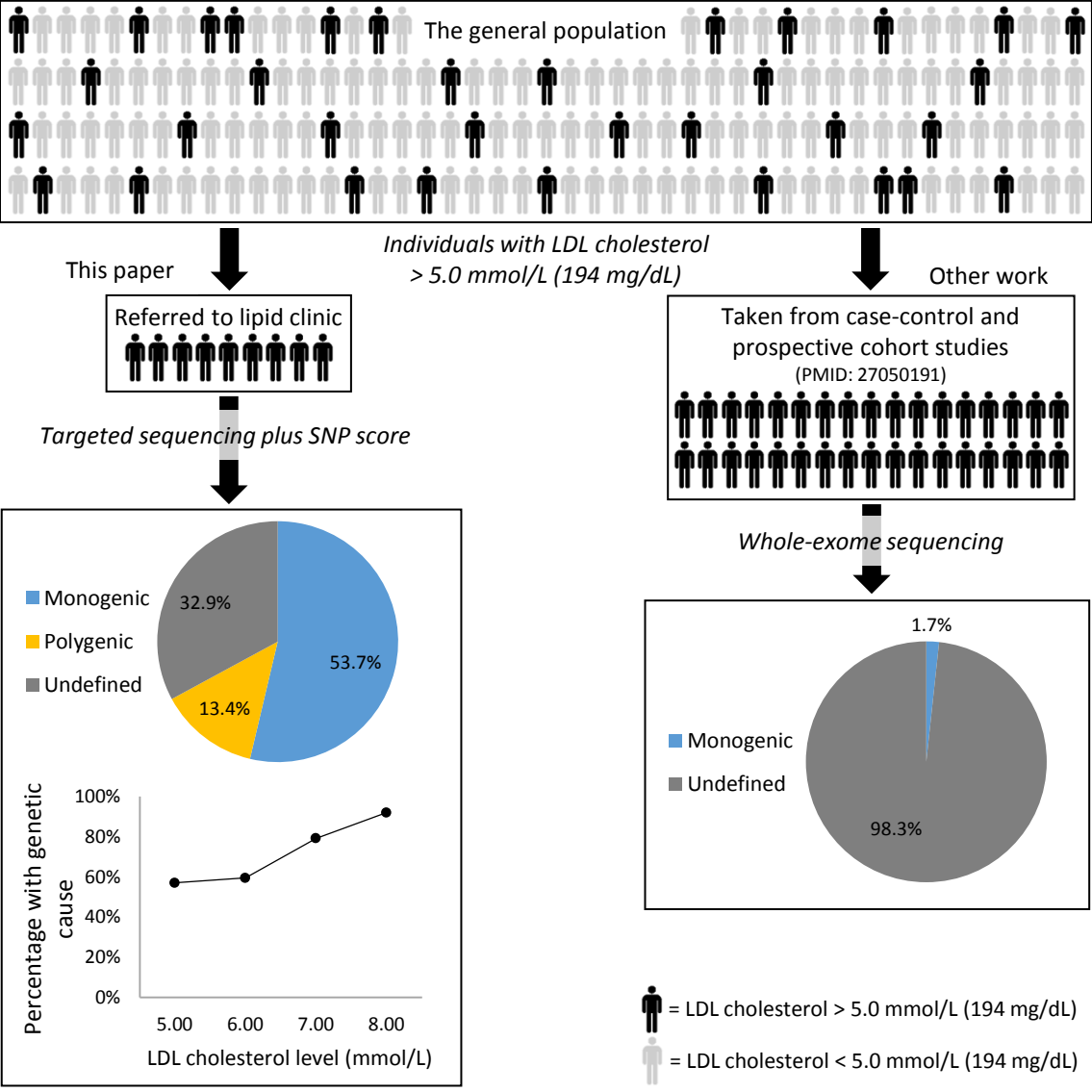
LDLR, LDLR	*LDLR* Intron 6 & Exon 9	*LDLR* c.941-4G>A; c.1285G>A	*LDLR* [IVS6-4G>A(;) p.V429M]	1	10422804; 2569482
LDLR, LDLR	*LDLR* Exons 11 & 17	*LDLR* c.1690A>C; c.2393_2401delTCCTCGTCT	*LDLR* p.[N564H(;) L799_F801del]	1	7550239; 9147888
LDLR, APOB	*LDLR* Exon 9; *APOB* Exon 26	*LDLR* c.1301C>T; *APOB* c.10580G>A	*LDLR* p.T434M; *APOB* p.R3527Q	1	11668627; 2563166
LDLR, APOB	*LDLR* Exon 10; *APOB* Exon 26	*LDLR* c.1408A>G; *APOB* c.10580G>A	*LDLR* p.S470G; *APOB* p.R3527Q	1	Novel; 2563166
LDLR, APOB	*LDLR* Intron 12; *APOB* Exon 26	*LDLR* c.1846-10G>T; *APOB* c.10580G>A	*LDLR* IVS12-10G>T; *APOB* p.R3527Q	1	Novel; 2563166
LDLR, APOB	*LDLR* Exon 13; *APOB* Exon 26	*LDLR* c.1868T>C; *APOB* c.10580G>A	*LDLR* p.I623T; *APOB* p.R3527Q	1	Novel; 2563166
LDLR, PCSK9	*LDLR* Exon 3; *PCSK9* Exon 5	*LDLR* c.259T>G; *PCSK9* c.709C>T	*LDLR* p.W87G; *PCSK9* p.R237W	1	2318961; 16465619
LDLR, PCSK9	*LDLR* Intron 8; *PCSK9* Exon 10	*LDLR* c.1187-10G>A; *PCSK9* c.1537A>G	*LDLR* IVS8-10G>A; *PCSK9* p.N513D	1	11668627; Novel

