The UCL low-density lipoprotein receptor gene variant database: pathogenicity update

Sarah Leigh, Marta Futema, Ros Whittall, Alison Taylor-Beadling, Maggie Williams, Johan T den Dunnen, Steve E Humphries

ABSTRACT

Background Familial hypercholesterolaemia (OMIM 143890) is most frequently caused by variations in the low-density lipoprotein receptor (LDLR) gene. Predicting whether novel variants are pathogenic may not be straightforward, especially for missense and synonymous variants. In 2013, the Association of Clinical Genetic Scientists published guidelines for the classification of variants, with categories 1 and 2 representing clearly not or unlikely pathogenic, respectively, 3 representing variants of unknown significance (VUS), and 4 and 5 representing likely to be or clearly pathogenic, respectively. Here, we update the University College London (UCL) LDLR variant database according to these guidelines.

Methods PubMed searches and alerts were used to identify novel LDLR variants for inclusion in the database. Standard in silico tools were used to predict potential pathogenicity. Variants were designated as class 4/5 only when the predictions from the different programs were concordant and as class 3 when predictions were discordant.

Results The updated database (http://www.lovd.nl/LDLR) now includes 2925 curated variants, representing 1707 independent events. All 129 nonsense variants, 337 small frame-shifting and 117/118 large rearrangements were classified as 4 or 5. Of the 795 missense variants, 115 were in classes 1 and 2, 605 in class 4 and 75 in class 3. 111/181 intronic variants, 4/34 synonymous variants and 14/37 promoter variants were assigned to classes 4 or 5. Overall, 112 (7%) of reported variants were class 3.

Conclusions This study updates the LDLR variant database and identifies a number of reported VUS where additional family and in vitro studies will be required to confirm or refute their pathogenicity.

INTRODUCTION

Familial hypercholesterolaemia (FH) is an autosomal co-dominant disorder with a reported frequency in Europe of between 1 in 217 in Denmark to 1 in 600 in Finland. FH results from defective clearance of atherogenic low-density lipoprotein cholesterol (LDL-C) particles from the blood; consequently, patients with FH have a 5–8 times higher-than-average risk of premature coronary heart disease (CHD), which can be significantly reduced with statin treatment. FH is caused predominantly by variants in the low-density lipoprotein receptor (LDLR) gene. Pathogenic changes in the LDLR gene result in either fewer or functionally impaired LDL receptors and consequently impaired LDL-C particle uptake. All recent guidelines for the management of FH (reviewed in ref. 6) have emphasised the clinical utility of identifying the genetic variant in index patients to confirm the FH diagnosis, support the early commencement of intensive statin therapy and to use the information to test unambiguously identify any relatives who may also carry the same variant, thereby allowing them to be offered lifestyle and therapeutic advice to reduce their CHD risk.

In the era of next-generation sequencing, there is an increasing demand for rapid genetic diagnosis; however, variant interpretation remains the most challenging part of the diagnostic process. The availability of an up-to-date open access locus-specific database for LDLR variants is a crucial tool for diagnostic and research laboratories. The UCL LDLR variant database embraces the principles of data sharing (variants and phenotypes), which ultimately will result in improved classification of novel LDLR variants (as pathogenic or not) as more data is collected. Clearly, accurate classification of such variants will be of benefit to patients and their relatives with regard to better diagnosis, prognosis, treatment and appropriate cascade testing. Since its inception in 1997, the UCL LDLR online variant database has undergone four major upgrades, with the addition of newly published variants, analysis and transition to improved web-based platforms.8–11 Predicting whether or not novel variants in LDLR are pathogenic may not always be straightforward, especially for synonymous and missense variants or for those occurring in intronic sequences or the promoter regions of the gene. Recently, guidelines for diagnostic laboratories reporting novel variants have been proposed by the Association for Clinical Genetic Science (ACGS), with classes 1 and 2 as clearly not or unlikely to be pathogenic, respectively, class 3 as variants of unknown significance (VUS), and 4 and 5 as likely to be or clearly pathogenic, respectively. The guidelines recommend a number of in silico programs for the prediction of the likely impact of missense variants on the resultant peptide, and intronic, missense and synonymous variants that may affect mRNA splicing. The guidelines state, “It is acceptable to predict the severity of an amino acid change using in-silico methods, but it is unacceptable to rely solely on these predictions to assign pathogenicity to a previously unclassified variant”. For variants affecting splicing, they state, “It is acceptable to assign nucleotide changes that disrupt the consensus dinucleotide splice site (+/-1 and +/-2) as clearly pathogenic requiring no further investigation. It is acceptable to use in-silico splice site prediction; however, it is unacceptable to...
base an unequivocal clinical interpretation solely on this line of evidence”. Therefore, in vitro studies or ex vivo examination of RNA is required to look for the presence of abnormal splice products. Here, we update the UCL LDLR variant database with the addition of variants in the literature since 2012 and classification of all variants according to the ACGS guidelines. This should provide a standardised set of data regarding the pathogenicity of reported LDLR gene variants that have been identified in subjects with the clinical phenotype of FH. We believe that this will be of value to diagnostic laboratories and to physicians requesting DNA tests of index cases with FH.

