

Title page

Manuscript title: Familial hypercholesterolemia: a single-nucleotide variant (SNV) in mosaic at the Low density lipoprotein receptor (*LDLR*).

Running title: mosaicism at the *LDLR*

Authors: Sonia Rodríguez-Nóvoa¹, Carmen Rodríguez-Jiménez¹, Concepción Alonso², Lara Rodríguez-Laguna³, Gema Gordo³, Víctor Martínez-Glez³ and Iluminada García Polo⁴.

Keywords: *LDLR*, hypercholesterolemia, mosaicism

Authors affiliation:

¹Department of Genetics of Metabolic Diseases and ³Vascular Malformations Section, Institute of Medical & Molecular Genetics (INGEMM), Hospital Universitario La Paz, IdiPAZ, Madrid, Spain; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER, U753); ² Department of Genetics, Hospital de La Princesa, Madrid, Spain; ⁴ Department of Internal Medicine, Hospital de la Princesa, Madrid, Spain.

Corresponding author:

Sonia Rodríguez-Nóvoa
Department of Genetics of Metabolic Diseases-INGEMM
Hospital Universitario La Paz
Paseo de la Castellana 261, 28046
Madrid, Spain
Tel: +34 91 727 72 17 ext. 249
e-mail: soniamaria.rodriguez@salud.madrid.org

Conflicts of interest: all authors declare no conflicts of interest.

Financial support: the study was supported by grant FIS15/00122, FEDER, ISCIII-Instituto de Salud Carlos III, Spain

Abstract

Familial hypercholesterolemia is most frequently caused by genetic variants in the *LDLR* gene. Most of *LDLR* pathogenic variants are missense, followed by splicing and deletion/insertions variants. Mosaicism is a genetic condition in which an individual shows more than one clone of cells with different genotypes. The objective of this article was the molecular characterization of a patient with hypercholesterolemia.

Genetic analysis of DNA from peripheral blood and saliva was performed by NGS, sanger sequencing and pyrosequencing technologies. NGS analysis detected the pathogenic variant *LDLR:c.1951G>T:p.(Asp651Tyr)* in 9%-12% of reads. The presence of the variant was confirmed by pyrosequencing analysis. The variant found was functional characterized using an in vitro model (CHO-IIdIA7 cells). Activity and expression of cell surface LDLR were measured by flow cytometry. Colocalization LDLR-Dil-LDL was detected by immunofluorescence. The LDLR activity showed 80% uptake, 50% binding and 53% expression of cell surface LDLR regarding wild type.

Herein, we report the first case of a mosaic single nucleotide variant affecting the *LDLR* gene in a patient with familial hypercholesterolemia. As has been described for other pathologies, mosaicism could be underestimated in FH and its detection will improve with the introduction of NGS technologies in the diagnostic routine.

1. Introduction

Different types of pathogenic variants in genes can cause genetic diseases. Most of them are missense variants that affect the coding region of genes. This type of variants consists in a change of aminoacid that alters or impairs the normal function of the resultant protein, therefore causing the disease phenotype. Until recently it was considered that pathogenic variants were always present in all the cells of an individual. However, it is increasingly evident that all individuals are actually complex mosaics composed of multiple genotypes acquired post-zygotically, and that in many cases this can produce a pathogenic effect. In the later cases, the variant must be present in the appropriate cell types and achieve a threshold in the number of cells affected so that the disease can be clinically detected. As a general rule, the earlier during embryonic development the mutation occurs, the greater the number of affected cells, and the greater the possibility of developing the disease. In diagnostic terms, not detecting variants in low mosaics, rather common with classical diagnostic methods, can lead to erroneous conclusions about the results of molecular tests, and therefore in the treatment and follow-up of patients [Forsberg et al. 2017].

