

ORIGINAL ARTICLE

Analysis of LDLR mRNA in patients with familial hypercholesterolemia revealed a novel mutation in intron 14, which activates a cryptic splice site

Mari Ann Kulseth¹, Knut Erik Berge¹, Martin Prøven Bogsrud² and Trond P Leren¹

Familial hypercholesterolemia (FH) is caused by a defective low-density lipoprotein receptor (LDLR), and > 1000 mutations in *LDLR* have been identified. However, in some patients with clinically defined FH, no mutation can be detected within the exons and adjacent intronic segments of the *LDLR*. We have analyzed RNA extracted from blood samples of patients with clinically defined FH and identified an aberrantly spliced mRNA containing an 81-bp insert from intron 14. The aberrant splicing was caused by a novel intronic mutation, c.2140+86C>G, which activated a cryptic splice site. Although the cryptic splice site does not completely surpass the normal splice site, the mutation was found to cosegregate with high cholesterol levels in a family, which supports the notion that c.2140+86C>G causes FH. The insertion of 81 bp in *LDLR* mRNA encodes an in-frame insertion of 27 amino acids in the LDLR. However, the insertion was found to hamper LDLR activity by preventing the receptor from leaving the endoplasmic reticulum, probably because of misfolding of the protein. In patients with clinically defined hypercholesterolemia, despite normal results from sequencing of exonic regions of the *LDLR* gene, characterization of the *LDLR* mRNA might identify the underlying genetic defect.

Journal of Human Genetics (2010) 55, 676–680; doi:10.1038/jhg.2010.87; published online 12 August 2010

Keywords: c.2140+86; cryptic splice site; familial hypercholesterolemia; FH; LDLR

INTRODUCTION

Dominantly inherited familial hypercholesterolemia (FH) is a frequent disorder of lipid metabolism caused by mutations in the low-density lipoprotein receptor (LDLR) gene. More than 1000 different mutations in the *LDLR* gene have been found to be associated with FH.¹ These mutations have been identified by genomic analysis, such as DNA sequencing of PCR products,² multiplex ligation-dependent probe amplification³ and single-strand conformation polymorphisms.⁴ Mutations introduce pre-mRNA splicing defects, premature stop codons or amino-acid substitutions in the LDLR protein, rendering the protein unable to fold correctly, bind LDL, be endocytosed or recycled back to the cell surface.

The gene encoding LDLR spans ~45 kb and is localized on chromosome 19p13.1–13.3. During pre-mRNA splicing, 18 exons are joined together to create an mRNA of 5.3 kb. Genetic analysis of the *LDLR* gene has mainly focused on exons and the flanking intron regions. Primers used for PCR amplification of the exons are usually located between 50 and 150 bp upstream and downstream of the exon–intron boundaries.

The prediction of the pathogenic role of sequence alterations in exons is straightforward for nonsense mutations, frameshift mutations or rearrangements that introduce premature stop codons, which cause synthesis of truncated proteins or most likely degradation of mRNA

by nonsense-mediated decay.⁵ The LDLR is highly conserved and seems to be vulnerable to missense mutations, deletions or insertions, which change the amino-acid composition of the protein. In the Human Gene Mutation database,¹ amino-acid substitutions at approximately half of the 860 residues that constitute the LDLR have been reported to be associated with FH. Whether the association is the result of amino-acid substitution or is caused by deleterious effects on the LDLR mRNA composition has not been systematically analyzed. Correct pre-mRNA splicing is a prerequisite for a functional LDLR. Pre-mRNA splicing defects might cause exon skipping, intron inclusions or activation of cryptic splice sites. Prediction of the pathogenic role of sequence alterations that might disrupt pre-mRNA splicing is less straightforward.