MATERIALS AND METHODS
Identification of LDLR variants
Newly reported variants were identified in the literature using the term ‘familial hypercholesterolemia mutation’ for PubMed searches and alerts. The nomenclature of variants was checked with Mutalzyer (https://mutalyzer.nl)15 to ensure that it adhered to Human Genome Variation Society recommendations (http://www.hgvs.org/mutnomen).14 In a number of cases, the nomenclature of the variants has been changed due to updated recommendations (eg, single base pair insertions now being described as duplications where one of the adjacent bases is the same as the inserted base and large rearrangements where formerly unknown breakpoints were described, eg, as follows: c.67+1?_941-1?del are now described as c.(67+1_68-1)_940-1_941-1?del). Nomenclature changes have also been made to correct inaccuracies in the literature in consultation with the originating authors wherever possible. All variants are described with respect to the reference sequences: LDLR LRG_274, based on NG_009060.1 with LRG_274_1 based on NM_000527.4. Variants were submitted to and made available through an LOVD-powered gene variant database15 that can be accessed via http://www.lovdl.nl/LDLR.

In silico prediction of variant pathogenicity
Nonsense substitutions, frame-shifting small and large rearrangements were not subjected to in silico analyses as they are accepted to be pathogenic (classes 4 and 5). The predicted effects of missense variants on LDLR function were assessed using the following open access software packages: (a) PolyPhen2 (HumDiv and HumVar) (http://genetics.bwh.harvard.edu/pph2/index.shtml),16 (b) SIFT (Sorting Intolerant From Tolerant) (http://sift.jcvi.org/www/SIFT_seq_submit2.html),17 Refined SIFT (http://sift.jcvi.org/www/SIFT_aligned_seqx_submit.html)10 (for LDLR amino acid sequences used, see online supplementary table S1; for results of Refined SIFT analysis, see online supplementary table S2) and (c) Mutation Taster (http://www.mutationtaster.org).18

The effects of intronic and some synonymous variants on gene splicing were assessed using Berkeley Drosophila Genome Project, Splice Site Prediction by Neural Network, with the minimum scores for 5’ and 3’ splice sites set at 0.01 (http://www.fruitfly.org/seq_tools/splice.html)19 and SplicePort (http://spliceport.cbcb.umd.edu/).20 Both programs give predictive scores for splice acceptor and donor sequences for wild-type and variant sequences. The pathogenic impact of LDLR variants in the promoter and 5’ untranslated region of the gene have recently been published.21 Where appropriate, structural conservation scores have been given to variants (see online supplementary table S3).11