Familial hypercholesterolemia (FH, MIM: 143890) is a genetic disorder characterized by an elevation in total cholesterol (TC) and low-density lipoproteins-cholesterol (LDL-C) leading to premature cardiovascular disease [Goldstein et al. 2001]. The worldwide prevalence of FH has been estimated in 1/200-300 [de Ferranti et al. 2016; Nordestgaard et al. 2013] and the main cause is the presence of pathogenic variants in the low density lipoprotein receptor gene (*LDLR*, MIM: 606945) which is involved in the LDL uptake. Up to date, more than 1700 variants have been reported in the *LDLR* gene and around 81% were classified as pathogenic or likely pathogenic [Leigh et al. 2017]. Almost all types of variants have been

described in *LDLR* being the missense the most frequent ones. Splicing variants, small and gross deletions and insertions have been described as well.

The increase in the number of reported variants in *LDLR* is due to the implementation of screening strategies to early detection of patients with FH and to the development of new molecular strategies. Conventional method for molecular diagnosis of FH is Sanger sequencing but due to its cost and time consuming, other methodologies have been developed such as the arrays-based methods [Tejedor et al. 2005]. The arrays methods are tailored to the most frequent mutations but this is a limitation because the less frequent or new variants are not detected which is an important methodology limitation. With the introduction of new technologies such as next generation sequencing (NGS), the success of the molecular diagnosis rate has increased. Despite this fact, a great number of patients remain without genetic confirmation. Traditionally, most of these patients are further classified as probably having a polygenic hypercholesterolemia.

Herein, we report a case of mosaicism in the *LDLR* gene in a patient with familial hypercholesterolemia.

2. Materials and methods.

2.1 Patients. The proband is a man 57 years old referred to our reference laboratory for genetic testing in 2016 because of a clinical suspicion of familial hypercholesterolemia. He has a personal history of hypercholesterolemia detected when he was 40, with the highest levels of cholesterol documented in 2008, CT: 282 mg/dL, LDL: 195 mg/dl. There is no history of tobacco or alcohol consumption. The patient had hypertension, currently under treatment. Different pathologies such as hypothyroidism, kidney disease, diabetes, or the consumption of drugs (progestogens, anabolic steroids) were discarded. At the age of 52, he

suffered acute myocardial infarction, needing percutaneous transluminal coronary angioplasty (PTCA), and implantation of a drug-eluting stent. At the age of 56, a coronary angiography showed severe stenosis in the first marginal obtuse.

As a family history, his father died at age of 77, probably due to cardiac insufficiency, the cholesterol levels are not known; his 93 years old mother is still alive, and she is not affected from dyslipidemia; and proband's sister shows normal levels of cholesterol. The patient has four children (33, 31, 30 and 27 years old). Both younger children had hypercholesterolemia detected at the age of 5.

Physical examination showed a height of 1.69 m, weight 74 kg, body-mass index (BMI) of 25.91; he had no arcus cornealis or xanthomas or xanthelasmas.

The patient was clinically classified as having a probably familial hypercholesterolemia according to Dutch Lipid Clinic Network Criteria (DLCNC), so the genetic study of familial hypercholesterolemia was performed.

2.2. Genetic analysis. The genomic DNA (gDNA) was extracted from blood using the Chemagen system and DNA extraction from whole saliva was carried out by Oragene™ commercial kit. Genetic analysis was performed by NGS using a customized panel of 198 genes. Preparation and exome enrichment steps were performed according to manufacturer's workflow (Roche Nimblegen) and it was sequenced using MiSeq system (Illumina, San Diego, California, USA). A subset of genes was chosen for analysis of hypercholesterolemia: *LDLR*, *APOB*, *PCSK9*, and *LDLRAP1*.

Multiplex ligation-dependent probe amplification (MLPA, MRC-Holland) was used for detection of large insertions/deletions in *LDLR* gene.

Sanger sequencing was used to confirm variants present in more than 15% of the reads in the deep sequencing data. Standard PCR and Sanger sequencing were performed using the 96-capillary ABI 3730xl ADN analyzer (Applied Biosystem, Foster, USA). The validation of variants detected in mosaics over 5% of the readings in the NGS was performed by the pyrosequencing technique: primers were designed using PyroMark software, and QIAGEN reagents and the Pyromark Q96 MD instrument (QIAGEN, USA) were used according to manufacturer's protocol.