Mutations throughout the *LDLR* exons and introns might cause pre-mRNA splicing defects, and might account for some unresolved cases of clinically defined hypercholesterolemia. Synonymous mutations and intronic mutations are likely to be overlooked during genetic analysis of the *LDLR* gene or are likely to be undetected, as point mutation screening of all intronic sequences is not yet feasible. Direct analysis of mRNA from a patient is the most reliable method for studying pre-mRNA splicing; as LDLR is expressed in leukocytes,⁶ total RNA isolated from blood samples can be used as the RNA source.

¹Department of Medical Genetics, Oslo University Hospital, Rikshospitalet, Oslo, Norway and ²Lipid Clinic, Oslo University Hospital, Rikshospitalet, Oslo, Norway
Correspondence: Dr MA Kulseth, Department of Medical Genetics, Oslo University Hospital, Forskningsveien 2b, Oslo, Rikshospitalet, N-0027 Oslo, Norway.
E-mail: mari.ann.kulseth@oslo-universitetssykehus.no

Received 20 May 2010; revised 15 June 2010; accepted 21 June 2010; published online 12 August 2010

We have established a reverse transcription (RT)-PCR screening assay for patients with clinically defined FH without any mutations detected in the *LDLR* gene in routine DNA sequencing and multiplex ligation-dependent probe amplification. We have identified a mutation located at position c.2140+86, which activates a cryptic splice site in intron 14 of the *LDLR* gene. The mutation was located outside the region covered by DNA sequencing. c.2140+86C>G causes an inclusion of 81 nucleotides that encode an in-frame insertion of 27 amino acids. The mutation was found to cosegregate with high cholesterol levels in a family.

MATERIALS AND METHODS

Patient samples

Patients with clinically defined FH and total cholesterol levels above 10 mmol l⁻¹ and triglyceride levels below 5 mmol l⁻¹ were subjected to screening by RT-PCR. No mutations were detected in these patients by routine genomic analysis of the *LDLR*. Moreover, neither the R3500Q mutation in the apolipoprotein B gene nor a PCSK9 gain-of-function mutation had been identified in these patients. Family members of the index patient were recruited to undergo segregation analysis. Written informed consent for molecular genetic analysis was obtained from all participants.

RNA isolation, RT-PCR

Blood samples were collected into PAXgene Blood RNA tubes and total RNA was extracted using the PAXgene Blood RNA kit (Qiagen, Valencia, CA, USA). RT-PCR was performed using 200 ng of total RNA and the Qiagen OneStep RT-PCR kit (Qiagen). The primers used are described in the study by Tveten *et al.*⁶ and designed to produce overlapping PCR products corresponding to exons 1–8, 3–10, 7–14, 11–17 and 13–18. The thermal cycling conditions were 50 °C for 30 min, 95 °C for 15 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s. The amplified products were analyzed on Criterion 5% TBE gels (Bio-Rad, Hercules, CA, USA), which were stained with ethidium bromide. To characterize the bands, the amplified products were cloned into the pCRII-TOPO vector (Invitrogen, Carlsbad, CA, USA). Subcloned fragments were reamplified using relevant primers and sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and analyzed on a 730 Genetic Analyzer (Applied Biosystems).

DNA analysis

DNA was extracted from EDTA-containing blood using a DNA extractor (Mole Genetics, Oslo, Norway). A DNA fragment consisting of *LDLR* exons 13 and 14 was amplified in a PCR using AmpliTaqGold, the upstream primer 5'-TGGCCTGTGTCTCATCCAGTGT-3' and the downstream primer 5'-GGCTCCTGTGCAGAAATCAGATCACC-3'. The thermal cycling conditions were 95 °C for 10 min, followed by 36 cycles of 95 °C for 30 s, 61 °C for 30 s and 72 °C for 60 s. The PCR product was purified using ExoSAP-IT (USB Corporation, Cleveland, OH, USA) according to the manufacturer's instructions. Version 3.1 of the BigDye terminator cycle-sequencing kit (Applied Biosystems) and the downstream primer were used for the sequencing reactions according to the manufacturer's instructions. The sequencing products were run on a Genetic Analyzer 3730 (Applied Biosystems) and analyzed using Seqscape version 2.5 software (Applied Biosystems).