Pathogenicity class was assigned to each variant according to the ACGS guidelines as follows: all nonsense substitutions and frame-shifting variants were assigned to classes 4 or 5 (likely to be or clearly pathogenic) as were exon-deleting large rearrangements. Intronic variants affecting residues ±1 or 2 from intron/exon boundaries were assigned to class 4 (or 5 where in vitro evidence revealed frame-shifting splicing events). Pathogenicity scores were assigned to the remaining variants by taking into account information available in the literature including family segregation, in vitro studies and an overall assessment of the predictions from in silico analyses. Furthermore, the reported frequency of variants in large population studies such as the 1000 Genomes22 and Exome Aggregation Consortium (ExAC) database23 was also taken into account, such that a minor allele frequency (MAF) >0.0002 was considered to be suggestive of a non-pathogenic variant as this is the frequency of the most common single FH-causing variant, the p.R3527Q located in APOB.24 For missense variants with no supporting evidence from the literature, variants with three or more consistent in silico predictions were assigned as either class 4 (likely to be pathogenic) or 2 (unlikely to be pathogenic) and those with inconsistent in silico predictions were assigned as class 3 (VUS).

RESULTS
A total of 2925 variants, representing 1707 unique events, have now been added to the UCL LDLR variant database on the LOVD3 platform (http://www.lovdl.nl/LDLR). The number of unique events reported on the database homepage differs from 1707 as the compiling software assumes that large rearrangements involving the same exons are identical, whereas they may well have different breakpoints as defined by the differing sizes of the fragments deleted or duplicated. For the purpose of analysis, the variants were subdivided into the following categories: promoter (n=37), intronic (n=181), synonymous substitutions (n=34), nonsense substitutions (n=129), missense substitutions (n=795), large rearrangements (>100 bp) (n=118), small frameshifting rearrangements (<100 bp) (n=337) and small non-frame-shifting rearrangements (<100 bp) (n=76). A summary of pathogenicity classifications for each variant category is shown in figure 1. Overall, 81% of variants were classified as pathogenic (n=800 class 4 and n=585 class 5) and 12% as non-pathogenic (n=26 class 1 and n=184 class 2), with 7% (n=112) being classed as VUSs.

All nonsense substitutions (1 class 4, 128 class 5), small frame-shifting rearrangements (3 class 4, 334 class 5) and 117/118 large rearrangements (22 class 4, 95 class 5) were considered to be pathogenic according to the criteria set out in the
ACGS guidelines. One large rearrangement was classified as a VUS as it was an in-frame duplication of exons 16–18, which may or may not impact on the correct transcription of the LDLR gene.

Missense substitutions represent the largest category of variants in the database (n=795), of which 605 (76%) were classified as likely pathogenic (class 4), 115 (15%) as clearly not or probably not pathogenic (classes 1 and 2) and 75 (9%) as VUSs. In vitro functional analysis was only available for 73 (9%) missense variants, and so the majority of the pathogenicity classifications were made using in silico predictions. In an attempt to gauge the reliability of the in silico tools used in this study, pathogenicity classifications based on in vitro functional studies and in silico predictions were compared where possible. The results were in agreement for 63/73 (86%) of the variants, and those with discordant results are shown in Table 1. Since the functional studies provide the more reliable evidence, we have assigned these variants to the categories suggested by the in vitro and not the in silico data.

From the total of 181 intronic variants, the 98 involving bases at positions ±1 or 2 from intron/exon boundaries were considered to be pathogenic. Seventeen of these were classed as clearly pathogenic (class 5) as in vitro evidence was available from the literature that demonstrated the disruption of normal mRNA splicing. In vitro functional analysis was available for 15 of the variants that involved bases beyond ±2 in the intron; as a result, 13 were classified as pathogenic (8 class 4, 5 class 5) (Table 2). Of the remaining 70 intronic variants 11 were classified as class 3 as in silico predictions suggested that they would disrupt normal splicing, but no in vitro evidence was available.

Only two of the in-frame small rearrangements (c.887_889delinsAGC, c.1659_1661delinsATACTTTC) were classified as clearly pathogenic as both resulted in the creation of in-frame termination codons. Of the remaining 74 variants in this category, 60 were assigned to class 4 and 13 to class 3 (VUS). Variant c.1120_1121delinsTC was assigned to class 2 because the resultant amino acid change, p.(Gly374ser), was predicted to be benign according to in silico analysis.