2.3. Bioinformatic analysis.

Bioinformatic analysis was performed using algorithms developed by our bioinformatics unit. Sequences were mapped to the CRCh37/hg19 human reference sequence and data bases used for analysis were Human Gene Mutation Database (HGMD® Public) (www.biobaseinternational.com/hgmd) from BIOBASE Corporation; Online Mendelian Inheritance in Man (www.omim.org); Gene Tests (www.genetests.org). Variant annotation was carried out with Ensembl's variant Effect Predictor Tool and was based on the transcripts LDLR-ENST00000558518, APOB-ENST00000233242, PCSK9- ENST00000302118, LDLRAP1-ENST00000374338. The *in silico* predictors of pathogenicity used were CADD (Combined Annotation Dependent Depletion), Polyphen (Polymorphism Phenotyping), MutAssesor, Fasthmm, and Vest. Scores of conservation used were Gerp2, PhasCons, Phylop. The splicing predictors used were MaxEntScan, NNSplice, GeneSplicer and Human Splicing Finder. Variants with minor allele frequency >1% were excluded from further analysis. The files were uploaded in BAM format for analysis using Alamut software (Interactive Biosoftware).

2.4. Functional characterization of the variant LDLR:c.1951G>T:p.(Asp651Tyr)

Cell culture and transfection: Chinese hamster ovary (CHO) *ldlA7* cells (kindly provided by Dr. Monty Krieger, Massachusetts Institute of Technology, Cambridge, MA) were utilized as the LDLR-deficient cell line [Kingsley and Krieger et al. 1984] [Krieger et al. 1983]. The cells were maintained in Ham's F-12 supplemented (Gibco, Life Technologies, Carlsbad, CA, USA) with 5% heat-inactivated foetal bovine serum (FBS; Gibco, Life Technologies, Carlsbad, CA, USA), 100 U/ml penicillin, 100µg/ml streptomycin, 2 mmol/L L-glutamine, Normocin 50mg/ml (InvivoGen, Toulouse France). Adherent CHO-*ldlA7* cells were grown in monolayer culture into 75 cm² flasks and incubated at 37°C under 5% CO₂ atmosphere.

The construct was generated into the expression vector *LDLR* NM_000527 Human cDNA ORF clone Human, C-GFPspark®-tag (HG10231-ACG Sino-biological, Wayne, USA). Site directed mutagenesis was carried out with QuikChange lightning site-directed mutagenesis kit (Agilent Technologies) according to the manufacturer's instructions. The plasmid sequence after mutagenesis was validated by Sanger sequencing.

CHO-*ldlA7* cells were seeded in 24 well plates at 15 x 10⁴/well and when they reached about 70-90% confluent were transiently transfected. Briefly, Lipofectamine 3000 (Invitrogen™ ThermoFisher-Scientific, Waltham, USA) was mixed with 2,5 µg of the LDLR expression plasmid previously mentioned. The cells were transfected with empty vector, wild type (WT), variant of interest (*c.1951G>T*) and control, according to the manufacturer's instructions. The mix was added to cells and LDLR activity was assessed 48h after transfection.

The class 2 variant *c.1646G>A;p.(Gly549Asp)* was used as a control since it was functionally confirmed in fibroblasts of a homozygous patient, and the activity of the receptor was < 2% [Hobbs et al. 1992] also this control has been validated in CHO-*ldlA7* getting the same results [Rodriguez-Jimenez et al. 2019]

LDLR activity and cell surface LDLR expression in transfected cells. Human LDL was isolated from a single donor and labelled with 1,1'-dioctadecyl-3,3,3,3'-tetramethylindocarbocyanine perchlorate (DiI; Molecular Probes, Life Technologies) [Calvo et al. 1998]. The transfected cells were incubated for 4 h at 4 °C and 37 °C for testing binding and uptake, respectively, with 20 µg of protein DiI-LDL/ml. Then cells were washed twice with phosphate-buffered saline-1% bovine serum albumin (PBS-2% BSA), and analysed by flow cytometry (BD FACSCanto™). Forward scatter and side scatter gates were established to exclude dead cells and cell debris, also 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; Invitrogen™, ThermoFisher-Scientific, Waltham, USA) solution was added directly to the samples to a final concentration of 0.1% in order to separate live cells of dead cells. The acquisition number of cells was set at 10⁴. All assays were performed in three independent experiments.

