

Analysis of publicly available *LDLR*, *APOB*, and *PCSK9* variants associated with familial hypercholesterolemia: application of ACMG guidelines and implications for familial hypercholesterolemia diagnosis

Joana Rita Chora, MSc^{1,2}, Ana Margarida Medeiros, MSc^{1,2}, Ana Catarina Alves, PhD^{1,2} and Mafalda Bourbon, PhD^{1,2}

Purpose: Familial hypercholesterolemia (FH) is an autosomal disorder of lipid metabolism presenting with increased cardiovascular risk. Although more than 1,700 variants have been associated with FH, the great majority have not been functionally proved to affect the low-density lipoprotein receptor cycle. We aimed to classify all described variants associated with FH and to establish the proportion of variants that lack evidence to support their pathogenicity.

Methods: We followed American College of Medical Genetics and Genomics (ACMG) guidelines for the classification, and collected information from a variety of databases and individual reports. A worldwide overview of publicly available FH variants was also performed.

Results: A total of 2,104 unique variants were identified as being associated with FH, but only 166 variants have been proven by

complete *in vitro* functional studies to be causative of disease. Additionally, applying the ACMG guidelines, 1,097 variants were considered pathogenic or likely pathogenic. Only seven variants were found in all five continents.

Conclusion: The lack of functional evidence for about 85% of all variants found in FH patients can compromise FH diagnosis and patient prognosis. ACMG classification improves variant interpretation, but functional studies are necessary to understand the effect of about 40% of all variants reported. Nevertheless, ACMG guidelines need to be adapted to FH for a better diagnosis.

Genet Med advance online publication 26 October 2017

Key Words: ACMG classification; familial hypercholesterolemia; FH diagnosis; functional studies

INTRODUCTION

Familial hypercholesterolemia (FH) is an autosomal disorder of lipid metabolism and one of the most common genetic disorders that confer increased cardiovascular risk.¹ Patients with FH have very high low-density lipoprotein (LDL) values from birth, so the early identification and treatment of these patients improves their prognosis.² However, the clinical diagnosis of FH has low specificity and sensitivity³ and only genetic diagnosis can confirm which individuals have a lifelong increased cardiovascular risk.⁴ The genetic diagnosis is made by finding a functional mutation in one of the three genes known to cause FH: low-density lipoprotein receptor gene (*LDLR*), apolipoprotein B gene (*APOB*), and proprotein convertase subtilisin/kexin type 9 gene (*PCSK9*). More than 1,700 variants have been identified in *LDLR* in patients with a clinical diagnosis of FH,⁵ but only a minority have been functionally proved to cause FH. This represents a huge gap for FH diagnosis. To overcome this common problem in the diagnosis of genetic disorders, i.e., the lack of functional evidence for variant pathogenicity, the American College of

Medical Genetics and Genomics (ACMG) published in 2015 an algorithm to classify all variants into classes that are helpful for the interpretation of the effect of the variant: pathogenic, likely pathogenic, variant of unknown significance (VUS), likely benign, and benign.⁶

In the present study we aimed to classify, following ACMG guidelines, all described variants associated with FH in various databases and individual reports to establish the proportion of variants that lack evidence to support their pathogenicity. A rigorous analysis of all functional studies performed for these variants has been performed and classified in two levels depending on the evidence obtained in each study. We also present the variant distribution throughout the world, highlighting the countries where more variants have been found and the ones with a higher percentage of variants functionally characterized. Additionally the misfits of ACMG algorithm to classify FH-associated variants have been identified so a specific algorithm for FH variant classification can be designed. Finally, an approach for the reporting of these variants to the clinician is discussed.

¹Unidade I&D, Grupo de Investigação Cardiovascular, Departamento de Promoção da Saúde e Doenças Não Transmissíveis, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal; ²University of Lisbon, Faculty of Sciences, BioISI—Biosystems & Integrative Sciences Institute, Lisbon, Portugal. Correspondence: Mafalda Bourbon (mafalda.bourbon@insa.min-saude.pt)

Submitted 16 March 2017; accepted 26 July 2017; advance online publication 26 October 2017. doi:10.1038/gim.2017.151

MATERIALS AND METHODS

Construction of an FH database

The FH database was constructed in the steps described below. All information included refers up to December 2015. **Supplementary Table S1** online lists the fields collected.