Site-directed mutagenesis

The QuikChange XL Site-directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) was used to insert 81 bp of the 5'-end of *LDLR* intron 14 between exons 14 and 15 in the pcDNA4-LDLR.⁷ To generate pcDNA4-LDLRins81 bp, the oligonucleotide 5'-CATGAGGAGCTGCCTCACAGGTGTGGCACACGCC TTGTTTCTGCGTCTGTGTCTCCAAGTCCCTCCTGAGCCTCTCTC TGCTCATCTGTCAAATGGAGGCTGAGGCTGCAGTGGCC-3' (81 bp insertion underlined) and its complementary oligonucleotide were obtained and PAGE purified by Medprobe (Oslo, Norway). The integrity of the pcDNA4-LDLRins81 bp plasmid was confirmed by DNA sequencing.

Computer analysis

The strength of the normal and cryptic donor splice site was calculated using algorithms developed by Burge laboratory,⁸ including MAXENT (maximum entropy), MDD (multiple dependence decomposition), a first-order MM (Markov's model) and a WMM (weight matrix model) using their web-based tool MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html), which scores short input sequences (9mer) containing a single splice site.

To screen for transposable elements in the *LDLR* DNA sequence, the RepeatMasker program (<http://www.repeatmasker.org>) created by Smit (AFA Smit *et al.*, unpublished data) was used.

Cell culture and transient transfection

T-REx CHO cells (Invitrogen) stably expressing a tetracycline repressor were maintained in complete medium: Ham's F-12 medium (Euroclone) supplemented with 10% (v/v) fetal calf serum (Euroclone, West York, UK), 50 Units ml⁻¹ penicillin (Euroclone), 50 µg ml⁻¹ streptomycin (Euroclone) and 10 µg ml⁻¹ blasticidin (Invitrogen) in a humidified atmosphere (37 °C, 5% CO₂). T-REx CHO cells were transiently transfected with pcDNA4-LDLR or pcDNA4-LDLRins81bp using Lipofectamine 2000 (Invitrogen). The cells were incubated with transfection mixture for 6 h at 37 °C; thereafter, the medium was changed to complete medium and the cells were incubated for another 48 h, the last 24 h with 1 µg ml⁻¹ tetracycline (Invitrogen).

Western blot analysis

Cells were lysed and subjected to western blot analysis as described in the study by Tveten *et al.*⁹ The membrane was cut in two. The upper part was immunostained with rabbit polyclonal anti-LDLR (Fitzgerald Industries Intl, North Acton, MA, USA; diluted 1:500 in 5% (w/v) nonfat dry milk (Bio-Rad) and 0.05% (v/v) Tween-20 (Sigma-Aldrich, St Louis, MO, USA)), whereas the lower part was immunostained with rabbit anti-tubulin (Nordic BioSite, Täby, Sweden; diluted 1:1000). After incubation at room temperature for 1 h, both were counterstained with horseradish peroxidase-conjugated anti-rabbit antibodies (diluted 1:1000; GE Healthcare, Waukesha, WI, USA). Signals were developed using SuperSignal West Dura Extended Duration Substrate (Pierce Biotechnology, Rockford, IL, USA), and ChemiDoc XRS (Bio-Rad) was used to detect the signals. The concentrations of the antibodies were optimized to achieve low background and a dose-dependent increase in signal intensity.