In order to quantify of cell surface LDLR expression, the transfected cells were incubated with an allophycocyanin-conjugated antibody against the human LDLR (clone 472413, R&D Systems) at 0.312 µg/mL or an IgG1 isotype control (Life Technologies) for 20 min at room temperature in the dark. Then the cells were washed, acquired and processed in the same conditions described above. All assays were performed in three independent experiments.

Colocalization of LDLR-Dil-LDL by immunofluorescence in transfected cells. CHO-*ldla7* were seeded in coverslips at 3x10⁴/well and transiently transfected with the constructs. After 48 hours the cells were incubated for 4 hours at 4°C or 37°C with 100µg of protein DiI-LDL/ml, fixed with 2% paraformaldehyde (PBS-2%PFA) for 10 min at room temperature and permeabilized with 0,1% TritonX-100 for 3 min at room temperature. Cells were incubated overnight at 4°C with primary antibody against GFP (Sino-Biological, Wayne, USA), fluorescent secondary antibody Alexa fluor-488 goat anti-mouse IgG (Invitrogen™,

ThermoFisher-Scientific, Waltham, USA) for 1 hour and finally with DAPI during 20 min at room temperature; after each incubation period the samples were washed three times with PBS-2%BSA. The coverslips were mounted on glass slides with ProLong (Invitrogen™, ThermoFisher-Scientific, Waltham, USA). Images were obtained with a Leica DM5500 B fluorescence microscope (Leica Microsystems, Germany).

3. Results

3.1 Genetic analysis.

NGS data from the proband's DNA blood sample was suitable for analysis after passing the quality parameters established in our laboratory: number of reads more than 30x in the 99% of the target bases. The results of the bioinformatics analysis of data from NGS showed the *c.1951G>T:p.(Asp651Tyr)* variant in 12% of the reads, located in exon 13 of the *LDLR* gene. This is a known pathogenic variant associated with FH [Humphries et al. 2006]. Sanger sequencing was negative for this variant. The presence of large deletions or insertions in the whole gene was discarded by MLPA.

When we performed a visual analysis of the bam file using Alamut software we observed three types of reads within exon 13. Close to our pathogenic variant of interest (*c.1951G>T*) we found the synonymous polymorphism *c.1959T>C* (rs5925), with a MAF/MinorAlleleCount for C=0.336/1682 according to dbSNP. This allowed us the visualization of three different compound genotypes: the first one included the pathogenic variant *c.1951G>T* plus the wild type allele for position *c.1959T* in approximately 12% of the reads. The second one carries the wild type allele for position *c.1951G* and the polymorphism *c.1959T>C* in 48% of the reads; and the third one carries the wild type allele for position *c.1951G* and the *c.1959T* in 40% of the reads (**Figure 1**).

A new blood sample from the proband was requested in order to confirm the results. After DNA extraction and NGS analysis of the new sample, the pathogenic variant *c.1951G>T* was observed in 9% of reads, confirming the previous results.

The variant *c.1951G>T*:p.(Asp651Tyr) has been previously described and associated with familial hypercholesterolemia [Humphries et al. 2006]. It affects a highly conserved nucleotide (phyloP: 5.69 [-14.1;6.4]) and a highly conserved amino acid (considering 11 species). There is also a large physicochemical difference between Asp and Tyr (Grantham dist.: 160 [0-215]). We use the *in silico* predictors of pathogenicity CADD, VEST, Sift, and Mutation taster and all of them classified the variant as deleterious. This variant is located at the protein domain LDLR class B repeat. Align GVGD: C15 (GV: 85.08 - GD: 92.25). This variant was classified as pathogenic accordingly to ACMG criteria for genetic variant classification [Nordestgaard et al. 2013].

We extended the genetic study to the proband's family. Two of the children with hypercholesterolemia were analyzed by NGS and they were heterozygous for the variant *c.1951G>T*:p.(Asp651Tyr). These results were confirmed by Sanger sequencing. The proband's wife and the other two children with normal levels of cholesterol showed the common allele at the position *c.1951* by Sanger sequencing. Family pedigree is shown in **Figure 2**.