As mentioned previously, across the database 112 variants (7% of the total) were classified as VUSs (class 3). With the exception of nonsense and frame-shifting small rearrangements, these variants occur in all other variant categories, 43% (n=16) of variants in the promoter region were class 3, as were 17% (n=13) of small non-frame-shifting rearrangements, 9% (n=75) of missense and 4% (n=8) intronic variants. Finally, only one synonymous and one large rearrangement were assigned to class 3.

Discussion

The ACGS practice guidelines for the evaluation of pathogenicity designate variants as classes 1 and 2 as clearly not or unlikely to be pathogenic, respectively, 3 as a VUS, and 4 or 5 as likely to be or clearly pathogenic, respectively. In general, we have adhered to these guidelines, but for FH there is such a strong a priori probability that an FH-causing variant will be found in the LDLR gene that we believe less stringent proof of functionality is reasonable. In particular, we believe that showing a novel variant with ‘likely pathogenic’ in silico predictions may be reported as class 4 if it is present in several unrelated patients with FH and is absent or at very low frequency (<MAF=0.0002) in sequence databases such as the 1000 Genomes and ExAC database (http://exac.broadinstitute.org). The finding in this study that in silico predictions matched the in vitro evidence in 63/73 (86%) missense variants lends support to this view.

Of the variants where the results were not in agreement (Table 1), five class 2 and three class 3 variants by in silico analyses were shown to be class 4 by in vitro functional studies. The variant c.226G>T, p.(Gly76Trp), resulted in strong pathogenic predictions from PolyPhen2, both SIFT analyses and Mutation Taster, and it might be expected that replacement of the highly conserved small glycine amino acid at position 76 with the larger tryptophan would be detrimental. However, this variant was reported in non-FH family members and displayed normal levels of LDLR expression, LDL-binding and internalisation.26 Similarly, c.769C>T, p.(Arg253Trp) (rs200990725), was found in non-FH samples (4 heterozygotes (Hz) in 1000 Genomes and 9 Hz in the ExAC) databases, and although the in silico analyses supported a pathogenic prediction, in vitro analysis revealed that there was no adverse effect on function.29 The two variants c.2389G>A and c.2389G>T, which both result in p.(Val797Met), were predicted to be probably not pathogenic by in silico analyses; however, in vitro studies show that both variants disrupt normal mRNA splicing as they affect the last nucleotide of exon 16.13 34 and so were deemed to be class 4. Thus, extra caution must be used with in silico analysis of variants close to intron/exon boundaries as predictions for in silico tools such as PolyPhen2 and SIFT are based on the effects of amino acid changes on the mature peptide rather than differences that the variant DNA may have on normal mRNA processing, and so such variants should also be analysed with the splice predicting programs also. Likewise, three synonymous variants (c.621C>T, p.(Gly207=), c.1216C>A, p.(Arg406=) and c.1813C>T, p.(Leu605=)) were shown to disrupt normal splicing by both in silico and in vitro analyses and c.1845G>A, p.(Glu615=) was predicted to do so by in silico analysis alone as it affected the last residue in exon 12 (no in vitro evidence is available for this variant at present). Although the majority of synonymous variants will not be pathogenic, closer examination of their impact is justified if they have not been reported in non-FH subjects, if they are present in unrelated patients with FH and if segregation with FH has been demonstrated.