Systematic collection of variants in LDLR, APOB, and PCSK9 associated with FH

Variant collection was performed as follows. The data were imported from the following public databases: Leiden Open (source) Variation Database versions 1–3,⁵ the Human Gene Mutation Database,⁷ and the Universal Mutation Database (UMD).⁸ Variants listed in the public databases in association with other diseases, such as hypocholesterolemia or hypobetalipoproteinemia, as well as loss-of-function variants in *PCSK9*, were not included. Only public users' Human Gene Mutation Database information was considered. Additionally, a PubMed search was conducted, using as keywords "familial hypercholesterolemia," "familial hypercholesterolaemia," "FH," "*LDLR*," "*APOB*," or "*PCSK9*," both separately and in combination. All papers published in the previous 10 years (2005–2015), written in English, were carefully analyzed, and any additional variant or additional information was added to our database.

Variant annotation

Data were pooled from the databases, duplicates were removed, and nomenclature was checked and corrected, when necessary, using the Mutalyzer Name Checker tool.⁹ All variants collected in papers were also checked with this software. The following transcripts were used in this work: *LDLR*: NM_000527.4, *APOB*: NM_000384.2, and *PCSK9*: NM_174936.3 (assembly: GRCh37).

Analysis of functional studies

Every paper with functional data cited in the public databases or found in the PubMed search was carefully analyzed concerning all information stated and validated by two or more team members. Functional studies were divided in two levels regarding their functional evidence (**Table 1**). Studies in heterologous cells or in homozygous patients' cells ascertain

more clearly the specific variant's effect in the *LDLR* cycle and so were considered level 1 studies. Studies in heterozygous patients' cells always have to take into account the interference of the wild-type allele and so the results are not so clear to interpret; also RNA studies without transcript quantification only inform about the different transcript species produced without clarifying the relative amount of mutant protein being produced, so these cases were considered level 2 studies. As in previous studies,¹⁰ we considered that results of less than 80% of normal activity (or less than 90% in homozygous patients' cells) were considered supportive of a damaging effect. All this information was added to the database.

In silico analysis

In silico analysis was performed for every variant using only open access software and included the following programs: Protein Variation Effect Analyzer,¹¹ Sorting Tolerant From Intolerant,¹² PolyPhen-2,¹³ and Mutation Taster¹⁴ for prediction of protein structure/function changes and evolutionary conservation; and Neural Network Splice Site Prediction Tool,¹⁵ MAXENTSCAN,¹⁶ FSPLICE,¹⁷ and Human Splicing Finder¹⁸ for prediction of splicing defects, when appropriate.

General considerations

Any variant that was wrongly written, either in one of the databases or in a paper, that was not possible to verify was excluded. Internal and yet unpublished information about functional studies, normolipidemic panel results, and novel variants from our lab were also included. All internal unpublished information has been submitted to ClinVar public database and is now live at the ClinVar website (<https://www.ncbi.nlm.nih.gov/clinvar/submitters/505909>).

Application of 2015 ACMG guidelines

The ACMG guidelines⁶ were applied to all variants identified following specific FH assumptions described in **Supplementary Tables S2** and **S3**. Very briefly, these assumptions included (i) nine criteria that were not considered for FH, namely, PS2, PM6, PM3, and BP1, because they contemplate

Table 1 Types of functional studies (FS) performed by gene

	Type of FS	<i>LDLR</i>	<i>APOB</i>	<i>PCSK9</i>	Total
Level 1	Homozygous patients' cell studies	71	–	–	71
	RNA studies with transcript quantification	19	–	–	19
	Heterologous cell studies	65	–	–	65
	Luciferase studies	19	–	1	20
	<i>APOB</i> and <i>PCSK9</i> complete <i>LDLR</i> cycle studies	–	7	18	25
	Total level 1 FS	174	7	19	200
Level 2	Heterozygous patients' cell studies	40	–	–	40
	RNA studies without transcript quantification	34	–	–	34
	Other/Proliferation studies	11	11	2	24
	Total level 2 FS	85	11	2	98
	Total	259	18	21	298

different disease mechanisms from the ones found in FH; PP5, BS1, and BP6, because the same information had been considered in other criteria and BP2 and BP5 were to be considered in a diagnosis context and not in this analysis and (ii) four criteria that were changed and/or divided into new criteria following suggested changes in the original algorithm, for example, the division of PS3/BS3 into the newly created PM7/BP8 for functional studies, as explained in “Analysis of functional studies” above, the division of PP1, and the new PM9 criteria for cosegregation with larger families and PS4, which has been changed to a moderate level.