Abbreviations: *LDLR*, gene encoding low density lipoprotein receptor, NM_000527; *APOB*, gene encoding apolipoprotein B, NM_000384; *PCSK9*, gene encoding proprotein convertase subtilisin/kexin type 9, NM_174936; PMID, PubMed unique identifier; ND, not determined; del, deletion; ins, insertion; *, termination codon; dup, duplication; fs, frameshift

Supplementary Table II. Single Nucleotide Polymorphisms Used in Construction of Weighted Genetic Risk Score

Chr:Start	rsID	Gene	Associated phenotype	Ref	Alt	LDL-C raising allele	Effect on LDL-C
1:55496039	rs11206510	*PCSK9*	LDL-C	T	C	T	0.09
1:109817590	rs12740374	*CELSR2*	LDL-C	G	T	G	0.23
2:21286057	rs515135	*APOB*	LDL-C	T	C	C	0.16
2:44073881	rs6544713	*ABCG8*	LDL-C	T	C	T	0.15
5:74655726	rs3846663	*HMGCR*	LDL-C	C	T	T	0.07
5:156398169	rs1501908	*TIMD4*	LDL-C	G	C	C	0.07
12:121388962	rs2650000	*HNF1A*	LDL-C	A	C	A	0.07
19:11202306	rs6511720	*LDLR*	LDL-C	G	T	G	0.26
19:19407718	rs10401969	*NCAN*	LDL-C	T	C	T	0.05
20:39228784	rs6102059	*MAFB*	LDL-C	C	T	C	0.06

Abbreviations: Chr, chromosome; Ref, reference allele; Alt, alternate allele; LDL-C, low-density lipoprotein cholesterol; PMID, PubMed unique identifier; *PCSK9*, gene encoding proprotein convertase subtilisin/kexin type 9; *CELSR2*, gene encoding cadherin EGF lag seven-pass G-type receptor 2; *APOB*, gene encoding apolipoprotein B; *ABCG8*, gene encoding ATP-binding cassette subfamily G member 8; *HMGCR*, gene encoding 3-hydroxy-3-methylglutaryl-CoA reductase; *TIMD4*, gene encoding T-cell immunoglobulin and mucin domains-containing protein 4; *HNF1A*, gene encoding hepatic nuclear factor-1 homeobox A; *LDLR*, gene encoding low density lipoprotein receptor; *NCAN*, gene encoding chondroitin sulfate proteoglycan 3; *MAFB*, gene encoding V-MAF musculoaponeurotic fibrosarcoma oncogene family protein B. Marker information and effect sizes, represented by beta coefficients, are from Kathiresan *et al.*(reference 7 from the Supplemental References).


The general population
Individuals with LDL cholesterol > 5.0 mmol/L (194 mg/dL)
This paper
Referred to lipid clinic
Targeted sequencing plus SNP score
Monogenic
Polygenic
Undefined
32.9%
53.7%
13.4%
Percentage with genetic cause
100%
80%
60%
40%
20%
0%
5.00
6.00
7.00
8.00
LDL cholesterol level (mmol/L)
Other work
Taken from case-control and prospective cohort studies
(PMID: 27050191)
Whole-exome sequencing
1.7%
Monogenic
Undefined
98.3%
= LDL cholesterol > 5.0 mmol/L (194 mg/dL)
= LDL cholesterol < 5.0 mmol/L (194 mg/dL)