As mentioned previously, it is generally accepted that intronic variants affecting bases ±1 and 2 from the intron/exon boundary will be pathogenic as these residues are highly conserved.44 The updated UCL LDLR database currently lists 98 such variants of which in vitro evidence was available for 17. The potential significance of variants beyond the immediate intron/exon boundary was examined for those variants with accompanying in vitro data (Table 2). Also, 11 variants within the first 12 bases of the intron/exon boundaries were found to be pathogenic by either inactivating the wild-type splice site resulting in exon skipping or by the creation of novel splice sites that were used preferentially. Furthermore, in vitro studies revealed that c.1359-31_1359-23delinsCGGCCG resulted in the removal of the invariant adenine at the consensus splicing branch site in intron 9, causing retention of intron 9 and use of cryptic splice sites in exon 10,39 and c.2140+86C>G resulted in the creation and
## Table 1  Low-density lipoprotein receptor (LDLR) missense variants with discordant *in silico* and *in vitro* functional studies

<table>
<thead>
<tr>
<th>DNA</th>
<th>Peptide</th>
<th>Refined SIFT</th>
<th>SIFT</th>
<th>PolyPhen 2 HumDiv</th>
<th>PolyPhen 2 HumVar</th>
<th>Mutation taster</th>
<th>Structural conservation</th>
<th>In silico prediction (ACGS classification)</th>
<th>In vitro classification (ACGS classification)</th>
<th>In vitro evidence (LDLR defect classification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.226G&gt;T</td>
<td>p.(Gly76Trp)</td>
<td>Not tolerated</td>
<td>Not tolerated</td>
<td>Probably damaging, 1</td>
<td>Probably damaging, 1</td>
<td>Disease causing</td>
<td>0.922</td>
<td>Likely to be pathogenic</td>
<td>Probably not pathogenic</td>
<td>Normal LDLR expression, LDL-binding and internalisation</td>
</tr>
<tr>
<td>c.344G&gt;A</td>
<td>p.(Arg115His)</td>
<td>Not tolerated</td>
<td>Not tolerated</td>
<td>Benign, 0.157</td>
<td>Benign, 0.128</td>
<td>Disease causing</td>
<td>0.54</td>
<td>VUS</td>
<td>Likely to be pathogenic</td>
<td>Retained in ER, 36% reduced LDLR on cell surface, 36% reduced LDL binding and internalisation (class 2)</td>
</tr>
<tr>
<td>c.768C&gt;G</td>
<td>p.(Asp256Glu)</td>
<td>Not tolerated</td>
<td>Not tolerated</td>
<td>Benign, 0.388</td>
<td>Benign, 0.329</td>
<td>Disease causing</td>
<td>0.721</td>
<td>VUS</td>
<td>Likely to be pathogenic</td>
<td>50–70% reduced LDL binding and internalisation, co-segregated with FH in large family (class 3)</td>
</tr>
<tr>
<td>c.1361C&gt;A</td>
<td>p.(Thr454Asn)</td>
<td>Not tolerated</td>
<td>Tolerated</td>
<td>Possibly damaging, 0.993</td>
<td>Possibly damaging, 0.916</td>
<td>Polymorphism</td>
<td>0.711</td>
<td>VUS</td>
<td>Likely to be pathogenic</td>
<td>Normal LDLR expression, LDL-binding and internalisation</td>
</tr>
<tr>
<td>c.1942T&gt;C</td>
<td>p.(Ser648Pro)</td>
<td>Tolerated</td>
<td>Benign, 0.12</td>
<td>Benign, 0.083</td>
<td>Benign, 0.07</td>
<td>Polymorphism</td>
<td>0.537</td>
<td>Likely to be pathogenic</td>
<td>Retained in ER, 80% reduced LDLR on cell surface, 75% reduced LDL binding and internalisation (class 2)</td>
<td></td>
</tr>
</tbody>
</table>

ACGS, Association for Clinical Genetic Science; FH, familial hypercholesterolaemia; LDL, low-density lipoprotein; SIFT, Sorting Intolerant From Tolerant; VUS, variants of unknown significance.
Genotype-phenotype correlations


use of a novel splice site in intron 14.42 Thus, there is the potential for variants deep into the introns to be pathogenic, and as with the synonymous variants, some of these intronic variants may warrant further examination. According to the ACGS guidelines, it is acceptable to predict that any variant that results in a premature truncation of the peptide either as a result of a nonsense variant or a reading shift change and will be pathogenic. However, in rearrangements where this is not the case it is more difficult to make a prediction; clearly deletion of whole exons and functional domains would be likely to have deleterious effects. Although changes in the peptide secondary and tertiary structures may result from the addition or removal of small numbers of amino acids, their effect is difficult to predict; furthermore, very few in vitro functional studies have been published for such variants, probably because they would be technically difficult and costly to perform and also because the assumption may be made that these variants are likely to be pathogenic. As advances are made in in vitro functional assays, it is hoped that more evidence will be provided to confirm or refute the pathogenicity of non-frame-shifting rearrangements.