Final classification of pathogenic, likely pathogenic, VUS, likely benign, or benign was assigned according to the published algorithm.⁶

Analysis of FH variants in the world

With all the information collected, several analyses were performed as the total number of variants and the total number of variants with functional results per country. Also, an analysis of the number of variants found per continent and number of variants common to all continents was conducted. It is important to note that only variants that were clearly stated to be found in a specific population were considered. In this study, we use “America” to refer to North and South America.

RESULTS

Variant analysis

Total variants

We performed an analysis of public databases and literature for every variant published associated with FH, in the genes *LDLR*, *APOB*, and *PCSK9*. The upper boxes in **Figure 1** show the number of variants in each gene reported in each database. All information was collected in a joint database and, after correcting nomenclature and filtering duplicates, we identified 1,860 different variants. The large majority (42%) of variants are shared only between the Human Gene Mutation Database and the Leiden Open (source) Variation Database, 28.7% of variants are present in only one of the databases, and 23.6% are shared between all three.

A PubMed search for papers published in 2005–2015, as well as an analysis of every paper referenced in the public databases, resulted in the identification of 255 additional variants (**Figure 1**). Adding to these, 46 as-yet-unpublished Portuguese variants brings the grand total to 2,161 different variants identified in FH patients: 1,915 in *LDLR*, 107 in *APOB*, and 139 in *PCSK9* (**Figure 1**). From all these, 57 variants present a minor allele frequency above 5% in either 1000 Genomes, the Exome Sequencing Project, or the Exome Aggregation Consortium and were therefore excluded from the following analysis; the final box in **Figure 1** shows the number of variants, 2,104 (1,894 *LDLR*, 97 *APOB*, 113 *PCSK9*), considered from this point on.

Types of variants

Variants were defined according to the kind of structural alteration that is observed at the DNA level—point