RESULTS

Blood samples collected from 30 unrelated patients with clinically defined hypercholesterolemia, but without any mutations detected by standard DNA analysis, were recruited for RNA analysis. Overlapping RT-PCR was performed as described by Tveten *et al.*⁶ RT-PCR amplification of RNA from an index patient revealed several longer amplicons in the RT-PCR of exons 11–17 and exons 13–18, with one major product being ~ 80 bp longer than the normal amplicon (Figure 1). Sequencing of RT-PCR products revealed that the major product contained an insertion of 81 bp from the 5'-end of intron 14. To identify the underlying mutation, we performed a PCR amplification of exons 13 and 14 from the patient's DNA using a downstream primer located further downstream in intron 14 (c.2140+220). DNA sequencing of the PCR product detected a novel intronic mutation c.2140+86C>G. This mutation is located outside the region covered by DNA sequencing of the *LDLR* in our laboratory, where the 5'-end of the downstream primer for amplification of exons 13 and 14 was located at c.2140+89.

The mutation c.2140+86C>G introduces a cryptic splice site by exchanging a cytosine to a guanine in the fifth position of the novel splice site. The substitution increases the similarity with the donor splice site consensus and causes an increase in the splice site score of the sequence located at positions +79 to +87 to the level of the normally used donor splice site located at the 5'-end of intron 14 (Table 1).

Segregation analysis was performed in the family of the index patient. Of 23 family members, 12 were shown to be heterozygous

for the mutation and c.2140+86C>G was found to cosegregate with high cholesterol levels (Figure 2a). Mutation carriers had a total cholesterol level of 8.5 mmol l⁻¹, compared with 5.3 mmol l⁻¹ in noncarriers, which represented a significant increase in total cholesterol ($P=0.000013$) (Figure 2b).

Among 550 index patients referred to us for genetic testing, c.2140+86C>G has been detected in 3 other index patients. In the family one of these patients, the mutation was found in six of seven tested members (Figure 3) and they all had LDL cholesterol levels above the 97th percentile. The aberrant splicing, causing the insertion of 81 bp from the 5'-end of intron 14, was confirmed in one member (II.2) of this family (results not shown).

Generation of a cryptic splice site might completely surpass the use of the normal splice site, or both may be used in various degrees. To establish the degree of detrimental effect caused by c.2140+86C>G on the splicing process, the RT-PCR products generated from II.2 (Figure 3) were cloned in a pCRII-TOPO vector to reveal whether

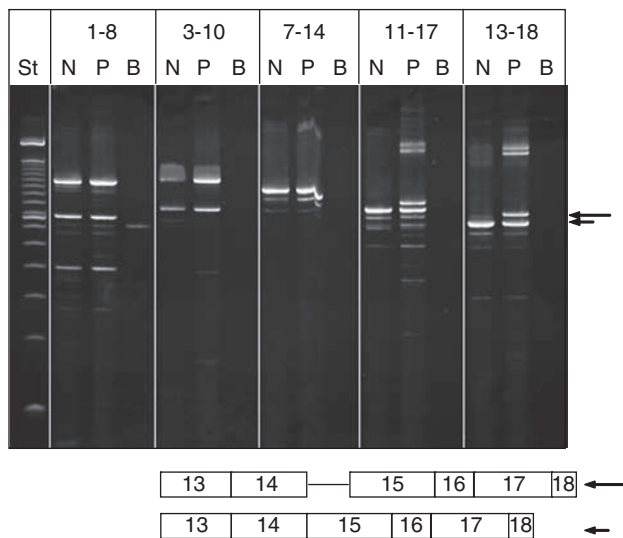


Figure 1 RT-PCR analysis of LDLR mRNA. Overlapping fragments were produced by using five different primer pairs spanning exons 1–8, 3–10, 7–14, 11–17 and 13–18. The left lane contains the 1 kb Plus DNA ladder obtained from Invitrogen. N, P and B denote normal, patient and blank, respectively. RT-PCR, reverse transcription PCR; LDLR, low-density lipoprotein receptor.

the mutant allele could give rise to correctly spliced LDLR mRNA. II.2 is heterozygous for a common synonymous single-nucleotide polymorphism in exon 13, c.1959T>C, and the linkage between this single-nucleotide polymorphism and the insertion of 81 bp was established by sequencing individual clones transfected with the pCRII-TOPO vector containing the RT-PCR amplicons. A total of 160 clones were isolated and subjected to DNA sequencing. Of them, 90 had a thymidine in position c.1959 and no insert, whereas 70 had a cytosine in position c.1959; of these 70 clones, 46 had an 81 bp insert from the 5'-end of intron 14, whereas 24 had no insert. The mutant allele contributed with 21% of normal transcripts. Thus, the allele containing the mutation c.2140+86C>G mainly gives rise to aberrantly spliced mRNA, but a significant production of normal mRNA might also be produced.