In order to confirm the mosaicism at position *c.1951* of the *LDLR* by an orthogonal method, we performed pyrosequencing analysis in the DNA extracted from peripheral blood from the proband, using as a positive control the older daughter, who was heterozygous for the variant *c.1951G>T* by NGS. The results showed that the proband is a mosaic with the pathogenic variant found in 20% of the reads in peripheral blood DNA. Heterozygous state was confirmed in the daughter. Since mosaicism can affect different tissues in a different

proportion, we also decide to analyze a saliva sample from the proband, using the same three technologies. NGS sequencing and pyrosequencing analysis showed a mosaic of 8% and 18% of the pathogenic variant respectively, whereas Sanger sequencing hardly showed the variant. **Figure 3** summarizes the results of blood and saliva at the position *c.1951* at the exon 13 of the *LDLR* using the different types of technologies.

3.2. Functional characterization of the variant LDLR:c.1951G>T;p.(Asp651Tyr)

Effect of the variant in the LDLR activity and cell surface LDLR expression in transfected cells. The variant p.(Asp651Tyr) showed 80% and 50% of Dil-LDL uptake and binding respectively and 50% expression of cell surface LDLR with respect to WT. The results revealed significant differences in binding and expression of cell surface LDLR (**Figure 4**).

Colocalization of LDLR-Dil-LDL in transfected cells. Immunofluorescence study revealed expression of LDLR in the variant p. (Asp651Tyr) and colocalization of LDLR-Dil-LDL. (**Figure 5**).

Discussion

In this study we report the first case of a mosaic single-nucleotide variant (SNV) affecting the most relevant gene involved in familial hypercholesterolemia, the *LDLR* gene. This phenomenon occurs when a *de novo* genetic variant appears post-zygotically causing a mosaicism in the embryo. The timing and location in which the variant occurs plays a key role in the distribution of the variant among tissues (that is, somatic, germline or gonosomal) and therefore, determine the final phenotype in the affected patient as well as the patterns of disease reoccurrence within families. Several biological specimens can be tested to detect mosaicism. In this study, we have analyzed two types of specimens, blood and saliva.

Although the liver is the target tissue for molecular diagnosis of hypercholesterolemia, since it is the main organ responsible for lipid homeostasis, it is not justified to obtain such a type of sample for genetic analyses due to the invasive nature of sampling. In any case, we assert that the proband's liver is affected to such an extent grade conferring the patient with hypercholesterolemia phenotype. The proband's germline tissue is also affected since two of his children are carriers of the pathogenic variant in a heterozygous form.

The presence of mutations in the form of somatic mosaicism is well known in pathologies that, in a classical way, are caused by constitutionally dominant mutations, and therefore compatible with life. As an example, the neurofibromatosis type I due to mutations in one of the most extensive genes in humans, the *NF1* gene. However, mutations in this gene can also occur in the form of somatic mosaicism giving rise to what is known as segmental NF1, since the clinical manifestations are limited to a part of the body [Biesecker et al. 2013]. Other examples of dominant diseases that may occur as mosaic include Duchenne myotonic dystrophy and the hereditary haemorrhagic telangiectasia.

Other pathologies present only in the form of mosaicism and are not heritable, probably because they are incompatible with life in germinal form. As an example we have the *PIK3CA* related overgrowth spectrum (PROS) [Keppler-Noreuil et al. 2015] or the McCune-Albright syndrome. A third form of mosaicism is the one causing germline mosaic disorders, in which dominant pathologies occur in several children of phenotypically healthy parents, due to the presence of a mutation in the germline tissue of one of the parents. The most reported pathology with this type of mechanism is osteogenesis imperfecta type II [Biesecker and Spinner. 2013].

The mosaic variant in the *LDLR* gene presented in this work belongs to the group of mosaicisms in dominant genes. However, in addition to the somatic mosaicism, the patient

transmitted the variant to two of his children, so his germ cells must be affected. This combination between somatic mosaicism and germinal mosaicism is known as gonosomal mosaicism. In this study, we performed the functional characterization of the variant p.(Asp651Tyr) since it was previously reported but lack of functional analysis. Our study showed an impairment of LDLR function.