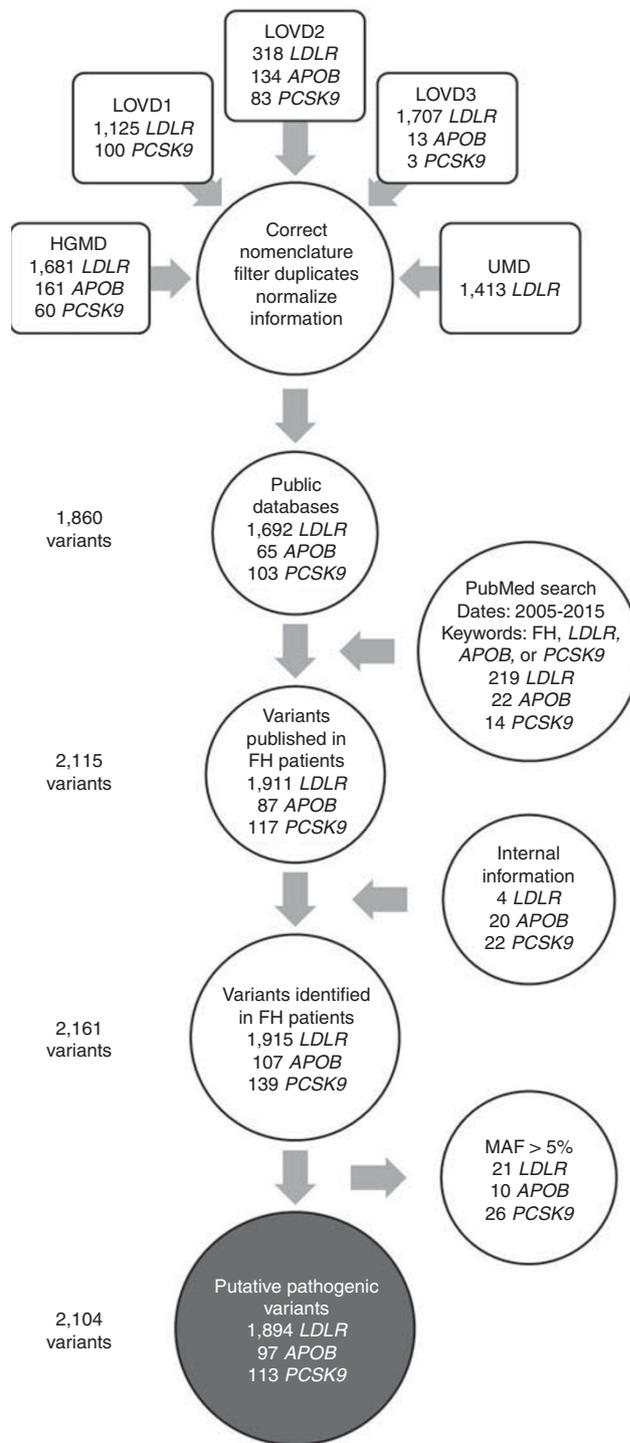


Figure 1 Analysis methodology and results. The top boxes correspond to the number of variants in public databases. Additional variants were added to the database collected from PubMed review and unpublished internal information. Variants with minor allele frequency >5% were excluded from further analysis. The numbers of variants analyzed in this paper are shown in the shaded box on the bottom. FH, familial hypercholesterolemia; HGMD, Human Gene Mutation Database;⁷ LOVD, Leiden Open (source) Variation Database versions 1 through 3;⁵ MAF, minor allele frequency; UMD, Universal Mutation Database.⁸

substitution, deletion, or insertion—or to the kind of effect that is expected to be observed in the RNA or protein produced/location of the variant, i.e., missense, frameshift, nonsense, in-frame deletion or insertion, synonymous, splicing, intronic, regulatory, or large rearrangements.

Structurally, point substitutions represent 68% of all variants (65% of *LDLR*, 97% of *APOB*, and 96% of *PCSK9* variants). Duplications and deletions of more than one exon (large rearrangements) are only described in *LDLR* and account for 9% of this gene's variants. Small deletions and insertions represent 26% of *LDLR*, 3% of *APOB*, and 4% of *PCSK9* variants.

The majority of FH variants described are missense, ranging from 46% of variants in *LDLR*, to 52% in *PCSK9*, to 83% in *APOB*. Null variants, including nonsense, frameshift, and large deletions, represent 33% of *LDLR* variants (Supplementary Figure S1). In *PCSK9*, intronic and synonymous variants account for 35% of variants.

The distribution of *LDLR* variants is shown in Supplementary Figure S2, except for large rearrangements, along the 18 exons and 6 protein domains. Exon 4, which codifies for the ligand binding domain, has the highest number ($n = 336$) of different variants described, most of them missense (53.3%). This is also the largest *LDLR* exon, with 381 nucleotides, but even when number of variants by nucleotide is considered, exon 4 remains the one with the highest number of variants, with 0.882 variants per nucleotide. This is also the exon with more variants functionally characterized and all with a pathogenic effect (Supplementary Table S4).

Variants with functional studies

Nonsense and frameshift variants as well as large deletions are considered null variants, and therefore inherently pathogenic. This means that 634 FH variants would not need functional studies to elucidate their pathogenicity, leaving 1,470 putative pathogenic variants (1,261 in *LDLR*, 96 in *APOB*, and 113 in *PCSK9*) in need of functional studies.

In total, we gathered functional study information for 392 variants. According to the evidence, we classified the type of functional study in two levels, with level 1 having the highest evidence (Table 1) and including 166 variants. Studies

performed in compound heterozygous cells were not considered for variant classification, due to the inability to distinguish between the two variants' effects. Studies in four variants were deemed inconclusive (Table 1) because of different results obtained in individuals with the same genotype. Some examples are variant c.259T > G/p.Trp87Gly in *LDLR*, referenced by Hobbs et al.¹⁹: (LDLR receptor activities were 25 to 55% of normal in subject 1, 30 to 50% in subject 2, and 65 to 100% in subject 3") and by Hattori et al.,²⁰ who reported a family possessing variant c.946_994del/p.(Asn316Leufs*38) in *LDLR* and presenting levels of LDLR protein ranging from 44% to 75% compared with wild-type, with levels of uptake between 77% and 91% compared with wild-type.

The 298 variants with the functional studies considered included 55 null variants, which means that only 16.5% (243/1,470) of variants needing functional study in the three genes have a clear pathogenicity assessment: 16% in *LDLR*, 19% in *APOB*, and 19% in *PCSK9*.

ACMG classification

The published ACMG criteria⁶ was used to classify all 2,104 variants considering all the information collected. The majority of criteria ($n = 15$) were applied as described in the original algorithm, nine criteria were not considered due to specificities of FH diagnosis, and four criteria were changed and split into new criteria as suggested in the original algorithm (Supplementary Table S2 and S3).