The inserted 81 bp from the 5'-end of intron 14 encodes an in-frame insertion of 27 amino acids. These amino acids are inserted at the end of EGF-repeat C in the LDLR. To examine the effect on the protein, 81 bp was inserted between exons 14 and 15 in the cDNA encoding human wild-type LDLR. Wild-type LDLR cDNA and LDLRins81 bp cDNA were transfected into CHO cells and cell lysates were analyzed on a western blot. As seen in Figure 4, the LDLR with the 27 amino-acid insert did not undergo oligosaccharide modifications in the Golgi apparatus, which increase the apparent molecular mass of the wild-type LDLR from ~120 to 160 kDa.¹⁰

DISCUSSION

By performing RT-PCR analysis on mRNA from individuals with clinically defined FH, we discovered a novel mutation located within intron 14 of the *LDLR*. The mutation, c.2140+86C>G, introduces a cryptic splice site that results in an in-frame insertion of 81 bp from the 5'-end of intron 14 in the LDLR mRNA. The cryptic splice site does not completely surpass the normal splice site. A small amount of normal transcript is produced from the mutant allele. However, the majority of transcripts produced from the mutant allele are abnormal and cosegregation of the mutation with high cholesterol levels in a family supports the notion that c.2140+86C>G causes FH. The insertion of 81 bp in the LDLR mRNA encodes an in-frame insertion of 27 amino acids in the LDLR. However, the insertion causes a retention of the receptor in the endoplasmic reticulum, probably caused by misfolding of the protein.

RT-PCR analysis on RNA from patients with a definite diagnosis, but without any identified disease-associated mutations after routine DNA sequencing, might solve some unsolved cases. Engel *et al.*¹¹

Table 1 (a) Normal and cryptic splice sites in LDLR intron 14 compared with the consensus donor splice site and (b) MaxEntScan scoring of the splice sites

| (a) | |
|------------------------|---|
| Normal splice site | CTGCCTCACAGgtgtggcacacgc |
| Intron 14 (+79 to +87) | CTGTCAAATGGgtacctaaggtc |
| Cryptic splice site | CTGTCAAATGGgtacgtcaaggtc (mutation in bold) |
| Consensus | CAGGTRAGT (r=a or r=g) |

| | Splice site | MAXENT | MDD | MM | WMM |
|------------------------|-------------|--------|-------|------|------|
| (b) | | | | | |
| Normal splice site | CAGGTGTGG | 6.34 | 10.38 | 5.59 | 7.81 |
| Intron 14 (+79 to +87) | TGGGTACCT | 1.54 | 5.68 | 3.92 | 1.93 |
| Cryptic splice site | TGGGTACGT | 7.70 | 11.88 | 8.49 | 5.77 |

Abbreviations: LDLR, low-density lipoprotein receptor; MAXENT, maximum entropy; MDD, multiple dependence decomposition; MM, Markov's model; WMM, weight matrix model.