The case described here confirms and highlights the fact that mosaicism does not always occur in clinically easily identifiable disorders. Therefore, we consider that in the FH the molecular confirmation of a mosaic pathogenic variant is especially useful in those patients with LDL-c borderline, and allows cascade screening in the family to further establish lifestyle changes or medication for early cardiovascular prevention. Although heterozygous FH frequency is around 1/200-250, a minor fraction of them is identified [Catapano et al. 2016], for this reason improving the molecular detection is of great interest. The use of new technologies such as the high-throughput deep sequencing allows the detection of mosaics. As we describe in our study it is possible that mosaicism can explain some of the cases of patients without molecular confirmation. This phenomenon can be suspected specially in those molecular unconfirmed cases with strong familial segregation, even more when less sensitive methods for analysis were previously used. Even if NGS was used for molecular diagnosis of FH, the bioinformatic treatment of the data should consider the possibility of a mosaicism so that these types of variants are not filtered prior to the variant calling. For example, in routine laboratories, a common variant call ratio used for filtering variants leaves out those that not reach the 20% of the reads. Therefore, the bioinformatics algorithms used for routine analysis need to improve in order to better detect mosaicism. Recent publications have addressed this topic [Huang et al. 2017]. On the other hand, another aspect to take into account is the reading depth usually used. Improving the deep coverage up to 1000x will increase the ability to detect low-level mosaicism. However, at present a relatively high cost

of the test may be a major limitation for large-scale application of this approach in diagnostic laboratory. Something that will probably be solved soon with the technological advances in the NGS platforms.

Sanger sequencing has been the gold standard technology used in most laboratories but it fails to identify cases of mosaicism. It is difficult to detect when the grade of mosaicism is low. Sanger sequencing possess less sensitivity compared with massive sequencing in detecting mosaicism. Recent publications state that the use of sanger sequencing is not adequate to validate NGS findings specially when the last one shows good quality parameters such as a good coverage of reads and quality [Beck et al. 2016; Sikkema-Raddatz et al. 2013; Strom et al. 2014]. The authors state that the use of sanger validation should be restricted for those cases in which the result of NGS has a serious implication for the probands and their families [Beck et al. 2016].

With the extensive use of massive sequencing techniques more genes are screened routinely providing ever more information. With adequate tools to process the information, it is possible to explore not only the main genes responsible for FH but other candidate genes. In our experience, certain patients remitted to our laboratory because suspicious of FH have genetic variants in other genes involved in cholesterol metabolism. It is possible that these candidate genes could explain the lipid disturbances in these patients, and mosaicism could also be present in them.

Nowadays, there is no doubt that next generation sequencing provides us with the better results, improving not only the detection of new mutations, but also allowing the identification of new genes and mosaicism. Further analysis is warranted to know the extent of the phenomenon of mosaicism not only in familial hypercholesterolemia but in other dyslipemias.

Financial support: the study was supported by grant FIS15/00122, FIS18/00917 ISCIII-Instituto de Salud Carlos III, Spain

Conflicts of interest: The authors have not conflicts of interest

Author contributions: SRN, analyzed the NGS data, design the study and write the manuscript; CRJ performed functional analysis, sanger analysis and figures; CA and IGP contributed to clinical management, and follow-up of the patient and review the manuscript; LRL and GG performed pyrosequencing analysis and figures; VM-G contributed to review the data and the manuscript. All authors have approved the final article.

REFERENCES

Beck, T. F., Mullikin, J. C., Program, N. C. S. & Biesecker, L. G. (2016). Systematic Evaluation of Sanger Validation of Next-Generation Sequencing Variants. *Clin Chem*, 62, 647-654.