Traditionally LDLR variants have been grouped into one of five classes based on their functional effects (class 1 Null, class 2 Transport defective, ie, retained in the ER, class 3 Binding defective, Class 4 Internalisation defective, Class 5 Recycling deficient).25 However, as more is understood about the different mechanisms that can impact on normal LDLR function, additional classes could usefully be added to this list. In 2001, Koivisto et al demonstrated that the cytoplasmic variant c.2531G>A, p.(Gly844Asp) interfered with the basolateral sorting of LDLR with the effect that the peptide was mistargeted to the apical surface of the cell, hence reducing the sorting of LDLR with the effect that the peptide was misstar
ted to the apical surface of the cell.49 50 It appears that c.2396T>G, p.(Leu799Arg) from the cell.49 50 It appears that c.2396T>G, p.(Leu799Arg) 

Table 2 Low-density lipoprotein receptor intronic variants affecting residues >±1 or 2

<table>
<thead>
<tr>
<th>Intron</th>
<th>Variant</th>
<th>Pathogenicity classification</th>
<th>In vitro evidence</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>c.190+2_190+3dup</td>
<td>Clearly pathogenic</td>
<td>cDNA sequencing revealed use of novel splice donor site (GT) at c.190+2_190+3, resulting in inclusion 2 bp from 5’ end of intron 2, a frame shift and premature truncation of the peptide—p.(Leu64Cysfs*34).15</td>
<td>Use of novel splice site</td>
</tr>
<tr>
<td>3</td>
<td>c.313+2dup</td>
<td>Likely to be pathogenic</td>
<td>RT-PCR sequencing revealed skipping of exon 3, resulting in an in-frame deletion—p.(Leu64_Pro105delinsSer).16</td>
<td>Inactivation of wild-type splice site</td>
</tr>
<tr>
<td>3</td>
<td>c.313+5G&gt;A</td>
<td>Likely to be pathogenic</td>
<td>cDNA sequencing revealed skipping of exon 3, resulting in an in-frame deletion—p.(Leu64_Pro105delinsSer).16</td>
<td>Inactivation of wild-type splice site</td>
</tr>
<tr>
<td>3</td>
<td>c.313+6T&gt;C</td>
<td>Likely to be pathogenic</td>
<td>cDNA sequencing revealed skipping of exon 3, resulting in an in-frame deletion—p.(Leu64_Pro105delinsSer).16</td>
<td>Inactivation of wild-type splice site</td>
</tr>
<tr>
<td>6</td>
<td>c.941-12G&gt;A</td>
<td>Likely not pathogenic</td>
<td>RT-PCR sequencing revealed abnormal splicing of intron 6 (no further information provided by authors).37</td>
<td>Unknown</td>
</tr>
<tr>
<td>7</td>
<td>c.1061-8T&gt;C</td>
<td>Clearly not pathogenic</td>
<td>RT-PCR sequencing reveals normal splicing of intron 7.34</td>
<td>Normal splicing</td>
</tr>
<tr>
<td>8</td>
<td>c.1186+5G&gt;A</td>
<td>Clearly pathogenic</td>
<td>RT-PCR sequencing revealed inclusion of intron 8, resulting in a frame shift and premature truncation of the peptide—p.(Gly396fs*26).30</td>
<td>Use of novel splice site</td>
</tr>
<tr>
<td>9</td>
<td>c.1359-31_1359-23delinsCGGCT</td>
<td>Clearly pathogenic</td>
<td>mRNA sequencing reveals retention of intron 9 and evidence that two additional transcripts are produced using cryptic splice sites in exon 10. Due to removal of the invariant A at consensus splicing branch site in intron 9.39</td>
<td>Use of novel splice sites</td>
</tr>
<tr>
<td>9</td>
<td>c.1359-5C&gt;G</td>
<td>Clearly pathogenic</td>
<td>mRNA sequencing revealed retention of intron 9, resulting in a frame shift and premature truncation of the peptide—p.(Ser65Argfs*1).34</td>
<td>Inactivation of wild-type splice site</td>
</tr>
<tr>
<td>10</td>
<td>c.1586+5G&gt;A</td>
<td>Likely to be pathogenic</td>
<td>RT-PCR sequencing revealed alternate splicing resulting in two abnormal mRNAs: (a) skipping of exon 10 and (b) inclusion of 22 novel amino acids from the activation of a cryptic splice site.30</td>
<td>Inactivation of wild-type splice site and use of cryptic splice site</td>
</tr>
<tr>
<td>12</td>
<td>c.1845+11C&gt;G</td>
<td>Likely to be pathogenic</td>
<td>RT-PCR revealed that approximately half the transcripts use the novel splice site, resulting in inclusion of 11 bp from 5’ end of intron 12 into the transcript, a frame shift and premature truncation of the peptide—p.(Glu615fs*3).31</td>
<td>Normal splicing and use of novel splice site</td>
</tr>
<tr>
<td>14</td>
<td>c.2140+5G&gt;A</td>
<td>Unlike likely to be pathogenic</td>
<td>RT-PCR sequencing reveals normal splicing of intron 14.44</td>
<td>Normal splicing</td>
</tr>
<tr>
<td>14</td>
<td>c.2140+86C&gt;G</td>
<td>Likely to be pathogenic</td>
<td>RT-PCR sequencing revealed creation of a novel splice donor site, resulting in inclusion of 81 bp from 5’ end of intron 14 and a 27aa insertion into peptide p.(Thr713_Glu714ins27). Resulting peptide fails to leave the ER.44</td>
<td>Use of novel splice site</td>
</tr>
<tr>
<td>15</td>
<td>c.2312-3C&gt;A</td>
<td>Likely to be pathogenic</td>
<td>cDNA sequencing reveals skipping of exon 16, predicted to result in an in-frame deletion—p.(Ala721_Ile726del).36</td>
<td>Inactivation of wild-type splice site</td>
</tr>
</tbody>
</table>
would thereby result in reducing the numbers of membrane-bound LDLR molecules for LDL-C clearance.