The majority of variants ($n = 986$) landed in the VUS category (Table 2 and Supplementary Table S5), mainly due to lack of sufficient information and/or evidence for it to be classified as either likely benign/benign or likely pathogenic/pathogenic. One variant, c.226 G > T/p.(Gly76Trp) in *LDLR* was classified as VUS because of a contradictory classification, fitting both the likely pathogenic and benign categories. A total of 705 variants were classified as pathogenic. Since all variants with minor allele frequency above 5% were excluded from this analysis, very few variants were classified as benign ($n = 12$) (Supplementary Table S6).

As described above, considering that all variants producing a null allele do not need functional proof to be cause of disease due to their deleterious effect on the LDLR protein, a

Table 2 Number of *LDLR*, *APOB*, and *PCSK9* variants associated with familial hypercholesterolemia with functional studies by ACMG classification

ACMG classification	No FS		Level 1 FS		Level 2 FS		Total
	Null	Others	Pathogenic	Benign	Pathogenic	Benign	
Benign	–	–	–	11	–	1	12
Likely benign	–	–	–	–	–	9	9
VUS	99	824	13	23	17	10	986
Likely pathogenic	–	323	43	–	26	–	392
Pathogenic	480	80	110	–	35	–	705
Total	579	1,227	166	34	78	20	2,104

ACMG, American College of Medical Genetics; FS, functional study; VUS, variant of unknown significance.

Null variants include nonsense, frameshift and single or multiexon deletions. Pathogenicity in functional studies was considered when a functional study had an LDLR activity result of less than 80% of normal, and for benign variants when a result of more than 80% of normal was obtained.



Figure 2 Familial hypercholesterolemia molecular heterogeneity worldwide. Only papers where patient origin was clearly stated were considered. Color scale indicates the number of different variants reported by country.

total of 1,470 variants needed functional evidence to be classified as pathogenic or benign. Functional studies were performed for 243/1,470 variants leaving 1,057 variants in *LDLR* (57.8% of total), 78 in *APOB* (80.4%), and 92 in *PCSK9* (81.4%) still needing functional evidence to be considered disease-causing. Using the ACMG classification with the described assumptions for FH these numbers are reduced to 824 variants in need of pathogenicity clarification: 655 variants in *LDLR* (34.6%), 77 in *APOB* (79.4%), and 92 in *PCSK9* (81.4%); still, about 40% of all variants found need functional evidence to be considered disease-causing mutations.

An interesting consideration drawn from the ACMG classification was that having a functional study was not enough to classify variants in the benign or pathogenic category. Of the 298 variants with level 1 or 2 functional studies, 63 did not have enough evidence to be classified by ACMG as anything other than VUS (Table 2). Furthermore, 43 variants with level 1 functional study were only classified as likely pathogenic despite their pathogenicity having been proven by complete *in vitro* studies. Additionally 99 null variants did not receive enough points in ACMG classification to be classified as pathogenic, and were instead classified as VUS. The majority are large deletions (>90%), which are known to affect the *LDLR* function but could not receive enough points in the algorithm to be classified as pathogenic.

FH variants across the world

The largest number of identified variants is found in Europe ($n = 1,491$), followed by Asia ($n = 332$), America ($n = 265$), Oceania ($n = 134$), and Africa ($n = 77$). No *APOB* or *PCSK9* variants were identified in African countries.

The number of different variants identified throughout the world by country can be observed in Figure 2. Since only

papers where patient origin was clearly stated were considered here, we believe this could be an underestimation. For 293 variants, no country was associated.

A large majority (80.6%) of variants with geographic information are exclusive to a single continent and only 352 cross continent borders (Supplementary Figure S3). Only seven *LDLR* variants are present on all five continents (no FH studies have been published in Antarctica), all of which are point substitutions: c.313+1G > A/p.Leu64_Pro105delinsSer, c.681C > G/p.(Asp227Glu), c.1222G > A/p.(Glu408Lys), c.1285G > A/p.(Val429Met), c.1432G > A/p.(Gly478Arg), c.2043C > A/p.(Cys681*), and c.2054C > T/p.Pro685Leu. These are classified as pathogenic ($n = 5$), likely pathogenic ($n = 1$, c.1222G > A/p.(Glu408Lys)), and VUS ($n = 1$, c.1432G > A/p.(Gly478Arg)).