- Biesecker, L. G. & Spinner, N. B. (2013). A genomic view of mosaicism and human disease. *Nat Rev Genet*, 14, 307-320.
- Calvo, D., Gomez-coronado, D., Suarez, Y., Lasuncion, M. A. & Vega, M. A. (1998). Human CD36 is a high affinity receptor for the native lipoproteins HDL, LDL, and VLDL. *J Lipid Res*, 39, 777-788.
- Catapano, A. L., Graham, I., De Backer, G., Wiklund, O., Chapman, M. J., Drexel, H., Hoes, A. W., Jennings, C. S., Landmesser, U., Pedersen, T. R., Reiner, Z., Riccardi, G., Taskinen, M. R., Tokgozoglu, L., Verschuren, W. M. M., Vlachopoulos, C., Wood, D. A., Zamorano, J. L., Cooney, M. T. & Group, E. S. C. S. D. (2016). 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *Eur Heart J*, 37, 2999-3058.
- De Ferranti, S. D., Rodday, A. M., Mendelson, M. M., Wong, J. B., Leslie, L. K. & Sheldrick, R. C. (2016). Prevalence of Familial Hypercholesterolemia in the 1999 to 2012 United States National Health and Nutrition Examination Surveys (NHANES). *Circulation*, 133, 1067-1072.
- Forsberg, L. A., Gisselsson, D. & Dumanski, J. P. (2017). Mosaicism in health and disease - clones picking up speed. *Nat Rev Genet*, 18, 128-142.
- Goldstein, J. L., H. H. Hobbs, M. S. Brown, Baudet A.L, Sly W.S., Valle D, Editors. (2001). Familial hypercholesterolemia. *The Metabolic and Molecular Bases of Inherited Disease*. New York, NY:McGraw-Hill, 2863–2913.
- Hobbs, H. H., Brown, M. S. & Goldstein, J. L. (1992). Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat*, 1, 445-466.
- Huang, A. Y., Zhang, Z., Ye, A. Y., Dou, Y., Yan, L., Yang, X., Zhang, Y. & Wei, L. (2017). MosaicHunter: accurate detection of postzygotic single-nucleotide mosaicism through next-generation sequencing of unpaired, trio, and paired samples. *Nucleic Acids Res*, 45, e76.
- Humphries, S. E., Cranston, T., Allen, M., Middleton-Price, H., Fernandez, M. C., Senior, V., Hawe, E., Iversen, A., Wray, R., Crook, M. A. & Wierzbicki, A. S. (2006). Mutational analysis in UK patients with a clinical diagnosis of familial hypercholesterolaemia: relationship with plasma lipid traits, heart disease risk and utility in relative tracing. *J Mol Med (Berl)*, 84, 203-214.
- Keppler-Noreuil, K. M., Rios, J. J., Parker, V. E., Semple, R. K., Lindhurst, M. J., Sapp, J. C., Alomari, A., Ezaki, M., Dobyys, W. & Biesecker, L. G. (2015). PIK3CA-related overgrowth spectrum (PROS): diagnostic and testing eligibility criteria, differential diagnosis, and evaluation. *Am J Med Genet A*, 167A, 287-295.
- Kingsley, D. M. & Krieger, M. (1984). Receptor-mediated endocytosis of low density lipoprotein: somatic cell mutants define multiple genes required for expression of surface-receptor activity. *Proc Natl Acad Sci U S A*, 81, 5454-5458.
- Krieger, M., Martin, J., Segal, M. & Kingsley, D. (1983). Amphotericin B selection of mutant Chinese hamster cells with defects in the receptor-mediated endocytosis of low density lipoprotein and cholesterol biosynthesis. *Proc Natl Acad Sci U S A*, 80, 5607-5611.
- Leigh, S., Futema, M., Whittall, R., Taylor-Beadling, A., Williams, M., Den Dunnen, J. T. & Humphries, S. E. (2017). The UCL low-density lipoprotein receptor gene variant database: pathogenicity update. *J Med Genet*, 54, 217-223.
- Nordestgaard, B. G., Chapman, M. J., Humphries, S. E., Ginsberg, H. N., Masana, L., Descamps, O. S., Wiklund, O., Hegele, R. A., Raal, F. J., Defesche, J. C., Wiegman, A., Santos, R. D., Watts, G. F., Parhofer, K. G., Hovingh, G. K., Kovanen, P. T., Boileau, C., Averna, M., Boren, J., Bruckert, E., Catapano, A. L., Kuivenhoven, J. A., Pajukanta, P., Ray, K., Stalenhoef, A. F., Stroes, E., Taskinen, M. R., Tybjaerg-Hansen, A. & European Atherosclerosis Society Consensus, P. (2013). 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*, 34, 3478-3490a.