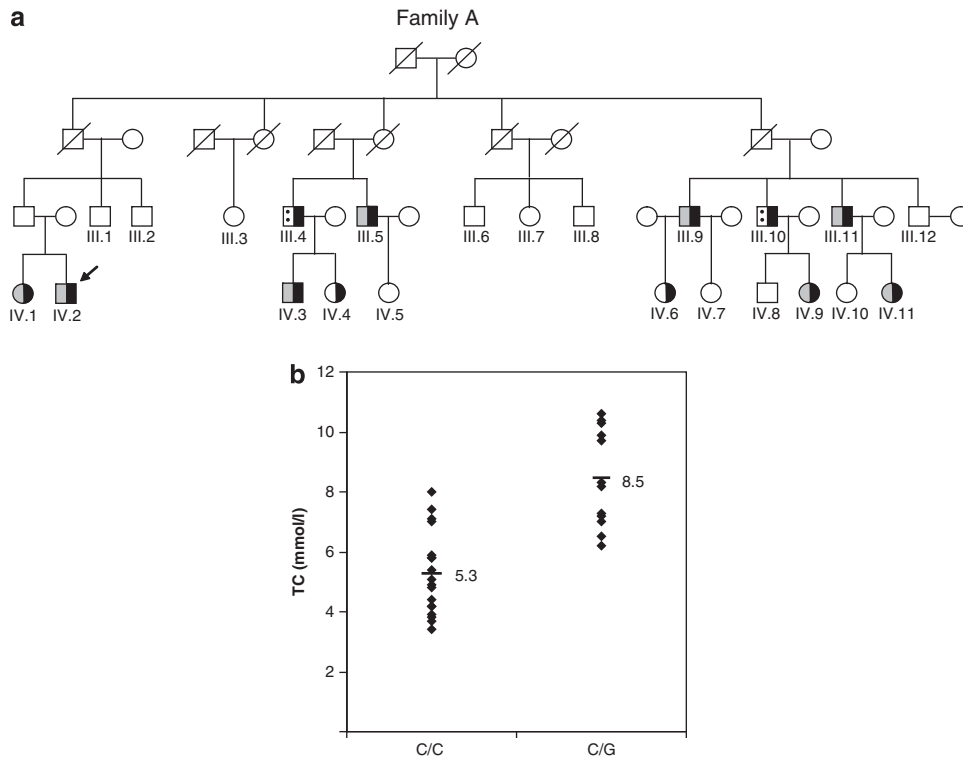


Figure 2 (a) Pedigree of family A. The index patient is indicated by an arrow. Black symbols, heterozygous for the *LDLR* mutation c.2140+86C>G; gray symbols, LDL cholesterol levels above the 97.5th percentile; dotted symbols, total cholesterol levels above 9.9 mmol l^{-1} (LDL cholesterol unknown). Family members who have been analyzed are denoted by a generation number. (b) Total cholesterol (TC) levels in family A. Homozygous for the normal allele: c.2140+86C/C; heterozygous for the mutant allele: c.2140+86C/G. *LDLR*, low-density lipoprotein receptor.

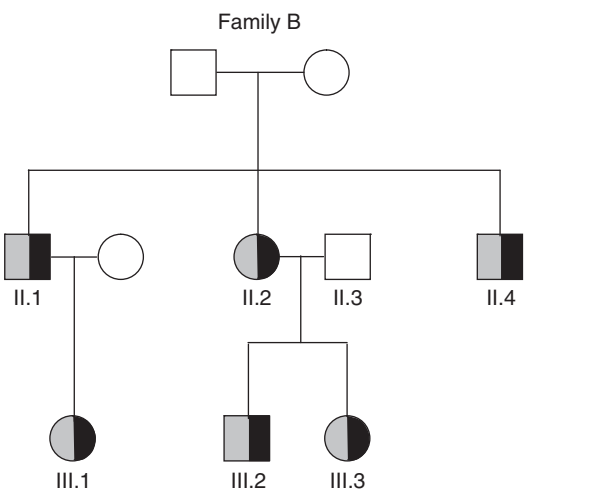


Figure 3 Pedigree of family B. Black symbols, heterozygous for the *LDLR* mutation c.2140+86C>G; gray symbols, LDL cholesterol levels above the 97.5th percentile. Family members who have been analyzed are denoted by a generation number. *LDLR*, low-density lipoprotein receptor.