The countries with the highest diversity of variants described are shown in Table 3. Expected number of FH patients per country (using World Bank 2014 population numbers and 1/500 FH prevalence²¹) is indicated for comparison. The only non-European country present, Japan, appears in 10th place, with 119 different variants published. The countries with the highest percentage of *LDLR* variants with functional assessment are Portugal (56.2%), Canada (44.4%), and Spain (43.0%) (Supplementary Table S7).

DISCUSSION

Although more than 2,000 variants have been reported in clinical FH patients, less than 10% have been validated as disease-causing by complete *in vitro* functional assays. This has serious implications for the genetic diagnosis of FH. Functional studies are of even greater importance for *APOB* and *PCSK9* genes, which are more polymorphic than *LDLR*. In fact, the number of *APOB* variants has increased lately due to the use of next-generation sequencing in genetic diagnosis

Table 3 Top 10 countries with the highest number of different familial hypercholesterolemia putative pathogenic variants described

Country	LDLR	APOB	PCSK9	Total	Expected number of FH patients ^a
United Kingdom	321	7	4	332	130,222
Netherlands	306	5	–	311	33,959
Italy	251	2	8	261	119,602
Portugal	137	40	39	216	20,609
Spain	205	1	5	211	92,129
France	147		12	159	129,336
Norway	129	2	14	145	10,544
Czech Republic	129	–	–	129	21,096
Germany	112	9	1	122	161,365
Japan	91	–	28	119	252,647

FH, familial hypercholesterolemia.

^aCalculated using World Bank 2014 and the proposed 1/500 prevalence for heterozygous FH.²¹ Of the 2,104 variants found associated with FH only 1,811 had a clear description of patient nationality so only those were considered for this analysis.

of FH, which offers the possibility of sequencing the whole coding sequence of *APOB* and splice regions not studied in routine diagnosis before. This is especially important because recently functional mutations have been found outside the conserved binding region of *APOB*,²² and also several rare variants have been found to be benign, showing the importance for *in vitro* validation for these variants so they can be considered disease-causing mutations.

Considering that all variants producing a null allele do not need functional proof to be disease-causing due to their deleterious effect on the LDLR protein, a total of 1,227 variants in either *LDLR*, *APOB*, or *PCSK9* still need functional evidence to be considered disease-causing. Using the ACMG classification with the specific assumptions for FH present here, this number is reduced to 824 variants that are still considered VUS; nevertheless this represents a significant improvement. However, still about 40% of all variants found in FH patients need functional evidence to be considered disease-causing mutations. This represents a vast gap for FH diagnosis and, consequently, for patient prognosis, as it has been shown that patients with FH have 16 times greater chance of developing premature coronary heart disease than dyslipidemic individuals without a functional mutation in the 3 genes causing FH.²³

As it is documented, only a few groups have the desire and/or ability to perform functional studies,^{10,24–27} even though some studies are not that difficult to perform. This must change in order to obtain the best FH diagnosis.

Although the ACMG guidelines have been demonstrated to be a valid tool for variant classification, having helped to clarify the classification of several variants for the genetic diagnosis of FH, several gaps in the algorithm have been identified: large deletions are classified as VUS, functional studies do not have the weight they should have on variant

classification, and some variants can be classified as benign and pathogenic at the same time. This highlights the need to develop specific guidelines for FH diagnosis covering these gaps and other specific criteria for FH.

For now we recommend that the results of the molecular diagnosis of FH should be reported only if solid functional evidence exists towards the pathogenicity of a variant, if the ACMG algorithm classifies the variant as pathogenic or likely pathogenic, or if the variant found is expected to produce a null protein. This will avoid misdiagnosis. The example of the variant c.806G > A/p.(Gly269Asp) (formerly known as Gly248Asp) exemplifies very well this problem; this variant was considered disease-causing in all countries where it was described—Spain,²⁸ the Netherlands,²⁹ and Italy (FH Rome-3)—until in 2008 we reported complete lack of cosegregation in a Portuguese family where later a pathogenic variant was found.³⁰ In 2012 the functional assay was performed³¹ and it proved beyond any doubt that this variant had little or no effect on the LDLR function, being considered as benign. This was one of the most common variants in Spain³¹ and was described in other countries as mentioned before; reports had to be withdrawn and new reports had to be sent to the clinician and it had to be explained to the patient that the cause of their dyslipidemia had, in fact, not been found yet. To avoid this kind of situation an effort should be made to perform functional assays for the remaining VUS in order for these variants to be reported to the clinician only with proof of pathogenic evidence. However, when a VUS is found and it cosegregates with the phenotype within the family, this result could be reported to the clinician/patient if, and only if, it is clearly explained that without further evidence this variant does not yet confirm the diagnosis of FH. In this situation, the lab should take it as their responsibility to provide an updated report when new evidence is published, changing the variant classification, for example, when the functional assay has been performed and reported in literature.