- Rodriguez-Jimenez, C., Pernia, O., Mostaza, J., Rodriguez-Antolin, C., De Dios Garcia-Diaz, J., Alonso-Cerezo, C., Garcia-Polo, I., Blanco, A., Lahoz, C., Arrieta, F., Beltran, L., Diaz de Bustamante, A., Garzon-Lorenzo, L., Alvarez-Sala, L. A., Asenjo, A., Ibanez de Caceres, I. & Rodriguez-Novoa, S. (2019). Functional analysis of new variants at the low-density lipoprotein receptor associated with familial hypercholesterolemia. *Hum Mutat*, 40, 1181-1190.
- Sikkema-Raddatz, B., Johansson, L. F., De Boer, E. N., Almomani, R., Boven, L. G., Van den Berg, M. P., Van Spaendonck-Zwarts, K. Y., Van Tintelen, J. P., Sijmons, R. H., Jongbloed, J. D. & Sinke, R. J. (2013). Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat*, 34, 1035-1042.
- Strom, S. P., Lee, H., Das, K., Vilain, E., Nelson, S. F., Grody, W. W. & Deignan, J. L. (2014). Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. *Genet Med*, 16, 510-515.
- Tejedor, D., Castillo, S., Mozas, P., Jimenez, E., Lopez, M., Tejedor, M. T., Artieda, M., Alonso, R., Mata, P., Simon, L., Martinez, A., Pocovi, M. & Spanish, F. H. G. (2005). Reliable low-density DNA array based on allele-specific probes for detection of 118 mutations causing familial hypercholesterolemia. *Clin Chem*, 51, 1137-1144.

Figure legends

Figure 1. NGS analysis of proband's blood. Visualization of BAM file using Alamut software focus on the exon 13 of *LDLR*. The arrow indicates the position of the genetic variant *LDLR:ex13:c.1951G>T:p.(Asp651Tyr)*.

Figure 2. Pedigree of the family showing the segregation of the variant *LDLR:ex13:c.1951G>T:p.(Asp651Tyr)*. The arrow indicates the proband individual; circle and square symbols represent women and men respectively; shadow filled symbols indicate the affected members with hypercholesterolemia. Line 1 below symbols correspond to the individual identification, line two indicates the age; line 3 indicates the *LDLR* genotype: p.(Asp651Tyr) or normal genotype at this position (wt).

Figure 3. Results of NGS, Sanger sequencing and pyrosequencing analysis performed in blood and saliva samples. NGS analysis of blood samples show the reads with the pathogenic variant in the proband and his daughter. Alamut software was used to visualize BAM file. **Sanger** sequencing corresponding to exon 13 of the *LDLR* shows that the variant is hardly detected in the proband whereas it is clearly detected in his daughter. **Pyrosequencing** analysis of proband and daughter blood shows approximately 20% and 50% of pathogenic variant, respectively. **Saliva** sample from the proband was analyzed using the same three technologies. NGS sequencing and pyrosequencing analysis showed a mosaic of 8% and 18% of the pathogenic variant respectively, whereas Sanger sequencing hardly shows the variant.

Figure 4. LDLR activity and cell surface LDLR expression in the transiently transfected CHO-*ldla7* cells, with WT, control p.(Gly549Asp) and variant p.(Asp651Tyr). (A) Uptake Dil-LDL. (B) Binding Dil-LDL. (C) Cell surface LDLR expression. Analysis performed by flow cytometry as described in Material and Methods. The specific median intensity of fluorescence of each sample was estimated after subtracting the median intensity of fluorescence of the empty vector. We represent the data as the percentage of each variant in relation to the WT. Each result correspond to mean \pm S.E.M. of the triplicate experiments compared with the WT. Statistical significance were determined by a two-tailed Student's t-test and a 95% confidence

interval. * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$, **** $p < 0,0001$ versus WT. The results were obtained in three independent experiments.

Figure 5. Expression of LDLR in the transiently transfected CHO-*ldlA7* cells, with WT, control p.(Gly549Asp) and LDLR variant p.(Asp651Tyr). Images of immunofluorescence show LDLR expression and colocalization of LDLR-Dil-LDL.