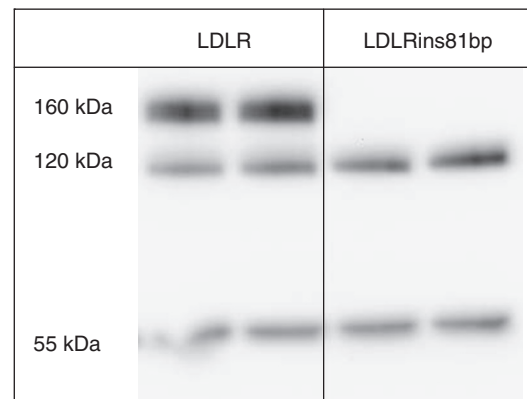


Figure 4 Western blot analysis of cell lysates from TRex-CHO cells transfected in duplicate with wild-type *LDLR* or *LDLR*ins81 bp. Cell lysates were prepared and equal amounts of protein were subjected to SDS-PAGE and western blot analysis, using an antibody against *LDLR* (upper part) and an antibody against tubulin (lower part). The experiment was conducted twice with similar results. *LDLR*, low-density lipoprotein receptor.

analyzed the RNA from five consecutive ornithine transcarbamylase-deficient patients in whom genomic ornithine transcarbamylase analysis had failed to show a disease-associated mutation. Ornithine transcarbamylase pre-mRNA splicing aberrations caused by deep intronic mutations were found in all five patients. As genetic testing improves and becomes more widely requested, RT-PCR analysis will also be important to interpret the clinical significance of unclassified sequence variants.¹² To perform RT-PCR analysis, an

appropriate RNA source has to be available. Peripheral blood is an easily accessible source. However, not all genes are expressed in leukocytes.¹³ The *LDLR* is expressed in leukocytes, although it fulfills its most important role in the liver, which is clearing lipoproteins from circulation.

Spurdle *et al.*¹⁴ emphasize the importance of determining the relative amounts of aberrantly spliced products generated by a variant allele to establish clinical relevance. Mutation

```

LDLR 47 CCTC-CTGAGCCTCTCTCTGCTCATCTGTCAAATGGG-----TACCT
          -      vv i v          v          -----
MIRb 115 CCTCTCTGAGCCTCAGTTTCCTCATCTGTAAAAATGGGATAATAAATACCT

LDLR 88 -----CAAGTTCGTTGTAAAGGACTCA-TGAG
          ----- i i i i v -
MIRb 165 ACCTCGCAGGGTTGTTGTGAGGATTAATGAG

```

Figure 5 Alignment of intron 14 (c.2140+47 to +112) and MIRb created by Repeat Masker (<http://www.repeatmasker.org>). Deletions, transversions and transitions are shown as '-', 'v' and 'i,' respectively. The insertion found in the aberrantly spliced LDLR mRNA is highlighted in gray and the mutation in bold. LDLR, low-density lipoprotein receptor.

c.2140+86C>G gives rise to several incorrect transcripts. However, in leukocytes, a significant amount (21%) of the normally spliced transcripts was produced from the mutant allele, indicating that the c.2140+86C>G mutant allele is not a null allele. However, the mutation cosegregated with high cholesterol levels in a family, which is strong evidence that c.2140+86C>G causes FH.

Alternative splicing is known to be tightly regulated in different tissues and developmental stages because of tissue-specific expression and/or activation of splicing factors.¹⁵ It is therefore likely that the tissue/developmental variation in splicing factors might also affect the outcome of a mutation and reduce or amplify the degree of aberrant splicing.¹⁴ The relative amount of aberrantly spliced LDLR mRNA was analyzed in blood samples. However, whether leukocytes are representative for hepatic tissue in this context is uncertain.