It is interesting to note that FH genetic diagnosis is happening almost all over the world, mainly within the scope of research projects,^{30,32–37} since only The Netherlands, Spain, and Uruguay have governmental approval for the identification of FH patients.^{38,39} The fact that most variants are restricted to a continent was also an interesting finding, suggesting that population migration occurs mostly within the same continent. However, there are several variants found on more than one continent, probably reflecting the discovery period especially by Portuguese and Spanish navigators in the 16th century. Only seven variants have been found across all five continents.

Finally, in this work a total of 1,894 *LDLR* variants were found reported in patients with a clinical diagnosis of FH, 187 more than the last review of the most complete FH database, the Leiden Open (source) Variation Database.⁵ This highlights the problem of database curation, which ideally would need a full-time person, for lifetime, to keep the databases updated but in most cases this is impossible without proper

funding. On the other hand, an updated database is crucial for the genetic diagnosis of a disease since a well-curated database is the most useful tool for variant interpretation.

Concluding, a map of the heterogeneity of all variants causing FH has been performed highlighting that a large number of variants do need *in vitro* functional validation to be considered disease-causing mutations. Without these studies the genetic diagnosis of FH is seriously compromised. To avoid misdiagnosis a positive result should be sent to the clinician only when enough evidence exists to support the pathogenicity of a variant. Although activity toward the genetic diagnosis of FH is seen almost all over the world, FH is still underdiagnosed and efforts should be made in both clinical and genetic identification of these patients. Since most countries have shown ability to perform the genetic diagnosis of FH, governmental approval and funding in each country for large-scale screening, as recommended by the World Health Organization in 1998,²¹ would be of utmost importance to improve FH patients' identification and prognosis. As discussed above, the laboratory tools have been developed and enhanced over the past 20 years and we are now more capable than ever to correctly identify these patients; moreover, several pharmacology treatments have been developed and shown to be effective for decreasing the elevated cardiovascular risk of these patients. It is clear now that implementation of these measures can be life-changing for these patients.

Limitations

An assumption was made in this study design that all papers published before 2005 were already included in the public databases references list. Any papers not in this condition were not included in this analysis.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

This work was partly supported by a grant from Genediag. J.R.C. and A.M.M. acknowledge their PhD fellowships funded by the Science and Technology Foundation (SFRH/BD/108503/2015 and SFRH/BD/113017/2015, respectively). The authors also acknowledge Eleanor Benton for editorial assistance.

DISCLOSURE

M.B. has received project grants from Alexion Pharmaceuticals. The other authors declare no conflict of interest.