It is clearly of great importance to be able to assess whether variants identified in clinical settings or as incidental findings in genomics projects are pathogenic or not. Although 93% (n=1595) of LDLR variants in the current upgrade of the database have been assigned to an ACGS pathogenicity category, 7% (n=112) remain as VUS. It is hoped that as more information becomes available from in vitro functional studies, the development of additional in silico tools and from the various genomics studies, it will be possible to determine the pathogenicity of these variants; indeed, the classification of other variants may also change as our knowledge increases.

In conclusion, the LDLR database provides a valuable resource to the research and clinical communities. We would like to encourage the ethos of data sharing and open access to resources, and so we urge researchers and clinicians to submit their variant data to the database via a link on the homepage (http://databases.lovd.nl/shared/genes/LDLR). While every effort has been made to ensure the accuracy of the data in this database, we accept that errors may have occurred and so we would be grateful if you could please inform us of any that you find.

Contributors SL: principal in collection and analysis of variants, and of manuscript preparation; MF: collation and analysis of variants, contributor in manuscript preparation; RW: responsibility for formatting data to be loaded onto the LOVD platform and responsible for maintenance and development of this database platform; SEH: project lead, giving direct guidance, advice and support for the study and assistance in manuscript preparation.

Funding British Heart Foundation BHF PG08/008.

Competing interests SEH holds a chair funded by the British Heart Foundation, and SEH and RW are supported by the BHF (PG08/008) and by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. MF is supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES
Genotype-phenotype correlations