Inclusion of intronic sequences in mRNA has been found to be an important cause of hereditary diseases, and the inclusion is often caused by activation of splice sites in transposable elements located within introns.¹⁶ The mutation c.2140+86C>G is located within a region containing a segment of a mammalian interspersed repeat (MIRb). The match obtained by Repeat Masker started at c.2140+47 and extended to c.2140+112 (Figure 5). Vorechovsky *et al.*¹⁶ found that transposable elements contributed to cryptic exon activation in ~50% of the mutation-induced intron inclusions in patients with hereditary diseases. Overall, 60% of the LDLR gene consists of transposable elements and, with the exception of the 3'-untranslated region, they are located within introns. Intronic mutations within these elements might be an underappreciated cause of FH.

By RT-PCR analysis of RNA from blood samples of individuals with clinically definite hypercholesterolemia, we have identified a novel

intronic mutation causing FH. When routine DNA sequencing fails to identify a disease-causing mutation, the next step in diagnostic investigation should be RNA analysis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Inger Esther Nossen and Nina Engstrøm for excellent technical assistance.

- 1 Stenson, P., Mort, M., Ball, E., Howells, K., Phillips, A., Thomas, N. *et al.* The Human Gene Mutation Database: 2008 update. *Genome Med.* **1**, 13 (2009).
- 2 Hobbs, H. H., Brown, M. S. & Goldstein, J. L. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum. Mutat.* **1**, 445–466 (1992).
- 3 Holla, O. L., Teie, C., Berge, K. E. & Leren, T. P. Identification of deletions and duplications in the low density lipoprotein receptor gene by MLPA. *Clin. Chim. Acta* **356**, 164–171 (2005).
- 4 Sozen, M., Whittall, R. & Humphries, S. E. Mutation detection in patients with familial hypercholesterolaemia using heteroduplex and single strand conformation polymorphism analysis by capillary electrophoresis. *Atheroscler. Suppl.* **5**, 7–11 (2004).
- 5 Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Sauliere, J., Wittkopp, N. & Izaurralde, E. mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett.* **581**, 2845–2853 (2007).
- 6 Tveten, K., Ranheim, T., Berge, K. E., Leren, T. P. & Kulseth, M. A. Analysis of alternatively spliced isoforms of human LDL receptor mRNA. *Clin. Chim. Acta* **373**, 151–157 (2006).
- 7 Sørensen, S., Ranheim, T., Bakken, K. S., Leren, T. P. & Kulseth, M. A. Retention of mutant low density lipoprotein receptor in endoplasmic reticulum (ER) leads to ER stress. *JBC* **281**, 468–476 (2006).
- 8 Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **11**, 377–394 (2004).
- 9 Tveten, K., Holla, O. L., Ranheim, T., Berge, K. E., Leren, T. P. & Kulseth, M. A. 4-Phenylbutyrate restores the functionality of a misfolded mutant low-density lipoprotein receptor. *FEBS J.* **274**, 1881–1893 (2007).
- 10 Tolleshaug, H., Goldstein, J. L., Schneider, W. J. & Brown, M. S. Posttranslational processing of the LDL receptor and its genetic disruption in familial hypercholesterolemia. *Cell* **30**, 715–724 (1982).
- 11 Engel, K., Nuoffer, J. M., Muhlhausen, C., Klaus, V., Largiadier, C. R., Tsiakas, K. *et al.* Analysis of mRNA transcripts improves the success rate of molecular genetic testing in OTC deficiency. *Mol. Genet. Metab.* **94**, 292–297 (2008).
- 12 Baralle, D., Lucassen, A. & Buratti, E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep.* **10**, 810–816 (2009).
- 13 Sullivan, P. F., Fan, C. & Perou, C. M. Evaluating the comparability of gene expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **141B**, 261–268 (2006).
- 14 Spurdle, A. B., Couch, F. J., Hogervorst, F. B., Radice, P. & Sinilnikova, O. M. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum. Mutat.* **29**, 1304–1313 (2008).
- 15 Wang, Z. & Burge, C. B. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802–813 (2008).
- 16 Vorechovsky, I. Transposable elements in disease-associated cryptic exons. *Hum. Genet.* **127**, 135–154 (2010).