REFERENCES

- Nordestgaard BG, Chapman MJ, Humphries SE, et al. 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;34:3478–3490.
- Catapano A, Graham I, De Backer G, et al. 2016 ESC/EAS guidelines for the management of dyslipidaemias. *Eur Heart J*. 2016;37:2999–3058.
- Henderson R, O'Kane M, McGilligan V, Watterson S. The genetics and screening of familial hypercholesterolaemia. *J Biomed Sci*. 2016;23:39.
- Santos RD. Phenotype vs. genotype in severe familial hypercholesterolemia. *Curr Opin Lipidol*. 2017;28:130–135.
- Leigh S, Futema M, Whittall R, et al. The UCL low-density lipoprotein receptor gene variant database: pathogenicity update. *J Med Genet*. 2017;54:217–223.
- Richards S, Aziz N, Bale S, et al. 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.
- Stenson PD, Mort M, Ball EV, Shaw K, Phillips AD, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet*. 2014;133:1–9.
- Villéger L, Abifadel M, Allard D, et al. The UMD-LDLR database: additions to the software and 490 new entries to the database. *Hum Mutat*. 2002;20:81–87.
- Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PEM. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Hum Mutat*. 2008;29:6–13.
- Benito-Vicente A, Alves AC, Etxebarria A, Medeiros AM, Martin C, Bourbon M. The importance of an integrated analysis of clinical, molecular, and functional data for the genetic diagnosis of familial hypercholesterolemia. *Genet Med*. 2015;17:980–988.
- Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015;31:2745–2747.
- Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–3814.
- Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249.
- Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11:361–362.
- Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol*. 1997;4:311–323.
- Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol*. 2004;11:377–394.
- FSPLICE: Find Splice Sites in Genomic DNA. <http://www.softberry.com/berry.phtml?topic=fssplice&group=programs&subgroup=gfind>. Accessed 10 February 2017.
- Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37:e67–e67.
- Hobbs, et al. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat*. 1992;1:445–466.
- Hattori H, Hirayama T, Nobe Y, et al. Eight novel mutations and functional impairments of the LDL receptor in familial hypercholesterolemia in the north of Japan. *J Hum Genet*. 2002;47:80–87.
- WHO Human Genetics Programme. *Familial hypercholesterolaemia (FH): report of a second WHO consultation*. World Health Organization: Geneva, 1999.
- Alves AC, Etxebarria A, Soutar AK, Martin C, Bourbon M. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet*. 2014;23:1817–1828.
- Khera A V., Emdin CA, Drake I, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med*. 2016;375:2349–2358.
- Cameron J, Holla OL, Laerdahl JK, et al. Characterization of novel mutations in the catalytic domain of the PCSK9 gene. *J Intern Med*. 2008;263:420–431.
- Silva S, Alves AC, Patel D, Malhó R, Soutar AK, Bourbon M. In vitro functional characterization of missense mutations in the LDLR gene. *Atherosclerosis*. 2012;225:128–134.
- Fasano T, Sun XM, Patel DD, Soutar AK. Degradation of LDLR protein mediated by "gain of function" PCSK9 mutants in normal and ARH cells. *Atherosclerosis*. 2009;203:166–171.
- Alves AC, Etxebarria A, Medeiros AM, et al. Characterization of the first PCSK9 gain of function homozygote. *J Am Coll Cardiol*. 2015;66:2152–2154.
- Mozas P, Castillo S, Tejedor D, et al. Molecular characterization of familial hypercholesterolemia in Spain: Identification of 39 novel and 77 recurrent mutations in LDLR. *Hum Mutat*. 2004;24:187–187.

29. FH Mutation Database. <http://www.jojo genetics.nl/wp/database/>. Accessed 10 March 2017.
30. Bourbon M, Alves AC, Medeiros AM, Silva S, Soutar AK. Familial hypercholesterolaemia in Portugal. *Atherosclerosis*. 2008;196:633–642.
31. Etxebarria A, Palacios L, Stef M, et al. Functional characterization of splicing and ligand-binding domain variants in the LDL receptor. *Hum Mutat*. 2012;33:232–243.
32. Jannes CE, Santos RD, de Souza Silva PR, et al. Familial hypercholesterolemia in Brazil: cascade screening program, clinical and genetic aspects. *Atherosclerosis*. 2015;238:101–107.
33. Marduel M, Carrié A, Sassolas A, et al. Molecular spectrum of autosomal dominant hypercholesterolemia in France. *Hum Mutat*. 2010;31:1811–1824.
34. Alharbi KK, Kashour TS, Al-Hussaini W, et al. Screening for genetic mutations in LDLR gene with familial hypercholesterolemia patients in the Saudi population. *Acta Biochim Pol*. 2015;62:559–562.
35. Vaca G, Vázquez A, Magaña MT, et al. Mutational analysis of the LDL receptor and APOB genes in Mexican individuals with autosomal dominant hypercholesterolemia. *Atherosclerosis*. 2011;218:391–396.
36. Grenkowitz T, Kassner U, Wuhle-Demuth M, et al. Clinical characterization and mutation spectrum of German patients with familial hypercholesterolemia. *Atherosclerosis*. 2016;253:88–93.
37. Taylor A, Wang D, Patel K, et al. Mutation detection rate and spectrum in familial hypercholesterolaemia patients in the UK pilot cascade project. *Clin Genet*. 2010;77:572–580.
38. Fouchier SW, Defesche JC, Umans-Eckenhausen MW, Kastelein JP. The molecular basis of familial hypercholesterolemia in The Netherlands. *Hum Genet*. 2001;109:602–615.
39. Santos RD, Bourbon M, Alonso R, et al. Clinical and molecular aspects of familial hypercholesterolemia in Ibero-American countries. *J Clin Lipidol*. 2017;11:160–166.